Historical Biogeography of Reptiles and Amphibians from the Lesser Sunda Islands of Indonesia

By

Sean Bryant Reilly

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Integrative Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jimmy A. McGuire, Chair
Professor Rauri C. Bowie
Professor Rosemary G. Gillespie

Fall 2016

Historical Biogeography of Reptiles and Amphibians from the Lesser Sunda Islands of Indonesia

by Sean Bryant Reilly

Abstract

Historical Biogeography of Reptiles and Amphibians from the Lesser Sunda Islands of Indonesia

by

Sean Bryant Reilly

Doctor of Philosophy in Integrative Biology

University of California, Berkeley

Professor Jimmy A. McGuire, Chair

The Lesser Sunda Archipelago, also known as Nusa Tenggara, lies in the southeastern portion of Indonesia and extends between Bali in the west, and New Guinea in the east. While the Lesser Sundas themselves are oceanic islands that have never been land bridged to a continent the islands on either side do. Bali and the other Greater Sunda Islands of Java, Sumatra, and Borneo become periodically land bridged with Asia during glacial maxima forming the Sunda Shelf. New Guinea and Aru become periodically land bridged to Australia during glacial maxima and form the Sahul Shelf. Given their current orientation, the Lesser Sundas may act as 'stepping stones' for animals and plants dispersing between the Sunda and Sahul Shelves and may act as a two-way filter for organisms dispersing between two of the world's great biogeographical realms. Alfred Russel Wallace's discovery of a pattern of clinal mixture of species from different biogeographical realms was a key insight leading to his identification of the Wallace Line and to his creation of the field of biogeography. Even though the Lesser Sundas played a critical role in the development of the field, this region has received little subsequent attention from historical biogeographers and our current understanding of Lesser Sunda biogeography has only modestly improved relative to what was known at the time of Wallace. The reptiles and amphibians of the Lesser Sundas represent a particularly interesting group of vertebrates from a biogeographical standpoint because they appear to show distributional patterns that are most consistent with a stepping-stone model of island colonization caused by the two-way filter zone. In Chapter 1, I review the geological and biogeographical literature for the Lesser Sundas and use these sources to formulate hypotheses concerning the colonization of the archipelago by rafting terrestrial vertebrates. In Chapters 2 through 4, I investigate the possibility that flying lizards, forest skinks, and fanged frogs have colonized the archipelago in a stepping-stone manner using a phylogenomic approach (using sequence data from mtDNA and hundreds of nuclear loci) whereby the relationships among island-specific lineages can be used to infer the sequence of island colonization. Flying lizards of the genus *Draco* form a monophyletic group that colonized the western Inner Arc islands of Lombok or Sumbawa from the Sunda Shelf around 10 million years ago when Lombok and Sumbawa first became land-positive. *Draco* continued expanding eastward through the Inner Arc until they reached Lembata, while a series of dispersal events from Flores south to Sumba, east to Timor, north to Wetar, west to Alor, and finally west to Pantar (the island immediately west of Lembata). The islands of Sumbawa and Flores contain multiple non-sister lineages that are parapatrically distributed and are exchanging migrants

within an island. Forest Skinks of the genus *Sphenomorphus* show relatively little morphological divergence across their range yet exhibit large levels of genetic divergence. The oldest lineages of *Sphenomorphus* within the Lesser Sundas occur on the islands of Lombok and Flores and they expanded eastward through the Inner Arc until they reached Pantar. But rather than reaching Alor from neighboring Pantar, *Sphenomorphus* dispersed from Flores south to Sumba, then east to Timor, Alor, and Wetar. There are multiple non-sister lineages of *Sphenomorphus* on Lombok, Flores, and Sumba, and estimates of migration between lineages within each island suggest that these lineages are not interbreeding. Fanged frogs of the genus *Limnonectes* have colonized the Inner Arc of the Lesser Sundas from the Sunda Shelf. It is possible that *Limnonectes kadarsani* and *L. dammermani* diverged *in situ* on Lombok after which *L. kadarsani* dispersed east all the way to Lembata. But rather although a tree topology consistent with a stepping-stone pattern of island colonization is suggested by the mtDNA data, the phylogenomic results suggest a leap-frog pattern where Lembata is derived from West Flores, and these two lineages are closer related to Sumbawa than they are to Eastern Flores. The parapatrically distributed lineages on Flores are experiencing asymmetrical gene flow with successful migrants moving from west to east. In summary, the oldest islands of the western Inner Arc tend to harbor the most divergent lineages for all three focal taxa, a pattern expected from lineages originating from the Sunda Shelf. In *Draco* and *Sphenomorphus*, the islands of the eastern Inner Banda Arc are colonized by way of the 'Sumba Route' where they disperse into the Outer Banda Arc island of Sumba and then move east to Timor, and finally north into the eastern Inner Arc. All three focal taxa show multiple non-sister lineages on some of the larger islands, suggesting either that multiple colonization events of a single island occurred, or possibly that formerly separated paleo-islands have since merged allowing for secondary contact of lineages that diverged in allopatry. These studies have shown that the biogeography of reptiles and amphibians within the Lesser Sundas is extremely complex. By examining biogeographical patterns across many co-distributed taxa these studies have the potential to provide insights into the geological history of the archipelago. From an evolutionary perspective, these studies highlight the presence of multiple independently evolving lineages within a currently described species occurring on the same island, which suggests that species diversity within reptiles and amphibians of the Lesser Sundas is underestimated.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

As an undergraduate and as a Masters student I always believed that the Museum of Vertebrate Zoology (MVZ) was the most prestigious institution for vertebrate biological research. I can still remember how happy I was when I learned that I had been accepted into the UC Berkeley graduate program and would be in the MVZ with Dr. Jim McGuire as my advisor. Shortly after arriving at Berkeley Jim and I became great friends and he helped me develop an incredible research project idea focusing on the Lesser Sunda Islands of Indonesia. I had been dreaming of travelling to Indonesia ever since I was in grade school, and now I had an opportunity to truly explore the region. Jim's level of support for me and my project made everything possible. He helped start off the project by paying for and leading an initial expedition to Sulawesi, Lombok, and Sumbawa that got my project off the ground. Jim then helped me write grants to get the project funded, which included multiple small grants, a National Geographic grant, and a full NSF grant for three years of salary, field work, and lab work. Jim and I ended up taking four separate two-month trips to Indonesia and we travelled to all of the major islands in the Lesser Sundas. I am forever indebted to Jim for his help with planning, packing, getting permits, travel logistics, communicating with locals, collaborating with Indonesian biologists, catching and preparing specimens, and so much more. I'm so thankful for everything he has taught me about how to be an evolutionary biologist, a field biologist, a museum biologist, a herpetologist, and about being an honest and hard working person. I am so thankful for all of his help and support and I look forward to a career of collaborations with him.

I am thankful to my dissertation committee members for their help in developing my project ideas. Rosemary Gillespie was always very insightful and helped me think about island biogeography from different perspectives. Rauri Bowie has always been so supportive of me and my research and he has helped encourage me along the way. Craig Moritz and David Wake were part of my qualifying exam committee and helped me develop as a biologist with our meetings. Craig Moritz would find time to meet with me whenever I needed, and he helped expand my views on evolutionary biology and conservation. David Wake helped change my understanding of what species are or aren't, and changed how I think about species formation.

Field work was certainly a large portion of this project and it could not have succeeded without a lot of hard work and help from a number of people. I am forever thankful to Alexander Stubbs for his help on the first three expeditions. Alexander is one of the smartest, friendliest, and hardest working people I have ever met and I learned so much from him on these adventures. Djoko Iskandar is one of our Indonesian collaborators and took part in three of the four expeditions. Djoko's knowledge of Indonesian reptiles and amphibians was indispensable, and this project certainly would not have been successful without his involvement. A number of other Indonesian students and herpetologists also helped an immense amount and include Umilaila Arafin, Gilang Ramadan, Evy Arida, and Amir Hamidy. Other participants on the expeditions include Ben Karin, Kristopher Harmon, Sarah Hykin, Luke Bloch, Jerome Fuchs, and Jim Patton. These people helped make my field experience an extremely enjoyable and memorable experience.

The community of people at the Museum of Vertebrate Zoology is what makes the MVZ so special. It's the open-door atmosphere and the diversity of biologists at the MVZ that create an

atmosphere that allows great research to flourish. I first want to thank the MVZ directors, Craig Moritz and Michael Nachman, for their leadership and support of me, my research, and my development as a scientist. I also want to thank Carol Spencer for her help dealing with accessioning specimens, obtaining tissue loans, and for being a wonderful supervisor during my time as a curatorial assistant. I want to thank Michelle Koo for her help with all things GIS, and for her collaboration on the red-bellied newt project. I also want to acknowledge all of the other MVZ faculty and staff that have so positively influenced my experience at the MVZ including Eileen Lacey, Ted Papenfuss, Chris Conroy, Carla Cicero, Anna Ippolito, Lelena Avila, Jim Patton, Monica Albe, Heather Brie Constable, Kat Corriveau, Terri Barclay, and many others. I am so thankful to all of the McGuire lab mates that either overlapped with me or who reached out to me such as Daniel Portik, Rayna Bell, Ammon Corl, Guinovere Wogan, Adam Leache, Matthew Fujita, Tom Devitt, Alison Davis, Sarah Hykin, Phillip Skipwith, Jeffrey Frederick, Jonathan Fong, Shobi Lawalata, Sarah Werning, Brian Lavin, Xiaoting Huang, Bryan Bach, Cynthia Wang, Alexander Stubbs, Benjamin Karin, Jackie Childers, Kristin Charles, Juli Goldenberg, and Sima Bouzid. I also want to thank all of the other graduate students, undergraduates, and post-docs in the MVZ that I have become friends with and who have helped me along the way: Charlotte Jennings, Sean Rovito, Sonal Singhal, Elaine Vo, Roberta Damasceno, Mike Holmes, Jesyka Melendez-Rosa, Yu Zeng, Ricardo Perera, Noelle Bittner, Jeremy Crawford, Darko Cotoras, Josh Penalba, Dan Rabosky, Kevin Rowe, Andrew Rush, Dan Wait, Rachel Walsh, Tali Hammond, Mallory Ballinger, Zachary Hanna, Jay McEntee, Dana Lin, Katya Mack, Taichi Suzuki, Luanne Wilson, Elizabeth Wommack, and many others. I would like to especially thank Lydia Smith for helping keep the evolutionary genetics lab so well organized and for helping me with lab work over the years. I would also like to thank Ke Bi for his help with the genomics lab work and bioinformatics pipeline that made the genomics part of this project possible. I want to thank David Wake for all of his support and help with my salamander research going back to my Masters work. David became my good friend, my collaborator, and my unofficial co-advisor. This is my MVZ family.

Over the course of my dissertation I spent a huge amount of time dealing with specimens, tissues, and lab work in the MVZ. I trained a number of undergraduate volunteers to help me with these tasks, and they completed an immense amount of work. I am forever indebted to the volunteers who helped me: Selena Ann, Stacy Shen, Vishruth Venkataraman, Stephanie Wong, Saachi Gupta, Amanda Radel, Jackie Childers, Kristopher Harmon, Chantelle Khambolja, Cynthia Wang, and Jennifer Lara.

I was first inspired to become a biologist by my parents. My mother, Leslie, and my father, Steve, have been so supportive of my choice to become a herpetologist, an idea I first had when I was seven years old. They have always been my foundation and have been there for me during the hard times and the great times. Thank you to my brother, Blake, and my sister, Rebecca, for your unconditional support and love. I especially want to thank my wife Michaela for all of her love and support. Whenever I got overwhelmed or stressed by academic life I could always depend on Michaela to be there for me and remind me what is important in life.

There are surely a number of people who are not mentioned here but deserve to be. Thank you to everybody who became my friend and helped me with this project.

CHAPTER 1

The Biogeography of the Lesser Sunda Islands of Indonesia: A Summary of the Archipelago's Geological History and Vertebrate Biogeography, with a Proposed Hypothesis-Testing Framework for Future Biogeographical Analysis

INTRODUCTION

Indonesia is one of the most biologically diverse countries in the world for both terrestrial and marine fauna. For example, the Indo-Australian archipelago makes up 1.3% of earth's surface yet it harbors 12% of the world's mammals, 16% of the world's reptiles, and 17% of the world's birds (Stone, 1994). Indonesia is an island nation containing over 17,000 islands, about 6,000 of which are inhabited by humans. The biodiversity of Indonesia stems from a number of factors including, 1) its large size spanning over 5,000 km from northern Sumatra to southeastern Papua, 2) its stable, tropical climate due to the position along the equator, 3) it contains thousands of islands, and 4) the country spans two (and perhaps 3) biogeographic realms (Goltenboth *et al*., 2006). This arrangement of islands has also generated cultural diversity within humans with more than 250 languages spoken by over 300 ethnic groups (Goltenboth *et al*., 2006).

The Lesser Sunda Archipelago, known in Indonesia as Nusa Tenggara, lies in the southeastern portion of Indonesia and occurs as a double arc of islands extending between Java and Bali in the west, and New Guinea in the east (Fig. 1). This island chain is one of the most geologically active and tectonically complex regions in the world (Rigg & Hall, 2011; Spakman & Hall, 2010). Given their current orientation, the Lesser Sundas may act as 'stepping stones' for animals and plants dispersing between the Greater Sunda Shelf (the Malay Peninsula, Borneo, Sumatra, Java, and Bali) and the Sahul Shelf (New Guinea, Australia, and their land-bridge islands). Thus, this island chain may act as a two-way filter for organisms dispersing between two of the world's great biogeographical realms, Asia and Australo-Papua. Alfred Russel Wallace's recognition of this striking pattern of a clinal mixture of species from different biogeographical realms was a key insight leading to his identification of the Wallace Line and, indeed, to his creation of the field of biogeography (Wallace, 1860, 1869, 1876). Given the critical role that the Lesser Sundas played in the development of the field, it is remarkable how little subsequent attention this region has received from historical biogeographers (Hall, 2009). Indeed, current understanding of Lesser Sunda biogeography has only modestly improved relative to what was known at the time of Wallace (Whitmore ,1987). The herpetofauna of the Lesser Sundas represent a particularly interesting group of vertebrates from a biogeographical standpoint because they appear to show distributional patterns that are most consistent with a stepping-stone model of island colonization caused by the two-way filter zone (Carlquist, 1965).

Below I summarize aspects of the geology, biogeography, and herpetofauna of the Lesser Sunda Islands. Then, I develop a number of biogeographical hypotheses regarding the timing and sequence of island colonization for rafting taxa that include reptiles and amphibians.

TECTONIC HISTORY OF THE LESSER SUNDA ISLANDS

Ernst Mayr (1944) noted that the zoogeography of the Malay Archipelago could not be

fully understood without knowledge of the region's geological history, an observation shaped by Wegener's (1915) theory of plate tectonics. There is a long history of geological study of eastern Indonesia dating back to colonial times. However, the tectonic history of the region is remarkably complicated, and theories on the geological evolution of the islands have changed extensively during the past few decades (Monk *et al*., 1996; Spakman & Hall, 2010; Hall, 2011). Below, I summarize the current understanding of the tectonics of the Lesser Sundas, with the goal of using this information to establish testable biogeographical hypotheses.

Geological Evolution of the Banda Arcs: Eastern Indonesia is the point of collision between four of the earth's major tectonic plates (Eurasian, Pacific, Philippines, and Australian), and is widely recognized as one of the most geologically complex and active regions in the world (Fig. 2). Bordered to the west by the Greater Sunda Shelf and to the east by the Sahul Shelf, the Lesser Sunda Islands can be considered oceanic islands in the sense that they have never been connected by land to continental Asia or continental Australo-Papua. The Sunda Shelf contains the islands of Bali, Java, Borneo, and Sumatra, and these islands have become connected to each other and to southeast Asia during the ~50 glacial maxima that have occurred over the last ~2 million years (Woodruff, 2010). The Lesser Sunda Islands occur as two geologically distinct island chains (see Fig. 3) termed the Inner and Outer Banda Arcs (Audley-Charles, 1987). The Banda Arcs include not only the Lesser Sunda Islands, but also the islands of southern Maluku to the north and east, which reflects the 180-degree curvature of this archipelago (Spakman & Hall, 2010). The biogeography of the Lesser Sundas and Maluku are largely independent of one another, as the constituent islands have never been in contact. The Lesser Sunda portion of the Inner Arc includes the islands of Lombok, Sumbawa, Komodo, Flores, Lembata, Pantar, Alor, Atauro, and Wetar. This Inner Arc is volcanic in origin, with many active volcanoes in the western islands including Lombok, Sumbawa, Flores, and Lembata. However, islands in the eastern part of the Inner Arc, such as Alor, Atauro, and Wetar have had no active volcanoes for the last 3 million years (Abbot & Chamaluan, 1981; Scotney *et al*., 2005), and some volcanoes in eastern Flores, such as Kelimutu, have been inactive for long periods of time and are considered extinct (Brouwer, 1939). In contrast, the islands of the Outer Banda Arc (including Sumba, Sabu, Rote, and Timor) are not volcanic, and all but Sumba are associated with uplift of the oceanic crust at the Australian continental plate boundary (Veevers, 1991; Spakman & Hall, 2010). Sumba is remarkable, in that it is thought to represent a piece of the Sunda (Asian) continental shelf that broke away nearly 50 Ma (Hall, 2011), although it was submerged beginning about 16 Ma and did not emerge to become sub-aerial once more until about 2 Ma (Fortuin *et al*., 1994; Audley-Charles, 2011).

The origin, nature of the crust, and age of the Inner and Outer Banda Arcs have long been controversial (Hall *et al*., 2011). However, a model for the tectonic origin of the region has recently emerged. Spakman and Hall (2010) posit that the Banda Arcs were formed by a complex 'Banda Arc Subduction Rollback' process that occurred at the triple-junction between the northward moving Australia Plate, the largely stationary Eurasia Plate, and the Pacific Plate. The process was initiated at about 23 Ma, when the Sula Spur, which sits immediately north of the Banda Embayment and resided at the northern edge of the rapidly northward-moving Australia Plate, collided with the SE Asian margin. For context, the Sula Spur includes present-day Seram, Ambon, Buru, Sula-Sanana, Peleng, and eastern components of Sulawesi. According to Spakman and Hall's (2010) model, this collision "broke and detached" the Australian continental slab that was subducting north of the Sula Spur while a transform fault east of Sumba

was changed into a west-dipping subduction fault. The result of this collision was that Sundaland began to rotate counter-clockwise, and the subduction fault east of Sumba began to migrate eastward (the initial stages of the 'subduction rollback' process). Beginning 15 Ma, the Banda Basin began opening as the subduction fault aligned with the Java Trench. As the Banda Basin opened, the islands of the Inner Banda Arc formed (beginning about 11 Ma) by volcanic action along the northern margin of the eastward expanding subduction zone. Beginning about 4 Ma, the eastern portion of the Banda forearc (the region between the volcanic arc and the oceanic trench) began to collide with the Australian plate margin, which terminated volcanic activity on the eastern Inner Banda Arc islands of Alor, Atauro, Wetar, and Romang (whereas volcanic activity has continued in the western sector of the Inner Banda Arc). In addition to terminating volcanic activity in the eastern Inner Arc, this collision initiated the uplift of the islands of the Outer Banda Arc by folding and thrusting along the margin of the Australian Plate. This uplift occurred at an exceptionally rapid pace – as much as 3 km uplift/million years in places (Roosmawati & Harris, 2009; Audley-Charles, 1986, 2011; Hall, 2011; Hall *et al*., 2011), allowing Sumba to rise from a depth of 5 km to an elevation of 1 km above sea-level beginning 4 Ma (Pirazzoli *et al*., 1993; Fortuin *et al*., 1997), and Timor to uplift by as much as 4 km over just the past 2 million years (Audley-Charles, 1986).

Even within the Inner Banda Arc there are major differences between the western and eastern Islands. The western islands of Lombok and Sumbawa lie above the eastern extent of the Java Trench where active subduction of the Indo-Australian plate is still occurring. There is an offset in the line of volcanoes between Sumbawa and Flores indicative of a major tectonic discontinuity that has been termed the "Sumba Fracture" (Audley-Charles, 1975). This discontinuity marks major differences between east and west Indonesia in both submarine morphology, the strike of the volcanic zone, and geological origin of islands of the outer Banda Arc.

Timing of Island Formation: For biogeographical purposes, it is essential to understand the patterns of connectivity and timing of subaerial emergence of the Lesser Sunda Islands. As described above, it is clear that the Inner Banda Arc islands predate the Outer Banda Arc islands, and were created by subduction driven volcanism beginning about 11 Ma. That said, there is little explicit information available regarding the timing of ultimate sub-aerial emergence of the Inner Banda Arc (Robert Hall, pers. comm.). Hall (2009) postulated the presence of emergent volcanoes where the western Inner Banda Arc islands of Lombok, Sumbawa, and Flores reside beginning about 10 Ma, with volcanoes and land appearing in the remainder of the Inner Banda Arc (Alor, Wetar) beginning about 5 Ma. The timing of emergence of the Outer Banda Arc islands are both easier to estimate (Hall, 2011), and younger, than the Inner Banda Arc Islands. Sumba and Timor are estimated to have become land-positive about 2 Ma (Hall, 2009; Audley-Charles, 2011), with Sabu being somewhat younger having initiated uplift from a depth of more than 2 km beginning about 1-2 Ma (Roosmawati & Harris, 2009; Rigg & Hall, 2011). The islands east of Timor are not well dated, but numerous geological studies point to a common thrust-folding mechanism along the Australia plate margin for Timor and Tanimbar, suggesting similar island ages (Audley-Charles, 1986; Charlton *et al*., 1991; Ishikawa *et al*., 2007; Kaneko *et al*., 2007). We use these rough relative ages of the Lesser Sunda Islands to inform our biogeographical hypotheses. To summarize, the oldest islands appear to be Lombok, Sumbawa, and Flores, which originated by about 10 Ma. The remaining islands of the Inner Banda Arc, including many of the islands stretching from Adonara to Damar, were apparently in place by 5

Ma. Finally, the Outer Banda Arc islands of Sumba, Sabu, Timor, and Tanimbar did not appear until about 2 Ma.

Effects of Sea Level Fluctuations: Given current channel depths, we can identify sets of major islands that likely merged into single sub-aerial islands during Pleistocene glacial maxima when sea-levels were up to 120 m lower than at present (Fairbanks, 1989; Bintanja *et al*., 2005). In particular, it is likely that (1) Lombok and Sumbawa, (2) Komodo, Rinca, Flores, Adonora, Solor, and Lembata, (3) Timor and Rote, and (4) Alor and Pantar merged during the ~50 glacial low-stand that have occurred over the past two million years, the most recent of which occured ~17,000 years ago, whereas the remaining major islands in the archipelago likely existed as distinct units throughout the Pleistocene, including Sumba, Sabu, and Wetar (Voris, 2000). The periodic joining of islands during glacial maxima does not necessarily imply that those islands should be considered as one unit when making biogeographical hypotheses. For example, the Greater Sunda Islands of Java, Sumatra, and Borneo become joined periodically yet they have many differences in their extant faunal compositions (Leonard *et al.*, 2015). Another factor to consider within the Lesser Sundas is the possibility of ancient islands that have sunk or eroded away, as Mayr (1944) noted that there is some evidence for the existence of additional islands in the Timor region, but the position, size, and chronology is unknown. I could find no additional studies confirming this, but one cannot rule this possibility out. This possibility also extends to land bridges that once connected islands. One interesting story comes from the islands of Lembata and Adonara, a story that I heard the locals tell when I was on Lembata, and that Brouwer (1939) also learned of. The story tells of a past land connection between northwest Lembata and northeast Adonara that sunk into the sea during an earthquake, possibly sometime in the seventeenth century.

Merging of Paleo-Islands: Although not specifically addressed in the geological literature, the possibility of volcanic islands in the Inner Banda Arc existing as separate paleo-islands in the past that have since merged must also be considered. The most likely islands to be relevant for this process include Sumbawa and Flores, which occur as long islands composed of multiple volcanoes. For example, one could imagine that the widely spaced volcanoes of Flores would have initially been separated by marine barriers, merging long after first becoming subaerial. Furthermore, Flores does contain a large north-south fault through its center that may represent the meeting of two paleo-islands (Katili, 1975).

VOLCANIC ACTIVITY IN THE LESSER SUNDAS

The strong tectonic activity of the Sunda Volcanic Arc extending into the Lesser Sundas region has resulted in a steady occurrence of natural disasters such as explosive volcanic eruptions, tsunamis, earthquakes, and landslides. Because many of the islands within the Lesser Sundas are small, the effects of these events can have large effects on the distribution and abundance of each islands fauna, and in some cases can lead to localized extinction (Monk *et al.*, 1997). There are approximately 24 active volcanoes on the Inner Banda Arc islands of Lombok (1), Sumbawa (2), Sangeang (1), Flores (15), Adonara (1), Lembata (3), and Pantar (1). The eastern Inner Arc Islands of Alor and Wetar have not had any volcanic activity for the past few million years.

Explosive volcanic eruptions can have major effects on the biodiversity and community

composition of small islands. One of the best-studied examples is of the Krakatau islands and the effects of the massive eruption in 1883, which sterilized the entire island group. However, this gave biologists a natural experiment to monitor the recolonization of the islands by plants, invertebrates, and vertebrates (Thornton & New, 1988; Thornton *et al.*, 1988; Whittaker *et al.*, 1989; Bush & Whittaker, 1991; Rawlinson & van Balen, 1992). While the Krakatau islands are much smaller than the major Lesser Sunda Islands they do show that large eruptions can cause local extinctions, and that recolonization can occur over short time scales.

Some of the largest volcanoes of the region that have erupted in the last 1,000 years are Gunung Agung on Bali, Gunung Rinjani on Lombok, and Gunung Tambora on Sumbawa. These volcanoes are the large due to their location in the beniof zone immediately north of the Java trench where active subduction of the Indo-Australian oceanic crust beneath the islands is occurring. Gunung Agung has had frequent major eruptions with approximately one every 100 yrs for the past 5,000 years, and the 1963 eruption had a major effect on global climate (Fontijn *et al.*, 2015). Gunung Rinjani is the second highest volcano in Indonesia at 3,726 meters and has been subject to some of the largest explosive eruptions during the Holocene including the massive eruption of 1257 (Emile-Geay *et al.*, 2008; Lavigne *et al.*, 2013). This eruption of Rinjani was responsible for the largest volcanic sulfur load released in the last 7,000 years with eight times more sulfur released than Krakatau and two times more than Tambora (Crowley, 2000). The Rinjani eruption deposited voluminous ashfall over most of the island, caused multiple pyroclastic flows that devastated the land and villages around the volcano, and left significant portions of Bali, Lombok, and Sumbawa sterile and uninhabitable for generations (Marrison, 1999). Almost 600 years later in 1815 Tambora volcano on Sumbawa erupted in what is known as the largest volcanic eruption in recorded human history spewing out nearly eight times as much fine ash into the atmosphere as Krakatau (Rampino & Self, 1982). Before the eruption Tambora towered to a height of nearly 4,200 meters, while after the eruption its elevation dropped to 2,850 meters due to the massive explosion of the blast (de Jong Boers, 1995). The eruption changed global climate for years afterward and killed over 130,000 people on the islands of Bali, Lombok and Sumbawa, though this is likely a vast underestimate of the total number of casualties from the region (de Jong Boers, 1995). After the eruption eastern Sumbawa was described as a lifeless moonscape, but after only 100 years the mid-elevation slopes had been re-forested with lush jungle (de Jong Boers, 1995). However, even to this day the faunal community has not recovered to the levels of other parts of Sumbawa (Trainor, 2002).

In summary, the large explosive volcanic eruptions of the Inner Banda Arc and Bali have impacted the faunal communities and vegetation of these islands. While there has certainly been localized extinction of nearly all animals in the immediate vicinity of these volcanoes, these eruptions and earlier ones may have also contributed to the extinction of species from entire islands or island groups. If complete insular extinction events have taken place, they should leave a recognizable phylogeographic signature.

WALLACE'S LINE AND THE BIOGEOGRAPHY OF WALLACEA

Alfred Russel Wallace was a British naturalist who spent nearly eight years, between 1854 until 1862, exploring the Malay Archipelago (including Malaysia, Singapore, and Indonesia). He collected more than 126,000 specimens, of which several thousand represented new species. Wallace divided the world into biogeographic realms, two of which were Australo-Papua and Southeast Asia. He noted that: "*South America and Africa, separated by the Atlantic,*

do not differ so widely as Asia and Australia", emphasizing the stark contrast in faunas between these adjacent realms (Wallace, 1860). However, the exact placement of the boundary between the Oriental and Australo-Papuan biogeographic realms was not known at the time. While Wallace had been exploring islands of the Lesser Sundas (Fig. 3), including Lombok, he arrived at Bali and noted:*"The islands of Bali and Lombok, situated at the East end of Java, are particularly interesting…they form the extreme points of the two great zoological divisions of the Eastern hemisphere; for although so similar in external appearance and in all physical features, they differ greatly in their natural productions"* (Wallace 1869). And later he emphasized the faunal dissimilarity of Bali and Lombok, which is captured in the quote: *"These islands differ far more from each other in their birds and quadrupeds than do England and Japan"* (Wallace, 1880).

Indeed, this boundary between Bali and Lombok, which also runs north between Borneo and Sulawesi (Fig. 4), became famous for demarcating the most abrupt faunal transition in the world and was termed "Wallace's Line" (Huxley, 1868; Lohman *et al.*, 2011). Wallace thought that these transition zone islands (Phillipines, Sulawesi, Maluku, and Lesser Sundas) had once been part of a Pacific-Australian continent. Wallace also recognized that there was a "*greatly increasing proportion of Australian forms and decreased proportion of Indian forms as we go from west to east*" (Wallace, 1869), suggesting that the Lesser Sunda Islands may have in fact been a transitional zone rather than belonging to the Australo-Papuan realm. However, even Wallace was much less certain about this line in his later writings due to the complexity of biogeographical patterns in the region. Weber (1902), Pelseneer (1904), Mertens (1930), and Brongersma (1936) insisted that Wallace's line was not the boundary between the Asian and Australo-Papuan realms. Van Kampen (1909) stated: "*Such a sharp boundary as Wallace drew it does not exist. Not only is there none where he drew it, but no such line exists anywhere in the archipelago*." Others such as Dickerson *et al*. (1928), Raven (1935), and Rensch (1936) defended the placement and significance of Wallace's Line. Indeed, there has been much debate on the significance of Wallace's Line, but probably the most striking feature of Wallace's line is that it separates a rich fauna from a relatively impoverished one, and a continental fauna from an insular fauna (Mayr, 1944).

After Wallace's initial exploration of the region many other naturalists and zoogeographers started to explore the region. The other great biogeographic line of the region is known as "Lydekker's Line", which was first recognized by Lydekker (1896) as a boundary between the Australian region and the Austro-Malayan region (Fig. 4). The true significance of Wallace's and Lydekker's lines are that Wallace's line essentially delimits the eastern boundary of the Sunda Shelf (which effectively follows the 120 m depth contour line) and Lydekker's Line follows the northwestern margin of the Sahul Shelf (which also effectively follows the 120 m depth contour line). The intervening islands were termed "Wallacea" by Dickerson (1928). According to Simpson (1977), Wallace's and Lydekker's Lines are the two biogeographical boundaries that are beyond dispute, but the problem is that there are thousands of islands between these two lines, many of which are separated by marine barriers of equivalent depth and width as those that separate the Sunda and Sahul shelves from their most proximate deep-water islands. While many species of Asian origin reach their eastern limit in Wallacea, and many Australo-Papuan species reach their western limit in Wallacea, a unique fauna has also evolved within the isolated oceanic islands of Wallacea (Whitten *et al*., 1987).

There are a number of other biogeographic lines drawn through this region including Muller's Line (Muller, 1846), Murray's Line (Murray, 1866), Huxley's Line (Huxley, 1868), and

Weber's Line (Mayr, 1944). These lines were drawn with respect to the taxa each researcher was focusing on at the time. Simpson (1977) reviewed the many biogeographic lines of Wallacea and concluded: "*It is fairly evident that there is no single faunal balance line for faunas as a whole, but separate balance lines for an undetermined number of groups*". Stresemann (1939) pointed out that Wallacea has four distinct regional faunas: Philippines, Sulawesi, Maluku, and Lesser Sundas. These four distinct regions have little in common except that they are in this transitional zone. The faunas of these transitional regions are not simply clinal mixtures of Oriental and Australian animals, because they include many endemics (regardless of their origins). However, the area of Wallacea is not united by that endemicity because most endemics are not widespread or uniformly distributed throughout or within the region. After examining the biogeographical lines and patterns of Wallacea, Simpson (1977) suggested: "*...let us not assign the intervening islands to any region, subregion, transitional or intermediate zone, or the like.*" But perhaps the most reasonable biogeographical line through Wallacea is Weber's Line, which was modified by Mayr (1944) to separate islands based on their proportion of Oriental or Australo-Papuan origin species. Weber's line runs south of Timor, then north between Babar and Tanimbar, and between Buru and the Sula Islands. Weber's Line is closer to Lydekker's Line than it is to Wallace's Line, thus illustrating the dominance of Oriental-origin fauna within western and central Wallacea. This preponderance of Oriental taxa in western and central Wallacea is likely the outcome of two geographical factors: First, the massive island of Sulawesi is a target of dispersal from the adjacent Sunda Shelf, and provides a natural dispersal corridor into Wallacea. Second, the western-most Lesser Sunda Islands are only narrowly separated from the Sunda Shelf, whereas the increasingly isolated eastern-most islands in this archipelago are much more distant from the Sahul Shelf.

The Lesser Sunda Islands are unique among the islands and archipelagos of Wallacea for many reasons. First, they essentially form an island peninsula extending east from the Sunda Shelf island of Bali, and the straits between islands are usually no greater than 20-30 km. This arrangement of the islands, in particular the Inner Banda Arc islands, resemble stepping-stones due to their highly linear arrangement and close proximity to one another. When a "stepping stone" island occurs between a source and the target it can increase the probability of successful colonization by many orders of magnitude (MacArthur & Wilson, 1967), and consequently the Lesser Sunda Islands are composed primarily a fauna of Asian origin. For example, Rensch (1936) analyzed bird and herpetofaunal data and showed that the Oriental fauna are dominant on each island as far eastward as Timor. Conversely, colonization of the Lesser Sundas from the east requires dispersal over much larger bodies of water. The initial distance between the Sahul shelf and the nearest islands is great (>200 km) and on average the distances between islands in the eastern portion of the Banda Arcs is also great (~50-100 km). Many of the Banda Arc islands east of Timor are small, creating small targets which lower the chances that dispersing or rafting taxa will land on their shores (MacArthur & Wilson, 1967). Furthermore, while colonization from the west or east are the dominant mechanisms that zoogeographers have envisioned taxa entering the archipelago, many taxa within the Lesser Sundas have patterns that are inconsistent with this idea. Auffenberg (1980) noted: "*The present distribution of at least the lower vertebrates, many invertebrates, and plants in the Lesser Sundas suggest a more complex distributional pattern than a simple two-directional east-west movement.*" Indeed, there is no simple biogeographic pattern that could be applied to the fauna of the Lesser Sundas, and the complex geological history, wind currents, ocean currents, and stochastic nature of rafting events had left a mosaic of species' geographic ranges and insular faunal communities.

Environmentally, the Lesser Sundas are unique within Wallacea due to the dramatic east-west climatic cline, with conditions becoming generally more xeric in the east and more tropical in the west. The islands are highly seasonal when compared to other regions of Indonesia, and have probably been seasonal for their entire existence (Heaney, 1991). During glacial maxima there would have been even less rain during monsoons, and seasonal forest and savannah would have expanded during these periods (Heaney, 1991). This environmental cline starts in Java however, and may help to explain the low faunal diversity of the Lesser Sundas. As stated above, most of the fauna of the Lesser Sundas dispersed from Java or Bali, but even the faunal diversity of Java is low when compared to other Greater Sunda Islands such as Sumatra or Borneo. Mayr (1944) suspected that Java had lower diversity likely due to high volcanic disturbances during the Pleistocene, which covered the island with lava and ash and likely caused localized extinctions. Java is also less humid and more peripheral, thus less accessible to colonists from the Asian mainland. The increase in aridity is more pronounced in the eastern portion of Java and in Bali where true tropical rainforest is replaced by monsoonal forest. The most characteristic Javan species are restricted to western Java, and even Bali has only a fraction of the faunal diversity that characterizes Java. As this trend of increasing aridity continues eastward into the Lesser Sundas, the combination of marine barriers, decreasing tropical habitat, and decreasing humidity prevents most of the tropical fauna of the Greater Sunda Islands from becoming established. In fact, each water barrier (especially Straits deeper than 120 m) between neighboring islands in the Banda Arcs contain faunal breaks. Mertens (1930) thought that other straits were even more efficient barriers such as the Bali Strait (Between Bali and Java) and the Sape Strait (Between Sumbawa and Flores). However, the effectiveness of each strait as a barrier and their overall contributions to the biogeographical patterns of the archipelago will not be fully understood until all of the distinct species of the region and their geographical ranges are described.

HERPETOFAUNA OF THE LESSER SUNDAS

Indonesia is suspected to harbor one of the greatest numbers of extant reptile and amphibian species in the world. In general the herpetofauna is poorly understood not only taxonomically, but also with respect to the basic biology, ecology, and geographic ranges of species. The Lesser Sundas are oceanic islands and this isolation has contributed to them being species-poor when compared to the Sunda and Sahul shelf islands of Sumatra, Borneo, Java, and New Guinea. One extensive survey of reptiles and amphibians was conducted by Mertens (1930), who surveyed and collected specimens from the western Inner Arc islands including Lombok, Sumbawa, Komodo, and western Flores. Mertens also visited Timor and concluded that Timor was one of the most disappointing places to visit herpetologically. Recent surveys of Timor-Leste have provided some insight on the diversity and distribution of herpetofauna of the region (Kaiser *et al.*, 2011; O'Shea *et al.*, 2012; Sanchez *et al.*, 2012; Kaiser *et al.*, 2013), but more work on other islands of the Lesser Sundas is needed. In general the herpetofauna of the Lesser Sundas is more Asiatic than Australian (Dunn, 1927a, 1927b; Dunn, 1928). Darlington (1957) emphasized that there was a gradual transition of Australian to Oriental reptiles and amphibians, a pattern that is seen in reptiles (Fig. 5, adapted from Carlquist, 1965; Whittaker & Fernández-Palacios, 2007). But many zoogeographers also noticed that within reptiles and amphibians there was not one biogeographic line between Asian and Australo-Papuan islands, and that it was "*advantageous to consider the validity of several superimposed biogeographic*

divisions" (Auffenberg, 1980). Below I summarize some of the studies on various reptile and amphibian groups.

Snakes: There are 29 species of terrestrial and semi-aquatic snakes that occur in the Lesser Sunda Islands (DeLang, 2011). The largest islands, however, contain less than 20 species with Lombok containing 18, Sumbawa 16, Flores 18, Sumba 12, and Timor 14 species. Insular endemism for snakes only occurs at the subspecies level, suggesting that snakes are either not as isolated as other taxa, have colonized the archipelago more recently, or are comprised of relatively cryptic species requiring further investigation. A study based on multiple snake species of the Lesser Sundas found that the snake fauna assemblage most closely resembles that of Bali, and that the major biogeographic break within snakes occurs at Weber's Line, between Babar and Tanimbar (How & Kitchener, 1997). Some species, such as *Trimeresurus insularis*, *Psammodynastes pulverulentus*, *Lycodon subcinctus*, and *Elaphe subradiata,* have likely remained isolated throughout the Pleistocene and have diverged morphologically from Sunda Shelf populations (How *et al.*, 1996). Species such as *Trimeresurus insularis* show a clear west to east pattern of colonization with the Lesser Sundas clade sister to Java (Malhotra & Thorpe, 2000), though the relationships within the Lesser Sundas are not well resolved (David *et al.*, 2003). Other species such as *Liasis macklottti* have dispersed from northern Australia into the Outer Banda Arc (Rawlings *et al.*, 2004). Perhaps the most perplexing biogeographical pattern within snakes belongs to *Daboia siamensis,* where populations occur on eastern Java, Sumbawa (pers. obs.), and Flores, but not on Bali or Lombok, and the closest related populations to these occur in China and Taiwan (Thorpe *et al.*, 2007).

Lizards: There are 39 species of lizards recorded from the Lesser Sunda Islands with only eight of these being endemic to the archipelago (Uetz & Hallerman, 2010). The most famous lizard of the Lesser Sundas, and the largest lizard in the world, is the Komodo Dragon (*Varanus komodoensis*), which occurs only on Komodo, Padar, Rinca, and Flores Islands in the center of the Inner Banda Arc and displays genetic differentiation among island populations (Ciofi *et al.*, 1999). Komodo Dragons dispersed to the Lesser Sundas from Australia, and used to occupy more islands in the past including Sumba and Timor where fossils have been found. Other lizard taxa known to have dispersed out of Australia include skinks of the genus *Glaphyromorphus* (Greer, 1990), and *Varanus timorensis, V. auffenbergi, and V. indicus* (Iskandar & Erdelen, 2006; Sweet & Pianka, 2007). Some skinks show little to no divergence among islands of the Lesser Sundas, including *Eutropis multifasciatus* and *Lamprolepis smaragdina* (Schmitt *et al.*, 2000). The wide-ranging Flores Forest Skink, *Sphenomorphus melanopogon,* occurs on a few small islands off of southwestern Java, but not on mainland Java or Bali, and also on nearly every major island in the Banda Arcs (Shea, 2012). Within geckos, *Cyrtodactylus laevigatus* and *C. darmandvillei* are endemic to the Lesser Sundas with the closest relative to *darmandvillei* being *C. jellesmae* in Sulawesi (Wood *et al.*, 2012). *Lepidodactylus lugubris* from the Lesser Sundas were once thought to be part of a wide-ranging species, but were recently shown to be endemic species not closely related to *L. lugubris* (Ota *et al.*, 2000). In flying lizards (Genus *Draco*), a west to east colonization pattern was inferred for the Lesser Sundas based on molecular phylogenetic analysis, with two endemic species recognized (McGuire & Heang, 2001).

Turtles and Crocodiles: Four species of non-marine turtles are native to the Lesser Sundas.

These include two species of Snake-neck turtles, *Chelodina mccordi* from Rote, and *Chelodina timorensis* from Timor, that represent Australo-Papuan elements. Two species of Asian origin occur in the archipelago, including the Southeast Asian Box Turtle (*Cuora amboinensis*) from Sumbawa and Timor, the Southeast Asian Softshell Turtle (*Amyda cartilaginea*). The only crocodile from the region is the Saltwater Crocodile (*Crocodylus porosus*) known from Lombok, Flores, Sumba, Rote, and Wetar (Monk *et al*., 1997; Trainor & Lesmana, 2000).

Amphibians: The only known amphibians from the Lesser Sundas are frogs. Salamanders do not occur anywhere in Indonesia, and caecilians terminate at Bali. Java and Bali contain 41 species of frogs and one species of caecilian, with one species of frog, *Oreophryne monticola*, occuring on Bali but not Java (Iskandar, 1998). Although it is assumed that undescribed species of frogs still occur in the Lesser Sundas, there are currently 18 species known from the region (Inger, 1999). These species represent the families Bufonidae (*Igneophrynus* and *Duttaphrynus*), Dicroglossidae (*Fejervarya*, *Limnonectes*, and *Occidozyga*), Hylidae (*Litoria*), Microhylidae (*Kaloula* and *Oreophryne*), Ranidae (*Papurana*), and Rhacophoridae (*Polypedates*) (AmphibiaWeb, 2015). Two of the species are endemic to a single island including *Occidozyga florianus* and *Oreophryne rookmaakeri*, both endemic to Flores. All species of frogs from the Lesser Sundas are on the IUCN Red List, with *Limnonectes dammermani* and *Oreophryne jeffersoniana* listed as "Near Threatened" and *Oreophryne monticola* listed as "Endangered" (IUCN, 2015). Very little research has been conducted on the frogs of the region, and thus almost nothing is known about their biogeographic history within the Lesser Sundas. One study found that the treefrog *Polypedates leucomystax* from Lombok Island was sister to populations in northern Philippines (Brown *et al.*, 2010).

BIOGEOGRAPHY OF OTHER FAUNA

Currently the extent and patterns of specific plant and animal groups are in a constant state of being updated, making any broad conclusions relatively speculative, though some generalizations can be made. Biogeographical units within the Lesser Sundas have been proposed based on general limits of endemic species and community composition of plants and animals (MacKinnon *et al.*, 1982). These biogeographical units have been used as a proxy to serve as conservation units and they generally correspond to island groups that become land-bridged to one another during glacial maxima. The bird fauna of the Lesser Sundas is the most well-studied group of vertebrates in the region and patterns of bird endemism are one of the major drivers of the boundaries of these biogeographical units. The Lesser Sundas contain two endemic genera of birds (*Buettikoferella:*Sylviidae and *Heleia:*Zosteropidae) both of which are restricted to Timor, and there are at least 12 passerines within the Lesser Sundas that are endemic (White & Bruce, 1986). The taxonomy of birds in the region, while much better understood than other groups, is still lagging behind other regions of the world. One study suggested that if current species limits were applied to the birds of the Lesser Sundas that dozens of new bird species might be recognized (White & Bruce, 1986).

Researchers have studied Lesser Sunda biogeography primarily by examining species distributions rather than by using phylogenetic or population genetic approaches (Halloway & Jardine, 1968; Michaux 1994, 2010), and their findings were generally consistent with the two-way filter model in the sense that the number of reptile, bird, and butterfly species of Asian origin do tend to decrease proportionately as you move eastward across the archipelago, while

Australo-Papuan taxa become more dominant (e.g., Whittaker & Fernández-Palacios, 2007). However, even these analyses are potentially confounded by our limited knowledge of species distributions, and especially species taxonomy for the region. For example, many taxa are treated as single widely distributed species spanning multiple islands in the archipelago despite the long periods of isolation experienced by these islands, and many taxa furthermore exhibit island-specific morphological differences. Furthermore, very few DNA sequence-based molecular phylogeography studies have been undertaken for Lesser Sunda taxa (but see Hisheh *et al.*, 1998 and Maharadatunkamsi *et al.*, 2003, both for the fruit bat, *Eonycteris spelaea)*. However, Kitchener and colleagues have published a series of allozyme and morphometrics studies of bats (Kitchener *et al*., 1993a, 1993b, 1997; Maharadatunkamsi *et al*., 2000; Schmitt *et al*., 1995, 2009), and small terrestrial mammals (Kitchener *et al*., 1994a, 1994b).

Some researchers have found the expected biogeographical pattern (given the prevelance of Asian origin fauna) of a west to east colonization pattern (such as for Fruit Bats, Hisheh *et al*., 1998), while other studies have found very different patterns. Hawkmoths have apparently colonized the Lesser Sundas from the east (by way of Tanimbar) and exhibit a pattern consistent with a west-to-east stepping-stone model to Lombok, where the Lombok Strait (Wallace's Line) becomes a biogeographic boundary (Beck *et al*., 2006). A similar pattern is seen in weevils, where all weevils on Bali are endemic and are derived from the Lesser Sundas, having diverged ~2-10 Ma (Tanzler *et al*., 2014), and weevil diversity and endemism in the archipelago has been vastly underestimated (Reidel *et al.*, 2014). Other taxa seem to have originated in the Lesser Sundas and spread outwards, as suggested for the Pacific Rat, which apparently originated on Flores (Thompson *et al.*, 2014), Weaver ants on Flores (Azuma, 2006), endemic ants on Timor (Andersen *et al*., 2013), and Wall-roosting Mouse-eared Bats (Wiantoro *et al.*, 2012). Colonization of the Lesser Sundas by way of Sulawesi is seen in Cicadas (Boer & Duffels, 1996), *Stegodon florensis* (van den Bergh *et al.*, 2001), and in other taxa as well (Vane-Wright, 1991). Multiple independent colonizations of the Inner Banda Arc are seen in flightless beetles from Sulawesi (Tanzler *et al.*, 2016). Although likely to be a rare case, colonization by way of the Philippines was proposed for *Ficedula* Flycatchers (Outlaw & Voelker, 2007).

The earliest evidence of hominins in the Lesser Sundas appears one million years ago (Brumm *et al.*, 2010). *Homo erectus* are estimated to have arrived in the archipelago during the late Pleistocene (Monk *et al.*, 1997), and their arrival corresponds to the disappearance of fauna such as the pygmy stegodont (*Stegodon sompoensis timorensis*) and giant tortoises (*Geochelone atlas*), as these and other large native fauna were likely hunted to extinction (Sondaar *et al.*, 1994). Fossils from a possible dwarf hominid species, *Homo floresiensis*, from the late Pleistocene (~74-95 thousand years ago) was recently discovered on the island of Flores (Brown *et al.*, 2004; Morwood *et al.*, 2005), though it is uncertain if *Homo floresiensis* represents a unique species (Tocheri *et al.*, 2007; Aiello, 2009) or pathological modern humans (Jacob *et al.*, 2006; Henneberg *et al.*, 2014). It is thought that both Austronesian and Papuan language groups began to infiltrate various regions of the archipelago starting around 6,000 years ago (O'Connor, 2007). Many of the larger mammal species that can be currently found in the region were introduced during this period (~6,000 years ago) such as palm civets (*Paradoxurus hermaphroditus*), long-tailed macaque monkeys (*Macaca fascicularis*), rusa deer (*Cervus timorensis*), squirrel (*Callosciurus* notatus), and possibly pangolin (*Manis javanica*) that were introduced from the west (Greater Sunda Islands), pigs (*Sus celebensis*) introduced from Sulawesi, and cuscus (*Phalanger orientalis*) which were brought from the east (Glover, 1971).

BIOGEOGRAPHICAL HYPOTHESES FOR COLONIZATION OF THE LESSER SUNDAS

Herein, we propose several alternative colonization models for the Lesser Sundas that are testable with phylogenetic/phylogeographic data. These models are based on the tectonic history of the archipelago, relative island ages, and the geographic proximity of each island to potential mainland source populations.

Null Hypothesis – Stepping-stone. The simplest explanation for the biogeography of the Lesser Sundas is a 'stepping-stone' model, whereby species have invaded the archipelago from the Sunda (west) or Sahul (east) Shelves and sequentially colonized the next nearest island across the chain (Fig. 6). If these islands indeed acted as 'stepping-stones' then we would expect pectinate phylogenies for focal taxa corresponding to the direction of dispersal. For example, taxa dispersing from Asia will theoretically reach more proximate islands such as Lombok first, before dispersing further along the chain. Thus, the lineage on Lombok should be most basal, followed by the lineage occurring on Sumbawa, and so on, with the most derived lineages occupying the most distant (eastward) islands such as Timor or Wetar. Correspondingly, taxa originating in Australo-Papua should exhibit the reverse pattern with eastern lineages representing the most basal divergences. We note that the stepping-stone model ignores island ages and depends on the current configuration of the Lesser Sundas Archipelago. Thus, it is most plausible in the context of recent arrivals to the archipelago.

Alternative Hypothesis 1 – Asian Invasion (Fig. 7a). The premise behind this model is that the westernmost islands of the Inner Banda Arc were the first to appear, and thus, the earliest invaders from the Sunda Shelf (west) could only have landed on Lombok, Sumbawa and Flores. Later, as more islands appeared, they could have been colonized sequentially beginning with the more eastern Inner Banda Arc islands (e.g., Alor, Wetar). Finally, the islands of the Outer Banda Arc (e.g., Sumba, Sabu, Timor, Tanimbar) appeared most recently and may have been invaded independently from the Inner Banda Arc. This model suggests a phylogenetic pattern involving an Asian outgroup, deepest divergences on the order of 5-10 million years, with the deepest ingroup splits occurring between Lombok, Sumbawa, and Flores, followed by divergences within the remaining Inner Arc islands, and with the most terminal divergences involving Outer Banda Arc populations.

Alternative Hypothesis 2 – Early Australo-Papuan Invasion (Fig. 7b). This model is most plausible for taxa that reached the Lesser Sundas from the Sahul Shelf (south and/or east) prior to uplift of the Outer Banda Arc Islands. With this model, the earliest arrivals would have had access to the same set of early-appearing islands as would Early Asian invaders (initially just Lombok, Sumbawa, and Flores; later the remaining Inner Arc Islands). At this time, the invasion of the Lesser Sundas from the Sahul Shelf would have required a much larger overwater dispersal event than would have been required of Asian invaders arriving from the Sunda Shelf, so there may not be many (or any) examples. Following successful colonization of the Inner Banda Arc, these lineages would have had opportunities to disperse to the Outer Banda Arc following uplift beginning 2 Ma. Phylogenetic patterns expected under this model would be the same as for the 'Asian Invasion' model, except that the outgroup would be of Australo-Papuan ancestry.

Alternative Hypothesis 3 – Timor Springboard (Fig. 7c). The 'Timor Springboard' model would be expected to apply to recent colonizers of the Lesser Sundas from the Sahul Shelf – taxa that invaded only after the Outer Banda Arc islands were already in place beginning 2 Ma. We propose that Timor would represent the most plausible entry point because it is geographically most proximate to the Sahul Shelf during glacial low-stands, and has a 575 km coastline paralleling the Australian coast. Once a Sahul Shelf-derived colonizer entered the Lesser Sundas by way of Timor, further dispersal into the archipelago could follow multiple paths including eastward toward Tanimbar, westward by way of Sabu toward Sumba, and northward into the Inner Banda Arc (from which further dispersal would be possible). Taxa fitting this model would be expected to be relatively young (less than 2 Ma), and to have Timor as a basal branching point. Note that this model differs from the East-West stepping stone model simply by having Timor serve as the initial entry point.

Alternative Hypothesis 4 – Selayar Bridge (Fig. 7d). Auffenberg (1980) proposed a biogeographical hypothesis that he termed the "Saleyer Bridge", whereby taxa from Sulawesi dispersed south into the middle of the Inner Banda Arc. This was not a true land bridge, but a time when islands to the north were close and allowed for southward colonization of the Lesser Sundas during the Pliocene, with a subsequent spread into the archipelago. Of course the reverse of this hypothesis is also possible, where taxa from the Lesser Sundas could have dispersed north into Sulawesi. Even at present day there are dozens of islands between the southwestern peninsula of Sulawesi and Flores that could act as stepping-stones including some major islands such as Selayar, Kayyadi, Tanahjampea, Kaloa, Bonerate, and Kalaotoa Islands. Of course the initial colonization of the Inner Banda Arc islands is bounded by the age of the islands, meaning that this colonization scenario could have been possible from approximately 10 Ma until present day.

Non-specific Models and Refinement of Hypotheses. Clearly, there are many plausible scenarios beyond those outlined in this set of hypotheses. It is possible that some lineages may have dispersed directly to interior islands from either the Sunda or Sahul Shelves. It is also possible that some lineages may have dispersed into the Lesser Sundas from northern Maluku, or even from southeastern Borneo. We should also not discount the possibility that the tectonic models that have served as the foundation for our hypothesis-generation are simply incorrect or oversimplified. We believe that it is quite likely that further study of the historical biogeography of the regions fauna will not only shed light on the biological processes in this region, but will also be informative with regard to both the tectonic history and the history of island connectivity in the region. Just as biogeographers look to the geological record to help them explain historical processes affecting the evolution of their focal taxa, geologists stand to benefit by consulting the biological record when attempting to elucidate the history of complex island archipelagos.

TESTING BIOGEOGRAPHICAL HYPOTHESES FOR LESSER SUNDA TAXA

Testing the biogeographical hypotheses outlined above will require biogeographical study of lineages, species, or species complexes that occur on multiple islands in the archipelago. An ideal focal taxon (or lineage) would be one that a) occurs on all or most of the major islands, b) the source population/species outside of the Lesser Sundas is known, and c) all populations/species within the Lesser Sundas form a monophyletic clade, or at least there is no

evidence of multiple invasions of the archipelago. Examination of the relationships of island-specific lineages to each other along with the timing of entry into the archipelago and the ages of island-specific lineages will allow certain hypotheses to be ruled out and others to be considered plausible. By examining many different taxa with a variety of life history strategies and that have entered the archipelago at many different time points in the last 10-12 million years we can begin to formulate some general biogeographical rules that govern the diversification of Lesser Sundas taxa.

A first glance into the biogeographical patterns of the taxa that have colonized the archipelago will require genetic analysis. A first pass approach should utilize mitochondrial DNA sequence data in a rooted, time-calibrated phylogenetic analysis. Mitochondrial DNA can easily be obtained from nearly any species of reptile and amphibian both rapidly and inexpensively, and will allow for the screening of many individuals for each focal group. Mitochondrial DNA also has a high mutation rate compared to nuclear DNA and a lower effective population size making this marker more sensitive for detecting historical isolation over short time periods. This approach will allow for a rough estimate of the ages of each species or lineage, determine if island-specific populations have been isolated for a sufficient amount of time to accumulate unique shared mutations (monophyly), and estimate approximate ages of any monophyletic island-specific lineages. Once appropriate focal taxa are identified using this mtDNA approach, individuals of each focal group should be chosen for further analysis using multiple nuclear loci. These individuals should be chosen to represent not only all of the islands that focal group occurs on, but to maximize the potential genetic variation within available samples by choosing samples with haplotypes that capture the greatest amount of genetic variation. Analyzing nuclear multi-locus sequence data in a phylogenetic framework will confirm if the mitochondrial signal accurately represents the evolutionary history of the focal group and may remedy some of the effects of incomplete lineage sorting or mito-nuclear discordance (Toews & Brelsford, 2012).

CONCLUSIONS

The geological history and processes of the Wallacean region are extremely complex and have been changing over the last 20 million years. This collection of oceanic islands, situated between Asia and Australo-Papua, has accumulated a diverse and divergent faunal assemblage along with numerous endemic species. While there is still disagreement with respect to the significance of certain water barriers as biogeographical lines, it is safe to assume that most biologists working in the region would agree that the biogeographical patterns are a result of a very complex history of island formation, island rearrangement, sweepstakes dispersal events, environmental factors, and island-specific faunal community interactions. The Lesser Sunda Islands are just one region of Wallacea but they are unique in their linear arrangement and in the fact that the ends of the archipelago are close to both the Sunda and Sahul shelves, allowing colonization from both ends. By utilizing sequence data from both mitochondrial DNA and genomic data in a phylogenetic framework from lineages that have colonized the Lesser Sunda Islands biogeographers will have the opportunity to identify the timing and sequence of island colonization. Once this information is known it will be imperative to link these biogeographical patterns to the process of species formation, and to help reconstruct the geological history of the region. The Lesser Sunda Islands, a geologically complex tropical oceanic archipelago influenced by multiple biogeographic realms, still hold many secrets for biologists to discover.

REFERENCES

Abbot, M.J. & Chamaluan F.H. (1981) Geochronology of some Banda Arc volcanics. *In*: Barber, A.J. and Wiryosujono, S. (eds) *The Geology and Tectonics of Eastern Indonesia.* Geological Research and Development Centre, Bandung, Special Publication, 2:253–268.

Aiello, L.C. (2010) Five years of Homo floresiensis. *American Journal of Physical Anthropology*, 142(2):167–179.

AmphibiaWeb (2015) Information on amphibian biology and conservation. [web application] Available: http://amphibiaweb.org/. (Accessed: 12 December 2015).

Andersen, A.N., Kohout, R.J. & Trainor, C.R. (2013) Biogeography of Timor and surrounding Wallacean Islands: endemism in ants of the genus Polyrhachis Fr. Smith. *Diversity*, 5(1):139–148.

Audley-Charles, M.G. (1975) The sumba fracture: A major discontinuity between eastern and western Indonesia. *Tectonophysics*, 26:213-228.

Audley-Charles, M.G. (1986) Rates of Neogene and Quaternary Tectonic Movements in the Southern Banda Arc Based on Micropalaeontology. *Journal of the Geological Society*, 143:161–175.

Audley-Charles, M.G. (1987) Dispersal of Gondwanaland: relevance to evolution of the Angiosperms *In*: Whitmore, T.C. (ed.) *Biogeographical Evolution of the Malay Archipelago* Oxford Monographs on Biogeography 4, Clarendon Press, Oxford, pp. 5–25.

Audley-Charles, M.G. (2011) Tectonic post-collision processes in Timor. *Geological Society, London, Special Publications*, 355:241–266.

Auffenberg, W. (1980) The herpetofauna of Komodo, with notes on adjacent areas. Bulletin of the Florida State Museum of Biological Sciences, 25(2):39–156.

Azuma, N., Ogata, K., Kikuchi, T. & Higashi, S. (2006) Phylogeography of Asian weaver ants, Oecophylla smaragdina. *Ecological Research*, 21(1):126–136.

Beck, J., Kitching, I.J. & Linsenmair, K.E. (2006) Wallace's line revisited: has vicariance or dispersal shaped the distribution of Malesian hawkmoths (Lepidoptera: Sphingidae)?. *Biological Journal of the Linnean Society*, 89(3):455–468.

Bintanja, R., van de Wal, R.S.W. & Oerlemans, J. (2005) Modelled atmospheric temperatures and global sea levels over the past million years. *Nature*, 437:125–128.

Brongersma, L.D. (1936) Some comments upon H. C. Raven's paper: "Wallace's Line and the distribution of Indo-Australian mammals." *Archives Neerlandaises Zoologie*, 2:240–256.

Brouwer, H.A. (1939) Exploration in the Lesser Sunda Islands. *The Geographical Journal* 154:1–10.

Brown, R.M., Linkem, C.W., Siler, C.D., Sukumaran, J., Esselstyn, J.A., Diesmos, A.C., Iskandar, D.T., Bickford, D., Evans, B.J., McGuire, J.A., Grismer, L., Supriatna, J. & Andayani, N. (2010) Phylogeography and historical demography of Polypedates leucomystax in the islands of Indonesia and the Philippines: Evidence for recent human-mediated range expansion? *Molecular Phylogenetics and Evolution*, 57(2):598–619.

Brown, P., Sutikna, T., Morwood, M.J., Soejono, R.P., Saptomo, E.W. & Due, R.A. (2004) A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia. *Nature*, 431(7012):1055–1061.

Brumm, A., Jensen, G.M., van den Bergh, G.D., Morwood, M.J., Kurniawan, I., Aziz, F. & Storey, M. (2010) Hominins on Flores, Indonesia, by one million years ago. *Nature*, 464(7289):748–752.
Bush, M.B. & Whittaker, R.J. (1991) Krakatau: colonization patters and hierarchies. *Journal of Biogeography*, 18:341–356.
Carlquist, S. (1965) *Island life: a natural history of the islands of the world*. Natural History Press, New York.
Charlton, T.R., De Smet, M., Samodra, H. & Kaye, S.J. (1991) The stratigraphic and structural evolution of the Tanimbar islands, eastern Indonesia. *Journal of Southeast Asian Earth Sciences*, 6:343–358.
Ciofi, C., Beaumontf, M.A., Swingland, I.R. & Bruford, M.W. (1999) Genetic divergence and units for conservation in the Komodo dragon Varanus komodoensis. *Proceedings of the Royal Society of London B: Biological Sciences*, 266(1435):2269–2274.
Crowley, T.J. (2000) Causes of climate change over the past 1000 years. *Science*, 289(5477):270–277.
Darlington, P.J. (1957) Zoogeography: The geographical distribution of animals. Harvard University Press (Cambridge): 675 p.
David, P., Vogel, G. & Vidal, N. (2003) On Trimeresurus fasciatus (Boulenger, 1896) (Serpentes: Crotalidae), with a discussion on its relationships based on morphological and molecular data. *Raffles Bulletin of Zoology*, 51(1):149–158.
De Boer, A.J. & Duffels, J.P. (1996) Historical biogeography of the cicadas of Wallacea, New Guinea and the West Pacific: a geotectonic explanation. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 124(1):153–177.
de Jong Boers, B. (1995) Mount Tambora in 1815: A volcanic eruption in Indonesia and its aftermath. *Indonesia*, 60:36–60.
de Lang, R. (2011) *The snakes of the Lesser Sunda Islands (Nusa Tenggara), Indonesia: a field guide to the terrestrial and semi-aquatic snakes with identification key*. Edition Chimaira.
Dickerson, R.E., Merrill, E.D., McGregor, R.C., Schultze, W., Taylor, E.H. & Herre, A.W. (1928) *Distribution of life in the Philippines* (No. 21). Bureau of Printing, Manila.
Dunn, E.R. (1927a) Results of the Douglas Burden expedition to the island of Komodo. II. Snakes from the East Indies. American Museum Novitates, 287:1–7.
Dunn, E.R. (1927b) Results of the Douglas Burden expedition to the island of Komodo. III. Lizards from the East Indies. American Museum Novitates, 288:1–13.
Dunn, E.R. (1928) Results of the Douglas Burden expedition to the island of Komodo. II. Frogs from the East Indies. American Museum Novitates, 315:1–9.
Emile-Geay, J.R., Seager, R., Cane, M.A., Cook, E.R. & Haug G.H. (2008) Volcanoes and ENSO over the past millennium. *Journal of Climatology*, 21(13):3134–3148.
Fairbanks, R.G. (1989) A 17,000 year glacio-eustatic sea level record: influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation. *Nature* 342:637–642.
Fontijn, K., Costa, F., Sutawidjaja, I., Newhall, C.G. & Herrin, J.S. (2015) A 5000-year record of multiple highly explosive mafic eruptions from Gunung Agung (Bali, Indonesia): implications for eruption frequency and volcanic hazards. *Bulletin of Volcanology*, 77:59.
Fortuin, A.R., Roep, T.B. & Sumosusastro, P.A. (1994) The Neogene sediments of east Sumba, Indonesia—products of a lost arc? *Journal of Southeast Asian Earth Sciences,* 9:67–79.
Fortuin, A.R., VanderWerff, W. & Wensink, H. (1997) Neogene basin history and

paleomagnetism of a rifted and inverted forearc region, on- and offshore Sumba, Eastern Indonesia. *Journal of Asian Earth Sciences* 15:61–88.
Glover, I.C. (1971) Prehistoric research in Timor. In *Aboriginal Man and Environment in Australia*. eds D.J. Mulvaney & J. Golson, pp. 158-181. Canberra: Australia National University.
Goltenboth, F., Timotius, K.H., Milan, P.P. & Margraf, J. (2006) *Ecology of Insular Southeast Asia: The Indonesian Archipelago*. Elsevier.
Greer, A.E. (1990) The Glaphyromorphus isolepis species group (Lacertilia: Scincidae): diagnosis of the taxon and description of a new species from Timor. *Journal of Herpetology*, 24(4):372–377.
Hall, R. (2009) Southeast Asia's changing paleogeography. *Blumea*, 54:148–161.
Hall, R. (2011) Australia-SE Asia collision: plate tectonics and crustal flow. Geological Society,London, Special Publications, 355:75–109.
Hall, R., Cottam, M.A. & Wilson, M.E.J. (2011) *The SE Asian Gateway: History and Tectonics of the Australia-Asia Collision.* Geological Society, London, Special Publications, 355:1–7.
Heaney, L.R. (1991) A synopsis of climatic and vegetational change in Southeast Asia. *Climatic Change*, 19:53–61.
Henneberg, M., Eckhardt, R.B., Chavanaves, S. & Hsü, K.J. (2014) Evolved developmental homeostasis disturbed in LB1 from Flores, Indonesia, denotes Down syndrome and not diagnostic traits of the invalid species Homo floresiensis. *Proceedings of the National Academy of Sciences USA*, 111(33):11967–11972.
Hisheh, S., Westerman, M. & Schmitt, L.H. (1998) Biogeography of the Indonesian archipelago: mitochondrial DNA variation in the fruit bat, Eonycteris spelaea. *Biological Journal of the Linnean Society*, 65(3):329–345.
Holloway, J.D. & Jardine, N. (1968) Two approaches to zoogeography: a study based on the distributions of butterflies, birds and bats in the Indo-Australian area. In *Proceedings of the Linnean Society of London*, 179(2):153–188.
Holt, B.G., Lessard, J.P., Borregaard, M.K., Fritz, S. A., Araújo, M.B., Dimitrov, D., Fabre, P., Graham, C.H., Graves, G.R., Jonsson, K.A., Nogues-Bravo, D., Wang, Z., Whittaker, R.J., Fjeldsa, J. & Rahbek, C. (2013) An update of Wallace's zoogeographic regions of the world. *Science*, 339(6115):74–78.
How, R.A. & Kitchener, D.J. (1997) Biogeography of Indonesian snakes. *Journal of Biogeography*, 24(6):725–735.
How, R.A., Schmitt, L.H. & Suyanto, A. (1996). Geographical variation in the morphology of four snake species from the Lesser Sunda Islands, eastern Indonesia. *Biological Journal of the Linnean Society*, 59(4):439-456.
Huxley, T.H. (1868) On the classification and distribution of the Alectoromorphae and Heteromorphae. Proccedings of the Zoological Society of London, 1868:294–319.
Inger, R.F. (1999) Distribution of amphibians in southern Asia and adjacent islands. *Patterns of Distribution of Amphibians: A global perspective*, 445–482.
Ishikawa, A., Kaneko, Y., Kadarusman, A. & Ota, T. (2007) Multiple generations of forearc mafic–ultramafic rocks in the Timor–Tanimbar ophiolite, eastern Indonesia. *Gondwana Research*, 11:200–217.
IUCN (2015) The IUCN Red List of Threatened Species. Version 2015–4. Available: http://www.iucnredlist.org. Accessed: 12 December 2015.

Jacob, T., Indriati, E., Soejono, R.P., Hsü, K., Frayer, D.W., Eckhardt, R.B., Kuperavage, A.J., Thorne, A. & Henneberg, M. (2006) Pygmoid Australomelanesian Homo sapiens skeletal remains from Liang Bua, Flores: population affinities and pathological abnormalities. *Proceedings of the National Academy of Sciences USA*, 103(36):13421–13426.

Kaiser, H., Carvalho, V.L., Ceballos, J., Freed, P., Heacox, S., Lester, B., Richards, S.J., Trainor, C. T., Caitlin, S. & O'Shea, M. (2011) The herpetofauna of Timor-Leste: a first report. *ZooKeys*, 109:19–86.

Kaiser, H., Sanchez, C., Heacox, S., Kathriner, A., Ribeiro, A.V., Soares, Z.A., de Araujo, L. L., Mecke, S. & O'Shea, M. (2013) First report on the herpetofauna of Ataúro Island, Timor-Leste. *Check List*, 9(4):752–762.

Kaneko, Y., Maruyama, S., Kadarusman, A., Ota, T., Ishikawa, M., Tsujimori, T., Ishikawa, A., Tsujimori, T., Ishikawa, A. & Okamoto, K. (2007) On-going orogeny in the outer-arc of the Timor-Tanimbar region, eastern Indonesia. *Gondwana Research*, 11:218–233.

Katili, J.A. (1975) Volcanism and plate tectonics in the Indonesian island arcs. *Tectonophysics*, 26(3):165–188.

Kitchener, D.J., Hisheh, S., Schmitt, L.H. & Maryanto, I. (1993a) Morphological and genetic variation in *Aethalops alecto* (Chiroptera, Pteropodidae) from Java, Bali, and Lombok Is, Indonesia. *Mammalia*, 57:255–272.

Kitchener, D.J., Schmitt, L.H., Hisheh, S., How, R.A., Cooper, N.K. & Maharadatunkamsi. (1993b) Morphological and genetic variation in the bearded tomb bats (Taphozous, Emballonuridae) of Nusa Tenggara, Indonesia. *Mammalia*, 57:63–83.

Kitchener, D.J., Hisheh, S., Schmitt, L.H. & Suyanto, A. (1994a) Shrews (Soricidae: Crocidura) from the Lesser Sunda Islands, and South-East Maluku, eastern Indonesia. *Australian Mammalogy*, 17:7–17.

Kitchener, D.J., Schmitt, L.H. & Maharadatunkamsi (1994b) Morphological and genetic variation in Suncus murinus (Soricidae, Crocodurinae) from Java, Lesser Sunda Islands, Maluku, and Sulawesi, Indonesia. *Mammalia*, 58:438–451.

Kitchener, D.J., Hisheh, S. & Schmitt, L.H. (1997) Morphological and Genetic Variation Among Island Populations of *Dobsonia Peronii* (Chiroptera: Pteropodidae) From the Lesser Sunda Islands, Indonesia. *Tropical Biodiversity*, 4:35–51.

Lavigne, F., Degeai, J.P., Komorowski, J.C., Guillet, S., Robert, V., Lahitte, P., Oppenheimer, C., Stoffel, M., Vidal, C.M., Pratomo, I., Wassmer, P., Hajdas, I., Sri Hadmoko, D. & de Belizal, E. (2013) Source of the great AD 1257 mystery eruption unveiled, Samalas volcano, Rinjani Volcanic Complex, Indonesia. *Proceedings of the National Academy of Sciences*, 110(42):16742–16747.

Leonard, J.A., Tex, R.J., Hawkins, M.T., Muñoz-Fuentes, V., Thorington, R. & Maldonado, J. E. (2015) Phylogeography of vertebrates on the Sunda Shelf: a multi-species comparison. *Journal of Biogeography*, 42(5):871–879.

Lohman, D.J., de Bruyn, M., Page, T., von Rintelen, K., Hall, R., Ng, P. K., Shih, H., Carvalho, G. R., & von Rintelen, T. (2011) Biogeography of the Indo-Australian archipelago. *Annual Review of Ecology, Evolution, and Systematics*, 42:205–226.

Lydekker, R. (1896) A geographical history of mammals. University Press, Cambridge.

MacArthur, R.H. & Wilson, E.O. (1967). *The theory of island biogeography* (Vol. 1). Princeton University Press.

MacKinnon, J., Beudels, R., Robinson, A. & Artha, M.B. (1982) *National Conservation Plan for Indonesia Vol IV: Nusa Tenggara*. UNDP/FAO National Parks Development Project. FAO, Bogor, Indonesia.

Maharadatunkamsi, Hisheh, S., Kitchener, D.J. & Schmitt, L.H. (2000) Genetic and morphometric diversity in Wallacea: geographical patterning in the horse shoe bat, *Rhinolophus affinis*. *Journal of Biogeography*, 27:193–201.

Maharadatunkamsi, Hisheh, S., Kitchener, D.J. & Schmitt, L.H. (2003) Relationships between morphology, genetics and geography in the cave fruit bat *Eonycteris spelaea* (Dobson, 1871) from Indonesia. *Biological Journal of the Linnean Society*, 79:511–522.

Malhotra, A., & Thorpe, R.S. (2004) A phylogeny of four mitochondrial gene regions suggests a revised taxonomy for Asian pitvipers (Trimeresurus and Ovophis). *Molecular Phylogenetics and Evolution*, *32*(1):83–100.

Marrison, G.E. (1999) Sasak and Javanese Literature of Lombok (KITLV, Leiden, Netherlands).

Mayr, E. (1944) Wallace's line in the light of recent zoogeographic studies. *The Quarterly Review of Biology*, 19(1):1–14.

McGuire, J.A. & Heang, K.B. (2001) Phylogenetic systematics of Southeast Asian flying lizards (Iguania : Agamidae : *Draco*) as inferred from mitochondrial DNA sequence data. *Biological Journal of the Linnean Society*, 72:203–229.

Mertens, R. (1930) Die Amphibien and Reptilien der inseln Bali, Lombok, Sumbawa and Flores. *Abhandlungen Senckenbergischen Naturforschenden Gesellschaft*, 42:115–344.

