GENETIC DIVERSITY OF CAPTIVE *Tomistoma schlegelii* REVEALED BY MICROSATELLITE AND mtDNA MARKERS AND FIELD SURVEY DATA AT REPORTED SITES OF OCCURRENCE IN PENINSULAR MALAYSIA

By

TARANJEET KAUR

A thesis submitted to the Faculty of Science,
Universiti Tunku Abdul Rahman,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy (Science)
May 2024

ABSTRACT

GENETIC DIVERSTIY OF CAPTIVE *Tomistoma schlegelii* REVEALED BY MICROSATELLITE AND mtDNA MARKERS AND FIELD SURVEY DATA AT REPORTED SITES OF OCCURRENCE IN PENINSULAR MALAYSIA.

Taranjeet Kaur

To contribute towards the on-going captive management of the Sunda gharial, *Tomistoma schlegelii*, a population-level genetic study was conducted using microsatellite and mitochondrial DNA (mtDNA) markers, alongside field surveys of confirmed occurrence sites. Microsatellite data identified six clusters (Clusters A to F), while mtDNA data revealed four haplotypes (H1, H2, H5, and the divergent H4 lineage) from 38 individuals. Nuclear genetic diversity within clusters was low (mean 0.25 ± 0.14 to 0.49 ± 0.11), and mtDNA diversity was detected only in Cluster B. Analysis of molecular variance (AMOVA) based on microsatellite data indicated moderate to substantial genetic differentiation. However, AMOVA for mtDNA data showed contrasting results for population differentiation and sex-biased dispersal analyses, depending on whether the divergent H4 haplotype was included or excluded. Coancestry analysis combining both markers identified Clusters B and E as important for breeding purposes. Separately, the daytime field surveys at recent capture of *T. schlegelii*, namely the Tengi River, Tekah/Nerus River, Sungkai River and Jengka River were carried out to record water parameters and geographical coordinates of capture site. The water parameters, including temperature, pH, salinity, total dissolved solids (TDS), and conductivity, revealed significant differences in pH and temperature between wet and dry seasons. Notably, the pH data aligned with

previous studies, which suggested a preference for neutral to mildly acidic rivers. Additionally, terrain analyses and Geographic Information System (GIS)-based assessments indicated that specific river features—such as elevations of 11–16 meters, smaller streams (1–3 Strahler Order), and slopes of 0° to 10°—were consistent across capture sites. These features, reportedly associated with nesting behaviour, combined with mtDNA AMOVA results, suggest potential movement of *T. schlegelii* females during the breeding season.

Keywords: *Tomistoma schlegelii,* microsatellite markers, mitochondrial DNA markers, field surveys, Geographic Information System (GIS)

Subject Area: QH426-470 Genetics

ACKNOWLEDGEMENT

I would like to express my gratitude to my supervisors, Prof. Dr. Alan Ong Han Kiat, Dr. Gideon Khoo, Dr. Khairul Nizam Abdul Maulud and Mr. Khairul Amirin (Department of Wildlife and National Parks (Perhilitan)) for their guidance and contribution through my postgraduate journey. I am grateful to UTAR for funding this study through UTARRF FMHS Research Grant (Vote 6200/A20) and the subsequent extension without which this work would not have started. I am equally grateful to Perhilitan for granting the special permits B-00004-16-16 and subsequently B-00973-16-19 to carry out this study involving the endangered *Tomistoma*.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The Sunda gharial, *Tomistoma schlegelii*, is a freshwater crocodilian occurring mainly in Malaysia and Indonesia, is one of the larger crocodilians found in peat and freshwater swamps. It is listed as endangered under criteria C 2a (i) in The IUCN Red List of Threatened Species and is also considered as "endangered" in Malaysia. It is a highly elusive crocodilian and the main threat to this species is habitat destruction due to wildfires and land conversion for agriculture and urbanization (IUCN, 2022).

Peat swamps are a critically endangered forested wetlands characterized by deep layers of peat soil and highly acidic waters, which support unique flora and fauna not found in other tropical forests in Asia (Jon *et al*., 2010). Due to the difficulty in accessing these ecosystems, their biodiversity is not well documented. Because of this, from the 36% of the historic peat swamps that remain till date in Southeast Asia, only 9% of the total peat swamps is classified as 'Protected Areas' by IUCN (Posa *et al*., 2011). The rapid disappearance of peat swamps is primarily due to deforestation for plantations, urbanization, logging, and drainage (Gunarso *et al*., 2013). As a mound nester, leaf litter and twigs have been reported as materials used by *T. schlegelii* for nest construction (Bezuijen *et al*., 1998; Bezuijen *et al*., 2001; Staniewicz *et al*., 2018, Lariman *et al*., 2021). Therefore, the land clearing mention above directly affects the breeding ecology of this species. Additionally, flooding caused by land conversion for agriculture and water irrigation has resulted in the displacement of *T. schlegelii*.

Due to ongoing habitat destruction and efforts to mitigate human-crocodilian conflict, most wild individuals of *T. schlegelii* are caught and held ex-situ by authorities. Over the years, the number of captive-held *T. schlegelii* has gradually increased in Peninsular Malaysia. Recent captives have been from various rivers, including the Dusun River, Selangor (2013), Tekah River, Terengganu (2014), Jengka River, Pahang (2014, 2015), and the Sungkai River, Perak (2016). These individuals were caught due to flooding, accidental entanglement with fishing equipment, or proximity to wildlife conservation centers. Despite being kept in captivity, there is a lack of conservation strategies and breeding plans for *T. schlegelii*. Currently, all wild caught individuals are placed indiscriminately (of any sex and age) and without consideration of their genetic variation in the Sungai Dusun Wildlife Conservation Centre. Similarly, in Zoo Negara, the breeding efforts initiated in 2002, without any knowledge of the genetic component, resulted in five hatchling that have survived till date. Although this is the only successful breeding in Peninsular Malaysia, the future genetic contribution to the total population is unclear as possible inbreeding would negatively contribute to the on-going conservation efforts.

So far, ecological studies in areas of confirmed occurrence or sightings in Peninsular Malaysia are lacking with previous studies focusing on peat swamps as habitats, compilation of historical report and unofficial sightings by locals (Sebastian, 1993; Stuebing *et al*., 2004; Stuebing *et al*., 2006; Simpson, 2014). The recent ecological study in Malaysia confirmed *T. schlegelii* occurrence in Sarawak. However, after the capture of an adult in the Samarahan River and two hatchlings in the Bunga-Baki River, subsequent observations were not detected

in the field surveys (Hassan *et al*., 2016). Similarly, studies employing Geographic Information System (GIS) have been mainly for mapping historical records and current data from Indonesia (Stuebing *et al*., 2006, Rodder *et al*., 2010). These were used to predict the current habitat protection status and identification of conservation areas. A study by Bonke *et al*. (2014), indicated homing instinct in juveniles when GIS was used to map its movement via telemetry. A subsequent study failed to recapture hatchlings in the mark-recapture attempt due to the elusive nature of this species and the difficulty in navigating its habitat (Stuebing, Sommerlad and Staniewicz, 2015). Therefore, till date, apart from the study by Bonke *et al*. (2014), there is no other study on the movement of *T. schlegelii*.

Additionally, genetic study at the population level is also lacking. Currently, intraspecific variation studies have been carried out using mitochondria DNA (mtDNA), mainly the control region and the partial ND 6-tRNA^{glu}-cyt *b* fragment (Kaur *et al*., 2013; Md Adzhar *et al*., 2017) and the D loop (Kurniati, 2003). These studies were able to identify haplotypes and their occurrence across Malaysia and Kalimantan. Kaur *et al*. (2013) identified a highly divergent lineage, the H4 haplotype in Peninsular Malaysia. Shafiei-Astani *et al*. (2015) was able to identify two clusters using a subsample of the Malaysian individuals from Kaur *et al*. (2013). These findings were obtained using universal ISSR markers and microsatellite markers developed by Chaeychomsri *et al*. (2008) and FitzSimmons (2001) for crocodiles, and Glenn *et al*. (1998) for the American alligator. Although the microsatellite markers were not specific for *Tomistoma*, 17 out of the 40 markers were able to amplify with ten loci showing

polymorphism. The Malaysian *T. schlegelii* population was separated into one consisting exclusively of Peninsular Malaysian individuals and the other with individuals from both Sarawak and Peninsular Malaysia (Shafiei-Astani *et al*., 2015). Another study carried by Chang *et al*. (2014) was for development of species-specific microsatellite markers from liver samples of this species. 18 new microsatellite markers by the hybridization capture method proved to be polymorphic on 19 individuals of unknown origins.

Protection of endangered species requires inputs from various fields, namely ecology and genetics. Ecological aspects are mostly protecting natural habitats and mitigating wildlife-human conflict, while genetical aspect involves identifying genetic diversity and population structure, and captive breeding programs (Hedrick, 2001; Galla *et al*, 2022). Genetic diversity is important especially in an endangered species to enable it to genetically respond to ecological variability. Indiscriminate placement of males and females of unknown genetic similarities or distance could result in higher inbreeding or outbreeding, defeating the purpose of holding the endangered species in rehabilitation centres and zoos (Hedrick, 2001; Presti and Wasko, 2014). To facilitate this, genetic markers have been developed so that the genetic variation can be quantified.

Microsatellites, also known as simple tandem repeats (STR), are tandem repeats of up to ten nucleotides found in both coding and non-coding regions of nuclear and organellar DNA prokaryotes and eukaryotes. They have a high degree of mutation (ranging from 10^{-6} to 10^{-2} per generation) due to slippage during

replication of DNA and recombination which creates variation in numbers and length of the alleles. The biparentally inherited alleles, homozygous or heterozygous, are easily detected by electrophoresis because microsatellites are co-dominant markers, hence their popularity as a marker for population level studies (Zane *et al*., 2002; Selkoe and Toonen, 2006; Vieira *et al*., 2016). Mitochondria DNA (mtDNA) markers are also commonly used marker for intraspecific variation due to the maternal inheritance, abundance, easy isolation, no recombination, small effective population size and a high mutation rate in general (Avise *et al*., 2004).

The combination of both microsatellite and mtDNA markers have proven to be useful in the African and American crocodiles by revealing genetic diversities, historical and contemporary gene flow, demography and population structure (Velo-Anton *et al*. 2014; Van Asch *et al.* 2019; Rossi *et al.* 2020; Sharma *et al.* 2021;). Additionally, a framework for establishing ecological networks was suggested for the Nile crocodile based on the dispersal of individuals and the source-sink dynamics of population of this vagile species (Velo-Anton *et al.* 2014)

Hence, due to the gradual increase of the captive-held *Tomistoma schlegelii*, the documentation of genetic variation, from both the maternal and nuclear markers, is necessary to make informed management and breeding plans for *T. schlegelii*. Similarly, since the recent capture sites were documented, investigating the river and habitat characteristic would help to narrow down the type, and variation of habitats preferred by this species in Peninsular Malaysia.

In this study, blood samples of 28 individuals from the previous mtDNA study (Kaur *et al*., 2013) and ten new individuals (including three with confirmed locations) were analysed to access their genetic variation at population level. The species-specific microsatellite markers (Chang *et al*., 2014) and the mtDNA markers for the partial ND 6-tRNA^{glu}-cyt *b* and the D-loop gene fragments (Kaur *et al*., 2013) were used for this purpose. Additionally, four confirmed locations of occurrence (the Ban Canal / Tengi River, Selangor; Sungkai River, Perak; Jengka River, Pahang and Tekah River, Terengganu) were accessed for water parameters, namely temperature, pH, salinity, total dissolved solids (TDS) and conductivity, using the descriptive and inferential analyses. Further physical similarities and differences between these were analysed using the terrain and distance analyses in GIS.

Therefore, the objectives of this study are:

1) Document the genetic diversity, population structure of *T. schlegelii* from species-specific microsatellite,

2) Document mtDNA diversity of clusters identified with microsatellite markers above,

3) Analyse and compare the water parameters, river features, terrain (elevation, slope, stream order) and distance to the nearest hub of the four confirmed sites of capture.

CHAPTER 2

LITERATURE REVIEW

2.1 *Tomistoma schlegelii*

2.1.1 Biology and Behaviour

The freshwater Sunda gharial, *Tomistoma schlegelii*, is a large crocodilian with body sizes ranging from 4.5-5.0 m in adult males while adult females are 2.5-3.0 m long (Bezuijen *et al*.,1998; Bezuijen *et al*., 2010). It was discovered in Borneo by Salomon Müeller in 1836 and named *Tomistoma schlegelii* in 1838 (Das, 2004). *Tomistoma* translated into Greek means 'sharp mouth' and was named based on the morphology of its sharp teeth along the long and slender snout (Bezuijen *et al*., 2001). Referring to these features, recently it has been loosely called the chopstick crocodile (Lariman *et al*., 2021) (Fig. 2.1). They are thought

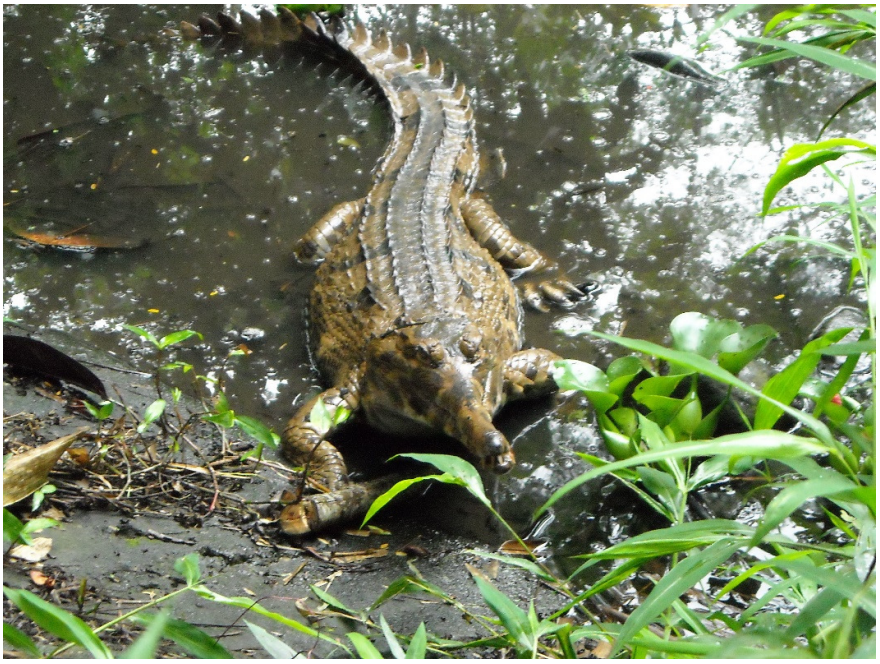


Fig. 2.1. Adult *T. schlegelii* in captivity.

to be more diurnal (Setiawan et al., 2019). However, in Peninsular Malaysia, the recent capture of wild *Tomistoma schlegelii* have been in the daytime when fishermen were setting traps.

Juveniles are darker than adults with retention of most of the juvenile patterns into adulthood. However, their dark underbelly turns lighter and finally white as they grow older (Bezuijen *et al*., 2001). Staniewicz, Youngprapakorn and Jones (2018) investigated the environmental factors like temperature, ambient light, background colour, and even stress for this physiological colour change and identified illumination as the only factor affecting the colour change.

Fig. 2.2. Underbelly of *T. schlegelii*. (A) A hatchling with dark underbelly. (B) Adult with white underbelly. (Photos by Taranjeet Kaur).

The main diet in juveniles include insects, arachnids, crustaceans, fishes and mammals (Staniewicz, 2011) while the diets in adults include shrimps, fish, monkeys and small deer (Bezuijen *et al*., 2001). Recently, Sideleau *et al*. (2022), reported *T. schlegelii* as preying on humans and suggested to investigate anthropogenic factors that could be causing this scenario.

Apart from a single study describing the anatomical and histological features of the phallic anatomy of adult male (Moore, Fitri and Augustine, 2020), information on the male and female biology is lacking for *T. schlegelii*. Mating in captivity have been observed from the month of November to February and April to June (Mathew *et al*., 2011). *Tomistoma schlegelii* is a mound nester and clutch sizes of 19 to 35 eggs have been reported (Bezuijen *et al*., 1998; Bezuijen *et al*., 2001; Mathew *et al*., 2011). The nests are always built on land, at a base of a tree with leaf litter and twigs (Bezuijen *et al*., 1998; Bezuijen *et al*., 2001; Staniewicz *et al*., 2018, Lariman *et al*., 2021). Hatchlings in captivity were reported in September (Mathew *et al*., 2011). In the wild, hatchlings were mostly seen in October to December after nesting season between the months of June to August (Staniewicz *et al*., 2018). Incubation period ranges from 72 to 100 days (Mathew *et al*., 2011; Staniewicz *et al*., 2018). Like all other crocodilian, the sex of hatchlings is temperature-dependent (Knotz, 2023), but the incubation temperature for male and female hatchlings has not been reported for this species.

2.1.2 Distribution and Habitat

Tomistoma schlegelii is found primarily in Sundaland about 5 degrees north and south of the equator, in freshwater or peat swamps that ranged across the land mass (Auliya *et al*., 2006; Bezuijen *et al*., 1998; 2002; Brochu, 2003; Stuebing *et al*., 2006). Although historically, the family Tomistominae was reported to be vastly distributed in the old world (Brochu, 1997), current reports have localised this species to mainly the peat swamps of Indonesia and Malaysia (Cox and Gombek, 1985; Sebastian, 1993; Bezuijen *et al*., 2002; Auliya *et al*., 2006;

Simpson, 2014) and Brunei and perhaps southern Thailand (Shaney *et al*., 2016) (Fig 2.3). In Sumatra and central Kalimantan, the rivers where this species occurs were reported as highly acidic with pH of 3-4 and associated with peat swamps (Bezuijen *et al*., 2002; Bezuijen *et al*., 2004; Auliya *et al*., 2016; Shaney *et al*., 2016; Setiawan *et al*., 2019). However, in east Kalimantan, *T. schlegelii* was found in the freshwater of Lake Mesangat, reported to have a pH of 7 (Lariman *et al*., 2021).

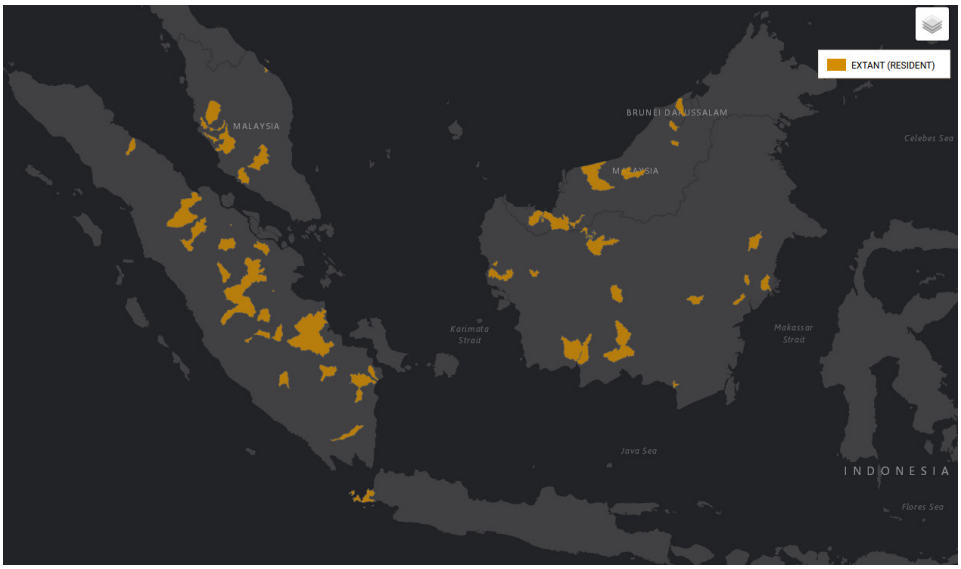


Fig. 2.3. Locations of areas of occurrence of *Tomistoma schlegelii*. Available at www.iucnredlist.org/species/21981/214287051. [Accessed 6 May 2024].

In Sumatra and central Kalimantan, the presence of the saltwater crocodile, *Crocodylus porosus* in the same habitat showed that there was habitat partitioning, with the *C. porosus* preferring coastline and brackish environments while *T. schlegelii* was restricted to freshwater and blackwater tributaries (Bezuijen *et al*., 2001; Auliya *et al*., 2006; Shaney *et al*., 2019). Similarly, in Lake Mesangat, the presence of the Siamese crocodile, *Crocodylus siamensis*, was mainly reported in open areas near floating vegetation while *T. schlegelii*

preferred the flooded forest and away from floating grass mats (Staniewicz *et al*., 2018).

2.1.3 Conservation Status of *T. schlegelii*

Tomistoma schlegelii was listed as Vulnerable in 2011 (IUCN, 2011) but had been upgraded to "Endangered" with an estimated of 2,400 individuals and a declining population trend (IUCN, 2022). The current distribution is in the freshwater and peat swamps of mostly Indonesia, Malaysia and Brunei. While its occurrence in Vietnam is uncertain, it is now extinct in Thailand. The challenge facing this species survival is mainly habitat decline as most of the habitat is undergoing conversion to agriculture land and other biological uses such as water management for irrigation for agriculture and dams, and human intrusions (IUCN, 2022). In Malaysia, it is totally protected by the Wildlife Conservation (Amendment) Act 2022 (wildlife.gov.my). In Peninsular Malaysia, due to the severe threats to its habitat, most individuals captured are placed in zoos and breeding centres. A compilation of 210 records from 1838 to 2004 across *T. schlegelii* range, recorded only 18 occurrences from Peninsular Malaysia (Stuebing *et al*., 2006). It was found that, despite still being present in Perak, Selangor and Pahang, *T. schlegelii* has become increasingly rare due to human density and land conversion.

Efforts towards conservation of this species have been limited. The Crocodile Specialist Group *Tomistoma* Task Force (CSG-TTF) formed since 2003 showed limited initiatives aside from fundraising for field research and raising international awareness to this species. Additionally, the ranching or farming of

Tomistoma schlegelii is not carried out as the skin has no commercial value, unlike other crocodilian species (Shaney *et al*., 2019).

2.1.4 Taxonomic Status of *T. schlegelii*

The taxonomic status of this species is still not resolved. *T. schlegelii* is the only surviving member of the family Tomistominae (Brochu, 1997; Brochu, 2007; Piras *et al*., 2007) and due to the gross similarity of its rostrum and jaws appearance, it is also called the false gharial. However, morphological characters and fossils placed *T. schlegelii* closer to *Crocodylus* (Brochu, 2003; Kobayashi *et al*., 2005; Piras *et al*., 2007). On the other hand, molecular data placed it as a sister taxon to *Gavialis* (Gatesy *et al*., 2003; Harshman *et al*., 2003; Janke *et al*., 2005; McAliley *et al*., 2006). A summary of this is depicted in Fig. 2.4.

Brochu and Sumrall (2020) recently suggested that perhaps previous morphological characters and fossils were based on overly conservative approach. This was due to accepting crocodilians as evolutionarily static "living fossils" while the discovery of cryptic species in the African continent crocodiles and in almost all Caimaninae suggest otherwise. These cryptic species have dermal features such as coloration and scalation that are not preserved in fossils. Another example of features not preserved in fossils is seen in the study on the penile morphology of *Tomistoma*. The glans, a site of species-specific novelties, showed similar ridge morphology to crocodiles while the tip of the glans resembles those of alligators (Moore, Fitri and Augustine, 2020). Nevertheless, the recent discovery of a Bronze Age crocodilian in China, demonstrating both gavialine and tomistomine features, lend support to the molecular data for a sister

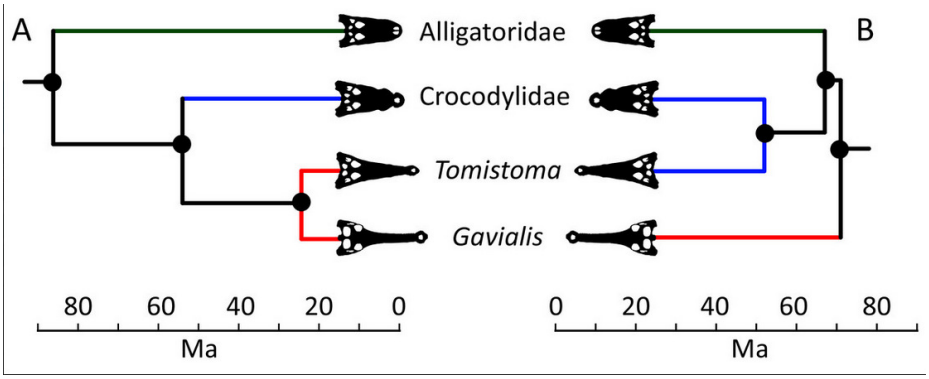


Fig. 2.4. Summary of crocodilians phylogenies based on genetical data and morphological characters. (A) Molecular phylogeny. (B) Morphological phylogeny. (Figure is from the study by Rio and Mannion, 2021).

taxon relationship (Iijima *et al*., 2022). Similarly, recent morphological-based phylogenetic analyses have started to place *Tomistoma* closer to *Gavialis* but stratigraphically early fossil Tomistominae like *Tomistoma dowsoni* is now placed closer to *Gavialis* than to *T. schlegelii* (Burke and Mannion, 2023). Although currently the morphological and molecular studies seem to complement each other more than before, with new discoveries of fossils, and better molecular technologies and understanding, the taxonomic relationship could probably change again.

The molecular data however had reported conflicting results on the common ancestor of the *Gavialis* / *Tomistoma* genera. Using the nuclear ribosomal DNA, Densmore and White (1991) found that *T. schlegelii* and *Gavialis* were closely related to *Alligator*, while recent studies have placed them closer to *Crocodylus* (Gatesy *et al*., 2003; Harshman *et al*., 2003; Janke *et al*., 2005; McAliley *et al*., 2006; Darlim *et al*., 2022).

2.2 Conservation Genetics

Conservation genetics, an interdisciplinary field of mainly genetics and ecology, ranges from prevention of species extinction, phylogenetic, and species delimitation and characterisation of genetic diversity has been applied to both naturally occurring and captive populations and species (Hedrick, 2001; Galla *et al*, 2022). Genetic diversity is important to be maintained especially in an endangered species to enable the species to genetically respond to ecological variations as loss of adaptability would increase susceptibility to extinction. Therefore, it is important to understand the nature of genetic variation, if there is selective advantage of an adaptive allele or selective disadvantage to deleterious alleles. These genetic variations are dependent on environment, population size, and genetic background of the species. Generally, habitat destruction and fragmentation would result in populations getting isolated. Depending on the genetic background, the species may not be able to adapt to the changing environment. Even if the species/population persists, due to the smaller sizes, they are predisposed to fixation of deleterious variations and inbreeding. Furthermore, the variation thought to be neutral could be adaptive or deleterious in different environment or evolutionary scenarios (Hedrick, 2001; Presti and Wasko, 2014, Willi *et al*., 2022).

Additionally, the knowledge of how this variation is distributed in a naturally occurring population is vital to understanding the demography and habitat viability so that better conservation efforts can be planned (Richardson and Whittaker, 2010). Previous analysis of static patterns of biodiversity components are getting replaced by methods that incorporate ecological, evolutionary and

landscape-level processes that generate the biodiversity hotspots but dynamic process such as migration and population sizes are still lacking (Kidd & Ritchie, 2006, Richardson and Whittaker, 2010, Bohm and Popescu, 2016). A Species Distribution Map (SDM), which includes data on species occurrence, distribution and habitat ranges, enables inferences about species niche and habitat barriers by utilising historical and current environment and species information. They also enable the prediction of the future species distribution, habitat ranges, and viability, allowing conservation and management efforts to begin (Richardson and Whittaker, 2010).

2.2.1 Conserving Captive Held Individuals

Apart from naturally occurring population, conservation genetics also involves species and individuals held in captivity. The primary focus of conservation efforts for captive wildlife is breeding for reintroduction into their habitats. However, this approach has numerous problems from establishing of self-sustaining captive populations, poor reintroduction success and high cost of maintenance (Bowkett, 2009). For a successful reintroduction program, a breeding plan based on genetic information, one which encompasses the maintenance of the evolutionary and functional genetic components is crucial. The evolutionary components provide a historical and adaptive perspective (mutation, gene flow, genetic drift, and natural selection) while the functional components are genetic traits. This ensures not only the conservation of the species in captivity but also provides a suitable stock for reintroduction purposes (Witzenberger and Hochkirch, 2011). A global survey of microsatellite studies on amphibians found that from 58 species that were reintroduced, 13 were able

to establish self-sustaining populations while another 18 were successfully breeding in the wild (Griffiths and Pavajeau, 2008).

A recent review on breeding using the studbook in 119 zoo populations, found that inbreeding was common (Witzenberger and Hochkirch, 2011). This issue arises from the assumption that founders were unrelated and that no inbreeding had occurred. In addition to closely related individuals in captivity, adaptation to captivity is another problem that could arise without proper consideration of the species' evolutionary history. Similarly, breeding with evolutionary significant unit would result in outbreeding. This poses additional challenges like managing the outbred populations with native populations, outbreeding depression and loss of local adaptation. Hybridisation is another setback for captive held wildlife and has been documented between the Bornean and Sumatran Orangutans (Witzenberger and Hochkirch, 2011).

In recent times, various genetic tools have been made available to document DNA polymorphism ranging from dominant markers that indicate the presence of an allele to co-dominant markers which detect homozygous and heterozygous individuals. Together with better statistical methods, these markers produce information that helps conservationist to understand the types of variation seen in the wild, in addition to information regarding subpopulation, demography, relatedness and founding effects (Presti and Wasko, 2014, Hedrick, 2001) and in estimates of relatedness using microsatellites especially for close relatives (Galla *et al*., 2022).

However, the approach of using just one molecular marker have led to misinterpretations. This is because markers like mtDNA could be interpreted wrongly as seen in the case of the wild primates, *M. mulatta* and *M. fascicularis*. A large overall genetic difference was documented in *M. mulatta* using the mtDNA restriction site analysis, but the allozyme polymorphism indicated that *M. fascicularis* was distinct due to population separated by water bodies (Melnick and Hoelzer, 1992). Relying solely on mtDNA, albeit highly polymorphic, resulted in an exaggerated division of populations into separate management units (MUs) due to anthropogenic population fragmentation followed by gene drift. Therefore, documenting the genetic diversity of captive populations and defining management units should incorporate multiple polymorphic genetic marker (Witzenberger and Hochkirch, 2011).

So far, the breeding approaches that have shown better results are the Minimizing Kinship approach, Maximum Avoidance of Inbreeding scheme and using a breeding circle (Caballero and Toro 2002; Witzenberger and Hochkirch, 2011). The first approach is breeding to minimise the overall level of relationships in the population. In such instances, pedigree-based analyses (inbreeding and kinship estimates) are also practiced with the same aim of maintaining genetic variability. Pedigrees-based analyses provide an estimate of kinship as the probability of alleles being identical-by-descent (IBD) from a common ancestor under the Mendelian inheritance assumption. For small population management, mean kinship, inbreeding coefficients and population-level gene diversity are usually analysed to track loss of founder alleles over time and minimize mean kinship and inbreeding. This enables holding areas to

maintain the evolutionary potential of the species. The challenge with endangered species is due to the depletion in sizes, most of the surviving individuals would have some level of co-ancestry (Galla *et al*., 2022). Maximum Avoidance of Inbreeding scheme involves equalizing family sizes in addition to mating the females to males of different subpopulations each year while breeding circles is where males are provided from each subpopulation for its neighbouring subpopulation during breeding season (Witzenberger and Hochkirch, 2011).

2.3 Measures to Document Genetic Polymorphism

DNA polymorphism can be documented with dominant markers that do not differentiate homozygotes and heterozygotes to co-dominant markers that do. Some examples of dominant markers are the Random Amplified Polymorphic DNA, RAPD; Amplified Fragment Length Polymorphism, ALFR and Inter Simple Sequence Repeats, ISSR) while the Restriction Fragment Length Polymorphism, RFLP; Simple Sequence Repeats, SSR and Single-nucleotide polymorphism, SNPs are co-dominant markers. Together with better statistical methods, these markers shed valuable insights on the types of variations seen in the wild, subpopulations, demography, relatedness and founding effects (Presti and Wasko, 2014; Hedrick, 2001).

Two examples of using the molecular markers above for conservation in crocodilians involved the use of 628 bp cytochrome *b*, cyt *b*, gene sequences and PCR-RFLP analysis of Indian crocodilians were able to provide authorities with an easy and rapid method of identifying the species of biologically degraded samples in their efforts to curb illegal poaching and smuggling (Meganathan,

Dubey and Haque, 2009). Similarly, Hassan, Koh and Abdullah (2008) analysed the cyt *b* gene fragment and the PCR-RFLP analysis of *C. porosus*, to help the local authorities with their enforcement efforts. At the population level, single nucleotide polymorphism markers, SNPs, used on the freshwater crocodile, *Crocodylus johnstoni*, at the northern Western Australia region found a low genetic diversity with strong population differentiation consistent with the three river basins, thus suggesting that the three river basins to be considered as separate management units. The study also indicated that microsatellites, or short tandem repeats (STR) were two to twelve times more informative for population comparisons and exclusion power when compared to SNPs (Cao *et al*, 2020). Additionally, relatedness of closely related individuals is also possible with the use of highly polymorphic microsatellites markers (Galla *et al*., 2022).

2.3.1 Microsatellites

Microsatellites are tandem repeats of up to ten nucleotides found in both coding and non-coding regions of nuclear and organellar DNA prokaryotes and eukaryotes. They are bi-parentally and co-dominantly inherited (Arif and Khan, 2009; Willi *et al*., 2022). They are found in random distribution within the genome due to historical processes, difference in constraints on coding and non-coding regions and functional roles of the repeat. The high degree of mutation (ranging from 10^{-6} to 10^{-2} per generation) is due to slippage during replication of DNA and recombination which creates variation in numbers and length of the alleles that can be detected by electrophoresis inherited (Arif and Khan, 2009; Willi *et al*., 2022). Mechanism for such variations are the stepwise mutation model (SMM) where mutations are due to insertion or deletion of repeat units.

The infinite allele model (IAM) is another model proposed for the variability of microsatellites in which mutations create an allelic state and infinite number of repeats not present in the population (Zane *et al*., 2002). Newer models such as the K-allele model which allows for identical probability of any alleles to mutate to any other and the two-phase, generalized stepwise (GSMM) have been proposed as well for the variability seen in microsatellites (Putman and Carbone, 2014).

Microsatellites are usually found as repeats of di, tri or tetra-nucleotides which differ in mutation rates from one taxon to another, with dinucleotides being the most widely used in genetic studies. Direct observations of microsatellite mutation in pedigrees were higher in dinucleotide repeats in natural population than tetra-nucleotides (Selkoe and Toonen, 2006). In crocodilians, the newly developed tetranucleotide markers were more polymorphic compared to the previous dinucleotide markers which were prone to stuttering as reported in *Caiman* (Amavet *et al*., 2015).

Microsatellite data can be analysed by both the frequency-based analyses like homozygosity, heterozygosity, the various fixation indexes and relatedness as well as distance-based analyses like phylogenetic trees, AMOVA, Mantel and the Principal Coordinate Analysis (PCoA) (Balloux and Lugon-Moulin, 2002; Alif and Khan, 2009; Willi *et al*., 2022).

2.3.1.1 Quantification of Variation

At the allelic level, diversity is reflected as number of effective alleles, polymorphic information index, private alleles and allelic richness while at the genotypic level parameters like observed heterozygosity (H_O) and expected heterozygosity (H_E) are the common ways of quantifying polymorphism. A locus is said to be homozygous when two identical alleles of a particular locus are inherited from each parent while in the case of heterozygotes, each parent contribute different alleles (Yunis and Ariola, 2013).

The Hardy-Weinberg Equilibrium is used to calculate the observed, H_O and expected, H_E heterozygosity frequencies when its assumptions are met using the formula, $p^2 + 2pq + q^2 = 1$, where $p^2 + q^2$ represent the homozygosity while 2pq represent the frequency of the heterozygous genotype. This calculation is based on a closed population, large enough that random mating occurs, with no natural selection and mutation affecting it. Therefore, it measures genetic polymorphism of a population in equilibrium. Deviations of any variable in the equation can either suggests genotyping error in the study or could also indicate other mechanism affecting the population such as purifying selection, inbreeding, population substructure and copy number variation (Chen, Cole and Grond-Ginsbach, 2017). Other measures have also been suggested especially in open populations or inbreeding in natural populations. The homozygosity by loci (HL) has been suggested for heterozygosity and inbreeding coefficients estimation in populations affected by migration and admixture of founders. Alternatively, internal relatedness (IR) has been proposed for populations with high inbreeding (Aparicio, Ortego and Cordero, 2006).

For microsatellite markers, before embarking on the DNA polymorphism analyses, it is suggested to test the markers for neutrality so that an unbiased interpretation of downstream analyses can be made (Chen, Cole and Grond-Ginsbach, 2017). Linkage disequilibrium is the tendency of alleles at linked loci to be inherited together. Ideally, genetic data obtained from a set of microsatellite markers should be in a random association. The level of linkage disequilibrium between a set of microsatellite markers can indicate an excess of either homozygotes or heterozygotes. This non-random association of microsatellite markers can be caused by selection, genetic recombination, mutation rate, genetic drift, mating system, population structure, and genetic linkage. The coefficient of linkage disequilibrium, *D*, can be used for this measure and its statistical significance tested using the Fisher's exact test and several resampling tests (Slatkin, 1994).

2.3.1.2 Fixation Indices and Genetic Differentiation

Mutations or variation between populations arise usually due to isolation by distance or adaption which decreases gene flow. The effect of this is more prominent in smaller populations as genetic drift may result in disappearance of a particular variant of gene which could have an adverse effect on its evolutionary potential. On the other hand, the purging of deleterious alleles can also occur (Kristensen *et al*., 2015). Meanwhile, a high gene flow between two populations may result in equivalent allelic frequencies, rendering them a single, effective population. Therefore, the level of gene flow has an important effect on the population, hence the importance of identifying population for better

management action (Balloux and Lugon-Moulin, 2002; Alif and Khan, 2009; Willi *et al*., 2022).

There are several ways of documenting population structure. The F_{IS}, known as the inbreeding coefficient, is a measure of gene correlation within individuals in a particular subpopulation, to check for random mating, which would indicate if samples came from one or several subpopulations. Analysing this can be done by pooling of samples to identify random mating which would not show a significant difference in the F_{IS} while incorporation of the other population(s) would increase the F_{IS} (Balloux and Lugon-Moulin, 2002). Software such as Genepop and FSTAT offer these analyses.

The F_{IS} measure was later expanded to hierarchical F-statistics to document genetic differentiation into a set of subpopulations comprising F_{IS}, F_{ST} and F_{IT}. The measure of interest is usually the F_{ST} which is the correlation of two alleles chosen randomly within subpopulation relative to the alleles at the total population level. Similar to the inbreeding coefficient (F_{IS}), the F_{ST} is a measure at the subpopulation level. When each subpopulation is identical, the F_{ST} value would be 0. As two subpopulations become increasing dissimilar to one another, the F_{ST} measure increases until 1, ultimately forming two separate subpopulation that are homozygous. Therefore, the F_{ST} is also known as the fixation index. The latter scenario of increasing homozygosity (or heterozygote deficiency) due to population structure is also known as the Wahlund effect (Balloux and Lugon-Moulin, 2002).

Population differentiation is affected by temporal and sex biases. When sampling a long-lived organism, often overlapping generations are pooled and treated as a subpopulation. This scenario is seen in crocodilian as reported by Muniz *et al*. (2019) and Van Asch *et al*. (2019). To overcome this, Kim and Sappington (2013) suggest sampling the same subpopulation over time and over generations. Similarly, for a sex-specific dispersal rate, sampling of only adults to detect for an overlap in the confidence levels would render the population as not showing any sex biased dispersal (Kim and Sappington, 2013). Incorporation of the SMM mutational pattern of microsatellite into the traditional F-statistic, known as the R-statistic, are usually reported together due to the ambiguity of the mutation mechanism of microsatellites (Balloux and Lugon-Moulin, 2002).

2.3.1.3 Population Structure and Differentiation

The program STRUCTURE is a Bayesian clustering method to infer the number of distinct populations (K) from which samples have been drawn and to infer the genetic ancestry of the individuals sampled, based on microsatellite genotypes at multiple loci (Kim and Sappington, 2013). The program accounts for the presence of Hardy–Weinberg or linkage disequilibrium (inclusion of weakly linked markers with some degree of non-independence (Falush *et al*., 2003) by introducing population structure and attempts to find population groupings that (as far as possible) are not in disequilibrium (Pritchard *et al*. 2000).

Population differentiation can also be investigated using Analysis of Molecular Variance (AMOVA) which partitions the genetic variation hierarchically. It estimates the proportion of genetic diversity within and between populations, or

among groups. It is calculated based on Euclidean distances between individuals in GenAlEx (Smouse, Banks and Peakall, 2017) and Arlequin (Excoffier and Lischer, 2010).

2.3.1.4 Parentage and Kinship

Parentage analyses is also popular with microsatellite markers because it is a co-dominant marker with high mutational properties resulting in multiple alleles at a locus even within a generation. This analysis provided remarkable insight to the mating behaviour and have identified a polyandrous mating system across crocodilians (Davis *et al*., 2001; Zucoloto *et al*., 2002; Lance *et al*., 2009; Zucoloto *et al*., 2009; Amavet *et al*., 2012;).

Parentage analyses can be performed in software like Gerud, which allows reconstruction of parental genotypes even when no parents are known, provided that the progeny array contains only full and half siblings (Jones and Ardren, 2003). Likelihood methods have also been designed in Cervus and ML Relate software which investigates alternative relationships that might exist between individuals, with both accommodating for genotyping errors and null alleles (Slate, Marshall and Pemberton 2000, Kalinowski *et al*., 2006).

Pedigree likelihood approach, which considers the likelihood of the entire pedigree structure and allows the simultaneous inference of parentage and sibship, is another way to carry out parentage analyses. This approach is applied in the software Colony (Jones and Wang, 2010). Coancestry coefficient approach in the Coancestry Software on the other hand estimate relatedness and

inbreeding coefficients from multilocus genotype data based also on likelihood method. It also allows for inbred individuals and account for genotyping errors (Wang, 2011).

In MolKin Software, the proportion of shared alleles between individuals is used for coancestry coefficients, r. By applying Malécot's (1948) definition of kinship to genetic markers, referred as identity-by-state instead of identity-by-descent, the molecular coancestry can be determined (Caballero and Toro 2002). MolKin has the bootstrapping procedure implemented which involves creating new datasets by randomly sampling individuals with replacement, so that the resulting datasets have the same size as the original, shown statistically to be typical of the variation that would be seen from collecting new datasets (Caballero and Toro 2002).

A similar software, Kin, is also based on Malecot's r. Population fission causes the average kinship between populations to become constant very quickly, causing between-population diversity to remain constant. This enables molecular coancestry information to be studied for the genetic relationships between populations, apart from accessing genetic diversity within and between populations (Tinker and Mather, 1993).

The molecular coancestry approach was proven to be more useful in recovering native genetic material in the Tambaqui fish farm in Brazil (Goncalves *et al*., 2019). An extinction risk assessment of native Italian chicken breeds found that at a high internal diversity (GD_W) should be preserved to maintain a high

heterozygous level as a short-term strategy while allelic richness and breed differentiation as a long-term strategy (Soglia *et al*., 2021). Allelic richness is important for a long-term evolutionary adaptivity (Petit, Mousadik and Pons., 1998) and both the gene and allelic diversities should be used to complement each other for conservation purposes (Alvarez *et al*., 2010).

2.3.1.5 Divergence Time Estimation

For the microsatellite data, the equation $(d\mu)^2 = 2w\tau$, where $d\mu$ is the genetic distance, w being the effective mutation rate while τ is the generations since divergence, is an estimator of divergence time since a population bifurcated. The assumption is that populations are at mutation-drift equilibrium and have constant effective size (Zhivotovsky, Goldstein and Feldman, 2001; Zhivotovsky, 2001). Newer estimates have emerged namely the T_D, involving microsatellites which considers mutation-drift disequilibrium and gene flow. Because this method needs the knowledge on the vicariance, a population that may be demonstrating a variation similar to that of an ancestral population can be detected by setting upper and lower bounds to such a vicariance event (Zhivotovsky, 2001). Using SNP, Avila-Cervantes *et al*. (2021) were able to identify divergence time of the American crocodile, *Crocodylus acutus*, by testing the demography models to the vicariance and historical colonization events.

