Genetic studies on prehistoric translocations of chickens in the Indo-Pacific

Michael James Bannister Herrera

Bachelor of Science (Biology)

Bachelor of Science (Archaeology)

Australian Centre for Ancient DNA

School of Biological Sciences

University of Adelaide

A thesis submitted for the degree of Doctor of Philosophy at The University of Adelaide

March 2015

Table of Contents

Contents

Thesis Abstract

The study conducted in this thesis examines the genetic population history of chickens in the Indo-Pacific region in order to infer the prehistoric human-mediated translocation of chickens and investigate whether the dispersal history of chickens in this region parallels the Austronesian expansion. The research focuses on chicken populations found in Island Southeast Asia, Pacific Ocean, and Indian Ocean – regions where Austronesian languages are spoken. The islands and archipelagos found in this region are separated by vast distances of ocean, thus the dispersal of chickens within this region is mediated only through human agency. The geographic distribution of genetic variation in chickens of this region is due only to humans translocating chickens during their voyages, thus this genetic information can be utilised to examine the expansion of the Austronesian-speaking people.

A genetic survey that spans two oceans is challenging, thus the study relied mostly on modern chicken DNA and available ancient DNA to reconstruct events that transpired several millennia ago. The use of modern DNA allowed comparison with reference sequences from across the globe, whereas ancient DNA allowed population continuity to be tested – *i.e.*, whether the modern specimen still represents past populations. The phylogeographic and population genetic analyses on these chickens provided unparalleled insights into the prehistoric translocation history of chickens in the Indo-Pacific region. These have allowed us to confirm the Philippine homeland of the Polynesian chickens and find the east African proximate population source for chickens in Madagascar. Furthermore, the study supports that chickens were dispersed into the Pacific along with the Austronesian expansion, but not in the Indian Ocean. The study also revealed original insights and highlights the complex

picture about the population history and human-mediated dispersals of chickens in the Indo-Pacific. This complexity is brought by the fact that the prehistoric translocation of chickens cannot be solely attributed to one dominant human group or expansion event that occurred in the region. Therefore, it is paramount to use archaeological and linguistic narratives to explain the genetics of chickens and reach the best inference possible about their history.

This research demonstrates the usefulness of using genetic studies on chickens in elucidating the origins and routes of prehistoric translocations and Austronesian expansion in the Indo-Pacific. This study advances our knowledge about prehistoric dispersal of chickens in the Indo-Pacific region and will precipitate exciting new avenues of research.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work, in the future, be used in a submission of any other degree or diploma in any university or other tertiary institution without prior approval of the University of Adelaide and where applicable, any partner institution responsible for the join-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provision of the Copyright Act 1968.

I also give permission for the digital version of this thesis to be made available on the web, via the University's digital repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

_____

Michael James Bannister Herrera

March 2015

Acknowledgements

Learning to become a proper researcher was indeed fraught with challenges. What is normally unbearable has become bearable because of the so many kind people who believed, supported, and encouraged me. I am forever grateful to have had the chance to work with all of you. I want to express my sincerest gratitude to them who have been a part in the creation of this work.

DR. JEREMY AUSTIN, my PhD adviser, for his unwavering guidance and patience, unfailing support, and constant stirring during the conduct of this research. As my mentor, he was an invaluable source of inspiration. His guidance was catalytic in the fruition of this work. Thank you for believing in me, Jeremy.

DR. VICKI THONSON, my co-supervisor, for all her able counsel and support during the conduct of this research. She patiently answered all of my questions as well as asked questions back to help me to better my work. She afforded her valuable time and guided me throughout my research. For which, I am truly thankful.

DR. ALAN COOPER, my co-supervisor, for all the guidance and suggestions during the planning stage of my research. He has always been a source of inspiration.

DR. JESSICA WADLEY, my mentor, for all the help and guidance that she afforded me. She patiently and carefully discussed concepts and ideas with me, which was helpful in the completion of my thesis. You have always been there, for which, I am very thankful.

DR. PHILIP PIPER, my external co-supervisor, thank you so much for supporting me throughout my PhD journey. I am forever indebted to you for your unwavering support and guidance.

DR. JULIEN SOUBRIER, DR. WOLFGANG HAAK, & KEIREN MITCHEL, thank you very much for providing me with technical support and assistance. I would not have been able to make it without you guys.