Michaux, B. (1994) Land movements and animal distributions in east Wallacea (eastern Indonesia, Papua New Guinea and Melanesia). *Palaeogeography, Palaeoclimatology, Palaeoecology*, 112(3):323–343.

Michaux, B. (2010) Biogeology of Wallacea: geotectonic models, areas of endemism, and natural biogeographical units. *Biological Journal of the Linnean Society*, 101(1):193–212.

Monk, K.A., Fretes, Y. & Reksodiharjo-Lilley, G. (1997) *The Ecology of Nusa Tenggara and Maluku*. Hong Kong: Periplus Editions Ltd. pp. 1–966.

Morwood, M.J., Brown, P., Sutikna, T., Saptomo, E.W., Westaway, K.E., Due, R.A., Roberts, G., Maeda, T., Wasisto, S. & Djubiantono, T. (2005) Further evidence for small-bodied hominins from the Late Pleistocene of Flores, Indonesia. *Nature*, 437(7061):1012–1017.

Muller, S. (1846) Uber den charakter der thierwelt auf den inseln des indischen archileps. *Archiv fur Naturgeschichte*, 12:109–128.

Murray, A. (1866) *The geographical distribution of mammals*. Day and Son, London.

O'Connor, S. (2007) New evidence from East Timor contributes to our understanding of earliest modern human colonisation east of the Sunda Shelf. *Antiquity*, 81(313):523–535.

O'Shea, M., Sanchez, C., Heacox, S., Kathriner, A., Carvalho, V.L., Ribeiro, A.V., Soares, Z.A., Araujo, L.D. & Kaiser, H. (2012) First update to herpetofaunal records from Timor-Leste. *Asian Herpetological Research*, 3(2):114–126.

Ota, H., Darevsky, I.S., Ineich, I. & Yamashiro, S. (2000) Reevaluation of the taxonomic status of two Lepidodactylus species (Squamata: Gekkonidae) from the Lesser Sunda Archipelago, Indonesia. *Copeia*, 2000(4):1109–1113.

Outlaw, D.C. & Voelker, G. (2008) Pliocene climatic change in insular Southeast Asia as an engine of diversification in Ficedula flycatchers. *Journal of Biogeography*, 35(4):739–752.

Pelseneer, P. (1904) La "Ligne de Weber," limite zoologique de l'Asie et de l'Australie. *Bulletin de la Classe des Sciences de Academie Royalc Belgique*, 1904:1001–1022.
Pirazzoli, P.A., Radtke, U., Hantoro, W.S., Jouannic, C., Hoang, C.T., Causse, C. & Best, M.B. (1993) A One Million-Year-Long Sequence of Marine Terraces on Sumba Island, Indonesia. *Marine Geology*, 109:221–236.
Rampino, M.R. & Self, S. (1982) Historic eruptions of Tambora (1815), Krakatau (1883), and Agung (1963), their stratospheric aerosols, and climate impact. *Quaternary Research*, 18:127–143.
Raven, H.C. (1935) Wallace's Line and the distribution of Indo-Australian mammals. *Bulletin of the Amererican Museum of Natural History*, 68:177–293.
Rawlings, L.H., Barker, D. & Donnellan, S.C. (2004) Phylogenetic relationships of the Australo-Papuan Liasis pythons (Reptilia: Macrostomata), based on mitochondrial DNA. *Australian Journal of Zoology*, 52(2):215–227.
Rawlinson, P.A. & van Balen, S. (1992) Colonization of the Krakatau Islands by vertebrates. *GeoJournal*, 28(2):225–231.
Rensch, B. (1936) Die Geschichte des Sundabogens. Eine tiergeographische Untersuchung. Pp. 1-318. Berlin, Borntraeger.
Riedel, A., Tänzler, R., Balke, M., Rahmadi, C. & Suhardjono, Y.R. (2014) Ninety-eight new species of Trigonopterus weevils from Sundaland and the Lesser Sunda Islands. *ZooKeys*, 467:1–162.
Rigg, J.W. & Hall, R. (2011) Structural and stratigraphic evolution of the Savu Basin, Indonesia. *In:* Hall, R., Cottam, M. A. &Wilson M. E. J. (eds), *The SE Asian Gateway: History and Tectonics of the Australia–Asia Collision*. Geological Society, London, Special Publications, 355:225–240.
Roomsawati, N. & Harris, R. (2009) Surface uplift history of the incipient Banda arc-continent collision: geology and synorogenic foraminifera of Rote and Savu Islands, Indonesia. *Tectonophysics*, 479:95–110.
Sanchez, C., Carvalho, V.L., Kathriner, A., O'Shea, M. & Kaiser, H. (2012) First report on the herpetofauna of the Oecusse District, an exclave of Timor-Leste. *Herpetology Notes*, 5:137–149.
Schmitt, L.H., Kitchener, D.J. & How, R.A. (1995) A genetic perspective of mammalian variation and evolution in the Indonesian Archipelago - biogeographic correlates in the fruit bat genus *Cynopterus*. *Evolution*, 49:399–412.
Schmitt, L.H., How, R.A., Hisheh, S., Goldberg, J. & Maryanto, I. (2000) Geographic patterns in genetic and morphological variation in two skink species along the Banda Arcs, southeastern Indonesia. *Journal of Herpetology*, 34:240–258.
Schmitt, L.H., Hisheh, S., Suyanto, A., Maharadatunkamsi, Newbound, C.N., Kitchener, D.J. & How, R.A. (2009) Crossing the Line: The impact of contemporary and historical sea barriers on the population structure of bats in southern Wallacea. *In:* Fleming, T.H., and Racey, P.A. (eds), *Island Bats: Evolution, Ecology, and Conservation.* The University of Chicago Press.
Scotney, P.M., Roberts, S., Herrington, R.J., Boyce, A.J. & Burgess, R. (2005) The development of volcanic hosted massive sulfide and barite gold orobodies on Wetar Island, Indonesia. *Mineralium Deposita*, 40:76–99.

Shea, G.M. (2012) On the identity of the type species of Sphenomorphus (Squamata: Scincidae): Lygosoma melanopogon Duméril and Bibron 1839, with a note on a new scalation character of the pes in Sphenomorphus. *Zootaxa*, 3490:1–29.

Simpson, G.G. (1977) Too many lines; the limits of the Oriental and Australian zoogeographic regions. *Proceedings of the American Philosophical Society*, 121(2):107–120.

Sondaar, P.Y., Van den Berg, G.D., Mubroto, B., Aziz, F., De Vos, J. & Batu, U.L. (1994) Middle Pleistocene faunal turnover and colonization of Flores (Indonesia) by *Homo erectus*. *Comptes rendus de l'Académie des sciences. Série 2. Sciences de la terre et des planètes*, 319(10):1255–1262.

Spakman, W. & Hall, R. (2010) Surface deformation and slab-mantle interaction during Banda Arc subduction rollback. *Nature Geoscience*, 3:562–566.

Stone, D. (1994) Tanah Air- Indonesia's Biodiversity. Editions Didier Millet, Singapore: 208 pp.

Stresemann, E. (1939) Die vogel von Celebes. *Zoogeographie. Journal of Ornithologie*, 87:312–425.

Sweet, S.S. & Pianka, E.R. (2007) Monitors, mammals and Wallace's Line. *Mertensiella*, 16:79–99.

Tänzler, R., Toussaint, E.F., Suhardjono, Y.R., Balke, M. & Riedel, A. (2014) Multiple transgressions of Wallace's Line explain diversity of flightless Trigonopterus weevils on Bali. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1782):20132528.

Tänzler, R., Van Dam, M.H., Toussaint, E.F., Suhardjono, Y.R., Balke, M. & Riedel, A. (2016) Macroevolution of hyperdiverse flightless beetles reflects the complex geological history of the Sunda Arc. *Scientific Reports*, 6.

Thomson, V., Aplin, K.P., Cooper, A., Hisheh, S., Suzuki, H., Maryanto, I., Yap, G. & Donnellan, S.C. (2014) Molecular genetic evidence for the place of origin of the Pacific rat, *Rattus exulans*. *PLoS One*, 9(3):e91356.

Thornton, I.W.B. & New, T.R. (1988) Krakatau invertebrates: the 1980s fauna in the context of a century of recolonization. *Philosophical Transactions of the Royal Society of London B*, 322:493–522.

Thornton, I.W.B., Zann, R.A., Rawlinson, P.A., Tidemann, C.R., Adikerana, A.S. & Widjoya, A.H.T. (1988) Colonization of the Krakatau Islands by vertebrates: equilibrium, succession, and possible delayed extinction. *Proceedings of the National Academy of Sciences USA*, 85:515–518.

Thorpe, R.S., Pook, C.E. & Malhotra, A. (2007) Phylogeography of the Russell's viper (*Daboia russelii*) complex in relation to variation in the colour pattern and symptoms of envenoming. *The Herpetological Journal*, 17(4):209–218.

Tocheri, M.W., Orr, C.M., Larson, S.G., Sutikna, T., Saptomo, E.W., Due, R.A., Djubiantono, T., Morwood, M.J. & Jungers, W.L. (2007) The primitive wrist of Homo floresiensis and its implications for hominin evolution. *Science*, 317(5845):1743–1745.

Toews, D.P. & Brelsford, A. (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, 21(16):3907–3930.

Trainor, C.R. (2002) Birds of Gunung Tambora, Sumbawa, Indonesia: effects of altitude, the 1815 cataclysmic volcanic eruption and trade. *Forktail*, 18:49–61.

Trainor, C. & Lesmana, D. (2000) Exploding volcanoes, unique birds, gigantic rats and elegant ikat: identifying sites of international conservation significance on Flores, East Nusa Tenggara. *Bogor: PKA/BirdLife International/WWF*.

Uetz, P. & Hallermann, J. (2010) The JCVI/TIGR reptile database.
Van den Bergh, G.D., de Vos, J. & Sondaar, P.Y. (2001) The Late Quaternary palaeogeography of mammal evolution in the Indonesian Archipelago. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 171(3):385–408.
Van Kampen, P.N. (1909) De zoogeographie van den Indiscne Archipel. *Bijbl. 3 en 4, Natuurk. Tijdschr. Ned. Indie.*
Vane-Wright, R.I. (1991) Transcending the Wallace line: do the western edges of the Australina region and the Australian plate coincide?. *Australian Systematic Botany*, 4(1):183–197.
Veevers, J.J. (1991) Phanerozoic Australia in the changing configuration of Proto-Pangea through Gondwanaland and Pangea to the present dispersed continents. *Australian Systematic Botany*, 4(1):1–11.
Voris, H.K. (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, 27(5):1153–1167.
Wallace, A.R. (1860) On the zoological geography of the Malay Archipelago. *Journal of the Linnean Society, London*, 14:172–184.
Wallace, A.R. (1869) *The Malay Archipelago: The Land of the Orangutan and the Bird of Paradise. A Narrative of Travels with Studies of Man and Nature.* MacMillan, London.
Wallace, A.R. (1876) *The Geographical Distribution of Animals: With a Study of the Relations of Living and Extinct Faunas as Elucidating the Past Changes of the Earth's Surface.* Macmillan, London.
Wallace, A.R. (1880) *Island life: the phenomena and causes of insular faunas and floras*. McMillan, London.
Weber, M. (1902) Der Indo-australische Archipel und die Geschichte seiner Tierwelt. 46pp. *Jena*.
Wegener, A. (1915) *The Origin of Continents and Oceans*. Friedrich Vieweg & Sohn.
White, C.M.N. & Bruce, M.D. (1986) *The birds of Wallacea (Sulawesi, the Moluccas and Lesser Sunda Islands, Indonesia). An annotated checklist.* British Ornithological Union Checklist No.7. London: Zoological Society of London.
Whitmore, T.C. (1987) *Biogeographical Evolution of the Malay Archipelago*. Clarendon Press.
Whittaker, R.J., Bush, M.B. & Richards, K. (1989) Plant recolonization and vegetation succession on the Krakatau Islands, Indonesia. *Ecological Monographs*, 59(2):59–123.
Whittaker, R.J. & Fernández-Palacios, J.M. (2007) *Island Biogeography: Ecology, Evolution, and Conservation*. Oxford University Press.
Whitten, A.J., Bishop, K.D., Nash, S.V. & Clayton, L. (1987) One or More Extinctions born Sulawesi, Indonesia? *Conservation Biology*, 1(1):42–48.
Wiantoro, S., Maryanto, I. & Abdullah, M.T. (2012) Phylogeny and phylogeography of Myotis muricola (Gray, 1846) (Chiroptera: Vespertilionidae) from the West and East of Wallace's Line inferred from partial mtDNA cytochrome b gene. *Pertanika Journal of Tropical Agricultural Science*, 35(2):271–292.
Wood, P.L., Heinicke, M.P., Jackman, T.R. & Bauer, A.M. (2012) Phylogeny of bent-toed geckos (Cyrtodactylus) reveals a west to east pattern of diversification. *Molecular Phylogenetics and Evolution*, 65(3):992–1003.
Woodruff, D.S. (2010) Biogeography and conservation in Southeast Asia: how 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining refugial-phase biodiversity. *Biodiversity and Conservation*, 19(4):919–941.

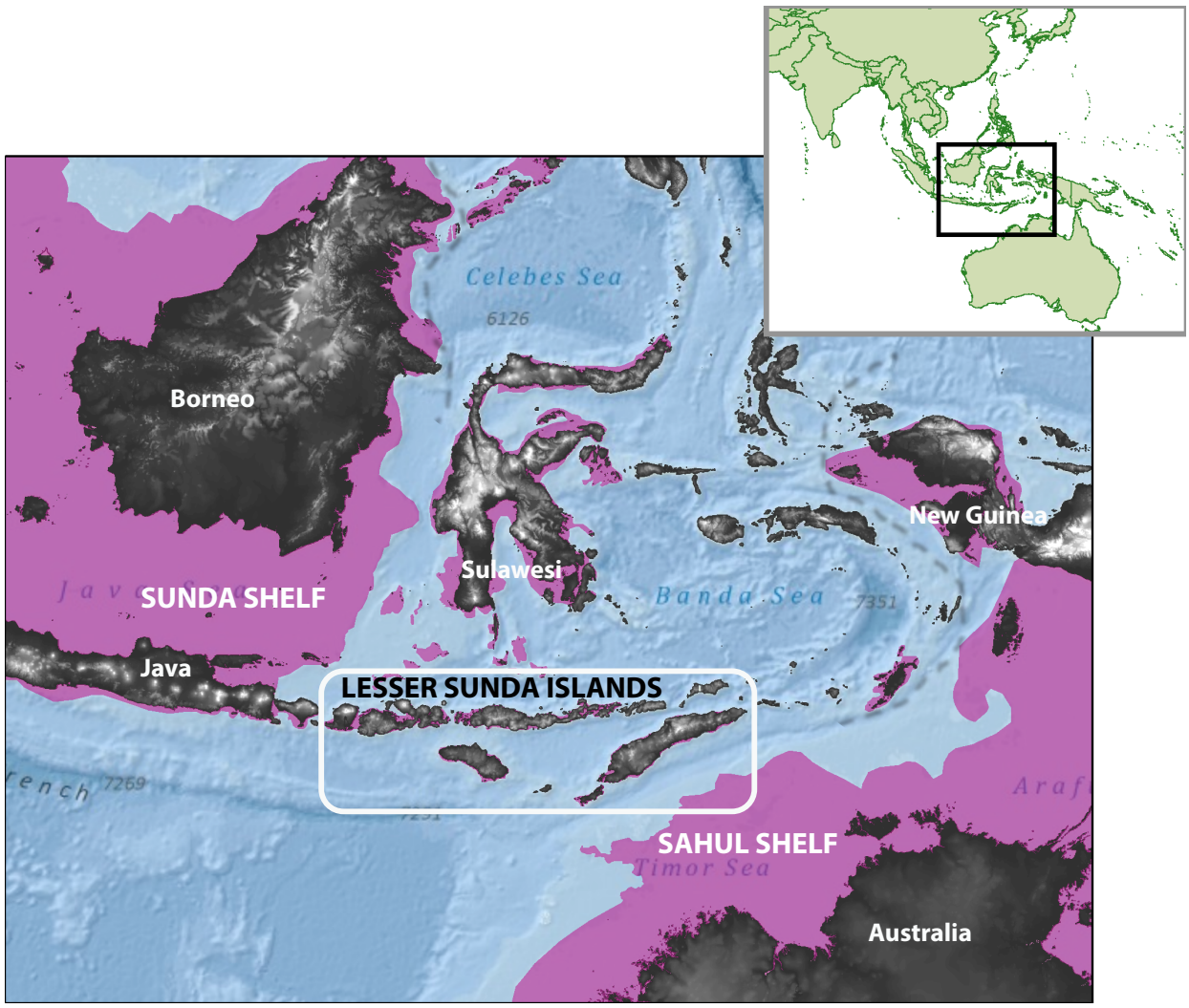


Figure 1. Map of eastern Indonesia. Islands are represented in black/white shading with white colors depicting higher elevation. Purple portions of the ocean represent areas that become land-positive during lower sea levels associated with glacial cycles.

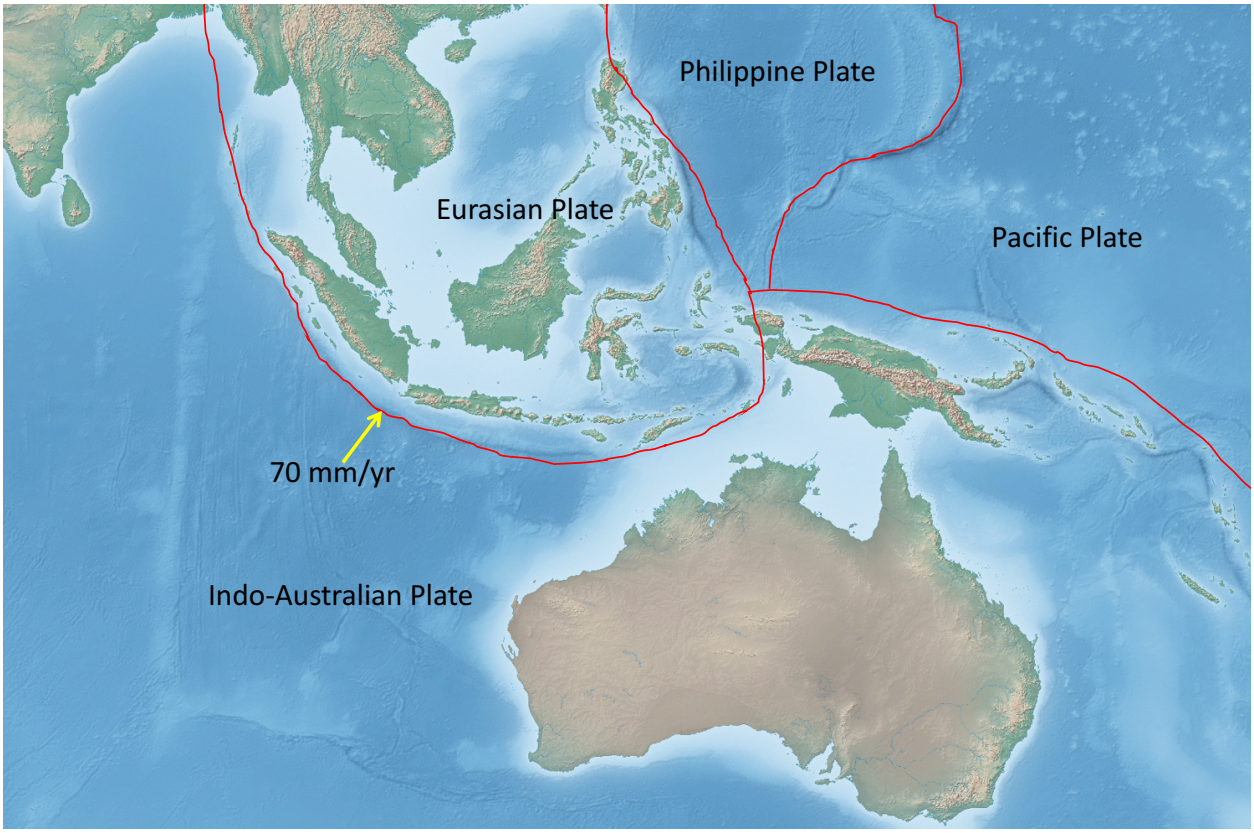


Figure 2. Map of the four major tectonic plates that occur in the Indo-Pacific region. Red lines represent plate boundaries.

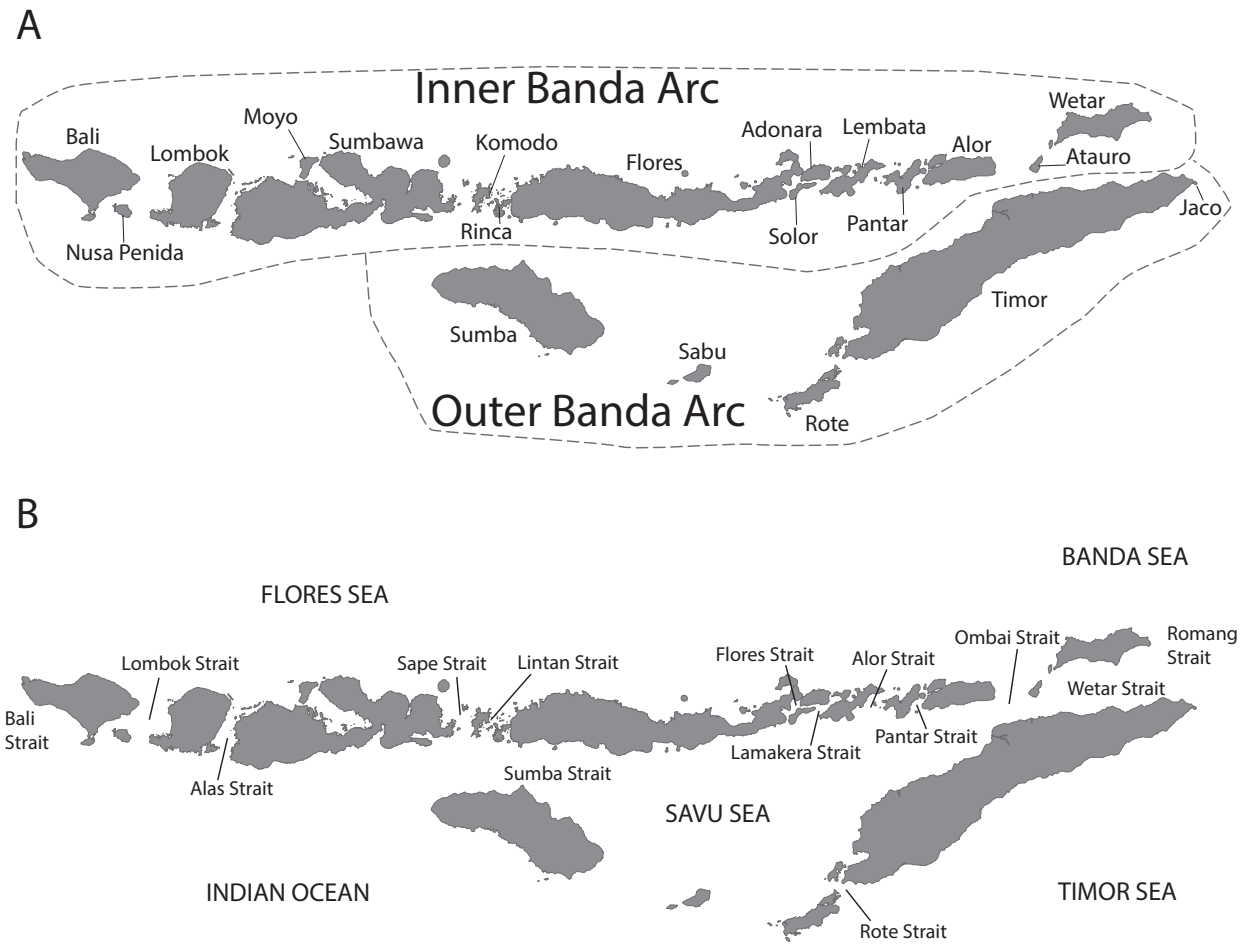


Figure 3. A) Map of the Lesser Sunda Islands and names of relevant islands. The Inner and Outer Banda Arcs are captured in dashed lines, and these arcs extend further to the east. B) Map of relevant oceanic basins and barriers.

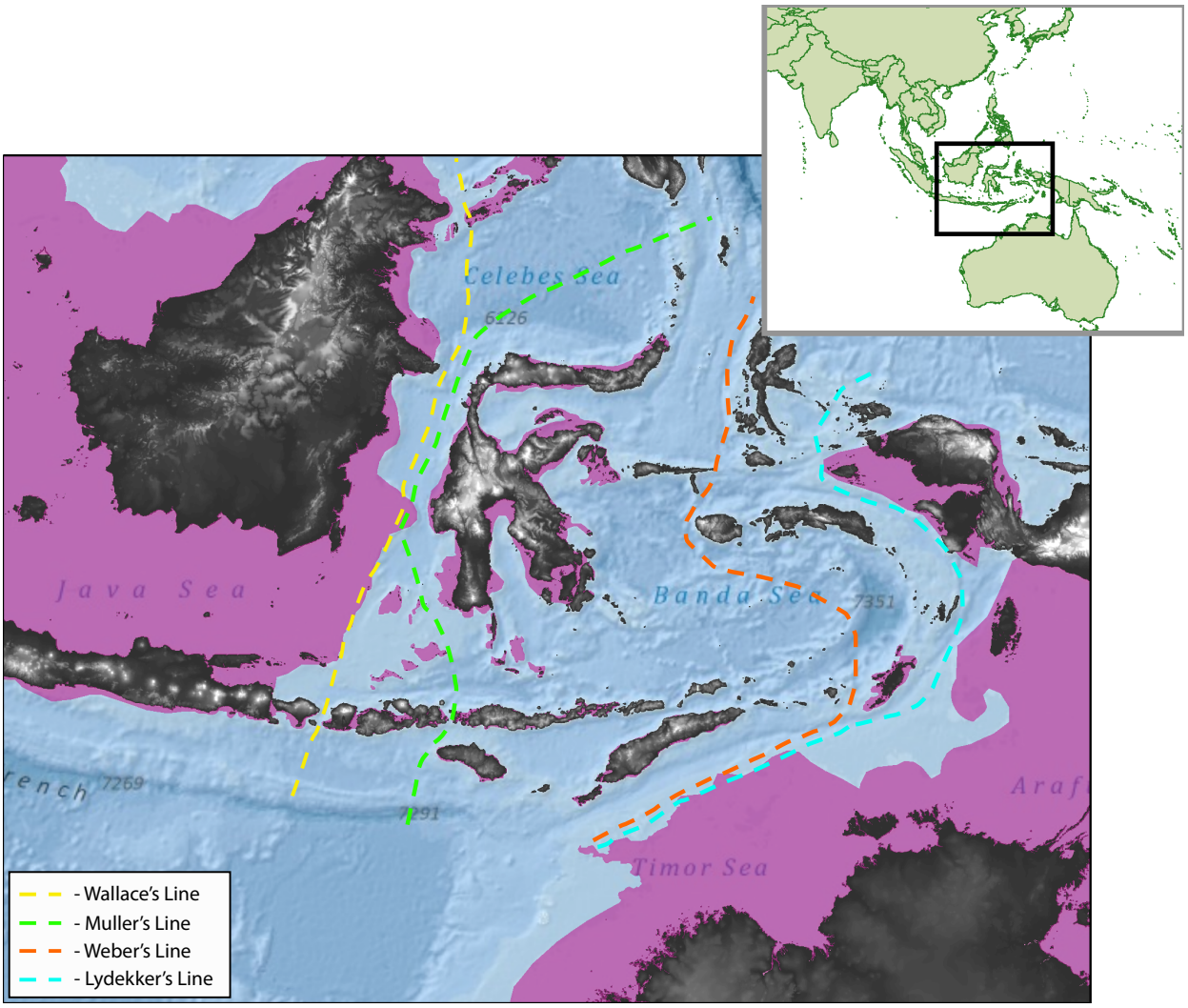


Figure 4. Relevant biogeographical lines drawn through eastern Indonesia.

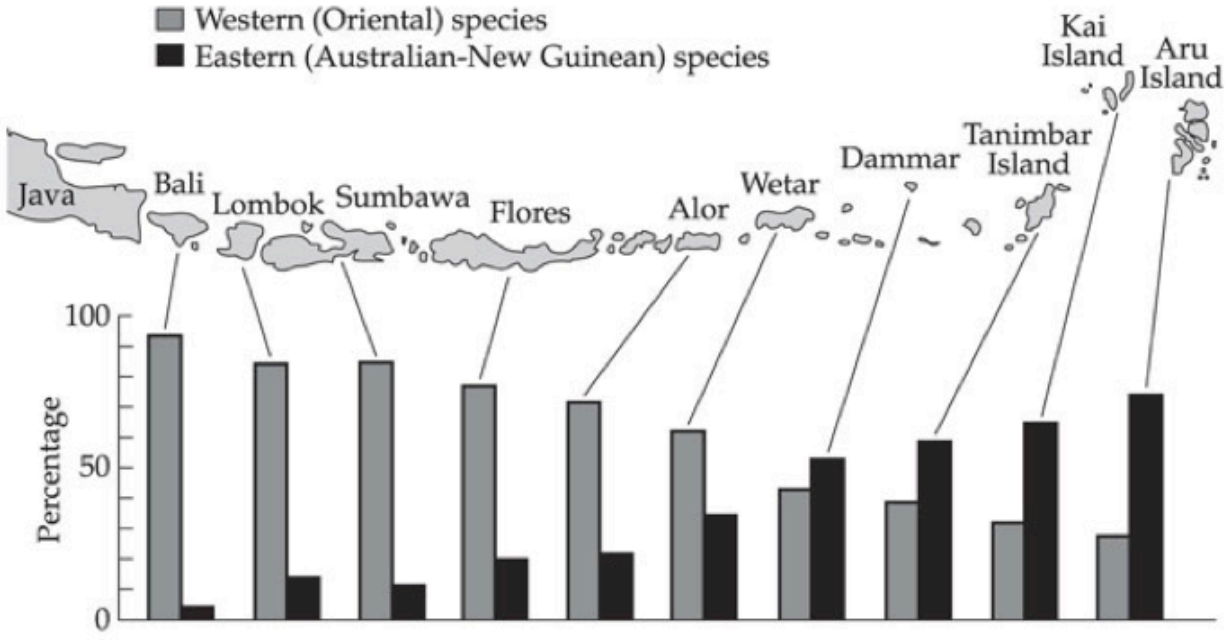


Figure 5. Proportions of reptile species that originate from either the Oriental or Australo-Papuan realms, starting with the Sunda Shelf island of Bali through the Banda Arc and ending at the Sahul Shelf Island of Aru (Carlquist, 1965).

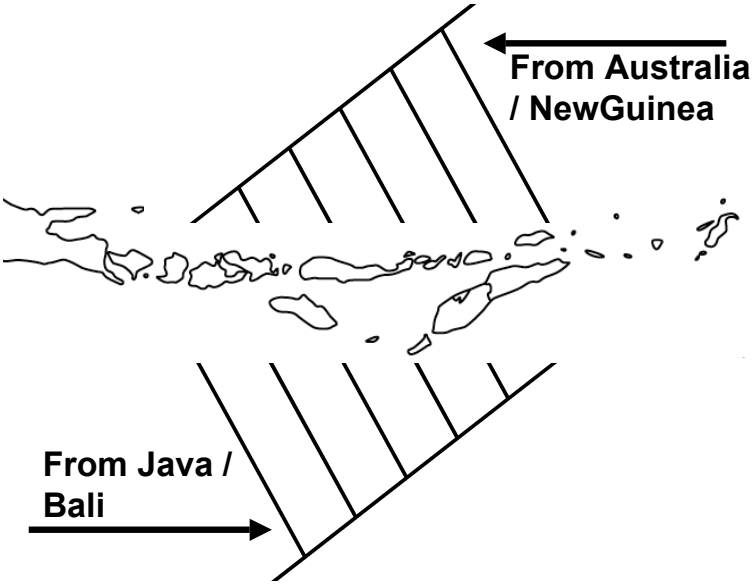


Figure 6. Expected pectinate topologies for taxa dispersing west to east (bottom) or from east to west (top) under the stepping-stone model of island colonization.

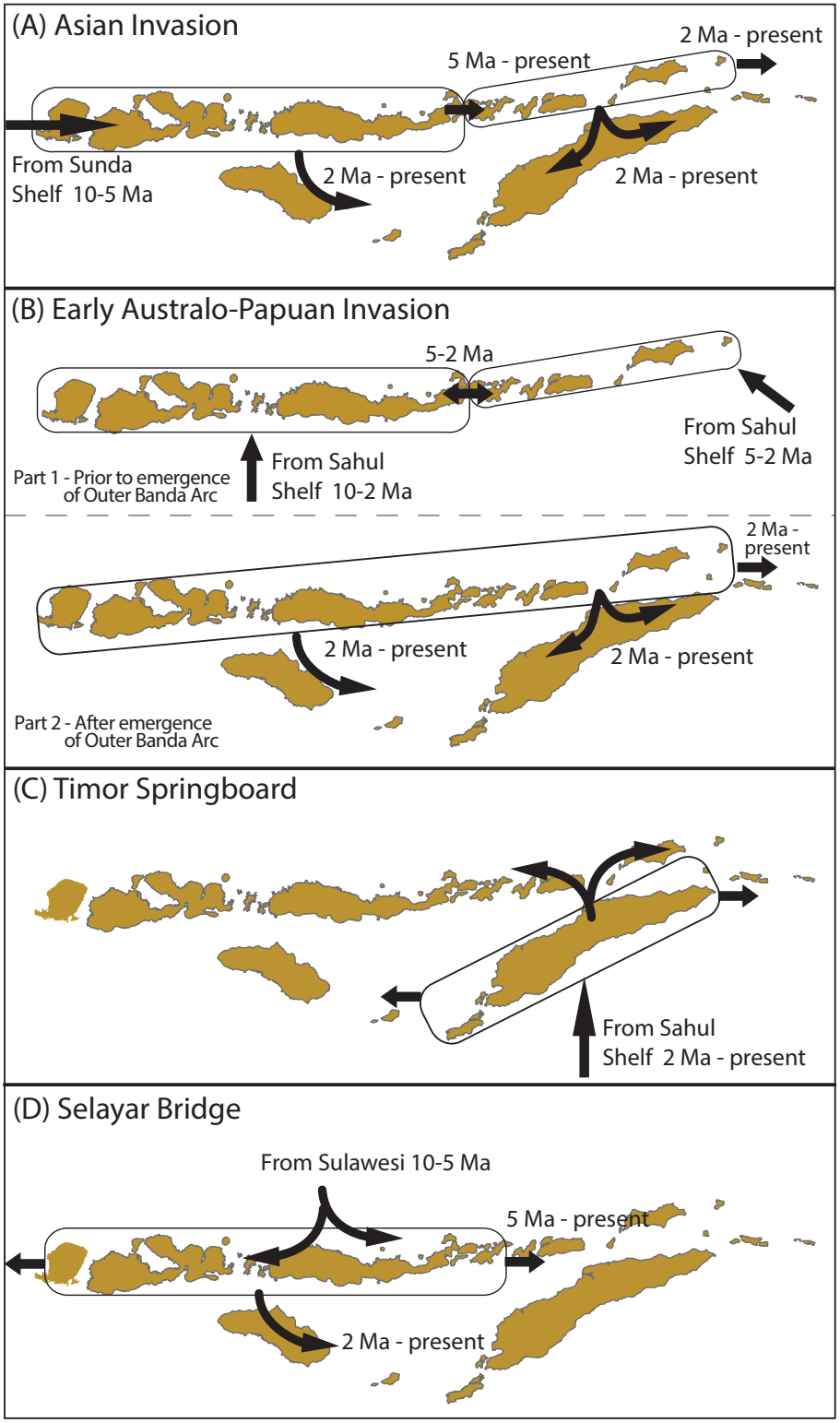


Figure 7. Island-age biogeographical hypotheses for reptiles and amphibians colonizing the Lesser Sunda Islands. Descriptions of each scenario are found in the text.

CHAPTER 2

Biogeographical History of Flying Lizards (Genus: *Draco*) From the Lesser Sunda Islands of Indonesia: Testing the Stepping-Stone Hypothesis Using a Phylogenomic Approach

ABSTRACT

Patterns of distribution of reptiles of Asian and Australo-Papuan origin suggest that the Lesser Sunda Islands may act as a two-way filter, and that this pattern is in part due to a stepping-stone model of island colonization. While no phylogenetic data for any reptile lineage supports a stepping-stone model of colonization, there are many appropriate reptile taxa that can be investigated in a phylogenetic framework in order to test this hypothesis. Flying lizards of the genus *Draco* occur throughout the Lesser Sunda Islands of Indonesia with two currently recognized species inhabiting the area. The biogeographical history of the *Draco* of the Lesser Sundas was initially inferred using phylogenetic analysis of mitochondrial *ND2* sequence data. The mitochondrial results confirm that the Lesser Sundas assemblage is monophyletic and split from its sister lineage, *D. volans* from Java and Bali, across Wallace's Line approximately 11 million years ago. Further mitochondrial data collected for 372 additional samples in this study recover island-specific lineages and a topology that is consistent with a general west-to-east colonization route. Lizards from the larger islands of Sumbawa, Flores, and Timor are paraphyletic and these lineages are parapatrically distributed. These results were then used to select 100 samples for further data collection using an exon-capture approach. After sequencing, aligning, and filtering four *Draco* transcriptomes a set of 709 nuclear genes (~1 Mb total target) was identified for an in-solution exon capture experiment. All genes were successfully captured and sequenced resulting in ~150X coverage of targeted regions. Both Maximum Likelihood and coalescent-based species tree phylogenies converged on a similar, well-supported topology that differed from the mitochondrial phylogeny in a number of ways, but still support an island-age influenced west-to-east colonization pattern whereby the eastern islands were colonized by way of Sumba. The SNP based species tree converged on a different topology than the other methods, however this analysis only considers a fraction of the data. *Draco* of the Lesser Sundas possibly represent nine genetically distinct populations or species as determined by genetic cluster analysis, with these clusters reflecting major clades in the genomic phylogenies. The parapatrically distributed lineages that occupy West and East Sumbawa, as well as the lineages that occupy West and East Flores, appear to be experiencing gene flow, with highly asymmetrical migration occurring on Flores. In summary, flying lizards appear to have arrived in the Lesser Sunda Islands shortly after they become land positive, and have had a relatively complex colonization history across the archipelago. Rather than reaching the Inner Banda Arc island Pantar Island from the neighboring Lembata, flying lizards reached Pantar by first dispersing to the Outer Banda Arc island of Sumba, then east to Timor, north to Wetar, west to Alor, and finally west to Pantar. The levels of genetic and morphological divergence among the allopatric populations of Lesser Sundas flying lizards suggest that species diversity is currently underestimated, and examination of specimens along with the genetic data presented here will be needed to inform a taxonomic revision of the complex.

INTRODUCTION

The field of biogeography aims to understand the distribution of life both currently and historically (Lomolino *et al.*, 2006). While biogeography typically pertains to the spatial patterns of defined taxonomic groups, communities, or ecosystems, recently many researchers have recently become interested in the patterns of geographic variation in genetic lineages within a species or group of closely related species. This field of research, known as phylogeography, has given biogeographers a new window into historical processes that may not be available in the fossil record or morphological traits. Phylogeography and historical biogeography can also shed light on the geological history of the region being examined by preserving a record of vicariance and dispersal events, population size changes, and migration between lineages or species (Avise 2000).

The Lesser Sunda Islands occur as two parallel, oceanic island arcs that extend west to east from Bali toward New Guinea in southeastern Indonesia. There are approximately 566 islands in the Lesser Sundas, many of which are small and 42 of which are inhabited by humans (Goltenboth *et al.* 2006). These islands are the result of complex tectonic processes associated with the collision of the Australian continental plate margin and the Banda Sea Plate. Subduction zones often result in double island-arc systems, with an outer arc forming close to the point of subduction and resulting from accretionary build-up of continental plate margin, and an inner arc forming more distant from the point of subduction via volcanic extrusion of superheated mantle and associated descending plate. The Lesser Sundas Archipelago exemplify this process, with the corresponding resultant island arcs termed the Inner and Outer Banda Arcs. The Inner Banda Arc is the more northern string of islands, all of which are volcanic. Its major islands include (from west to east) Lombok, Sumbawa, Komodo, Flores, Lembata, Pantar, Alor, and Wetar, with additional small islands extending further to the east. The Outer Banda Arc is the more southern string of non-volcanic islands, each of which is composed of uplifted continental crust that is correspondingly overlain with marine deposits. The major islands of the Outer Banda Arc include (from west to east) Sumba, Sabu, Rote, and Timor, with additional islands extending further east. Both the Inner and Outer Band Arcs extend eastward beyond the margins of the Lesser Sundas Archipelago, curving northward and ultimately westward to include Damar, Romang, and Banda (Inner Arc), as well as Tanimbar, Kei, Seram, and Buru (Outer Arc). The oldest islands are those of the western Inner Banda Arc such as Lombok and Sumbawa, which are estimated to have become land-positive approximately 10-12 million years ago (Hall 2009). The next oldest islands are those of the central and eastern Inner Banda Arc such as Flores, Alor, and Wetar which are thought to have become land-positive ~3-7 Million years ago. The Outer Banda Arc contains the youngest islands, and are thought to have become land positive approximately 1-3 million years ago (Hall 2009; Audley-Charles 2011).

The Lesser Sunda Islands have been colonized by reptilian fauna originating from the Greater Sunda Shelf Islands (namely Java and Bali) in the west, Australia to the south, New Guinea to the east, and Sulawesi to the north. Drawing on biotic communities from such a diverse set of source regions, representing three biogeographic realms, has created complex biogeographical patterns within the archipelago. These patterns have intrigued and perplexed biologists, yet most of the colonization scenarios that have been proposed for Lesser Sundas taxa have been highly speculative rather than the result of quantitaive analysis. The first step to understanding how reptile fauna, and likely many other terrestrial vertebrates, colonized the archipelago is to identify a monopyletic reptile lineage that has colonized most or all of the major

Lesser Sunda islands. Here, I examine the biogeographical history of a fascinating group of lizards, flying lizards of the Genus *Draco*, that has colonized nearly every major island within the Lesser Sunda Archipelago providing a unique opportunity to examine the timing and sequence of island colonization in this poorly understood region.

Biology and Taxonomy of Lesser Sundas Draco: The flying lizards of the genus *Draco* are widespread ranging from southwest India to Southeast Asia, the Greater Sunda Shelf, the Philippines, as well as most of the major islands of Wallacea. The genus contains at least 45 species, though this number will likely increase. Flying lizards get their name from their specialized gliding locomotor behavior that they use to move move between trees, primarily in dipterocarp-dominated forests. They are diurnally active and typically found in forest-edge or disturbed habitats, and are commonly observed on exposed trunks and branches of trees with low leaf density. All species of *Draco* are arboreal and contain wing-like patagial membranes (Fig. 1), which are supported by elongated thoracic ribs, as well as expandable throat lappets that are supported by the hyoid apparatus (McGuire & Dudley 2011). The patagia, along with the dewlap in males, also serve as important display structures for mating and courtship. Patagial and dewlap coloration are species specific in most *Draco* species, and many species are also sexually dichromatic. Their diet consists primarily of ants and arboreal termites, but they will also eat other small invertebrates (Inger, 1983).

Flying lizards are one of the most widely distributed groups of lizards in the Lesser Sunda Islands and can be found on every major island except Sabu (Fig. 2). The taxonomy of the Lesser Sundas' flying lizards has fluctuated over the last two centuries (see Box 1), but there are currently two species recognized, each of which is present on multiple islands, but with distributions that are non-overlapping. The Lesser Sundas populations were once thought to be part of a wide ranging *Draco volans* (now restricted to Java and Bali), were subsequently described as distinct subspecies *D. v. timoriensis* and *D. v. boschmai*, and were subsequently elevated to species status by McGuire and Heang (2001). McGuire and Heang (2001) elevated *D. boschmai* and *D. timoriensis* because they are clearly diagnosable from *D. volans, D. sumatranus,* and other flying lizards on the basis of morphology (presence of an enlarged series of keeled paravertebral scales and different color patterns). However, they also noted that populations on different Lesser Sunda islands are quite distinct in terms of coloration, and in need of further taxonomic revision. McGuire and Heang (2001) found *D. volans* to be the sister taxon of the Lesser Sundas assemblage. *Draco boschmai* occupies the Inner Banda Arc islands of Lombok, Sumbawa, Moyo, Komodo, Rinca, Flores, Adonara, and Lembata, as well as the Outer Banda Arc island of Sumba (Fig. 2). The allopatrically distributed *Draco timoriensis* occurs on the Inner Banda Arc islands of Alor and Wetar, as well as the Outer Banda Arc islands of Timor, Semau, and Rote (Fig. 2). *Draco timorensis* is distinguished from *D. boschmai* based on the presence of a row of enlarged, keeled scales along each side of the vertebral line, and a large spine-like tubercle in the nuchal region. Both *D. boschmai* and *D. timorensis* exhibit substantial intraspecific variation in color and scale characters between populations (Mertens 1930; McGuire and Heang, 2001), and within *boschmai* there are differences in color as well between populations on Flores + Sumba and Lombok + Sumbawa (Musters 1983; McGuire and Heang, 2001)

The only molecular study to include multiple *Draco* samples from the Lesser Sundas was that of McGuire & Heang (2001), who analyzed 1,120 bp of the mitochondrial *ND2* gene and found that flying lizards from the Lesser Sundas form a monophyletic assemblage that is sister to

Boschma's Flying Dragon (*Draco boschmai*; Hennig, 1936)
Geographic Range: Lombok, Sumbawa, Komodo, Rinca, Flores, Adonara, Sumba
Type Locality: Maumere, Flores; Holotype=ZMA11025
-*Draco volans timorensis* (Kuhl 1820)
-*Draco timoriensis* (De Rooij 1915)
-*Draco volans reticulatus* (Mertens 1930, Forcart 1953, Darevsky 1964, Auffenberg 1980)
-*Draco volans boschmai* (Hennig 1936, Wermuth 1967, Musters 1983)
-*Draco boschmai* (McGuire & Heang 2001)

Draco boschmai is also sexually dichromatic and exhibit substantial inter-island variation in color pattern. In some populations both the dorsal and ventral patagium are entirely suffused with melanic pigments. Females from these populations have patagia characterized by large pale spots on a dark background and lack melanic pigments on the ventral surface of the patagium. In other populations neither the males nor females have melanic pigments on the dorsal or ventral surfaces of the patagium.

Timor Flying Dragon (*Draco timoriensis*; Kuhl, 1820)
Geographic Range: Timor, Rote, Semau, Alor, Wetar
No Type Specimen
-*Draco timoriensis* (Kuhl 1820, Gray 1831, Dumeril & Bibron 1837, McGuire & Heang 2001)
-*Draco timorensis* (Gray 1845, Boulenger 1885, Lidth de Jeude 1895, Wandolleck 1900, Werner 1910, Barbour 1912, De Rooy 1915, Dunn 1927, Manacas 1956)
-*Draco volans timoriensis* (Hennig 1936, Wermuth 1967, Musters 1983)

Draco timoriensis is sexually dichromatic with males containing a dorsal patagium coloration of bright yellow with diffuse gray lateral bands, and the ventral patagium lacking melanic pigments. Females from Timor and Rote exhibit a dorsal patagium coloration of black or dark brown with light horizontal striations, while the ventral patagium of females is saturated with melanic pigments. However, *timoriensis* females from Alor and Wetar have ventral patagium that either lack melanic pigments or have a few scattered dark spots.

Box 1. Taxanomic history and characteristics of *Draco boschmai* and *D. timoriensis*.

a *D. volans* + *D. sumatranus* clade. This study also revealed that *D. timoriensis* is nested within *D. boschmai*. The authors' note that both species are composed of several diagnosable, allopatric lineages and further taxonomic changes are necessary after a thorough evaluation is completed. Additionally, a few major islands such as Pantar contain flying lizards that are not currently assigned to either species and thus clearly warrant examination.

Transcriptome-based Exon Capture: In recent decades historical biogeographical patterns have been revealed using molecular data in a phylogenetic framework. Mitochondrial DNA sequence data has been the primary marker of choice for this type of study due to its haploid nature, small effective population size compared to nuclear DNA, and a high information content relative to nuclear loci of comparable size. However, mtDNA is maternally inherited and may not account for male biased gene flow, and there is an increasing body of literature that has highlighted the prevalence of mito-nuclear discordance whereby mtDNA patterns do not reflect true evolutionary relationships (McGuire *et al.*, 2007; Toews & Brelsford, 2012). It has become widely accepted that the best way to reconstruct a molecular phylogeny, or to make population

demographic inferences, is to analyze genetic data representing many unlinked nuclear loci. This has been accomplished recently by the development of genomic methods that allow for screening of hundreds or thousands of loci, combined with improved analytical software for analyzing such massively multi-locus data. For example, multi-species coalescent phylogenetic methods can resolve species phylogenies at both shallow and deep divergence scales. However, obtaining informative sequence data representing orthologous loci across population level or species level divergences remains a major hurdle for genome-scale studies of organisms for which full-genome sequence data are unavailable.

Exon capture experiments hybridize genomic libraries to short single-stranded oligonucleotide probes that are complimentary to targeted regions of the genome. The design of exon-capture probes relies on existing genomic sequence data from the taxa of interest (or from closely related taxa). Among the more accessible genomic resources that can be developed for non-model organisms are *de novo* assembled transcriptomes. A transcriptome represents sequences for the full suite of mRNA molecules that were being expressed at the time of collection in one or more tissue samples. By using transcriptome sequences to develop tiled probes, the pooled capture of barcoded libraries is a cost-effective method for capturing many independent orthologous loci across multiple individuals (Bi *et al.,* 2012; Bi *et al.,* 2013).

In this study I examine the biogeographical history of the flying lizards (the *Draco boschmai/timoriensis* clade) of the Lesser Sunda Islands using DNA sequence data. The sequence data was analyzed for phylogenetic relationships, population structuring, and migration between divergent lineages. The major goals of these analyses are to estimate the colonization history of the Lesser Sundas Islands colonization by flying lizards to test biogeographical hypotheses (outlined in Chapter 1) regarding the orientation and age of the islands themselves. Because flying lizards dispersed into the archipelago from the Greater Sunda Islands in the west we expect a pectinate phylogeny where the basal branch would contain all samples from Lombok, followed by a Sumbawa Branch, and each subsequent branch would be an island or islands that are immediately east of the previous lineage. However, if island age has played a large part in the sequence of island colonization we might see a non-pectinate topology whereby the youngest islands such as Sumba and Timor are nested within the island immediately north of them.

MATERIALS & METHODS

Summary: As a first-pass approach mitochondrial DNA was sequenced from all available samples of Flying Lizards from the Lesser Sundas to identify major clades within and among islands, as well as to obtain a rough estimate of the age of the Lesser Sundas clade and subclades. These data were then used to choose samples for transcriptome sequencing and barcoded genomic library development. After aligning transcripts and identifying orthologous loci that meet a number of criteria for length and information content probes were developed for use in an in-solution exon-capture experiment. Pooled barcoded sample libraries were then hybridized to the probes to capture desired genomic fragments for enrichment and sequencing.

Sample Collection: Flying lizard specimens were collected from the field including *Draco volans* from Java and Bali, *D. boschmai* from Lombok, Sumbawa, Flores, Lembata, and Sumba, and *D. timoriensis* from Timor, Rote, Pantar, Alor, and Wetar islands. Most of these samples and tissues

were collected on four separate expeditions to the Lesser Sunda Islands undertaken in the years 2010-2013. Liver tissue was dissected from euthanized lizards and either stored in RNALater, or flash frozen in liquid nitrogen. Specimens were given field tags in the catalogs of Jimmy A. McGuire (JAM#), Sean B. Reilly (SBR#), or Alexander L. Stubbs (ALS#). The tissues were divided in half and deposited in both the Museum of Vertebrate Zoology at UC Berkeley (and subsequently will receive MVZ catalog numbers) and the Museum Zoologicum Bogoriense (which are given MZB catalog numbers). The formalin-fixed specimens were divided equally between the MVZ and MZB collections.

MtDNA Data Collection: DNA was extracted from liver tissue using standard salt extraction techniques or by using the DNeasy kit. DNA extractions were then diluted to concentrations suitable for PCR-amplification (~20-60 ng/uL). We sequenced the *ND2* gene for 372 flying lizards, 363 of which are from the Lesser Sunda Islands (Table 1). All sequence data was collected using standard PCR-amplification using the primers METf.1: AAGCAGTTGGGCCCATRCC) and ALAr.2m: AAAGTGTCTGAGTTGCA- TTCRG (Macey *et al.*, 1997). PCR reactions contained 18.3 µL water, 2.5 µL of 10X buffer, 1.5 µL magnesium chloride, 1.5 µL dNTPs (2 µM), 0.6 µL of each primer, 0.2 µL Taq polymerase, and 1µL genomic DNA at concentration of 20-40 ng/µL. PCR products were cleaned by using ExoSAP-IT (USB, Cleveland, OH) before being labeled with fluorescent-dye nucleotides through cycle sequencing reactions for both forward and reverse primers. Ethanol precipitation was used to clean cycle sequencing products, which were sequenced on an ABI 3730 sequencer (Applied Biosystems, Foster City, CA, USA). Raw sequence reads were combined in Codoncode Aligner 3.5.2 (CodonCode Corporation, Dedham, MA, USA) and aligned with MUSCLE (Edgar, 2004).

MtDNA Data Analyses: The sequence alignment of 1,092 bp was imported into JMODELTEST V2.1.4 (Darriba *et al.*, 2012) to determine the best-fit model of sequence evolution (HKY+I+G) as supported by the program BEAST V1.8 (Drummond & Rambaut, 2007). A BEAST run was conducted using the uncorrelated relaxed clock model and a coalescent constant size tree prior with a uniform distribution. A rule of thumb 1% rate of sequence evolution per million years (which corresponds to 2% divergence per million years for any two lineage comparison) was applied to obtain a rough approximation of timing of entry into the archipelago, as well as the ages of island-specific lineages. A preliminary run was carried out to determine the appropriate number of generations that would result in ESS values for each parameter greater $\geq$ 200, as viewed in TRACER V1.6 (Rambaut & Drummond, 2009). Once the appropriate run length was determined, two separate runs of 100 million generations were carried out, sampling every 10,000 generations for a total of 10,000 saved samples per run. A burnin of 10% was removed from each of the two runs and the remaining 18,000 trees were combined to create a 50% majority rule consensus tree. The tree was rooted using the outgroup *Draco beccarii* from Sulawesi (LSUMZ81223). Nodal support was assessed using posterior probability values. A Maximum Likelihood analysis was also undertaken using the program RAXML (Stamatakis, 2014). The default GTR+I+G model of sequence evolution was applied, and node support was assessed with 100 bootstrap replicates.

Transcriptome Sequencing: Total RNA was extracted from four flying lizard samples (JAM 11504, *D. beccarii*, Sulawesi; JAM 12091, *D. boschmai*, Sumbawa; JAM 12477, *D. boschmai*, Flores; JAM 12741, *D. boschmai*, Lembata) using the RNEasy Protect Mini Kit (Qiagen) and protocol. Samples were evaluated using a BioAnalyzer 2100 RNA Pico chip (Agilent), with RIN

scores greater than 8 except for one sample. Sequencing libraries were prepared using half reactions of the TruSeq RNA Library Preparation Kit V2 (Illumina), beginning with Poly-A selection for samples with high RIN scores (> 8.0) and Ribo-Zero Magnetic Gold (Epicentre) ribosomal RNA removal for samples with low RIN scores (< 8.0). Libraries were pooled and sequenced on an Illumina HiSeq2500 with 100 bp paired-end reads. Transcriptomic data were cleaned following Singhal (2013). Cleaned data were assembled using TRINITY (Grabherr *et al*., 2011) and annotated with *Anolis carolinensis* (Ensembl) as a reference genome using reciprocal BLASTX (Altschul *et al*., 1997) and EXONERATE (Slater & Birney, 2005). Annotated transcripts were compared from the four individuals to search for orthologs via BLAST (Altschul *et al*., 1990). Mitochondrial loci were removed from the transcripts. Only transcripts with a GC content between 40%-70% were kept because extreme GC content causes reduced capture efficiency for the targets (Bi *et al*., 2012). All the bioinformatics pipelines for transcriptome data processing and annotation are available at https://github.com/CGRL-QB3-UCBerkeley/DenovoTranscriptome.

Marker Develoment: Annotated and filtered contigs from all transcripts were aligned to identify shared markers. Markers under 300 bp were discarded and markers greater than 1,000 bp were cut down to a maximum length of 1,000 bp. The remaining genes were examined for repetitive elements, short repeats, and low complexity regions, which are problematic for probe design and capture. The four sets of transcripts were screened using the REPEATMASKER Web Server (Smit *et al*., 2015), which resulted in the masking of repetitive elements or low complexity regions. To be conservative, if any of the four transcripts for a gene contained masked sites, that gene was removed from the final marker set. The *D. boschmai* transcripts were 0.4-0.5% divergent on average. The resulting 3,727 markers from the three *D. boschmai* transcripts were compared to identify the variability of each marker as determined by the number of polymorphic sites. The markers were sorted according to the number of variable sites per locus. All invariable loci and loci with only a single variable site were discarded, as well as the top 5% of the most variable loci. A total of 1,200 of the resulting 2,578 candidate loci were integrated into an Agilent MicroArray chip design. The estimated target size of the combined loci was approximately 1,021,000 bp. Pipelines for marker development are available at https://github.com/CGRL-QB3-UCBerkeley/MarkerDevelopment Pylogenomics.

Sample Library Preperation: A total of 96 *Draco* samples were initially chosen for library preparation. These samples were picked by examining the mitochondrial tree to maximize the genetic diversity from all of the islands (Table 1). The DNA was quantified by Qubit DNA BR assay (Life Technologies) and 1500 ng total DNA was diluted in 100 µl of ultrapure H_2O. A Bioruptor UCD-200 (Diagenode) was used to sonicate the samples on a low setting for 15 minutes, using 30s on/30s off cycling. For each sonicated sample, 5 µl of product was run on a 1% gel at 100V for 45 min to ensure fragments were appropriately sized (200–500 bp). Individual genomic libraries were prepared following the protocol of Meyer and Kircher (2010), with slight modifications, including the use of at least 1,500 ng total DNA for library preparation (rather than 500 ng) to remedy the possibility of decreased library diversity resulting from a large genome size. We used 7-9 cycles of post-adapter ligation PCR to enrich the libraries and incorporate a 7bp P7 index. The resulting 50 µl of amplified library product had an average concentration of 30 ng/µl measured by a Nanodrop 1000 spectrophotometer (Thermo Scientific), producing an average yield of 1,500 ng total library DNA.

Agilent Microarray: The 1,200 loci identified from the transcripts were combined with 5 positive control loci obtained by Sanger sequencing (*CMOS, BDNF, PNPLA8, RAB5,* and *ADIPOR2*) for a total of 1,205 loci. These loci were included from all four transcripts so that there were four copies of each gene present to ensure unbiased hybridization across Lesser Sundas flying lizard lineages as well as outgoups. The 96 PCR-enriched libraries were normalized to 15 ng/uL and 18 uL of each library were added to a single library pool. The library pool was placed in a vacuum centrifuge and dried down to a volume of 138 uL. Blocking oligos, chicken COT-1 DNA, Agilent blocking reagent, and Agilent hybridization buffer were added to the library pool to a volume of 520 uL. This mixture was heat shocked at 95°C for 3 minutes followed by 37°C for 30 min, and then 490 uL of this mixture was added to the Agilent slide. This chip was hybridized at 65°C for 65 hours at 12 rpm in the array oven. After hybridization, the captured DNA was eluted following the protocol from Hodges *et al.* (2009), and dried down to a volume of 235 uL. Six enrichment PCR reactions were performed, each with 37 uL elute, 10 uL Phusion HF Buffer, 0.4 uL dNTPs (25 mM), 1 uL IS5 primer, 1 uL IS6 primer, and 0.5 uL Phusion Polymerase. These enrichment PCRs were then pooled and run on a BioAnalyzer to check the concentration and fragment size distribution. The pooled enriched elute and the original pooled libraries were run on a qPCR using 3 positive (*CMOS, BDNF,* PNPLA8) and 3 negative (*Glor2, KIAA,* RAB7A) control primers to look for enrichment of the positive control loci and depletion of the negative control loci. The enriched library mixture was then sequenced on a single lane of an Illumina HiSeq 2500 sequencer at UC Berkeley.

Marker Refinement: The data returned from the Illumina sequencer was of high quality. However, the average coverage was 8X with some loci having very low coverage. Additionally, 491 of the loci were invariable within the Lesser Sundas assemblage suggesting that some of the variable sites detected between transcripts were sequencing errors. There were 1,634 contigs (exons) representing 709 genes that were variable within the Lesser Sundas samples, and these were chosen for a follow-up experiment. These contigs contained both exonic and intronic sequences. We designed 120 bp probes with 3X tiling across two of the individuals sequenced with the Agilent chip (SBR199, *D. timoriensis*, Timor; JAM7032, *D. walkeri*, Sulawesi) for a total of 44,964 unique probes. These probes were manufactured by MYBaits as part of their in-solution exon-capture kit.

MYBAits In-Solution Exon Capture: Two libraries that failed in the Agilent experiment were removed, and six new samples from East Timor (Timor-Leste) were included for a total of 100 samples to be screened with the MYBait kit (Table 1). These 100 samples represent a widespread sampling scheme that contain all sampled island populations as well as different populations within larger islands (Fig. 3). Libraries were pooled in equal amounts with each pool containing 6 individuals. Pools were determined by grouping closely related samples to minimize competition for probe binding. MYbaits capture reactions were performed following the v2.3.1 manual with some modifications. For each capture reaction library master mix, the pooled libraries were vacuum dried at 45°C for 60 min and re-suspended in ultrapure H_2O, then combined with 1.66 µl each of salmon sperm COT-1, human COT-1, chicken COT-1, and xGEN blocking oligos. The combined volume of water for DNA resuspension and volume of blocking oligonucleotides totaled 6.5 µl. The hybridization reaction proceeded at 65°C for 24-28 hours. Individual capture reactions were purified using streptavidin-coated magnetic beads and post-

capture products were PCR-amplified using four independent reactions of 14 cycles each. These reactions were resuspended in 11 µl of ultrapure H_2O, and had an average concentration of 4-7 ng/µl, as measured by Qubit. Purified PCR products from the same capture were combined and quantified using a BioAnalyzer 2100 DNA-1000 chip. The combined post-capture amplified products ranged from 3-9 ng/µl, and the average product size was ~370 bp. The combined post-capture libraries were sequenced on one lane of an Illumina HiSeq2500 with 100 bp paired-end reads.

Data Pipeline: Raw sequence data were cleaned following Singhal (2013) and Bi *et al*. (2012). Raw fastq reads were filtered using TRIMMOMATIC (Bolger *et al*., 2014) and CUTADAPT (Martin, 2011) to trim adapter contaminations and low quality reads. BOWTIE2 (Langmead & Salzberg, 2012) was used to align the data to *Escherichia coli* (NCBI: 48994873) to remove potential bacterial contamination. Exact duplicates were eliminated as well as low complexity sequences using a custom script. Overlapping paired reads were also merged using FLASH (Magoč & Salzberg, 2011) and COPE (Liu *et al*., 2012) to avoid inflated coverage estimates in the overlapping region. The resulting cleaned reads of each individual specimen were *de novo* assembled using ABYSS (Simpson *et al*., 2009). Individual raw assemblies were generated using a wide range of k-mers (21, 31, 41, 51, 61 and 71) and we then used CD-HIT-EST (Li & Godzik, 2006), BLAT (Kent, 2002), and CAP3 (Huang & Madan, 1999) to cluster and merge all raw assemblies into final, less-redundant assemblies. BLASTN (evalue cutoff = 1e-10, similarity cutoff = 70) was used to compare the target sequences with the raw assemblies of each individual in order to identify the set of contigs that were associated with targets (in-target assemblies). A self-BLASTN (evalue cutoff =1e-20) was run to compare the assemblies against themselves to mask any regions from a contig that matched other regions from other contigs. For each matched contig EXONERATE (http://www.genome.iastate.edu/bioinfo/resources/manuals /exonerate/exonerate.man.html) was used to define protein-coding and flanking regions. Flanking sequences were retained if they were within 250 bp of a coding region. Finally, all discrete contigs that were derived from the same reference transcript were joined together with Ns based on their relative BLAST hit positions to the reference. Most of the final in-target assemblies contain multiple contigs, and each includes both coding regions and flanking sequences if captured.

Cleaned sequence data were then aligned to the resulting individual-specific in-target assemblies using NOVOALIGN (Li & Durbin, 2009) and only reads that mapped uniquely to the reference were retained. The programs Picard (http://broadinstitute.github.io/picard/) and GATK (McKenna *et al*., 2010) were used to perform re-alignment. Finally, the program SAMTOOLS/BCFTOOLS (Li *et al*., 2009) was used to generate individual consensus sequences by calling genotypes and incorporating ambiguous sites in the in-target assemblies. A consensus base was only kept when the site depth was above 10X. Sites were masked within a 5 bp window around an indel. Sites were also filtered out where more than two alleles were called. Then FASTQ were converted to FASTA using seqtk (https://github.com/lh3/seqtk), and putative repetitive elements and short repeats were masked using REPEATMASKER (Smit et al. 2015) with "vertebrata metazoa" as a database. Markers were removed if more than 80% of the bases were Ns. The read depth of each individual marker was calculated and loci were filtered out if the depth fell outside the 99th percentile of the statistics. Markers were also eliminated if the individual heteozygosity fell outside the 99th percentile of the statistics. The final filtered assemblies of each individual specimen were aligned using MAFFT (Katoh & Standley, 2013).

Alignments were then trimmed using TRIMAL (Capella-Gutierrez *et al*., 2009). Alignments were removed if more than 25% missing data (Ns) are present in 25% of the samples, or if the proportion of shared polymorphic sites in any locus was greater than 20%. The bioinformatic pipelines of sequence capture data processing are available at https://github.com/CGRL-QB3-UCBerkeley/denovoTargetCapturePhylogenomics.

Evaluation of Data: To evaluate capture efficiency, the average per-base sequence depth (or coverage) was calculated separately for the exon sequences and for the flanking sequences for each sample. The coverage at each base pair site for either data set was inferred using SAMTOOLS (Li *et al*., 2009). The per base pair coverage estimates for all sequences (exon or flanking) associated with each transcript (up to 709 genes) were averaged, resulting in a set of average coverage estimates across loci. The resulting output of the set of average coverage estimates was used to infer the median, upper and lower quartiles, and range of coverage estimates using samples or genes as factors. These calculations were performed and automated across samples using python scripts and the output was visualized in R (Portik *et al.*, 2016). Differences in the levels of coverage were examined using pooling size as a factor.

The resulting alignments of exon-only data and flanking region data were evaluated for taxon number, sequence length, percentage of missing data, and proportion of informative sites. These results were visualized in R, and the relationship between the number of informative sites and alignment length was investigated using a simple linear regression. The relationship between phylogenetic distance and missing data was also investigated using a simple linear regression. The percentage of missing data was calculated from the final concatenated alignment of exon-only loci that passed multiple post-processing filters, including a minimum length of 100 bp, no more than 80% missing data per sequence in alignments, and no more than 25% total missing data across an alignment. These filters were enforced using a custom alignment refinement python script for all alignments. All custom python scripts for sequence capture performance evaluation are available on github (Portik *et al.*, 2016).

Principle Components of Molecular Co-Variance: Principle components analysis of genetic co-variance was conducted using the *adegenet* package for R (Jombart, 2008). The results of the top three components are plotted against one another to visualize spatial clustering of individual samples.

Phylogenomic Analyses: Phylogenomic analyses were undertaken using three approaches: a concatenated Maximum Likelihood analysis, a summary multispecies coalescent analysis, and a full multispecies coalescent analysis. These analyses are described below. The concatenated maximum likelihood analysis using an alignment of all sequence data was analyzed with RAXML (Stamatakis, 2014) under the GTR+I+G model of sequence evolution. Nodal support was assessed with 100 bootstrap replicates.

The summary multispecies coalescent analysis was undertaken using ASTRAL-II (Mirarab *et al.*, 2014). For this analysis, individual gene trees were first estimated for each of the 709 loci using RAXML and the GTR+I+G model of evolution. The gene trees were then used as input files for ASTRAL-II. For this analysis each individual was treated as a "species" because it was unclear where the species boundaries lie within the system and a direct comparison with the topology of the RAXML tree was desired. To look at the information content of our loci, we randomly sampled gene trees in bins of 5, 10, 20, 50, 100, 200, and 500 gene trees (one per

randomly selected locus) to re-construct summary multispecies coalescent trees in ASTRAL-II. This procedure was repeated 10 times for each bin and the average Robinson-Foulds distance was calculated in comparison with the full 709-locus supertree. The Robinson-Foulds calculation is a measure of the difference in topology between two trees.

A full species tree approach was also applied to the data using the program SNAPP as implemented in BEAST2 (Bouckaert *et al.*, 2014). This program assumes unlinked SNP data, so the dataset consisted of one randomly chosen informative SNP per locus. The program also requires designation of samples to a species *a priori*, so we treated each clade from our mtDNA phylogeny as a lineage. The program was run for 1,000,000 generations with the first 100,000 generations removed as burn-in. The resulting trees are superimposed on top of one another as a tree "cloud" to visualize the uncertainty of the topology, and a consensus tree is also overlaid.

Population Structure: The program STRUCTURE (Pritchard *et al.*, 2000) was used to assess population differentiation. For this analysis, one informative SNP was randomly chosen per locus from within the Lesser Sundas assemblage (89 individuals of *D. boschmai* and *D. timoriensis*). Each STRUCTURE run requires an assumption of the number of populations (K values) and these are then evaluated post-hoc to determine the best fit number of populations given the data. As a first pass, I performed analyses with K=1 to K=20, with 10 replicates per K-value. Each run had a burn-in of 50,000 generations followed by 50,000 generations retained for analysis. The results were then imported into STRUCTURE HARVESTER (Earl & VonHoldt, 2012) to determine the most likely number of populations as determined by the Delta K method. Given the results of the first pass, a second batch of analyses was performed with a 100,000 generation burn-in and a 100,000 generation run for K=2 up to K=13 to examine the sequential division of the assemblage for each assumed number of populations.

Inter-Island Demographics: Demographic analysis utilized the flanking sequence from each locus because these regions are presumably not under selection, though they are linked to exonic regions that are likely under selection. These data were analyzed with the program G-PHOCS (Gronau *et al.*, 2011), which is an isolation-with-migration program capable of making inferences from genomic sequence data representing unlinked neutrally evolving loci. This program estimates unscaled effective population sizes of the extant populations as well as their ancestor, a population divergence times, and migration rates between extant populations. This analysis was run to estimate relative rates of migration between divergent lineages of *D. boschmai* that are parapatrically distributed on the islands of Sumbawa and Flores. These analyses each compared two populations: West Sumbawa vs East Sumbawa, and West Flores vs East Flores + Lembata. An initial run of 500,000 generations was used to assess convergence of the parameters, followed by a separate run of 1,000,000 generations. After removing 100,000 generations from each of the two runs as burn-in, the remaining 1,300,000 generations were combined in Tracer (Rambaut & Drummond, 2009) to assess the posterior distribution of the demographic parameters.

A mutation rate of 2.2×10^{-9} mutations/site/year was used to convert unscaled parameter estimates into real world values (Kumar & Submaranian, 2002). A mutation rate was also calculated by using an unpublished time-calibrated *ND2* phylogeny for *Draco* to put a rough date on a genus-wide phylogeny utilizing the same genomic loci used in this study, and back calculating the rate (4×10^{-10}) given the level of divergences between clades (McGuire pers. com.). All values for Theta and Tau given by the program G-PhoCS are scaled by 10^{-4}.

Demographic parameter estimates were converted to estimates of effective population sizes (individuals) by dividing the scaled Theta estimate by the mutation rate, then dividing that value by 4 (because diploid organisms will have an effective population size of 4 at any given locus). The population divergence time in years was calculated by dividing the scaled Tau estimate by the mutation rate. Migration rate estimates were converted to Migrants per Generation by multiplying the migration estimate by the converted effective population size estimate for the population receiving migrants, then dividing that value by the number of generations that have passed (in years) since divergence. A generation time of 1 year was used to convert migration rates.

RESULTS

MtDNA Phylogeny: Both the Maximum Likelihood (Fig. 4) and Bayesian (Fig. 5) phylogenetic analyses of the mtDNA data recovered the same well-supported clades that are either allopatrically or parapatrically distributed (Fig. 6). In both analyses, the Lesser Sundas assemblage was found to be monophyletic and sister to a clade containing *D. sumatranus* (from the Malay Peninsula, Sumatra, and Borneo) and *D. volans* (from Java and Bali). The split between the Lesser Sundas and the *D. sumatranus* + *D. volans* clade was estimated to have occurred approximately 11.4 Ma (95% CI = 9.5-13.6 Ma), and the most recent common ancestor of the Lesser Sundas flying lizards clade was dated at approximately 8.7 Ma (95% CI = 7.1-10.1 Ma). In both phylogenies, the basal lineage was a clade containing Lombok and Western Sumbawa, each of which are monophyletic and appear to have split from one another approximately 6.7 Ma (95% CI = 5.0-8.4 Ma). The next major split is between an Eastern Flores + Lembata clade and a clade containing all remaining island populations. The timing of this divergence was estimated at approximately 5.4 Ma (95% CI = 4.4-6.3 Ma). This was followed by a split between Eastern Sumbawa and all remaining populations ~ 4.4 Ma. In the Bayesian phylogeny, the next split was between Western Flores and all remaining populations ~3.7 MA (Fig. 5). The Maximum Likelihood phylogeny does not resolve the relationships between clades well beyond this point (Fig. 4). The relationships between the populations on Wetar, Alor + Pantar, Rote, West Timor, East Timor1, East Timor2, and Sumba are not well supported in either the Bayesian or Maximum Likelihood phylogenies, though the support for the monophyly of each of these clades is high.

The geographical position of the mitochondrial split within Sumbawa is at the narrow isthmus that connects the large western and eastern halves of the island (Fig. 6). The Split within Flores occurs in the center of the island just east of the town of Bajawa (Fig. 6). The boundary between the West Timor clade and the East Timor2 clade is not well-defined due to lack of sampling, and the boundary between the East Timor1 and East Timor2 clades is situated in western Timor-Leste (Fig. 6).

Exon-Capture Data Characteristics: The total alignment of both the targeted and flanking regions from the 709 genes included 967,361 bp. One library failed to sequence so the final dataset is for 99 flying lizards, including 10 outgroup samples and 89 ingroup samples. The 10 outgroup samples include eight closely related members of the *D. volans* group (five *Draco volans* from Bali, one *D. volans* from Java, one *D. sumatranus* from Sumatra, and one *D. modigliani* from Enggano Island), as well as two more distantly related members of the *D. lineatus* group (one *D. walkeri* and one *D. beccarii*, both from Sulawesi). The ingroup sample

was composed of 35 *D. timoriensis* from 5 islands, and 54 *D. boschmai* from 5 islands. The average coverage for the targeted regions was >150X, while the flanking regions had approximately 100X coverage on average (Fig. 7). However, the average coverage for each individual library was highly variable for both the targeted (Fig. 8) and flanking (Fig. 9) regions.