2.3.1.6 Crocodilians Population Studies

In the Morelet's crocodile, *Crocodylus moreletti*, a high expected heterozygosity levels, H_E, of 0.49 did not show any genetic differentiation by the F and R

statistics (Dever *et al*., 2002). However, the STRUCTURE analysis showed population structure following the river systems sampled when re-analysed in such a configuration. Together with the Neighbour Joining tree, two locations, the New River and the New River lagoon, were identified as source population displaying a "mainland" while the other localities acted as "island" in the island model of gene flow. A similar scenario was seen in Nile crocodile, *Crocodylus niloticus*, across Africa and Madagascar, where STRUCTURE found population corresponding to seven river drainages while the distance based AMOVA and F and R statistics found three subpopulations (Hekkala *et al*., 2010). Natal philopatry and ecological constraint were suggested as possible factors contributing to the demographically fragmented population in Madagascar and across Africa, which also corresponded to evolutionary lineages that are distinct, similar to an earlier mtDNA study of (Hekkala, 2004).

In contrast, the distance and Bayesian based analyses were congruent for the population structure of the black caiman, *Melanochushus niger*. The allele variation indicated a significant population structure, supported by STRUCTURE and the Nei's genetic distance tree that populations were structured according to the ecologically different habitat probably due to selection (de Thoisy *et al*., 2006). This trend was also noted in the opportunistic species, the spectacled caiman, *Caiman crocodilus*, co-existing with *M. niger*. By comparing spatial distribution of the microsatellite alleles with ecological data, the geographic expansion of this species was hypothesised to be from the central Amazon region to the western (Ecuador) and to the eastern (French Guiana) populations because oldest alleles and highest genetic diversity is

expected to be at the centre of species distribution (de Thoisy *et al*., 2006). Another example of congruence across the analyses was in the case of the smooth-fronted caiman, *Paleosuchus trigonatus*. The low levels of H_E, the weak but significant F_{ST} (0.049) in AMOVA analysis and two cluster in the STRUCTURE analysis using locprior function showed a shallow population structure of this species in Brazil (Bittencourt *et al*., 2019). These results were also congruent with the mtDNA findings of Hekkala (2004).

Using an approach at the allelic level, Bashyal *et al*. (2014) were able to differentiate the two populations of the American crocodile, *Crocodylus acutus* from Panama and Mexico, and of *C. moreletii* from Belize and Mexico by the difference in their allelic richness and number of unique alleles. However, as both the allelic composition parameters are affected by sampling sizes, a larger sampling size was suggested for future studies. As *C. acutus* has a wide region of occurrence, microsatellites were used to identify if the population could be structured along the Pacific coast (Mauger *et al*., 2017). This study found that individual populations were diverse as 81% of molecular variance was within subpopulations suggesting that *C. acutus* in Pacific Costa Rica are not panmictic. However, the inbreeding coefficients (F_{IS}) for the five populations ranged between 0.096-0.179 which was due to heterozygote deficiency caused by sampling mainly juveniles in localities with small population.

The findings above were able to impact the conservation of the respective crocodilians. For instance, the black water populations of *M. niger* in Brazil and French Guiana should be treated as a single unit due to the gene flow between

these two populations (de Thoisy *et al*., 2006) while the findings of *Crocodylus acutus* indicated that the individual subpopulations should be treated separately (Mauger *et al*., 2017).

2.3.1.7 Crocodilians Paternity Studies

Paternity studies across crocodilian were not only able to identify parents and offspring, but also suggest mating systems and behaviour which are difficult to document by ecological and behavioural studies due to their amphibious life style. Multiple paternity was seen across crocodilians and most researchers have suggested this to be a means of increasing effective population sizes naturally and reduce inbreeding and could be used in the management and conservation of crocodilians (Lance *et al*., 2009; Amavet *et al*., 2012; Oliveira *et al*., 2014; Ojeda *et al*., 2017).

Using nest samples, Davis *et al*. (2001) detected multiple paternity in American alligators, *Alligator mississippiensis*. In two cases, fidelity towards males was observed. Primary males showed a significant loss of paternity in clutches involving three sires compared to clutches with only two sires. However, secondary males maintained their paternity despite the presence of a third sire. Lance *et al*. (2009) conducted a four-year study, which revealed that males mated with multiple females within a single breeding season as well. Studies on paternity revealed fidelity to nesting sites in the spectacled caiman *Caiman crocodilus* as reported by de Oliveira *et al*. (2010). In the captive broad-snouted caiman, *Caiman latirostris*, while females displaying maternal nest-guarding behaviour were confirmed to be mothers, two other females exhibiting similar

behaviour were not maternal parents. This discrepancy might be attributed to factors such as captivity, social adaptation, or potential behavioural anomalies (Zucoloto *et al*., 2009).

Apart from Alligatoridae, multiple paternity was also detected in *Crocodylus* (McVay *et al*., 2008; Lewis *et al*., 2013; Budd *et al*., 2015). In *Crocodylus porosus*, multiple paternity was attributed not only to females mating with multiple partners but also to a promiscuous mating system. Evidence of males engaging in multiple matings was detected when two wild-bred clutches from nearby nests were found to share the same paternal parent. Interestingly, despite the higher frequency of mate encounters in denser communal dens, the occurrence of multiple paternity was lower, possibly due to increased aggression. A higher hatching success was not detected unlike previous studies (Lewis *et al*., 2013).

2.3.1.8 *Tomistoma schlegelii* Studies

Population studies on this species have been carried out using 17 *T. schlegelii* samples from Peninsular Malaysia and Sarawak (Shafie-Astani *et al*., 2015). In addition to the 40 microsatellite markers specific for crocodiles (FitzSimmons, 2001; Chaeychomsri *et al*., 2008) and the American alligator (Glenn *et al*. 1998), 45 Inter Simple Sequence Repeat (ISSR) markers were also incorporated. Both datasets were able to differentiate populations of *T. schlegelii*. For the microsatellites, the observed and expected heterozygosity were between 0.588-1 and 0.470-0.891, respectively (Shafiei-Astani *et al*., 2015).

Species specific microsatellite markers were developed from the preserved liver samples of a *Tomistoma* using the hybridization capture method (Chang *et al*., 2014). From the randomly picked colonies, 68 out of 111 had microsatellite repeats. After tagging the primer pairs and performing PCR, 18 pairs were identified as being polymorphic when 19 individuals were screened. The number of alleles ranged from 2 to 16 and observed heterozygosity from 0.105 to 0.684. The probability for excluding false parental pairs was 99.9999 %. These species-specific markers have not been tested in populations with known or likely geographical regions because the 19 individuals tested in Chang *et al*. (2014) were of unknown origins.

2.3.2 Mitochondrial DNA (mtDNA)

The mitochondrial DNA (mtDNA) genome is a widely used genetic marker, favoured for its phylogenetically informative traits such as maternal inheritance, high abundance, ease of isolation, lack of recombination, small effective population size, and generally high mutation rate (Zouros *et al*., 1994; Burzynski *et al*., 2003). Being haploid, the coalescence of neutral mtDNA genes (tracing back to a common ancestor) is closely linked to the effective population size, making mtDNA particularly useful for demographic studies (Avise *et al*., 1987; DeSalle, Schierwater, and Hadrys, 2017). Recently, with advancements in next-generation sequencing (NGS), mtDNA has been proposed as an environmental DNA (eDNA) marker. Its high copy number increases the likelihood of detection in environmental samples (DeSalle, Schierwater, and Hadrys, 2017).

Based on the complete mtDNA genome sequencing of the saltwater crocodile, *Crocodylus porosus*, the crocodilian mitochondria DNA (mtDNA) comprised of two ribosomal RNAs, 22 transfer RNAs, about 13 genes and a control region. Although the gene order is conserved among crocodilians, the arrangements of tRNAs genes differs from other vertebrates (Li *et al*., 2007). The $tRNA^{Phe}$ is placed between the control region and the 12S rDNA while the tRNAs arrangements after NADH4 is $tRNA^{Ser}$ followed by $tRNA^{His}$, followed by the common arrangement of $tRNA^{Leu}$ and NADH5 (Li *et al*., 2007; Boore, 1999).

According to Ray and Densmore (2002), 57% of variability within crocodilian was detected in the mtDNA control region. Additionally, the control region also consists of mtDNA variable number tandem repeats (VNTRs) believed to be responsible in the replication of the mtDNA and the transcription of RNA (Taanman, 1999; Avise *et al*., 2004; Bogenhagen, 2012; DeSalle, Schierwater and Hadrys, 2017). In *Tomistoma schlegelii*, the PCR amplification of the region between the conserved sequences box (CSB) III and the 12S rDNA revealed band sizes of up to 3.1 kb consisting multiple VNTRs patterns (Kaur and Ong, 2011). Unique sequence repeats the mtDNA VNTRs were detected in *Alligator* with a chimeric aspect of the 5' and 3' in its repeat sequences, while the *Crocodylus* repeat motifs ('TAGG' motif) (Ray and Densmore, 2003).

2.3.2.1 mtDNA Markers in Crocodilians Studies

Various mtDNA genes have been utilised for population level studies of crocodilians. Meganathan *et al*. (2013) analysed the cytochrome oxidase subunit 1 (COI) gene for forensic identification of the Indian mugger, *Crocodylus*

palustris, *C. porosus*, and the Indian gharial, *Gavialis*, which are threatened by illegal poaching. This marker was found to be discriminatory enough for future use in enforcement efforts. This is because the 600 bp of the COI gene is known to be sensitive enough for all taxa identification. As a short, conserved DNA sequence that exhibits high interspecific variation, it is thus a DNA bar-coding tool for species delimitation (Hwang and Kim, 1999; Rubinoff and Holland, 2005). In *Osteolaemus*, together with nuclear lactate dehydrogenase A, these genes were able to identify species of the dwarf crocodile, the slender snouted crocodile, *Mecistops*, *Crocodylus* and hybrids (Shirley *et al*., 2015). Similarly, using both the COI and cyt *b* gene, Balaguera-Reina *et al*. (2020) were able to identify the geographical origins of two unidentified new *C. acutus* individuals in San Andres Island, Colombia.

The D-loop is another marker used in crocodilian study for population and species level studies. At the population level, the D-loop was also able to show that there was no clear genetic structure for *C. porosus* in the Northern Territory although three haplotypes seemed restricted to northern Australia individuals (Luck *et al*., 2012). Separately, D loop with the combination of ND4 and the nuclear C-*mos* detected hybridization in *C. porosus* and *C. mindorensis* (Tabora *et al*., 2012).

The partial ND 6-tRNA^{glu}-cyt *b* gene is another marker used in crocodilians. Ray *et al*. (2000) found high variations in *Osteolaemus* while McAliley *et al*. (2006) found low levels of diversity in *Crocodylus*. This gene fragment was also used in the *T. schlegelii* study and although the nucleotide variation was higher in the

D-loop, this fragment was able to detect a haplotype (H3) which was conserved at the D-loop (Kaur *et al*., 2013).

2.3.2.2 mtDNA studies on *T. schlegelii*

Kurniati (2003) used the D-loop to access the genetic diversity and structure of the Kalimantan population using eleven wild caught individuals. Over the 451 bp gene, four haplotypes were detected namely, A, B, C, D. The haplotype network separated the four haplotypes to two regions, the central and east Kalimantan, and the western Kalimantan, with two haplotypes in each divide. The A haplotype was found in eastern Kalimantan while the central Kalimantan had both the A and B haplotypes, with the latter detected only in the Mapan River. Haplotypes C and D were found in western Kalimantan. The haplotype and nucleotide diversity were high with the overall haplotype diversity, h = 0.5714 and nucleotide diversity, π = 0.2326. The genetic divergence within the two geographical region was similar at 0.2% but between regions was 1.53%. The between region divergence was equated to 20 million years ago (MYA) based on the time of the Schwaner Mountains barrier formation.

On a larger geographical region of occurrence covering Sumatra, Peninsular Malaysia, Sarawak and Kalimantan, six haplotypes were found, H1, H2, H3, H4, H5, H6 (Kaur *et al*., 2013). This study was carried out using the partial ND 6-$tRNA^{glu}$-cytochrome (cyt *b*) gene and the cyt *b*-control region (CR) (Fig. 2.5).

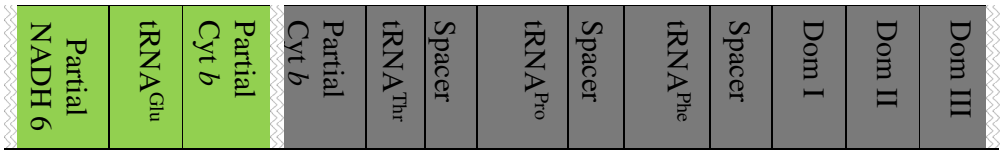


Fig. 2.5. Schematic diagram showing the two fragments of mtDNA genes used by Kaur *et al*., 2013. Green ND6-cyt *b* fragment. Gray: cyt *b*-CR fragment.

Peninsular Malaysia had the highest *h* and π consisting of four haplotypes the H1, H2, H4 and H5. Haplotype H3 was identified in samples originating from Sumatra, while haplotypes H6 was seen in east Kalimantan. The remaining haplotypes were all found in Peninsular Malaysia. Only haplotype H2 was seen in more than one region, in Peninsular Malaysia and Sarawak. The divergent haplotype, H4, was not parsimoniously connected at the 95% connection limit in TCS and was placed as a basal and reciprocally monophyletic clade in the phylogenetic trees. Additionally, the genetic distance between the H4 haplotype with the other haplotypes was 2.4%. Although this measure is well below the interspecific distance (15% genetic distance between *Gavialis* and *Tomistoma*), it was comparable with the genetic distance of hybrids in *Crocodylus* (1.6-4.1%) (Kaur *et al*., 2013).

Due to the vast geographical coverage of this study, the sample with uncertain origins in the data set were able to be assigned to the geographical areas (Fig. 2.6). Subsequently using similar mtDNA region above (cyt *b*-CR), Md Adzhar *et al*. (2017) reported to have found another haplotype in Sarawak. So far, the mtDNA study on this species have found geographical and ancient river systems (Pleistocene) connection with haplotype H2 (Sarawak haplotype) occurrence in both Sarawak and Peninsular Malaysia (Kaur *et al*., 2013; Kurniati, 2003).

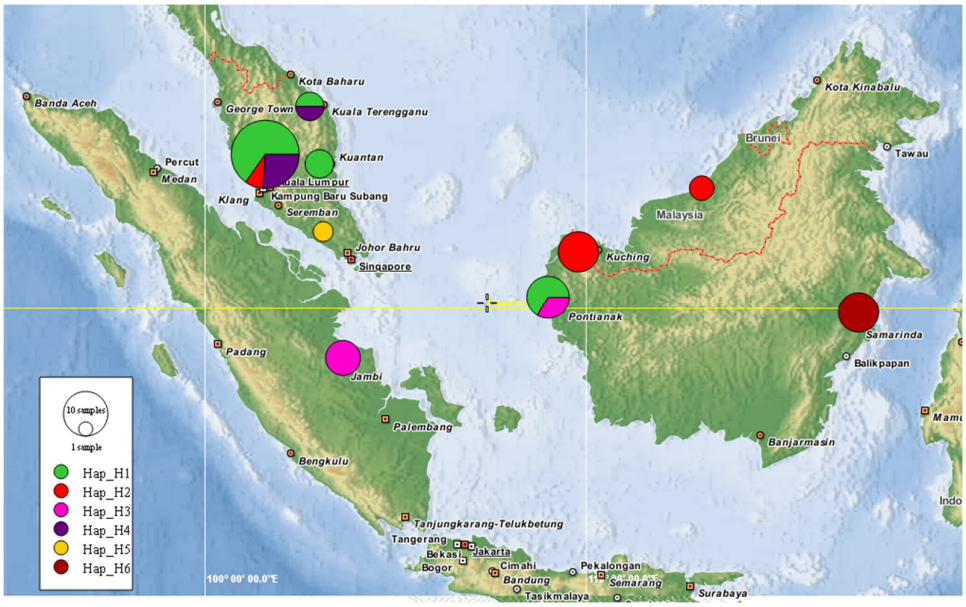


Fig. 2.6: Haplotype distribution of *Tomistoma schlegelii* from the study of Kaur *et al*. (2013). Legend indicates colour of the haplotypes H1, H2, H3, H4, H5 and H6.

2.3.2.3 Divergence Time

Because mtDNA is maternally inherited, the ancestral mtDNA sequence could be inferred from the surviving populations or species by considering the tips of the taxonomic tree as a basic operational taxonomic unit (OTU). An OTU on a gene tree is similar fragment(s) of gene(s) while a population or species is considered an OTU on the population tree. Alternatively, correctional statistic could be applied to the mean pairwise mtDNA divergence between two gene/ populations/ species. However, the mtDNA variability must be neutral in estimation of molecular clocks and genetic distance estimates which are usually found in the non-coding region of the mtDNA (D-loop) or base substitution at silent protein-coding regions (Avise *et al*., 2004; 1989).

At the intraspecific level, lineages may not evolve independently from the last common node due to natural selection, genetic drift and population sizes. Therefore, usually for intraspecific data the formula used for a molecular clock is $d_A = 2\lambda t$ by Nei (1987); where λ is the substitution rate per nucleotide site per year, d_A being the mean distance between individuals, populations or lineages within groups and t as the divergence time. Using the whole mtDNA genome, the divergence time estimation has been carried out for Crocodylia, analysed with a variety of software programs (Janke *et al*, 2005). Different evolutionary rates were taken into consideration as it is reasonable to assume that the rate of evolution of two lineages change independently following their last common ancestral node. This study estimated the divergence times of between *Tomistoma* / *Gavialis* and *Crocodylus* at approximately 80 million years before present (MYBP) while the split between *Gavialis* and *Tomistoma* approximately 42 MYBP. However, a subsequent study by Roos, Aggarwal and Janke (2007) reported a divergence time of approximately 22-28 MYBP for *Gavialis* and *Tomistoma*.

2.4 Gender-Biased Dispersal and Philopatry

As mtDNA is maternally inherited, it has also been used to infer patterns of differential migration between sexes. A more homogenised population is seen in female-biased dispersal as the migrating female haplotype is passed onto the population. Alternatively, if population structure is analysed using the maternal AMOVA or F-Statistic, and a specific pattern is observed, a female philopatry is concluded (Escorza - Treviñ and Dizon, 2000). In loggerhead sea turtles, low nucleotide diversity observed in the D-loop at a beach near a dense nesting site

of Cabo Verde archipelago was attributed to female philopatry, as evidenced by the weak yet significant genetic differentiation among the beaches (Baltazar-Soares *et al*., 2020). In green turtles, both female and male philopatry were identified when a significantly high F_{ST} was observed for both sexes, accompanied by non-significant differences in their mtDNA haplotype frequencies. Previously, it was believed that only female philopatry existed in this species, based on higher gene flow and non-significant population structure estimates derived from fragment restriction and microsatellite analyses, as opposed to mtDNA analyses. (Fitz Simmons *et al*., 1997).

Meanwhile, in frill neck lizards, a juvenile male-biased dispersal was detected using the mtDNA ND4 and ND2 genes, amplified fragment length polymorphism (AFLP) and mark-recapture data. The presence of mixed nuclear genotypes in juvenile males, alongside the distinct mitochondrial haplotypes, suggested male-biased dispersal patterns (Ujvari, Dowton and Madsen, 2008). Similarly, a study using a 10 years mark-recapture data with genetic analyses comprising nine microsatellite markers detected male-biased dispersal in the slateygrey snakes. The low genetic differentiation detected with the F_{ST} and a low relatedness coefficient (r) in males and a stronger genetic structuring in females using the spatial autocorrelation analyses supported the mark-recapture data (Dubey *et al*., 2008).

Therefore, it is important to use a combination of genetic markers as seen in the studies of green turtle where mtDNA markers were able to also suggest a male and female philopatry (Fitz Simmons *et al*., 1997). Similarly, the field

observation using the mark-recapture techniques and genetic markers were also powerful in elucidating movement of sexes and at different life stages (Dubey *et al*., 2008; Ujvari, Dowton and Madsen, 2008). The knowledge on sex-biased dispersal or philopatry helps in the conservation efforts by understanding the population dynamics, genetic structure, and social systems in various species (Williams and Rabenold, 2005)

2.5 Field Surveys in Crocodilians

Crocodilian surveys, namely night and day surveys, are carried out to detect and monitor populations of crocodilian and relative densities (Cox and Gombek, 1985; Webb and Manolis, 2006). For an effective conservation plan, it is essential to gather comprehensive information about their needs and threats, which can be derived from ecological studies and population trends. These are carried out using direct (night spotlight counts, day counts and aerial counts for individuals) and indirect methods (footprints, basking prints, faecal droppings and nest counts). The detection of any significant unplanned population declines with ongoing monitoring program would enable management aims and objectives to change over time as new information becomes available (Webb and Manolis, 2006).

2.5.1 Night Spotlight Surveys

The night spotlight survey is the most widely used methods for crocodilian detection (Pacheco, 1994). Most night surveys employ the modified night spotlight count technique adopted from Messel *et al*. (1981). This technique is used to detect crocodilians because the tapetum of the crocodilian will reflect

light when torched in the dark and individuals can be seen as a 'red eyeshine', known as eyes only (EO), from a 0.5 km distance (Kar and Bustard, 1990). The team usually consists of three people (observer, recorder and boatman). An observer shines the torch from one side of the bank to the other at a 45° angle and spots the eyeshine. The recorder notes down all sightings and related observation to those sightings such as location and animal characteristics. The boat driver navigates to maintain the boat at middle of the river throughout the survey. The boat is usually a small (about 15 feet) flat bottom boat with a 15-45 horsepower outboard engine (Ross, 1998). The start and end points of each survey are important for future replication (Thorbjarnarson, Platt and Khaing, 2000).

Although it has the advantage of detecting crocodilian easily, this technique is generally used for monitoring trends of populations over time rather than the number of total populations in every year or season (Webb and Manolis, 2006) and for estimation of total population using a mark recapture technique by the Petersen estimate (Moore, 1994). This is because the night survey method is affected by methodological (types of boat, torches and different observers), environmental (cloud cover, water level, seasons, wind speed, wave height, floating vegetation, water and ambient temperatures and types of habitat) and biological (different age group and species, wariness of the individual or population and the densities of the population) factors (Bayliss *et al*., 1986; Pacheco, 1996).

Therefore, this method of survey is conducted using strict and standardised survey procedures with respect to time, location, meteorological conditions, habitat conditions, observer bias, counting procedure and equipment used (Moore, 1994). In addition to that, a standardised body size classification is maintained over the surveys for a more precise population structure inferences over time (Pacheco, 1994, Moore, 1994). The night survey technique has been reported to be difficult for detecting *Tomistoma* due to the inaccessible habitat in addition to the secretive nature of the species (Stanwieswz *et al*., 2018).

2.5.2 Day Surveys

The day survey methodology is similar to night spotlight surveys except it is carried out during the day, usually in areas known to be basking sites of crocodilians. Therefore, it is used for both direct and indirect counts, in addition to note ecological factors involved with sightings during the night surveys. In the case of *T. schlegelii*, this approach was able to describe the type of habitat and river characteristics which is peat swamps with acidic water (Bejuizen *et al*., 1998; Auliya et. al, 2006; Stanwieswz, 2011). A telemetry study for home range of *T. schlegelii* was carried out in Central Kalimantan which suggested possible homing instinct in this species. It is the first successful transmitter attachment and radio tracking of *T. schlegelii*. It was revealed that home range could be larger in the wet seasons when the peat swamp is inundated as opposed to ranges seen during the dry season where its habitat ranged between 0.103-0.34 ha (Bonke *et al*., 2014).

When surveying *Crocodylus porosus* and *C. novaeguineae*, Montague (1983) found day survey counts to be 12.9 times fewer than night spotlight counts and that the method was a more realistic measure for larger crocodiles. This is because 73% of slides detected in the day surveys were from animals measuring 10 inches in width in contrast to 23.1% of the same size detected in the night. Factors that affect the day surveys are similar to the night spotlight surveys namely methodological (types of boat and observers), environmental (day length, availability of sunlight, tide, floating vegetation, seasons, water and ambient temperatures, river width and sinuosity and types of habitat) and biological (different age group and species) factors (Baylis *et al*., 1986; Thorbjarnarson, Platt and Khaing, 2000). The probability of sighting decreases with increasing density of bank vegetation, the decreasing river width and increase river sinuosity (Bayliss *et al*., 1986).

2.5.3 Distribution of *T. schlegelii*

2.5.3.1 Malaysia

Webb and Manolis (2006) in their report to CITES Secretariat, stated that in areas with little survey data on population and where it is believed that only a small proportion of the population can be sighted, local knowledge about real abundance is crucial. It was further suggested to combine local knowledge with topographical maps and aerial photographs to gain insights into historical and current distribution of *T. schlegelii* (Webb and Manolis, 2006). This has been practiced in Malaysia when a compilation based mostly on interviews, museum records, captive records and personal communications with state wildlife

officers have also been documented (Sebastian, 1993; Stuebing *et al*., 2004; 2006).

During a preliminary survey in preparation for a conservation and management policy of crocodile resources in Sarawak, Cox and Gombek (1985) spotted three individuals of this species at the Ensengai Baki River although no other sightings were recorded in other localities. The first field survey in Peninsular Malaysia was carried out for two weeks at Tasek Bera with no crocodiles spotted (Simpson *et al*., 1998). However, this report did highlight that the habitat in and around the lake area (primarily peat swamp) was intact for the species (Fig. 2.3). Another survey was carried out in the peat swamp forest of Selangor and Pahang with no positive sighting of any crocodiles (Simpson, 2014).

A recent study in western Sarawak after an adult was caught in the Samarahan River while two hatchlings were caught in the Bunga-Baki River, Serian in early 2015 (Hassan *et al*., 2016). The Kepayang River, Ulu Sebuyau, was also included as a carcass was found here and sighting of live individuals were reported by locals. The day surveys at the end of the wet season found the river pH ranged from 4.2-6.7. The Samarahan and Serian study sites were close to peat swamps and the Ulu Sebuyau site had brown to black stagnant water, which is characteristic of peat swamp localities. The temperature ranged from 26.67-27.77°C. During the survey, no sightings of *Tomistoma* was detected and Hassan *et al*. (2016) attributed this to the secretive and wary nature of this species, methodological reasons such as surveys done only during day light and the short

duration of both boat and foot surveys, and environmental factor such as inaccessible habitats.

2.5.3.2 Indonesia

Information based on repeated surveys and spotting of *T. schlegelii* individuals in Sumatra found that this species was primarily in peat swamps with pH ranging between 3.5 to 4.0 (Bezuijen *et al*., 1998; 2002; 2004). Information gathered during these surveys were nest characteristics and breeding periods between June to September and predation of nests by pigs (Bezuijen *et al*., 1998; 2001). Additionally, mortality of *T. schlegelii* individuals due to drowning by fishing equipment was also noted (Bezuijen *et al*., 2001). Shaney *et al*. (2016) resurveyed areas in Bezuijen (2001) study in Sumatra in addition to Riau Province and found a declining population of this species from 0.02 croc / km over the years (Bezuijen *et al*., 2002; 2004) to none in the lower reaches of the Merang River. Similarly, with Simpang Melaka Creek, densities reduced from 0.03 croc / km to no sightings in 2016. However, an increase was seen at the Air Hitam Laut River and Shaney *et al*. (2016) contributed these findings to an intact habitat. Subsequently, one confirmed sighting was detected upstream of Sembilang National Park (TNS) Area (Benu River) in South Sumatra Province using both direct and indirect surveys. Five other sightings of eyes only (EO) were suspected to be *T. schlegelii* (Setiawan *et al*., 2019).

Similarly, a study in Central Kalimantan on relative densities, movement patterns and home ranges also found that peat swamps and acidic waters (pH4) were the preferred habitat of *T. schlegelii* (Fig. 2.3). The population here showed

a higher density due to the ongoing habitat protection and enforcement measurements taken for the Orangutans management at Tanjung Putting (Auliya *et al*., 2006). In east Kalimantan, the mark-recapture study detected movement of juveniles during rainy season from August to October in Lake Mesangat (Stanwieswz, 2011; Stanwieswz *et al*., 2018). Additionally, both the *Tomistoma* and *C. siamensis* shared the habitat although the latter was found more at the southern stretches of the lake. Stuebing, Sommerlad and Staniewicz (2015) found that after capturing and tagging the juveniles, recapture was difficult due to the highly secretive behaviour of this species in combination with the preference of a habitat which is usually inaccessible.

2.6 Geographic Information System (GIS)

Geographic Information System (GIS) is a computerized system for capturing, storing, verifying, and displaying data on a map allowing for easier / clearer visualisation and analyses of patterns and relationships. It comprises spatial distribution of the studied subject (usually in vector or raster layers) and databases with a geographical component to store the attributes of the subject (Rosca *et al*., 2020). Vector data is usually point data (coordinates), lines (transect lines), and polygons (habitat ranges). It has additional data attributes consisting of information like the number of individuals sampled at a location, the name of a river displayed as a line, or the type of habitat represented by a polygon. Raster meanwhile are continuous matrices of grid cells, with each cell containing a single value summarizing the landscape feature it represents (Bohm and Popescu, 2016).

Terrain analysis is the classification of the surface of the Earth by qualitative and quantitative description. It is carried out by analysing the Digital Elevation Models (DEMs) in a GIS software. It gives the ability to quantitatively characterise the land surface in terms of slope gradient and curvature without disruption of land cover features. It is mainly used in hydrological modelling and analysis of watersheds and drainage networks, landslide hazard assessment and ecological studies. The assumption using the DEMs is that the landform position and shape influences the distribution of soils and soil properties in the landscape (MacMillan, Moon and Coupé, 2007; Schillaci, Braun and Kropáček 2015; Xiong, Tang and Strobl, 2022). The steps involved in analysing the DEMs usually consists of the overlay analysis, modelling, buffering and network analysis (Congalton and Green, 1992).

GIS and spatial analyses for studying ecology and conservation of wildlife is becoming popular as it can help delineate habitat patches, evaluate their size, shape, and connectivity, therefore aiding conservation efforts. As the ability of animals to migrate and disperse is affected by characteristics of the landscape across space, connectivity measures specific to the species of interest is necessary as habitat suitability could vary with developmental stage. The physical and structural connectivity analyses can be carried out using Euclidean (or straight-line) distance between patches while the leastcost path analysis calculates a cost surface based on habitat qualities that impede or facilitate movement of a species (Bohm and Popescu, 2016).

The field of phylogeography has predominantly focused on the genetic component while the geographical component has often been superficially explored, typically as a secondary or qualitative comparison. Incorporation of GIS-based approaches to the genetic data helps to reduce the inaccuracy which arise from the stochasticity of gene-lineage coalescence. Various shortcomings based on assumptions like spatial homogeneity can be addressed using algorithms in GIS to include barriers and human activities (Kidd and Ritchie, 2006). Richards, Carstens and Lacey Knowles (2007) illustrated in detail the methodology of combining the distribution and coalescent-based modelling techniques to give a powerful means of detecting and inferring biogeographical patterns and causes. Although the species distribution is based in the present, the inferences can be traced to the past using paleo distributions and projected to the future (Richards, Carstens and Lacey Knowles, 2007).

2.6.1 GIS in Crocodilian Studies

GIS has been incorporated across crocodilian studies to get a better understanding of the historical and current crocodilian populations, trends and habitat status. The historical distribution of *Crocodylus* in Java Island, mapped from reliefs seen in temples and fossils findings, was able to indicate the presence of *C. siamensis* when compared to its current habitat (Wibowo, 2020). GIS was used for the visualization of the distribution of telemetry studies on movements and spatial use of crocodilian worldwide (Mascarenhas-Junior *et al*., 2023). In Australia, the satellite tracking carried out on translocated large *C. porosus* males (52-126 km away from origin), mapped against river and coastline vector layers in ArcGIS found that all three males showed purposeful

homing travel after spending some time in the new location to perhaps interpret local cues and direction to travel back. This information is important as it proves that translocation of 'nuisance' individual from one place to another is futile (Read *et al*., 2007).

Identification of habitat and factors affecting the distribution of crocodilians has been carried out across the globe (Salem, 2013; Brackhane *et al*., 2018; Binaday *et al*., 2021). In the Philippines, using parameters important in the nesting ecology for *C. porosus* such as slope (as the most important) followed by temperature and precipitation, five habitats in Palawan Island were identified as intact for this species (Binaday *et al*., 2021). Similarly, at the Nasser Lake in Egypt, a habitat suitability map based on nesting preference records from field surveys was generated using GIS (Salem, 2013).

In Timor-Leste, GIS was used to develop a crocodile management plan for *C. porosus*. Areas of occurrence, current locations of crocodile attacks, satellite imagery and STRM DEM (from STRM 90m Digital Elevation Database v 4.1 and Geographic Information Group (GIG) Timor-Leste) were overlaid to identify areas of human-crocodile conflict. The map would be useful for wildlife managers to make informed conservation decisions (Brackhane *et al*., 2018). A similar approach for *C. niloticus* at the Tekeze River Dam in Ethopia was carried out. Overlaying the GPS coordinates obtained from field surveys and habitat preferences such as preferred riverbanks, shallow water depth and rocky ground along the main rivers and tributaries to the satellite imagery (Landsat TM5) was not only able to identify the type of land utilization at the dam but also identify

communal activities and floods as factors impacting this species and its habitat (Adugna *et al*., 2017)

2.6.2 GIS in *T. schlegelii* Study

Rodder *et al*. (2010) used 118 georeferenced records of *T. schlegelii* populations throughout its known range from Stuebing *et al*. (2006) and online databases (GBIF, www.gbif.org; HerpNet, www.herpnet.org) to investigate whether the established reserves within Southeast Asia effectively protected *T. schlegelii* using Species Distribution Models (SDMs) derived from climate niches. The assessment found the remaining habitats suitable for *T. schlegelii* were peat or freshwater swamp forest, followed by fragmented riverine forests and saltwater swamp forests. Major continuous habitats were found situated on the east coast of Sumatra and the south-western coast of Borneo of which only small parts are currently protected applying IUCN standards. The paper urged for more ecological studies to be carried out before the remaining peat swamps are destroyed.

CHAPTER 3

MATERIALS AND METHODS

3.1 Genetic Study

3.1.1 Sampling Sites

Blood samples used for this study were from ten new individuals sampled from the Sungai Dusun Wildlife Conservation Centre, Zoo Kemaman and Taman Buaya Langkawi. Of these ten samples (TK056-TK067), three are from confirmed locations. Additionally, 28 samples (TK002-TK030) from a previous study (Kaur *et al*, 2013) were retrieved from the Faculty of Medicine and Health Sciences, Universiti Tunku Abdul Rahman and incorporated in this study (Table 3.1). The locality of TK014 and TK016 is indistinguishable as either one could be from Perak or Terengganu.

3.1.2 Blood Collection

Each individual was physically restrained by securing the snout with a rope and immobilised by four to six men. The snout and limbs were secured to facilitate a lateral approach for caudal vein venepuncture (Fig 3.1). Blood collection site was cleaned with water, followed by 70 % alcohol. An 18 G spinal needle (SPINOCAN, Germany) and a 20 ml syringe containing approximately 1ml of 10 % EDTA were used to draw 10 ml of whole blood. Samples were immediately placed into a 15 ml Falcon tube containing additional 1-2 ml of 10 % EDTA. Pressure was applied at the site of blood collection to stop the bleeding. The crocodile was then released.

Table 3.1: The blood collection localities and geographical origins, and sex for all 38 Malaysian *Tomistoma* samples.

No.	Sample	Sampling location	Sampling Date	Origin	Sex
1	TK 002	Matang Wildlife Centre	Nov 13, 2006	Sadong, Sarawak	F
2	TK 003	Matang Wildlife Centre	Nov 14, 2006	Runjing, Sarawak	F
3.	TK 004	Matang Wildlife Centre	Nov 14, 2006	Sadong, Sarawak	M
4.	TK 005	Matang Wildlife Centre	Nov 14, 2006	Sadong, Sarawak	F
5.	TK 006	Zoo Negara	Dec 5, 2006	Selangor	F
6.	TK 007	Zoo Negara	Dec 5, 2006	Selangor	M
7.	TK 008	Zoo Negara	Dec 5, 2006	Selangor	F
8.	TK 009	Zoo Negara	Dec 5, 2006	Selangor	M
9.	TK 010	Zoo Melaka	Dec 18, 2006	Selangor	F
10.	TK 011	Zoo Melaka	Dec 18, 2006	Selangor	F
11.	TK 012	Zoo Melaka	Dec 18, 2006	Selangor	F
12.	TK 013	Zoo Melaka	Dec 18, 2006	Selangor	F
13.	TK 014	Zoo Taiping	Jan 25, 2007	#Perak/Terengganu	F
14.	TK 015	Zoo Taiping	Jan 25, 2007	Selangor	M
15.	TK 016	Zoo Taiping	Jan 25, 2007	#Perak/Terengganu	F
16.	TK 017	Mini Zoo Temerloh	Feb 23, 2007	Pahang	F
17.	TK 018	Mini Zoo Temerloh	Feb 23, 2007	Pahang	F
18.	TK 019	Mini Zoo Temerloh	Feb 23, 2007	Pahang	F
19.	TK 020	Mini Zoo Temerloh	Feb 23, 2007	Pahang	F
20.	TK 021	Miri Crocodile Farm	May 4, 2007	Kuching	F
21.	TK 022	Miri Crocodile Farm	May 4, 2007	Miri	F
22.	TK 023	Miri Crocodile Farm	May 4, 2007	Kuching	F
23.	TK 024	Zoo Negara	Nov 9, 2007	Zoo Bred (2004)	F
24.	TK 025	Zoo Negara	Nov 9, 2007	Zoo Bred (2004)	F
25.	TK 026	Zoo Negara	Nov 9, 2007	Zoo Bred (2004)	F
26.	TK 027	Zoo Negara	Nov 9, 2007	Zoo Bred (2004)	F
27.	TK 028	Zoo Negara	Nov 9, 2007	Zoo Bred	F
28.	TK 030	Zoo Negara	Nov 9, 2007	Unknown	F
29.	TK056	*SDWCC	Apr 4, 2017	Selangor	F

30.	TK058	*SDWCC	Apr 4, 2017	[1]Sungkai, Perak	F
31.	TK060	*SDWCC	Apr 4, 2017	[2]Ban Canal, Selangor	F
32.	TK061	*SDWCC	Apr 4, 2017	Pulau Pinang/Terengganu	F
33.	TK062	*SDWCC	Apr 4, 2017	Pulau Pinang/Terengganu	F
34.	TK063	*SDWCC	Apr 4, 2017	Johor / Zoo Melaka	M
35.	TK064	Zoo Kemaman	Oct 8, 2019	[3]Tekah, Terengganu	F
36.	TK065	Zoo Kemaman	Oct 9, 2019	Kelantan	F
37.	TK066	Taman Buaya Langkawi	Aug 8, 2019	Individual purchased from Singapore	F
38.	TK067	Taman Buaya Langkawi	Aug 8, 2019	Individual purchased from Singapore	F

M: Male, F: Female.

[1] Caught in Aug 2016, coordinates: 3.950083333,101.2249722

[2] Caught in Apr 2013, coordinates: 3.692369,101.342225

[3] Caught in June 2014, coordinates: 5.293138889,102.9951944

*Sungai Dusun Wildlife Conservation Centre.

Indistinguishable individuals

Fig 3.1. Physical restraint of *Tomistoma* for blood collection. (A) *Tomistoma* immobilised by six men sitting on the dorsal of the crocodile. (B) Secured snout and limbs of a restrained *Tomistoma*. (C) Blood collection from the lateral approach of the caudal vein.

3.1.3 Sample Storage

Whole blood samples were stored on ice and immediately transported to the laboratory and stored in a -20°C freezer (Haier, China).

3.1.4 Primer Selection

3. 1.4.1 Microsatellite DNA Primers

Eighteen microsatellite markers designed specifically for *Tomistoma* by Chang *et al*. (2014) were used in the present study. Details of these markers are summarised in Table 3.2 with primers and loci codes listed following Chang *et al*. (2014).

3.1.4.2 mtDNA Primers

The mtDNA primers used in this study were based on Kaur *et al*. (2013) which included Set I and Set II amplifying the partial cytochrome *b*-control region (cyt *b*-CR) fragment and Set III targeting the ND 6-tRNA^{glu}-cyt *b* gene fragment (Table 3.3).

3.1.5 Laboratory Analysis

3.1.5.1 DNA Extraction

Total DNA was extracted following Kaur *et al*. (2013) which is a modified phenol chloroform method. An extraction buffer containing 400 g/ml proteinase K, 1% SDS and 50 µL of red blood cells (RBCs) in a final volume of 600 uL was incubated at 37°C overnight. An equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) and then chloroform: isoamyl alcohol (24:1) were used to separate DNA from the proteins via centrifugation at 12,000rpm for

Table 3.2: Details of microsatellite markers used in this study.

Primer name	5'- Forward primer (gttt) Reverse primer - 3'	Product size	Repeat Motif
TM01	GGAAACAGCTATGACCATCAGTTGTCAATGCCCACAC gtttGTGGTGTGTATGTGAAATGGC	149 - 157	$(AGAT)_9$
TM02	GGAAACAGCTATGACCATCCGTGCCCTGCTTAACTAATG gtttCAGAGACAGAGATTAGCGGC	161 - 177	$(AGAT)_{11}$
TM03	GGAAACAGCTATGACCATGTGCCCATTCTGCTCTAC gtttGGTTTGCCTCACATTCACAG	186 - 194	$(AGAT)_8$
TM04	GGAAACAGCTATGACCATTGAAGCCTCCAACATTCTTTC gtttCCTGCATTTGAGGATGGCTG	191 - 215	$(AGAT)_{14}$
TM05	GGAAACAGCTATGACCATCCATGATTAAGATGAAACACCAC gtttACTAAGGGCTTGTGGTTCCTC	225 - 237	$(AGAT)_{12}$
TM06	GGAAACAGCTATGACCATGCTATCTGGAACCCTGTGATG gtttAGTCTGCGCTTCAAGAATGTG	266 - 278	$(AGAT)_8$
TM07	GGAAACAGCTATGACCATAAGCACTTGGAGGAGAACATG gtttCTGACTGCTCCTAATCACCAG	305 - 333	$(AGAT)_{12}$
TM08	GGAAACAGCTATGACCATGTGCAGGTAATTACAGTATTGCAG gtttAATGATCTGCTCTGCCAGGG	389 - 397	$(AGAT)_6$
TM09	GGAAACAGCTATGACCATCCACCTTTGTCTCTGTTCACC gtttAGTGTCATGAGCAGGGATCC	390 - 410	$(AGAT)_{14}$
TM10	GGAAACAGCTATGACCATAAGCACATTCAGGCACATGTAG gtttGGTCAGCTCAGCATTTAGGC	411 - 427	$(AGAT)_{10}$
TM11	GGAAACAGCTATGACCATGACCGATATATCAGTGCACC gtttGCCTGCTAACCCTCATAGTC	412 - 428	$(AGAT)_{10}$
TM12	GGAAACAGCTATGACCATTGTTCCACCATTTCCTGTCC gtttGATCTGGAGAAGCTCTGGAGG	488 - 504	$(AGAT)_{14}$
TM13	GGAAACAGCTATGACCATAAGAAGCGGCAAGGAAGTTC gtttAACCCACTCCTCCACTACAC	438 - 466	$(AGAT)_8$
TM15	GGAAACAGCTATGACCATGAGTTGACCCTGCCAAATG gtttGGGCTTGGCGTTCATTACTC	470 - 526	$(AGAT)_7$
TM19	CAGTCGGGCGTCATCAGGAGCATTGGATCTGAAGCTG	430 - 442	$(AGAT)_6$

TM21	gtttGTCCTATGTTGTGCCATCCC CAGTCGGGCGTCATCAGCTGTGGACAATGGCATCTC gtttCTGTTGCATGACCTTGAGACC	470 - 490	$(ACAG)_{13}$
TM23	GGAAACAGCTATGACCATAATAAGCACATATGAACAGGTG gtttGGCTGGATTTCTGGTTCTTGG	207 - 275	$(ACAG)_{36}$
TM24	GGAAACAGCTATGACCATGGCTCACTATACTTCACAGAT gtttACTCTCAACAAACCTACCTGT	286 - 439	$(ACAG)_{26}$

[Footnote]: Primer names and details were based on Chang *et al*., (2014) without modification.

Table 3.3: Details of mtDNA primers used in this study.