DR. JAIME GONGORA, thank you very much for providing the valuable Indonesian chicken samples, for they were a critical component of the research. Thank you as well for the guidance and support.

I also express my thanks to all the people who have in part allowed me to accomplish this work; this will include museum people (American Museum of Natural History, National Museum of Natural History) and all the people who helped me during my sampling trip.

I would also like to extend my gratitude to my all friends who helped me put some cheer in my face during trying times.

Finally, I would like to express my deepest thanks to my family. Although very far, they were always there when I needed a shoulder to lean on.

List of Tables

List of Figures

CHAPTER 1: Introduction

Introduction

The human history of Island Southeast Asia (ISEA) and the Pacific region is characterised by two waves of migration into and across the region. The first of these dispersal events occurred during the late Pleistocene (Dunn *et al.* 2005), as part of the first wave of modern human dispersal out of Africa by people who spoke a diverse group of languages spoken today in Papua New Guinea (referred to as Papuan languages). This language group is predominantly found in New Guinea, in the lesser Sunda Chain, the Bismarck Archipelago, and with the Solomon Islands as the easternmost fringe. The eastern most fringe of the Papuan languages demarcates Near Oceania from Remote Oceania (Figure 1-1). A second group, known as the Austronesians, with more Asian ancestry arrived in the region during the Holocene (Bellwood *et al.* 1995).

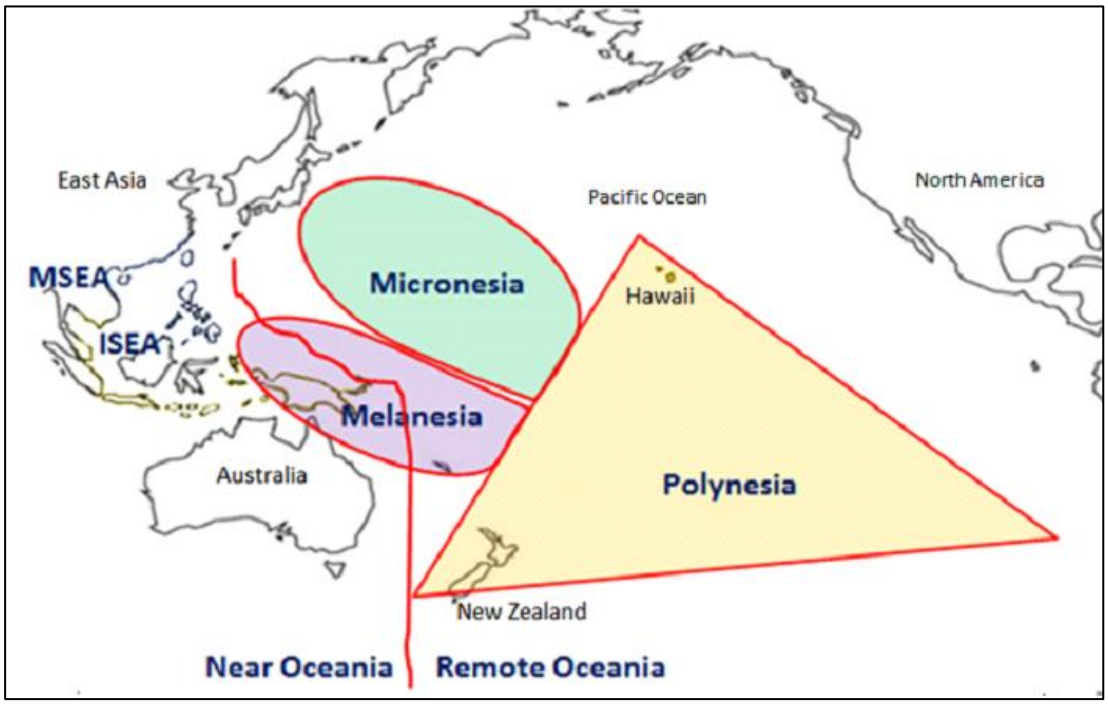


Figure 1-1. Map of Southeast Asia and the Pacific regions showing the demarcation between Near and Remote Oceania, and Micronesia, Melanesia and Polynesia (after Hurles *et al*. 2003)

The Austronesian expansion is unique compared to other prehistoric migrations because it was not made solely over land but mainly across vast distances of open water in the Pacific and the Indian Ocean. This remarkable feat shaped the population history of not only the human groups who made this journey, but also the range of faunal species that they brought with them. These species were predominantly domesticates such as pigs (*Sus scrofa*), dogs (*Canis spp.*), and chickens (*Gallus gallus*). The geographical patterns of genetic diversity in these organisms can provide insights into the history they shared with the people making these transoceanic migrations. Of the transported domesticates, chickens were the most ubiquitous part of the initial colonisation process of the Pacific and thus provide an unparalleled opportunity to test the initial aspects of the prehistoric expansions of the Austronesians.

