Phylogeography of the Komodo monitor *Varanus komodoensis* (Reptilia: Varanidae) inferred from mitochondrial DNA Control Region I and the implications for *in situ* management plans

Doctoral thesis

Dissertation

zur Erlangung der Doktorwürde (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Evy Ayu Arida

aus

Jogjakarta in Indonesien

Bonn, im Januar 2011

Gedruckt mit Unterstützung des Deutschen Akademischen Austauschdienstes

Angefertig mit der Genehmigung von
Mathematisch-Wissenschaftlichen Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

Die Arbeit wurde am
Zoologisches Forschunginstitut und Museum Alexander Koenig
(ZFMK), Bonn durchgeführt

Referent: Prof. Dr. Wolfgang Böhme

Koreferent: Prof. Dr. Steven F. Perry

Tag der Mündlichen Prüfung: 10. 03. 2011

Erscheinungsjahr: 2011

für meine Familie:
Mum, Dad, und Daz

Summary

The Komodo monitor (*Varanus komodoensis*) is the largest living lizard in the world. However, it was discovered by scientists only about one hundred years ago, and is now vulnerable to extinction in the wild. The total population size is currently estimated to be about 2,300 individuals, while habitat degradation associated with human activities seems to have accelerated the decrease in population size over time. Worldwide, the wild populations of the Komodo monitor occur only in the Lesser Sunda Islands, in southern Wallacea. Located between the Asian and Australian biogeographic realms, the degree of endemism in Wallacea is relatively high. Given the relatively small population size, limited distribution, and habitat degradation, a sound conservation programme needs to be designed to protect the existing wild populations of the Komodo monitor. To help design the programmes for conservation, I investigated the population genetic structure across the distribution range by applying a molecular phylogeographic approach. Eleven (11) mitochondrial DNA (mtDNA) lineages were determined in this study, and the phylogenetic relationships among these lineages formed two major clades. This genealogy was rooted with the Lace monitor (*Varanus varius*), the Australian sister species of the Komodo monitor. Further, I assessed the extent of the population genetic structure across the entire population using AMOVA, which is a method in population genetics to evaluate the diversity of haplotypes in hierarchical subdivisions. To illustrate the relationships among populations distributed across the current range of the Komodo monitor, I used Statistical Parsimony method. Finally, I discussed the resulting population genetic structure with regard to palaeogeographical and palaeontological data, i.e. Pleistocene sea level fluctuations and the fossil Komodo monitors recovered in Australasia to infer the historical processes involved in shaping the current state of maternal population structure. The results from my study can serve as basic data that complement the existing information on ecology, behaviour, and population genetics, which are among the important components to devise management programmes for conservation.

The Komodo National Park in the province of Nusa Tenggara Timur, Indonesia, is established to protect the extant populations of the Komodo monitor. Within the National Park, populations occur on four islands, i.e. Komodo, Rinca, Gili Motang, and Nusa Kode. Beyond the National Park, two Nature Reserves are established on the western and northern coasts of Flores. In particular, the population on Flores is facing

extinction threats from habitat degradation and loss aggravated by the poaching of the adult's main prey, i.e. the Timor deer (*Cervus timorensis*), resulting in a significantly reduced distribution range. A population used to occur on Padar, a small island that lies between Komodo and Rinca. Padar was completely burnt by a wild fire about thirty (30) years ago. This population is now considered extinct. Two other small populations occur on the small islands of Gili Motang and Nusa Kode. Because small island populations often exhibit a lower genetic diversity, which is associated with reduced population viability, these areas may call for prioritisation in the conservation management plan. Therefore, an identification of areas with lower genetic diversity is one of the main foci in this study.

To help identify priority areas for conservation, I characterised the level of genetic diversity across the extant populations using three hundred and sixty six (366) mtDNA Control Region I (CRI) sequences of the Komodo monitor. Based on nucleotide substitution, sequence length, and the position of single indels, I determined eleven (11) haplotypes. The geographic distribution of these haplotypes representing maternal lineages is non-random. More specifically, three regions with unique maternal lineage compositions were identified. These regions will be referred to as the Western, Central, and Eastern regions. The Western region includes Komodo, whereas the Central region is composed of populations distributed on the islands of Rinca, Gili Motang, and Nusa Kode, as well as a population on the western coast of Flores. The Eastern region consists a single population on the northern coast of Flores. An assessment of the genetic diversity across island populations revealed a higher level of genetic diversity in larger island populations and low level of genetic diversity in small island populations. The number of haplotypes found on the larger islands is greater than that on the small islands. Komodo, Rinca, and Flores harbour four (4), six (6), and four (4) haplotypes, respectively. By contrast, only one haplotype was found on both Gili Motang and Nusa Kode. Haplotype diversity (h) and nucleotide diversity (π_n) are also greater in larger island populations than those on small island populations.

For a more detailed assessment on the populations of the Komodo monitor aimed at helping the conservation management, I investigated the phylogenetic relationships among the eleven haplotypes as well as the geographic distribution of these maternal lineages. Three phylogenetic methods were applied to infer the relationships among haplotypes and the results revealed two clades, an Eastern clade and a Central-Western

clade. The Central-Western haplotype group clusters nine (9) lineages representing all individuals distributed in the Central and Western regions, the otherwise Eastern. On the other hand, the Eastern haplotype group consists two (2) lineages distributed mainly on the northern coast of Flores. It is interesting to note, that one individual sampled from a population on the western coast of Flores (the Central region) expressed haplotype H10, which was otherwise found only in the Eastern region. Haplotype divergence across the whole population seems to be relatively small, with a range of 0-2%. Pairwise *p* distances among all haplotypes indicate that the Eastern haplotypes are the most divergent. The non-random geographic distribution of these haplotypes was tested using AMOVA. I explored two alternative hypotheses of population genetic structure. Grouping haplotypes according to each of the five islands they were from, explained only about 54% of the total molecular variance. Alternatively, grouping the data according to the three regions, i.e. Western, Central, and Eastern accounted for about 90% of the total molecular variance. I generated a haplotype network using the Statistical Parsimony method to illustrate the relationships among populations distributed across regions. In the network, the populations in the Western and Central regions are largely characterised by the distribution of the Western and Central haplotypes. These two populations are connected by a hypothetical intermediate haplotype, which is not found in this study. On the other hand, the population in the Eastern region is characterised exclusively by the Eastern haplotypes and is separated entirely from both the Western and Central regions. In brief, these results suggest a strong genetic structure of three subdivisions, each of which is characterised by its distinctive haplotype composition. The highly divergent Eastern region is characterised exclusively by the Eastern haplotypes. However, one of these two haplotypes, i.e. H10 is shared between the Eastern and Central regions, possibly due to an event of migration between these regions or as a consequence of habitat fragmentation. Similarly, the Western region is also a unique subdivision. This Western region comprises the whole population on Komodo, in which a total of four (4) haplotypes were found. One of these typical Western haplotypes, i.e. H3 is shared with a population on Rinca (the Central region). The coexistence of this lineage in the Western and Central regions may indicate an event of dispersal between Komodo and Rinca. It is interesting to note, that the Central region shares both the Western and Eastern lineages. Apart from these shared haplotypes, the Central region harbours the highest number of unique haplotypes, i.e.

five (5) haplotypes among all three regions. Therefore, the Central region is a distinct subdivision and also an important source of haplotype diversity.

Concordant with the genetic structure of three regions revealed in this study, three separate units of conservation may be assigned to maintain the level of genetic diversity and the associated evolutionary history shaping this structure. Despite the likelihood of dispersal among the three regions, the level of genetic differentiation among these regions seems to be maintained through a process of isolation related with geographic distance and seawater barrier. In particular, the presence of seawater may become a more effective barrier to dispersal when combined with an increased distance between islands. This result is corroborated by the estimates of nuclear gene flow among islands in a previous population genetic study. In addition, the Pleistocene sea level fluctuations were probably involved in shaping the genetic structure of the current population through multiple exposures of land bridges, which facilitated migrations. Therefore, in order to maintain mtDNA diversity and the associated evolutionary history for each of the three regions, Komodo should be managed separately from the islands in the Central region and Flores North. In contrast, a plan to augment the isolated small population on Gili Motang should consider transferring individuals from a source with a genetic similarity within the Central region. Based on the mtDNA CRI haplotype distribution in this study, Gili Motang harbours only one maternal lineage, i.e. haplotype H5 that is shared with the neighbouring locations in the Central region, for example Flores West. Thus, an augmentation programme for Gili Motang should transfer one or more individuals with haplotype H5 from Florest West, rather than transferring those with a divergent haplotype such as haplotype H10, or individuals from Flores North. Nevertheless, phylogeography is only one element to consider for planning conservation management. Further studies are still needed to design a robust conservation programme, for instance an investigation on the mating system aimed at predicting the number of yearly recruits. Besides, several caveats in the current study are yet to be addressed. The drawbacks with the small sample size for the Eastern region, the deficiency of sample from the northwestern peninsula of Rinca, and the phylogenetic analyses that disregard the variation in sequence length should be resolved in the future studies.

Zusammenfassung

Der Komodowaran (*Varanus komodoensis*) gehört der Familie Varanidae an und ist die größte heute lebende Echse mit einer Gesamtlänge von bis zu 3 Meters. Trotzdem wurde er erst vor etwa 100 Jahren entdeckt und ist bereits heute vom Aussterben bedroht. Die Gesamtpopulationsgröße wird derzeit auf ungefähr 2300 Individuen geschätzt; Habitatzerstörung einhergehend mit anthropogenen Aktivitäten scheinen den Rückgang der Populationsgröße mit der Zeit beschleunigt zu haben. Weltweit kommen wilde Komodowaranpopulationen nur auf den Kleinen Sunda Inseln in der südlichen Wallacea vor. Zwischen den asiatischen und australischen biogeographischen Regionen gelegen, ist der Grad an Endemismus in Wallacea relativ hoch. Angesichts der relativ geringen Populationsgröße, des begrenzten Verbreitungsgebiets und der Habitatzerstörung muß ein solides Erhaltungsprogramm entwickelt werden, um die wildlebenden Komodowarane zu schützen. Mit der Absicht die Ausarbeitung von Naturschutzprogrammen zu unterstützen, habe ich die populationsgenetische Struktur von *V. komodoensis* über das gesamte Verbreitungsgebiet mittels eines molekularen, phylogeographischen Ansatzes untersucht. Elf verschiedene mitochondriale Abstammungslinien (mtDNA) wurden in dieser Studie ermittelt und die phylogenetischen Beziehungen zwischen diesen Linien sprechen für zwei distinkte Hauptgruppen. Diese Genealogie wurde mit dem Buntwaran (*Varanus varius*), der australischen Schwesterart des Komodowarans, gewurzelt. Ferner wurde die populationsgenetische Struktur über die gesamte Population mittels einer AMOVA, einer populationsgenetischen Methode, welche die Diversität an Haplotypen in hierarchischen Untergruppen aufteilt, evaluiert. Um die verwandtschaftlichen Beziehungen zwischen den Populationen des heutigen Verbreitungsgebietes zu illustrieren, wurde die Statistische Parsimonie-Methode benutzte. Schließlich wurde die daraus resultierende populationsgenetische Struktur von *V. komodoensis* im Hinblick auf paläogeographische und paläontologische Daten, d. h. pleistozäne Meeresspiegelfluktuationen und fossile Komodowarane in Australasien, diskutiert, um mutmaßliche, historische Prozesse auszumachen, die bei der Ausprägung der gegenwärtigen, maternalen Populationsstruktur involviert gewesen sein könnten. Die Ergebnisse meiner Studie können als grundlegende Daten dienen, um vorhandene Informationen über Ökologie, Verhalten und Populationsgenetik des Komodowarans zu

ergänzen, die wichtige Komponenten bei der Entwicklung von Managementprogrammen für den Naturschutz darstellen.

Der Komodo Nationalpark in der Provinz Nusa Tenggara Timur, Indonesien, wurde errichtet, um die dort lebenden Komodowarane zu schützen. Innerhalb des Parks finden sich Populationen auf den vier Inseln Komodo, Rinca, Gili Motang und Nusa Kode. Außerhalb des Parks wurden an der West- und Nordküste von Flores zwei Naturreservate etabliert. Vor allem die Population auf Flores ist durch Habitatzerstörung und -verlust bedroht. Diese Bedrohungssituation wird durch das Wildern der Hauptnahrungsquelle der Adulten, des Timorhirsches (*Cervus timorensis*), verschärft und resultiert in ein reduziertes Verbreitungsgebiet auf dieser Insel. Eine weitere Population existierte auf Padar, einem kleinen Eiland, das zwischen Komodo und Rinca liegt. Padar wurde von einem Feuer vor ungefähr 30 Jahren vollständig verwüstet. Seither gilt diese Inselpopulation als ausgestorben. Zwei andere, kleine Populationen kommen auf den Inseln Gili Motang und Nusa Kode vor. Da kleine Inselpopulationen oft eine geringe genetische Diversität aufweisen, was mit einer reduzierten Überlebensfähigkeit assoziiert ist, erfordern diese Gebiete Prioritätensetzung im Naturschutzmanagementplan. Deshalb ist einer der Hauptaugenmerke dieser Studie, Gebiete mit geringer genetischer Diversität zu identifizieren.

Um Naturschutzgebiete mit Priorität zu identifizieren, wurden insgesamt 366 Sequenzen der mtDNA Kontrollregion I (CRI) des Komodowarans analysiert, um das Ausmaß der genetischen Diversität innerhalb der verbliebenen Populationen zu charakterisieren. Basierend auf Nukleotidsubstitutionen, Sequenzlänge und der Position einzelner Indels wurden elf unterschiedliche Haplotypen identifiziert. Die geographische Verteilung dieser mtDNA Haplotypen, die maternale Entwicklungslinien repräsentieren, ist nicht zufällig. Genauer gesagt konnten drei Regionen mit einzigartigen Zusammensetzungen maternaler Linien identifiziert werden. Diese Regionen werden im Folgenden als West, Zentral und Ost bezeichnet. Die Region West umfaßt lediglich Komodo Island, während die Zentral-Region aus den Populationen der Inseln Rinca, Gili Motang, Nusa Kode sowie der Population an der Westküste von Flores besteht. Die Region Ost hingegen besteht nur aus einer einzigen Population an der Nordküste von Flores. Es wurde eine Abschätzung der genetischen Diversität über alle Inselpopulationen durchgeführt und herausgefunden, dass größere Populationen generell eine höhere genetische Diversität aufweisen als kleine Inselpopulationen. Die Anzahl

verschiedener Haplotypen ist auf den größeren Inseln höher als die auf kleinen Inseln. Komodo, Rinca und Flores beherbergen jeweils vier, sechs und vier verschiedene Haplotypen. Im Gegensatz dazu wurde nur ein Haplotyp auf Gili Motang und Nusa Kode gefunden. Haplotypendiversität (h) und Nukleotiddiversität (π_n) sind ebenfalls höher in größeren Inselpopulationen als auf kleinen Inseln.

Für eine detailliertere Beurteilung der Komodowaranpopulationen zu Naturschutzmanagementvorhaben wurden die phylogenetischen Verwandtschaftsbeziehungen zwischen den Haplotypen und der geographischen Verbreitung dieser maternalen Entwicklungslinien untersucht. Phylogenetische Methoden kamen zur Anwendung, um die Verwandtschaftsverhältnisse zwischen den einzelnen Haplotypen zu analysieren. Diese Analyse resultierte in zwei Hauptgruppen, eine östliche und eine zentral-westliche. Die Haplotypengruppierung Zentral-West vereint neun Entwicklungslinien, die alle Individuen der Regionen Zentral und West umfassen. Demgegenüber besteht die Haplotypengruppe Ost aus zwei Linien, die hauptsächlich an der Nordküste von Flores verbreitet sind. Ein Exemplar, das an der Westküste von Flores besammelt wurde, wies Haplotyp H10 auf, der ansonsten nur in der Region Ost gefunden wurde. Die Haplotypendivergenz innerhalb der gesamten Metapopulation scheint relativ gering zu sein und beträgt zwischen 0 und 2%. Paarweise *p*-Distanzen aller Haplotypen zeigen, dass östliche Haplotypen am stärksten divergieren. Die nicht-zufällige Verteilung dieser Haplotypen wurde mit einer AMOVA getestet. Zwei alternative Hypothesen der genetischen Populationsstruktur wurden gegeneinander gestellt. Während die Gruppierung der Haplotypen nach ihrer Inselherkunft lediglich 54% der beobachteten molekularen Varianz erklärten konnte, war die Gruppierung der Daten nach den drei Regionen West, Zentral und Ost für fast 90% der gesamten molekularen Varianz verantwortlich. Ein Haplotypennetzwerk basierend auf Statistischen Parsimoniemethode wurde generiert, um die Beziehungen zwischen den Populationen, die über das Verbreitungsgebiet vorkommen, zu illustrieren. In diesem Netzwerk sind die Populationen der Regionen West und Zentral größtenteils durch die Verteilung der westlichen und zentralen Haplotypen charakterisiert, die durch einen hypothetischen, intermediären Haplotyp miteinander verbunden sind. Demgegenüber ist der Population der Region Ost ausschließlich durch östliche Haplotypen charakterisiert und ist vollständig von den beiden Regionen West und Zentral separiert. Kurz gesagt legen diese Ergebnisse eine genetische Struktur aus drei Untergruppen nahe, von denen

jede durch eine unterschiedliche Haplotypenzusammensetzung charakterisiert ist. Die Region Ost ist hoch divergent, da sie ausschließlich aus östlichen Haplotypen besteht. Jedoch ist einer dieser beiden Haplotypen (H10) den Regionen Ost und Zentral gemein, möglicherweise aufgrund von Migration zwischen diesen beiden Regionen oder als Konsequenz von Habitatfragmentierung. Gleichfalls stellt die Region West eine eigenständige Untereinheit dar. Diese Untereinheit umfaßt die Insel Komodo, für welche insgesamt vier Haplotypen festgestellt wurden. Einer dieser typischen westlichen Haplotypen (H3) kommt auch in der Population auf Rinca (Zentral Region) vor. Die Co-Existenz dieser Linie in den Regionen West und Zentral kann auf ein Dispersionsereignis zwischen Komodo und Rinca hindeuten. In diesem Zusammenhang ist es interessant festzustellen, dass sich die Region Zentral sowohl westliche als auch östliche Linien teilt. Abgesehen von diesen geteilten Haplotypen beherbergt die Zentral-Region die höchste Anzahl endemischer Haplotypen aller drei Regionen. Aus diesem Grund ist diese Region eine wichtige Quelle für Haplotypendiversität und unterscheidet sich somit von den anderen beiden untersuchten Regionen.

Übereinstimmend mit der genetischen Struktur der drei Regionen, wie sie in dieser Studie aufgezeigt werden konnte, sollten auch drei getrennte Naturschutzeinheiten für den Erhalt der genetische Diversität von *V. komodoensis* festgelegt werden. Trotz eines geringen Anteils von Migration zwischen den drei Regionen, scheint die genetische Differenzierung erhalten zu bleiben bedingt durch geographische Isolation und Wasserbarrieren. Besonders Seewasserbarrieren scheinen mit größer werdender Distanz zwischen den Inseln effektiver zu sein bei der Einschränkung von Migration. Dieses Ergebnis stimmt mit den Genfluß-Abschätzungen zwischen den Inselpopulationen des Komodowarans in einer früheren Studie überein. Durch mehrfaches Trockenfallen von Landbrücken, die Migration vereinfachten, waren Fluktuationen der pleistozänen Meeresspiegelstände wahrscheinlich involviert in die Gestaltung der genetischen Struktur der heutigen Populationen. Um die mtDNA-Diversität innerhalb jeder der drei Regionen aufrechtzuerhalten, sollte Komodo Island daher separat von den Inseln der Region Zentral und Nord-Flores verwaltet werden. Im Gegensatz hierzu sollte ein Programm zur Aufstockung der kleinen, isolierten Population auf Gili Motang Individuen mit einem ähnlichen genetischen Ursprung berücksichtigen, wie z. B. West-Flores. Basierend auf den in dieser Studie analysierten mtDNA CRI-Sequenzen beherbergt Gili Motang lediglich eine maternale Linie, die es

mit den benachbarten Inseln in der Region Zentral (wie z. B. West-Flores) gemeinsam hat. Daher sollte ein Auffrischungsprogramm für Gili Motang nicht Individuen mit unterschiedlichen Haplotypen vermischen, wie z. B. von Nord-Flores. Dennoch ist die Phylogeographie von *V. komodoensis* nur ein Aspekt, der bei der Planung von Artenschutzmaßnahmen berücksichtigt werden muß. Weitere Studien, wie etwa Paarungsystem zur Vorhersage der jährlichen Nachkommenschaft, sind notwendig, um ein robustes Naturschutzprogramm zu erstellen. Daneben müssen einige Vorsicht der aktuellen Untersuchung thematisiert werden. Die geringe Stichprobengröße der Region Ost, ein Mangel an Proben von der nordwestlichen Halbinsel Rincas und phylogenetische Methoden, die nicht Sequenzlängenunterschiede als molekulare Merkmale in die Analyse einbeziehen, sollten in zukünftigen populationsgenetischen Studien des Komodowarans beseitigt werden.

Table of contents

List of figures

List of tables

List of matrices

Chapter 1

INTRODUCTION

1.1. Phylogenetic relationships and biogeography of extant monitor lizards

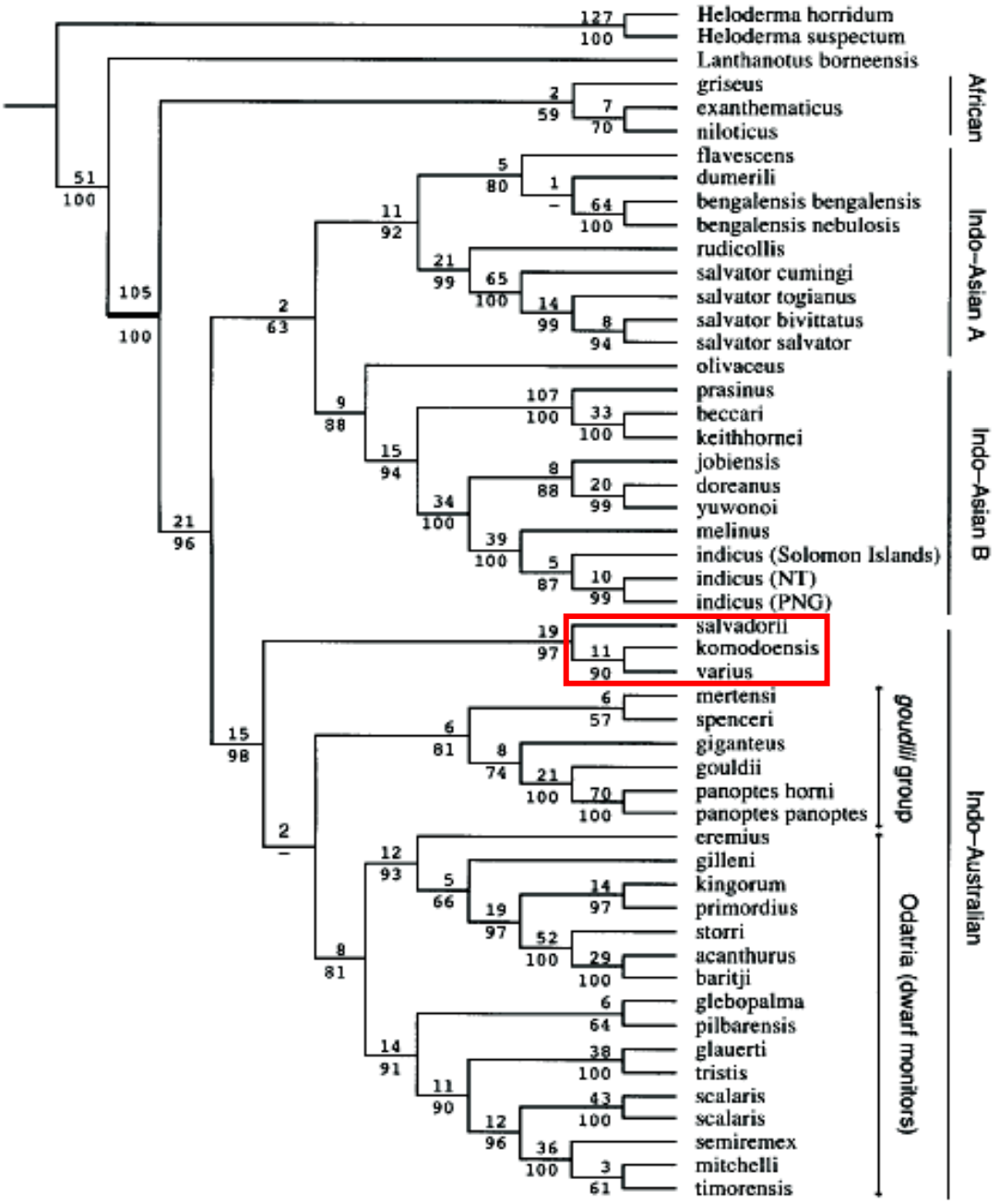
Monitor lizard is the familiar term used to identify a group of small to giant size lizards taxonomically classified under the Superfamily Varanoidea. The word “monitor” may refer to the peculiar behaviour of these lizards to stand on their hind limbs and monitor their surroundings. Monitor lizards are currently distributed in the warmer areas in Africa, Asia, many islands in the Pacific, and Australia. Historically, they were distributed over a much larger area: fossils have been found, for instance in Europe and North America, and other areas that constituted the ancient continents of Laurasia and Gondwana. *Saniwides mongoliensis* is the oldest fossil monitor lizard found in Laurasian fragment. Known from its remains found in Mongolia, *Saniwides* was dated to the Late Cretaceous, ~80 Ma (megaanum; Molnar, 2004b). *Saniwa* is an example of a more recent varanid from outside the current distribution. Fossils have been found in Oligocene - Eocene deposits in North America and Europe and dated to 30-45 Ma (Estes, 1983). These fossils found in Laurasian deposits are relatively old compared to those found in Gondwanan deposits. A fossil monitor lizard found in Egypt has been recently attributed to an early African *Varanus* and was dated to ~30 Ma in the Oligocene (Smith *et al.*, 2008). However, the oldest *Varanus* is reported from the Late Eocene (37 Ma) Egypt and this fossil may indicate the emergence of this genus in Africa, and thus on Gondwanaland (Holmes *et al.*, 2010). Another fossil *Varanus* was recovered in Miocene Songhor, Kenya dated to 19.5-19.9 Ma (Rage and Bailon, 2005). A more certain case is *Varanus rusingensis*, also from Kenya, dates to ~17.8 Ma (Clos, 1995). The earliest evidence of the genus in Australia comes from the Middle Miocene of South Australia (Estes, 1984). A more well known example, the Giant Roamer (*Varanus priscus*, also called *Megalania prisca*), is known from Pliocene central Australia dated to ~4.5 Ma (Molnar, 2004a). The genus *Varanus* seems to appear in Gondwana earlier than in Laurasia. A fossil *Varanus* from Artenay, France was dated to 17-18 Ma, and a fossil from Middle Miocene Kazakhstan is also described as *Varanus* (Rage and Bailon, 2005).

The genus *Varanus* is the largest living monitor lizard group and the sole surviving member of the family Varanidae. Currently, a total of fifty-eight (58) species of *Varanus* are taxonomically recognised and almost half of them occur in Australia. Therefore, global species diversity of monitor lizard is the greatest in Australia, where twenty-four (24) species are distributed. There are eight (8) species distributed in Asia, thirteen (13) species in Wallacea, thirteen (13) species in New Guinea and islands in the Pacific, and four (4) species in Africa (Böhme, 2003). The closest living relative to *Varanus* is the Earless monitor (*Lanthanotus borneensis*). This species is only known from the island of Borneo and is the only member of the family Lanthanotidae (Pianka, 2004). Other living relatives of *Varanus* are two species within the family Helodermatidae. The Gila Monster (*Heloderma suspectum*) and the Mexican Beaded lizard (*Heloderma horridum*) are the surviving members of this family that includes a total of six genera (Beck, 2004). The fossil helodermatids of *Lowesaurus*, *Primaderma*, *Paraderma*, *Eurheloderma*, and *Estesia* were found in north America, Europe, and Asia (Nydam, 2000).

Phylogenetic relationship among species of extant monitor lizards is generally resolved since the work of Ast (2001), which was based on mitochondrial DNA (Figure 1). In this phylogeny, the six major monophyletic clades seem to reflect the current geographic distribution. The African clade is basal (outgroup) to all other *Varanus* included in the study and each of the Asian and Australian lineages forms two subclades. Bridging the Asian and Australian subclades is a lineage of three species including the Komodo monitor (*Varanus komodoensis*). It is interesting to note that some Asian and Australian species are distributed in the Indonesian Archipelago. Consequently, the Asian clade was determined to be Indo-Asian, whereas the Australian clade was labelled as Indo-Australian (J. Ast, personal communication). More interestingly, this phylogeny shows a topology that suggests a faunal transition in Wallacea, which is situated in the eastern part of the Indonesian Archipelago (Figure 2). Wallacea is a biogeographic realm that is geographically located between the Oriental and Australian realms. In particular, southern Wallacea harbours both Oriental and Australian varanids, for instance the Asian Water monitor (*Varanus salvator*), and the Komodo monitor. The Water monitor is phylogenetically affiliated with Asian species in the Indo-Asian clade, whereas the Komodo monitor is nested in the Indo-Australian clade (Figure 1). Besides these two species, a few other Asian and Australian monitor

lizards also occur in Wallacea, for example the Togian Water monitor (*V. togianus*) in the Togian Islands near Sulawesi, the Spotted Tree monitor (*V. timorensis*) in Timor Island in the Lesser Sundas, as well as the Panay monitor (*V. mabitang*) and Gray's monitor (*V. olivaceus*) in the Philippines (Böhme, 2003).

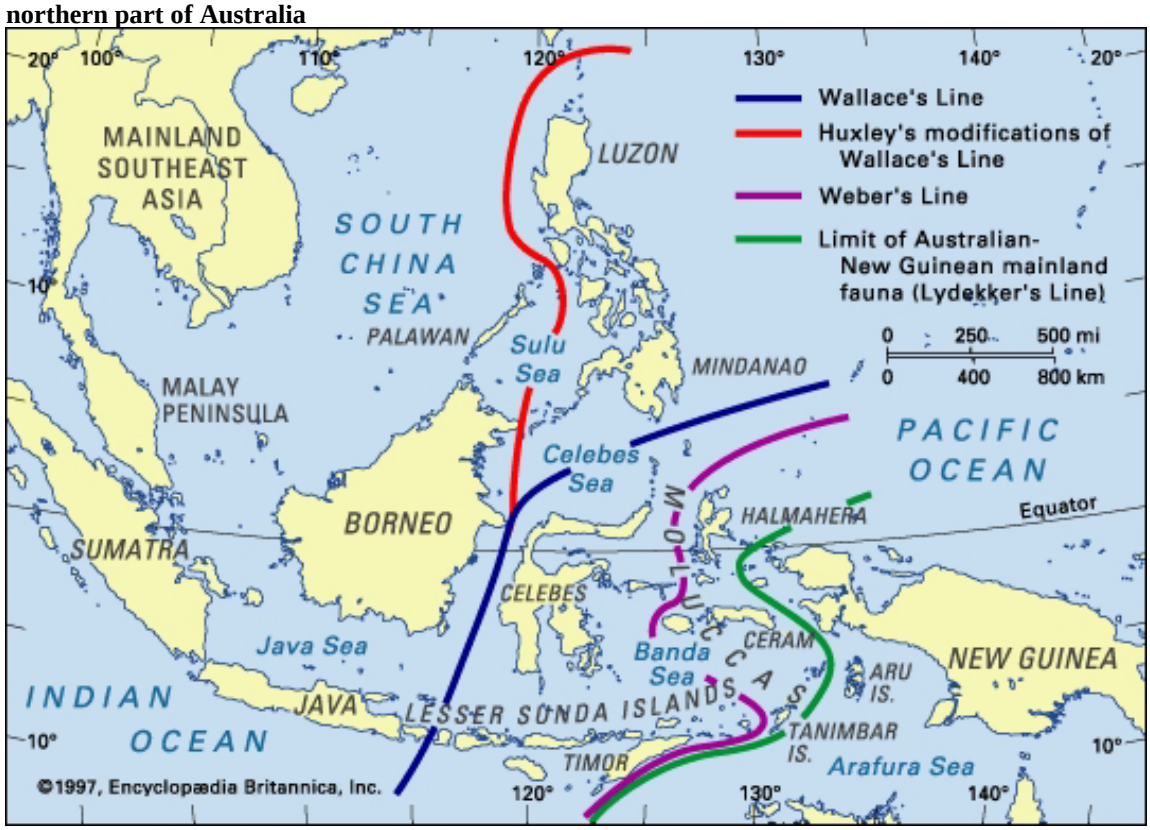
Figure 1. Phylogeny of extant *Varanus* based on mtDNA sequences (after Ast, 2001)



Wallacea is particularly an interesting region for the biogeography of *Varanus*. Besides the co-occurrence of Asian and Australian varanids in this region, endemic varanids are also present in this region, for instance the Komodo monitor in the Lesser

Sundas and the Togian Water monitor in the Togian Islands. Nevertheless, vertebrate endemism in Wallacea is common. For instance, the dwarf Lowland Anoa buffalo (*Bubalus depressicornis*), the Pygmy tarsier (*Tarsius pumilus*), the Maleo bird (*Macrocephalon maleo*), the Flores shrew, (*Suncus mertensi*), the Red-knobbed hornbill (*Aceros cassidix*), the Sulawesi toad (*Bufo celebensis*), and the Roti Island snake-necked turtle (*Chelodina mccordi*). Figure 2 shows Wallacea, which is defined by two biogeographical lines, i.e. Wallace's Line in the west and Lydekker's Line in the east. This region encompasses Sulawesi and the neighbouring smaller islands, Maluku Islands, Tanimbar Islands, and the Lesser Sunda Arc. Following a modification of Wallace's Line by Thomas Henry Huxley in 1868, Wallacea was redefined to include most of the Philippine Islands that are situated in the north of Sulawesi. Note that Palawan Island lies west of Huxley's modification line. For a review on the zoogeographical boundaries in the Indonesian Archipelago, a relatively detailed reference is given in Simpson (1977).

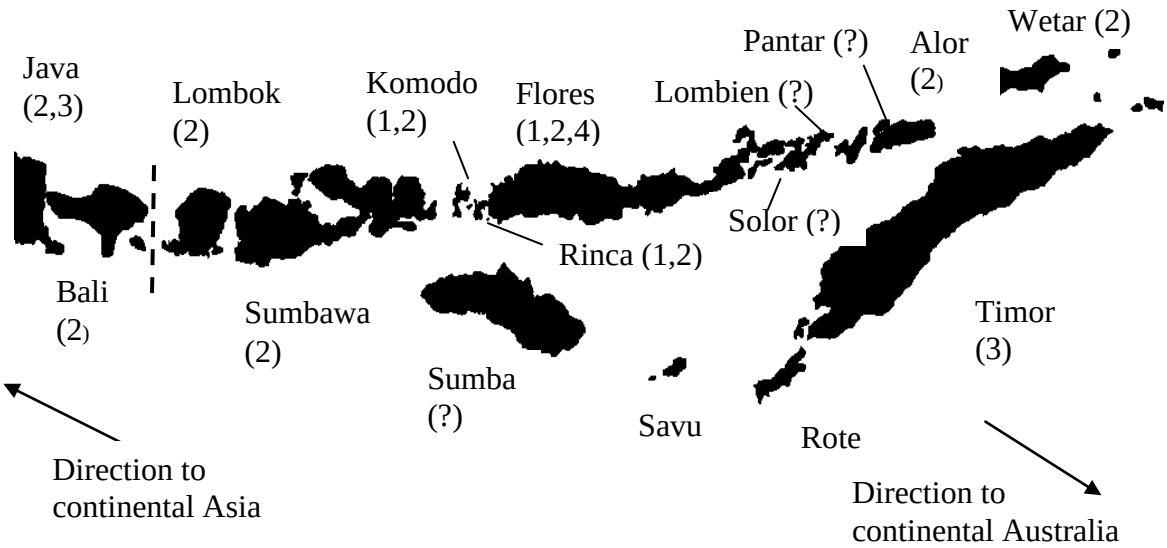
Figure 2. The biogeographic realm of Wallacea with reference to Southeast Asia, New Guinea, and northern part of Australia



Four species of monitor lizards, i.e. *V. komodoensis*, *V. salvator*, *V. timorensis*, and *V. auffenbergi* are currently known from the Lesser Sunda Islands in southern

Wallacea (Böhme, 2003). All these, but the Asian Water monitor (*V. salvator*), are endemic to the Lesser Sunda Islands. In spite of their current distribution in the Lesser Sundas, the Komodo monitor and the Spotted Tree monitor are phylogenetically closer to Australian varanids (Figure 1). Thus, it is probable that varanid radiation progressed from Australia to southern Wallacea. Indeed, the Lesser Sundas are situated on the margin between the Eurasian Plate and the Indian-Australian Plate (Hall, 2002), suggesting a varanid radiation influenced by plate tectonics. A long period of faunal radiation in Wallacea is generally associated with some geological processes such as plate tectonics and multiple volcanic eruptions, which commonly occur since the Cenozoic (~25 Ma) in Southeast Asia through to Southwest Pacific (Hall, 1998, 2001). Consequently, it is reasonable to hypothesize a scenario of varanid colonisation from Australia to the southern Wallacea. On the other hand, the Asian Water monitor, *Varanus salvator*, is also distributed in the Lesser Sundas (Figure 3). This species is the most widespread monitor lizard in the world, with a distribution range reaching as far west as Sri Lanka and as far east as the Lesser Sunda Islands, Sulawesi, and the Philippines (Auffenberg, 1981, Shine *et al.*, 1996). It seems that the Water monitor does not disperse beyond Wallacea, although its ability to cross a water barrier might give rise to its successful colonisation on islands. Therefore, varanid dispersal in the Lesser Sunda seems to come from two directions. This bi-directionality of varanid radiation into the Lesser Sunda Islands reflects the importance of this region as a transition zone, which demarcates the distribution of Asian and Australian varanid lizards.

Figure 3. The distribution of the Komodo monitor (1) and the Asian water monitor (2) on major islands in the Lesser Sunda. Fossil varanids reported from around the region include *V. bolkayi* (3) and *V. hooijeri* (4). Dashed line is Lombok Strait that marks the distinction of Greater Sunda region in the west of the line from the Lesser Sunda region eastward of the line.



Several factors such as geography, species dispersal ability, species adaptation to novel environment, and species mode of life may be related to the co-occurrence of Asian and Australian varanids in the Lesser Sunda Islands. Dispersal ability and adaptability more than geographic barrier can affect distribution range of a varanid species, such as in the case of the Asian Water monitor. Besides its excellent swimming ability, the Water monitor seems to adapt well in different habitats, even in an altered environment (Gaulke *et al.*, 1999). On the other hand, the distribution of the Bengal monitor, *V. bengalensis*, seems to be more limited, probably due to its apparent terrestrial nature. Populations of Bengal monitor can be found as far west as Iran and as north as Afghanistan, while the island of Java in Indonesia is both its easternmost and southernmost distribution (Auffenberg, 1994). The distribution of this species no further east than Java may indicate the presence of a geographical barrier to dispersal eastward of Java, which is crossed by the Asian Water monitor. Similarly, many monitor lizards are restricted to Australia and New Guinea. Some species such as Gould's monitor (*V. gouldi*) and the arboreal Black-headed monitor (*V. tristis*) are restricted to the Australian continent, where a relatively wide range of prey choice is available. Although they may be adaptive to many different habitats and are distributed on the peripheries of the continent (King and Green, 1999), dispersal out of Australia seems to be hampered by the presence of sea. Further, a species' mode of life rather than a geographic setting may be more related to its tendency to disperse, such as in the endemic Crocodile monitor (*V. salvadorii*) in New Guinea. The arboreal Crocodile monitor does not seem to occur in Australia, in spite of the geographic proximity of this island to Australia and the presence of historical land bridges between the two lands during periods of continental glaciation in the late Pleistocene (Voris, 2000). One can speculate that the arboreal mode of life in this species may have limited its distribution in the New Guinean rainforests. Apparently, the Crocodile monitor did not disperse to Australia because of an unsuitable environmental condition of the land bridges, which is thought to be similar to that of a tropical savanna (Torgersen *et al.*, 1988).

1.2. The Lesser Sunda Islands and hypotheses for global monitor lizard radiation

Being the transition zone for the Asian and Australian radiations of monitor lizards, the Lesser Sunda region is a critical location for inferring dispersal events of these lizards from Asia into Australia and *vice versa*. The distribution of Komodo monitor in this region may be related to the possible dispersal of its ancestor from either

Australia or Asia, each of which could be a region of origin for this species. On the other hand, the distribution of an Asian species, i.e. the Asian Water monitor in this region clearly indicates an Asian radiation through an over-water dispersal to the easternmost island within the arc of islands, i.e. Wetar (Figure 3). An inference of dispersal events in the Lesser Sunda region can also be used to elucidate a possible scenario of monitor lizard radiation in Wallacea. The fossil *V. hooijeri* from Flores dated to the Holocene has similar features to the Philippine endemic *V. olivaceus* (Brongersma, 1958). This character similarity between extinct and extant varanids may be used to postulate a north-south dispersal and radiation in Wallacea. The Lesser Sunda Islands is one of the key-regions in global varanid radiation. This chain of islands in the southern Wallacea is an important region to infer a recent dispersal of monitor lizards from and to Asia in a phylogeographic framework. Another key-region in global varanid radiation is the shared land margin between continental Asia and continental Africa, through which Asian varanids may have dispersed recently to Africa and *vice versa* (Arida and Böhme, 2010). The patterns of varanid distribution and dispersal in the Lesser Sunda Islands can be used in addition to phylogeny for verifying either of the two competing hypotheses on the origin of global monitor lizard radiation. A Laurasian hypothesis assumes an origin of monitor lizards in central Asia and the subsequent radiation through the Indonesian Archipelago to Australia and Pacific islands. On the contrary, a Gondwanan theory predicts the emergence of varanids on Gondwanaland. Following a vicariance of the populations in Africa, India, and Australia, some monitor lizards occurring in Australia subsequently radiate through the Indonesian Archipelago. Either scenario involves a radiation route via the Indonesian Archipelago, for which the Lesser Sunda Islands serve as a transitional zone. Therefore, the patterns of species distribution and dispersal in this region are suggestive of the origin of their radiation.

The ancestors for the endemic varanids in the Lesser Sundas may have come from Asia or Australia and speciation patterns across these islands are presumably a result of vicariant radiations and dispersals. A radiation of varanids in the Lesser Sundas by vicariance may be postulated, with an assumption that the there is a considerably long period of isolation between two populations. On the other hand, over-water dispersal may explain a speciation process across these islands, some of which are volcanic. Based on the Laurasian hypothesis of varanid radiation, the endemic varanids of the Lesser Sundas may have evolved on these islands from their Asian ancestor. This

ancestral form dispersed through the Indonesian Archipelago and might even reach Australia as early as Late Miocene (~10 Ma), at the time the Lesser Sunda Arc had been fully developed and was similar to its present position (Hall, 2002). This dispersal route of monitor lizards through the chains of islands between mainland Asia and Australia is consistent with molecular phylogenetic data (e.g. Fuller *et al.*, 1998, Amer and Kumazawa, 2008). In contrast, an Australian ancestor for the Lesser Sunda endemic varanids would be consistent with the Gondwanan hypothesis. This scenario for varanid radiation from Australia to Asia is presented based on two lines of evidence: the primitive morphology of Australian varanids and the significant divergence among Asian and Australian species inferred from immunological data (Hutchinson and Donnellan, 1993). In addition, the diversity of varanid lizard species in the world is the greatest in Australia, suggesting a diversification on this continent. Following this diversification in Australia, varanids radiate to Asia by dispersing across the Indonesian Archipelago. This view of Gondwanan origin for varanid lizards radiation is supported by a molecular phylogenetic dating, which gives an estimate of ~120 Ma since the separation of two clades representing Asian and Australian *Varanus*. This divergence time indicates a varanid rafting to Southeast Asia on a smaller fragment that separated from the northern parts of Australia or on the Indian subcontinent, and refutes the hypothesis of dispersal events from Southeast Asia to Australia and the Pacific islands about 25 Ma (Schulte *et al.*, 2003). Nevertheless, fossil evidence indicates the origin of early monitor lizards in central Asia (Estes, 1983), although the genus *Varanus* may have emerged on Gondwanaland. These theories of global varanid radiation have been explored in a phylogenetic study involving both extant and extinct Platynotan species. The results of this study suggested, that the Laurasian hypothesis is largely the most parsimonious. However, the alternative Gondwanan theory cannot be rejected because the genus *Varanus* is closely related to Serpentes, which is thought to originate in Gondwana (Peppin, 1999).

The competing hypotheses of global varanid radiation may be relevant as background information to deduce the pattern of dispersal for Komodo monitor in southern Wallacea. The Laurasian hypothesis of monitor lizard radiation from Asia may support a hypothesis of west-to-east dispersal in the Lesser Sunda Islands. The ancestral Komodo monitor might have emerged in Asia and dispersed to the Lesser Sunda Islands. Indeed, fossil *Varanus bolkayi* from the mid-Pleistocene Trinil (East Java) resembles the

Komodo monitor (Hooijer, 1972). This evidence is consistent with a historical distribution of this species in Java, but see Auffenberg (1981) for rebuttal. Another fossil of the Komodo monitor was found on Flores, one of the islands in the Lesser Sunda. The fossil teeth were dated to mid-Pleistocene, between 0.85-0.90 Ma (Morwood *et al.*, 1999, van den Bergh *et al.*, 2001). On the contrary, the discovery of fossil vertebrae identified as belonging to the Komodo monitor in Australia has provided a strong evidence for the origin of the Komodo monitor in Australia (Hocknull *et al.*, 2009). Additionally, two vertebrae of monitor lizard dated to the Pleistocene Atambua, in western Timor, may represent a subspecies of Komodo monitor (Hooijer, 1972). This fossil evidence from Australia and Timor implies the emergence of ancestral Komodo monitor in Australia, as well as providing support for the Gondawanan theory of varanid radiation. Despite this controversy, it is interesting to note that the Komodo monitor may represent a shift between Asian and Australian varanids. The phylogenetic position of the Komodo monitor (Figure 1) reflects molecular characteristics that are transitional between Asian and Australian varanids. Parallel to its position in the phylogeny of *Varanus*, current distribution of the Komodo monitor in southern Wallacea may indicate an evolutionary process that occured in the intermediate environment between Asian and Australian biogeographic regions.

A scenario of Komodo monitor dispersal in Australasia may be evaluated in light of the sequence of fossils found in this region. To date, the fossil vertebrae from Pliocene Queensland, Australia, is the oldest record for the Komodo monitor beyond its current distribution range. Younger fossils were also recovered from Pleistocene deposits in Australia (Hocknull *et al.*, 2009), Timor (Hooijer, 1972), and Flores (Morwood *et al.*, 1999, van den Bergh *et al.*, 2001). It seems that the Komodo monitor dispersed to the Lesser Sunda Islands during the Pleistocene. This scenario of dispersal from Australia seems to be corroborated by phylogeny. The Komodo monitor is phylogenetically closer to extant Australian than to Asian varanids (Ast, 2001, Schulte *et al.*, 2003). According to this phylogeny, the Komodo monitor is sister to the Australian endemic Lace monitor (*V. varius*) and also a close relative to the New Guinean endemic Crocodile monitor (*V. salvadorii*). A period of evolution for these three *Varanus* species or their ancestor in Australasia may be explained by the presence of land connections among islands in the Lesser Sundas as well as between New Guinea and Australia during the Pleistocene (Voris, 2000). In addition, the evidence of a giant varanid species

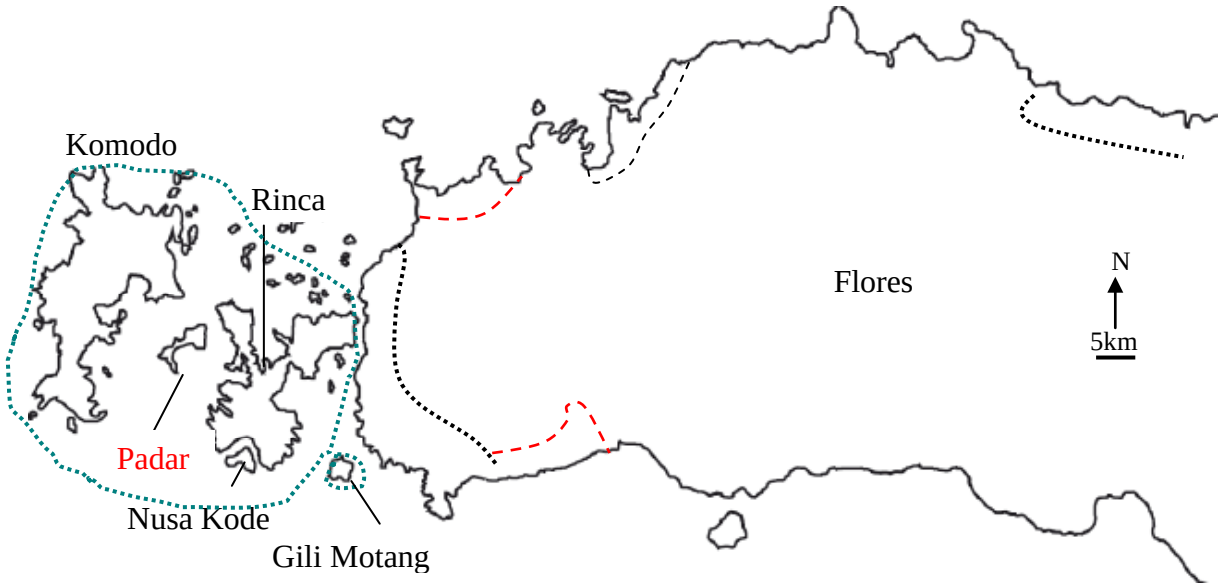
from Plio-Pleistocene Australia seems to be concordant with the timing of dispersal scenario for the Komodo monitor from Australia. Two fossils of the Giant Roamer (*V. priscus*) were dated to at least 4.5 Ma in the early Pliocene (Molnar, 2004a) and ~0.19 Ma in the late Pleistocene (Pianka and King, 2004). This species of giant varanid from Central Australia is regarded as a close relative of the Komodo monitor (Conrad *et al.*, 2008). Moreover, an examination of the cranial osteology of the Giant Roamer indicates some characteristics that suggest this species as sister to the Komodo monitor (Head *et al.*, 2009). The Giant Roamer is thought to have been widespread on the Australian continent and probably a giant ancestral for the more modern varanids. The body size of *V. priscus* is about one-third larger than the modern adult *V. komodoensis* (Wroe, 2002). The Komodo monitor could actually be an example of island dwarfism, despite the fact that it is the largest living lizard. Island dwarfism has been also found in a species of elephant from the Lesser Sundas. The extinct Pygmy Elephant (*Stegodon florensis*) known from their remains found on Flores is estimated to be much smaller than the modern elephants (van den Bergh *et al.*, 2009). Therefore, it is likely that the ancestral Komodo monitor arrived in the Lesser Sundas during the Pleistocene and subsequently evolved to a smaller size. Nevertheless, this hypothetical scenario of the dispersal from Australia is based largely on fossils from Australia. It is important to note, that fossil Komodo monitors from Southeast Asia seem to be scarce to counterbalance the existing theory of Australian dispersal. The Southeast Asian tropics are generally not the best location for fossilisation (Kidwell and Flessa, 1996), partly because of the climate and to some extent the growing human settlements.