Marker	Primer set	Primer name	Primer sequence (5' to 3')	Fragment size
cyt *b*-CR region	Set I	L14930	AGCGGGCAAAATAGAAAACTGA	701 bp
		H15630	ATAGAGATGCCGGGATTACGAA	
	Set II	L15516	TATCACTCTCATGTACTCTTCTTG	417 bp
		H15932	AGCTATTTTCATTTTTATTTCTAT	
ND 6-tRNA^{glu}-cyt *b*.	Set III	L13457ND6	CCCCACAAACACCAACCAA	570 bp
		H14027CB	CGACGGATGCGAAGGCTATG	

[Footnote]: Primer names and details were based on Kaur *et al*. (2013) without modification.

5 mins. Chilled pure ethanol was the used to precipitate the DNA from the solution by gently inverting the tubes for approximately three to five minutes. The ethanol was discarded, and the DNA was subsequently dried using a vacuum desiccator to remove any traces of ethanol. The DNA was eluted in 100 µL of 10 mM Tris-HCl buffer overnight at 4°C before final storage at -20ºC.

3.1.5.2 DNA Quantification

Extracted DNA was checked for quantity and purity at absorbances of 260 nm and 280 nm (A260/A280) with NanoPhotometer UV/Vis spectrophotometer (Implen, Germany). DNA bands were also visualised using gel electrophoresis by running 1 µL of the eluted DNA with 1 uL of loading dye (Thermo Fisher Scientific, USA) and 1 µL of a 1 kb ladder (Thermo Fisher Scientific, USA) on a 1% agarose gel premixed with GelRed (Biotium, Fremont, USA). The gel was visualised with UV Bioimaging system (UVP, Upland, USA) after running the gel electrophoresis at 70V for an hour.

3.1.5.3 Polymerase Chain Reaction (PCR) Using Microsatellite Markers

Samples included for PCR were all the 38 samples in Table 3.1. The optimum PCR conditions in a 25 uL final volume for all the primers were generally Taq DNA Polymerase (Invitrogen, Thermo Fisher Scientific, USA) at 0.5 U, $MgCl_2$ at 2 mM, each primer at 0.5µM, 0.2 mM dNTPs mix (Invitrogen) and 50-150 ng of DNA. Similarly, PCR cycles were initial denaturation at 95°C for 2 min, followed by 25 cycles of 95°C for 30 sec for denaturation, annealing temperature for 30 sec, and 72°C for the elongation time, with a final elongation at 72°C for 5 min. Differences in the PCR parameters are shown in Table 3.4.

Table 3.4: Optimised PCR parameters and conditions for the microsatellite markers.

Markers	Annealing temperature (°C)	Elongation time (sec)
TM02	55	11
TM03	55	12
TM04	58	13
TM05	55	15
TM07	55	20
TM08	57	24
TM10	57	26
TM11	55	26
TM12	55	31
TM13	53	28
TM15	55	32
TM21	56	30
TM23	57	17
TM24	50	26

3.1.5.4 PCR Using mtDNA Markers

Samples used for PCR were the ten new individuals in Table 3.1 (TK056-TK067) as sequences from the other 28 individuals are available in GenBank (accession numbers HM593977- HM594032, HM594033- HM594088). Since both the cyt *b*-CR and the ND 6-tRNA^{glu}-cyt *b* primers were from the study by Kaur *et al*. (2013), the optimised PCR conditions in that study were followed: Taq DNA Polymerase (Invitrogen, Thermo Fisher Scientific, USA) at 0.5 U, MgCl2 at 1 mM, 0.2 mM dNTPs (Invitrogen) and 50ng of DNA. The variations in the PCR parameters are listed in Table 3.5. PCR cycles also followed Kaur *et al*. (2013) namely the initial denaturation at 95°C for 2 min, followed by 35 cycles of 95°C for 30 s for denaturation, annealing temperature for 30 s, and 72°C for the elongation time, with a final elongation at 72°C for 5 min.

Table 3.5. Optimised PCR parameters and conditions for the mtDNA primer sets.

Primer set	Primer conc. (μM)	Annealing temp. (°C)	Elongation (sec)
Set I	0.3	54	45
Set II	0.3	55	25
Set III	0.2	55	35

3.1.6 Gel Electrophoresis

For microsatellite PCR products, polyacrylamide gel electrophoresis (PAGE) was also optimised for gel concentration and running time (Table 3.6). All runs were at the constant voltage of 70V. All samples were subjected to PAGE twice, first without clustering of samples, followed by the second run whereby visually similar band sizes were arranged side by side in a gel for easier confirmation of band size in downstream analysis. The second run was carried out with at least 3 lanes for the 50 bp DNA Ladder (New England Biolabs Inc. (NEB), USA) placed at the far left and right of the gel lanes and a third placed in the middle lane of the gel for better precision in normalisation of the band sizes against the DNA ladder.

For mtDNA PCR products, a 1% agarose gel premixed with GelRed was run for 40 mins at a constant voltage of 70V. Target bands were accessed for any smearing by running 5 uL of PCR product mixed with 1 uL of loading dye and 1 uL of a 1 kb ladder (Thermo Fisher Scientific, USA) to determine if gel purification was needed before sequencing.

Table 3.6: Optimised PAGE conditions.

Markers	PAGE concentration (%)	Electrophoresis time (hours)
TM02	12	6
TM03	12	4
TM04	12	4
TM05	12	4.5
TM07	6	2
TM08	6	3
TM10	8	4.5
TM11	8	4.5
TM12	6	2.5
TM13	10	5.5
TM15	8	4.5
TM21	8	4
TM23	10	6
TM24	10	6

3.1.7 DNA Purification

Before sequencing, the PCR products were purified using SV spin columns kit (GeneAll, Korea), following manufacturer protocols. The gel purification method was used when there was smearing and primer dimer formation. Bands of interest were excised from a 1% agarose gel and diluted with the GB Buffer at a 1:3 ratio before incubation at 50°C. This mixture was then centrifuged and a final additional centrifuge step to remove gel traces was carried out. The column was then washed with the NW Buffer and centrifuged. All centrifugation steps were carried out at 12,000 rpm and for 30 s. An additional centrifuge step at 13,500 rpm for 1 min was carried out to remove traces of the buffer before adding the Elution Buffer.

For clear bands without smearing, the direct purification method was utilised. The PCR products were mixed with the PB buffer at a 1:5 ratio and subsequently washed with the NW buffer and centrifuged at 12,000rpm for 30 s. The subsequent steps were similar to the gel purification method.

3.1.8 Sequencing

For both the microsatellite and mtDNA, the same primers for PCR amplification were used for DNA sequencing. Sequencing in both directions was carried out by outsourcing to NHK Bioscience Solution (Malaysia). All the new samples in this study (TK056-TK067) were sequenced for all three mtDNA primer sets (Table 3.7). For microsatellite markers, one samples for each primer set was sequenced for confirmation of repeat size and sequences. Samples that produced single band were selected because double bands failed to separate adequately on the agarose gel, even when the gel concentration was increased from 1 to 5%.

Table 3. 7: Details of samples and markers used for sequencing.

No.	Sample	Markers			
		Microsatellite	mtDNA		
		Locus	Set I	Set II	Set III
1	TK 004	TM02	-	-	-
2	TK 005	TM03, TM05, TM08	-	-	-
3.	TK 008	TM10	-	-	-
4.	TK 009	TM07	-	-	-
5.	TK 010	TM11	-	-	-
6.	TK 011	TM23	-	-	-
7.	TK 020	TM12, TM13, TM15, TM21	-	-	-
8.	TK 023	TM24	-	-	-
9.	TK 024	TM04	-	-	-
10.	TK 056	-	✓	✓	✓
11.	TK 058	-	✓	✓	✓
12.	TK 060	-	✓	✓	✓
13.	TK 061	-	✓	✓	✓
14.	TK 062	-	✓	✓	✓
15.	TK 063	-	✓	✓	✓
16.	TK 064	-	✓	✓	✓
17.	TK 065	-	✓	✓	✓
18.	TK 066	-	✓	✓	✓
19.	TK 067	-	✓	✓	✓

3.1.9 Microsatellite Analyses

3.1.9.1 Scoring and Scoring Errors

Band sizes were scored with GelJ v 2 (Heras *et al*., 2015) using the DICE method while tolerance and similarity percentage was left as default. In GelQuest v 3.4.3.0 (Sequentix Digital DNA Processing, Klein Raden, Germany), band sizes were obtained by extracting and transforming the lanes of the gel images into a trace data curve and selecting base sizes as the parameter.

A check for scoring errors was carried out using MICRO-CHECKER v 2.2.3 (van Oosterhout *et al*., 2007) to detect large allele drop-out, stuttering and null alleles. The analysis was carried out at 95% confidence levels with 10,000 randomisations assuming (i) a single Malaysian population, (ii) Peninsular Malaysia and Sarawak as separate population and (iii) clustering computed by STRUCTURE v 2.3.4 (Pritchard, Stephens and Donnelly, 2000).

3.1.9.2 Population Structure

As the exact locations of all except three individuals (TK058, TK060 and TK064) were unknown (Table 3.1), STRUCTURE v 2.3.4 (Pritchard, Stephens and Donnelly, 2000) was used to identify subsets of the whole data based on allele frequency differences, as this program assigns individuals to their respective subpopulations based on Bayesian clustering approach.

The analysis was done using the admixture model and assuming different values of F_{ST} for different populations. The burnin was set at 100,000 with the MCMC reps at 1,000,000 while numbers of clusters (k) set at minimum of 1 to maximum

of 10 and the number of iterations as 10. The analyses were carried out considering the dataset without any population assignment and without using the locprior options.

Structure Harvester Web v 0.6.94 (http://taylor0.biology.ucla.edu/structure) was used to visualise results from STRUCTURE and the best K was determined using the method by Evano implemented in this program (Earl and vonHoldt, 2012). The program CLUMPAK (CLUMPAK server (tau.ac.il)) was used for graphical presentation of the results by STRUCTURE.

Apart from the Bayesian approach above, the distance-based analysis was carried out using POPULATIONS v 1.2.32 (Langella, 2015) to first generate a distance matrix and subsequently a dendogram for individuals. The Nei *et al*. (1983) distance, D_a, was selected as it assumes genetic differences due to genetic drift and mutation and generates reliable population trees than other distance measures for microsatellite DNA data. The Neighbor-Joining (NJ) algorithm which allows for unequal rates of evolution was used for the dendogram and bootstrapping was perform with 1,000 randomisations.

3.1.9.3 Linkage Disequilibrium

Deviations from random association between genotypes, Linkage Disequilibrium (LD) by Weir (1996), was carried out using Genepop v. 4.7.5 (Rousset, 2008) with Markov chain parameters set as 10,000 dememorizations with 10,000 iterations for each of the 100 batches. The deviations from LD at pairs of loci in each sample was tested with the G-statistic. FSTAT v 2.9.4

(Goudet, 2003) was used for the gametic disequilibrium (*D'*), which is less dependent on allelic frequencies so that an unbiased estimator can be obtained. The testing of significance was by the log-likelihood ratio G-statistic and the estimation of the p value was done by randomisation of the data set at 10,000, adjusted at nominal level 1% and 5% (by sequential Bonferroni correction).

3.1.9.4 Genetic Diversity Analyses

The number of alleles (N_A), number or effective alleles (N_E), number of private alleles (P_A), observed and expected heterozygosity (H_o and H_e) were generated in GenAlEx v 6.51b2 (Peakall and Smouse, 2006). Additionally, these were also represented in a graphical chart, together with the number of locally common alleles occurring in 50% or less of the populations (No. LComm Alleles < = 50%) and number of different alleles with a frequency > = 5% (Na. Freq. > = 5%). The mean F_{IS} and the Hardy Weinberg Equilibrium (HWE) was also analysed with GenAlEx. Molkin v 3.0 (Gutiérrez *et al*., 2005) was used to compute the Polymorphic Informative Content (PIC). These analyses were carried out assuming separate population following the clusters generated by STRUCTURE.

Because genetic subdivisions can lead to deficit of heterozygotes in the total sample (Wahlund effect, Wahlund, 1928), tests for deviations from Hardy-Weinberg equilibrium (HWE) was carried out using the permutation test (set at 10,000) implemented in the program FSTAT. This program was also used for allelic richness (A_R), which is the number of alleles independent of sample size, calculated per locus per clusters and overall samples for comparisons between

populations. Additionally, the inbreeding coefficient (F_{IS}) was also generated with FSTAT. Tests were carried out at the locus and sample level with 10,000 permutations. The proportion of randomization that gave a larger F_{IS}-value than the observed was used to accept the null hypothesis.

3.1.9.5 Genetic Differentiation

Weir and Cockerham estimators of F-Statistics which measures population structure based on inbreeding was carried out in FSTAT, as it corrects for the unequal and small sample size for unbiased comparisons between clusters or populations of different sizes. The pairwise F_{ST} was also carried out following results of STRUCTURE in FSTAT with 10,000 permutations for the nominal p value of 5%. The AMOVA analysis was carried out using GenAlEx for R_{ST} statistic considering the clusters identified by STRUCTURE and subsequently without considering the cluster comprising the most divergent haplotype, H4.

3.1.9.6 Sex-Biased Dispersal

The assignment-based test in GenAlEx was used for the individual level analysis. The significance of the corrected assignment index was determined with a nonparametric Mann-Whitney U test in GenAlEx. The dispersal test was run considering all samples as Malaysia and subsequently only for Peninsular Malaysian samples.

3.1.9.7 Relatedness

Relatedness was estimated in FSTAT using Hamilton's (1971) relatedness, relat, an estimator which is strictly equivalent to Queller and Goodnight (1989). It is

the average relatedness of individuals within clusters compared to the whole population. Randomisation was carried out with 10,000 permutations.

Relatedness was also carried out using genetic similarity by the maximum likelihood estimates in ML-Relate (Kalinowski *et al*., 2006) which accommodates null alleles. The presence and the frequency of null alleles were tested with U statistic using the Monte-Carlo randomization test (Guo and Thompson, 1992). Four characterizations of relationships namely the parent-offspring (PO); full-siblings (FS); half-siblings (HS) and unrelated (U) were tested at the 95% confidence level with 1,000,000 permutations. The relationships categorized as PO or FS were accepted as kinship.

Molkin program was used for the molecular coancestry matrix and within cluster molecular coancestry of clusters according to Caballero and Toro (2002). These measures were weighted by the PIC value.

3.1.9.8 Gain / Loss Analyses

The gain / loss analysis was carried out by removing one cluster at a time from the data set in Molkin. The maximum overall Nei's (1987) gene diversity (GD) was analysed with the bootstrap option due to unequal sample sizes of the clusters. 100 samples of 100 individuals per population were generated in the bootstrap procedure. The Petit, Mousadik and Pons (1998) diversity, which uses the Hurlbert's (1971) estimator was used to assess the contribution to the total allelic richness to the cluster. This was carried out with rarefaction of number of alleles per locus (k = four copies, the minimum in this dataset).

3.1.10 mtDNA Analyses

3.1.10.1 Single Contig And Sequence Identity

Any mismatch between the forward and reverse sequences generated on the DNA chromatogram was corrected manually in Finch TV v 1.4.0 (Geospiza Inc., USA). For the control region, a single contig was obtained when the primer set I and II sequences were overlapped using DNA Dragon v 1.5.1 (SequentiX) program. Default parameters were used under the assemble function which were full search option with minimum overlap diagonal length as 80 bases or higher, 80% identity of complete overlapping fragment and maximum number of consecutive gaps to insert as 8 bases.

The nucleotide BLAST search for gene homology was carried out using default options in the GenBank website for the control region and the ND 6-tRNA^{glu}-cyt *b* fragment separately.

3.1.10.2 Multiple Sequence Alignment and Concatenation Of The ND 6-tRNA^{glu}-cyt *b* And The Single Contig Of Control Region

A multiple sequence alignment (MSA) was generated and trimmed for the cyt *b*-CR region and the ND 6-tRNA^{glu}-cyt *b* using the ClustalW program in MEGA X v10.0.5 (Kumar *et al*., 2018). DNA sequences from Kaur *et al*., (2013) for samples TK002-TK038 (GenBank Accession for cyt *b*-CR: HM593977 to HM594032, ND 6-tRNA^{glu}-cyt *b*: HM594033 to HM594088) and other relevant sequences available in GenBank (GenBank Accession MT554040, KR674082 and AJ810454) were also incorporated into each dataset.

The ND 6-tRNA^{glu}-cyt *b* and the cyt *b*-CR region datasets were converted into nexus files for the subsequent concatenation of the mtDNA fragments using DnaSP v 6 (Rosaz *et al*., 2017), to facilitate data comparison against Kaur *et al*., (2013). The concatenated dataset was used for all subsequent analyses.

3.1.10.3 Genetic Diversity Analyses

The program DnaSP was used to generate the haplotype diversity, *h*, and nucleotide diversity, π, at the clusters and the entire data set level. Both *h* and π and their variances were estimated following Nei (1987).

3.1.10.4 AMOVA Analyses for Population Structure

The AMOVA analysis for mtDNA data was carried out using GenAlEx by treating clusters as populations using adults only. This was repeated by excluding the most divergent haplotype, H4. This additional analysis was carried out as it was identified as divergent from the other haplotypes and showed a higher intraspecific divergence compared to other crocodilians (Kaur *et al*., 2013).

3.1.10.5 AMOVA Analyses for Sex-Biased Dispersal

A separate AMOVA analysis was carried out using adult females only. Similar to population analyses, clusters were treated as populations, and the analysis was repeated by excluding haplotype H4. Due to the low representation of males in each cluster, the AMOVA analysis for males in clusters could not be carried out.

3.1.11 Haplotype network

Haplotype networks were generated using median joining network, which is a distance-based approach was selected as it can infer unsampled sequences to the nodes in the total length of the network, with PopArt (Leigh & Bryant, 2015). This network comprised the concatenated dataset (Table 3.1) and *T. schlegelii* sequences available in GenBank for comparison (GenBank Accession number MT554040 from Pan *et al*., 2021; GenBank Accession number AJ810454 from Janke *et al*., 2005).

A second median joining network was carried out using the above sequences, truncated to 451 bp of the D-loop. This was to compare the findings of Kurniati (2003) which comprised of eleven individuals from the west, central and east Kalimantan. The sequence from Md Adzhar *et al*. (2017) (GenBank Accession number KR674082) was also incorporated into this analysis. The summary of samples used for the networks and subsequent phylogenetic trees is in Table 3.8.

3.1.12 Selection of Best-Fit Models of Nucleotide Substitutions

MEGA X was used to identify the for the concatenated dataset, ND 6-tRNA^{glu}-cyt *b* and the cyt *b*-CR using default settings; Maximum Likelihood statistical method using all sites and Automatic Neighbour-joining tree). Additionally, jModelTest v 2.1.10 (Darriba *et al*., 2012) was also used using default settings; NNI branch swapping, each with Bio-NJ as a starting tree, to compare the model with that of MEGA X. The analysis was carried out using both the Bayesian Information Criterion (BIC) and the Akaike Information Criterion,

corrected (AICc). The former is a stricter criterion to identification nucleotide substitution models (Posada, 2008).

3.1.13 Phylogenetic Trees

3.1.13.1 The Maximum Likelihood Tree

MEGA X was used to generate the Maximum Likelihood (ML) tree. Using all sites and the model and rates generated by jModeltest, all other options were left as default. The tree was generated using all available samples in GenBank (Pan *et al*., 2021; Kaur *et al*., 2013; Janke *et al*., 2005) with the ten new samples in this study. Test of phylogeny was done using the bootstrap option with 100 replicates with *Gavialis* (GenBank Accession AJ810455 and NC008241) and *Crocodylus porosus* (NC008143) as outgroups.

Table 3.8: List of *T. schlegelii* samples used for the networks and phylogenetic trees analysed.

Analysis (Data type)	Sample size	Sample
Haplotype network (Concatenated dataset)	40	This study (TK056-TK067), Kaur *et al*., 2013 (TK002-TK030), Janke *et al*., 2005 (AJ810454), Pan *et al*., 2021 (MT554040).
Haplotype network (D loop)	52	This study (TK056-TK067), Kaur *et al*., 2013 (TK002-TK030), Janke *et al*., 2005 (AJ810454), Pan *et al*., 2021 (MT554040). Md Adzhar *et al*., 2017 (KR674082), *Helen, 2003
Maximum likelihood tree and Divergence time tree (Concatenated dataset)	68	This study (TK056-TK067), Kaur *et al*., 2013 (all samples), Janke *et al*., 2005 (AJ810454), Pan *et al*., 2021 (MT554040)

*Sequences presented as multiple sequence alignment with the published article. Brackets: Sample labels

3.1.13.2 Estimation of Haplotype Divergence Time

Divergence of lineages was carried out using BEAST (Bayesian evolutionary analysis sampling trees) v 2.6.7 (Bouckaert *et al*., 2014) package for evolutionary inference from molecular sequences. The program Beauti in this package was used to prepare the file for BEAST. The concatenated data was set as non-coding as both the NADH 6 and cyt *b* are partial sequences. Site model was set following results of MEGA and jModeltest, leaving all other parameters in the 'Site model' as default. The clock model selected was the random local clock (which considers branch rate for each branch in the tree) and all other parameters set as default. Other sequence available from GenBank for *Tomistoma* (GenBank Accession MT554040 and AJ810454) were also incorporated. Sequences of *Gavialis* (GenBank Accession AJ810455 and NC008241) *Crocodylus porosus* (NC008143) were also incorporated as the most recent common ancestor in absence of a fossil.

The calibrated Yule tree prior was selected because a constant lineage birth rate is assumed for each branch in the tree. The genera *Gavialis* and *Tomistoma* were treated as monophyletic with a log normal distribution centred about 22.2 million years with a 95% probability range covering the current consensus estimate of the date of the most recent common ancestor of these two genera which is 18-27 Mya (Roos *et al*., 2007). A log normal distribution with a 95% probability range covering the current consensus estimate of the date of the most recent common ancestor for the *Crocodylus* and *Gavialis*/ *Tomistoma* split was set at 80 million years (Janke *et al*., 2005). The MCMC chain was set at 50,000,000, with a burnin of 1,000.

The file was saved and run in BEAST. The tree file generated was then run in Tree Annotator program in this package with a 10% burnin and the target tree type set as Maximum Clade Credibility Tree. Mean heights was selected for the node heights which sets ages of each node in the tree to the mean height across the entire sample of trees for that clade. The final tree was visualised using FigTree v 1.4.4 (Rambaut, 2018).

3.2 Field Study

3.2.1 Site Selection

Based on records from the Department of Wildlife and National Parks, Peninsular Malaysia (PERHILITAN) (Khairul Amirin, pers. comm.), four rivers in Peninsular Malaysia were selected for the study: (i) Sungkai River, Perak; (ii) the Nerus River and its tributary, the Tekah River, Terengganu; (iii) the Dusun River and its tributary, the Tengi River, Selangor; and (iv) the Jengka River, Pahang. These sites were chosen due to recent captures of T. schlegelii from the wild (Fig. 3.2).

The start and end points of the surveys are listed in Table 3.9. The study sites were primarily located within palm oil plantations, with some sections passing through secondary forest. The lower reaches of Ban Canal and the Tengi River were surrounded by peat swamp, however, recent land conversion in the upper canal region has repurposed areas for livestock farming and palm oil plantations. The delineation of the Jengka River survey area around Kampung Bukit Lada was based on information provided by a boatman following the capture of an individual in 2014 (Fig. 3.3). While this individual was placed in captivity (either

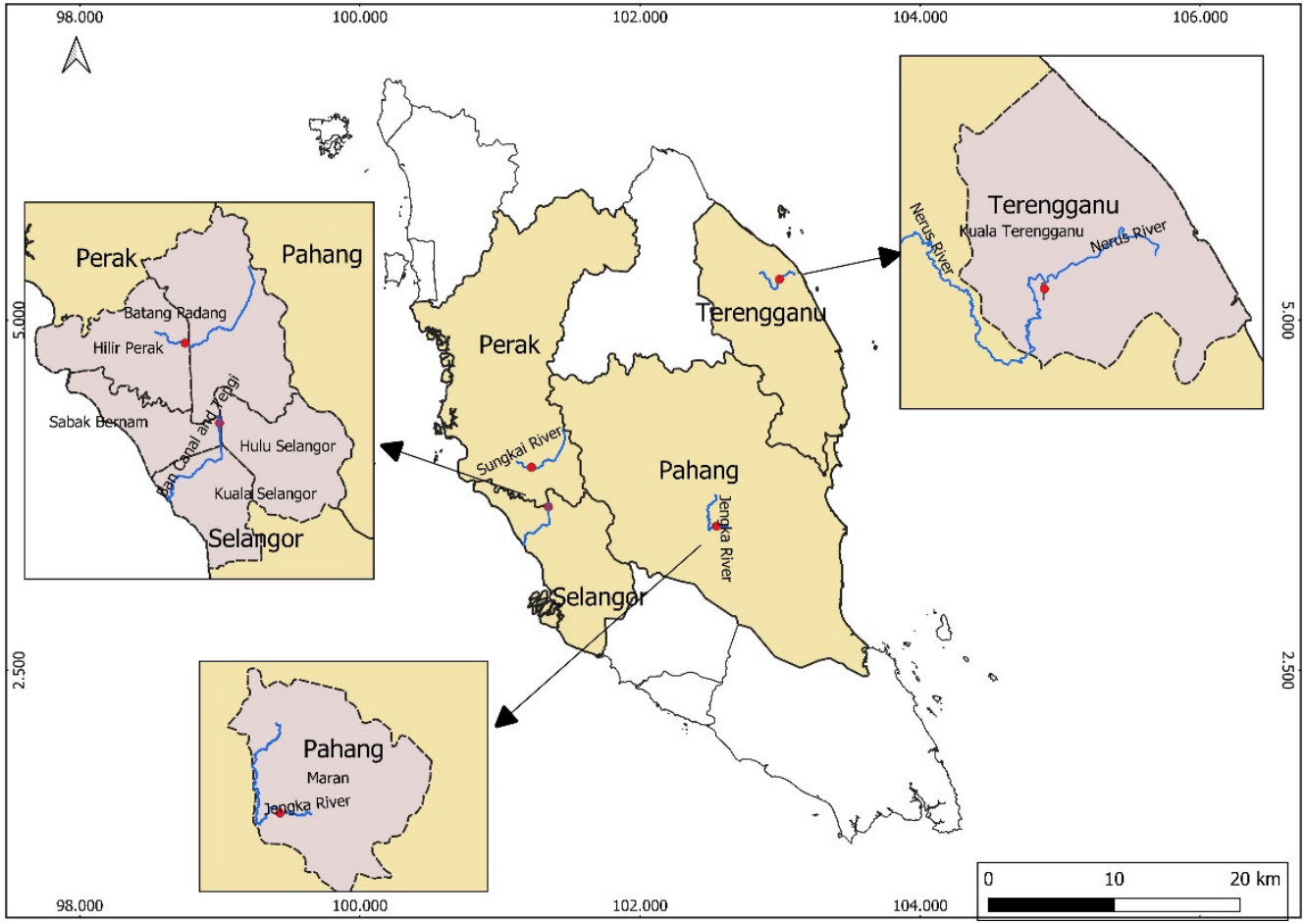


Fig. 3.2. The rivers sampled in Peninsular Malaysia. Red dot: Capture locations of *T. schlegelii*.

Table 3.9: Details of survey areas of the rivers selected in this study.

Site / river	River		Survey area	
	Start of river	End of river	Start of survey	End of survey
Nerus	5.3369, 102.85608	5.326273, 103.09887	5.322472, 103.042778	5.269139, 102.989000
Tekah	5.299188, 102.997198	5.290898, 102.995697	5.299667, 102.996111	5.293139, 102.995194
Sungkai	4.149025, 101.462277	3.989173, 101.119704	3.948722, 101.187278	3.952972, 101.284972
Dusun /Ban Canal	3.693667, 101.341194	3.574417, 101.351861	3.693667, 101.341194	3.574417, 101.351861
Tengi	3.573250, 101.351778	3.481778, 101.221417	3.573250, 101.351778	3.481778, 101.221417
Jengka	3.789650, 102.533325	3.517233, 102.609868	3.530750, 102.542556	3.529250, 102.556806

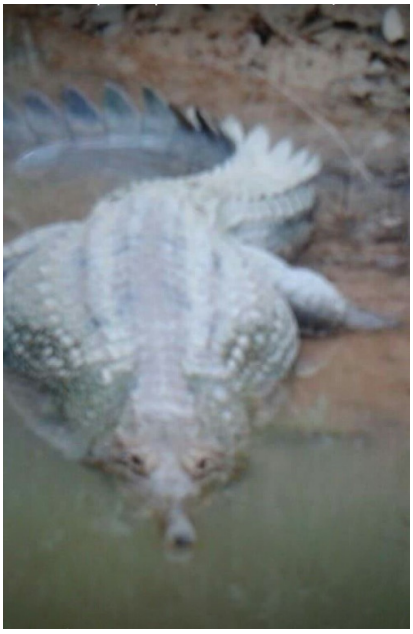


Fig. 3.3. Photo of an adult *Tomistoma* captured in 2014 from the Jengka River near the cemetery (3.529871, 102.547297). Photo provided by Tarmizi Razak.

TK061 or TK062, Table 3.1), another individual, captured on land at coordinates 3.528496, 102.556276 after the 2015 floods subsided, was immediately released into the Pahang River upon capture.

3.2.2 River Surveys

The rivers were surveyed during day and night in both the dry and wet seasons. For the dry season, surveys were carried out in April (inter-monsoonal period after the northeast monsoon) and the wet season surveys were in September during the southwest monsoon (Wong *et al*., 2009). The duration of each survey was 2-4 days. The boatmen in Sungai Dusun and Jengka were experienced fishermen well-versed in navigating the rivers. Meanwhile, the villagers who caught the *Tomistoma* in Terengganu and Perak, also regular fishermen were hired for their expertise. Surveys in Sungkai River were carried out on motorbike due to the navigational challenge faced when using a boat.

3.2.2.1 Day Surveys

Surveys were conducted on foot and motorbikes for the river in Perak and using fibreglass boat and sampan (not more than 3 meters) for the rivers in Selangor and Terengganu. A kayak (2 meters) was used for the river in Pahang. Surveys began at 9 am and ended at approximately 2 pm. The GPS unit (Garmin GPS eTrex® Basic) was left turned on throughout the survey to maintain the satellite signal and reduce recording error.

3.2.2.2 Data Collection

The targeted GPS sites and water parameter measurements were approximately 0.5 km from each site, except for the Sungkai River. In addition to this, locations and measurements were also noted at areas where a *Tomistoma* had been spotted or caught before by the locals and PERHILITAN staff and at the junctions of streams and tributaries draining into the main river.

The water parameters recorded was water temperature (◦C), pH, salinity (ppm), total dissolved solids (TDS) (ppm) and conductivity (µS), using Multi-Parameter Testr 35 Series. These parameters were measured three times; at both sides of the riverbank and in between the two banks for each site.

Observation of activities at and by the river was noted together with the type of vegetation at the riverbanks and the characteristics of the riverbank. Photos of these were taken using a digital camera (Nikon Coolpix) and by hand phone (HTC one M9). Photos of the capture sites were also taken.

3.2.2.3 Night Surveys

The modified methodology for night spotlight surveys (Bayliss, 1987) was employed for this purpose. Surveys were carried out an hour after night fall (8-8.30 pm). Similar to the day surveys, a boat was used to survey the rivers in Selangor and Terengganu while a kayak was used for surveys in Pahang. In Selangor, surveys were done by paddling the boat (with engine off) from start point to end point (down river). The engine was turned on to return to the start point upriver. In Terengganu, the boat engine was on for both directions as this river is tidal and wide which makes paddling impractical. Surveys in Pahang were done using the kayak without navigation back to the starting point. The Sungkai River in Perak was surveyed on foot as this river is difficult to navigate using boat and kayak and therefore was only done at the location around and at the recent *Tomistoma* capture (100 meters up and downriver to that point).

3.2.3 Statistical Test

The data collected was analysed using JASP v 0.16.3 (University of Amsterdam) for descriptive analyses (mean and standard error) for each water parameter, in each river and season separately. The same software was also used for the inferential analyses. Paired t-test was used to test for significant difference in means of each water parameter in between seasons. When there was a significant difference, the one tailed t-test was carried out to identify which season the mean differed in. Before carrying out the paired t-test, normality was checked with Shapiro-Wilks test ($p>0.005$).

The variance of each parameter and seasons was also analysed between all the rivers using ANOVA and the non-parametric Kruskal - Wallis. Test assumptions were checked prior to that. The Levene's test was used to check homogeneity of variance and normality was checked using the Q-Q plot. If the Levene's test was not significant with a normal distribution by the Q-Q plot, the standard ANOVA was carried out. If the ANOVA was significant, then the post hoc test was done to identify which factors and parameters were significantly different. When the Levene's test showed non-homogeneity of variances and the Q-Q plot showed a normal distribution, the Welch ANOVA was carried out and if significant, the Games-Howell post hoc test was carried out to identify which factors differed. When both assumptions were violated, the Welch ANOVA was still used as the ANOVA has been stated to be robust to departures from population normality (Najdi and Ahad, 2019) together with the non-parametric Kruskal-Wallis.

3.2.4 Geographic Information System

Terrin analyses which comprised the elevation, slope and Strahler order of rivers and capture sites was carried out using the software QGIS v 3.34.1 (QGIS.org, 2023). Similarly, the distance to the nearest hub of the capture sites was also carried out with QGIS.

3.2.4.1 Distance to the Nearest Hub

The country, states, district boundaries and rivers shapefiles were obtained from Diva-GIS (https://www.diva-gis.org/gdata) and were reprojected into state grids (GDM 2000 / Selangor Grid for Tengi River, GDM 2000 / Perak Grid for Sungkai River, GDM 2000 / Pahang Grid for Jengka River, GDM 2000 /

Terengganu Grid for Nerus River). Point data of villages, towns, city and neighbourhoods were obtained from Open Street Map using OSM data tool of area surrounding the rivers. Polygons of Protected Areas (PAs) were downloaded from protected planet website (www.protectedplanet.net) for terrestrial and inland waters protected areas of the states surveyed. These together with the coordinates of section of river surveyed and the confirmed site of capture were reprojected into the respective coordinate reference system (GDM 2000 respectively).

The distance to the nearest settlement (villages, town, neighbourhood and city) and PAs was generated using the vector analyses in QGIS. The distance to nearest hub (by lines) was carried out for PAs while distance to nearest hub (by points) was used for settlements comprising of villages, neighbourhoods and township. These distances are based on feature centres in a direct line.

3.2.4.2 Terrain Analyses

The digital elevation models (DEMs) of the capture sites and areas surrounding the respective rivers were downloaded using the STRM downloader plugins in QGIS from Earthdata (https://urs.earthdata.nasa.gov). For the river elevation, the plugin QChainage was used to plot every 1 kilometre (km) from the beginning to the end of the river, except for the Tekah River which was every 0.1 km. The point sampling tool was used to extract the elevation and DataPlotly plugins was used to generate the bar plots of elevation against the distance of each km. For the elevation of the exact point of capture, the profile tool plugin was used as it plots terrain profile along selected polylines (river).

The Strahler Order of the rivers / streams around or at the capture site was obtained by using the SAGA plugins for terrain analyses. The fill analyses were carried out using Wang and Liu with default settings (minimum slope of 0.1 degree). The Strahler Order was generated using the fill DEM and the raster calc function was used to estimate the best Strahler order for the site of capture. Subsequently the Channel Network and Drainage Basin function was used to generate the channel network by setting the threshold to the Strahler Order obtained by raster calc.

CHAPTER 4

RESULTS

4.1 Genetic Study

4.1.1 DNA Extraction and Quantification

The extracted DNA (Appendix A) had a purity between 1.7-2.0 (A260 / A280) and concentration ranging between 120,000 to 140,000 ng/μL which was subsequently diluted to 50 ng/μL for downstream procedures (Table 4.1).

4.1.2 PCR

For the microsatellite primers, TM1, TM6 and TM19 failed to optimise showing multiple bands while TM9 had only one band of the same size throughout the samples and therefore were not included in the downstream analyses. PCR for the mtDNA cyt *b*-CR region and the ND 6-tRNA^{glu}-cyt *b* was unproblematic.

4.1.3. DNA Sequencing

4.1.3.1 Microsatellite

For microsatellite primers, good sequencing results with clear peaks were seen in all primer sets except in TM7. The marker TM7, produced double bands in all individuals in the sample set (Fig. 4.1). This led to poor separation of the bands on the agarose gel even when concentration of the gel was increased (APPENDIX B).

For marker TM23, although good sequencing was obtained from both forward and reverse sequencing, the 5' of the reverse sequencing was not able to generate

Table 4.1: The list of extracted samples, DNA purity and concentration.

No.	Sample	Spectrophotometry A260/A280	DNA Concentration (ng/µL)
1	TK 002	1.9	119,873
2	TK 003	2.0	131,000
3.	TK 004	1.9	120,111
4.	TK 005	1.9	127,844
5.	TK 006	1.9	139,120
6.	TK 007	1.7	125,120
7.	TK 008	2.0	129,300
8.	TK 009	1.7	119,900
9.	TK 010	1.8	136,284
10.	TK 011	1.8	127,239
11.	TK 012	2.0	137,199
12.	TK 013	1.7	119,249
13.	TK 014	1.7	126,890
14.	TK 015	1.9	139,000
15.	TK 016	1.7	128,222
16.	TK 017	1.7	140,060
17.	TK 018	2.0	126.487
18.	TK 019	2.0	135,629
19.	TK 020	2.0	139,929
20.	TK 021	1.9	128,115
21.	TK 022	1.7	126,298
22.	TK 023	1.7	120,200
23.	TK 024	1.8	136,000
24.	TK 025	1.9	137,496
25.	TK 026	1.9	126,290
26.	TK 027	2.0	129,389
27.	TK 028	2.0	120,741
28.	TK 030	1.8	119,700
29.	TK 056	1.8	133,556
30.	TK 058	1.8	139,940
31.	TK 060	1.9	134,100
32.	TK 061	1.8	139,734
33.	TK 062	1.8	138,649
34.	TK 063	1.9	127,931
35.	TK 064	1.8	133,167
36.	TK 065	1.9	139,950
37.	TK 066	1.8	136,114
38.	TK 067	2.0	138,973

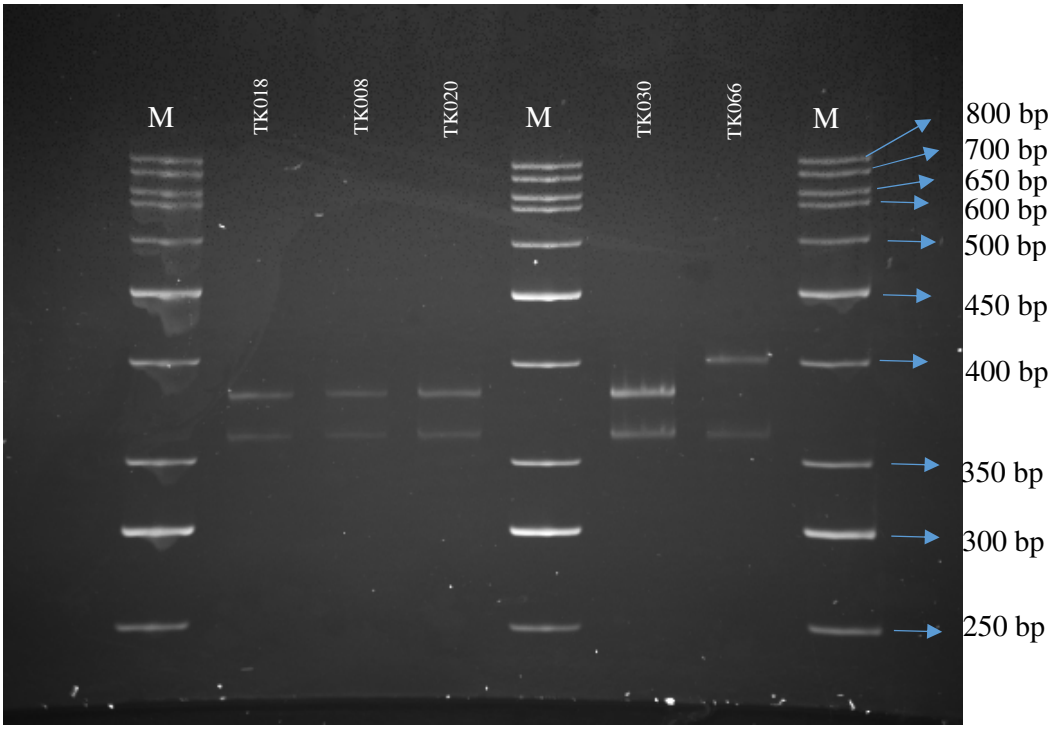


Fig. 4.1. Double bands for TM07 on the optimised PAGE gel showing different migration rates and separation of the heterozygous bands. Labels on lanes: *T. schlegelii* samples. M: 50 bp DNA ladder.

clear and strong peaks. Therefore, perfect alignment of both the forward and reverse sequencing could not be achieved for the entire fragment (APPENDIX C). The ten new samples produced good sequencing results for all three mtDNA primer sets.

4.1.3.2 mtDNA

Contig assembly was straightforward with minimal noise and nucleotide identity recorded from both the forward and reverse sequences of each sample.

4.1.4 Microsatellite Genetic Analyses

4.1.4.1 Band Size Scoring

Both GelJ and GelQuest normalises; firstly, each gel with the 3 lanes of DNA ladder on the gel. Subsequently, normalisation is across all the gels for each marker (APPENDIX D). The band sizes are in APPENDIX E. The band sizes were similar for most markers to the previous study (Chang *et al*., 2014) with slight differences for markers TM10, TM13, TM15 and TM21 (Table 4.2). Markers TM7 and TM23 differed from the previous study due to direct sequencing problem described in section 4.1.3.1 above. All the sequences (except TM7) had more than 90% similarity with sequences from Chang *et al*. (2014) using BLAST search.

Table 4.2: Band sizes of microsatellite markers used in this study compared to Chang *et al*. (2014).

No	Marker	This study		Previous study		BLAST
		Allele size	Repeat motif	Allele size	Repeat motif	score (%)
1	TM02	160-176	*$(AGAT)_{11}$	161-177	$(AGAT)_{11}$	100
2	TM03	192-212	*$(AGAT)_{8}$	186-194	$(AGAT)_{8}$	95.1
3	TM04	204-244	$(AGAT)_{14}$	191-215	$(AGAT)_{14}$	98.4
4	TM05	256-464	*$(AGAT)_{12}$	225-237	$(AGAT)_{12}$	100
5	TM07	376-408	$(AGAT)_{4}$	305-333	$(AGAT)_{12}$	79
6	TM08	412–420	$(AGAT)_{4}$	389-397	$(AGAT)_{6}$	97
7	TM10	508-536	$(AGAT)_{11}$	411-427	$(AGAT)_{10}$	98.4
8	TM11	452-468	$(AGAT)_{8}$	412-428	$(AGAT)_{10}$	91
9	TM12	504-520	$(AGAT)_{11}$	488-504	$(AGAT)_{14}$	96.6
10	TM13	552-564	*$(AGAT)_{8}$	438-466	$(AGAT)_{8}$	99.6
11	TM15	352-364	$(AGAT)_{6}$	470-526	$(AGAT)_{7}$	99.2
12	TM21	216-340	$(ACAG)_{12}$	470-490	$(ACAG)_{13}$	96.9
13	TM23	300-440	$(ACAG)_{29}$	207-275	$(ACAG)_{36}$	99.6
14	TM24	300-376	$(ACAG)_{23}$	286-439	$(ACAG)_{26}$	98.3

* Identical number of repeats

4.1.4.2 Identification of Genotyping Errors

MICRO-CHECKER detected that the TM13 had significant shortage of heterozygote genotypes in all populations under consideration with alleles of one repeat unit difference while TM7, TM10, TM23 and TM24 did not show any signs of homozygote deficit due to null allele (Table 4.3). When null alleles were present, it was due to homozygotes excess for most allele size classes with or without significant shortage of heterozygote genotypes with alleles of one repeat unit difference. When no null alleles were present, large allele dropouts or scoring error due to stuttering were also not detected (Table 4.3). Cluster C could not be analysed due to insufficient data.

Table 4.3: Results of MICRO-CHECKER program without considering region, following geographical regions and STRUCTURE clusters.