The average number of samples per alignment was 98 out of the 99 total samples (Fig. 10a, Fig. 11a). The Number of informative sites has a relatively linear relationship with the length of the loci, with very few outlier loci (Fig. 10b), and there is on average 5% informative sites per alignment (Fig 11c). There is no clear relationship between the alignment length and the percentage of gaps in each alignment (Fig. 10c). The final length of the contigs ranged from 100 bp up to ~5,500 bp (Fig 11b), and the percent of missing data was no higher than 25% after the additional filtering step (Fig. 11d).

Principle Components Analysis: The majority of the genetic co-variance was explained by the top three components (Fig. 12d). When PC1 is plotted against PC2 the samples fall into three main clusters that roughly correspond to *Draco volans* (upper right)*, D. boschmai* (lower right), and *D. timoriensis* (lower left), with the one exception that samples from Sumba (currently ascribed to *D. boschmai*) group with *D. timoriensis* samples (Fig. 12a). When PC3 is compared to either PC1 or PC2, we see fine-scale differentiation of the *boschmai* lineages (Fig. 12b-c). None of the components appear to support fine-scale clustering of island-specific populations within *D. timoriensis*.

Phylogenomic Trees: The Maximum Likelihood phylogeny produced by RAxML recovered many of the same clades as the mitochondrial phylogeny, but with some differences and with generally higher support for clade relationships (Fig. 13). In this phylogeny, the Lesser Sundas assemblage is recovered as monophyletic and sister to *D. volans* from Java and Bali (non-parametric bootstrap proportion (BP) = 100). Just as in the mtDNA phylogenies, the basal split is between a Lombok + West Sumbawa clade (BP=100) and all other populations (BP=100). The next major split is between Eastern Sumbawa and all other populations (BP=100), followed by a split between a Flores + Lembata clade and all other populations (BP=100), with Lembata nested within Flores. The remaining clade includes Sumba as sister to *D. timoriensis*. Within *D. timoriensis,* the major split is between an Inner Banda Arc assemblage (Wetar, Alor, Pantar) and an Outer Banda Arc assemblage (Timor, Rote), with Pantar nested within an Alor clade, and Rote nested within a Timor clade.

The summary multispecies coalescent phylogeny produced by ASTRAL-II converged on nearly the same topology as the Maximum Likelihood tree with respect to the relationships of the major lineages (Fig. 14). The exceptions are that the ASTRAL tree has Lombok nested within West Sumbawa, all West Flores samples form a monophyletic group, and Alor and Pantar are sister lineages. The Robinson-Foulds distances (%) suggest that the topology of the tree improves dramatically with the addition of loci up to 100 loci, and that with even 500 loci, the recovered topology is not identical to that recovered with the full 709-gene dataset (Fig. 15).

The species tree topology recovered by SNAPP differs in a number of ways from the Maximum Likelihood tree and the multispecies coalescent tree. The main differences are that East Sumbawa is sister to Flores + Lembata, and that the Wetar samples comprise the most basal lineage of *D. timoriensis* (Fig. 16).

Population Structure: The most likely number of populations as determined by the Delta K value

was 2 (Fig. 17), and these populations correspond to *D. boschmai* and *D. timoriensis* with the caveat that Sumba was grouped with *D. timoriensis* (Fig. 18). However, there was population structure recovered with up to K=9, and after that there was a maximum of nine genetic clusters returned regardless of the value assumed for K.
With K set to 3, Lombok + West Sumbawa formed a cluster, Flores + Lembata formed a cluster, and *D. timoriensis* + Sumba formed a cluster. In this analysis, the population from East Sumbawa was inferred to have experienced admixture from both the Lombok + West Sumbawa and the Flores + Lembata clusters. When K was increased to 4, East and West Sumbawa were united into a single cluster, Lombok represented a second distinct cluster, and Flores + Lembata and *D. timoriensis* + Sumba remained distinct clusters. In this analysis, the West Flores cluster was inferred to have experienced some admixture with the Sumbawa cluster. With K increased to 5, the results were similar to those with K=4 except that East Flores and Lembata were found to represent distinct clusters. When K was increased to 6, East and West Sumbawa were split into additional distinct clusters. With K set to 7, Sumba was added as a cluster distinct from *D. timoriensis*. With K set to 8 and 9, West Flores was identified as distinct from central Flores, and Alor + Pantar were separated into distinct clusters.

Demographic Analyses: Coalescent analysis of flanking (non-coding) sequence data recovered a strong signal of population expansion (~10X) for both the Sumbawa and Flores lineages when compared to their ancestral populations (Table 3, Fig. 19c-d). Divergence of the Flores lineages occurred after divergence of the Sumbawa lineages (Table 3, Fig. 19a-b), consistent with the topologies of both the mtDNA and nuclear phylogenies. The Eastern and Western Sumbawa lineages appear to have experienced limited gene flow, with relatively equal rates in both directions (Fig. 19e). Western Flores and Eastern Flores + Lembata also appear to have exchanged genes but with approximately 6X more gene flow from East to West than in the opposite direction (Fig. 19f).

DISCUSSION

Flying lizards occur on nearly all of the major islands within the Lesser Sundas archipelago, and display high levels of morphological variation between allopatric island-specific populations. This variation was used to describe two species, each endemic to the Lesser Sundas. *Draco boschmai* is described from the Inner Banda Arc islands of Lombok eastward to Adonara, also occurring on the Outer Banda Arc island of Sumba. *Draco timoriensis* is described from the Inner Banda Arc islands of Alor and Wetar as well as the Outer Banda Arc islands of Timor and Rote. Phylogenetic and population genetic analysis of both mitochondrial DNA and a 709-gene nuclear dataset have now revealed a detailed biogeographical history of this assemblage, shedding light on both the timing and sequence of island colonization. Inter-island phylogenetic breaks, range clarifications for the described species, complex dispersal routes, and discordance between mitochondrial and nuclear datasets are discussed below.

Mitochondrial Phylogeography: Our mitochondrial phylogenies examined over 350 individuals from ~80 localities spread across 11 major islands in the Lesser Sundas and Bali. The most basal clade within the Lesser Sundas contains lineages from Lombok and western Sumbawa. The fact that these regions are home to the oldest Lesser Sundas *Draco* clade is not surprising given that Lombok and Sumbawa are thought to be the oldest islands in the archipelago, and the age of this

clade is consistent with the estimated ages of the islands at around 10 Ma. Lombok and Sumbawa are also the islands closest to Java and Bali, so they would be expected to be the first islands colonized via overwater dispersal from the Greater Sunda Shelf. The next oldest lineages are those of Eastern Flores + Lembata, and then Eastern Sumbawa. The ages of these lineages are also consistent with the estimated ages of the islands, though the topology would suggest a long-distance rafting event to eastern Flores occurred before colonization of western Flores. The occurrence of deeply divergent non-sister lineages on Sumbawa and Flores suggest that each may be be composite islands formed from once-separated paleo-islands that have since merged. Given that these two islands are each composed of linearly arranged series of volcanoes, it is not hard to imagine these now-contiguous landmasses first arising as widely-separated volcanic islands that subsequently merged as the original volcanoes enlarged or perhaps as additional volcanoes appeared to fill the gaps between older ens. Movement and collision of islands may have also played a role.

In the time-calibrated phylogeny, Western Flores is strongly supported as the sister to *D. timoriensis* + Sumba populations, though this relationship is not recovered in the MrBayes tree. While there are many well-supported lineages in this more-derived clade, the relationships between them are not well supported suggesting that multiple colonization events occurred during a small timeframe. Interestingly, Timor is represented by three lineages, which roughly correspond to Western Timor, central Timor, and Eastern Timor. While this scenario could arise by multiple colonization events of Timor, it could also represent isolation by distance plus multiple dispersals from Timor to the surrounding islands. If there were phylogeographic structure on Timor (which is likely given the size of the island) then dispersals from different lineages to other islands would render Timor populations polyphyletic.

Genomic Phylogeography: While our mitochondrial dataset was useful for recovering major clades and their distributions, it was unable to accurately determine the relationships between them hindering inference of the sequence of island colonization. Analysis of nearly 1 million base pairs of sequence data from over 700 independent nuclear loci was able to resolve these relationships, and while the topology of these genomic trees was similar to the mtDNA phylogeny in many ways, there are also many differences. Similar to the mtDNA phylogeny, our genomic phylogenies, both ML and supertree, recover the Lesser Sundas as a monophyletic assemblage sister to *D. volans*, with a Lombok + West Sumbawa lineage sister to the rest of the Lesser Sundas. The supertree topology indicates that Lombok is nested within West Sumbawa suggesting that West Sumbawa may have been the first region colonized followed by westward range expansion (or overwater dispersal) to Lombok. Lombok and Sumbawa are separated by the shallow Alas Strait which becomes a land-bridge during glacial maxima which would have allowed multiple opportunities for dispersal between these two islands, even as recently as ~20,000 years ago. However, given that Lombok and West Sumbawa are monophyletic in both mtDNA and nDNA trees, gene flow and migration between these two lineages was not detected. This lack of dispersal between the islands is hard to explain, but it could be due to unsuitable habitat in the land-bridged region or competitive exclusion.

In contrast to the mtDNA phylogeny, which places Eastern Flores + Lembata as sister to the rest of the Lesser Sundas, our genomic phylogenies place Eastern Sumbawa as sister to the rest of the Lesser Sundas. This topology makes more sense given the age and placement of the islands, suggesting that a dispersal event occurred from Eastern Sumbawa to the neighboring island of Flores, rather than the mtDNA pattern which suggested a long-distance dispersal to

Eastern Flores followed by a back-colonization of Eastern Sumbawa.

Our genomic phylogenies group all Flores samples together, with Lembata nested within and sister to far Eastern Flores. The pectinate topology suggests that colonization of the western portion of the island occurred first followed by eastward range expansion all the way to Lembata. The islands off the east coast of Flores include Solor, Adonara, and Lembata, which are all separated by very shallow straits which form land bridges during glacial maxima. There is a story told by the people of Lembata that there was once a land bridge between Lembata and Adonara as recently as a few hundred years ago that fell into the sea during a large earthquake. This connectivity between Lembata and Flores helps explain their close relationship, though even in nuclear genes the Lembata samples are monophyletic suggesting that their physical isolation has lasted long enough for lineage sorting to take place.

The genomic phylogenies suggest that Sumba was then colonized from Flores, which lies directly north of Sumba. While Sumba is thought to be a relatively young island, it is highly genetically distinct suggesting a long period of isolation. Flying lizards on Sumba are sister to *D. timoriensis*, implying that a long distance over water dispersal event took place across the Banda Sea to Timor. From Timor, *Draco* then dispersed north to Wetar, then from Wetar they dispersed west to Alor and Pantar. Rote was also colonized from Timor, as the Rote clade is nested within the Timor clade.

The SNP-based species tree has a number of topological disagreements with the ML tree and supertree. Rather than East Sumbawa being sister to all islands east of Sumbawa, it is sister to the Flores +Lembata clade. Another disagreement is within *D. timoriensis* where Wetar is the most basal clade, rather than sister to Alor + Pantar as in the ML and supertree. The final disagreement involves Rote as sister to all Timor samples, rather than being nested within Timor as in the ML and supertree. While a coalescent species tree approach is a very powerful tool for analyzing large multi-locus datasets, this tree suffers from a lack of data when compared to the ML and supertree approaches because it only utilizes one polymorphic site per locus. Additionally, the species tree approach requires that samples be designated to a species/population *a priori*, which can constrain the tree topology.

Population Structure: The Delta K method suggests that there is most likely two populations within Lesser Sundas *Draco* that roughly correspond to *D. boschmai* and *D. timoriensis*, with the exception that Sumba (classified as *boschmai*) is grouped with *timoriensis*. This result is not surprising given that in all phylogenies Sumba is more closely related to *timoriensis* than it is to any *boschmai* population. However, there is a strong signal of population structuring up to K=9 populations, after which the program only returns 9 population clusters regardless of the value of K. The clusters returned are the same as the major clades recovered in the genomic phylogenies for the most part. As the number of K is increased, the clusters returned are the deepest clades within the phylogeny, followed by more derived clades. The program is essentially taking a vertical slice through the phylogeny from root to tips and sequentially recovering the next deepest splits in the tree. At K=9 most major clades from the phylogeny are returned as distinct clusters with the exception that Timor, Rote, and Wetar are grouped together.

Patterns of admixture between genetic clusters are also apparent when examining the bar plots, especially between the three major Flores clusters (West, Central, and East + Lembata). The pattern shows that central Flores is a distinct cluster with little influence from the western or eastern populations, although individuals from both the western and eastern populations have some central Flores ancestry. There is also a small influence of Eastern Sumbawa ancestry in a

few individuals from Western Sumbawa, though only one individual from Eastern Sumbawa shows any Western Sumbawa ancestry.

Demographic Analyses: Demographic analyses were performed for two pairs of parapatrically-distributed lineages, one on Sumbawa and one on Flores. The current effective population sizes of all four populations are comparable, and the ancestral effective population sizes for both islands are small in comparison to current population sizes. This pattern of a very small ancestral population size is expected in an oceanic island system, as new islands are likely colonized by one or a few individuals, which would result in a strong bottleneck due to the decreased genetic diversity of the founders when compared to the source population. The similar current effective population sizes also make sense as the area of the regions occupied by each lineage are similar in size.

The divergence times are also similar for both the Sumbawa and Flores lineage pairs, with the Sumbawa divergence being slightly older than the Flores divergence. In both the mtDNA and genomic phylogenies, the split between the Sumbawa populations is older than between the Flores populations supporting the relative values of these divergence time estimates.

The estimates for migration between both lineage pair comparisons returned non-zero values suggesting that there has been gene flow since divergence in both pairs. There is low, but symmetrical, gene flow between the West and East Sumbawa lineages while there is highly asymmetrical gene flow between West and East Flores with a higher migration rate from East to West than from West to East. This finding is supported by the STRUCTURE results that show an influence of Central and Eastern Flores populations in all West Flores individuals, but little to no influence of Western Flores ancestry within the Central or Eastern clusters.

Biogeography and Taxonomy of Lesser Sundas Draco: The most unexpected outcomes of this study include the paraphyly of Sumbawa, the colonization of the *D. timoriensis* inhabited islands by way of Sumba, and the sharp genetic boundary at the Alor Strait separating Pantar and Lembata Islands.

The sharp boundary between West and East Sumbawa lineages occurs at the narrow isthmus that connects the two halves of the island, and it is unclear what historical processes have led to this divergence. The possible causes of this phylogenetic break include: 1) multiple colonization events of the western and eastern portions of the island, 2) colonization and subsequent divergence on two separate islands that have since merged allowing for secondary contact, and 3) restricted gene flow at the narrow central Sumbawa isthmus and allo-parapatric divergence between the two lineages. The possibility of Sumbawa representing two merged paleo-islands should be given more attention by geologists. While a major explosive eruption of Tambora volcano occurred 200 years ago, this eruption apparently did not cause flying lizards to become locally extinct as is evident in the ancient lineage restricted to eastern Sumbawa. As the ash fallout and impact from the eruption was greatest in the northern and central portions of Sumbawa it is likely that flying lizards from eastern Sumbawa took refuge along the south and east coastal areas.

It is intriguing that the eastern Inner Banda Arc islands of Pantar and Alor (which become connected during glacial maxima) were not colonized by way of Lembata, which lies approximately 20 km to the west. Instead, a highly unlikely scenario of dispersal events from Flores to Sumba (~80 km), from Sumba to Timor (~300 km), from Timor to Wetar (~50 km), and from Wetar to Alor + Pantar (~80 km) resulted in the colonization of Pantar Island. This

hypothetical scenario is illustrated in Figure 20, though it should be noted that there are many other island colonization scenarios that could produce a tree with the same topology. This result highlights the stochastic and unpredictable nature of rafting events between oceanic islands that are influenced by many factors including the orientation of the island shore from which dispersers are swept into the sea from, ocean current and wind currents that change during seasons and climate cycles, the distance and size of target islands, and the persistence of newly arrived colonists to islands that may already be inhabited (eg. competitive exclusion). Competitive exclusion must play a role within this system because there are few or no instances of multiple colonization events on any one island, yet every major island has flying lizards. The likelihood that every island only receives dispersing flying lizards once is low, and if competitive exclusion did not exist we would expect to see individuals from any one island being placed all over the phylogeny mixed with other island lineages. From personal observations in the field, flying lizards seem to be filling the same ecological niche on every island, and many island-specific populations show morphological divergence among each other such as body size and coloration. This would make it difficult for newly arrived dispersers to become established on an already colonized island because they would be recognized as foreign, and would be competing for the same resources.

The genomic results presented here suggest that the Lesser Sundas *Draco* assemblage is in need of taxonomic revision due to the multiple independently evolving lineages within the group, though it is not clear at this point how many independent evolutionary lineages merit species status. Within *D. boschmai* there may be somewhere between 4-6 candidate species occurring on: 1) Lombok, 2) West Sumbawa, 3) East Sumbawa, 4) Flores + Lembata, and 5) Sumba. Specimens from each of these candidate species require morphological examination. Populations on the islands of Komodo and Rinca have not been examined here and may also be distinct. Within *D. timoriensis* there may be up to three species with populations from Timor + Rote, Wetar, and Alor + Pantar each requiring morphological examination.

CONCLUSIONS

Flying lizards of the Genus *Draco* are found on every major island in the Lesser Sunda archipelago except Sabu and colonized these oceanic islands by over-water dispersal events. Genetic samples were obtained from Bali and from 10 of the major Lesser Sunda Islands including many localities from the larger islands of Sumbawa, Flores, Sumba, and Timor. An initial analysis of mtDNA from over 350 samples confirmed that the Lesser Sundas assemblage is monophyletic and split from *Draco volans* from Java and Bali over 10 million years ago. Most of the smaller islands form monophyletic clades while the larger islands of Sumbawa, Flores, and Timor are paraphyletic, either from multiple colonizations of those islands or because of dispersal events from those islands to other islands. The genetic relationships of these clades were explored further by developing a 709-gene dataset derived from transcriptome sequences. These 709 genes were sequenced for 99 flying lizards from throughout the archipelago. The genomic phylogenies were able to resolve the relationships of these lineages and provide a clearer picture of the sequence that the islands were colonized. While the tree is somewhat consistent with a stepping-stone model of island colonization there are many aspects of the tree topology that do not agree with such a scenario. In particular, the fact that *Draco* dispersed south to Sumba, then east to Timor, then north to Wetar, and finally back west to Alor + Pantar is unexpected. Given the close distance between Lembata and Pantar it is remarkable that the

easternmost islands of the Inner Banda Arc were not colonized by way of Lembata. Additionally, the major phylogenetic break at the narrow isthmus of central Sumbawa warrants further study. The genetic divergence between populations on either side of Sumbawa is the greatest of any two populations within the Lesser Sunda Islands, and this could have been due to the merging of two paleo-islands to create Sumbawa. In summary, the biogeographical history of Lesser Sundas flying lizards is complex and will not be fully understood until we better understand the geological history of the archipelago, and *Draco boschmai* (and possibly *Draco timoriensis*) certainly represents multiple distinct species.

ACKNOWLEDGEMENTS

I would like to thank Jimmy A. McGuire, Djoko Iskandar, Alexander Stubbs, Ben Karin, Kristopher Harmon, Umilaila Arafin, Evy Arida, Gilang Ramadan, Amir Hamidy, Luke Bloch, Sarah Hykin, Jerome Fuchs, and Jim Patton for their help with the field collection of specimens and tissues. I thank Hinrich Kaiser for tissues from East Timor. I thank Ke Bi and Lydia Smith for laboratory support, Carol Spencer and Jackie Childers for accessioning of specimens, and Michelle Koo for GIS map support, and Ke Bi for bioinformatics support. Funding was provided by the MVZ, the Department of Integrative Biology at UC Berkeley, the National Geographic Foundation, and the National Science Foundation. Collection of specimens was carried out under UC Berkeley IACUC # A3084-01.

REFERENCES

Aljanabi, S. & Martinez, I. (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, 25:4692–4693.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, 215:403–410.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25(17):3389–3402.

Audley-Charles, M.G. (1987) Dispersal of Gondwanaland: relevance to evolution of the Angiosperms *In*: Whitmore, T.C. (ed.) *Biogeographical Evolution of the Malay Archipelago* Oxford Monographs on Biogeography 4, Clarendon Press, Oxford, pp. 5–25.

Audley-Charles, M.G. (2011) Tectonic post-collision processes in Timor. *Geological Society, London, Special Publications*, 355:241–266.

Auffenberg, W. (1980) The herpetofauna of Komodo, with notes on adjacent areas. *Bulletin of the Florida State Museum, Biological Sciences*, 25(2):39–156.

Avise, J.C. (2000) *Phylogeography: the history and formation of species*. Harvard university press.

Bi, K., Vanderpool, D., Singhal, S., Linderoth, T., Moritz, C., & Good, J.M. (2012) Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics*, 13(1):403.

Bi, K., Linderoth, T., Vanderpool, D., Good, J. M., Nielsen, R., & Moritz, C. (2013) Unlocking the vault: next-generation museum population genomics. *Molecular Ecology*, 22(24):6018–6032.

Bolger, A.M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114–2120.
Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 10(4), p.e1003537.
Boulenger, G.A. (1885) Catalogue of the Lizards in the British Museum (Nat. Hist.) I. Geckonidae, Eublepharidae, Uroplatidae, Pygopodidae, Agamidae. London: 450 pp.
Capella-Gutierrez, S., Silla-Martinez, J.M. & Gabaldon, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25:1972–1973.
Darevsky, I.S. (1964) Die Reptilien der Inseln Komodo, Padar und Rintja im Kleinen Sunda-Archipel, Indonesien. *Senckenbergiana Biologica*, 43:563–576.
Darriba, D., Taboada, G.L., Doallo, R., & Posada, D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9(8):772–772.
de Rooij, N. (1915) The reptiles of the Indo- Australian Archipelago. I. Lacertilia, Chelo- nia, Emydosauria. E. J. Brill Ltd., Leiden, The Netherlands.
Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7(1):214.
Earl, D.A. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2):359–361.
Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5):1792–1797.
Forcart, L. (1953) Die Amphibien und Reptilien von Sumba, ihre zoo- geographischen Beziehungen und Revision der Unterarten von *Typhlops polygrammicus*. *Verh. Naturf. Ges. Basel*, 64:356–388.
Grabherr, M., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. & Regev, A. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 15:644–652.
Gronau, I., Hubisz, M. J., Gulko, B., Danko, C. G. & Siepel, A. (2011) Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, 43:1031–1034.
Hall, R. (2009) Southeast Asia's changing paleogeography. *Blumea*, 54:148–161.
Hennig, W. (1936) Revision der Gattung Draco (Agamidae). Temminckia, Leiden 1: 153–220.
Hodges, E., Rooks, M., Xuan, Z., Bhattacharjee, A., Gordon, D.B., Brizuela, L., McCombie, W.R. & Hannon, G.J. (2009) Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively parallel sequencing. *Nature Protocols*, 4(6):960–974.
Huang, X. & Madan, A. (1999) CAP3: a DNA sequence assembly program. *Genome Research*, 9:868–877.
Jombart, T. (2008) *adegenet*: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11):1403–1405.
Katoh, K. & Standley, D.M. (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, 30:722–780.
Kent, W. (2002) BLAT–the BLAST-like alignment tool. Genome Research, 12:656–664.

Kuhl, H. (1820) Beiträge zur Zoologie und vergleichenden Anatomie. Hermannsche Buchhandlung, Frankfurt, 152 pp.
Kumar, S. & Subramanian, S. (2002) Mutation rates in mammalian genomes. *Proceedings of the National Academy of Sciences*, 99(2):803–808.
Langmead, B. & Salzberg, S. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9:357–359.
Li, W. & Godzik, A. (2006) CD-Hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22:1658–1659.
Li, H. & Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25:1754–1760.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, 25:2078–2079.
Liu, B., Yuan, J., Yiu, S-M., Li, Z., Xie, Y., Chen, Y., Shi, Y., Zhang, H., Li, Y., Lam, T-W. & Luo, R. (2012) COPE: an accurate k-mer-based pair-end reads connection tool to facilitate genome assembly. *Bioinformatics*, 28:2870–2874.
Lomolino, M.V., Riddle, B.R. & Brown, J.H. (2006) *Biogeography*. Sunderland, MA: Sinauer Associates.
Macey, J.R., Larson, A., Ananjeva, N.B., Fang, Z. & Papenfuss, T.J. (1997) Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. *Molecular Biology and Evolution,* 14:91–104.
Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17:10–12.
Magoč, T. & Salzberg, S. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, 27:2957–2963.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. & DePristo, M.A. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20:1297–1303.
Mertens, R. (1930) Die Amphibien und Reptilien der Inseln Bali, Lombok, Sumbawa und Flores. Abhandlungen der Senckenbergischen Natur- forschenden Gesellschaft, 42:115–344.
Meyer, M. & Kircher, M. (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, 2010(6), pdb.prot5448.
Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S. & Warnow, T. (2014) ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics*, 30(17):541–548.
Musters, C.J.M. (1983) Taxonomy of the genus Draco L. (Agamidae, Lacertilia, Reptilia). *Zoologische Verhandelingen*, 199:1–120.
McGuire, J.A. & Heang, K.B. (2001) Phylogenetic systematics of Southeast Asian flying lizards (Iguania: Agamidae: Draco) as inferred from mitochondrial DNA sequence data. *Biological Journal of the Linnean Society*, 72:203–229.
McGuire, J.A., Linkem, C.W., Koo, M.S., Hutchison, D.W., Lappin, A.K., Orange, D.I., Lemos-Espinal, J., Riddle, B.R. & Jaeger, J.R. (2007) Mitochondrial Introgression and Incomplete Lineage Sorting through Space and Time: Phylogenetics of Crotaphytid Lizards. *Evolution*, 61(12):2879–2897.

McGuire, J.A., & Dudley, R. (2011) The biology of gliding in flying lizards (Genus *Draco*) and their fossil and extant analogs. *Integrative and Comparative Biology*, 51(6):983–990.
Portik, D.M., Smith, L.L. & Bi, K. (2016) An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (Class: Amphibia, Order: Anura). *Molecular Ecology Resources*, 16(5):1069–1083.
Pritchard, J.K., Stephens, M., & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155(2):945–959.
Rambaut, A. & Drummond, A.J. (2009) Tracer version 1.5.0. *URL http://tree.bio.ed.ac.uk/software/tracer/[accessed 1 July 2013]*.
Simpson, J., Wong, K., Jackman, S., Schein, J., Jones, S. & Birol, I. (2009) ABySS: a parallel assembler for short read sequence data. *Genome Research*, 19:1117–1123.
Singhal, S. (2013) *De novo* transcriptomic analyses for non-model organisms: an evaluation of methods across a multi-species data set. *Molecular Ecology Resources*, 13:403–416.
Slater, G.S. & Birney, E. (2005) Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, 6:31.
Smit, A.F.A., Hubley, R. & Green, P. (2015) *RepeatMasker Open-4.0*. [http://www.repeatmasker.org]
Stamatakis, A. (2014) RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9):1312–1313.
Toews, D.P., & Brelsford, A. (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, 21(16):3907–3930.
Wermuth, H. (1967) *Liste der rezenten Amphibien und Reptilien*, *In* R. Mertens, W. Hennig, and H. Wermuth (eds), Das Tierreich, Vol. 86. Walter de Gruyter and Co., Berlin.

TABLES

Table 1. Summary statistics for transcriptomes sequenced from four flying lizards.

Catalog #	Species	Locality	Total Length (bp)	Total Contigs	Mean	Max	>2000bp	>1000bp	GC%
JAM11504	*D. walkeri*	Sulawesi	8,715,915	8,154	1,068	16,179	1,042	3,119	48.41
JAM12091	*D. boschmai*	East Sumbawa	14,681,065	9,449	1,553	14,911	2,533	5,810	48.61
JAM12477	*D. boschmai*	West Flores	14,301,600	9,368	1,526	14,944	2,430	5,674	49.07
JAM12741	*D. boschmai*	Lembata	12,431,086	8,625	1,441	14,884	2,099	5,086	48.71

Table 2. Locality information for genetic samples. X=exon capture sample.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
JAM11506	*Draco*	*boschmai*	Lombok	-8.67180	116.16745	Lombok	
JAM11507	*Draco*	*boschmai*	Lombok	-8.67180	116.16745	Lombok	X
JAM11508	*Draco*	*boschmai*	Lombok	-8.67180	116.16745	Lombok	
JAM11509	*Draco*	*boschmai*	Lombok	-8.67180	116.16745	Lombok	
JAM11510	*Draco*	*boschmai*	Lombok	-8.67180	116.16745	Lombok	X
JAM11511	*Draco*	*boschmai*	Lombok	-8.67180	116.16745	Lombok	
JAM11512	*Draco*	*boschmai*	Lombok	-8.67180	116.16745	Lombok	
JAM11513	*Draco*	*boschmai*	Lombok	-8.67180	116.16745	Lombok	
JAM11514	*Draco*	*boschmai*	Lombok	-8.67180	116.16745	Lombok	
JAM11515	*Draco*	*boschmai*	Lombok	-8.67180	116.16745	Lombok	X
JAM11516	*Draco*	*boschmai*	Lombok	-8.67180	116.16745	Lombok	
JAM11528	*Draco*	*boschmai*	Lombok	-8.53411	116.24258	Lombok	
JAM11529	*Draco*	*boschmai*	Lombok	-8.52801	116.27552	Lombok	
JAM11530	*Draco*	*boschmai*	Lombok	-8.53189	116.24080	Lombok	
JAM11531	*Draco*	*boschmai*	Lombok	-8.53189	116.24080	Lombok	
JAM11539	*Draco*	*boschmai*	Lombok	-8.66991	116.16084	Lombok	
JAM11540	*Draco*	*boschmai*	Lombok	-8.66991	116.16084	Lombok	
JAM11541	*Draco*	*boschmai*	Lombok	-8.66991	116.16084	Lombok	
JAM11542	*Draco*	*boschmai*	Lombok	-8.66991	116.16084	Lombok	
JAM11543	*Draco*	*boschmai*	Lombok	-8.66991	116.16084	Lombok	
JAM11544	*Draco*	*boschmai*	Lombok	-8.66991	116.16084	Lombok	
JAM11545	*Draco*	*boschmai*	Lombok	-8.66991	116.16084	Lombok	
JAM11546	*Draco*	*boschmai*	Lombok	-8.66991	116.16084	Lombok	
JAM11547	*Draco*	*boschmai*	Lombok	-8.66991	116.16084	Lombok	
JAM11548	*Draco*	*boschmai*	Lombok	-8.66991	116.16084	Lombok	
JAM11639	*Draco*	*boschmai*	Sumbawa	-8.43816	117.29297	West Sumbawa	
JAM11640	*Draco*	*boschmai*	Sumbawa	-8.43816	117.29297	West Sumbawa	
JAM11641	*Draco*	*boschmai*	Sumbawa	-8.43816	117.29297	West Sumbawa	
JAM11642	*Draco*	*boschmai*	Sumbawa	-8.43816	117.29297	West Sumbawa	X
JAM11643	*Draco*	*boschmai*	Sumbawa	-8.43816	117.29297	West Sumbawa	
JAM11644	*Draco*	*boschmai*	Sumbawa	-8.43816	117.29297	West Sumbawa	
JAM11645	*Draco*	*boschmai*	Sumbawa	-8.43816	117.29297	West Sumbawa	
JAM11646	*Draco*	*boschmai*	Sumbawa	-8.43816	117.29297	West Sumbawa	
JAM11647	*Draco*	*boschmai*	Sumbawa	-8.43816	117.29297	West Sumbawa	
JAM11649	*Draco*	*boschmai*	Sumbawa	-8.58800	117.28601	West Sumbawa	X
JAM11650	*Draco*	*boschmai*	Sumbawa	-8.58467	117.29019	West Sumbawa	
JAM11651	*Draco*	*boschmai*	Sumbawa	-8.58537	117.28937	West Sumbawa	X
JAM11652	*Draco*	*boschmai*	Sumbawa	-8.57435	117.31339	West Sumbawa	
JAM11653	*Draco*	*boschmai*	Sumbawa	-8.70333	117.40506	West Sumbawa	
JAM11655	*Draco*	*boschmai*	Sumbawa	-8.71719	117.39417	West Sumbawa	
JAM11656	*Draco*	*boschmai*	Sumbawa	-8.71719	117.39417	West Sumbawa	X
JAM11657	*Draco*	*boschmai*	Sumbawa	-8.71719	117.39417	West Sumbawa	
JAM11669	*Draco*	*boschmai*	Sumbawa	-8.58349	117.65656	West Sumbawa	
JAM11670	*Draco*	*boschmai*	Sumbawa	-8.58349	117.65656	West Sumbawa	
JAM11671	*Draco*	*boschmai*	Sumbawa	-8.58349	117.65656	West Sumbawa	
JAM11677	*Draco*	*boschmai*	Sumbawa	-8.56787	117.66972	West Sumbawa	
JAM11678	*Draco*	*boschmai*	Sumbawa	-8.56787	117.66972	West Sumbawa	
JAM11679	*Draco*	*boschmai*	Sumbawa	-8.56787	117.66972	West Sumbawa	
JAM11680	*Draco*	*boschmai*	Sumbawa	-8.56787	117.66972	West Sumbawa	
JAM11681	*Draco*	*boschmai*	Sumbawa	-8.56787	117.66972	West Sumbawa	X
JAM11682	*Draco*	*boschmai*	Sumbawa	-8.58349	117.65656	West Sumbawa	
JAM11710	*Draco*	*boschmai*	Sumbawa	-8.61458	116.85419	West Sumbawa	X

Table 2 cont.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
JAM11711	*Draco*	*boschmai*	Sumbawa	-8.61458	116.85419	West Sumbawa	X
JAM11742	*Draco*	*boschmai*	Sumbawa	-8.59253	118.28566	East Sumbawa	
JAM11743	*Draco*	*boschmai*	Sumbawa	-8.59253	118.28566	East Sumbawa	
JAM11744	*Draco*	*boschmai*	Sumbawa	-8.59253	118.28566	East Sumbawa	X
JAM11745	*Draco*	*boschmai*	Sumbawa	-8.59253	118.28566	East Sumbawa	
JAM11746	*Draco*	*boschmai*	Sumbawa	-8.59253	118.28566	East Sumbawa	
JAM11747	*Draco*	*boschmai*	Sumbawa	-8.59253	118.28566	East Sumbawa	
JAM11769	*Draco*	*boschmai*	Flores	-8.494136	119.879587	West Flores	
JAM11770	*Draco*	*boschmai*	Flores	-8.494136	119.879587	West Flores	
JAM11771	*Draco*	*boschmai*	Flores	-8.494136	119.879587	West Flores	
JAM11772	*Draco*	*boschmai*	Flores	-8.494136	119.879587	West Flores	
JAM11793	*Draco*	*boschmai*	Flores	-8.53961	119.92802	West Flores	
JAM11796	*Draco*	*boschmai*	Flores	-8.51781	119.8894	West Flores	X
JAM11797	*Draco*	*boschmai*	Flores	-8.51781	119.8894	West Flores	
JAM11799	*Draco*	*boschmai*	Flores	-8.494136	119.879587	West Flores	
JAM11800	*Draco*	*boschmai*	Flores	-8.494136	119.879587	West Flores	X
JAM11801	*Draco*	*boschmai*	Flores	-8.494136	119.879587	West Flores	
JAM11802	*Draco*	*boschmai*	Flores	-8.494136	119.879587	West Flores	
JAM11803	*Draco*	*boschmai*	Flores	-8.494136	119.879587	West Flores	
JAM11804	*Draco*	*boschmai*	Flores	-8.494136	119.879587	West Flores	
JAM11906	*Draco*	*boschmai*	Lombok	-8.53912	116.53959	Lombok	
JAM11907	*Draco*	*boschmai*	Lombok	-8.53912	116.53959	Lombok	
JAM11908	*Draco*	*boschmai*	Lombok	-8.53912	116.53959	Lombok	
JAM11909	*Draco*	*boschmai*	Lombok	-8.53912	116.53959	Lombok	
JAM11910	*Draco*	*boschmai*	Lombok	-8.53912	116.53959	Lombok	X
JAM11911	*Draco*	*boschmai*	Lombok	-8.53912	116.53959	Lombok	
JAM11912	*Draco*	*boschmai*	Lombok	-8.53912	116.53959	Lombok	
JAM11935	*Draco*	*boschmai*	Lombok	-8.26857	116.43265	Lombok	X
JAM11936	*Draco*	*boschmai*	Lombok	-8.26857	116.43265	Lombok	X
JAM11937	*Draco*	*boschmai*	Lombok	-8.26857	116.43265	Lombok	
JAM11939	*Draco*	*boschmai*	Lombok	-8.26857	116.43265	Lombok	X
JAM12091	*Draco*	*boschmai*	Sumbawa	-8.82686	118.01466	East Sumbawa	
JAM12092	*Draco*	*boschmai*	Sumbawa	-8.82686	118.01466	East Sumbawa	
JAM12093	*Draco*	*boschmai*	Sumbawa	-8.82686	118.01466	East Sumbawa	
JAM12097	*Draco*	*boschmai*	Sumbawa	-8.440202	118.301486	East Sumbawa	
JAM12098	*Draco*	*boschmai*	Sumbawa	-8.440202	118.301486	East Sumbawa	
JAM12099	*Draco*	*boschmai*	Sumbawa	-8.440202	118.301486	East Sumbawa	
JAM12100	*Draco*	*boschmai*	Sumbawa	-8.440202	118.301486	East Sumbawa	X
JAM12101	*Draco*	*boschmai*	Sumbawa	-8.440202	118.301486	East Sumbawa	
JAM12114	*Draco*	*boschmai*	Sumbawa	-8.43721	118.30004	East Sumbawa	X
JAM12115	*Draco*	*boschmai*	Sumbawa	-8.43721	118.30004	East Sumbawa	
JAM12118	*Draco*	*boschmai*	Sumbawa	-8.54473	118.35419	East Sumbawa	
JAM12119	*Draco*	*boschmai*	Sumbawa	-8.54473	118.35419	East Sumbawa	
JAM12121	*Draco*	*boschmai*	Sumbawa	-8.54473	118.35419	East Sumbawa	
JAM12123	*Draco*	*boschmai*	Sumbawa	-8.54473	118.35419	East Sumbawa	
JAM12124	*Draco*	*boschmai*	Sumbawa	-8.54473	118.35419	East Sumbawa	
JAM12127	*Draco*	*boschmai*	Sumbawa	-8.49214	118.52927	East Sumbawa	
JAM12165	*Draco*	*boschmai*	Sumbawa	-8.54700	118.46947	East Sumbawa	
JAM12168	*Draco*	*boschmai*	Sumbawa	-8.63729	118.47919	East Sumbawa	
JAM12169	*Draco*	*boschmai*	Sumbawa	-8.63729	118.47919	East Sumbawa	
JAM12176	*Draco*	*boschmai*	Sumbawa	-8.54700	118.46947	East Sumbawa	X
JAM12177	*Draco*	*boschmai*	Sumbawa	-8.54700	118.46947	East Sumbawa	

Table 2 cont.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
JAM12196	*Draco*	*boschmai*	Sumbawa	-8.49236	118.52563	East Sumbawa	
JAM12197	*Draco*	*boschmai*	Sumbawa	-8.49236	118.52563	East Sumbawa	
JAM12198	*Draco*	*boschmai*	Sumbawa	-8.49236	118.52563	East Sumbawa	
JAM12200	*Draco*	*boschmai*	Sumbawa	-8.67775	118.68163	East Sumbawa	
JAM12201	*Draco*	*boschmai*	Sumbawa	-8.67775	118.68163	East Sumbawa	
JAM12237	*Draco*	*boschmai*	Sumbawa	-8.35	118.782	East Sumbawa	
JAM12238	*Draco*	*boschmai*	Sumbawa	-8.35	118.782	East Sumbawa	
JAM12239	*Draco*	*boschmai*	Sumbawa	-8.35	118.782	East Sumbawa	
JAM12240	*Draco*	*boschmai*	Sumbawa	-8.35	118.782	East Sumbawa	X
JAM12241	*Draco*	*boschmai*	Sumbawa	-8.35	118.782	East Sumbawa	
JAM12242	*Draco*	*boschmai*	Sumbawa	-8.35	118.782	East Sumbawa	
JAM12243	*Draco*	*boschmai*	Sumbawa	-8.35	118.782	East Sumbawa	
JAM12244	*Draco*	*boschmai*	Sumbawa	-8.35	118.782	East Sumbawa	
JAM12254	*Draco*	*boschmai*	Sumbawa	-8.67	118.906	East Sumbawa	
JAM12256	*Draco*	*boschmai*	Sumbawa	-8.49715	118.80463	East Sumbawa	
JAM12258	*Draco*	*boschmai*	Sumbawa	-8.505	118.793	East Sumbawa	
JAM12259	*Draco*	*boschmai*	Sumbawa	-8.505	118.793	East Sumbawa	X
JAM12260	*Draco*	*boschmai*	Sumbawa	-8.505	118.793	East Sumbawa	
JAM12266	*Draco*	*boschmai*	Sumbawa	-8.66500	118.07012	East Sumbawa	X
JAM12267	*Draco*	*boschmai*	Sumbawa	-8.66500	118.07012	East Sumbawa	
JAM12321	*Draco*	*boschmai*	Sumbawa	-8.76763	118.60463	East Sumbawa	
JAM12322	*Draco*	*boschmai*	Sumbawa	-8.76763	118.60463	East Sumbawa	
JAM12323	*Draco*	*boschmai*	Sumbawa	-8.76689	118.60458	East Sumbawa	
JAM12345	*Draco*	*boschmai*	Sumbawa	-8.68169	118.92421	East Sumbawa	X
JAM12355	*Draco*	*boschmai*	Flores	-8.733694	121.335462	Central Flores	
JAM12357	*Draco*	*boschmai*	Flores	-8.78828	121.67159	Central Flores	
JAM12358	*Draco*	*boschmai*	Flores	-8.78828	121.67159	Central Flores	
JAM12359	*Draco*	*boschmai*	Flores	-8.78616	121.67315	Central Flores	
JAM12360	*Draco*	*boschmai*	Flores	-8.78616	121.67315	Central Flores	
JAM12361	*Draco*	*boschmai*	Flores	-8.78616	121.67315	Central Flores	
JAM12362	*Draco*	*boschmai*	Flores	-8.78616	121.67315	Central Flores	
JAM12363	*Draco*	*boschmai*	Flores	-8.78616	121.67315	Central Flores	
JAM12364	*Draco*	*boschmai*	Flores	-8.78616	121.67315	Central Flores	X
JAM12365	*Draco*	*boschmai*	Flores	-8.78616	121.67315	Central Flores	
JAM12366	*Draco*	*boschmai*	Flores	-8.78616	121.67315	Central Flores	X
JAM12367	*Draco*	*boschmai*	Flores	-8.78616	121.67315	Central Flores	
JAM12368	*Draco*	*boschmai*	Flores	-8.78616	121.67315	Central Flores	
JAM12374	*Draco*	*boschmai*	Flores	-8.7803	121.6692	Central Flores	
JAM12375	*Draco*	*boschmai*	Flores	-8.78029	121.66988	Central Flores	
JAM12400	*Draco*	*boschmai*	Flores	-8.495767	119.897378	West Flores	X
JAM12422	*Draco*	*boschmai*	Flores	-8.59810	119.961020	West Flores	
JAM12423	*Draco*	*boschmai*	Flores	-8.59810	119.961020	West Flores	
JAM12424	*Draco*	*boschmai*	Flores	-8.59810	119.961020	West Flores	
JAM12425	*Draco*	*boschmai*	Flores	-8.59810	119.961020	West Flores	X
JAM12429	*Draco*	*boschmai*	Flores	-8.68671	120.26979	West Flores	X
JAM12448	*Draco*	*boschmai*	Flores	-8.80738	120.59010	West Flores	
JAM12449	*Draco*	*boschmai*	Flores	-8.80738	120.59010	West Flores	
JAM12450	*Draco*	*boschmai*	Flores	-8.80738	120.59010	West Flores	
JAM12451	*Draco*	*boschmai*	Flores	-8.80738	120.59010	West Flores	
JAM12452	*Draco*	*boschmai*	Flores	-8.80738	120.59010	West Flores	
JAM12453	*Draco*	*boschmai*	Flores	-8.80738	120.59010	West Flores	
JAM12454	*Draco*	*boschmai*	Flores	-8.80738	120.59010	West Flores	

Table 2 cont.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
JAM12455	*Draco*	*boschmai*	Flores	-8.80738	120.59010	West Flores	
JAM12456	*Draco*	*boschmai*	Flores	-8.80738	120.59010	West Flores	X
JAM12457	*Draco*	*boschmai*	Flores	-8.80738	120.59010	West Flores	
JAM12458	*Draco*	*boschmai*	Flores	-8.80738	120.59010	West Flores	
JAM12476	*Draco*	*boschmai*	Flores	-8.75715	120.98904	West Flores	
JAM12477	*Draco*	*boschmai*	Flores	-8.71399	121.02392	West Flores	
JAM12478	*Draco*	*boschmai*	Flores	-8.71399	121.02392	West Flores	
JAM12529	*Draco*	*boschmai*	Flores	-8.83634	121.68251	Central Flores	X
JAM12530	*Draco*	*boschmai*	Flores	-8.83634	121.68251	Central Flores	X
JAM12531	*Draco*	*boschmai*	Flores	-8.83634	121.68251	Central Flores	X
JAM12532	*Draco*	*boschmai*	Flores	-8.83634	121.68251	Central Flores	
JAM12533	*Draco*	*boschmai*	Flores	-8.83634	121.68251	Central Flores	
JAM12545	*Draco*	*boschmai*	Flores	-8.775186	121.949508	Central Flores 2	X
JAM12546	*Draco*	*boschmai*	Flores	-8.775186	121.949508	Central Flores	X
JAM12555	*Draco*	*boschmai*	Flores	-8.29509	123.01810	East Flores	
JAM12556	*Draco*	*boschmai*	Flores	-8.29509	123.01810	East Flores	
JAM12557	*Draco*	*boschmai*	Flores	-8.29509	123.01810	East Flores	
JAM12558	*Draco*	*boschmai*	Flores	-8.29509	123.01810	East Flores	
JAM12559	*Draco*	*boschmai*	Flores	-8.29509	123.01810	East Flores	X
JAM12576	*Draco*	*boschmai*	Flores	-8.27238	122.99255	East Flores	
JAM12577	*Draco*	*boschmai*	Flores	-8.27238	122.99255	East Flores	
JAM12578	*Draco*	*boschmai*	Flores	-8.27238	122.99255	East Flores	
JAM12579	*Draco*	*boschmai*	Flores	-8.27238	122.99255	East Flores	
JAM12583	*Draco*	*boschmai*	Flores	-8.298918	123.017705	East Flores	
JAM12584	*Draco*	*boschmai*	Flores	-8.298918	123.017705	East Flores	
JAM12585	*Draco*	*boschmai*	Flores	-8.298918	123.017705	East Flores	
JAM12586	*Draco*	*boschmai*	Flores	-8.298918	123.017705	East Flores	
JAM12587	*Draco*	*boschmai*	Flores	-8.298918	123.017705	East Flores	
JAM12588	*Draco*	*boschmai*	Flores	-8.298918	123.017705	East Flores	
JAM12589	*Draco*	*boschmai*	Flores	-8.298918	123.017705	East Flores	X
JAM12590	*Draco*	*boschmai*	Flores	-8.298918	123.017705	East Flores	
JAM12591	*Draco*	*boschmai*	Flores	-8.298918	123.017705	East Flores	
JAM12599	*Draco*	*boschmai*	Flores	-8.10555	122.52252	East Flores	
JAM12600	*Draco*	*boschmai*	Flores	-8.10555	122.52252	East Flores	
JAM12608	*Draco*	*boschmai*	Flores	-8.10555	122.52252	East Flores	
JAM12609	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	X
JAM12610	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12611	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12612	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12613	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	X
JAM12614	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12615	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12616	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12617	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12626	*Draco*	*boschmai*	Lembata	-8.272336	123.458282	Lembata	
JAM12627	*Draco*	*boschmai*	Lembata	-8.272336	123.458282	Lembata	X
JAM12645	*Draco*	*boschmai*	Lembata	-8.471636	123.350576	Lembata	
JAM12646	*Draco*	*boschmai*	Lembata	-8.471636	123.350576	Lembata	
JAM12648	*Draco*	*boschmai*	Lembata	-8.471636	123.350576	Lembata	
JAM12649	*Draco*	*boschmai*	Lembata	-8.471636	123.350576	Lembata	
JAM12718	*Draco*	*boschmai*	Lembata	-8.536959	123.448057	Lembata	
JAM12719	*Draco*	*boschmai*	Lembata	-8.536959	123.448057	Lembata	

Table 2 cont.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
JAM12720	*Draco*	*boschmai*	Lembata	-8.536959	123.448057	Lembata	
JAM12721	*Draco*	*boschmai*	Lembata	-8.536959	123.448057	Lembata	
JAM12722	*Draco*	*boschmai*	Lembata	-8.536959	123.448057	Lembata	
JAM12723	*Draco*	*boschmai*	Lembata	-8.536959	123.448057	Lembata	
JAM12726	*Draco*	*boschmai*	Lembata	-8.536959	123.448057	Lembata	
JAM12727	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12728	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12729	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12730	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12731	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	X
JAM12732	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12733	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12734	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12735	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12736	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12737	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	X
JAM12738	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12739	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12740	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12741	*Draco*	*boschmai*	Lembata	-8.39496	123.39997	Lembata	
JAM12800	*Draco*	*timorensis*	Timor	-10.15731	123.58482	West Timor	
JAM12802	*Draco*	*timorensis*	Timor	-10.15731	123.58482	West Timor	
JAM12804	*Draco*	*timorensis*	Alor	-8.19708	124.60590	Alor	X
JAM12805	*Draco*	*timorensis*	Alor	-8.19708	124.60590	Alor	
JAM12806	*Draco*	*timorensis*	Alor	-8.19708	124.60590	Alor	
JAM12807	*Draco*	*timorensis*	Alor	-8.19708	124.60590	Alor	
JAM12808	*Draco*	*timorensis*	Alor	-8.19708	124.60590	Alor	
JAM12809	*Draco*	*timorensis*	Alor	-8.19708	124.60590	Alor	
JAM12810	*Draco*	*timorensis*	Alor	-8.19708	124.60590	Alor	
JAM12811	*Draco*	*timorensis*	Alor	-8.19708	124.60590	Alor	X
JAM12812	*Draco*	*timorensis*	Alor	-8.19708	124.60590	Alor	
JAM12813	*Draco*	*timorensis*	Alor	-8.19708	124.60590	Alor	
JAM12814	*Draco*	*timorensis*	Alor	-8.19708	124.60590	Alor	X
JAM12815	*Draco*	*timorensis*	Alor	-8.19708	124.60590	Alor	
JAM12816	*Draco*	*timorensis*	Alor	-8.19708	124.60590	Alor	
JAM12824	*Draco*	*timorensis*	Alor	-8.16993	124.59335	Alor	
JAM12825	*Draco*	*timorensis*	Alor	-8.16993	124.59335	Alor	X
JAM12883	*Draco*	*timorensis*	Alor	-8.19602	124.69077	Alor	
JAM13002	*Draco*	*timorensis*	Alor	-8.17922	124.55897	Alor	
JAM13003	*Draco*	*timorensis*	Alor	-8.17922	124.55897	Alor	X
JAM13004	*Draco*	*timorensis*	Alor	-8.17922	124.55897	Alor	
JAM13005	*Draco*	*timorensis*	Alor	-8.17922	124.55897	Alor	
JAM13025	*Draco*	*boschmai*	Flores	-8.21648	122.97288	East Flores	X
JAM13074	*Draco*	*boschmai*	Sumba	-10.02181	120.05845	Southwest Sumba	
JAM13075	*Draco*	*boschmai*	Sumba	-10.02298	120.06001	Southwest Sumba	X
JAM13076	*Draco*	*boschmai*	Sumba	-10.02298	120.06001	Southwest Sumba	X
JAM13160	*Draco*	*boschmai*	Sumba	-10.02102	120.05795	Southwest Sumba	X
JAM13161	*Draco*	*boschmai*	Sumba	-10.02334	120.06014	Southwest Sumba	X
JAM13162	*Draco*	*boschmai*	Sumba	-10.02334	120.06014	Southwest Sumba	
JAM13163	*Draco*	*boschmai*	Sumba	-10.02334	120.06014	Southwest Sumba	
JAM13199	*Draco*	*boschmai*	Sumba	-10.01874	120.05257	Southwest Sumba	
JAM13222	*Draco*	*boschmai*	Sumba	-10.01721	120.05066	Southwest Sumba	X

Table 2 cont.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
JAM13223	*Draco*	*boschmai*	Sumba	-10.01721	120.05066	Southwest Sumba	X
JAM13224	*Draco*	*boschmai*	Sumba	-10.01721	120.05066	Southwest Sumba	
JAM13225	*Draco*	*boschmai*	Sumba	-10.01721	120.05066	Southwest Sumba	
JAM13226	*Draco*	*boschmai*	Sumba	-10.01721	120.05066	Southwest Sumba	
JAM13229	*Draco*	*boschmai*	Sumba	-10.01721	120.05066	Southwest Sumba	X
JAM13231	*Draco*	*boschmai*	Sumba	-10.01401	120.04813	Southwest Sumba	
JAM13274	*Draco*	*boschmai*	Sumba	-10.08773	120.75008	East Sumba	X
JAM13290	*Draco*	*boschmai*	Sumba	-10.23209	120.52045	East Sumba	
JAM13291	*Draco*	*boschmai*	Sumba	-10.22961	120.52826	East Sumba	X
JAM13292	*Draco*	*boschmai*	Sumba	-10.22961	120.52826	East Sumba	
JAM13293	*Draco*	*boschmai*	Sumba	-10.22961	120.52826	East Sumba	
JAM13299	*draco*	*boschmai*	Sumba	-10.21165	120.61823	East Sumba	X
JAM13313	*Draco*	*boschmai*	Sumba	-9.65087	119.75185	East Sumba	
JAM13314	*Draco*	*boschmai*	Sumba	-9.65087	119.75185	East Sumba	
JAM13315	*Draco*	*boschmai*	Sumba	-9.65087	119.75185	East Sumba	
JAM13316	*Draco*	*boschmai*	Sumba	-9.65087	119.75185	East Sumba	
JAM13317	*Draco*	*boschmai*	Sumba	-9.65087	119.75185	East Sumba	X
JAM13320	*Draco*	*boschmai*	Sumba	-9.59077	119.57867	East Sumba	X
JAM13321	*Draco*	*boschmai*	Sumba	-9.59077	119.57867	East Sumba	
JAM13322	*Draco*	*boschmai*	Sumba	-9.64868	119.73212	East Sumba	
JAM13470	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13471	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13472	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13473	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13474	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13475	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	X
JAM13476	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13477	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13478	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13479	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13480	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	X
JAM13481	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13482	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13483	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13484	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13485	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13486	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	X
JAM13487	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13488	*Draco*	*timorensis*	Timor	-10.24853	123.65707	West Timor	
JAM13518	*Draco*	*timorensis*	Timor	-10.262283	123.76075	West Timor	
JAM13522	*Draco*	*timorensis*	Timor	-10.25825	123.79923	N/A	X
JAM13557	*Draco*	*timorensis*	Timor	-10.26617	123.56583	West Timor	
JAM13607	*Draco*	*timorensis*	Timor	-10.03867	123.92876	West Timor	X
JAM13608	*Draco*	*timorensis*	Timor	-10.03867	123.92876	West Timor	X
JAM13609	*Draco*	*timorensis*	Timor	-10.03869	123.93173	West Timor	
JAM13677	*Draco*	*timorensis*	Wetar	-7.92642	126.41137	Wetar	
JAM13678	*Draco*	*timorensis*	Wetar	-7.92847	126.40781	Wetar	X
JAM13679	*Draco*	*timorensis*	Wetar	-7.92642	126.41137	Wetar	
JAM13721	*Draco*	*timorensis*	Wetar	-7.92642	126.41137	Wetar	
JAM13722	*Draco*	*timorensis*	Wetar	-7.92642	126.41137	Wetar	X
JAM13723	*Draco*	*timorensis*	Wetar	-7.92642	126.41137	Wetar	
JAM13752	*Draco*	*timorensis*	Wetar	-7.92642	126.41137	Wetar	X

Table 2 cont.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
JAM13809	*Draco*	*timorensis*	Wetar	-7.92772	126.42298	Wetar	X
JAM13874	*Draco*	*timorensis*	Wetar	-7.92972	126.42298	Wetar	X
JAM13904	*Draco*	*timorensis*	Wetar	-7.92612	126.41111	Wetar	
JAM13906	*Draco*	*timorensis*	Wetar	-7.91965	126.39785	Wetar	X
JAM13979	*Draco*	*timorensis*	Pantar	-8.35595	124.2547	Pantar	X
JAM13980	*Draco*	*timorensis*	Pantar	-8.35595	124.2547	Pantar	X
JAM13981	*Draco*	*timorensis*	Pantar	-8.35392	124.25347	Pantar	
JAM13982	*Draco*	*timorensis*	Pantar	-8.35392	124.25347	Pantar	X
JAM13983	*Draco*	*timorensis*	Pantar	-8.35392	124.25347	Pantar	
JAM13984	*Draco*	*timorensis*	Pantar	-8.35392	124.25347	Pantar	
JAM13985	*Draco*	*timorensis*	Pantar	-8.35392	124.25347	Pantar	
JAM13986	*Draco*	*timorensis*	Pantar	-8.35591	124.25462	Pantar	
JAM13987	*Draco*	*timorensis*	Pantar	-8.35591	124.25462	Pantar	X
JAM14054	*Draco*	*timorensis*	Pantar	-8.35392	124.25347	Pantar	X
JAM14055	*Draco*	*timorensis*	Pantar	-8.35392	124.25347	Pantar	
ALS21	*Draco*	*volans*	Bali	-8.4380975	115.401722	Bali	X
ALS22	*Draco*	*volans*	Bali	-8.4380975	115.401722	Bali	X
ALS24	*Draco*	*volans*	Bali	-8.4380975	115.401722	Bali	X
ALS32	*Draco*	*volans*	Bali	-8.4380975	115.401722	Bali	X
ALS35	*Draco*	*volans*	Bali	-8.4380975	115.401722	Bali	X
ALS36	*Draco*	*volans*	Bali	-8.4380975	115.401722	Bali	
SBR128	*Draco*	*timorensis*	Rote	-10.88277	122.83071	Rote	X
SBR129	*Draco*	*timorensis*	Rote	-10.88277	122.83071	Rote	X
SBR130	*Draco*	*timorensis*	Rote	-10.88277	122.83071	Rote	X
SBR132	*Draco*	*timorensis*	Rote	-10.87308	122.84415	Rote	
SBR133	*Draco*	*timorensis*	Rote	-10.87308	122.84415	Rote	
SBR134	*Draco*	*timorensis*	Rote	-10.86902	122.84645	Rote	
SBR135	*Draco*	*timorensis*	Rote	-10.86045	112.92305	Rote	X
SBR136	*Draco*	*timorensis*	Rote	-10.86045	112.92305	Rote	X
SBR137	*Draco*	*timorensis*	Rote	-10.86045	112.92305	Rote	
SBR138	*Draco*	*timorensis*	Rote	-10.86045	112.92305	Rote	X
SBR139	*Draco*	*timorensis*	Rote	-10.87308	122.84415	Rote	
SBR180	*Draco*	*timorensis*	Rote	-10.87308	122.84415	Rote	
SBR181	*Draco*	*timorensis*	Rote	-10.87308	122.84415	Rote	
SBR199	*Draco*	*timorensis*	Timor	-9.8218	124.31039	N/A	X
USNM579037	*Draco*	*timorensis*	East Timor	-8.833	126.383	East Timor	X
USNM579039	*Draco*	*timorensis*	East Timor	-8.833	126.383	East Timor	
USNM579040	*Draco*	*timorensis*	East Timor	-9.198	124.371	Central Timor	
USNM579041	*Draco*	*timorensis*	East Timor	-9.198	124.371	Central Timor	
USNM579042	*Draco*	*timorensis*	East Timor	-9.198	124.371	Central Timor	X
USNM579043	*Draco*	*timorensis*	East Timor	-9.198	124.371	Central Timor	
USNM579298	*Draco*	*timorensis*	East Timor	-8.533	126.167	East Timor	
USNM579299	*Draco*	*timorensis*	East Timor	-9.01	125.65	East Timor	
USNM579300	*Draco*	*timorensis*	East Timor	-9.01	125.65	East Timor	
USNM579301	*Draco*	*timorensis*	East Timor	-9.316	125.25	Central Timor	X
USNM579302	*Draco*	*timorensis*	East Timor	-8.85	125.6	Central Timor	
USNM579303	*Draco*	*timorensis*	East Timor	-9.316	125.25	Central Timor	
USNM579304	*Draco*	*timorensis*	East Timor	-9.316	125.25	Central Timor	X
USNM579490	*Draco*	*timorensis*	East Timor	-8.433	126.967	East Timor	X
USNM579491	*Draco*	*timorensis*	East Timor	-8.35	127.05	East Timor	
USNM579492	*Draco*	*timorensis*	East Timor	-8.783	125.45	East Timor	
USNM579711	*Draco*	*timorensis*	East Timor	-8.55	125.533	East Timor	X

Table 2 cont.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
WAM107005	*Draco*	*timorensis*	Timor	-10.1666	123.6	West Timor	
WAM105619	*Draco*	*timorensis*	Rote	-10.7333	123.1	Rote	
WAM101714	*Draco*	*boschmai*	Sumba	-9.4333	119.35	East Sumba	
WAM104530	*Draco*	*boschmai*	Flores	-8.5833	120.5	West Flores	
WAM105108	*Draco*	*boschmai*	Lembata	-8.4333	123.3667	Lembata	
WAM105107	*Draco*	*boschmai*	Lembata	-8.4333	123.3667	Lembata	
WAM98623	*Draco*	*boschmai*	Sumbawa	-8.5833	117.2889	West Sumbawa	
LSUMZ81223	*Draco*	*beccarii*	Sulawesi	N/A	N/A	N/A	
LSUMZ81441	*Draco*	*volans*	Java	N/A	N/A	N/A	
TNHC56733	*Draco*	*sumatranus*	Sumatra	N/A	N/A	N/A	
JAM2079	*Draco*	*volans*	Java	N/A	N/A	N/A	X
JAM4281	*Draco*	*modigliani*	Enggano	N/A	N/A	N/A	X
JAM7032	*Draco*	*walkeri*	Sulawesi	N/A	N/A	N/A	X
JAM9054	*Draco*	*beccarii*	Sulawesi	N/A	N/A	N/A	X
SZL102	*Draco*	*sumatranus*	Sumatra	N/A	N/A	N/A	X

Table 3. Converted G-PHOCS demographic parameters for parapatrically distributed *Draco* lineages on Sumbawa and Flores islands.

		Mammalian mutation rate			***ND2* scaled mutation rate**		
		2.2 X 10^-9			4 X 10^-10		
Parameter	**Comparison**	**Mean**	**95% CI Low**	**95% CI High**	**Mean**	**95% CI Low**	**95% CI High**
Effective Population Size (Individuals)	West Sumbawa	2,388,636	2,302,273	2,477,273	13,137,500	12,662,502	13,625,002
	East Sumbawa	3,240,909	3,128,409	3,354,545	17,825,000	17,206,250	18,449,998
	Sumbawa Ancestor	293,182	276,136	311,364	1,612,501	1,518,748	1,712,502
Population Divergence Time (Years)	Sumbawa	1,159,090	1,140,909	1,254,455	6,375,000	6,275,000	6,899,503
Migration Rate (Migrants per generation)	West to East	8.1	6.0	9.9	8.1	6.0	9.9
	East to West	6.8	4.8	8.3	6.8	4.8	8.3
Effective Population Size (Individuals)	West Flores	3,567,045	3,392,045	3,744,318	19,618,748	18,656,248	20,593,749
	East Flores + Lembata	3,370,454	3,250,000	3,490,909	18,537,497	17,875,000	19,120,000
	Flores Ancestor	344,318	325,000	360,227	1,893,749	1,787,500	1,981,249
Population Divergence Time (Years)	Flores	1,040,909	1,022,727	1,059,091	5,725,000	5,625,000	5,825,000
Migration Rate (Migrants per generation)	West to East	9.4	3.8	15.2	9.4	3.8	15.2
	East to West	61.7	48.8	75.7	61.7	48.8	75.7

FIGURE LEGENDS

Figure 1. Flying lizards from the Lesser Sunda Islands. Upper left panel demonstrates the cryptic coloration against tree bark. Upper right panels show males with the dewlap retracted and extended. The lower four panels demonstrate some of the variation in color pattern of the underside of the patagia. (Photos: S. Reilly)

Figure 2. Distribution map of *Draco* species within the Lesser Sunda Islands. Some islands are shaded gray because there are no flying lizards recorded (Nusa Penida, Sabu, Atauro) and other islands have flying lizards but they are not assigned to either *D. boschmai* or *D. timoriensis* (Solor, Pantar).

Figure 3. Dot localities of *Draco* samples used in the exon-capture experiment.

Figure 4. Maximum Likelihood phylogeny of the *ND2* mitochondrial gene produced by the software package RAXML. Numbers at nodes represent bootstrap support. Colored shapes next to each clade name correspond to localities in Figure 6.

Figure 5. Time-calibrated Bayesian phylogeny of the *ND2* mitochondrial gene produced by the software package BEAST. Branch lengths are in units of time and node bars represent 95% confidence intervals. Numbers at nodes represent posterior probability support. Colored shapes next to each clade name correspond to localities in Figure 6.

Figure 6. Localities of major mitochondrial clades as shown in Figures 4-5.

Figure 7. Average coverage of the 709 loci targeted in the exon-capture experiment. Coverage refers to the number of unique Illumina sequence reads that map to any given site after duplicate reads are removed. The target regions correspond to the exonic sequences derived from transcriptomes for which the probes were designed to bind. The flanking regions are the introns that lie on either side of each exon.

Figure 8. The average coverage of the targeted exons for each library.

Figure 9. The average coverage of the flanking regions for each library.

Figure 10. Alignment length summary plots for the 709 nuclear loci plotting the number of individuals (A), the number of informative sites (B), and the percentage of gaps (C) for each alignment length.

Figure 11. Frequency bar plots showing the number of taxa per alignment (A), the distribution of sequence lengths (B), the percent of informative sites per gene (C), and the percent missing data per gene (D).

Figure 12. Principle components analysis of genetic co-variance from the 709-gene dataset. Colors of each dot correspond to the islands/regions colored in the map. A) PC1 vs PC2, B) PC1 vs PC3, C) PC2 vs PC3, and D) the percent of variance explained by each component.

Figure 13. Maximum Likelihood phylogeny of the concatenated nuclear dataset produced in RAxML. Numbers at nodes represent bootstrap support and colored bars for clades correspond to the map.

Figure 14. Supertree of 709 individual RAxML gene trees produced by ASTRAL. Colors correspond to the map above.

Figure 15. Robinson-Foulds % distances for ASTRAL supertrees using subsets of 5, 10, 20, 50, 100, 200, and 500 loci replicated 10 times each. The lower the Robinson-Foulds % distance, the more similar the topology is to the full 709 gene ASTRAL supertree.

Figure 16. SNP-based species tree produced by SNAPP utilizing one informative SNP per gene.

Figure 17. Comparisons of STRUCTURE results from K=1 up to K=20 shown as A) Delta K values and B) the mean of the Ln probability of the data.

Figure 18. STRUCTURE population clustering results for K=2 up to K=9.

Figure 19. Unconverted demographic parameter estimate distributions produced by G-PHOCS analyses of the flanking sequence data. A) The population divergence time estimates (tau) for the split between Western and Eastern Sumbawa populations, B) and for Western Flores and Eastern Flores + Lembata populations (red bars represent the 95% confidence intervals). C) The effective population size estimates for the Sumbawa ancestor (red), the West Sumbawa population (black), and the East Sumbawa population (purple). D) The effective population size estimates for the Flores + Lembata ancestor (red), the Western Flores population (black), and the East Flores + Lembata population (purple). E) Migration rates from East Sumbawa to West Sumbawa (black) and from West Sumbawa to East Sumbawa (purple). F) Migration rates from East Flores + Lembata to West Flores (purple) and from West Flores to East Flores + Lembata (black).

Figure 20. A biogeographical hypothesis for the colonization of the Lesser Sunda Islands by *Draco volans* from Bali derived from the topology of the 709-gene phylogenies. Arrows indicate a dispersal event, dashed lines indicate a vicariance event, and shaded islands indicate islands that are uninhabited by *Draco* at that time period.

FIGURES

Figure 1.

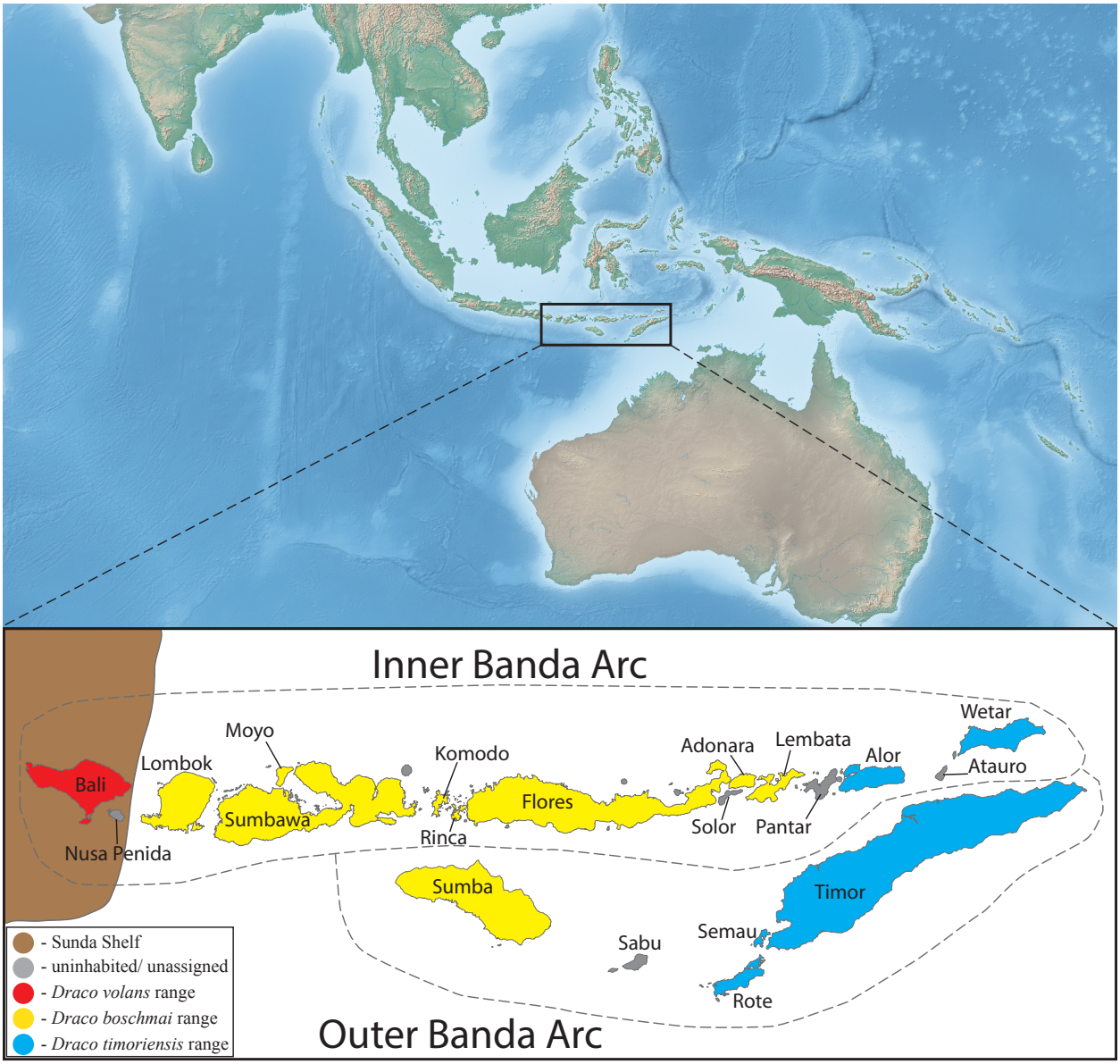

Inner Banda Arc
Moyo
Komodo
Adonara
Lembata
Wetar
Lombok
Alor
Atauro
Bali
Sumbawa
Flores
Solor
Pantar
Rinca
Nusa Penida
Sumba
Timor
Semau
Sabu
Rote
- Sunda Shelf
- uninhabited/ unassigned
- Draco volans range
- Draco boschmai range
- Draco timoriensis range
Outer Banda Arc

Figure 2.

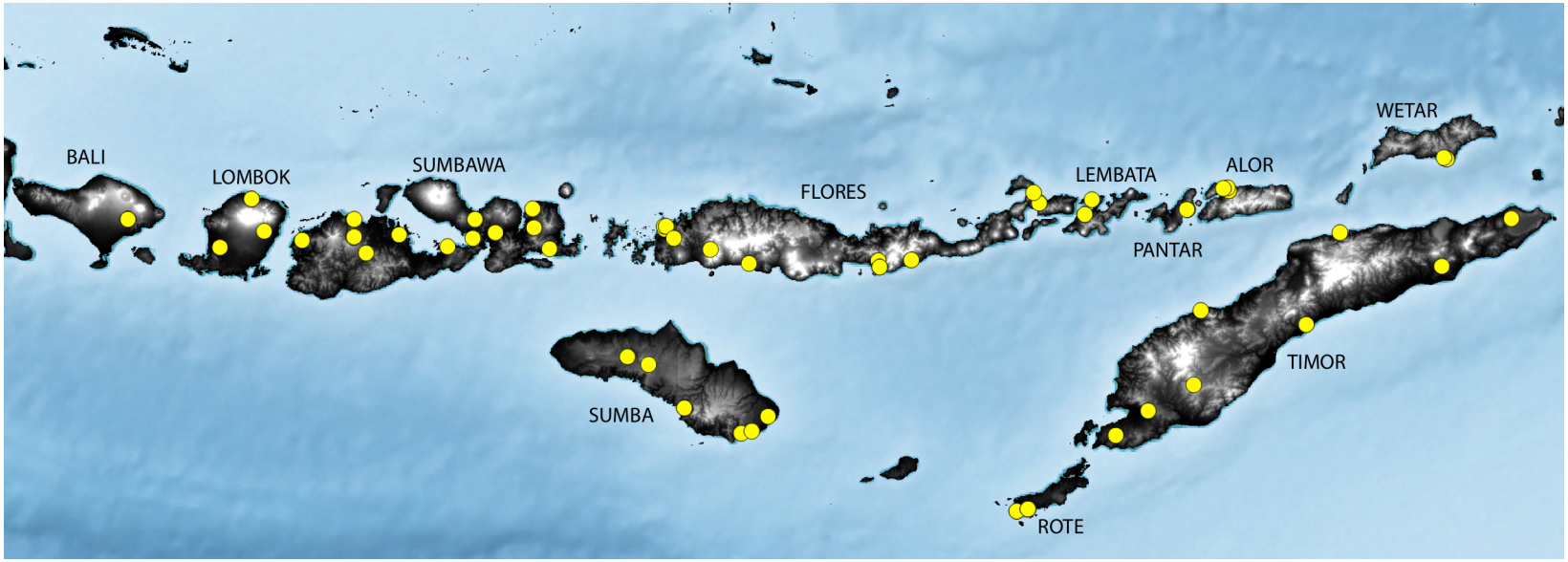

BALI
LOMBOK
SUMBAWA
FLORES
LEMBATA
ALOR
WETAR
PANTAR
TIMOR
SUMBA
ROTE

Figure 3.