1.3. Komodo monitor: the restricted distribution and population status

The Komodo monitor is currently known to occur only on five islands of various sizes in the Lesser Sunda Islands, Indonesia. The islands of Komodo, Rinca, Nusa Kode, and Gili Motang are situated within the Komodo National Park, whereas Flores lies just outside the eastern border of the national park. Komodo and Rinca are relatively much larger in size compared to Nusa Kode and Gili Motang (Figure 4). Komodo monitors have also been observed on the small island of Padar, which lies between Komodo and Rinca. In addition, the monitors have been reported to occur on Nusa Mbarapu, an islet that lies just off the southeastern coast of Komodo (Auffenberg, 1981). The population on Padar is thought to have become extinct following a sharp population decline caused by a series of fires that burnt most of the island's vegetation in the 1980s. Along with

fire, a significant depletion of prey species on this island may have exacerbated the decrease in population size (Sastrawan and Ciofi, 2002).

Figure 4. Current distribution range of the Komodo monitor in the Komodo National Park (within dotted blue curves) and Flores. Dashed black line shows reported occurrence of lizards in locations not sampled in this study, whereas dashed red lines indicate past occurrence of monitors reported by Auffenberg (1981). Dotted black line show study sites on Flores.



Current population of wild Komodo monitor is considered vulnerable due to the risk of extinction in the medium-term future (IUCN, 2007). Population size has been decreasing for about the past thirty years, while occupancy area is estimated to be less than 2,000 km^2 (Table 1). The Komodo monitor is a top predator within its distribution range. Adult animals prey on the Timor deer (*Cervus timorensis*), the Water buffalo (*Bubalus bubalis*), and wild boar (*Sus scrofa*). The monitors are threatened by a loss of their main prey, i.e. the Timor deer, due to the unregulated hunting practices of this ungulate. Human activities such as chasing the deers with the help of dogs and savannah burning have been reported to take place both in the National Park and on Flores. As a consequence of these hunting practices, the dogs remaining in the park area are becoming feral, thus posing a threat as invasive competitors for Komodo monitors. Moreover, forests and savannahs have been burnt to open agricultural land and to establish human settlements, resulting in the declining habitat and reduced habitat quality. In spite of this severe habitat fragmentation, the depletion of prey on islands was considered as a major threat to the population persistence (Goodwin *et al.*, 1997).

Table 1. The size of Komodo monitor populations across islands and the corresponding island size. The different area sizes for Flores represent the extent of study areas on the island.

Source	Komodo	Padar	Rinca	Flores	Gili Motang	Nusa Kode	Nusa Mbarapu	Total
Pfeffer (1959)	400-500	100	400-500	500	n/a	n/a	n/a	1,500-2,000
Auffenberg (1981)	2,348	60	792	2,448	40	25	n/a	5,713 (+1,500) #
	393 km^2	13.5 km^2	278 km^2	400 km^2	11.3 km^2	9.6 km^2	0.6 km^2	1,092.5 km^2
PKA (1991-1996)	1,687	n/a	1,110	66**	32	n/a	n/a	2,969
Ciofi & De Boer (2004)	1,150	0	1,110	69*	32	14	n/a	2,375
	340 km^2	20 km^2	210 km^2	70 km^2	10 km^2	7 km^2	n/a	657.0 km^2

*Study conducted on Wae Wuul and Wolo Tado Reserves
**Estimate from Wae Wuul only
Estimate of yearly hatchling recruit

The populations on Nusa Kode (7 km^2) and Gili Motang (10 km^2) may be of particular conservation concern, because a small population is susceptible to reduced genetic diversity and local extinction (Frankham, 1996b). The population sizes for these islands are estimated to be smaller than fifty individuals and these figures are much smaller than those for the larger islands (Table 1). A previous study on the population genetics of wild Komodo monitor using nuclear markers indicated a low genetic diversity for Gili Motang, while an observation on allele fixation in this small island population seems to warrant a further investigation for inbreeding depression (Ciofi and Bruford, 1999). If genetic diversity is low enough that it results in allele fixation, the small population on Gili Motang might suffer from an inbreeding depression that can lead to extinction. Thus, a sufficient amount of gene flow is necessary for Gili Motang to maintain its genetic diversity. However, the frequency of migration between Gili Motang and the nearby Flores seems to be low to allow for an adequate gene flow. Based on genotypic data, one individual was estimated to migrate every two generations (Ciofi *et al.*, 1999, Ciofi and Bruford, 1999). This relatively low frequency of migration between Gili Motang and Flores is probably related to the distance between the two islands. Parallel to the low genetic diversity on Gili Motang, the mean body mass of individuals on this island was relatively small compared with that of the population on Rinca. Likewise, the relative abundance on Gili Motang was lower than that on Rinca (Jessop *et al.*, 2007). Therefore, it is important to recognise a sign of inbreeding for Gili

Motang, in order to conserve the population from extinction. On the other hand, a very little information is currently available for the small population on Nusa Kode. An earlier study reported that the genetic divergence between Nusa Kode and the nearby population on Rinca was relatively low, suggesting a higher level of gene flow between the two populations (Ciofi, 2002). Nevertheless, more detailed information is needed to ensure an accurate assessment for this small island population.

Besides the restriction on only five islands, the extant populations of Komodo monitor are generally vulnerable to extinction due to a reduction in population size associated with disturbed environment. Given an estimate of population size of ~3,000 individuals in the 1990s (Table 1), about a quarter of this size is expected to mate and contribute in the maintenance of a viable population. The number of potential breeders is calculated as the effective population size (*Ne*), which ranged from 20-27% of the population sizes across study sites on four islands (Ciofi and Bruford, 1999). Thus, about 750 individuals are potentially effective to support a continuing population as a whole. This estimate of *Ne* is an example of population genetic data that is useful to evaluate population characteristics. In addition to the available data from population genetic study, information on the genetic diversity and population structuring is essential to assess population status for conservation.

1.4. Genetic diversity and population divergence

The long term survival of a species depends on the maintenance of its genetic variation, which often correlates with population fitness and is generally quantified at molecular level as genetic diversity (Reed and Frankham, 2003). The assessment of genetic diversity has been facilitated by the availability of various molecular markers and computer programmes designed to measure variations at molecular level. For a sound assessment of genetic diversity, the combination of molecular markers of different classes, e.g. mitochondrial and nuclear genes is generally a better approach than the use of one type of marker. A combination of different markers with different characteristics is expected to yield DNA sequences that provide adequate variations reflecting independent events during the course of population dynamics (Wan *et al.*, 2004). Knowledge of the genetic diversity of a species may be applied to the identification of areas for conservation priority. Moreover, variations in the level of genetic diversity among populations may imply a significant degree of genetic divergence that shapes the structure of the entire population. Thus, recommendations for a conservation programme

may be formulated based on the level of genetic diversity and the structure of the population as a whole.

The level of divergence among populations is calculated as Wright's *F*-statistics. These measures of population differentiation are commonly applied for allelic pairs of heterozygous nuclear DNA sequences screened for a number of loci, such as the microsatellites in the nuclear DNA (Charlesworth, 1998, Balloux and Lougon-Moulin, 2002). There are three indices in the statistics, i.e. F_{IT} (total fixation index), F_{ST} (fixation index), and F_{IS} (inbreeding coefficient). Each of these indices represents the degree of heterozygosity in a ranked population. F_{IT} measures the proportion of heterozygosity among all individuals in the whole population, while F_{ST} indicates heterozygosity level among subpopulations. On the other hand, F_{IS} is used to assess heterozygosity among individuals within a subpopulation. In addition, F_{IS} can be regarded as a coefficient of inbreeding. The absolute heterozygosity for each of the population ranks, i.e. within-subpopulation, among subpopulations, and total population, is determined based on allele frequencies. The absolute heterozygosity for within-subpopulation (H_I) is calculated as the average of observed allele frequencies of individuals within a given subpopulation, whereas the absolute heterozygosity for among subpopulations (H_S) is calculated as the average of expected allele frequencies among subpopulations with regard to Hardy-Weinberg Equilibrium (HWE). The absolute heterozygosity for the total population (H_T) is computed as the average of expected allele frequencies in the total population under HWE. The expected allele frequency distribution of HWE is given as follows:

$$1 = p^2 + 2pq + q^2$$
$$(AA) \quad (Aa) \quad (aa)$$

where p^2, $2pq$, q^2 are the frequencies of genotypes AA, Aa, aa, respectively in zygotes of any generation, and p and q are the allele frequencies in gametes in previous generation, and p+q=1 (Hartl and Clark, 1997).

Unfortunately, there are many assumptions that have to be satisfied for Wright's differentiation indices to be valid. In order to reach HWE, organisms are expected to reproduce sexually and to mate randomly, to have two alleles in their diploid loci, and to have identical allele frequencies in male and female individuals. Furthermore, population has to be very large with non-overlapping generations. Migration, mutation, and selection are assumed negligible. Although these assumptions are very restrictive,

the theory is meant to be a generalised model of population divergence. If two populations diverge, no other evolutionary forces are expected to be at work apart from those imposed by the reproduction process of the organism itself. Further, Wright's F_{ST} can also be interpreted as an inbreeding index, because the accumulation of homozygous alleles is a result of reduced heterozygosity, i.e. heterozygosity= 1- homozygosity. A higher value of F_{ST} may be interpreted as a lower chance for inbreeding to occur, whereas a higher chance of inbreeding may be indicated by a low F_{ST} value (Hartl and Clark, 1997). Furthermore, F_{ST} may be regarded as an indirect measure of relative population fitness. For example, a population with a low F_{ST} is thought to have a low level of fitness, because it has a larger chance for inbreeding. And a population that has a larger chance for inbreeding is expected to have a low genetic diversity. Thus, the evolutionary potential of a population is measured in terms of its genetic diversity that is linked to a chance for inbreeding. To summarize, a given F_{ST} value is expected to demonstrate not only the level of divergence between populations, but it also indicates the level of inbreeding that is related to population fitness.

To date, F_{ST} has been used in population genetic studies aimed at species conservation, because this measure of heterozygosity may be used to evaluate population fitness (Reed and Frankham, 2003). In addition, the degree of genetic divergence computed in F_{ST} can also be related to the effective number of migrants, through the equation:

$$F_{ST} = 1/ (1+ 4N_e m_e)$$

where N_e is effective population size or the potential breeders that can maintain viable population and m_e is the effective migration rate assuming island model of migration (Balloux and Lougon-Moulin, 2002). Thus, F_{ST} links together the level of divergence among populations, the degree of inbreeding, population fitness, and the effective number of migrants between populations. Given this link between population parameters, F_{ST} is useful and convenient for an application in studies aiming to define population structure and population status for conservation.

An analogous measure of population differentiation to Wright's *F*-statistics is φ-statistics, which accommodates molecular data from mitochondrial DNA (mtDNA) sequences and restriction sites. The three indices in the statistics, i.e. φ_{ST}, φ_{SC}, and φ_{CT} can be calculated using AMOVA (Analyses of Molecular Variance) from a matrix of

squared distances between pairs of sequences in a given sample. Alternative distances that can be applied in the method include the mean number of restriction site differences, nucleotide diversity, and patristic distance along a given network. AMOVA is a direct modification of ANOVA (Analyses of Variance), which is a statistical measure of dispersion among samples. In ANOVA, one assumes that the samples are normally distributed. However, this assumption is inappropriate for molecular data. Consequently, the significance of variance components and φ-statistics in AMOVA are tested using permutation, which eliminates the assumption of normal distribution (Excoffier *et al.*, 1992). φ-statistics and *F*-statistics have been applied in many studies describing the pattern of divergence among populations and the subsequent population structure. For example, a genetic structure was demonstrated in the population of Asian elephants sampled from seven countries. A moderate level of genetic differentiation was found in the whole population with F_{ST}= 0.29 *P*= 0.0001 (Fleischer *et al.*, 2001). In the populations of Southeast Asian seahorses, the level of differentiation varies across species, ranging from moderate to very divergent, with φ_{ST}= 0.19-0.81 *P*= 0.0001 (Lourie *et al.*, 2005).

For many organisms, the level of genetic divergence and the subsequent genetic structuring in a population may be associated with its geographic position and the presence of barrier to dispersal. Generally, the distribution of populations on islands is expected to show a degree of genetic structuring, although a strong genetic structure may also be found in continuous array of genotypes, for instance a continental population with geographic barrier to dispersal such as mountains or rivers (Avise *et al.*, 1987). In addition, genetic structure can also be found in populations of species with considerable dispersal ability, such as the Australian magpie (Toon *et al.*, 2007) and sea turtles (Bowen and Karl, 2007). On the other hand, social organisation can be influential in the genetic structuring of a population, because it inhibits random mating among individuals within the population. The degree of population subdivision in several social mammals such as the Black-tailed Prairie dog (*Cynomys ludovicianus*) and the Red Howler monkey (*Alouatta seniculus*) have been shown to be influenced considerably by their social organisation (Storz, 1999). Regardless of the factors shaping the genetic structure of a population, an increased total genetic diversity may indicate the presence of structure. The total genetic diversity in a structured population is enhanced due to the accumulation of local genetic diversity found in the substructures (Amos and Harwood,

1998). From a conservation management perspective, knowledge of genetic diversity in structured populations is fundamental to evaluating population status.

1.5. Identification of conservation units

The identification of units for conservation is generally based on the levels of genetic divergence. The Evolutionary Significant Units (ESUs) and Management Units (MUs) are among those proposed to categorise conservation priorities. ESUs are genetically divergent populations that have been isolated in the past and evolve independently. ESUs can be recognised from the deep divergence between mtDNA lineages and the significant more recent divergence of nuclear genes. The genetic divergence based on mtDNA is expected to show a reciprocal monophyly, whereas the divergence at nuclear DNA may be used to verify the pattern of genetic divergence inferred using mtDNA. Additionally, nuclear marker can demonstrate the level of gene flow connecting populations at a more recent time, which otherwise cannot be determined from a maternal gene flow inferred using mtDNA. On the other hand, MUs are characterised by nuclear or mitochondrial divergence regardless of reciprocal monophyly. The criteria to characterise ESUs seem to be very restrictive, because monophyly may not be shown in studies at population level. Therefore, a common sense approach is needed in the application (Moritz, 1994). It is important to emphasize, that these criteria should be regarded as diagnostics for recognising ESUs rather than criteria to define ESUs (Paetkau, 1999). The recognition of ESUs is intended to preserve ecological and evolutionary processes. Because ESUs evolve independently in separate geographic locations and involve different ecological processes, they become the sources of genetic diversity across the whole population. Thus, maintenance of genetic diversity across ESUs implies a conservation of the evolutionary and ecological processes that shape a population. On the other hand, the degree of genetic divergence between MUs is less than that between the larger ESUs. Consequently, MUs may be considered as ecological units for population monitoring within an ESU (Moritz, 1999). ESUs and MUs have been determined for many threatened vertebrates, including a species of Rainbowfish in Australia (Zhu *et al.*, 1998), the Asian elephant (Fleischer *et al.*, 2001), the Giant tortoises of the Galapagos Islands (Beheregaray *et al.*, 2003), and the Grey parrot in Príncipe Island (Melo and O'Ryan, 2007). Notwithstanding the number of threatened species, for which ESUs have been already identified, the criteria for identifying ESUs still remain challenged by several aspects such as the exclusive use

of molecular data (Crandall *et al.*, 2000) and the implementation in the assessment of conservation status for a given species (Green, 2005). Although the matter of ESUs has remained debatable for more than a decade, current biological studies aimed at making conservation recommendations should take into account recent advances in molecular techniques. The use of molecular data is particularly important in the identification of ESUs because DNA polymorphisms can provide a relatively rapid assessment of population structure and demography. Moreover, molecular data may serve to complement other available data, such as ecology and behaviour. Nevertheless, two considerations should be taken when only molecular data set is available to characterise units for conservation. First, ESUs should be identified based on a substantial reproductive isolation between them. Second, ESUs should demonstrate a representation of important component(s) in the evolutionary inheritance of the species (Waples, 1991).

An identification of appropriate conservation units is essential to devising a management plan for the Komodo monitor. A review (Frankham, 1996a) revealed that the cases of extinction in recent reptiles are dominated by species inhabiting islands and that low genetic diversity is often exhibited by island endemics. Given these indicators, the restricted distribution of Komodo monitor to five islands in southern Wallacea implies that a sound conservation management plan is long overdue. A previous study on the population genetics of Komodo monitor has shown that the island of Komodo should be regarded as an ESU because the population on this island retained a level of genetic diversity that makes up an evolutionary potential, which should be managed separately. Komodo harboured the highest number of private alleles and was shown to be highly isolated. More than 20% of all variations found on this island were not shared with the other three island populations (Ciofi *et al.*, 1999). Nevertheless, this population genetic study could not resolve a structure among populations distributed on islands eastward of Komodo. Therefore, a genetic structure across the whole population still has to be determined to identify units for conservation. Moreover, the status for the small island populations is yet to be verified. It is important to assess the genetic diversity, divergence, and eventually the status for the small island populations because small populations are thought to be the most vulnerable to the effect of genetic drift (Lacy, 1987). Gili Motang is a small island with a considerably high level of isolation. The distance between Gili Motang and the nearest island of Flores is about 2.5 km (Figure 4). Based on microsatellite data, Gili Motang has been found to be moderately distinct

from all other sampling locations (F_{ST} = 0.20-0.56 P<0.01) and to exhibit a fairly low within-population genetic diversity. This level of differentiation suggests that dispersal between Gili Motang and the other populations is fairly limited. It has been thought that the degree of isolation for Gili Motang is related to the strong sea current in the area, which may add to the significant allele fixation due to genetic drift in this small island population (Ciofi and Bruford, 1999).

1.6. Phylogeographic studies of *Varanus*

The field of phylogeography deals with the extent to which the geographic distribution of organisms is linked to their phylogenetic relationships. Studies in phylogeography combine intraspecific genealogy (intraspecific phylogeny), geography, and dispersal to infer the distribution of lineages over space. An influential factor affecting the geographic distribution of an organism is its ability to disperse, although behavioural factors such as site-fidelity (philopatry) and social system may offset dispersal. Furthermore, dispersal may also be influenced by a geographic setting, for example the presence of a water body as a barrier to dispersal for terrestrial species (Avise, 2004). Other factors that may influence the distribution of lineages over geography include geological events, e.g. plate tectonics that may facilitate dispersal (Bossuyt and Milinkovitch, 2001) and the emergence of volcanic islands. The Mangrove monitor (*Varanus indicus*) is an example of varanid species, of which geographic distribution is influenced by volcanism. This species has expanded its distribution onto Long Island in the northeast of New Guinea, which emerged as a result of volcanic activity. The Mangrove monitor is a tramp species, which can colonise a new area relatively easily, but compete poorly in an established community (Cook *et al.*, 2001). However, a re-evaluation of the *V. indicus* spp. complex revealed a number of new varanid species within the last decade (Böhme, 2003), suggesting a call for studies aimed at clarifying the relationships among island populations as well as at determining speciation patterns in this species complex. Phylogeography may be applied to unravel the relationships among these varanid populations, by linking the species phylogeny and their current distribution on islands with their dispersal pattern. Moreover, phylogeography may also be performed to shed light on the possible patterns of diversification that leads eventually to speciation in this group. Indeed, phylogeographic studies of many species have shown deep mtDNA divergence that corresponds to their

geographic distribution (Avise *et al.*, 1987), and such a relationship may be shown for the Mangrove monitor.

At present, studies on the phylogeography of varanid lizards are only rarely found in the literature. Among these studies is an investigation on Rosenberg's monitor (*V. rosenbergi*) populations in southern Australia that revealed a population structure comprising five ESUs (Smith *et al.*, 2007). One out of these five conservation units was found to correspond to the threatened populations in the southeast. Based on partial ND4 gene and three tRNA genes, the results of this study showed thirteen haplotypes across the distribution range of Rosenberg's monitor in Australia. These haplotypes were clustered in five monophyletic clades, each of which represents one ESU that is located in disjunct populations in Western Australia, New South Wales, mainland South Australia, Kangaroo Island, and Sir Joseph Banks Islands. It was discovered in the study that larger distribution breaks among mainland populations were consistent with those of other reptiles in Australia. In addition, the latter two island populations off the South Australian coast were grouped in two separate clades. Another phylogeographic study on monitor lizards discusses the relationships among populations of the Asian Water monitor (*V. salvator*) in Sulawesi and its satellite islands. The results of this ongoing study are expected to contribute a significant knowledge on the genetic structure and varanid diversity in Wallacea (A. Koch, personal communication).

The characterisation of varanid populations has received minimal attention from the scientific community (e.g. Smith *et. al.*, 2007 and A. Koch, personal communication). Currently, many populations of varanid lizards seem to face threats from commercial exploitations such as the harvest for skin and pet trade (Shine *et al.*, 1996, Pernetta, 2009). Informed conservation efforts are essential to ensure population viability and a characterisation of population genetic structure would be an effective method for recognising areas of conservation priorities, as exemplified in the study for Rosenberg's monitor (Smith *et al.*, 2007). Moreover, a systematic knowledge on the genetic diversity and relationships among the Asian Water monitor populations is fundamental to help in the monitoring programmes for sustainable harvest. One explanation to the very small number of phylogeographic studies on varanid lizards may be attributable to technical difficulties, which include labour-intensive sampling of very agile animals capable of escape from traps. Some species are even capable of multiple methods of dispersal. Thus, sufficient knowledge and effort are needed to establish an

efficient trapping scheme for data collection. Furthermore, one would need a considerable amount of time and energy to deal with the permit for collecting samples or specimens, because most species of monitor lizards are under law protection. Nevertheless, the population status of monitor lizards should still be evaluated preferably using phylogeographic methods, in order to ascertain sufficient knowledge for formulating priorities in conservation management practices.

Phylogeography is an appropriate approach to infer the geographic distribution of lineages and to characterise population genetic structure for conservation. However, the use of phylogeographic information for conservation may still be maximised by incorporating additional information such as environmental alteration, including change in climate at local scale. The current pattern of lineage distribution and population structure may change faster than expected with a change in climate within the distribution range of a species. A changing climate may enhance dispersal, which in turn influences lineage distribution and the subsequent population structure. On the contrary, a population may adapt to the local climate than dispersing because of the availability of food. Climate is a key factor in the distribution of monitor lizards in general. All extant varanids are found in areas of relatively warm climate in Africa, Asia, and Australia, suggesting a geographic delimitation by temperature. Interestingly, varanids do not occur in the South American tropics. One explanation for this is the event of continental drift that started separating South America from Africa in the Cretaceous. The ancestors of modern varanids were radiating from Asia to Africa during the Cretaceous (Estes, 1983). Nevertheless, South America had already drifted away from Africa, and a colonisation from Africa into South America was unlikely. Thus, monitor lizard distribution generally ecompasses warmer regions, with an exception of South America. It is interesting to note, that some varanid lizards inhabit extreme habitats such as the deserts, where the climate can be particularly hot. The Desert monitor (*V. griseus*) is distributed in the Sahara in northern Africa and some arid areas in Asia. Seven species of varanids, i.e. *V. brevicauda, V. tristis, V. gilleni, V. caudolineatus, V. eremius, V. giganteus, V. gouldi* co-occur in the Great Victoria Desert in Australia. These arid-adapted varanids are thought to have an extraordinary adaptation to climate. The high temperatures in the deserts that are lethal for many species seem to be within the range of tolerance for these varanids (Pianka, 1994). One may speculate that the climate in the deserts precludes these varanids either to select a suitable ecological niche in this arid

region, for example areas with relatively dense bush growth, or to shift their physiological regime. A comparative phylogeographic study may be conducted to shed light on the presence of a genetic structure for desert varanids, which may be explained by climatic adaptation across multiple species.

A phylogeographic study of the Komodo monitor may provide an insight on the genetic structuring and its restricted distribution in the Lesser Sunda Islands. Although the genetic structure and limited occurrence of the Komodo monitor may be linked to some climatic factors, a comparative approach should be applied to investigate the possibility of climate as a factor that shapes its genetic structure and limits its distribution. It is interesting to consider, that the limited distribution of the Komodo monitor may be driven by a narrow tolerance to climatic variation. Nevertheless, a similarity in climate in the areas beyond its current distribution range may reflect a larger distribution for this varanid in the past. The present climate in the Lesser Sunda is heterogeneous, with most islands in this region are characterised by seasonal dry periods that vary from five to eight (5-8) months per year (Monk *et al.*, 1997). Apparently, a similar climate has prevailed in the past in a wider area in the Australasian region. In the Lesser Sunda Islands, a semi-arid condition occurred in the Last Glacial Period (LGP), about 18,000 years ago. During this period, a putative dry savannah corridor stretched from the present-day Indochina through to the Lesser Sunda Islands. This open habitat was presumably a barrier to dispersal for rainforest species, but otherwise a route for early hominid dispersal to Australia (Bird *et al.*, 2005). Thus, it is possible that the Komodo monitor was once also distributed in more northern areas than the Lesser Sunda Islands. However, there is no evidence so far to support the possibility that the Komodo monitor has occurred on this historical savannah corridor. Therefore, it remains to be investigated, whether semi-arid climate is one of the factors that delimit its current distribution range in the Lesser Sunda Islands. In addition, the heterogenous climate in this region may also influence the genetic structure of the Komodo monitor population.

1.7. Aim of Study

The aim of my study is to determine the genetic structure of wild Komodo monitor population using mitochondrial DNA (mtDNA) sequences. Results of this study are expected to provide some information on the genetic characteristics of the population from molecular data, including the levels of genetic diversity and genetic divergence. This study is also a part of a larger collaborative research project, which is aimed at assessing the status of wild Komodo monitor population, in order to help in the design of conservation management. Therefore, the outcomes from this study may complement the information obtained previously from several investigations on ecology (e.g. Auffenberg, 1981, Sastrawan and Ciofi, 2002, Jessop *et al.*, 2007) and population genetics based on nuclear markers (Ciofi *et al.*, 1999, Ciofi and Bruford, 1999). I used the Control Region I (CRI), one of the major non-coding regions within the mitochondrial genome of the Komodo monitor, to evaluate the genetic diversity, genetic divergence, and genetic structure of the extant Komodo monitor population in the Lesser Sunda Islands. Small island populations are my particular focus, owing to the concerns on small isolated island populations of endemic species from conservation and scientific perspectives. I constructed a genealogy of mtDNA haplotypes to describe relationships among maternal lineages and evaluated two population genetic structures using phylogeographic methods, which include Analyses of Molecular Variance (AMOVA) and a statistical parsimony analysis. In light of paleogeographic and paleontological data of Southeast Asia and Australasia, I discussed putative ancestral radiation of the Komodo monitor in Australasia.

Chapter 2

MATERIALS AND METHODS

2.1. Samples and laboratory processing

Tissue samples were collected during a period of over ten years (1994-2005) by the Komodo monitor field research team in a larger research project with an objective to provide an assessment of population status for conservation. Sampling period was broken down into two subperiods, i.e. Subperiod I in 1994-1998 and Subperiod II in 2001-2005. Tissue samples collected in the Subperiod I are stored at the Department of Evolutionary Biology, University of Florence in Italy, whereas those collected during Subperiod II are deposited at Museum Zoologicum Bogoriense (MZB) in Cibinong, Indonesia. My data set comprises a total of three hundred and sixty six (366) whole blood samples. Sixty (60) of these samples were collected during the Subperiod I, whereas three hundred and six (306) samples were collected during the Subperiod II. Additionally, six (6) samples of the Lace monitor were added to make up the out-group population. These samples were three (3) whole DNA extracts from the University of Florence, Italy and three (3) muscle or organ tissues sampled from frozen specimens available at the Zoologisches Forschungsmuseum Alexander Koenig (ZFMK) in Bonn, Germany. A list of samples with details on the location of capture, sampling date, field code, and body condition is attached in Appendix A. The whole DNA for this study was extracted using Phenol-Chloroform method and the target sequences were obtained from the DNA template amplified using standard Polymerase Chain Reaction (PCR) technique. I conducted both extraction and PCR processes at the Molecular Genetics Laboratory at the University of Florence, Italy, as well as at the Genetics Laboratory at the Museum Zoologicum Bogoriense (MZB), Indonesia. Further, I performed a genetic analysis for PCR products at the Molecular Genetics Laboratory at the University of Florence, Italy. The extraction and PCR protocols are reported in Appendices B and C, respectively.

2.2. The genetic marker: Control Region I (CRI)

I used the Control Region I (CRI) as the genetic marker in my study. The Control Region (also named D-loop region) is a non-coding segment of the vertebrate mitochondrial genome. The Control Region consists of conserved sequence as well as

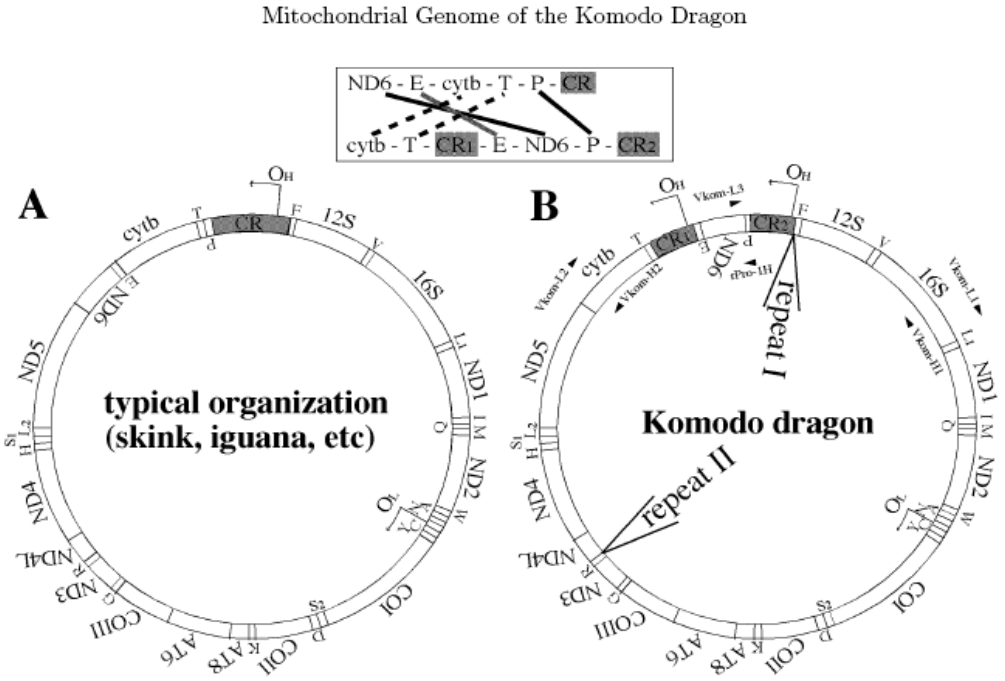
variable blocks and is estimated to evolve four to five times faster than the entire mitochondrial genome. The Control Region is involved in the initiation of DNA replication and transcription (Taberlet, 1996b). Within the vertebrate Control Region, short tandem repeats of variable lengths are often found. These short nucleotide sequences repeated in tandem are called as Variable Number of Tandem Repeats (VNTRs), which generate size heteroplasmy, a phenomenon of differential lengths among mtDNA molecules within an individual animal. VNTRs are related to the termination of transcription in mtDNA molecule (Lunt *et al.*, 1998). It has been suggested, that size heteroplasmy may occur in the Control Region, due to an accumulation of insertions or deletions (indels) caused by the absence of coding constraints in this region. In other words, the lack of coding constraints is associated with base substitution frequencies in this segment (Brehm *et al.*, 2003).

The Komodo monitor possesses a double Control Region within its mitochondrial genome, as also the case in many species of snakes. The typical vertebrate Control Region is flanked with Proline tRNA ($tRNA^{Pro}$) and Phenylalanine tRNA ($tRNA^{Phe}$). Due to an extensive gene reshuffling and a duplication of the Control Region, the mitochondrial genome of the Komodo monitor has a new gene arrangement. However, the Control Region II still retains the typical flanking genes, i.e. $tRNA^{Pro}$ and $tRNA^{Phe}$ (Figure 5). The two Control Regions are very similar in their sequence but do differ in their length. The length of the Control Region I (CRI) sequence is 727 base pairs (bp), whereas the Control Region II (CRII) is 849-bp long. The gene rearrangement in the Komodo monitor mtDNA is likely to take place through a slipped-strand mechanism, by which nucleotides are incorrectly paired during replication. Further, the occurrence of mispaired nucleotides may be facilitated by the presence of tandem repeats that increases the chance of replication errors within the Control Region. In addition, slipped-strand mispairing is also suggested as a means to maintain the double-copy of the Control Region (Kumazawa and Endo, 2004).

The Komodo monitor mitochondrial genome contains all the thirty-seven (37) genes characteristic of vertebrates (Figure 5). Nevertheless, the organisation of these genes is different from that commonly found in vertebrates, due to the aforementioned rearrangement. A tandem repeat can be found at the 3' end of the CRII and consists of at least six (6) units of 115-bp long repeat, which continues near the 5' end of the phenylalanine tRNA ($tRNA^{Phe}$) for at least another six (6) times. This region of tandem

repeat is called as Tandem Repeat Region I. An unusually long tandem repeat is inserted between Arginine tRNA ($tRNA^{Arg}$) and ND4L gene. This long region of tandem repeat is named Tandem Repeat Region II, which contains 15 units of 106-bp long repeat that makes up a total length of about 1.6 kilo base pairs (kbp). In addition, a common tandem-repeat region in vertebrate mtDNAs can also be found in the Komodo monitor mtDNA genome, i.e. at the 5' end of each Control Region. This 36-bp repeat unit extends from the $tRNA^{Thr}$ flanking the CRI and continued four (4) times. An identical unit starts just after the $tRNA^{Pro}$ flanking the CRII and continues five (5) times. Each of these sequences of minor tandem repeat region within the CRI and CRII is termed Tandem Repeat Region III (Kumazawa and Endo, 2004).

Figure 5. Structures of typical mtDNA gene arrangement in Squamates and gene rearrangement in the Komodo monitor mtDNA (after Kumazawa & Endo, 2004)



The choice of mtDNA Control Region I (CRI) as the molecular marker in this study is aimed at detecting variable nucleotide sites that are assumed to result from neutral mutation. A previous study investigating the pattern of population divergence among populations of the Komodo monitor used Simple Sequence Repeats (SSRs) or microsatellite as a marker, of which variations are also assumed to be neutral. Therefore, the assumption of neutral substitution in this study is consistent with the previous study focusing on population genetics, which I refer to in the discussion part of this thesis.

2.3. Laboratory techniques

2.3.1. Genomic DNA extraction

I performed a standard Phenol-Chloroform method to extract the whole-DNA from whole-blood samples. A small amount of a whole-blood sample was digested with Proteinase K overnight. This procedure was followed with an addition of one volume of phenol: chloroform solution and then another volume of chloroform, in order to remove proteins. The extracted DNA pellet was then washed with alcohol and resuspended in Tris-HCl EDTA solution (Sambrook *et al.*, 1989). Following an extraction procedure, the whole-DNA was ready to be amplified or it can be temporary stored at 4°C for immediate applications. A detailed extraction protocol is given in Appendix B.

2.3.2. Polymerase Chain Reaction (PCR)

Polymerase Chain Reaction is a standard automatic cyclic reaction for amplifying a length of DNA strand. A small quantity of template DNA is used in the reaction and amplification is done by polymerase enzyme that is isolated from *Thermus aquaticus*, a species of bacterium originally found in a hot spring in Yellowstone National Park, Wyoming, USA (Brock, 1994). The use of this enzyme marked the start of PCR automation, in that the thermophilic enzyme amplifies stretches of DNA at a certain range of temperature (~75°-80°C), which limits its function to only during the amplification step. This enzyme is inactive at temperatures beyond the range mentioned above, for example during a denaturation step, where amplification must not function. Principally, there are three steps in the automated PCR, i.e. denaturation of double helix template at high temperature, primer annealing on target location of a DNA helix at a lower temperature, and primer extension (amplification) at a temperature between that of denaturation and annealing temperatures (Saiki *et al.*, 1988). The details of the PCR protocol are presented in Appendix C.

A DNA primer (also called DNA oligonucleotide) is a short stretch of nucleotide sequence of about 15-30 bp that is designed to anneal on a targeted location of a complementary single-stranded DNA. During the course of a PCR, primers are amplified into a much longer sequence by DNA polymerase. Primers can be designed in pairs of forward and reverse short sequences, each of which starts before and after the target sequence, respectively. The primer pair of the Komodo monitor mtDNA CRI starts within the cytochrome b gene (forward primer) and within the ND6 gene (reverse

primer). This design amplifies about 700 bp of the target sequence. As a note, the Control Region II (CRII) has a very similar sequence but it has different flanking genes (Figure 5). Thus, the CRI sequences were specifically targeted using the above primers, which should not work to amplify CRII sequences. Additionally, it is important to design a primer that starts beyond the target sequence, because this design can maximise the possibility of obtaining a complete sequence of the target region.

2.3.3. Gel Electrophoresis

Gel electrophoresis is a technique that is originally used to visualise differences in protein mobility. Proteins, e.g. enzymes of different allelic forms can be determined using this relatively rapid and cost-efficient technique. The technique has also been used to define broad-scale phylogeographic patterns in allozyme studies. However, this technique is less suitable for fine-scale population studies that require details in nucleotide site variations (Parker *et al.,* 1998). I applied a gel electrophoresis technique to verify results from extraction and amplification procedures. This technique incorporates two steps, i.e. an infusion of DNA extract or PCR product into agarose gel stained with Ethidium Bromide followed by a charging of this gel with electricity. I mixed DNA extract or PCR product with a dye (Blue Bromophenol 6x) to facilitate product loading into each well in the gel. Charged with electrical current, several DNA extracts or PCR products ran through different lanes on the gel and the fluorescent DNA bands were observed on the gel illuminated with ultra violet light. To size these bands, I used dyed DNA ladder of 100-bp scale. The details on gel electrophoresis protocol are presented in Appendix D.

2.3.4. Automated Sequencing

Automated sequencing of mtDNA genes has been commonly applied to determine molecular variation at nucleotide level. The automated dye-terminator (fluorescent sequencing) method is nowadays widely applied due to the reliability of the method and the relatively low-cost process. This sequencing technique makes use four different fluorescent dyes for each base, i.e. Adenosine, Guanine, Cytosine, and Thymine (Takumi *et al.,* 1997). This method has been developed following earlier sequencing methods, i.e. chemical degradation (Maxam and Gilbert, 1977) and chain-termination (Sanger *et al.,* 1977). The automated fluorescent sequencing process starts with a “Cycle sequencing” reaction, by which a small quantity of the double-stranded

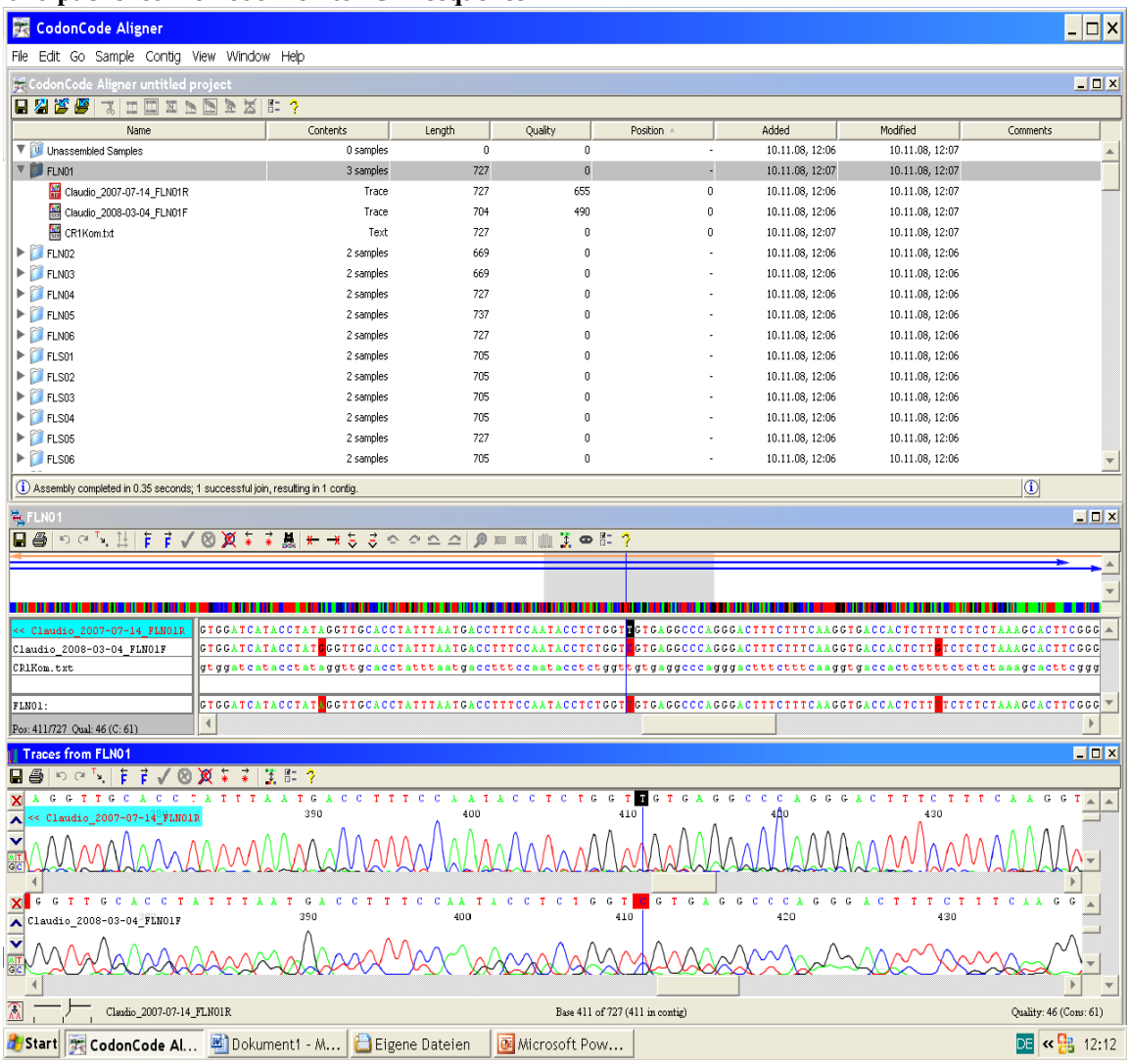
PCR products is separated and amplified by PCR to increase the sensitivity of the sequencing process. The primers for Cycle sequencing are designed to amplify strands from two opposite directions, as it is in the case of standard PCR. However, each of the primers is applied separately for each sample to ascertain a sequencing of each strand of the double helix at a time. Additionally, PCR products may be purified to remove chemical excess before a Cycle sequencing reaction. Following a Cycle sequencing reaction, DNA duplicates were precipitated using isopropanol to ensure an optimal sequencing process. I sequenced my PCR products using an ABI 3100 Genetic Analyzer™ automated sequencer equipped with sixteen (16) capillaries, which took about two and a half hours for running sixteen (16) samples each time. More details of my sequencing protocol are presented in Appendix E.

2.4. Sequence editing and haplotype analysis

DNA sequences were manually edited and analysed using a computer programme for DNA sequence alignment/ assembly, CodonCode Aligner™ version 2.0, to collapse for haplotypes. Figure 6 (next page) is a screen shot from CodonCode Aligner™ that shows the three windows I used to edit sequence assemblies for each individual animal in my data set. The uppermost window shows a list of sequence assemblies for each individual animal, and the window in the middle shows the forward and reverse nucleotide sequences as well as the relative position of a nucleotide site within the length of a CRI sequence. Finally, the window at the bottom shows an electropherogram of these forward and reverse sequences.

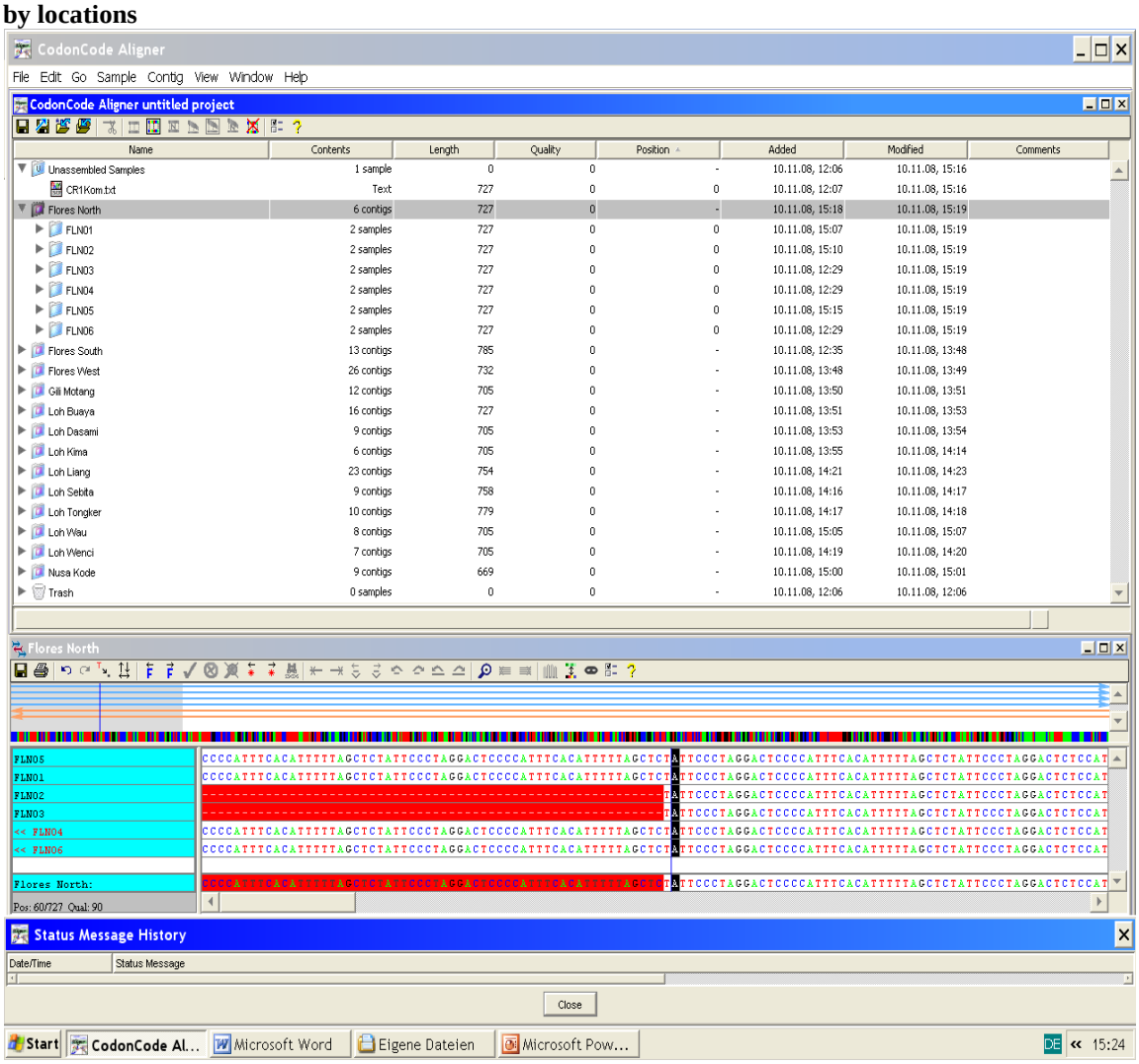
All forward and reverse sequences were assembled to their respective label, which represent an individual animal. I edited each of these sequence assemblies manually by referring to its electropherogram to ascertain a correct base calling. Following this editing process, a single continuous sequence (contig) for every individual animal was obtained. I verified all these contigs using a published Komodo monitor CRI sequence (Kumazawa and Endo, 2004) and determined the possible variations in sequence length.

Figure 6. A screenshot showing sequence editing process using references from electropherograms and published Komodo monitor CRI sequence



Following the sequence editing and verification process, I carried out an analysis to determine haplotypes. Figure 7 (next page) shows two windows I used to define haplotypes. The uppermost window shows a list of contig assemblies in sampling locations, while the middle window shows all contigs from a sampling location. I grouped all contigs of individual animal into their corresponding sampling location. All these contigs in the group were compared to each other and the distinct sequences could be seen immediately in this step. Further, I compared all the distinct sequences from all sampling locations to collapse for haplotypes. Identical sequences were removed and only the unique ones were retained, resulting in eleven different sequences, which represent variations in my entire data set. These eleven sequences were subsequently designated as haplotypes that I put together in an assembly and converted this assembly into a “Nexus” file for further analyses.