Locus	No region	Peninsular Malaysia	Cluster A (Sar)	Cluster B	Cluster D	Cluster E	Cluster F
TM2	+	+	-	-	-	-	-
TM3	+	+	+	-	-	+	+
TM4	+*	+*	-	-	-	-	-
TM5	+*	+*	-	-	-	+*	-
TM7	-	-	-	-	-	-	-
TM8	+*	+*	-	-	-	-	-
TM10	-	-	-	-	-	-	-
TM11	+*	+*	-	-	-	+*	+
TM12	+*	+*	+	+	-	+*	+
TM13	+*	+*	+*	-	+*	+*	+*
TM15	+*	+	-	-	-	-	-
TM21	+	+	-	-	-	-	-
TM23	-	-	-	-	-	-	-
TM24	-	-	-	-	-	-	-

+: presence of null alleles
*: highly significant shortage of heterozygote genotypes with alleles of one repeat unit difference.
Sar: Sarawak.

4.1.4.3 Population Structure Using Bayesian Approach

When STRUCTURE analysis was run without considering any population, six (K=6) clusters were detected, A- F. (Fig. 4.2 and 4.3). The inclusion or removal

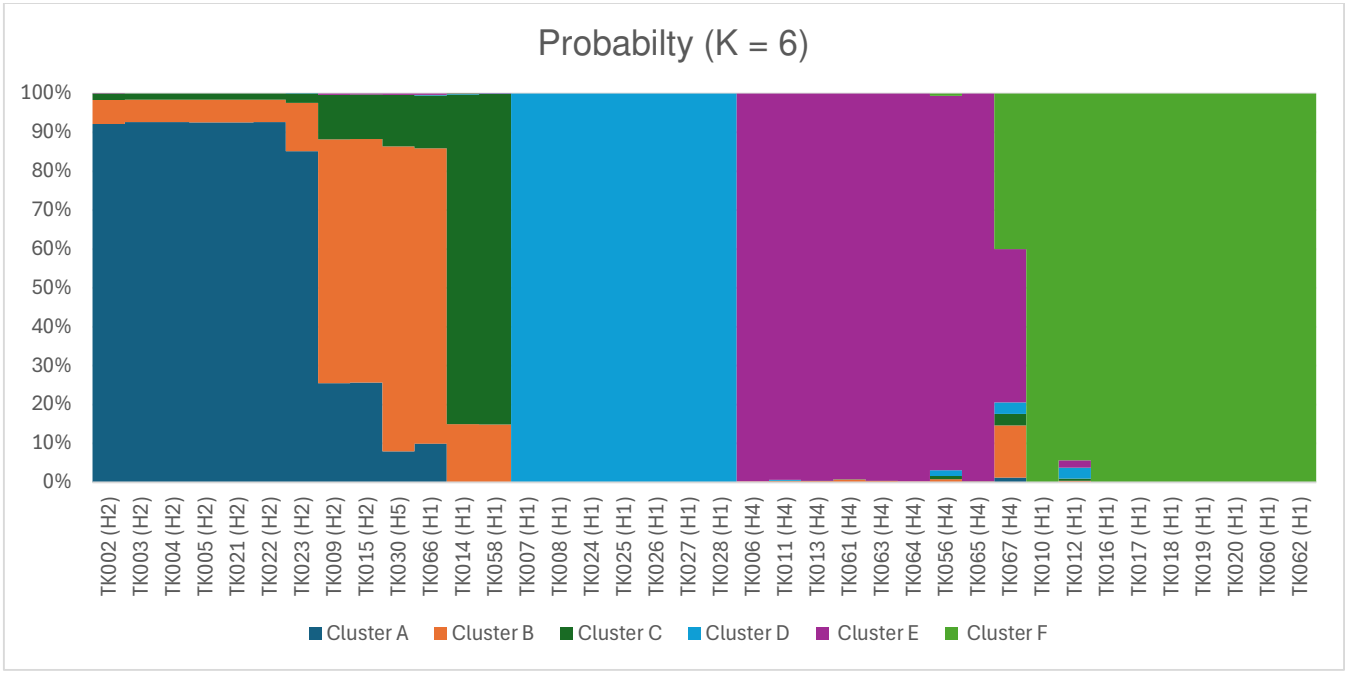


Fig. 4.2. Graphical HARVESTER results showing the clusters identified by STRUCTURE (Cluster A-F).
X axis: *T. schlegelii* individuals with their maternal lineage in brackets. Y axis: The membership percentage for clusters.

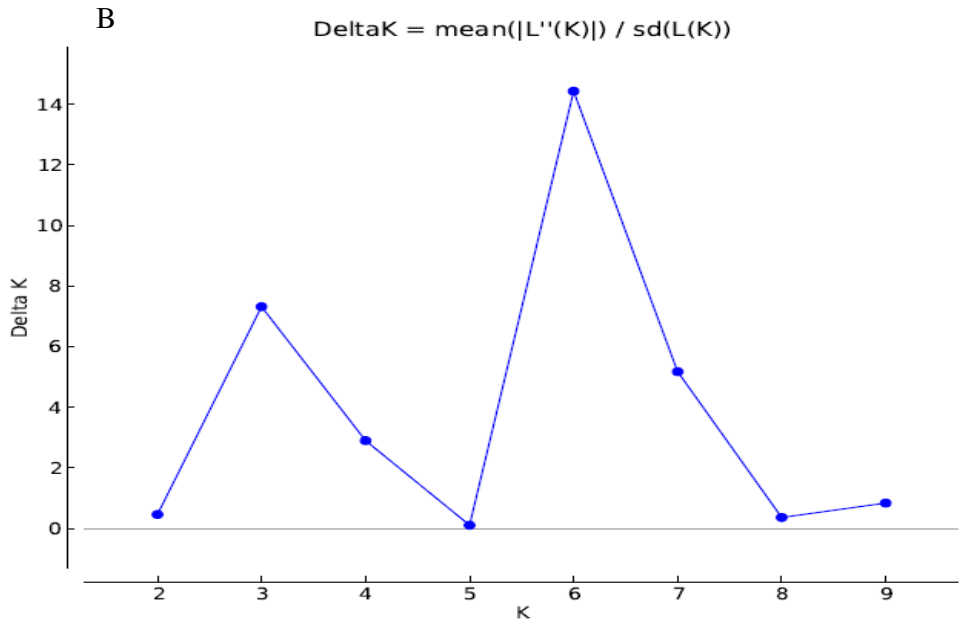


Fig 4.3. Evano's Delta *K* (ΔK) graph for STRUCTURE results. X axis: The best K. Y axis: Delta K.

of TM13 did not affect the results. Geographically, Cluster A and D consisted of only one geographical area/ state. Cluster A had individuals from Sarawak exclusively (Table 3.1). These individuals were sampled from Kuching (wild caught) and Miri (farm bred) in the previous study (Kaur *et al*., 2013). Cluster B comprised four individuals from Peninsular Malaysia whereby two of these were sampled from Selangor while two were of unknown origin. Cluster D were individuals from the known family unit (individual TK007, TK008, TK024, TK025, TK026, TK027, TK028) whose parents (TK007 and TK008) are from Selangor.

The individual TK058 from Perak and TK014 formed Cluster C. The latter was held in a pen with another adult female of similar size post capture. As microchip tagging of the two individuals (TK014 and TK016) occurred after a lapse of time, identification of the individuals from Terengganu and Perak was not possible due to similarities in size and physical characteristics. Therefore, it is uncertain whether Cluster C exclusively consists of individuals from Perak Cluster E consists of individuals from Selangor and Terengganu primarily, while Cluster F consisted of individuals from Selangor, Terengganu, Pahang, and of unknown origins. Except for cluster B, all other clusters comprised of individuals of only one mtDNA haplotype (Fig. 4.2).

4.1.4.4 Population Structure Using Distance Approach

The distance-based tree showed five main groupings (Fig. 4.4). Clade I and II were represented by the common Peninsular Malaysian haplotype, H1, as described by Kaur *et al*., (2013). Similar to STRUCTURE results, Clade III was

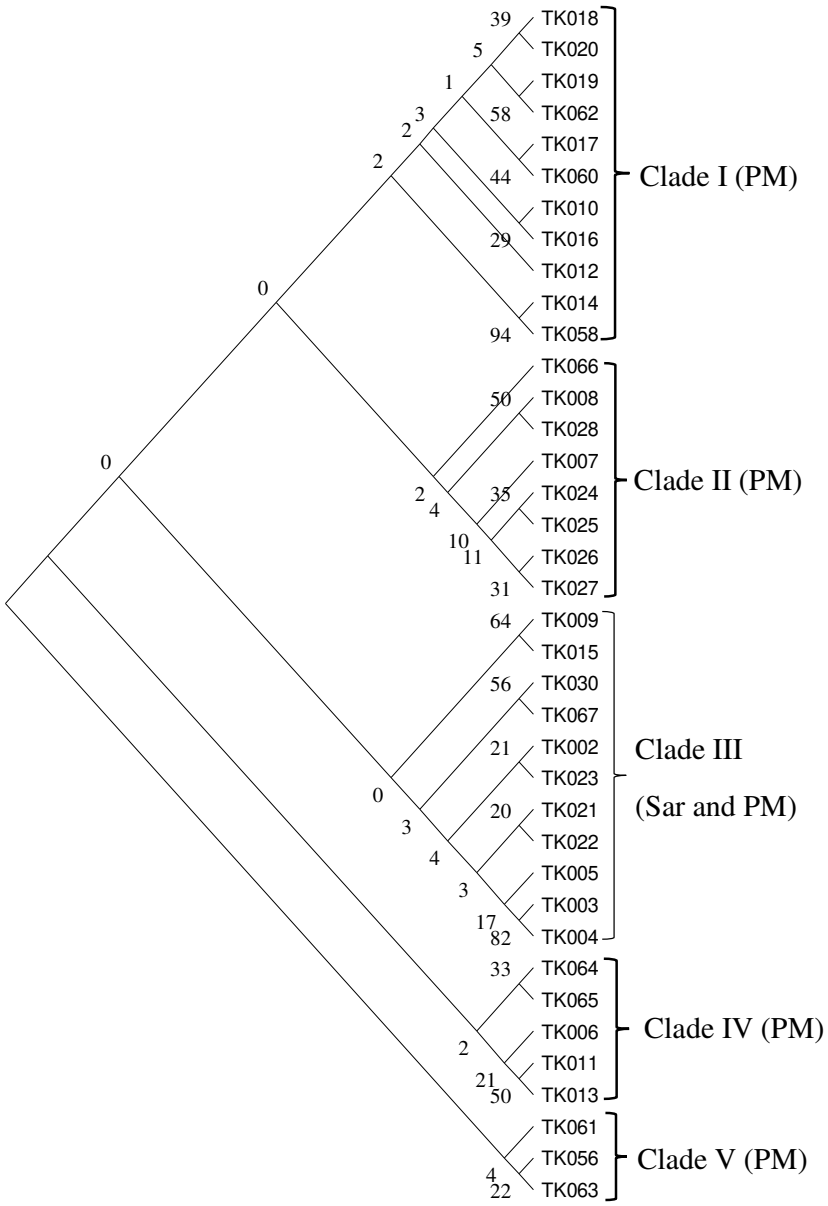


Fig.4.4. Neighbour Joining (NJ) tree based on Nei's D_a (1983) genetic distance for the microsatellite data. TK002-TK067: *T. schlegelii* samples. PM: Peninsular Malaysia. Sar: Sarawak. I-V: clades.

represented by haplotype H1, H2 and H5, while Clade IV and V were represented by the unique haplotype H4. Unlike the results from STRUCTURE, the distance-based analysis showed only Clade V consisted of individuals from

Selangor exclusively, while all other clades had a mix representation of states / geographical areas. Additionally, clades showed low bootstrap values.

4.1.4.5 Linkage Disequilibrium

For the STRUCTURE results, after correction using Fisher's method in Genepop, six pairs of loci showed linkage disequilibrium, namely locus TM11 with TM2 and TM3, locus TM12 with TM3 and TM5, locus TM10 with TM15 and locus TM4 with TM23 with p values ranging from 0.0015-0.0095. However, when using FSTAT, no linkage was seen. When using the Nei's distance-based results, after correction using Fisher's method, only two pairs were not linked namely the locus TM5 with TM11 and TM7 with TM23. Due to this and the low bootstrap values in the NJ tree, the STRUCTURE clustering were used for downstream analyses. Additionally, STRUCTURE results were differentiated into the Sarawak and Peninsular Malaysia region.

4.1.4.6 Microsatellite Genetic Diversity

With exception of loci TM8 and TM15, the Polymorphic Informative Content (PIC) was relatively high for all other loci indicating that the markers were informative (Table 4.4). FSTAT results, after randomisation, showed at least seven loci were in HWE namely TM2, TM7, TM10, TM15, TM21, TM23, T24. The locus TM3, TM5, TM12 and TM13 were not in HWE when overall and within sample randomisation was carried out. Significant shortages of heterozygotes were only seen in locus TM13 in all clusters. The inclusion and exclusion of TM13 in downstream analyses did not affect the results. TM23 had the highest allelic richness in all clusters except Cluster D (Table 4.4). TM24

Table 4.4: Allelic and genotypic diversities for each locus.

Locus	PIC	A_R	P_A	H_O		H_E		F_{IS}		HWE p-values	HWE p-values
			(in clusters)	Mean	SE	Mean	SE	Mean	SE	(*F*, overall samples)	(*f*, within samples)
TM2	0.640	2.585	B	0.460	0.156	0.467	0.129	0.009	0.145	0.0005>0.00050	0.7046>0.52400
TM3	0.637	2.569	-	0.152	0.117	0.513	0.042	0.704	0.233	0.0001	0.0001
TM4	0.774	3.034	A, B, C	0.243	0.064	0.408	0.111	0.403	0.023	0.0001	0.0001>0.00000
TM5	0.690	2.742	B, C, D	0.142	0.071	0.437	0.032	0.626	0.206	0.0001	0.0001
TM7	0.650	2.603	C	0.597	0.163	0.516	0.109	-0.171	0.236	0.0225>0.02250	0.5992>0.43250
TM8	0.257	1.563	A	0.048	0.048	0.189	0.122	0.781	0.126	0.0001	0.0021>0.00000
TM10	0.612	2.497	-	0.391	0.122	0.474	0.106	0.211	0.146	0.0005>0.00050	0.041 >0.01890
TM11	0.638	2.578	A, D, F	0.119	0.119	0.283	0.096	0.682	0.260	0.0001	0.0002>0.00000
TM12	0.660	2.636	-	0.048	0.030	0.440	0.121	0.737	0.190	0.0001	0.0001
TM13	0.408	1.961	C	0.000	0.000	0.460	0.028	1.000	0.000	0.0001	0.0001
TM15	0.261	1.570	-	0.159	0.083	0.240	0.102	0.349	0.187	0.0019>0.00190	0.0168>0.00260
TM21	0.650	2.616	B, C	0.516	0.053	0.565	0.042	0.067	0.101	0.0036>0.00360	0.1901>0.10260
TM23	0.893	3.516	A, B, D, E, F	0.901	0.046	0.773	0.037	-0.191	0.113	0.2311>0.23110	0.8044>0.62510
TM24	0.873	3.423	A, B, C, D, E, F	0.809	0.065	0.719	0.041	-0.139	0.106	0.01210>0.01210	0.56490>0.38810

PIC, Polymorphic Informative Content
A_R, allelic richness
P_A, Private alleles
H_O, Observed heterozygosity
H_E, Expected heterozygosity
F_{IS}, Inbreeding coefficient
HWE, Hardy Weinberg Equilibrium
SE: standard error

had the second highest allelic richness overall and in most clusters except Cluster D and Cluster E while TM8, TM13 and TM15 had the lowest allelic richness. These loci showed a similar trend for the H_O and H_E although the latter parameter in locus TM13 was comparable to the other loci. Both TM23 and TM24 had private alleles in almost all clusters. Locus TM3, TM10, TM12 and TM15 did not show private alleles in any clusters. The inbreeding coefficient, F_{IS}, was high indicating an excess of homozygotes as revealed by both GenAlEx and FSTAT in most loci. In locus TM13, the F_{IS} was 1 while locus TM7, TM23 and TM24 showed an excess in heterozygotes.

When considering mean values for the separate clusters, the highest mean for the number of alleles, N_A, and effective alleles, N_E, was seen in Cluster B, 3.714 ± 0.412 and 3.048 ± 0.348 respectively followed by Cluster A and F while Cluster C had the lowest N_A = 1.714 ± 0.221 and N_E = 1.643 ± 0.190 (Table 4.5). The mean for private alleles, P_A, was highest in Cluster A (0.643 ± 0.308) followed by Cluster B and Cluster C (0.429 ± 0.137) despite the highest number of loci (six) showing private alleles were in the latter. The lowest P_A was in Cluster E (0.214 ± 0.155). Summary of these is in the graphical chart as Fig.4.5.

The mean H_O for clusters was low ranging from 0.250 ± 0.144 to 0.49 ± 0.066 while H_E was low to moderate ranging from 0.277 ± 0.078 to 0.611 ± 0.044 with Cluster D having the highest H_O followed by Cluster B, F and E. The mean F_{IS} ranged from 0.07 ± 0.15 (Cluster D) to 0.53 ± 0.10 (Cluster B). With FSTAT, the F_{IS} ranged from 0.05-0.54. At the adjusted nominal of 5%, the p = 0.0006 was equal to the randomisations carried out for Cluster A, B, E but was not

Table 4.5: Allelic and genotypic diversities for each cluster.

CLUSTER	Diversity parameter	Mean	SE	Overall F_{IS}
A	N_A	3.214	0.585	0.45 ***
	N_E	2.429	0.43	
	H_O	0.276	0.07	
	H_E	0.449	0.072	
	F_{IS}	0.354	0.114	
B	N_A	3.714	0.412	0.543 ***
	N_E	3.048	0.348	
	H_O	0.329	0.08	
	H_E	0.611	0.044	
	F_{IS}	0.533	0.102	
C	N_A	1.714	0.221	0.147 ns
	N_E	1.643	0.19	
	H_O	0.25	0.114	
	H_E	0.277	0.078	
	F_{IS}	0.143	0.234	
D	N_A	2.786	0.261	0.046 ns
	N_E	2.193	0.215	
	H_O	0.49	0.107	
	H_E	0.475	0.058	
	F_{IS}	0.066	0.154	
E	N_A	3.071	0.45	0.454 ***
	N_E	2.385	0.3	
	H_O	0.304	0.086	
	H_E	0.507	0.057	
	F_{IS}	0.423	0.136	
F	N_A	3.286	0.597	0.329 ***
	N_E	2.423	0.393	
	H_O	0.317	0.101	
	H_E	0.459	0.071	
	F_{IS}	0.414	0.130	

N_A, number of alleles,
N_E, number of effective alleles,
F_{IS} (ns = not significant, * P<0.05, ** P<0.01, *** P<0.001)
SE: standard error

significant in Cluster C and D (p = 0.0006 < 0.1048 and 0.3119 respectively).

Only Cluster C and D were not significant at the 1% nominal value (p = 0.0006 > 0.0001 respectively) suggesting inbreeding in clusters A, B, E and F.

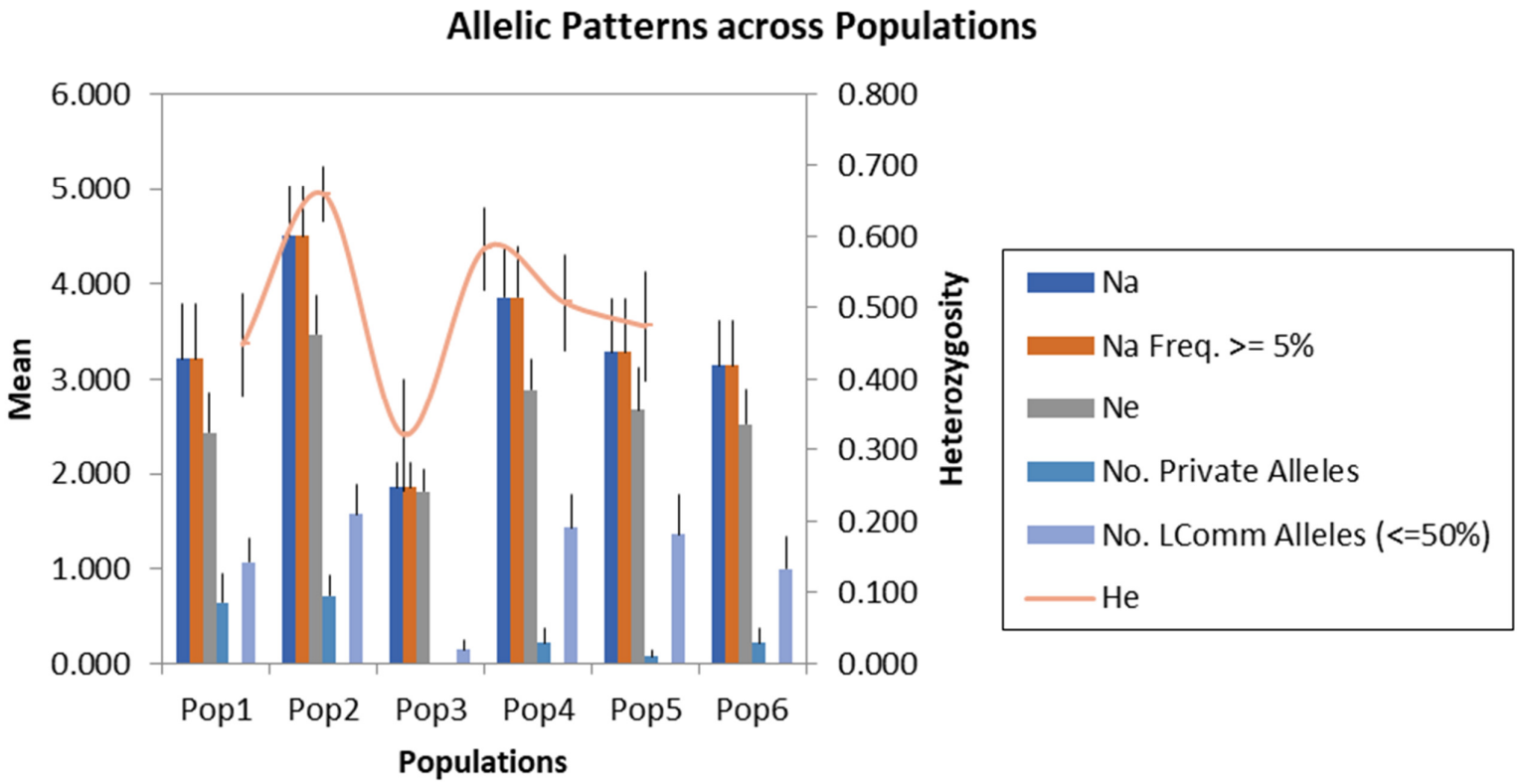


Fig. 4.5. Chart of allelic pattern across clusters. Na: Number of alleles. Na Freq. >=5%: number of different alleles having a frequency of 5% or more. Ne: number or effective alleles. No. Private Alleles: Number of private alleles. No. LComm Alleles (<=50%): locally common alleles occurring in 50% or less of the populations. He: heterozygosity.

4.1.4.7 Microsatellite Genetic Differentiation

The theta (Θ) in the F-Statistic was 0.24 (0.24 ±0.043) while the *F* and *f* were 0.52 (0.518 ±0.094) and 0.368 (0.364 ±0.112), respectively, indicating a moderate cluster differentiation. The pairwise F_{ST} supported the differentiation between clusters except for Cluster C which was not significantly differentiated with all other clusters, and Cluster B with Cluster E after randomisation (Table 4.6). The AMOVA analyses showed moderate (when H4 haplotype / Cluster E not considered) to substantial differentiation (when H4 haplotype / Cluster E was considered) (Table 4.7). The six clusters showed that the among-group variation was 30% and the R_{ST} was significant (p value < 0.01) indicating that clusters differed from each other. The within population variation was just 7% and the non-significant R_{IS} indicated individuals within the clusters were similar with low variation but were significantly different when compared to the total population (R_{IT} = 0.38, p value < 0.01).

Table 4.6: The pairwise F_{ST} when considering six clusters (below diagonal) and their significance (above diagonal).

Cluster	A	B	C	D	E	F
A	0.000	**0.00067**	0.02533^{NS}	**0.00067**	**0.00133**	**0.00200**
B	0.1482	0.000	0.04800^{NS}	**0.00267**	0.00800^{NS}	**0.00200**
C	0.4288	0.2086	0.000	0.02400^{NS}	0.02533^{NS}	0.01667^{NS}
D	0.3157	0.1767	0.3326	0.000	**0.00067**	**0.00067**
E	0.2792	0.0951	0.2825	0.1843	0.000	**0.00067**
F	0.3342	0.2478	0.2705	0.2045	0.206	0.000

*Adjusted nominal level (5%) for multiple comparisons is 0.003333.
Bold font: significant F_{ST}.
NS: not significant.

Table 4.7: AMOVA and its variance according to clusters.

Source	df	SS	MS	Est. Var.	%	R-Statistics	Value	p-value
All six clusters								
Among clusters	5	62571.790	12514.358	842.350	30%	Rst	**0.305**	0.001
Within clusters	32	67917.315	2122.416	198.445	7%	Ris	0.103	0.089
Within Indiv	38	65570.000	1725.526	1725.526	62%	Rit	**0.376**	0.001
Total	75	196059.105		2766.321	100%			
Without H4 / cluster E								
Among clusters	4	48068.969	12017.242	828.303	28%	Rst	**0.276**	0.001
Within clusters	25	61603.565	2464.143	295.538	10%	Ris	0.136	0.059
Within Indiv	30	56192.000	1873.067	1873.067	62%	Rit	**0.375**	0.001
Total	59	165864.533		2996.908	100			

df: degree of freedom.
SS: Sum of Squares.
MS: Mean Squares.
Est. Var.: Estimated Variance.
Bold font: significant parameters.

4.1.4.8 Sex-Biased Dispersal

The assignment index distribution and its nonparametric Mann-Whitney U test indicated a non-significant finding. There were more females with negative corrected assignment index (AIc) but the lower value of mean AIc (mAIc) was seen in the male group led to inconclusive findings (APPENDIX F).

4.1.4.9 Relatedness

The relat, based on Queller and Goodnight (1989) which is the average relatedness within cluster compared to the whole data set was 0.317 ± 0.05. The ML Relate uncovered one PO (TK06 and TK61) and 38 FS relationships (Table 4.8). These relationships were between individuals within the clusters. All relationships were seen within the states except for individual TK017 from Pahang in Cluster F which had a FS relationship with two cluster members from Selangor (TK010 and TK060). Similarly, TK064 from Terengganu in Cluster E had a FS relationship with two cluster members from Selangor (TK011 and TK013). The FS relationships for individuals of unknown origins are listed in Table 4.8 as well.

A higher weighted molecular coancestry was observed in clusters A, C and F (0.543, 0.687 and 0.507 respectively) compared to individuals from other clusters (Table 4.9). Meanwhile, Cluster B and E had the lowest weighted molecular coancestry (0.362 and 0.441 respectively).

Table 4.8. The FS relationship obtained when analysis carried out without specifying clusters.

CLUSTER	INDIVIDUAL (FS)		States
A	TK004	TK003	Sarawak-Sarawak
	TK005	TK002	Sarawak-Sarawak
	TK005	TK003	Sarawak-Sarawak
	TK005	TK004	Sarawak-Sarawak
	TK021	TK004	Sarawak-Sarawak
	TK021	TK005	Sarawak-Sarawak
	TK022	TK021	Sarawak-Sarawak
	TK023	TK021	Sarawak-Sarawak
B	TK015	TK009	Selangor-Selangor
C	TK058	TK014	Sungkai-Sungkai / Terengganu
D	TK024	TK008	Selangor-Selangor
	TK025	TK024	Hatchling from TK008
	TK026	TK024	Hatchling from TK008
	TK026	TK025	Hatchling from TK008
	TK027	TK008	Hatchling and maternal parent
	TK027	TK024	Hatchling from TK008
	TK027	TK025	Hatchling from TK008
	TK027	TK026	Hatchling from TK008
	TK028	TK008	Hatchling and maternal parent
	TK028	TK024	Hatchling from TK008
	TK028	TK027	Hatchling from TK008
E	TK063	TK011	Kulai / Zoo Melaka-Selangor
	TK64	TK011	Tekah River (Terengganu)-Selangor
	TK64	TK013	Tekah River (Terengganu)-Selangor
	TK64	TK061	Tekah River (Terengganu)-Pulau Pinang / Setiu (Terengganu)
	TK64	TK063	Tekah River (Terengganu)-Kulai / Zoo Melaka
	TK65	TK061	Kelantan-Pulau Pinang / Setiu
	TK65	TK063	Kelantan-Kulai / Zoo Melaka
	TK65	TK064	Tekah River (Terengganu)-Kelantan
F	TK060	TK017	Dusun River (Selangor)-Pahang
	TK062	TK017	Pulau Pinang / Setiu-Pahang
	TK062	TK019	Pulau Pinang / Setiu-Pahang
	TK062	TK060	Pulau Pinang / Setiu-Dusun River (Selangor)
	TK016	TK010	Sungkai / Terengganu-Selangor
	TK017	TK010	Pahang-Selangor
	TK019	TK017	Pahang-Pahang
	TK020	TK016	Pahang-Sungkai / Terengganu
	TK020	TK018	Pahang-Pahang

4.1.4.10 Gain/Loss

The gain/loss analysis showed that when removing Cluster B and D, the contribution to the Nei's (1987) gene diversity at all the three levels (the within and between cluster diversity and across the clusters) decreased, rendering them more important for the diversity of the population from the nuclear gene perspective (Table 4.10). Cluster D showed a decrease of 0.003% for within cluster to 0.1% for total loss of diversity, while a larger magnitude of contribution was seen in Cluster B (0.1% for the between clusters, 3.1% and 3.2% for the internal and total diversities).

At the allelic level, clusters A, C and D showed that removal of these would result in the loss of allelic richness (based on Petit, Mousadik and Pons (1998) internal diversity) at the between cluster and total population level while the internal allelic diversity is not lost (a negative value). Exclusion of Cluster B would result at the within cluster and total population allelic loss. Removal of Cluster E on the other hand would result in the loss of allelic richness at all levels.

Table 4.9: Weighted Molecular Coancestry (in bold) and Weighted Coancestry Matrix among the clusters.

CLUSTERS	A	B	C	D	E	F
A	**0.543044**					
B	0.2646223	**0.3622322**				
C	0.113048	0.1634546	**0.687411**			
D	0.2177017	0.2177742	0.214458	**0.4843794**		
E	0.2209655	0.2553293	0.1938793	0.2714653	**0.4408691**	
F	0.2344289	0.1788382	0.2986401	0.3072755	0.2705331	**0.5072786**

Table 4.10: Gain and Loss analyses on the genetic diversities of each cluster when a cluster is removed.

Clusters	GD	GD_W	GD_B	Loss/Gain	Pe_Int Diversity	Pe Divergence	Petit (4) (%)
A	0.6349413	0.8955303	-4.598048	-3.702518	-0.163525	3.428425	3.2649
B	0.6384625	***-3.066417***	***-0.1020529***	***-3.16847***	5.851453	-2.182212	3.669241
C	0.6526951	1.671192	-2.681099	-1.009906	-4.771122	6.547727	1.776604
D	0.658726	***-0.003069047***	***-0.09216668***	***-0.09523572***	-0.2958438	0.3562465	0.06040268
E	0.6575397	-1.231562	0.9563967	-0.275165	***0.4397821***	***0.09342926***	***0.5332114***
F	0.6679124	1.734806	-0.4368106	1.297996	-1.060744	-0.4614582	-1.522202

GD: Nei's gene diversity after excluding the cluster.
GD: Nei's gene diversity after excluding the cluster.
GD_W: contribution at the within cluster level.
GD_B: contribution at the between cluster level.
Loss/Gain: total contribution to GD.
Pe_Int Diversity: contribution to within-cluster allelic diversity after rarefaction.
Pe Divergence: contribution to between-cluster allelic diversity after rarefaction.
Petit (4) (%): total contribution to allelic diversity after rarefaction to 4 copies.
Bold and italics font: important cluster at all levels for the gene and allelic diversity.

4.1.5 mtDNA Genetic Analyses

4.1.5.1 Genetic Diversities

The total length for the control region (CR) fragment was 874 bp while the ND 6-tRNA^{glu}-cyt *b* was 428 bp resulting in the concatenated data of 1,302 bp length. A total of 32 parsimony variable sites were detected in the ten new samples which consisted of the H1 and H4 haplotypes (Table 4.11).

The nucleotide diversity for the samples in this study was $\pi = 0.00538 \pm 0.00627$ and haplotype diversity, $h = 0.8 \pm 0.02688$ for the whole of Peninsular Malaysia. Only Cluster B showed genetic diversities with $\pi = 0.00813 \pm 0.00126$ and $h = 0.553 \pm 0.073$. Cluster B comprised of individuals from three mtDNA lineages, the H1 (1 individual), H5 (1 individual) and H2 (2 individuals) haplotypes. Clusters C, D and F comprised of individuals with mtDNA haplotype H1, the common Peninsular Malaysian lineage, as reported by Kaur *et al*. (2013). Cluster A consisted of the common Sarawak mtDNA haplotype, H2. Similarly, Cluster E consisted exclusively of individuals of H4 lineage.

Comparison over a 451 bp of the D-loop in Kurniati (2003) found that the haplotype A, C and D of that study were similar to haplotype H5 / H6, haplotype H1 / H3 and haplotype H2 respectively (Fig. 4.6). The haplotype B (Mapan) from Kurniati (2003) differed from all other haplotypes with a transversion at the 15522 bp of the mtDNA genome (in Domain II of the D-loop). This single transversion is the only variation between the Mapan haplotype and the H5 and H6 haplotypes.

Table 4.11: List of haplotypes according to STRUCTURE clusters, mtDNA haplotype diversity (h) and diversity (π) within clusters and geographical distribution.

Cluster	Individuals	Haplotype	mtDNA ($h \pm$ SD, $\pi \pm$ SD)	Region
A	TK 002	H2	-	
	TK 003	H2		Sarawak
	TK 004	H2		
	TK 005	H2		
	TK 021	H2		
	TK 022	H2		
	TK 023	H2		
B	TK 009	H2	0.553 ± 0.073,	Peninsular
	TK 015	H2	0.00813 ± 0.00126	Malaysia
	TK 030	H5		
	*TK066	H1		
C	TK 014	H1	-	Peninsular
	*TK058	H1		Malaysia
D	TK 007	H1	-	Peninsular
	TK 008	H1		Malaysia
	TK 024	H1		
	TK 025	H1		
	TK 026	H1		
	TK 027	H1		
	TK 028	H1		
E	TK 006	H4	-	Peninsular
	TK 011	H4		Malaysia
	TK 013	H4		
	*TK056	H4		
	*TK061	H4		
	*TK063	H4		
	*TK064	H4		
	*TK065	H4		
F	TK 010	H1	-	Peninsular
	TK 012	H1		Malaysia
	TK 016	H1		
	TK 017	H1		
	TK 018	H1		
	TK 019	H1		
	TK 020	H1		
	*TK060	H1		
	*TK062	H1		
	*TK067	H1		

* Samples collected in this study.
SD: standard deviations.

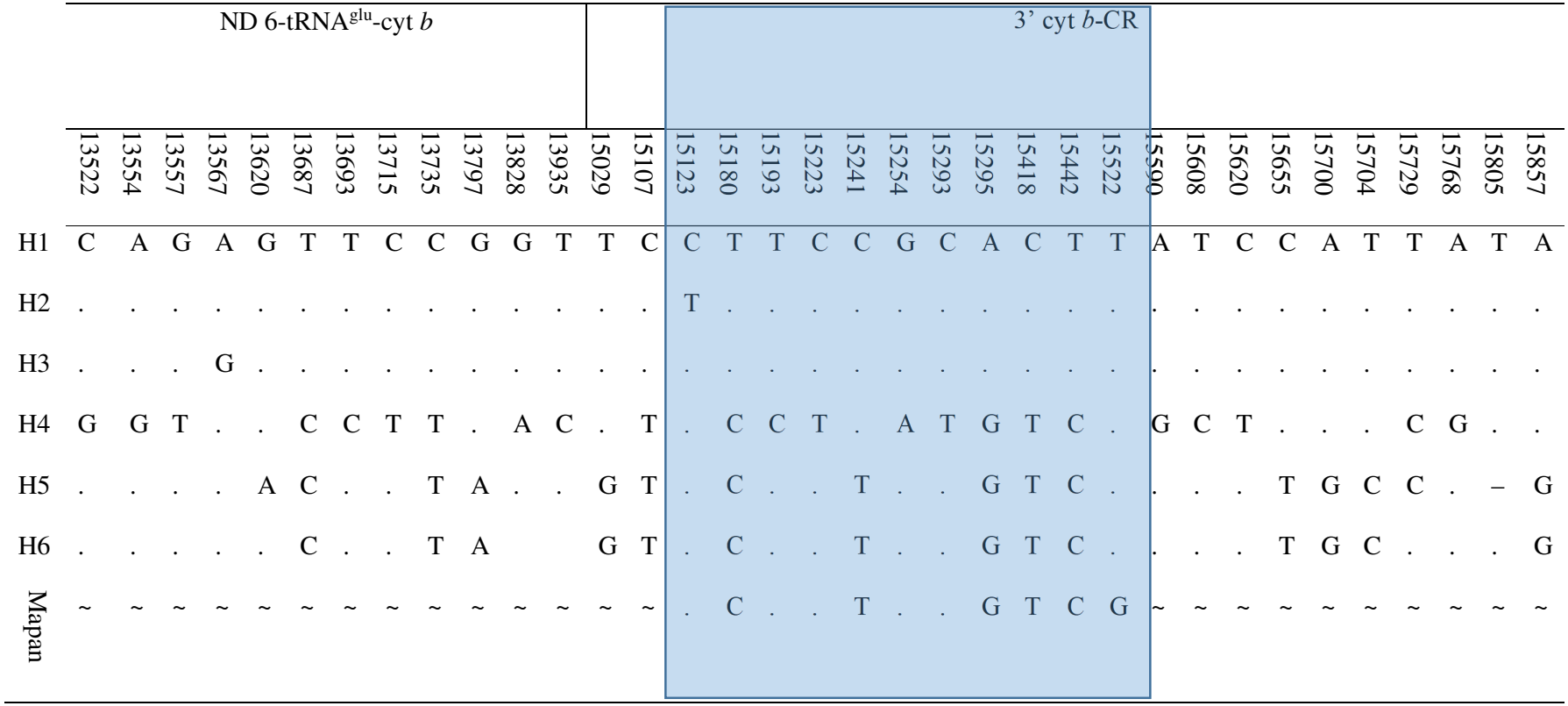


Fig. 4.6. Schematic diagram indicating the polymorphic sites obtained using mtDNA concatenated data from the present dataset and Kaur *et al*. (2013). Blue highlight: The 451 bp of the CR region by Kurniati (2003). H1-H6: The six haplotypes described by Kaur *et al*. (2013). Mapan: The haplotype B from Kurniati (2003) labelled here as 'Mapan', the location of origin of haplotype B.

The H1 and H3 are conserved at the CR fragment with a single transition seen in H2 haplotype. At the ND 6-tRNA^{glu}-cyt *b* fragment, the H1 and H2 haplotype were conserved but had a single transition with haplotype H3. The haplotype described by Md Adzhar *et al*. (2017) was identified as the H2 while the GenBank sequence MT554040 was the H1 haplotype.

4.1.5.2 Population Structure

The mtDNA data showed a substantial population differentiation using AMOVA when considering all adults, with more than 90% of variation seen between cluster (PhiPT = 0.91, p value < 0.01) (Table 4.12). However, no differentiation was seen when the H4 haplotype was not considered.

4.1.5.3 Sex-Biased Dispersal

When females only were considered for the AMOVA analysis, a significant PhiPT indicated female philopatry (Table 4.12). When the Cluster E (H4 individuals) were excluded in the AMOVA analyses, the PhiPT was not significant indicating dispersal. The AMOVA could not be carried out for the males in the clusters as only Cluster B had more than one male.

4.1.5.4 Haplotype Network

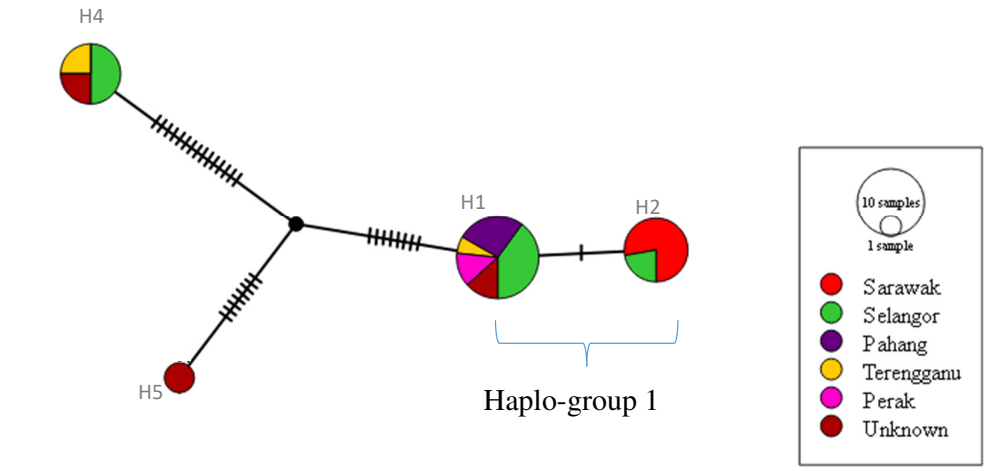
The median joining network generated with concatenated dataset and the 451 bp of D-loop are similar with three genealogical relationships (Fig 4.7). Both the networks placed the H4 node (H4 haplotype) separately from all other nodes. Meanwhile, the H1 and H2 haplotypes were on the same evolutionary connection in both networks. For the D-loop network, the node labelled H2

Table 4.12: AMOVA and its variance according to populations and clusters.

Source	df	SS	MS	Est. Var.	%	PhiPT	Value	p-value
(Adults – six clusters)								
Among Pops	5	150.095	30.019	4.792	91%	PhiPT	**0.912**	0.001
Within Pops	32	14.800	0.463	0.463	9%			
Total	37	164.895		5.254	100%			
(only females – six clusters)								
Among Pops	4	124.296	31.074	6.001	92%	PhiPT	**0.921**	0.001
Within Pops	22	11.333	0.515	0.515	8%			
Total	26	135.630		6.516	100%			
(adults - without H4)								
Among Pops	4	7.933	1.983	0.241	29%	PhiPT	0.290	0.074
Within Pops	25	14.8	0.592	0.592	71%			
Total	29	22.73		0.833	100%			
(only females, without H4)								
Among Pops	3	3.287	1.096	0.059	7%	PhiPT	0.070	0.087
Within Pops	19	14.8	0.779	0.779	93%			
Total	22	18.09		0.838	100%			

df: degree of freedom.
SS: Sum of Squares.
MS: Mean Squares.
Est. Var.: Estimated Variance.
PhiPT: Fixation index.
Bold font: Significant parameters.

(A)



(B)

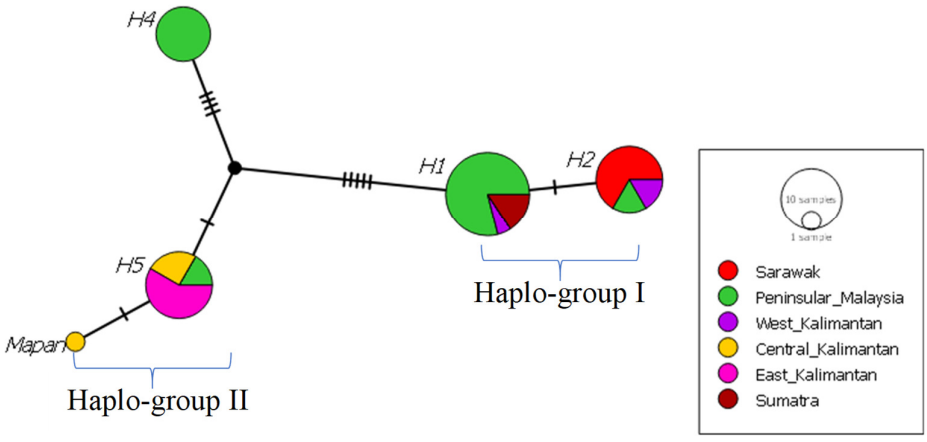


Fig. 4.7. Median Joining Network. (A) Network using concatenated mtDNA data (1,302bp) with the samples in this study (N=40) (B) Network using a 451bp of the mtDNA D-loop (N=52). Node sizes reflects the number of individuals. Node colour represents geographical locations. Mutational steps are represented as edges on the branches.

comprised individuals with the H2 and H3 lineage as they are conserved over the 451 bp (Fig.4.6). Due to being different by only 1 bp with haplotype H1 and on the same evolutionary branch, these lineages form a haplogroup (Haplo-group I) separated from other haplotypes. Haplo-group I comprised individuals from Sumatra, Peninsular Malaysia, Sarawak and west Kalimantan.