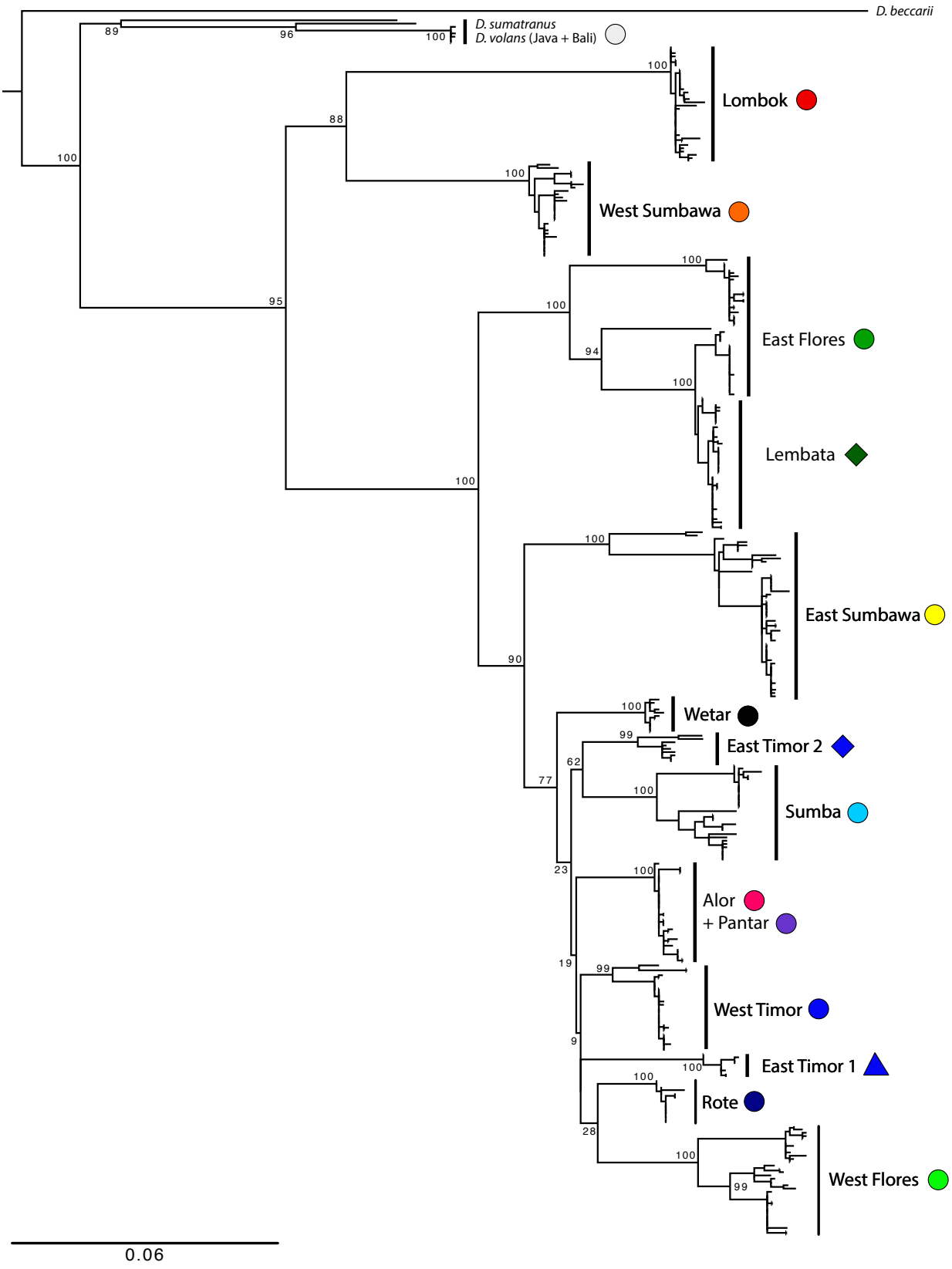

D. beccarii
D. sumatranus
D. volans (Java + Bali)
Lombok
West Sumbawa
East Flores
Lembata
East Sumbawa
Wetar
East Timor 2
Sumba
Alor
+ Pantar
West Timor
East Timor 1
Rote
West Flores
0.06

Figure 4.

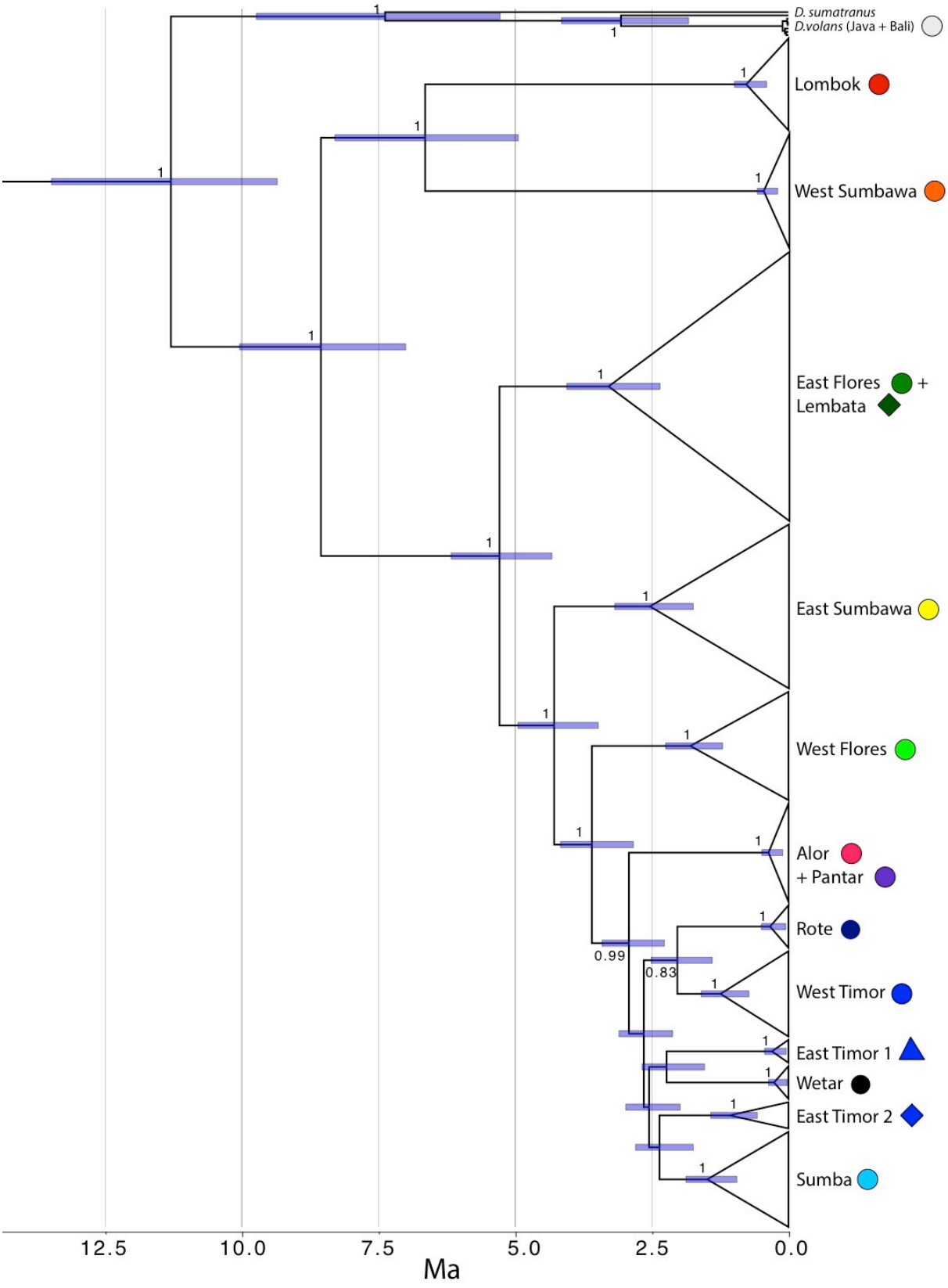

D. sumatranus
D.volans (Java + Bali)
Lombok
West Sumbawa
East Flores +
Lembata
East Sumbawa
West Flores
Alor
+ Pantar
Rote
West Timor
East Timor 1
Wetar
East Timor 2
Sumba
0.99
0.83
12.5
10.0
7.5
5.0
2.5
0.0
Ma

Figure 5.

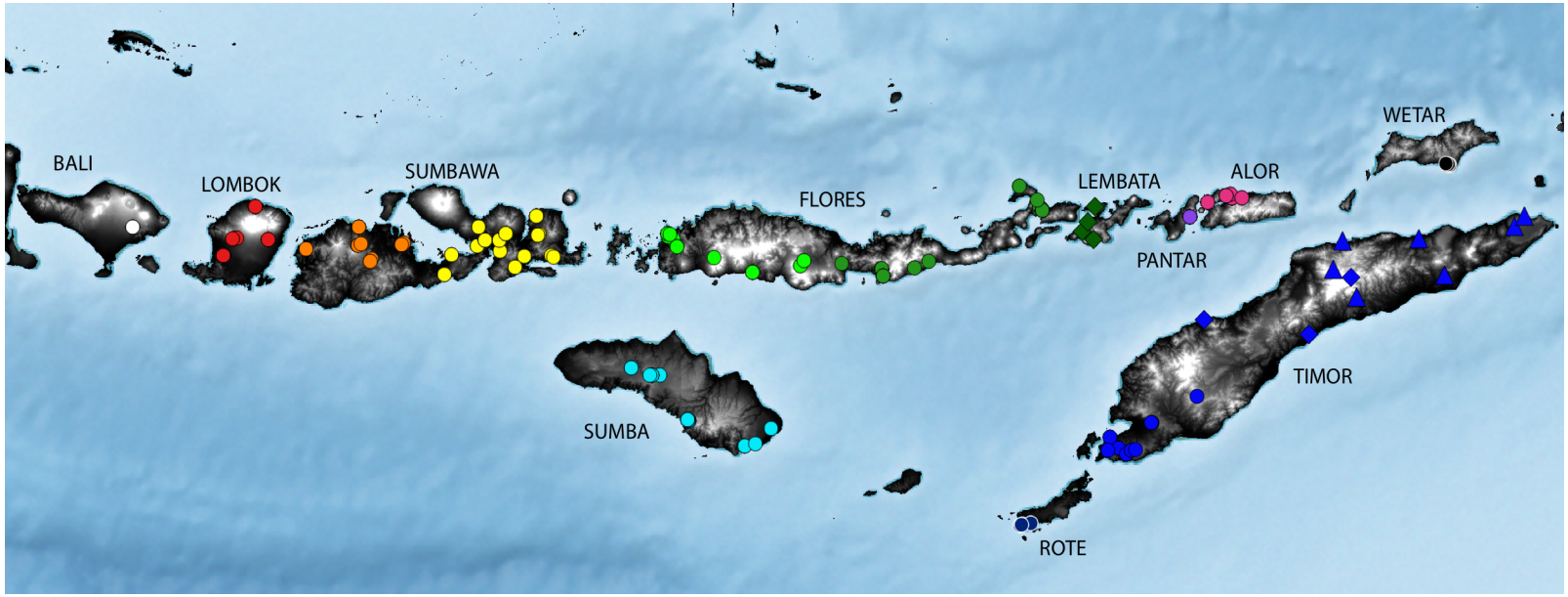

BALI
LOMBOK
SUMBAWA
FLORES
LEMBATA
ALOR
WETAR
PANTAR
TIMOR
SUMBA
ROTE

Figure 6.

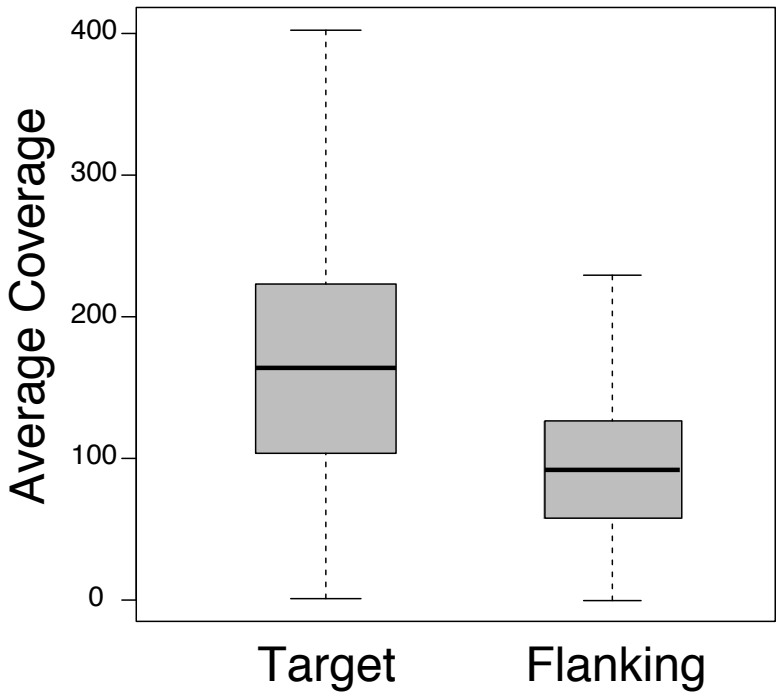


Figure 7.

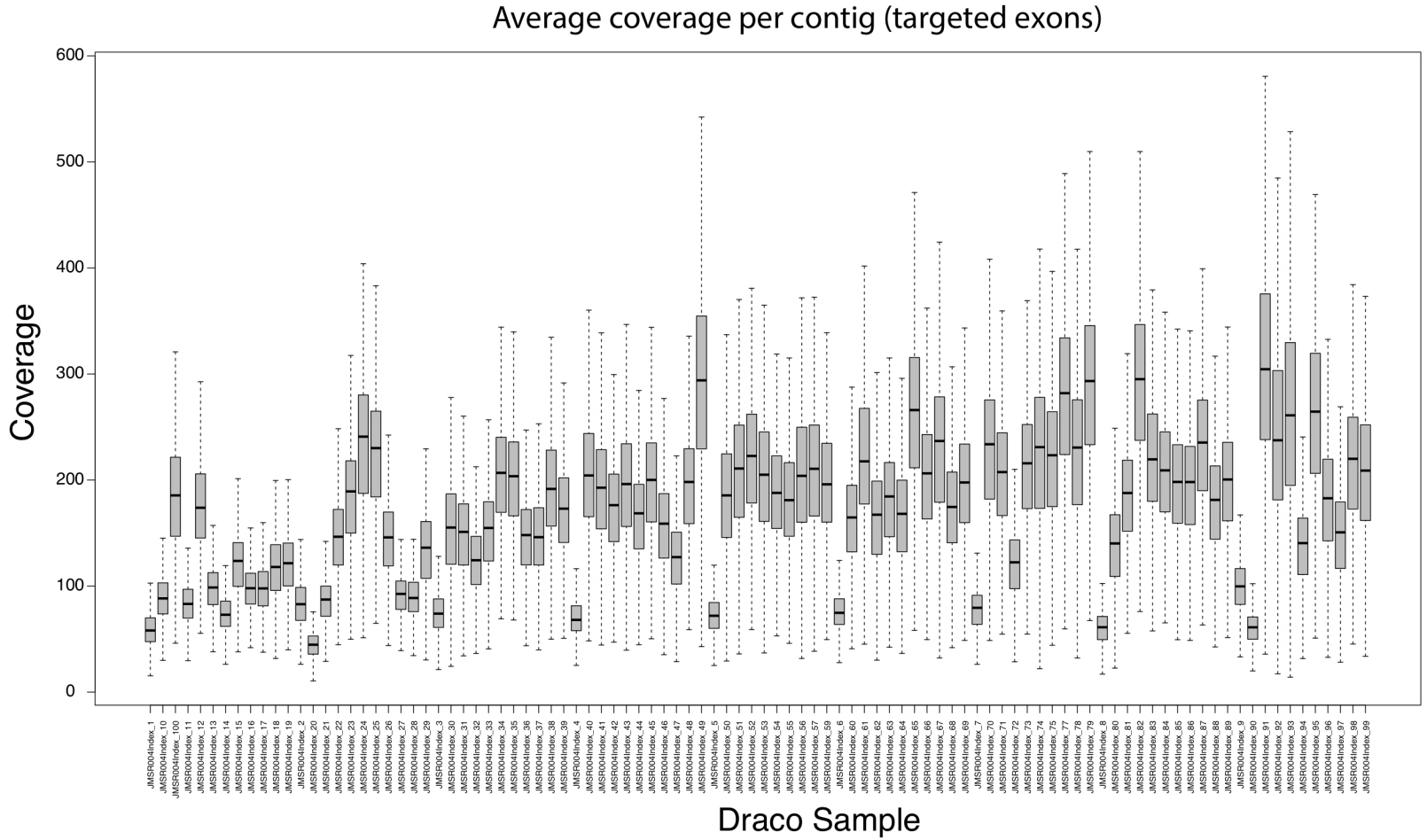

Average coverage per contig (targeted exons)
Coverage
600
500
400
300
200
100
0
Draco Sample

Figure 8.

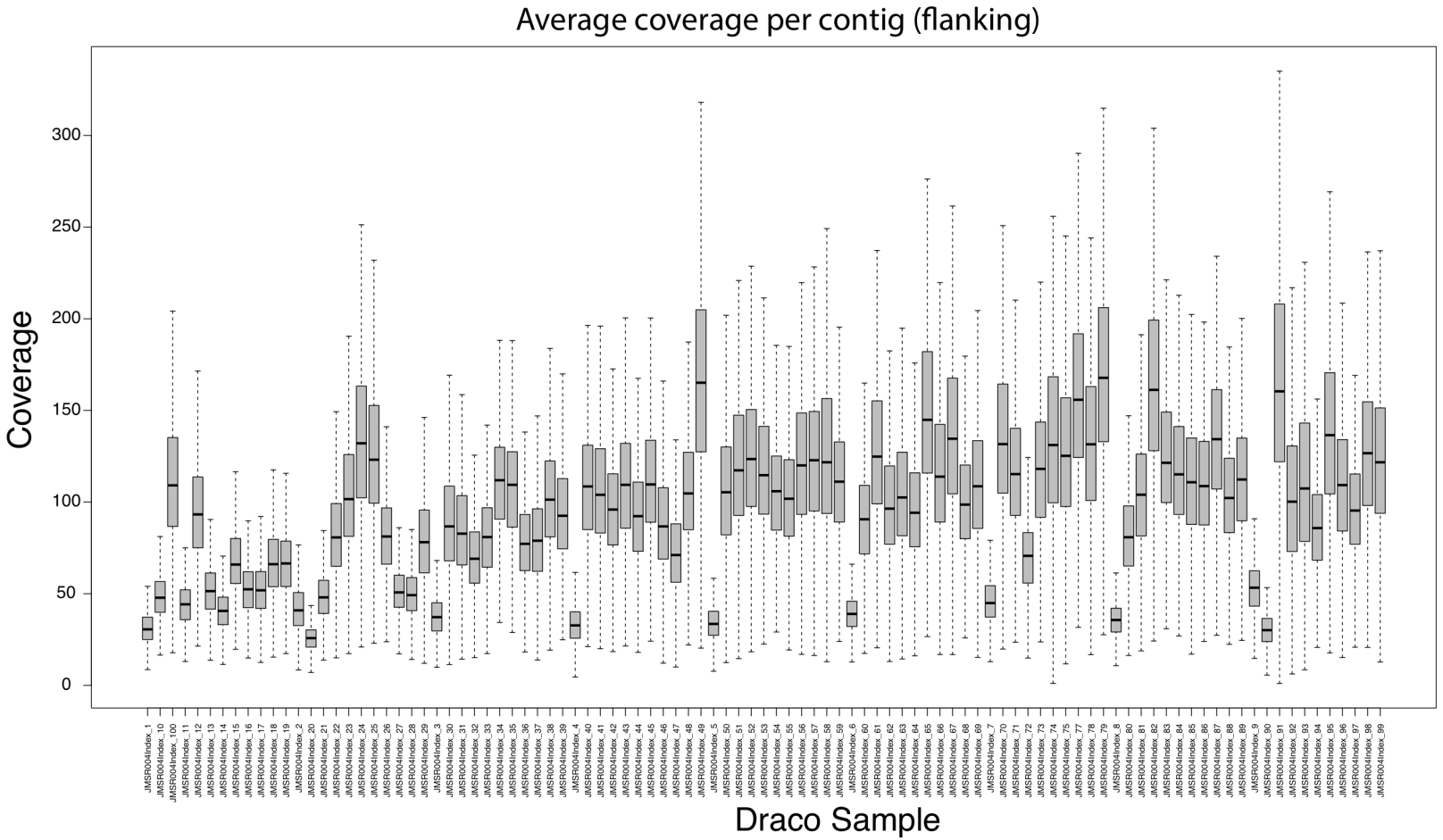

Average coverage per contig (flanking)
Coverage
0
50
100
150
200
250
300
Draco Sample

Figure 8.

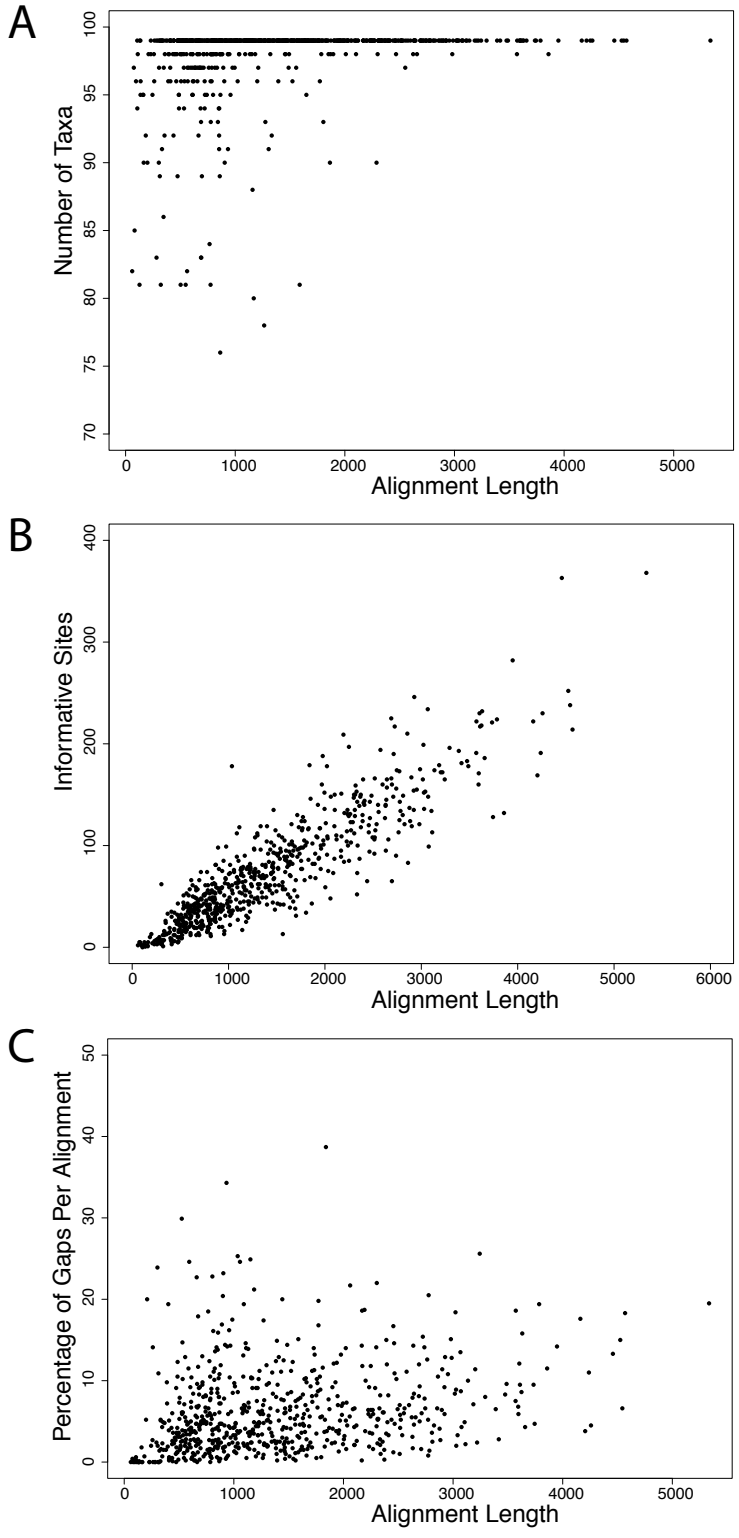

A
Number of Taxa
70
75
80
85
90
95
100
0
1000
2000
3000
4000
5000
Alignment Length
B
Informative Sites
0
100
200
300
400
0
1000
2000
3000
4000
5000
6000
Alignment Length
C
Percentage of Gaps Per Alignment
0
10
20
30
40
50
0
1000
2000
3000
4000
5000
Alignment Length

Figure 10.

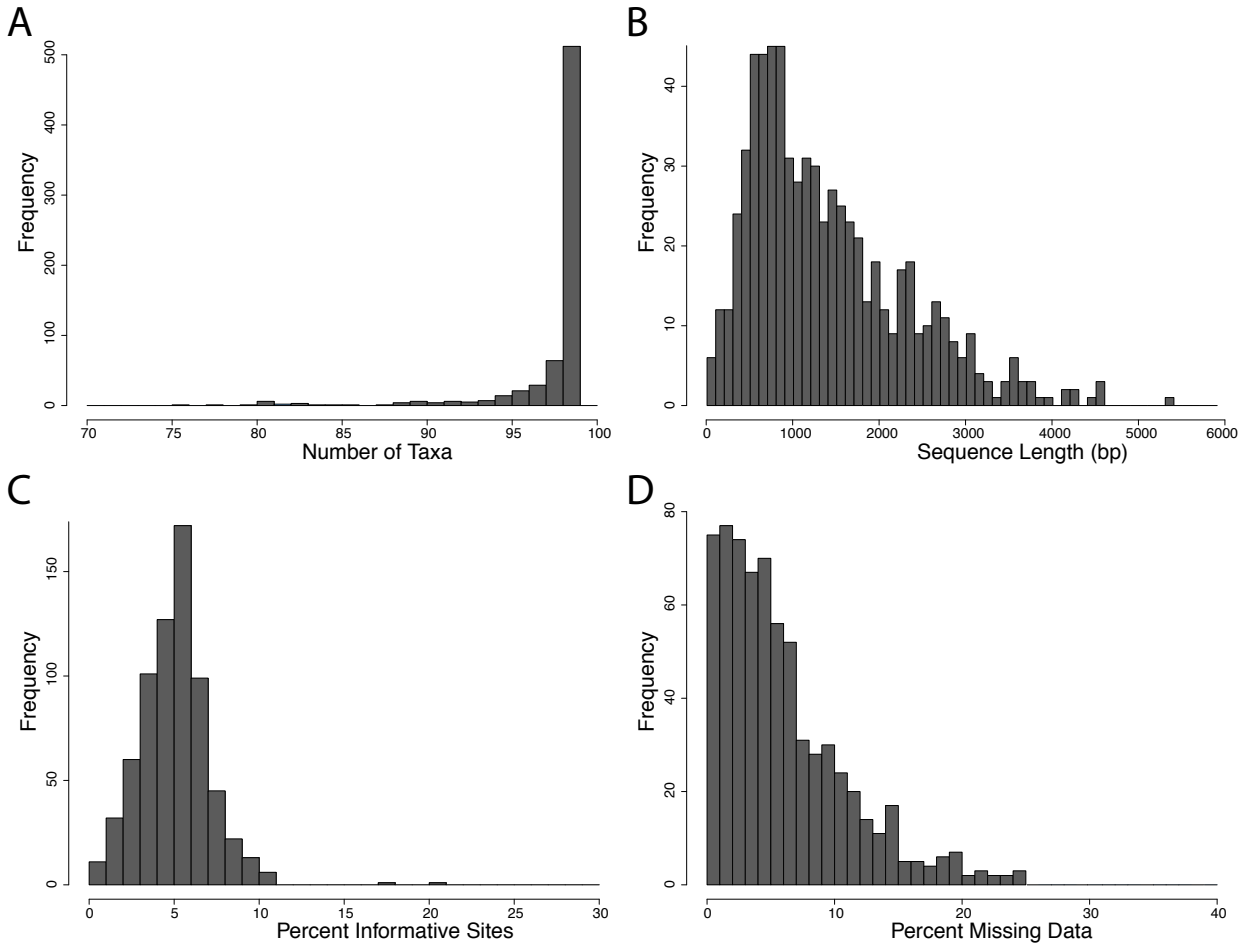

A
Frequency
0
100
200
300
400
500
70
75
80
85
90
95
100
Number of Taxa
B
Frequency
0
10
20
30
40
0
1000
2000
3000
4000
5000
6000
Sequence Length (bp)
C
Frequency
0
50
100
150
0
5
10
15
20
25
30
Percent Informative Sites
D
Frequency
0
20
40
60
80
0
10
20
30
40
Percent Missing Data

Figure 11.

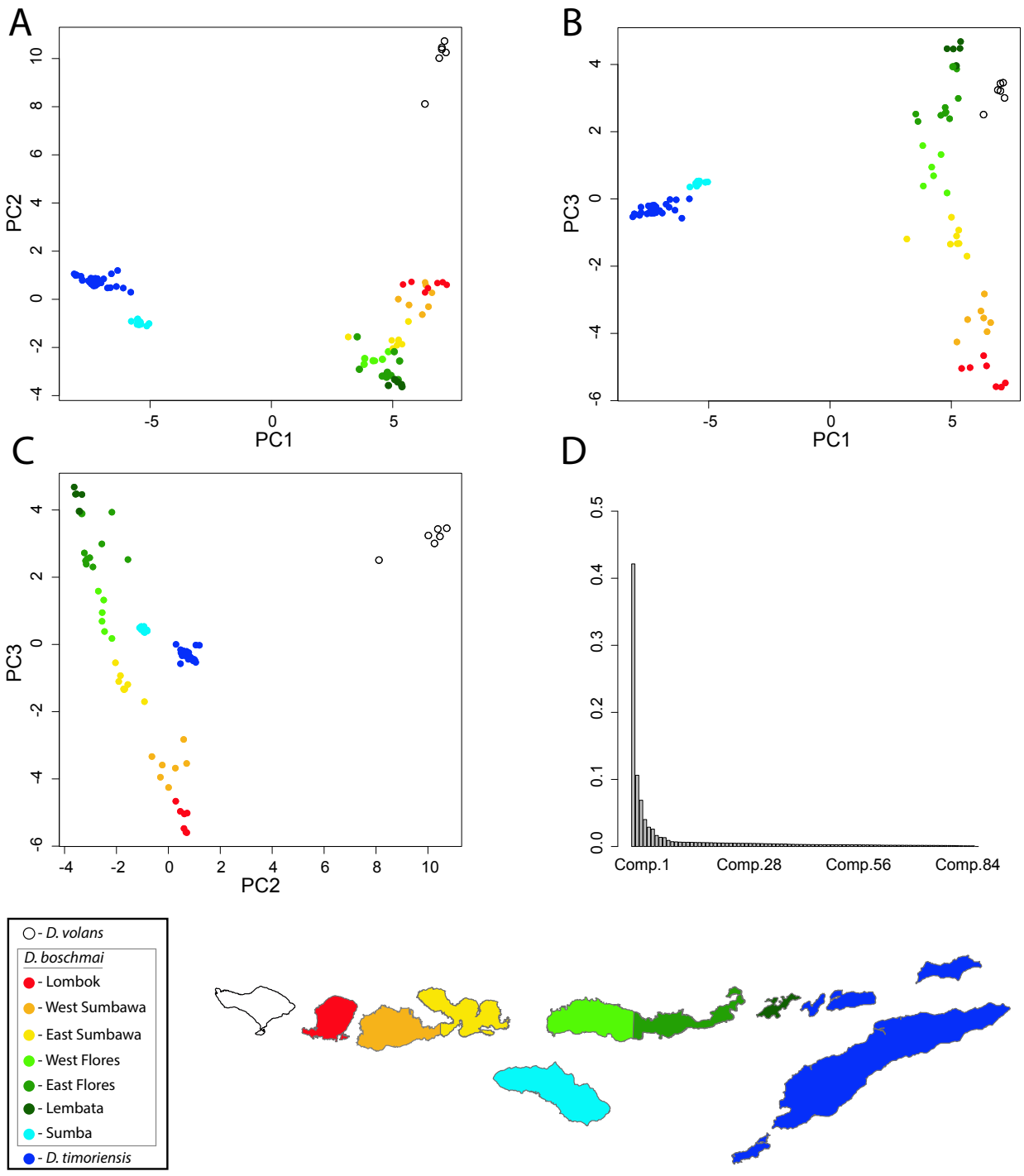

A
PC2
PC1
B
PC3
PC1
C
PC3
PC2
D
Comp.1
Comp.28
Comp.56
Comp.84
- D. volans
D. boschmai
- Lombok
- West Sumbawa
- East Sumbawa
- West Flores
- East Flores
- Lembata
- Sumba
- D. timoriensis

Figure 12.

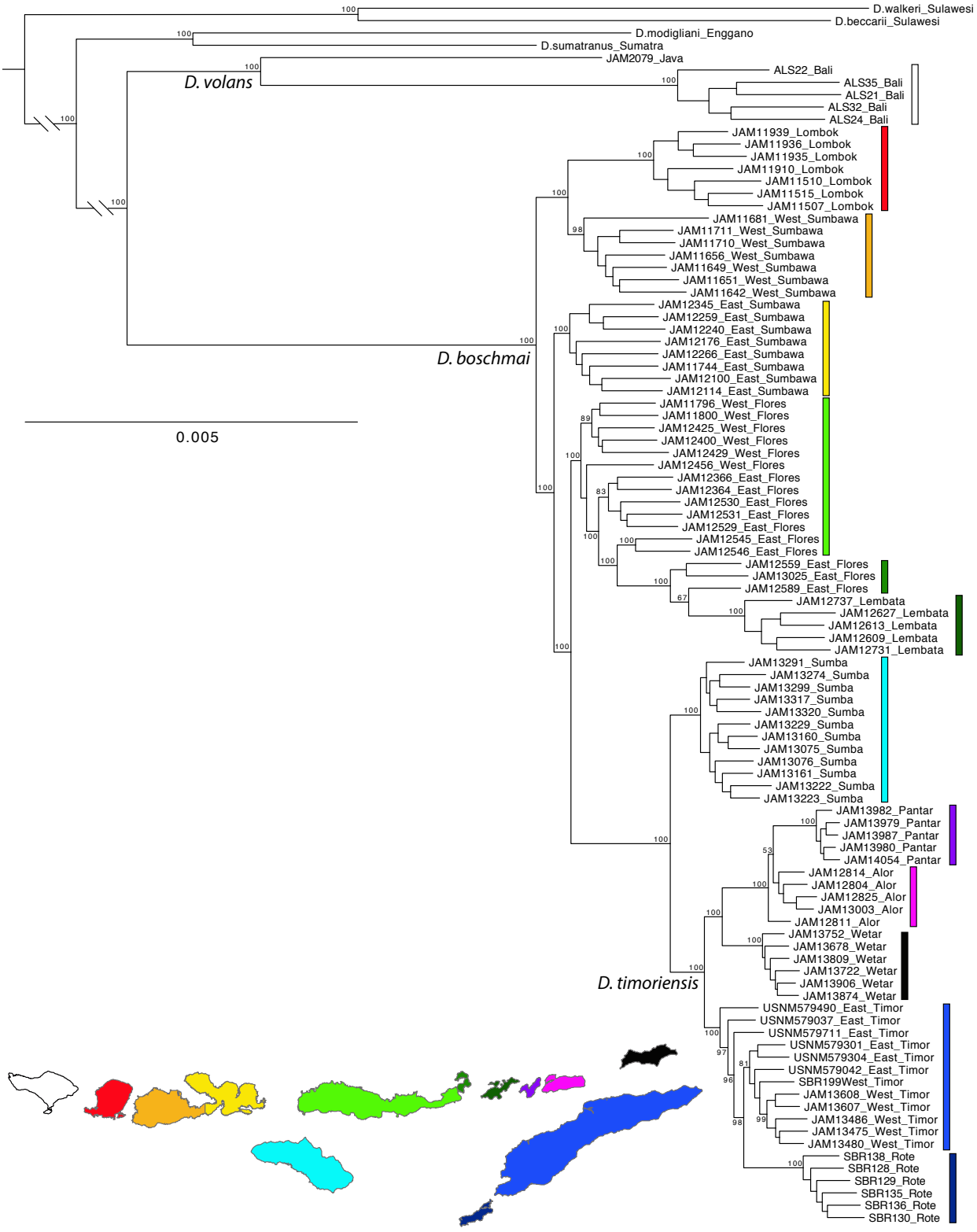

D.walkeri_Sulawesi
D.beccarii_Sulawesi
D.modigliani_Enggano
D.sumatranus_Sumatra
JAM2079_Java
D. volans
ALS22_Bali
ALS35_Bali
ALS21_Bali
ALS32_Bali
ALS24_Bali
JAM11939_Lombok
JAM11936_Lombok
JAM11935_Lombok
JAM11910_Lombok
JAM11510_Lombok
JAM11515_Lombok
JAM11507_Lombok
JAM11681_West_Sumbawa
JAM11711_West_Sumbawa
JAM11710_West_Sumbawa
JAM11656_West_Sumbawa
JAM11649_West_Sumbawa
JAM11651_West_Sumbawa
JAM11642_West_Sumbawa
JAM12345_East_Sumbawa
JAM12259_East_Sumbawa
JAM12240_East_Sumbawa
JAM12176_East_Sumbawa
JAM12266_East_Sumbawa
JAM11744_East_Sumbawa
JAM12100_East_Sumbawa
JAM12114_East_Sumbawa
D. boschmai
0.005
JAM11796_West_Flores
JAM11800_West_Flores
JAM12425_West_Flores
JAM12400_West_Flores
JAM12429_West_Flores
JAM12456_West_Flores
JAM12366_East_Flores
JAM12364_East_Flores
JAM12530_East_Flores
JAM12531_East_Flores
JAM12529_East_Flores
JAM12545_East_Flores
JAM12546_East_Flores
JAM12559_East_Flores
JAM13025_East_Flores
JAM12589_East_Flores
JAM12737_Lembata
JAM12627_Lembata
JAM12613_Lembata
JAM12609_Lembata
JAM12731_Lembata
JAM13291_Sumba
JAM13274_Sumba
JAM13299_Sumba
JAM13317_Sumba
JAM13320_Sumba
JAM13229_Sumba
JAM13160_Sumba
JAM13075_Sumba
JAM13076_Sumba
JAM13161_Sumba
JAM13222_Sumba
JAM13223_Sumba
JAM13982_Pantar
JAM13979_Pantar
JAM13987_Pantar
JAM13980_Pantar
JAM14054_Pantar
JAM12814_Alor
JAM12804_Alor
JAM12825_Alor
JAM13003_Alor
JAM12811_Alor
JAM13752_Wetar
JAM13678_Wetar
JAM13809_Wetar
JAM13722_Wetar
JAM13906_Wetar
JAM13874_Wetar
D. timoriensis
USNM579490_East_Timor
USNM579037_East_Timor
USNM579711_East_Timor
USNM579301_East_Timor
USNM579304_East_Timor
USNM579042_East_Timor
SBR199West_Timor
JAM13608_West_Timor
JAM13607_West_Timor
JAM13486_West_Timor
JAM13475_West_Timor
JAM13480_West_Timor
SBR138_Rote
SBR128_Rote
SBR129_Rote
SBR135_Rote
SBR136_Rote
SBR130_Rote

Figure 13.

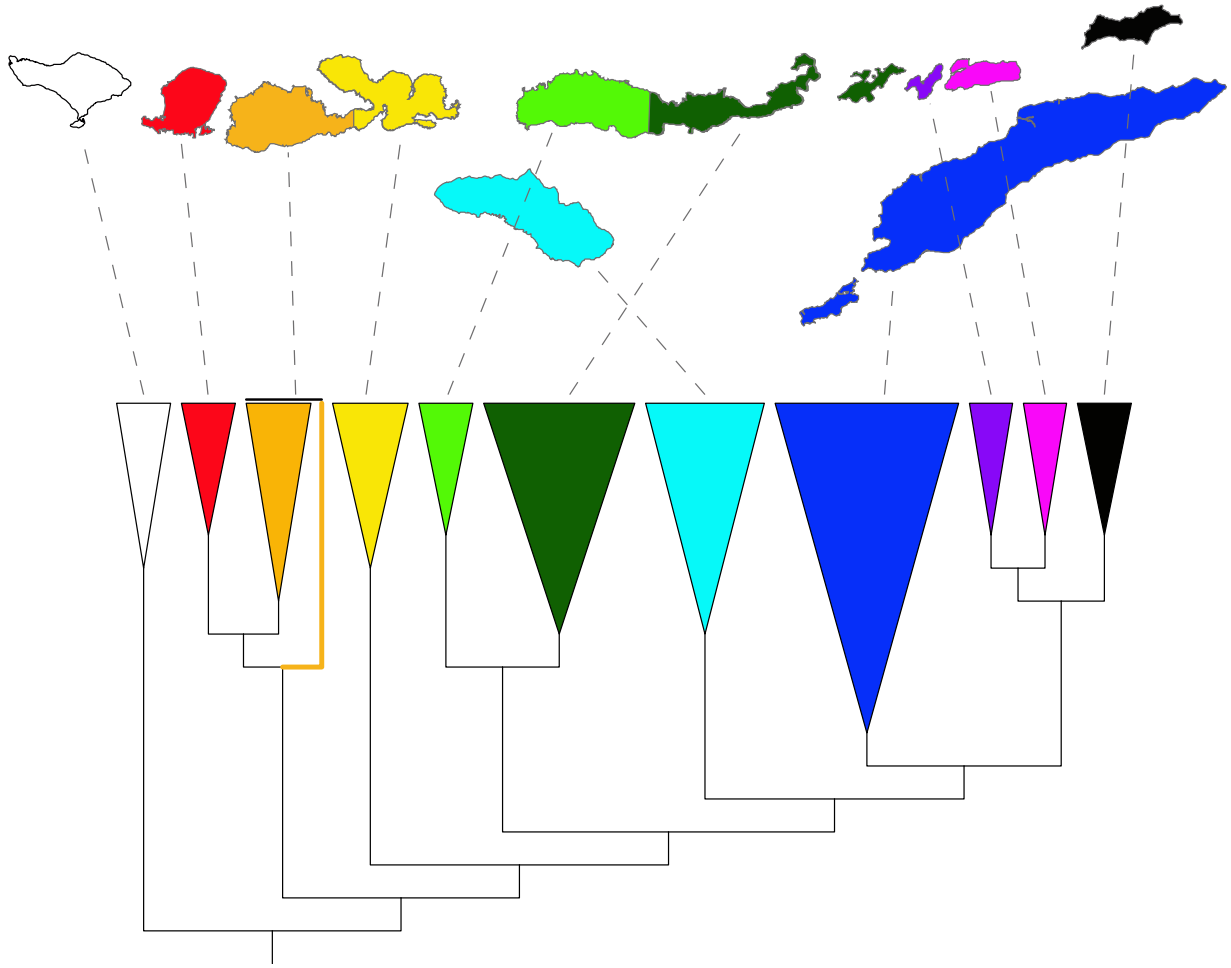


Figure 14.

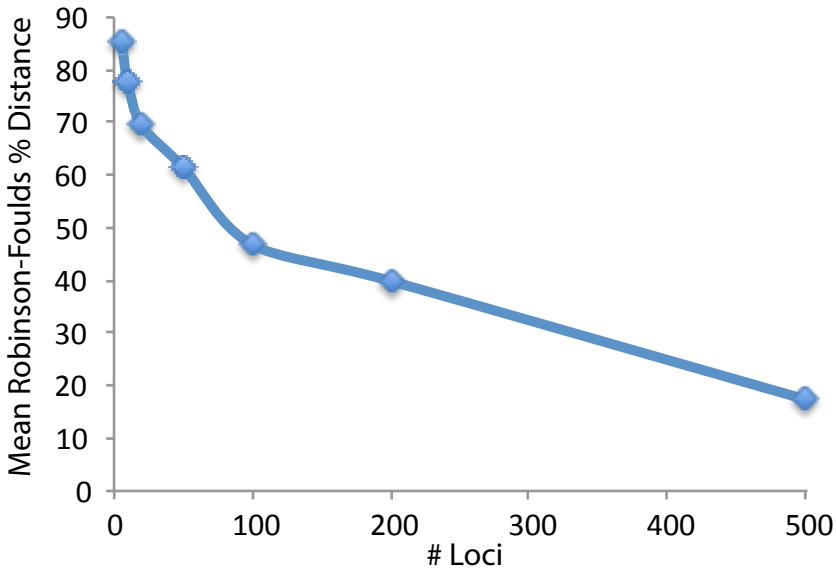

Mean Robinson-Foulds % Distance
90
80
70
60
50
40
30
20
10
0
0
100
200
300
400
500
Loci

Figure 15.

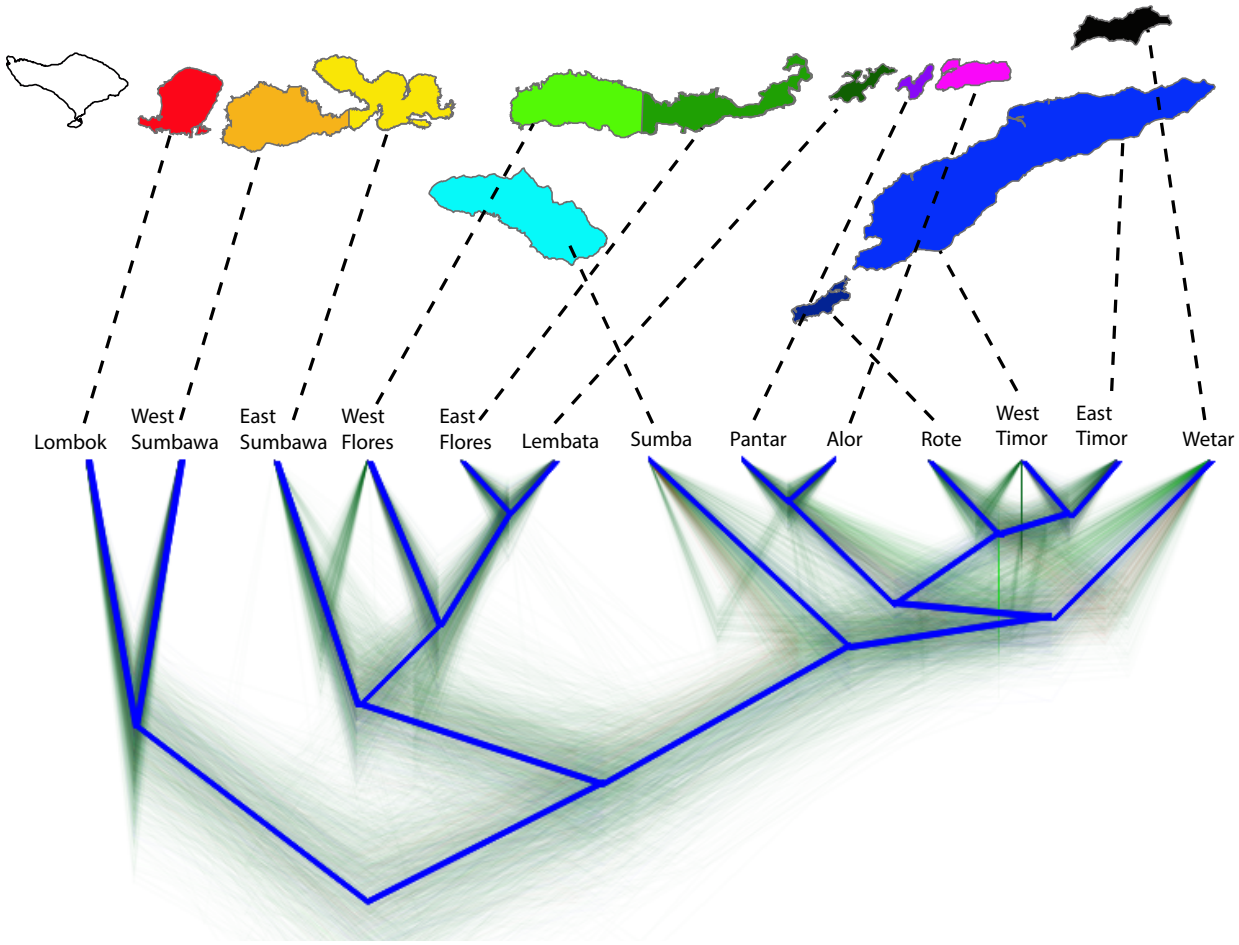

Lombok
West Sumbawa
East Sumbawa
West Flores
East Flores
Lembata
Sumba
Pantar
Alor
Rote
West Timor
East Timor
Wetar

Figure 16.

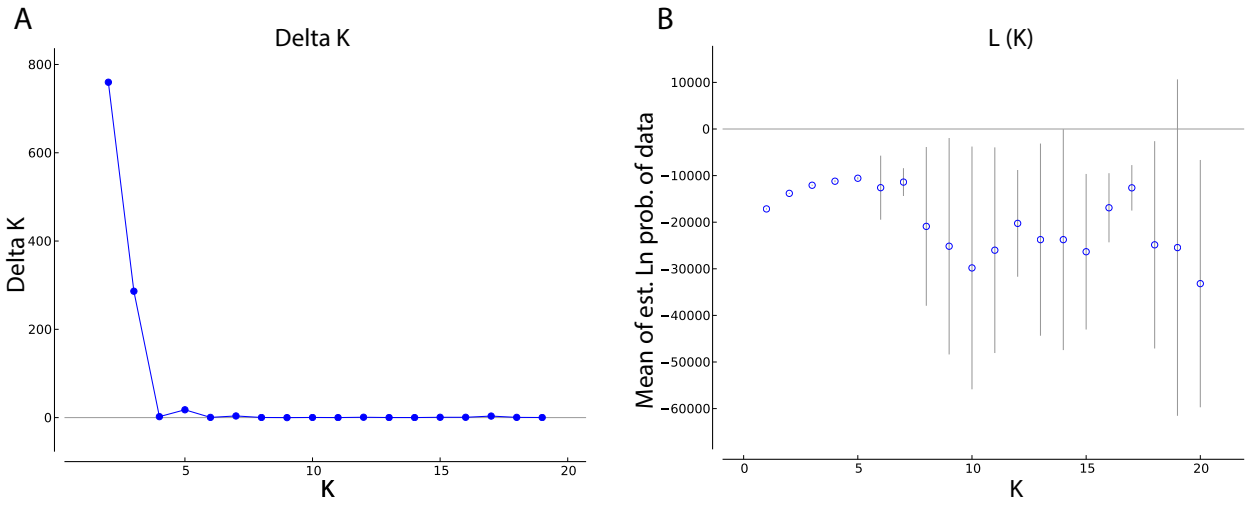

A
Delta K
Delta K
800
600
400
200
0
5
10
15
20
K
B
L (K)
Mean of est. Ln prob. of data
10000
0
−10000
−20000
−30000
−40000
−50000
−60000
0
5
10
15
20
K

Figure 17.

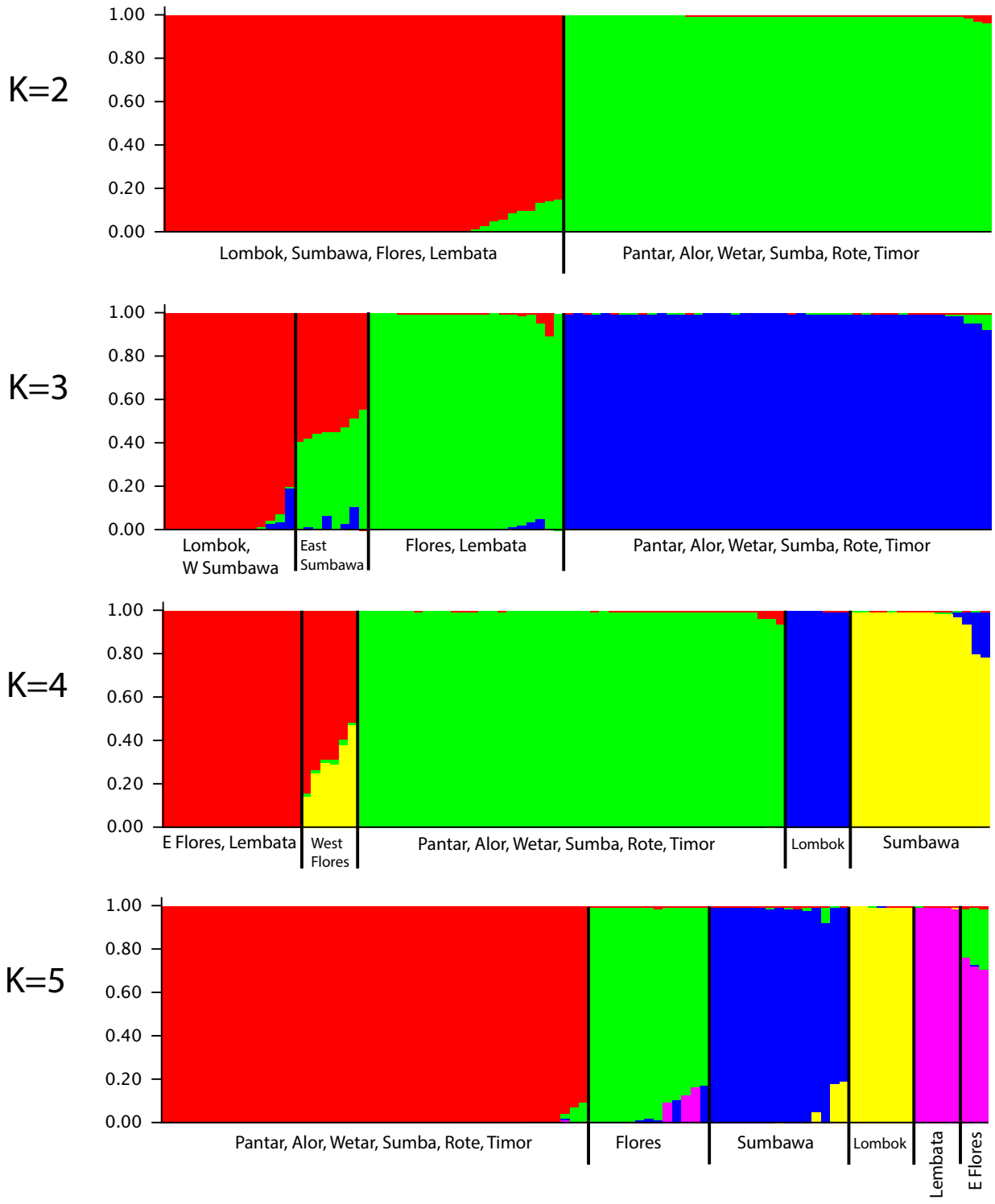

K=2
1.00
0.80
0.60
0.40
0.20
0.00
Lombok, Sumbawa, Flores, Lembata
Pantar, Alor, Wetar, Sumba, Rote, Timor
K=3
1.00
0.80
0.60
0.40
0.20
0.00
Lombok, W Sumbawa
East Sumbawa
Flores, Lembata
Pantar, Alor, Wetar, Sumba, Rote, Timor
K=4
1.00
0.80
0.60
0.40
0.20
0.00
E Flores, Lembata
West Flores
Pantar, Alor, Wetar, Sumba, Rote, Timor
Lombok
Sumbawa
K=5
1.00
0.80
0.60
0.40
0.20
0.00
Pantar, Alor, Wetar, Sumba, Rote, Timor
Flores
Sumbawa
Lombok
Lembata
E Flores

Figure 18.

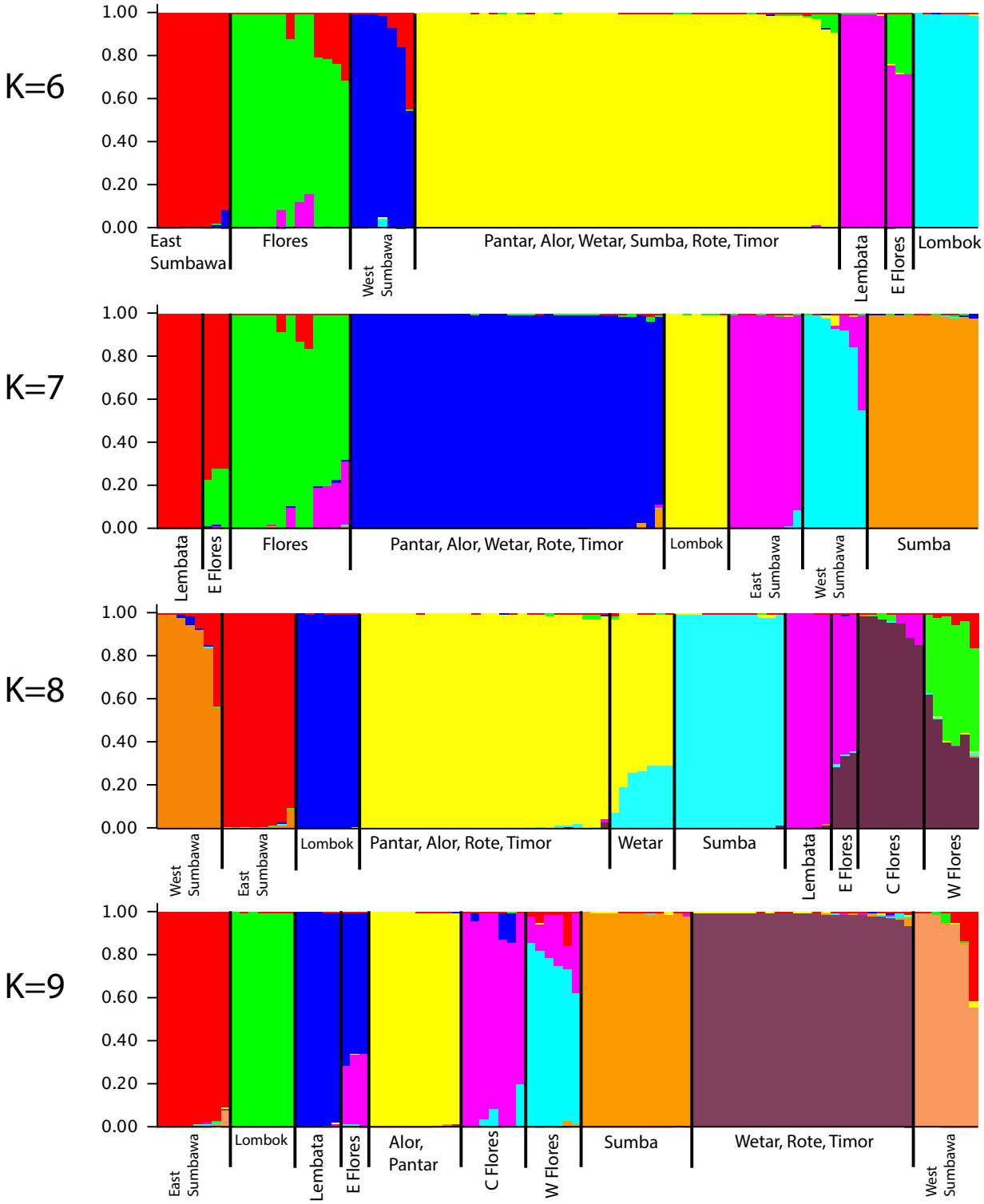


Figure 18 cont.

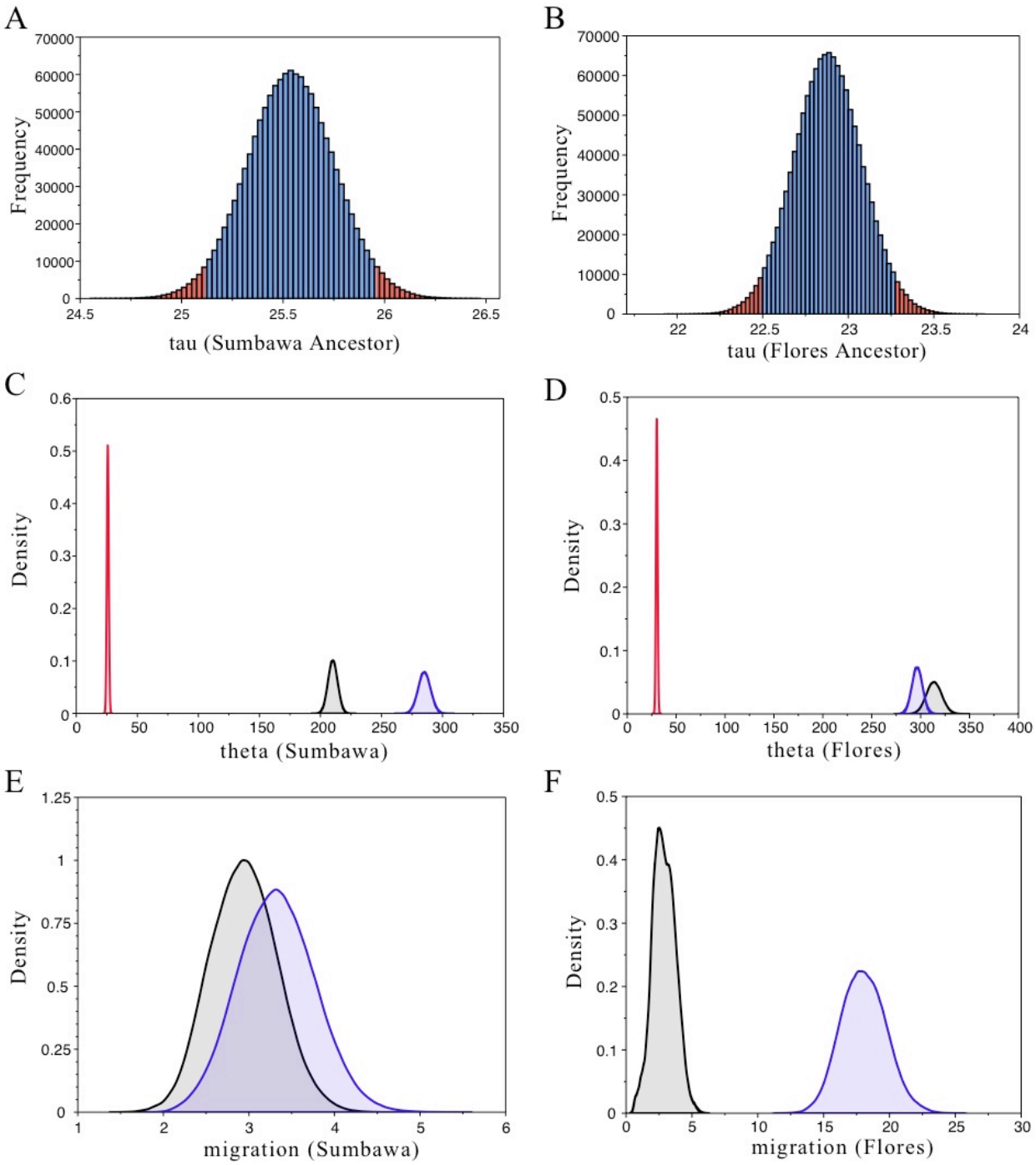

A
Frequency
tau (Sumbawa Ancestor)
B
Frequency
tau (Flores Ancestor)
C
Density
theta (Sumbawa)
D
Density
theta (Flores)
E
Density
migration (Sumbawa)
F
Density
migration (Flores)

Figure 19.

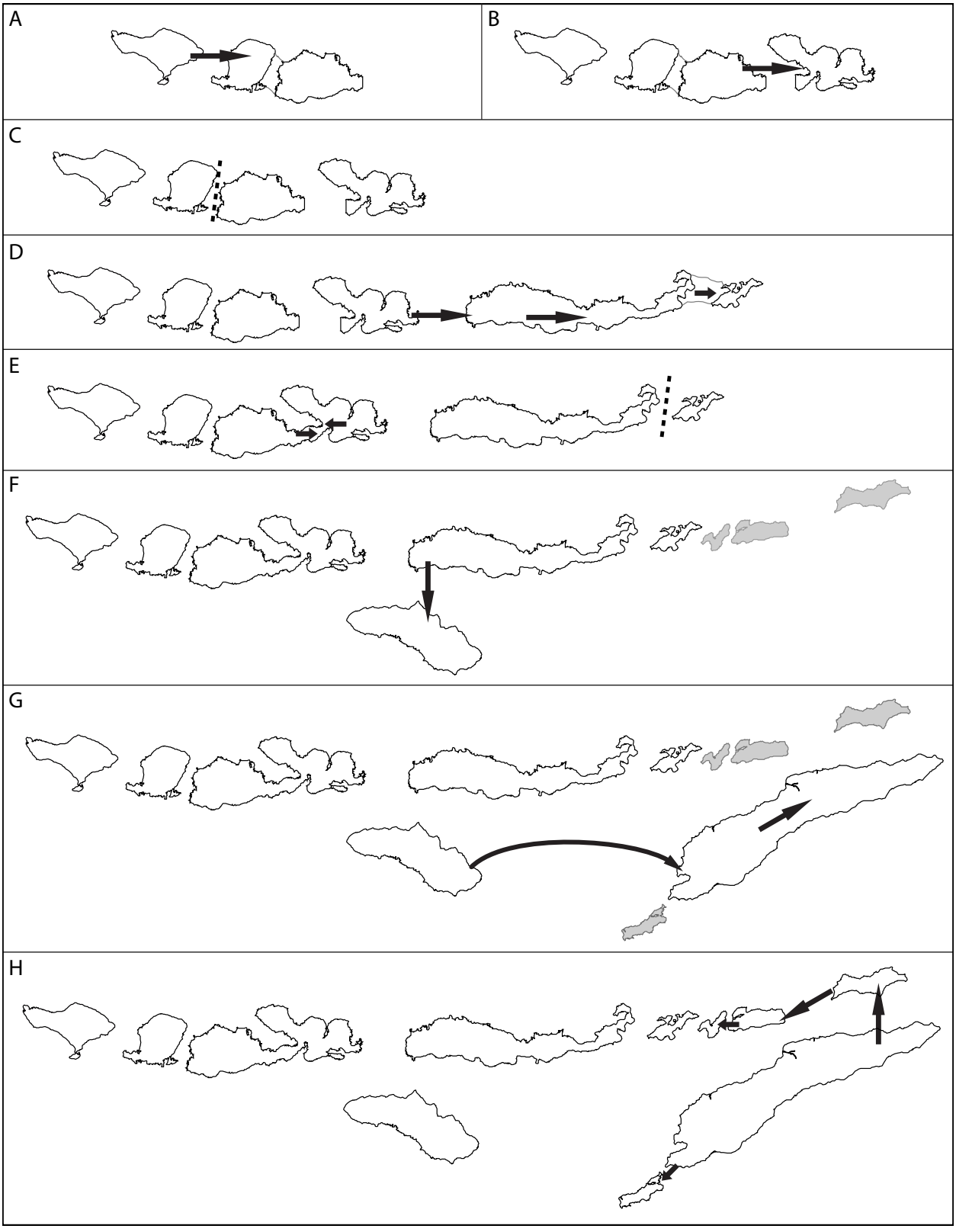


Figure 20.

CHAPTER 3

Biogeographical History of Forest Skinks (Genus: *Sphenomorphus*) From the Lesser Sunda Islands of Indonesia

ABSTRACT

The Lesser Sunda Islands of Indonesia are an archipelago of tropical oceanic islands that harbor a highly endemic faunal community. The geological history of this archipelago is extremely complex and still poorly understood with respect to details relevant for biogeographical studies. The lizard species *Sphenomorphus melanopogon* (Flores Forest Skink) has a range that includes all of the major Lesser Sunda Islands as well as a few small islands off of West Java and many of the Banda Arc islands of Maluku province. Previous examination of the morphological variation of *S. melanopogon* showed variation in color pattern and size between populations and suggested that Flores Island might contain multiple independently evolving lineages. The biogeographical history of *S. melanopogon* was initially inferred using phylogenetic analysis of mitochondrial DNA (*ND4*) from 237 specimens from the Lesser Sundas. The results showed deep divergences between lineages (up to 20 million years), with some islands containing multiple non-sister lineages that are parapatrically or sympatrically distributed. Transcriptome sequencing of three individuals identified thousands of nuclear genes, 1,200 of which were chosen for probe development based on their level of interspecific variation useful for phylogenetic analysis. Using the mtDNA phylogeny as a guide, 104 *S. melanopogon* samples representing all the major mtDNA clades were chosen for an exon-capture experiment targeting the ~1,200 nuclear genes. This approach produced high coverage data with low levels missing data, as well as a large amount of non-coding flanking sequence data that was used for phylogenetic and demographic analyses. Phylogenomic analysis revealed that nearly all mtDNA lineages were also monophyletic nuclear lineages with the exception of samples from Timor Island, which were found to represent multiple lineages in the mtDNA tree but a single monophyletic grouping in the nuclear tree. The relationships of these major nDNA lineages differs from the mtDNA tree, and there is strong genetic structuring of populations up to an assumed number of 9 populations (or species). Demographic analysis suggests deep divergence times and little to no gene flow between parapatrically and sympatrically occurring lineages on Lombok, Flores, and Sumba Islands, suggesting that these lineages represent distinct species. The patterns of island colonization are complex and provide insight into dispersal patterns during the early stages of the archipelago's geological development. Overall, the islands may be older than is currently estimated, and some islands such as Flores may have represented multiple islands in the past that have since become merged allowing for secondary contact between long-diverged insular lineages. A re-examination of the morphological variation between the nuclear DNA lineages described here will be needed to formally describe the many species within this complex. This study has shown that the biogeographical history of the Lesser Sundas archipelago is extremely complex and that biodiversity within this region is vastly underestimated.

INTRODUCTION

Oceanic islands are ideal systems for studying the evolution and diversification of life because their remote nature allows for the isolation and subsequent divergence of populations (Whittaker & Fernandez-Palacios, 2007; Losos & Ricklefs, 2010). Island archipelagos are clusters of islands that usually share a similar geological origin. Populations of a particular species that have colonized multiple islands within an archipelago often have different evolutionary histories, are subject to slightly different environments, and interact with island-specific communities of plants and animals (Whittaker & Fernandez-Palacios, 2007; Parent *et al.*, 2008). Some of the archipelagos that have been most influential to our understanding of island biogeography and diversification include the Hawaiian Islands (Roderick & Gillespie, 1998; Lerner *et al.*, 2011) and the Galapagos Islands (Caccone *et al.*, 2002; Parent *et al.*, 2008; Grant & Grant, 2008). While the Hawaiian and Galapagos archipelagos have taught us a great deal about the evolutionary processes affecting population divergence and species formation, these archipelagos are quite isolated from the mainland, and were formed by a relatively simple accumulation of volcanoes resulting from a hotspot under the ocean crust (Whittaker & Fernandez-Palacios, 2007). Studies of more geologically complex archipelagos that lie closer to continental source populations, such as the Philippines (Jones & Kenedy, 2008; Brown *et al.*, 2013) and the islands of Wallacea (Esselstyn *et al.*, 2010), are providing new insights into how complex island colonization routes and levels of genetic divergence correlate with morphological and ecological divergence.

The Lesser Sunda Islands are a group of approximately 15 larger and hundreds of smaller oceanic islands that lie in southeastern Indonesia. The archipelago is extremely geologically complex and includes both volcanic and non-volcanic islands, older (~10-12 MA) as well as younger (~1 MA) islands, and a climatic gradient with more tropical climates in the west and drier climates as one moves east. The islands occur as two parallel island arcs, the Inner and Outer Banda Arc, extending in a west-east direction between the Greater Sunda Islands of Java and Bali in the west and New Guinea in the East (Fig. 1). For most groups of terrestrial fauna, the Lesser Sundas are relatively species poor when compared to land-bridged islands lying on the Sunda and Sahul shelves (Simpson, 1977). These land-bridged Islands (Such as Bali in the west or Aru in the East) periodically become connected with Asia or Australia/New Guinea during glacial maximum when the sea level drops, and during this time terrestrial taxa are able to colonize these islands (Monk *et al*., 1997). However, oceanic islands such as the Lesser Sundas recruit their terrestrial fauna only via overwater dispersal events, which are much less common.

While the Lesser Sundas are thought to be relatively species poor, the archipelago contains a number of wide-ranging species that contain populations on multiple islands, and it's possible that the various island populations within species have been isolated from each other. Most of the work in the Lesser Sundas is still that of compiling inventories and descriptions of species despite the fact that numerous expeditions have been made to the archipelago (Monk *et al.*, 1997). The Lesser Sundas contain a number of species that are endemic to the archipelago, making the archipelago a distinct biogeographic province within the greater biogeographic realm of Wallacea (Stresemann, 1939). Because straits with a depth of over 120 meters separate many of the islands within the Lesser Sundas, we can assume that many of the Lesser Sunda Islands have never been land-bridged to one another and that wide-ranging taxa separated by these straits may represent clades composed of multiple single-island species or incipient species. There is a lack of genetic studies focusing on the biogeographic patterns of taxa from the Lesser Sundas,

likely because of the difficulty of conducting field surveys of the region and the lack of fresh tissue for genetic analyses in museum collections. Focused phylogeographic studies of wide-ranging species will likely uncover many cryptic species, thereby increasing the number of recognized species in the archipelago as well as the proportion of endemic lineages.

Terrestrial lizards have been an important group used to study diversification and biogeography of island systems (Camargo *et al.*, 2010). For example, phylogeographical studies of Galapagos lava lizards (Kizirian *et al.*, 2004; Jordan & Snell, 2008; Benavides *et al.*, 2009) and Iguanas (Tzika *et al.*, 2008) along with many other studies have revealed the timing and sequence of island colonization. Skinks of the genus *Sphenomorphus* make up an extremely diverse clade that occurs throughout Asia, Wallacea, Australo-Papua, and the western Pacific Islands, demonstrating that this clade of skinks is well-adapted to over-water dispersal and colonization of islands. The nominate species for the genus is *Sphenomorphus melanopogon*, which occurs throughout southern Wallacea and on two small islands off of southwest Java, with the type locality being Timor (Dumeril & Bibron, 1839). Most of the Lesser Sunda Islands are inhabited by *S. melanopogon*, and this region makes up the majority of their range. While there are two other species of *Sphenomorphus* described from the Lesser Sunda Islands (such as *S. striolatus* from Flores and *S. vanheurni* from nearby Bali) it is unclear if these species are closely related to, or perhaps even derived from, *S. melanopogon*. *Sphenomorphus melanopogon*, once commonly known as Flores Forest Skinks, are relatively small skinks with various shades of brown and variable color patterns (Fig. 2). They occur from sea level up to approximately 1,200 meters elevation, and are commonly seen in more arid environments such as monsoon forest, along creek beds, and in ecotonal areas (Auffenberg, 1980). *Sphenomorphus melanopogon* used to be composed of four subspecies described on the basis of their divergent color patterns and geographical restriction to certain islands. However, a recent revision of the species has synonymized all subspecies due to the presence of clinal variation in color pattern and overlapping scale characteristics (Shea 2012). Shea (2012) suggests that to address any taxonomic implications of the morphological variation that further work on the Lesser Sundas *Sphenomorphus* is needed that incorporates genetic data, with a special focus on the island of Flores, a region that contained substantial diversity in color pattern and size.