Figure 7. A screenshot showing the process of collapsing for haplotypes from sequences assembled by locations



2.5. Phylogenetic analyses

I analysed the relationships among Komodo monitor haplotypes using parsimony method implemented in PAUP* 4.0b10 (Swofford, 2001). This method is based on a conjecture that phylogenetic relationship among organisms is best explained using as few assumptions, e.g. nucleotide changes as possible. Within a set of sequences, the parsimonious sites are those having the least number of changes to construct a tree. These sites are termed as parsimony-informative sites and the most parsimonious tree is the shortest among all possible trees. As a note, a parsimony analysis does not allow for backward mutation and parallel substitution or homoplasy (Nei and Kumar, 2000). I ran my Nexus file in PAUP* for a heuristic tree search using the stepwise tree sampling procedure with random addition sequence, and the Tree-Bisection-Reconnection (TBR) branch swapping algorithm. The Bootstrap support values for 50%-majority-rule

consensus tree were obtained by one thousand (1,000) tree reiterations and shown as a percentage of the same topology sampled. Additionally, the pairwise genetic distances among haplotypes were calculated as *p* distance and a distance corrected using Kimura two-parameter (K2P) model of transitions and transversions. I regressed the two genetic distances to test for possible substitution saturation, which may occur for rapidly evolving region such as the Control Region.

A different approach to reconstruct a phylogeny is the likelihood method, which uses the probability of nucleotide changes, regardless of the number of changes within nucleotide sites. This method is based on an assumption that nucleotide substitutions follow a pattern that can be modelled. Thus, the phylogenetic relationships among organisms are not always best explained by the most parsimonious assumption. Rather, a correct model for nucleotide substitutions is indispensable to infer the correct topology (Nei and Kumar, 2000). In other words, the topology that has the maximum likelihood for a certain model is chosen as the best explanation for the phylogeny in question. To select an appropriate substitution model, I used the computer programme jModelTest (Posada, 2008), a user-friendly interface of ModelTest (Posada and Crandall, 1998). jModelTest uses an algorithm for maximum likelihood method called "Phyml" and the software can suggest one or two among eighty eight (88) nucleotide substitution models using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and the hierarchical Likelihood Ratio Test (hLRT). The programme is free and can be downloaded from the website: http://darwin.uvigo.es (Posada, 2008). I conducted a heuristic search for Maximum Likelihood (ML) trees to reconstruct a consensus ML tree of the Komodo monitor haplotypes and outgroup sequences in PAUP* with one tousand (1,000) Bootstrap replicates to obtain values for node support.

I also used a Bayesian inference of phylogeny implemented in MrBayes version 3.1.2 to find an alternative topology using a different approach. The Bayesian tree was obtained using a Bayesian probability method, which is based on the Bayes' theorem that explains a relation between probabilities of two stochastic events. One event has a probability of presence, i.e. posterior/conditional probability, given the probability of the presence of the other event, i.e. prior/marginal probability (see Appendix G for an expression of the Bayesian Theorem). Thus, the Bayesian inference of phylogeny assumes that the best-fit model of evolution has a greater probability than the other model being compared. There are two types of prior (marginal probability) for the

analysis in Mr.Bayes, i.e. priors for unconstrained parameters such as transition/transversion, branch length, site-specific substitution rates, and fossil age distribution, and also priors for constrained parameters such as stationary nucleotide frequencies and reversible mutation. The priors for unconstrained parameters in MrBayes are the exponential and gamma distributions, whereas the prior for constrained parameters is a uniform Dirichlet distribution. In practice, however, a prior is not always available for analyses and therefore the default prior probability density in MrBayes is a flat Dirichlet distribution (e.g. the value for all bases is 1.0). For the definitions of the three probability distributions, please refer to texts on statistical probablity for example Evans *et al*. (2000). The posterior probability in MrBayes is the sample of trees obtained during the analysis. Some of these trees are assumed to be non-representative and therefore discarded during the analysis as burn-ins (Huelsenbeck and Ronquist, 2001).

2.6. Phylogeography methods

I applied the Analysis of Molecular Variance (AMOVA) to determine the amount of molecular dispersion among populations of the Komodo monitor. The magnitude of the dispersion was computed as variance components for three hierarchies, each of which contributed to the total dispersion across the whole population. The smallest unit of variation is the lowest hierarchy, e.g. sampling location, in which dispersion is measured among individuals. When sampling locations are pooled together within a group, the dispersion is then calculated among these sampling locations. The group of sampling locations is the next hierarchy and is also a larger source of variation. The group of all groups of sampling locations is the highest in the hierarchy, and the dispersion among groups of sampling locations may be expected to make the largest variation across the entire population. The sum of dispersions calculated for each hierarchy makes up the total variation in the data set and is given in percentage (100%). Differentiation indices are also calculated in AMOVA. The differentiation indices or φ-statistics computed in AMOVA are the measures of divergence between a given pair of population units. The total divergence is given by φ_{ST}, which calculates the divergence among all pairs of the smallest units in the whole population, e.g. sampling locations. On the other hand, φ_{SC} and φ_{CT} are the measures of divergence at lower hierarchies, i.e. within a group of sampling locations and among groups, respectively. I generated AMOVA tables and the subsequent differentiation indices using Arlequin version 3.1, a software package for population genetics analysis. In addition, I conducted a correlation

test between genetic and geographic distance matrices using Mantel's test that is also available in Arlequin (Excoffier *et al.*, 2006).

I generated a haplotype network to illustrate the relationships among populations that otherwise may not be observable in a phylogenetic tree. The network is estimated from all sequences in the data set, from which haplotypes can be collapsed by allowing for an alignment gap to be treated as a fifth character or as a missing datum. The probability of parsimony is calculated for pairwise differences among all sequences until the probability exceeds a connection limit that can be set by the user from 90% to 99%. Otherwise, an exact number of differences can also be set to calculate a parsimony probability. A haplotype network is also called cladogram estimation or statistical parsimony and is implemented in TCS version 1.21 (Clement *et al.*, 2000).

2.7. Divergence Time Estimation

The relative age of a given population is indispensable to estimate the extent of its divergence. A method to assess divergence time between two monophyletic clades of a molecular phylogeny is implemented in the package BEAST (Bayesian Evolutionary Analysis by Sampling Trees). Two tasks can be done using this method, i.e. phylogenetic reconstruction and evolutionary hypotheses testing. I applied this programme to reconstruct an alternative genealogy of the Komodo monitor haplotypes and to estimate a divergence time between two given clades in the resulting genealogy. The divergence among haplotypes and the time of separation between a given pair of monophyletic clades in the genealogy does not necessarily reflect a separation between populations. Instead, the genetic divergence among populations can be inferred using phylogeographic parameters, i.e. percent genetic variation and the associated differentiation indices. Nonetheless, a divergence time analysis can provide an estimate of time since the separation between two lineages, i.e. haplotypes, or the relative age of a haplotype group. An estimate of the relative age of a given lineage relies on some information that contain a time frame, such as a fossil record or a geological event that can be used as a calibration point. Because fossil record is not always available for a given species, an estimate of geological time can be a very useful substitute (Drummond and Rambaut, 2007).

I used the fossil teeth of the Komodo monitor discovered on Flores and dated to ~0.90 Ma (Morwood, 2001, van den Bergh *et al.*, 2001) to calibrate one of the nodes in

the resulting genealogy from BEAST. An ample of knowledge on the geological events in the Lesser Sunda, e.g. Pleistocene Glacial Maxima (~0.18 Ma) and a series of island emergence within the region during the Miocene (50-5 Ma) may also be used as calibration points. Additionally, a reconstruction of Pleistocene sea levels in Southeast Asia that provides estimates of historical connection among islands in this region can be useful to infer possible migrations among populations of the Komodo monitor. For example, the islands of Flores and Komodo were estimated to be connected at 75m below the current sea level during the period of Glacial Maxima (Voris, 2000). Because of the emergence of land bridge between Flores and Komodo, a migration between the two islands may then be postulated.

Chapter 3

RESULTS

3.1. The Control Region I (CRI) sequence characteristics

3.1.1. Length variations

Following the sequence editing and assembly, I found many ambiguous sites near the 5' end (near $tRNA^{Thr}$) in the majority of sequences in my data set. A removal of these ambiguous sites resulted in aligned sequences of different length. I confirmed that all Komodo monitor sequences in my data set were the Light-strand (L-strand) sequences of the Komodo monitor CRI by referring to the mtDNA CRI sequence published by Kumazawa and Endo (2004). Three (3) types of sequence varying in length were found in my alignment, i.e. maximum-length sequence (727 bp), medium-length sequence (705 bp), and minimum-length sequence (669 bp). Apparently, the shorter sequence types contain multiple indels (insertions or deletions) at the 5' end, i.e. medium-length sequences have a 22 bp indel and minimum-length sequences have a 58 bp indel. The majority of the Komodo monitor sequences in this data set (~90%) are of the shorter types. Of these shorter sequences, fifty-eight (58) sequences (~17.7%) are minimum-length and two hundred and seventy (270) sequences (~82.33%) are medium in length. Only thirty-eight (38) sequences across all samples (~10%) are maximum-length. The six sequences of the Lace monitor, *V. varius*, were all 652-bp long, and thus 47 bp shorter than the shortest type of the Komodo monitor sequences. In the multiple sequence alignment encompassing all the Komodo monitor and Lace monitor sequences in my data set, I found two types of multiple indels in the Lace monitor sequences. The shorter type is 5-bp long, and the longer type is 71-bp long. Both multiple indels are located near the 5' end. Multiple indels have been found in other vertebrate Control Region sequences, for instance 86-bp multiple indels in the Roe deer, *Capreolus capreolus*, 47-bp multiple indels in the domestic cow, *Bos taurus* (Douzery and Randi, 1997), 74-83-bp multiple indels in the Sturgeon fishes (Ludwig *et al.*, 2000), 49-bp multiple indels in the Morelet's crocodile, *Crocodylus moreletii* (Ray and Densmore, 2003), and 39-bp multiple indels in the Western brook lamprey, *Lampetra richardsoni* (White and Martin, 2009). In my data set, the length of one unit of multiple indels corresponds to the absence of a partial or complete 36-bp repeat motif. This

phenomenon may have been a consequence of a tandem repeat loss that is reported in the Sturgeon fish to be associated with length heteroplasmy (Buroker *et al.*, 1990).

The tandem repeat motif in my data set was 36-bp long (Figures 8a and 8b) and I found a total of four (4) repeat units within the maximum-length CRI sequences, for example FLN01 (Flores North) and FLW01 (Flores West). The first repeat unit (R1) is started within the $tRNA^{Thr}$ for fourteen (14) nucleotide sites (Kumazawa and Endo, 2004), but this part could not be retrieved in my study without ambiguous alignment. Thus, only the remaining 22 bp of R1 could be retained for an unambiguous alignment. The three (3) other repeats, i.e. R2, R3, and R4, continued directly after R1. In most cases, R2 sequences were very similar to R3 sequences, whereas R4 sequences were less similar to both R2 and R3 sequences. At the end of R4 sequences, there were substitutions by Thymine (e.g. nucleotide positions 127, 128, and 130) that made up most of the changes in this repeat unit (Figure 8b).

3.1.2. Nucleotide substitutions and Control Region I (CRI) sequence structure

Variable sites were found to occur more often near the two extremes of Komodo monitor CRI sequences, i.e. near 5' end (up to 135^{th} nucleotide site) and near 3' end (starting from 566^{th} nucleotide site). Between these two points, there was only a single variable nucleotide site, i.e. at 334^{th} base position (Figure 8a). The pattern of variable nucleotide sites distribution near both ends of Control Region sequences has also been described for several vertebrates, e.g. sturgeon (Brown *et al.*, 1993), some mammals (Sbisà *et al.*, 1997), and vultures (Roques *et al.*, 2004). Figure 8a shows an approximation of three domains commonly described for vertebrate Control Region. There were more variable sites in Domain I (ETAS domain) than in either Domain II or Domain III. Eleven (11) variable sites and two sites with single indels were present within Domain I, and four (4) variable sites were observed in Domain II (conserved domain), and two (2) other variable sites were found in Domain III (CSB domain). In total, there were seventeen (17) variable sites across CRI sequences of Komodo monitor in this data set. Additionally, the characteristic boxes in vertebrate Control Region, i.e. ETAS1, CSB1, and CSB2 were found conserved, as previously identified in both Komodo monitor Control Region I (CRI) and Control Region II (CRII) sequences (Kumazawa and Endo, 2004).

The variable sites in the the Komodo monitor CRI sequences occur within the repeat region, i.e. positions 1-130 and also beyond the repeat region, i.e. positions 131-727 (Figure 8b). Nine (9) variable sites and two (2) single indel sites were situated within the tandem repeat region, whereas six (6) other variable sites were found beyond this region (see details in Figure 8a). The positions of these variable sites within each repeat unit seem to be non-random, with one or two variable sites occurring at similar positions. By counting nucleotide sites from the first position in each repeat unit, I found two (2) of the four variable sites in R1 were located at positions 17 and 25. The two (2) variable sites within R2 and R3 were found at positions 16 and 24, whereas only one variable site was found in R4, at position 16. The number of variable sites in R1 still remains to be determined awaiting available flanking $tRNA^{Thr}$ sequence. However, I was able tp observe the total number of variable sites for R1, R2, and R3. A total of four (4) variable sites and one single indel site were present in R2, two (2) variable sites and one indel site were present in R3, and a total of two (2) variable sites were present in R4. Figure 8b shows the positions of all variable sites and indel sites within the eleven representative sequences of the Komodo monitor. The non-random positions of variable sites in each of the four repeat units are also indicated in this figure.

Only eleven (11) out of the total seventeen (17) variable sites found in the data set were used for the overall genetic distance analysis, due to an adjustment for sequence length and the presence of indels. The four (4) variable sites within R2 fell within the 58-bp multiple-indel site and all these variable sites were consequently ignored, because all sites parallel to this multiple-indel site were discarded, in order to adjust for a comparable sequence length. In other words, all sequences were adjusted to 669 bp (the minimum length), resulting in the removal of four variable sites. The presence of two single indels at positions 49 and 86, each of which corresponds to a position in R2 and R3, respectively, has reduced this sequence length further to 668 bp for overall sequence comparison. Hence, the number of variable sites used to calculate genetic distance depends on the length of the sequences being compared. For example, sequence FLW01 (727 bp) and sequence LBU0042 (704 bp) are aligned. Twenty-three (23) sites will be ignored to calculate genetic distance between these sequences, due to the presence of 22-bp multiple indel and one single indel at position 86 (Figure 8a). Nevertheless, if two medium-length sequences were aligned, I included the variable sites in R2 to calculate genetic distance, because these sites represent a level of variability and cannot simply be

discarded. Besides nucleotide substitutions, a variation in sequence length also shows an important level of variability. A variation in sequence length is a class of mutation that becomes increasingly an important molecular character in phylogenetics (Müller, 2006). Because a variation in sequence length seem to result from the loss or gain of one multiple-indel or more, it is important to find a method to properly code this array of gaps in an alignment. Multiple indels have been thought to have originated as a single indel, and accordingly, they can be regarded as a single mutation event (Simmons and Ochoterena, 2000). Thus, single and multiple indels can be regarded as mutations of a different class other than that of nucleotide substitutions, both of which are two important molecular variations in phylogenetic analysis (Kawakita *et al.*, 2003).

As a summary of the results in this section, some statistics of the sequence characteristics are presented as follows: 1) The number of polymorphic sites in the current data set ranged from 11-17; 2) All substitutions across all comparable nucleotide sites were transitions; 3) The base proportion in the alignment was A= 24.21%, T= 34.43%, G= 14.83%, and C= 26.53%, which shows a bias toward AT changes (58.64%) rather than GC changes (41.36%) and a deficiency in Guanine; 4) The overall mean number of pairwise differences among sequences after a final sequence length adjustment to 668 bp was 1.199 ± 0.769; and 5) The mean nucleotide diversity per site was 0.0018 ± 0.0013.

Figure 8a. A schematic illustration of the Komodo monitor mtDNA Control Region I (L-strand) sequence structure and the locations of variable nucleotide sites. Dashed vertical lines show characteristic segments and spacers. Grey area shows a part of $tRNA^{Thr}$, where the first tandem repeat unit starts. Numbers below the diagram show sequence lengths of characteristic segments (underlined) and spacers in base pairs (bp), and bold prints show characteristic boxes of the vertebrate Control Region.

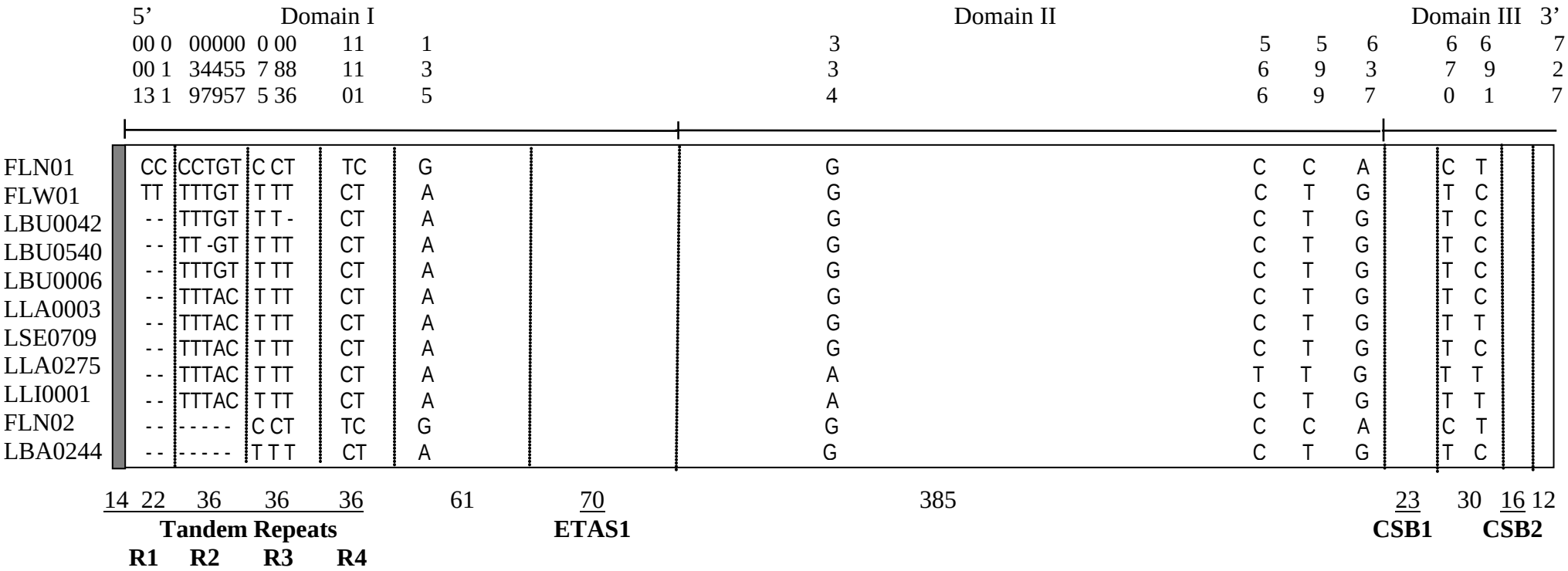
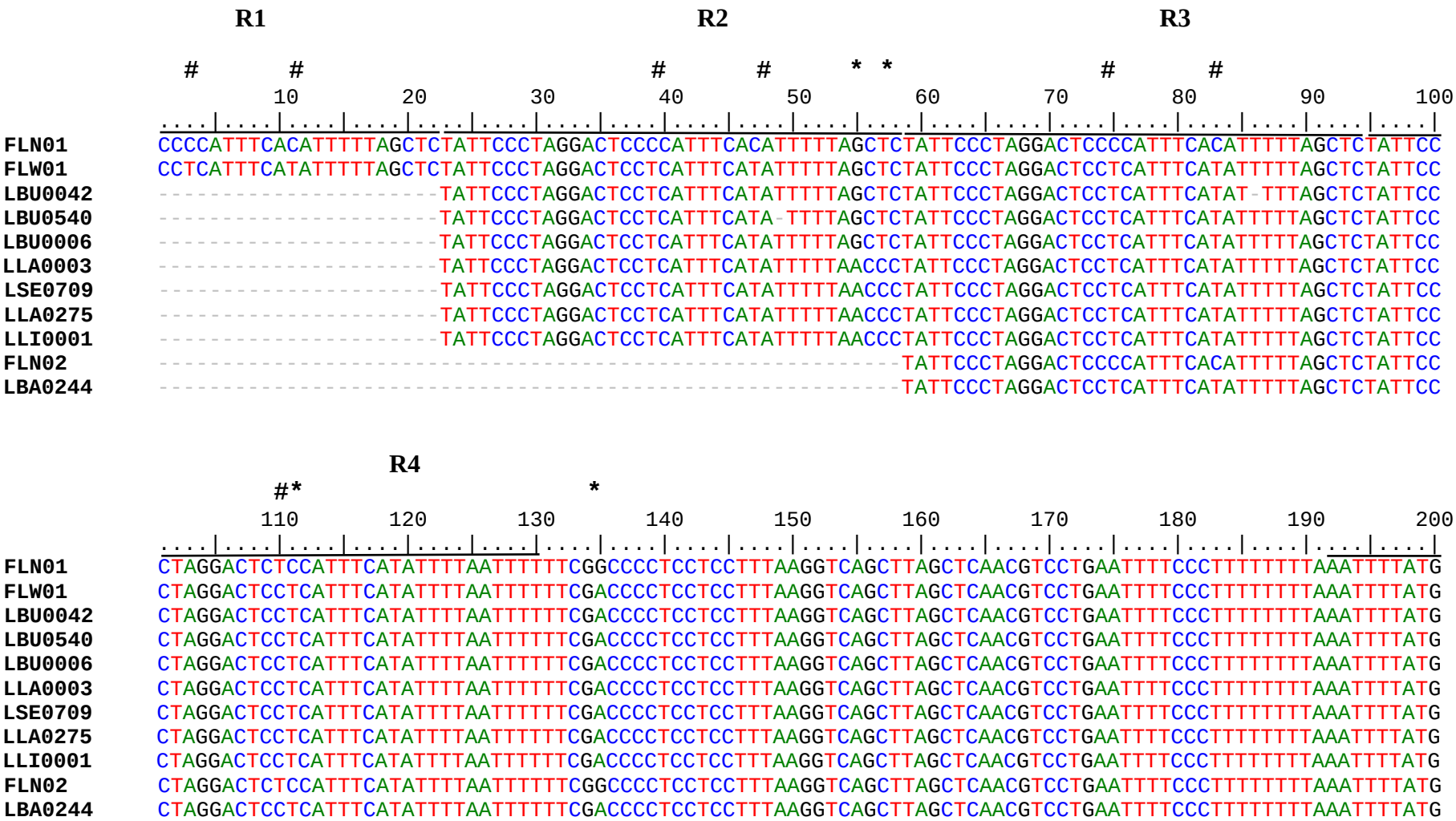
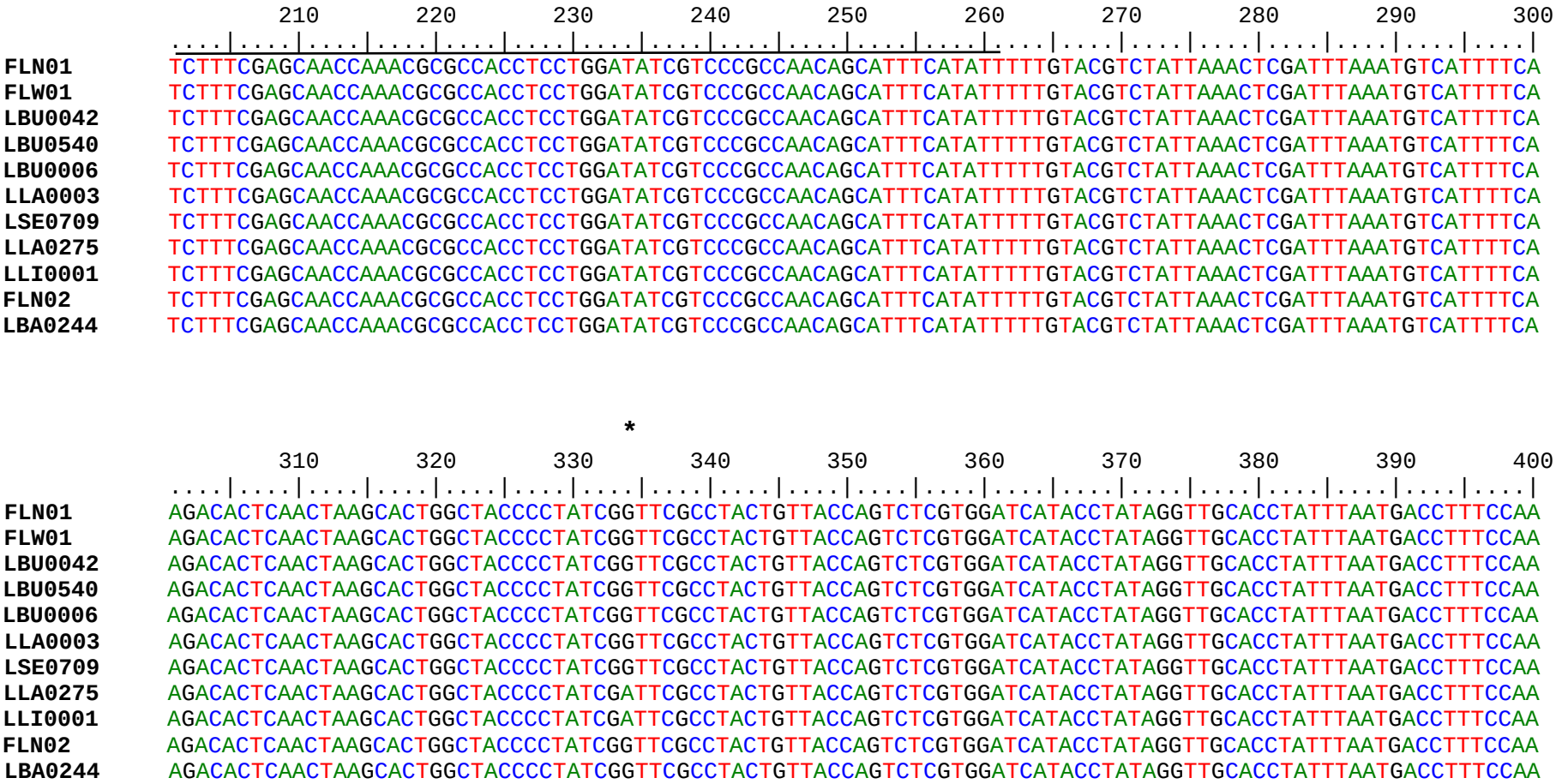


Figure 8b. An alignment of eleven representative sequences of the Komodo monitor. Variable sites are indicated by # for those occurring at similar positions in each repeat unit and other variable sites are indicated by *. Characteristic sequence boxes, i.e. ETAS and CSB are determined as in the published CRI of the Komodo monitor (Kumazawa and Endo, 2004).

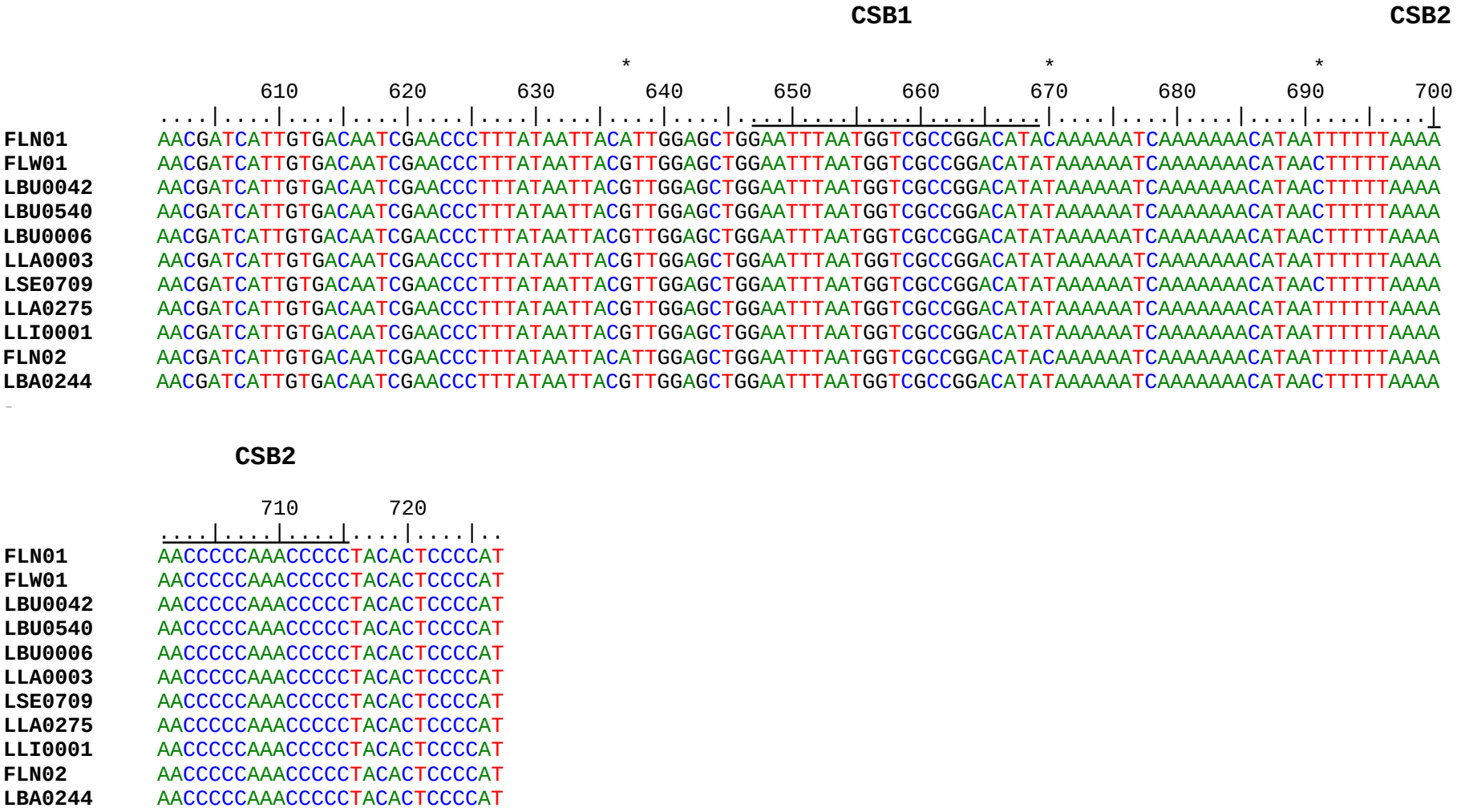


ETAS1


210 220 230 240 250 260 270 280 290 300
FLN01 TCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTCA
FLW01 TCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTCA
LBU0042 TCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTCA
LBU0540 TCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTCA
LBU0006 TCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTCA
LLA0003 TCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTCA
LSE0709 TCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTCA
LLA0275 TCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTCA
LLI0001 TCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTCA
FLN02 TCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTCA
LBA0244 TCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTCA
*
310 320 330 340 350 360 370 380 390 400
FLN01 AGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCAA
FLW01 AGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCAA
LBU0042 AGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCAA
LBU0540 AGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCAA
LBU0006 AGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCAA
LLA0003 AGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCAA
LSE0709 AGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCAA
LLA0275 AGACACTCAACTAAGCACTGGCTACCCCTATCGATTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCAA
LLI0001 AGACACTCAACTAAGCACTGGCTACCCCTATCGATTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCAA
FLN02 AGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCAA
LBA0244 AGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCAA

```
                410       420       430       440       450       460       470       480       490       500
           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
FLN01      TACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTCA
FLW01      TACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTCA
LBU0042    TACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTCA
LBU0540    TACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTCA
LBU0006    TACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTCA
LLA0003    TACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTCA
LSE0709    TACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTCA
LLA0275    TACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTCA
LLI0001    TACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTCA
FLN02      TACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTCA
LBA0244    TACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTCA

                                                          *                                       *
                510       520       530       540       550       560       570       580       590       600
           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
FLN01      TAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACCA
FLW01      TAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACTA
LBU0042    TAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACTA
LBU0540    TAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACTA
LBU0006    TAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACTA
LLA0003    TAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACTA
LSE0709    TAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACTA
LLA0275    TAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATTGCATGTTCGTCCGGGGTCAGCTTTACATTTACTA
LLI0001    TAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACTA
FLN02      TAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACCA
LBA0244    TAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACTA
```


CSB1
CSB2
610
620
630
640
650
660
670
680
690
700
FLN01 AACGATCATTGTGACAATCGAACCCTTTATAATTACATTGGAGCTGGAATTTAATGGTCGCCGGACATACAAAAAATCAAAAAAAACATAATTTTTTAAAA
FLW01 AACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAAACATAACTTTTTAAAA
LBU0042 AACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAAACATAACTTTTTAAAA
LBU0540 AACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAAACATAACTTTTTAAAA
LBU0006 AACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAAACATAACTTTTTAAAA
LLA0003 AACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAAACATAATTTTTTAAAA
LSE0709 AACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAAACATAACTTTTTAAAA
LLA0275 AACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAAACATAATTTTTTAAAA
LLI0001 AACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAAACATAATTTTTTAAAA
FLN02 AACGATCATTGTGACAATCGAACCCTTTATAATTACATTGGAGCTGGAATTTAATGGTCGCCGGACATACAAAAAATCAAAAAAAACATAATTTTTTAAAA
LBA0244 AACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAAACATAACTTTTTAAAA
CSB2
710
720
FLN01 AACCCCCAAACCCCCTACACTCCCCAT
FLW01 AACCCCCAAACCCCCTACACTCCCCAT
LBU0042 AACCCCCAAACCCCCTACACTCCCCAT
LBU0540 AACCCCCAAACCCCCTACACTCCCCAT
LBU0006 AACCCCCAAACCCCCTACACTCCCCAT
LLA0003 AACCCCCAAACCCCCTACACTCCCCAT
LSE0709 AACCCCCAAACCCCCTACACTCCCCAT
LLA0275 AACCCCCAAACCCCCTACACTCCCCAT
LLI0001 AACCCCCAAACCCCCTACACTCCCCAT
FLN02 AACCCCCAAACCCCCTACACTCCCCAT
LBA0244 AACCCCCAAACCCCCTACACTCCCCAT

3.1.3. Comparative sequence characteristics

The presence of multiple indels in the Komodo monitor CRI sequences is probably a unique event. In this study, I observed a different pattern of mutiple indels in the Lace monitor sequences. Indels were absent within the alignment that included exclusively the Lace monitor sequences; however, multiple and single indels were present when they were aligned with representative sequences (haplotypes) of the Komodo monitor. In total, there were ten (10) indel events in the alignment that included eleven (11) Komodo monitor haplotypes and six (6) Lace monitor sequences. Six (6) of these events were multiple indels and four (4) events were single indels. Four (4) multiple indels were present near the 5' end and two (2) others were present near the 3' end of the aligned sequences. The size of the multiple indels and the position of gaps denoting the absence of nucleotides in the alignment for each species were different. A 5-bp multiple gaps within the Lace monitor sequences overlapped with both the 22-bp and 58-bp multiple gaps present in the Komodo monitor sequences, whereas a 70-bp multiple gaps in the Lace monitor sequences overlapped with one end of the 58-bp multiple gaps present in the Komodo monitor sequences. All these multiple indels were observed near the 5' end. Two (2) shorter multiple indels were found near the 3' end, i.e. 3-bp multiple indels at positions 698-700 and 2-bp multiple indels at positions 724-725. Further, two (2) single indels were found in Domain I, i.e. at positions 49 and 86, and two other single indels were observed in Domain III, i.e. at positions 163 and 687 (see Appendix I for multiple sequence alignment for the Komodo monitor haplotypes and six Lace monitor sequences).

A total of fifty-two (52) variable sites were observed in the multiple sequence alignment for the Komodo monitor and the Lace monitor, which comprised seventeen (17) sequences with 731 nucleotide sites. Both transitional and transversional changes were observed among interspecific pairwise alignments, whereas only transitional changes were observed among intraspecific pairwise alignments. Additionally, mean pairwise transition/ transversion (ti/tv) across the whole data set was relatively small, with about two transitions occured for every transversional change (1.980 ± 0.184).

3.2. Haplotype frequency distribution

I defined haplotypes for the Komodo monitor based on sequence length, single indel position, and nucleotide substitutions. Eleven (11) haplotypes were resulted from a total of

three hundred and thirty six (366) CRI sequences of the Komodo monitor examined during the study. All the eleven haplotypes were consistent with those listed in the haplotype database that I generated using the population genetics software Arlequin version 3.1. Four (4) sequences sampled from Komodo, four (4) from Rinca, and three (3) from Flores were found distinct among all sequences within the current data set. These distinct sequences were designated accordingly as haplotypes. As a note, the eleven representative sequences presented in Figure 8b above are essentially these distinct sequences (haplotypes). Further, I coded these distinct sequences as haplotypes H1-H11 shown in Table 2 below.

Table 2. Haplotype frequency distribution across islands. The number of haplotypes found on each island is indicated within brackets.

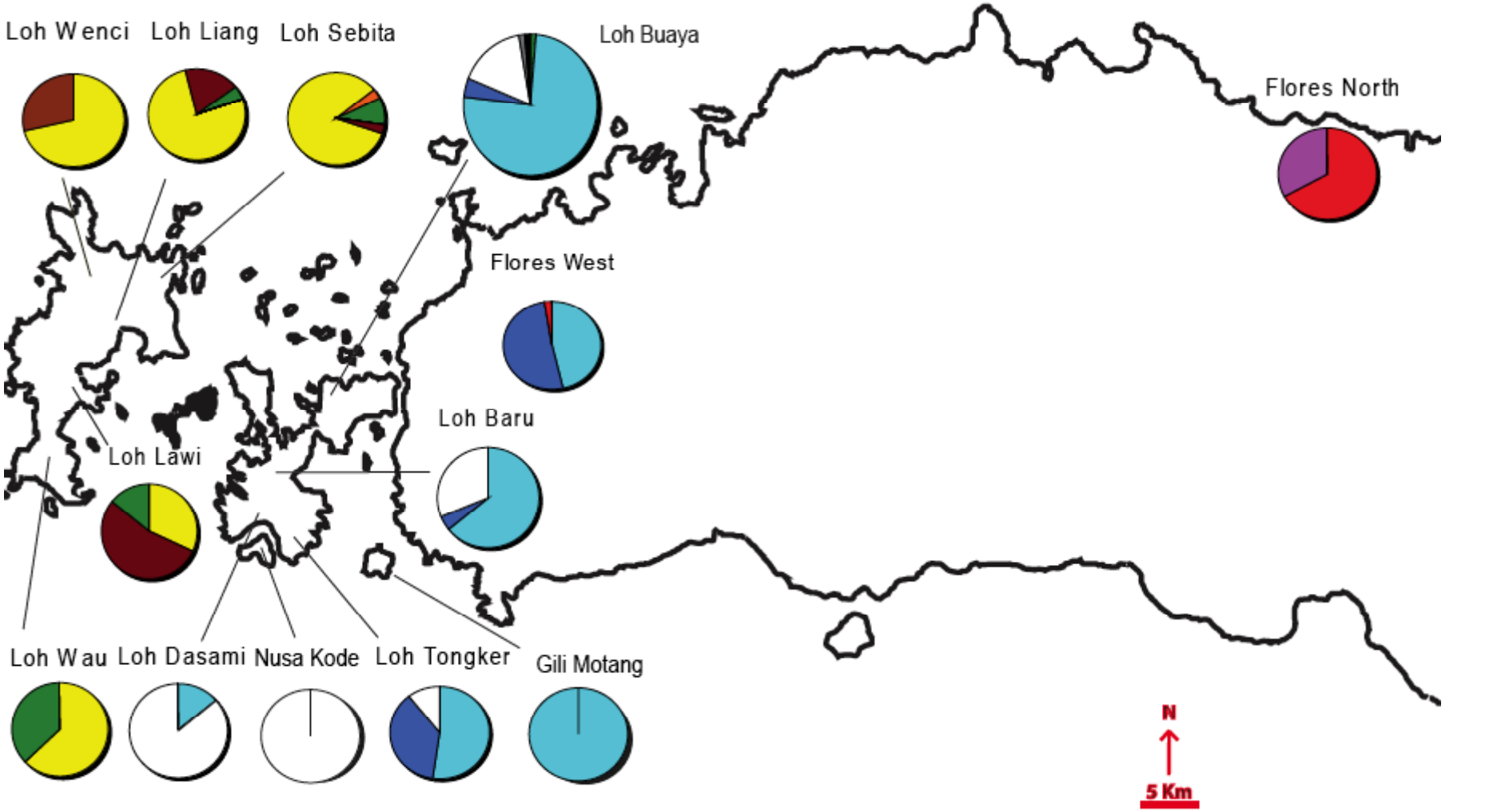
Haplotype	Reference Sequence	Length (bp)	Island					Total individuals per haplotype
			Komodo (4)	Rinca (6)	Flores (4)	Gili Motang (1)	Nusa Kode (1)	
H1	LLI0001	705	78	0	0	0	0	78
H2	LLA0003	705	26	0	0	0	0	26
H3	LLA0275	705	12	1	0	0	0	13
H4	LSE0709	705	1	0	0	0	0	1
H5	LBU0006	705	0	105	18	27	0	150
H6	FLW01	727	0	13	20	0	0	33
H7	LBA0244	669	0	44	0	0	12	56
H8	LBU0042	704	0	1	0	0	0	1
H9	LBU0540	704	0	1	0	0	0	1
H10	FLN01	727	0	0	5	0	0	5
H11	FLN02	669	0	0	2	0	0	2
Total individuals per island			117	165	45	27	12	366

Haplotype H5 was the most common haplotype across the current data set, with a total of one hundred and fifty (150) individuals sampled from four islands. On the other hand, haplotype H1 was the most common in Komodo and the second most frequently sampled in the whole data set. Further, four (4) rare haplotypes were discovered in three separate locations on three different islands of the larger size. These rare haplotypes were haplotypes H4, H8, H9, and H11. Haplotype H4 was found on Komodo, haplotypes H8 and H9 were found on Rinca, and haplotype H11 was found on Flores. A single individual animal represented each of these haplotypes, except haplotype H11, which was sampled from two individuals.

Overall, the larger islands harbour a larger number of haplotypes than smaller islands do. Among all the four (4) haplotypes from Komodo, three (3) were exclusively found on this

island. On the other hand, haplotype H3 was shared with one location on Rinca, i.e. Loh Buaya. Among a total of six (6) haplotypes discovered on Rinca, four (4) haplotypes were shared with another island or more. Only one haplotype was found on each of the small islands of Gili Motang and Nusa Kode. All these haplotypes found on the small islands can also be found on the neighbouring larger islands. Intuitively, I grouped haplotypes H1, H2, H3, and H4 as “Western” haplotypes, and haplotypes H5, H6, H7, H8, and H9 as “Central” haplotypes. Further, I set haplotypes H10 and H11 in “Eastern” haplotype group (Figure 9).

Figure 9. The haplotype frequency distribution shown being overlaid upon the geographic locations across the extant distribution range for the Komodo monitor. The pie chart for Loh Buaya is enlarged, in order to show all haplotype contents in this location.



3.3. Genetic divergence

3.3.1. Genetic distance

I evaluated the genetic distances among pairs of haplotypes in terms of *p* distance. *p* distance is a measure of an uncorrected genetic distance or sequence divergence, which can be calculated from the number of nucleotide changes per site between two sequences and is given in the equation below:

$$p = n_d / n$$

where n_d is the number of nucleotide differences between two sequences and *n* is the total number of nucleotide sites examined. Because there were variations in length in the entire alignment, the total number of nucleotide sites differed from one pairwise alignment to another. However, this does not seem to make a great difference in the resulting *p* distance values, because the number of differences for a pair of sequences is proportional to their length. In other words, longer sequences tend to have more variable sites, but they also have more nucleotide sites (*n*) to divide. Table 3 shows *p* distances (uncorrected distances) and the total character differences between pairs of the Komodo monitor haplotypes. As previously mentioned in section 3.1.3, no transversional mutation was found in the intraspecific sequence alignment for both the Komodo monitor and the Lace monitor. On the other hand, transversions were observed among pairwise interspecific alignments of the Komodo monitor and the Lace monitor sequences.

Because one type of substitution may occur much more frequently than the other to a point of substitution saturation, models of nucleotide substitutions were commonly used to correct distance measures. The Jukes-Cantor model has one mutation rate, which means that transitions and tranversions arise at an equal frequency. On the other hand, the Kimura two-parameter (K2P) model differentiates the rate of transitional changes (α) and transversional changes (2β) per site per year. This model can be applied to sequences with any initial nucleotide frequencies (Nei and Kumar, 2000). I observed a linear relationship in the plot of *p* distances against Jukes-Cantor corrected distances for pairs of the Komodo monitor and the Lace monitor sequences. No large differences were found between *p* distances (uncorrected) and K2P distances (model-corrected). The maximum difference between *p* distances and K2P distances for pairs of the Komodo monitor haplotype and the Lace monitor sequence is 0.0078 (data not shown). Similarly, a linear relationship was shown in the plot for *p* distances

and Jukes-Cantor distances as well as in the plot for *p* distances and K2P distances for the Komodo monitor haplotypes. Therefore, substitution saturation was unlikely to occur among all pairs of sequences in this data set. This suggests that there were observable divergences among paired sequences, despite the seemingly lower level. Additionally, a considerable level of interspecific sequence divergence among pairs of the Komodo monitor haplotypes and the Lace monitor sequences were observed. The mean pairwise *p* distance for the Komodo monitor haplotype and a sequence of the Lace monitor was 8.06%, and the mean pairwise K2P-corrected distance was 8.60%.

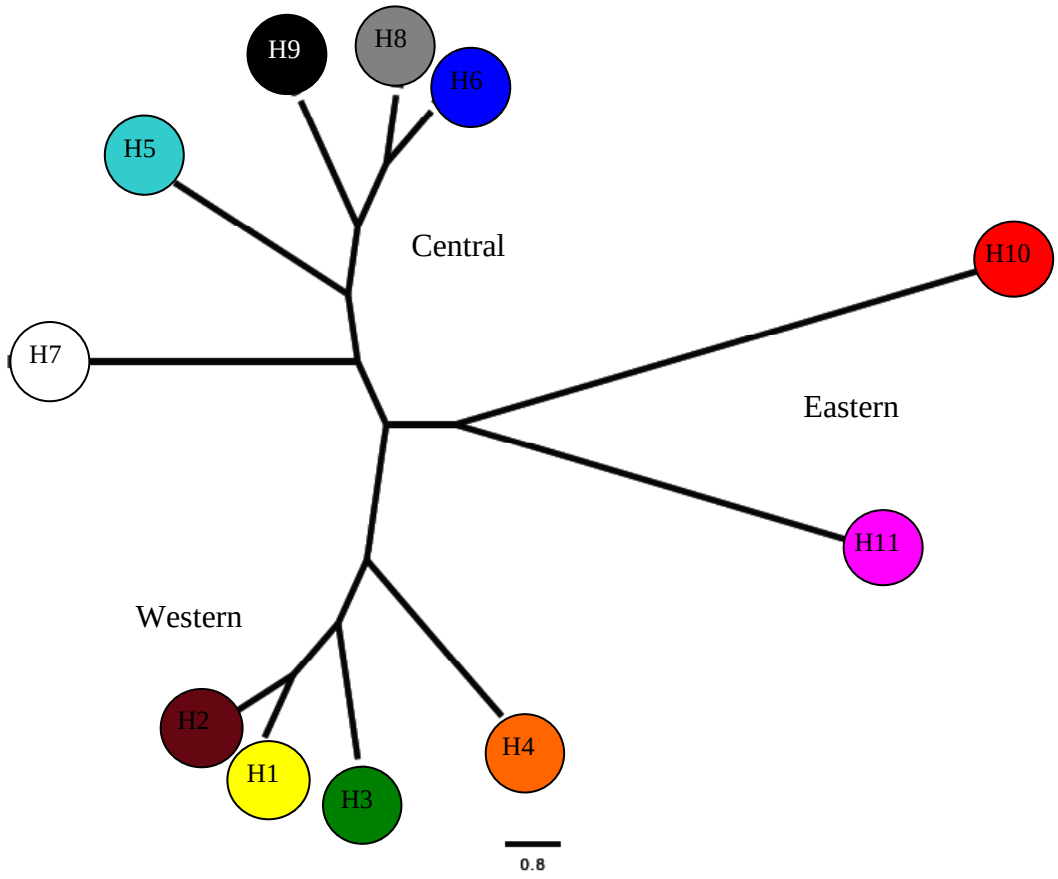
Among all the Komodo monitor haplotypes, the mean pairwise *p* distance was 0.68%, which is comparable to mean pairwise *p* distances for other reptilian mtDNA sequences, for instance 0.58% in the Leatherback turtle (*Dermochelys coriacea*) Control Region sequences (Dutton *et al.*, 1999) and 0.64% in the Olive Sea snake (*Aipysurus laevis*) ND4 sequences (Lukoschek *et al.*, 2007). The range of pairwise sequence divergence among haplotypes of the Komodo monitor is reflected as *p* distances that span from 0-2%, which is comparable to the range of mtDNA Control Region sequence divergence in some turtles of the genus *Glyptemys* (Rosenbaum *et al.*, 2007, Amato *et al.*, 2008). Further, the greatest sequence divergence in terms of the number of nucleotide difference among haplotypes was found between haplotypes H3 (Komodo) and H10 (Flores North), with fourteen (14) nucleotide changes were present along the 705-bp aligned sequences (Matrix 1). No mutation was found among the Central haplotypes, i.e. haplotypes H5-H9 as well among the Eastern haplotypes, i.e. haplotypes H10-H11. A maximum of three (3) mutations were observed among pairs of the Western haplotypes. The fact that there was no observed nucleotide change among the Central haplotypes as well as among the Eastern haplotypes clearly indicates that those haplotypes set in the Central and Eastern groups differed to one another only in their sequence length. It is also interesting to observe, that the Western haplotype H4 and the Central haplotype H7 differed also only by sequence length, i.e. by one repeat unit of 36 bp. Other than the presence of multiple indels, there was no variable site observed between the haplotypes H4 and H7 represented by the sequences LSE0709 and LBA0244 (Figure 8a).