Similarly, H5 haplotype was also on a separate evolutionary trajectory. In the D-loop network, the node labelled H5 comprised individuals with the H5 and H6 lineages as they were also conserved for over the 451 bp. The separation of these with Mapan haplotype was also by 1 bp and formed a separate haplogroup (Haplo-group II). The Haplo-group II comprised Peninsular Malaysia, central and east Kalimantan individuals. The H4 node was on a separate evolutionary branch differentiated from the other two haplogroups by 15 bp for the concatenated dataset, and 4 bp for the D-loop dataset.

4.1.5.5 Best-Fit Models of Nucleotide Substitutions

Both MEGA and jModelTest identified the HKY model as the best fit (Table 4.13). However, jModelTest also indicated HKY + I (0.86) for the cyt *b*-CR marker using AICc criterion. Since the BIC is a more stringent criterion and all other analyses supported the HKY model, the HKY model was ultimately used for downstream analyses.

Table 4.13: The Best-Fit Models of Nucleotide Substitutions identified using individual and concatenated datasets.

Software	Dataset	Model	Parameter	BIC	AICc	lnL	Invariant
MEGA	ND 6-tRNA^{glu}-cyt *b*	HKY	69	1922.578816	1399.847474	-630.5902194	-
	cyt *b*-CR	HKY	137	4265.766	3034.981	-1380.170	-
	Concatenated data	HKY	69	4622.017179	4022.075672	-1941.92855	-
jModetltest	ND 6-tRNA^{glu}-cyt *b*	HKY	136	2103.5351	1679.5494	639.7472	-
	cyt *b*-CR	HKY+1	135	-	2937.1714	1308.77	0.86
		HKY	134	3528.5819	-	1310.49	-
	Concatenated data	HKY	70	4385.8147	4031.8734	1941.8993	-

HKY: Hasegawa-Kishino-Yano.
BIC: Bayesian Information Criterion
AICc: Akaike Information Criterion, corrected
lnL: Maximum Likelihood value

4.1.5.6 The Maximum Likelihood (ML) Tree

The Maximum Likelihood tree showed that *T. schlegelii* haplotypes were separated with strong bootstrap confidence (98%). The H4 formed a clade of its own from the other lineages. While the H5 and H6 haplotype clustered into a clade of their own, they were place on separate branches with strong bootstrap (98%). Similarly, the clade consisting of H1, H2 and H3 lineages also placed the haplotypes on separate branches with high values. All the individuals with H6 lineage were sampled in Kalimantan while the H5 lineage in Peninsular Malaysia. While haplotype H2 is represented by Peninsular Malaysia and Sarawak individuals, the H1 and H3 lineages consisted of Peninsular Malaysia and Sumatra individuals respectively.

4.1.5.7 Divergence Time Tree

The divergence time detected the three main clades (one composing only H4 lineage, the second comprising H5 and H6 and the third comprising H1, H2 and H3 lineages) separated from pre-Early Pliocene Epoch. The H4 lineage diverged at approximately 6.4 million years ago (Mya) (Fig. 4.9) while the H1, H2 and H3 haplotypes diverged from the H5 and H6 approximately at 4.84 Mya. Within clades, all three showed that the divergence between H5 with H6; and among the H1, H2 and H3 lineages was within the Pleistocene era (1.3-1.8 Mya). The divergence time shown here are estimates as they are estimated on sequences of living crocodilians without any fossil / morphological calibration and hence are only to indicate the relative divergence of the haplotypes.

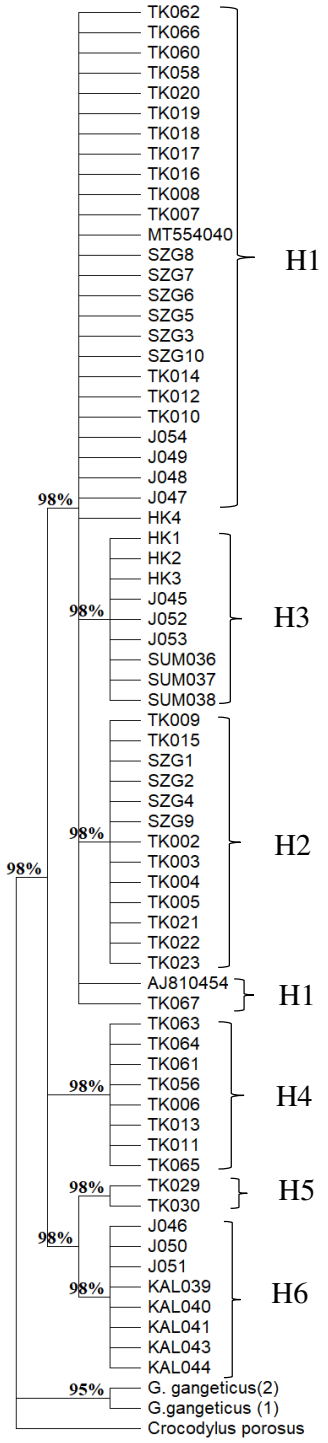


Fig. 4.8. The Maximum likelihood tree generated using concatenated mtDNA sequences. H1-H6: haplotype H1-H6. Samples are marked indicating origins; TK002-TK067: Malaysia. KAL: east Kalimantan. J: Java. SUM: Sumatra. Samples of unknown origins labelled following holding localities; SZG: Singapore Zoo. HK: Hong Kong Wildlife Park. MT554040 and AJ810454 are *T. schlegelii* sequences from GenBank. G. *gangeticus* and *Crocodylus porosus* are outgroups listed in Table 3.8. Percentage: Bootstrap confidence.

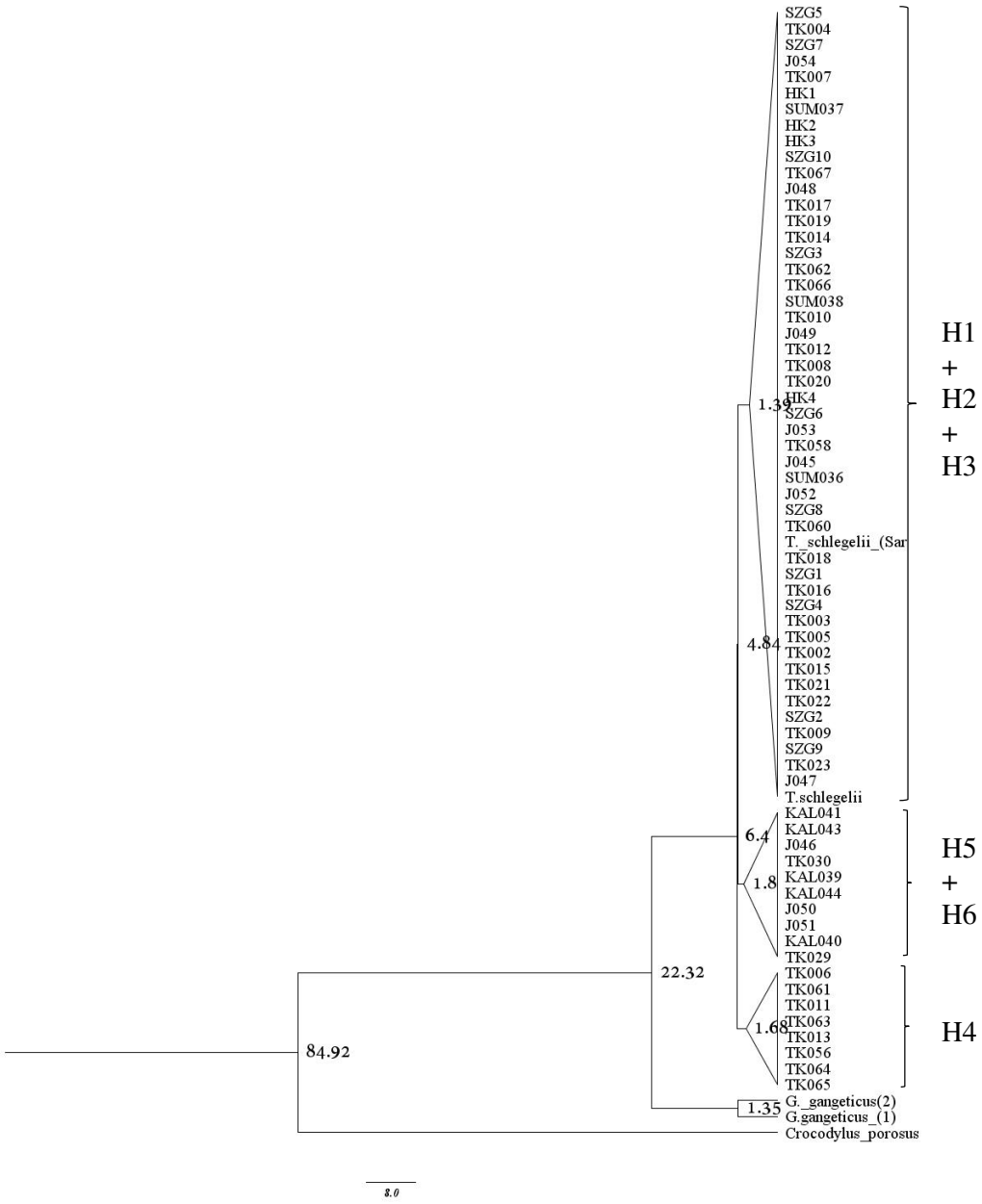


Fig. 4.9. The cladogram showing divergence time of the *T. schlegelii* haplotypes at the respective nodes. H1-H6: haplotype H1-H6. Samples are marked indicating origins; TK002-TK067: Malaysia. KAL: east Kalimantan. J: Java. SUM: Sumatra. Samples of unknown origins labelled following holding localities; SZG: Singapore Zoo. HK: Hong Kong Wildlife Park. MT554040 and AJ810454 are *T. schlegelii* sequences from GenBank. G. *gangeticus* and *Crocodylus porosus* are outgroups listed in Table 3.8. Divergence time indicated on branches. Scale bar: Time in million year ago, Mya.

4.2 Field Study

4.2.1 Day and Night Surveys

The number of sampling sites and total length of rivers surveyed are listed in Table 4.14. The Nerus and its tributary, the Tekah River, were the only tidal rivers in this study. The riverbanks of the Sungkai and Jengka Rivers were similar with very steep banks ranging from 2-5 m high (wet and dry seasons) (Fig. 10A) and clearer waters during the dry season as compared to murky waters in the rainy season (Fig 4.10B). The Selangor and Terengganu rivers had gradual banks with water colour being constant between the two seasons surveyed. The former two were also more difficult to survey compared to rivers in Terengganu and Selangor. This is due to river obstructions mainly due to rocks and fallen trees (Fig. 4.11). Only the Tengi River was slightly darker coloured compared to the rest.

Table 4.14: Length of rivers surveyed and the number of sampling sites.

Rivers / tributaries	Length (approx., km)	No. of sampling sites	Survey length (km)
Nerus	105	6	6.9
Tekah	2	2	0.2
Sungkai	81.6	7	17.44
Dusun / Ban Canal	15	5	15
Tengi	43	10	20.45
Jengka	84	10	2.5

Except for Sungkai River, the main fishing methods were by means of nets, “bubu” (traditional rattan fish trap) and sampan (a small boat) (Fig. 4.12A). The fishing in Sungkai River is done using fishing lines and nets (Fig. 4.12B). There is also livestock farming seen in Dusun River. A poultry and cattle farm were opened at the upper and lower reaches of the Ban Canal, respectively (Fig. 4.12C).

Fig. 4.10. Photos of river banks and water type. (A) Steep river banks of the Sungkai River in the dry and wet season. (B) Jengka River turns murky in the wet season.

Fig. 4.11. River obstructions and colouration. (A) Rocks at the river base during the dry season in the Jengka River. (B) Fallen trees in both dry and wet seasons. (C) The dark colouration of the Tengi River compared to the clear water of the Jengka River in (A). (D) Smaller river of lower pH draining into the Tengi River.

Fig. 4.12. Human activities in the rivers surveyed. (A) The traditional fish trap “bubu” destroyed in what the fishermen believed was due to a crocodile in Tengi River. (B) Fishing net used in all four rivers. (C) Cattle left to forage along the Ban Canal, also shows a low slope of the riverbank.

The final data and sampling sites presented here are the sites where three sampling replicates were achieved in both season (APPENDIX G). For Sungkai River, one site with just one sampling replicate in the dry month was also included to increase the sampling sites for this river. Similarly, in Jengka River, three sites with just two replicates in the dry months were included for the same purpose.

No *Tomistoma* was spotted in all the rivers during the night surveys. In the Nerus River, a sub adult *C. porosus* was identified at (5.322472, 103.042778) in both seasons.

4.2.2 Statistical Tests

4.2.2.1 Descriptive Analyses

The Tengi and Tekah rivers were the most acidic (pH 5.4-5.7) followed by the Nerus River. The Jengka River was the closest to neutral (pH 7-7.5) followed by the Ban Canal and Sungkai River (Table 4.15). The pH for the Jengka River was not acidic for both seasons, while the Tengi River was acidic in both seasons without much difference between the readings (pH 5.7). The Ban Canal was slightly acidic for both seasons while the Sungkai River was slightly acidic only in the dry season (6.69 ± 0.057). The Terengganu rivers were more acidic in the rainy months (6.1-6.6 in dry months versus 5.3-5.9 on the wet months). Despite

Table 4.15: Mean and standard error of each water parameter for the two seasons according to rivers.

River	Season	Temperature (°C)	pH	Conductivity	TDS	Salinity
	Dry	30.653 ± 0.513	7.520 ± 0.091	34.445 ± 0.763	22.558± 0.423	21.637± 0.488
Jengka, Pahang	Wet	29.670 ± 0.172	7.028 ± 0.011	35.117± 0.660	22.860± 0.360	22.477± 0.436
Sungkai, Perak	Dry	28.393 ± 0.371	6.69 ± 0.057	34.933 ± 0.582	24.813 ± 0.421	24.507 ± 0.27
	Wet	30.047 ± 0.451	7.034 ± 0.111	33.053 ± 0.537	23.3 ± 0.411	24.027 ± 0.269
Ban Canal, Selangor	Dry	29.05 ± 0.151	6.77 5 ± 0.278	39.925 ± 5.714	28.333 ± 4.067	26.75 ± 2.339
	Wet	28.75 ± 0.127	6.702 ± 0.123	50.75 ± 13.951	35.992 ± 9.859	31.292 ± 5.881
Tengi, Selangor	Dry	29.697 ± 0.403	5.723 ± 0.286	37.073 ± 1.993	26.337 ± 1.422	25.753 ± 0.732
	Wet	28.83 ± 0.291	5.696 ± 0.324	42.313 ± 2.2	30.043 ± 1.559	27.687 ± 0.877
Nerus, Terengganu	Dry	30.142 ± 0.226	6.281 ± 0.084	25.663 ± 0.253	18.2 ± 0.166	21.767 ± 0.584
	Wet	27.446 ± 0.261	5.881 ± 0.053	24.721 ± 1.111	17.55 ± 0.814	20.079 ± 0.447
Tekah, Terengganu	Dry	29.322 ± 0.16	6.153 ± 0.055	27.244 ± 1.258	19.433 ± 0.925	21.767 ± 0.517
	Wet	27.567 ± 0.084	5.383 ± 0.266	24.111 ± 0.144	17.144 ± 0.08	20.033 ± 0.058

TDS: Total dissolved solids.

being tidal rivers, the Nerus and Tekah rivers had similar salinity to the other rivers. The pH, temperature and salinity of the capture sites of *Tomistoma* in the Tekah River were also similar for the sighting site of *C. porosus* in the Nerus River (Appendix G).

4.2.2.2 Inferential Analyses (Paired t-test)

Most of the water parameter data were not significant ($p > 0.05$) when tested with Shapiro Wilks test, except for the Ban Canal and Tengi River in which the conductivity, TDS and salinity deviated from normality (Table 4.16). The Ban Canal was the only location that did not show a significant p value for the t-test and therefore results are not shown below.

Table 4.16: List of significant parameters after the Shapiro-Wilks (w) test.

Location	Measure 1-Measure 2	w	p
Sungkai, Perak	Temp_D - Temp_W	0.942	0.679
Jengka, Pahang	Temp_D - Temp_W	0.905	0.25
	pH_D - pH_W	0.81	0.019
Tekah, Terengganu	Temp_D - Temp_W	0.98	0.726
Nerus, Terengganu	Temp_D - Temp_W	0.886	0.214
	pH_D - pH_W	0.859	0.118
	Salinity_D - Salinity_W	0.928	0.502
Tengi, Selangor	Temp_D - Temp_W	0.94	0.554
	Cond_D - Cond_W	0.645	< .001
	TDS_D - TDS_W	0.651	< .001
	Salinity_D - Salinity_W	0.625	< .001

D: water parameters measured in the dry season.
W: water parameters measured in the rainy season.
Measure 1: water parameters in the dry season.
Measure 2: water parameters in the wet season.
Temp: temperature.
Cond: conductivity.
TDS: total dissolved solids.

Temperature was the only parameter that was significantly different in both season for the rest of the rivers surveyed. While all parameters showed a significantly higher means in the dry season, the water parameters of the Tengi and Sungkai rivers were significantly higher in the wet season (Table 4.17 and Table 4.18).

4.2.2.3 Inferential Analyses (ANOVA)

When season was selected as a factor, the Levene's test results showed that only the temperature data deviated ($p = 0.02$) but the pH data did not violate homogeneity of variance ($p = 0.237$) and therefore the Welch ANOVA was carried out for temperature data.

The pH of the rivers showed significant difference in the standard ANOVA, $F(1, 94) = 5.981$, $p = .016$, $\eta^2 = .060$ (Table 4.18). The standard post hoc test indicated a significantly higher pH in the dry season ($M = 6.545$, $SD = 0.768$) as compared to the wet season ($M = 6.142$, $SD = 0.845$) for all the rivers (Table 4.19)

There was significant difference in the Welch ANOVA for temperature, $F(1, 75) = 30.094$, $p < .001$, $\eta^2 = .243$ (Table 4.18). Similar to the pH above, The Games-Howell post-hoc test indicated a significantly higher temperature between the dry ($M = 30.616$, $SD = 2.357$) as compared to the wet season ($M = 28.463$, $SD = 1.356$) for all the rivers (Table 4.20).

Table 4.17: The standard paired t-test showing only significant results.

Sites							95% CI for Mean Difference			
	Measure 1 - Measure 2	t	df	p	Mean Difference	SE Difference	Lower	Upper	Cohen's d	SE Cohen's d
Tekah	Temp_D - W	22.571	2	0.002	1.756	0.078	1.421	2.09	13.031	0.682
Nerus	Temp_D - W	11.044	7	< .001	2.696	0.244	2.119	3.273	3.905	1.032
	pH_D - W	4.594	7	0.003	0.4	0.087	0.194	0.606	1.624	0.659
	Salinity_D - W	2.428	7	0.046	1.688	0.695	0.044	3.331	0.859	0.552
Sungkai	Temp_D -W	-3.874	4	0.018	-1.653	0.427	-2.838	-0.468	-1.732	0.725
Ban / Tengi	Temp_D -W	7.223	9	< .001	0.867	0.12	0.595	1.138	2.284	0.074

Measure 1: parameters in the dry season (D).
Measure 2: parameters in the wet season (W).
t: t-test value.
df: degree of freedom.
p: significant value of t test (<0.05).
SE: standard error.
Temp: temperature.

Table 4.18: List of parameters tested significant with the Wilcoxon signed-test, W.

						95% CI for Hodges-Lehmann Estimate			
Sites	Measure 1 - Measure 2	W	z	p	Hodges-Lehmann Estimate	Lower	Upper	Rank-Biserial Corr	SE Rank-Biserial Corr
Jengka	pH_D -W	55	2.803	0.002	0.507	0.285	0.682	1	0.342
Belara	Temp_D -W	36	2.521	0.008	7.517	6.867	10.367	1	0.377
Tengi	Cond_D - W	0	-2.803	0.002	-4.45	-10.083	-2.533	-1	0.342
	TDS_D - W	0	-2.803	0.002	-3.133	-7.267	-1.667	-1	0.342
	Salinity_D - W	0	-2.803	0.002	-1.55	-4.067	-0.817	-1	0.342
	Cond_D - W	0	-2.803	0.002	-4.45	-10.083	-2.533	-1	0.342

Measure 1: water parameters in the dry season (D).
Measure 2: water parameters in the wet season (W).
p: significant value of t test (<0.05).
z: measure of standard deviations.
Hodges-Lehmann Estimate: median difference between two seasons.
Corr: correlation.

Table 4.19: The list of significant standard ANOVA, F, between seasons for water parameters that did not deviate from normality.

Water parameter	Homogeneity Correction	Cases	Sum of Squares	df	Mean Square	F	p	η^2
Temperature	None	season	111.227	1.000	111.227	30.094	< .001	0.243
		Residuals	347.425	94.000	3.696			
	Welch	season	111.227	1.000	111.227	30.094	< .001	0.243
		Residuals	347.425	75.042	4.630			
pH	None	season	3.901	1	3.901	5.981	0.016	0.060
		Residuals	61.307	94	0.652			

df: degree of freedom.
p: significant value (<.0.05).
η^2: eta squared.

Table 4.20: Mean between season for temperature and pH.

Water parameter	Season	N	Mean	SD	SE	Coefficient of Variation
Temperature	dry	48	30.616	2.357	0.340	0.077
	wet	48	28.463	1.356	0.196	0.048
pH	dry	48	6.545	0.768	0.111	0.117
	wet	48	6.142	0.845	0.122	0.138

N: sample size.
SD: standard deviations of mean.
SE: standard error of mean.

Significant differences in all water parameters were recorded in between river locations. The Welch ANOVA was carried out as all data violated the Levene's test ($p < 0.001$) (Table 4.21). The Games-Howell post-hoc test showed that in the case of temperature, the mean for Ban Canal was significantly lower than the mean for the Jengka River (Table 4.22).

For pH, the Ban Canal was significantly higher (closer to neutral) than the Terengganu rivers and the Tengi River which the canal drains into. The pH for Jengka and Sungkai rivers was also significantly higher in the former and both these rivers had mean pH significantly higher than the Terengganu Rivers (more acidic).

The conductivity of the Nerus River was significantly lower than the Jengka and Sungkai rivers. Similar findings were seen with the mean variation of TDS as in conductivity. For salinity, the mean salinity of the Tengi River was significantly higher than Sungkai River and both means were significantly higher than all other rivers except for the Ban Canal (Table 4.22).

Table 4.21: The list of significant water parameters between locations tested with ANOVA, F.

Water parameter	Homogeneity Correction	Cases	Sum of Squares	df	Mean Square	F	p	η^2
temperature	Welch	waypoints	54.407	6	9.068	3.477	0.009	0.119
		Residuals	404.244	31.30	12.917			
pH	Welch	waypoints	35.135	6	5.856	30.23	< .001	0.539
		Residuals	30.073	29.34	1.025			
Conductivity	Welch	waypoints	4127.067	6	687.845	44.275	< .001	0.49
		Residuals	4294.672	29.75	144.349			
TDS	Welch	waypoints	2029.295	6	338.216	33.328	< .001	0.486
		Residuals	2142.759	29.03	73.813			
Salinity	Welch	waypoints	694.502	6	115.75	21.226	< .001	0.455
		Residuals	833.174	30.30	27.496			

TDS: total dissolved solids.
df: degree of freedom.
p: significant value (<.0.05).
η^2: eta squared.

Table 4.22: Mean between locations for water parameters that showed significant difference.

Water parameter	waypoints	N	Mean	SD	SE	Coefficient of Variation
Temperature	BAN CANAL	8	28.9	0.303	0.11	0.01
	JENGKA	20	30.16	1.28	0.29	0.042
pH	BAN CANAL	8	6.738	0.4	0.14	0.059
	JENGKA	20	7.274	0.322	0.07	0.044
	NERUS	16	6.081	0.281	0.07	0.046
	SUNGKAI	10	6.862	0.26	0.08	0.038
	TEKAH	6	5.768	0.516	0.21	0.09
	TENGI	20	5.71	0.94	0.21	0.165
Conductivity	BAN CANAL	8	45.34	20.57	7.27	0.454
	JENGKA	20	34.78	2.223	0.5	0.064
	NERUS	16	25.19	2.256	0.56	0.09
	SUNGKAI	10	33.99	1.542	0.49	0.045
	TEKAH	6	25.68	2.207	0.9	0.086
	TENGI	20	39.69	6.998	1.57	0.176
TDS	BAN CANAL	8	32.16	14.55	5.14	0.452
	JENGKA	20	22.71	1.219	0.27	0.054
	NERUS	16	17.88	1.64	0.41	0.092
	SUNGKAI	10	24.06	1.186	0.38	0.049
	TEKAH	6	18.29	1.614	0.66	0.088
	TENGI	20	28.19	4.971	1.11	0.176
Salinity	BAN CANAL	8	29.02	8.635	3.05	0.298
	JENGKA	20	22.06	1.488	0.33	0.067
	NERUS	16	20.92	1.666	0.42	0.08
	SUNGKAI	10	24.27	0.622	0.2	0.026
	TEKAH	6	20.9	1.107	0.45	0.053
	TENGI	20	26.72	2.676	0.6	0.1

TDS: total dissolved solids.
N: sample size.
SD: standard deviations of mean.
SE: standard error of mean.

4.2.3 GIS Analyses

4.2.3.1 Distance to the Nearest Hub

The Sungkai River capture site was found to be furthest from the nearest Protected Areas (PAs) when compared to the other 3 sites, with both Jengka and Tekah rivers having similar distances (3.8 km). The Sungai Dusun Wildlife Conservation Centre, Selangor was the only PA in which a female was caught. All sites are within 10 km of settlements (Table 4.23). The maps for these analyses are shown in Fig. 4.13.

4.2.3.2 Terrain Analyses

The elevation at the site of capture of *Tomistoma* ranged from 11-32 meters (Table 4.24). The graphs using DataPlotly and the Terrain Profile are depicted in Fig. 4.14 and 4.15. The Strahler Order of 3 and below (Table 4.24) indicates that this species was captured closer to where smaller streams discharge into the main river. The overlay with the river vector layer is in Fig. 4.16. The slope was also low in all the rivers surveyed (Table 4.24). Slope maps are in Fig.4.17.

Table 4.23: Table showing distance of capture / sighting site to the nearest settlement and nearest Protected Area (PA).

River	Type of locality	Name	Distance (km)
Sungkai	Village	Kg. Changkat Sulaiman	7.93
	PA	Sg. Dusun Wildlife Reserve	40.08
	PA	Bukit Tapah Water Catchment Forest	40.4
Ban Canal / Tengi	Village	Felda Soeharto	5.36
	PA	Within SDWRCC	0
Jengka	Village	Kampung Awah	3.8
	PA	Jengka Water Catchment Forest	18. 37
Tekah	Village	Kampung Sg. Ikan	3.84
	PA	Bukit Kesing Water Catchment Forest	13.97

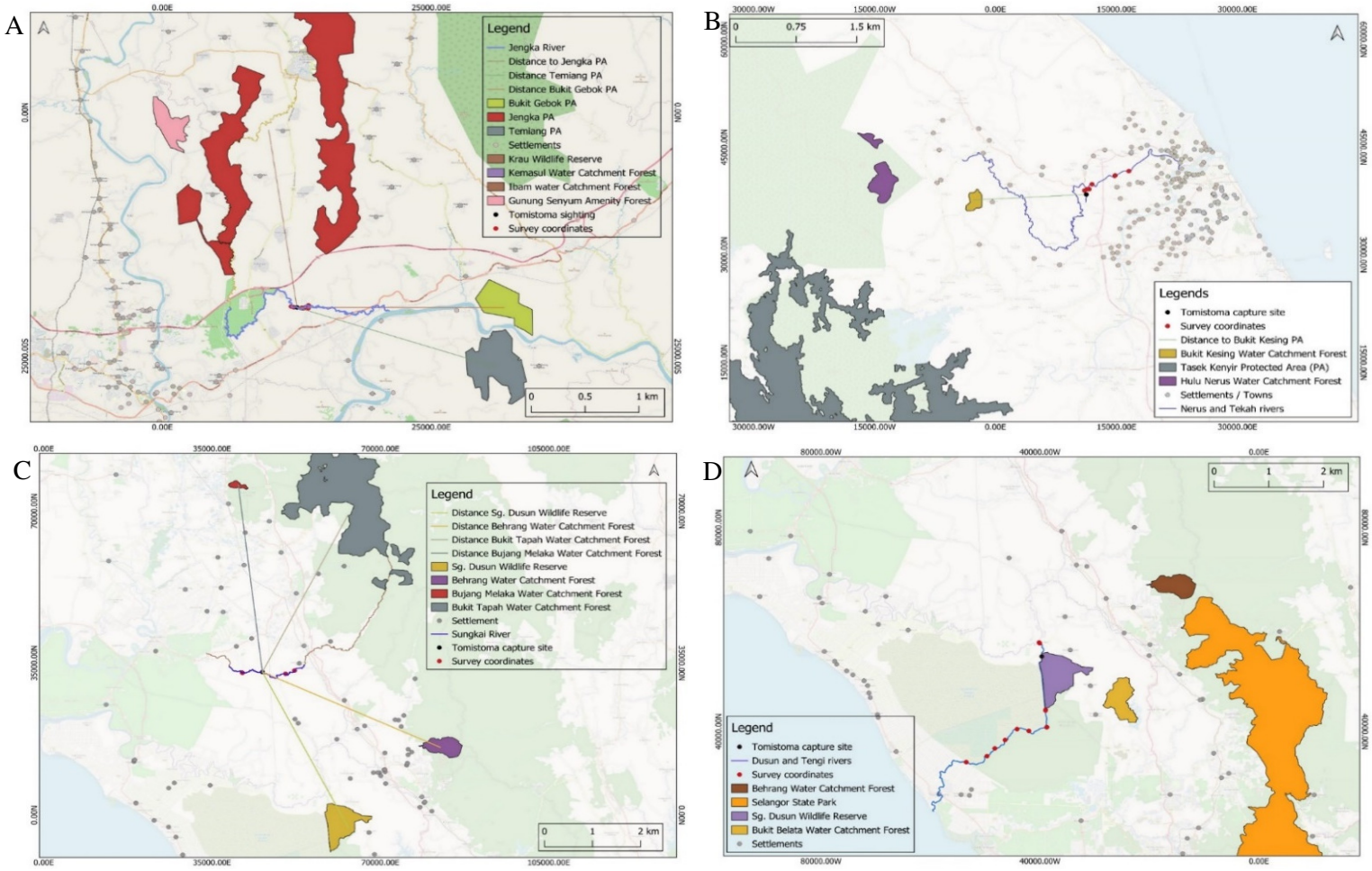


Fig. 4.13. Results of nearest hub analyses by (line and point) indicating locations of the *T. schlegelii* capture sites relative to the protected areas (PAs) and human settlements. PA: indicated in coloured boxes in Legend. (A) Jengka River, Pahang. (B) Tekah River, Terengganu. (C) Sungkai River, Perak. (D) Ban Canal, Selangor within the Sungai Dusun Wildlife Reserve.

Table 4.24: Terrain analyses of the capture sites and whole rivers.

Rivers	Elevation (m)			Slope (º)		Strahler Order
	Whole river	River surveyed	Capture site	Whole river	Capture site	Capture site
Dusun River	2-20	2-20	16	0-14	0-10	Between 2 and 3
Sungkai River	6-1731	13-40	27	0. 33-39.63	0-10	Between 1 and 3
Jengka River	26-72	33-37	32	0-29.89	0-10	Between 1 and 2
Tekah River	1-30	7 -21	11	0.34-19.42	0-10	1

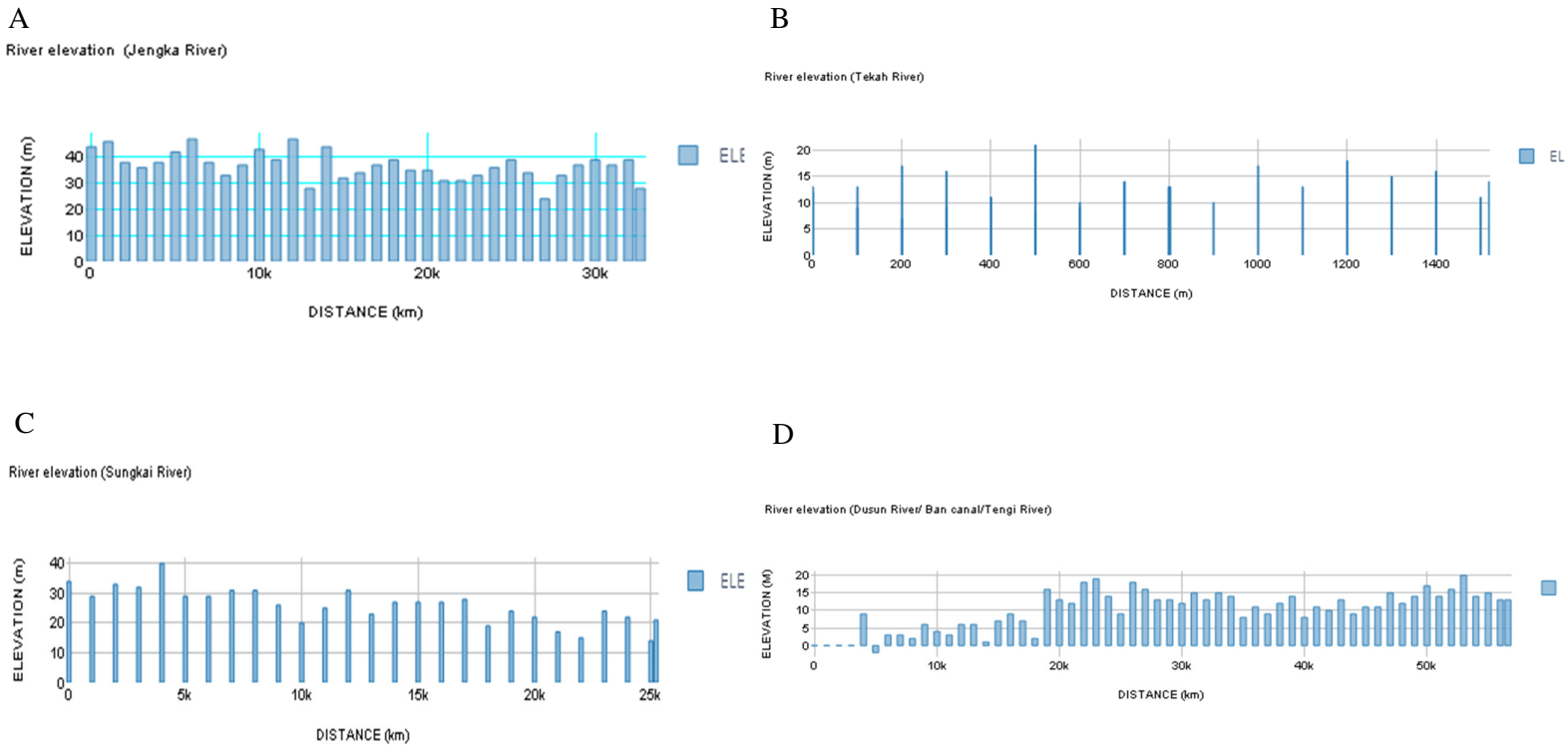


Fig. 4.14. Results of DataPlotly showing river elevation across the surveyed area. (A) Jengka River results. (B) Tekah River. (C) Sungkai River. (D) Dusun River, Ban Canal, Tengi River.

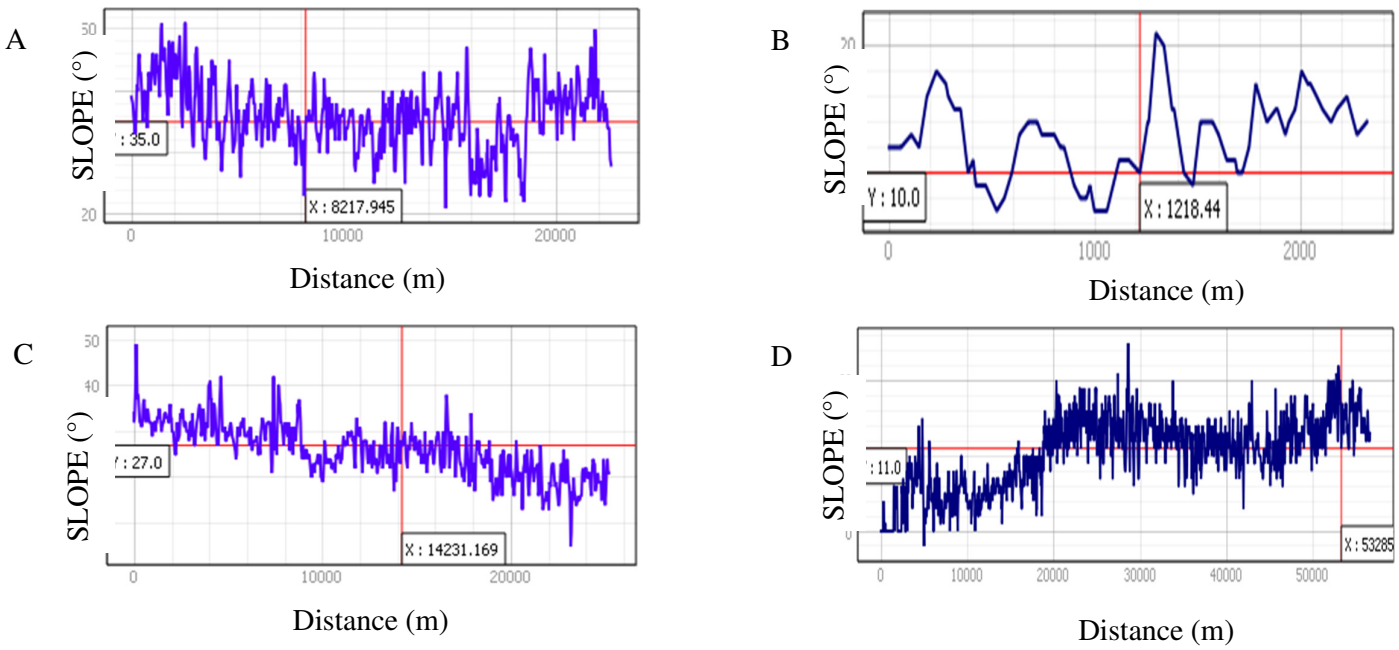


Fig. 4.15. Results of terrain analyses. A: Jengka River results. B: Tekah River. C: Sungkai River. D: Dusun River, Ban Canal, Tengi River. Terrain analyses carried out by sampling every 1 km except for Tekah River which was carried out every 0.1 km.

Fig. 4.16. Results of Strahler Order analyses of the *T. schlegelii* capture sites. Range of Strahler Order is from 1-6, with lower number indicating smaller streams.

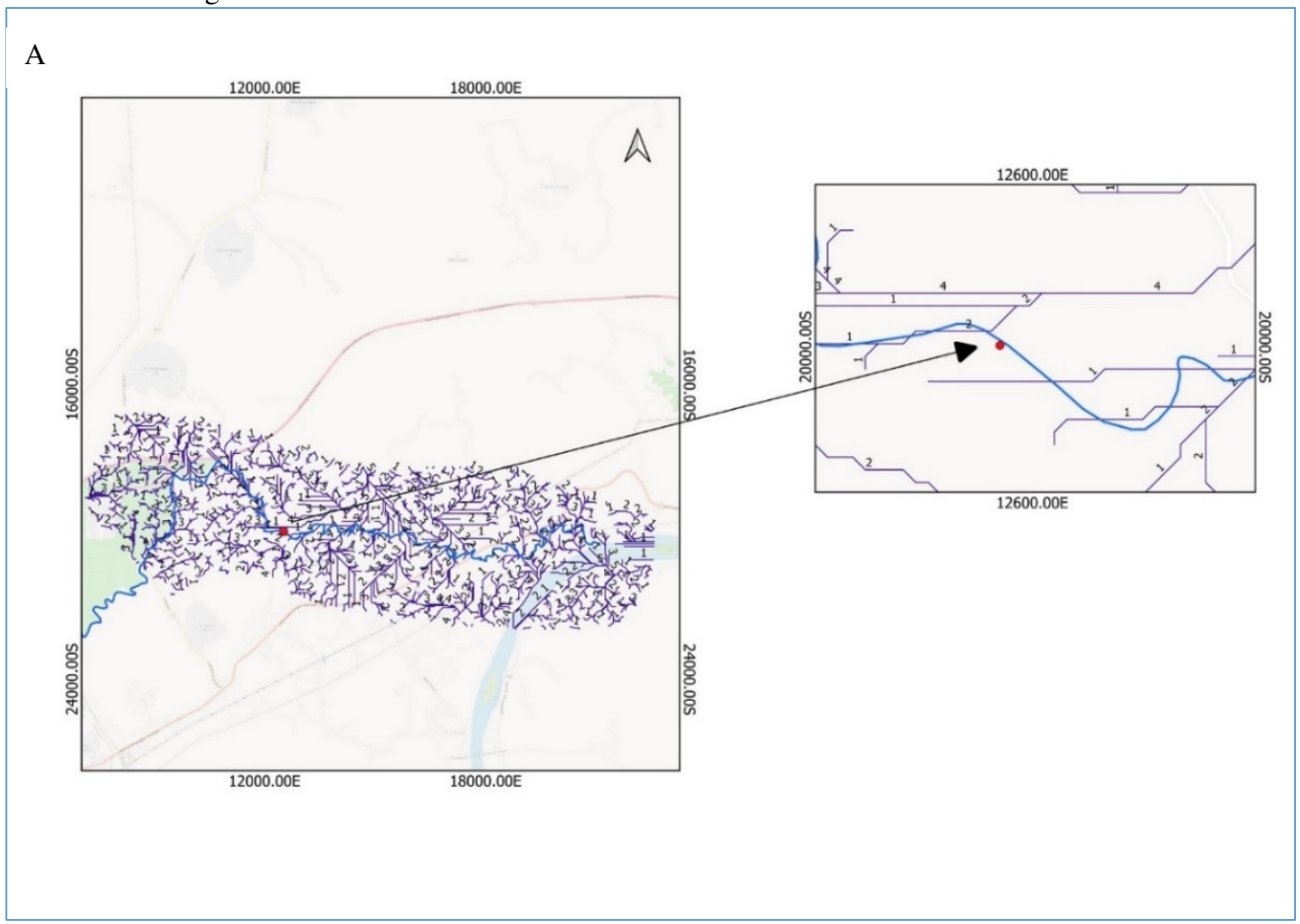


Fig. 4.16 (A) Strahler Order of the Jengka River.