In the last decade, large scale genomic datasets for non-model organisms have become easier and less costly to obtain. The most common types of genomic data available for non-model organisms include ultraconserved elements or UCEs (Faircloth *et al.*, 2012), restriction site associated markers or RAD markers (Baird *et al.*, 2008), parallel tagged amplicon sequencing (Lemmon *et al.*, 2012), and transciptome based exon-capture (Bi *et al.*, 2012). These types of data each have their own pros and cons. UCE and RAD loci are typically very short in length, and for data analyses usually one informative variable site per locus is used. RAD loci are useful for studying a group of closely related populations or species, but as individuals become more divergent they share fewer homologous loci due to mutations in restriction enzyme binding sites. UCEs are useful for examining groups of more divergent taxa but because of the conserved nature of the loci may not be particularly informative for shallow evolutionary time scales (but see Brumfield lab paper). The exon-capture approach allows for the screening of many independent sequence loci, with roughly equal coverage from a divergent set of taxa. The loci targeted for an exon-capture experiment can be carefully evaluated prior to the design of the capture array, thereby allowing for the selection of loci that are optimized in terms of information content for the needs of a study. Thus, targeted exon-capture loci can be more informative than UCE or RAD loci (Bi *et al.*, 2012). In addition, while the exon-capture

approach targets exons, it also captures a large quantity of intronic and flanking sequence, which are particularly appropriate for demographic analyses.

Phylogeographic studies of taxa that have colonized multiple islands of the archipelago also have the potential to test biogeographical hypotheses regarding the timing and entry into the archipelago, as well as the sequence of island colonization within the archipelago. In this study, I utilize an exon-capture approach to collect sequence data for hundreds of independent nuclear loci from *Sphenomorphus melanopogon* populations in the Banda Arc Islands of Wallacea with the goals of 1) determining if *Sphenomorphus melanopogon* is monophyletic with respect to other species of *Sphenomorphus*, 2) determining the relative timing of entry into the Lesser Sundas, 3) estimating the age of island-specific lineages, 4) determine if *S. melanopogon* is closer related to Sunda or Sahul Shelf species of Sphenomorphus, and 5) using the biogeographical patterns to test hypotheses regarding the island colonization models described in Chapter 1.

MATERIALS & METHODS

Sample Collection: *Sphenomorphus melanopogon* specimens were collected from the field from Lombok, Sumbawa, Flores, Lembata, Pantar, Alor, Wetar, and Sumba islands. Samples from East Timor were obtained from Hinrich Kaiser, who has led several expeditions to Timor-Leste (depositing specimens in the Museum of Comparative Zoology at Harvard, and the United States National Museum). Samples and tissues were collected on four separate expeditions to the Lesser Sunda Islands that occurred between 2010 and 2013. Liver tissue was dissected from euthanized lizards and either stored in RNALater, or flash frozen in liquid nitrogen. Specimens were given field tags in the catalogs of Jimmy A. McGuire (JAM#), Alexander L. Stubbs (ALS#), or Ben R. Karin (BRK#). Tissues were divided in half at the time of preparation in the field, with samples deposited in both the Museum of Vertebrate Zoology at UC Berkeley (to subsequently receive MVZ catalog numbers) and the Museum Zoologicum Bogoriense (which are given MZB catalog numbers). The formalin fixed specimens are divided equally between the MVZ and MZB collections.

MtDNA Data Collection: DNA was extracted from liver tissue using standard salt extraction techniques or by using the DNeasy kit. DNA extractions were then diluted to concentrations suitable for PCR-amplification (~20-60 ng/uL). The *ND4* gene was sequenced for 247 skinks, 237 of which are from the Lesser Sunda Islands (Table 1). All sequence data was collected using standard PCR-amplification using the primers ND4 and LEU (Arevalo *et al.* 1994). PCR reactions contained 18.3 µL water, 2.5 µL of 10X buffer, 1.5 µL magnesium chloride, 1.5 µL dNTPs (2 µM), 0.6 µL of each primer, 0.2 µL Taq polymerase, and 1µL genomic DNA at concentration of 20-40 ng/µL. PCR products were cleaned using ExoSAP-IT (USB, Cleveland, OH) before being labeled with fluorescent-dye nucleotides through cycle sequencing reactions for both forward and reverse primers. Ethanol precipitation was used to clean cycle sequencing products, which were sequenced on an ABI 3730 sequencer (Applied Biosystems, Foster City, CA, USA). Raw sequence reads were combined in CODONCODE ALIGNER 3.5.2 (CodonCode Corporation, Dedham, MA, USA), aligned with MUSCLE (Edgar, 2004), and then manually checked and edited.

MtDNA Data Analyses: The sequence alignment of 932 bp was imported into JMODELTEST V2.1.4 (Darriba *et al.*, 2012) to determine the best-fit model of sequence evolution (GTR+G) as

supported by the program BEAST V1.8 (Drummond & Rambaut, 2007). A BEAST run was conducted using the uncorrelated relaxed clock model and a coalescent constant size tree prior with a uniform distribution. A rule of thumb 1% rate of sequence evolution per million years (which corresponds to 2% divergence per million years for any two lineage comparison) was applied to obtain a rough approximation of timing of entry into the archipelago, as well as the ages of island-specific lineages. A preliminary run was carried out to determine the appropriate number of generations needed to achieve ESS values $\geq$ 200, as calculated in TRACER V1.6 (Rambaut & Drummond, 2009). Once the appropriate run length was determined, two separate runs of 50 million generations were carried out, sampling every 5,000 generations for a total of 10,000 saved generations per run. A burn-in of 10% was removed from each of the two runs and the remaining 18,000 trees were combined using LogCombiner to create a 50% majority rule consensus tree. The tree was rooted using the outgroups *Sphenomorphus tagapayo* and *S. indicus*, which were shown to be some of the closest relatives to *S. melanopogon* in an unpublished phylogeny (C. Linkem, pers. comm.). Nodal support was assessed using posterior probability values. A Maximum Likelihood approach was also taken using the program RAXML (Stamatakis, 2014). The default GTR+I+G model of sequence evolution was applied, and nodal support was assessed with 100 bootstrap replicates.

Transcriptome Sequencing: Total RNA was extracted from three *Sphenomorphus melanopogon* samples (JAM 12180, Sumbawa; JAM 12480, Flores; JAM 12652, Lembata) using the RNEasy Protect Mini Kit (Qiagen) and protocol. Samples were evaluated using a BioAnalyzer 2100 RNA Pico chip (Agilent), with RIN scores greater than 8. Sequencing libraries were prepared using half reactions of the TruSeq RNA Library Preparation Kit V2 (Illumina), beginning with Poly-A selection for samples with high RIN scores (> 8.0) and Ribo-Zero Magnetic Gold (Epicentre) ribosomal RNA removal for samples with low RIN scores (< 8.0). Libraries were pooled and sequenced on an Illumina HiSeq2500 with 100 bp paired-end reads. Transcriptomic data were cleaned following Singhal (2013). Cleaned data were assembled using TRINITY (Grabherr *et al*., 2011) and annotated with *Anolis carolinensis* (Ensembl) as a reference genome using reciprocal BLASTX (Altschul *et al*., 1997) and EXONERATE (Slater & Birney, 2005). Annotated transcripts were compared from the three individuals to search for orthologs via BLAST (Altschul *et al*., 1990). Mitochondrial sequences were removed from the transcripts. Only transcripts with a GC content between 40%-70% were kept because extreme GC content causes reduced capture efficiency for the targets (Bi *et al*., 2012). All the bioinformatics pipelines for transcriptome data processing and annotation are available at https://github.com/CGRL-QB3-UCBerkeley/DenovoTranscriptome.

Marker Development: Annotated and filtered contigs from all transcripts were aligned to identify shared markers. Markers under 300 bp were discarded and markers greater than 1,000 bp were cut down to a maximum length of 1,000 bp. The remaining genes were examined for repetitive elements, short repeats, and low complexity regions, which are problematic for probe design and capture. The three sets of transcripts were screened using the REPEATMASKER Web Server (Smit *et al*., 2015), which resulted in the masking of repetitive elements or low complexity regions. To be conservative, if any of the three transcripts for a gene contained masked sites, that gene was removed from the final marker set. The three *Sphenomorphus melanopogon* transcripts were 1.6% divergent on average. The resulting 5,103 markers from the three *Sphenomorphus* transcripts were compared to identify the variability of each marker as determined by the number

of polymorphic sites. The markers were sorted according to the number of variable sites per locus. All invariable loci and loci with only a single variable site were discarded, as well as the top 5% of the most variable loci. A total of 1,199 of the resulting 3,619 candidate loci were integrated into a MYBaits probe design. The estimated target size of the combined loci was approximately 1,090,000 bp. Pipelines for marker development are available at https://github.com/CGRL-QB3-UCBerkeley/MarkerDevelopment Pylogenomics.

Sample Library Preperation: A total of 104 *Sphenomorphus* samples were chosen for library preparation (Table 1). These samples were picked by examining the mitochondrial tree to maximize the genetic diversity from all of the islands and to include some outgroups for rooting of phylogenetic trees (Table 1). The DNA was quantified by Qubit DNA BR assay (Life Technologies) and 1500 ng total DNA was diluted in 100 µl of ultrapure H_2O. A Bioruptor UCD-200 (Diagenode) was used to sonicate the samples on a low setting for 15 minutes, using 30s on/30s off cycling. For each sonicated sample, 5 µl of product was run on a 1% gel at 100V for 45 min to ensure fragments were appropriately sized (200–500 bp). Individual genomic libraries were prepared following Meyer and Kircher (2010), with slight modifications, including the use of at least 1,500 ng total DNA for library preparation (rather than 500 ng) to remedy the possibility of decreased library diversity resulting from a large genome size. We used 7-9 cycles of post-adapter ligation PCR to enrich the libraries and incorporate a 7bp P7 index. The resulting 50 µl of amplified library product had an average concentration of 30 ng/µl as measured by a Nanodrop 1000 spectrophotometer (Thermo Scientific), producing an average yield of 1,500 ng total library DNA.

MYBaits In-Solution Exon Capture: Libraries were pooled in equal amounts with each ingroup pool containing 6 individuals and the outgroup pool containing 4 individuals. Pools were selected by grouping closely related samples to minimize competition for probe binding. MYbaits capture reactions were performed following the v2.3.1 manual with some modifications. For each capture reaction library master mix, the pooled libraries were vacuum dried at 45°C for 60 min and re-suspended in ultrapure H_2O, then combined with 1.66 µl each of salmon sperm COT-1, human COT-1, chicken COT-1, and xGEN blocking oligos. The combined volume of water for DNA resuspension and volume of blocking oligonucleotides totaled 6.5 µl. The hybridization reaction proceeded at 65°C for 24-28 hours. Individual capture reactions were purified using streptavidin-coated magnetic beads, and post-capture products were PCR amplified using four independent reactions of 14 cycles each. These reactions were resuspended in 11 µl of ultrapure H_2O, and had an average concentration of 4-7 ng/µl, as measured by Qubit. Purified PCR products from the same capture were combined and quantified using a BioAnalyzer 2100 DNA-1000 chip. The combined post-capture amplified products ranged from 3-9 ng/µl, and the average product size was ~370 bp. The combined post-capture libraries were sequenced on one lane of an Illumina HiSeq2500 with 100 bp paired-end reads.

Data Pipeline: Raw sequence data were cleaned following Singhal (2013) and Bi *et al*. (2012). Raw fastq reads were filtered using TRIMMOMATIC (Bolger *et al*., 2014) and CUTADAPT (Martin, 2011) to trim adapter contamination and low quality reads. BOWTIE2 (Langmead & Salzberg, 2012) was used to align the data to *Escherichia coli* (NCBI: 48994873) to remove potential bacterial contamination. Exact duplicates were eliminated as well as low complexity sequences using a custom script. Overlapping paired reads were also merged using FLASH (Magoč &

Salzberg, 2011) and COPE (Liu *et al*., 2012) to avoid inflated coverage estimates in overlapping regions. The resulting cleaned reads of each individual specimen were *de novo* assembled using ABYSS (Simpson *et al*., 2009). Individual raw assemblies were generated using a wide range of k-mers (21, 31, 41, 51, 61 and 71), and I then used CD-HIT-EST (Li & Godzik, 2006), BLAT (Kent, 2002), and CAP3 (Huang & Madan, 1999) to cluster and merge all raw assemblies into final, less-redundant assemblies. BLASTN (evalue cutoff = 1e-10, similarity cutoff = 70) was used to compare the target sequences with the raw assemblies of each individual in order to identify the set of contigs that were associated with targets (in-target assemblies). A self-BLASTN (evalue cutoff =1e-20) was run to compare the assemblies against themselves to mask any regions from a contig that matched other regions from other contigs. For each matched contig EXONERATE (http://www.genome.iastate.edu/ bioinfo/resources/manuals/exonerate/exonerate.man.html) was used to define protein-coding and flanking regions. Flanking sequences were retained if they were within 250 bp of a coding region. Finally, all discrete contigs that were derived from the same reference transcript were joined together with Ns based on their relative BLAST hit positions to the reference. Most of the final in-target assemblies contain multiple contigs, and each includes both coding regions and flanking sequences if captured.

Cleaned sequence data were then aligned to the resulting individual-specific in-target assemblies using NOVOALIGN (Li & Durbin, 2009) and only reads that mapped uniquely to the reference were retained. The programs Picard (http://broadinstitute.github.io/picard/) and GATK (McKenna *et al*., 2010) were used to perform re-alignment. Finally, the program SAMTOOLS/BCFTOOLS (Li *et al*., 2009) was used to generate individual consensus sequences by calling genotypes and incorporating ambiguous sites in the in-target assemblies. A consensus base was only kept when the site depth was above 10X. Sites were masked within a 5 bp window around an indel. Sites were also filtered out where more than two alleles were called. Then FASTQ were converted to FASTA using seqtk (https://github.com/lh3/seqtk), and putative repetitive elements and short repeats were masked using REPEATMASKER (Smit et al. 2015) with "vertebrata metazoa" as the database. Markers were removed if more than 80% of the bases were Ns. The read depth of each individual marker was calculated and loci were filtered out if the depth fell outside the 99th percentile of the statistics. Markers were also eliminated if the individual heteozygosity fell outside the 99th percentile of the statistics. The final filtered assemblies of each individual specimen were aligned using MAFFT (Katoh & Standley, 2013). Alignments were then trimmed using TRIMAL (Capella-Gutierrez *et al*., 2009). Alignments were removed if more than 25% missing data (Ns) were present in 25% of the samples, or if the proportion of shared polymorphic sites in any locus was greater than 20%. The bioinformatic pipelines of sequence capture data processing are available at https://github.com/CGRL-QB3-UCBerkeley/denovoTargetCapturePhylogenomics.

Evaluation of Data: To evaluate capture efficiency, the average per-base sequence depth (or coverage) was calculated separately for the exon sequences and for the flanking sequences of each sample. The coverage at each base pair site for either data set was inferred using SAMTOOLS (Li *et al*., 2009). The per base pair coverage estimates for all sequences (exon or flanking) associated with each gene (1,154 genes were retained after all filtering) were averaged, resulting in a set of average coverage estimates across loci. The resulting output of the set of average coverage estimates was used to infer the median, upper and lower quartiles, and range of coverage estimates using samples or genes as factors. These calculations were performed and automated across samples using python scripts and the output was visualized in R. Differences

in the levels of coverage were examined using pooling size as a factor. Cite Dan's paper

The resulting alignments of exon-only data and flanking region data were evaluated for taxon number, sequence length, percentage of missing data, and proportion of informative sites. These results were visualized in R, and the relationship between the number of informative sites and alignment length was investigated using a simple linear regression. The relationship between phylogenetic distance and missing data was also investigated using a simple linear regression. The percentage of missing data was calculated from the final concatenated alignment of exon-only loci that passed multiple post-processing filters, including a minimum length of 100 bp, no more than 80% missing data per sequence in alignments, and no more than 25% total missing data across an alignment. These filters were enforced using a custom alignment refinement python script for all alignments. All custom python scripts for sequence capture performance evaluation are available on github (https://github.com/dportik/).

Phylogenomic Analyses: The concatenated alignment of all sequence data was analyzed with RAxML (Stamatakis, 2014) under the GTR+I+G model of sequence evolution. Nodal support was assessed with 100 bootstrap replicates.

Individual gene trees for each of the 1,154 genes were also generated with RAxML. These gene trees were used as the input files for a supertree approach as implemented in ASTRAL-II (Mirarab *et al.*, 2014). For this analysis, each individual was treated as a "species" because it was unclear where the species boundaries lie within the system and direct comparisons with the topology of the RAxML tree were desired. To look at the efficacy of our loci, we randomly sampled gene trees in bins of 5, 10, 20, 50, 100, 200, and 500 trees to re-construct supertrees in ASTRAL. This was repeated 10 times for each bin and the average Robinson-Foulds distance was calculated in comparison with the full 1,154-locus supertree. The Robinson-Foulds calculation is a measure of the difference in topology between two trees (Robinson & Foulds, 1981).

A species tree approach was also applied to the data with the program SNAPP as implemented in BEAST2 (Bouckaert *et al.*, 2014). This program analyzes unlinked SNP data so the dataset consisted of one randomly chosen informative SNP per locus. The program also requires that you designate your samples to a species *a priori*, so we treated each clade from our mtDNA phylogeny as a group. The program was run for 1,000,000 generations with the first 100,000 generations removed as burn-in. Convergence was assessed using TRACER V1.6 (Rambaut & Drummond, 2009) to confirm that ESS values were greater than 200. The resulting trees are displayed on top of each other as a tree "cloud" to visualize the uncertainty of the topology, and a consensus tree is also overlaid.

Population Structure: One informative SNP per locus was randomly chosen from within the Lesser Sundas assemblage (92 individuals of *Sphenomorphus melanopogon* from the Lesser Sundas) to create an input file for the program STRUCTURE (Pritchard *et al.*, 2000). The program was run with a 50,000 generation burn-in and a 50,000 generation run from K=1 up to K=15 populations with 10 replicates per K. The results were then imported into STRUCTURE HARVESTER (Earl, 2012) to determine the most likely number of populations as determined by the Delta K method. Then the program was run with a 100,000 generation burnin and a 100,000 generation run for K=2 up to K=12 to examine the sequential division of the assemblage for each assumed number of populations.

Inter-Island Demographics: Demographic analysis utilized the flanking sequence from 1,049 loci because these regions are presumably not under selection, though they are linked to exonic regions under selection. These data were analyzed with the program G-PHOCS (Gronau *et al.*, 2011), which is an isolation-with-migration program that is capable of dealing with genomic sequence data from unlinked neutrally evolving loci. This program estimates the effective population sizes of the extant populations as well as their ancestor, population divergence times, and migration rates between extant populations. This analysis was run to estimate relative rates of migration between divergent lineages of *Sphenomorphus melanopogon* that are parapatrically or sympatrically distributed on the islands of Lombok, Sumba, and Flores. The analyses for Lombok and Sumba islands each compared two populations: Lombok1 vs Lombok2 lineages, and South Sumba vs East Sumba lineages. The Lombok analysis combined two runs after a burn-in of 100,000 generations was removed from each run for a combined dataset of 1,514,854 generations with all ESS values greater than 400. The Sumba analysis was slower to run and required four separate runs to be combined after a burnin of 100,000 generations was removed from each run for a combined dataset of 1,242,513 generations with all ESS values greater than 600.

The Flores analysis compared three lineages: West Flores, East Flores1, and East Flores2 lineages. The Flores analysis considered migration bands between all three lineages resulting in six migration bands: two bands between West Flores and East Flores1, two bands between West Flores and East Flores2, and two bands between East Flores1 and East Flores2. The Central Flores lineage was not included due to its extremely divergent nature, small sample size (2 individuals), and to allow the program to finish runs in a reasonable amount of time. For the Flores analysis, three runs were combined due to the slow runtime of such parameter-rich analyses. After removal of 200,000 generations of burn-in from each run, the remaining generations were combined for a total of just over 1 million generations. This combined dataset was viewed in Tracer (Rambaut & Drummond, 2009) to assess the posterior distribution of the demographic parameters and confirm that all ESS values were greater than 200.

A mutation rate of 2.2×10^{-9} mutations/site/year was used to convert parameter estimates (Kumar & Submaranian, 2002). All values for Theta and Tau given by the program G-PHOCS are scaled by 10^{-4}. Demographic parameters estimates were converted to estimates of effective population sizes (individuals) by dividing the scaled Theta estimate by the mutation rate, then dividing that value by 4 (because diploid organisms will have an effective population size of 4 at any given locus). The population divergence time in years was calculated by dividing the scaled Tau estimate by the mutation rate. Migration rate estimates were converted to Migrants per Generation by multiplying the migration estimate by the converted effective population size estimate for the population receiving the gene flow, then dividing that value by the number of generations that have passed (in years) since divergence. While there is not a published estimate of the generation time for *Sphenomorphus melanopogon,* we choose to use a generation time of 1 year. However, the age at which sexual maturity is reached ranges from 1-4 years in studies of other skinks (Brooks, 1967; Vitt & Cooper, 1986; Blomberg & Shine, 2001; Wapstra *et al.*, 2001).

RESULTS

MtDNA Phylogeny: By incorporating sequences from the Lesser Sundas into a larger *Sphenomorphus ND4* data matrix, it was confirmed that *Sphenomorphus melanopogon* from the

Lesser Sundas are monophyletic, and most closely related to species from Malaysia (Pers. comm. Charles Linkem). Both the Maximum Likelihood (Fig. 3) and Bayesian (Fig. 4) phylogenies converged on nearly the same topology with respect to the relationships of major clades. The geographic distribution of these clades can be seen in Figure 5. Though poorly supported, the ML phylogeny places a Lombok clade as sister to the rest of the Lesser Sundas assemblage, while the Bayesian phylogeny places two individuals from central Flores as basal to the Lesser Sundas assemblage. For five islands, all *Sphenomorphus* samples were placed together as monophyletic in both phylogenies, including Sumbawa, Lembata, Pantar, Alor, and Wetar. For the remaining islands, intra-island populations were found to represent paraphyletic assemblages relative to other island populations. For example, samples from Lombok form two non-sister clades that appear to be sympatrically distributed. Samples from Sumba form two non-sister clades that do not appear to overlap geographically, with samples representing one clade obtained in the tropical forest of southwest Sumba and samples representing a second clade obtained from the dry eastern coast. However, we did not obtain dense sampling across the entire island, and improved sampling will be needed to determine the full ranges of each Sumba clade. All but one sample from Timor form one of the most genetically distinct clades, which is sister to the lineage representing the Inner Arc island of Alor. However, one sample from Timor's interior mountainous region is recovered as the sister lineage of the eastern Sumba clade. The island of Flores appears to contain the greatest *Sphenomorphus* genetic diversity within the Lesser Sundas, with four clades recovered. The easternmost end of Flores appears to contain two sympatrically distributed species level lineages, while most of Flores is comprised by a single clade (the West Flores clade) and two samples from the Bajawa region are recovered as one of the most basal lineages. Populations on Kur, Banda, and Ai islands in Maluku Province, which represent the extreme eastern end of the Inner Banda Arc, were found to be deeply nested within this otherwise Lesser Sundas assemblage, with the Kur population placed as the sister lineage of Wetar, and the Banda Islands as sister to Kur + Wetar.

Because the exact source of the Lesser Sundas *Sphenomorphus* is not known the timing of entry into the archipelago was unable to be accurately estimated. According to an unpublished phylogeny of *Sphenomorphus* the Lesser Sundas clade is closely related to species from Malaysia such as *S. indicus* (pers. comm. Charles Linkem). We estimated that the Lesser Sundas clade split from its sister clade in Malaysia ~33.7 Ma (95% CI ~23-48 MA). However, we did not include any *Sphenomorphus* samples from Java or Bali which could be the source of the Lesser Sundas clade. The first divergence event within the Lesser Sundas appears to be between a Lombok1 lineage and a Central Flores lineage nearly 20 Ma (95% CI ~15-26 MY). The oldest lineages are from the western islands of the Inner Banda Arc, such as Lombok, Sumbawa, Flores, and Lembata. In general, the younger lineages are from the eastern Inner Banda Arc islands of Pantar, Alor, Wetar, Kur and the Banda Islands, along with the Outer Banda Arc islands of Sumba and Timor.

The results of the mtDNA analysis were used to choose samples for an exon-capture experiment. The 100 ingroup samples were chosen to represent all major lineages and sampled island populations from the Lesser Sundas and Maluku (Fig. 6).

Exon-Capture Data Characteristics: A total of 1,153 out of the original 1,200 targeted loci passed the filtering stages. The total concatenated alignment of all 1,153 loci consisting of both targeted exons and flanking introns was 1,801,789 base pairs of sequence data. All 104 libraries were successfully sequenced and retained for analyses. The average coverage of the targeted

regions was approximately 82X after duplicate reads were removed, though individual sample coverage ranged from ~25X up to ~140X (Fig. 7). The flanking regions had approximately 35X average coverage after duplicates were removed, with individual sample coverages ranging from ~10X to ~50X (Fig. 8).

More than 500 of the 1,153 alignments contained all 104 samples and the average number of samples per alignment was 103 (Fig. 9b, Fig. 10a). The number of informative sites has a relatively linear relationship with the length of the loci with very few outlier loci (Fig. 9a), and there is on average 10% informative sites per alignment (Fig. 10c). There is no clear relationship between the alignment length and the percentage of gaps in each alignment (Fig. 9c). The final length of the contigs ranged from 100 bp up to ~4,200 bp (Fig 10b), and the percent of missing data was no higher than 25% after the additional filtering step (Fig. 10d).

Phylogenomic Trees: The Maximum Likelihood analysis of the concatenated target + flanking dataset produced a well-supported tree with every major node receiving a bootstrap value of 100 (Fig. 11). The coalescent analysis implemented in ASTRAL-II also converged on the same topology as the Maximum Likelihood tree (Fig. 12). Interestingly, *Sphenomorphus* from the Lesser Sundas are more closely related to *S. meyeri* from Aru/New Guinea than they are to *S. indicus* from China, *S. textus* from Sulawesi, or *S. maculatus* from Cambodia. Like the mtDNA Maximum Likelihood tree, the phylogenomic tree has a Lombok1 lineage as sister to the rest of the Lesser Sundas assemblage, followed by a lineage consisting of two individuals from the Bajawa region of western central Flores. The next major clade consists of three sub-clades: Lombok 2, Sumbawa, and West Flores. While the mtDNA phylogeny places Sumbawa as the basal lineage of this clade with Lombok 2 sister to West Flores, the genomic ML tree places Lombok 2 as basal with Sumbawa sister to West Flores. Moving further up the tree, the East Flores 1 clade is sister to the remainder of the complex. The next clade consists of the Pantar, Lembata, and East Flores 2 lineages, with a highly distinct Pantar lineage sister to Lembata + East Flores 2 lineages that are each well supported as monophyletic but only weakly genetically divergent from one another. Proceeding up the tree, the next two major branches are both from Sumba, followed by a clade composed of relatively weakly divergent (yet well-supported) lineages from Alor, Timor, Wetar, and the Mollucan islands of Kur, Ai, and Banda. The more basal Sumba branch represents samples from the tropical forest of Laiwangi Wanggameti National Park and thus comprises a southwest Sumba lineage. The second Sumba branch includes samples from the dry eastern coast of Sumba, thereby forming an east Sumba lineage. Sister to the east Sumba lineage is a divergent clade that includes the terminal branches of the tree, including Alor, Timor, Wetar, and the eastern Maluku islands of Kur, Banda, and Ai.

The SNP based species tree produced by SNAPP also converged on a topology similar to the Maximum Likelihood and supertree topologies (Fig. 13). Because there were no outgroups included in the SNAPP analysis, the program set the root at a different branch of the tree. However, if the SNAPP tree is rooted on the branch leading to the Lombok 1 lineage, then this tree is identical to the other phylogenomic trees.

Population Structure: The most likely number of populations from the STRUCTURE analyses, as determined by Delta K values, was two (Fig. 14a). The highest probability of the data however occurs when K=7 (Fig. 14b). When K is set to two, one of the genetic clusters include samples from Lombok, Sumbawa, and West and Central Flores, with East Flores 1 showing admixture with the second cluster (Fig. 15). The second cluster contains East Flores 2, Lembata, Pantar,

Alor, Wetar, Sumba, Timor, and Maluku. From K=3 up to K=9 there is more fractal population structure revealed with each increase of K. The recovered groupings represent monophyletic lineages from the phylogenies for the most part. Even at K=9 there is still very distinct genetic clusters showing almost no evidence of gene flow between clusters. The K=9 clusters correspond to Lombok 1, Central Flores, Lombok 2, Sumbawa, West Flores, East Flores 1, Pantar, East Flores 2 + Lembata, Sumba, and Timor + Alor + Wetar (Maluku not included? Say so at the beginning of this section or in the Methods if its not noted there already). Above K=9 the program continued to return nine primary genetic clusters, and therefore no results are shown for cluster schemes involving K values greater than nine.

Demographic Analyses: The parameters in all three analyses of the Lombok, Sumba, and Flores lineages returned robust parameter estimates as seen in the narrow confidence intervals for each parameter (Fig. 16, 17). The rescaled parameter estimates are presented in Table 3.

The comparison of the two *Sphenomorphus melanopogon* lineages that occur on Lombok Island revealed that the effective population size of the Lombok 1 lineage (~1.2 million individuals) is slightly larger than that of the Lombok 2 lineage (~1 million individuals), though these estimates are very similar (Fig. 16; Table 3). The ancestral population size is estimated at approximately 1.3 million individuals, and the two extant lineages are estimated to have diverged from one another ~5.7 million years ago. Since these Lombok populations diverged, gene flow has been either non-existent or extremely low as the estimates of migrants per generation was estimated as approximately one migrant every 10,000 generations), even when considering the upper 95% confidence values.

The comparison of the two lineages that occur on Sumba revealed that the South Sumba lineage has a much larger effective population size than does the East Sumba lineage (~2.2 vs ~0.5 million individuals, respectively; Fig. 16, Table 3). The ancestral population of the two lineages was estimated to include roughly 0.8 million individuals, with the two extant lineages having diverged from one another ~2.3 million years ago. Since their divergence, the two lineages have been exchanging very low numbers of migrants each generation (~0.03 in each direction). The 95% confidence interval for migration does not contain zero in both directions suggesting a non-zero level of migration, though not enough to prevent lineage divergence.

The comparison of three of the Flores lineages revealed that the West Flores lineage has a much larger effective population size (~3 million individuals) when compared to the East Flores 1 (~0.6 million individuals) and East Flores 2 (~0.8 million individuals) lineages (Fig. 17). The ancestor of the two East Flores lineages is estimated to have had ~0.65 million individuals, with the East Flores lineages having diverged roughly 4.4 million years ago. The ancestor of all three Flores lineages is estimated to have had ~1.2 million individuals, with the West Flores lineage having diverged from the East Flores Ancestor lineage ~5.7 million years ago. All six migration parameters were estimated to be well below 1 migrant per generation (~0.0004-0.005), and four of these bands include a value of 0 migrants per generation in their confidence intervals.

DISCUSSION

Forest skinks described as *Sphenomorphus melanopogon* are found on most major islands within the Inner and Outer Banda Arcs of southeastern Indonesia. There is marked variation in color pattern among island populations that once served as the basis for recognizing subspecies, though those have since been synonymized. The enigmatic *Sphenomorphus striolatus* is only

known from a few islands including Flores, but very little work has focused on this species and it is difficult to distinguish from *S. melanopogon*. Given the comprehensive sampling of *Sphenomorphus* from the Lesser Sundas for this study, it is likely that some of our samples (such as the Central Flores clade) represent *S. striolatus* and future work will determine if any of the distinct genetic lineages recovered here represent this species. This study employed phylogenetic analysis of mitochondrial DNA from 237 ingroup samples, and phylogenomic and demographic analysis of a 1,153-gene nuclear dataset for 100 ingroup samples. The historical biogeographical implications are discussed below with respect to the timing and sequence of island colonization, the historical geological evolution of the Lesser Sundas Archipelago, and the taxonomic status of island-specific lineages.

Mitochondrial Phylogeography: While our Maximum Likelihood and Bayesian phylogenetic estimates from the mitochondrial data differ with respect to the inferred basal lineage of *Sphenomorphus melanopogon*, they recovered the same relationships among all other lineages. By analyzing this dataset along with other *ND4* sequences from another study (C. Linkem pers. com.), we find that *S. melanopogon* is most closely related to Asian *Sphenomorphus*. Thus, it is not surprising that the basal lineages occur on Lombok and Flores. Because the ages of the basal lineages are quite old with respect to the age of the archipelago, the source of *S. melanopogon* could have come from any of the surrounding Indonesian islands because the islands of Lombok and Flores were colonized at a time when the other Lesser Sunda Islands (except Sumbawa) had not yet formed. However, it is still unclear where the source population or species occurred and our lack of samples of *Sphenomorphus* species from Java and Bali prevent us from inferring a dispersal event from these islands across Wallace's Line into the Inner Banda Arc.

The ages of the species and of many of the island lineages are older than the islands themselves. If the mutation rate used to calibrate our time-tree is too low it would inflate the divergence time estimates. If the mutation rate used here (~2% lineage divergence per million years) is doubled (similar to what is observed in birds), then our inference would be that the most recent common ancestor existed ~10 Ma, which is roughly the age of the oldest islands in the western Inner Banda Arc. This would also place the age of the younger Outer Banda Arc island lineages at about 2-3 million years old, which is similar to the estimated timing that these islands are thought to have become sub-aerial.

One interesting finding from the mtDNA phylogenies is that four of the major islands contain more than one lineage. Lombok contains two distinct sympatric lineages that differ from each other by more than 15% sequence divergence at the *ND4* gene. It is likely that the more derived Lombok lineage nested within the Sumbawa/Flores clade is the result of a back colonization event that occurred after significant time allowed for genetic incompatibility to accumulate between this lineage and the older *Sphenomorphus* lineage that initially colonized Lombok. The island of Sumba also contains two distinct lineages that are not sister taxa, and the younger age of these lineages is expected given that Sumba is thought to have become sub-aerial only ~2 Ma. These two Sumba lineages were found on different sides of the island and it is not clear if their ranges overlap or come into contact. All but one sample from the island of Timor group together in the phylogeny (East Timor 2 lineage), and the one divergent haplotype (East Timor 1) that is sister to samples from East Sumba was collected from the same locality as a sample belonging to the East Timor 2 lineage indicating sympatry. The most intriguing island within the range of *S. melanopogon* is Flores, which contains four distinct lineages. The West Flores lineage occupies the majority of the island from the western coast eastward to at least the

Ende region. The Central Flores lineage was only found at one locality (two individuals) from the Bajawa region, and West Flores lineage individuals were found just a few kilometers to the north suggesting that these two lineages likely come into contact or are even sympatric. One of the most unexpected results from Flores involves the East Flores 1 and 2 lineages that co-occur in extreme eastern Flores near Larantuka. These two lineages are sister to each other and estimated to have diverged from one another nearly 5-9 Ma. Samples from Lembata are nested within East Flores 2 lineage suggesting that Lembata (which has a land-bridge connection to Flores) was colonized from this lineage recently.

Genomic Phylogeography: Both the concatenated Maximum Likelihood and coalescent supertree analyses converged on the same topology. Almost all distinct mtDNA lineages were recovered as distinct nuclear DNA lineages with the exception of the East Timor 1 sample that is grouped with all other Timor samples in the phylogenomic analyses. Over 1,000 nuclear loci confirm the distinctiveness of these lineages, supporting the idea that *Sphenomorphus melanopogon* represents a species complex, especially when considering the multiple lineages on Lombok, Flores, and Sumba that have remained genetically distinct despite the opportunity to exchange genes. The topology of the tree suggests a complex colonization scenario that is not consistent with a stepping stone model of island colonization. The topology is consistent with an island-age model of island colonization (See Chapter 1), with the caveat that there appears to be either multiple colonization events of some islands (Lombok and Flores), presumably *in situ* divergence within some islands (Flores and Sumba), and a faster mitochondrial mutation rate.

One interesting outcome of the genomic phylogenies concerns the possible source or sister taxon of *Sphenomorphus melanopogon.* While a genus-wide unpublished analysis found that *S. melanopogon* was most closely related to Asian species such as *S. indicus* and *S. maculatus* (C. Linkem pers. com.), this study found that *S. melanopogon* is more closely related to *Sphenomorphus meyeri* from Aru. Aru Island lies in the eastern portion of Maluku province, though it is not part of the Banda Arc formations, and lying on the Sahul Shelf it becomes landbridged with Papua during glacial maxima. The name *S. melanopogon* has been applied to *Sphenomorphus* from New Guinea and Aru, and it was not until recently that *S. meyeri* was formally described as a species distinct from *S. melanopogon* (Shea 2012). A recent squamate-wide phylogenetic analysis (Pyron *et al.*, 2013) found that *S. melanopogon* is sister to *S. jobiensis* and *S. muelleri*, both of which occur on New Guinea. This is consistent with a model in which *S. melanopogon* invaded the western Lesser Sundas via long-distance dispersal from the Sahul Shelf, presumably before the more eastern islands formed. It is not clear if New Guinea/Aru were colonized from the Lesser Sundas, or if the Lesser Sundas were colonized from New Guinea/Aru. However, the sampling of outgroups for the genomic portion of this study is extremely limited and more comprehensive sampling of the genus will be needed to formulate any meaningful biogeographical scenarios for the genus *Sphenomorphus* as a whole.

In several respects, the topology of the phylogenomic tree is quite different from that of the mtDNA tree. The clade containing the Lombok 2, Sumbawa, and West Flores lineages was recovered with Lombok 2 as sister to Sumbawa + West Flores, rather than Sumbawa sister to Lombok 2 + West Flores as in the mtDNA tree. This scenario could be interpreted as a secondary "stepping stone" colonization pattern involving dispersal from Lombok 2 to Sumbawa to West Flores. The phylogenomic tree also places the East Flores 1 lineage as sister to a clade composed of the eastern Inner Banda Arc islands (plus East Flores 2), the Outer Banda Arc Islands, and the islands of Maluku. Another major difference between the phylogenomic and mtDNA trees

involves the placement of individuals from Pantar Island, which is recovered as sister to East Flores 2 + Lembata. This result is surprising given that Alor and Pantar are expected to be connected during glacial maxima, and suggests a possible westward "stepping stone" model involving dispersal first from Pantar to Lembata and then from Lembata to East Flores. The two lineages from Sumba were widely separated on the mtDNA tree, but were recovered in the phylogenomic analyses as consecutive branches leading to a Timor + Alor + Wetar + Maluku clade. This topology strongly suggests that the Timor Island Group (Timor + Alor + Wetar) was colonized by way of Sumba. This colonization scenario is unexpected due to the large dispersal distance between Sumba and each of these islands. I would have predicted a more direct colonization pathway to Alor and Wetar by way of the Inner Banda Arc, where the straits between islands are relatively narrow. This same colonization scenario also applies to Flying Lizards (See Chapter 2) and suggests that the arrangement of the islands, particularly Sumba, may have been dramatically different in the past. One reconstruction of the Lesser Sundas Archipelago, inferred Sumba and Timor to be extremely close to one another 3 Ma, perhaps even connected (Burrett *et al.*, 1991). The Timor group lineages and Maluku are all very closely related and suggests a very recent colonization of these islands. Most surprising is the close relationship between the eastern Maluku islands of Kur, Banda, and Ai to samples from Wetar. If the Maluku islands were naturally colonized, then this suggests a number of successful dispersal events to extremely small and isolated islands over a short period of time. During the northwest monsoons (November to March), water flows from the Flores Sea eastward into the Banda Sea, and then flows northeast with the current passing Wetar Island (Salm & Halim, 1984). Another possibility is that the Maluku islands were colonized as a result of human introduction, possibly through the movement of nutmeg trees or other spice trees (where eggs could be deposited), either by the Bandanese who inhabited the islands as early as 8,000 years ago, by the Javanese who traded extensively in the region, or by Europeans in the last few hundred years (Monk *et al.*, 1997).

Population Structure: The main conclusions that can be taken from the population structure plots is that there is defined genetic structuring of populations. Regardless of the number of populations assumed, the analyses usually return clusters that correspond to island lineages defined by the genomic phylogeny, or clades of island lineages. Even up to an assumed K of nine populations, the analyses return nearly pure blocks for each cluster suggesting genetic isolation of these populations or species. The population on Pantar Island, while closely related to East Flores 2 + Lembata in the phylogeny, shows some ancestry from the Timor islands group (Timor + Alor + Wetar) and this is likely a result of the periodic merging of Alor and Pantar when sea levels drop during glacial maxima.

Demographic Analyses: In general, the confidence intervals around all estimated demographic parameters were narrow suggesting a strong signal in the genomic data. The main reason for running such analyses was to determine the extent of gene flow between distinct lineages that occupy the same island, and thus have the potential for gene flow between them.

The two most divergent lineages that co-occur are the lineages on Lombok, which are estimated to have had a population divergence event around 5.5 Ma. The current effective population sizes are roughly equal, and are both much larger than the ancestral population size, an expected outcome for colonization of oceanic islands where founding populations are expected to be quite small. The migration estimates between the Lombok lineages is essentially

zero, even when considering the upper 95% confidence interval, and suggest complete genetic isolation of these lineages.

The two most recently diverged lineages that potentially co-occur are the lineages on the younger island of Sumba. The effective population size of the population on the xeric eastern coastline is estimated to be small compared to the southwestern coastal population that occurs in tropical forest habitat. The population divergence time of approximately 2 Ma is comparable to the estimated timing of sub-aerial emergence of the island. In this scenario, one could imagine that as the Sumba block was uplifting that multiple peaks of the island first emerged as distinct islands, which were only connected after further uplift created land bridges between them. At this point the populations would have potentially been separated long enough to diverge genetically. The migration estimates are extremely low, suggesting only a few migrants in either direction every hundred generations. This level of gene flow would not be sufficient to prevent divergence into distinct species. Given enough time, gene flow will likely cease.

For the comparison of Flores populations, the Central Flores lineage was not included due to its high level of divergence in nuclear loci and its impediment on the progress of the analysis. The effective population size of the Western lineage was estimated to be much greater than either of the Eastern lineages, a result that makes sense given the much larger range size of the Western lineage. The population divergence time between the Western lineage and the ancestor of the Eastern lineages is similar to the divergence time estimated for the Lombok lineages, nearly 5.5 Ma, while the Eastern lineages are estimated to have diverged just over 4 Ma. All of the migration estimates between lineages are low, with no genes entering into the Western lineage. The only migration estimate that does not contain zero migration in the posterior distribution is the estimate of migration from West Flores into the East Flores 2 lineage, though even this estimate suggests less than one migrant every hundred generations. The most surprising estimates are those between the sympatrically occurring and more recently diverged East Flores lineages that are genetically isolated. These results suggest that these lineages are evolving independently of one another and warrant elevation to full species.

Taxonomy of Lesser Sundas Sphenomorphus: It is clear that the species diversity of *Sphenomorphus* from the Lesser Sundas is vastly underestimated. Genomic data suggest that within the Lesser Sundas there could be as many as 10-12 distinct species, some with deep population divergences within them (e.g. West Flores). While it still needs to be confirmed with morphological characters, our two highly divergent Central Flores samples might represent *Sphenomorphus striolatus*. Because the type locality for *Sphenomorphus melanopogon* is Timor, the Timor + Alor + Wetar + Maluku clade would retain this name. I would not split the Timor + Alor + Wetar + Maluku clade into separate species because their divergence within the tree is not great, and they remain as a distinct unit in the population structure analyses even up to K=9 and greater. However, populations on Timor, Alor, and Wetar are reciprocally monophyletic in both mtDNA and nDNA and have very little opportunity for gene exchange suggesting that they may also deserve designation as evolutionary significant units, and perhaps they do represent newly formed species. While still preliminary, I predict that new names will have to be assigned to at least 9 other lineages including: Lombok1, Lombok2, Sumbawa, West Flores, East Flores1, East Flores2 + Lembata, South Sumba, East Sumba, and Pantar. These lineages are highly divergent, have been genetically isolated for millions of years, and do not merge even when they come into contact.

Given that a recent study examined morphological characters from museum specimens of *Sphenomorphus melanopogon* from throughout their range (Shea 2012), it may be possible to use this information along with the genetic data to describe these new species. One possible difficulty to this approach is that Shea (2012) did not find enough morphological divergence to warrant recognizing additional species, though if he had known the extent of genetic isolation between populations it may have helped him know which sets of samples to compare. However, Shea knew that there was diversity within the complex and that genetic data would be needed to inform any taxonomic revision. Particularly important specimens that should be examined include the sets of sympatrically and parapatrically occurring lineages from Lombok, Flores, and Sumba. While only a crude observation at present, I have noticed differences in color pattern and body size among the Flores lineage specimens, differences in color pattern and hemipene morphology between Lombok lineage specimens, and differences in color pattern between Sumba lineage specimens. A formal morphological analysis of the complex in light of these genomic results is greatly needed.

CONCLUSIONS

There is significant genetic divergence among lineages of *Sphenomorphus* skinks within the Lesser Sundas, and populations in Maluku province are relatively young and derived from Wetar Island populations. The oldest lineages occur on the oldest islands in the western Inner Arc of the archipelago, and the supposedly young (~2 MY) island of Sumba contains two independently evolving lineages that diverged from each other just over 2 million years ago. Lombok contains the oldest lineage in the complex that is sympatric with another lineage that split from it over 5 million years ago, with negligible gene flow since their divergence. The island of Flores contains the most diversity with four independently evolving lineages that diverged over four million years ago. The complex patterns do not support a stepping-stone model of colonization and suggest a model influenced by island-age, with the caveat that there have been multiple colonization events of some islands, and that the populations from the Timor island group (Timor, Alor, and Wetar) along with Maluku populations are derived from Sumba rather than from the Inner Arc. Given the relatively small strait separating Pantar and Alor, along with the fact that these two islands likely become periodically merged during glacial maxima, one would expect Alor to have been colonized by populations from Pantar. Other forces must be influencing colonization patterns in the archipelago such as ocean and wind currents, possible rearrangement of the islands over time, or the merging/splitting of islands. A taxonomic revision of the *Sphenomorphus* of the Lesser Sundas is needed, with at least 9-11 species occurring in the complex. This study shows that the Lesser Sundas have a complex geological history that has helped produce a diverse insular fauna and that more studies of other wide-ranging taxa from the archipelago are needed to fully understand the geological evolution of the archipelago as well as to realize the true biodiversity of the region.

ACKNOWLEDGEMENTS

I would like to thank Jimmy A. McGuire, Djoko Iskandar, Alexander Stubbs, Ben Karin, Kristopher Harmon, Umilaila Arafin, Evy Arida, Gilang Ramadan, Amir Hamidy, Luke Bloch, Sarah Hykin, Jerome Fuchs, and Jim Patton for their help with the field collection of specimens and tissues. I thank Hinrich Kaiser for tissues from East Timor. I thank Ke Bi and Lydia Smith

for laboratory support, Carol Spencer and Jackie Childers for accessioning of specimens, and Michelle Koo for GIS map support, and Ke Bi for bioinformatics support. Funding was provided by the MVZ, the Department of Integrative Biology at UC Berkeley, the National Geographic Foundation, and the National Science Foundation. Collection of specimens was carried out under UC Berkeley IACUC # A3084-01.

REFERENCES

Arèvalo, E., Davis, S.K. & Sites, J.W. (1994) Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the Sceloporus grammicus complex (Phrynosomatidae) in central Mexico. *Systematic Biology*, 43(3):387–418.

Auffenberg, W. (1980) *The herpetofauna of Komodo, with notes on adjacent areas*. University of Florida.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, 215:403–410.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25(17):3389–3402.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A. & Johnson, E.A. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS One*, 3(10):e3376.

Benavides, E., Baum, R., Snell, H.M., Snell, H.L. & Sites Jr, J.W. (2009) Island biogeography of Galápagos lava lizards (Tropiduridae: Microlophus): species diversity and colonization of the archipelago. *Evolution*, 63(6):1606–1626.

Bi, K., Vanderpool, D., Singhal, S., Linderoth, T., Moritz, C. & Good, J.M. (2012) Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics*, 13(1):403.

Blomberg, S.P. & Shine, R. (2001) Modeling life history strategies with capture–recapture data: evolutionary demography of the water skink Eulamprus tympanum. *Austral Ecology*, 26(4):349–359.

Bolger, A.M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30:2114–2120.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*(4):e1003537.

Brooks, G.R. (1967) Population ecology of the ground skink, *Lygosoma laterale* (Say). *Ecological Monographs*, 37(2):72–87.

Brown, R.M., Siler, C.D., Oliveros, C.H., Esselstyn, J.A., Diesmos, A.C., Hosner, P.A., Linkem, C.W., Barley, A.J., Oaks, J.R., Sanguila, M.B. & Welton, L.J. (2013) Evolutionary processes of diversification in a model island archipelago. *Annual Review of Ecology, Evolution, and Systematics*, 44:411–435.

Burrett, C., Berry, D.R. & Varne, R. (1991) Asian and South-western Pacific continental terranes derived from Gondwana, and their biogeographic significance. *Australian Systematic Botany*, 4:13–24.

Caccone, A., Gentile, G., Gibbs, J.P., Fritts, T.H., Snell, H.L., Betts, J. & Powell, J.R. (2002) Phylogeography and history of giant Galápagos tortoises. *Evolution*, 56(10):2052–2066.

Camargo, A., Sinervo, B. and Sites, J.W. (2010) Lizards as model organisms for linking phylogeographic and speciation studies. *Molecular Ecology*, 19(16):3250–3270.
Capella-Gutierrez, S., Silla-Martinez, J.M. & Gabaldon, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25:1972–1973.
Darriba, D., Taboada, G.L., Doallo, R. & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9(8):772–772.
Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7(1):214.
Duméril, A.M.C. & Bibron, G. (1839) *Erpétologie générale* (Vol. 5). Roret.
Earl, D.A. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2):359–361.
Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5):1792–1797.
Esselstyn, J.A., Oliveros, C.H., Moyle, R.G., Peterson, A.T., McGuire, J.A. & Brown, R.M. (2010) Integrating phylogenetic and taxonomic evidence illuminates complex biogeographic patterns along Huxley's modification of Wallace's Line. *Journal of Biogeography*, 37(11):2054–2066.
Faircloth, B.C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, R.T. & Glenn, T.C. (2012) Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, 61(5):717–726.
Grabherr M., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z., Mauceli E., Hacohen N., Gnirke A., Rhind N., di Palma F., Birren B.W., Nusbaum C., Lindblad-Toh K., Friedman N. & Regev A. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 15:644–652.
Grant, P.R. & Grant, B.R. (2011) *How and why species multiply: the radiation of Darwin's finches*. Princeton University Press.
Gronau, I., Hubisz, M.J., Gulko, B., Danko, C.G. & Siepel, A. (2011) Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, 43:1031–1034.
Huang, X. & Madan, A. (1999) CAP3: a DNA sequence assembly program. *Genome Research*, 9:868–877.
Jones, A.W. & Kennedy, R.S. (2008) Evolution in a tropical archipelago: comparative phylogeography of Philippine fauna and flora reveals complex patterns of colonization and diversification. *Biological Journal of the Linnean Society*, 95(3):620–639.
Jordan, M.A. & Snell, H.L. (2008) Historical fragmentation of islands and genetic drift in populations of Galápagos lava lizards (*Microlophus albemarlensis* complex). *Molecular Ecology*, 17(5):1224–1237.
Katoh, K. & Standley, D.M. (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, 30:722–780.
Kent, W. (2002) BLAT–the BLAST-like alignment tool. Genome Research, 12:656–664.
Kizirian, D., Trager, A., Donnelly, M.A. & Wright, J.W. (2004) Evolution of Galapagos Island lava lizards (Iguania: Tropiduridae: Microlophus). *Molecular Phylogenetics and Evolution*, 32(3):761–769.

Kumar, S. & Subramanian, S. (2002) Mutation rates in mammalian genomes. *Proceedings of the National Academy of Sciences*, 99(2):803-808.
Langmead, B. & Salzberg, S. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9:357–359.
Lemmon, A.R., Emme, S.A. and Lemmon, E.M. (2012) Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, 61(5):727–744.
Lerner, H.R., Meyer, M., James, H.F., Hofreiter, M. & Fleischer, R.C. (2011) Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. *Current Biology*, 21(21):1838–1844.
Li, H. & Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25:1754–1760.
Li, W. & Godzik, A. (2006) CD-Hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22:1658–1659.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, 25:2078–2079.
Liu, B., Yuan, J., Yiu, S-M., Li, Z., Xie, Y., Chen, Y., Shi, Y., Zhang, H., Li, Y., Lam, T-W. & Luo, R. (2012) COPE: an accurate k-mer-based pair-end reads connection tool to facilitate genome assembly. *Bioinformatics*, 28:2870–2874.
Losos, J.B. & Ricklefs, R.E. (2009) *The theory of island biogeography revisited*. Princeton University Press.
Magoč, T. & Salzberg, S. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, 27:2957–2963.
Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17:10–12.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. & DePristo, M.A. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20:1297–1303.
Meyer, M. & Kircher, M. (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, 2010(6), pdb.prot5448.
Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M. S. & Warnow, T. (2014) ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics*, 30(17):541-548.
Monk, K.A., de Fretes, Y. & Reksodiharjo-Lilley, G. (1997) *The Ecology of Nusa Tenggara and Maluku*. The Ecology of Indonesia Series, Volume 5. Periplus eds., Singapore.
Parent, C.E., Caccone, A. & Petren, K. (2008) Colonization and diversification of Galápagos terrestrial fauna: a phylogenetic and biogeographical synthesis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 363(1508):3347–3361.
Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155(2):945–959.
Pyron, R.A., Burbrink, F.T. & Weins, J.J. (2013) A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evolutionary Biology*, 13:93.

Rambaut, A. & Drummond, A.J. (2009) Tracer version 1.5.0. *URL http://tree.bio.ed.ac.uk/software/tracer/[accessed 5 July 2015]*.
Robinson, D.F. & Foulds, L.R. (1981) Comparison of phylogenetic trees. *Mathematical Biosciences*, 53(1):131–147.
Roderick, G.K. & Gillespie, R.G. (1998) Speciation and phylogeography of Hawaiian terrestrial arthropods. *Molecular Ecology*, 7(4):519–531.
Salm, R.V. & Halim, M. (1984) Marine conservation data atlas. Planning for the survival of Indonesia's seas and coasts. *IUCN/WWF*, Bogor, Indonesia.
Shea, G.M. (2012) On the identity of the type species of Sphenomorphus (Squamata: Scincidae): *Lygosoma melanopogon* Duméril and Bibron 1839, with a note on a new scalation character of the pes in Sphenomorphus. *Zootaxa*, 3490:1–29.
Simpson, G.G. (1977) Too many lines; the limits of the Oriental and Australian zoogeographic regions. *Proceedings of the American Philosophical Society*, 121(2):107–120.
Simpson, J., Wong, K., Jackman, S., Schein, J., Jones, S. & Birol, I. (2009) ABySS: a parallel assembler for short read sequence data. *Genome Research*, 19:1117–1123.
Singhal, S. (2013) *De novo* transcriptomic analyses for non-model organisms: an evaluation of methods across a multi-species data set. *Molecular Ecology Resources*, 13:403–416.
Slater, G.S. & Birney, E. (2005) Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, 6:31.
Smit, A.F.A., Hubley, R. & Green, P. (2015) *RepeatMasker Open-4.0*. [http://www.repeatmasker.org]
Stamatakis, A. (2014) RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9):1312–1313.
Stresemann, E. (1939) Die vogel von Celebes. *Zoogeographie. Journal of Ornithologie*, 87:312–425.
Tzika, A.C., Rosa, S.F., Fabiani, A., Snell, H.L., Snell, H.M., Marquez, C., Tapia, W., Rassmann, K., Gentile, G. & Milinkovitch, M.C. (2008) Population genetics of Galápagos land iguana (genus *Conolophus*) remnant populations. *Molecular Ecology*, 17(23):4943–4952.
Vitt, L.J. & Cooper Jr, W.E. (1986) Skink reproduction and sexual dimorphism: *Eumeces fasciatus* in the southeastern United States, with notes on *Eumeces inexpectatus*. *Journal of Herpetology*, 20(1):65–76.
Wapstra, E., Swain, R. & O'Reilly, J.M. (2001) Geographic variation in age and size at maturity in a small Australian viviparous skink. *Copeia*, 2001(3):646–655.
Whittaker, R.J. & Fernández-Palacios, J.M. (2007) *Island biogeography: ecology, evolution, and conservation*. Oxford University Press.

TABLES

Table 1. Locality information for samples used in genetic analyses. Samples used in the exon-capture experiment are denoted by an "X" in the ExonCap column.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
JAM11533	*Sphenomorphus*	*melanopogon*	Lombok	-8.53220	116.39875	Lombok1	X
JAM11614	*Sphenomorphus*	*melanopogon*	Lombok	-8.85640	116.2829	Lombok2	X
JAM11615	*Sphenomorphus*	*melanopogon*	Lombok	-8.85640	116.2829	Lombok2	
JAM11616	*Sphenomorphus*	*melanopogon*	Lombok	-8.85640	116.2829	Lombok2	X
JAM11617	*Sphenomorphus*	*melanopogon*	Lombok	-8.85640	116.2829	Lombok2	X
JAM11618	*Sphenomorphus*	*melanopogon*	Lombok	-8.85640	116.2829	Lombok2	X
JAM11619	*Sphenomorphus*	*melanopogon*	Lombok	-8.85640	116.2829	Lombok2	X
JAM11620	*Sphenomorphus*	*melanopogon*	Lombok	-8.85640	116.2829	Lombok2	X
JAM11638	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.40759	117.2084	Sumbawa	X
JAM11654	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.70333	117.4051	Sumbawa	
JAM11658	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.53714	117.463	Sumbawa	
JAM11660	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.53714	117.463	Sumbawa	X
JAM11661	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.53714	117.463	Sumbawa	
JAM11662	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.58349	117.6566	Sumbawa	
JAM11663	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.58349	117.6566	Sumbawa	
JAM11664	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.58349	117.6566	Sumbawa	
JAM11665	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.58349	117.6566	Sumbawa	
JAM11666	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.58349	117.6566	Sumbawa	X
JAM11667	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.58349	117.6566	Sumbawa	
JAM11668	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.58349	117.6566	Sumbawa	X
JAM11773	*Sphenomorphus*	*melanopogon*	Flores	-8.49414	119.8796	West Flores	
JAM11779	*Sphenomorphus*	*melanopogon*	Flores	-8.5453	119.9106	West Flores	
JAM11780	*Sphenomorphus*	*melanopogon*	Flores	-8.5453	119.9106	West Flores	
JAM11781	*Sphenomorphus*	*melanopogon*	Flores	-8.5453	119.9106	West Flores	
JAM11784	*Sphenomorphus*	*melanopogon*	Flores	-8.5453	119.9106	West Flores	
JAM11785	*Sphenomorphus*	*melanopogon*	Flores	-8.5453	119.9106	West Flores	
JAM11786	*Sphenomorphus*	*melanopogon*	Flores	-8.5453	119.9106	West Flores	
JAM11788	*Sphenomorphus*	*melanopogon*	Flores	-8.53961	119.928	West Flores	
JAM11789	*Sphenomorphus*	*melanopogon*	Flores	-8.53961	119.928	West Flores	X
JAM11790	*Sphenomorphus*	*melanopogon*	Flores	-8.53961	119.928	West Flores	
JAM11791	*Sphenomorphus*	*melanopogon*	Flores	-8.53961	119.928	West Flores	
JAM11792	*Sphenomorphus*	*melanopogon*	Flores	-8.53961	119.928	West Flores	
JAM11902	*Sphenomorphus*	*melanopogon*	Lombok	-8.53912	116.5396	Lombok2	X
JAM11903	*Sphenomorphus*	*melanopogon*	Lombok	-8.53912	116.5396	Lombok1	X
JAM11904	*Sphenomorphus*	*melanopogon*	Lombok	-8.53912	116.5396	Lombok1	X
JAM11905	*Sphenomorphus*	*melanopogon*	Lombok	-8.53912	116.5396	Lombok1	
JAM11925	*Sphenomorphus*	*melanopogon*	Lombok	-8.50704	116.5601	Lombok1	
JAM11926	*Sphenomorphus*	*melanopogon*	Lombok	-8.50704	116.5601	Lombok1	
JAM11927	*Sphenomorphus*	*melanopogon*	Lombok	-8.50704	116.5601	Lombok1	X
JAM11928	*Sphenomorphus*	*melanopogon*	Lombok	-8.50704	116.5601	Lombok1	X
JAM11929	*Sphenomorphus*	*melanopogon*	Lombok	-8.50704	116.5601	Lombok1	
JAM11930	*Sphenomorphus*	*melanopogon*	Lombok	-8.50704	116.5601	Lombok1	X
JAM11931	*Sphenomorphus*	*melanopogon*	Lombok	-8.50704	116.5601	Lombok1	

Table 1 cont.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
JAM11932	*Sphenomorphus*	*melanopogon*	Lombok	-8.50704	116.5601	Lombok1	X
JAM12179	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.48660	118.66402	Sumbawa	
JAM12180	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.48660	118.66402	Sumbawa	
JAM12190	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.48660	118.66402	Sumbawa	X
JAM12229	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.382	118.29	Sumbawa	
JAM12230	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.382	118.29	Sumbawa	
JAM12231	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.382	118.29	Sumbawa	
JAM12233	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.382	118.29	Sumbawa	X
JAM12234	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.382	118.29	Sumbawa	
JAM12235	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.382	118.29	Sumbawa	
JAM12236	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.382	118.29	Sumbawa	X
JAM12248	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.34	118.908	Sumbawa	
JAM12269	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.66500	118.0701	Sumbawa	X
JAM12302	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.50065	119.0363	Sumbawa	
JAM12343	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.59961	119.0123	Sumbawa	X
JAM12344	*Sphenomorphus*	*melanopogon*	Sumbawa	-8.68169	118.9242	Sumbawa	X
JAM12376	*Sphenomorphus*	*melanopogon*	Flores	-8.698	121.814	West Flores	X
JAM12401	*Sphenomorphus*	*melanopogon*	Flores	-8.49577	119.8974	West Flores	
JAM12402	*Sphenomorphus*	*melanopogon*	Flores	-8.49577	119.8974	West Flores	X
JAM12403	*Sphenomorphus*	*melanopogon*	Flores	-8.49577	119.8974	West Flores	
JAM12404	*Sphenomorphus*	*melanopogon*	Flores	-8.49577	119.8974	West Flores	
JAM12405	*Sphenomorphus*	*melanopogon*	Flores	-8.49577	119.8974	West Flores	
JAM12406	*Sphenomorphus*	*melanopogon*	Flores	-8.49577	119.8974	West Flores	
JAM12408	*Sphenomorphus*	*melanopogon*	Flores	-8.54537	119.9107	West Flores	X
JAM12410	*Sphenomorphus*	*melanopogon*	Flores	-8.54537	119.9107	West Flores	
JAM12411	*Sphenomorphus*	*melanopogon*	Flores	-8.54537	119.9107	West Flores	X
JAM12464	*Sphenomorphus*	*melanopogon*	Flores	-8.80738	120.59010	West Flores	
JAM12465	*Sphenomorphus*	*melanopogon*	Flores	-8.80738	120.59010	West Flores	X
JAM12466	*Sphenomorphus*	*melanopogon*	Flores	-8.80738	120.59010	West Flores	X
JAM12479	*Sphenomorphus*	*melanopogon*	Flores	-8.71399	121.0239	Central Flores	X
JAM12480	*Sphenomorphus*	*melanopogon*	Flores	-8.71399	121.0239	Central Flores	X
JAM12481	*Sphenomorphus*	*melanopogon*	Flores	-8.60773	121.0818	West Flores	X
JAM12534	*Sphenomorphus*	*melanopogon*	Flores	-8.83634	121.6825	West Flores	X
JAM12560	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	X
JAM12561	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	
JAM12562	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	
JAM12563	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	
JAM12564	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	
JAM12565	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	X
JAM12566	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 1	X
JAM12567	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	
JAM12568	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	
JAM12569	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	

Table 1 cont.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
JAM12570	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	X
JAM12571	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 1	X
JAM12572	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	
JAM12573	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	
JAM12574	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	
JAM12575	*Sphenomorphus*	*melanopogon*	Flores	-8.27238	122.9926	East Flores 2	X
JAM12595	*Sphenomorphus*	*melanopogon*	Flores	-8.24706	122.9808	East Flores 2	
JAM12597	*Sphenomorphus*	*melanopogon*	Flores	-8.13754	122.8857	East Flores 1	X
JAM12601	*Sphenomorphus*	*melanopogon*	Flores	-8.10555	122.5225	East Flores 2	
JAM12602	*Sphenomorphus*	*melanopogon*	Flores	-8.10555	122.5225	East Flores 2	X
JAM12603	*Sphenomorphus*	*melanopogon*	Flores	-8.10555	122.5225	East Flores 1	
JAM12604	*Sphenomorphus*	*melanopogon*	Flores	-8.10555	122.5225	East Flores 1	X
JAM12605	*Sphenomorphus*	*melanopogon*	Flores	-8.10555	122.5225	East Flores 2	
JAM12606	*Sphenomorphus*	*melanopogon*	Flores	-8.10555	122.5225	East Flores 1	X
JAM12607	*Sphenomorphus*	*melanopogon*	Flores	-8.10555	122.5225	East Flores 2	
JAM12652	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	
JAM12655	*Sphenomorphus*	*melanopogon*	Lembata	-8.53691	123.4481	Lembata	X
JAM12656	*Sphenomorphus*	*melanopogon*	Lembata	-8.53691	123.4481	Lembata	
JAM12657	*Sphenomorphus*	*melanopogon*	Lembata	-8.53691	123.4481	Lembata	
JAM12658	*Sphenomorphus*	*melanopogon*	Lembata	-8.53691	123.4481	Lembata	
JAM12659	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	
JAM12660	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	
JAM12661	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	
JAM12662	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	
JAM12663	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	X
JAM12664	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	
JAM12665	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	X
JAM12666	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	
JAM12667	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	X
JAM12668	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	
JAM12669	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	
JAM12670	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	X
JAM12671	*Sphenomorphus*	*melanopogon*	Lembata	-8.56556	123.43	Lembata	X
JAM12672	*Sphenomorphus*	*melanopogon*	Lembata	-8.53691	123.4481	Lembata	
JAM12673	*Sphenomorphus*	*melanopogon*	Lembata	-8.53691	123.4481	Lembata	
JAM12674	*Sphenomorphus*	*melanopogon*	Lembata	-8.53691	123.4481	Lembata	
JAM12675	*Sphenomorphus*	*melanopogon*	Lembata	-8.53691	123.4481	Lembata	
JAM12717	*Sphenomorphus*	*melanopogon*	Lembata	-8.53691	123.4481	Lembata	
JAM12830	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12831	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12833	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	X
JAM12834	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12835	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	

Table 1 cont.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
JAM12836	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12837	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12838	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12839	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12840	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12841	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12842	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	X
JAM12843	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12844	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12845	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12846	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12849	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	X
JAM12850	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	X
JAM12851	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12852	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	
JAM12853	*Sphenomorphus*	*melanopogon*	Alor	-8.16993	124.5934	Alor	X
JAM13009	*Sphenomorphus*	*melanopogon*	Alor	-8.17922	124.559	Alor	
JAM13010	*Sphenomorphus*	*melanopogon*	Alor	-8.17922	124.559	Alor	
JAM13011	*Sphenomorphus*	*melanopogon*	Alor	-8.17922	124.559	Alor	X
JAM13012	*Sphenomorphus*	*melanopogon*	Alor	-8.17922	124.559	Alor	X
JAM13142	*Sphenomorphus*	*melanopogon*	Sumba	-10.021	120.058	South Sumba	X
JAM13149	*Sphenomorphus*	*melanopogon*	Sumba	-10.0251	120.0575	South Sumba	X
JAM13167	*Sphenomorphus*	*melanopogon*	Sumba	-10.0222	120.0586	South Sumba	X
JAM13172	*Sphenomorphus*	*melanopogon*	Sumba	-10.0383	120.0578	South Sumba	X
JAM13201	*Sphenomorphus*	*melanopogon*	Sumba	-10.0324	120.0579	South Sumba	
JAM13202	*Sphenomorphus*	*melanopogon*	Sumba	-10.0383	120.0578	South Sumba	
JAM13203	*Sphenomorphus*	*melanopogon*	Sumba	-10.0383	120.0578	South Sumba	
JAM13204	*Sphenomorphus*	*melanopogon*	Sumba	-10.021	120.058	South Sumba	X
JAM13206	*Sphenomorphus*	*melanopogon*	Sumba	-10.0383	120.0578	South Sumba	X
JAM13207	*Sphenomorphus*	*melanopogon*	Sumba	-10.0324	120.0579	South Sumba	X
JAM13267	*Sphenomorphus*	*melanopogon*	Sumba	-9.74842	120.5836	East Sumba	
JAM13268	*Sphenomorphus*	*melanopogon*	Sumba	-9.74883	120.582	East Sumba	X
JAM13269	*Sphenomorphus*	*melanopogon*	Sumba	-9.74883	120.582	East Sumba	X
JAM13270	*Sphenomorphus*	*melanopogon*	Sumba	-9.74883	120.582	East Sumba	X
JAM13271	*Sphenomorphus*	*melanopogon*	Sumba	-9.74883	120.582	East Sumba	X
JAM13278	*Sphenomorphus*	*melanopogon*	Sumba	-10.0877	120.7501	East Sumba	
JAM13279	*Sphenomorphus*	*melanopogon*	Sumba	-10.0877	120.7501	East Sumba	
JAM13282	*Sphenomorphus*	*melanopogon*	Sumba	-10.1372	120.7126	East Sumba	
JAM13305	*Sphenomorphus*	*melanopogon*	Sumba	-10.2117	120.6182	East Sumba	X
JAM13306	*Sphenomorphus*	*melanopogon*	Sumba	-10.2117	120.6182	East Sumba	X
JAM13307	*Sphenomorphus*	*melanopogon*	Sumba	-10.2117	120.6182	East Sumba	X
JAM13652	*Sphenomorphus*	*melanopogon*	Wetar	-7.92823	126.4086	Wetar	
JAM13662	*Sphenomorphus*	*melanopogon*	Wetar	-7.92487	126.4079	Wetar	

Table 1 cont.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
JAM13663	*Sphenomorphus*	*melanopogon*	Wetar	-7.92487	126.4079	Wetar	X
JAM13664	*Sphenomorphus*	*melanopogon*	Wetar	-7.92487	126.4079	Wetar	
JAM13666	*Sphenomorphus*	*melanopogon*	Wetar	-7.92487	126.4079	Wetar	X
JAM13667	*Sphenomorphus*	*melanopogon*	Wetar	-7.92487	126.4079	Wetar	
JAM13668	*Sphenomorphus*	*melanopogon*	Wetar	-7.92487	126.4079	Wetar	X
JAM13669	*Sphenomorphus*	*melanopogon*	Wetar	-7.92487	126.4079	Wetar	X
JAM13670	*Sphenomorphus*	*melanopogon*	Wetar	-7.92487	126.4079	Wetar	X
JAM13680	*Sphenomorphus*	*melanopogon*	Wetar	-7.92847	126.4078	Wetar	
JAM13854	*Sphenomorphus*	*melanopogon*	Wetar	-7.92696	126.408	Wetar	X
JAM13881	*Sphenomorphus*	*melanopogon*	Wetar	-7.92481	126.4073	Wetar	
JAM13937	*Sphenomorphus*	*melanopogon*	Pantar	-8.35549	124.2544	Pantar	X
JAM13938	*Sphenomorphus*	*melanopogon*	Pantar	-8.35549	124.2544	Pantar	X
JAM13939	*Sphenomorphus*	*melanopogon*	Pantar	-8.35124	124.2507	Pantar	
JAM13973	*Sphenomorphus*	*melanopogon*	Pantar	-8.35268	124.2516	Pantar	
JAM13974	*Sphenomorphus*	*melanopogon*	Pantar	-8.35268	124.2516	Pantar	
JAM13975	*Sphenomorphus*	*melanopogon*	Pantar	-8.35268	124.2516	Pantar	
JAM13976	*Sphenomorphus*	*melanopogon*	Pantar	-8.35124	124.2507	Pantar	
JAM13977	*Sphenomorphus*	*melanopogon*	Pantar	-8.35124	124.2507	Pantar	X
JAM13978	*Sphenomorphus*	*melanopogon*	Pantar	-8.35392	124.2535	Pantar	X
JAM14017	*Sphenomorphus*	*melanopogon*	Pantar	-8.35268	124.2516	Pantar	X
JAM14018	*Sphenomorphus*	*melanopogon*	Pantar	-8.35268	124.2516	Pantar	X
JAM14019	*Sphenomorphus*	*melanopogon*	Pantar	-8.35268	124.2516	Pantar	
JAM14020	*Sphenomorphus*	*melanopogon*	Pantar	-8.35268	124.2516	Pantar	
JAM14021	*Sphenomorphus*	*melanopogon*	Pantar	-8.35124	124.2507	Pantar	
JAM14022	*Sphenomorphus*	*melanopogon*	Pantar	-8.35124	124.2507	Pantar	
JAM14023	*Sphenomorphus*	*melanopogon*	Pantar	-8.35124	124.2507	Pantar	
JAM14024	*Sphenomorphus*	*melanopogon*	Pantar	-8.35268	124.2516	Pantar	
JAM14025	*Sphenomorphus*	*melanopogon*	Pantar	-8.35268	124.2516	Pantar	
JAM14026	*Sphenomorphus*	*melanopogon*	Pantar	-8.35392	124.2535	Pantar	
JAM14038	*Sphenomorphus*	*melanopogon*	Pantar	-8.35268	124.2516	Pantar	
JAM14039	*Sphenomorphus*	*melanopogon*	Pantar	-8.35268	124.2516	Pantar	
JAM14063	*Sphenomorphus*	*melanopogon*	Pantar	-8.35124	124.2507	Pantar	
JAM14064	*Sphenomorphus*	*melanopogon*	Pantar	-8.35268	124.2516	Pantar	
USNM579230	*Sphenomorphus*	*melanopogon*	Timor	-8.7813	125.4547	Timor	X
USNM579234	*Sphenomorphus*	*melanopogon*	Timor	-8.7813	125.4547	Timor	X
USNM579359	*Sphenomorphus*	*melanopogon*	Timor	-9.0333	126.0785	Timor	
USNM579360	*Sphenomorphus*	*melanopogon*	Timor	-9.0333	126.0785	Timor	
USNM579477	*Sphenomorphus*	*melanopogon*	Timor	-8.4191	126.9783	Timor	
USNM579478	*Sphenomorphus*	*melanopogon*	Timor	-8.4191	126.9783	Timor	X
USNM579482	*Sphenomorphus*	*melanopogon*	Timor	-8.4763	127.1749	Timor	
USNM579483	*Sphenomorphus*	*melanopogon*	Timor	-8.4772	127.1748	Timor	
USNM579484	*Sphenomorphus*	*melanopogon*	Timor	-8.4189	127.1358	Timor	
USNM579485	*Sphenomorphus*	*melanopogon*	Timor	-8.4189	127.1358	Timor	

Table 1 cont.