Matrix 1. Pairwise sequence differences among the Komodo monitor CRI haplotypes. Below diagonal are *p* distances, and above diagonal are total nucleotide differences between haplotype pairs. Sequence length of each haplotype is presented in brackets.

	H1 (705)	H2 (705)	H3 (705)	H4 (705)	H5 (705)	H6 (727)	H7 (669)	H8 (705)	H9 (705)	H10 (727)	H11 (669)
H1 (705)	-	1	1	2	4	4	2	4	4	13	9
H2 (705)	0.0014	-	2	1	3	3	1	3	3	12	8
H3 (705)	0.0014	0.0028	-	3	5	5	3	5	5	14	10
H4 (705)	0.0028	0.0014	0.0043	-	2	2	0	2	2	13	9
H5 (705)	0.0057	0.0043	0.0071	0.0028	-	0	0	0	0	11	9
H6 (727)	0.0056	0.0042	0.0071	0.0028	0.0000	-	0	0	0	13	9
H7 (669)	0.0030	0.0015	0.0045	0.0000	0.0000	0.0000	-	0	0	9	9
H8 (705)	0.0057	0.0043	0.0071	0.0028	0.0000	0.0000	0.0000	-	0	11	9
H9 (705)	0.0057	0.0043	0.0071	0.0028	0.0000	0.0000	0.0000	0.0000	-	11	9
H10 (727)	0.0185	0.0171	0.0199	0.0185	0.0157	0.0179	0.0136	0.0157	0.0157	-	0
H11 (669)	0.0135	0.0120	0.0150	0.0135	0.0135	0.0136	0.0135	0.0135	0.0135	0.000	-

I generated a Neighbour Joining (NJ) tree based on the previously calculated *p* distances among pairs of haplotypes to estimate a clustering pattern. Figure 10 shows the resulting tree with three clusters that correspond to the three haplotype-groups determined earlier intuitively based on their relative geographic position, i.e. Western, Central, and Eastern groups. The haplotypes H1-H4, which can be found mostly on Komodo, were nested in the Western cluster. On the other hand, haplotypes H5-H9 distributed on Rinca, Flores, Gili Motang, and Nusa Kode were clustered in the Central group. The Eastern cluster comprised haplotypes H10 and H11 that were distributed exclusively on Flores.

Figure 10. A Neighbour Joining (NJ) tree of eleven haplotypes shows three clusters that correspond to three geographic groups: Western, Central, and Eastern.



3.3.2. Diversity indices: gene diversity and nucleotide diversity

Having known that some haplotypes differed only in their sequence length, I computed Nei's gene diversity (*h*) to assess the probability that any two randomly chosen sequences drawn from a sample were different. The calculation of *h* is based on haplotype frequencies (Nei, 1973), and therefore it may disregard nucleotide substitutions as the

characters to determine haplotypes. On the other hand, there were many sequences in this data set that varied solely due to nucleotide changes, for instance those within the Western haplotype group. Therefore, I also computed nucleotide diversity (π_n) for all pairs of sequences across sampling sites to evaluate the variations in nucleotide substitutions. Nucleotide diversity is a molecular diversity index that is used to quantify the average number of nucleotide differences per site among pairs of DNA sequences. I performed all computations of diversity indices in Arlequin version 3.1. Table 3 shows the estimates for *h* and π_n, with their values rounded to the nearest 0.0001. Due to the small values, π_n were presented in percentage.

Table 3. Haplotype diversity and nucleotide diversity across sampling locations. Sample size for each location is indicated in brackets. The proportion of variable sites is given as the number of variable nucleotide sites in aligned sequences after a length adjustment.

Island	Sampling site	No. of haplotypes	*h* ± SD	Proportion of variable sites	% π_n ± SD
Komodo	Loh Wenci/ LWE (7)	2	0.4762 ± 0.1713	1/705	0.07 ± 0.0007
	Loh Sebita/ LSE (31)	4	0.2946 ± 0.1020	3/705	0.05 ± 0.0006
	Loh Liang/ LLI (43)	3	0.3832 ± 0.0796	2/705	0.06 ± 0.0006
	Loh Lawi/ LLA (28)	3	0.6111 ± 0.0596	2/705	0.11 ± 0.0009
	Loh Wau/ LWA (8)	2	0.5357 ± 0.1232	1/705	0.08 ± 0.0008
	Komodo pooled	4	0.4999 ± 0.0432	3/705	0.08 ± 0.0007
Rinca	Loh Buaya/ LBU (74)	6	0.4043 ± 0.0645	3/669	0.01 ± 0.0003
	Loh Baru/ LBA (58)	3	0.5027 ± 0.0493	0/669	0
	Loh Dasami/ LDS (14)	2	0.2637 ± 0.1360	0/669	0
	Loh Tongker/ LTK (19)	3	0.6082 ± 0.0698	0/669	0
	Rinca pooled	6	0.5208 ± 0.0330	3/669	0.005 ± 0.0002
Flores	Flores West/ FLW (39)	3	0.5371 ± 0.0306	11/705	0.08 ± 0.0007
	Flores North/ FLN (6)	2	0.5333 ± 0.1721	0/669	0
	Flores pooled	4	0.6424 ± 0.0392	9/669	0.40 ± 0.0023
Nusa Kode	Nusa Kode/ NSK (12)	1	0	0/669	0
Gili Motang	Gili Motang/ GMO (27)	1	0	0/705	0

Haplotype diversity (*h*) may better reflect the extent of genetic diversity in my data set, because it allows a measurement that does not base solely on nucleotide changes as a source of DNA polymorphism. On the other hand, nucleotide diversity (π_n) is a useful measure of genetic diversity, when only nucleotide substitution is available as a source of molecular variation. Indeed, indels are generally ignored in phylogenetic analyses due to the uncertainty to include this type of mutation as a molecular character, as well as the limited availability of phylogenetic methods that allows alignment gaps to be incorporated in the analysis (Kawakita *et al.*, 2003). Nevertheless, inclusion of indels as molecular characters may enhance the amount of genetic variability for characterising a population genetic structure (Pearce, 2006). Therefore, the use of multiple indels seems yet to be evaluated in more details in a separate study before its applications, including in phylogeographic studies. By taking both diversity measures into perspective, one can evaluate the extent of molecular diversity in a broader sense, in which length variations generated by multiple indels may be seen as an indication of the relatively low frequency in nucleotide substitutions. However, for the purpose of phylogenetic analyses in this study, the variable nucleotide sites occurring within two tandem-repeats, i.e. R1 and R2 were removed, due to an adjustment of sequence length to the minimum (669 bp). Additionally, there was also an uncertainty in coding the variable sequence length.

3.4. Phylogenetic analyses

I applied three phylogenetic methods to determine genealogical relationships among the eleven Komodo monitor haplotypes. There was a general branching pattern, although three different consensus tree topologies were found. All the three topologies showed monophyly for the Komodo monitor and two separate subclades corresponding to the Eastern and Western-Central haplotype groups were nested within this monophyletic clade for the Komodo monitors In particular, the Western-Central clade was a cluster of all Komodo monitor haplotypes found in this study, excepting the Eastern haplotypes. The Bootstrap supports for the Maximum Parsimony (MP) and Maximum Likelihood (ML) trees were variable. Some of these values appeared to be low (< 70%), whereas a total support was given for the monophyly of the Komodo monitor clade in both trees. Similarly, the posterior probabilities for the Bayesian tree were generally low. The six Lace monitor sequences used as outgroups apparently represented two different clades with three specimens in each clade.

Additionally, the MP tree topology showed a different resolution for the relationships among Lace monitors from the other two topologies of the ML and Bayesian consensus trees.

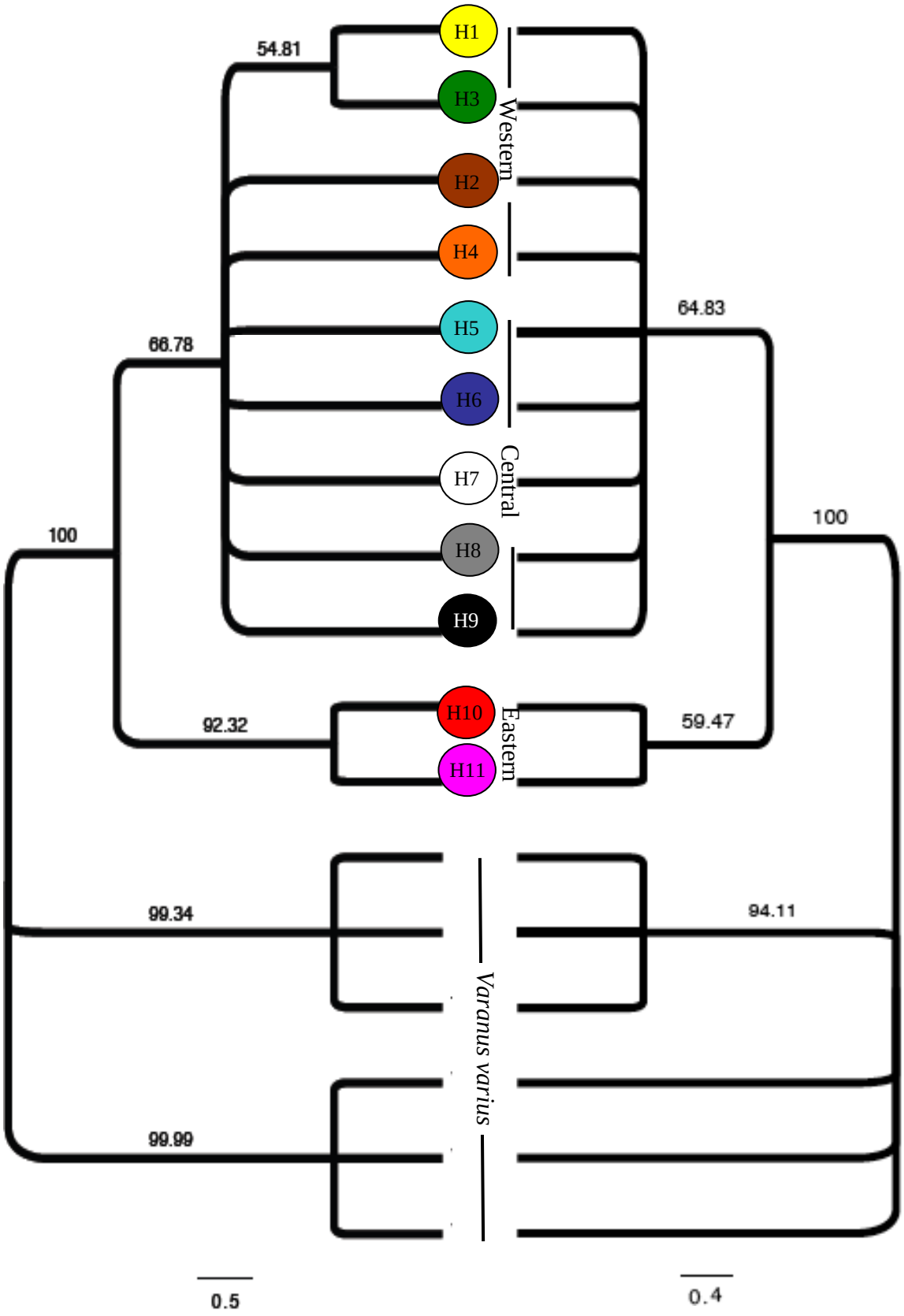
3.4.1. Maximum Parsimony and Maximum Likelihood trees

Out of seven hundred and thirty one (731) alignment sites, seventy-two (72) sites were parsimony informative and three (3) sites were parsimony uninformative, leaving almost 90% of the characters constant. All nucleotides were weighted equal, due to the small transition/ transversion ratio. A heuristic tree search was conducted with the starting trees obtained by stepwise addition. To find all the alternative parsimony trees, I chose a random addition-sequence and a Tree-Bisection-Reconnection (TBR) tree swapping methods. Seven hundred and forty eight (748) trees were resulted in one tree island and one thousand Bootstrap iterations were performed to generate a consensus tree. Figure 11 shows the consensus topology from all MP trees with two Komodo monitor subclades. The Eastern subclade comprised haplotypes H10 and H11 as previously grouped in the NJ tree, and the Western-Central subclade consisted of all haplotypes from the Western and Central groups. There was a higher Bootstrap value for the Eastern subclade than that of the Western-Central subclade. Interestingly, two haplotypes from Komodo Island were clustered together, albeit with a low Bootstrap value, and this cluster was nested in the Western-Central subclade.

A topology similar to that of the MP consensus tree was found in the ML consensus tree. The node supports for this tree seemed to be also relatively low. I generated the ML trees using PAUP* with a nucleotide substitution model of Hasegawa-Kishino-Yano (HKY 1985). This model describes six possible nucleotide substitutions, of which two are of transitional types, i.e. A⇔G and C⇔T, and four changes are of transversional types, i.e. A⇔C, A⇔T, G ⇔C and G⇔T. All the base frequencies in this model are assumed to be unequal (Hasegawa *et al.*, 1985). The HKY model was statistically selected based on the AIC and BIC in the computer programme jModelTest, and the evolutionary rate was estimated to follow a Gamma distribution, i.e. HKY+G using the same programme. The Likelihood scores were computed using Phyml (Guindon and Gascuel, 2003) also in jModelTest with default settings. All the eighty-eight (88) models were fitted to the data with 100% confidence interval. The maximum likelihood estimates for this model was expressed as negative natural-log likelihoods (–lnL= 2031.014) and there were thirty-seven free parameters (K=37) for this model. Further, the HKY model was chosen with AIC=4136.028 and BIC=4306.021,

which gave an estimate of the amount of information being lost, when the nucleotide substitution in the data set was compared to the HKY+G model. The differences between the given values of AIC or BIC with their minimum (i.e. ΔAIC and ΔBIC, respectively) were zero, indicating HKY+G as the best fit to describe the evolutionary pattern of nucleotide sequences in my data set. A Gamma shape parameter of less than 1 (α= 0.115) indicates an L-shaped distribution, which means that most nucleotide sites were invariable and some others changed at a higher rate. The base frequencies were estimated at A= 0.237, C= 0.272, G= 0.147, T= 0.344, and the rate categories were set at four (4), because four different nucleotides were assumed to change at four different rates. The transition/ transversion ratio estimated using PAUP* was 50.938 and the estimate of Kappa= was 93.761. The number of distinct data patterns, i.e. all possible ML trees under this model was 62.

Figure 11. Cladogram of the Komodo monitor haplotypes with six Lace monitor sequences as outgroups. The topologies of consensus MP and ML trees are similar.

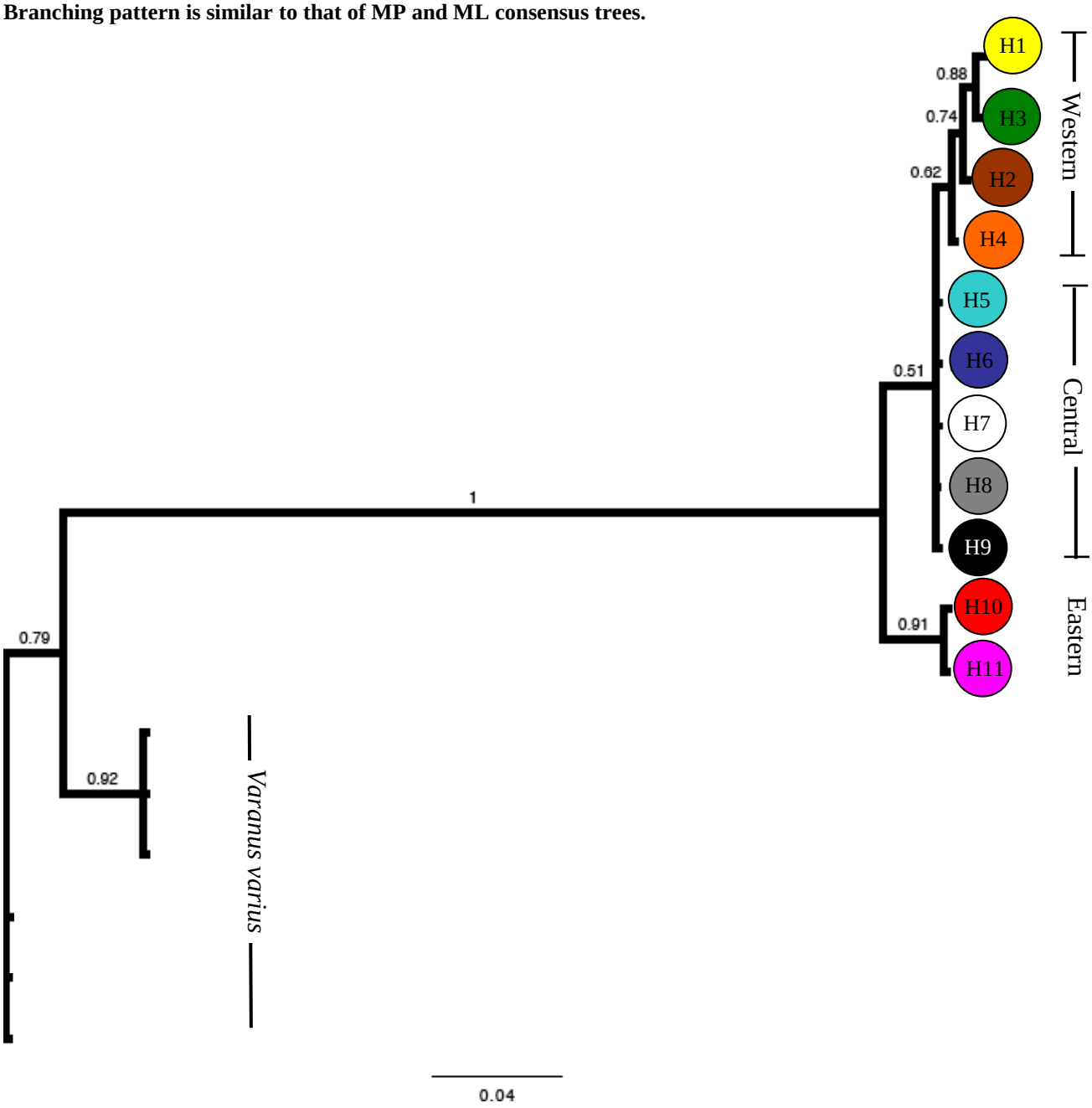


3.4.2. Bayesian Inference

The Bayesian consensus tree was reconstructed using the HKY+G model that was also applied for the ML analysis. This model reflects an overall substitution model for the data set, which was one type of transitional change and one type of transversional change that fitted best with a Gamma distribution. Therefore, the likelihood parameters were set with a number of substitution types of two (Nst= 2) and Gamma rates for substitution across nucleotide sites. I ran one million (1,000 000) Markov Chain Monte Carlo (MCMC) replicates with a default setting of one (1) cold chain and three (3) hot chains. The trees were sampled every one thousand (1,000) generations and a convergence was reached when the average standard deviation of split frequencies was less than 0.01. The first two hundred (200) trees were discarded as burn-ins. Finally, a 50% majority-rule consensus tree and the posterior probability estimates were obtained from the remaining eight hundred (800) trees.

Despite the low posterior probabilities for many nodes, the Bayesian consensus tree showed a similar branching pattern to the MP and ML consensus trees. There was a relatively high support for the Eastern subclade (~90%), whereas a support for the larger clade containing the Central and Western haplotypes was minimal (~50%). Additionally, four haplotypes distributed predominantly on Komodo, i.e. H1, H2, H3, and H4 were nested in a separate clade within this larger clade, although the posterior probability value for this clade was low. Overall, a monophyly for the Komodo monitor was supported well (100%), and two distinct clades of the Lace monitor were positioned at the root of the tree. Similarly, a monophyly for Komodo monitor was also shown in the MP and ML analyses, with a slight difference in the root topology of the MP tree. Figure 12 shows the topology of the Bayesian consensus tree, which is rooted with six sequences of the Komodo monitor sister species, i.e. the Lace monitor (*Varanud varius*) from Australia.

Figure 12. Phylogram of the Komodo monitor haplotypes with the Lace monitor sequences as outgroups. Branching pattern is similar to that of MP and ML consensus trees.



3.5. Population genetic structures

3.5.1. Analysis of Molecular Variance (AMOVA)

The genetic structure of the entire Komodo monitor population was determined by evaluating the amount of molecular variance distributed among population subdivisions. I assessed two possible population structures, i.e. at island and region levels using the Analysis of Molecular Variance (AMOVA) that I ran in Arlequin. I proposed a population structure at

an island-level, in which the five island populations on Komodo, Rinca, Flores, Gili Motang, and Nusa Kode were regarded separately as five different subdivisions. Nevertheless, the whole population of the Komodo monitor could also be structured at a region-level, in which sampling locations on islands were grouped in three separate regions, i.e. Western, Central, and Eastern. I evaluated this population subdivision in three regions based on the geographic distribution of the three haplotype-groups shown in Figures 9 and 10. All sampling locations on Komodo were included in the Western region, and all sampling locations on the islands of Rinca, Gili Motang, Nusa Kode, as well as a location in the western coast of Flores, i.e. Florest West, were grouped in the Central region. Flores North was the only location sampled in the Eastern region. Table 4 shows the results from AMOVA. A slightly over the half of the total molecular variation was accounted for by the variance components among island populations (Design I). By comparison, almost 90% of the total molecular variation was explained by the variance components among regions (Design II). These results show that the whole population of the Komodo monitor was significantly structured, and a stronger structure is shown by the higher differentiations among regions.

Table 4. Analysis of Molecular Variance (AMOVA) of two population structure designs. Design I shows a genetic partitioning among islands, and Design II shows the amount of molecular variance partitioned among regions.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Design I				
Among islands	4	129. 119	0.407	53.81
Among sampling sites within an island	8	49.320	0.238	31.43
Within sampling sites	353	39.392	0.112	14.76
Total	365	217.831	0.757	100.00
Design II				
Among regions	2	173.358	1.032	89.14
Among sampling sites within a region	10	5.080	0.014	1.22
Within sampling sites	353	39. 392	0.112	9.64
Total	365	217.831	1.158	100.00

I assessed the level of genetic diversity among islands and among regions based on the differentiation indices generated in AMOVA. In the island structure, the differentiation indices were relatively high in all hierarchies. The genetic diversity within sampling locations was fairly high (φ_{ST}= 0.852, P< 0.01), whereas the genetic diversity among populations within an island (φ_{SC}= 0.680, P< 0.05) and that among islands (φ_{CT}= 0.538, P< 0.01) were

relatively lower. The significance of the statistics was obtained after a permutation test with one thousand and twenty three (1023) pseudoreplicates. Similar to the φ-statistic results for the island structure, I found a high level of differentiations for the region structure. The genetic diversity for within-sampling locations across the distribution range was the greatest (φ_{ST}= 0.903), and that among region populations was relatively high (φ_{CT}= 0.891). However, the within-region genetic diversity was relatively the lowest (φ_{SC}= 0.113). All these φ-statistic values for the region structure were significant at P< 0.01. It is important to note, that both island and region patterns show the highest differentiation within sampling sites, suggesting a high genetic diversity within local populations.

3.5.2. Genetic divergences at population, island, and region levels

The amount of genetic divergence within a population, i.e. sampling location was computed as the mean number of pairwise sequence differences (π), which measures the net differences among all pairs of sequences in a given population. This molecular index differs from nucleotide diversity (π_n), in that the latter quantifies the amount of variability per nucleotide site across all sequences in a given population. On the other hand, the mean number of pairwise difference (π) measures the total of variable nucleotide sites between paired sequences in a population. In short, π_n is a measure of genetic diversity, whereas π is a measure of genetic divergence. Matrix 2 shows the mean number of pairwise sequence differences (π) within each of the thirteen (13) sampling locations and the corrected mean number of pairwise sequence differences between sampling locations. The corrected π values for comparisons between sampling locations were obtained by subtracting a given value of between-sampling location π with the average of pairwise within-sampling location π. This correction can be written as: π_{XY}- (π_X + π_Y)/2. In addition, the degree of genetic divergence was also demonstrated as the extent of genetic differentiation between pairs of sampling locations across the whole geographic distribution. This pairwise genetic differentiation was calculated as pairwise φ_{ST} shown in Matrix 3. Note that these values for pairwise φ_{ST} were computed based solely on the covariance of pairwise distances. Thus, the variations in sequence length were not incorporated in the computation.

3.5.2.1. Divergences at population level

A sampling location is the lowest unit in a population hierarchy. The genetic divergence in terms of the mean number of pairwise differences (π) among individuals at this level was the highest in Loh Lawi (Komodo), followed by Flores West (Flores) and Loh Wau (Komodo). The genetic divergence in three locations on Rinca, i.e. Loh Baru, Loh Dasami, and Loh Tongker, as well as in Flores North and in the two small islands was zero. This figure shows that the amount of genetic distinction within these populations could not be quantified in terms of nucleotide sequence divergence, because the haplotypes found in these locations differed only by the length of their nucleotide sequence. Three of such haplotype were shared among the above three locations on Rinca and the small islands. Similarly, the two haplotypes found in Flores North also differed only in their sequence length.

The genetic divergence between pairs of populations was found maximal between Flores North and Loh Wau (Komodo). There seems to be a trend of a higher divergence between Flores North and any population located westward. In addition, the genetic divergence between Loh Sebita (Komodo) and each of the two small islands were also among the largest. In contrast, a zero genetic divergence was found among all locations on Rinca (Matrix 2). It becomes more interesting to note, that the level of genetic divergence measured as the genetic differentiation between populations, i.e. pairwise φ_{ST}, shows a similar trend to the degree of genetic divergence among pairs of populations measured as π. The level of genetic differentiation (pairwise φ_{ST}) was relatively high between Flores North and each of the other twelve (12) sampling locations, whereas pairwise φ_{ST} values among populations distributed on Rinca were very close to zero, i.e. indicated as negative. A low level of differentiation was also found among pairs of populations distributed in the Central region, including those between Flores West and each the two small islands. Interestingly, a few pairs of populations distributed on Komodo also show a lower level of genetic divergence and differentiation, e.g. Loh Lawi-Loh Wenci, and Loh Liang-Loh Sebita (Matrices 2 and 3).

Matrix 2. Mean number of pairwise differences (π) within sampling location (in bold) and the standard deviations (in italics). Below diagonal are corrected π among populations, and above diagonal are indicators of statistical significance/ non-significance for corrected π, at $P \leq 0.05$. Statistically significant π values are indicated by a plus sign (+), whereas statistically non-significant π values are indicated by a minus sign (-).

Island	Sampling sites	Komodo					Rinca				Flores		Nusa Kode	Gili Motang
		LWE	LSE	LLI	LLA	LWA	LBU	LBA	LDS	LTK	FLW	FLN	NSK	GMO
Komodo (117) 0.561 ± 0.462	LWE (7)	**0.476**	-	-	-	+	+	+	+	+	+	+	+	+
		0.464												
	LSE (31)	0.019	**0.370**	-	+	+	+	+	+	+	+	+	+	+
			0.364											
	LLI (43)	-0.027	0.008	**0.401**	-	+	+	+	+	+	+	+	+	+
				0.379										
	LLA (28)	0.035	0.205	0.113	**0.770**	+	+	+	+	+	+	+	+	+
					0.580									
	LWA (8)	0.661	0.043	0.104	0.294	**0.536**	+	+	+	+	+	+	+	+
						0.493								
Rinca (165) 0.036 ± 0.101	LBU (74)	1.430	1.716	1.610	1.179	2.043	**0.081**	-	-	-	-	+	-	-
							0.154							
	LBA (58)	1.476	1.815	1.660	1.222	2.107	0.000	**0.000**	-	-	-	+	-	-
								0.000						
	LDS (14)	1.476	1.815	1.660	1.222	2.107	0.000	0.000	**0.000**	-	-	+	-	-
									0.000					
	LTK (19)	1.476	1.815	1.660	1.222	2.107	0.000	0.000	0.000	**0.000**	-	+	-	-
										0.000				
Flores (45) 2.802 ± 1.508	FLW (39)	1.425	1.765	1.610	1.171	2.056	0.271	-0.000	-0.000	-0.000	**0.564**	+	-	-
											0.470			
	FLN (6)	8.476	8.880	8.660	8.222	9.107	8.973	9.000	9.000	9.000	8.538	**0.000**	+	+
												0.000		
Nusa Kode	NSK (12)	1.476	9.000	1.660	1.222	2.107	0.000	0.000	0.000	0.000	-0.000	9.000	**0.000**	-
													0.000	
Gili Motang	GMO (27)	1.476	9.000	1.660	1.222	2.107	0.000	0.000	0.000	0.000	-0.000	9.000		**0.000**
														0.000

Matrix 3. Pairwise φ_{ST} among sampling locations (below diagonal) and statistical significance/ non-significance (above diagonal) at $P \leq 0.05$. Statistically significant φ_{ST} values are indicated by a plus sign (+), whereas non-significant φ_{ST} values are indicated by a minus sign (-). Sample size for each sampling location is given between brackets.

Island	Sampling location	Komodo					Rinca				Flores		Nusa Kode	Gili Motang
		LWE	LSE	LLI	LLA	LWA	LBU	LBA	LDS	LTK	FLW	FLN	NSK	GMO
Komodo	LWE (7)	-	-	-	-	+	+	+	+	+	+	+	+	+
	LSE (31)	0.061	-	-	+	-	+	+	+	+	+	+	+	+
	LLI (43)	-0.057	0.019	-	+	+	+	+	+	+	+	+	+	+
	LLA (28)	0.026	0.269	0.175	-	+	+	+	+	+	+	+	+	+
	LWA (8)	0.233	0.113	0.209	0.282	-	+	+	+	+	+	+	+	+
Rinca	LBU (74)	0.929	0.914	0.891	0.816	0.945	-	-	-	-	-	+	+	-
	LBA (58)	0.971	0.971	0.907	0.832	0.973	-0.003	-	-	-	-	+	-	-
	LDS (14)	0.908	0.875	0.845	0.699	0.919	-0.036	0.000	-	-	-	+	-	-
	LTK (19)	0.926	0.887	0.855	0.725	0.934	-0.025	0.000	0.000	-	-	+	-	-
Flores	FLW (39)	0.755	0.807	0.789	0.665	0.813	0.008	0.010	-0.032	-0.021	-	+	-	-
	FLN (6)	0.970	0.965	0.960	0.926	0.967	0.992	1.000	1.000	1.000	0.954	-	+	+
Nusa Kode	NSK (12)	0.899	0.870	0.838	0.687	0.910	-0.042	0.000	0.000	0.000	-0.039	1.000	-	-
Gili Motang	GMO (27)	0.944	0.901	0.870	0.757	0.949	-0.016	0.000	0.000	0.000	-0.010	1.000	0.000	-

3.5.2.2. Divergences at island level

The mean number of pairwise difference at the next hierarchical level, i.e. a group of sampling locations within an island, was found the highest on Flores. Komodo and Rinca also show a relatively high genetic divergence as measured by π, whereas no genetic divergence was found within the small island populations (Matrix 2). A zero divergence within the islands of Gili Motang and Nusa Kode was shown due to the presence of only one haplotype on these two small islands. The level of genetic divergence and the degree of differentiation among islands are given in Matrix 4. The genetic divergence at island level was the highest between Komodo and each of the small islands (π= 1.583, *P*= 0.000), which is comparable to the divergence between Komodo and Rinca (π= 1.560, *P*= 0.000). Further, three pairs of island population in the Central region, i.e. Rinca-Gili Motang, Rinca-Nusa Kode, and Gili Motang-Nusa Kode apparently show almost no genetic divergence, because the majority of haplotypes distributed on these islands differ only in their sequence length. Concordant with this trend in genetic divergence, the level of differentiation between Komodo and each of the small islands were relatively high (φ_{ST}= 0.775 and 0.753). It is important to note, that the differentiation between Komodo and Rinca was the largest (φ_{ST}= 0.860, *P*= 0.000). A lesser extent of differentiation was present between Flores-Rinca and Flores-Gili Motang. Finally, there was no differentiation between Rinca-Gili Motang, Rinca-Nusa Kode, as well as between the two small islands, all of which are grouped in the Central region.

Matrix 4. Pairwise φ_{ST} among islands (below diagonal) and mean number of pairwise sequence differences (π) among islands (above diagonal). Underlined values are statistically significant at $P \leq 0.05$, whereas values not underlined indicate a statistical non-significance. Sample size for each island is given in brackets.

Island	Komodo	Rinca	Flores	Nusa Kode	Gili Motang
Komodo (117)	-	1.560	1.465	1.583	1.583
Rinca (165)	0.860	-	0.189	0.000	0.000
Flores (45)	0.580	0.278	-	0.191	0.191
Nusa Kode (12)	0.753	-0.043	0.056	-	0.000
Gili Motang (27)	0.775	-0.018	0.102	0.000	-

3.5.2.3. Divergences at region level

Matrix 2 shows that the genetic divergence within a region was the largest in the Western region, which included all populations on Komodo (π= 0.561 ± 0.462). On the other

hand, no genetic divergence was found in Flores North, i.e. the Eastern region. A genetic divergence at the highest population hierarchy was determined among three regions, i.e. Western, Central, and Eastern that made up the structure of the Komodo monitor population (Matrix 5). It is interesting to observe in the data, that the divergence among regions was relatively the highest between the Eastern and Central regions (π= 8.918, *P*= 0.000). However, the divergence between the Eastern and Western regions was also as high (π= 8.599, *P*= 0.000). Apparently, the genetic divergence between the Western and Central regions was significantly the least (π= 1.560, *P*= 0.000), suggesting a higher sequence similarity among populations on Komodo and those in the Central region. Further, the population in Flores North that represents the Eastern region in this study was the most distinct population, based on my observation on the sequence divergence data only. The same trend applies to the genetic divergence measured in pairwise φ_{ST}s among regions. The pairwise population differentiation indicated the largest genetic divergence between the Eastern and Central regions (φ_{ST}= 0.989, *P*= 0.000). This figure is fairly close to the differentiation between the Eastern and Western regions (φ_{ST}= 0.941, *P*= 0.000). Finally, the level of divergence between the Western and Central regions was also the least (φ_{ST}= 0.863, *P*= 0.000).

Matrix 5. Pairwise φ_{ST} among regions (below diagonal) and mean number of pairwise sequence differences (π) among regions (above diagonal). Underlined values are statistically significant at *P* values ≤ 0.05, whereas values not underlined indicate statistical non-significance. Sample size for each island is given in brackets.

Region	Western	Central	Eastern
	Komodo	Rinca, Flores West, Nusa Kode, Gili Motang	Flores North
Western (117)	-	1.560	8.599
Central (243)	0.863	-	8.918
Eastern (6)	0.941	0.989	-

3.5.3. Shared haplotypes and lineage divergences among islands and regions

The genetic divergences (π and pairwise φ_{ST}) among islands and regions have shown a degree of distinction for the population on Komodo, i.e. the Western region, from those on the islands eastward. This distinction for Komodo may suggest an interesting evolutionary history on this island population. Komodo was characterised by a total of four (4) haplotypes. Three of these haplotypes, i.e. H1, H2, and H4 were not shared with any other island, whereas the haplotype H3 was also found on Rinca. Haplotype H3 was found in four (4) of

the five (5) sampling locations on Komodo, as well as in Loh Buaya, on Rinca (Figure 9). The distribution of haplotype H3 on Komodo and Rinca may indicate an event of migration between these islands. Likewise, an inter-island migration might have taken place between Rinca and Nusa Kode, Rinca and Gili Motang, Rinca and Flores, as well as between Flores and Gili Motang. Furthermore, a migration might have also occurred between two locations on Flores, i.e. Flores West and Flores North. These putative migrations among populations across islands may subsequently affect the degree of divergence among islands as well as among regions. To quantify the extent of haplotype divergence between islands and between regions, I computed *p* distances from all pairs of different haplotypes found across island populations and region populations. The mean of these pairwise *p* distances can be used as an index of haplotype divergence between two island populations and/or two region populations. In particular, within-island and within-region haplotype divergence indices were computed as the mean of all *p* distances from all pairs of different haplotypes present on an island and on a region, respectively. Similarly, I computed between-island and between-region haplotype divergence indices as the mean of all *p* distances from all pairs of different haplotypes present in two given islands and two given regions, respectively. Matrix 6 shows the haplotype divergence indices for four islands, i.e. Komodo, Rinca, Nusa Kode, and Gili Motang and for the three regions I defined previously. The two locations on Flores, i.e. Flores West and Flores North were treated separately in the Central and Eastern region, respectively.

Matrix 6. The mean uncorrected distances among pairs of different haplotypes (pairwise *p* distances) among islands and regions computed as haplotype divergence indices (%) and the standard deviation.

Region	Island/Location	Western	Central				Eastern
		Komodo	Rinca	Flores West	Nusa Kode	Gili Motang	Flores North
Western	Komodo	0.24 ± 0.0012					
Central	Rinca	0.44 ± 0.0020	0.22 ± 0.0033				
	Flores West	0.95 ± 0.0068	0.70 ± 0.0085	1.12 ± 0.0098			
	Nusa Kode	0.23 ± 0.0019	0.09 ± 0.0020	0.45 ± 0.0079	-		
	Gili Motang	0.50 ± 0.0018	0.14 ± 0.0032	0.79 ± 0.0111	0	-	
	Total	0.44 ± 0.0020	0.63 ± 0.0072				1.46 ± 0.0015
Eastern	Flores North	1.60 ± 0.0029	1.51 ± 0.0021	1.21 ± 0.0070	1.36 ± 0.0001	1.46 ± 0.0016	0

It was not surprising to find out that the haplotype divergence among regions was larger than that of among islands. Evidently, there was a tendency that some haplotypes were shared among islands within a region than among locations or islands in two different regions. For example, each of the haplotypes H5 and H6 was shared among three islands in the Central region. Haplotype H7 was shared between two islands, which are also grouped in the Central region. On the other hand, haplotype H3 was shared between two islands grouped in two different regions and haplotype H10 was shared between two locations situated in two separate regions. Further, haplotype divergence was the largest between the Western and Eastern regions (0.016 ± 0.0029), whereas that between the Western and Central regions (0.004 ± 0.0020) was relatively lower. Interestingly, the haplotype divergence between the Central and Eastern regions (0.015 ± 0.0015) seemed to be comparable to that between the Western and Eastern regions. Unfortunately, the sample size for the Eastern haplotypes was too small that it did not permit a statistcal test, e.g. Student's t-test to perform a meaningful comparison of two means from two different samples. However, these results seem to show a greater distinction for the Eastern region from the other regions, because of the two most divergent haplotypes found in this region (Matrix 1, Figure 9). Further, the haplotype divergence between the islands of Komodo and Flores (0.010 ± 0.0061, data not shown in Matrix 6 above) was the largest among all the between-island comparisons. The within-island haplotype divergence for Flores (0.012 ± 0.0070) was the greatest among all the within-island indices. Indeed, it has been shown from the pairwise *p* distances of all haplotypes found in this study, that the two haplotypes distributed in Flores North were the most divergent (Matrix 1). Moreover, the current pattern of haplotype distribution shows no evidence of migration to Flores North that may carry a third haplotype. Thus, a higher level of haplotype divergence may be expected for Flores North. On the contrary, no haplotype divergence was expected for the small island populations of Nusa Kode and Gili Motang, due to the presence of only a single haplotype on each of these islands. No haplotype divergence was also expected between Nusa Kode and Gili Motang, because the two haplotypes found on these islands differed only by their sequence length. Likewise, a haplotype divergence could not be detected within the population in Flores North, i.e. the Eastern region, because no variable site was found in the aligned haplotypes found in this region.

Using the haplotype divergence indices in Matrix 6, I estimated a raw divergence time among lineages grouped in islands and regions. This estimate may imply a time since the

most recent separation between two island populations or region populations, because it takes into account a putative migration event, which was inferred from the shared haplotypes in the two populations. Based on the substitution rates for mtDNA sequences of 2% per site per million years (Brown *et al.*, 1976), as well as a mtDNA Control Region evolutionary rate of 2.5% per site per million years determined for humans (Parsons *et al.*, 1997), I computed a raw time since the most recent separation between two geographically-grouped lineages. In addition, the estimates of divergence time were also calculated for a much faster rate of mtDNA Control Region, which is thought to evolve four to five times faster than all the other sequences of mtDNA (Taberlet, 1996b). The method that I used to date the population divergence here is similar to the method that was applied to date the divergence time for the populations of the Common European adder, *Vipera berus* (Ursenbacher *et al.*, 2006). Table 5 shows my estimates of the raw divergence time among maternal lineages grouped geographically in islands and regions. Note that Komodo and Flores North may be regarded as the Western and Eastern region, respectively.

Table 5. Estimates of raw divergence time (in thousand years) among lineages grouped in islands and regions based on the haplotype divergence indices presented in Matrix 6. Figures in brackets are the range estimates calculated from the standard deviation of the indices.

Island/ location	Substitution rate (per million years)			
	2%	2.5%	8%	10%
Komodo-Rinca	220 (120-320)	176 (96-256)	55 (30-80)	44 (24-64)
Komodo-Flores North	800 (655-945)	640 (524-756)	200 (164-236)	160 (131-189)
Rinca-Flores North	755 (650-860)	604 (520-688)	189 (162-215)	151 (130-172)
Rinca-Nusa Kode	45 (0-145)	36 (0-116)	11 (0-36)	9 (0-29)
Flores North-Flores West	605 (255-955)	484 (204-764)	151 (64-239)	121 (51-191)
Flores West-Gili Motang	395 (0-950)	316 (0-760)	99 (0-237)	79 (0-190)

A refined estimation of population divergence is not as simple as I demonstrated here. However, two of the four important components in the calculation of divergence time between two populations, i.e. genetic distance and migration with gene flow were considered in the method I used in this study. Two other features required to properly estimate population divergence time are the ancestral and descendant effective population sizes (Nei, 1972, Arbogast *et al.*, 2002). Nevertheless, my rough estimates of population divergence

seem to produce interesting results. First, the estimates of divergence time based on the haplotype divergence index fell within a time scale of less than one million years, presumably due to the the low genetic divergence among haplotypes. Even a time calibration using the slowest substitution rate resulted in an estimate of only nearly one million years, e.g. maximum separation time between Komodo and Flores North (the Western and Eastern regions). Second, the divergence times estimated from the haplotype divergence indices seem to be in accord with the earliest fossil evidence of the Komodo monitor in the Lesser Sundas. The oldest fossil Komodo monitor in these islands were the teeth found on Flores dated to ~900,000 years (Morwood *et al.*, 1999, van den Bergh *et al.*, 2001). Interestingly, the estimates of the divergence time between the most divergent lineages, i.e. those distributed in Flores North and in Komodo were as great as ~800,000 years and these figures fall within the time frame since the evidence of the Komodo monitor's first appearance in the Lesser Sundas. Third, the raw estimates of divergence time based on the index of haplotype divergence apparently showed an association between time and geographic distance among locations. For example, the time estimates for the separation between Komodo and Rinca ranged from ~44,000 to 220,000 years ago, whereas those between Rinca and the nearby Nusa Kode was shown to be much more recent, i.e. ~9,000-45,000 years ago. These figures seem to be consistent with the current geographic distances between islands in this example.

3.5.4. Relationship between genetic and geographic distances

I evaluated the relationship between genetic and geographic distances using Mantel's test available in Arlequin version 3.1. The genetic and geographic distance matrices were generated from pairwise φ_{ST} between sampling locations and pairwise straight-line distance between sampling locations, respectively to test for a correlation between haplotypes and their distribution across islands and regions. I obtained the coordinates for sampling locations from Panoramio™, a worldwide online database of photographic images accessible at http://www.panoramio.com/map/#lt=-8.4&ln=119.35&z=0. This database is linked to the web-based global positioning tool Google Earth™, by which I determined the geographic distances between sampling locations. All photographs published in the online database were taken on location and labelled by the author (Aganto Seno, personal communication) during a series of regular field surveys conducted by a team from Balai Kehutanan dan Sumberdaya Alam (BKSDA) of the Indonesian Ministry of Forestry. Matrix 7 shows the straight-line distances between pairs of sampling locations measured in kilometer.

In the Arlequin input file, the pairwise values of φ_{ST} were assigned as matrix Y, which had a mean of 0.551. The geographic distance matrix was assigned as matrix X1, and the mean for this matrix was 41.373 km. The results from Mantel's test showed a significant positive correlation between genetic and geographic distances. This correlation explains that geographic distance determines the pattern of the genetic distance, given the data matrices. Further, the matrix Y of pairwise φ_{ST} values was determined by matrix X1 of geographic distances by 37% (Mantel's test r^2= 0.370, r= 0.608, *P*< 0.01, 1000 permutations, regression equation: y= 0.37x + 0.006). In addition, the correlation between the two matrices was increased to almost 61% after a log-transformation of geographic distances (Mantel's test r^2= 0.603, r= 0.777, *P*< 0.01, 1000 permutations, regression equation: y= 0.603x + 0.844). Therefore, geographic distance seems to be an influential factor in the haplotype distribution for the Komodo monitor. However, a distance of seawater may be more effective to explain the current haplotype distribution across islands for this terrestrial varanid.

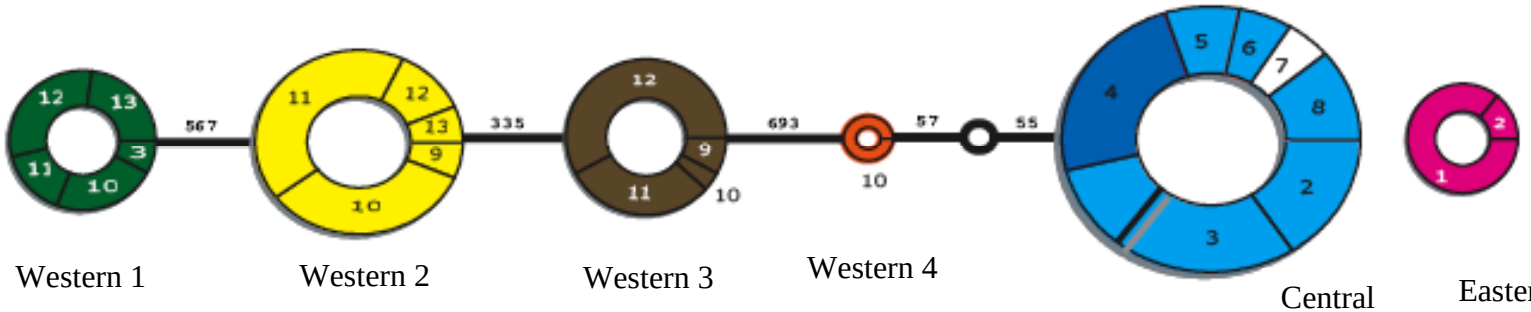
Matrix 7. Geographic distances between sampling locations (km) were measured as the shortest straight line connecting two given locations.