B

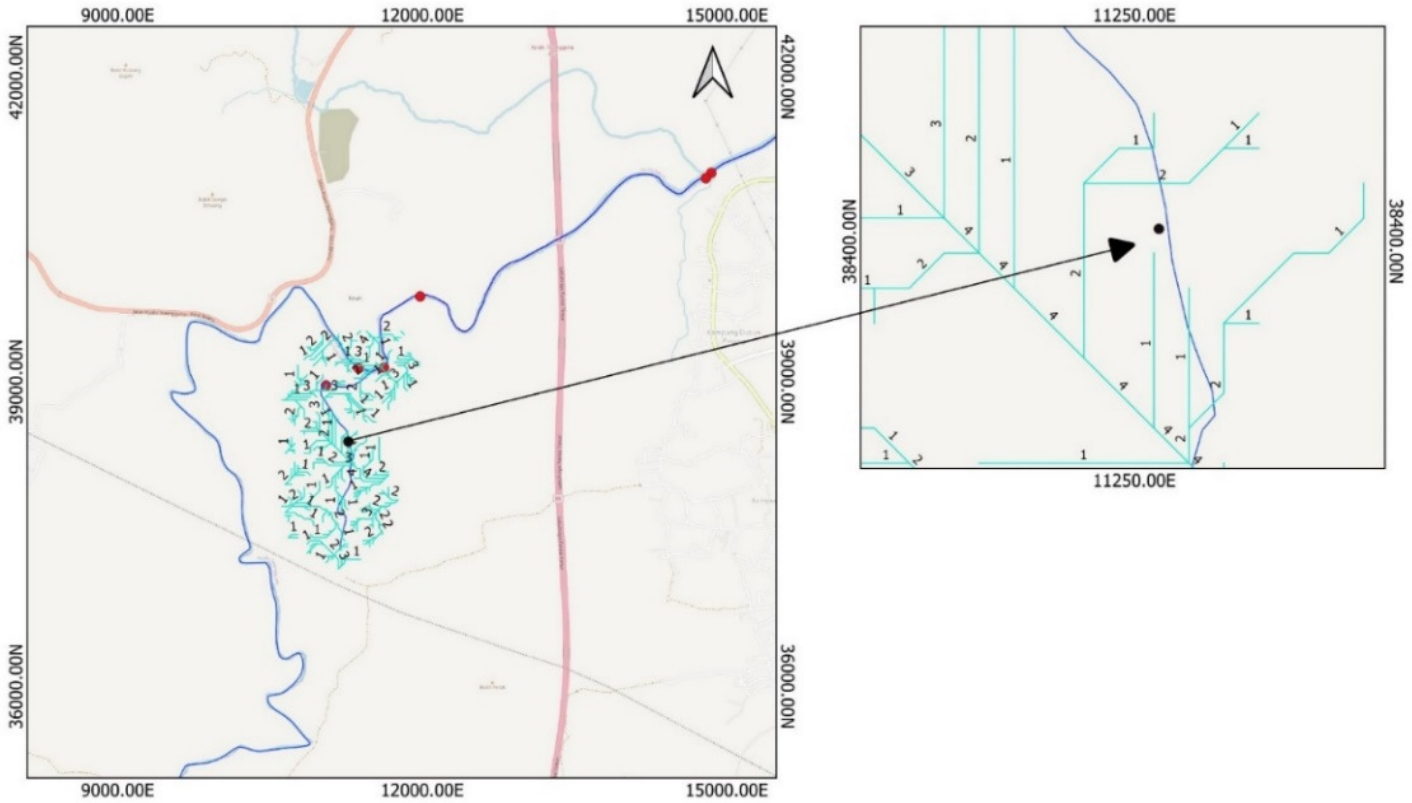


Fig. 4.16 (B) Strahler Order of the Tekah River.

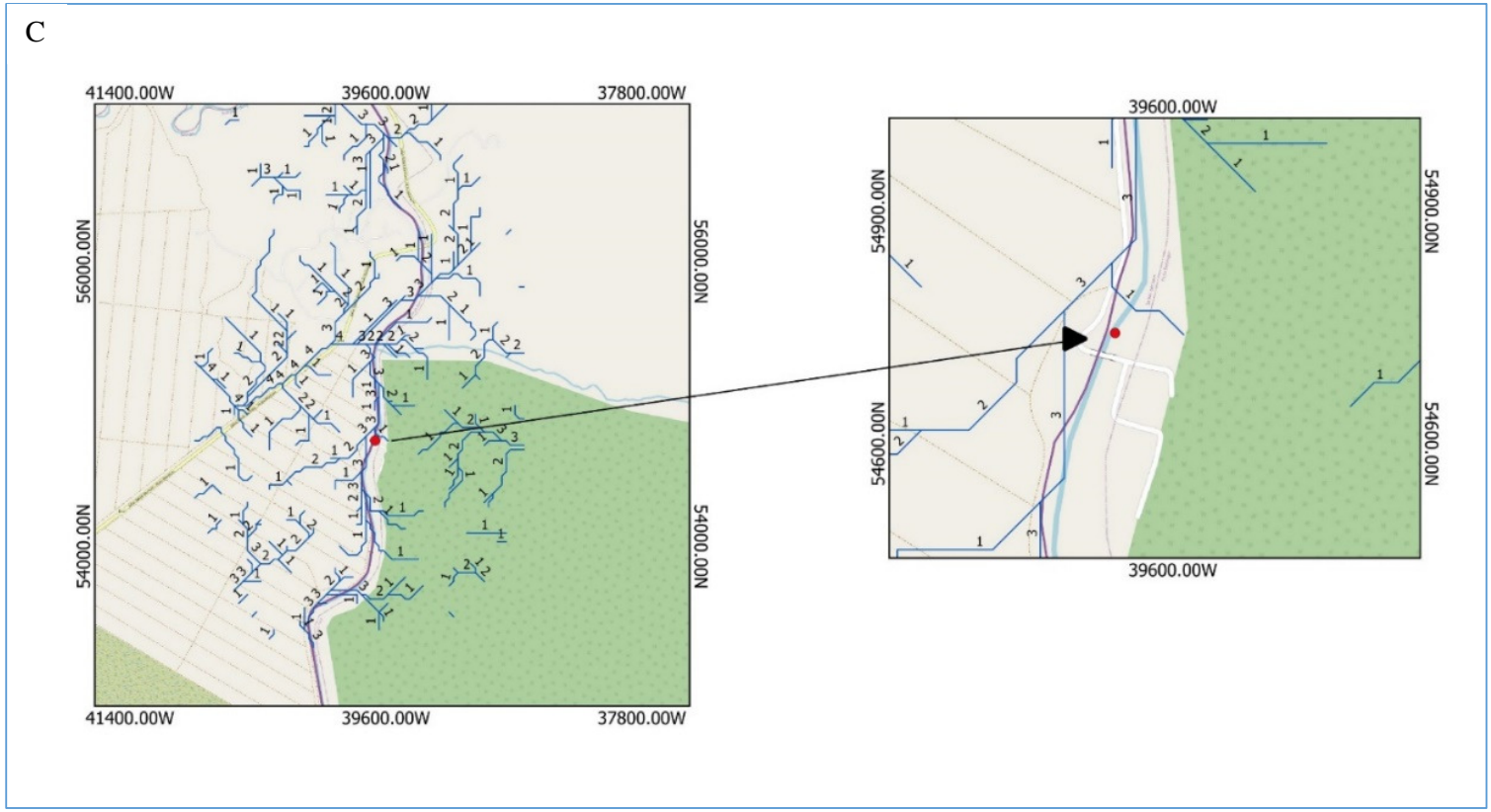


Fig. 4.16 (C) Strahler Order of the Sungkai River.

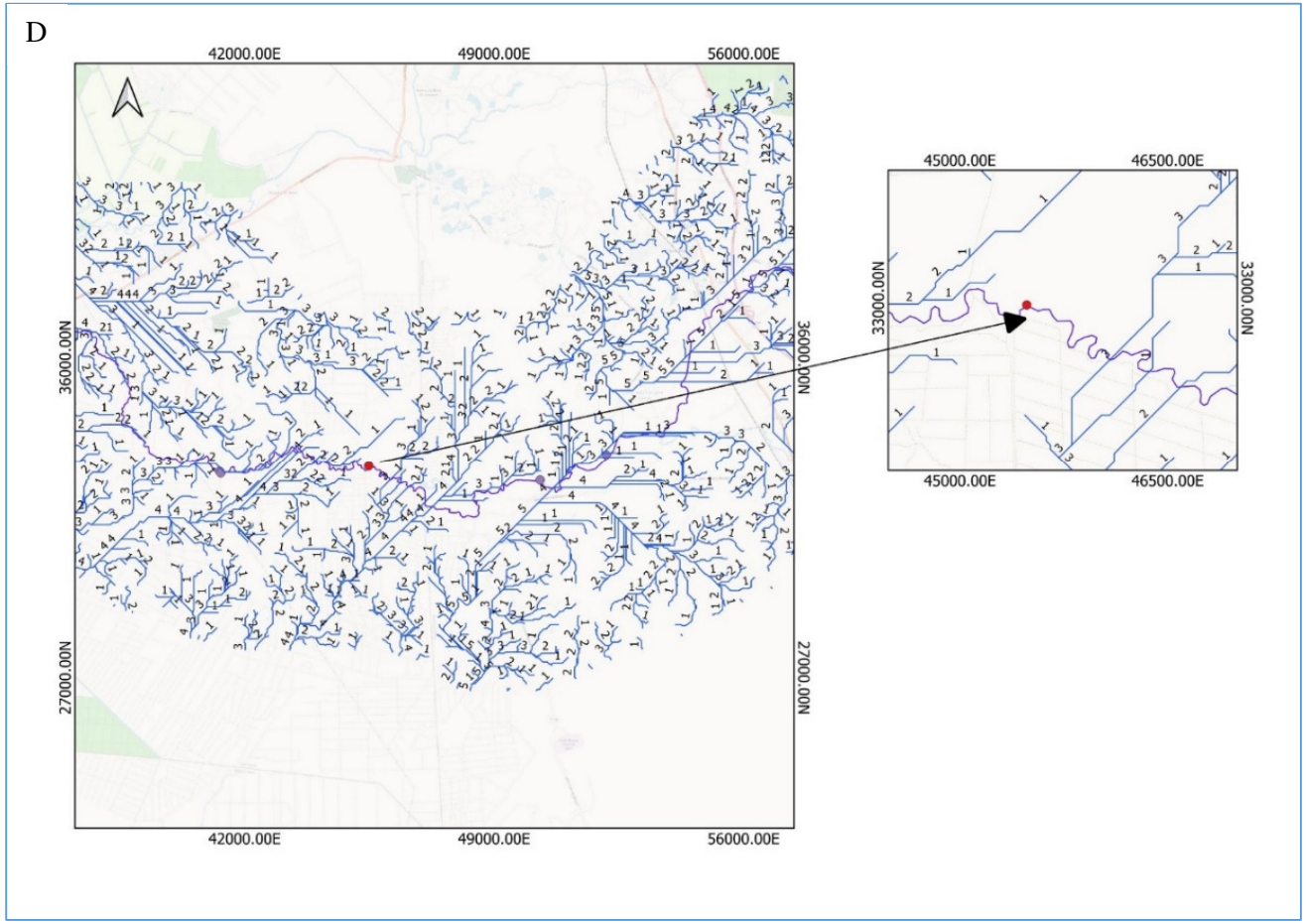


Fig4.16 (D) Strahler Order of the Ban Canal.

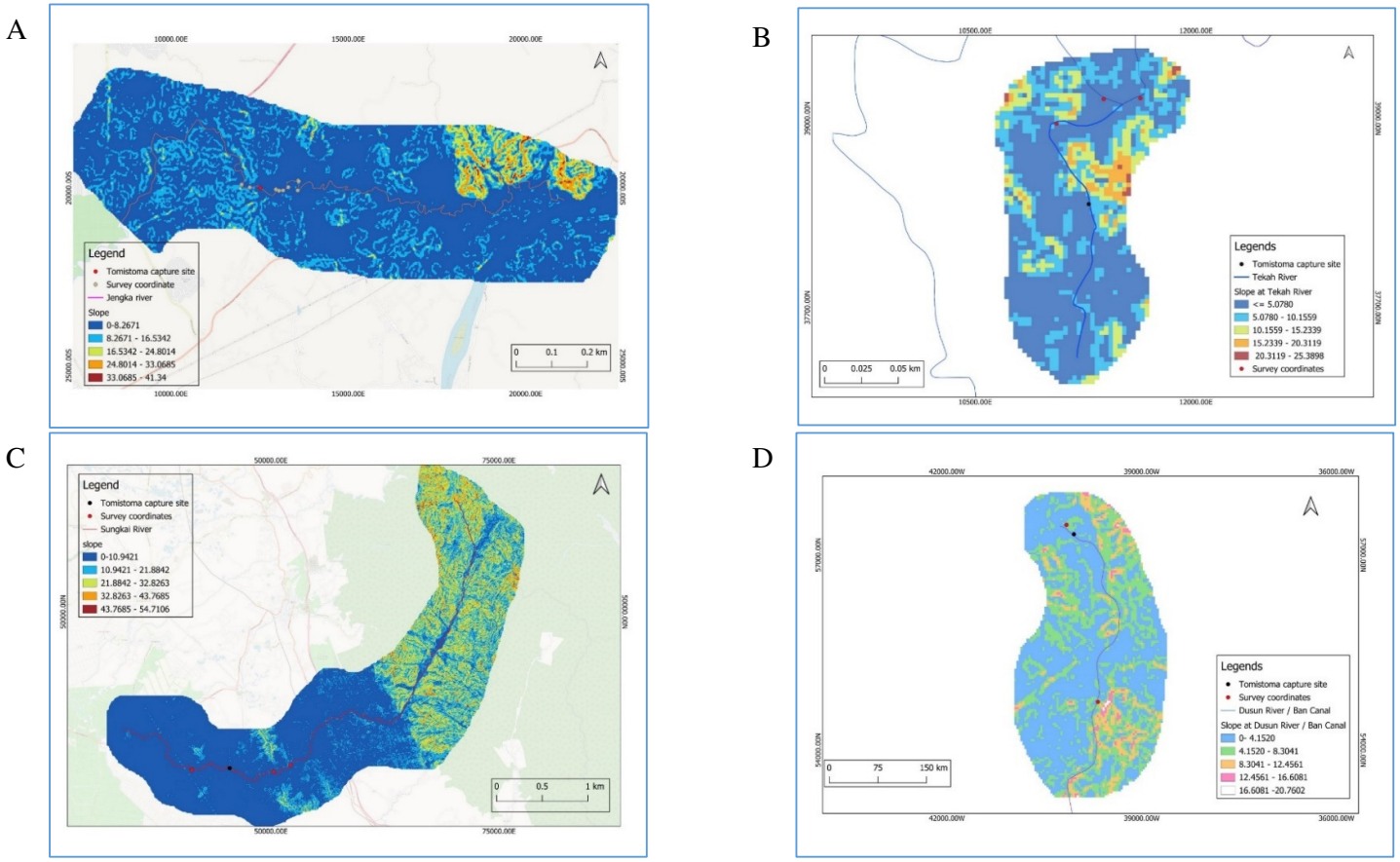


Fig. 4.17: Maps showing river slopes and gradient details of rivers surveyed. (A) Tekah River. (B) Sungkai River. (C) Dusun River. (D) Ban Canal.The gradient of slopes is represented by colored boxes in the Legends, with blue boxes indicating a slope range of 0–10° across all maps (A–D).

CHAPTER 5

DISCUSSION

5.1 Effectiveness of Microsatellite and mtDNA Markers

The microsatellite markers by Chang *et al*. (2014) proved to be informative with a high Polymorphic Informative Content (PIC) in most loci. Although a lower heterozygosity was detected in the present dataset, the Hardy-Weinberg Equilibrium (HWE) results for significant departures ($p<0.05$) due to excess of homozygotes were similar to Chang *et al*. (2014). However, after Bonferroni corrections more loci were in HWE and no linkage disequilibrium was detected when the current dataset was analysed following the STRUCTURE results of K=6, compared to Chang *et al*. (2014) proving to be suitable for population level study of *Tomistoma*.

The ND 6-tRNA^{glu}-cyt *b* and the cyt *b*-CR markers used in the present study have been proven to be informative by Kaur *et al*. (2013). In this study, only H1 and H4 were detected with five individuals each from the ten new samples. Together with the microsatellite markers, the genetic analyses were able to differentiate the four mtDNA haplotypes into six clusters, with the separation of the Sarawak H2 lineage from the Peninsular Malaysia H2 lineage. The H1 lineage was differentiated into three exclusive clusters, Clusters C, D and F and a single state of origin was attributed to Clusters A (Sarawak), D (Selangor) and perhaps Cluster C (Perak). Furthermore, the microsatellite analyses supported the findings of Kaur *et al*. (2013) on the H4 haplotype with the separate clustering of this lineage into Cluster E exclusively.

In contrast, the combination of using microsatellite markers of other crocodilians (*Crocodylus* spp. and *Alligator mississippiensis*) and ISSR were only able to differentiate the *Tomistoma schlegelii* population into Peninsular Malaysia and Sarawak (Shafiei-Astani *et al*., 2015). Previous crocodilians studies have also shown that information at population and individual level was not satisfactory due to low levels of polymorphism using nonspecific microsatellites markers, (*Caiman crocodilus*-de Oliveira *et al*., 2010; *A. sinensis*-Jing *et al*., 2009). Additionally, species-specific dinucleotide markers were found to be prone to scoring errors due to stutter making the continuation of datasets for long term projects problematic for *A. mississippiensis* (Subalusky *et al*., 2012) and unable to detect multiple paternity in *A. sinensis* (Wu *et al*., 2012) resulting in the development of tetranucleotides markers. As all the markers by Chang *et al*., (2014) for *T. schlegelii* were tetranucleotide, no such problems were encountered in the current dataset. In fact, the maternal parent (TK008) and all her offsprings clustered together with one adult male (TK007) (Cluster D) in the current dataset. Furthermore, private alleles and distinct allelic patterns (Fig. 4.6) were seen in every cluster including Cluster C which consisted of only two individuals.

5.2 Genetic Diversity Based on Microsatellite and mtDNA Markers

The observed heterozygosity levels within clusters (H_O = 0.25-0.5) and for all locus (H_O = 0-0.6), except TM23 and TM24, was low. As this is the first study of *T. schlegelii* using microsatellite markers specific for this species, the findings of this study are compared with other crocodilian studies employing similar methods. The low diversity reported in this study is similar with previous

crocodilian studies with observed heterozygosity, H_O, range of 0.4-0.6 (Alligatoridae-Glenn *et al*., 1998; Muniz *et al*., 2019; Wu *et al*., 2012; Xu *et al*., 2005; *Crocodylus*-Milian-Garcia *et al*., 2015; *Gavialis*-Sharma *et al*., 2021). Several reasons have been described for a low diversity level in crocodilians. Glenn *et al*. (1998) explained it as a "signature" at the nuclear level diversity due to a prehistorical population bottleneck or a selective event that eliminated much of the variation from the crocodilian genome. Muniz *et al*. (2019) suggested that traits like longevity, generation overlap and late maturation contributed to this "signature". However, heterozygosity is lowered when population is structured or due to inbreeding or sampling closely related individuals as indicated by the positive F_{IS} in this study. This is supported by the FS and PO relationships found between individuals within each cluster and the Queller and Goodnight (1989) estimator = 0.317 ± 0.05. Additionally, the microsatellite data AMOVA showed low variations among individuals and a non-significant R_{IS} within clusters. Although the relat, equivalent to the Queller and Goodnight (1989) estimator, was high, it was still lower than that reported in *Alligator sinensis* (0.5 ± 0.25) which Xu *et al*. (2005) indicated was due the small founder population at the breeding centre and the nature reserve.

Previously, Kaur *et al*. (2013) reported distribution of mtDNA haplotypes in Malaysia and Indonesia and detected the occurrence of four haplotypes in Peninsular Malaysia (H1, H2, H4 and H5). In contrast, only H1 and H4 haplotypes were detected from the ten new samples in this study with mtDNA diversities only observed in Cluster B.

Inferences about mtDNA variability in Cluster B are limited due to the unconfirmed origins of some sampled individuals. Cluster B consists of multiple haplotypes and possibly mtDNA subpopulations, but only two individuals have known origins (TK009 and TK015, both from Selangor). The H5 haplotype individual (TK30) was surrendered to the national zoo following the 2007 floods, which primarily affected Johor, Pahang, Terengganu, and Kelantan. The H1 haplotype individual (TK66), purchased from Singapore, also has uncertain origins, as it is believed that the living specimens in Singapore were originally sourced from Peninsular Malaysia, although the specific states are unknown (Abraham Matthews, pers. comm., 2008). Comparisons with similar diversity patterns in *Crocodylus suchus* and *C. niloticus* have attributed mtDNA diversity to historical gene flow and connectivity (Van Asch *et al*., 2019; Velo-Anton *et al*., 2014). Similarly, Kaur *et al*. (2013) proposed that ancient river systems facilitated the distribution of *T. schlegelii* haplotypes across regions. In contrast, increased mtDNA diversity in *Caiman crocodilus fuscus* sensu stricto was linked to factors such as metapopulations and the introduction of new individuals (Diaz-Moreno *et al*., 2021).

5.3 Population Structure Using Microsatellite and mtDNA

While the microsatellite markers were more sensitive at the population level than mtDNA (AMOVA showed conflicting results with the inclusion or exclusion of the divergent haplotype, H4), the explanation for the findings at the population level is limited due to the uncertain origins of some specimens of the captive *T. schlegelii*. The significantly differentiated AMOVA using microsatellite data (R_{ST} and R_{IT}) was congruent with most of the pairwise F_{ST}. The latter indicated

Cluster C as the only cluster not significantly differentiated with all other clusters. Spatial connectivity remains a possibility for this species, as seen by the non-significance differentiation of Cluster C in the pairwise F_{ST} analysis. The Sungkai River, in which TK058 (Cluster C) was caught, is close to the Bernam River. Individual TK060 (H1 haplotype) from Cluster F was caught in the Dusun Rivers, Selangor which is a tributary of the Bernam River. Similarly, the Cluster D comprise of individuals from Selangor state. Possibility for the non-significance differentiation with Cluster B and E could be explained by spatial connectivity as well as Selangor state is the likely origin of 50% of Cluster B and E members. Compared to Cluster C, the non-significant differentiation between Cluster B and E was at a smaller scale (pairwise $F_{ST} = 0.0951$, $p = 0.008 > 0.003$). Due to the small sample size of each cluster, the bottleneck and migration analyses could not be carried out and the spatial connectivity possibility is based on field data / geographical representation. Additionally, individuals in Clusters E and F from both divide of the Titiwangsa Ranges showed FS relationship within the respective clusters. The individual TK017 from Pahang had a FS relationship with individuals from Selangor, namely TK010 and TK060 while for Cluster E, individual TK064 caught in the Tekah River, Terengganu had FS relationship with TK011 and TK013, both individuals likely from Selangor.

In Peninsular Malaysia, freshwater fish surveys also did not show a west and east divide suggesting that the geographical barrier may not be a factor in their distribution and instead a north-southern crescent faunal line was suggested (Ng *et al.*, 2016). In this data set of *T. schlegelii*, the population seemed to show a

trend similar to the above where the east and west Peninsular Malaysia shows genetic connection for this species.

Some of the population differentiation trends seen across crocodilian have been due to either geographical barriers (Hekkala *et al*., 2010), ecological preferences (Farias *et al*., 2004; de Thoisy *et al*., 2006; Vasconcelos *et al*., 2008) or distance (Glenn *et al*., 1998; Dever *et al*., 2002; Rossi *et al*., 2020). Apart from suggesting flooding as a mechanism of contact between populations, terrestrial movement of crocodilian have been reported in *Crocodylus moreletii* (Dever *et al*., 2002) and the west African crocodile, *C. suchus* (Velo-Anton *et al*., 2014) and *Osteolaemus* (Smolensky *et al*., 2015). The Cameroon Volcanic Line did not limit the distribution of O. *tetraspis* to the west of this barrier while further investigation was suggested in the case of *C. suchus* when one individual was confirmed to have dispersed over long distance on land (Velo-Anton *et al*., 2014). The mtDNA indicated a historical gene flow and connectivity while microsatellite data show low levels of movements between sub-basins indicating a metapopulation system resembling a source-sink dynamic (Velo-Anton *et al*., 2014). Similarly, the combination of microsatellite AMOVA and mtDNA AMOVA (excluding the H4 lineage) may indicate the presence of a metapopulation structure influenced by female-mediated gene flow. Stronger evidence for this hypothesis could be achieved with a larger sample size, which would enhance the reliability of bottleneck and migration analyses. The combination of the microsatellite AMOVA and mtDNA AMOVA (including the H4 lineage is discussed in 5.4)

5.4 mtDNA Haplotype Distribution and Divergence

The divergence time of the three main clades began pre-Early Pliocene Epoch (Fig. 4.9), followed by further divergence within these clades during the mid-Pleistocene Epoch (1.4–1.8 million years ago). This pattern aligns with the haplotype network, which depicts three distinct evolutionary relationships among haplotypes in the median-joining networks (Fig. 4.7). The H4 lineage, placed on its own evolutionary branch, is consistent with findings by Kaur *et al*. (2013), where this haplotype was placed on a separate branch in the minimum spanning network and was not parsimoniously connected in the TCS network tree. Interestingly, the H4 haplotype, observed on both sides of the Titiwangsa range, has only been identified in Peninsular Malaysia thus far. The divergence since before the onset of Early Pliocene Epoch could probably be the reason it has not been detected from other regions thus far. However, this would be better supported with a bigger sampling size across Malaysia and Indonesia. The separate placement from other haplogroups in the median-joining network (Fig. 4.7), and its early divergence time, supports the suggestion by Kaur *et al*. (2013) to further investigate this lineage as a potential management or evolutionary significant unit. Furthermore, microsatellite data (from a nuclear gene perspective) grouped individuals with this lineage exclusively into Cluster E (Fig. 4.3), which showed significant differentiation in both microsatellite and mtDNA AMOVA analyses. Additionally, the H4 haplotype displayed a genetic distance comparable to other hybridising crocodilian species (*Crocodylus acutus*, *C. rhombifer*, *C. moreletii*) and greater than the genetic distance within *Osteolaemus* species (Kaur *et al*., 2013).

Based on the concatenated data (Fig. 4.7), the H2 haplotype occurs in Sarawak and Peninsular Malaysia (Kaur *et al*., 2013). Additionally, the results of Kurniati (2003) (Fig. 4.6) and the D-loop median joining network (based on 451 bp) demonstrated that the H1 and H2 haplotypes were also present in West Kalimantan (Figs. 4.7 and 4.8). This suggests that Haplogroup I is the most widespread across the region of occurrence for this species. As noted earlier (Section 5.2), Kaur *et al*. (2013) attributed historical gene flow to ancient river systems during periods when the Sunda Shelf was exposed. This is supported by the divergence time (1.39 million years ago) within this clade in the divergence cladogram (Fig 4.9). Two major ancient river systems, the Malacca River System (draining western Peninsular Malaysia and eastern Sumatra) and the North Sunda River System (draining eastern Peninsular Malaysia and western Borneo), are thought to have facilitated such gene flow (Voris, 2000). Studies of frogs, snakes, and freshwater fishes have similarly highlighted Peninsular Malaysia as a region of mitochondrial DNA diversity overlap due to Pleistocene-era river connectivity (Inger and Voris, 2001; Tan *et al*., 2012).

With only three individuals from Sumatra (H3 haplotype) included in this study, there remains a possibility that the H1 and H2 haplotypes also occur in Sumatra, potentially due to ancient river systems. This warrants further investigation with a larger sample size across the island. Notably, both the H1 and H2 haplotypes are found in Selangor State, a region in Peninsular Malaysia that represents the last landmass to be separated from Sumatra around 17,000 years ago (Voris, 2000).

5.5 Possible Movement of Females During Breeding Season

The recent three wild caught *Tomistoma schlegelii* in Peninsular Malaysia (August-Sungkai River, Perak; April-Dusun River, Selangor; and June-Tekah River, Terengganu) were all captured during the breeding months of April-August (Mathew *et al*., 2011; Staniewicz *et al*., 2018), perhaps in search of males or nesting sites. The skewed female-to-male ratio (35 females to 3 males) among captive *T. schlegelii* in Malaysia prevented the FSTAT analysis of sex-biased dispersal from being conducted due to insufficient male representation at the population level. However, the mtDNA AMOVA analyses reveal contrasting findings depending on the inclusion (female biased philopatry) or exclusion of the H4 haplotype (female dispersal), supported by capture information and the GIS analysis findings.

A female philopatry (mtDNA AMOVA considering the H4 haplotype) is probable as the Sungkai River has seen adult females caught in the early 2000s (TK014/TK016) and again in August 2016 (TK058). Although homing instinct was indicated when a young *T. schlegelii* (< 1.5 m) returned to its capture site after 17 days (Bonke *et al*., 2014), studies on movement of adults are lacking. So far studies on the mating behaviour of *T. schlegelii* has been on nest characteristics, breeding month and period, and acoustic signals during mating (Bezuijen *et al*., 1998; Bezuijen *et al*., 2001; Staniewicz, Youngprapakorn and Jones, 2018; Staniewicz *et al*., 2018; Lariman *et al*., 2021).

A more homogenize population is expected in a female biased dispersal analysed with the maternally inherited mtDNA, due to the migrating female introducing

her lineage in addition to the offspring migrating to their own niche (Escorza-Treviño and Dizon, 2000). A female-biased dispersal (mtDNA AMOVA not considering the H4 haplotype) is also a possibility as the capture of an adult female adjacent to the holding enclosure at the Sungai Dusun Wildlife Conservation Centre could be due to searching for a male to mate as there was a male kept at the centre at that time. The centre embarked on a breeding / conservation initiative for this species circa 2010. Usually, sightings of *T. schlegelii* or nests have been downriver of the Tengi River. Meanwhile, the female caught in Terengganu (June) in fishing equipment was the first encounter for the fishermen here who are well familiar with the identification of *Crocodylus porosus*.

Usually, dispersal would enable an individual to escape competition for food source, breeding mate, nesting locations and access to unrelated mating partner to reduce effects of inbreeding. In contrast, philopatry is usually seen as means to seek security as females settle closer to kins or relatives while familiarity to nesting sites and food source increases survival (Clutton-Brock and Lucas, 2012). Due to this, in *C. porosus*, females are thought to not move far from suitable nesting sites (Lewis *et al*., 2013). Furthermore, limited dispersal (<50 km) detected with spatial analyses supported the kinship analyses of reproductive philopatry in *C. porosus* (Lloyd-Jones *et al*., 2023). Male migration and female philopatry was suggested when a higher mtDNA F_{ST} compared to microsatellites were documented for the American crocodile, *C. acutus* (Rodriguez *et al*., 2011). Similarly, philopatry has been more commonly

reported in other crocodilians as well (Hekkala *et al*., 2010; Rodriguez *et al*., 2011; Shirley *et al*., 2013; Hekkala *et al*., 2015).

In contrast, dispersal was suggested for *C. suchus* population in southern Mauritanian mountains (Velo-Anton *et al*., 2014). This was concluded when the metapopulation resembled a source-sink dynamic which relies on individuals' dispersal ability to genetically "rescue" groups of isolated individuals. This was supported by the persistence of relict populations across the Sahara-Sahel, and carcasses found in drier periods believed to be of individuals dispersing during the rainy periods which partially connected the subpopulations. However, as philopatry is more common in crocodilians, Velo-Anton *et al*. (2014) also suggested factors such as ontogenetic shifts and phenotypic plasticity at the intra-individual level, and behavioural traits at the inter-individual level for these movements. Ontogenetic shifts were also reported in *C. porosus* when telemetry-derived data of female nesters was compiled over a 10-year period sites. As females gained nesting experiences, affinity towards certain nesting sites and nest guarding was seen in contrast to smaller nesters which either moved to and from nesting sites, or within home ranges and nesting sites (Baker *et al*., 2019).

Without proper representation of individuals for genetic analysis and lack of behavioural studies, direct and indirect breeding knowledge of *T. schlegelii* is limited. The small sample size in this study limits the findings, which include microsatellite data with incomplete field records and terrain analyses (discussed in Section 5.8 below), to merely suggestive of potential movements by females

during the mating season. Whether due to dispersal, philopatry, inter- and intra-individual variability will need to be investigated in the future.

5.6 Conservation Strategy

5.6.1 Identification of Important Clusters for Breeding

Where prior information such as locality or population of origin is lacking or incomplete (such as the captive held *T. schlegelii* in Peninsular Malaysia), pedigree analyses between individuals relies on molecular markers, using the molecular coancestry rather than genealogical coancestry with the aim of maintaining or maximising genetic variability (Toro *et al*., 2011). Understanding the genetic diversity (within and between populations) and population structure from both nuclear and mtDNA variations is necessary to design a proper breeding plan. As stated by Galla *et al*. (2022), the surviving endangered species usually have some level of coancestry due to the remaining small and usually structured populations. This was seen in this study where the weighted within-cluster molecular coancestry was high ranging from 0.364 to 0.608 while the weighted molecular coancestry distance ranged from 0.11 to 0.31.

In this study, the removal of cluster A, C and F resulted in an increase at the internal diversity and reduction of allelic richness which is based on the Petit, Mousadik and Pons (1998) internal diversity (Table 4.10). This indicates that breeding individuals in this cluster would favour inbred populations which is also supported by the lower heterozygosity (Table 4.5, Fig. 4.5). Incorporating Cluster B which had maximum overall Nei's (1987) gene diversity, the lowest weighted within cluster molecular coancestry and the highest number of

effective private alleles into the breeding plan would not only maintain and maximise the nuclear gene diversities, it would also prevent the loss of mtDNA diversity as well as this is the only cluster that showed mtDNA diversity. Although maximising the gene diversity usually also maximises the allelic richness (Caballero and Toro, 2002), this is not a rule (Alvarez *et al*., 2010) as seen in Cluster B in this study. Cluster E proved important for allelic richness as its removal from breeding would result in loss of diversity at all levels. In addition to that, the divergent maternal lineage, H4 (all Cluster E individuals), would also be maintained in the captive population.

In current captive situation, there is movement or exchange between breeding facilities and zoos in Peninsular Malaysia. Due to the endangered status of *T. schlegelii*, breeding programs should be carried out in as many holding centres as possible. Except for Cluster F which had a negative Petit, Mousadik and Pons. (1998) diversity and therefore rendering it the least important for the diversity at the allelic level, all other clusters showed some level of contribution (gene and allelic diversities) either at the within and between clusters and total population (Table 8). With the availability of the microsatellite and mtDNA data of these captive individuals from this study, these holding area could test the between individual molecular coancestry coefficients for proper assignments of males to females for a better breeding plan.

A review of 188 publications using microsatellites markers, studbooks, theoretical work and genetic effects of management strategies in ex-situ conservation in vertebrates, found that long-term captive breeding led to a loss

of expected heterozygosity and number of alleles (Witzenberger and Hochkirch, 2011). Additionally, it was found that to minimise loss of genetic diversity, a minimum number of 15 unrelated founders and a captive population size of 100 individuals was needed for a self-sustaining captive population which resembles the wild populations. This estimate of founders was also found to be able to retain 90% of the natural genetic diversity which would then be suitable for future reintroduction programmes (Witzenberger and Hochkirch, 2011). As the numbers of captive *Tomistoma* increase over time due to habitat destruction, the approach of starting with a minimum of 15 unrelated individuals as founders should be considered.

5.6.2 Multiple Paternity

Multiple mating strategy should be incorporated for the captive *T. schlegelii* in Peninsular Malaysia as the available captive and natural population data of other crocodilians suggests that multiple paternity occurs commonly. A study carried out on *Alligator sinensis* within the Anhui Province found 52% of the nests had two paternal origins while another 8% had three with the remaining nest having single paternity in 50 nests each from the second and third filial generation (F2 and F3) collected a decade apart. The hatchlings within the clutches also showed genetic variations which was suggested could help with the diversity of this endangered species (Wang *et al*., 2017). In natural population of *Crocodylus porosus* (Lewis *et al*., 2013) and *A. mississippiensis* (Lance *et al*., 2009), it is thought that this mating strategy was to ensure fertilization of eggs and avoid infertility due to sperm depletion in the dominant male as it mates with several females. In *A. mississippiensis*, this mating strategy is supported by home ranges

and radio telemetry studies that show home range overlap during breeding. In natural populations of *Caiman latirostris*, polyandry was attributed for the successful recovery of natural population after population bottleneck with increase diversity and gene flow from nearby populations (Ojeda *et al*., 2017). With most clusters in this study showing significant inbreeding, this approach should be given serious consideration when planning for captive breeding program for *T. schlegelii*.

The only successful breeding of captive held *Tomistoma* was in Zoo Negara where three males shared an enclosure with a female. With improvements made in diet and enclosure enrichment, the female laid 19 eggs of which eleven were fertile producing seven hatchlings (Mathew *et al*., 2011). The inbreeding coefficient, F_{IS}, was the lowest (Cluster D-one male and female and five hatchlings) and not significant (Table 4). The possibility of the family unit being the least inbred from all the cluster is probably due to multiple paternity. From the mtDNA data as well, although the adult male and female in this cluster belonged to the H1 lineage, they differed at the mtDNA variable tandem number repeats (VTNRs), which are stably inherited from the maternal parent (Kaur and Ong, 2011). This indicates that the adult males and female in this cluster are not maternal siblings.

In addition to that, the gain / loss analyses (Table 8) also identified this cluster as an important contributor of the Nei's (1987) gene diversity at all levels, albeit in a much smaller magnitude compared to Cluster B. The genetic data and information about the breeders in Zoo Negara, together with the confirmed

multiple paternity seen across crocodilian (Lance *et al*., 2009; Hu and Wu, 2010; Amavet *et al*., 2012; Budd *et al*., 2015; Wang *et al*., 2017) suggests that this approach would be beneficial in future breeding plans for *T. schlegelii* to increase diversity and prevent inbreeding in a single clutch / nesting season.

The theory of sperm storage is another possibility for the multiple paternity observed in crocodilians. In reptiles, this phenomenon is more common in turtles and squamates. A female Western Diamond-backed Rattlesnake held in isolation was able to produce two litters post capture. The litters, one and six years apart, upon further genetic analysis, were confirmed due to paternal contribution instead of facultative parthenogenesis (Levine, Schuett and Booth, 2021). For crocodilians, this was first suggested by Davenport (1995) in the Dwarf Caiman, *Paleosuchus palpebrosus*. However, Davis *et al*. (2001) questioned the necessity for such a mechanism when crocodilians breed once per mating season. Subsequently, Gist *et al*. (2008) reported the ability to store sperms in the oviduct glands of *A. mississippiensis* within the same breeding season while Lance *et al*. (2009) attributed paternity of same mates in subsequent breeding season to mate fidelity rather than sperm storage. Evidence for sperm storage in successive years and competition from newer supply of sperm in the current breeding season is lacking in *C. porosus* (Lewis *et al*., 2013) while sperm storage in *Ca. latirostris* and *C. moreletii* remains to be determined (McVay *et al*., 2008; Amavet *et al*., 2012). In the case of *Tomistoma*, this possibility is remote from the sample of this study. The adults in Zoo Negara have been placed together since the early 1990s and only bred in 2003 and 2004, while in Sungai Dusun Wildlife Conservation Centre, breeding females held with a mature male have

only laid unfertilised eggs including the adult female that was caught outside of the breeding centre in 2013.

5.7 Field Surveys

Although the paired t-test indicated just temperature was different between seasons, the ANOVA indicated the pH was also different between seasons. Additionally, conductivity, salinity and TDS were also different between rivers (Table 4.21). Hassan *et al*., 2016 indicated that this species might be able to tolerate a wider range of river quality when assessments of the Kepayang River reported a water quality (WQ) of Class II as compared to the other two rivers which had a Class I WQ. This possibility is supported by the habitat range of this species from peat or freshwater swamp forest, to fragmented riverine and saltwater swamp forests (Rodder *et al*., 2010).

From the surveys and areas covered in this study, the locations of the recently caught *T. schlegelii* are mostly in freshwater rivers except in the Tengi River which is drained by smaller non-permanent streams that have higher acidic measures (pH 3-4) and black water which appear during periods after inundation of water (rainy season). An acidic pH was also seen in the Terengganu rivers. When compared to earlier studies in Indonesia (Bezuijen *et al*., 1998; 2002; 2004; Auliya *et al*., 2006; Stanwieswz, 2018) the pH of 3-4 is much lower than the rivers surveyed here but comparable to two out of the three rivers surveyed in Sarawak (Hassan *et al*., 2016). Both the Kepayang River and the Samarahan River (pH between 5.6- 6.7) with similar land use as in Peninsular Malaysia (pass

through or are near oil palm plantations), seems to have similar pH levels, while the relatively intact Bunga-Baki River was more acidic (pH 4.2).

The pH of ground water can be lowered by organic acids from decomposition of vegetation, which is a characteristic of peat, or washed down carbon from land (Gasim *et al*., 2007). A comparison with the study by Alssgeer *et al*. (2017) of the Nerus River which found that the pH range (6.32 and 6.06 for the dry and wet season, respectively) was comparable to this study (6.32 and 5.88 for dry and wet season, respectively). However, the recorded salinity was higher (ranging between 47-66) when compared to this study (ranging between 21.77-20.08). The difference in the salinity could be due to the inclusion of sampling sites further downstream of the Nerus River by Alssgeer *et al*. (2017). Freshwater rivers could have higher salinity due to disturbance of the natural hydrological cycle by irrigation and saline discharges from dryland as well (Hart *et al*., 1990). The same reason could explain the salinity of the Ban Canal and the Tengi River as they are periodically flushed by the Bernam River Water Plant. Additionally, as the Sungkai Rivers is surrounded by oil palm plantations, fertilizers of agriculture land could also affect the soil salinity and subsequently the salinity of rivers (Melles, Cañedo-Argüelles and Derry, 2023).

The night surveys in the Nerus River detected a juvenile *Crocodylus porosus* at the downstream of the survey location. The water parameters between the Nerus and Tekah rivers did not show a difference in the ANOVA analysis indicating that both these crocodilians are able to co-exist in the same habitat albeit the Tekah was significantly less disturbed by human activity. This habitat sharing

phenomenon was also seen in habitat thought to be 'intact' for *T. schlegelii* in Central Kalimantan (*C. porosus*) and East Kalimantan (*C. siamensis*) (Auliya *et al*., 2006; Stanwieswz, 2018). Usually, sympatry in crocodilian is seen between a species with a geographically narrower habitat range with one that has a wider range like *C. porosus*. In the Mesangat Lake, Indonesia, although both *T. schlegelii* and *C. siamensis* are freshwater crocodilian and mound nesters, the nesting season differed between the two (Stanwieswz *et al*., 2018).

5.8 GIS and Terrain Analyses

The terrain analyses of the areas and rivers of the four recent captive *T. schlegelii* also supports movements of females due to breeding instincts. The elevation of 11-32 m and slope of 0-10° in the current dataset seem to corroborate the breeding and nesting ecology reported for other crocodilians. Adugna *et al*. ((2017) and Binaday *et al*. (2012) identified elevation and slope (2-18°) as important for the construction of nests as it enables the guarding female to easily access the river during the breeding season. These were two of the factors incorporated amongst others when habitat suitability maps were generated for *Crocodylus niloticus* and *C. porosus* (Adugna *et al*., 2017; Binaday *et al*., 2012). Similarly, Strahler order in the range of 1-3 in the current dataset indicated the location of capture was in smaller rivers/streams. Although Strahler order for rivers have not been reported in the past for *T. schlegelii*, recent occurrences in Indonesia and Sarawak had been described to be in smaller rivers. In Sarawak, the latest capture of hatchlings and an adult between 2014-2015 were from smaller rivers of between 2-10 m in width namely the Bunga-Baki River and a tributary of the Samarahan River (Hassan *et al*., 2016). Similarly, in Indonesia,

T. schlegelii was found to occur only in the smaller river at the upper tributaries of the Air Hitam Laut River (Shaney *et al*., 2016). Additionally, eggshells were also found here supporting the earlier observation by Bezuijen *et al*. (1998) that nesting sites were away from main and permanent waterway.

The nearest hub showed that all the females, except the one caught adjacent to the Sungai Dusun Wildlife Conservation Centre enclosure, were caught closer to human settlements compared to protected areas (PAs). These settlements are in previously forested land converted for agriculture (oil palm plantations). Being a highly elusive species, the presence of these individuals in closer proximity to human settlements and activities, combined with the mtDNA AMOVA (considering the H4 haplotype), is suggestive of a possible female philopatry in this species. The adult females could have return to their previous nesting sites when the habitat was still intact. Therefore, the ongoing threat of habitat diminishment continue to threaten its nesting locations. The potential female philopatry of *T. schlegelii*, supported by the terrain analyses inferred in the current study suggests that fragmented forest and swamps need to be protected from conversion for urbanization and agriculture use. This would not only serve to maintain a wild population but also mitigate any potential human-crocodile conflict that may arise as seen by the distance to the nearest hub analyses. The data obtained from the GIS analyses could be used as a basis to identify similar terrains to the four rivers surveyed to serve as a refuge which may perhaps reduce the continuation of translocation of wild individuals into captivity.

A three-point scoring system based on abundance, nesting habitat and potential for effective conservation potential was suggested for conservation activities of *T. schlegelii* (Stuebing *et al*., 2006). Similarly, Rodder *et al*. (2010) also accessed conservation areas based on occurrence data, however, genetic component and captive scenarios should also be incorporated in the conservation efforts as well (Miller *et al*., 2006; Koskela *et al*., 2013). Apart from a possible female philopatry of *T. schlegelii*, the mtDNA genetic diversity and microsatellite population structure suggests that Peninsular Malaysia should be considered equally as high priority for conservation efforts and in future conservation strategies for this species.