Sample Number	Genus	Species	Island	Latitude	Longitude	MtDNA Clade	ExonCap
USNM579486	*Sphenomorphus*	*melanopogon*	Timor	-8.4189	127.1358	Timor	
USNM579487	*Sphenomorphus*	*melanopogon*	Timor	-8.7813	125.4547	Timor	
USNM579488	*Sphenomorphus*	*melanopogon*	Timor	-8.7812	125.4548	Timor	
USNM579489	*Sphenomorphus*	*melanopogon*	Timor	-8.7813	125.4547	Timor	
USNM579765	*Sphenomorphus*	*melanopogon*	Timor	-8.783	125.455	Timor	X
USNM579766	*Sphenomorphus*	*melanopogon*	Timor	-8.783	125.455	Timor	
USNM580529	*Sphenomorphus*	*melanopogon*	Timor	-9.0333	126.0785	Timor	
USNM580530	*Sphenomorphus*	*melanopogon*	Timor	-9.0333	126.0785	Timor	
USNM580534	*Sphenomorphus*	*melanopogon*	Timor	-9.0317	126.0767	Timor	
USNM580535	*Sphenomorphus*	*melanopogon*	Timor	-9.0317	126.0767	Timor	
USNM580536	*Sphenomorphus*	*melanopogon*	Timor	-9.0317	126.0767	Timor	
USNM580537	*Sphenomorphus*	*melanopogon*	Timor	-9.0317	126.0767	Timor	
USNM580538	*Sphenomorphus*	*melanopogon*	Timor	-9.0317	126.0767	Timor	X
USNM580539	*Sphenomorphus*	*melanopogon*	Timor	-8.7812	125.4548	Timor	
USNM580540	*Sphenomorphus*	*melanopogon*	Timor	-8.6088	126.3823	Timor	X
USNM580541	*Sphenomorphus*	*melanopogon*	Timor	-8.6088	126.3823	Timor	
USNM581140	*Sphenomorphus*	*melanopogon*	Timor	-8.7811	125.4548	Timor	
USNM581141	*Sphenomorphus*	*melanopogon*	Timor	-8.7811	125.4548	Timor	
MCZ-R-192865	*Sphenomorphus*	*melanopogon*	Jaco	-8.41536	127.3126	Timor	X
MCZ-R-192866	*Sphenomorphus*	*melanopogon*	Jaco	-8.41536	127.3126	Timor	X
MCZ-R-192918	*Sphenomorphus*	*melanopogon*	Timor	-8.41553	126.9839	Timor	X
MCZ-R-192919	*Sphenomorphus*	*melanopogon*	Timor	-8.41553	126.9839	Timor	X
ALS874	*Sphenomorphus*	*melanopogon*	Kur	-5.3434	131.995	Kur	X
ALS875	*Sphenomorphus*	*melanopogon*	Kur	-5.3434	131.995	Kur	X
ALS1016	*Sphenomorphus*	*melanopogon*	Kur	-5.3434	131.995	Kur	X
BRK65	*Sphenomorphus*	*melanopogon*	Banda Besar	-4.5541	129.928	Banda	X
BRK66	*Sphenomorphus*	*melanopogon*	Banda Besar	-4.5541	129.928	Banda	X
BRK67	*Sphenomorphus*	*melanopogon*	Banda Besar	-4.5541	129.928	Banda	
BRK141	*Sphenomorphus*	*melanopogon*	Ai	-4.5259	129.774	Banda	X
BRK142	*Sphenomorphus*	*melanopogon*	Ai	-4.5259	129.774	Banda	
ALS275	*Sphenomorphus*	*meyeri*	Aru	N/A	N/A	N/A	X
JAM11309	*Sphenomorphus*	*textus*	Sulawesi	N/A	N/A	N/A	X
CAS214892	*Sphenomorphus*	*indicus*	China	N/A	N/A	outgroup	
MVZ236749	*Sphenomorphus*	*indicus*	China	N/A	N/A	N/A	X
MVZ258402	*Sphenomorphus*	*maculatus*	Cambodia	N/A	N/A	N/A	X
KU308926	*Sphenomorphus*	*tagapayo*	Philippines	N/A	N/A	outgroup	

Table 2. Summary statistics for transcriptomes sequenced from three specimens of *Sphenomorphus melanopogon*.

Catalog #	Species	Locality	Total Length (bp)	Total Contigs	Mean Length (bp)	Max Contig Length (bp)	>2000bp	>1000bp	GC%
JAM12180	*S. melanopogon*	Sumbawa	13,773,597	8,832	1,559	16,953	2,422	5,203	48.74
JAM12480	*S. melanopogon*	Flores	15,999,020	9,185	1,741	16,982	2,943	5,891	48.83
JAM12652	*S. melanopogon*	Lembata	14,467,744	9,233	1,566	11,688	2,659	5,826	48.92

Table 3. Converted G-PHOCS demographic parameters for specific within-island lineage comparisons of *Sphenomorphus melanopogon*.

Parameter	Comparison	Mean	95% CI Low	95% CI High
Effective Population Size (Individuals)	Lombok 1	1,226,136	1,198,863	1,254,545
	Lombok 2	967,045	943,181	988,636
	Lombok Ancestor	1,297,727	1,203,409	1,392,045
Population Divergence Time (Years)	Lombok Ancestor	5,654,545	5,504,545	5,804,545
Migration Rate (Migrants per generation)	Lombok 1 -> Lombok 2	0.0002	0	0.0011
	Lombok 2 -> Lombok 1	0.0001	0	0.0003
Effective Population Size (Individuals)	South Sumba	2,206,818	2,146,591	2,267,045
	East Sumba	525,000	510,227	538,636
	Sumba Ancestor	837,500	794,318	880,681
Population Divergence Time (Years)	Sumba Ancestor	2,272,727	2,200,000	2,350,000
Migration Rate (Migrants per generation)	South Sumba -> East Sumba	0.0337	0.0127	0.0554
	East Sumba -> South Sumba	0.0356	0.0109	0.0630
Effective Population Size (Individuals)	West Flores	3,069,772	3,012,727	3,126,250
	East Flores 1	635,000	617,045	653,977
	East Flores 2	813,863	791,250	836,136
	East Flores Ancestor	647,727	503,636	769,886
	Flores Ancestor	1,156,818	1,075,000	1,242,272
Population Divergence Time (Years)	East Flores Ancestor	4,381,364	4,210,000	4,584,545
	Flores Ancestor	5,689,091	5,565,909	5,807,727
Migration Rate (Migrants per generation)	West Flores -> East Flores 1	0.0015	0	0.0044
	East Flores 1 -> West Flores	0.0009	0	0.0045
	West Flores -> East Flores 2	0.0052	0.0007	0.0106
	East Flores 2 -> West Flores	0.0027	0	0.0080
	East Flores 1 -> East Flores 2	0.0004	0	0.0020
	East Flores 2 -> East Flores 1	0.0051	0.0001	0.0113

FIGURE LEGENDS

Figure 1. A) *Sphenomorphus melanopogon* from Alor island. B) Ventral pigmentation and coloration of a male *Sphenomorphus melanopogon* from Alor island. C) Male *Sphenomorphus melanopogon* from Rinca island. Note the red coloration of the head. D) Female *Sphenomorphus melanopogon* from Rinca island. E) Attaching sticky traps to the base of a tree on Wetar island to capture *Sphenomorphus melanopogon* and other herpetofauna. F) Prime *Sphenomorphus melanopogon* habitat on Lombok island. There were approximately 8-10 *Sphenomorphus melanopogon* near the base of this large tree.

Figure 2. Maps of the distribution of *Sphenomorphus melanopogon*. A) The general region of Indonesia where *Sphenomorphus melanopogon* can be found. B) The range limits of *Sphenomorphus melanopogon* are enclosed by the yellow dashed line and notable island populations distant from the Lesser Sundas are shown with a yellow dot. Note the islands of Deli and Tinjil off of southwest Java. C) Map of the Lesser Sunda Islands with islands known to be inhabited by *Sphenomorphus melanopogon* shaded yellow and islands where they are not recorded from shaded gray.

Figure 3. Distribution of *Sphenomorphus melanopogon* samples used in exon-capture genetic analysis.

Figure 4. Maximum Likelihood phylogeny of the mitochondrial *ND4* gene produced in RAxML. Numbers at nodes represent bootstrap support. Colored symbols beside designated lineages correspond to the symbols in map Figure 6.

Figure 5. Time-calibrated Bayesian phylogeny of the mitochondrial *ND4* gene produced in BEAST. Numbers at nodes represent posterior probability values. Colored symbols beside designated lineages correspond to the symbols in map Figure 6.

Figure 6. Map of major mitochondrial lineages within the Lesser Sunda Islands. Colored symbols correspond to clades designated in Figures 4 and 5.

Figure 7. The average coverage of the targeted exons for each library. Coverage is calculated after duplicated reads are removed.

Figure 8. The average coverage of the flanking regions for each library. Coverage is calculated after duplicated reads are removed.

Figure 9. Alignment length summary plots for the 1,153 nuclear loci plotting the number of informative sites (A), the number of individuals (B), and the percentage of gaps (C) for each alignment length.

Figure 10. Frequency bar plots showing the number of taxa per alignment (A), the distribution of sequence lengths (B), the percent of informative sites per gene (C), and the percent missing data per gene (D).

Figure 11. Maximum Likelihood phylogeny of the concatenated nuclear dataset produced in RAxML. Numbers at nodes represent bootstrap support and colored bars for clades correspond to the map.

Figure 12. Supertree of 1,153 individual RAxML gene trees produced by ASTRAL. Colors correspond to the map above. The plot on the lower right represents Robinson-Foulds % distances for ASTRAL supertrees using subsets of 5, 10, 20, 50, 100, 200, and 500 loci replicated 10 times each. In general, the lower the Robinson-Foulds % distance the more similar the topology is to the full 1,153 gene ASTRAL supertree.

Figure 13. SNP-based species tree produced by SNAPP utilizing one informative SNP per gene.

Figure 14. Comparisons of STRUCTURE results from K=1 up to K=15 shown as A) Delta K values and B) the mean of the Ln probability of the data.

Figure 15. STRUCTURE population clustering results for K=2 up to K=9.

Figure 16. Unconverted demographic parameter estimate distributions produced by G-PHOCS analyses of the flanking sequence data. A) Marginal probability distributions for effective population size of Lombok 1 (black), Lombok 2 (purple), and ancestral lineages (red). B) Marginal probability distributions for effective population size of the South Sumba (black), East Sumba (purple), and ancestral lineages (red). C) Marginal probability distributions for the population divergence time between the Lombok 1 and Lombok 2 lineages. D) Marginal probability distributions for the population divergence time between the South Sumba and East Sumba lineages. E) Estimates of the migration rates between the Lombok lineages. F) Estimates of the migration rates from the South to East Sumba lineage (black) and from the East to South Sumba lineages (purple).

Figure 17. Unconverted demographic parameter estimate distributions produced by G-PHOCS analyses of the flanking sequence data for the Flores lineages. A) Marginal probability distributions for effective population size of West Flores (black), East Flores 1 (purple), East Flores 2 (red), East Flores Ancestor (orange), and Flores Ancestor (green) lineages. B) Marginal probability distributions for the population divergence time between the East Flores lineages (black) and between West Flores and the East Flores Ancestor (purple). C) Estimates of the migration rates between lineages.

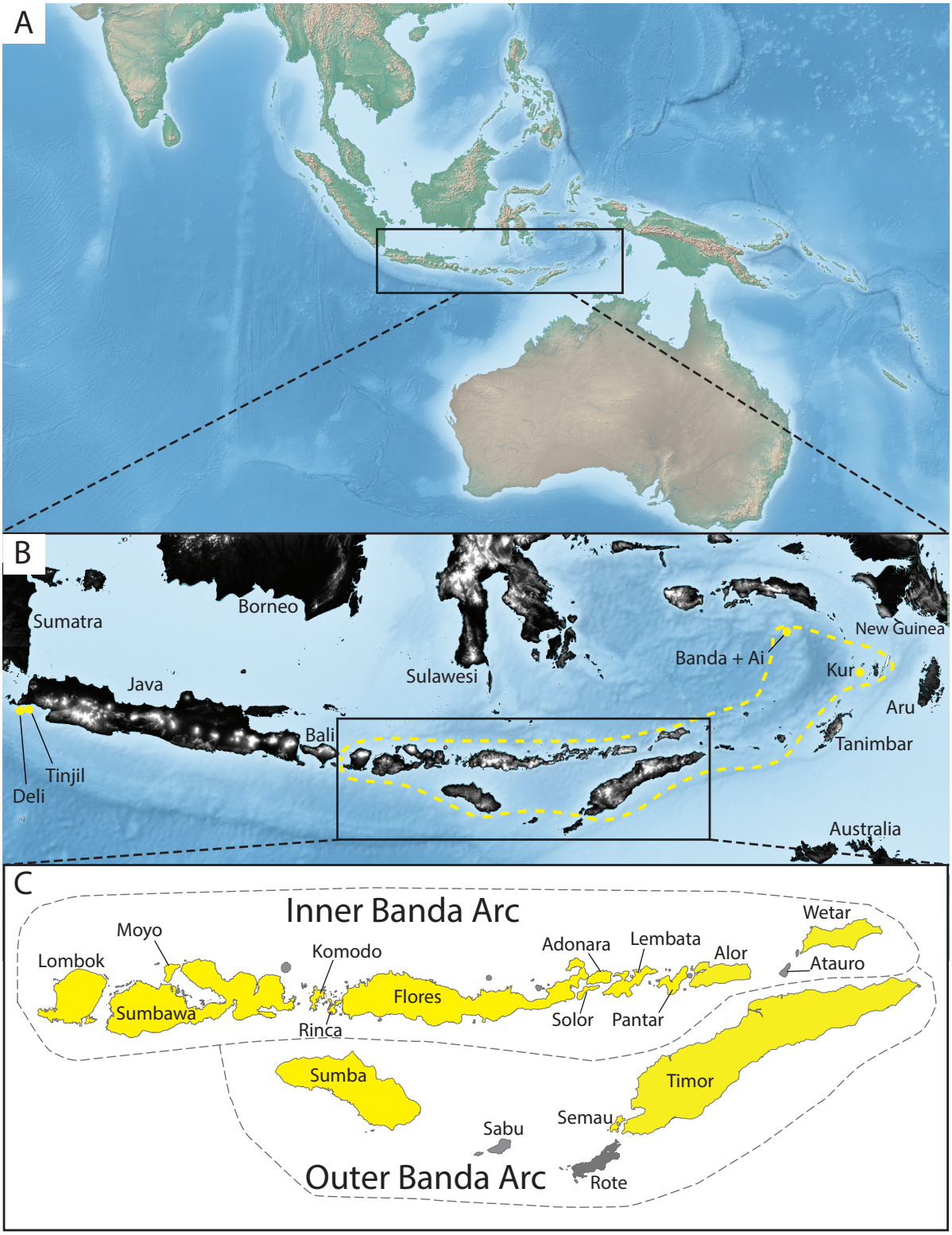

A
B
Sumatra
Borneo
Java
Sulawesi
Bali
Deli
Tinjil
Banda + Ai
New Guinea
Kur
Aru
Tanimbar
Australia
C
Inner Banda Arc
Lombok
Moyo
Sumbawa
Komodo
Rinca
Flores
Adonara
Lembata
Alor
Solor
Pantar
Wetar
Atauro
Sumba
Timor
Semau
Sabu
Rote
Outer Banda Arc

Figure 1.

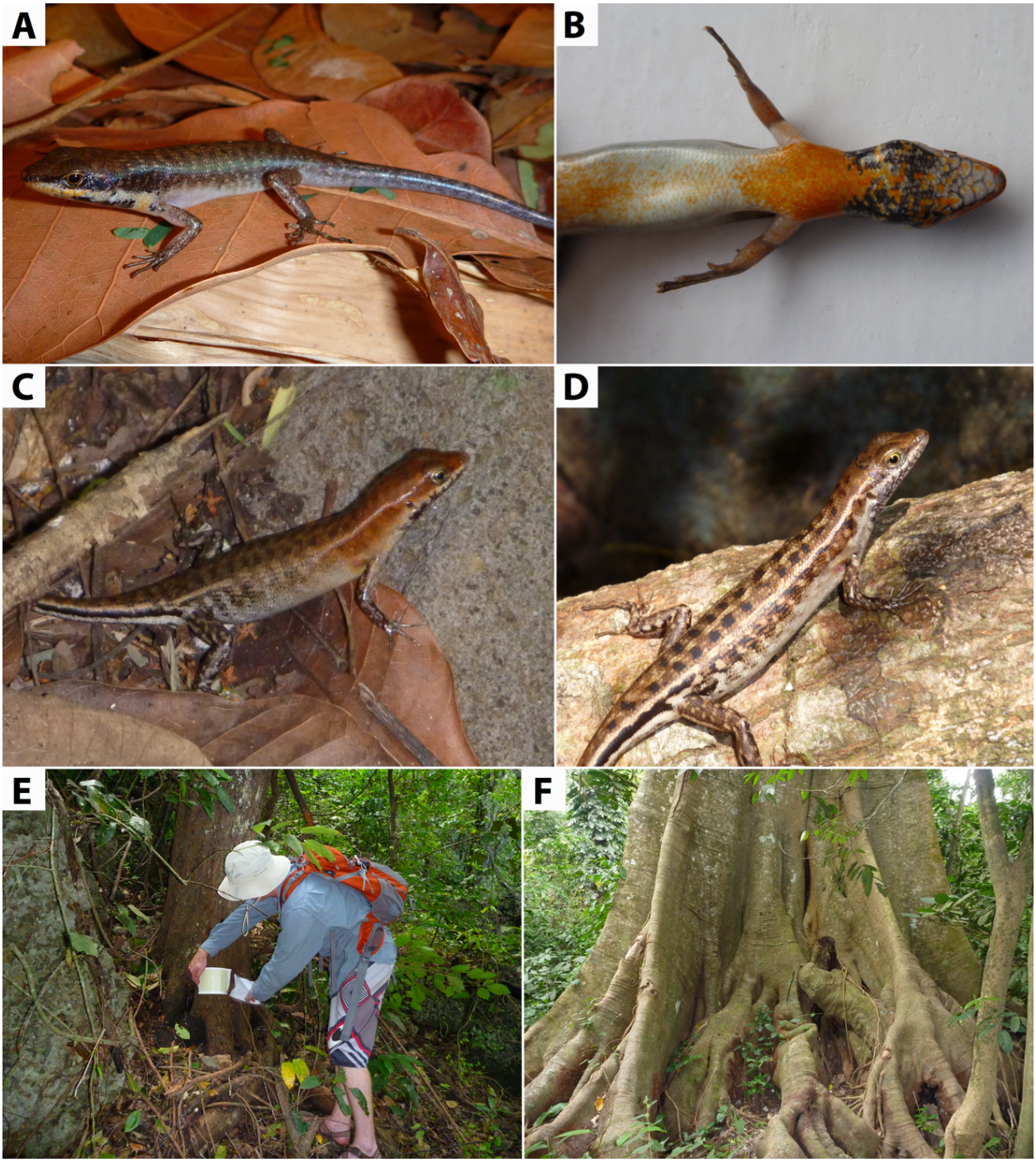


Figure 2.

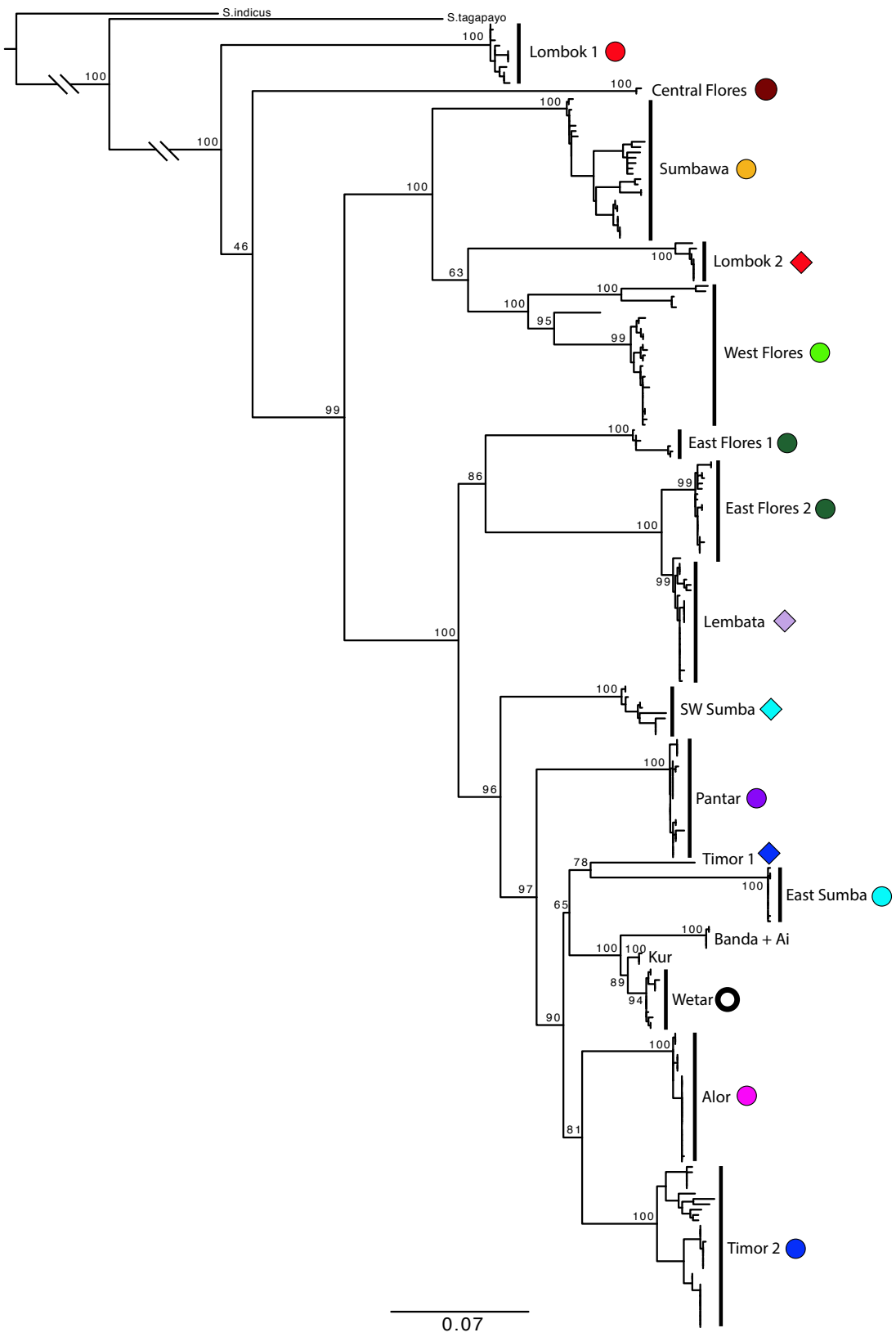

S.indicus
S.tagapayo
Lombok 1
Central Flores
Sumbawa
Lombok 2
West Flores
East Flores 1
East Flores 2
Lembata
SW Sumba
Pantar
Timor 1
East Sumba
Banda + Ai
Kur
Wetar
Alor
Timor 2
0.07

Figure 3.

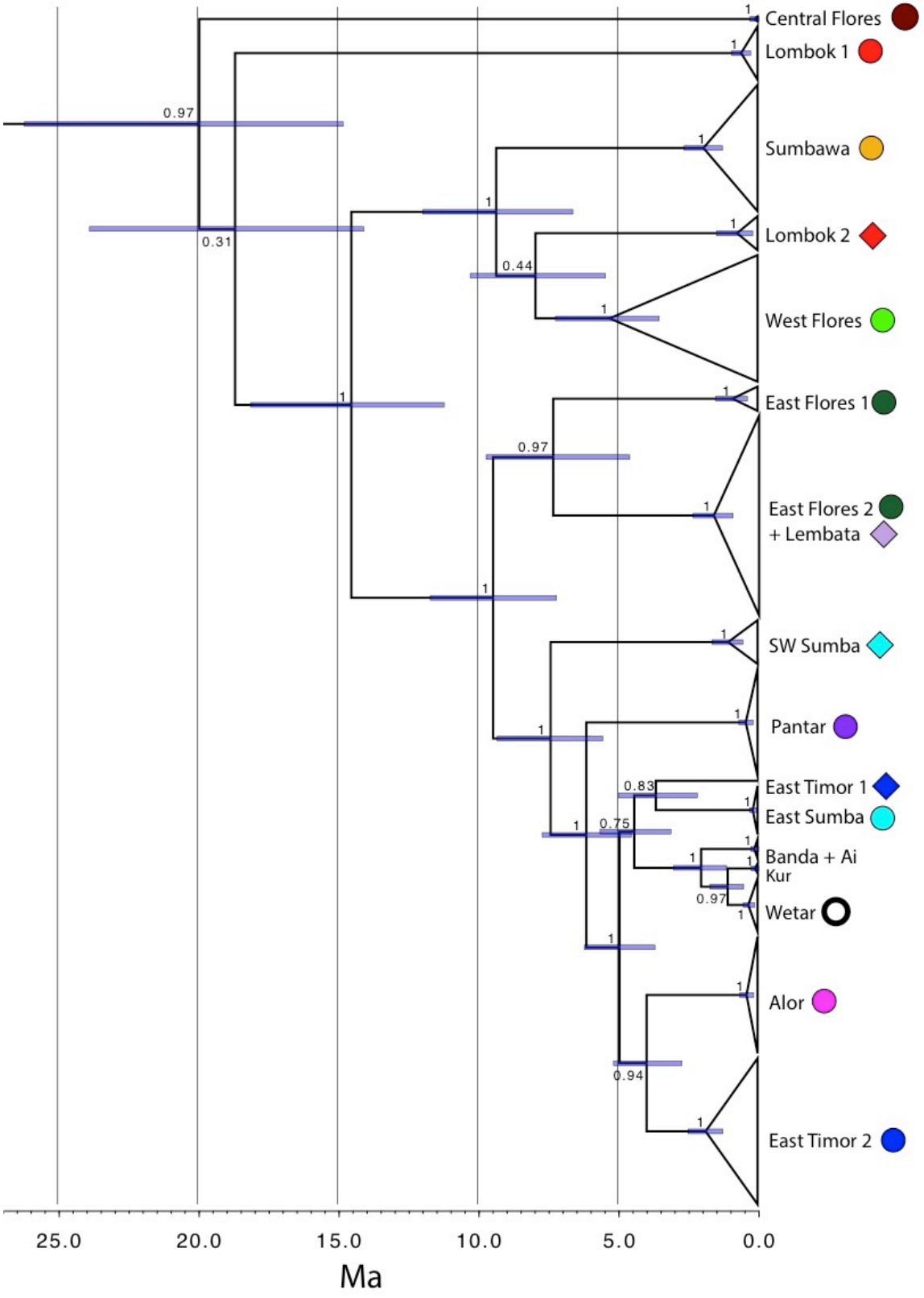

Central Flores
Lombok 1
Sumbawa
Lombok 2
West Flores
East Flores 1
East Flores 2
+ Lembata
SW Sumba
Pantar
East Timor 1
East Sumba
Banda + Ai
Kur
Wetar
Alor
East Timor 2
0.97
0.31
0.44
0.97
0.83
0.75
0.97
0.94
25.0
20.0
15.0
10.0
5.0
0.0
Ma

Figure 4.

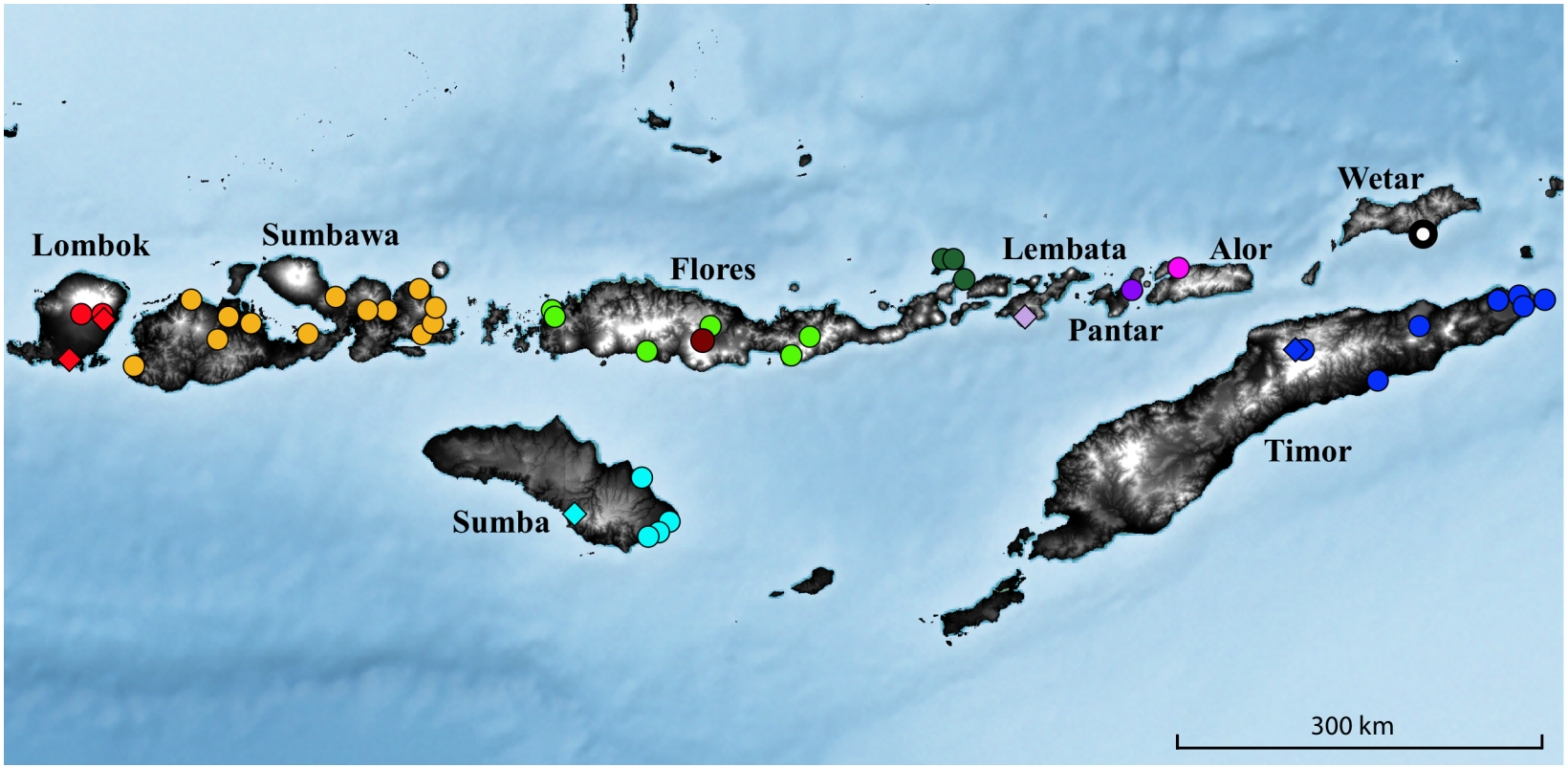

Lombok
Sumbawa
Flores
Lembata
Alor
Wetar
Pantar
Timor
Sumba
300 km

Figure 5.

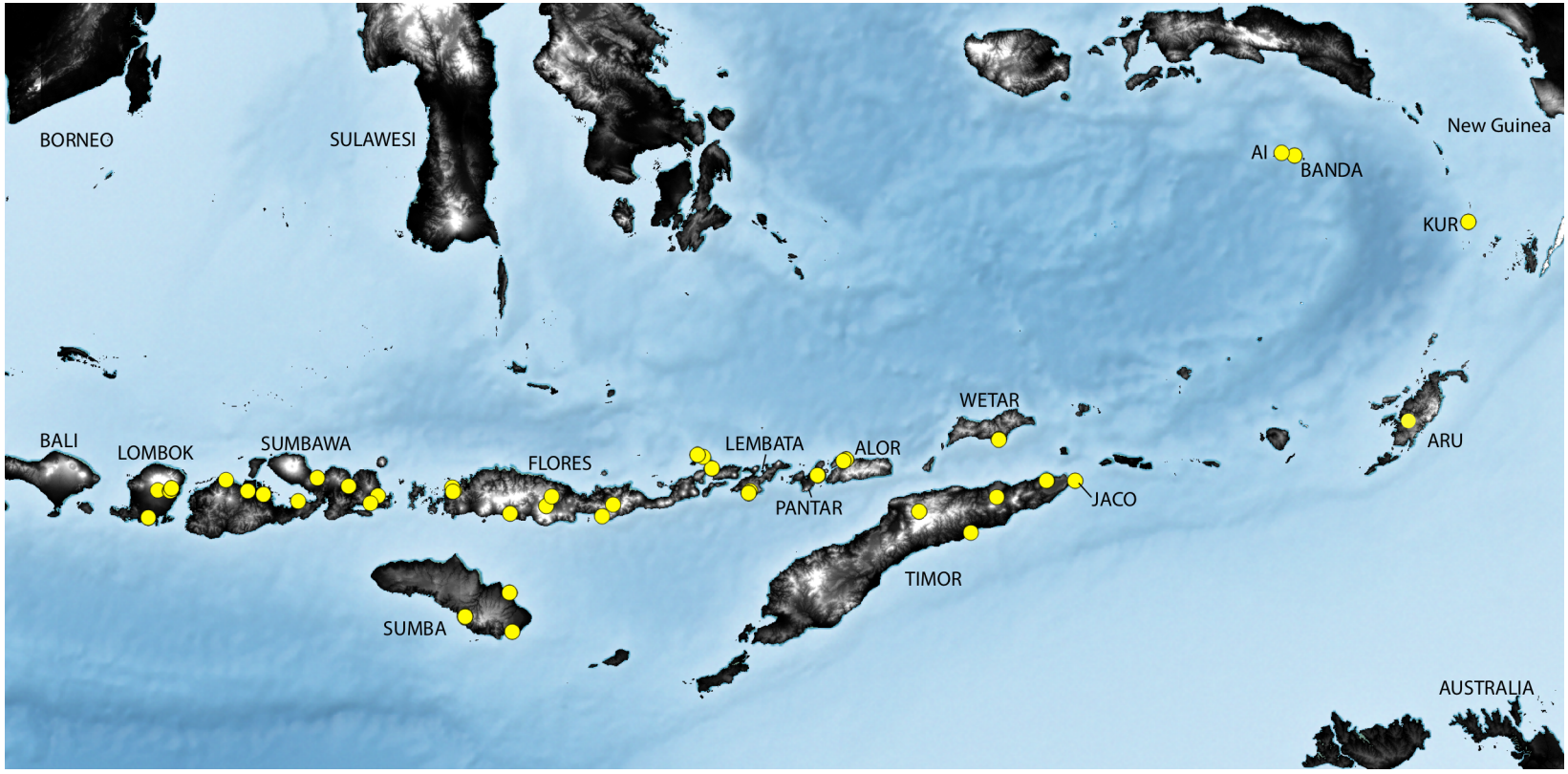

BORNEO
SULAWESI
New Guinea
AI
BANDA
KUR
BALI
LOMBOK
SUMBAWA
FLORES
LEMBATA
ALOR
WETAR
PANTAR
JACO
ARU
TIMOR
SUMBA
AUSTRALIA

Figure 6.

Figure 7.

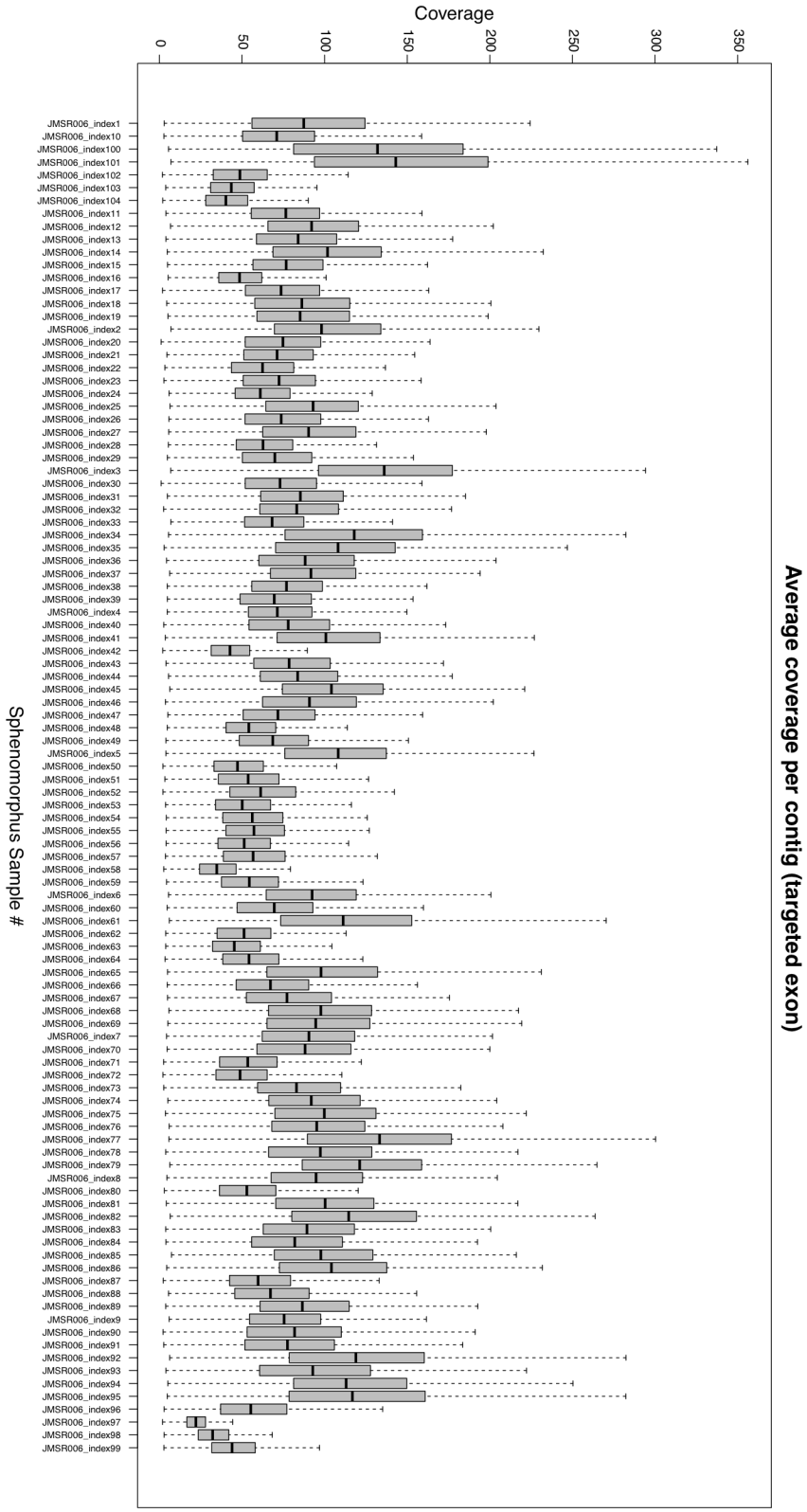
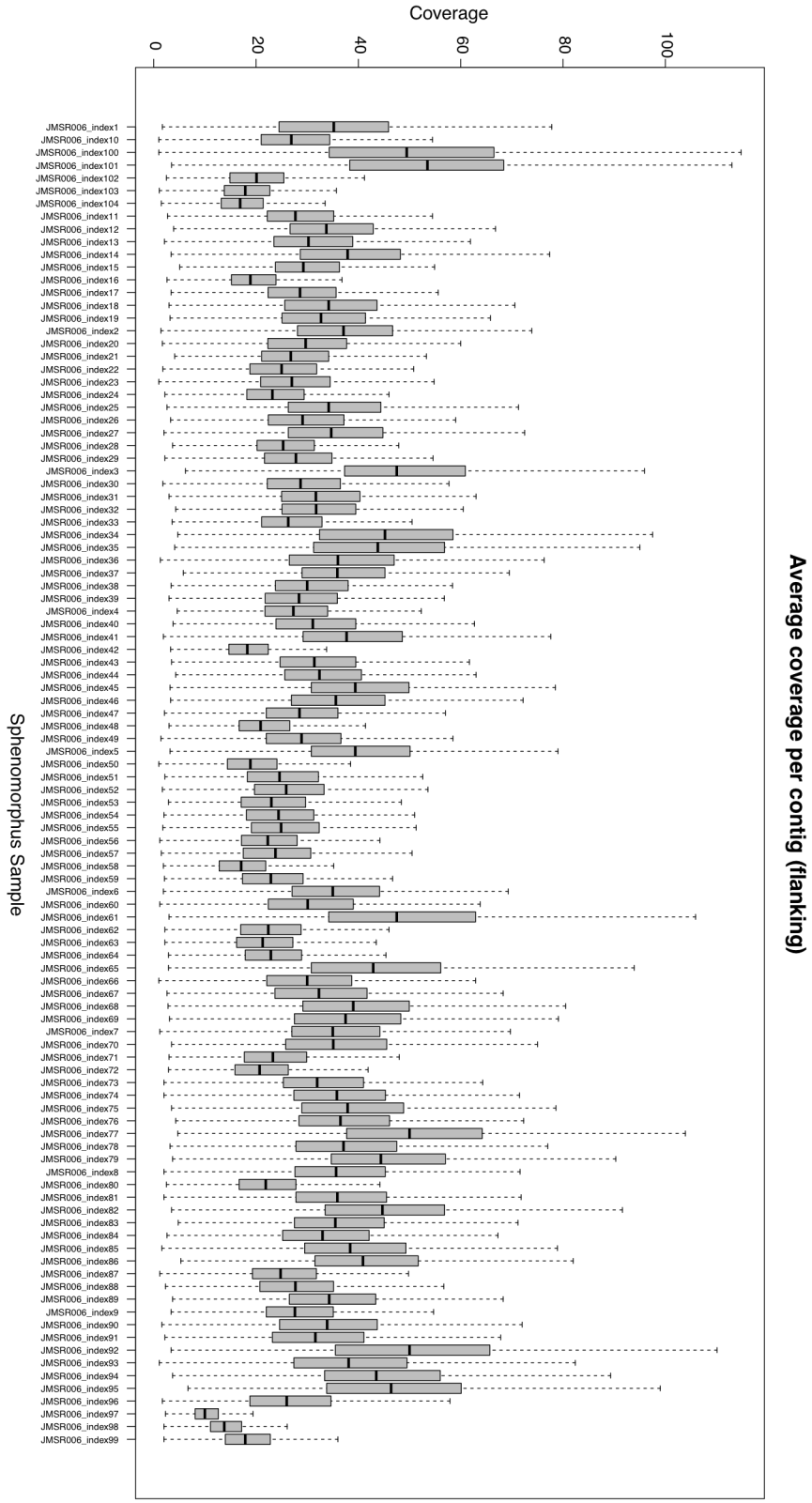

Average coverage per contig (targeted exon)
Coverage
0
50
100
150
200
250
300
350
Sphenomorphus Sample #
JMSR006_index1
JMSR006_index10
JMSR006_index100
JMSR006_index101
JMSR006_index102
JMSR006_index103
JMSR006_index104
JMSR006_index11
JMSR006_index12
JMSR006_index13
JMSR006_index14
JMSR006_index15
JMSR006_index16
JMSR006_index17
JMSR006_index18
JMSR006_index19
JMSR006_index2
JMSR006_index20
JMSR006_index21
JMSR006_index22
JMSR006_index23
JMSR006_index24
JMSR006_index25
JMSR006_index26
JMSR006_index27
JMSR006_index28
JMSR006_index29
JMSR006_index3
JMSR006_index30
JMSR006_index31
JMSR006_index32
JMSR006_index33
JMSR006_index34
JMSR006_index35
JMSR006_index36
JMSR006_index37
JMSR006_index38
JMSR006_index39
JMSR006_index4
JMSR006_index40
JMSR006_index41
JMSR006_index42
JMSR006_index43
JMSR006_index44
JMSR006_index45
JMSR006_index46
JMSR006_index47
JMSR006_index48
JMSR006_index49
JMSR006_index5
JMSR006_index50
JMSR006_index51
JMSR006_index52
JMSR006_index53
JMSR006_index54
JMSR006_index55
JMSR006_index56
JMSR006_index57
JMSR006_index58
JMSR006_index59
JMSR006_index6
JMSR006_index60
JMSR006_index61
JMSR006_index62
JMSR006_index63
JMSR006_index64
JMSR006_index65
JMSR006_index66
JMSR006_index67
JMSR006_index68
JMSR006_index69
JMSR006_index7
JMSR006_index70
JMSR006_index71
JMSR006_index72
JMSR006_index73
JMSR006_index74
JMSR006_index75
JMSR006_index76
JMSR006_index77
JMSR006_index78
JMSR006_index79
JMSR006_index8
JMSR006_index80
JMSR006_index81
JMSR006_index82
JMSR006_index83
JMSR006_index84
JMSR006_index85
JMSR006_index86
JMSR006_index87
JMSR006_index88
JMSR006_index89
JMSR006_index9
JMSR006_index90
JMSR006_index91
JMSR006_index92
JMSR006_index93
JMSR006_index94
JMSR006_index95
JMSR006_index96
JMSR006_index97
JMSR006_index98
JMSR006_index99

Figure 8.


Coverage
0
20
40
60
80
100
JMSR006_index1
JMSR006_index10
JMSR006_index100
JMSR006_index101
JMSR006_index102
JMSR006_index103
JMSR006_index104
JMSR006_index11
JMSR006_index12
JMSR006_index13
JMSR006_index14
JMSR006_index15
JMSR006_index16
JMSR006_index17
JMSR006_index18
JMSR006_index19
JMSR006_index2
JMSR006_index20
JMSR006_index21
JMSR006_index22
JMSR006_index23
JMSR006_index24
JMSR006_index25
JMSR006_index26
JMSR006_index27
JMSR006_index28
JMSR006_index29
JMSR006_index3
JMSR006_index30
JMSR006_index31
JMSR006_index32
JMSR006_index33
JMSR006_index34
JMSR006_index35
JMSR006_index36
JMSR006_index37
JMSR006_index38
JMSR006_index39
JMSR006_index4
JMSR006_index40
JMSR006_index41
JMSR006_index42
JMSR006_index43
JMSR006_index44
JMSR006_index45
JMSR006_index46
JMSR006_index47
JMSR006_index48
JMSR006_index49
JMSR006_index5
JMSR006_index50
JMSR006_index51
JMSR006_index52
JMSR006_index53
JMSR006_index54
JMSR006_index55
JMSR006_index56
JMSR006_index57
JMSR006_index58
JMSR006_index59
JMSR006_index6
JMSR006_index60
JMSR006_index61
JMSR006_index62
JMSR006_index63
JMSR006_index64
JMSR006_index65
JMSR006_index66
JMSR006_index67
JMSR006_index68
JMSR006_index69
JMSR006_index7
JMSR006_index70
JMSR006_index71
JMSR006_index72
JMSR006_index73
JMSR006_index74
JMSR006_index75
JMSR006_index76
JMSR006_index77
JMSR006_index78
JMSR006_index79
JMSR006_index8
JMSR006_index80
JMSR006_index81
JMSR006_index82
JMSR006_index83
JMSR006_index84
JMSR006_index85
JMSR006_index86
JMSR006_index87
JMSR006_index88
JMSR006_index89
JMSR006_index9
JMSR006_index90
JMSR006_index91
JMSR006_index92
JMSR006_index93
JMSR006_index94
JMSR006_index95
JMSR006_index96
JMSR006_index97
JMSR006_index98
JMSR006_index99
Sphenomorphus Sample
Average coverage per contig (flanking)

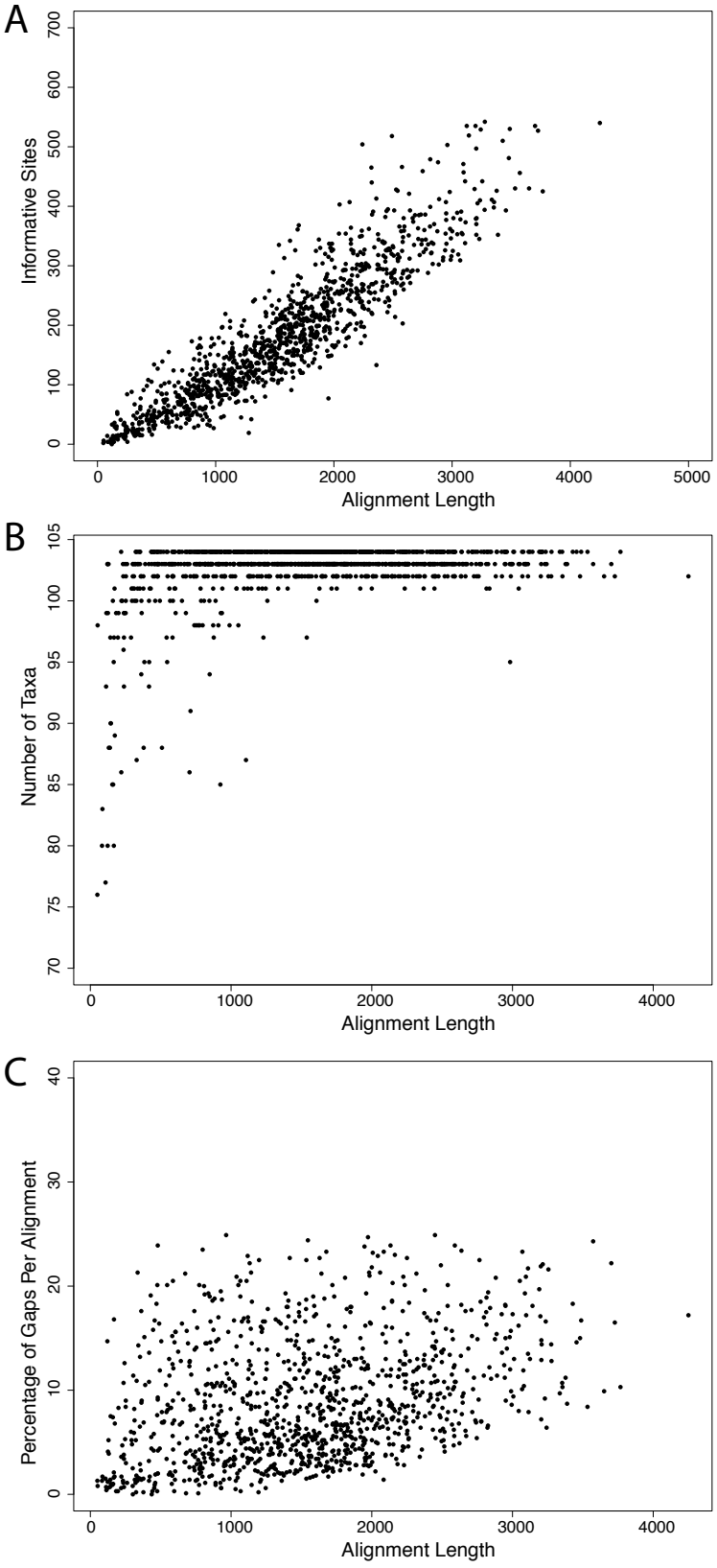

A
Informative Sites
0
100
200
300
400
500
600
700
0
1000
2000
3000
4000
5000
Alignment Length
B
Number of Taxa
70
75
80
85
90
95
100
105
0
1000
2000
3000
4000
Alignment Length
C
Percentage of Gaps Per Alignment
0
10
20
30
40
0
1000
2000
3000
4000
Alignment Length

Figure 9.

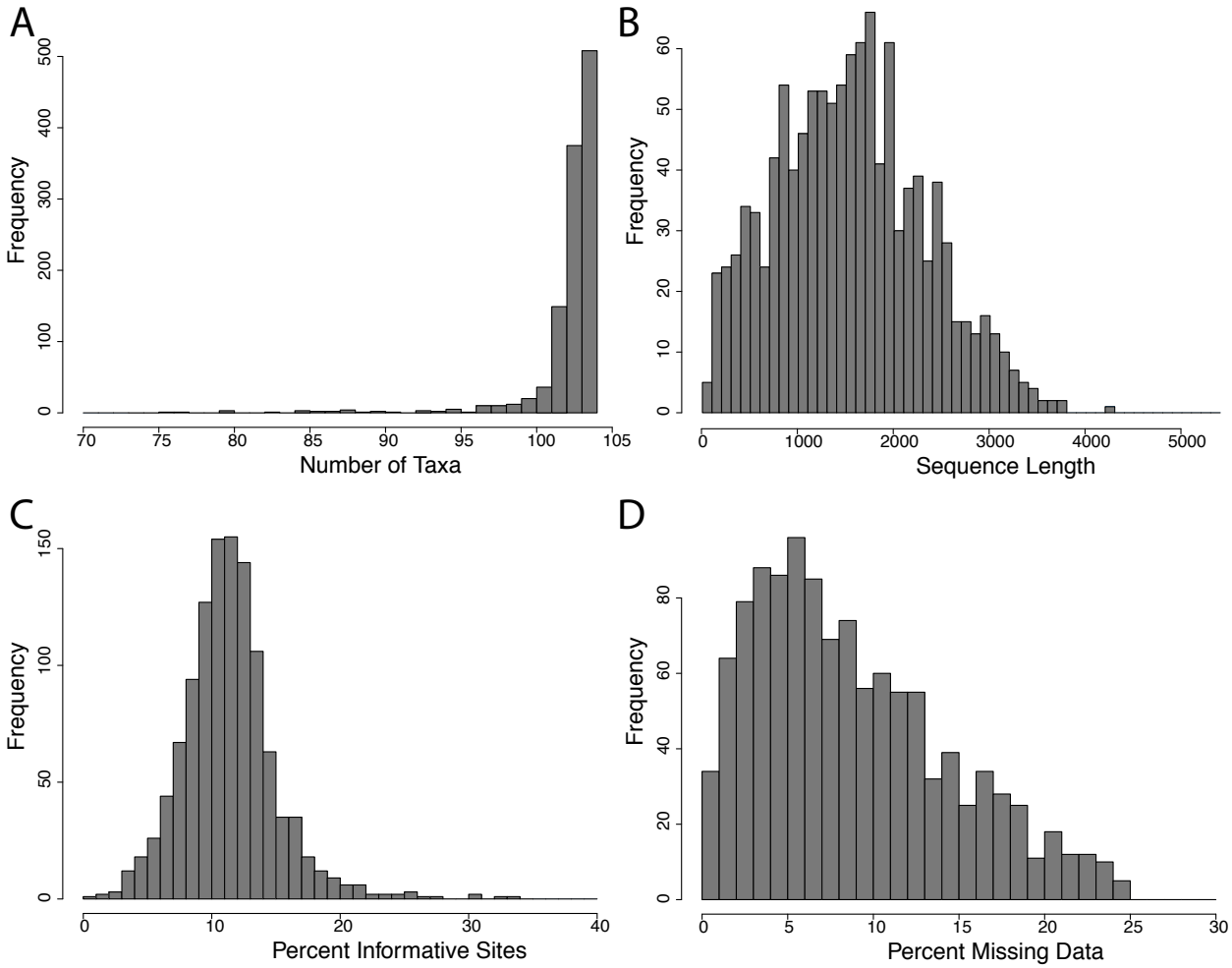

A
Frequency
0
100
200
300
400
500
70
75
80
85
90
95
100
105
Number of Taxa
B
Frequency
0
10
20
30
40
50
60
0
1000
2000
3000
4000
5000
Sequence Length
C
Frequency
0
50
100
150
0
10
20
30
40
Percent Informative Sites
D
Frequency
0
20
40
60
80
0
5
10
15
20
25
30
Percent Missing Data

Figure 10.

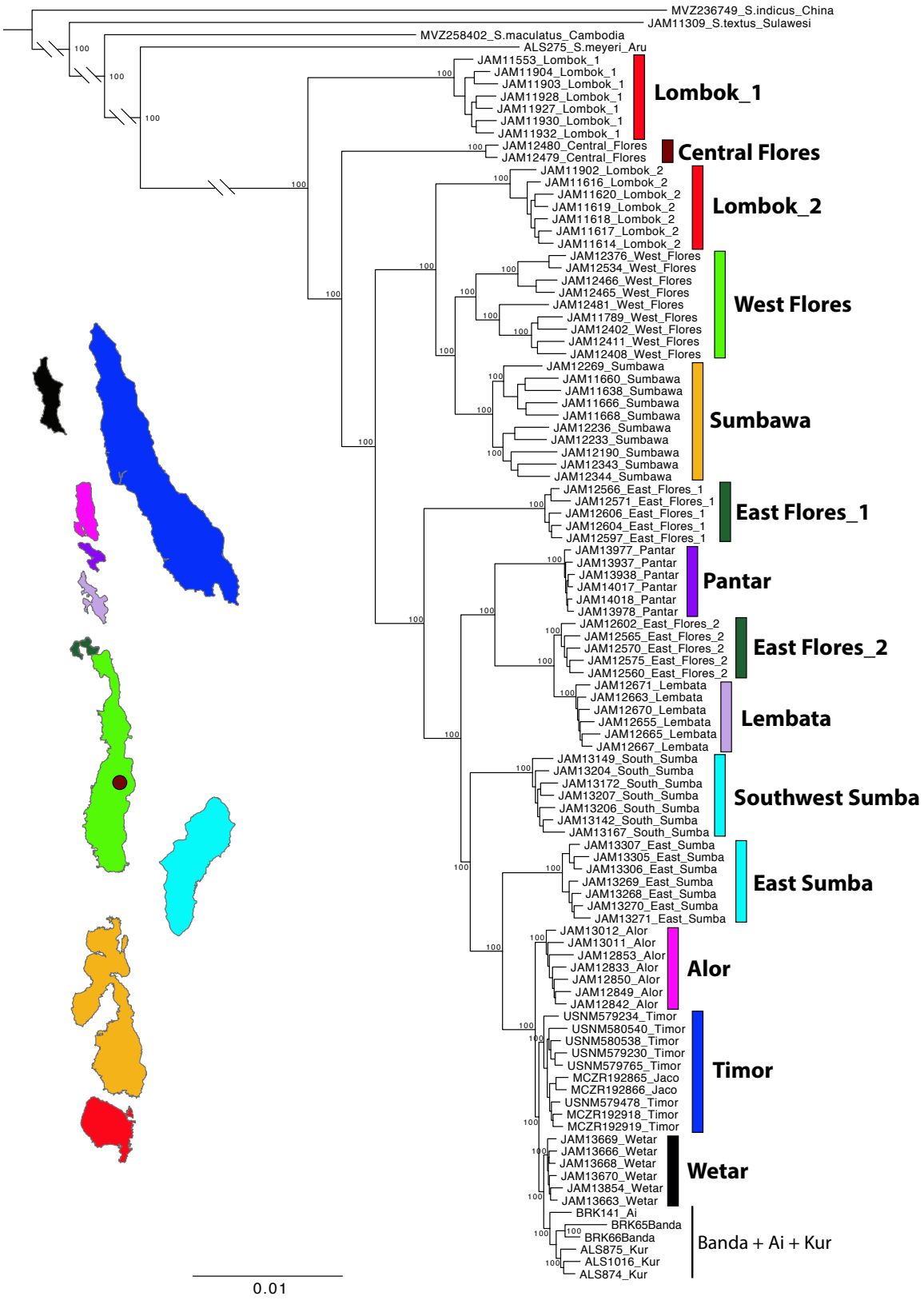

MVZ236749_S.indicus_China
JAM11309_S.textus_Sulawesi
MVZ258402_S.maculatus_Cambodia
ALS275_S.meyeri_Aru
JAM11553_Lombok_1
JAM11904_Lombok_1
JAM11903_Lombok_1
JAM11928_Lombok_1
JAM11927_Lombok_1
JAM11930_Lombok_1
JAM11932_Lombok_1
Lombok_1
JAM12480_Central_Flores
JAM12479_Central_Flores
Central Flores
JAM11902_Lombok_2
JAM11616_Lombok_2
JAM11620_Lombok_2
JAM11619_Lombok_2
JAM11618_Lombok_2
JAM11617_Lombok_2
JAM11614_Lombok_2
Lombok_2
JAM12376_West_Flores
JAM12534_West_Flores
JAM12466_West_Flores
JAM12465_West_Flores
JAM12481_West_Flores
JAM11789_West_Flores
JAM12402_West_Flores
JAM12411_West_Flores
JAM12408_West_Flores
West Flores
JAM12269_Sumbawa
JAM11660_Sumbawa
JAM11638_Sumbawa
JAM11666_Sumbawa
JAM11668_Sumbawa
JAM12236_Sumbawa
JAM12233_Sumbawa
JAM12190_Sumbawa
JAM12343_Sumbawa
JAM12344_Sumbawa
Sumbawa
JAM12566_East_Flores_1
JAM12571_East_Flores_1
JAM12606_East_Flores_1
JAM12604_East_Flores_1
JAM12597_East_Flores_1
East Flores_1
JAM13977_Pantar
JAM13937_Pantar
JAM13938_Pantar
JAM14017_Pantar
JAM14018_Pantar
JAM13978_Pantar
Pantar
JAM12602_East_Flores_2
JAM12565_East_Flores_2
JAM12570_East_Flores_2
JAM12575_East_Flores_2
JAM12560_East_Flores_2
East Flores_2
JAM12671_Lembata
JAM12663_Lembata
JAM12670_Lembata
JAM12655_Lembata
JAM12665_Lembata
JAM12667_Lembata
Lembata
JAM13149_South_Sumba
JAM13204_South_Sumba
JAM13172_South_Sumba
JAM13207_South_Sumba
JAM13206_South_Sumba
JAM13142_South_Sumba
JAM13167_South_Sumba
Southwest Sumba
JAM13307_East_Sumba
JAM13305_East_Sumba
JAM13306_East_Sumba
JAM13269_East_Sumba
JAM13268_East_Sumba
JAM13270_East_Sumba
JAM13271_East_Sumba
East Sumba
JAM13012_Alor
JAM13011_Alor
JAM12853_Alor
JAM12833_Alor
JAM12850_Alor
JAM12849_Alor
JAM12842_Alor
Alor
USNM579234_Timor
USNM580540_Timor
USNM580538_Timor
USNM579230_Timor
USNM579765_Timor
MCZR192865_Jaco
MCZR192866_Jaco
USNM579478_Timor
MCZR192918_Timor
MCZR192919_Timor
Timor
JAM13669_Wetar
JAM13666_Wetar
JAM13668_Wetar
JAM13670_Wetar
JAM13854_Wetar
JAM13663_Wetar
Wetar
BRK141_Ai
BRK65Banda
BRK66Banda
ALS875_Kur
ALS1016_Kur
ALS874_Kur
Banda + Ai + Kur
0.01

Figure 11.

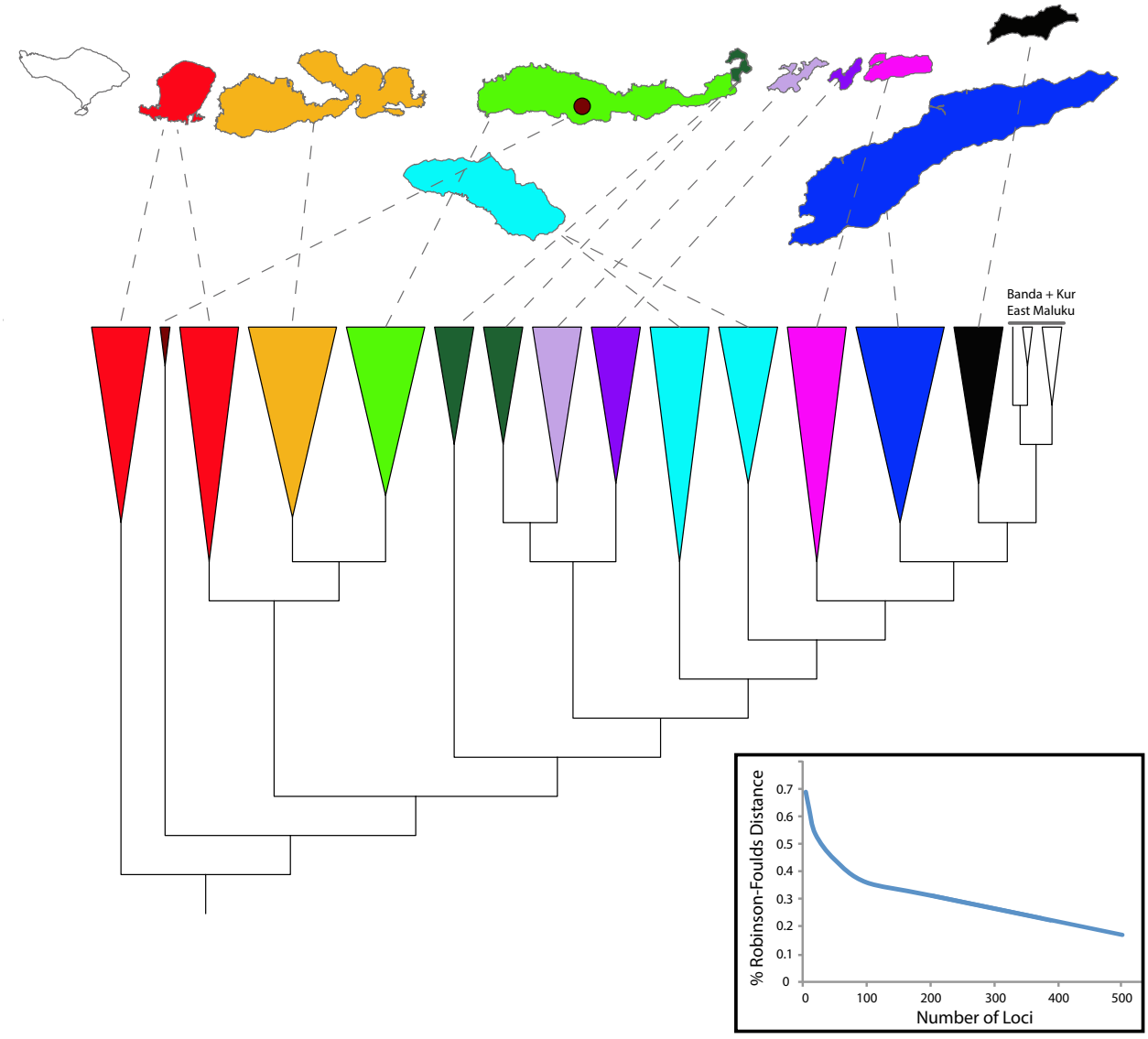

Banda + Kur
East Maluku
% Robinson-Foulds Distance
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
100
200
300
400
500
Number of Loci

Figure 12.

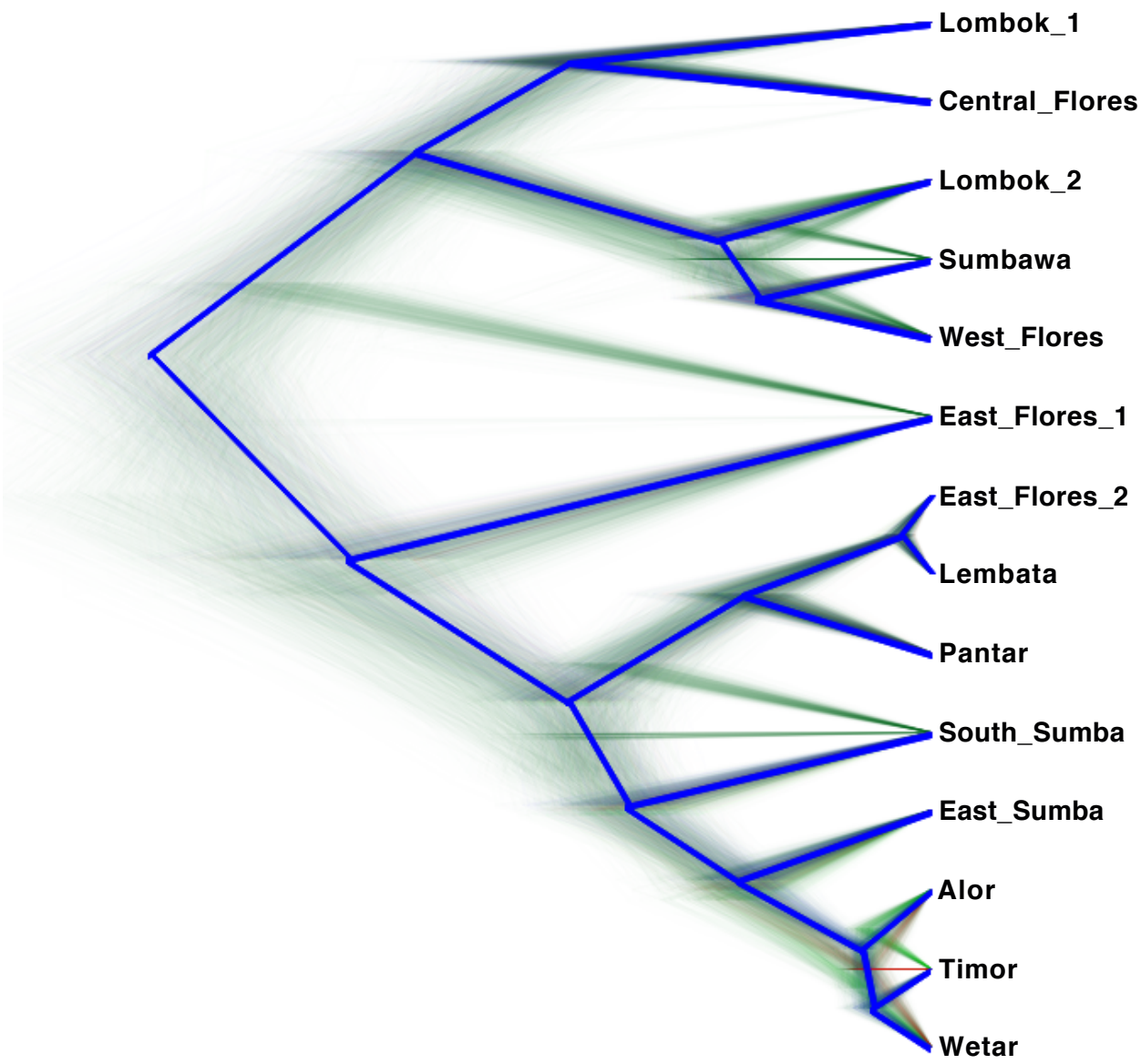

Lombok_1
Central_Flores
Lombok_2
Sumbawa
West_Flores
East_Flores_1
East_Flores_2
Lembata
Pantar
South_Sumba
East_Sumba
Alor
Timor
Wetar

Figure 13.

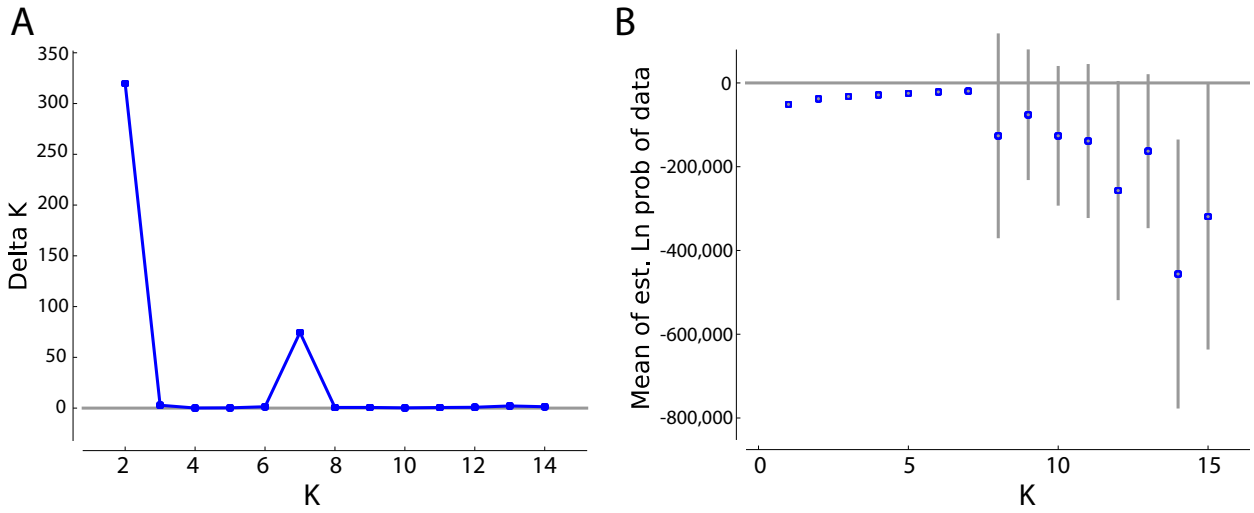

A
Delta K
350
300
250
200
150
100
50
0
2
4
6
8
10
12
14
K
B
Mean of est. Ln prob of data
0
-200,000
-400,000
-600,000
-800,000
0
5
10
15
K

Figure 14.

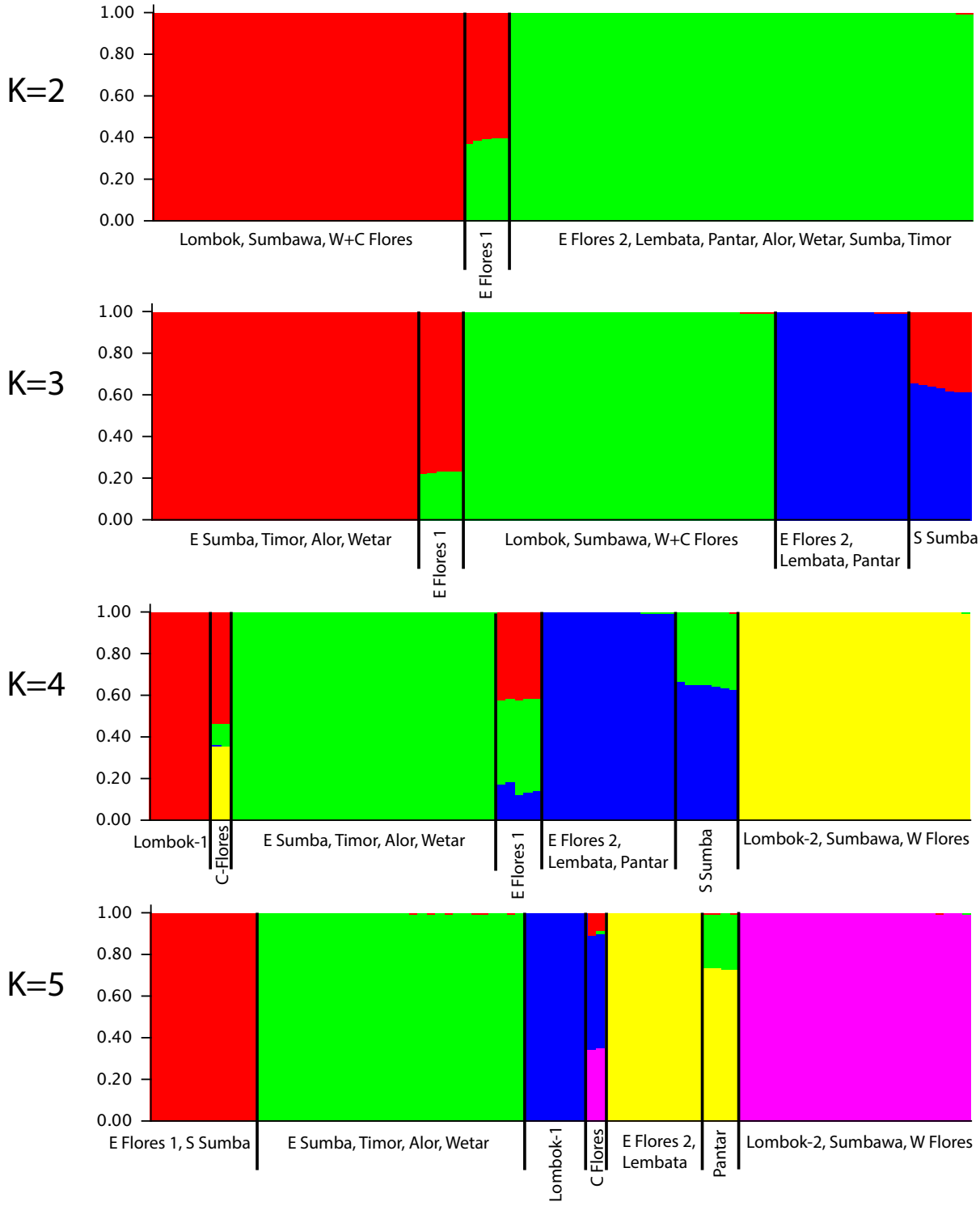

K=2
1.00
0.80
0.60
0.40
0.20
0.00
Lombok, Sumbawa, W+C Flores
E Flores 1
E Flores 2, Lembata, Pantar, Alor, Wetar, Sumba, Timor
K=3
1.00
0.80
0.60
0.40
0.20
0.00
E Sumba, Timor, Alor, Wetar
E Flores 1
Lombok, Sumbawa, W+C Flores
E Flores 2,
Lembata, Pantar
S Sumba
K=4
1.00
0.80
0.60
0.40
0.20
0.00
Lombok-1
C-Flores
E Sumba, Timor, Alor, Wetar
E Flores 1
E Flores 2,
Lembata, Pantar
S Sumba
Lombok-2, Sumbawa, W Flores
K=5
1.00
0.80
0.60
0.40
0.20
0.00
E Flores 1, S Sumba
E Sumba, Timor, Alor, Wetar
Lombok-1
C Flores
E Flores 2,
Lembata
Pantar
Lombok-2, Sumbawa, W Flores

Figure 15.