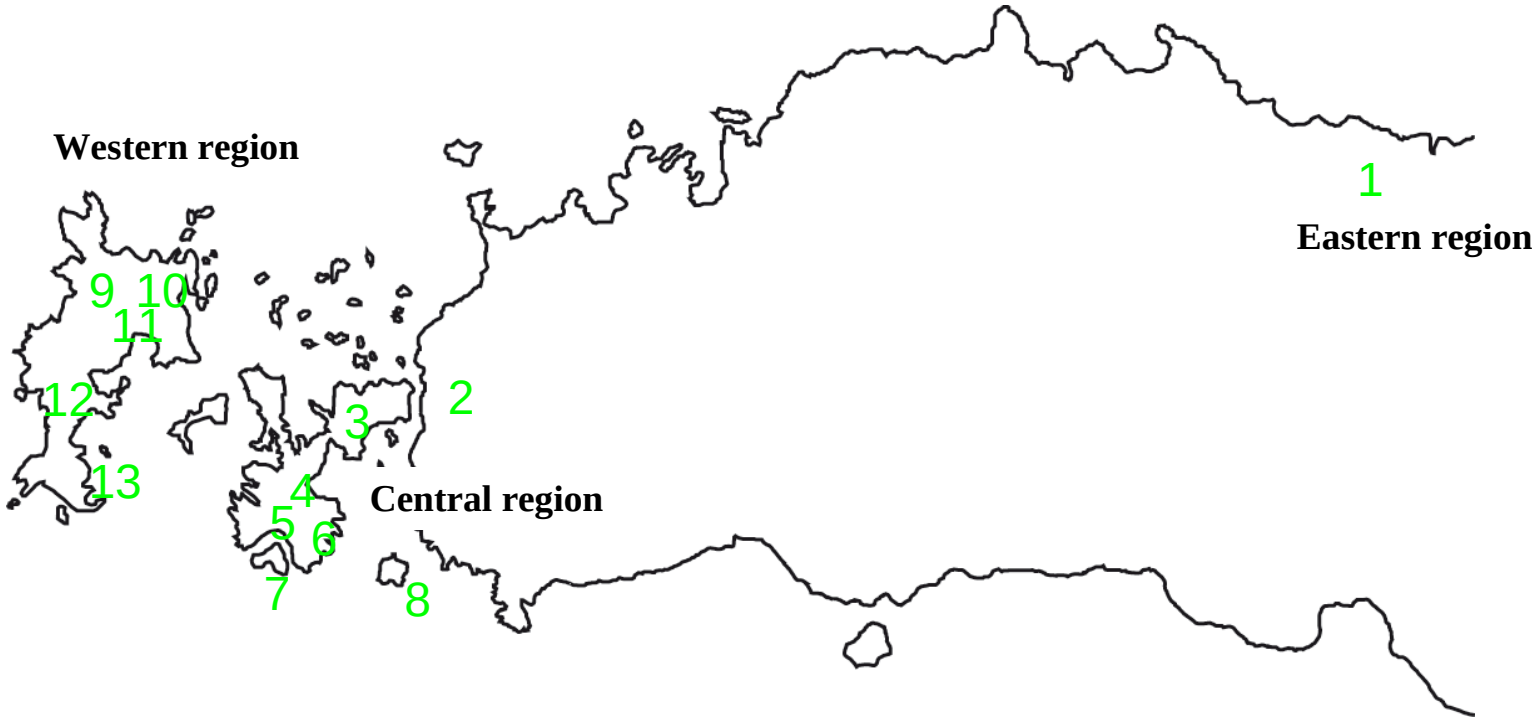
Island	Sampling locations	Komodo					Rinca				Flores		Nusa Kode	Gili Motang
		LWE	LSE	LLI	LLA	LWA	LBU	LBA	LDS	LTK	FLW	FLN	NSK	GMO
Komodo	Loh Wenci	-												
	Loh Sebita	12.82	-											
	Loh Liang	9.24	6.71	-										
	Loh Lawi	8.49	17.20	11.02	-									
	Loh Wau	19.69	21.20	15.71	11.76	-								
Rinca	Loh Buaya	35.92	23.59	26.47	35.70	31.53	-							
	Loh Baru	36.30	26.53	27.26	34.35	27.51	8.49	-						
	Loh Dasami	39.50	30.61	30.80	35.74	27.34	14.80	6.89	-					
	Loh Tongker	40.74	30.66	31.85	38.62	31.25	11.54	4.74	7.02	-				
Flores	Flores West	45.02	31.82	36.73	47.20	44.75	14.36	22.23	27.72	22.63	-			
	Flores North	147.17	135.23	141.11	151.48	150.54	119.01	124.72	128.87	122.79	105.41	-		
Nusa Kode	Nusa Kode	38.69	30.49	29.96	34.46	25.50	16.31	7.78	1.55	8.56	28.92	130.86	-	
Gili Motang	Gili Motang	48.57	38.12	40.18	46.70	38.73	17.14	12.42	12.05	8.16	23.82	118.53	21.20	-

3.6. Haplotype network

I generated a network to evaluate the relationships among all eleven Komodo monitor haplotypes using TCS (Clement *et al.*, 2000). In this analysis, I treated alignment gaps as "missing" and therefore, indels were not considered as sequence variation. The connection limit for parsimony probability was set at 95%, i.e. as the default in this programme. All the four Western haplotypes I defined earlier in the Section 3.2, i.e. H1, H2, H3, and H4 were shown in the network. This result confirms the absence of sequence length differences among these haplotypes. On the other hand, all the five Central haplotypes, i.e. H5, H6, H7, H8, and H9 were regarded as a single haplotype and labelled as "Central" in this network (Figure 13). The collapsing of all Central haplotypes into one was due to the presence of three variations in sequence length, i.e. full-length (727 bp), medium-length (705 bp), and short-length (669 bp). In addition, the single indels detected at two different positions within two of the three Central haplotypes with medium-length sequences, i.e. H8 and H9 were also ignored. Thus, haplotypes H8 and H9 were also collapsed into "Central". Similarly, the two Eastern haplotypes were also collapsed into a single haplotype called "Eastern" in this network, confirming their difference only in sequence size. Despite this discrepancy between haplotypes defined using TCS and those I defined originally in section 3.2, the haplotype network seems to be useful to illustrate a closer relationship between the Western and Central haplotypes. A result from this analysis shows a connection between all the four Western haplotypes to all the five Central haplotypes. This connection is shown as a possibly unsampled haplotype, which is depicted as a small black circle in the network (Figure 13). In addition, the relatively distant position between the Eastern haplotypes and all the other haplotypes in this study is also depicted in the network. The Eastern haplotypes are not connected with the rest of the network, suggesting a relatively great sequence divergence between these haplotypes and all the other haplotypes determined in this study (see Matrix 1 for *p* distances).

Figure 13. Haplotype network depicting relationships among haplotypes distributed in 13 sampling locations. Colours indicate original haplotype definition based on sequence length, nucleotide substitution, and single indel positions. Numbers in green and numbers beside or in the circles indicate sampling location (1=FLN, 2=FLW, 3=LBU, 4=LBA, 5=LDS, 6=LTK, 7=NSK, 8=GMO, 9=LWE, 10=LSE, 11=LLI, 12=LLA, 13=LWA). Numbers beside black lines indicate nucleotide positions that distingusih a haplotype from the others.





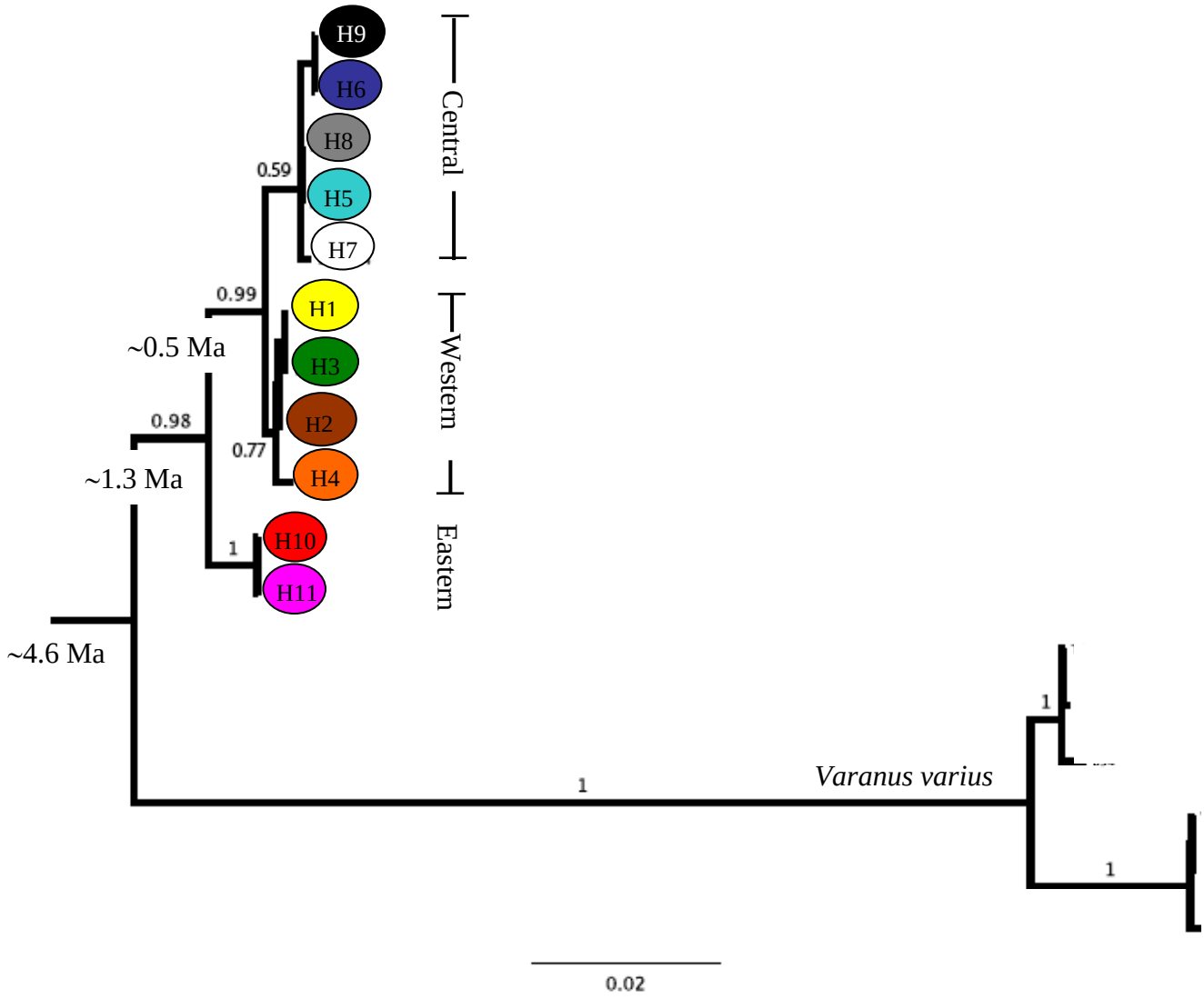
3.7. Divergence time estimates

I ran an analysis of divergence time using the computer programme BEAST to estimate the age of the putative haplotype groups for the Komodo monitor. These groups are defined earlier in this study by their geographic distribution (Figure 9) and NJ tree (Figure 10). I generated a separate Bayesian tree in BEAST to infer a bifurcating topology of the eleven Komodo monitor haplotypes. One hundred thousand (100,000) trees were sampled from the Markov chain generated using BEAST to reconstruct a Bayesian consensus topology and its subsequent posterior probabilities (Figure 14). This topology was very similar to the Bayesian tree topology inferred using Mr. Bayes (Figure 12). The nucleotide substitution model applied in this analysis is HKY+G, and the molecular clock model was set as a relaxed clock with an uncorrelated Exponential distribution. To calibrate the node that splits into the Komodo monitor and the Lace monitor, I used fossil vertebrae of the Komodo monitor from Australia dated to ~3.8 Ma as the zero offset of the Exponential prior. In addition, I used fossil teeth of the Komodo monitor found on Flores dated to ~0.9 Ma (Morwood *et al.*, 1998) to calibrate the node that splits into the Eastern and Central-Western clades of the Komodo monitor.

The segregation between the Komodo monitor and the Lace monitor was estimated to occur about 4.6 Ma (mean: 4.595 Ma, 95% HPD range: 3.800-6.144 Ma). Interestingly, this estimate is consistent with the age of the fossil Giant Roamer (*V. priscus*) found in Australia (Molnar, 2004a). It is also important to note, that the separation between the Eastern and Central-Western haplotype groups was estimated to have occurred since slightly over one million years ago (mean: 1.261 Ma, 95% HPD range: 0.900-1.951 Ma). In addition, the Western and Central clades were estimated to split since around 500,000 years ago (mean: 0.499 Ma, 95% HPD range: 0.116-1.034 Ma). This estimate for the divergence between the Western and Central clades predates the earliest known glacial period that is estimated to occur about 250, 000 years ago (Voris, 2000), suggesting a relatively long divergence among female lineages distributed on Komodo and those on the eastern islands. If a gene flow occurred between the populations on Komodo and Rinca, as indicated by the distribution of haplotype H3 on both islands (Figure 9), female migrants might have reached or come from either of these islands since at least half of a million years ago. It is possible that a female migration occurred from Komodo to Rinca because haploytpe H3 seems to be more common on Komodo than on Rinca. Thus, one may expect a rare female migration from the

population with a higher frequency of haplotype distribution to the population with lower frequency of haplotype distribution. The frequency of migration between Komodo and Rinca is relatively low, i.e. one individual animal every ~67-191 years (Ciofi and Bruford, 1999) and females tend to disperse less than males do (Ciofi and De Boer, 2004). Otherwise, both island populations seem to retain haplotype H3 that was once evolved on either island.

Figure 14. Divergence time estimates based on the relaxed clock model. The Central-Western and Eastern node was calibrated with the fossil teeth of the Komodo monitor from Flores dated to 0.9 Ma. The node splitting into the Komodo monitor and the Lace monitor was calibrated with the fossil vertebrae from Australia dated to 3.8 Ma.



Chapter 4

DISCUSSION

4.1. Variations among the Komodo monitor Control Region I (CRI) sequences

The vertebrate mitochondrial DNA Control Region sequences have been found highly polymorphic, for example in mammals (Saccone *et al.*, 1991), fish (Brown *et al.*, 1993), and birds (Ruokonen and Kvist, 2002). Therefore, this rapidly evolving non-coding region is useful to infer the relationships among populations of a species and to construct a phylogeny of closely related species (Taberlet, 1996a). In the alignment of mitochondrial DNA Control Region I (CRI) sequences of the Komodo monitor, there are two major types of polymorphism. One source of polymorphism in this data set is the variation in nucleotide substitutions, for which the overall extent was expressed as the mean number of pairwise differences across all samples (π= 1.199 ± 0.769). The other source of polymorphism in this data set is the difference in sequence length. The CRI sequences for the Komodo monitor vary largely due to the presence of contiguous multiple indels, which are denoted as the alignment gaps near the flanking $tRNA^{Thr}$ (the 5' end of the L-strand). These long indels are apparently generated by a variation in the number of tandemly repeated sequences termed as the Variable Number of Tandem Repeats (VNTRs). I assume, that the VNTRs result from a process of DNA mispairing prior to replication called illegitimate elongation (Buroker *et al.*, 1990). In this process, a tandem repeat can fold into a structurally stable hairpin loop, which facilitates incorrect pairings of H- and L-strands before the course of DNA replication. This misalignment of DNA strands can eventually result in either a gain or loss of a repeat unit. Similarly, a sequence length variation can also be generated by an intrahelical slipped-strand mispairing followed by an interhelical unequal crossing-over during a DNA replication (Levinson and Gutman, 1987). Interestingly, this variation in sequence length can be found within an individual animal and this phenomenon is called length heteroplasmy, which has been discovered in some vertebrates, including percid fish (Nesbø *et al.*, 1998) and crocodiles (Ray and Densmore, 2003). Nevertheless, a heteroplasmic state of the vertebrate mitochondrial DNA still remains to be investigated in many more species, especially reptilian species. In bats, a length heteroplasmy was reported to be consistently passed on from some female parents to their offsprings (Wilkinson and Chapman, 1991). In spite of this discovery,

the possibility for length heteroplasmy to alter the functional capacity of a mitochondrion may still need to be assessed. Otherwise, a length heteroplasmy may be considered to evolve in a neutral fashion, because the Control Region is a non-coding segment.

A length heteroplasmy can also occur in the Komodo monitor CRI sequences because of the presence of VNTRs (Figures 8a and 8b). However, the variation in sequence length due to these VNTRs can be regarded as a source of polymorphism, for two reasons. First, the variation in the number of tandem repeats may counterbalance the lower level of nucleotide substitutions, as reported in the Chihuahuan Spotted whiptail (*Cnemidophorus exsanguis*). The sequence divergence in this teiid lizard was relatively low, whereas the Control Region sequences vary in three different lengths (Moritz and Brown, 1986). Second, these tandem repeats are thought to be common in the vertebrate mitochondrial DNA Control Region and the mutational process of these repeated sequences is similar to nuclear microsatellites, which are useful to infer a population structure (Lunt *et al.*, 1998).

I discuss the two types of polymorphism in the Komodo monitor CRI sequences in more details, in two steps. First, I evaluate the structure of these sequences and the extent of nucleotide substitutions observed in the alignment of all the eleven haplotypes. Second, I assess the importance of length polymorphism for characterising the haplotypic forms of these sequences. In addition, I highlight the significance of a future study investigating the evolution of varanid mtDNA Control Region for an application in phylogeographic studies.

In general, the structure of the Komodo monitor CRI sequences was consistent with the structure of the mitochondrial Control Region in mammals, which is more commonly described for vertebrates. The variable sites in the mammalian Control Region are more frequently found in Domain I (ETAS domain) and Domain III (CSB domain), than in Domain II (Central Conserved domain). As a note, Domain I and Domain III flank Domain II. This pattern of variable sites distribution has been shown in several mammalian orders, for instance primates, rodents, and carnivores (Sbisà *et al.*, 1997). In addition, some other vertebrates also demonstrate this pattern of variable sites distribution in their Control Region sequences, for example sturgeons (Brown *et al.*, 1993), the Swordfish (Rosel and Block, 1996), some guillemot birds (Kidd and Friesen, 1998, Roques *et al.*, 2004), some vultures (Roques *et al.*, 2004), and also some crocodiles (Ray and Densmore, 2003). Further, the principal basis for the division of mtDNA Control Region sequence into three domains is the number of variable sites, i.e. the frequency of nucleotide substitution along the sequence. It

seems that no precise points of reference, such as the sequences flanking a domain, are used to delimit the three sections in an mtDNA Control Region sequence. The pattern of variable sites distribution in the present alignment of the Komodo monitor CRI sequences is generally congruent with that of the other vertebrates, where the nucleotide sites in the middle part of the sequence is less variable than those in the flanking regions. I allocated three domains in the Komodo monitor CRI sequences arbitrarily, based on the positions of ETAS1 and CSB1 boxes. The purpose of this allocation was to assess a pattern of the variable site distribution. More specifically, I assessed the possibility of the two characteristic boxes above to be used as the flanking segments for the Central Conserved domain. In Figure 8b, the ETAS domain ends at the last nucleotide site of the ETAS box (position 261). This domain harbours the greatest number of variable sites. The Central Conserved Domain starts just after the ETAS box (nucleotide position 262) and it ends just before the CSB1 box (nucleotide position 646). This domain contains fewer variable sites than the ETAS domain. Finally, the CSB domain starts in the CSB1 box (position 647) and it harbours the least variable sites. This pattern of variable site distribution demonstrates an inconsistency with the pattern reported for vertebrates, because there are more variable sites in the Central Conserved domain than in the CSB domain. Thus, the ETAS1 and CSB1 boxes in the Komodo monitor CRI seem to be unsuitable to delimit domains in the mitDNA Control Region. Nevertheless, an alignment of the Control Region sequences from multiple varanid lizard species is still needed, in order to make a more detailed evaluation on their general sequence structure. Furthermore, the duplicated state of some varanid Control Regions may be an important feature to estimate the rate of nucleotide substitutions, which generate variable sites within each domain. As a note, a copy of the Control region is thought to evolve in concert with its duplicate, as in the case of the duplicated Control Region of the Amazon parrots (Eberhard *et al.*, 2001) and snakes (Kumazawa *et al.*, 1996).

I found two discrepancies between the Komodo monitor CRI sequence structure and that reported from mammals. First, only one ETAS box seems to be present in the Komodo monitor CRI sequences (Kumazawa and Endo, 2004), whereas two ETAS boxes are commonly identified in mammals (Sbisà *et al.*, 1997). Second, two CSB boxes are observable in the Komodo monitor's CRI, while up to three CSB boxes can be found in some mammals. It is interesting to note, that the ETAS and CSB boxes are associated with a regulation function of mtDNA replication and transcription (Sbisà *et al.*, 1997). Thus, the presence or

absence of these boxes may be related to a mitochondrial genome evolution in vertebrates. Nevertheless, only a little is known on the structure and evolution of the mtDNA Control Region sequences in reptiles (Brehm *et al.*, 2003), especially in varanid lizards. Therefore, an in-depth study to characterise the structure and the subsequent evolution of varanid Control Regions is necessary to allow for the best use of this potential marker in population genetic, phylogenetic, and phylogeographic studies for these lizards.

More attention should also be paid on the phenomenon of length heteroplasmy in the mtDNA Control Region from a molecular evolution perspective, in addition to the evolution of variable nucleotide sites in this region. The variation in sequence length was a significant source of DNA polymorphism in my data set. Indeed, nucleotide substitutions across the three domains were apparently low. I found eleven haplotypes from a total of three hundred and sixty six (366) CRI sequences of the Komodo monitor, with some of these haplotypes differ only in their sequence length. The two haplotypes found in North Flores, i.e. H10 and H11 differ by 58 bp (~1.6 repeat units), while three (3) out of all the five (5) Central haplotypes differ to each other only in their sequence length. Among these three Central haplotypes, haplotype H6 has the full-length sequence (727 bp), whereas haplotype H7 has the shortest length (669 bp). Being the longest and shortest variants, these two haplotypes differ by 58 bp. On the other hand, haplotype H5 has a medium-length sequence (705 bp). It differs from haplotype H6 by 22 bp, and from haplotype H7 by 36 bp (a single repeat unit). I consider the other two haplotypes in this group, i.e. H8 and H9 as medium in length, although they differ from haplotype H5 by a single indel. This additional single indel further shortened the two sequences for haplotypes H8 and H9 to 704 bp (Figures 8a and 8b). It is interesting to note, that all the Western haplotypes are medium in length (705 bp), and they differ to one another by one to three (1-3) bases. I consider that the length of a sequence is an important variation to distinguish the Eastern and Central haplotypes, because there was a lack of observable nucleotide substitution between pairs of haplotypes in each of the Eastern and Central haplotype groups. On the contrary, all the Western haplotypes were determined by the variation in their nucleotide sites, due to the absence of variable sequence length among these sequences. In brief, a trade-off for one particular type of molecular variation apparently occurs among mtDNA CRI sequences of the Komodo monitor. The presence of a variable sequence length seems to be at the cost of the absence of variable nucleotide sites, and *vice versa*.

To summarize this discussion on the sequence variations, I present the general pattern of genetic divergence among haplotypes. The genetic divergence among haplotypes was calculated as the number of pairwise nucleotide difference (π). The number of variable sites between pairs of haplotypes was higher between the Eastern haplotypes and those from the Central or Western group. In particular, the Eastern haplotype H10 had the largest number of variable sites, when it was aligned with any of the Western haplotypes. Additionally, an alignment of haplotype H10 and the Central haplotype H6 also yielded a large number of variable sites (Matrix1). Interestingly, the number of variable sites was also increased, when two haplotypes from two different haplotype groups were aligned. Indeed, there seems to be a trend in the number of variable sites among pairwise alignments across haplotype groups. The number of variable sites between the Western and Eastern haplotypes was always among the highest, whereas the number of variable sites between the Western and Central haplotypes was constantly among the lowest. Parallel to this trend in the number of variable sites among haplotype groups, the genetic divergence measured as pairwise *p* distances was also the highest between the Western and Eastern haplotypes, while pairwise *p* distances were the lowest between the Western and Central haplotypes.

4.2. Genetic diversity among populations on islands

Eleven haplotypes were discovered in my data set, which comprises more than three hundred mtDNA CRI sequences. The number of haplotypes distributed across islands seems to vary with island size. In other words, there was a tendency of a greater genetic diversity for larger island-populations than for smaller island-populations. For example, only a single haplotype was found in each of the small islands of Gili Motang and Nusa Kode. In contrasts, there were six (6) haplotypes found on Rinca, four (4) on Komodo, and four (4) on Flores. The number of haplotypes found in a sampling location ranged from 1-6 (mean: 2.692 ± 1.3156). To evaluate the degree of genetic diversity at three population levels, i.e. sampling location, island, and region, I used two population genetic parameters. First, haplotype diversity (*h*) is a diversity measure that I applied to assess the variation in mtDNA CRI sequences without ignoring the differences in sequence length. I computed *h* for all sampling locations as a measure of genetic diversity at a “population” level. I would use the word “population” to refer to the smallest unit in the hierarchy, i.e. sampling location, unless indicated otherwise. Further, *h*s at population level were pooled together to determine the degree of genetic diversity at island and region levels. Second, nucleotide diversity (π_n) was

computed as a measure of genetic diversity using the information from variable nucleotide sites. I also computed π_n at three hierarchical levels, i.e. population, island, and region.

The genetic diversity among islands seems to be the greatest on Flores (h= 0.6424 ± 0.0392, π_n= 0.0040 ± 0.0023). Four haplotypes were present in Flores and two of these haplotypes were also common in Rinca, the neighbouring island located at a minimum distance of ~0.30 km (Figure 15). The haplotypes H5 and H6 were found in Flores West and H5 was also found on the small island of Gili Motang, which is located at a nearest distance of ~2.50 km from the western coast of Flores. The other two haplotypes, i.e. H10 and H11, were distributed exclusively on Flores and were the most divergent among all the other haplotypes in this study (Matrix 1). In addition, the population on Komodo was fairly diverse (h= 0.4999 ± 0.0432, π_n= 0.0008 ± 0.0007). Komodo was also the most distinct among all the other island populations, with only one of its four haplotypes was found on another island. The haplotype H3 was shared with Loh Buaya, on Rinca, and was represented by one individual. Moreover, the population in Loh Lawi was the most genetically diverse among all the thirteen populations sampled for this study (h= 0.6111 ± 0.0596, π_n= 0.0011 ± 0.0009).

A contrasting level of genetic diversity to that calculated for Flores and Komodo comes from the populations on the small islands. A zero genetic diversity was found in the populations on Gili Motang and Nusa Kode, due to the lack of sequence variations both in nucleotide substitutions and sequence length. This deficiency in genetic diversity may suggest a common phenomenon of a reduced genetic variability in small isolated populations on islands, such as in the case of the Black-footed Rock-wallaby in Barrow Island, Western Australia (Eldridge *et al.*, 1999) and the Common shrew in the Scottish Islands (White and Searle, 2006), but see e.g. Salgueiro *et. al*. (2004) and Lawrence *et. al*. (2008) for volant dispersers. The haplotypes H5 and H7 found on Gili Motang and Nusa Kode were shared with the neighbouring islands of Flores and Rinca, respectively. Both haplotypes were common on Flores and Rinca, which may habe been the source populations for the small islands. On the contrary, the genetic diversity of Rinca was relatively high (h= 0.5208 ± 0.0330) and was comparable to those for Komodo and Flores. Additionally, five (5) of the six (6) haplotypes found in Rinca were variable in their sequence length. However, the genetic diversity measured in terms of nucleotide diversity for Rinca seemed to be very low (π_n= 0.00005 ± 0.0002). Similarly, the nucleotide diversity in the Flores North population seemed to be lacking, due to the absence of variable sites in the Eastern haplotypes.

In summary, the level of genetic diversity for island population seems to be higher in Flores, Komodo, and Rinca. On the contrary, the level of genetic diversity in Gili Motang and Nusa Kode is much lower. However, the population in the isolated part of Flores, i.e. Flores North seems to have almost no genetic diversity. Unfortunately, the sample size for Flores North was very small, i.e. six individual animals, that the current assessment on the level of genetic diversity in this population might have been hindered. By adding more samples for Flores North in the future, I expect to gain a different picture of the level of genetic diversity in this population.

4.3. Geographic distribution of haplotypes and relationships among lineages

Apparently, phylogeographic methods are useful to elucidate the mosaic pattern of haplotype distribution over space for the Komodo monitor. I observed a non-random pattern of haplotype distribution in three geographic regions, i.e. Western, Central, and Eastern. Each of these regions was distinct in their haplotype composition. The island of Komodo harboured all the Western haplotypes, i.e. H1, H2, H3, and H4, although haplotype H3 was also distributed on the northern part Rinca (Table 2 and Figure 9). The Central haplotypes, i.e. H5, H6, H7, H8, and H9 were distributed on Rinca, Nusa Kode, Gili Motang, as well as in Flores West, whereas the Eastern haplotypes H10 and H11 were distributed exclusively on Flores, especially on the northern coastal areas. Haplotype H10 was also found in a location on the western coast, i.e. Flores West. This mosaic pattern of haplotype distribution in three regions was tested for a population genetic structure. However, the distribution of haplotypes on five islands was also tested for a population genetic structure, because some haplotypes were available only on a particular island. For example, haplotypes H1, H2, and H4 were available only on Komodo, while haplotypes H10 and H11 were distributed only on Flores. Thus, a population genetic structure was tested based on the observed mosaic pattern of three regions, as well as the limited distribution of some haplotypes on a particular island. It is interesting to note, that most of the haplotypes found on Komodo (75%) were not shared with other populations, whereas Rinca and Flores shared four (~66%) and two (50%) of their haplotypes and with other islands.

Having known that there was a pattern of haplotype distribution over a geographic space, I examined the phylogenetic relationships of the eleven haplotypes distributed across islands. The relationships among these haplotypes were inferred from three consensus topologies resulted from three phylogenetic analyses. The branching pattern of these

consensus trees is generally very similar. All the Komodo monitor haplotypes were nested in a clade with an absolute node support and the trees were rooted with the Lace monitor sequences. The Eastern haplotypes branched out of the clade for the Komodo monitor and all the other nine (9) haplotypes were grouped in the other larger clade, suggesting a larger genetic divergence between the Eastern lineages and those distributed in the Central and Western regions. However, a monophyly could not be resolved for the Central and Western haplotype groups. The MP and ML consensus trees showed a state of polytomy, and the node support for the Western clade in the Bayesian consensus tree was considerably low. In particular, the phylogenetic relationships among the Western and Central haplotypes appear to be closer, because most of these haplotypes differ by relatively few nucleotides and the Western haplotype H4 differ to the Central haplotype H7 only in their sequence lengths (Figure 8a, Matrix 1). In short, these results showed that the Central and Western haplotypes are relatively more closely related, while a more distant relationship was demonstrated between the Eastern group and both the Western and Central groups.

4.4. Population structures inferred from CRI sequence variations

I assessed two alternatives for the genetic structure of the entire Komodo monitor population using AMOVA. All results of genetic structure analyses were calculated solely based on the pairwise nucleotide differences. I used pairwise nucleotide differences in the current assessment of population genetic structure, because the variation in the nucleotide sites is consistent with the phylogenetic methods I used to analyse the relationships among haplotypes. Moreover, the use of variable nucleotide sites is a major source of genetic diversity estimations. The variance components from sequence data were generated based on the sums of squared pairwise nucleotide differences in AMOVA, although the variation in sequence length may be also explained. In AMOVA, the variation in the sequence length may be coded the absence or presence of each of the four repeat units.

The AMOVA results showed that a population subdivision of three regions was better resolved than a subdivision of five islands. Indeed, this resolution was indicated by the non-random distribution of haplotypes in the Western, Central, and Eastern regions. The major genetic structure in the entire population of Komodo monitor with three-region subdivision was explained by 89.14% of the total genetic variation. To a lesser extent, a genetic structure is also reflected from the distribution of maternal lineages on five islands. This genetic structure was explained by the variance components among islands by 53.81%.

An island may become a subdivision in the entire population, because the haplotype composition across islands was apparently distinctive. Despite the possible dispersal events among these islands, which I inferred from the haplotypes shared among islands (Figure 9), some islands seemed to retain some particular haplotypes. For example, Komodo retains 75% of its haplotypes and therefore, its composition of four haplotypes is different from that on the other islands. Based on this data, one may expect that the population on this island is one of the subdivisions in the entire population. Indeed, Komodo by itself is a separate region, i.e. Western, or a subdivision in the major genetic structure besides the other two regions (φ_{CT}= 0.891, P< 0.01). Additionally, a less significant degree of genetic divergence among islands was indicated (φ_{CT}= 0.538, P< 0.01), which renders Komodo a distinctive island population.

Given the genetic structure of three regions, Flores North was the only location in the Eastern region. The population on Flores was assigned into two regions, i.e. Flores North in the Eastern region and Flores West in the Central region. The level of differentiation among regions indicated a significantly large divergence between the Eastern and Central regions (φ_{ST}= 0.989, P= 0.000), and the nearest geographic distance between these regions, i.e. between Flores North and Flores West was more than 100 km (Matrix 5). The haplotype H10 was shared between these two distantly separated populations on Flores, and this may be explained by either a fragmentation of continuous habitat or a migration event in the recent past. A habitat fragmentation seems to occur along the northern and western coasts, possibly due to the growing number of human settlements in this area since the mid-1980s. The prevailing government-sponsored transmigration and voluntary migrations between the mid-1980s and mid-1990s in Indonesia have probably initiated the movement of people to Flores (Tirtosudarmo, 2009), resulting in the growing size of human population on this island. The occurrence of the Komodo monitor on the northwestern coasts of Flores (Figure 4) was reported in the previous field observation on the behavioural ecology of the Komodo monitor (Auffenberg, 1981). This population seems to disappear recently at around the same time as the population density in Wae Wuul Reserve on the northwestern is reduced due to the increasing areas for agriculture and human settlements in this area (Ciofi, 2002). The Eastern haplotypes would be commonly distributed along the northern and western coasts, if a continuous population had occurred in this area. Thus, the distribution of haplotype H10 in Flores North and Flores West might indicate a remaining habitat patch for the Komodo monitor on Flores. Alternatively, an event of migration to a suitable habitat with gene flow

may explain the presence of haplotype H10 in the two populations on Flores. A migration could occur between Flores North and Flores West via an inland route, which may be assumed from fossils. The remains of the Komodo monitor were discovered in Liang Bua cave, which is situated inland in the western part of Flores (van den Bergh *et al.*, 2007). In addition, a few other fossils of the Komodo monitor were recovered from a palaeontological site in Tangi Talo, in the central part of Flores (Morwood *et al.*, 1998). The locations of these fossils may suggest an inland route of a migration event between the areas in the western coasts and those near the central part of the island. Nevertheless, these fossils may indicate a continuous distribution of the Komodo monitor from the western part through to the central part of Flores. A deduction for a habitat fragmentation or an event of migration with gene flow currently cannot be made without ambiguity, because of the absence of samples from the northwestern coasts. By sampling in these areas, one may be able to discover a haplotype that demonstrates the link between the populations in Flores North and Flores West.

Although Komodo, which lies in the Western region, seems to be the most distinctive region within the major genetic structure for the Komodo monitor, the Central and Eastern regions seem to be as unique. The level of genetic differentiation between Komodo and the other island populations were the largest (range φ_{ST}= 0.580 - 0.860), and a large proportion (75%) of the haplotypes on Komodo were not shared with other islands. These results are corroborated by population genetic data, which demonstrate the high distinction of Komodo among all populations (range F_{ST}= 0.324- 0.557, P< 0.01) and more than 20% of the alleles found in Komodo were not shared with other populations (Ciofi and Bruford, 1999). It is interesting to note, that Komodo was highly differentiated from Rinca, the nearest island located in the Central region (φ_{ST}= 0.860, P= 0.000). On the other hand, Komodo was less differentiated from Flores (φ_{ST}= 0.580, P= 0.000), which is situated farther than Rinca. In the present geography, Komodo and Rinca are separated by seawater at a nearest distance of ~5 km, while the minimum distance between Komodo and Flores is about five times greater (Figure 15). A migration is likely to occur between Komodo and Rinca, although it may be rare. Komodo and Rinca shared haplotype H3, which was found in one individual in northern Rinca (Table 2, Figure 9). Thus, the significant level of differentiation between Komodo and Rinca may indicate a very limited molecular similarity, which might have resulted from a long period of isolation with a rare migration between the two islands. Similarly, a long period of isolation with a rare event of migration between Komodo and Flores may explain

the extent of differentiation between these islands. However, Komodo and Flores does not seem to share any haplotype, which may be used to infer an event of migration between the two islands. Again, I should probably emphasize a broader sampling in the northwestern coasts of Flores, where a shared haplotype between Komodo and Flores might be found. In addition, a lower scale of genetic differentiation between Rinca and Flores (φ_{ST}= 0.278, $P \leq$ 0.05) seems to make sense. A relatively narrow strait separates these islands at a minimum distance of ~0.30 km (Figure 15), which may facilitate migrations. I discuss the level of genetic divergence and the pattern of migrations among islands in more details in Section 4.5.

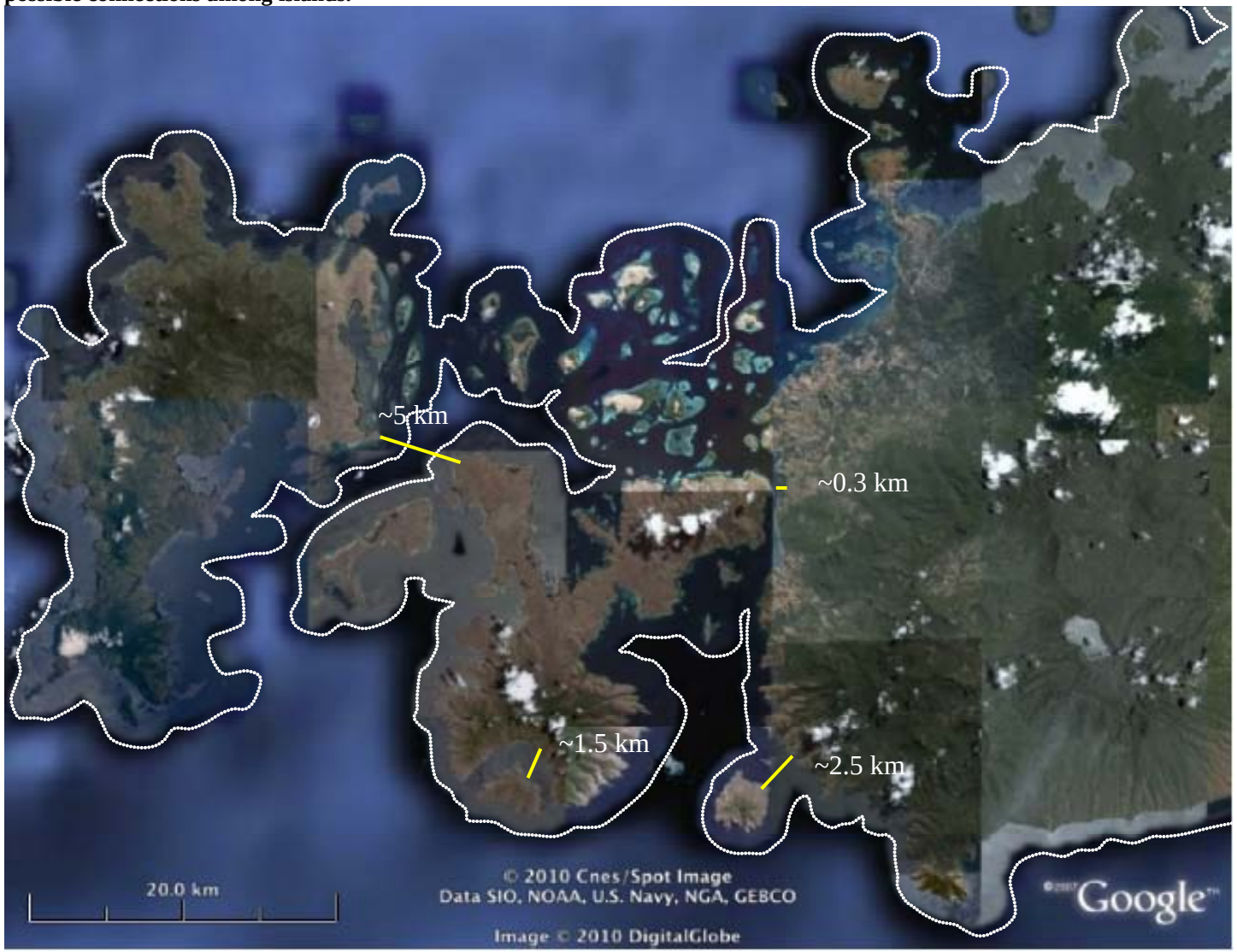
In spite of the high level of genetic divergence between Komodo and the eastern islands, the Central and Eastern regions also showed a high level of distinctive genetic characteristics. Only one haplotype (~14%) of all those, i.e. seven (7) haplotypes distributed in the Central region was shared with Komodo, and three (~43%) were shared with Flores. However, only one (~14%), i.e haplotype H10 was shared with the Eastern region. On the other hand, the Eastern region shared 50% of its haplotypes with the Central region and none was shared with Komodo. The larger islands, i. e. Komodo, Rinca, and Flores are important sources of genetic diversity, as indicated by the proportion of genetic diversity they contribute in the total molecular variation. However, the two small island populations in the Central region were genetically impoverished, which can be determined by the lack of genetic diversity (Table 3) and unique haplotypes (Figure 9). The depletion of genetic variation in these small island populations is corroborated by the low allelic diversity and the fixation of alleles that are inferred from microsatellite loci for Gili Motang (Ciofi and Bruford, 1999). These results on the relatively low genetic diversity on small island populations are consistent with the susceptibility of small island populations to the effects of genetic drift (Frankham, 1996b).

The lack of genetic diversity and the absence of unique haplotype in the small islands were also reflected in the partitioning of molecular variation in five islands (Table 4). About a third (31.43%) of the total variation was due to the genetic partitioning among populations within islands, suggesting a significant genetic diversity within islands. Evidently, the three larger islands harbour two to four (2-4) unique haplotypes, which contributed as a proportion in the total genetic variation. Thus, the seemingly high genetic differentiation across all islands is demonstrated only for the larger islands. Similarly, the level of genetic differentiation among populations within islands also seems to be high (φ_{SC}= 0.680, P< 0.05).

However, this high level of molecular variation within islands was apparently calculated only from the differences for the larger islands, i.e. Komodo, Rinca, and Flores, because there was a zero genetic diversity in the small island populations. I assume that the extent of molecular variation and genetic differentiation at within-island level should also increase, should the degree of genetic diversity in the small island populations be higher than currently observed. Furthermore, all haplotypes on the small islands in the Central region were shared with at least one larger island. Nusa Kode shared haplotype H7 with all the four populations on Rinca, while Gili Motang shared haplotype H5 with all populations in the Central region, but Nusa Kode. Because of the lower genetic diversity in the small island populations and the presence of two haplotypes commonly found in the larger neighbouring islands, i.e. Rinca and Flores, the populations on Gili Motang and Nusa Kode might have been colonised from these larger islands and therefore, they can be regarded as a part of the larger subdivision, i.e. the Central region. Indeed, a relatively little variation, i.e. ~1.2% was partitioned among populations within regions (Table 4), which reflected a high similarity among haplotypes distributed in a region, including those in the Central region.

To summarize, the Komodo monitor population is strongly structured in three regions. A less strong genetic structure in islands emerged from the distinctive haplotype composition in the three larger islands. The haplotype network illustrates the interconnections among populations across islands that concord with the structure in three regions. All the populations in the Western and Central regions were connected and the connection was depicted as a hypothetical haplotype, which was not recorded in my data set. The Western haplotype H4 is one of the three private haplotypes in Komodo that seems to be another intermediate haplotype connecting the Western and Central regions. A single individual represented haplotype H4, i.e. sequence ID LSE0709 sampled in Loh Sebita, in the northeastern Komodo. On the other hand, the Eastern region and the Western and Central regions seems to disconnect, suggesting a considerable divergence between the population in the easternmost of the range and all those to its west. This level of divergence was equally demonstrated in the MP tree topology, where both the Eastern haplotypes formed a separate clade from the Central and Western haplotypes (Figure 11). The closer connection between the Western and Central populations was represented by a one-step mutation, whereas the disconnection of the Eastern haplotype group from the others demonstrated the divergence between the population in the Eastern region and those in the Western and Central regions.

Figure 15. Satellite imagery of islands in the Komodo National Park and western coast of Flores, showing relative positions of islands to one another. The approximate minimum distances between islands are indicated. Dashed lines indicate submerged contour of islands outlined on satellite imagery, showing possible connections among islands.



4.5. Historical dispersal and vicariance

A hypothesis to explain the presence of haplotype H3 on Komodo and Rinca is an event of dispersal between these islands during the Pleistocene, because a recent dispersal between Komodo and Rinca seems to be highly unlikely. Currently, the shortest sea distance between Komodo and Rinca is ~5 km (Figure 15). An adult Komodo monitor can travel in average for less than 1 km per day on land, although a male individual may travel up to 5 km in a day (Ciofi *et al.*, 2007). Komodo monitors also tend to stay on land, despite their ability to swim in saltwater. A previous observation reported that the monitors were able to swim for a limited distance of about 2 km or less in saltwater (Burden, 1928). These monitors were released into the sea from a boat and they were inclined to go back to the island where they were taken from. Thus, a dispersal event involving a travel through a long sea distance may be impractical for the Komodo monitor. However, this terrestrial monitor lizard seems to have dispersed across the Lesser Sundas, most likely during the Pleistocene. The islands in the Lesser Sundas were interconnected during three periods of continental glaciation since 250,000 years ago. The sea level in this region was low and shallow shelfs were exposed, creating land bridges among islands. Komodo and the eastern islands were connected at 75-120 m below the present sea level for ~60,000 years (Voris, 2000). It is possible, that this connection between Komodo and the islands eastward has facilitated dispersals. Figure 15 shows a hypothetical land bridge between Komodo, Rinca, and Flores, which I drew from the shallower parts around the islets north of Rinca. An event of dispersal between Komodo and Rinca may have occurred through this land bridge. Indeed, haplotype H3 was found in the northern parts of Komodo and Rinca (Figure 9). Addtionally, Statistical Parsimony also suggested a link between the two island populations through a rare haplotype, i.e. haplotype H4/Western 4, which was found in the northeastern Komodo (Figure 13). However, the time of dispersal cannot be inferred without an estimate of divergence time between the two island populations. The populations on Komodo and Rinca are estimated to separate since 44,000 years ago, after a dispersal of a female individual with haplotype H3. On the other hand, the two island populations may have been separated for a longer time, possibly since a deglaciation period of 220,000 years ago (Table 5). All these estimates of divergence time for Komodo and Rinca indicated a time frame in the Late Pleistocene, and more specifically between two glaciation periods, i.e. between 250,000-135,000 years ago and between 135,000-18,000 years ago (Voris, 2000). Therefore, a female individual with haplotype H3

might have migrated from Komodo to Rinca sometime ago during a glaciation period, when the land bridge to the north of Rinca was present. These estimates of divergence time should be taken as preliminary results, because they were calculated simply from the mean uncorrected pairwise *p* distances. A further analysis of divergence time among populations is yet to be performed by incorporating some population parameters such as the female effective population size. Additionally, a sample from the northwestern peninsula of Rinca is essential to corroborate the putative connection between this location and the northeastern part of Komodo.

Although a dispersal theory may explain the presence of haplotype H3 on Komodo and Rinca, a different biogeographic scenario should be imposed to explain the segregation between lineages found on these islands. A state of population vicariance may have occurred following glaciation periods, when the sea level rose and the land bridge to the north of Rinca was inundated. Populations on Komodo and Rinca were separated by seawater during this time and lineages distributed on these islands started to diverge in isolation. I estimated the segregation between lineages distributed on Komodo and Rinca to have occurred during a period of deglaciation, probably since 220,000 or 176,000 years ago. On the other hand, the divergence between Komodo and Rinca may have occurred more recently, since 55,000 or 44,000 years ago. These estimates of divergence time were based on the assumption of different evolutionary rates for mtDNA Control Regions, which may be as slow as 2% and as fast as 10% (Table 5). A similar event of vicariance seems to affect the small islands, which apparently are isolated from their neighbouring islands. However, the degree of isolation for Nusa Kode and Gili Motang is much less than the isolation of Komodo from Rinca. The geographic distance between these small islands and the larger islands in the Central region is relatively short (Figure 15). Both of these small islands shared their only haplotype with the neighbouring larger islands, suggesting an origin of the small island populations from the larger islands. Gili Motang shared a haplotype with the nearest population in Flores West, and the nearest distance between this small island and the southwestern coast of Flores is approximately 2.5 km. Likewise, Nusa Kode shared the same haplotype with the nearest population in Loh Dasami, which lies on Rinca about 1.5 km to its north (Figure 15). One may expect, that the sea distance between Gili Motang and southwestern Flores is a barrier to dispersal, although occasional crossings may also be expected. I estimated the divergence of maternal lineages between Gili Motang and Flores West to occur since at least 79,000 years

ago (Table 5). This result seems to be in contrast with a migration rate estimate based on the genotypic data from microsatellite loci for Gili Motang (Ciofi and Bruford, 1999). Gili Motang is expected to receive one female or male migrant every two generations, i.e. approximately equals to 24 years. Nonetheless, Gili Motang is greatly differentiated (F_{ST} = 0.225) from the population in the southwestern Flores (Ciofi, 2002), suggesting a relatively rare migration into this small island population. On the other hand, the population on Nusa Kode seems to maintain a regular gene flow from Rinca. The times estimated for the divergence of maternal lineages distributed on Nusa Kode and Rinca were more recent than those for the separation between the maternal lineages found in Gili Motang and Flores West. I estimated the divergence time for Nusa Kode and Rinca of 9,000- 45,000 years ago (Table 5). These time estimates far predate thosefor the separation between Gili Motang and Flores West. Furthermore, the level of genotypic differentiation between Nusa Kode and Loh Dasami was very low (F_{ST} = 0.0001), suggesting a homogenous allele frequency in the two populations due to gene flow (Ciofi, 2002). Therefore, the degree of isolation for Gili Motang is evidently higher than that for Nusa Kode. In addition, the higher degree of isolation for Gili Motang can be related to the mitochondrial and nuclear genetic diversity impoverishment in this small island population, whereas the less isolated population in Nusa Kode seems to maintain its nuclear genetic diversity by a gene flow from Loh Dasami in southern Rinca despite the low mitochondrial genetic diversity.

It was not surprising to observe, that the most frequent dispersal events among islands seemed to occur between Rinca and Flores. The number of haplotypes shared between Rinca and Flores was the double of that between the other islands. Two haplotypes were shared between Rinca and Flores, i.e. haplotypes H5 and H6, whereas only one haplotype was shared between the other pairs of islands (Table 2, Figure 9). It may be hypothesized, that the close proximity of Rinca to Flores may facilitate migrations between them. The nearest distance between the two islands is a sea channel of ~0.30 km. Thus, it is possible that Komodo monitors are able to cross this sea distance by swimming or floating. Indeed, an estimate of gene flow based on a nuclear data set suggested a relatively frequent migration between Rinca and Flores. One female or male individual was expected to migrate every two to fifteen (2-15) years between Rinca and Flores (Ciofi and Bruford, 1999). Given this estimate of migration frequency as well as the relatively short sea distance between Rinca and Flores, an over-water crossing seems to be possible for Komodo monitors. This inference

of dispersal between Rinca and Flores is on the contrary to that between Rinca and Komodo. As I discussed earlier in this section, the presence of a land bridge to the north of Rinca is postulated for Rinca and Komodo because of the relatively large sea distance between these islands, i.e minimum of ~5 km. Moreover, two strong evidences corroborate the hypothesis of the Komodo monitor dispersal through the northern land bridge in the Late Pleistocene, i.e. the distribution of haplotype H3 and the presence of the intermediate haplotype H4 in the northern parts of these islands. Compared to the expected frequency of migration between Rinca and Flores, a relatively rare migration was estimated to occur between Rinca and Komodo. One individual was estimated to disperse between Rinca and Komodo every sixty seven to one hundred and ninety (67-190) years (Ciofi and Bruford, 1999).