Background to the Austronesian expansion

The Austronesian expansion is considered one of the most extensive prehistoric human migrations. Our current knowledge of the Austronesian expansion is inferred from the geographic distribution of Austronesian languages spoken today. Belonging to a single language family, Austronesian dialects are spoken from Madagascar in the western Indian Ocean to Easter Island in the eastern Pacific, and as far north as Hawaii (Bellwood 1991). Examination of the distribution of the Austronesian languages has shown that Formosa, otherwise known as modern day Taiwan, exhibits the highest level of diversity (Gray *et al*. 2009) - nine out of ten Austronesian language sub-families are found in Taiwan (Diamond 2000). This high linguistic diversity suggests Taiwan is where the expansion of the Austronesian-speaking people originated. One of these ten sub-families, Malayo-Polynesian (MP),

is thought to have developed in the Philippines after a population expansion out of Taiwan, spreading through Indonesia and continuing into Oceania. This theory is known as the 'Out-of-Taiwan' hypothesis. Thus, all Austronesian languages spoken in the Philippines, Indonesia, the Pacific Islands, and Madagascar belong to the Malayo-Polynesian subfamily. It is thought that a group of the Austronesian-speaking people moved out of Taiwan around 4500 years ago (ya), colonised the Philippines by 4000 ya and then dispersed rapidly throughout the rest of ISEA (Bellwood 1995; Bellwood et al. 2007). At around 3500 ya this expansion reached the Bismarck archipelago and formed the foundational culture in the Pacific called the Lapita Cultural Complex (LCC), whose name comes from a distinctive style of pottery (Spriggs 1995; Kirch 1997). The descendants of this culture later went on to colonise the eastern-most fringes of Remote Oceania.

The colonisation of Oceania by Austronesian speakers is considered as the fastest (Wilmshurst *et al.* 2011) and largest expansion in prehistoric times (Bellwood 1991). The expansion into Remote Oceania only happened after the LCC had time to develop in the Bismarck Archipelago and Solomon Islands in Melanesia (Spriggs 1995; Kirch 1997). The Lapita, equipped with advanced sea-faring skills, went on to reach and occupy most islands of the Pacific, including Hawaii to the north and Easter Island (and possibly South America) in the east. Considered as the ancestors of the Polynesians, the Lapita people are believed to have spoken languages of the Oceanic branch of the Austronesian language family (Gray *et al*. 2009).

The expansion of the Austronesian can also be observed from linguistic studies of populations in the Indian Ocean. Sailing from ISEA, Austronesian-

speakers reached Madagascar by ca. 50-500 A.D. (Dewar & Wright 1993), however the Austronesian language spoken in Madagascar, Malagasy, also contains both Sanskrit and African terms (Adelaar 1995). This presumably resulted from the interactions that Austronesian mariners had with populations found around the Indian Ocean rim. The Austronesian-speaking people appear to have possessed the complex navigational skills and a cultural background that allowed them to successfully find and settle islands in the Pacific, and furthermore, to reach Madagascar.

The genetic landscape of the Austronesian dispersal

Linguistics analysis demonstrates that ISEA, the Pacific and Madagascar have a shared history of Austronesian languages. However, the rapid expansion of the Austronesians is also reflected in the genetics of their descendants still living in the region today. Thus, like linguistic evidence, genetics can also be used to infer past human dispersals and interactions. Within ISEA and the Pacific, several genetic studies have described how the expansion of the Austronesian speakers influenced the genetic landscape in the region (Su *et a*l. 2000; Capelli *et al*. 2001; Friedlaender *et al*. 2008; Soares *et al*. 2011). Key to these investigations is finding the genetic origins and makeup of populations that settled Remote Oceania and how their genetic constitution differ depending on whether they interacted with the original Papuan inhabitants of islands they colonised along the way in ISEA and Near Oceania.