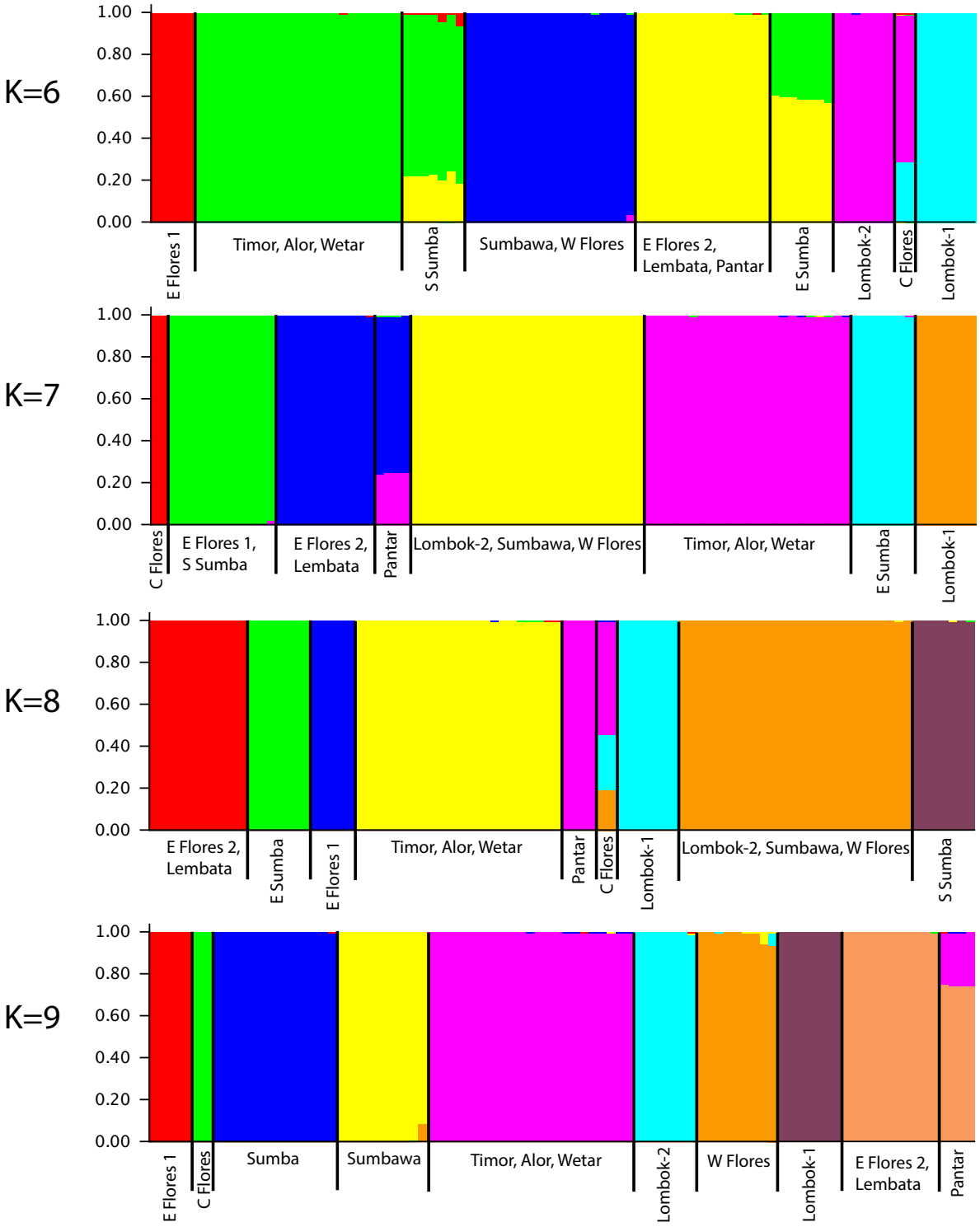


Figure 15 cont.

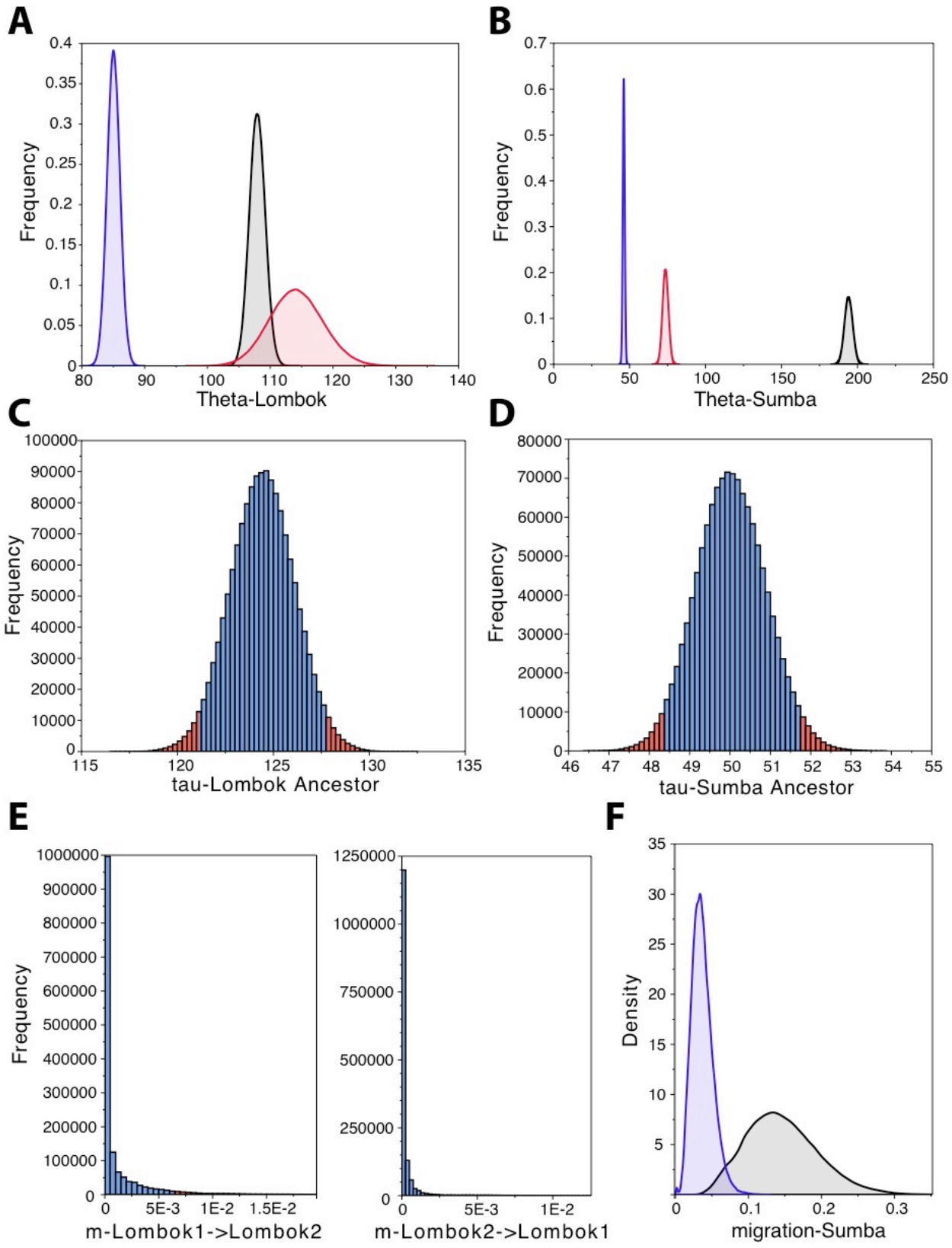

A
Frequency
Theta-Lombok
B
Frequency
Theta-Sumba
C
Frequency
tau-Lombok Ancestor
D
Frequency
tau-Sumba Ancestor
E
Frequency
m-Lombok1->Lombok2
m-Lombok2->Lombok1
F
Density
migration-Sumba

Figure 16.

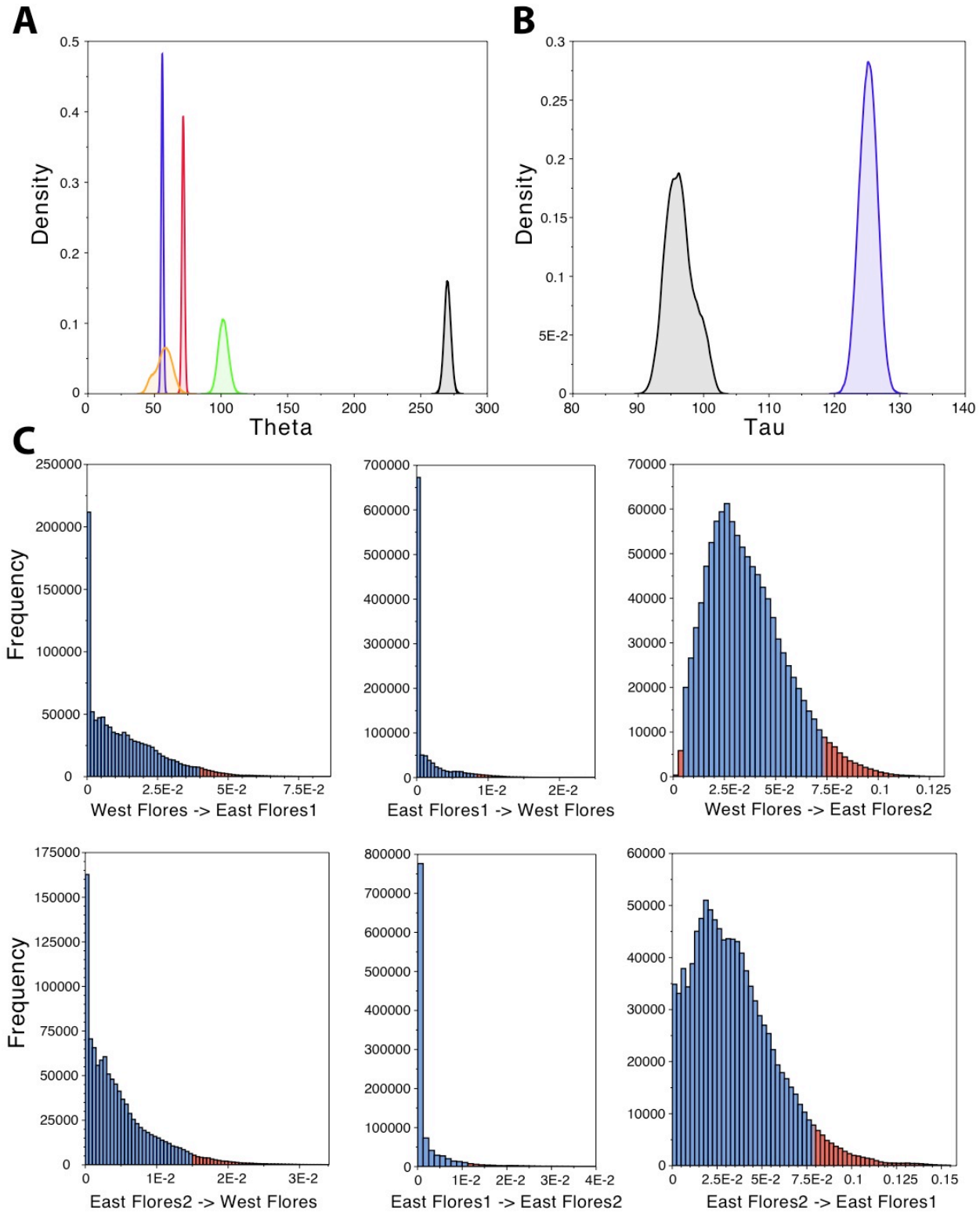

A
Density
Theta
B
Density
Tau
C
Frequency
West Flores -> East Flores1
East Flores1 -> West Flores
West Flores -> East Flores2
Frequency
East Flores2 -> West Flores
East Flores1 -> East Flores2
East Flores2 -> East Flores1

Figure 17.

CHAPTER 4

Biogeographical History of Fanged Frogs (Genus: *Limnonectes*) From the Lesser Sunda Islands of Indonesia: A Phylogenomic Approach

ABSTRACT

By studying the biogeographical history of taxa that have colonized isolated oceanic islands, we can elucidate how the timing and sequence of island colonization affects population and species divergence. Amphibians are poor dispersers over oceanic barriers due to their poor osmotic tolerance, yet they have colonized many oceanic islands close to mainland sources. Frogs of the genus *Limnonectes* are distributed throughout Southeast Asia and occur on many of the oceanic islands of the Wallacean region. The Lesser Sunda Islands occur as a parallel double-island arc that runs west-to-east between Bali and New Guinea, and are considered a distinct region both biologically and geologically. Two species (*L. dammermani* and *L. kadarsani*) inhabit the Inner Arc islands of the Lesser Sundas including Lombok, Sumbawa, Flores, Adonara, and Lembata. Previous analyses suggested that both of these Lesser Sundas species were derived from *L. microdiscus* on the neighboring islands of Java and Bali. Tissue samples and specimens of *Limnonectes* frogs were collected from multiple localities within the islands of Lombok, Sumbawa, Flores, and Lembata. We collected mitochondrial sequence data as well as exome-capture nuclear sequence data for this assemblage. Phylogenetic analysis of mtDNA data for 153 samples from the Lesser Sundas placed *L. dammermani* from Lombok as sister to *L. kadarsani* from Lombok, thereby rendering *L. kadarsani* paraphyletic. The Bayesian mtDNA phylogeny has a pectinate topology that is consistent with a stepping-stone model of island colonization, while the Maximum Likelihood topology suggests dispersal from Sumbawa to Lembata, then back to Flores. The exon-capture experiment successfully captured 974 nuclear genes from 48 *Limnonectes* samples with high coverage and low levels of missing data. Maximum Likelihood and coalescent phylogenetic analysis of the genomic dataset converged on a similar topology containing the same major lineages as the mtDNA phylogeny but differing with respect to relationships among those lineages. In contrast with the mtDNA tree, the genomic topology places *L. dammermani* and *L. kadarsani* as monophyletic sister taxa. Within the more widespread *L. kadarsani,* the Lombok lineage is basal while Sumbawa and Lembata lineages are nested within a Flores clade. Analysis of genetic structure also recovered each major lineage within *L. kadarsani* as a distinct cluster, and supported a model with minimal admixture when K=5 except for two samples in central Flores. Demographic analyses of the East and West Flores populations provided an estimated divergence time of ~1.9 MA, with roughly 10 times as much migration from west to east as in the opposite direction. While the Bayesian mtDNA phylogeny suggested a stepping-stone model of island colonization, the genomic phylogeny suggests a much more complicated pattern that would require leap-frog dispersals to more distant islands. The topology of all trees is consistent with an *in situ* speciation event on Lombok, though further sampling and analyses are needed to confirm this.

INTRODUCTION

How do island archipelagos acquire their constituent faunas? The answer is very much dependent on the geological and tectonic processes that gave rise to the islands in question. For oceanic islands, which have never been connected by dry land to continental source populations, the entire fauna will have either arrived directly by overwater dispersal or will represent the outcome of in situ diversification of a lineage that itself arrived by overwater dispersal. Of course, there are many factors that determine faunal diversity on oceanic islands, such as the distance from a source, the number of such sources, the age of the island, the size of the island, habitat heterogeneity on the island, the stability of the island, and many others (MacArthur & Wilson, 1967; Simberloff, 1974; Losos & Ricklefs, 2009). Notably, oceanic islands are rarely colonized by amphibians, which is in stark contrast when compared with other vertebrate groups such as reptiles, mammals, and birds (Darwin, 1859; Blaustein *et al.*, 1994; but see Vences *et al.*, 2003; Pyron, 2014). One of the major proposed explanations for the paucity of amphibians on oceanic islands is their poor osmotic tolerance, which causes loss of water balance when in contact with sea water (Balinsky, 1981; Duellman & Trueb, 1986). Frogs are usually tied to cool and moist habitats or clean streams (Duellman & Trueb, 1986), and the prevalence of these habitats on oceanic islands will greatly influence the survival and species richness of frogs (Ricklefs & Lovette, 1999).

Wallacea is a region containing hundreds of oceanic islands that lie between Southeast Asia and Australia (Dickerson, 1928). Many of the larger islands in Wallacea are inhabited by anurans that have colonized the islands by over-water dispersal, though some species have been introduced by humans (AmphibiaWeb, 2016). Wallacea can be divided into four regions that are distinct in both their geological history and their faunal diversity (Stresemann, 1939). One of these regions includes the Lesser Sunda Islands, which occur as two parallel island arcs that stretch between Bali in the west and New Guinea in the east. The Lesser Sundas contain 18 described species of frogs and toads and no species of salamanders or caecilians, though it is thought that the frog diversity is an underestimate (Inger, 1999). The Lesser Sundas contain a wide variety of habitats but in general they are hotter and drier than other parts of Indonesia, which is not optimal for many of the tropical species of frogs from the surrounding regions of the Greater Sunda Islands (Java, Sumatra, Borneo), Sulawesi, Maluku, and New Guinea. In fact, some of the hotter and drier large islands of the Lesser Sundas such as Lembata and Sabu contain only one species of frog, and some islands such as Pantar or Atauro are not known to be inhabited by frogs at all (Kaiser *et al.*, 2013; AmphibiaWeb, 2016). Most of the species of frogs in the Lesser Sundas are on the IUCN Red List as their ranges are small, and their abundances are low (IUCN, 2016). Although very little research has been conducted investigating the biogeographical history and genetic structure of frogs within the Lesser Sundas, there is an increasing body of genetic literature focusing on frogs that have colonized oceanic islands or archipelagos (Brown & Guttman, 2002; Vences *et al.*, 2003; John Measey *et al.*, 2007; Brown *et al.*, 2010; Brown & Siler, 2014; Bell *et al.*, 2015a; Bell *et al.*, 2015b). The Lesser Sundas are a unique system in which to examine biogeographical patterns within frogs because these oceanic islands are close enough to continental areas to be colonized by frogs, and their linear arrangement between these two sources (Asia and Australo-Papua) allows for some hypotheses to be made regarding the sequence of island colonization.

Fanged frogs of the genus *Limnonectes* (Family: Dicroglossidae) are common throughout Southeast Asia, and two species occur in the Lesser Sundas (AmphibiaWeb, 2016). Both species

are restricted to the western islands of the Inner Banda Arc, with *L. dammermani* reported from Lombok, Sumbawa, and Flores, and *L. kadarsani* reported from Lombok, Sumbawa, Flores, Adonara, and Lembata (Iskandar & Mumpuni, 2004a; Iskandar & Mumpuni, 2004b). Research on *Limnonectes* from other parts of Indonesia, such as the Greater Sunda Islands, Sulawesi, and the Philippines have shown that these frogs are not only able to colonize oceanic islands, but that they diverge ecologically and morphologically, and form new species that are sometimes sympatrically distributed (Evans *et al.*, 2003; McLeod, 2010; Setiadi *et al.*, 2011; Iskandar *et al.*, 2014). The two species of *Limnonectes* in the Lesser Sundas are most closely related to *L. microdiscus* of adjacent Java and Bali (pers. comm., J. McGuire). Given that these lineages arrived from the west and have only colonized the Inner Banda Arc, they make an excellent study system that can be used to test the stepping-stone hypothesis of island colonization. In this model, the island nearest to the source (i.e., Lombok) is predicted to have been colonized first from adjacent Bali, followed by the next closest island (Sumbawa) and so on until they reach the island furthest from the source which would represent the most derived lineage (in this case it would be Flores for *L. dammermani,* and Lembata for *L. kadarsani*). Additionally, this system is ideal for determining if *L. dammermani* and *L. kadarsani* diverged in allopatry or if they possibly diverged parapatrically within an island.

A common approach to answer the questions posed above regarding the sequence of island colonization and mode of speciation is to utilize mitochondrial DNA sequence data in a phylogenetic analysis. While this method can certainly tell us a lot about the matrilineal history of lineages, the mitochondrial phylogeny does not always represent the true evolutionary history, and it can tell us nothing about genomic population structure or historical demographics such as gene flow or population divergence times. New methods for quickly and inexpensively obtaining genomic sequence data from non-model organisms are now available such as ddRADs (Peterson *et al.*, 2012), ultra-conserved elements or UCEs (Bejerano *et al.*, 2004), anchored-tag enrichment (Lemmon *et al.*, 2012), and transcriptome-based exon-capture (Bi *et al.*, 2012; Bi *et al.*, 2013). The ddRAD and UCE methods return thousands of very short, conserved loci from which researchers typically choose one variable site per locus for analysis. Anchored Tag enrichment amplifies hundreds of longer loci but the method is costly, and lab work must be outsourced. Transcriptome-based exon-capture methods appear to provide the best type of datasets for phylogenetic and demographic inference as they can screen hundreds to thousands of independent loci that can be quite long (>1,000 bp) with roughly equal capture efficiency across divergent samples characterized by up to 15% nuclear divergence (pers. comm., Ke Bi). Additionally, the exon-capture method returns a large amount non-coding intron sequence data flanking the exons that can be used for historical demographic analyses that require neutrally evolving loci.

In this study the biogeographical history of *Limnonectes* from the Lesser Sunda Islands is inferred using both mitochondrial DNA from a large number of samples (~150) and approximately 1,000 transcriptome-derived nuclear loci for a subset of samples (n=48). From a taxonomic standpoint, the monophyly of these species will be tested and any cryptic lineages shown to be independently evolving will be recommended for elevation to full species. The genetic structure, timing, and sequence of island colonization will be estimated to test biogeographical hypotheses outlined in Chapter 1, with the “stepping-stone” and “Asian invasion” models being the most relevant for this system, as the other models deal with dispersal from Australo-Papua and Sulawesi.

MATERIALS & METHODS

Sample Collection: Limnonectes specimens were collected from the field and included *L. microdiscus* from Java, *L. dammermani* from Lombok, and *L. kadarsani* from Lombok, Sumbawa, Flores, and Lembata islands. These samples and tissues were collected on two separate expeditions to the Lesser Sunda Islands that occurred during 2010 and 2011. Liver tissue was dissected from euthanized frogs and either stored in RNALater, or flash frozen in liquid nitrogen. Specimens were given field tags in the catalog of Jimmy A. McGuire (JAM#). The tissues were divided in half and deposited in both the Museum of Vertebrate Zoology at UC Berkeley (and subsequently will receive MVZ catalog numbers) and the Museum Zoologicum Bogoriense (which are given MZB catalog numbers). The formalin fixed voucher specimens were divided equally between the MVZ and MZB collections.

MtDNA Data Collection: DNA was extracted from liver tissue using standard salt extraction techniques (Aljanabi & Martinez, 1997) or using the DNeasy kit. DNA extractions were then diluted to concentrations suitable for PCR-amplification (~20-60 ng/uL). The *16S* gene was sequenced for 153 frogs from the Lesser Sunda Islands and for 3 *Limnonectes microdiscus* from Java to serve as an outgroup (Table 1). All sequence data was collected using standard PCR-amplification protocols using the primers 16sc-L (5'-GTRGGCCTAAAAGCAGCCAC-3') and 16sd-H (5'-TCCGGTCTGAACTCAGATGACGTAG-3') (Evans *et al.*, 2003). PCR reactions contained 18.3 µL water, 2.5 µL of 10X buffer, 1.5 µL magnesium chloride, 1.5 µL dNTPs (2 µM), 0.6 µL of each primer, 0.2 µL Taq polymerase, and 1µL genomic DNA at concentration of 20-40 ng/µL. PCR products were cleaned using ExoSAP-IT (USB, Cleveland, OH) before being labeled with fluorescent-dye nucleotides through cycle sequencing reactions for both forward and reverse primers. Ethanol precipitation was used to clean cycle sequencing products, which were sequenced on an ABI 3730 sequencer (Applied Biosystems, Foster City, CA, USA). Raw sequence reads were combined in Codoncode Aligner 3.5.2 (CodonCode Corporation, Dedham, MA, USA) and aligned with MUSCLE (Edgar, 2004).

MtDNA Data Analyses: The sequence alignment of 874 bp was imported into JMODELTEST V2.1.4 (Darriba *et al.*, 2012) to determine the best-fit model of sequence evolution (HKY +G) that is supported by the program BEAST V1.8 (Drummond & Rambaut, 2007). A BEAST run was conducted using the uncorrelated relaxed clock model and a coalescent constant size tree prior with a uniform distribution. A rule of thumb 1% rate of sequence evolution per million years (which corresponds to 2% divergence per million years for any two-lineage comparison) was applied to obtain a rough approximation of timing of entry into the archipelago, as well as the ages of island-specific lineages. A preliminary run was carried out to determine the appropriate number of generations required to achieve ESS values for each parameter that were greater than 200, as viewed in TRACER V1.6 (Rambaut & Drummond, 2009). Once the appropriate run length was determined, two separate runs of 100 million generations were carried out, sampling every 10,000 generations for a total of 10,000 saved generations per run. A burnin of 10% was removed from each of the two runs and the remaining 18,000 trees were combined to create a 50% majority rule consensus tree. The tree was rooted using the outgroup *L. microdiscus* from Java. Nodal support was assessed using posterior probability values. A Maximum Likelihood approach was also estimated using the program RAXML (Stamatakis, 2014). The default GTR+I+G model of sequence evolution was applied, and nodal support was assessed with 100 bootstrap replicates.

Transcriptome Sequencing: Total RNA was extracted from two *L. kadarsani* samples (JAM12330, Sumbawa; JAM13047, Lembata) using the RNEasy Protect Mini Kit (Qiagen) and protocol. Samples were evaluated using a BioAnalyzer 2100 RNA Pico chip (Agilent), with RIN scores greater than 7. Sequencing libraries were prepared using half reactions of the TruSeq RNA Library Preparation Kit V2 (Illumina), beginning with Poly-A selection for samples with high RIN scores (> 8.0) and Ribo-Zero Magnetic Gold (Epicentre) ribosomal RNA removal for samples with low RIN scores (< 7.0). Libraries were pooled and sequenced on an Illumina HiSeq2500 with 100 bp paired-end reads. Transcriptomic data were cleaned following Singhal (2013). Cleaned data were assembled using TRINITY (Grabherr *et al*., 2011) and annotated with *Anolis carolinensis* (Ensembl) as a reference genome using reciprocal BLASTX (Altschul *et al*., 1997) and EXONERATE (Slater & Birney, 2005). Annotated transcripts were compared from the two individuals to search for orthologs via BLAST (Altschul *et al*., 1990). Mitochondrial loci were removed from the transcripts. Only transcripts with a GC content between 40%-70% were kept because extreme GC content causes reduced capture efficiency for the targets (Bi *et al*., 2012). All the bioinformatics pipelines for transcriptome data processing and annotation are available at https://github.com/CGRL-QB3-UCBerkeley/DenovoTranscriptome.

Marker Develoment: Annotated and filtered contigs from all transcripts were aligned to identify shared markers. Markers under 300 bp were discarded and markers greater than 1,000 bp were reduced to a maximum length of 1,000 bp. The remaining genes were examined for repetitive elements, short repeats, and low complexity regions, which are problematic for probe design and capture. The two sets of transcripts were screened using the REPEATMASKER Web Server (Smit *et al*., 2015), which resulted in the masking of repetitive elements or low complexity regions. To be conservative, if either of the two transcripts for a gene contained masked sites, that gene was removed from the final marker set. The two *L. kadarsani* transcripts were 0.43% divergent on average. The resulting markers from the two transcripts were compared to identify the variability of each marker as determined by the number of polymorphic sites. The markers were sorted according to the number of variable sites per locus. All invariable loci and loci with only a single variable site were discarded, as well as the top 5% of the most variable loci. A total of 1,200 of the 2,912 candidate loci were integrated into a MYBaits in-solution exon-capture kit using only the JAM12330 transciptome as a reference. The estimated target size of the combined loci was approximately 1,030,000 bp and the probes were tiled at 3X across the reference. Four of the 1,200 loci identified from the transcripts were positive control loci confirmed by Sanger sequencing (*RAB5A, ADIPOR2, WNK1,* and *WASH4P*). Pipelines for marker development are available at https://github.com/CGRL-QB3-UCBerkeley/MarkerDevelopment Pylogenomics.

Sample Library Preparation: A total of 48 *Limnonectes* samples were initially chosen for library preparation. These samples were picked by examining the mitochondrial tree to maximize the genetic diversity from all of the islands (Table 1). The DNA was quantified by Qubit DNA BR assay (Life Technologies) and 1500 ng total DNA was diluted in 100 µl of ultrapure H_2O. A Bioruptor UCD-200 (Diagenode) was used to sonicate the samples on a low setting for 15 minutes, using 30s on/30s off cycling. For each sonicated sample, 5 µl of product was run on a 1% gel at 100V for 45 min to ensure fragments were appropriately sized (200–500 bp). Individual genomic libraries were prepared following Meyer and Kircher (2010), with slight modifications, including the use of at least 1,500 ng total DNA for library preparation (rather

than 500 ng) to remedy the possibility of decreased library diversity resulting from a large genome size. We used 7-9 cycles of post-adapter ligation PCR to enrich the libraries and incorporate a 7bp P7 index. The resulting 50 µl of amplified library product had an average concentration of 30 ng/µl measured by a Nanodrop 1000 spectrophotometer (Thermo Scientific), producing an average yield of 1,500 ng total library DNA.

MYBAits In-Solution Exon Capture: Libraries were pooled in equal amounts with each pool containing 4 individuals. Pools were determined by grouping closely related samples to minimize competition for probe binding. MYbaits capture reactions were performed following the v2.3.1 manual with some modifications. For each capture reaction library master mix, the pooled libraries were vacuum dried at 45°C for 60 min and re-suspended in ultrapure H_2O, then combined with 1.66 µl each of salmon sperm COT-1, human COT-1, chicken COT-1, and xGEN blocking oligos. The combined volume of water for DNA resuspension and volume of blocking oligonucleotides totaled 6.5 µl. The hybridization reaction proceeded at 65°C for 24-28 hours. Individual capture reactions were purified using streptavidin-coated magnetic beads, and post-capture products were PCR-amplified using four independent reactions of 14 cycles each. These reactions were resuspended in 11 µl of ultrapure H_2O, and had concentrations between 7-34 ng/µl, as measured using a Qubit. Purified PCR products from the same capture were combined and quantified using a BioAnalyzer 2100 DNA-1000 chip. The combined post-capture amplified products ranged from 8-30 ng/µl, and the average product size was ~370 bp. The combined post-capture libraries were sequenced on one lane of an Illumina HiSeq2500 with 100 bp paired-end reads.

Data Pipeline: Raw sequence data were cleaned following Singhal (2013) and Bi *et al*. (2012). Raw fastq reads were filtered using TRIMMOMATIC (Bolger *et al*., 2014) and CUTADAPT (Martin, 2011) to trim adapter contamination and low quality reads. BOWTIE2 (Langmead & Salzberg, 2012) was used to align the data to *Escherichia coli* (NCBI: 48994873) to remove potential bacterial contamination. Exact duplicates were eliminated as well as low complexity sequences using a custom script. Overlapping paired reads were also merged using FLASH (Magoč & Salzberg, 2011) and COPE (Liu *et al*., 2012) to avoid inflated coverage estimates in the overlapping region. The resulting cleaned reads of each individual specimen were *de novo* assembled using ABYSS (Simpson *et al*., 2009). Individual raw assemblies were generated using a wide range of k-mers (21, 31, 41, 51, 61 and 71) and then CD-HIT-EST (Li & Godzik, 2006), BLAT (Kent, 2002), and CAP3 (Huang & Madan, 1999) were used to cluster and merge all raw assemblies into final, less-redundant assemblies. BLASTN (e-value cutoff = 1e-10, similarity cutoff = 70) was used to compare the target sequences with the raw assemblies of each individual in order to identify the set of contigs that were associated with targets (in-target assemblies). A self-BLASTN (e-value cutoff = $1e^{-20}$) was run to compare the assemblies against themselves to mask any regions from a contig that matched other regions from other contigs. For each matched contig, EXONERATE (http://www.genome.iastate.edu/bioinfo/resources/manuals/exonerate/exonerate.man.html) was used to define protein-coding and flanking regions. Flanking sequences were retained if they were within 250 bp of a coding region. Finally, all discrete contigs that were derived from the same reference transcript were joined together with Ns based on their relative BLAST hit positions to the reference. Most of the final in-target assemblies contain multiple contigs, and each includes both coding regions and flanking sequences if captured.

Cleaned sequence data were then aligned to the resulting individual-specific in-target assemblies using NOVOALIGN (Li & Durbin, 2009) and only reads that mapped uniquely to the reference were retained. The programs Picard (http://broadinstitute.github.io/picard/) and GATK (McKenna *et al*., 2010) were used to perform re-alignment. Finally, the program SAMtools/BCFtools (Li *et al*., 2009) was used to generate individual consensus sequences by calling genotypes and incorporating ambiguous sites in the in-target assemblies. A consensus base was only kept when the site depth was above 10X. Sites were masked within a 5 bp window around an indel. Sites were also filtered out if more than two alleles were called. FASTQ were converted to FASTA using seqtk (https://github.com/lh3/seqtk), and putative repetitive elements and short repeats were masked using REPEATMASKER (Smit et al. 2015) with "vertebrata metazoa" as a database. Markers were removed if more than 80% of the bases were Ns. The read depth of each individual marker was calculated and loci were filtered out if the depth fell outside the 99th percentile of the statistics. Markers were also eliminated if the individual heteozygosity fell outside the 99th percentile of the statistics. The final filtered assemblies of each individual specimen were aligned using MAFFT (Katoh & Standley, 2013). Alignments were then trimmed using TRIMAL (Capella-Gutierrez *et al*., 2009). Alignments were removed if more than 25% missing data (Ns) were present in 25% of the samples, or if the proportion of shared polymorphic sites in any locus was greater than 20%. The bioinformatic pipelines employed here for processing sequence capture data are available at https://github.com/CGRL-QB3-UCBerkeley/denovoTargetCapturePhylogenomics.

Evaluation of Data: To evaluate capture efficiency, the average per-base sequence depth (or coverage) was calculated separately for the exon sequences and for the flanking sequences of each sample. The coverage at each base pair site for either data set was inferred using SAMtools (Li *et al*., 2009). The per base pair coverage estimates for all sequences (exon or flanking) associated with each transcript (up to 974 genes for this experiment) were averaged, resulting in a set of average coverage estimates across loci. The resulting output of the set of average coverage estimates was used to infer the median, upper and lower quartiles, and range of coverage estimates using samples or genes as factors. These calculations were performed and automated across samples using python scripts with the output visualized in R. Differences in levels of coverage were examined using pooling size as a factor.

The resulting alignments of exon-only data and flanking region data were evaluated for taxon number, sequence length, percentage of missing data, and proportion of informative sites. These results were visualized in R, and the relationship between the number of informative sites and alignment length was investigated using a simple linear regression. The relationship between phylogenetic distance and missing data was also investigated using a simple linear regression. The percentage of missing data was calculated from the final concatenated alignment of exon-only loci that passed multiple post-processing filters, including a minimum length of 100 bp, no more than 80% missing data per sequence in alignments, and no more than 25% total missing data across an alignment. These filters were enforced using a custom alignment refinement python script for all alignments. After all filtering, 974 of the 1,200 genes were kept for analyses. All custom python scripts for sequence capture performance evaluation are available on github (https://github.com/dportik/).

Phylogenomic Analyses: The concatenated alignment of all sequence data (Target + Flanking) was subjected to Maximum Likelihood phylogenetic analysis using RAxML (Stamatakis, 2014)

under the GTR+I+G model of sequence evolution. Nodal support was assessed with 100 bootstrap replicates. Individual gene trees for each of the 974 genes were also estimated using RAxML. These gene trees served as input files for a supertree approach as implemented in ASTRAL-II (Mirarab *et al.*, 2014). For this analysis each individual was treated as a "species" because it was unclear where the species boundaries lie within this system and because direct comparisons with the topology of the RAxML tree were desired.

Population Structure: One informative SNP per locus was randomly chosen from within the Lesser Sundas assemblage (38 individuals of *L. kadarsani*) to create an input file for the program STRUCTURE (Pritchard *et al.*, 2000). The program was run for 40,000 generations (with the first 20,000 generations constituting burn-in) for K=1 through K=6 populations and with 10 replicates per K. The results were then imported into STRUCTURE HARVESTER (Earl, 2012) to determine the most likely number of populations as determined by the Delta K method. The program was then run for 100,000 generations (50,000 generations as burnin) for K=2 up to K=5 to examine the sequential division of the assemblage for each assumed number of populations.

Inter-Island Demographics: Demographic analyses utilized the flanking sequence from each locus because these regions are presumably not under selection, though they are linked to exonic regions under selection. These data were analyzed with the program G-PHOCS (Gronau *et al.*, 2011), which is an isolation-with-migration program that is capable of dealing with genomic sequence data from unlinked neutrally evolving loci. This program estimates the effective population sizes of the extant populations as well as their ancestor, population divergence times, and migration rates between extant populations. This analysis was run to estimate relative rates of migration between divergent lineages that are parapatrically distributed on the island of Flores. An initial run of 200,000 generations was used to assess convergence of the parameters, followed by a run of ~595,000 generations. After removing 59,000 generations, the remaining 536,000 generations were visualized in TRACER (Rambaut & Drummond, 2009) to assess the posterior distribution of the demographic parameters.

A mutation rate of $2.2 * 10^{-9}$ mutations/site/year was used to convert parameter estimates (Kumar & Submaranian, 2002). All values for Theta and Tau given by the program G-PHOCS are scaled by 10^{-4}. Demographic parameter estimates were converted to estimates of effective population sizes (individuals) by dividing the scaled Theta estimate by the mutation rate, then dividing that value by 4 (because diploid organisms will have an effective population size of 4 at any given locus). The population divergence time in years was calculated by dividing the scaled Tau estimate by the mutation rate. Migration rate estimates were converted to Migrants per Generation by multiplying the migration estimate by the converted effective population size estimate for the population receiving the gene flow, then dividing that value by the number of generations that have passed (in years) since divergence. A generation time of 1 year was used to convert migration rates.

RESULTS

MtDNA Phylogeny: Both the Maximum Likelihood and Bayesian phylogenies produced similar topologies when rooted with *Limnonectes microdiscus* from Java Island (Fig. 3,4). Clade localities for both trees can be visualized in Figure 5. Both phylogenies contain a basal split between all samples from Lombok versus those from Sumbawa, Flores, and Lembata. This

branching arrangement suggests that *L. kadarsani* is paraphyletic, with *L. kadarsani* from Lombok more closely related to *L. dammermani* from Lombok than to other *L. kadarsani* from islands to the east. In addition, the relationships among the remaining populations of *L. kadarsani* are structured, indicating that samples from Sumbawa are sister to those from Flores + Lembata. The major difference between the two phylogenies is that the ML tree suggests that populations on Flores form a monophyletic group that is nested within a Lembata clade, whereas the Bayesian tree places Lembata as a monophyletic assemblage nested within a Flores clade. The Bayesian tree (Fig. 4) is time-calibrated and indicates the split between the Lesser Sundas clade and *L. microdiscus* to be ~9 MYBP (95% CI low=2.95, high=14.07). The split between all Lombok samples and the rest of the Lesser Sundas was estimated to be more recent at ~4 MY (95% CI low=2.26, high=5.68). The Lombok populations of *L. dammermani* and *L. kadarsani* are estimated to have diverged from one another ~2.2 MY (95% CI low=1.04, high=3.28), while the rest of the Lesser Sundas (Sumbawa + Flores + Lembata) populations began diverging from one another just less than 1 MY ago according to this analysis.

Exon-Capture Data Characteristics: The total alignment of both the targeted and flanking regions from the 974 genes was 1,235,981 bp. The average coverage for the targeted regions was ~50X, while the flanking regions had approximately 25X coverage on average. However, the average coverage for each individual library was highly variable for both the targeted (Fig. 6) and flanking (Fig. 7) regions.

The average number of taxa per alignment was 47 out of the 48 samples (Fig. 8c, Fig. 9a). The number of informative sites has a relatively linear relationship with locus length (Fig. 8a), and there is on average 7% informative sites per alignment (Fig. 9c). There is no clear relationship between the alignment length and the percentage of gaps in each alignment (Fig. 8b). The final length of the contigs ranged from 100 bp up to ~4,500 bp (Fig. 9b), and the percent of missing data was no higher than 25% after the additional filtering step (Fig. 9d).

Phylogenomic Trees: Both the concatenated Maximum Likelihood (Fig. 10) and the coalescent supertree (Fig. 11) converged on the same topology with respect to the relationships between major lineages. All major lineages, as well as the major nodes in the tree are well supported with bootstrap support of 100 (Fig. 10). Both described species within the Lesser Sundas, *L. dammermani* and *L. kadarsani,* are monophyletic sister taxa. Within *L. kadarsani,* the basal lineage is on Lombok, which is sister to the rest of the Lesser Sundas. Within the remaining lineages, East Flores is most basal, followed by a Sumbawa lineage, which is sister to a West Flores + Lembata lineage with Lembata nested within West Flores.

Population Structure within L. kadarsani: The Delta K analysis finds that the most likely number of populations within *L. kadarsani* is 2 (Fig. 12a). The highest mean estimate of the Ln probability of the data is at K=5 (Fig. 12b). The population structure bar plots show a high level of genetic structuring from K=2 up to K=5, after which the program returned only five meaningful clusters corresponding to the results of the K=5 analysis (Fig. 12c). The Lombok population is the most distinctive and represents a pure gene pool block in each analysis. With K=3, Lembata becomes a pure population cluster, Sumbawa + East Flores becomes a distinct population cluster, and Western/Central Flores is shown as admixed between the previous two clusters. With K=4, Sumbawa becomes a distinct cluster, Lembata + East Flores becomes a distinct cluster, and Western/Central Flores becomes a mostly distinct cluster with two

individuals showing admixture with Lembata + East Flores.

Demographic Analyses: The effective population size of the Western Flores population was estimated to be nearly three times larger than the East Flores population with approximately 1.8 million and 0.5 million individuals respectively (Table 3, Fig. 13a). The most recent common ancestor of the Flores populations was estimated at approximately 0.6 million individuals. The Flores populations were estimated to have diverged from one another approximately 1.9 million years ago (Table 3, Fig. 13b). Migration appears to be unidirectional with approximately 2.6 migrants per generation from West Flores into East Flores, and roughly 0.2 migrants per generation from East Flores into West Flores (Table 3, Fig. 13c).

DISCUSSION

There are currently two species of *Limnonectes* known from the Lesser Sundas (excluding Bali) and they are both restricted to the western islands of the Inner Banda Arc. *Limnonectes kadarsani* is known to occur on the islands of Lombok, Sumbawa, Flores, Adonara, and Lembata (Iskandar & Mumpuni, 2004b), and we were able to collect samples of this species from each of these islands except Adonara. Despite not sampling Adonara, our materials from near Larantuka on extreme eastern Flores are likely to be similar to those from Adonara given the narrow, shallow channel that separates these islands. With this sampling, we were able to infer the historical biogeography of *L. kadarsani*. The second species in the Lesser Sundas, *L. dammermani,* also is said to occur on multiple islands in the Inner Arc, including Lombok, Sumbawa, and Flores (Iskandar & Mumpuni, 2004a). However, we have doubts about the occurrence of this species outside of Lombok Island (the type locality), and suspect that samples referred to this species from other islands may be misidentified *L. kadarsani*.

Our initial analysis of mitochondrial DNA showed a very interesting pattern with *L. kadarsani* from Lombok recovered as sister to *L. dammermani* from Lombok suggesting *in situ* speciation and paraphyly of *L. kadarsani*. However, the exome-capture data strongly support a sister taxon relationship between *L. kadarsani* and *L. dammermani*, which we infer represents a relatively ancient case of mitochondrial capture via hybridization. This scenario involves colonization of islands east of Lombok (Sumbawa or Flores), followed by allopatric divergence and the formation of *L. kadarsani*. Then after a back colonization of Lombok, the *L. kadarsani* population would have captured the *L. dammermani* mtDNA (Fig. 14). We argue that the introgression event was not recent but rather took place on the order of *millions* of years ago because the sequence divergence between the mt haplotypes is ~4%, corresponding to roughly 2 million years. Such a pattern has been observed in other studies (see McGuire et al. 2007; Toews & Brelsford, 2012). Another possible case of more recent mtDNA capture (~0.2 million years ago) involves the East Flores population which contain mtDNA haplotypes similar to populations immediately to the west and east , West Flores and Lembata, yet East Flores individuals are distinct in nDNA and are sister to a clade containing Sumbawa + West Flores + Lembata.

The results of the 974 nuclear gene dataset, with over 1.2 million base pairs of sequence data per sample, are likely to better estimate the evolutionary history of the Lesser Sundas *Limnonectes* than are the mtDNA data. When choosing loci relevant for the exon-capture experiment it was confirmed that they span multiple chromosomes and are widely distributed within chromosomes. The phylogeographic patterns suggested by the genomic phylogeny are

different from the mitochondrial patterns and suggest a complicated biogeographical history. Both described species were inferred to be monophyletic, but this does not rule out the possibility of an *in situ* speciation event on Lombok. This pattern could have been created by either an *in situ* speciation event on Lombok followed by a dispersal event from *L. kadarsani* eastward (Fig. 15), or it could be the result of an eastward dispersal even by *L. dammermani* followed by allopatric speciation and a westward re-colonization of Lombok by a newly formed *L. kadarsani* (Fig. 14). The genomic phylogenetic relationships of the major *L. kadarsani* lineages from Sumbawa, Flores, and Lembata are well supported and do not match the mitochondrial relationships. Rather than Sumbawa representing the basal lineage of this group the samples from the Larantuka area of East Flores were found to be basal, suggesting that East Flores and Lembata are not closely related as suggested by the mtDNA phylogeny. This result raises the question as to where the Adonara Island population fits as it is equally close to Lembata and East Flores and likely becomes land-bridged to both periodically.

One factor that can allow for *in situ* speciation within a small island is the occurrence of environmental heterogeneity, including elevation, soil, precipitation, and many other abiotic features that can allow populations to become locally adapted and subsequently diverge. Lombok contains the second tallest volcano in all of Indonesia, Gunung Rinjani, which has steep slopes and rises to over 3,700 meters. It has been shown that the composition of plants changes drastically with elevation on the slopes of Gunung Rinjani (Dossa *et al.*, 2013) and this elevational habitat zoning could have contributed to parapatric or allo-parapatric divergence of *Limnonectes* on Lombok. Lombok also has been the site of many large volcanic eruptions, with one of the largest eruptions in the Holocene occurring in the year 1257 that caused massive pyroclastic flows, deposited large quantities of ash over the entire region, and left most of Lombok uninhabitable (Marrison, 1999; Lavigne *et al.*, 2013). Given that volcanic activity has consistently devastated Lombok over the course of its existence, there may be extinct lineages of *Limnonectes* that would obscure the inference of the biogeographical history of these frogs.

Assuming that *L. dammermani* and *L. kadarsani* diverged in allopatry and have come into secondary contact on Lombok, there are many possible colonization scenarios that are consistent with the genomic topology. One colonization scenario that could produce the patterns seen in the genomic phylogeny would involve an initial dispersal from the Lombok to Flores followed by divergence and the formation of *L. kadarsani*, a re-colonization of Lombok from East Flores, then dispersal from East Flores to Sumbawa, dispersal from Sumbawa to West Flores, and finally dispersal from West Flores to Lembata (Fig. 16). This scenario involves four long distance leap-frog dispersal events with one from Lombok to East Flores, one from East Flores back to Lombok, one from East Flores to Sumbawa, and another from West Flores to Lembata. Another scenario that could produce the same topology involves the same dispersal from Lombok to East Flores, a back colonization of Lombok from East Flores, followed by a dispersal event from East Flores into West Flores, then from West Flores to Sumbawa, and finally from West Flores to Lembata (Fig. 17). This scenario requires three leap-frog dispersal events with one from Lombok to East Flores, one from East Flores to Lombok, and one from West Flores to Lembata. An additional possible colonization scenario would involve dispersal from Lombok to Sumbawa, followed by allopatric species formation of *L. kadarsani,* back colonization of Lombok from Sumbawa, then a dispersal from Sumbawa to East Flores, followed by another dispersal event from Sumbawa to West Flores, and finally a dispersal from West Flores to Lembata (Fig. 18). This scenario involves only two leap-frog dispersal events, but also requires at least three dispersal events from Sumbawa to other islands. Distinguishing between

these scenarios in any convincing way is difficult, but the data strongly reject the null model involving a simple stepping stone process from west to east.

Genetic clustering results also support the distinctiveness of the five major lineages within *L. kadarsani*: Lombok, Sumbawa, West + Central Flores, East Flores, and Lembata. While it is unclear if any of these lineages represents distinct species, it is evident that the Lombok population is the most divergent and has been isolated for the longest interval. At present, each lineage should be treated as a management unit for conservation given that gene flow is either extremely low or absent, at least between lineages on separate islands. The analysis of migration between the two major lineages within Flores shows a clear signal of unidirectional gene flow from West to East but not from East to West. This pattern of gene flow indicates that the western Flores population is essentially genetically isolated from eastern Flores. Interestingly, the population divergence time estimate from the genomic data (~1.9 MY) is much older than the mitochondrial divergence estimate (<0.5 MY). If the genomic phylogeny is in fact correct, then the mtDNA patterns could be the result of mitochondrial introgression and subsequent replacement, either from West Flores or from Lembata into East Flores. Samples from Adonara and from additional localities in East Flores will be critical for resolving this question.

CONCLUSIONS

In this study, the biogeographical history of *Limnonectes* fanged frogs from the Lesser Sunda Islands of Indonesia was inferred using both a mitochondrial dataset and a 974 gene nuclear dataset. Both datasets and all analyses recovered the same major lineages within *L. kadarsani*, and *L. dammermani* was only found to occur on Lombok though further sampling may confirm that *L. dammermani* also occurs on Sumbawa and Flores as suggested by some sources. The mitochondrial phylogenies render *L. kadarsani* polyphyletic placing the Lombok population of *L. kadarsani* as sister to *L. dammermani*, to the exclusion of all other *kadarsani* populations. This suggests that the islands east of Lombok were colonized by the ancestor of *L. dammermani* and *L. kadarsani*, and after this initial eastward jump there was an *in situ* speciation event on Lombok. However, the nuclear dataset produced trees with a different topology than the mtDNA and recover each species as monophyletic sister lineages. But because the basal lineage within *L. kadarsani* is from Lombok it does not rule out an *in situ* divergence event on Lombok, rather it suggests that dispersal to the islands east of Lombok came from the *L. kadarsani* lineage. While there are a number of possible colonization scenarios that could produce the topology seen in the genomic phylogenies it is clear that it was not a stepping-stone model of island colonization as suggested by the Bayesian mtDNA data. If one considers the fewest leap-frog dispersal events (the most parsimonious model) to be most likely, there are still multiple scenarios that remain valid. Drawing on biogeographic patters from other Lesser Sunda taxa (see Ch. 2 & 3) we see a repeated pattern of within-island divergence or non-sister lineages occurring parapatrically or sympatrically within an island. While biogeographical studies of other taxa are needed to understand the common factors affecting divergence of populations among the Lesser Sundas, this study has shown a possible example of *in situ* speciation, multiple long distance leap-frog dispersal events, and provides another example of parapatrically distributed non-sister lineages occurring on Flores.

ACKNOWLEDGEMENTS

I would like to thank Jimmy A. McGuire, Djoko Iskandar, Alexander Stubbs, Ben Karin, Kristopher Harmon, Umilaila Arafin, Evy Arida, Gilang Ramadan, Amir Hamidy, Luke Bloch, Sarah Hykin, Jerome Fuchs, and Jim Patton for their help with the field collection of specimens and tissues. I thank Ke Bi and Lydia Smith for laboratory support, Carol Spencer and Jackie Childers for accessioning of specimens, and Michelle Koo for GIS map support, and Ke Bi for bioinformatics support. Funding was provided by the MVZ, the Department of Integrative Biology at UC Berkeley, the National Geographic Foundation, and the National Science Foundation. Collection of specimens was carried out under UC Berkeley IACUC # A3084-01.

REFERENCES

Aljanabi, S. & Martinez, I. (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, 25:4692–4693.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, 215:403–410.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17):3389–3402.

AmphibiaWeb (2016) Information on amphibian biology and conservation. Available:

Balinsky, J.B. (1981) Adaptation of nitrogen metabolism to hyperosmotic environment in amphibia. *Journal of Experimental Zoology,* 215:335–350.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S. & Haussler, D. (2004) Ultracon- served elements in the human genome. *Science*, 304:1321–1325.

Bell, R.C., Drewes, R.C., Channing, A., Gvoždík, V., Kielgast, J., Lötters, S., Stuart, B.L. & Zamudio, K.R. (2015a) Overseas dispersal of *Hyperolius* reed frogs from Central Africa to the oceanic islands of São Tomé and Príncipe. *Journal of Biogeography*, 42(1):65–75.

Bell, R.C., Drewes, R.C. & Zamudio, K.R. (2015b) Reed frog diversification in the Gulf of Guinea: overseas dispersal, the progression rule, and in situ speciation. *Evolution*, 69(4):904–915.

Bi, K., Vanderpool, D., Singhal, S., Linderoth, T., Moritz, C., & Good, J.M. (2012) Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics*, 13(1):403.

Bi, K., Linderoth, T., Vanderpool, D., Good, J. M., Nielsen, R., & Moritz, C. (2013) Unlocking the vault: next-generation museum population genomics. *Molecular Ecology*, 22(24):6018–6032.

Blaustein, A.R., Wake, D.B. & Sousa, W.P. (1994) Amphibian declines: judging stability, persistence, and susceptibility of populations to local and global extinctions. *Conservation Biology*, 8:60–71.

Bolger, A.M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30:2114–2120.

Brown, R.M. & Guttman, S.I. (2002) Phylogenetic systematics of the *Rana signata* complex of Philippine and Bornean stream frogs: reconsideration of Huxley's modification of Wallace's Line at the Oriental–Australian faunal zone interface. *Biological Journal of the Linnean Society*, 76(3):393–461.

Brown, R.M. & Siler, C.D. (2014) Spotted stream frog diversification at the Australasian faunal zone interface, mainland versus island comparisons, and a test of the Philippine 'dual-umbilicus' hypothesis. *Journal of Biogeography*, 41(1):182–195.

Brown, R.M., Linkem, C.W., Siler, C.D., Sukumaran, J., Esselstyn, J.A., Diesmos, A.C., Iskandar, D.T., Bickford, D., Evans, B.J., McGuire, J.A., Grismer, L., Supriatna, J. & Andayani, N. (2010) Phylogeography and historical demography of *Polypedates leucomystax* in the islands of Indonesia and the Philippines: Evidence for recent human-mediated range expansion? *Molecular Phylogenetics and Evolution*, 57(2):598–619.

Capella-Gutierrez, S., Silla-Martinez, J.M. & Gabaldon, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25:1972–1973.

Darwin, C. (1859) *The origin of species*. London: John Murray.

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

De Queiroz, A. (2005) The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology & Evolution*, 20(2):68–73.

Dickerson, R.E., Merrill, E.D., McGregor, R.C., Schultze, W., Taylor, E.H. & Herre, A.W. (1928). *Distribution of life in the Philippines* (No. 21). Bureau of Printing.

Dossa, G.G.O., Paudel, E., Fujinuma, J., Yu, H., Chutipong, W., Zhang, Y. & Paz, S. (2013) Factors Determining Forest Diversity and Biomass on a Tropical Volcano, Mt. Rinjani, Lombok, Indonesia. *PLoS ONE*, 8(7):e67720.

Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*(1):214.

Duellman, W. & Trueb, L. (1986) *Biology of amphibians*. McGraw-Hill, New York.

Earl, D.A. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2):359–361.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5):1792–1797.

Evans, B.J, Brown, R.M., McGuire, J.A., Supriatna, J., Andayani, N., Diesmos, A., Iskandar, D., Melnick, D.J. & Cannatella, D.C. (2003) Phylogenetics of fanged frogs: testing biogeographical hypotheses at the interface of the Asian and Australian faunal zones. *Systematic Biology*, 52(6):794–819.

Grabherr, M., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. & Regev, A. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 15:644–652.

Gronau, I., Hubisz, M. J., Gulko, B., Danko, C.G. & Siepel, A. (2011) Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, 43:1031–1034.

Huang, X. & Madan, A. (1999) CAP3: a DNA sequence assembly program. *Genome Research*, 9:868–877.

Inger, R.F. (1999) Distribution of amphibians in southern Asia and adjacent islands. *Patterns of distribution of amphibians: a global perspective*. Pp.445–482.

Iskandar, D. & Mumpuni (2004a) *Limnonectes dammermani*. The IUCN Red List of Threatened Species 2004: e.T58331A11768050. Downloaded on 15 September 2016.

Iskandar, D. & Mumpuni (2004b) *Limnonectes kadarsani*. The IUCN Red List of Threatened Species 2004: e.T58342A11769525. Downloaded on 15 September 2016.
Iskandar, D.T., Evans, B.J. & McGuire, J.A. (2014) A novel reproductive mode in frogs: a new species of fanged frog with internal fertilization and birth of tadpoles. *PLoS ONE*, 9(12):e115884.
John Measey, G., Vences, M., Drewes, R.C., Chiari, Y., Melo, M. & Bourles, B. (2007) Freshwater paths across the ocean: molecular phylogeny of the frog *Ptychadena newtoni* gives insights into amphibian colonization of oceanic islands. *Journal of Biogeography*, 34(1):7–20.
Kaiser, H., Sanchez, C., Heacox, S., Kathriner, A., Ribeiro, A. V., Soares, Z. A., de Araujo, L.L., Mecke, S. & O'Shea, M. (2013) First report on the herpetofauna of Ataúro Island, Timor-Leste. *Check List*, 9(4):752–762.
Katoh, K. & Standley, D.M. (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, 30:722–780.
Kent, W. (2002) BLAT–the BLAST-like alignment tool. *Genome Research*, 12:656–664.
Kumar, S. & Subramanian, S. (2002) Mutation rates in mammalian genomes. *Proceedings of the National Academy of Sciences USA*, 99(2):803–808.
Langmead, B. & Salzberg, S. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9:357–359.
Lavigne, F., Degeai, J.P., Komorowski, J.C., Guillet, S., Robert, V., Lahitte, P., Oppenheimer, C., Stoffel, M., Vidal, C.M., Pratomo, I., Wassmer, P., Hajdas, I., Sri Hadmoko, D. & de Belizal, E. (2013) Source of the great AD 1257 mystery eruption unveiled, Samalas volcano, Rinjani Volcanic Complex, Indonesia. *Proceedings of the National Academy of Sciences*, 110(42):16742-16747.
Lemmon, A.R., Emme, S.A. & Lemmon, E.M. (2012) Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, 61(5):727–744.
Li, W. & Godzik, A. (2006) CD-Hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22:1658–1659.
Li, H. & Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25:1754–1760.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, 25:2078–2079.
Liu, B., Yuan, J., Yiu, S-M., Li, Z., Xie, Y., Chen, Y., Shi, Y., Zhang, H., Li, Y., Lam, T-W. & Luo, R. (2012) COPE: an accurate k-mer-based pair-end reads connection tool to facilitate genome assembly. *Bioinformatics*, 28:2870–2874.
Losos, J.B., & Ricklefs, R.E. (2009) *The theory of island biogeography revisited*. Princeton University Press. Princeton, NJ.
MacArthur, R.H. & Wilson, E.O. (1967) *Theory of Island Biogeography.* Princeton University Press. Princeton, NJ.
Marrison, G.E. (1999) Sasak and Javanese Literature of Lombok (KITLV, Leiden, Netherlands).
Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17:10–12.
Magoč, T. & Salzberg, S. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, 27:2957–2963.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. & DePristo, M.A. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20:1297–1303.

McLeod, D.S. (2010) Of least concern? Systematics of a cryptic species complex: *Limnonectes kuhlii* (Amphibia: Anura: Dicroglossidae). *Molecular Phylogenetics and Evolution*, 56(3):991-1000.

Meyer, M. & Kircher, M. (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, 2010(6):pdb.prot5448.

Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S. & Warnow, T. (2014) ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics*, 30(17):541–548.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S. & Hoekstra, H.E. (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, 7(5):e37135.

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

Pyron, R.A. (2014) Biogeographic analysis reveals ancient continental vicariance and recent oceanic dispersal in amphibians. *Systematic Biology*, 63(5):779–797.

Rambaut, A. & Drummond, A.J. (2009) Tracer version 1.5.0. Available: http://tree.bio.ed.ac.uk/software/tracer/. (Accessed: 01 July 2013).

Ricklefs, R.E. & Lovette, I.J. (1999) The roles of island area per se and habitat diversity in the species–area relationships of four Lesser Antillean faunal groups. *Journal of Animal Ecology*, 68(6):1142–1160.

Setiadi, M.I., McGuire, J.A., Brown, R.M., Zubairi, M., Iskandar, D.T., Andayani, N., Supriatna, J. & Evans, B.J. (2011) Adaptive radiation and ecological opportunity in Sulawesi and Philippine fanged frog (*Limnonectes*) communities. *The American Naturalist*, 178(2):221-240.

Simberloff, D.S. (1974) Equilibrium theory of island biogeography and ecology. *Annual Review of Ecology and Systematics*, 1974:161–182.

Simpson, J., Wong, K., Jackman, S., Schein, J., Jones, S. & Birol, I. (2009) ABySS: a parallel assembler for short read sequence data. *Genome Research*, 19:1117–1123.

Singhal, S. (2013) *De novo* transcriptomic analyses for non-model organisms: an evaluation of methods across a multi-species data set. *Molecular Ecology Resources*, 13:403–416.

Slater, G.S. & Birney, E. (2005) Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, 6:31.

Smit, A.F.A., Hubley, R. & Green, P. (2015) *RepeatMasker Open-4.0*. Available: http://www.repeatmasker.org. (Accessed: 01 January 2016).

Stamatakis, A. (2014) RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9):1312–1313.

Stresemann, E. (1939) Die vogel von Celebes. *Zoogeographie*. *Journal of Ornithologie*, 87:312–425.

Toews, D.P., & Brelsford, A. (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, 21(16):3907–3930.

Vences, M., Vieites, D.R., Glaw, F., Brinkmann, H., Kosuch, J., Veith, M. & Meyer, A. (2003) Multiple overseas dispersal in amphibians. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1532):2435–2442.

Whittaker, R.J. & Fernández-Palacios, J.M. (2007) *Island biogeography: ecology, evolution, and conservation*. Oxford University Press.

TABLES

Table 1. Summary statistics for transcriptomes sequenced from two *Limnonectes kadarsani*.

Catalog #	Species	Locality	Length (bp)	Total Contigs	Mean Length (bp)	N25	N50	N90	GC%
JAM12330	*L. kadarsani*	Sumbawa	5,405,216	5,432	995	2117	1414	477	46.8
JAM13047	*L. kadarsani*	Lembata	5,286,948	5,192	1,018	1968	1390	509	46.77

Table 2. Locality information for genetic samples. X=exon capture sample.

Sample Number	Genus	Species	Island	Latitude (S)	Longitude (E)	ExonCap
JAM11297	*Limnonectes*	*torajae*	Sulawesi	-	-	X
JAM11299	*Limnonectes*	*modestus*	Sulawesi	-	-	X
JAM11563	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	X
JAM11564	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	
JAM11565	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	
JAM11566	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	X
JAM11567	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	X
JAM11568	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	
JAM11569	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	X
JAM11570	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	
JAM11571	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	X
JAM11572	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	
JAM11573	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	X
JAM11574	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	
JAM11575	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	
JAM11576	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	
JAM11577	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	X
JAM11578	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	
JAM11579	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	X
JAM11580	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	
JAM11583	*Limnonectes*	*kadarsani*	Lombok	-8.53159	116.39878	X
JAM11585	*Limnonectes*	*dammermanni*	Lombok	-8.53159	116.39878	X
JAM11592	*Limnonectes*	*dammermanni*	Lombok	-8.53159	116.39878	X
JAM11593	*Limnonectes*	*dammermanni*	Lombok	-8.53159	116.39878	X
JAM11683	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11684	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11685	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11686	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11687	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11688	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11689	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	X
JAM11690	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11691	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11692	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11693	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11694	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11695	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11696	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11697	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11698	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11699	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	

Table 2 cont.

Sample Number	Genus	Species	Island	Latitude (S)	Longitude (E)	ExonCap
JAM11701	*Limnonectes*	*kadarsani*	Sumbawa	-8.57435	117.31339	
JAM11732	*Limnonectes*	*kadarsani*	Sumbawa	-8.72371	117.38319	X
JAM11733	*Limnonectes*	*kadarsani*	Sumbawa	-8.72371	117.38319	
JAM11734	*Limnonectes*	*kadarsani*	Sumbawa	-8.72371	117.38319	
JAM11735	*Limnonectes*	*kadarsani*	Sumbawa	-8.72371	117.38319	
JAM11736	*Limnonectes*	*kadarsani*	Sumbawa	-8.72371	117.38319	
JAM11737	*Limnonectes*	*kadarsani*	Sumbawa	-8.72371	117.38319	
JAM11738	*Limnonectes*	*kadarsani*	Sumbawa	-8.72371	117.38319	
JAM11739	*Limnonectes*	*kadarsani*	Sumbawa	-8.72371	117.38319	
JAM11740	*Limnonectes*	*kadarsani*	Sumbawa	-8.72371	117.38319	X
JAM11969	*Limnonectes*	*dammermani*	Lombok	-8.40266	116.53959	X
JAM11970	*Limnonectes*	*dammermani*	Lombok	-8.40266	116.53959	X
JAM11971	*Limnonectes*	*dammermani*	Lombok	-8.40266	116.53959	X
JAM11972	*Limnonectes*	*dammermani*	Lombok	-8.40266	116.53959	X
JAM11973	*Limnonectes*	*kadarsani*	Lombok	-8.30545	116.40701	X
JAM11974	*Limnonectes*	*kadarsani*	Lombok	-8.30545	116.40701	
JAM11975	*Limnonectes*	*kadarsani*	Lombok	-8.30545	116.40701	
JAM11976	*Limnonectes*	*kadarsani*	Lombok	-8.30545	116.40701	
JAM12043	*Limnonectes*	*kadarsani*	Sumbawa	-9.02763	116.81893	
JAM12049	*Limnonectes*	*kadarsani*	Sumbawa	-9.02763	116.81893	X
JAM12050	*Limnonectes*	*kadarsani*	Sumbawa	-9.02763	116.81893	
JAM12052	*Limnonectes*	*kadarsani*	Sumbawa	-9.050508	116.863726	
JAM12128	*Limnonectes*	*kadarsani*	Sumbawa	-8.49571	118.49123	
JAM12129	*Limnonectes*	*kadarsani*	Sumbawa	-8.49571	118.49123	
JAM12130	*Limnonectes*	*kadarsani*	Sumbawa	-8.49571	118.49123	
JAM12131	*Limnonectes*	*kadarsani*	Sumbawa	-8.49571	118.49123	X
JAM12205	*Limnonectes*	*kadarsani*	Sumbawa	-8.74106	118.60437	
JAM12208	*Limnonectes*	*kadarsani*	Sumbawa	-8.74106	118.60437	X
JAM12309	*Limnonectes*	*kadarsani*	Sumbawa	-8.76643	118.60496	X
JAM12310	*Limnonectes*	*kadarsani*	Sumbawa	-8.76643	118.60496	
JAM12311	*Limnonectes*	*kadarsani*	Sumbawa	-8.76643	118.60496	
JAM12312	*Limnonectes*	*kadarsani*	Sumbawa	-8.76643	118.60496	X
JAM12347	*Limnonectes*	*kadarsani*	Flores	-8.67925	120.3015	
JAM12348	*Limnonectes*	*kadarsani*	Flores	-8.67925	120.3015	
JAM12349	*Limnonectes*	*kadarsani*	Flores	-8.67925	120.3015	
JAM12350	*Limnonectes*	*kadarsani*	Flores	-8.67925	120.3015	
JAM12351	*Limnonectes*	*kadarsani*	Flores	-8.67925	120.3015	X
JAM12352	*Limnonectes*	*kadarsani*	Flores	-8.67925	120.3015	
JAM12353	*Limnonectes*	*kadarsani*	Flores	-8.67925	120.3015	X
JAM12378	*Limnonectes*	*kadarsani*	Flores	-8.70574	121.77329	
JAM12396	*Limnonectes*	*kadarsani*	Flores	-8.72257	121.75079	X

Table 2 cont.

Sample Number	Genus	Species	Island	Latitude (S)	Longitude (E)	ExonCap
JAM12397	*Limnonectes*	*kadarsani*	Flores	-8.72257	121.75079	
JAM12415	*Limnonectes*	*kadarsani*	Flores	-8.594198	119.968570	
JAM12416	*Limnonectes*	*kadarsani*	Flores	-8.59810	119.961020	
JAM12417	*Limnonectes*	*kadarsani*	Flores	-8.59810	119.961020	
JAM12418	*Limnonectes*	*kadarsani*	Flores	-8.59810	119.961020	
JAM12419	*Limnonectes*	*kadarsani*	Flores	-8.59810	119.961020	
JAM12420	*Limnonectes*	*kadarsani*	Flores	-8.59810	119.961020	
JAM12421	*Limnonectes*	*kadarsani*	Flores	-8.59810	119.961020	
JAM12432	*Limnonectes*	*kadarsani*	Flores	-8.59027	120.43418	
JAM12433	*Limnonectes*	*kadarsani*	Flores	-8.59027	120.43418	
JAM12434	*Limnonectes*	*kadarsani*	Flores	-8.59027	120.43418	
JAM12435	*Limnonectes*	*kadarsani*	Flores	-8.59027	120.43418	
JAM12436	*Limnonectes*	*kadarsani*	Flores	-8.59027	120.43418	
JAM12437	*Limnonectes*	*kadarsani*	Flores	-8.59027	120.43418	
JAM12438	*Limnonectes*	*kadarsani*	Flores	-8.59027	120.43418	
JAM12439	*Limnonectes*	*kadarsani*	Flores	-8.59027	120.43418	
JAM12440	*Limnonectes*	*kadarsani*	Flores	-8.594198	119.968570	X
JAM12441	*Limnonectes*	*kadarsani*	Flores	-8.594198	119.968570	
JAM12442	*Limnonectes*	*kadarsani*	Flores	-8.594198	119.968570	
JAM12443	*Limnonectes*	*kadarsani*	Flores	-8.594198	119.968570	
JAM12444	*Limnonectes*	*kadarsani*	Flores	-8.594198	119.968570	
JAM12445	*Limnonectes*	*kadarsani*	Flores	-8.594198	119.968570	X
JAM12446	*Limnonectes*	*kadarsani*	Flores	-8.594198	119.968570	
JAM12500	*Limnonectes*	*kadarsani*	Flores	-8.68687	121.03889	
JAM12501	*Limnonectes*	*kadarsani*	Flores	-8.68687	121.03889	
JAM12502	*Limnonectes*	*kadarsani*	Flores	-8.68687	121.03889	
JAM12503	*Limnonectes*	*kadarsani*	Flores	-8.68687	121.03889	
JAM12504	*Limnonectes*	*kadarsani*	Flores	-8.68687	121.03889	
JAM12505	*Limnonectes*	*kadarsani*	Flores	-8.68687	121.03889	
JAM12506	*Limnonectes*	*kadarsani*	Flores	-8.68687	121.03889	
JAM12507	*Limnonectes*	*kadarsani*	Flores	-8.68687	121.03889	
JAM12508	*Limnonectes*	*kadarsani*	Flores	-8.68687	121.03889	
JAM12509	*Limnonectes*	*kadarsani*	Flores	-8.68687	121.03889	
JAM12510	*Limnonectes*	*kadarsani*	Flores	-8.68687	121.03889	
JAM12511	*Limnonectes*	*kadarsani*	Flores	-8.68687	121.03889	X
JAM12512	*Limnonectes*	*kadarsani*	Flores	-8.75954	121.70090	
JAM12513	*Limnonectes*	*kadarsani*	Flores	-8.75954	121.70090	
JAM12514	*Limnonectes*	*kadarsani*	Flores	-8.75954	121.70090	
JAM12515	*Limnonectes*	*kadarsani*	Flores	-8.75954	121.70090	
JAM12516	*Limnonectes*	*kadarsani*	Flores	-8.75954	121.70090	
JAM12517	*Limnonectes*	*kadarsani*	Flores	-8.75954	121.70090	

Table 2 cont.