It is interesting to note, that seawater seems to be a barrier to dispersal for the Komodo monitor. Nevertheless, the sea distance is more likely of a barrier to dispersal for this terrestrial varanid lizard than the presence of seawater *per se*. Indeed, the results from the Mantel's test showed a significant influence of geographic distance in the distribution of maternal lineages across sampling locations. The correlation between geographic and genetic distance matrices was significantly positive ($r= 0.777$, $P< 0.001$, $y= 0.603x + 0.844$), suggesting a pattern of Isolation By Distance (IBD). This pattern of population structure in my study may be interpreted as the mosaic arrangement of haplotypes in two or more isolated locations, which are given rise by a geographic distance over land, sea, or both. The Mantel's test significant correlation may indicate the role of dispersal and vicariance in distributing maternal lineages across islands. On one hand, dispersal homogenises the molecular variations among islands, while on the other, vicariance operates to enhance the total molecular variation across all the five island populations. Thus, the opposing forces of dispersal and vicariance seem to have shaped the current pattern of haplotype distribution across island populations. Apparently, the level of isolation on islands for the terrestrial Komodo monitor is determined by the repeated seawater fluctuations in the Pleistocene (Table 5). The sea distance between islands was reduced during the periods of low sea level, whereas the high sea levels increased the distance of sea between islands. In the case of shallow shelves with a bathymetry of less than 200 m such as the shelf between the islands of Komodo and Flores (Sprintall *et al.*, 2003), islands were interconnected during the low sea level and separated in isolation during the high sea level. This interconnection among islands in the Lesser Sunda (Figure 15) is likely to have facilitated migrations for the Komodo

monitors, while the increasing sea distance between these islands seems to generate a barrier to their readily limited movement (Ciofi *et al.*, 2007). These Pleistocene dispersal and vicariance scenarios are evidently substantiated by the observed IBD pattern. In particular, the extent of isolation on islands seems to become apparent beyond a certain threshold of sea distance, which may be close to about two (2) kilometres. This figure is given based on the approximate maximum sea distance swam by the monitors in saltwater (Burden, 1928), as well as the greater level of isolation for Gili Motang from the nearest population on the southwestern coast of Flores of ~2.5 km sea distance. Nonetheless, it is also important to note that the land distance between two populations may equally contribute in the distribution of maternal lineages across the distribution range. For example, the relatively great distance (>100 km) between Flores North and Flores West may influence the distribution of a haplotype from either location in the other.

The current pattern of lineage distribution across the whole population of the Komodo monitor in the Lesser Sunda Islands cannot be explained either by dispersal or by vicariance alone. A vicariance scenario has been generally accepted to explain the distribution of fauna in Wallacea, for example the distribution of some fauna in Sulawesi, including the grasshoppers of the genus *Chitaura* (Butlin *et al.*, 1998), as well as some bufonid toads and also some monkeys of the genus *Macaca* (Evans *et al.*, 2003). On the other hand, dispersal seems to have played an equally important role in the distribution of some other taxa in Wallacea such as birds (Jønsson *et al.*, 2004). Some studies have indicated that neither dispersal nor vicariance can effectively explain the pattern of faunal distribution in Wallacea (Steppan *et al.*, 2003), due to the complex patterns of biological and geological evolutions. The faunal distribution across Wallacea seems to be intertwined with some geological processes that shaped its geography, especially during the Pleistocene. Some of the islands in this region are of continental fragment in origin and might have shifted to their current positions in Wallacea while transporting continental biotas. On the other hand, some other islands emerged as a result of volcanic eruptions (Hall, 1998, 2001). Consequently, the existing biotas on these volcanic islands might have evolved from their ancestors that first colonised these islands. In the Komodo monitor, the vicariance scenario in the Pleistocene seems to be consistent with the divergence and separation time between island populations, e.g. between Komodo and Rinca, and between Komodo and Flores. On the other hand, an event of dispersal can explain the presence of a rare sample of a Western haplotype in the

Central region. Further, an inter-island dispersal between Komodo and Rinca is postulated, for which a land connection between these islands is necessary. Not only was there a land bridge between Komodo and Rinca, a larger land connection (Figure 15) was likely to emerge between Komodo and Flores during glaciations. This land bridge connecting Komodo, Rinca, and Flores seems to have facilitated the dispersal between Komodo and the eastern islands. Unfortunately, the lack of samples from the northwestern coast of Flores has hindered an inference of a direct dispersal between Komodo and Flores to be made. Similarly, the possibility of dispersal between populations on the northern and western coasts of Flores could not be assessed without samples from this area. Additionally, a short-distance dispersal over the sea may explain the distribution of haplotypes on some islands, e.g. the dispersal between southern Rinca and Nusa Kode.

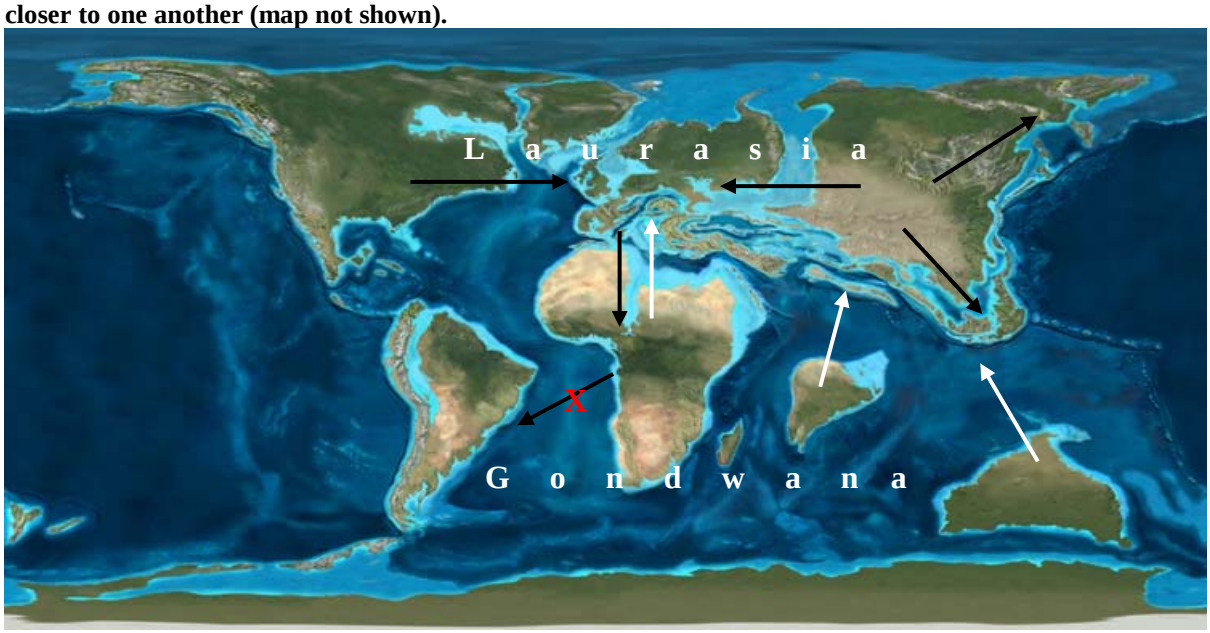
The time estimates for the dispersal and vicariance among island populations of the Komodo monitor seem to be relatively recent, i.e. in the Late Pleistocene. Albeit raw, these estimates based on the divergence among the geographically grouped maternal lineages are apparently consistent with the evolutionary history of this species in Australasia. I estimated the time since the divergence between the Eastern and Central-Western haplotype groups to have occurred since about 1.2 Ma. In particular, the Western and Central haplotype groups were estimated to emerge about 500,000 years ago. Extending this estimate to the root of the Bayesian consensus tree, I found that the Komodo monitor has evolved in the Lesser Sunda Islands quite recently. The Komodo monitor might have split from its living sister species, i.e. the Lace monitor, only since about 4.5 Ma. This estimate is slightly older than the age estimate for the fossil Komodo monitor from Australia dated to about 3.8 million years ago (Hocknull *et al.*, 2009). Thus, the ancestral Komodo monitor might have arrived in the Lesser Sunda during the mid-Pleistocene by a rare dispersal from Australia.

4.6. Biogeographic hypotheses on the origin of the Komodo monitor

A hypothesis on the origin of the Komodo monitor in Australia may be linked to the Gondwanan and Laurasian theories of monitor lizard radiation. The hypothesis of an ancestral Komodo monitor dispersal from Australia is consistent with the Gondwanan source of monitor lizard radiation, and may be called as an “east-to-west” theory. On the other hand, the ancestor of the Komodo monitor might have come from Asia. This “west-to-east” dispersal scenario is concordant with the Laurasian theory of monitor lizard origin in Asia and the subsequent distribution throughout Australasia. Figure 16 illustrates the competing

hypotheses on the origin and dispersal routes of monitor lizards. The map in Figure 16 shows the relative positions of Laurasian and Gondwanan continental fragments, during the transition time of the Late Cretaceous to the Tertiary period (Blakey, 2010).

Figure 16. Schematic illustration of the hypothetical routes of varanid dispersal: Laurasian and Gondwanan theories set two different evolutionary time frames of varanid diversification. Map shows the geography during the Late Cretaceous-Tertiary periods (~65 Ma), a time frame for the onset of varanid dispersal based on the Laurasian theory (black arrows). The Gondwanan dispersal (white arrows) sets an earlier origin of varanids to ~150 Ma, when fragments of this supercontinent were situated relatively closer to one another (map not shown).



Several lines of evidence indicate an origin of the Komodo monitor in Australia. The fossil records of the Komodo monitor from Australia (Hocknull *et al.*, 2009) and from the island of Timor (Hooijer, 1972) evidently indicate the distribution of this species as far as Australia in the recent past. A closer phylogenetic relationship between the Komodo monitor and varanid lizards from Australia corroborates the fossil record, suggesting the emergence of the ancestral Komodo monitor in the continental Australia. Indeed, the species diversity of the extant monitor lizards is the greatest in Australia, which indicates a diversification of monitor lizards, including the Komodo monitor, on this Gondwanan fragment. The scenario of an Australian origin for the Komodo monitor becomes more plausible by the presence of some Australian elements in the faunal composition of the Lesser Sundas (Michaux, 1994). In contrast to this scenario, the Komodo monitor might have evolved in the Lesser Sunda from an Asian ancestor, and subsequently dispersed to Australia. However, only one fossil Komodo monitor is so far reported from Sundaland. A fossil vertebrae of *Varanus bolkayi* from Trinil, East Java, is thought to be a subspecies of the Komodo monitor, implying the

westernmost occurrence of this species worldwide (Hooijer, 1972). Furthermore, there is no strong evidence to support this hypothesis of the Komodo monitor dispersal from Java. The Lombok Strait is a major barrier to dispersal from Java to the Lesser Sundas for most vertebrate fauna, including the Komodo monitor. This strait that lies between the islands of Bali and Lombok is essentially the southern part of the Wallace's line (Figure 2), through which only a few vertebrates are known to have successfully crossed to colonise the Lesser Sundas. A species of human, *Homo erectus*, was reported to be able to use a watercraft to cross this deep channel (Morwood *et al.*, 1998), whereas an ancestral form of a Pygmy elephant species, *Stegodon florensis*, might have used the islands of Lombok and Sumbawa as stepping stones for its dispersal from Java to Flores (van den Bergh *et al.*, 2007).

The best scenario for the origin of the Komodo monitor seems to be consistent with the Gondwanan theory of monitor lizard radiation from Australia. It is likely, that the ancestral Komodo monitor came from Australia. The dispersal from Australia might have taken place during the past one million years ago, when the geographic and climatic changes in Australasia were generally more influential in the evolution of fauna in this region (Hall, 1998). The divergence time between the Komodo monitor and its Australian sister species, the Lace monitor, was estimated to have occurred in the Pliocene (~4.6 Ma) in this study. However, a migration of the Komodo monitor from Australia to the Lesser Sundas might have occurred only much later after speciation. This view is corroborated by the independent estimates of time since the separation of the Komodo monitor lineages grouped in islands. These estimates indicate a fairly recent time of approximately less than one million years ago since the separation of the most divergent lineage groups in the Eastern and Western regions (Table 5). During the past one million years, the sea level was fluctuating and the geography and climate in Australasia were changed multiple times. The lowered sea level during this period of time has facilitated a migration of the ancestral Komodo monitor from Australia to the Lesser Sunda Islands, via Timor. Because the sea level was low, the sea distance between Australia and Timor was shortened by the exposure of the northwestern edge of the Australian plate. Nevertheless, Timor and Australia have never been connected by a land bridge due to depth of the Timor Trough. This deep channel is formed between Timor and the northwestern part of the Australian continental plate since as early as the Late Miocene (~10 Ma) with the depth range of 1,500-3,200 m (Veevers, 1971). Similarly, the islands of Bali and Lombok have never been connected by a land bridge, due to the depth of the Lombok

Strait (Voris, 2000). This strait lies between the islands of Bali and Lombok with a typical depth of 800-1,000 m and a north-to-south sea current (Murray *et al.*, 1990). Thus, the Timor Trough is a comparable barrier to the Lombok Strait for the dispersal of the Komodo monitor to the Lesser Sundas.

The theory of the ancestral Komodo monitor migration from Australia cannot fully explain the event of dispersal to Timor without postulating a hypothetical sea-crossing event through the Timor Sea. The frequent cyclones commonly happened in Australasia during the continental glacial periods (Verstappen, 1980) are likely the events that have sent the Komodo monitors off the northwestern coasts of continental Australia to Timor. As this sea-crossing dispersal has to be postulated for the scenario of a migration from Australia, the same event of sea-crossing dispersal can also be assumed for the theory of dispersal from Sundaland to the Lesser Sundas. Nonetheless, this west-to-east dispersal seems to be unlikely, because the Komodo monitor is phylogenetically closer to both the extant and extinct Australian varanids than to Asian varanids. In addition, there is no reliable fossil evidence of the Komodo monitor so far from Sundaland. Therefore, the east-to-west theory for the Komodo monitor origin of dispersal seems to be more likely than the west-to-east scenario, which is consistent with the Laurasian theory of varanid lizard origin of dispersal.

In general, the dispersals of varanid lizards in the Lesser Sunda Islands seem to be multidirectional. The Lesser Sundas seems to be the meeting point of Asian and Australian varanids. Yet, the distribution and dispersal of varanid lizards in this region remain to be some of the challenging themes in vertebrate evolutionary history in Wallacea. Despite the notion of the influential geological processes in the distribution of fauna in this region, very little is known about the population divergence of varanid species in this biogeographic realm. The Asian water monitor, *Varanus salvator*, also occurs in the Lesser Sundas. The Water monitor is widely distributed in Sundaland and Wallacea, including a few small islands in the Flores Sea, in the north of Flores (E. Arida, personal observation). The Water monitor co-occurs with the Komodo monitor on Flores, while its distribution on Komodo and Rinca (C. Ciofi, personal communication) may still need to be confirmed in a more detailed survey. Given its aquatic nature, the Water monitor may have dispersed to the Lesser Sundas through a northern route from Sulawesi and/or through a western route from Bali. This species is known to occur in the Lesser Sundas as east as on Wetar (Figure 3), suggesting a reverse directionality of dispersal to that of the Komodo monitor. A study on the extent of

population divergence for the Water monitor in the Lesser Sundas should be conducted to better understand the multidirectionality of varanid dispersals in Australasia. Additionally, a survey on the distribution of two understudied species, i.e. *V. timorensis* and *V. auffenbergi* should also be conducted to complement the current knowledge on the diversity of varanid lizards in Wallacea. Finally, a comprehension on the species diversity and the evolutionary processes that give rise to the pattern of varanid distribution in Wallacea may signal for conservation awareness.

4.7. General discussion and implications for *in situ* conservation programmes

The phylogeographic pattern for the Komodo monitor populations was described using variations in the mtDNA Control Region I (CRI) sequences. The phylogenetic and population genetic analysis in this study were performed based on the nucleotide substitutions, although a different type of DNA polymorphism was observed in the data set. The sequence length variations are a source of DNA polymorphism that may be incorporated in phylogeographic analyses, pending more detailed data on the evolution of the mtDNA CRI sequences. In the current study, I applied three phylogenetic methods and conventionally treated all alignment gaps as missing data to construct a consensus genealogy. In other words, the variations in sequence length were ignored in the phylogenetic analyses. These variations were also ignored in the population genetic analysis to demonstrate a genetic partitioning among island and region populations. The Komodo monitor CRI sequences show three major length differences, which result from the presence of contiguous indels of three different sizes. Currently, it is impossible to code these multiple indels for phylogenetic analysis. The flanking sequence at the 5' end of my CRI sequences still remains to be validated, because two flanking sequences should be present to introduce an indel or an alignment gap (Simmons and Ochoterena, 2000). More importantly, the evolution of contiguous multiple indels in the current CRI sequences is yet inadequately understood for a phylogenetic application. The presence or absence of these indels is apparently an effect of the number of tandem repeats that varies among sequences. Neither the tandem repeats nor such contiguous multiple indels present in the Komodo monitor CRI were observed in the Lace monitor Control Region sequences. Therefore, a separate study on the evolution of sequence length variation in the mitochondrial genome of varanid lizards is necessary to determine its significance as a molecular character for phylogenetic analyses.

The results from phylogenetic analyses showed two monophyletic clades of eleven maternal lineages. A closer relationship between the Western and Central haplotypes was demonstrated in all the three tree topologies, while the Eastern haplotypes formed a separate monophyletic clade. The Western and Central haplotypes were distributed on Komodo and the islands in the Central region, respectively, whereas the Eastern haplotypes were found almost exclusively in a location in northern Flores. It is important to note, that the Eastern haplotyes were the most divergent among all haplotypes defined in this study. In other words, the number of pairwise differences between an Eastern haplotype and a given haplotype from either the Western or Central group was among the largest of all pairwise comparisons. Indeed, all the three consensus trees showed a high node support for the split between the Eastern and the Central-Western clades, suggesting a monophyly for both clades and a significantly large genetic distance between them. Further, a population genetic analysis of AMOVA showed two different patterns of molecular variation partitioning across the whole population. A strong partition was shown among three separate regions of distinct haplotype diversity. On the other hand, a less strong partition was also demonstrated among five island populations. Thus, the observed mosaic distribution of the eleven maternal lineages was concordant with a subdivision of the whole population in the Western, Central, and Eastern regions. Although the larger portion of the total molecular variation was accounted for by a subdivision of the entire population into three regions, the distribution and diversity of maternal lineages on each island was unique. For example, the island of Komodo essentially represented the Western region.

The pattern of haplotype distribution and genetic divergence among the three regions may be related to the sea level fluctuations in the Late Pleistocene. The molecular clock estimates demonstrated the approximate times since the divergence of haplotypes grouped in islands, and these divergence time estimates coincided with the timing of the Late Pleistocene sea level fluctuations. Assuming no convergent evolution, the co-occurrence of a haplotype on two or more islands suggested a historical movement among islands that were presumably facilitated by a land bridge that repeatedly emerged since the past 250,000 years in the Late Pleistocene. The Komodo monitor does not seem to be a very adept swimmer, especially for a long distance. Therefore, a land bridge between two relatively distant islands should be postulated to explain the observed pattern of haplotype distribution. In the Pleistocene, islands in the Lesser Sundas were possibly interconnected by land bridges during the periods

of low water. The exact depth of the channel between the islands in the Komodo National Park and Flores seems to be currently unavailable, but the estimated bathymetry of 200 m or less in this region (Sprintall *et al.*, 2003) suggests a possible land connection among these islands during the periods of Pleistocene low sea level. These island connections are probably similar to those among the islands in Sundaland, which occurred as the shallow shelves of about 200 m were exposed by the lowered sea level (Voris, 2000). In addition, the historical dispersal among islands possibly also occurred without the presence of a land bridge; however, this dispersal might occur as the sea distance between these islands was reduced during the lowered sea level associated with glaciation periods.

Given the relationships among haplotypes and population structure of the Komodo monitor, an assessment of the level of genetic diversity may bring an insight onto the characteristic details across the populations in the Lesser Sunda Islands. The level of genetic diversity among island populations seems to follow a trend. Although there was a discrepancy between the two measures of genetic diversity used in this study, the larger island populations tend to show a larger genetic diversity than the small island populations do. Due to the length variations for some sequences, haplotype diversity (h) and nucleotide diversity (π_n) measures seem to show a different level of genetic diversity. The island population on Flores as a whole, showed a relatively high level of genetic diversity. Haplotype diversity (h= 0.6424 ± 0.0392) was the highest in Flores, and the average nucleotide substitutions per site among these haplotypes (π_n= 0.40 ± 0.0023) were also the greatest. Indeed, the pairwise differences among the haplotypes distributed on Flores were the greates among all pairwise comparisons (Matrix 1). Flores is the largest island within the current distribution range of the Komodo monitor, where habitat degradation associated with growing human settlements, unsustainable practices of agriculture, and wildlife poaching may become a serious threat to the survival of the Komodo monitor on this island (Ciofi, 2002). Two nature reserves, i.e. Wae Wuul in the western coast and Wolo Tado in the northern coast have been established to protect the existing populations on Flores. However, a stronger protection seems to be necessarily enforced to maintain the relatively high genetic diversity on this island population. More importantly, the stronger protection for Flores is needed to counterbalance the declining population on this island. A relatively recent study indicated the absence of monitors in the southern and northwestern coasts of Flores (Ciofi and De Boer, 2004), where a population was previously observed (Auffenberg, 1981). A

regular series of field surveys on the western half of Flores are inevitable to monitor the remaining populations, in order to gain an insight into the most current distribution of individuals in this area on Flores. A similar level of genetic diversity to the population on Flores was found for the island populations on Komodo and Rinca. On the contrary, the level of genetic diversity for Gili Motang and Nusa Kode was the lowest due to the lack of molecular variations, both in nucleotide substitution and sequence length. The lack of genetic diversity in these isolated small island populations is consistent with the low level of allelic diversity and heterozygosity inferred from nuclear data (Ciofi and Bruford, 1999). Because a reduced genetic diversity are often found in many populations of reptilian endemics distributed on islands (Frankham, 1996a), a maintenance of the existing genetic diversity in these small island populations is required to prevent local extinctions (Frankham, 1996b).

Three separate units for conservation may be assigned to maintain the level of genetic diversity for the Komodo monitor population in the Lesser Sundas. Based on the phylogeographic analyses in this study, the whole population is structured in three regions. The composition of haplotypes and the relationships among these haplotypes distributed within each of the three regions apparently determine their unique genetic features. Komodo is a separate region in the west, where only a quarter of its total number of maternal lineages was shared with a nearby island, i.e. Rinca. The level of genetic exchange between Komodo and the islands eastwards was very low (Ciofi and Bruford, 1999), suggesting rare migration events and the subsequent isolation of this island population from those eastwards. The Central region equally merits a separate unit. This region is composed of three islands as well as a location on the western coast of Flores. A total of six haplotypes were discovered on Rinca, and most of them were shared with the other islands, including the two small islands. The level of genetic exchange among these islands in the Central region was apparently more frequent than that between Komodo and the islands eastwards (Ciofi and Bruford, 1999). Yet, the genetic composition in the Central region is distinct to the other two regions. Likewise, the Eastern region is a discrete subdivision in the whole population of the Komodo monitor. This region is represented by a single population in Flores North, which is geographically and genetically distant from all the others westwards. The Eastern region is regarded a separate unit for conservation, due to its considerable level of genetic divergence from the other regions. This mosaic pattern of haplotype distribution in three regions seems to be associated with the extent of geographic distance separating them. In other words, the larger a

geographic distance separating two populations, e.g. population at region level, the more divergent their haplotypes are.

A designation of three Management Units (MUs) for conservation is necessary to prevent a further contraction of the Komodo monitor population. The existing populations in the Lesser Sunda Islands, Indonesia, are thought to be a relict of a larger distribution range in the past that might have spaned from Australia in the east to eastern Java in the west. To help *in situ* conservation programmes, these populations should be managed with regard to the subdivision in three regions, in order to maintain the existing genetic diversity and the evolutionary processes involved in shaping this phylogeographic pattern. In other words, the level of genetic divergence among these three regions and the associated historical events should be preserved to counterbalance the effects of habitat degradation on islands. The population management programmes may include an augmentation of the small island populations and a reintroduction to an extinct population. A suitable source population for an augmentation for each of the small island populations can be readily identified from the phylogeographic pattern described in this study. For an augmentation programme for Gili Motang, one should consider to transfer individuals of similar genetic material to those currently found in Gili Motang, for example Flores West or Loh Tongker in the Central region. Because Gili Motang shared the haplotype H5 with both of the proposed population sources, the individuals to be transferred from either Flores West or Loh Tongker should be characterised as having the haplotype H5. On the other hand, a source population for a reintroduction to Padar cannot be readily identified in this study. Despite its geographic proximity to the western coasts of Rinca, no data are currently available to deduce the lineages distributed on this island in the past.

An identification of three MUs based on a phylogeographic pattern was performed in this study for the Komodo monitor. This study is the first to perform a phylogeographic analyses for the Komodo monitor using a relatively large sample size. The current phylogenetic relationships of all mtDNA haplotypes described in this study failed to show a reciprocally monophyletic relationship for the haplotypes distributed on Komodo and those in the Central region. Based on the phylogeny alone, Komodo and the Central region cannot be recognised as two separate sources of genetic diversity. However, these two regions were demonstrated to retain a significant proportion of the total molecular variations. Moreover, most of the haplotypes found on Komodo were unique to this island, whereas two rare

haplotypes were found on Rinca, in the Central region. Komodo also showed a significant distinction in the level of allelic diversity and heterozygosity from the islands eastwards (Ciofi and Bruford, 1999). Thus, Komodo may be designated as one MU, besides the Central and Eastern regions. Furthermore, Komodo might be recognised as an ESU, provided that there was a considerable level of reproductive isolation from the other populations (Waples, 1991). The Central and Eastern regions can be recognised as two separate MUs, because they also showed a significant level of distinction, i.e. in the haplotype composition and phylogenetic divergence. A further assessment on the reproductive isolation for Flores North should be conducted to characterise an ESU for the Eastern region by supporting the reciprocal monophyly demonstrated in this study for the Eastern clade.

The integration of phylogeographic pattern in a conservation management plan is indispensable, because this pattern may be used to infer the evolutionary history of a population. In addition to the knowledge on the level of genetic diversity, the evolutionary processes that give rise to the current phylogeographic pattern may shed light the current state of a population, for which conservation priorities may be given. The application of molecular techniques can accelerate the procedures for identifying the phylogeographic pattern of a population, including the construction of a genealogy as well as the quantification of genetic divergence among subpopulations. Nevertheless, further considerations should be taken to address the complexity of some populations with a distribution over islands. Such considerations include the demographic trend and several ecological factors, e.g. prey abundance and habitat quality. In addition, a future study on the impact of climate change on the distribution of the Komodo monitor across the Lesser Sundas may be needed to assess the relationships between the reduced distribution area and a changing environment. Nevertheless, the putative human activities accelerating the rate of habitat degradation should be reduced to a minimum, in order to preserve this species from extinction. Habitat degradation seems to be common in Flores, presumably because there is insufficient control for poaching and unsustainable agriculture. On the other hand, a compensatory subsistence for local inhabitants such as a paid involvement in a sustainable ecotourism or conservation project is probably a more effective method to reduce poaching and habitat degradation. Additionally, an awareness campaign in zoos around the world displaying Komodo monitors would be effective to reach a larger public to deliver the message on the importance of conserving this species and its habitats.

Chapter 5

CONCLUSIONS

A study on the phylogeography of the wild populations of the Komodo monitor was conducted as a part of a larger collaborative project among institutions in Indonesia, Italy, and Germany. The results of this study indicate a designation of three Management Units (MUs) in the distribution range of the Komodo monitor in the Lesser Sunda Islands, Indonesia. The designation of three MUs may be extended to an application of this result for an *in situ* conservation management plan. However, further studies are still needed to implement an appropriate strategy to protect the Komodo monitor from extinction in the medium-term future. Some of the key conclusions in this study are presented as follows:

1. Eleven (11) haplotypes are identified from the data set of three hundred and sixty six (366) mtDNA CRI sequences and the distribution of these haplotypes across islands is non-random. The larger islands of Komodo, Rinca, and Flores harbour a larger number of haplotypes and each of the small islands of Gili Motang and Nusa Kode harbours one haplotype. A higher level of genetic diversity is found in the larger island populations, whereas a genetic diversity impoverishment seems to occur in the small island populations.

2. The phylogenetic relationship among the eleven maternal lineages of the Komodo monitor demonstrates a major branching pattern of two clades, i.e. Central-Western and Eastern. The Central-Western clade encompasses all haplotypes distributed in the Central and Western regions, and the Eastern clade groups all haplotypes distributed in the Eastern region.

3. The population of the Komodo monitor is strongly structured in three regions, whereas a population structure in five islands is shown to a lesser degree. These results suggest a higher level of genetic divergence among the three regions and a higher level of genetic similarity among five island populations.

4. The positive correlation between geographic and genetic distances indicates a level of isolation among populations. Similarly, the larger sea distance (> 2 km) is likely a barrier to dispersal among islands for the terrestrial Komodo monitor.

5. An exposure of the land bridge connecting Komodo and the islands eastwards during the Late Pleistocene is proposed to explain a dispersal event between islands.

References

Amato, M. L., R. J. Brooks, and J. Fu. 2008. A phylogeographic analysis of populaions of the wood turtle (*Glyptemys insculpta*) throughout its range. Molecular Ecology **17**:570-581.

Amer, S. A. M. and Y. Kumazawa. 2008. Timing of a mtDNA gene rearrangement and intercontinental dispersal of varanid lizards. Genes, Genetics and Systematics **83**:275-280.

Amos, W. and J. Harwood. 1998. Factor affecting levels of genetic diversity in natural populations. Philosophical Transactions of the Royal Society London B **353**:177-186.

Arbogast, B. S., S. V. Edwards, J. Wakeley, P. Beerli, and J. B. Slowinski. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. Annual Review of Ecology and Systematics **33**:707-740.

Arida, E. and W. Böhme. 2010. The origin of *Varanus*: when fossils, morphology, and molecules alone are never enough. Biawak **4**:117-124.

Ast, J. C. 2001. Mitochondrial DNA Evidence and Evolution in Varanoidea (Squamata). Cladistics **17**:211-226.

Auffenberg, W. 1981. The behavioral ecology of the Komodo monitor. University of Florida Presses, Gainesville.

Auffenberg, W. 1994. The Bengal monitor. University Presses of Florida, Gainesville, Florida.

Avise, J. C. 2004. Molecular markers, natural history, and evolution. Sinauer Associates Inc., Sunderland.

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

Balloux, F. and N. Lougon-Moulin. 2002. The estimation of population differentiation with microsatellite markers. Molecular Ecology **11**:155-165.

Beck, D. D. 2004. Overview of the Family Helodermatidae for Varanophiles. In: E. R. Pianka, D. R. King, and R. A. King, editors. Varanoid lizards of the world. Indiana University Press, Bloomington.

Beheregaray, L. B., C. Ciofi, A. Caccone, J. P. Gibbs, and J. R. Powell. 2003. Genetic divergence, phylogeography, and conservation units of giant tortosies from Santa Cruz and Pinzón, Galápagos Islands. Conservation Genetics **4**:31-46.

Bird, M. I., D. Taylor, and C. Hunt. 2005. Palaeoenvironment of insular Southeast Asia during the last glacial period: a savanna corridor in Sundaland? Quaternary Science Reviews **24**:2228-2242.

Blakey, R. 2010. Paleogeography. Geology Department, Northern Arizona University, USA, Flagstaff, Arizona.

Böhme, W. 2003. Checklist of the living monitor lizards of the world (family Varanidae). Zoologische Verhandelingen Leiden **341**.

Bossuyt, F. and M. C. Milinkovitch. 2001. Amphibians as indicators of early tertiary "out-of-India" dispersal of vertebrates. Science **292**:93-95.

Bowen, B. W. and S. A. Karl. 2007. Population genetics and phylogeography of sea turtles. Molecular Eoclogy **16**:4886-4907.

Brehm, A., D. J. Harris, C. Aves, J. Jesus, F. Thomarat, and L. Vicente. 2003. Structure and evolution of the mitochondrial DNA complete control region in the lizard *Lacerta dugesii* (Lacertidae, Sauria). Journal of Molecular Evolution **56**:46-53.

Brock, T. D. 1994. Life at high temperatures. Yellowstone Association for Natural Science, History & Education, Inc., Yellowstone National Park, Wyoming.

Brongersma, L. D. 1958. On an extinct species of the genus *Varanus* (Reptilia, Sauria) from the island of Flores. Zoologische Mededelingen **36**:113-125.

Brown, J. R., A. T. Beckenbach, and M. J. Smith. 1993. Intraspecific DNA sequence variation of the mitochondrial control region of white sturgeon (*Acipenser transmontanus*). Molecular Biology and Evolution **10**:326-341.

Brown, W. M., M. George, and A. C. Wilson. 1976. Rapid evolution of animal mitochondrial DNA. Proceeding of National academy of Sciences USA **76**:1967-1971.

Burden, W. D. 1928. Results of the Douglas Burden expedition to the island of Komodo: observations on the habits and distribution of *Varanus komodoensis* Ouwens. American Museum Novitates **316**:1-10.

Buroker, N. E., J. R. Brown, T. A. Gilbert, P. J. O'Hara, A. T. Beckenbach, W. K. Thomas, and M. J. Smith. 1990. Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model. Genetics **124**:157-163.

Butlin, R. K., C. Walton, K. A. Monk, and J. R. Bridle. 1998. Biogeography of Sulawesi grasshoppers, genus *Chitaura*, using DNA sequence data. Pages 355-359 in: R. Hall and J. D. Holloway, editors. Biogeography and Geological Evolution of SE Asia. Backhuys, Leiden, The Netherlands.

Charlesworth, B. 1998. Measures of divergence between populations and the effect of forces that reduce variability. Molecular Biology and Evolution **15**:538-543.

Ciofi, C. 2002. Conservation genetics. Pages 129-164 in: J. B. Murphy, C. Ciofi, C. de La Panouse, and T. Walsh, editors. Komodo dragons biology and conservation. Smithsonian Institution Press, Washington.

Ciofi, C., M. A. Beaumont, I. R. Swingland, and M. W. Bruford. 1999. Genetic divergence and units for conservation in the Komodo dragon *Varanus komodoensis*. Proceeding of The Royal Society London B **266**:2269-2274.

Ciofi, C. and M. W. Bruford. 1999. Genetic structure and gene flow among Komodo dragon populations inferred by microsatellite loci analysis. Molecular Ecology **8**:S17-S30.

Ciofi, C. and M. E. De Boer. 2004. Distribution and conservation status of the Komodo monitor (*Varanus komodoensis*). Herpetological Journal **14**:99-107.

Ciofi, C., J. Puswati, D. Winana, M. E. De Boer, C. Guido, and P. Sastrawan. 2007. Preliminary analysis of home range structure in the Komodo monitor, *Varanus komodoensis*. Copeia **2**:462-470.

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology **9**:1657-1660.

Clos, L. M. 1995. A new species of *Varanus* (Reptilia: Sauria) from the Miocene of Kenya. Journal of Vertebrate Paleontology **15**:254-267.

Conrad, J. L., O. Rieppel, and L. Grande. 2008. Re-assessment of varanid evolution based on new data from *Saniwa ensidens* Leidy, 1870 (Squamata, Reptilia). American Museum Novitates.

Cook, S., R. Singidan, and I. W. B. Thornton. 2001. Colonization of an island volcano, Long Island, Papua New Guinea, and an emergent island, Motmot, in its caldera lake. IV. Colonization by non-avian vertebrates. Journal of Biogeography **28**:1353-1363.

Crandall, K. A., O. R. P. Bininda-edmonds, G. M. Mace, and R. K. Wayne. 2000. Considering evolutinary processes in conservation biology. Trends in Ecology and Evolution **15**:290-295.

Douzery, E. and E. Randi. 1997. The mitochondrial control region of Cervidae: evolutionary patterns and phylogenetic content. Molecular Biology and Evolution **14**:1154-1166.

Drummond, A. J. and A. Rambaut. 2007. BEAST: Bayesian Evolutionary Analysis by Sampling Trees. BMC Evolutionary Biology 7 **7**:214.

Dutton, P. H., B. W. Bowen, D. W. Owens, A. Barragan, and S. K. Davis. 1999. Global phylogeography of the leatherback turtle (*Dermochelys coriacea*). Journal of Zoology **248**:397-409.

Estes, R. 1983. The fossil record and early distribution of lizards. Museum of Comparative Zoology Harvard University, Massachusetts.

Estes, R. 1984. Fish, ampbibians and reptiles from the Etadunna Formation, Miocene of South Australia. Australian Zoologist **21**:335-343.

Evans, B. J., J. Supriatna, N. Andayani, M. I. Setiadi, D. C. Cannatella, and D. J. Melnick. 2003. Monkeys and toads define areas of endemism on Sulawesi. Evolution **57**:1436-1443.

Evans, M., N. Hastings, and J. B. Peacock. 2000. Statistical distributions. 3rd edition. John Wiley & Sons, Inc., New York.

Excoffier, L., G. Laval, and S. Schneider. 2006. Arlequin an integrated software package for population genetics. Bern.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics **131**:479-491.

Fleischer, R. C., E. A. Perry, K. Muralidharan, E. E. Stevens, and C. M. Wemmer. 2001. Phylogeography of the Asian elephant (*Elephas maximus*) based on mitochondrial DNA. Evolution **55**:1882-1892.

Frankham, R. 1996a. Do island populations have less genetic variation than mainland populations? Heredity **78**:3111-3127.

Frankham, R. 1996b. Relationship of genetic variation to population size in wildlife. Conservation Biology **10**:1500-1508.

Fuller, S., P. Baverstock, and D. King. 1998. Geographic origins of Goannas (Varanidae): A molecular perspective. Molecular Phylogenetics and Evolution **9**:294-307.

Gaulke, M., W. Erdelen, and F. Abel. 1999. A radio-telemetric study of the Water monitor lizard (*Varanus salvator*) in North Sumatra, Indonesia. Mertensiella (Advances in Monitor Research II) **11**:63-78.

Goodwin, H. J., I. J. Kent, K. T. Parker, and M. J. Walpole. 1997. Tourism, conservation and sustainable development: Volume III Komodo National Park Indonesia. Durrel Institute of Conservation and Ecology (DICE), Institute of Mathematics and Statistics (IMS), University of Kent.

Green, D. M. 2005. Designatable units for status assessment of endangered species. Conservation Biology **19**:1813-1820.

Hall, R. 1998. The plate tectonic of Cenozoic SE Asia and the distribution of land and sea. Pages 99-131 in: R. Hall and J. D. Holloway, editors. Biogeography and geological evolution of SE Asia. Backhuys Publishers, Leiden, The Netherlands.

Hall, R. 2001. Cenozoic reconstruction of SE Asia and the SW Pacific: changing patterns of land and sea. In: I. Metcalfe, J. M. B. Smith, M. Morwood, and I. D. Davidson, editors. Faunal and floral migrations and evolution in SE Asia-Australasia. A.A. Balkema (Swets & Zeitlinger Publishers), Lisse.

Hall, R. 2002. Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer-based reconstruction, model and animations. Journal of Asian Earth Sciences **20**:353-431.

Hartl, D. L. and A. G. Clark. 1997. Principles of population genetics. Sinauer Associates, Inc., Sunderland, Massachussetts.

Hasegawa, M., H. Kishino, and K. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution **22**:160-174.

Head, J. J., P. M. Barrett, and E. J. Rayfield. 2009. Neurocranial osteology and systematic relationships of *Varanus* (*Megalania*) prisca Owen, 1859 (Squamata:Varanidae). Zoological Journal of the Linnean Society **155**:445-457.

Hocknull, S. A., P. J. Piper, G. D. van den Bergh, R. A. Due, M. J. Morwood, and I. Kurniawan. 2009. Dragon's paradise lost: palaeobiogeography, evolution and extinction of the largest-ever terrestrial lizards (Varanidae). Pages 1-15 PLoS ONE.

Holmes, R. B., A. M. Murray, Y. S. Attia, E. L. Simons, and P. Chatrath. 2010. Oldest known *Varanus* (Squamata: Varanidae) from the Upper Eocene and Lower Oligocene of Egypt: support for an African origin of the genus. Palaentology **53**:1099-1110.

Hooijer, D. A. 1972. *Varanus* (Reptilia, Sauria) from the Pleistocene of Timor. Zoologische Mededelingen Museum Leiden **47**:445-447.

Huelsenbeck, J. P. and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics **17**:754-755.

Hutchinson, M. N. and S. C. Donnellan. 1993. Biogeography and the phylogeny of the Squamata. In: C. J. Glasby, G. J. B. Ross, and P. L. Beesley, editors. Fauna of Australia 2A:1-24. Australian Government Publishing Service, Canberra.

IUCN. 2007. Red List of Threatened Species. *Varanus komodoensis*. World Conservation Monitoring Centre.

Jessop, T. S., T. Madsen, C. Ciofi, M. J. Imansyah, D. Purwandana, H. Rudiharto, A. Arifiandy, and J. A. phillips. 2007. Island differences in population size structure and catch per unit effort and their conservation implications for Komodo dragons. Biological Conservation **135**:247-255.

Jønsson, K. A., M. Irestedt, J. Fuchs, P. G. P. Ericson, L. Christidis, R. C. K. Bowie, J. A. Norman, E. Pasquet, and J. Fjeldså. 2004. Explosive avian radiations and multi-directional dispersal across Wallacea: evidence from the Campephagidae and other Crown Corvida (Aves). Molecular Phylogenetics and Evolution **47**:221-236.

Kawakita, A., T. Sota, J. S. Ascher, M. Ito, H. Tanaka, and M. Kato. 2003. Evolution and phylogenetic utility of alignment gaps within intron sequences of three nuclear genes in Bumble bees (*Bombus*). Molecular Biology and Evolution **20**:87-92.

Kidd, M. G. and V. L. Friesen. 1998. Sequence variation in the guillemot (Alcidae: *Cepphus*) mitochondrial control region and its nuclear homolog. Molecular Biology and Evolution **15**:61-70.

Kidwell, S. M. and K. W. Flessa. 1996. The quality of the fossil record: population, species, and communities. Annual Review of Earth and Planetary Sciences **24**:433-464.

King, D. and B. Green. 1999. Monitors The biology of varanid lizards. 2nd edition. Krieger Publishing Company, Malabar.

Kumazawa, Y. and H. Endo. 2004. Mitocodrial genome of the Komodo dragon: efficient sequencing method with reptile-oriented primers and novel gene rearrangements. DNA Research **11**:115-125.

Kumazawa, Y., H. Ota, M. Nishida, and T. Ozawa. 1996. Gene rearrangement in snake mitochondrial genomes: highly concerted evolution of control-region-like sequences duplicated and inserted into a tRNA gene cluster. Molecular Biology and Evolution **13**:1242-1254.

Lacy, R. C. 1987. Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. Conservation Biology **1**:143-158.

Lawrence, H. A., G. A. Taylor, C. D. Millar, and D. M. Lambert. 2008. High mitochondrial and nuclear genetic diversity in one of the world's most endangered seabirds, the Chatham Island Taiko (*Pterodroma magentae*). Conservation Genetics **9**:1293-1301.

Levinson, G. and G. A. Gutman. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Molecular Biology and Evolution **43**:203-221.

Lourie, S. A., D. M. Green, and A. C. J. Vincent. 2005. Dispersal, habitat differences, and comparative phylogeography of Southeast Asian seahorses (Sygnathidae: *Hippocampus*). Molecular Ecology **14**:1073-1094.

Ludwig, A., B. May, L. Debus, and I. Jenneckens. 2000. Heteroplasmy in the mtDNA control region of Sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*). Genetics **156**:1933-1947.

Lukoschek, V., M. Waycott, and H. Marsh. 2007. Phylogeography of the Olive sea snake, *Aipysurus laevis* (Hydrophiinae) indicates Pleistocene range expansion around northern Australia but low contemporary gene flow. Molecular Ecology **16**:3406-3422.

Lunt, D. H., L. E. Whipple, and B. C. Hyman. 1998. Mitochondrial DNA variable number tandem repeats (VNTRs): utility and problems in molecular ecology. Molecular Ecology **7**:1441-1455.

Maxam, A. M. and W. Gilbert. 1977. A new method for sequencing DNA. Proceeding of the National Academy of Sciences **74**:560-564.

Melo, M. and C. O'Ryan. 2007. Genetic differentiation between Príncipe Island and mainland populations of the grey parrot (*Psittacus erithacus*), and implications for conservation. Molecular Ecology **16**:1673-1685.

Michaux, B. 1994. Land movements and animal distribution in east Wallacea (eastern Indonesia, Papua New Guinea and Melanesia). Palaeogeography, Palaeoclimatology, Palaeoecology **112**:323-343.

Molnar, R. E. 2004a. Dragons in the dust: the paleobiology of the giant monitor lizard Megalania. Indiana University Press, Bloomington.

Molnar, R. E. 2004b. The long and honorable history of monitors and their kin. In: E. R. Pianka, D. King, and R. A. King, editors. Varanoid lizards of the world. Indiana University Press, Bloomington.

Monk, K., Y. De Fretes, and G. Reksodiharjo-Lilley. 1997. The ecology of Nusa Tenggara and Maluku. Oxford University Press, Oxford.

Moritz, C. 1994. Defining 'Evolutionary Significant Units' for conservation. Trends in Ecology and Evolution **9**:373-375.

Moritz, C. 1999. Conservation units and translocations: strategies for conserving evolutionary processes. Hereditas **130**:217-228.

Moritz, C. and W. M. Brown. 1986. Tandem duplication of D-loop and ribosomal RNA sequences in lizard mitochondrial DNA. Science **233**:1425-1427.

Morwood, M. 2001. Early hominid occupation of Flores, East Indonesia, and its wider significance. In: Faunal and floral migrations and evolution in SE Asia-Australasia: 38th U.S. Rock Mechanics Symposium. CRC Press, Washington.

Morwood, M. J., F. Aziz, P. O'Sullivan, Nasruddin, D. R. Hobbs, and A. Raza. 1999. Archaeological and paleontological research in central Flores, east Indonesia: results of fieldwork 1997-98. Antiquity **73**:273-286.

Morwood, M. J., P. B. O'Sullivan, F. Aziz, and A. Raza. 1998. Fission-track ages of stone tools and fossils on the east Indonesian island of Flores. Nature **392**:173-176.

Müller, K. 2006. Incorporating information from length-mutational events into phylogenetic analysis. Molecular Phylogenetics and Evolution **38**:667-676.

Murray, S. P., A. Dharma, J. C. Kindle, H. E. Hurlburt, and W. A. Kuperman. 1990. Characteristics of circulation in an Indonesian archipelago strait from hydrography, current measurements and modelling results. Naval Oceanographic and Atmospheric Research Laboratory, Stennis Space Center, Mississippi.

Nei, M. 1972. Genetic distance between populations. American Naturalist **106**:283-292.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences of the United States of America **70**:3321-3323.

Nei, M. and S. Kumar. 2000. Molecular evolution and phylogenetics. Oxford University Press, New York.

Nesbø, C. L., M. O. Arab, and K. S. Jakobsen. 1998. Heteroplasmy, length and sequence variation in the mtDNA control region of three percid fish species (*Perca fluviatilis*, *Acerina cernua*, *Stizotedion lucioperca*). Genetics **148**:1907-1919.

Nydam, R. L. 2000. A new taxon of Helodermatid-like lizard from the Albian-Cenomanian of Utah. Journal of Vertebrate Paleontology **20**:285-294.

Paetkau, D. 1999. Using genetics to identify intraspecific conservation units: a critique of current methods. Conservation Biology **13**:1507-1509.

Parker, P. G., A. A. Snow, S. M. D., G. C. Booton, and P. A. Fuerst. 1998. What can molecules tell us about populations: choosing and using a molecular marker. Ecology **79**:361-382.

Parsons, T. J., D. S. Muniec, K. Sullivan, N. Woodyatt, R. Alliston-Greiner, M. R. Wilson, D. L. Berry, K. A. Holland, V. W. Weedn, P. Gill, and M. M. Holland. 1997. A high observed substitution rate in the human mitochondrial DNA control region. Nature Genetics **15**:363-368.

Pearce, J. M. 2006. Minding the gap: frequency of indels in mtDNA control region sequence data and influence on population genetic analyses. Molecular Ecology **15**:333-341.

Peppin, D. J. 1999. The origin of Monitor lizards based on a review of the fossil evidence. Mertensiella Advances in Monitor Research II:11-22.

Pernetta, A. P. 2009. Monitoring the trade: using CITES database to examine the global trade in live monitor lizards (*Varanus* spp.). Biawak **3**:37-45.

Pianka, E. R. 1994. Comparative ecology of *Varanus* in the Great Victoria Desert. Australian Journal of Ecology **19**:395-408.

Pianka, E. R. 2004. *Lanthanotus borneensis*. In: E. R. Pianka, D. R. King, and R. A. King, editors. Varanoid lizards of the world. Indiana University Press, Bloomington.

Pianka, E. R. and D. R. King. 2004. Introduction. In: E. R. Pianka, D. R. King, and R. A. King, editors. Varanoid lizards of the world. Indiana University Press, Bloomington.

Posada, D. 2008. jModelTest: Phylogenetic model averaging. Molecular Biology and Evolution **25**:1253-1256.

Posada, D. and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14**:817-818.

Rage, J.-C. and S. Bailon. 2005. Amphibians and squamate reptiles from the late early Miocene (MN 4) of Béon 1 (Montréal-du-Gers, southwestern France). Geodiversitas **27**:413-441.

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

Reed, D. H. and R. Frankham. 2003. Correlation between fitness and genetic diversity. Conservation Biology **17**:230-237.

Roques, S., J. A. Godoy, J. J. Negro, and F. Hiraldo. 2004. Organization and variation of the mitochondrial control region in two vulture species, *Gypaetus barbatus* and *Neophron percnopterus*. Journal of Heredity **95**:332-337.