The preferred genetic markers are mitochondrial DNA (mtDNA) and the non-recombining region of the Y chromosome (NRY). These two non-recombining genetic markers have a uniparental pattern of inheritance and they retain records of

genetic diversity over time (Underhill & Kivisild 2007). Furthermore, the mtDNA accumulates mutations at a faster rate, which is important for investigating recent evolutionary histories such as the expansions in the Pacific (Hagelberg *et al*. 2008).

The predominant human mitochondrial lineage in Remote Oceania is characterised by four mutations within the hypervariable segment 1 (HVS1) of the mtDNA control region, called the Polynesian motif (Melton *et al*. 1995). This motif is nested within the variation of haplogroup B4a, which can be found in Near Oceania, Indonesia and Taiwan (Soares *et a*l. 2011). Other mitochondrial lineages found in the Pacific region include haplogroups P and Q but these occur only at low frequencies. Haplogroups P and Q are observed in populations in New Guinea but not in Taiwan and very infrequently elsewhere in ISEA (Kayser *et al*. 2006; Tabbada *et al.* 2010). This suggests that in Remote Oceania both the Papuan and the Austronesian speakers contributed to the current mitochondrial diversity of the region. Interestingly, the mtDNA lineage in Madagascar also falls within the diversity of the Polynesian motif.

The NRY also illustrates the same story as mtDNA for populations in Remote Oceania. Both Papuan speakers in New Guinea and Austronesian speakers contribute to the genetic composition of male populations in Remote Oceania (Kayser *et al.* 2006). The most frequent lineage in Remote Oceania is haplogroup C2, which can be found in the Lesser Sunda Chain (Cox *et al*. 2007). Surprisingly the Y chromosome lineage frequently found in Taiwan (O1a) is not observed in Remote Oceania and is only found infrequently in ISEA and Near Oceania.

Characterising the genetic history of Austronesian-speaking people has profoundly improved our understanding of the underlying prehistoric processes that led them to where they are now. But most of these studies relied heavily on the analyses of non-recombining genetic markers that do not account for historical admixture events (Lipson *et a*l. 2014). A study involving genome-wide nuclear SNPs demonstrated that the most parsimonious explanation for the Austronesian ancestry is one that involves movements from Taiwan to ISEA (Lipson *et a*l. 2014). However, Austronesian populations outside of Taiwan also have varying levels of admixture (Tofanelli *et al*. 2009; Cox *et al*. 2010; Xu *et al.* 2012; Pierron *et al.* 2014). Populations in ISEA, for example, have levels of admixture coming from Negrito, Melanesian, and Austro-Asiatic speakers from mainland Southeast Asia (MSEA) (Tabbada *et al.* 2010).

The advent of agriculture in Island Southeast Asia & the Pacific

The expansion of the Austronesian-speaking people is suggested to have co-occurred with the development and spread of agriculture in ISEA. After the initial development of agriculture in various domestication centres of the world (Figure 1-2), subsequent expansions into adjacent regions occurred (Bellwood 1995; Diamond & Bellwood 2003; Bellwood 2007). This not only caused a fundamental shift in subsistence strategy from hunting and gathering to lifestyles that were more sedentary but it also resulted in the transmission of genes, language, and the knowledge of farming. The development of agriculture is believed to have led to shifts in population demographics not only on the continents but also in ISEA and the Pacific.