Sample Number	Genus	Species	Island	Latitude (S)	Longitude (E)	ExonCap
JAM12518	*Limnonectes*	*kadarsani*	Flores	-8.75954	121.70090	
JAM12519	*Limnonectes*	*kadarsani*	Flores	-8.75954	121.70090	
JAM12520	*Limnonectes*	*kadarsani*	Flores	-8.75954	121.70090	X
JAM12521	*Limnonectes*	*kadarsani*	Flores	-8.75954	121.70090	
JAM12522	*Limnonectes*	*kadarsani*	Flores	-8.75954	121.70090	
JAM12638	*Limnonectes*	*kadarsani*	Lembata	-8.38791	123.40071	
JAM12639	*Limnonectes*	*kadarsani*	Lembata	-8.38791	123.40071	
JAM12640	*Limnonectes*	*kadarsani*	Lembata	-8.38791	123.40071	
JAM12641	*Limnonectes*	*kadarsani*	Lembata	-8.38791	123.40071	
JAM12642	*Limnonectes*	*kadarsani*	Lembata	-8.38791	123.40071	
JAM12643	*Limnonectes*	*kadarsani*	Lembata	-8.38791	123.40071	X
JAM12644	*Limnonectes*	*kadarsani*	Lembata	-8.38791	123.40071	
JAM12750	*Limnonectes*	*kadarsani*	Lembata	-8.38181	123.40195	
JAM12751	*Limnonectes*	*kadarsani*	Lembata	-8.38181	123.40195	
JAM12752	*Limnonectes*	*kadarsani*	Lembata	-8.38181	123.40195	
JAM12753	*Limnonectes*	*kadarsani*	Lembata	-8.38181	123.40195	X
JAM12754	*Limnonectes*	*kadarsani*	Lembata	-8.38181	123.40195	X
JAM12755	*Limnonectes*	*kadarsani*	Lembata	-8.38181	123.40195	
JAM12792	*Limnonectes*	*kadarsani*	Lembata	-8.38181	123.40195	
JAM12793	*Limnonectes*	*kadarsani*	Lembata	-8.38181	123.40195	
JAM12794	*Limnonectes*	*kadarsani*	Lembata	-8.38181	123.40195	
JAM12795	*Limnonectes*	*kadarsani*	Lembata	-8.38181	123.40195	X
JAM12797	*Limnonectes*	*kadarsani*	Lembata	-8.38181	123.40195	X
JAM13030	*Limnonectes*	*kadarsani*	Flores	-8.21648	122.97288	X
JAM13031	*Limnonectes*	*kadarsani*	Flores	-8.21648	122.97288	X
JAM13032	*Limnonectes*	*kadarsani*	Flores	-8.21648	122.97288	
JAM13033	*Limnonectes*	*kadarsani*	Flores	-8.21648	122.97288	X
JAM13046	*Limnonectes*	*kadarsani*	Lembata	-8.471636	123.350576	X
JAM13047	*Limnonectes*	*kadarsani*	Lembata	-8.471636	123.350576	X
JAM13048	*Limnonectes*	*kadarsani*	Lembata	-8.471636	123.350576	
JAM13049	*Limnonectes*	*kadarsani*	Lembata	-8.471636	123.350576	X
JAM13050	*Limnonectes*	*kadarsani*	Lembata	-8.471636	123.350576	
MVZ254313	*Limnonectes*	*microdiscus*	Java	-6.83936	106.92757	X
MVZ254314	*Limnonectes*	*microdiscus*	Java	-6.83936	106.92757	X
MVZ254316	*Limnonectes*	*microdiscus*	Java	-6.83936	106.92757	X

Table 3. Converted G-PHOCS demographic parameters for parapatrically distributed *Limnonectes* lineages on Flores island.

Parameter	Comparison	Mean	95% Low	95% High
Effective Population Size (individuals)				
	West Flores	1,793,409	1,730,455	1,857,045
	East Flores	464,205	441,705	487,045
	Flores Ancestor	602,386	553,977	650,909
Divergence Time (years)				
	Flores Ancestor	1,905,455	1,818,636	1,992,273
Migration (migrants per generation)				
	West to East	2.57	2.23	2.91
	East to West	0.21	0.03	0.39

FIGURE LEGENDS

Figure 1. *Limnonectes kadarsani* from Lombok (A), Sumbawa (B), and *L. dammermani* from Lombok (C and D).

Figure 2. Map of southern Indonesia with land represented in gray scale and ocean in shades of blue. The yellow box outlines the zoomed inset below of the western inner arc islands of the Lesser Sundas with yellow dots representing sample localities used in both mtDNA and nDNA analyses, and yellow diamonds representing localities with only mtDNA.

Figure 3. Maximum Likelihood phylogeny of the *16S* mitochondrial gene for *Limnonectes microdiscus, L. dammermani,* and *L. kadarsani*. Numbers at nodes represent bootstrap support. The branch lengths are scaled by the number of mutations, with the scale bar in the lower left representing the percent mutations.

Figure 4. Time calibrated Bayesian phylogeny of the *16S* mitochondrial gene for *Limnonectes microdiscus, L. dammermani,* and *L. kadarsani*. Bars at nodes represent the 95% confidence intervals for node ages (in millions of years), numbers at nodes represent posterior probability support, and the colored shapes to the right of clades correspond to the map in Figure 5.

Figure 5. Map of Lesser Sundas *Limnonectes* samples in the mtDNA phylogeny. The color and shape of each locality correspond to the lineages defined in Figures 3 and 4.

Figure 6. Plots of the average coverage for the targeted genomic regions in each of the 48 samples. Black horizontal line represents the average coverage for that individual, darker bars represent the 95% confidence interval, and the dashed lines represent the range of coverage.

Figure 7. Plots of the average coverage for the flanking regions in each of the 48 samples. Black horizontal line represents the average coverage for that individual, darker bars represent the 95% confidence interval, and the dashed lines represent the range of coverage.

Figure 8. Scatter plots of the 974 nuclear gene dataset where each dot represents one nuclear gene. A) The number of informative sites plotted against the length of the alignment. B) The percentage of gaps plotted against the alignment length. C) The number of taxa per alignment plotted against the alignment length.

Figure 9. Frequency plots of the 974 nuclear gene dataset. A) The number of taxa present in each gene alignment. B) The distribution of sequence lengths. C) The distribution of percent informative sites. D) The distribution of percent missing data.

Figure 10. Maximum Likelihood tree based on the concatenated nuclear dataset produced by RAXML.

Figure 11. Coalescent species tree of 974 nuclear gene trees produced by ASTRAL-II. Colors of the tree correspond to the colored regions of the map.

Figure 12. Population structure results for *Limnonectes kadarsani.* A) Delta K estimates for the most probable number of populations (K) as determined by STRUCTURE HARVESTER. B) Mean estimates of the Ln Probability of the Data for K=1-6. C) Population structure bar plots for an estimated number of populations K=2-5. Each bar represents one individual frog, and the proportion of each color represents the percent of ancestry.

Figure 13. Unconverted posterior probability distributions for demographic analysis of West versus East Flores lineages. A) Theta posterior distributions for effective population size. Purple=East Flores, Red=Flores Ancestor, Gray=West Flores. B) Tau posterior distribution for the population divergence time between West and East Flores lineages. C) Migration posterior distributions for West into East Flores (Gray color) and for East into West Flores (Purple color).

Figure 14. General allopatric speciation colonization scenario for *Limnonectes* from the Lesser Sunda Islands consistent with the topology of the genomic trees. Gray shaded islands indicate the presence of *Limnonectes*. A) Initial colonization of Lombok from either Java or Bali. B) Dispersal of *L. dammermani* from Lombok into eastern islands. C) Allopatric speciation on eastern islands resulting in the formation of *L. kadarsani*. D) Back-colonization of Lombok by *L. kadarsani*.

Figure 15. General "*in situ*" speciation colonization scenario for *Limnonectes* from the Lesser Sunda Islands consistent with the topology of the genomic trees. Gray shaded islands indicate the presence of *Limnonectes*. A) Initial colonization of Lombok from either Java or Bali. B) *In situ* speciation on Lombok. C) Dispersal into the eastern islands by *L. kadarsani*.

Figure 16. Hypothetical allopatric speciation colonization scenario #1. Gray shaded islands indicate the presence of *Limnonectes*. A) Initial colonization of Lombok from either Java or Bali. B) allopatric species formation of *L. dammermani*. C) Dispersal from Lombok to East Flores and subsequent allopatric species formation of *L. kadarsani*. D) Back colonization of Lombok from East Flores. E) Dispersal from East Flores to Sumbawa, and then from Sumbawa to West Flores. F) Dispersal from West Flores to Lembata. G) Current distribution.

Figure 17. Hypothetical allopatric speciation colonization scenario #2. Gray shaded islands indicate the presence of *Limnonectes*. A) Initial colonization of Lombok from either Java or Bali. B) allopatric species formation of *L. dammermani*. C) Dispersal from Lombok to East Flores and subsequent allopatric species formation of *L. kadarsani*. D) Back colonization of Lombok from East Flores. E) Range expansion of East Flores into West Flores and subsequent dispersal from West Flores to Sumbawa. F) Dispersal from West Flores to Lembata. G) Current distribution.

Figure 18. Hypothetical allopatric speciation colonization scenario #3. Gray shaded islands indicate the presence of *Limnonectes*. A) Initial colonization of Lombok from either Java or Bali. B) allopatric species formation of *L. dammermani*. C) Dispersal from Lombok to Sumbawa followed by allopatric divergence and formation of *L. kadarsani*. D) Back colonization of Lombok from Sumbawa. E) Dispersal from Sumbawa to East Flores. F) Dispersal from Sumbawa to West Flores followed by dispersal from West Flores to Lembata. G) Current distribution.

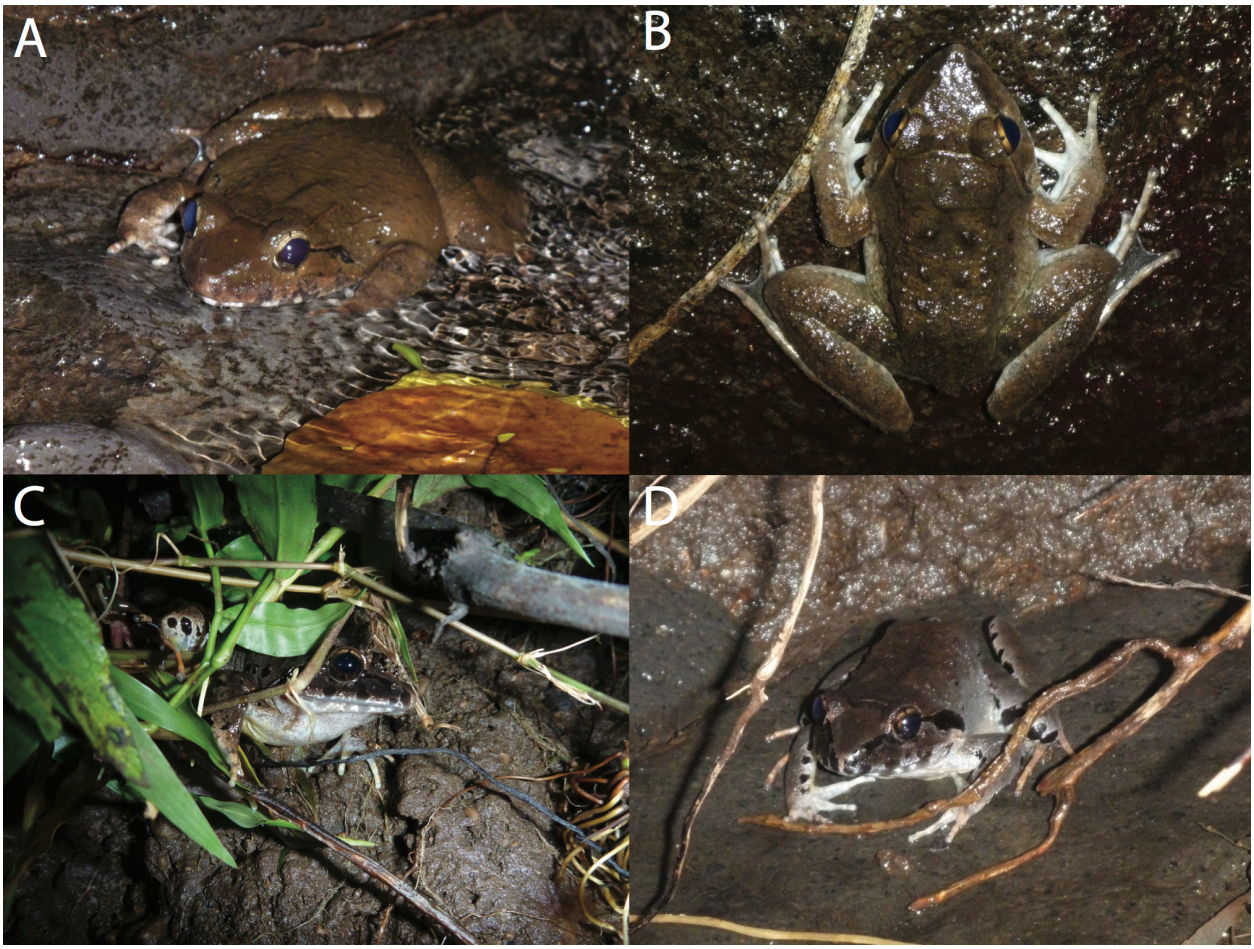


Figure 1.

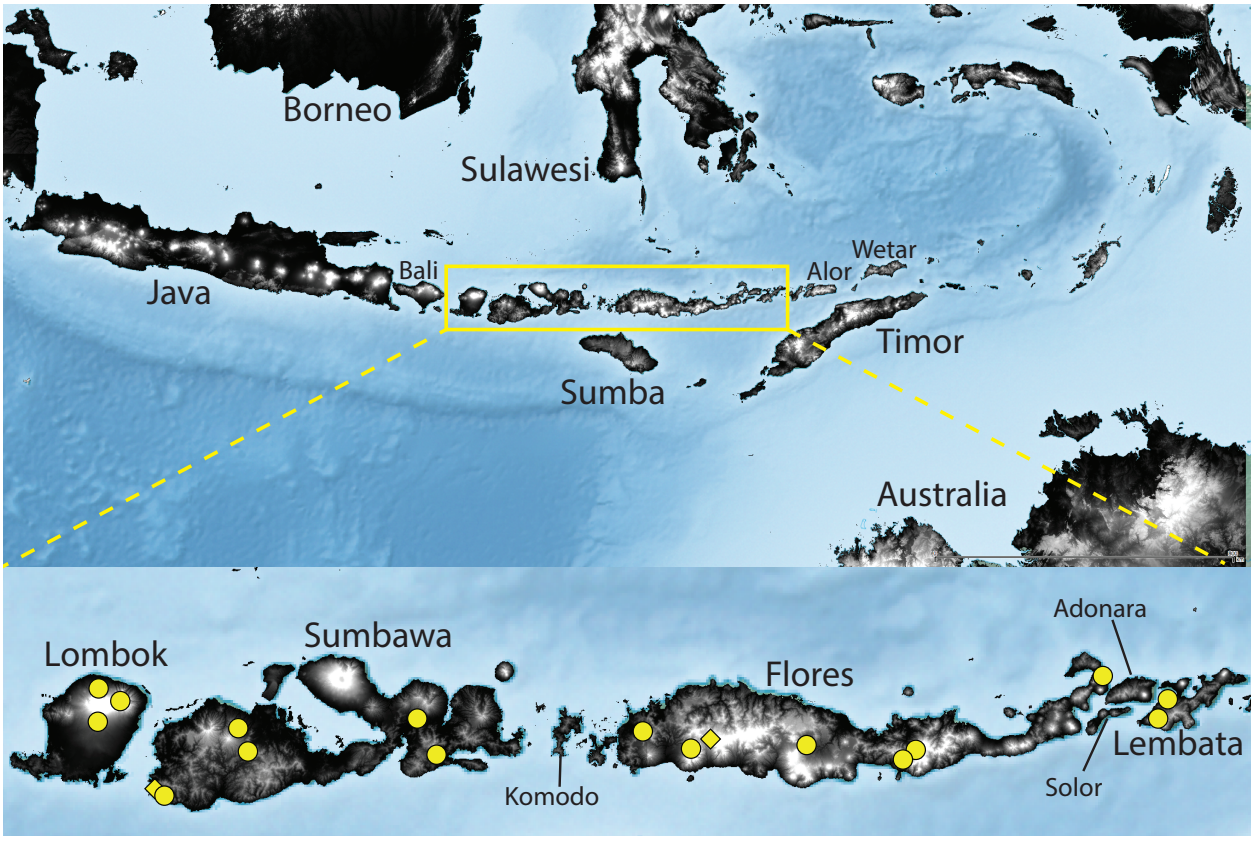

Borneo
Sulawesi
Java
Bali
Wetar
Alor
Timor
Sumba
Australia
Lombok
Sumbawa
Flores
Adonara
Lembata
Solor
Komodo

Figure 2.

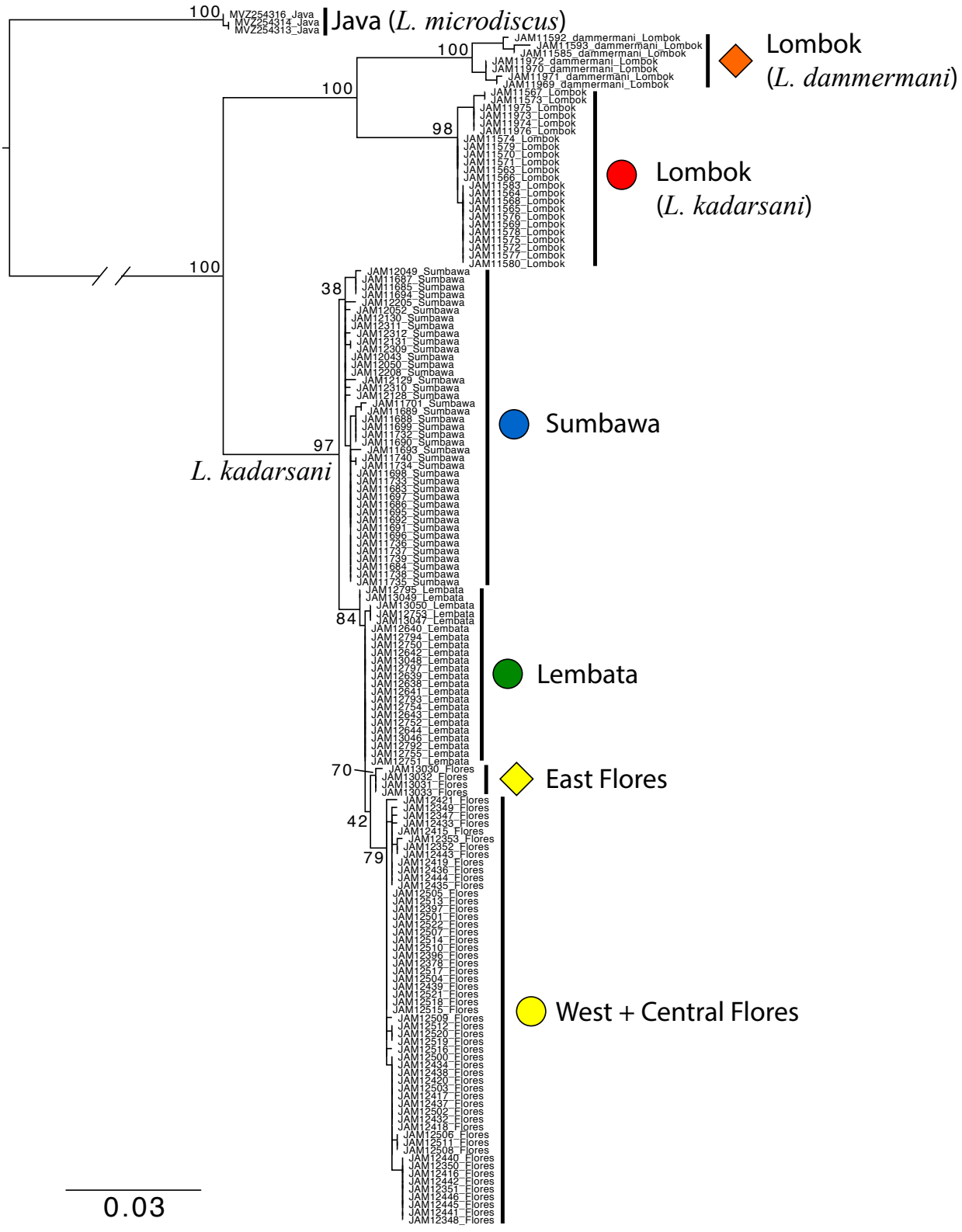

Java (L. microdiscus)
Lombok (L. dammermani)
Lombok (L. kadarsani)
Sumbawa
Lembata
East Flores
West + Central Flores
L. kadarsani
0.03

Figure 3.

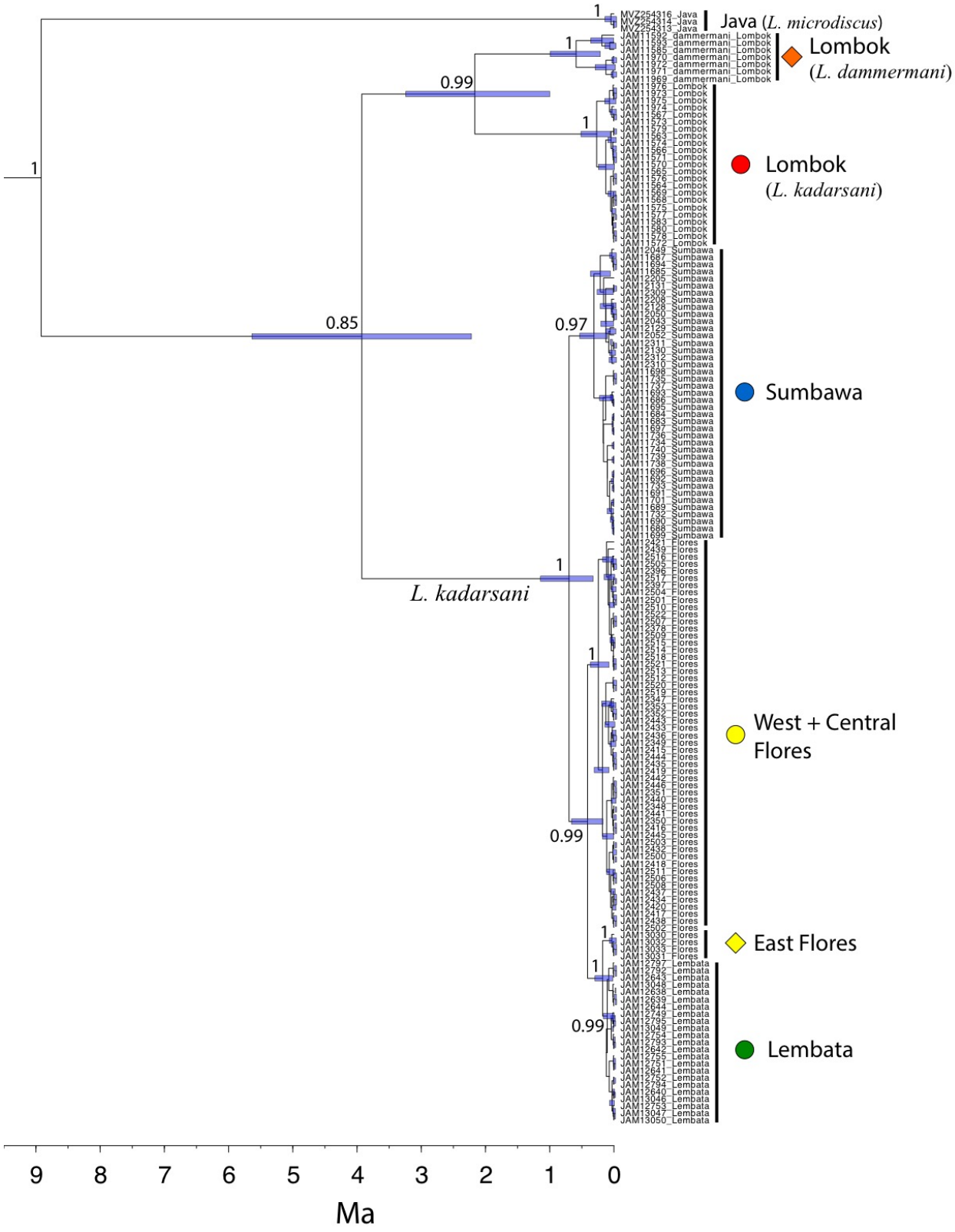

Java (L. microdiscus)
Lombok
(L. dammermani)
Lombok
(L. kadarsani)
Sumbawa
West + Central
Flores
East Flores
Lembata
L. kadarsani
1
0.99
0.85
0.97
9
8
7
6
5
4
3
2
1
0
Ma

Figure 4.

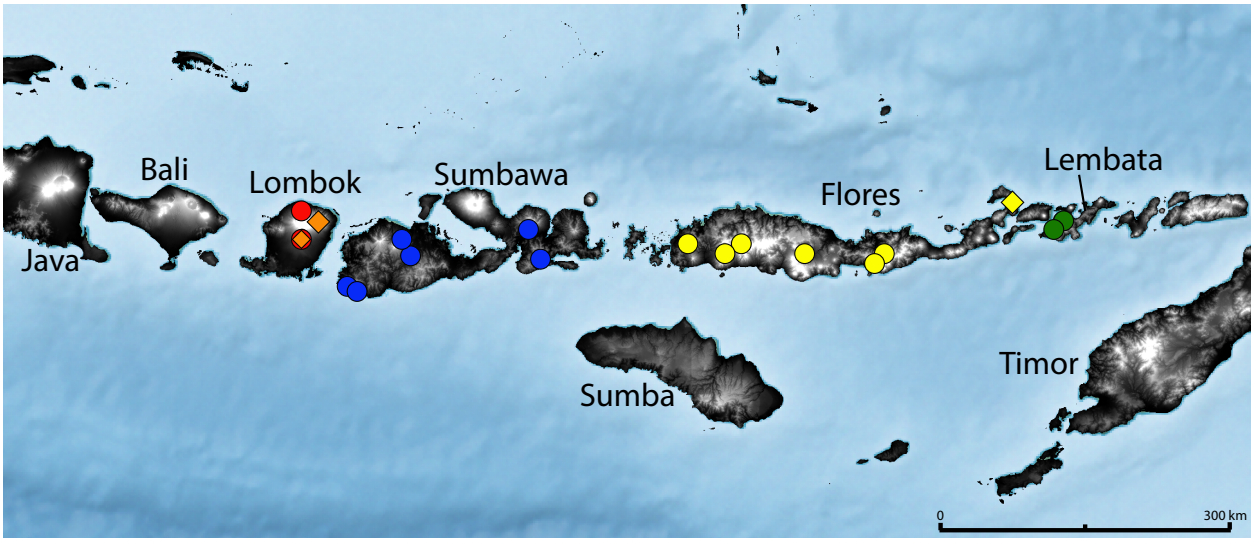

Java
Bali
Lombok
Sumbawa
Flores
Lembata
Timor
Sumba
0
300 km

Figure 5.

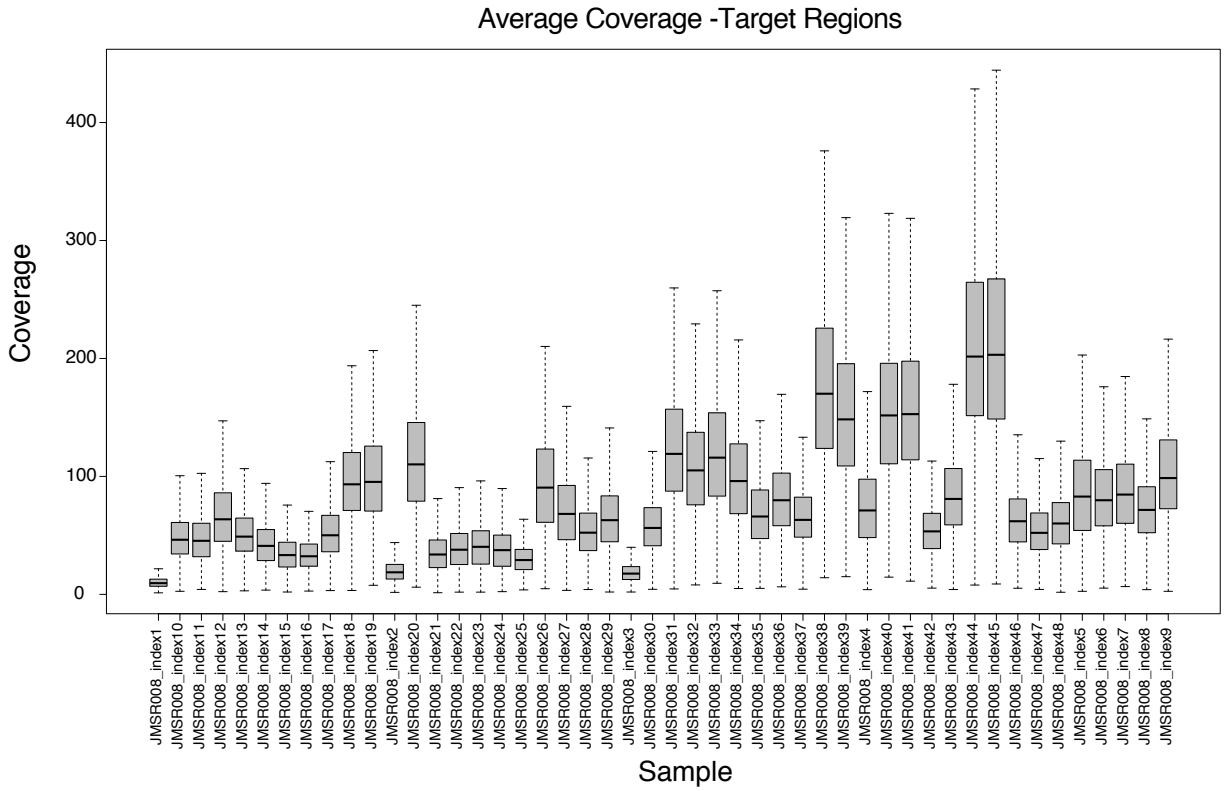

Average Coverage -Target Regions
Coverage
0
100
200
300
400
JMSR008_index1
JMSR008_index10
JMSR008_index11
JMSR008_index12
JMSR008_index13
JMSR008_index14
JMSR008_index15
JMSR008_index16
JMSR008_index17
JMSR008_index18
JMSR008_index19
JMSR008_index2
JMSR008_index20
JMSR008_index21
JMSR008_index22
JMSR008_index23
JMSR008_index24
JMSR008_index25
JMSR008_index26
JMSR008_index27
JMSR008_index28
JMSR008_index29
JMSR008_index3
JMSR008_index30
JMSR008_index31
JMSR008_index32
JMSR008_index33
JMSR008_index34
JMSR008_index35
JMSR008_index36
JMSR008_index37
JMSR008_index38
JMSR008_index39
JMSR008_index4
JMSR008_index40
JMSR008_index41
JMSR008_index42
JMSR008_index43
JMSR008_index44
JMSR008_index45
JMSR008_index46
JMSR008_index47
JMSR008_index48
JMSR008_index5
JMSR008_index6
JMSR008_index7
JMSR008_index8
JMSR008_index9
Sample

Figure 6.

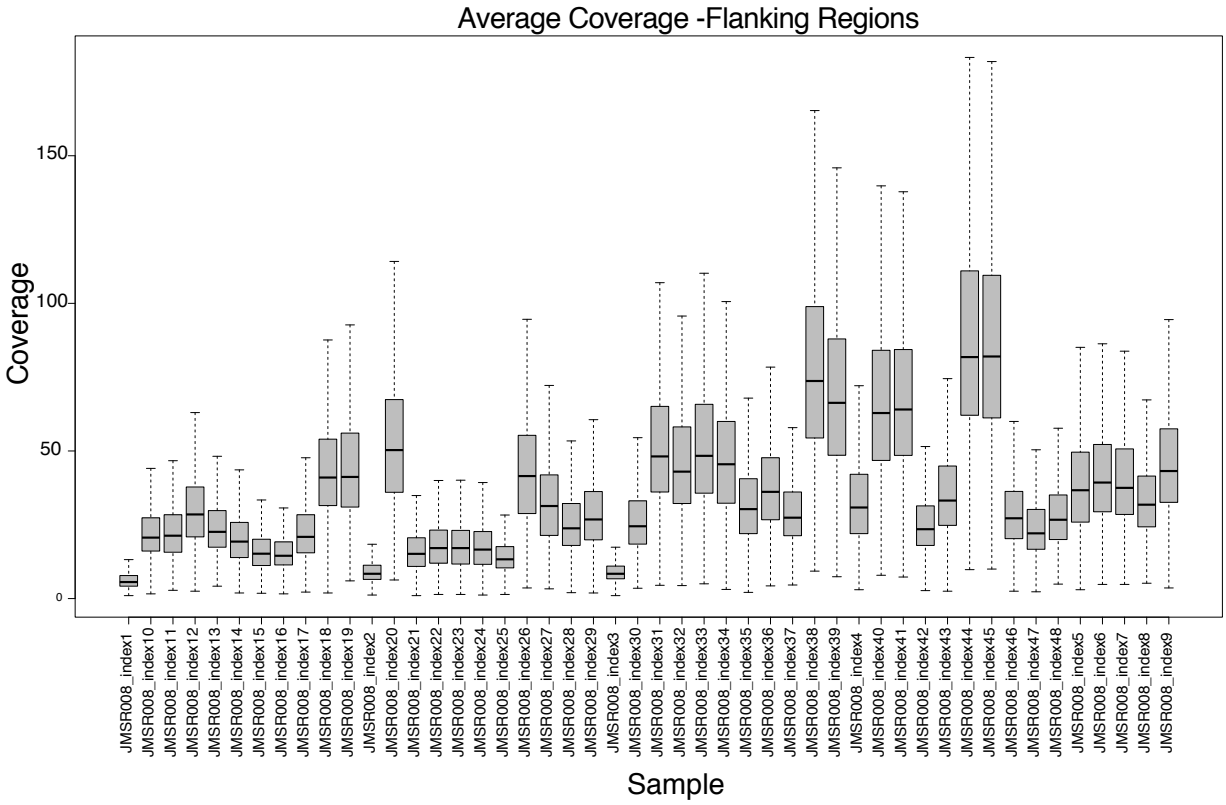

Average Coverage -Flanking Regions
Coverage
150
100
50
0
JMSR008_index1
JMSR008_index10
JMSR008_index11
JMSR008_index12
JMSR008_index13
JMSR008_index14
JMSR008_index15
JMSR008_index16
JMSR008_index17
JMSR008_index18
JMSR008_index19
JMSR008_index2
JMSR008_index20
JMSR008_index21
JMSR008_index22
JMSR008_index23
JMSR008_index24
JMSR008_index25
JMSR008_index26
JMSR008_index27
JMSR008_index28
JMSR008_index29
JMSR008_index3
JMSR008_index30
JMSR008_index31
JMSR008_index32
JMSR008_index33
JMSR008_index34
JMSR008_index35
JMSR008_index36
JMSR008_index37
JMSR008_index38
JMSR008_index39
JMSR008_index4
JMSR008_index40
JMSR008_index41
JMSR008_index42
JMSR008_index43
JMSR008_index44
JMSR008_index45
JMSR008_index46
JMSR008_index47
JMSR008_index48
JMSR008_index5
JMSR008_index6
JMSR008_index7
JMSR008_index8
JMSR008_index9
Sample

Figure 7.

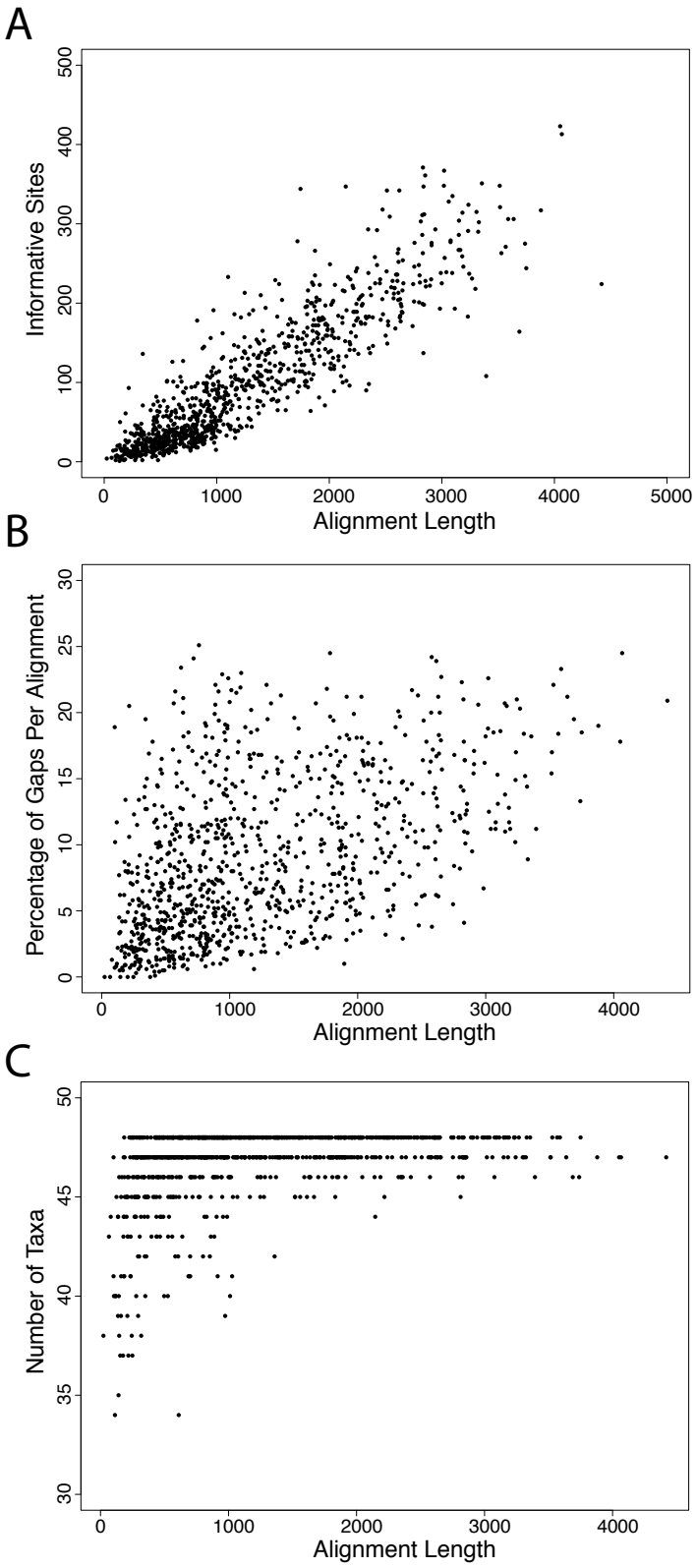

A
Informative Sites
0
100
200
300
400
500
0
1000
2000
3000
4000
5000
Alignment Length
B
Percentage of Gaps Per Alignment
0
5
10
15
20
25
30
0
1000
2000
3000
4000
Alignment Length
C
Number of Taxa
30
35
40
45
50
0
1000
2000
3000
4000
Alignment Length

Figure 8.

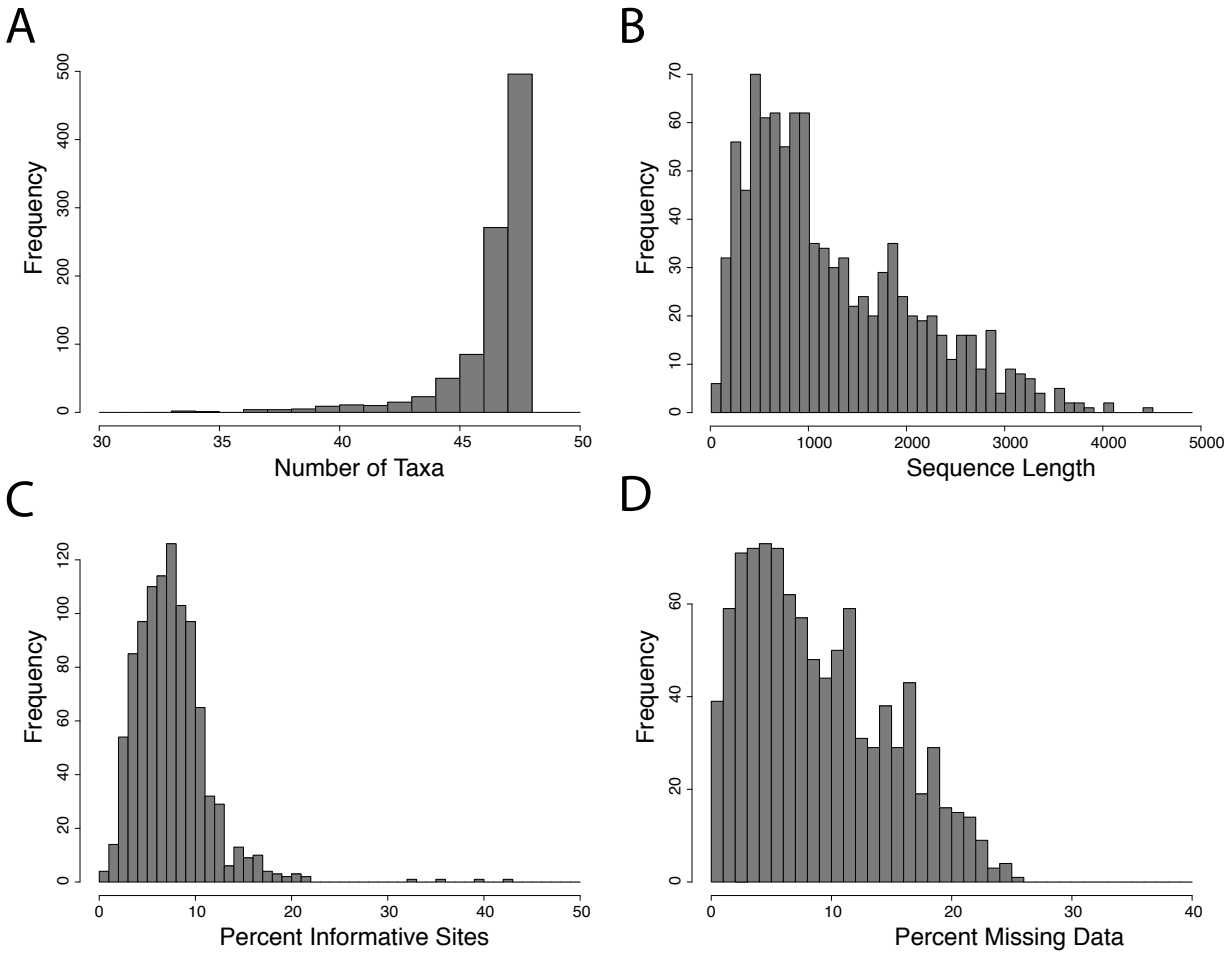

A
Frequency
Number of Taxa
B
Frequency
Sequence Length
C
Frequency
Percent Informative Sites
D
Frequency
Percent Missing Data

Figure 9.

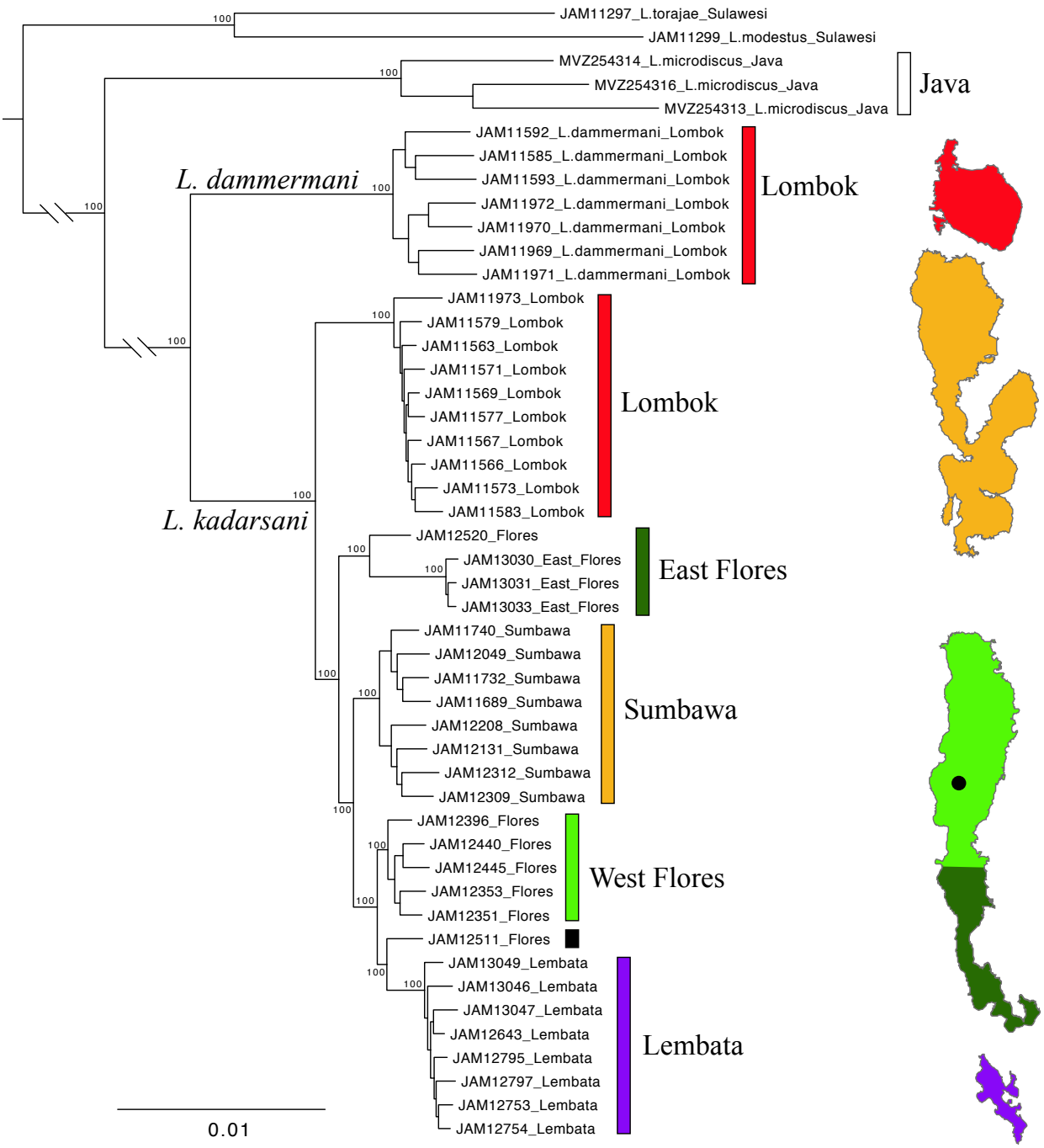

JAM11297_L.torajae_Sulawesi
JAM11299_L.modestus_Sulawesi
MVZ254314_L.microdiscus_Java
MVZ254316_L.microdiscus_Java
MVZ254313_L.microdiscus_Java
Java
L. dammermani
JAM11592_L.dammermani_Lombok
JAM11585_L.dammermani_Lombok
JAM11593_L.dammermani_Lombok
JAM11972_L.dammermani_Lombok
JAM11970_L.dammermani_Lombok
JAM11969_L.dammermani_Lombok
JAM11971_L.dammermani_Lombok
Lombok
JAM11973_Lombok
JAM11579_Lombok
JAM11563_Lombok
JAM11571_Lombok
JAM11569_Lombok
JAM11577_Lombok
JAM11567_Lombok
JAM11566_Lombok
JAM11573_Lombok
JAM11583_Lombok
Lombok
L. kadarsani
JAM12520_Flores
JAM13030_East_Flores
JAM13031_East_Flores
JAM13033_East_Flores
East Flores
JAM11740_Sumbawa
JAM12049_Sumbawa
JAM11732_Sumbawa
JAM11689_Sumbawa
JAM12208_Sumbawa
JAM12131_Sumbawa
JAM12312_Sumbawa
JAM12309_Sumbawa
Sumbawa
JAM12396_Flores
JAM12440_Flores
JAM12445_Flores
JAM12353_Flores
JAM12351_Flores
West Flores
JAM12511_Flores
JAM13049_Lembata
JAM13046_Lembata
JAM13047_Lembata
JAM12643_Lembata
JAM12795_Lembata
JAM12797_Lembata
JAM12753_Lembata
JAM12754_Lembata
Lembata
0.01

Figure 10.

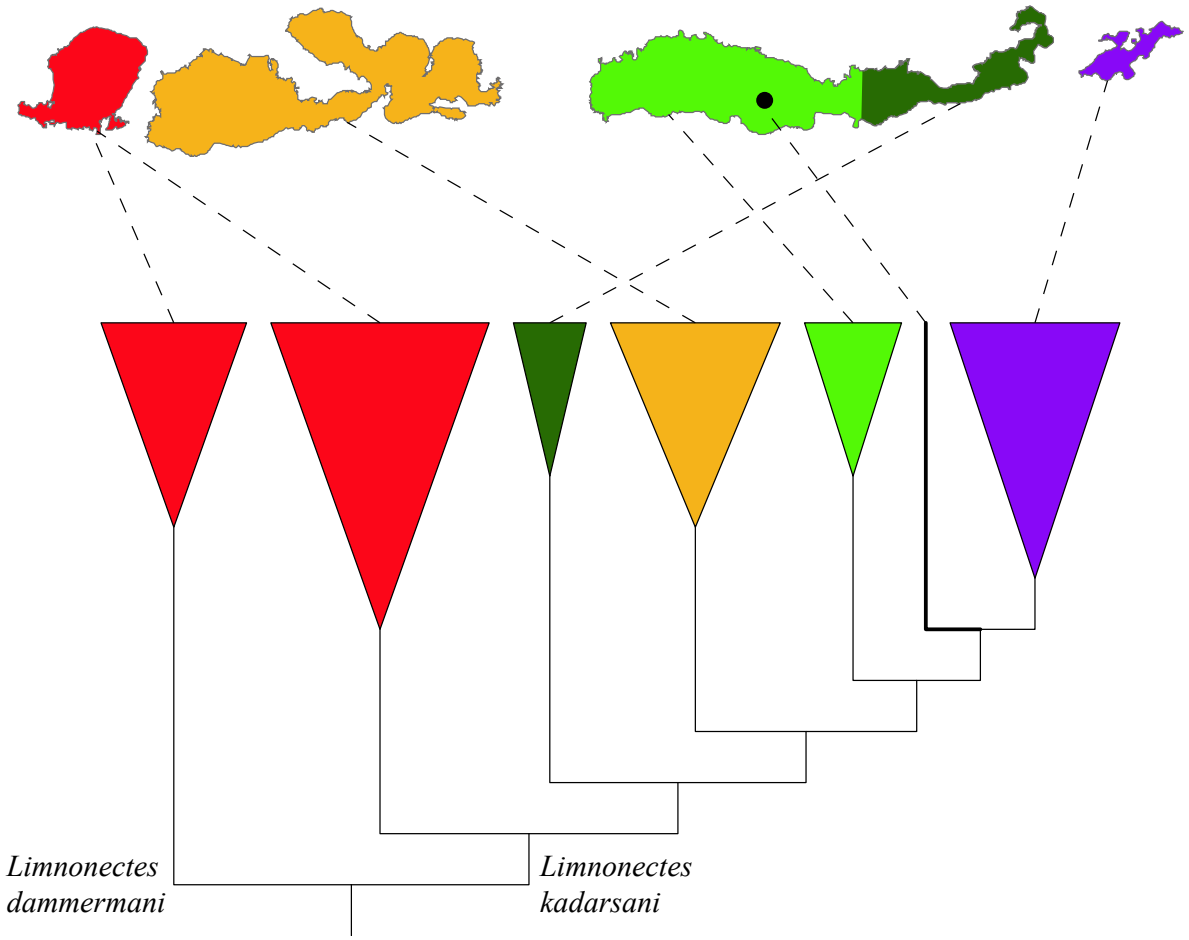


Figure 11.

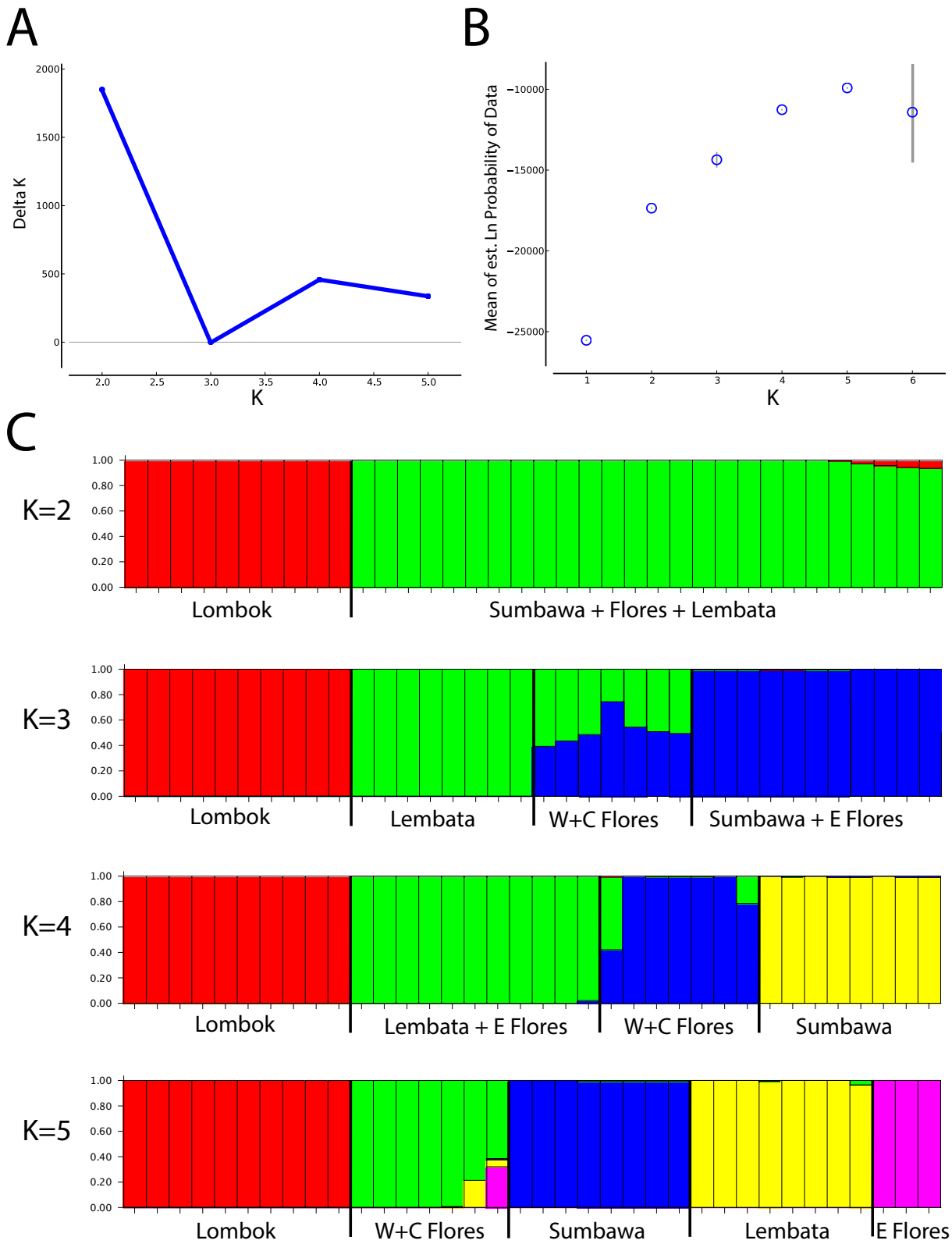

A
Delta K
2000
1500
1000
500
0
2.0
2.5
3.0
3.5
4.0
4.5
5.0
K
B
Mean of est. Ln Probability of Data
−10000
−15000
−20000
−25000
1
2
3
4
5
6
K
C
K=2
1.00
0.80
0.60
0.40
0.20
0.00
Lombok
Sumbawa + Flores + Lembata
K=3
Lombok
Lembata
W+C Flores
Sumbawa + E Flores
K=4
Lombok
Lembata + E Flores
W+C Flores
Sumbawa
K=5
Lombok
W+C Flores
Sumbawa
Lembata
E Flores

Figure 12.

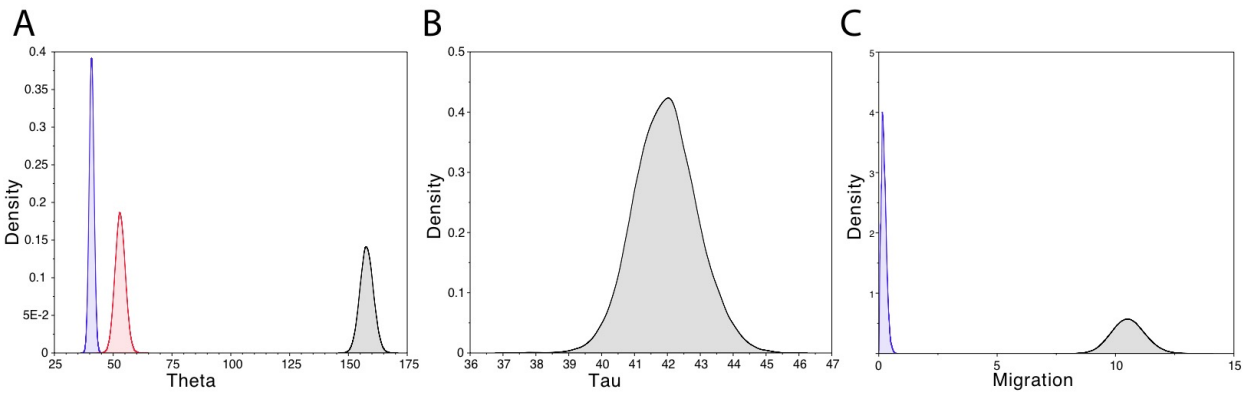

A
B
C
Density
Theta
Tau
Migration
0.4
0.35
0.3
0.25
0.2
0.15
0.1
5E-2
0
25
50
75
100
125
150
175
0.5
0.4
0.3
0.2
0.1
36
37
38
39
40
41
42
43
44
45
46
47
5
4
3
2
1
0
5
10
15

Figure 13.

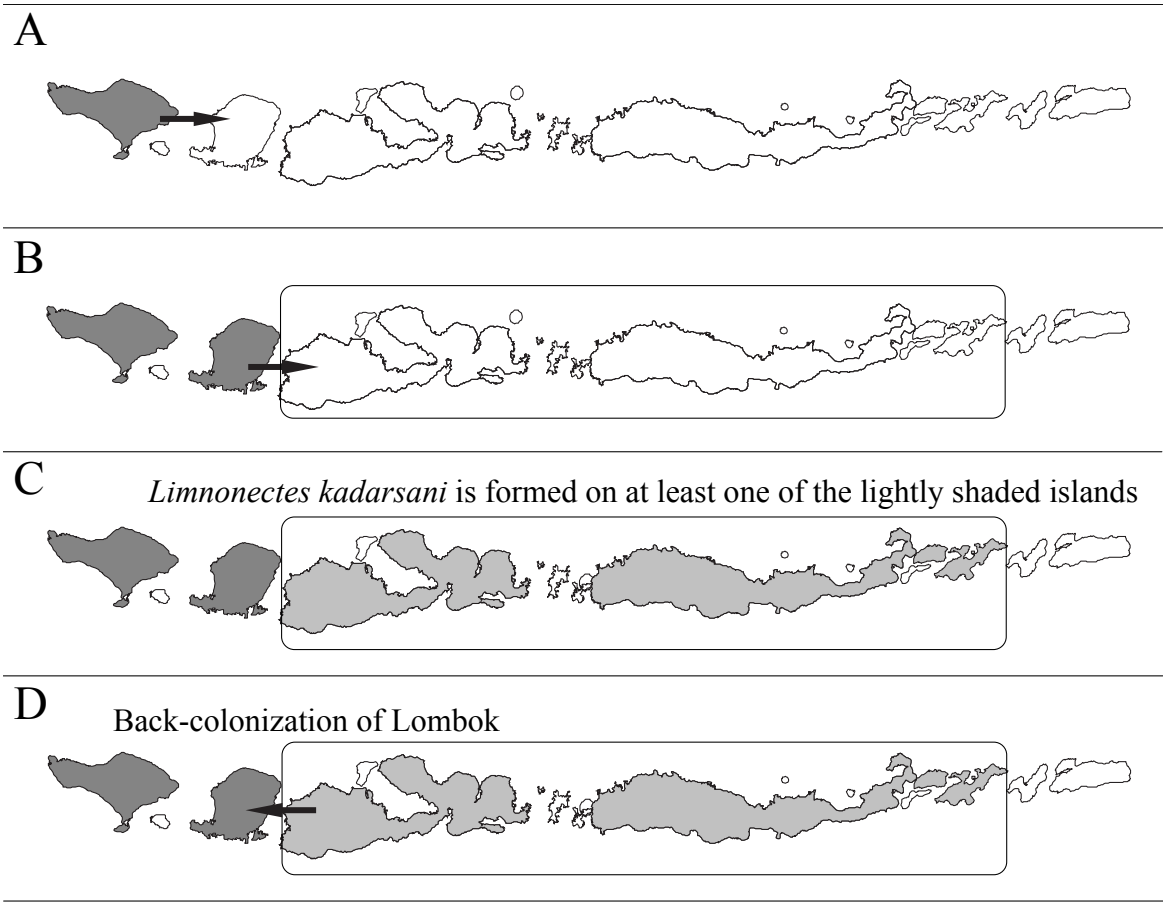

A
B
C
Limnonectes kadarsani is formed on at least one of the lightly shaded islands
D
Back-colonization of Lombok

Figure 14.

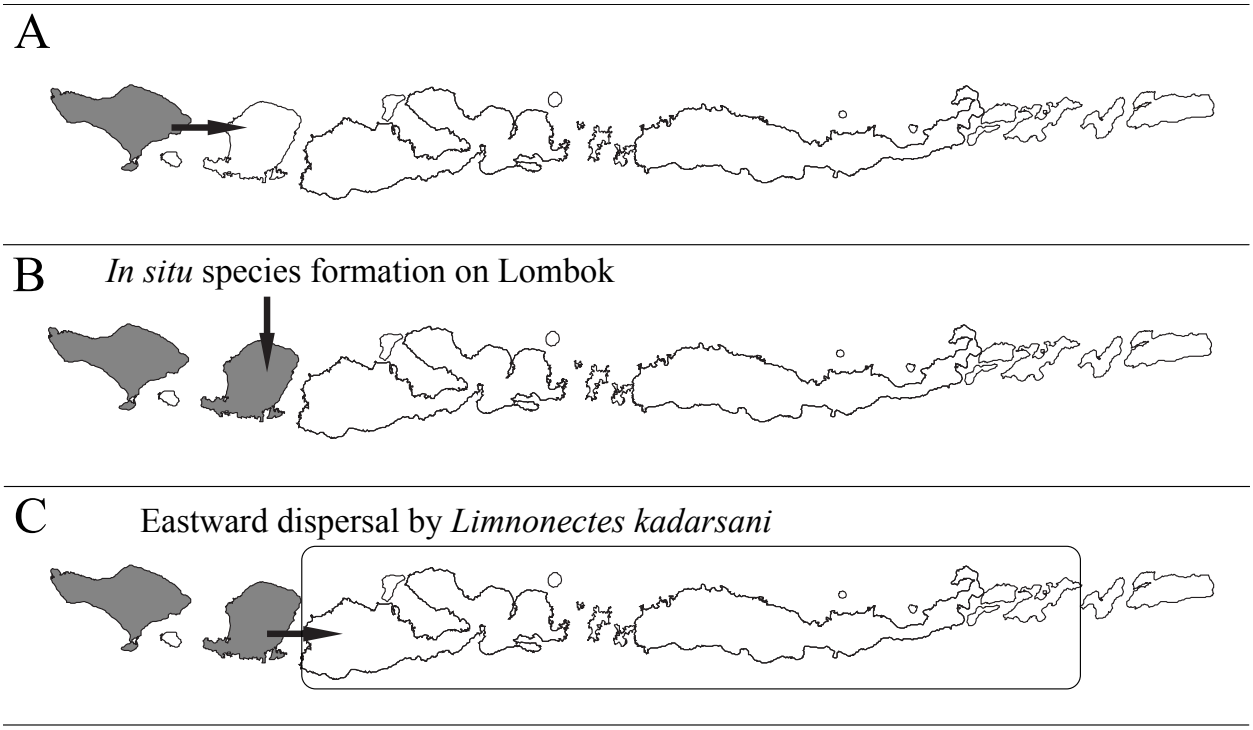

A
B
In situ species formation on Lombok
C
Eastward dispersal by *Limnonectes kadarsani*

Figure 15.

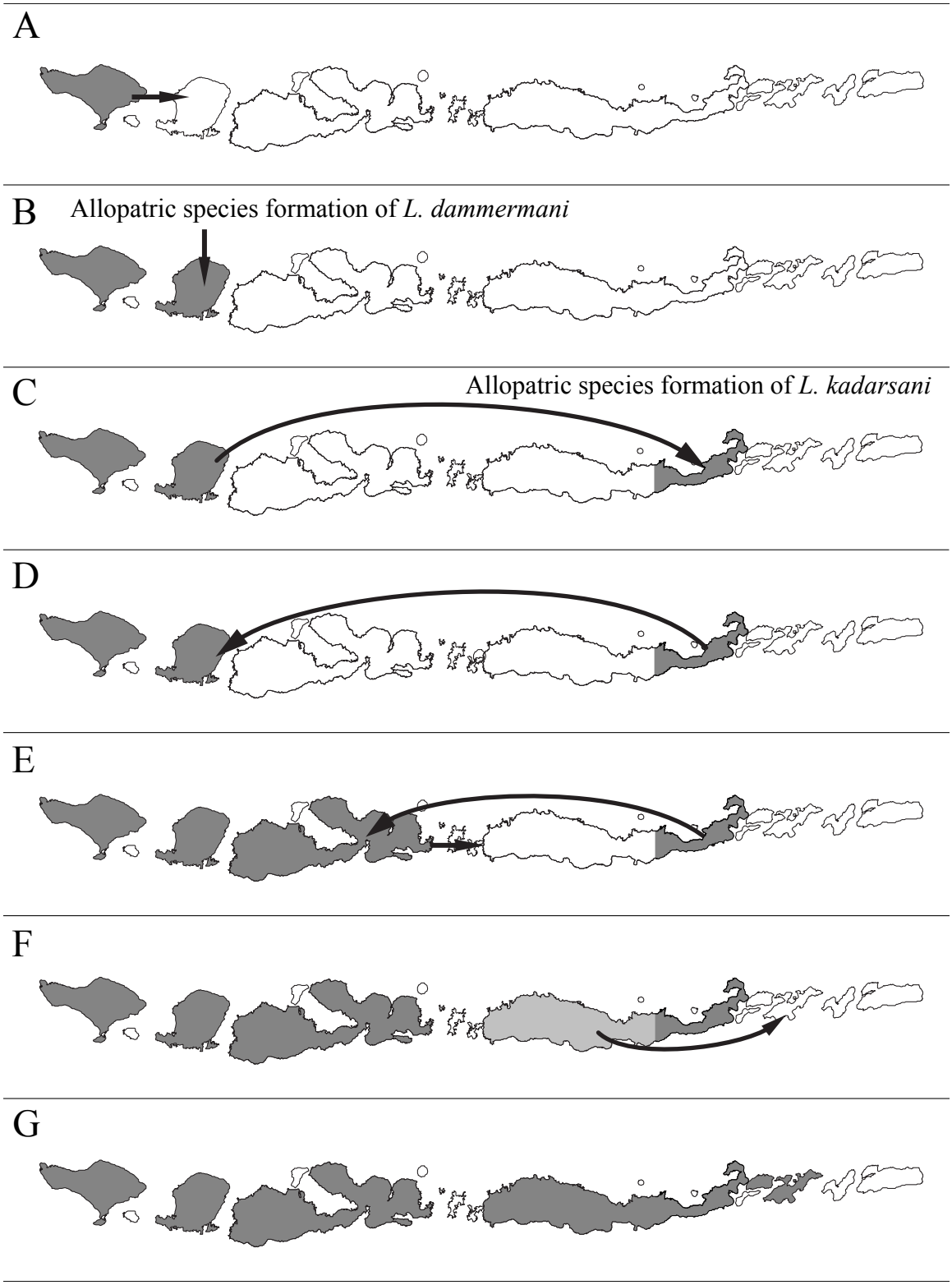

A
B
Allopatric species formation of *L. dammermani*
C
Allopatric species formation of *L. kadarsani*
D
E
F
G

Figure 16.

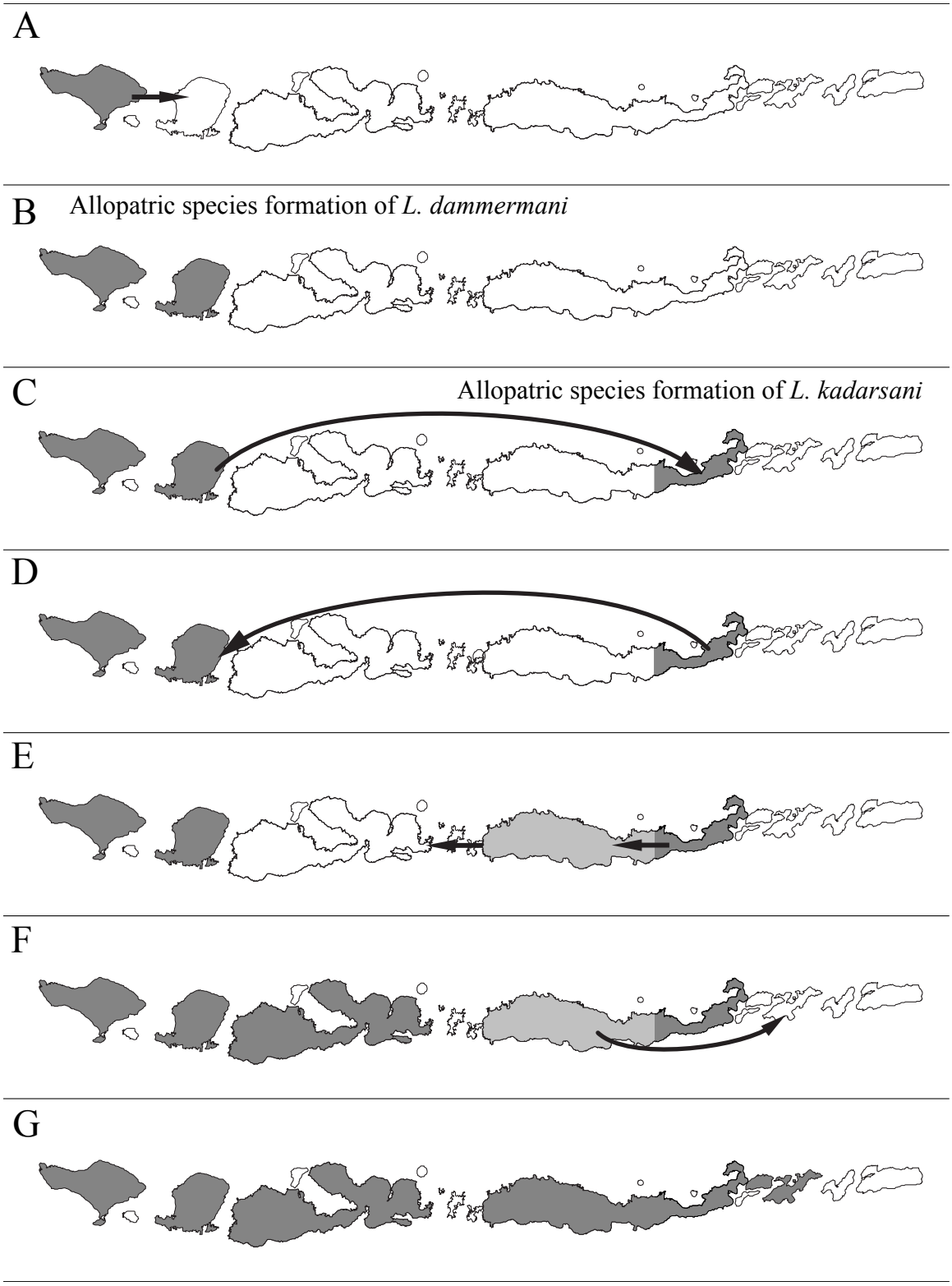

A
B
Allopatric species formation of *L. dammermani*
C
Allopatric species formation of *L. kadarsani*
D
E
F
G

Figure 17.

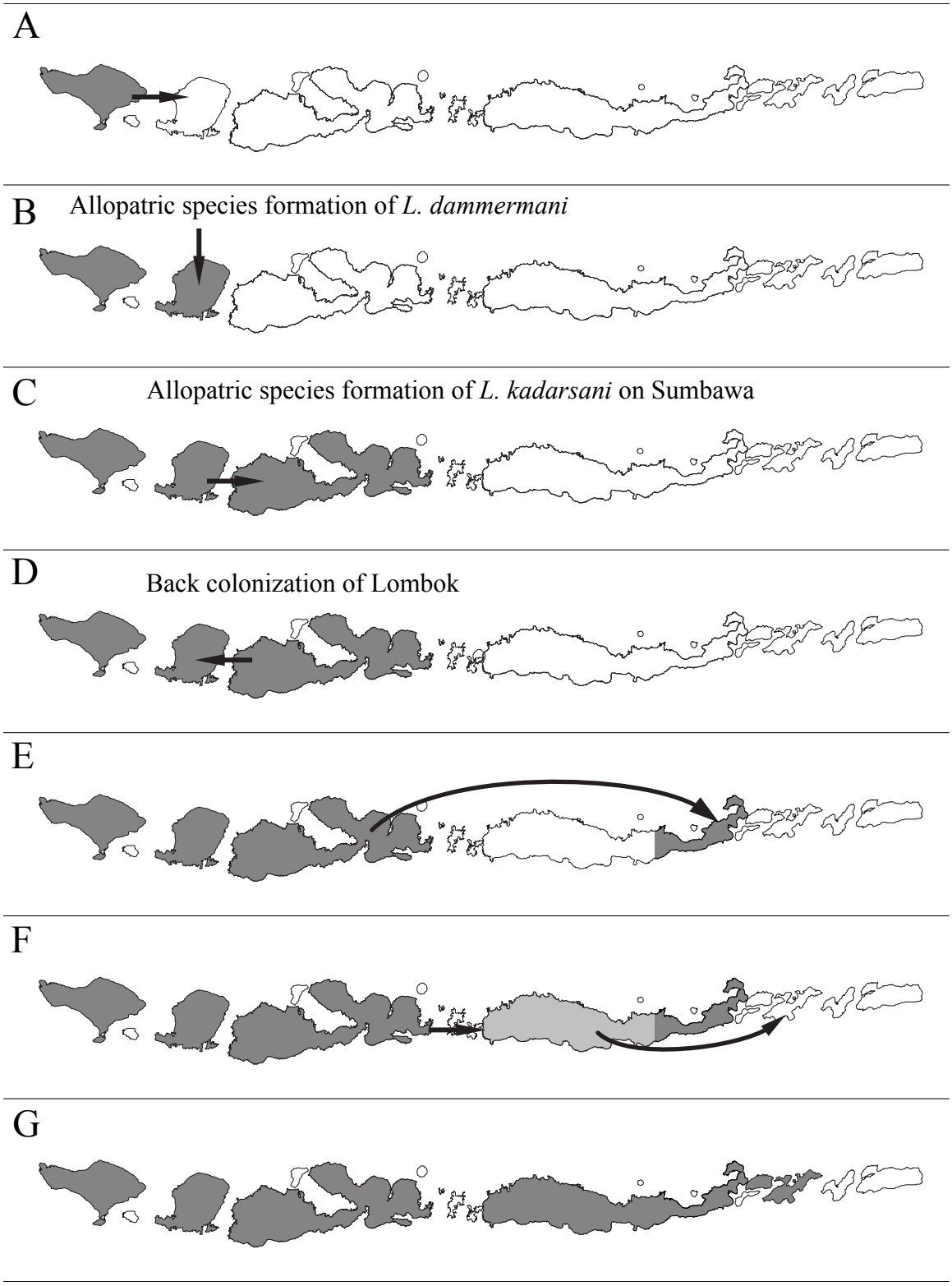

A
B
Allopatric species formation of *L. dammermani*
C
Allopatric species formation of *L. kadarsani* on Sumbawa
D
Back colonization of Lombok
E
F
G

Figure 18.