Rosel, P. E. and B. A. Block. 1996. Mitochondrial control region variability and global population structure in the swordfish, *Xiphias gladius*. Marine Biology **125**:11-22.

Rosenbaum, P. A., J. M. Robertson, and K. R. Zamudio. 2007. Unexpectedly low genetic divergences among populations of the threatened bog turtle (*Glyptemys muhlenbergii*). Conservation Genetics **8**:331-342.

Ruokonen, M. and L. Kvist. 2002. Structure and evolution of the avian mitochondrial control region. Molecular Phylogenetics and Evolution **23**:422-432.

Saccone, C., G. Pesole, and E. Sbisá. 1991. The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. Journal of Molecular Evolution **33**:83-91.

Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487-491.

Salgueiro, P., M. M. Coelho, J. M. Palmeirim, and M. Ruedi. 2004. Mitochondrial DNA variation and population structure of the island endemic Azorean bat (*Nyctalus azoreum*). Molecular Ecology **13**:3357-3366.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2nd edition. CSH Laboratory Press, Cold spring Harbor, NY.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proceeding of the National Academy of Sciences USA **74**:5463-5467.

Sastrawan, P. and C. Ciofi. 2002. Population distribution and home range. In: J. B. Murphy, C. Ciofi, C. de La Panouse, and T. Walsh, editors. Komodo dragons biology and conservation. Smithsonian Institution Press, Washington.

Sbisà, E., F. Tanzariello, A. Reyes, G. Pesole, and C. Saccone. 1997. Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. Gene **205**:125-140.

Schulte, J. A., J. Melville, and A. Larson. 2003. Molecular phylogenetic evidence for ancient divergence of lizard taxa on either side of Wallace's Line. Proceeding of Royal Society London B **270**:597-603.

Shine, R., P. Harlow, J. S. Keogh, and Boeadi. 1996. Commercial harvesting of giant lizards: the biology of Water monitors *Varanus salvator* in southern Sumatra. Biological Conservation **77**:125-134.

Simmons, M. P. and H. Ochoterena. 2000. Gaps as characters in sequence-based phylogenetic analyses. Systematic Biology **49**:369-381.

Simpson, G. G. 1977. Too many lines: The limits of the Oriental and Australian zoogeographic regions. Proceedings of the American Philosophical Society **121**:107-120.

Smith, K. T., B.-A. S. Bhullar, and P. A. Holroyd. 2008. Earliest African record of *Varanus* stem-clade from Oligocene Egypt. Journal of Vertebrate Paleontology **28**:909-913.

Smith, W., I. A. W. Scott, and J. S. Keogh. 2007. Molecular phylogeography of Rosenberg's goanna (Reptilia: Varanidae: *Varanus rosenbergi*) and its conservation status in New South Wales. Systematics and Biodiversity **5**:361-369.

Sprintall, J., J. T. Potemra, S. L. Hautala, N. A. Bray, and W. W. Pandoe. 2003. Temperature and salinity variability in the exit passages of the Indonesian Throughflow. Deep-Sea Research II **50**:2183-2204.

Steppan, S. J., C. Zawadzki, and L. R. Heaney. 2003. Molecular phylogeny of the endemic Philippine rodent *Apomys* (Muridae) and the dynamics of diversification in an oceanic archipelago. Biological Journal of the Linnean Society **80**:699-715.

Storz, J. F. 1999. Genetic consequences of mammalian social structure. Journal of Mammalogy **80**:553-569.

Swofford, D. L. 2001. Phylogenetic Analysis Using Parsimony and other methods (PAUP*). Sinauer Associates, Inc., Sunderland.

Taberlet, P. 1996. The use of mitochondrial DNA control region sequencing in conservation genetics. Pages 125-142 in: T. B. Smith and R. K. Wayne, editors. Molecular genetic approaches in conservation. Oxford University Press, New York.

Takumi, T., H. Fujiwake, and Y. Kurachi. 1997. A dye terminator method for automated DNA sequencing using four fluorescent dideoxynucleosides and thermal cycling. Analytical Science **13**:735-739.

Toon, A., P. B. Mather, A. M. Baker, and K. L. Durrant. 2007. Pleistocene refugia in arid landscape: analysis of a widely distributed Australian passerine. Molecular Ecology **16**:2525-2541.

Torgersen, T., J. Luly, P. De Deckker, M. R. Jones, D. E. Searle, A. R. Chivas, and W. J. Ullman. 1988. Late quaternary environments of the Carpentaria Basin, Australia. Palaeogeograhy, Palaeoclimatology, Palaeoecology **67**:245-261.

Ursenbacher, S., M. Carlsson, V. Helfer, H. Tegelström, and L. Fumagalli. 2006. Phylogeography and Pleistocene refugia of the adder (*Vipera berus*) as

inferred from mitochondrial DNA sequence data. Molecular Ecology **15**:3425-3437.

van den Bergh, G. D., R. D. Awe, M. J. Morwood, T. Sutikna, Jatmiko, and E. W. Saptomo. 2007. The youngest *Stegodon* remains in southeast Asia from the Late Pleistocene archaelogical site Liang Bua, Flores, Indonesia. Quaternary International **182**:16-48.

van den Bergh, G. D., J. de Vos, F. Aziz, and M. J. Morwood. 2001. Elephantoidea in the Indonesian region: new *Stegodon* findings from Flores. The World of Elephants- International Congress, Rome.

van den Bergh, G. D., H. J. M. Meijer, R. A. Due, M. J. Morwood, K. Szabó, L. W. van den Hoek Ostende, T. Sutikna, E. W. Saptomo, P. J. Piper, and K. M. Dobney. 2009. The Liang Bua faunal remains: a 95 k.yr. sequence from Flores, East Indonesia. Journal of Human Evolution **57**:527-537.

Veevers, J. J. 1971. Shallow stratigraphy and structure of the Australian continental margin beneath the Timore Sea. Marine Geology **11**:209-249.

Verstappen, H. T. 1980. Quaternary climatic changes and natural environment in SE Asia. GeoJournal **4**:45-54.

Voris, H. K. 2000. Maps of the Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. Journal of Biogeography **27**:1153-1167.

Wan, Q.-H., H. Wu, T. Fujihara, and S.-G. Fang. 2004. Which genetic marker for which conservation genetics issue? Electrophoresis **25**:2165-2176.

Waples, R. S. 1991. Pacific salmon, *Onchorynchus* spp., and definition of "species" under the endangered species act. Marine Fisheries Review **53**:11-22.

White, M. M. and H. R. Martin. 2009. Structure and conservation of tandem repeats in the mitochondrial DNA control region of the Least brook lamprey (*Lampetra aepyptera*). Journal of Molecular Evolution **68**:715-723.

White, T. A. and J. B. Searle. 2006. Genetic diversity and population size: island populations of the Common shrew, *Sorex araneius*. Molecular Ecology **16**:2005-2016.

Wilkinson, G. S. and A. Chapman, M. 1991. Length and sequence variation in Evening bat D-loop mtDNA. Genetics **128**:607-617.

Wroe, S. 2002. A review of terrestrial mammalian and reptilian carnivore ecology in Australian fossil faunas, and factors influencing their diversity: the myth of reptilian domination and its broader ramifications. Australian Journal of Zoology **50**:1-24.

Zhu, D., S. Degnan, and C. Moritz. 1998. Evolutionary distinctiveness and status of the endangered Lake Eacham Rainbowfish (*Melanotaenia eachamensis*). Conservation Biology **12**:80-93.

Appendices

Appendix A: Samples data

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
1	FLN01	PO1	Flores	Pota	1994-1998	Not tagged				
2	FLN02	PO2	Flores	Pota	1994-1998	Not tagged				
3	FLN03	PO3	Flores	Pota	1994-1998	Not tagged				
4	FLN04	RI1	Flores	Riung	1994-1998	Not tagged				
5	FLN05	RI2	Flores	Riung	1994-1998	Not tagged				
6	FLN06	RI3	Flores	Riung	1994-1998	Not tagged				
7	FLW01	WW1	Flores	Wae Wuul	1994-1998	Not tagged				
8	FLW03	WW3	Flores	Wae Wuul	1994-1998	Not tagged				
9	FLW04	WW4	Flores	Wae Wuul	1994-1998	Not tagged				
10	FLW05	WW5	Flores	Wae Wuul	1994-1998	Not tagged				
11	FLW06	WW6	Flores	Wae Wuul	1994-1998	Not tagged				
12	FLW07	WW7	Flores	Wae Wuul	1994-1998	Not tagged				
13	FLW08	WW8	Flores	Wae Wuul	1994-1998	Not tagged				
14	FLW09	WW9	Flores	Wae Wuul	1994-1998	Not tagged				
15	FLW10	WW10	Flores	Wae Wuul	1994-1998	Not tagged				
16	FLW11	WW11	Flores	Wae Wuul	1994-1998	Not tagged				
17	FLW12	WW12	Flores	Wae Wuul	1994-1998	Not tagged				
18	FLW13	WW13	Flores	Wae Wuul	1994-1998	Not tagged				
19	FLW14	WW14	Flores	Wae Wuul	1994-1998	Not tagged				
20	FLW15	WW15	Flores	Wae Wuul	1994-1998	Not tagged				
21	FLW16	WW16	Flores	Wae Wuul	1994-1998	Not tagged				
22	FLW17	WW17	Flores	Wae Wuul	1994-1998	Not tagged				
23	FLW18	WW18	Flores	Wae Wuul	1994-1998	Not tagged				
24	FLW19	WW19	Flores	Wae Wuul	1994-1998	Not tagged				

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
25	FLW20	WW20	Flores	Wae Wuul	1994-1998	Not tagged				
26	FLW21	WW21	Flores	Wae Wuul	1994-1998	Not tagged				
27	FLW22	WW22	Flores	Wae Wuul	1994-1998	Not tagged				
28	FLW23	WW23	Flores	Wae Wuul	1994-1998	Not tagged				
29	FLW24	WW24	Flores	Wae Wuul	1994-1998	Not tagged				
30	FLW25	WW25	Flores	Wae Wuul	1994-1998	Not tagged				
31	FLW26	WW26	Flores	Wae Wuul	1994-1998	Not tagged				
32	FLW27	WW27	Flores	Wae Wuul	1994-1998	Not tagged				
33	FLW28	LE1	Flores	Lenteng	1994-1998	Not tagged				
34	FLW29	LE2	Flores	Lenteng	1994-1998	Not tagged				
35	FLW30	LE3	Flores	Lenteng	1994-1998	Not tagged				
36	FLW31	LE4	Flores	Lenteng	1994-1998	Not tagged				
37	FLW32	LE5	Flores	Lenteng	1994-1998	Not tagged				
38	FLW33	LE6	Flores	Lenteng	1994-1998	Not tagged				
39	FLW34	LE7	Flores	Lenteng	1994-1998	Not tagged				
40	FLW35	LE8	Flores	Lenteng	1994-1998	Not tagged				
41	FLW36	LE9	Flores	Lenteng	1994-1998	Not tagged				
42	FLW37	LE10	Flores	Lenteng	1994-1998	Not tagged				
43	FLW38	LE11	Flores	Lenteng	1994-1998	Not tagged				
44	FLW39	LE12	Flores	Lenteng	1994-1998	Not tagged				
45	FLW40	LE13	Flores	Lenteng	1994-1998	Not tagged				
46	LWA01	LWA1	Komodo	Loh Wau	1994-1998	Not tagged				
47	LWA02	LWA2	Komodo	Loh Wau	1994-1998	Not tagged				
48	LWA03	LWA3	Komodo	Loh Wau	1994-1998	Not tagged				
49	LWA04	LWA4	Komodo	Loh Wau	1994-1998	Not tagged				
50	LWA05	LWA5	Komodo	Loh Wau	1994-1998	Not tagged				
51	LWA06	LWA6	Komodo	Loh Wau	1994-1998	Not tagged				

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
52	LWA07	LWA7	Komodo	Loh Wau	1994-1998	Not tagged				
53	LWA08	LWA8	Komodo	Loh Wau	1994-1998	Not tagged				
54	LWE01	LWE1	Komodo	Loh Wenci	1994-1998	Not tagged				
55	LWE02	LWE2	Komodo	Loh Wenci	1994-1998	Not tagged				
56	LWE03	LWE3	Komodo	Loh Wenci	1994-1998	Not tagged				
57	LWE04	LWE4	Komodo	Loh Wenci	1994-1998	Not tagged				
58	LWE05	LWE5	Komodo	Loh Wenci	1994-1998	Not tagged				
59	LWE06	LWE6	Komodo	Loh Wenci	1994-1998	Not tagged				
60	LWE07	LWE7	Komodo	Loh Wenci	1994-1998	Not tagged				
61	LBU0005A	K005A	Rinca	Loh Buaya	21-Mar-05	00-063A-9978	78,10	174,50	8,45	Juvenile
62	LBU0006	K006	Rinca	Loh Buaya	23-Oct-02	00-063A-8DB2	70,00		7,40	Juvenile
63	LBU0010	K010	Rinca	Loh Buaya	20-Mar-05	00-0639-BF58	77,10	162,50	8,15	Juvenile
64	LBU0012	K012	Rinca	Loh Buaya	21-Apr-03	00-063A-0EC4	107,75	222,00	29,40	Adult
65	LBU0013	K013	Rinca	Loh Buaya	24-Mar-05	00-063A-20DC	99,80	215,30	19,55	Subadult
66	LBU0018	K018	Rinca	Loh Buaya	30-Mar-03	00-063A-50C3	129,00	262,00	49,60	Adult
67	LBU0019	K019	Rinca	Loh Buaya	20-Apr-03	00-063A-2ADE	94,85	208,50	15,60	Subadult
68	LBU0021	K021	Rinca	Loh Buaya	23-Oct-02	00-063A-8381	94,00		19,00	Adult
69	LBU0022	K022	Rinca	Loh Buaya	20-Apr-03	00-0639-E5EC	71,00	162,50	6,00	Juvenile
70	LBU0025	K025	Rinca	Loh Buaya	07-Nov-02	00-063A-2CFF			66,80	Adult
71	LBU0028	K028	Rinca	Loh Buaya	05-Oct-04	00-063A-7607	88,60	195,40	14,15	Subadult
72	LBU0033	K033	Rinca	Loh Buaya	06-Oct-04	00-063A-0E5B	145,10	186,70	67,00	Adult
73	LBU0039	K039	Rinca	Loh Buaya	08-Oct-04	00-063A-30BC	93,70	181,40	16,60	Subadult
74	LBU0042	K042	Rinca	Loh Buaya	27-Apr-03	00-063A-AB29	93,50	208,00	15,75	Subadult
75	LBU0046	K046	Rinca	Loh Buaya	23-Oct-02	00-063A-AA06	149,00		70,20	Adult
76	LBU0183	K183	Rinca	Loh Buaya	05-Apr-03	00-0643-8258	147,50	287,50	72,00	Adult
77	LBU0186	K186	Rinca	Loh Buaya	11-Apr-03	00-0643-951D	56,50	123,50	3,20	Juvenile
78	LBU0187	K187	Rinca	Loh Buaya	11-Apr-03	00-0643-89AC	140,25	278,50	78,80	Adult

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
79	LBU0188A	K188A	Rinca	Loh Buaya	15-Apr-03	00-0643-71A3	109,50	190,00	25,60	Subadult
80	LBU0188B	K188B	Rinca	Loh Buaya	15-Apr-03	00-0643-71A3	109,50	190,00	25,60	Subadult
81	LBU0189	K189	Rinca	Loh Buaya	17-Apr-03	00-0643-774C	94,50	189,00	17,60	Subadult
82	LBU0191	K191	Rinca	Loh Buaya	17-Apr-03	00-0643-6E32	133,50	273,50	53,60	Adult
83	LBU0192	K192	Rinca	Loh Buaya	19-Apr-03	00-0643-5227	132,75	255,80	54,40	Adult
84	LBU0193	K193	Rinca	Loh Buaya	18-Apr-03	00-0643-9717	146,20	293,30	62,40	Adult
85	LBU0194	K194	Rinca	Loh Buaya	02-Apr-03	00-0643-A80C	32,55	83,60	0,50	Juvenile
86	LBU0197	K197	Rinca	Loh Buaya	19-Apr-03	00-0643-414F	55,25	133,70		Juvenile
87	LBU0199	K199	Rinca	Loh Buaya	05-Oct-04	00-0643-9806	105,45	211,50	23,70	Adult
88	LBU0204	K204	Rinca	Loh Buaya	20-Apr-03	00-0643-988B	95,35	211,00	15,45	Subadult
89	LBU0206	K206	Rinca	Loh Buaya	20-Apr-03	00-0643-84FF	108,60	219,00	26,70	Subadult
90	LBU0208	K208	Rinca	Loh Buaya	25-Mar-05	00-0643-9491	119,95	249,50	32,25	Adult
91	LBU0210	K210	Rinca	Loh Buaya	22-Apr-03	00-0643-80B9	53,75	132,00	2,25	Juvenile
92	LBU0211	K211	Rinca	Loh Buaya	07-Oct-04	00-0643-A834	69,75	159,90	5,65	Juvenile
93	LBU0212	K212	Rinca	Loh Buaya	22-Apr-03	00-0643-6FC7	37,25	97,50	0,65	Juvenile
94	LBU0213	K213	Rinca	Loh Buaya	21-Mar-05	00-0643-9C56	65,00	152,10	4,80	Juvenile
95	LBU0214	K214	Rinca	Loh Buaya	07-Oct-04	00-0643-7E5F	69,10	159,90	5,60	Juvenile
96	LBU0215	K215	Rinca	Loh Buaya	22-Apr-03	00-0643-8C24	108,50	221,00	30,90	Subadult
97	LBU0218	K216	Rinca	Loh Buaya	05-Oct-04	00-0643-9831	56,75	130,50	2,85	Juvenile
98	LBU0219	K217	Rinca	Loh Buaya	21-Apr-03	00-0643-741E	138,00	267,00		Adult
99	LBU0411	K411	Rinca	Loh Buaya	06-Oct-03	00-0643-7DA0	76,85	176,20	8,70	Juvenile
100	LBU0505	K505	Rinca	Loh Buaya	05-Apr-04	00-064C-AE31	112,70	249,00	34,50	Adult
101	LBU0519	K519	Rinca	Loh Buaya	28-Mar-05	00-064C-EF9E	57,95	142,00	3,10	Juvenile
102	LBU0522	K522	Rinca	Loh Buaya	27-Mar-05	00-064C-F198	52,65	132,40	2,40	Juvenile
103	LBU0527	K527	Rinca	Loh Buaya	05-Apr-04	00-064C-BEB8	71,90	171,60	6,70	Juvenile
104	LBU0536	K536	Rinca	Loh Buaya	05-Apr-04	00-064D-0E22	82,50	148,50	9,70	Subadult
105	LBU0537	K537	Rinca	Loh Buaya	05-Apr-04	00-064D-10D0	48,90	116,00	2,00	Juvenile

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
106	LBU0539	K539	Rinca	Loh Buaya	05-Apr-04	00-064E-45DB	54,15	124,70	2,85	Juvenile
107	LBU0540	K540	Rinca	Loh Buaya	06-Apr-04	00-064D-1AE8	65,95	137,20	4,65	Juvenile
108	LBU0544	K544	Rinca	Loh Buaya	05-Apr-04	00-064D-5508	127,35	241,90	47,50	Adult
109	LBU0546	K546	Rinca	Loh Buaya	08-Oct-04	00-064E-430F	54,45	130,70	2,30	Juvenile
110	LBU0547	K547	Rinca	Loh Buaya	07-Apr-04	00-064E-9DC9	111,10	228,90	33,50	Adult
111	LBU0549	K549	Rinca	Loh Buaya	07-Apr-04	00-064D-064F	54,20	121,60	3,10	Juvenile
112	LBU0551	K551	Rinca	Loh Buaya	09-Apr-04	00-064D-0420	70,45	158,00	6,55	Juvenile
113	LBU0555	K555	Rinca	Loh Buaya	10-Apr-04	00-064E-913A	127,85	252,30	50,00	Adult
114	LBU0700	K700	Rinca	Loh Buaya	06-Oct-04	00-0643-5586	82,20	180,90	9,20	Juvenile
115	LBU0702	K702	Rinca	Loh Buaya	06-Oct-04	00-064C-A604	52,50	125,50	4,60	Juvenile
116	LBU0703	K703	Rinca	Loh Buaya	07-Oct-04	00-0643-3C4B	57,25	98,60	2,65	Juvenile
117	LBU0704	K704	Rinca	Loh Buaya	05-Oct-04	00-063A-8381	98,95	208,30	24,50	Adult
118	LBU0721	K721	Rinca	Loh Buaya	06-Oct-04	00-0643-88A9	86,25	176,40	12,60	Subadult
119	LBU0729	K729	Rinca	Loh Buaya	21-Mar-05	00-064D-2209	83,00	192,40	10,80	Subadult
120	LBU0804	K804	Rinca	Loh Buaya	27-Mar-05	00-064D-17C7	110,20	213,50	23,80	Subadult
121	LBU0807	K807	Rinca	Loh Buaya	06-Oct-04	00-063A-50C3	134,85	270,00	54,00	Adult
122	LBU0809	K809	Rinca	Loh Buaya	23-Mar-05	00-064E-8DF4	61,60	152,20	3,98	Juvenile
123	LBU0810	K810	Rinca	Loh Buaya	07-Oct-04	00-064C-AD5F	38,40	87,20	0,75	Juvenile
124	LBU0958	K958	Rinca	Loh Buaya	19-Mar-05	00-064C-F41D	18,00	45,50	0,10	Hatchling
125	LBU0959	K959	Rinca	Loh Buaya	22-Mar-05	00-064C-B246	107,95	222,50	28,10	Adult
126	LBU0961	K961	Rinca	Loh Buaya	23-Mar-05	00-064D-3C5E	51,20	125,30	2,10	Juvenile
127	LBU0962	K962	Rinca	Loh Buaya	24-Mar-05	00-066F-E7CE	54,30	139,00	6,08	Juvenile
128	LBU0963	K963	Rinca	Loh Buaya	19-Mar-05	00-064E-211B	18,55	46,90	0,12	Hatchling
129	LBU0964	K964	Rinca	Loh Buaya	25-Mar-05	00-066B-EF97	43,15	112,10	1,33	Juvenile
130	LBU0965	K965	Rinca	Loh Buaya	23-Mar-05	00-0643-9DE9	80,10	155,50	8,93	Juvenile
131	LBU0966	K966	Rinca	Loh Buaya	19-Mar-05	00-064E-2CE0	18,05	43,90	0,09	Hatchling
132	LBU0969	K969	Rinca	Loh Buaya	19-Mar-05	00-064C-BA83	19,00	46,80	0,09	Hatchling

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
133	LBU0972	K972	Rinca	Loh Buaya	24-Mar-05	00-0670-089C	53,30	128,20	4,98	Juvenile
134	LBU0975	K975	Rinca	Loh Buaya	19-Mar-05	00-064C-AB45	17,65	42,90	0,09	Hatchling
135	LSE0003	K003B	Komodo	Loh Sebita	18-Feb-05	00-064D-4848	16,95	39,40	0,05	Hatchling
136	LSE0007	K007	Komodo	Loh Sebita	03-Nov-02	00-063A-2BA7	114,00	237,80	29,60	Adult
137	LSE0016	K016	Komodo	Loh Sebita	05-Nov-02	00-063A-24A2	92,00	211,00	19,80	Adult
138	LSE0024	K024	Komodo	Loh Sebita	04-Nov-02	00-0639-F3E7	107,00	219,20	21,20	Adult
139	LSE0044	K044	Komodo	Loh Sebita	02-Nov-02	00-063A-2E1E	102,00	222,50	21,60	Adult
140	LSE0047	K047	Komodo	Loh Sebita	04-Nov-02	00-063A-28B6	118,00	233,00	35,80	Adult
141	LSE0048	K048	Komodo	Loh Sebita	04-Nov-02	00-063A-ACFA	55,00	131,00	2,10	Juvenile
142	LSE0131	K131	Komodo	Loh Sebita	19-Mar-03	00-0643-41D6	18,25	33,80	0,07	Hatchling
143	LSE0133	K133	Komodo	Loh Sebita	20-Mar-03	00-0643-853F	18,75	44,70	0,10	Hatchling
144	LSE0135	K135	Komodo	Loh Sebita	19-Mar-03	00-0643-9F21	19,90	46,60	0,10	Hatchling
145	LSE0138	K138	Komodo	Loh Sebita	19-Mar-03	00-0643-A2D2	18,95	33,00	0,09	Hatchling
146	LSE0139	K139	Komodo	Loh Sebita	19-Mar-03	00-0643-95E4	18,60	44,00	0,10	Hatchling
147	LSE0140	K140	Komodo	Loh Sebita	19-Mar-03	00-0643-A49F	18,55	44,30	0,11	Hatchling
148	LSE0142	K142	Komodo	Loh Sebita	19-Mar-03	00-0643-A930	19,85	46,70	0,11	Hatchling
149	LSE0144	K144	Komodo	Loh Sebita	19-Mar-03	00-0643-A936	19,90	47,00	0,10	Hatchling
150	LSE0153	K153	Komodo	Loh Sebita	19-Mar-03	00-0643-7213	16,45	38,40	0,07	Hatchling
151	LSE0154	K154	Komodo	Loh Sebita	19-Mar-03	00-0643-9095	20,55	46,50	0,11	Hatchling
152	LSE0155	K155	Komodo	Loh Sebita	19-Mar-03	00-0643-7F21	19,90	45,20	0,11	Hatchling
153	LSE0156	K156	Komodo	Loh Sebita	20-Mar-03	00-0643-A71F	18,25	41,50	0,09	Hatchling
154	LSE0157	K157	Komodo	Loh Sebita	19-Mar-03	00-0643-7CA0	17,95	42,30	0,10	Hatchling
155	LSE0159	K159	Komodo	Loh Sebita	19-Mar-03	00-0643-A3BC	19,00	45,70	0,10	Hatchling
156	LSE0162	K162	Komodo	Loh Sebita	20-Mar-03	00-0643-A3B7	18,20	42,50	0,08	Hatchling
157	LSE0163	K163	Komodo	Loh Sebita	20-Mar-03	00-0643-8F93	19,60	49,00	0,09	Hatchling
158	LSE0168	K168	Komodo	Loh Sebita	20-Mar-03	00-0643-A404	20,00	47,30	0,13	Hatchling
159	LSE0172	K172	Komodo	Loh Sebita	20-Mar-03	00-0643-9DE7	19,20	45,60	0,11	Hatchling

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
160	LSE0670	K670	Komodo	Loh Sebita	11-Sep-04	00-064E-7DEF	106,40	120,50	20,80	Adult
161	LSE0677	K677	Komodo	Loh Sebita	12-Mar-05	00-064D-10CD	18,15	42,50	0,09	Hatchling
162	LSE0709	K709	Komodo	Loh Sebita	12-Sep-04	00-064C-E86C	66,50	159,70	4,40	Juvenile
163	LSE0711	K711	Komodo	Loh Sebita	13-Sep-04	00-064C-C378	93,25	198,20	11,85	Subadult
164	LSE0719	K719	Komodo	Loh Sebita	09-Sep-04	00-063A-2E1E	105,90	229,00	21,75	Adult
165	LSE0720	K720	Komodo	Loh Sebita	16-Sep-04	00-064D-0765	35,45	89,80	0,65	Juvenile
166	LLI0001	K001B	Komodo	Loh Liang	16-Feb-05	00-064C-B6D4	17,65	41,40	0,08	Hatchling
167	LLI0002A	K002A	Komodo	Loh Liang	28-Feb-03	00-063A-9592	154,75	304,50	81,50	Adult
168	LLI0002B	K002B	Komodo	Loh Liang	16-Feb-05	00-064D-2C40	19,35	45,90	0,09	Hatchling
169	LLI0005	K005B	Komodo	Loh Liang	02-Mar-03	00-063A-8015	72,25	167,50	6,20	Juvenile
170	LLI0017	K017	Komodo	Loh Liang	26-Oct-02	00-063A-22B2	98,00	194,00	18,80	Adult
171	LLI0036A	K036A	Komodo	Loh Liang	28-Feb-03	00-063A-81F6	153,50	276,00	105,00	Adult
172	LLI0040	K040	Komodo	Loh Liang	25-Oct-02	00-063A-309A	108,00	217,50	25,00	Subadult
173	LLI0041	K041	Komodo	Loh Liang	27-Feb-03	00-063A-0A09	152,00	296,00	42,00	Adult
174	LLI0045	K045	Komodo	Loh Liang	27-Mar-03	00-0639-FC93	76,00	155,00	8,20	Juvenile
175	LLI0049	K049	Komodo	Loh Liang	28-Feb-03	00-0639-E012	152,00	301,00	79,00	Adult
176	LLI0050	K050	Komodo	Loh Liang	06-Mar-03	00-063A-092F	100,50	216,00	20,85	Subadult
177	LLI0056	K056	Komodo	Loh Liang	28-Feb-03	00-063A-24B9	62,00	153,00	4,90	Juvenile
178	LLI0143	K143	Komodo	Loh Liang	15-Mar-03	00-063A-1342	108,50	221,60	18,80	Adult
179	LLI0164	K164	Komodo	Loh Liang	21-Mar-03	00-0643-9D9C	17,65	27,30	0,08	Hatchling
180	LLI0169	K169	Komodo	Loh Liang	21-Mar-03	00-0643-A53B	17,75	43,20	0,08	Hatchling
181	LLI0171	K171	Komodo	Loh Liang	21-Mar-03	00-0643-8F5E	19,10	47,40	0,11	Hatchling
182	LLI0178	K178	Komodo	Loh Liang	21-Mar-03	00-0643-84DB	18,80	44,00	0,10	Hatchling
183	LLI0190	K190	Komodo	Loh Liang	21-Mar-03	00-0643-737B	16,55	39,40	0,07	Hatchling
184	LLI0500	K500	Komodo	Loh Liang	06-Mar-04	00-0643-7E76	20,65	50,00	0,11	Hatchling
185	LLI0502	K502	Komodo	Loh Liang	17-Mar-04	00-064D-0CE2	62,35	149,70	4,15	Juvenile
186	LLI0503	K503	Komodo	Loh Liang	11-Mar-04	00-0643-88A1	41,60	104,70	1,25	Juvenile

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
187	LLI0506	K506	Komodo	Loh Liang	18-Mar-04	00-064C-BF58	143,35	275,80	62,30	Adult
188	LLI0513	K513	Komodo	Loh Liang	07-Mar-04	00-0643-4045	64,90	148,30	4,80	Juvenile
189	LLI0514	K514	Komodo	Loh Liang	10-Mar-04	00-0643-981A	105,10	220,00	26,00	Subadult
190	LLI0524	K524	Komodo	Loh Liang	17-Mar-04	00-064D-0DD6	20,65	51,20	0,11	Hatchling
191	LLI0526	K526	Komodo	Loh Liang	18-Mar-04	00-064C-CD4D	80,55	177,50	10,05	Juvenile
192	LLI0532	K532	Komodo	Loh Liang	13-Mar-04	00-0643-6085	51,25	128,20	2,05	Juvenile
193	LLI0662	K662	Komodo	Loh Liang	18-Aug-04	00-064D-1238	105,35	204,90	22,60	Adult
194	LLI0674	K674	Komodo	Loh Pinda	17-Aug-04	00-064E-1FBD	126,05	252,00	50,00	Subadult
195	LLI0905	K905	Komodo	Loh Liang	28-Feb-05	0-064D-4C0E	18,30	44,00	0,10	Hatchling
196	LLI0906	K906	Komodo	Loh Liang	04-Mar-05	00-063C-1383	20,35	50,00	0,11	Post-Hatchling
197	LLI0907	K907	Komodo	Loh Liang	28-Feb-05	00-063D-DEFB	18,60	42,30	0,08	Hatchling
198	LLI0908	K908	Komodo	Loh Liang	02-Mar-05	00-063D-B3B7	51,35	129,40	2,05	Juvenile
199	LLI0909	K909	Komodo	Loh Liang	10-Mar-05	00-063B-AC73	55,45	135,10	2,40	Juvenile
200	LLI0911	K911	Komodo	Loh Liang	05-Mar-05	00-064C-BBB0	53,65	133,50	2,10	Juvenile
201	LLI0916	K916	Komodo	Loh Liang	09-Mar-05	00-064D-2055	56,15	135,50	2,90	Juvenile
202	LLI0917	K917	Komodo	Loh Liang	28-Feb-05	00-063D-958E	65,75	153,00	5,05	Juvenile
203	LLI0919	K919	Komodo	Loh Liang	02-Mar-05	00-0643-42D9	75,00	162,60	7,75	Subadult
204	LLI0921	K921	Komodo	Loh Liang	10-Mar-05	00-066D-7650	61,80	150,70	4,20	Juvenile
205	LLI0922	K922	Komodo	Loh Liang	09-Mar-05	00-064C-DBCE	56,35	140,00	3,25	Juvenile
206	LLI0923	K923	Komodo	Loh Liang	06-Mar-05	00-066D-6E67	63,40	153,70	4,55	Juvenile
207	LLI0926	K926	Komodo	Loh Liang	11-Mar-05	00-064D-4AD3	51,10	122,70	2,40	Juvenile
208	LLI0927	K927	Komodo	Loh Liang	12-Mar-05	00-064E-13D9	17,50	42,30	0,09	Hatchling
209	LLA0003	K003A	Komodo	Loh Lawi	30-Oct-02	00-063A-A5B3	80,00	182,20	8,40	Juvenile
210	LLA0011	K011	Komodo	Loh Lawi	01-Nov-02	00-063A-8639	101,00	209,00	21,20	Adult
211	LLA0023	K023	Komodo	Loh Lawi	03-Jun-05	00-063A-75E0	87,60	195,90	12,95	Subadult
212	LLA0026	K026	Komodo	Loh Lawi	30-Oct-02	00-063A-8B03	52,00	128,10	1,80	Juvenile
213	LLA0035	K035	Komodo	Loh Lawi	31-Oct-02	00-063A-7F1E	101,00	215,00	16,20	Subadult

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
214	LLA0271	K271	Komodo	Loh Lawi	10-Jun-03	00-0643-A798	49,65	121,20	1,60	Juvenile
215	LLA0272	K272	Komodo	Loh Lawi	02-Jun-05	00-0643-9F50	136,40	273,80	51,30	Adult
216	LLA0273	K273	Komodo	Loh Lawi	11-Jun-03	00-0643-A2DE	143,55	276,00	56,00	Adult
217	LLA0274	K274	Komodo	Loh Lawi	11-Jun-03	00-0643-7AFF	94,55	203,80	13,40	Subadult
218	LLA0275	K275	Komodo	Loh Lawi	11-Jun-03	00-0643-723A	127,75	253,00	34,80	Adult
219	LLA0282	K282	Komodo	Loh Lawi	08-Jun-03	00-0643-81A7	139,00	278,50	60,80	Adult
220	LLA0288	K288	Komodo	Loh Lawi	07-Jun-03	00-0643-95AC	43,50	109,00	1,15	Juvenile
221	LLA0545	K545	Komodo	Loh Lawi	06-Jun-04	00-064E-8F8B	48,90	119,60	1,95	Juvenile
222	LLA0550	K550	Komodo	Loh Lawi	02-Jun-05	00-0643-97B8	59,45	141,50	2,68	Juvenile
223	LLA0556	K556	Komodo	Loh Lawi	01-Jun-04	00-063A-75E0	82,95	188,00	13,30	Subadult
224	LLA0561	K561	Komodo	Loh Lawi	03-Jun-04	00-064C-AEAA	39,85	86,20	1,20	Juvenile
225	LLA0563	K563	Komodo	Loh Lawi	06-Jun-04	00-063A-A5B3	81,25	186,70	10,15	Subadult
226	LLA0564A	K564A	Komodo	Loh Lawi	02-Jun-05	00-064E-7405	52,85	129,10	1,98	Juvenile
227	LLA0564B	K564B	Komodo	Loh Lawi	06-Jun-04	00-064E-80BA	43,75	115,00	1,45	Juvenile
228	LLA0566	K566	Komodo	Loh Lawi	08-Jun-04	00-064D-3380	118,25	251,00	35,00	Adult
229	LLA0569	K569	Komodo	Loh Lawi	05-Jun-04	00-064C-D708	42,90	111,00	1,30	Juvenile
230	LLA0570	K570	Komodo	Loh Lawi	04-Jun-04	00-064C-AF81	65,10	151,00	5,15	Juvenile
231	LLA0574	K574	Komodo	Loh Lawi	07-Jun-04	00-064C-AD82	133,50	275,00	51,40	Adult
232	LLA0579	K579	Komodo	Loh Lawi	05-Jun-04	00-064E-8630	45,30	114,50	1,70	Juvenile
233	LLA0585	K585	Komodo	Loh Lawi	08-Jun-04	00-064D-234F	81,25	182,00	11,65	Subadult
234	LLA0588	K588	Komodo	Loh Lawi	03-Jun-05	00-064E-AE67	52,75	127,50	21,13	Juvenile
235	LLA0673	K673	Komodo	Loh Lawi	14-Aug-04	00-064C-C528	51,65	129,40	1,80	Juvenile
236	LLA0675	K675	Komodo	Loh Lawi	14-Aug-04	00-064C-ADD9	100,25	216,00	21,65	Adult
237	LBA0243	K243	Rinca	Loh Baru	12-Jul-03	00-0643-99D1	105,75	183,40	27,30	Subadult
238	LBA0244	K244	Rinca	Loh Baru	12-Jul-03	00-0643-967E	28,20	70,90	0,30	Post Hatchling
239	LBA0249	K249	Rinca	Loh Baru	29-Jun-05	00-0643-40C9	130,60	269,60	42,40	Adult
240	LBA0255	K255	Rinca	Loh Baru	11-Jul-03	00-0643-8EE3	65,85	155,80	4,75	Juvenile

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
241	LBA0257	K257	Rinca	Loh Baru	11-Jul-03	00-0643-362E	41,95	109,00	1,20	Juvenile
242	LBA0621	K621	Rinca	Loh Baru	29-Jun-04	00-064D-470D	68,15	158,40	5,95	Juvenile
243	LBA0622	K622	Rinca	Loh Baru	26-Jun-05	00-064C-CE4B	61,90	145,90	3,88	Juvenile
244	LBA0623	K623	Rinca	Loh Baru	02-Jul-04	00-064C-F0EB	82,95	178,40	11,40	Subadult
245	LBA0625	K625	Rinca	Loh Baru	28-Jun-04	00-0643-9A19	100,50	207,00	20,10	Subadult
246	LBA0626	K626	Rinca	Loh Baru	29-Jun-04	00-064D-3DA9	79,85	177,60	8,40	Juvenile
247	LBA0627	K627	Rinca	Loh Baru	01-Jul-05	00-064D-34D4	83,55	185,00	11,60	Subadult
248	LBA0628	K628	Rinca	Loh Baru	28-Jun-04	00-064C-C892	44,65	111,60	0,95	Juvenile
249	LBA0629	K629	Rinca	Loh Baru	29-Jun-04	00-064E-5DC9	121,65	249,50	41,60	Adult
250	LBA0630	K630	Rinca	Loh Baru	30-Jun-04	00-064C-B432	69,45	165,30	6,05	Juvenile
251	LBA0631	K631	Rinca	Loh Baru	27-Jun-05	00-064C-EF16	60,70	150,20	3,95	Juvenile
252	LBA0633	K633	Rinca	Loh Baru	25-Jun-05	00-0643-91A3	69,60	155,00	6,73	Juvenile
253	LBA0635	K635	Rinca	Loh Baru	06-Jul-04	00-064D-4424	90,65	172,50	16,85	Subadult
254	LBA0636	K636	Rinca	Loh Baru	30-Jun-04	00-0643-8E73	69,15	106,40	5,00	Juvenile
255	LBA0637	K637	Rinca	Loh Baru	01-Jul-05	00-064E-5978	88,95	187,60	14,90	Subadult
256	LBA0638	K638	Rinca	Loh Baru	24-Jun-05	00-0643-7AE8	60,30	150,20	3,73	Juvenile
257	LBA0639	K639	Rinca	Loh Baru	28-Jun-04	00-0643-877F	116,80	229,30	34,00	Adult
258	LBA0640	K640	Rinca	Loh Baru	25-Jun-05	00-064E-BF38	99,00	215,50	21,55	Subadult
259	LBA0641	K641	Rinca	Loh Baru	29-Jun-04	00-0643-89A5	75,85	173,30	8,80	Juvenile
260	LBA0643	K643	Rinca	Loh Baru	04-Jul-04	00-0643-8FAF	111,95	236,50	27,90	Subadult
261	LBA0644	K644	Rinca	Loh Baru	02-Jul-04	00-064C-DA8F	138,65	280,00	64,40	Adult
262	LBA0645	K645	Rinca	Loh Baru	06-Jul-04	00-0643-6FAB	117,60	247,20	35,00	Adult
263	LBA0646	K646	Rinca	Loh Baru	02-Jul-04	00-064D-1FEA	69,65	165,50	6,55	Juvenile
264	LBA0647	K647	Rinca	Loh Baru	06-Jul-04	00-064D-1962	94,15	203,00	16,80	Subadult
265	LBA0649	K649	Rinca	Loh Baru	30-Jun-05	00-064C-A904	65,95	137,00	4,03	Juvenile
266	LBA0650	K650	Rinca	Loh Baru	02-Jul-04	00-064E-5308	111,00	229,90	33,95	Adult
267	LBA0651	K651	Rinca	Loh Baru	30-Jun-04	00-064C-AEA1	78,30	167,70	9,45	Juvenile

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
268	LBA0654	K654	Rinca	Loh Baru	04-Jul-04	00-0643-8CDF	60,50	147,80	3,30	Juvenile
269	LBA0656	K656	Rinca	Loh Baru	02-Jul-05	00-064D-0BB1	74,40	169,50	7,03	Juvenile
270	LBA0657	K657	Rinca	Loh Baru	30-Jun-04	00-064D-679D	88,90	196,30	13,70	Subadult
271	LBA0658	K658	Rinca	Loh Baru	04-Jul-04	00-064C-AFDC	46,40	114,50	1,25	Juvenile
272	LBA0659	K659	Rinca	Loh Baru	02-Jul-04	00-064C-C434	100,00	197,50	21,45	Subadult
273	LBA0661	K661	Rinca	Loh Baru	05-Jul-04	00-064D-2261	59,00	134,60	3,35	Juvenile
274	LBAV043	V043	Rinca	Loh Baru	24-Jun-05	00-0643-962B	69,05	138,30	5,68	Juvenile
275	LBAV046	V046	Rinca	Loh Baru	27-Jun-05	00-066D-5ACF	44,30	112,60	1,40	Juvenile
276	LBAV047	V047	Rinca	Loh Baru	26-Jun-05	00-066C-ABBD	44,15	115,00	1,38	Juvenile
277	LBAV049	V047	Rinca	Loh Baru	26-Jun-05	00-066D-6D3F	121,75	249,00	34,10	Adult
278	LBAV050	V050	Rinca	Loh Baru	25-Jun-05	00-066F-F404	58,80	142,50	3,10	Juvenile
279	LBAV051	V051	Rinca	Loh Baru	25-Jun-05	00-066F-F791	133,85	265,50	37,50	Adult
280	LBAV052	V052	Rinca	Loh Baru	27-Jun-05	00-0670-018B	69,80	164,60	7,08	Subadult
281	LBAV053	V053	Rinca	Loh Baru	30-Jun-05	00-0648-29B9	96,20	217,70	20,80	Subadult
282	LBAV054	V054	Rinca	Loh Baru	27-Jun-05	00-066D-76AB	47,15	118,50	1,50	Juvenile
283	LBAV055	V055	Rinca	Loh Baru	25-Jun-05	00-0670-DFCB	51,85	130,90	2,33	Juvenile
284	LBAV056	V056	Rinca	Loh Baru	26-Jun-05	00-0643-54F8	69,40	164,00	5,98	Juvenile
285	LBAV057	V057	Rinca	Loh Baru	29-Jun-05	00-066D-5E54	53,50	129,40	2,63	Juvenile
286	LBAV058	V058	Rinca	Loh Baru	25-Jun-05	00-066D-65F0	67,75	161,50	4,73	Juvenile
287	LBAV060	V060	Rinca	Loh Baru	27-Jun-05	00-066B-E51E	53,90	131,60	2,60	Juvenile
288	LBAV061	V061	Rinca	Loh Baru	29-Jun-05	00-066B-FDDC	55,65	135,00	2,38	Juvenile
289	LBAV062	V062	Rinca	Loh Baru	29-Jun-05	00-066D-4722	43,95	112,40	1,43	Juvenile
290	LBAV064	V064	Rinca	Loh Baru	30-Jun-05	00-0648-5E2F	48,55	118,10	1,60	Juvenile
291	LBAV066	V066	Rinca	Loh Baru	03-Jul-05	00-0643-7E06	63,25	153,00	4,72	Juvenile
292	LBAV068	V068	Rinca	Loh Baru	02-Jul-05	00-0643-7E17	63,65	150,80	5,08	Juvenile
293	LBAV069	V069	Rinca	Loh Baru	02-Jul-05	00-0634-9527	49,20	116,40	1,63	Juvenile
294	LBAV070	V070	Rinca	Loh Baru	02-Jul-05	00-0643-93D4	66,95	157,20	5,33	Juvenile

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
295	LTK0250	K250	Rinca	Loh Tongker	05-Jul-03	00-0643-6FA9	53,40	140,00	3,30	Juvenile
296	LTK0251	K251	Rinca	Loh Tongker	07-Jul-03	00-0643-9F98	97,20	204,20	18,25	Subadult
297	LTK0259	K259	Rinca	Loh Tongker	07-Jul-03	00-0643-771D	92,00	203,00	14,70	Subadult
298	LTK0260	K260	Rinca	Loh Tongker	05-Jul-03	00-0643-886D	91,20	197,20	13,50	Subadult
299	LTK0262	K262	Rinca	Loh Tongker	18-Jun-05	00-0643-83BD	65,25	19,25	5,28	Subadult
300	LTK0600	K600	Rinca	Loh Tongker	20-Jun-04	00-0643-74FA	64,15	154,00	4,60	Juvenile
301	LTK0603	K603	Rinca	Loh Tongker	20-Jun-04	00-064D-3858	104,20	212,50	26,55	Subadult
302	LTK0606	K606	Rinca	Loh Tongker	20-Jun-04	00-064C-F78E	53,70	133,10	2,60	Juvenile
303	LTK0610	K610	Rinca	Loh Tongker	20-Jun-04	00-064D-217C	96,25	204,50	17,30	Subadult
304	LTK0618	K618	Rinca	Loh Tongker	23-Jun-04	00-064E-3134	47,80	115,00	2,00	Juvenile
305	LTKV014	V014	Rinca	Loh Tongker	18-Jun-05	00-0669-62D9	40,55	99,40	1,15	Juvenile
306	LTKV019	V019	Rinca	Loh Tongker	18-Jun-05	00-066C-6194	40,40	105,20	1,05	Juvenile
307	LTKV021	V021	Rinca	Loh Tongker	19-Jun-05	00-066C-8755	43,65	113,00	1,30	Juvenile
308	LTKV026	V026	Rinca	Loh Tongker	19-Jun-05	00-0669-96EC	51,05	115,70	2,10	Juvenile
309	LTKV040	V040	Rinca	Loh Tongker	22-Jun-05	00-066F-EFBD	123,75	261,40	39,50	Adult
310	LTKV041	V041	Rinca	Loh Tongker	21-Jun-05	00-066D-664F	54,95	133,10	2,70	Juvenile
311	LTKV042	V042	Rinca	Loh Tongker	22-Jun-05	00-066B-FE39	43,00	106,50	1,05	Juvenile
312	LTKV044	V044	Rinca	Loh Tongker	22-Jun-05	00-066D-72A0	87,65	179,70	14,40	Subadult
313	LTKV048	V048	Rinca	Loh Tongker	23-Jun-05	00-066D-6076	52,40	120,80	2,28	Juvenile
314	LDS0221	K221	Rinca	Loh Dasami	28-Apr-03	00-0643-7B1E	69,15	157,00	4,80	Juvenile
315	LDS0520	K520	Rinca	Loh Dasami	26-Mar-04	00-064D-267D	49,15	123,00	2,00	Juvenile
316	LDS0521	K521	Rinca	Loh Dasami	29-Mar-04	00-064C-CACB	70,50	150,00	6,05	Juvenile
317	LDS0528	K528	Rinca	Loh Dasami	01-Apr-04	00-064D-3E2D	51,95	126,70	2,20	Juvenile
318	LDS0529	K529	Rinca	Loh Dasami	31-Mar-04	00-064E-361E	49,15	121,10	1,75	Juvenile
319	LDS0530	K530	Rinca	Loh Dasami	29-Mar-04	00-064C-D0FC	103,35	222,00	21,00	Subadult
320	LDS0533	K533	Rinca	Loh Dasami	26-Mar-04	00-0643-7B11	47,10	113,80	1,35	Juvenile
321	LDS0538	k538	Rinca	Loh Dasami	01-Apr-04	00-064C-B7E6	75,00	167,90	7,60	Juvenile