5.9 Limitations of Study and Recommendations

There are four main limitations in this study. The first is the unknown capture location of most *T. schlegelii* individuals from Peninsular Malaysian with only three from confirmed sites. This hinders proper inferences at the population level analyses. The second is the skewed sex ratios of captive held individuals restricting the male mtDNA AMOVA analyses for a better understanding of the population structure and perhaps movements of *T. schlegelii*. Third is the divergent H4 haplotype. As this dataset and analyses supports Kaur *et al*. (2013) findings that the H4 haplotype is divergent, it is unclear if it could further be a different management unit and hence the uncertainty on how to consider it in statistical analyses for better conclusions. Lastly, the lack of sightings of this crocodilian in the river surveys restricted ecological data for *T. schlegelii* for Peninsular Malaysia.

At the regional level, comprehensive studies should be conducted to document both nuclear gene and mitochondrial diversity to understand the population structure of *T. schlegelii* across wild and farmed populations as well as holding areas globally, with emphasis on accuracy of specimen origin. Populations exhibiting signs of inbreeding should be identified to improve breeding strategies utilizing available genetic information.

The ecological data for this species is minuscule in Peninsular Malaysia. Future studies using telemetry devices would be beneficial to understand if there is habitat preferences / adaptation in addition to documenting movement of the adult throughout the year to shed light on the breeding behaviour of *T. schlegelii*.

Physical barriers or lack thereof, as well as distance and mechanism for differentiation should also be considered. This is particularly relevant in clusters F and E, where full sibling relationships were observed across geographical barriers, highlighting the importance of these factors. It is recommended that existing conservation efforts be updated to iincorporated genetic data by the relevant authorities and bodies.

6.0 CONCLUSION

All the objectives in this study were achieved. For the first objective, microsatellite markers detected six clusters namely Cluster A-F that had a low H_O (0.28-0.49) and H_E (0.28-0.61) within clusters. For the second objective, genetic diversity from the concatenated mtDNA dataset was only seen in Cluster B ($h = 0.553 \pm 0.073$; $\pi = 0.0081 \pm 0.0013$).

Despite the low diversity documented using microsatellite markers, they were more sensitive at the population level with the identification of six clusters compared to the concatenated mtDNA data detecting four haplotypes for the Malaysian populations. Although clusters were represented with closely related individuals, the combination of microsatellite and concatenated mtDNA datasets were able to suggest important clusters (Cluster B and E) for future breeding program for *T. schlegelii* using the coancestry approach. So far, local zoos have not incorporated this approach for any of their captive held species, relying only on pedigree data.

For the third objective, the findings of water parameter (pH 5.7-7.5) of the rivers in this study concurred with river pH reported in previous studies. However, the GIS analyses is the first to describe the terrain (low elevation of 10-32 m; slope 0-10° and Strahler order ≤ 3) of confirmed capture sites of adult female *T. schlegelii* in Peninsular Malaysia. These findings concurred with reported breeding / nesting ecology of crocodilians. The combination of the distance to the nearest hub analyses (≤ 10 km from human settlements), mtDNA AMOVA

analyses and capture information (breeding period of the year), were suggestive of movements of females for breeding purposes. To increase the chances of positive sightings of *T. schlegelii* in future field studies, the GIS findings in this study could be used as a basis to identify survey areas during the breeding months of March-August.

REFERENCES

Adugna, C., Kiros, S., Dejene, T., Asmelash, T. and Hadgu, K.M., 2017. Distribution and habitat suitability of Nile crocodile (*Crocodilus niloticus*, L. 1768) in Tekeze River Dam, Tigray, Ethiopia. *International Journal of Biodiversity and Conservation*, 9(12), pp.350-362.

Alssgeer, H.M.A., Gasim, M.B., Azid, A. and Alabyad, L.O.M., 2017. The evaluation of physico-chemical and biological characteristics of water quality between wet and dry seasons of Nerus River, Kuala Terengganu, Malaysia. *Journal of Fundamental and Applied Sciences*, 9(2S), pp.563-582.

Álvarez, I., Royo, L.J., Pérez-Pardal, L., Fernández, I., Payeras, L., and Goyache, F., 2010. Assessing losses of genetic variability in the endangered Mallorquí horse. *Czech Journal of Animal Science*, 55(10), pp.456–462.

Amavet, P.S., Vilardi, J.C., Rueda, E.C., Larriera, A., and Saidman, B.O., 2012. Mating system and population analysis of the broad-snouted caiman (*Caiman latirostris*) using microsatellite markers. *Amphibia-Reptilia*, 33(1), pp.83-93.

Amavet, P.S., Rueda, E.C., Siroski, P.A., Larriera, A., and Saidman, B.O., 2015. Isolation and characterization of new microsatellite markers for application in population genetic studies of Caiman latirostris and related species. *Amphibia-Reptilia*, 36(2), pp.175-180.

Aparicio, J.M., Ortego, J. and Cordero, P.J., 2006. What should we weigh to estimate heterozygosity, alleles, or loci? *Molecular Ecology*, 15(14), pp.4659-4665.

Arif, I.A. and Khan, H.A., 2009. Molecular markers for biodiversity analysis of wildlife animals: a brief review. *Animal Biodiversity and Conservation*, 32(1), pp.9-17.

Auliya, M., Schwedick, B., Sommerlad, R., Brend, S. and Samedi, 2006. A short-term assessment of the conservation status of *Tomistoma schlegelii* (Crocodylia: Crocodylidae) in Tanjung Puting National Park (Central Kalimantan, Indonesia). [online] Available at: <URL: http//tomistoma.org> [Accessed 29 March 2008].

Avila-Cervantes, J., Arias, C., Venegas-Anaya, M., Vargas, M., Larsson, H.C. and McMillan, W.O., 2021. Effect of the Central American Isthmus on gene flow and divergence of the American crocodile (*Crocodylus acutus*). *Evolution*, 75(2), pp.245-259.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A. and Saunders, N.C., 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, 18(1), pp.489-522.

Avise, J.C., 2004. Molecular markers, Natural History and Evolution (2nd ed.). Sunderland, Massachusetts: Sinauer Associates.

Balaguera-Reina, S.A., Moncada-Jimenez, J.F., Prada-Quiroga, C.F., Hernandez-Gonzalez, F., Bolaños-Cubillos, N.W., Farfán-Ardila, N., Garcia-Calderón, L.M. and Densmore III, L.D., 2020. Tracking a voyager: mitochondrial DNA analyses reveal mainland-to-island dispersal of an American crocodile (*Crocodylus acutus*) across the Caribbean. *Biological Journal of the Linnean Society*, 131(3), pp.647-655.

Balloux, F. and Lugon-Moulin, N., 2002. The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, 11(2), pp.155-165.

Baltazar-Soares, M., Klein, J. D., Correia, S. M., Reischig, T., Taxonera, A., Roque, S. M., ... and Eizaguirre, C. 2020. Distribution of genetic diversity reveals colonization patterns and philopatry of the loggerhead sea turtles across geographic scales. *Scientific reports*, 10(1), pp.18001.

Bashyal, A., Gross, B.A., Venegas-Anaya, M.D., Lowrance, F. and Densmore III, L.D., 2014. Assessment of microsatellites in estimating inter-and intraspecific variation among Neotropical Crocodylus species. Genetics and Molecular Research. [online] Available at: <DOI http://dx.doi.org/10.4238/2014.July.25.2> [Accessed 18 July 2019]

Bayliss, P., Webb, G.J.W., Whitehead, P.J., Dempsey, K. and Smith, A., 1986. Estimating the Abundance of Saltwater Crocodiles, Crocodylus-Porosus Schneider, in Tidal Wetlands of the Northern Territory-a Mark-Recapture Experiment to Correct Spotlight Counts to Absolute Numbers, and the Calibration of Helicopter and Spotlight Count. *Wildlife Research*, 13(2), pp.309-320.

Bayliss, P., 1987. Survey methods and monitoring within crocodile management programmes. In: J.W. Webb, C. Manolis and P. J. Whitehead, eds. 1987. *Wildlife Management: Crocodiles and Alligators*. Winnellie, Northern Territory. Surrey Beatty and Sons Pry Limited. pp.157-175.

Bezuijen, M.R., Hartoyo, P., Elliot, M. and Baker, B.A. Wildlife Management International Pty Limited (1998). *Project Tomistoma. Second report on the ecology of the false gharial (Tomistoma schlegelii) in Sumatra. Darwin*. [online] Available at: < https://www.researchgate.net/publication/237448745> [Accessed 29 March 2007].

Bezuijen, M.R., Webb, G.J. and Hartoyo, P., 2001. Peat swamp forest and the false gharial *Tomistoma schlegelii* (Crocodilia, Reptilia) in the Merang River, eastern Sumatra, Indonesia. *Oryx*, 35(4), pp.301-307.

Bezuijen, M.R., Wibowo, P. and Wirawijaya, H., Wildlife Management International Pty. Limited/Wetlands International-Indonesia Program (2002). *Assessment of the management and conservation of the Merang River as habitat*

for the false gharial. [online] Available at: <http://www.iucncsg.org/365_docs/attachments/protarea/Tomi-278d5918.pdf> [Accessed 29 March 2008].

Bezuijen, M.R., Suryansyah, Huda, I., Andriyono, S., Pratihno, P., Potess, F. and Sommerlad, R., 2004. False gharial (*Tomistoma schlegelii*) surveys in West Kalimantan, Indonesia in 2004. [online] Available at: <URL: http//tomistoma.org> . [Accessed 29 March 2009].

Bezuijen, M.R., Shwedick, B.M., Sommerlad, R., Stevenson, C. and Steubing, R.B., 2010 Tomistoma *schlegelii*. In: S.C. Manolis and C. Stevenson. Ed. *Crocodiles. Status Survey and Conservation Action Plan*. Darwin. Crocodile Specialist Group. 133-138.

Binaday, J.W.B., Manalo, R.I., Bragais, M.A., Palao, L.K.M., Mojica, L.R.P., Nicopior, O.B.S. and Baltazar, P.C., 2021. Determining Habitat Suitability for the Identification of Priority Conservation Sites for Indo-Pacific crocodile (*Crocodylus porosus* Schneider, 1801) in the Philippines. *Philippine Journal of Science*, 150(S1),pp.333-344.

Bittencourt, P.S., Campos, Z., de Lima Muniz, F., Marioni, B., Souza, B.C., Da Silveira, R., de Thoisy, B., Hrbek, T. and Farias, I.P., 2019. Evidence of cryptic lineages within a small South American crocodilian: the Schneider's dwarf caiman, *Paleosuchus trigonatus* (Alligatoridae: Caimaninae). PeerJ, 7, p.e6580. [online] Available at: < https://doi.org/10.7717/peerj.6580 > [Accessed 16 March 2023].

Bogenhagen, D.F., 2012. Mitochondrial DNA nucleoid structure. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 1819(9-10), pp.914-920.

Böhm, M. and Popescu, V.D., 2016. Landscape ecology, biogeography, and GIS methods. *Reptile Ecology and Conservation: A Handbook of Techniques*, *298*. [online] Available at: < https://discovery.ucl.ac.uk/id/eprint/1495953/1/Bohm_Landscape%20ecology.book.chapter.pdf > [Accessed 13 Nov 2023].

Bonke, R., Ihlow, F., Böhme, and W., Rödder, D., 2014. Movement patterns of *Tomistoma schlegelii* in the Sekonyer Kanan River (Tanjung Puting National Park, Central Kalimantan, Indonesia): preliminary range size estimates. *Salamandra*, 50(1), pp.40-52.

Bouckaert, R., *et al*., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biol*ogy, 10:e1003537. [online] Available at: < doi: 10.1371/journal.pcbi.1003537 > [Accessed 14 January 2017].

Bowkett, A.E., 2009. Recent captive-breeding proposals and the return of the ark concept to global species conservation. *Conservation Biology*, 23(3), pp.773-776.

Brackhane, S., Xavier, F.M., Gusmao, M. and Fukuda, Y., 2018. Habitat mapping of the saltwater crocodile (*Crocodylus porosus*) in Timor-Leste. *Herpetological Review*, 49(3), pp.439-441.

Brochu, C.A., 1997. Morphology, fossils, divergence timing, and the phylogenetic relationships of *Gavialis*. *Systematic Biology*, 46 (3), pp.479 – 522.

Brochu, C.A., 2003. Phylogenetic approaches toward crocodylian history. *Annual Review of Earth Planetary Sciences*, 31, pp.357 – 397.

Brochu, C.A., 2007. Systematics and taxonomy of Eocene Tomistomine crocodylians from Britian and Northern Europe. *Palaeontology*, 50 (4), pp.917 – 928.

Brochu, C.A. and Sumrall, C.D., 2020. Modern cryptic species and crocodylian diversity in the fossil record. *Zoological Journal of the Linnean Society*, 189(2), pp.700-711.

Budd, K.M., Spotila, J.R., and Mauger, L.A., 2015. Preliminary Mating Analysis of American Crocodiles, *Crocodylus acutus*, in Las Baulas, Santa Rosa, and Palo Verde National Parks, Guanacaste, Costa Rica. *South American Journal of Herpetology*, 10(1), pp.4-9.

Burke, P.M. and Mannion, P.D., 2023. Neuroanatomy of the crocodylian *Tomistoma dowsoni* from the Miocene of North Africa provides insights into the evolutionary history of gavialoids. *Journal of Anatomy*, 243(1), pp.1-22.

Burzyński, A., Zbawicka, M., Skibinski, D.O. and Wenne, R., 2003. Evidence for recombination of mtDNA in the marine mussel *Mytilus trossulus* from the Baltic. *Molecular Biology and Evolution*, 20(3), pp.388-392.

Caballero, A. and Toro, M.A., 2002. Analysis of genetic diversity for the management of conserved subdivided populations. *Conservation Genetics*, 3, pp.289-299.

Cao, R., Somaweera, R., Brittain, K., FitzSimmons, N.N., Georges, A., and Gongora, J., 2020. Genetic structure and diversity of Australian freshwater crocodiles (*Crocodylus johnstoni*) from the Kimberley, Western Australia. *Conservation Genetics*, 21, pp.421-429.

Chaeychomsri, W., Chaeychomsri, S., Tuntirungkij, M., Tabthipwon, P., Noparatnaraporn, N., and Siripholvat, V., 2008. Characterization of microsatellite markers for the Siamese crocodile and amplification in the closely related genus Crocodylus. *Agriculture and Natural Resources*, 42(4), pp.682-692.

Chang, Y.C., Li, S.H., Chang, M.H., and Chen, S.L., 2014. Isolation and characterization of 18 microsatellite loci for an endangered freshwater crocodile *Tomistoma schlegelii*. *Conservation Genetics Resources*, 6, pp.495-498.

Chen, B., Cole, J.W. and Grond-Ginsbach, C., 2017. Departure from Hardy Weinberg equilibrium and genotyping error. *Frontiers in Genetics*, 8, p.300354. [online] Available at: <https://doi.org/10.3389/fgene.2017.00167> [Accessed 29 March 2024].

Clutton-Brock, T.H. and Lukas, D., 2012. The evolution of social philopatry and dispersal in female mammals. *Molecular Ecology*, 21(3), pp.472-492.

Congalton, R.G. and Green, K., 1992. The abcs of GIS. *Journal of Forestry*, 90(11), pp.13-20.

Cox, J. and Gombek, F., 1985. *A preliminary survey of the crocodile resource in Sarawak, East Malaysia*. World Wildlife Fund Malaysia.

Darlim, G., Lee, M.S., Walter, J. and Rabi, M., 2022. The impact of molecular data on the phylogenetic position of the putative oldest crown crocodilian and the age of the clade. *Biology Letters*, 18(2), p.20210603. [online] Available at: <https://doi.org/10.1098/rsbl.2021.0603> [Accessed 27 December 2023].

Darriba, D., Taboada, G.L., Doallo, R. and Posada, D., 2012. jModelTest 2: more models, new heuristics, and high-performance computing. *Nature Methods*, 9(8), p.772.

Davenport, M., 1995. Evidence of possible sperm storage in the caiman, *Paleosuchus palpebrosus*. *Herpetological Review*, 26(1), pp.14-14.

Davis, L.M., Glenn, T.C., Elsey, R.M., Dessauer, H.C., and Sawyer, R.H., 2001. Multiple paternity and mating patterns in the American alligator, *Alligator mississippiensis*. *Molecular Ecology*, 10(4), pp.1011-1024.

de Oliveira, D.P., Farias, I.P., Marioni, B., Campos, Z. and Hrbek, T., 2010. Microsatellite markers for mating system and population analyses of the spectacled *Caiman crocodilus* (Linnaeus 1758). *Conservation Genetics Resources*, 2, pp.181-184.

DeSalle, R., Schierwater, B. and Hadrys, H., 2017. MtDNA: The small workhorse of evolutionary studies. *Frontiers in Bioscience*, 22, pp.873-887.

de Thoisy, B., Hrbek, T., Farias, I.P., Vasconcelos, W.R., and Lavergne, A., 2006. Genetic structure, population dynamics, and conservation of black caiman (*Melanosuchus niger*). *Biological Conservation*, 133(4), pp.474-482.

Dever, J.A., Strauss, R.E., Rainwater, T.R., McMurry, S.T., and Densmore III, L.D., 2002. Genetic diversity, population subdivision, and gene flow in Morelet's crocodile (*Crocodylus moreletii*) from Belize, Central America. *Copeia*, 2002(4), pp.1078-1091.

Díaz-Moreno, D.M., Hernández-Gonzalez, F., Moncada-Jimenez, J.F., Mora, C., Prada, C., Jiménez-Alonso, G., and Balaguera-Reina, S.A., 2021. Molecular

characterization of the spectacled caiman (*Caiman crocodilus*) in the upper Magdalena River basin, Colombia: demographic and phylogeographic insights. *Systematics and Biodiversity*, 19(8), pp.1040-1048.

Dubey, S., Brown, G. P., Madsen, T. and Shine, R. 2008. Male-biased dispersal in a tropical Australian snake (*Stegonotus cucullatus*, Colubridae). *Molecular Ecology*, 17(15), pp. 3506-3514.

Earl, D. A. and VonHoldt, B. M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, pp.359-361.

Escorza-Treviño, S. and Dizon, A.E., 2000. Phylogeography, intraspecific structure and sex-biased dispersal of Dall's porpoise, Phocoenoides dalli, revealed by mitochondrial and microsatellite DNA analyses. *Molecular Ecology*, 9(8), pp.1049-1060.

Excoffier, L. and Lischer, H.E., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), pp.564-567.

Falush, D., Stephens, M., and Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, 164(4), pp.1567-1587.

FitzSimmons, N. N., Limpus, C. J., Norman, J. A., Goldizen, A. R., Miller, J. D. and Moritz, C. (1997). Philopatry of male marine turtles inferred from mitochondrial DNA markers. *Proceedings of the National Academy of Sciences*, 94(16), pp.8912-8917.

Galla, S.J., Brown, L., Couch-Lewis, Y., Cubrinovska, I., Eason, D., Gooley, R.M., Hamilton, J.A., Heath, J.A., Hauser, S.S., Latch, E.K. and Matocq, M.D., 2022. The relevance of pedigrees in the conservation genomics era. *Molecular Ecology*. 31, pp.41–54.

Gasim, M.B., Ismail, B.S., Toriman, E., Mir, S.I. and Chek, T.C., 2007. A physico-chemical assessment of the Bebar River, Pahang, Malaysia. *Global Journal of Environmental Research*, 1(1), pp.7-11.

Gatesy, J., Amato, G., Norell, M., DeSalle, R. and Hayashi, C., 2003. Combined support for wholesale taxic atavism in Gavialine crocodylians. *Systematic Biology*, 52 (3), pp.403-422.

Gist, D.H., Bagwill, A., Lance, V., Sever, D.M., and Elsey, R.M., 2008. Sperm storage in the oviduct of the American alligator. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, 309(10), pp.581-587.

Glenn, T.C., Dessauer, H.C., and Braun, M.J., 1998. Characterization of microsatellite DNA loci in American alligators. *Copeia*, pp.591-601.

Gonçalves, R.A., Santos, C.H.A., de Sá Leitão, C.S., de Souza, É.M.S. and de Almeida-Val, V.M.F., 2019. Genetic basis of *Colossoma macropomum* broodstock: Perspectives for an improvement program. *Journal of the World Aquaculture Society*, 50(3), pp.633-644.

Goudet, J. 2003. Fstat (ver. 2.9.4), a program to estimate and test population genetics parameters. [online] Available at: <http://www.t-de-meeus.fr/Programs/Fstat294.zip> [Accessed 5 May 2021].

Griffiths, R.A. and Pavajeau, L., 2008. Captive breeding, reintroduction, and the conservation of amphibians. *Conservation Biology*, 22(4), pp.852-861.

Grosso, A.R., Bastos-Silveira, C., Coelho, M.M. and Dias, D., 2006. *Columba palumbus* Cyt b-like Numt sequence: comparison with functional homologue and the use of universal primers. *Folia Zoologica*, 55(2), p.131. [online] Available at: <https://www.researchgate.net/profile/Cristiane-Bastos-Silveira/publication/228342532> [Accessed 29 March 2024].

Gunarso, P., Hartoyo, M.E., Agus, F. and Killeen, T.J., 2013. Oil Palm and Land Use Change in Indonesia, Malaysia, and Papua New Guinea. T. J., Killeen and J. Goon. [online] Available at: < file:///C:/Users/TARA/Downloads/4_oil_palm_and_land_use_change_Gunarso_et_al.pdf > [Accessed 16 August 2016].

Gutiérrez, J. P., Royo, L. J., Álvarez, I., and Goyache, F., 2005. MolKin v2. 0: a computer program for genetic analysis of populations using molecular coancestry information. *Journal of Heredity*, 96(6), pp.718-721.

Hammond, R.L., Lawson Handley, L.J., Winney, B.J., Bruford, M.W. and Perrin, N., 2006. Genetic evidence for female-biased dispersal and gene flow in a polygynous primate. *Proceedings of the Royal Society B: Biological Sciences*, 273(1585), pp.479-484.

Hariyadi, I., Machrizal, R., Dimenta, R.H., Khairul, K., Hasibuan, R. and Gultom, H.S.B. 2019. Fish biodiversity in false gharial habitat (*Tomistoma schlegelii* Müller, 1838) in Labuhan Batu district. In IOP Conference Series: Earth and Environmental Science. 348(1), 1-6. [online] Available at: <https://iopscience.iop.org/article/10.1088/1755-1315/348/1/012027/pdf> [Accessed 29 March 2008].

Harshman, J., Huddleston, C.J., Bollback, J.P., Parsons, T.J. and Braun, M., 2003. True and false gharials: A nuclear gene phylogeny of Crocodylia. *Systematic Biology*, 52 (3), pp.386 – 402.

Hart, B.T., Bailey, P., Edwards, R., Hortle, K., James, K., McMahon, A., Meredith, C. and Swadling, K., 1990. Effects of salinity on river, stream and wetland ecosystems in Victoria, Australia. *Water Research*, 24(9), pp.1103-1117.

Hassan, R., Koh, H.E. and Abdullah, N.S.S., 2008. Potential molecular markers for estuarine crocodiles (*Crocodylus porosus*) from Sarawak. [online] Available at: <https://ir.uitm.edu.my/id/eprint/58826/1/58826.PDF> [Accessed 4 June 2018]

Hassan, R., Ahmad, R., Md Adzhar, M.A.A., Abdul Gani, M.I.Z., Ayob, A., and Zainudin, R., 2016. Notes on the wild *Tomistoma* populations in western Sarawak, Malaysian Borneo. *International Journal of Ecology*, 2016. [online] Available at: <https://doi.org/10.1155/2016/4357623> [Accessed 4 June 2018]

Hedrick, P.W., 2001. Conservation genetics: where are we now? *Trends in Ecology and Evolution*, 16(11), pp.629-636.

Hekkala, E.R., 2004. *Conservation genetics at the species boundary: case studies from African and Caribbean crocodiles (Genus: Crocodylus)*. PhD. Columbia University.

Hekkala, E.R., Amato, G., DeSalle, R., and Blum, M.J., 2010. Molecular assessment of population differentiation and individual assignment potential of Nile crocodile (*Crocodylus niloticus*) populations. *Conservation Genetics*, 11, pp.1435-1443.

Heras, J., Domínguez, C., Mata, E., Pascual, V., Lozano, C., Torres, C., Zarazaga, M., 2015. GelJ–a tool for analyzing DNA fingerprint gel images. *BMC Bioinformatics*. 16(1), pp.1-8.

Hinlo, M.R.P., *et al*., 2014. Population genetics implications for the conservation of the Philippine Crocodile *Crocodylus mindorensis* Schmidt, 1935 (Crocodylia: Crocodylidae). *Journal of Threatened Taxa*, 6(3), pp.5513-5533.

Hu, Y., and Wu, X.B., 2010. Multiple paternity in Chinese alligator (Alligator sinensis) clutches during a reproductive season at Xuanzhou Nature Reserve. *Amphibia-Reptilia*, 31(3), pp.419-424.

Huang, L., and Wang, Y.Q., 2004. SSR polymorphism of *Alligator sinensis* and conservation strategy of genetic diversity. *Acta Genetica Sinica*, 31(2), pp.143-150.

Hwang, U.W. and Kim, W., 1999. General properties and phylogenetic utilities of nuclear ribosomal DNA and mitochondrial DNA commonly used in molecular systematics. *The Korean Journal of Parasitology*, 37(4), pp.215. [online] Available at: <10.3347/kjp.1999.37.4.215>[Accessed 4 March 2023]

Iijima, M., Qiao, Y., Lin, W., Peng, Y., Yoneda, M. and Liu, J. 2022. An intermediate crocodylian linking two extant gharials from the Bronze Age of China and its human-induced extinction. Proceedings of the Royal Society B. 289: 20220085 [Accessed 3rd April 2025].

IUCN, 2011: The IUCN Red List of Threatened Species. Version 2009.1. [online] Available at: < URL: www.iucnredlist.org.> [Accessed 3rd August 2018].

IUCN, 2022. The IUCN Red List of Threatened Species v 2023-1. [online] Available at: < URL: www.iucnredlist.org.> [Accessed 6 May 2024].

Janke, A., Gyllberg, A., Hughes, S., Aggarwal, R.K. and Arnason, U. 2005. Mitogenomic analyses place the Gharial (*Gavialis gangeticus*) on the crocodile tree and provide pre-K / T divergence times for most crocodilians. *Journal of Molecular Evolution*, 61, pp.620 – 626.

Jing, W., Wang, X.L., Lan, H. and Fang, S.G. 2009. Eleven novel microsatellite markers for the Chinese alligator (*Alligator sinensis*). *Conservation Genetics*, 10, pp.543-546.

Jones, A.G. and Ardren, W.R., 2003. Methods of parentage analysis in natural populations. *Molecular Ecology*, 12(10), pp.2511-2523.

Jones, O.R. and Wang, J., 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, 10(3), pp.551-555.

Kalinowski, S.T., Wagner, A.P., and Taper, M.L., 2006. ML-Relate: a computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes*, 6(2), pp.576-579.

Kar, S.K. and Bustard, H.R., 1990. Results of a pilot saltwater crocodile *Crocodylus porosus* Schneider restocking in Bhitarkanika wildlife sanctuary, Orissa. *Journal of the Bombay Natural History Society*, 87(2), pp.195-200.

Kaur, T. and Ong, A.H., 2011. Heteroplasmy, length, and sequence characterization of the mitochondrial control region in *Tomistoma schlegelii*. *Biochemical Genetics*, 49, pp.562-575.

Kaur, T., Japning, J.R.R., Sabki, M.S., Sidik, I., Chong, L.K., and Ong, A.H., 2013. Genetic diversity of *Tomistoma schlegelii* inferred from mtDNA markers. *Biochemical Genetics*, 51, pp. 275-295.

Kidd, D.M. and Ritchie, M.G., 2006. Phylogeographic information systems: putting the geography into phylogeography. *Journal of Biogeography*, 33(11), pp.1851-1865.

Kim, K.S. and Sappington, T.W., 2013. Microsatellite data analysis for population genetics. In: S. K. Kantartzi, ed. 2013. *Microsatellites: Methods and protocols*. Springer Science+Business Media pp.271-295.

Kobayashi, Y. and Tomida, Y., 2005. Anatomy of *Toyotomaphimeia machikanensis* (Crocodylia) from the Pleistocene of Japan and reassessment of its phylogenetic status. *Journal of Vertebrate Palaeontology*, 25, pp.78A–79A.

Koskela, J., Lefèvre, F., Schueler, S., Kraigher, H., Olrik, D.C., Hubert, J., Longauer, R., Bozzano, M., Yrjänä, L., Alizoti, P. and Rotach, P., 2013. Translating conservation genetics into management: Pan-European minimum requirements for dynamic conservation units of forest tree genetic diversity. *Biological Conservation*, 157, pp.39-49.

Kristensen, T.N., Hoffmann, A.A., Pertoldi, C. and Stronen, A.V., 2015. What can livestock breeders learn from conservation genetics and vice versa? *Frontiers in Genetics*, 6, p.121860. [online] Available at: <http://www.frontiersin.org/Genetics/editorialboard> [Accessed 4 February 2023]

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35(6), pp.1547.

Kurniati, H., 2003. Control Region-mtDNA Heterogeneity of Kalimantan False Gharial (Tomistoma schlegelii) Population: A Preliminary Study. *Berkala Penelitian Hayati*, 9(1), pp.19-25.

Lance, S.L., Tuberville, T.D., Dueck, L., Holz-Schietinger, C., Trosclair Iii, P.L., Elsey, R.M., and Glenn, T.C., 2009. Multiyear multiple paternity and mate fidelity in the American alligator, *Alligator mississippiensis*. *Molecular Ecology*, 18(21), pp.4508-4520.

Langella, O., 2015. Populations 1.2. 32. Population Genetic Software. Gif-sur-Yvette, France: Laboratoire Evolution, Génomes et Spéciation.

Lariman, D. ., and Trimurti, S., 2021. Study of false gharial (*Tomistoma schlegelii*) habitat characteristics in Mesangat wetlands, East Kutai, East Kalimantan. *International Journal of Science and Research*, 10(12), pp.649-653.

Leigh, J. W., and Bryant, D., 2015. POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6(9), pp.1110-1116.

Levine, B.A., Schuett, G.W., and Booth, W., 2021. Exceptional long-term sperm storage by a female vertebrate. *Plos One*, 16(6), e0252049. [online] Available at: <https://doi.org/10.1371/journal.pone.0252049>[Accessed 18March 2024].

Lewis, J. L., FitzSimmons, N.N., Jamerlan, M.L., Buchan, J.C., and Grigg, G.C., 2013. Mating systems and multiple paternity in the estuarine crocodile (*Crocodylus porosus*). *Journal of Herpetology*, 47(1), pp.24-33.

Li, Y., Wu, X., Ji, X., Yan, P. and Amato, G. 2007. The complete mitochondrial genome of salt-water crocodile (Crocodylus porosus) and phylogeny of crocodilians. Journal of Genetics and Genomics, 34(2), pp.119-128.

Luck, N.L., Thomas, K.C., Morin-Adeline, V.E., Barwick, S., Chong, A.Y., Carpenter, E.L., Wan, L., Willet, C.E., Langford-Salisbury, S.M., Abdelsayd, M. and Ang, R.A., 2012. Mitochondrial DNA analyses of the saltwater crocodile

(*Crocodylus porosus*) from the Northern Territory of Australia. *Australian Journal of Zoology*, 60(1), pp.18-25.

MacMillan, R.A., Moon, D.E. and Coupé, R.A., 2007. Automated predictive ecological mapping in a forest region of BC, Canada, 2001–2005. *Geoderma*, 140(4), pp.353-373.

Mascarenhas-Junior, P.B., Sousa Correia, J.M. and Simões, P.I., 2023. Tracking crocodylia: a review of telemetry studies on movements and spatial use. *Animal Biotelemetry*, 11(1), pp.21. [online] Available at: < https://doi.org/10.1186/s40317-023-00333-2 > [Accessed 17 March 2024]

Mathew, A., Ganesan, M., Majid, R.A., and Beastall, C., 2011. Breeding of False Gharial (*Tomistoma schlegelii*) at Zoo Negara, Malaysia. [online] Available at: < https://www.zoonegara.my/RPFalseGharial.pdf.> [Accessed September 25, 2012].

Mauger, L.A., Velez, E., Cherkiss, M.S., Brien, M.L., Mazzotti, F.J. and Spotila, J.R., 2017. Conservation genetics of American crocodile, *Crocodylus acutus*, populations in Pacific Costa Rica. *Nature Conservation*, 17, pp.1-17.

McAliley, L.R., Willis, R.E., Ray, D.A., White, P.S., Brochu, C.A., and Densmore, L.D., 2006. Are crocodiles really monophyletic? – Evidence for subdivisions from sequence and morphological data. *Molecular Phylogenetics and Evolution*, 39, pp.16 – 32.

McVay, J. D., *et al*., 2008. Evidence of multiple paternity in Morelet's crocodile (*Crocodylus moreletii*) in Belize, CA, inferred from microsatellite markers. Journal of Experimental Zoology Part A: *Ecological Genetics and Physiology*, 309(10), pp.643-648.

Md Adzhar, M.A.A., and Hassan, R., 2017. Relationships among *Tomistoma schlegelii* in Malaysia based on Cyt *b*-control region gene analysis. *International Journal of Zoology*, 2017. [online] Available at: < https://doi.org/10.1155/2017/5431041 [Accessed 4 June 2018]

Meganathan, P.R., Dubey, B. and Haque, I., 2009. Molecular identification of crocodile species using novel primers for forensic analysis. *Conservation Genetics*, 10, pp.767-770.

Meganathan, P.R., Dubey, B., Jogayya, K.N. and Haque, I., 2013. Identification of Indian crocodile species through DNA barcodes. *Journal of Forensic Sciences*, 58(4), pp.993-998.

Melles, S.J., Cañedo-Argüelles, M. and Derry, A.M., 2023. Documenting the impacts of increasing salinity in freshwater and coastal ecosystems: Introduction to the special issue. *Limnology and Oceanography Letters*. [online] Available at: <doi: 10.1002/lol2.10307>[Accessed 4 June 2018]

Melnick, D.J. and Hoelzer, G.A., 1992. Differences in male and female macaque dispersal lead to contrasting distributions of nuclear and mitochondrial DNA variation. *International Journal of Primatology*, 13, pp.379-393.

Messel, H., 1981. *The Blyth-Cadell Rivers System Study and the Status of Crocodylus porosus in Tidal Waterways of Northern Australia: Methods for Analysis and Dynamics of a Population of C. porosus* (Vol. 1). Elsevier Science and Technology.

Milián-García, Y., *et al*., 2015. Genetic evidence of hybridization between the critically endangered Cuban crocodile and the American crocodile: implications for population history and in situ/ex situ conservation. *Heredity*, 114(3), pp.272-280.

Miller, R.M., Rodríguez, J.P., Aniskowicz-Fowler, T., Bambaradeniya, C., Boles, R., Eaton, M.A., Gärdenfors, U., Keller, V., Molur, S., Walker, S. and Pollock, C., 2006. Extinction risk and conservation priorities. *Science*, 313(5786), pp.441-441.

Montague, J.J., 1983. Influence of water level, hunting pressure and habitat type on crocodile abundance in the Fly River drainage, Papua New Guinea. *Biological Conservation*, 26(4), pp.309-339.

Moore, C.T., 1994. *Single-site and regional surveys of crocodilians or crocodilian sign: design and inference considerations for estimation of population trend. Crocodiles*. Pattaya, Thailand, 1994. IUCN. Gland, Switzerland.

Moore, B.C., Fitri, W.N. and Augustine, L., 2020. Crocodylian conservation and evolution insights from an anatomical and histological examination of phalli from male false gharial (*Tomistoma schlegelii*). *Anatomia, Histologia, Embryologia*, 49(3), pp.390-401.

Muniz, F.L., Ximenes, A.M., Bittencourt, P.S., Hernández-Rangel, S. M., Campos, Z., Hrbek, T., and Farias, I. P., 2019. Detecting population structure of *Paleosuchus trigonatus* (Alligatoridae: Caimaninae) through microsatellites markers developed by next generation sequencing. *Molecular Biology Reports*, 46(2), pp.2473-2484.

Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York, pp. 512.

Ojeda, G.N., Amavet, P.S., Rueda, E.C., Siroski, P.A., and Larriera, A., 2017. Mating system of *Caiman yacare* (Reptilia: Alligatoridae) described from microsatellite genotypes. *Journal of Heredity*, 108(2), pp.135-141.

Oliveira, D.P., Marioni, B., Farias, I.P. and Hrbek, T., 2014. Genetic evidence for polygamy as a mating strategy in *Caiman crocodilus*. *Journal of Heredity*, 105(4), pp.485-492.

Pacheco, L. F. 1994. Estimating crocodilian abundance in forest lagoons. In: Group, Proceedings of the 12th Working Meeting of the IUCN-SSC Crocodile Specialist. Crocodiles, pp. 241–244. IUCN, Gland.

Pacheco, L.F., 1996. Wariness of caiman populations and its effect on abundance estimates. *Journal of Herpetology*, 30(1), pp.123-126.

Pan, T., Miao, J.S., Zhang, H.B., Yan, P., Lee, P.S., Jiang, X.Y., Ouyang, J.H., Deng, Y.P., Zhang, B.W. and Wu, X.B., 2021. Near-complete phylogeny of extant Crocodylia (Reptilia) using mitogenome-based data. *Zoological Journal of the Linnean Society*, 191(4), pp.1075-1089.

Peakall, R. and Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), pp.288-295.

Petit, R.J., El Mousadik, A. and Pons, O., 1998. Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, 12(4), pp.844-855.

Piras, P., Delfino, M., Del Favero, L. and Kotsakis, T., 2007. Phylogenetic position of the crocodylian *Megadontosuchus arduini* and Tomistomine palaeobiogeography. *Acta Palaeontologica Polonica*, 52 (2), pp.315 – 328.

Presti, F.T. and Wasko, A.P., 2014. A review of microsatellite markers and their application on genetic diversity studies in parrots. *Open Journal of Genetics*, pp.69-77.

Pritchard, J.K., Stephens, M., and Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155(2), pp.945-959.

Posa, M.R.C., Wijedasa, L.S. and Corlett, R.T. (2011). Biodiversity and conservation of tropical peat swamp forests. *BioScience*, 61(1), pp.49-57.

Putman, A.I. and Carbone, I., 2014. Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecology and Evolution*, 4(22), pp.4399-4428.

Rambaut, A., 2018. FigTree–Tree Figure Drawing Tool Version v. 1.4. 4. Institute of Evolutionary Biology, University of Edinburgh: Edinburgh.

Ray, D.A., White, P.S., Duong, H.V., Cullen, T. and Densmore, L.D., 2000. High levels of genetic variability in West 125 African dwarf crocodiles *Osteolaemus tetraspsis tetraspis* In C. Grigg (ed). 2000. *Crocodilian Biology and Evolution*. Chipping Norton: Surrey Beatty and Sons. 58 – 63.

Ray, D.A. and Densmore, L. 2002. The crocodilian mitochondrial control region: general structure, conserved sequences, and evolutionary implications. *Journal of Experimental Zoology*, 294(4), pp.334-345.

Ray, D.A. and Densmore, L.D. 2003. Repetitive sequences in the crocodilian mitochondrial control region: poly-A sequences and heteroplasmic tandem repeats. *Molecular Biology and Evolution*, 20(6), pp.1006-1013.

Read, M.A., Grigg, G.C., Irwin, S.R., Shanahan, D. and Franklin, C.E., 2007. Satellite tracking reveals long distance coastal travel and homing by translocated estuarine crocodiles, *Crocodylus porosus*. PLoS one, 2(9), p.e949. [online] Available at: < [online] Available at: < [Accessed 1 December, 2020]

Richards, C.L., Carstens, B.C. and Lacey Knowles, L., 2007. Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. *Journal of Biogeography*, 34(11), pp.1833-1845.

Richardson, D.M. and Whittaker, R.J., 2010. Conservation biogeography–foundations, concepts, and challenges. *Diversity and Distributions*, 16(3), pp.313-320.

Rödder, D., Engler, J.O., Bonke, R., Weinsheimer, F. and Pertel, W., 2010. Fading of the last giants: an assessment of habitat availability of the Sunda gharial *Tomistoma schlegelii* and coverage with protected areas. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 20(6), pp.678-684.

Rodriguez, D., Forstner, M.R., Moler, P.E., Wasilewski, J.A., Cherkiss, M.S. and Densmore III, L.D., 2011. Effect of human-mediated migration and hybridization on the recovery of the American crocodile in Florida (USA). *Conservation Genetics*, 12, pp.449-459.

Roos, J., Aggarwal, R.K. and Janke, A., 2007. Extended mitogenomic phylogenetic analyses yield new insight into crocodylian evolution and their survival of the Cretaceous–Tertiary boundary. *Molecular Phylogenetics and Evolution*, 45(2), pp.663-673.

Rosca, A., Juca, I., Timbota, O., Belin, V., Bertici, R. and Herbei, M., 2020. Methods for digitalizing information from analogic support and creating GIS databases. *Research Journal of Agricultural Science*, 52(4), pp.104-112.

Ross, J.P., 1998. *Crocodiles: status survey and conservation action plan* Ed. 2. Gland, Switzerland. IUCN, The World Conservation Union, Publications Services Unit.

Rossi, N.A., *et. al*, 2020. High levels of population genetic differentiation in the American crocodile (*Crocodylus acutus*). Plos One, 15(7), e0235288. [online] Available at: < https://doi.org/10.1371/journal.pone.0235288 [Accessed 12 December, 2021]

Rousset, F., 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, 8(1), pp.103-106.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., and Sánchez-Gracia, A., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, 34(12), pp.3299-3302.

Rubinoff, D. and Holland, B.S., 2005. Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic Biology*, 54(6), pp.952-961.

Salem, A.H.I., 2013. Habitat vulnerability for the Nile crocodile (*Crocodylus niloticus*) in Nasser Lake (Egypt). *Transylvanian Review of Systematical and Ecological Research*, 15(1), pp.19-32.

Schillaci, C., Braun, A. and Kropácek, J., 2015 Terrain Analysis and Landform Recognition. In: *Geomorphological Techniques*. London British Society for Geomorphology. Chap. 2, 4. 2.

Sebastian, A.C., 1993. The crocodilians of Malaysia. *Crocodiles, Proceedings of the Second Regional Meeting of the Crocodile Specialist Group*. Gland. 1993. IUCN, Gland, Switzerland.

Selkoe, K. A., and Toonen, R. J., 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, 9(5), pp.615-629.

Setiawan, A., Safri, A., Prima, B., Kadarisman, R., Indrianti, W. and Oktaberi, R., 2019. Recent Observation of False Gharial (*Tomistoma schlegelii*) in the Sembilang Natonal Park (TNS) Area, Banyuasin Regency, South Sumatera Province. *Biovalentia*: *Biological Research Journal*, 5(1), pp.8-13.

Shafiei-Astani, B., *et al*., 2015. Molecular genetic variation and structure of Southeast Asian crocodile (*Tomistoma schlegelii*): comparative potentials of SSRs versus ISSRs. *Gene*, 571(1), pp.107-116.

Shaney, K.J., Hamidy, A., Arida, A., Arimbi, A., Sismanto, P., Erwan, P. and Smith, E.N., 2016. *A Report on the Status of the False Gharial (T. schlegelii) in Sumatra, 2016*. Crocodile Specialist Group Report. [online] Available at: < http://www.iucncsg.org/365_docs/attachments/protarea/f20b70f0a8b97200af0dd496c2a60c95.pdf > [Accessed 17 March 2018].

Shaney, K.J., Hamidy, A., Walsh, M., Arida, E., Arimbi, A. and Smith, E.N., 2019. Impacts of anthropogenic pressures on the contemporary biogeography of threatened crocodilians in Indonesia. *Oryx*, 53(3), pp.570–581

Shirley, M.H., 2013. *Hierarchical processes structuring crocodile populations in Central Africa*. PhD. University of Florida.

Shirley, M.H., Villanova, V.L., Vliet, K.A., and Austin, J.D., 2015. Genetic barcoding facilitates captive and wild management of three cryptic African crocodile species complexes. *Animal Conservation*, 18(4), pp.322-330.

Sideleau, B., Staniewicz, A., Syah, M. and Shaney, K.J., 2022. An analysis of tomistoma (*Tomistoma schlegelii*) attacks on humans. *Marine and Freshwater Researc*h, 73(11), pp.1331-1338.

Simpson, B.K., Lopez, A., Latif, S. and Yusoh, A., 1998. *Tomistoma (Tomistoma schlegelii) at Tasek Bera, Peninsular Malaysia.* In Crocodiles. Proceedings of the 14th Working Meeting of the Crocodile Specialist Group, IUCN-The World Conservation Union, 1998 (pp. 32-45). Gland: IUCN/SSC Crocodile Specialist Group.