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
322	LDS0542	K542	Rinca	Loh Dasami	31-Mar-04	00-064D-4CE3	103,25	222,80	22,30	Subadult
323	LDS0543	K543	Rinca	Loh Dasami	30-Mar-04	00-064C-DA15	56,75	131,00	3,79	Juvenile
324	LDS0974	K974	Rinca	Loh Dasami	14-Apr-05	00-066B-E6E1	55,05	133,90	2,48	Juvenile
325	LDS0976	K976	Rinca	Loh Dasami	15-Apr-05	00-066D-7C7B	42,45	106,80	1,08	Juvenile
326	LDS0977	K977	Rinca	Loh Dasami	12-Apr-05	00-0670-1076	61,35	141,70	4,38	Juvenile
327	LDS0981	K981	Rinca	Loh Dasami	15-Apr-05	00-066C-C0C9	47,95	121,20	1,53	Juvenile
328	GMO0001	K001A	Gili Motang	Gili Motang	27-Nov-02	00-063A-1ADA	67,00	161,50		Subadult
329	GMO0004	K004	Gili Motang	Gili Motang	26-Nov-02	00-063A-2A68	70,00	162,00	6,80	Juvenile
330	GMO0015	K015	Gili Motang	Gili Motang	25-Nov-02	00-063A-8D35	66,00	155,00	4,75	Juvenile
331	GMO0020	K020	Gili Motang	Gili Motang	22-Nov-02	00-063A-0672	90,00	199,00	12,60	Subadult
332	GMO0029	K029	Gili Motang	Gili Motang	27-Nov-02	00-063A-A15F	73,00	172,00	5,80	Juvenile
333	GMO0031	K031	Gili Motang	Gili Motang	22-Nov-02	00-0639-D89D	99,00	148,00	11,40	Subadult
334	GMO0032	K032	Gili Motang	Gili Motang	27-Nov-02	00-063A-97D9	61,00	152,00	3,45	Juvenile
335	GMO0034	K034	Gili Motang	Gili Motang	24-Nov-02	00-063A-8F2D	82,00	187,50	11,20	Subadult
336	GMO0043	K043	Gili Motang	Gili Motang	02-Dec-02	00-063A-865A	72,00	166,50	5,80	Juvenile
337	GMO0051	K051	Gili Motang	Gili Motang	28-Nov-02	00-063A-8739	81,00	191,00	10,80	Subadult
338	GMO0052	K052	Gili Motang	Gili Motang	29-Nov-02	00-063A-3597	62,00	153,50	3,30	Juvenile
339	GMO0053	K053	Gili Motang	Gili Motang	24-Nov-02	00-063A-A76E	86,00	192,00	11,60	Subadult
340	GMO0054	K054	Gili Motang	Gili Motang	29-Nov-02	00-063A-11EF	81,00	151,30	9,20	Juvenile
341	GMO0055	K055	Gili Motang	Gili Motang	25-Nov-02	00-063A-0080	104,00	225,70	21,30	Subadult
342	GMO0058	K058	Gili Motang	Gili Motang	29-Nov-02	00-063A-8B14	72,00	171,70	6,60	Juvenile
343	GMO0059	K059	Gili Motang	Gili Motang	24-Nov-02	00-0639-ED13	71,00	174,00	7,00	Juvenile
344	GMO0060	K060	Gili Motang	Gili Motang	27-Nov-02	00-063A-11D5	77,00	177,50	7,00	Juvenile
345	GMO0061	K061	Gili Motang	Gili Motang	29-Nov-02	00-063A-213D	77,00	176,50	7,00	Juvenile
346	GMO0062	K062	Gili Motang	Gili Motang	26-Nov-02	00-063A-293D	76,00	171,50	6,30	Juvenile
347	GMO0063	K063	Gili Motang	Gili Motang	25-Nov-02	00-063A-04B8	94,00	203,20	15,50	Subadult
348	GMO0064	K064	Gili Motang	Gili Motang	28-Nov-02	00-063A-917F	66,00	158,50	4,70	Juvenile

No.	Sequence	Field Code	Island	Location	Sampling Date	Tag Number	SVL	TL	Weight	Age
349	GMO0069	K069	Gili Motang	Gili Motang	01-Dec-02	00-063A-06F0	82,00	174,00	11,20	Subadult
350	GMO0676	K676	Gili Motang	Gili Motang	24-Aug-04	00-064D-02AE	51,65	130,30	1,80	Juvenile
351	GMO0710	K710	Gili Motang	Gili Motang	29-Aug-04	00-064C-E807	81,75	189,80	10,80	Subadult
352	GMO0713	K713	Gili Motang	Gili Motang	31-Aug-04	00-064C-FD26	117,25	239,30	24,40	Subadult
353	GMO0718	K718	Gili Motang	Gili Motang	01-Sep-04	00-064C-CE6D	68,65	158,90	5,35	Juvenile
354	GMO0723	K723	Gili Motang	Gili Motang	26-Aug-04	00-064E-1E48	99,60	216,00	18,60	Subadult
355	NSK0263	K263	Nusa Kode	Nusa Kode	30-Jun-03	00-0643-71F4	72,35	170,60	7,00	Juvenile
356	NSK0265	K264	Nusa Kode	Nusa Kode	30-Jun-03	00-0643-4047	70,90	158,80	5,80	Juvenile
357	NSK0266	K266	Nusa Kode	Nusa Kode	28-Jun-03	00-0643-3BF9	85,80	164,00	10,35	Subadult
358	NSK0267	K267	Nusa Kode	Nusa Kode	27-Jun-03	00-0643-968F	95,35	205,10	15,85	Subadult
359	NSK0268F	K268	Nusa Kode	Nusa Kode	27-Jun-03	00-0643-93E4	70,10	164,20	7,60	Juvenile
360	NSK0269F	K269	Nusa Kode	Nusa Kode	26-Jun-03	00-0643-7E75	72,00	119,70	6,80	Juvenile
361	NSK0982	K982	Nusa Kode	Nusa Kode	16-Apr-05	00-066D-4CE2	70,05	161,50	5,33	Juvenile
362	NSK0984	K984	Nusa Kode	Nusa Kode	17-Apr-05	00-0671-0044	60,55	113,90	3,58	Juvenile
363	NSK0985	K985	Nusa Kode	Nusa Kode	19-Apr-05	00-066F-F89A	68,80	151,50	5,83	Juvenile
364	NSK0987	K987	Nusa Kode	Nusa Kode	16-Apr-05	00-066B-E126	69,55	164,10	5,93	Juvenile
365	NSK0988	K988	Nusa Kode	Nusa Kode	17-Apr-05	00-066B-F949	81,50	182,00	9,78	Subadult
366	NSK0990	K990	Nusa Kode	Nusa Kode	17-Apr-05	00-066D-705D	47,40	120,20	1,58	Juvenile

Appendix B: Genomic DNA extraction protocol

Phenol-Chloroform-IsoAmyl-Alcohol Method

Extraction Buffer

Final mix	Stock solution
Tris-HCl 100 mM	1 M pH 8.0
EDTA 5 mM	0.5 M pH 8.0
NaCl 100 mM	5 M
SDS 0.5%	10%

PCIAA solution (Phenol-Chloroform-IsoAmyl Alcohol)

Phenol	:25 proportion
Chloroform	:24 proportion
IsoAmyl Alcohol	:1 proportion

TE solution (Tris-HCl EDTA)

Tris-HCl 100 mM

EDTA 1 mM

1. Prepare digestion mix in a 1.5 ml microcentrifuge tube: 600 µl Extraction Buffer, 90 µl blood sample, and 10 µl Proteinase K
2. Put the tube on shaking incubator at 37°C and leave overnight
3. Prepare in a new 1.5 ml microcentrifuge tube: 700 µl digestion mix (1.) and 700 µl PCIAA solution
4. Shake mix 3. for 1 minute, then centrifuge for 5 minute at 13,000 rpm.
5. Prepare in a new 1.5 ml microcentrifuge tube: supernatant of mix 3. and a volume of Chloroform-IsoAmyl Alcohol
6. Shake mix 5. For 30 seconds and centrifuge for 5 minutes at 13,000 rpm
7. Prepare in a new 2.0 ml microcentrifuge tube: 1/3 supernatant of mix 5. and double volume of cold ethanol 100%
8. Mix gently and centrifuge for 3 minutes at 13,000 rpm. Check for DNA pellet

9. If DNA pellet is visible, throw supernatant and dry pellet. Resuspend in 100 µl TE
10. If DNA pellet is invisible, add 30 µl NaCl 3M and mix
11. Centrifuge mix 10. for 3 minutes at 12,000 rpm
12. Throw supernatant and add 500 µl alcohol 70% and mix
13. Centrifuge mix 12. for 3 minutes at 12,000 rpm
14. Throw away supernatant and dry pellet. Resuspend in 100 µl TE

Appendix C: Polymerase Chain Reaction (PCR) protocol

Cocktail mix (volume per PCR tube: 9 μl)

Final mix	Stock solution
H_2O	
Taq Buffer 1X	10X
$MgCl_2$ 1 mM	50 mM
KCRL Forward 0.25 μM	10 μM
KCRL Reverse 0.25 μM	10 μM
dNTPS 10 μM	10 mM
Taq Recombinant 0.5 U	5U/μl

Primers

KCRL Forward	GCGCCTATTTTCTCCTATTCCT
KCRL Reverse	GGGAGGGTTCTTGTAGTTGAAG

Product size: 939 bp

PCR Conditions

Denaturation	95°C for 5 minutes	
Denaturation	95°C for 30 seconds	⎫
Annealing	52°C for 30 seconds	⎬ X 35
Elongation	72°C for 1 minute	⎭
Elongation	72°C 10 minutes	

Appendix D: Gel Electrophoresis protocol

Agarose gel 1% (36 well)

TBE 5X 100 ml

Agarose powder 1 g

Ethidium Bromide 0.5 μl

1. Cast hot Agarose mix on gel dock for 36 wells for about 30 minutes
2. Place dock with gel on electrophoretic machine
3. Prepare for mixing: 3 μl PCR product and 2 μl loading dye (Blue di Bromophenol)
4. Inject mix 3. in wells but leave a few wells for DNA ladder (100 bp interval)
5. Connect with electricity source and run 100 V for about 30 minutes
6. Check gel with ultraviolet light and save picture

Appendix E: Automated Sequencing

Preparation protocols:

1) DNA purification, cycle sequencing, and product precipitation with Sodium acetate

2) Direct cycle sequencing with diluted DNA template and product precipitation with Isopropanol

Protocol 1

***DNA purification with* Exosap™**

1. Prepare in 0.2 ml PCR tube: 2.5 µl PCR product and 1 µl Exosap™ enzyme
2. Centrifuge mix to remove bubbles
3. Load in thermo cycler (PCR machine)

Purification PCR Conditions

Start	37°C 15 minutes
Purification	80°C 15 minutes
Cooling	4°C ∞

Cycle Sequencing

Final mix

2.5 µl Better Buffer (BSA, dNTPs, Mg)

1.5 µl Primer (KCRL For/ Rev)

0.5 µl Big Dye

1-1.5 µl H_2O

Volume of mix per sample: 5.5 µl (volume of purified DNA: 1.5-2 µl)

Cycle sequencing PCR Conditions

Denaturation	96°C for 10 seconds	⎫
Annealing	50°C for 5 seconds	⎬ X 35
Elongation	60°C for 4 minutes	⎭
Cooling	4°C ∞	

Precipitation of cycle sequencing PCR products

1. Prepare in 0.2 ml PCR tube: add 1.5 µl Milli Q water to the 7.5 µl cycle sequencing PCR product

2. Add 2.5 µl EDTA 100 mM and 2.0 µl Sodium acetate 3 M
3. Prepare in 1.5 ml tube: transfer all 13.5 µl from 1. and 2.
4. Add 53 µl ethanol 95% (kept at -20°C) and shake 3-4 times
5. Incubate at room temperature for 15 minutes
6. Throw away supernatant
7. Add 70 µl ethanol 70% (kept at -20°C) and shake 3-4 times
8. Centrifuge at 13,000 for 5 minutes
9. Throw away supernatant
10. Leave pellet in oven at 60°C to eliminate all remaining ethanol
11. Resuspend pellet in 15 µl High Density Formamide (HDF)
12. Transfer resuspended samples to 96-well sequencing plate
13. Put rubbery lid on
14. Denature at 95°C for 3 minutes in thermo cycler
15. Put sequencing plate on cooling rack for at least 5 minutes before transferring to sequencer

Protocol 2

Cycle Sequencing of diluted PCR product (1:10)

Final mix

2 µl Buffer 5X

1 µl Primer (KCRL For/ Rev) 3.2 µM

0.5 µl Big Dye v 3.1?

1.5 µl H_2O

Volume of mix per sample: 5 µl (volume of DNA: 5 µl)

Cycle sequencing PCR Conditions

Denaturation	96°C for 10 seconds	⎫
Annealing	50°C for 5 seconds	⎬ X 35
Elongation	60°C for 4 minutes	⎭
Cooling	4°C ∞	

Precipitation of cycle sequencing PCR products with isopropanol

1. Add 90 µl of 63% isopropanol to each sample

2. Vortex for 30 seconds
3. Precipitate for 15 minutes at room temperature
4. Centrifuge (Eppendorf 5804) at 3,700 rpm for 45 minutes
5. Tip over plate on paper towel to remove supernatant
6. Place inverted plate on paper towel in centrifuge and spin at 1,000 for 5 seconds
7. Add 150% isopropanol and pour off immediately
8. Repeat step 6
9. Dry pellet at 40°C for 4 minutes in thermo cycler
10. Add 15 μl High Density Formamide (HDF)
11. Vortex for 30 seconds
12. Denature at 95°C for 3 minutes
13. Place on ice for at least 5 minutes
14. Load samples on sequencer plate

Running samples in sequencer (ABI 3100 Genetic Analyser)

1. Check polymer supply in the large syringe (one run of 16 samples needs 50 μl polymer)
2. Prepare to run sequencer by clicking on "Run 3100 data Collection"
3. Wait until all four buttons on the Service Console turn green
4. Open GA 3100 Plate Manager to create file folder and label all samples
5. Check Pin Valve: closed: to add polymer
   opened: to run samples
6. Check buffer (1X TBE)
7. Put sequencing plates with samples on A and B trays accordingly
8. Run analysis
9. Retrieve sequences by checking signals and electropherograms
10. Transfer sequences to a disk

Appendix F: Alignment programme

CodonCode Aligner™ version 2.0

Codon Code Corporation 2002-2007 available at http://www.codoncode.com/aligner/index.htm

This user-friendly alignment programme is available for use in Windows and Mac OS X, where sequence assembly and editing can be done by referring to known sequence (sequence reference) as well as sequence traces. By referring to known sequence, mutational changes can be detected as red-marked bases. CodonCode Aligner™ generates consensus sequence automatically using sequence quality, which is considered to be more accurate than automatic editing based on sequence majority. This feature saves time for manual editing, because one has to look only at regions that have poor quality. The demo mode of CodonCode Aligner™ is fully functional, except for saving and printing. A licensed server is therefore necessary to accommodate these functions.

I found multiple deletions of 22 bp or 58 bp at the beginning of some sequences e.g. FLN02 in the picture below. These dashes represent deletions that are aligned and not missing bases resulted from reading errors commonly found at sequence ends. Two important features in this programme to note regarding identification of chunk deletions at beginning of sequences are:

1. The programme uses local alignment to build Contigs (instead of end-to-end alignment). Thus, bases with poor signals at both ends of sequences have no influence in the alignment process.
2. Bases of unaligned ends are marked in grey and ignored when calculating consensus, although one can clip these sites manually before assembly.

The first window in the picture shows an assembly being analysed i.e., two sequences of forward and reverse FLN02 and two sequence references i.e. Control Region I of Komodo dragon and a sequence that include the whole cytochrome b gene, tRNA that encodes Threonine, the CRI, tRNA that encodes Glutamate, and ND6 gene all downloaded from GenBank.

The second window shows bases contents of those sequences assembled for analysis. Both FLN02 sequences show 22 bp-deletion and these deletions are aligned with both reference sequences. Please notice that these sites are not marked in grey, as in the case of unaligned bases. These deletion sites are also included in the consensus

sequence (Contig 1). The third window shows the traces of both FLN02 sequences. The vertical straight blue line shows a base chosen that is located just before the CRI starts, i.e. within $tRNA^{Thr}$. The two reference sequences were removed from the assembly and the sites beyond CRI were excluded from the analysis.

Appendix G: Softwares for phylogenetic and phylogeographic analyses

Phylogenetic Analyses

PAUP* (Phylogenetic Analyses Using Parsimony and other methods)

Swofford, David L. (2001) PAUP* version 4.0b10. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts

PAUP* is a software package for inference of evolutionary trees for use in Macintosh, Windows, UNIX/VMS, or DOS-based formats. Input file of PAUP* is in NEXUS format, which can be created using nucleotide sequence editors such as CodonCode Aligner, ClustalX2, and BioEdit. Commands are categorised into blocks of TAXA, CHARACTERS, ASSUMPTIONS, SETS, TREES, CODONS, and DISTANCES. In Windows application, one has to type commands based on these blocks ended with a semicolon. Using Macintosh, one can simply choose from the menu.

Parsimony principle: preference for the least complex explanation for an observation (In phylogenetics, the preferred phylogenetic tree is the one that requires the smallest number of evolutionary changes)

A simplified pipeline to generate an evolutionary tree based on parsimony on PAUP* in Windows is given below:

1) Creating NEXUS file of sequences data with all sequences identifiers (names) that will appear on the tree

2) Executing file to retrieve basic information on the data

3) Logging analysis into an output file

4) Setting assumptions e.g., character weight, sequence deletion, character exclusion

5) Searching for trees with default parsimony method (heuristic search)

6) Saving and displaying trees (also possibly printing in low resolution)

7) Creating consensus tree

8) Running bootstrap analysis for statistical support

Maximum Likelihood principle: the probability of nucleotide substitution gives rise to the topology that has the maximum likelihood given an appropriate model

To generate an evolutionary tree based on maximum likelihood method on PAUP* in Windows, follow these steps:

1) Creating NEXUS file of sequences data with all sequences identifiers (names) that will appear on the tree with maximum likelihood block generated using ModelTest

2) Executing file to retrieve basic information on the data

3) Logging analysis into an output file

4) Change optimality criterion to likelihood

5) Searching for trees

6) Saving and displaying trees (also possibly printing in low resolution)

7) Creating consensus tree

8) Running bootstrap analysis for statistical support

MrBayes

Huelsenbeck, John P. and Fredrik Ronquist (2001) MrBayes version 3.1.2 Bayesian inference of phylogenetic trees. Bioinformatics 17: 754-755

MrBayes is a programme for the Bayesian inference of phylogeny that uses command line interface and can be used in all platforms. It is a free programme that is optimised for speed and can be downloaded from http://morphobank.ebc.uu.se/mrbayes. Input file for MrBayes is standard Nexus file of aligned nucleotide or amino acid sequences. The two most important commands in MrBayes are “lset” that sets the parameters of the likelihood model and “mcmc” that sets parameters of Markov Chain Monte Carlo (MCMC) as well as initiates the analysis. By default, there are four output files automatically generated by MrBayes, i.e., summaries of parameters from two runs (.p files) and two tree files (.t files). Consensus tree is filed under .con file, which can be opened using tree viewing programme such as FigTree or TreeView.

Bayesian Theorem: Relation between conditional (posterior) and marginal (prior) probabilities of stochastic events of A and B and is expressed in the statement below:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

The phylogenetic tree chosen using this method is the one that has the maximum posterior probability given a prior

Simplified commands for running MrBayes are described below:

1) Creating Nexus file

2) Executing input file on MrBayes

3) Logging analysis to a file

4) Setting parameters of likelihood model

5) Setting parameters of MCMC

6) Running analysis

7) Saving parameter values and trees

8) Retrieval cladograms and phylograms

jModelTest

Posada, David (2008) jModelTest: phylogenetic model averaging. Molecular Biology and Evolution 25: 1253-1256

I used jModelTest to select a substitution model for phylogenetic analysis using Maximum Likelihood method. The programme implements five different selection tests i.e., Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Decision theory performance-based Test (DT), and hierarchical and dynamical Likelihood Ratio Tests (hLRT and dLRT) to choose the most appropriate model of evolution. The programme supersedes ModelTest, a model selector programme with 56 different models to assess. jModelTest has 88 models that result from combination of eleven nucleotide substitution schemes and base frequencies, rates of substitution among sites and proportion of invariable sites. Details on the selection tests are given below.

Hierarchical Likelihood Ratio Test (hLRT) is a statistic to test the goodness of fit. Models are tested in pairs and hierarchically nested. The algorithm of this test is given below:

$$\delta = -2 \log \Delta \qquad (1)$$

$$\Delta = \frac{\max [\, L_0 \,(\text{NullModel} \mid \text{Data})]}{\max [\, L_1 \,(\text{AlternativeModel} \mid \text{Data})]} \qquad (2)$$

where L_0: likelihood under Null hypothesis (a simple model), L_1: likelihood under alternative hypothesis (a more complex model), and Δ is always equal to or greater than zero, since L_1 is always equal to or higher than L_0

Chosen model is the one that follows χ^2 distribution, therefore:

$\delta = -2 \log \Delta$

$\delta = -2 \log (L_0 / L_1)$

$\delta = 2 (\ln L_0 - \ln L_1)$

δ = df (degrees of freedom) of χ^2 distribution

df: # of parameters free for the alternative model can have (difference in # of parameters between Null- and Alternative models)

But, estimate of χ^2 used to represent the distribution of LRT is problematic, in that it only represents one side of the likelihood. It has been suggested that using 50% of each likelihood i.e., of Null- and Alternative models (χ_0^2 and χ_1^2) may solve the problem.

Akaike Information Criterion (AIC) is an unbiased estimator of information quantity (Kullback-Leibler), which is a measure of lost information when a model is used to approximate full reality. Basically, the test estimates the distance between models and sample. A minimum value of AIC means shorter distance from full reality. Algorithm:

$AIC_i = -2 \ln L_i + 2 K_i$

where L_i= max. likelihood of data and K_i= number of free parameters of i-th model

To compare and rank candidate models, one can simply draw differences between given AIC values and the minimum. Interpretation:

Δi= 1-2, models have substantial support

Δi= 3-7, models have less support

Δi>10, models have no support

Bayesian Information Criterion (BIC) estimates information quantity as AIC does, but penalizes overparameterized data heavier than AIC does.

$BIC_i = -2 \ln L_i + 2 K_i \log n$

where L_i= max. likelihood of data and K_i= number of free parameters of *i*-th model and *n*= number of characters, e.g. characters within sequences

Decision theory performance-based Test (DT) is an approach to select a model based on its phylogenetic performance, which is measured as the expected error on the branch length estimates weighted by their BIC. The best model chosen using this approach is the one that minimises a risk function and generally results in a better accuracy in branch length estimates.

Simplified steps for testing models on MacIntosh:

1. Execute input file (e.g., Komodo_1998.nex) on PAUP*, log on, and check for base frequency
2. Execute Modelblock (.nex) on PAUP* to create raw log likelihood from 56 models incorporated in the block
3. Retrieve likelihood file "model.scores" and change file name into input file name i.e., Komodo_1998.scores
4. Open ModelTest and choose input file in Step 3. and output format to appear on Console
5. Retrieve model(s) chosen by hLRTs and AIC
6. Proceed on using MrModelTest to choose model using BIC

Phylogeographic inference

ARLEQUIN

Excoffier, L., G. Laval and S. Schneider (2005) Arlequin version 3.1: An integrated software package population genetics data analysis. Evolutionary Bioinformatics Online 1: 47-50

Arlequin is software for extracting genetic and demographic information from samples of several data types, either in haplotypic or genotypic form. The basic data types are: DNA sequences, RFLP data, microsatellite data, standard data, and allele frequency data. The input file for Arlequin is called "Project File" and one can create a Project File from scratch or use Project File Wizard to create an outline of a Project File. To keep track all operations during Arlequin session, a log file

(Arlequin_log.txt) is created automatically and will be displayed on a web browser when Internet connection is available.

The general procedure for data analyses using Arlequin is as follows:

1) Project File setting

Project File is set up using Arlequin format (*.arp) containing profile and data sections. Each data type requires different setting (examples available in the user's manual); however nucleotide sequences can be directly analysed when treated as standard data.

2) Project loading in Arlequin and calculation setting

A project has to be manually loaded and Arlequin will show if a project is either valid or invalid. When a project is valid, one can see the properties on the project tab, whereas when a project is invalid, one can see a warning in the result directory. Calculation can be set using calculation menu, which includes general settings for polymorphism control and haplotype definition, as well as settings for genetic structure, haplotype inference, and linkage disequilibrium.

3) Project running and result directory retrieval

Once a project is running, one can stop at any time. Nevertheless, a complete result of a large file requires only about half a second and this can immediately be found as a file with the same name of the project name with *.res extension. Results can also be viewed at any time as a web file for example Firefox or Internet Explorer document, regardless of Internet connection.

Some statistical algorithms used in Arlequin:

1) Gene diversity (H) and variance of gene diversity V(H)

$$\hat{H} = \frac{n}{n-1}\left(1 - \sum_{i=1}^{k} p_i^2\right)$$

$$\mathrm{V}(\hat{H}) = \frac{2}{n(n-1)}\left\{2(n-2)\left[\sum_{i=1}^{k} p_i^3 - \left(\sum_{i=1}^{k} p_i^2\right)^2\right] + \sum_{i=1}^{k} p_i^2 - \left(\sum_{i=1}^{k} p_i^2\right)^2\right\}$$

n= the total number of sequences in a population, *k*=the number of haplotypes, and p_i= the sample frequency of *i*-th haplotype. For haploid data such as mtDNA data, gene diversity can be used as a measure of haplotype diversity within a population. Defined haplotypic mtDNA sequences given in the Arlequin's haplotype database generated using DNA data file format are used to create haplotype frequency data file format. I used this haplotype frequency data file to compute haplotype diversity among sequences within sampling locations.

2) Mean number of pairwise differences among sequences (π) for a given population and variance of mean pairwise differences among sequences V (π)

$$\hat{\pi} = \frac{n}{n-1}\sum_{i=1}^{k}\sum_{j=1}^{k} p_i p_j \hat{d}_{ij}$$

$$\mathrm{V}(\hat{\pi}) = \frac{3n(n+1)\hat{\pi} + 2(n^2+n+3)\hat{\pi}^2}{11(n^2-7n+6)}$$

where *n*= the total number of sequences in a population, *k*= the number of haplotypes, p_i and p_j= the sample frequency of respective *i*-th and *j*-th haplotype, and d_{ij}= the estimate of mutations having occurred since the divergence of haplotype *i* from haplotype *j*. For each population, only polymorphic nucleotide sites are taken into the calculation. Thus, mean pairwise differences among sequences measures the net differences among sequences within a population.

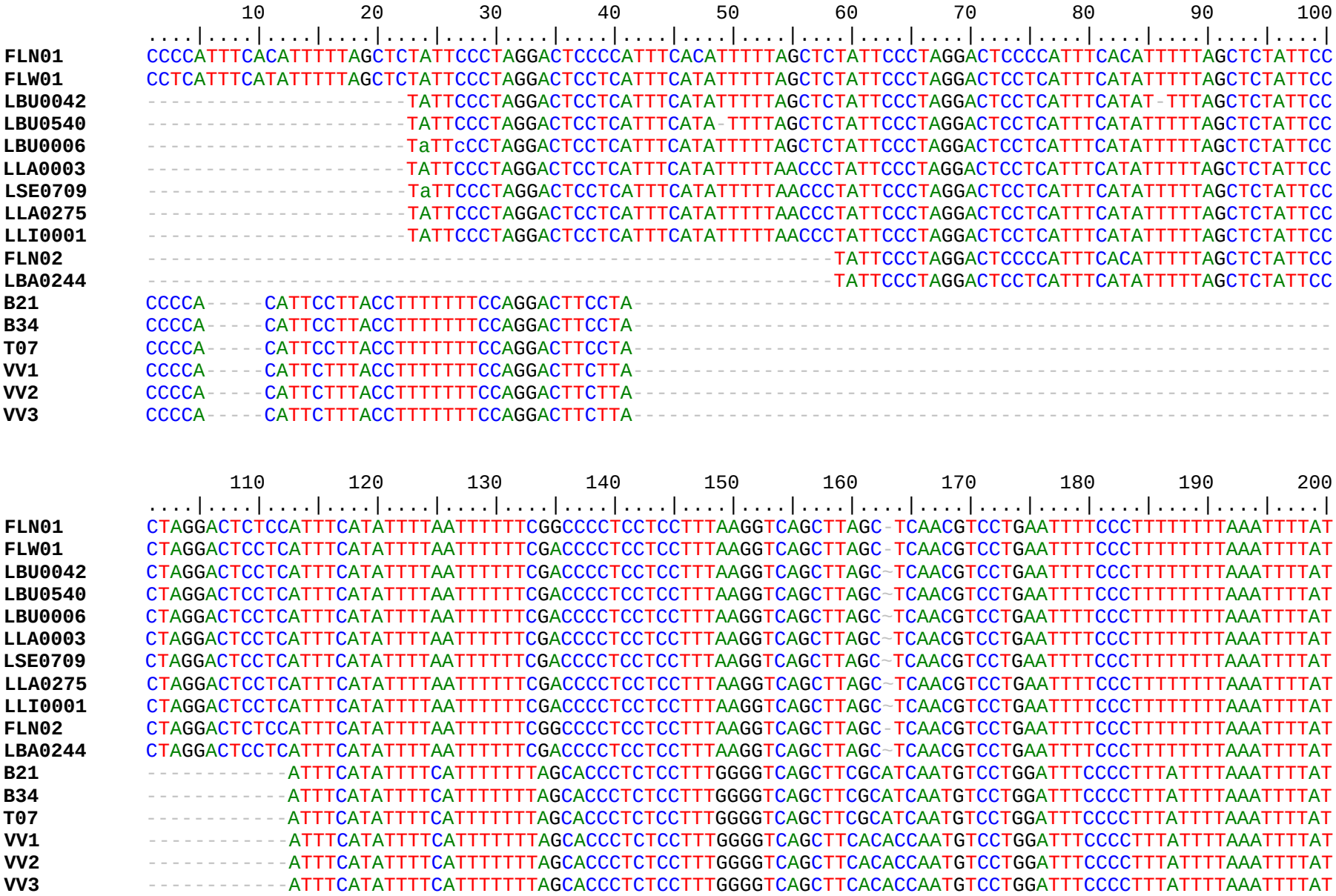
3) Nucleotide diversity (π_n) or average gene diversity across all usable nucleotide sites for a given sample and variance of nucleotide diversity V (π_n)

$$\hat{\pi}_n = \frac{\sum_{i=1}^{k}\sum_{j<i} p_i p_j \hat{d}_{ij}}{L}$$

$$\mathrm{V}(\hat{\pi}_n) = \frac{n+1}{3(n-1)L}\hat{\pi}_n + \frac{2(n^2+n+3)}{9n(n-1)}\hat{\pi}_n^2$$

where *n*= the total number of sequences in a population, *k*= the number of haplotypes, p_i and p_j= the sample frequency of respective *i*-th and *j*-th haplotype, d_{ij}= the estimate of mutations having occurred since the divergence of haplotype *i* from haplotype *j*, and *L*= the number of all usable nucleotide sites (missing data are not considered in the analysis).

Appendix H: Multiple Sequence Alignment



```
                210       220       230       240       250       260       270       280       290       300
           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
FLN01      GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTC
FLW01      GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTC
LBU0042    GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTC
LBU0540    GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTC
LBU0006    GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTC
LLA0003    GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTC
LSE0709    GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTC
LLA0275    GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTC
LLI0001    GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTC
FLN02      GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTC
LBA0244    GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATATCGTCCCGCCAACAGCATTTCATATTTTTGTACGTCTATTAAACTCGATTTAAATGTCATTTTC
B21        GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATTTCGTCCCGCCAACGACATTTCAAATTTTTGTACGTCTCTTAAACTCGATTTAAGTGTCATTTTT
B34        GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATTTCGTCCCGCCAACGACATTTCAAATTTTTGTACGTCTCTTAAACTCGATTTAAGTGTCATTTTT
T07        GTCTTTCGAGCAACCAAACGCGCCACCTCCTGGATTTCGTCCCGCCAACGACATTTCAAATTTTTGTACGTCTCTTAAACTCGATTTAAGTGTCATTTTT
VV1        GTCTTTCGAGCAACCAAACGCGTCACCTCCTGGATTTCGTCCCGCCAACGACATTTCAAATTTTTGTACGTCTTTTAAACTCGATTTAAGTGTCATTTTT
VV2        GTCTTTCGAGCAACCAAACGCGTCACCTCCTGGATTTCGTCCCGCCAACGACATTTCAAATTTTTGTACGTCTTTTAAACTCGATTTAAGTGTCATTTTT
VV3        GTCTTTCGAGCAACCAAACGCGTCACCTCCTGGATTTCGTCCCGCCAACGACATTTCAAATTTTTGTACGTCTTTTAAACTCGATTTAAGTGTCATTTTT

                310       320       330       340       350       360       370       380       390       400
           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
FLN01      AAGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
FLW01      AAGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
LBU0042    AAGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
LBU0540    AAGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
LBU0006    AAGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
LLA0003    AAGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
LSE0709    AAGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
LLA0275    AAGACACTCAACTAAGCACTGGCTACCCCTATCGATTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
LLI0001    AAGACACTCAACTAAGCACTGGCTACCCCTATCGATTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
FLN02      AAGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
LBA0244    AAGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGATCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
B21        AAGACACTCAACTAAGCACTGGCTACCCCTATCGATTCGCCTACTGTTACCAGTCTCGTGGTCCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
B34        AAGACACTCAACTAAGCACTGGCTACCCCTATCGATTCGCCTACTGTTACCAGTCTCGTGGTCCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
T07        AAGACACTCAACTAAGCACTGGCTACCCCTATCGATTCGCCTACTGTTACCAGTCTCGTGGTCCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
VV1        AAGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGTCCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
VV2        AAGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGTCCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
VV3        AAGACACTCAACTAAGCACTGGCTACCCCTATCGGTTCGCCTACTGTTACCAGTCTCGTGGTCCATACCTATAGGTTGCACCTATTTAATGACCTTTCCA
```

```
                410       420       430       440       450       460       470       480       490       500
            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
FLN01       ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
FLW01       ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
LBU0042     ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
LBU0540     ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
LBU0006     ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
LLA0003     ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
LSE0709     ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
LLA0275     ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
LLI0001     ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
FLN02       ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
LBA0244     ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
B21         ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
B34         ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
T07         ATACCTCTGGTTGTGAGGCCCAGGGACTTTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
VV1         ATACCTCTGGTTGTGAGGCCCAGGGACTCTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
VV2         ATACCTCTGGTTGTGAGGCCCAGGGACTCTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC
VV3         ATACCTCTGGTTGTGAGGCCCAGGGACTCTCTTTCAAGGTGACCACTCTTTTCTCTCTAAAGCACTTCGGGTTGGGTGAATCTCAGGAGCTTATCACCTC

                510       520       530       540       550       560       570       580       590       600
            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
FLN01       ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACC
FLW01       ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACT
LBU0042     ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACT
LBU0540     ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACT
LBU0006     ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACT
LLA0003     ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACT
LSE0709     ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACT
LLA0275     ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATTGCATGTTCGTCCGGGGTCAGCTTTACATTTACT
LLI0001     ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACT
FLN02       ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACC
LBA0244     ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTTTTTTTTTGGGTGGATCTCAGATCGCATGTTCGTCCGGGGTCAGCTTTACATTTACT
B21         ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCTTTTTCTCTTTTTTTTTAGGTGGATCTCAGATAGCATGTTCGTCCGGGGTCAGCTTTACATTCATT
B34         ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCTTTTTCTCTTTTTTTTTAGGTGGATCTCAGATAGCATGTTCGTCCGGGGTCAGCTTTACATTCATT
T07         ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCTTTTTCTCTTTTTTTTTAGGTGGATCTCAGATAGCATGTTCGTCCGGGGTCAGCTTTACATTCATT
VV1         ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTCTTTTTTTAGGTGGATCTCAGATAGCATGTTCGTCCGGGGTCAGCTTTACATTCATC
VV2         ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTCTTTTTTTAGGTGGATCTCAGATAGCATGTTCGTCCGGGGTCAGCTTTACATTCATC
VV3         ATAACTACGGTCACCCTGGTATAGGCGCCTTCAGCCTTTTTCTCTTTTTTTAGGTGGATCTCAGATAGCATGTTCGTCCGGGGTCAGCTTTACATTCATC
```

```
                610       620       630       640       650       660       670       680       690       700
           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
FLN01      AAACGATCATTGTGACAATCGAACCCTTTATAATTACATTGGAGCTGGAATTTAATGGTCGCCGGACATACAAAAAATCAAAAAAA-CATAATTTTTTAA
FLW01      AAACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAA-CATAACTTTTTAA
LBU0042    AAACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAA~CATAACTTTTTAA
LBU0540    AAACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAA~CATAACTTTTTAA
LBU0006    AAACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAA~CATAactttttaa
LLA0003    AAACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAA~CATAATTTTTTAA
LSE0709    AAACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAA~CATAACTTTTTAA
LLA0275    AAACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAA~CATAATTTTTTAA
LLI0001    AAACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAA~CATAATTTTTTAA
FLN02      AAACGATCATTGTGACAATCGAACCCTTTATAATTACATTGGAGCTGGAATTTAATGGTCGCCGGACATACAAAAAATCAAAAAAA-CATAATTTTTTAA
LBA0244    AAACGATCATTGTGACAATCGAACCCTTTATAATTACGTTGGAGCTGGAATTTAATGGTCGCCGGACATATAAAAAATCAAAAAAA~CATAACTTTTTAA
B21        AAACGATCATTGTGACAACCGAACCCTTTATAATTACGTTGGAGCTGGATTTTAATGGTCGCCGGACATATAAAAATCAAAAAAAATCATAATTTTT---
B34        AAACGATCATTGTGACAACCGAACCCTTTATAATTACGTTGGAGCTGGATTTTAATGGTCGCCGGACATATAAAAATCAAAAAAAATCATAATTTTT---
T07        AAACGATCATTGTGACAACCGAACCCTTTATAATTACGTTGGAGCTGGATTTTAATGGTCGCCGGACATATAAAAATCAAAAAAAATCATAATTTTT---
VV1        AAACGATCATTGTGACAACCGAACCCTTTATAATTACGTTGGAGCTGGATTTTAATGGTCGCCGGACATATAAAAATCAAAAAAAATCATAATTTTT---
VV2        AAACGATCATTGTGACAACCGAACCCTTTATAATTACGTTGGAGCTGGATTTTAATGGTCGCCGGACATATAAAAATCAAAAAAAATCATAATTTTT---
VV3        AAACGATCATTGTGACAACCGAACCCTTTATAATTACGTTGGAGCTGGATTTTAATGGTCGCCGGACATATAAAAATCAAAAAAAATCATAATTTTT---

                710       720       730
           ....|....|....|....|....|....|.
FLN01      AAAACCCCCAAACCCCCtACACT--CCCCAT
FLW01      AAAACCCCCAAACCCCCTACACT--CCCCAT
LBU0042    AAAACCCCCAAACCCCCTACACT~~CCCCAT
LBU0540    AAAACCCCCAAACCCCCTACACT~~CCCCAT
LBU0006    aaaaccccccaaaccccctacact~~ccccat
LLA0003    AAAACCCCCAAACCCCCTACACT~~CCCCAT
LSE0709    AAAACCCCCAAACCCCCTACACT~~CCCCAT
LLA0275    AAAACCCCCAAACCCCCTACACT~~CCCCAT
LLI0001    AAAACCCCCAAACCCCCTACACT~~CCCCAT
FLN02      AAAACCCCCAAACCCCCTACACT--CCCCAT
LBA0244    AAAACCCCCAAACCCCCTACACT~~CCCCAT
B21        AAAACCCCCTTTCCCCCAACATTTTCCCCAC
B34        AAAACCCCCTTTCCCCCAACATTTTCCCCAC
T07        AAAACCCCCTTTCCCCCAACATTTTCCCCAC
VV1        AAAACCCCCTTTCCCCCAACATTTCTCTCAC
VV2        AAAACCCCCTTTCCCCCAACATTTCTCTCAC
VV3        AAAACCCCCTTTCCCCCAACATTTCTCTCAC
```

Appendix I: List of Haplotypes and Their Collapsed Sequences

No	Haplotype	Sequence Reference	Length	Intraspecific indels (site #)	Collapsed Sequences
1	H1	LLI0001	705	1-22	LLI002A, B, 0017, 0036A, 0040-41, 0050, 0056. 0143, 0169, 0171, 0178, 0190, LLI0500, 0506, 0513-14, 0524, 0532, 0662, 0905-09, 0917, 0919, 0921, 0923, LLI0926-27; LWA01-02, 04, 06, 08; LWE03-07; LLA0011, 0023, 0273, 0288, 0545, 0556, 0564A, 0574, 0673; LSE0016, 0024, 0044, 0047-48, 0131, 0133, 0135, 0138-0140, 0142, 0144, LSE0153-0157, 0159, 0162-63, 0168, 0172, 0677, 0719-20
2	H2	LLA0003	705	1-22	LLA0026, 0035, 0271, 0274, 0550, 0563, 0564B, 0566, 0569-70, 0585, 0588, LLA0675; LWE01-02; LLI0005, 0502-03, 0526, 0674, 0911, 0916, 0922; LSE0711
3	H3	LLA0275	705	1-22	LLA0282, 0561, 0579; LSE0003, 0007, 0670; LLI0049, 0164; LWA03, 05, 07; LBU0411
4	H4	LSE0709	705	1-22	-
5	H5	LBU0006	705	1-22	LBU0005A, 0010, 0012, 0018-19, 0021-22, 0028, 0033, 0039, 0188A, B, 0189, LBU0191-94, 0197, 0199, 0204, 0206, 0208, 0210-15, 0218-19, 0522, 0527, 0537, LBU0539, 0544, 0546, 0549, 0555, 0700, 0703, 0721, 0729, 0804, 0807, 0809-10, LBU0958-59, 0961, 0963, 0966, 0969, 0972, 0975; FLW07, 12, 17, 19-21, 24, 27; GMO0001, 0004, 0015, 0020, 0029, 0031-32, 0034, 0043, 0051-55, 0058-64, GMO0069, 0676, 0710, 0713, 0718, 0723;

					LDS0538, 0542; LBA0243, 0249, 0255, 0623, 0625-30, 0633, 0635, 0639-40, 0644-45, 0650, 0654, 0656-59, LBAV043, V047, V050, V052-53, V055-56, V058, V061-62, V064, V066, V068-70; LTK0250, 0262, 0600, 0603, 0610, 0618, V014, V040, V042, V048
6	H6	FLW01	727	-	FLW01, 03-06, 08-11, 13-16, 18, 22-23, 25, 32, 34, 36 LBU0025, 0046, 0962; LBA0622, LBAV051, V054; LTK0251, 0606, LTKV019, V021, V041, V044
7	H7	LBA0244	699	1-58	LBA0257, 0621, 0631, 0636-38, 0641, 0643, 0646-47, 0649, 0651, 0661, LBAV046, V049, V057, V060; LBU0013, 0183, 0186-87, 0505, 0519, 0536, 0551, 0702, 0704, 0964-65 LDS0221, 0520-21, 0528-30, 0533, 0543, 0974, 0976, 0977, 0981; LTK0259-60; NSK0263, 0265-67, 0268F, 0269F, 0982, 0984-85, 0987-88, 0990
8	H8	LBU0042	705	1-22, 81	-
9	H9	LBU0540	705	1-22, 49	-
10	H10	FLN01	727	-	FLN01, 04-06; FLW26
11	H11	FLN02	699	1-58	FLN02, 03

Appendix J: Haplotype frequency distribution in populations, islands, and regions

Island	Sampling location	Western (4)				Central (5)					Eastern (2)		Total per sampling location	Total per island
		H1 (705)	H2 (705)	H3 (705)	H4 (705)	H5 (705)	H6 (727)	H7 (699)	H8 (705)	H9 (705)	H10 (727)	H11 (699)		
Komodo	Loh Wenci	5	2	0	0	0	0	0	0	0	0	0	7	117
	Loh Sebita	26	1	3	1	0	0	0	0	0	0	0	31	
	Loh Liang	33	8	2	0	0	0	0	0	0	0	0	43	
	Loh Lawi	9	15	4	0	0	0	0	0	0	0	0	28	
	Loh Wau	5	0	3	0	0	0	0	0	0	0	0	8	
Rinca	Loh Buaya	0	0	1	0	56	3	12	1	1	0	0	74	165
	Loh Baru	0	0	0	0	37	3	18	0	0	0	0	58	
	Loh Dasami	0	0	0	0	2	0	12	0	0	0	0	14	
	Loh Tongker	0	0	0	0	10	7	2	0	0	0	0	19	
Flores	Flores North	0	0	0	0	0	0	0	0	0	4	2	6	45
	Flores West	0	0	0	0	18	20	0	0	0	1	0	39	
Nusa Kode		0	0	0	0	0	0	12	0	0	0	0	12	12
Gili Motang		0	0	0	0	27	0	0	0	0	0	0	27	27
Total per haplotype		78	26	13	1	123	33	83	1	1	5	2	366	
Total per region		118				241					7			

Appendix K: Geographic coordinates for sampling locations

No.	Location/Island	South	East	Photo Reference
1	Pota (FLN) / Flores	8° 20’ 19.56”	120° 45’ 41.72”	-
2	Riung (FLN)/ Flores	8° 25’ 8.60”	121° 1’ 57.36”	Pelk, A.
3	Wae Wuul (FLW)/ Flores	8° 35’ 42.28”	119° 49’ 54.00”	Seno, Aganto
4	Loh Buaya (LBU)/ Rinca	8° 39’ 25.02”	119° 43’ 24.41”	Seno, Aganto
5	Loh Baru (LBA)/ Rinca	8° 43’ 42.13”	119° 41’ 34.64”	Seno, Aganto
6	Loh Tongker (LTK)/ Rinca	8° 45’ 14.05”	119° 43’ 27.42”	Seno, Aganto
7	Loh Dasami (LDS)/ Rinca	8° 47’ 16.66”	119° 40’ 17.43”	Seno, Aganto
8	Nusa Kode (NSK)/ Nusa Kode	8° 25’ 8.60”	119° 49’ 54.00”	Seno, Aganto
9	Gili Motang (GMO)/ Gili Motang	8° 47’ 45.83”	119° 46’ 55.45”	Seno, Aganto
10	Loh Sebita (LSE)/ Komodo	8° 32’ 7.75”	119° 32’ 55.07”	Seno, Aganto
11	Loh Wenci (LWE)/ Komodo	8° 31’ 10.91”	119° 25’ 54.84”	-
12	Loh Liang (LLI)/ Komodo	8° 34’ 14.69”	119° 29’ 47.90”	Seno, Aganto
13	Loh Lawi (LLA)/ Komodo	8° 36’ 11.11”	119° 24’ 20.76”	Seno, Aganto
14	Loh Wau (LWA)/ Komodo	8° 42’ 1.80”	119° 26’ 24.35”	Seno, Aganto

Appendix L: Divergence Time Estimation

BEAST

Drummond, A. J. and A. Rambaut. 2007. BEAST: Bayesian Evolutionary Analysis by Sampling Trees. BMC Evolutionary Biology 7:214

Bayesian Evolutionary Analyses by Sampling Trees (BEAST) is a computer programme for reconstructing phylogenies as well as testing evolutionary hypotheses. Divergence time is also of concern for molecular dating in speciation process as well as measurable evolving population for example, ancient DNA. Input file for this programme is an XML (Extended Mark-up Language) that can be created using BEAUti (Bayesian Evolutionary Analysis Utility) from a Nexus file. There are six (6) steps to create a BEAST XML file that can be done by working on six (6) panels appearing in BEAUti:

1) Importing a Nexus file that will appears in Data panel
2) Setting up taxon subsets from sequence data in Taxa panel
3) Setting a substitution model, site heterogeneity model, and molecular rate variation model in Model panel
4) Setting priors for all the parameters in the model in Prior panel
5) Tuning and weighting parameter values for Markov Chain Monte Carlo (MCMC) chain in Operator panel
6) Setting the number of generations the MCMC algorithm will run in MCMC panel

Once BEAST XML input file is created, one can subsequently run this file in BEAST. However, a tree created using MrBayes can be used as a starting tree in BEAST by a simple copy-and-paste of the Newick tree from .con file in MrBayes to the BEAST XML file. A slight adjustment in the part that describes the starting tree is necessary. Further, BEAST output can be viewed using LogCombiner or Tracer that are included in the software package.

Acknowledgements

This thesis is another step in my life that I have truly indebted from my parents for letting me reaching higher and farther. I extend special thanks to my colleagues Sri Sulandari and M. Syamsul Arifin Zein at Genetics Laboratory of MZB-LIPI whose endorsement and friendship have made my study possible. I would also pass my gratitude to my supervisors, Wolfgang Böhme and Steven Perry for all their support during the whole period of my study. I thank my co-supervisor, Claudio Ciofi, who has let me messing around in his laboratory in Florence, Italy. I thank Ursula Bott for her efficiency, which helped me a lot in dealing with the all-weather German bureaucracy and Chiara Natali for having babysat me during my stay in Italy. Besides, a bunch of thanks to those who have made my life as a tropical student in Germany easier and more interesting: André Koch, to whom I have asked a series of silly to impossible questions; Thomas Wesener, for keeping me updated with many happenings at ZFMK and beyond; Amir Hamidy, whose unbeatable patience of listening to my complaints should win a Nobel Prize for peace; Patrick "Bobo" Tchoua, for his friendship that has inspired me for being *scientifique sans frontier*, and Côme Linsoussi for each and every critical question he asked me. I thank Yanti Mirdayanti and her family for helping me getting my English back in track and Christian Kleißle for helping me improving my German for survival. Karen Meusemann, Julia L. Schwarzer, and Carola Greve helped me coping with scientific and non-scientific life in and out of ZFMK. Barbara Albertoni and Jeffrey S. Hannam helped me see the different faces of Europe. Besides, I owe Pungki Lupiningdyah for sharing some quality time of living in "harsh" Europe. Finally, I extend my most sincere and unabridged gratitude to Jeremy A. Miller for sharing every bits and pieces of passionately living nerdy with attitude. Financial support for me was provided by the German Agency for Academic Exchange (DAAD) under the Fellowship for Young Academics and Scientists scheme, whereas Laboratory of Evolutionary Biology, University of Florence, Italy and Genetics Laboratory at MZB-LIPI, Cibinong, Indonesia made materials and laboratory resources available for me. Wolfgang Peter made specimens of the Lace monitor available for me from the Tiergarten Straubing in Bavaria, Germany.