Simpson, B.K., 2014. *Status Assessment of Tomistoma in Peninsular Malaysia: Peat Swamp Forests of Selangor and Pahang.* International Union for Conservation of Nature. IUCN Species Survival Commission (SSC), Crocodile Specialist Group, Tomistoma Task Force. [online] Available at: <https://policycommons.net/artifacts/1368862/status-assessment-of-tomistoma-in-peninsular-malaysia/1983036> [Accessed 12 April 2016].

Slate, J., Marshall, T. and Pemberton, J., 2000. A retrospective assessment of the accuracy of the paternity inference program CERVUS. *Molecular Ecology*, 9(6), pp.801-808.

Slatkin, M., 1994. Linkage disequilibrium in growing and stable populations. *Genetics*, 137(1), pp.331-336.

Smolensky, N.L., Hurtado, L.A., and Fitzgerald, L.A., 2015. DNA barcoding of Cameroon samples enhances our knowledge on the distributional limits of putative species of *Osteolaemus* (African dwarf crocodiles). *Conservation Genetics*, 16, pp.235-240.

Smouse, P.E., Banks, S.C. and Peakall, R., 2017. Converting quadratic entropy to diversity: Both animals and alleles are diverse, but some are more diverse than others. PLoS One, 12(10), p.e0185499., [online] Available at: <https://doi.org/10.1371/journal.pone.0185499 > [Accessed 9 March 2020]

Soglia, D., *et al*., 2021. Genetic diversity of 17 autochthonous Italian chicken breeds and their extinction risk status. *Frontiers in Genetics*, 12, 715656. [online] Available at: <https://doi.org/10.3389/fgene.2021.715656> [Accessed 9 October 2023]

Spice, E.K., Goodman, D.H., Reid, S.B. and Docker, M.F., 2012. Neither philopatric nor panmictic: microsatellite and mtDNA evidence suggests lack of natal homing but limits to dispersal in Pacific lamprey. *Molecular Ecology*, 21(12), pp.2916-2930.

Staniewicz, A. 2011. *Diet and demography of Tomistoma schlegelii in Mesangat Lake, East Kalimantan, Indonesia* (On-line pdf). EAZA IUCN/SSC Southeast Asia Campaign. [online] Available at: <http://www.southeastasiacampaign.org/download/publications/2011%20Mesangat%20Tomistoma%20report.pdf.> [Accessed 6 September 2016]

Staniewicz, A., Behler, N., Dharmasyah, S., and Jones, G., 2018. Niche partitioning between juvenile sympatric crocodilians in Mesangat Lake, East Kalimantan, Indonesia. *Raffles Bulletin of Zoology*, 66, pp.528–537.

Staniewicz, A., Youngprapakorn, U. and Jones, G., 2018. First report of physiological color change in a crocodilian. *Copeia*, 106(2), pp.264-267.

Stuebing, R., Sah, M., Lading, E. and Jong, J., 2004. The status of *Tomistoma schlegelii* (Mueller) in Malaysia. In Proceedings of the 17th Working Meeting of the IUCN-SSC Crocodile Specialist Group (Vol. 1, pp. 136-140).

Stuebing, R.B., Bezuijen, M.R., Auliya, M. and Voris, H.K., 2006. The current and historic distribution of *Tomistoma schlegelii* (the False Gharial)(Müller, 1838)(Crocodylia, Reptilia). *The Raffles Bulletin of Zoology*, 54(1), pp.181-197.

Stuebing, R., Sommerlad, R. and Staniewicz, A., 2015. Conservation of the Sunda gharial *Tomistoma schlegelii* in Lake Mesangat, Indonesia. *International Zoo Yearbook*, 49(1), pp.137-149. [online] Available at: < DOI:10.1111/izy.12068 > [Accessed 29 August 2016]

Taanman, J.W., 1999. The mitochondrial genome: structure, transcription, translation, and replication. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 1410(2), pp.103-123.

Tabora, J.A.G., Hinlo, M.R.P., Bailey, C.A., Lei, R., Pomares, C.C., Rebong, G., Van Weerd, M., Engberg, S.E., Brennemann, R.A. and Louis Jr, E., 2012. Detection of *Crocodylus mindorensis* x *Crocodylus porosus* (Crocodylidae) hybrids in a Philippine crocodile systematics analysis. *Zootaxa*, 3560(1), pp.1-31.

Thorbjarnarson, J., Platt, S.G. and Khaing, U.S.T., 2000. A population survey of the estuarine crocodile in the Ayeyarwady Delta, Myanmar. *Oryx*, 34(4), pp.317-324.

Tinker, N.A. and Mather, D.E., 1993. KIN: Software for computing kinship coefficients. *Journal of Heredity*, 84(3), pp.238-238.

Toro, M.Á., García-Cortés, L.A., and Legarra, A., 2011. A note on the rationale for estimating genealogical coancestry from molecular markers. *Genetics Selection Evolution*, 43, pp.1-10.

Ujvari, B., Dowton, M. and Madsen, T. 2008. Population genetic structure, gene flow and sex-biased dispersal in frillneck lizards (*Chlamydosaurus kingii*). *Molecular Ecology*, 17(15), pp.3557-3564.

Van Asch, B., *et al*, C., 2019. Phylogeography, genetic diversity, and population structure of Nile crocodile populations at the fringes of the southern African distribution. *Plos One*, 14(12), e0226505. [online] Available at: [online]

Available at: <https://doi.org/10.1371/journal.pone.0226505> [Accessed 11 October 2022]

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. and Shipley, P.F., 2007. Microchecker version 2.2. 3 University of Hull.

Vasconcelos, W.R., Hrbek, T., Da Silveira, R., De Thoisy, B., Ruffeil, L.A.A. D. S., and Farias, I. P., 2008. Phylogeographic and conservation genetic analysis of the black caiman (*Melanosuchus niger*). *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, 309(10), pp.600-613.

Velo-Anton, G., Godinho, R., Campos, J.C., and Brito, J.C., 2014. Should I stay or should I go? Dispersal and population structure in small, isolated desert populations of West African crocodiles. *PLoS One*, 9(4), e94626. [online] Available at: < DOI: 10.1371/journal.pone.0094626 > [Accessed 12 December, 2021]

Vieira, M.L.C., Santini, L., Diniz, A. L., and Munhoz, C.D.F., 2016. Microsatellite markers: what they mean and why they are so useful. *Genetics and Molecular Biology*, 39, pp.312-328.

Voris, H.K., 2000. Maps of Pleistocene Sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, 27(5), pp.1153-1167.

Wang, H., *et al*., 2017. Multiple paternity: A compensation mechanism of the Chinese alligator for inbreeding. *Animal Reproduction Science*, 187, pp.124-132.

Wang, J., 2011. COANCESTRY: a program for simulating, estimating, and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources*, 11(1), pp.141-145.

Webb, G.J.W. and Manolis, S.C., 2006. Guidelines on the Harvesting and Management of Wild Crocodilian Populations and the Determination of "Detriment" within the Context of CITES. Report to CITES Secretariat.

Wibowo, A., 2020. In search of Makara, an Extirpated Freshwater Crocodile *Crocodylus siamensis* (Schneider, 1801) of Java Island, Using Temple Reliefs, Ψ AIC and Habitat Modeling. [online] Available at: <https://doi.org/10.20944/preprints202011.0595.v1 > [Accessed 16 June, 2022]

Williams, D. A. and Rabenold, K. N. 2005. Male-biased dispersal, female philopatry, and routes to fitness in a social corvid. *Journal of Animal Ecology*, 74(1), 150-159.

Witzenberger, K.A. and Hochkirch, A., 2011. Ex situ conservation genetics: a review of molecular studies on the genetic consequences of captive breeding programmes for endangered animal species. *Biodiversity and Conservation*, 20, pp.1843-1861.

Wu, Q., Wu, X., and Glenn, T., 2012. Fourteen novel microsatellite loci in the Chinese alligator (*Alligator sinensis*) isolated via 454 pyrosequencing. *Conservation Genetics Resources*, 4, pp.729-732.

Xiong, L., Li, S., Tang, G. and Strobl, J., 2022. Geomorphometry and terrain analysis: Data, methods, platforms, and applications. *Earth-Science Reviews*, 233, p.104191., [online] Available at: <https://doi.org/10.1016/j.earscirev.2022.104191> [Accessed 3 Nov 2023].

Xu, Q., Fang, S., Wang, Z., and Wang, Z., 2005. Microsatellite analysis of genetic diversity in the Chinese alligator (*Alligator sinensis*) Changxing captive population. *Conservation Genetics*, 6, pp.941-951.

Yunis, E. and Arriola, J.G. 2013. Homozygosity. In: S. Maloy, and K. Hughes, eds. 2013. Brenner's Encyclopedia of Genetics. Academic Press. pp. 522-526.

Zane, L., Bargelloni, L., and Patarnello, T., 2002. Strategies for microsatellite isolation: a review. *Molecular Ecology*, 11(1), pp.1-16.

Zhivotovsky, L.A., 2001. Estimating divergence time with the use of microsatellite genetic distances: impacts of population growth and gene flow. *Molecular Biology and Evolution*, 18(5), pp.700-709.

Zhivotovsky, L.A., Goldstein, D.B. and Feldman, M.W., 2001. Genetic sampling error of distance (δμ) 2 and variation in mutation rate among microsatellite loci. *Molecular Biology and Evolution*, 18(12), pp.2141-2145.

Zouros, E., Oberhauser Ball, A., Saavedra, C. and Freeman, K.R., 1994. An unusual type of mitochondrial DNA inheritance in the blue mussel Mytilus. *Proceedings of the National Academy of Sciences*, 91(16), pp.7463-7467.

Zucoloto, R.B., Verdade, L.M., and Coutinho, L.L., 2002. Microsatellite DNA library for Caiman latirostris. *Journal of Experimental Zoology*, 294(4), pp.346-351.

Zucoloto, R.B., Verdade, L.M., Villela, P., Regitano, L. C., and Coutinho, L.L., 2009. Parentage test in broad-snouted caimans (*Caiman latirostris*, Crocodylidae) using microsatellite DNA. *Genetics and Molecular Biology*, 32, pp.874-881.

APPENDIX A

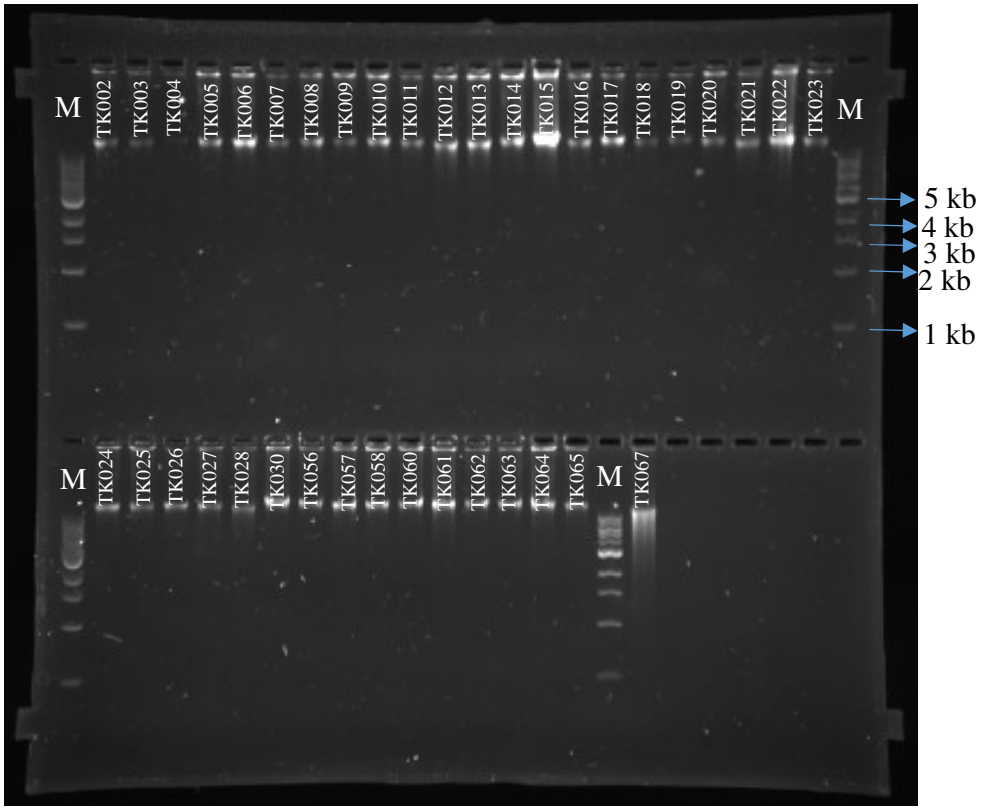


Fig. 1. Extracted DNA bands on a 1% agarose gel. TK002-TK067: T. *schlegelli* sample. M: 1 kb DNA ladder.

APPENDIX B

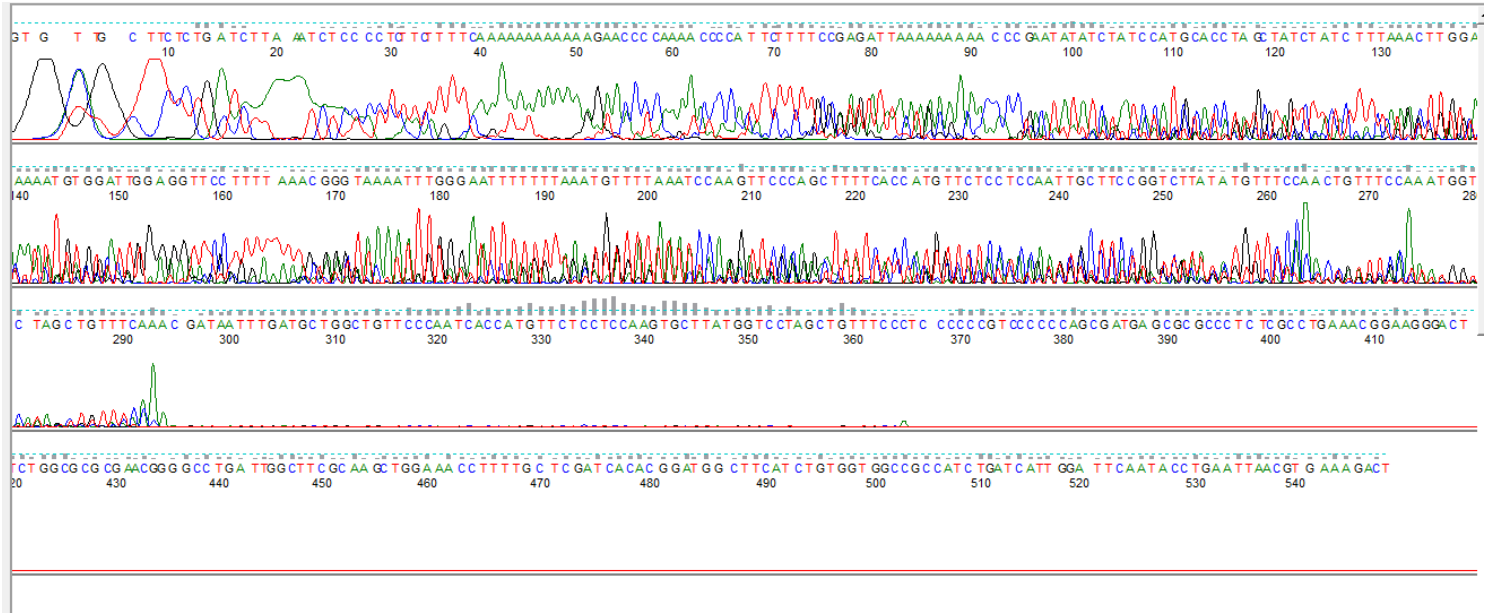


Fig. 1. DNA electropherogram noise sequencing of primer TM07 due due to incomplete separation of double bands on the agarose gel gel

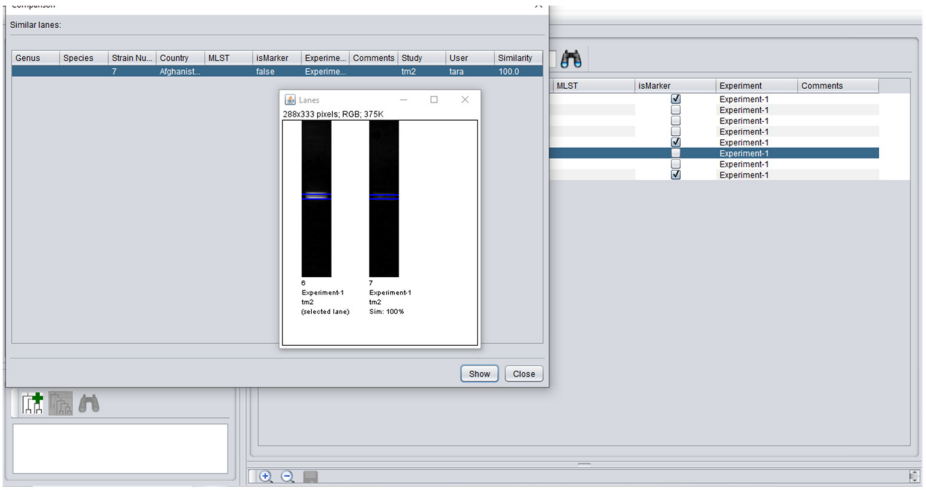
```
!23/11F ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
!23/11R ACCCCTTTTT TTGTTTGTTT GGGTGGATTT TTTTTTGTTT GATGGAGGGA GGGAGGAAAG AAAATAAATA CGAGACAGGA

!23/11F -----TCCAT GGAACTGTGA CAGACAGACA GACAGACAGA CAGACAGACA GACAGACAGA CAGACAGACA GACAGACAGA
!23/11R AGTGAA.AGA .AC.GACA.. .......... .......... .......... .......... .......... ..........

!23/11F CAGACAGACA GACAGACAGA CAGACAGATT TATTTTCTTG ACTCCCATGC TGGTTAATTA AATTTCATTT TCTTATCTAT
!23/11R .......... .......... .......C.. ....C..C.. ..A....... ..C.G..A.C ..A...G... ..AG....T.

!23/11F TTTGTTTGAT TAATAAATAT CTGTATGTAT TAAAAAAAAA GAAACAAGCA GGCTGACTGA GTGAAATATA TACTTCCTTC
!23/11R ....CA.... A...C..A.. ...G..T.TA A.....GT.. ........T. ......G..C A...G..... C...G.AG..

!23/11F CCCCCCCCCC ACCAACCAAC AACCAAACAT ACATCCAAAC AAACAA
!23/11R TATT------ .AT...T.T. ---------- ---------- ------
```

Fig. 1. Low sequence identity between the forward and reverse primers of TM23 due to poor sequencing of 3' of the reverse primer.

APPENDIX D

MICROSATELLITE BAND SCORING

A



B

Fig 1. Screen shot of band scoring and sizing. (A) Example of scoring in GelJ program. (B) Example of sizing of bands in GelQuest program.

APPENDIX E

Table 1: Band sizes of the microsatellite markers used in this study. TK002 – TK067 are blood samples from *T. schlegelii* used in the microsatellite study. Cluster A – F are results of STRUCTURE which assigned the samples to their respective clusters.

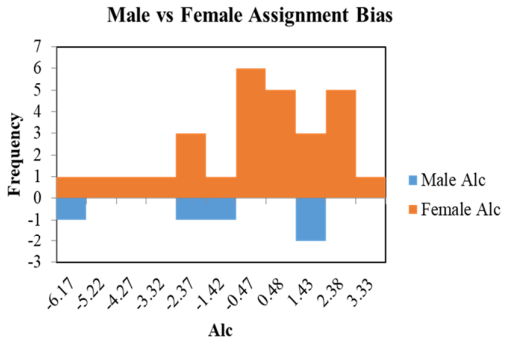
No.	Ind	Cluster	TM2		TM3		TM4		TM5		TM7		TM8		TM10	
1.	TK02	A	164	164	204	204	220	228	272	272	376	376	420	420	508	536
2.	TK03	A	164	172	192	192	228	228	272	284	376	376	412	412	536	536
3.	TK04	A	164	164	192	192	228	228	272	284	376	376	420	420	536	536
4.	TK05	A	164	164	192	192	228	228	272	272	376	376	416	416	536	536
5.	TK21	A	164	164	196	196	228	228	272	272	376	376	412	420	536	536
6.	TK22	A	164	164	196	196	220	228	272	284	376	376	412	420	508	508
7.	TK23	A	164	164	204	204	232	232	272	272	376	376	416	416	536	536
8.	TK30	B	160	160	196	196	212	232	268	268	376	376	416	416	536	536
9.	TK15	B	164	170	196	196	224	240	268	268	376	376	420	420	508	536
10.	TK09	B	164	170	192	192	224	224	268	268	380	408	420	420	508	536
11.	TK67	B	162	170	192	192	232	232	272	464	376	376	416	416	520	520
12.	TK66	B	164	168	192	204	208	208	272	272	408	408	420	420	508	508
13.	TK14	C	174	174	196	196	204	204	280	280	404	408	420	420	520	520
14.	TK58	C	174	174	192	192	204	204	256	256	404	408	420	420	520	520
15.	TK24	D	164	168	192	204	240	240	280	280	380	408	420	420	512	520
16.	TK25	D	164	172	192	204	240	240	280	280	380	408	420	420	520	536
17.	TK26	D	164	168	192	204	240	240	280	280	380	408	420	420	520	536
18.	TK27	D	164	168	192	204	240	240	280	280	380	408	420	420	520	536
19.	TK28	D	164	168	192	204	208	244	276	276	376	388	420	420	512	520
20.	TK07	D	168	172	192	192	208	244	280	280	388	408	420	420	512	536
21.	TK08	D	164	176	204	204	240	240	276	276	376	376	420	420	520	520
22.	TK06	E	164	168	212	212	244	244	272	272	380	408	420	420	508	508
23.	TK11	E	164	168	204	204	208	244	272	272	380	408	420	420	520	536
24.	TK13	E	160	160	204	204	244	244	272	272	380	408	420	420	520	536
25.	TK56	E	160	160	204	204	240	240	268	268	388	408	420	420	520	536
26.	TK61	E	164	164	212	212	240	240	268	268	408	408	420	420	508	508
27.	TK63	E	164	170	204	204	208	244	268	268	380	408	420	420	520	520

28.	TK64	E	164	164	212	212	208	244	272	272	380	408	420	420	520	520
29.	TK65	E	164	164	212	212	244	244	272	272	376	376	420	420	520	536
30.	TK16	F	176	176	204	204	228	228	280	280	380	408	420	420	512	512
31.	TK17	F	164	176	204	204	228	228	280	280	376	388	420	420	520	536
32.	TK18	F	164	176	204	204	228	228	280	280	376	376	420	420	508	508
33.	TK19	F	174	174	204	204	228	228	272	284	376	388	420	420	520	520
34.	TK20	F	176	176	204	204	228	228	280	280	376	376	420	420	520	536
35.	TK12	F	164	176	204	204	224	240	280	280	380	408	420	420	520	520
36.	TK60	F	164	176	212	212	228	228	280	280	376	388	420	420	520	536
37.	TK62	F	174	174	212	212	228	228	272	284	376	388	420	420	520	520
38.	TK10	F	174	174	204	204	228	228	280	280	380	408	420	420	520	536

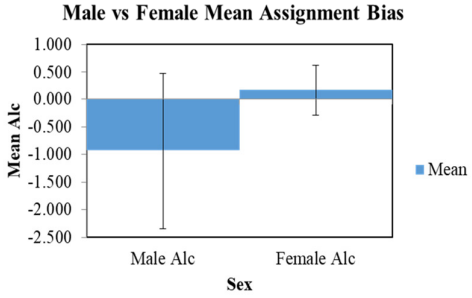
Table 1-continued

No.	Ind	Cluster	TM11		TM12		TM13		TM15		TM21		TM23		TM24	
1.	TK02	A	468	468	520	520	564	564	352	364	228	228	320	356	340	340
2.	TK03	A	468	468	516	516	564	564	360	360	228	232	336	340	316	344
3.	TK04	A	468	468	516	516	564	564	360	360	228	228	336	340	316	362
4.	TK05	A	468	468	516	516	560	560	352	364	228	236	404	440	316	376
5.	TK21	A	468	468	504	504	560	560	360	360	228	228	384	404	316	328
6.	TK22	A	468	468	516	516	560	560	360	364	232	236	384	384	348	360
7.	TK23	A	468	468	504	512	564	564	360	360	228	228	384	420	328	328
8.	TK30	B	456	456	512	512	560	560	360	360	228	228	336	336	332	336
9.	TK15	B	456	456	516	516	560	560	352	364	232	236	384	404	316	328
10.	TK09	B	456	456	516	516	564	564	352	364	232	236	356	404	316	336
11.	TK67	B	456	456	520	520	560	560	360	360	228	228	308	336	316	364
12.	TK66	B	452	452	504	504	560	560	360	360	216	216	344	360	340	340
13.	TK14	C	452	452	520	520	552	552	360	360	236	340	356	384	348	356
14.	TK58	C	452	452	520	520	564	564	360	360	232	232	384	404	326	356
15.	TK24	D	452	464	504	504	564	564	360	360	228	228	320	336	328	336
16.	TK25	D	452	464	504	504	564	564	360	360	228	232	300	336	320	328
17.	TK26	D	452	464	504	504	560	560	360	360	224	232	300	320	320	328
18.	TK27	D	452	464	504	504	560	560	360	360	228	232	320	336	316	336
19.	TK28	D	452	456	504	504	560	560	360	360	228	228	320	336	328	328
20.	TK07	D	464	464	504	512	560	560	352	352	228	232	300	328	316	320
21.	TK08	D	452	452	504	504	564	564	360	360	228	228	320	336	328	328
22.	TK06	E	456	456	512	512	560	560	360	364	228	232	308	336	300	300
23.	TK11	E	456	456	512	512	560	560	360	360	228	232	320	320	336	336
24.	TK13	E	456	456	512	512	560	560	360	360	224	228	312	324	300	336
25.	TK56	E	452	452	516	516	564	564	360	360	232	232	320	320	332	332
26.	TK61	E	452	452	516	516	560	560	360	360	232	232	304	336	300	336
27.	TK63	E	452	452	516	516	564	564	360	360	228	232	324	340	300	336
28.	TK64	E	452	452	504	504	560	560	360	360	232	236	324	344	300	336
29.	TK65	E	452	452	516	516	564	564	360	360	232	236	308	324	300	336
30.	TK16	F	452	452	504	504	560	560	360	360	224	224	336	340	332	340

31.	TK17	F	452	452	504	504	560	560	360	360	232	232	300	348	324	340
32.	TK18	F	452	452	512	512	564	564	360	360	224	232	300	320	332	356
33.	TK19	F	460	460	516	516	564	564	360	360	224	232	300	348	332	340
34.	TK20	F	460	460	512	512	564	564	360	360	224	224	336	384	336	340
35.	TK12	F	452	452	512	512	560	560	360	360	224	232	344	352	324	348
36.	TK60	F	452	452	504	504	564	564	360	360	232	232	300	348	328	336
37.	TK62	F	452	452	512	512	560	560	360	360	224	232	300	348	332	340
38.	TK10	F	452	452	504	504	560	560	360	360	232	232	304	384	328	340



(A)

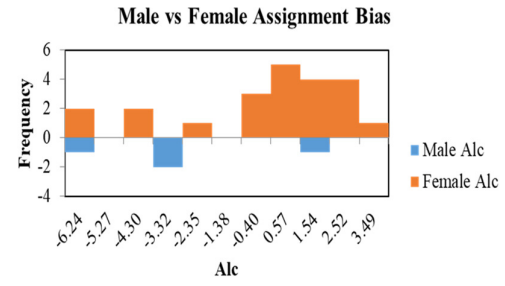


(B)

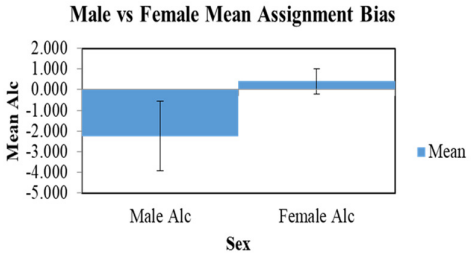
Outcomes of U-test	
R1 for Male Alc	69
R2 for Female Alc	492
N1	5
N2	28
U1	54
U2	86
T1	69
Mean T1	85
SD T1	19.916
Z	-0.803
Two-Tailed Prob. for |Z|	0.422 (ns)

(C)

Fig.1. Sex-biased analysis conducted for the total number of samples. (A) Assignment index showing more females with negative values. (B) Mean corrected assignment index showed more males instead of females with negative values (C) The non-parametric Mann-Whitney's U-test results, showed a non-significant result ($p = 0.422 > 0.05$). R: Sum of ranks of the respective AIc. N1: Sample size for males. N2: Sample size for females. U1: Mann-Whitney test results for males. U2: Mann-Whitney test results for females. T1: the rank selected for the Mann-Whitney U's test. Mean T1: mean of the tested rank. SD T1: standard deviation of the tested ranks. Z: z-score.



(A)



(B)

Outcomes of U-test	
R1 for Male AIc	35
R2 for Female AIc	316
N1	4
N2	22
U1	25
U2	63
T1	35
Mean T1	54
SD T1	14.071
Z	-1.350
Two-Tailed Prob. for |Z|	0.177 (ns)

(C)

Fig.2. Sex-biased analysis carried out considering adults only. (A) Assignment index showing more females with negative values. (B) Mean corrected assignment index showed more males instead of females with negative values (C) The non-parametric Mann-Whitney's U-test results, showed a non-significant result ($p= 0.177 > 0.05$). R: Sum of ranks of the respective AIc. N1: Sample size for males. N2: Sample size for females. U1: Mann-Whitney test results for males. U2: Mann-Whitney test results for females. T1: the rank selected for the Mann-Whitney U's test. Mean T1: mean of the tested rank. SD T1: standard deviation of the tested ranks. Z: z-score

5 kb
4 kb
3 kb
2 kb

1 kb

APPENDIX G

Table 1: The coordinate of capture sites and means of water parameters in the dry and wet season.

Location	Coordinates	Water parameter	Checkpoints (Dry, Wet)	Season	
				Dry	Wet
Pahang (Jengka River)	3.530750, 102.542556	Temp	(3, 3)	31.96667	28.8
		pH	(3, 3)	7.716667	6.963333
		Cond	(3, 3)	31.56667	31.6
		TDS	(3, 3)	20.73333	21.1
		Salinity	(3, 3)	20.83333	20.3
	3.530028, 102.544639	Temp	(2, 3)	32.95	29.3
		pH	(2, 3)	7.89	7.073333
		Cond	(2, 3)	40.25	38.26667
		TDS	(2, 3)	25.55	25.06667
		Salinity	(2, 3)	25.4	23.96667
	*3.529889, 102.547306	Temp	(3, 3)	33.43333	30.9
		pH	(3, 3)	8.096667	7.066667
		Cond	(3, 3)	34.1	31.53333
		TDS	(3, 3)	23.13333	21.13333
		Salinity	(3, 3)	21.26667	20.03333
	3.529528, 102.551250	Temp	(3, 3)	30.16667	29.93333
		pH	(3, 3)	7.37	7.03
		Cond	(3, 3)	35.36667	35.8
		TDS	(3, 3)	23.06667	22.9
		Salinity	(3, 3)	23.03333	23.33333
	3.529194, 102.552028	Temp	(3, 3)	30.83333	29.36667
		pH	(3, 3)	7.553333	7.006667
		Cond	(3, 3)	33.83333	35.5
		TDS	(3, 3)	22.2	23.06667
		Salinity	(3, 3)	21.53333	22.26667
	3.529250, 102.553056	Temp	(2, 3)	30.3	29.73333
		pH	(2, 3)	7.325	7.053333
		Cond	(2, 3)	35.95	35.43333
		TDS	(2, 3)	22.95	23
		Salinity	(2, 3)	21.8	21.9
	3.530083, 102.554417	Temp	(3, 3)	29.66667	29.46667
		pH	(3, 3)	7.323333	7.033333
		Cond	(3, 3)	34.4	36
		TDS	(3, 3)	22.76667	22.96667
		Salinity	(3, 3)	21.13333	23.2

	Coordinates	Water parameter	Checkpoints (Dry, Wet)	Season	
				Dry	Wet
	3.531750, 102.556917	Temp	(3, 3)	29.16667	29.8
		pH	(3, 3)	7.266667	7.006667
		Cond	(3, 3)	32.6	35.1
		TDS	(3, 3)	21.26667	22.86667
		Salinity	(3, 3)	20.4	23.33333
	3.531583, 102.557056	Temp	(2, 3)	29.15	29.6
		pH	(2, 3)	7.32	7.04
		Cond	(2, 3)	32.85	36.73333
		TDS	(2, 3)	21.45	23.66667
		Salinity	(2, 3)	20.3	23.8
	3.529250, 102.556806	Temp	(3, 3)	28.9	29.8
		pH	(3, 3)	7.336667	7.003333
		Cond	(3, 3)	33.53333	35.2
		TDS	(3, 3)	22.46667	22.83333
		Salinity	(3, 3)	20.66667	22.63333
Perak (Sungkai River)	3.948722, 101.187278	Temp	(3, 3)	27.33333	30.03333
		pH	(3, 3)	6.506667	7.37
		Cond	(3, 3)	36.13333	34.13333
		TDS	(3, 3)	25.83333	24.26667
		Salinity	(3, 3)	24.53333	24.6
	3.948250, 101.187583	Temp	(1, 3)	29.4	30.33333
		pH	(1, 3)	6.84	7.006667
		Cond	(1, 3)	36	33.2
		TDS	(1, 3)	25.6	22.66667
		Salinity	(1, 3)	25.2	23.2
	3.950083, 101.224972	Temp	(3, 3)	29	30.83333
		pH	(3, 3)	6.78	7.066667
		Cond	(3, 3)	35.3	32.43333
		TDS	(3, 3)	24.83333	23.1
		Salinity	(3, 3)	24.93333	24.26667
	* 3.946444, 101.268250	Temp	(3, 3)	28.33333	30.7
		pH	(3, 3)	6.673333	7.053333
		Cond	(3, 3)	34.16667	31.33333
		TDS	(3, 3)	24.23333	22.23333
		Salinity	(3, 3)	24.2	23.6
	3.952972, 101.284972	Temp	(3, 3)	27.9	28.33333
		pH	(3, 3)	6.65	6.673333
		Cond	(3, 3)	33.06667	34.16667
		TDS	(3, 3)	23.56667	24.23333
		Salinity	(3, 3)	23.66667	24.46667

	Coordinates	Water parameter	Checkpoints (Dry, Wet)	Season	
				Dry	Wet
Selangor (Ban Canal)	3.693667, 101.341194	Temp	(3, 3)	29.16667	28.66667
		pH	(3, 3)	7.283333	6.896667
		Cond	(3, 3)	34.2	36.8
		TDS	(3, 3)	24.3	26.13333
		Salinity	(3, 3)	24.4	25.4
	3.692361, 101.342222	Temp	(3, 3)	29.2	28.73333
		pH	(3, 3)	7.02	6.886667
		Cond	(3, 3)	34.23333	37.1
		TDS	(3, 3)	24.3	26.3
		Salinity	(3, 3)	24.43333	25.5
	* 3.669083, 101.345556	Temp	(3, 3)	29.23333	29.1
		pH	(3, 3)	6.8	6.65
		Cond	(3, 3)	34.2	36.5
		TDS	(3, 3)	24.2	25.96667
		Salinity	(3, 3)	24.4	25.33333
	3.574417, 101.351861	Temp	(3, 3)	28.6	28.5
		pH	(3, 3)	5.996667	6.373333
		Cond	(3, 3)	57.06667	92.6
		TDS	(3, 3)	40.53333	65.56667
		Salinity	(3, 3)	33.76667	48.93333
(Tengi River)	3.573250, 101.351778	Temp	(3, 3)	29.43333	28.66667
		pH	(3, 3)	6.333333	6.536667
		Cond	(3, 3)	37.36667	56.96667
		TDS	(3, 3)	26.46667	40.46667
		Salinity	(3, 3)	25.93333	33.9
	3.543639, 101.353861	Temp	(3, 3)	27.43333	27.1
		pH	(3, 3)	4.03	3.786667
		Cond	(3, 3)	48.8	49.36667
		TDS	(3, 3)	34.5	35.03333
		Salinity	(3, 3)	30	30.16667
	3.543694, 101.354000	Temp	(3, 3)	27.36667	27.2
		pH	(3, 3)	3.993333	3.756667
		Cond	(3, 3)	48.76667	49.26667
		TDS	(3, 3)	34.9	34.93333
		Salinity	(3, 3)	30.1	30.13333
	3.543472, 101.353778	Temp	(3, 3)	29.93333	29
		pH	(3, 3)	6.126667	6.076667
		Cond	(3, 3)	34.86667	39.43333
		TDS	(3, 3)	24.76667	27.93333
		Salinity	(3, 3)	24.86667	26.46667

	Coordinates	Water parameter	Checkpoints (Dry, Wet)	Season	
				Dry	Wet
	3.537000, 101.324389	Temp	(3, 3)	30.13333	29.3
		pH	(3, 3)	6.133333	6.22
		Cond	(3, 3)	34.26667	39.13333
		TDS	(3, 3)	24.33333	27.76667
		Salinity	(3, 3)	24.7	26.43333
	3.540222, 101.304944	Temp	(3, 3)	30.23333	29.36667
		pH	(3, 3)	6.186667	6.163333
		Cond	(3, 3)	33.86667	38.53333
		TDS	(3, 3)	24.06667	27.36667
		Salinity	(3, 3)	24.53333	26.23333
	3.521028, 101.285278	Temp	(3, 3)	30.56667	29.46667
		pH	(3, 3)	6.103333	6.163333
		Cond	(3, 3)	33.16667	37.66667
		TDS	(3, 3)	23.6	26.73333
		Salinity	(3, 3)	24.36667	25.9
	3.506056, 101.268361	Temp	(3, 3)	30.9	29.46667
		pH	(3, 3)	6.113333	6.106667
		Cond	(3, 3)	33.23333	37.26667
		TDS	(3, 3)	23.56667	26.46667
		Salinity	(3, 3)	24.43333	25.73333
	3.492472, 101.255528	Temp	(3, 3)	30.33333	29.26667
		pH	(3, 3)	6.15	6.09
		Cond	(3, 3)	33.1	36.86667
		TDS	(3, 3)	23.5	26.3
		Salinity	(3, 3)	24.23333	25.63333
	3.481778, 101.221417	Temp	(3, 3)	30.63333	29.46667
		pH	(3, 3)	6.06	6.063333
		Cond	(3, 3)	33.3	38.63333
		TDS	(3, 3)	23.66667	27.43333
		Salinity	(3, 3)	24.36667	26.26667
Terengganu (Nerus River)	**5.322472, 103.042778	Temp	(3, 3)	30.8	27.96667
		pH	(3, 3)	6.34	5.7
		Cond	(3, 3)	25.43333	31.06667
		TDS	(3, 3)	18.03333	22.26667
		Salinity	(3, 3)	21.16667	22.86667
	5.317056, 103.027444	Temp	(3, 3)	30.16667	28.3
		pH	(3, 3)	6.146667	5.976667
		Cond	(3, 3)	26.96667	27.76667
		TDS	(3, 3)	19	19.7
		Salinity	(3, 3)	25.76667	20.46667

		Water parameter	Checkpoints (Dry, Wet)	Season	
				Dry	Wet
	5.316583, 103.026972	Temp	(3, 3)	31.03333	28.33333
		pH	(3, 3)	6.326667	6.01
		Cond	(3, 3)	26	24.26667
		TDS	(3, 3)	18.46667	17.16667
		Salinity	(3, 3)	21.76667	20.3
	5.306056, 103.001583	Temp	(3, 3)	30.33333	27.63333
		pH	(3, 3)	6.266667	6.1
		Cond	(3, 3)	24.7	23.46667
		TDS	(3, 3)	17.53333	16.6
		Salinity	(3, 3)	20.93333	19.73333
	5.299722, 102.998361	Temp	(3, 3)	28.93333 333	27.43333
		pH	(3, 3)	6.523333 333	5.95
		Cond	(3, 3)	24.8	24.3
		TDS	(3, 3)	17.63333 333	17.23333
		Salinity	(3, 3)	20.6	20.06667
	5.282361, 102.986889	Temp	(3, 3)	30.06667	26.7
		pH	(3, 3)	6.416667	5.82
		Cond	(3, 3)	25.76667	22.33333
		TDS	(3, 3)	18.3	15.83333
		Salinity	(3, 3)	21.3	19.1
	5.274111, 102.990639	Temp	(3, 3)	30	26.6
		pH	(3, 3)	6.456667	5.786667
		Cond	(3, 3)	25.8	22.06667
		TDS	(3, 3)	18.3	15.66667
		Salinity	(3, 3)	21.3	19
	5.269139, 102.989000	Temp	(3, 3)	29.8	26.6
		pH	(3, 3)	5.77	5.703333
		Cond	(3, 3)	25.83333	22.5
		TDS	(3, 3)	18.33333	15.93333
		Salinity	(3, 3)	21.3	19.1
(Tekah River)	5.299667, 102.996111	Temp	(3, 3)	29.1	27.46667
		pH	(3, 3)	6.256667	5.91
		Cond	(3, 3)	24.73333	24.4
		TDS	(3, 3)	17.6	17.3
		Salinity	(3, 3)	20.73333	20.13333

	Coordinates	Water parameter	Checkpoints (Dry, Wet)	Season	
				Dry	Wet
	5.298139, 102.993222	Temp	(3, 3)	29.63333	27.73333
		pH	(3, 3)	6.07	5.19
		Cond	(3, 3)	28.36667	23.96667
		TDS	(3, 3)	20.13333	17.1
		Salinity	(3, 3)	22.3	20.03333
	* 5.293139, 102.995194	Temp	(3, 3)	29.23333	27.5
		pH	(3, 3)	6.133333	5.05
		Cond	(3, 3)	28.63333	23.96667
		TDS	(3, 3)	20.56667	17.03333
		Salinity	(3, 3)	22.26667	19.93333
(Belara River)	5.286000, 103.019389	Temp	(3, 3)	33.6	26.3
		pH	(3, 3)	6.83	5.513333
		Cond	(3, 3)	30.03333	23.33333
		TDS	(3, 3)	21.36667	17
		Salinity	(3, 3)	23.6	19.53333
	5.289111, 103.021056	Temp	(3, 3)	39.06667	25.7
		pH	(3, 3)	7.096667	5.216667
		Cond	(3, 3)	30.73333	22.73333
		TDS	(3, 3)	21.96667	16.13333
		Salinity	(3, 3)	25.1	19
	5.290750, 103.020250	Temp	(3, 3)	33.96667	26.6
		pH	(3, 3)	6.38	5.316667
		Cond	(3, 3)	30.4	23.26667
		TDS	(3, 3)	21.6	16.56667
		Salinity	(3, 3)	23.83333	19.46667
	5.291111, 103.020028	Temp	(3, 3)	34.66667	26.4
		pH	(3, 3)	6.763333	5.263333
		Cond	(3, 3)	29.86667	22.73333
		TDS	(3, 3)	21.23333	16.1
		Salinity	(3, 3)	23.96667	19.2
	5.294278, 103.018833	Temp	(3, 3)	35.3	27.03333
		pH	(3, 3)	6.433333	5.263333
		Cond	(3, 3)	29.56667	22.83333
		TDS	(3, 3)	21.66667	16.03333
		Salinity	(3, 3)	23.83333	19.36667
	5.295611, 103.021722	Temp	(3, 3)	34.96667	28.36667
		pH	(3, 3)	6.37	5.363333
		Cond	(3, 3)	29.7	24.86667
		TDS	(3, 3)	21.06667	17.8
		Salinity	(3, 3)	23.93333	20.6

Coordinates	Water parameter	Checkpoints (Dry, Wet)	Season	
			Dry	Wet
5.297500, 103.022889	Temp	(3, 3)	33.43333	26.66667
	pH	(3, 3)	6.343333	5.236667
	Cond	(3, 3)	29.83333	22.56667
	TDS	(3, 3)	21.2	16.1
	Salinity	(3, 3)	23.66667	19.26667
5.299167, 103.023722	Temp	(3, 3)	33.8	26.66667
	pH	(3, 3)	6.243333	5.206667
	Cond	(3, 3)	30.06667	24
	TDS	(3, 3)	21.4	17.1
	Salinity	(3, 3)	23.8	19.83333

*Locations of recent *Tomistoma* capture

**Location of C. porosus sightings