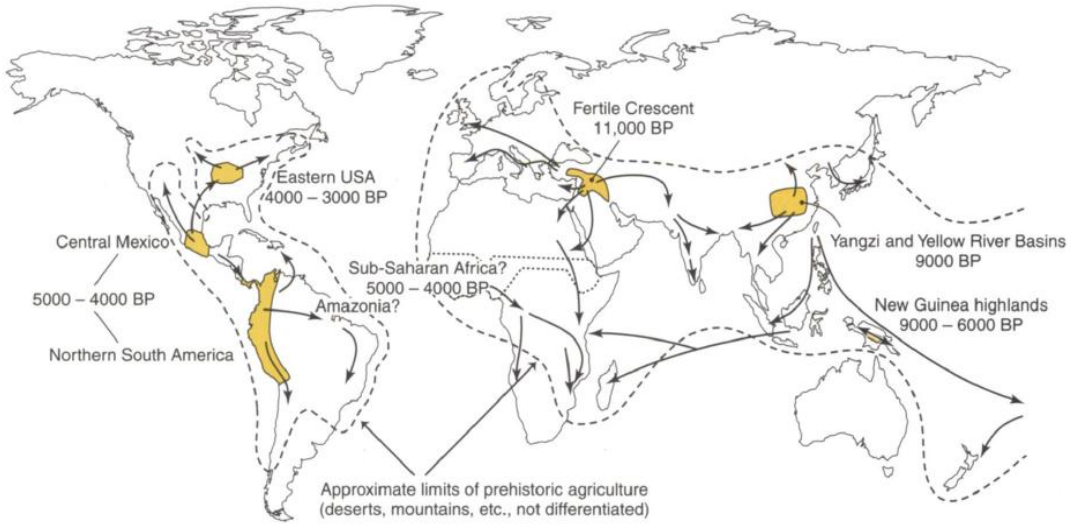


Figure 1-2. Worldwide map showing the regions where agriculture developed (Diamond & Bellwood 2000).

In Asia, two centres for the origin of agriculture have been identified: South China and New Guinea. The development of agriculture in South China led to human population expansions across much of Southeast Asia. In MSEA, the cultures that expanded were the Austroasiatic-speaking populations of northern Vietnam, who proceeded southwards to Laos, south Vietnam, Cambodia, and Thailand (Higham 2004; Higham, Higham 2009). The agricultural expansion from South China is also believed to have reached Taiwan and from there led to the 'Out-of-Taiwan' expansion of the Austronesian-speaking people into ISEA. The development of agriculture in South China is commonly used as a basis for understanding the development of agriculture in Southeast Asia. Agriculture also developed independently in the highlands of New Guinea, which is also believed to be a part of the development of agriculture in ISEA (Donohue & Denham 2010; Denham 2013).

In ISEA, the most common hypothesis suggests that the agricultural revolution in the region originated from Taiwanese practices rather than from those

in MSEA (Bellwood *et al*. 2007). This theory follows the same timeline and direction as the linguistic evidence, and these two lines of evidence are known collectively as the language-farming hypothesis (Diamond & Bellwood 2003). Archaeological evidence suggests that the colonisation of Taiwan around 6000 ya occurred after the development of agriculture in the Yellow River Basin in South China (Kirch 1997; Bellwood 2007; Chi & Hung 2010).

The early and independent rise of agriculture that occurred in the neighbouring region of New Guinea is an important consideration in the development of agricultural practices in ISEA and is unlikely to have occurred in isolation (Denham *et al*. 2003; Denham 2004). In contrast to the cereal-based farming that spread from South China or Taiwan, farming in New Guinea was oriented more on exploiting tuberous plants and trees. Domesticated plants that can trace their origins to ISEA and the New Guinea region include banana (*Musa* spp.), sugarcane (*Saccharum officinarum*), and yam (*Dioscorea* spp.) (Grivet *et al.* 2004; Lebot 2009; Perrier *et al.* 2011). Although, it is clear that ISEA and New Guinea had their own agricultural practices, they probably did not involve the domestication of agricultural animals (Denham 2013).

The animal domesticates in ISEA and the Pacific

Several species of animals and plants were transported by the Austronesian populations as they expanded into ISEA and the Pacific. Archaeological evidence shows that domesticated animal species such as pigs, dogs, chickens, and a range of cereal crops, tubers, and tree fruits were intentionally carried during this dispersal (Bellwood 1997; Bellwood & Hiscock 2005). The translocation of these species is

suggested to be an essential colonisation behaviour that forms part of the adaptive strategies for long distance sea crossings (Anderson & O'Connor 2008; Matisoo-Smith & Robins 2009) and for the re-creation of subsistence economies on newly colonised islands (Kirch 1997). This translocation also included other commensal species such as the Pacific rat (*Rattus exulans*) (Matisoo-Smith 1994) and a lizard species (*Lipinia noctua*) (Austin 1999).

Pigs, dogs, and chickens are intimately linked with the dispersal of the Pacific migrants as they lack the natural ability to disperse over water and were likely deliberately translocated, in effect they share a parallel history with humans (Matisoo-Smith 1994). This forms the central basis of a "commensal model", an approach that essentially uses the association of human-translocated species to infer prehistoric human expansions in much the same way that human genetics trace human origins and dispersals (Matisoo-Smith 1994; Matisoo-Smith *et al.* 2009). Thus, tracing spatial and temporal distribution of these domesticates and examining the genetic variations and diversity of their descendants across their relevant geographic range can provide an excellent proxy to infer the origin and expansion history of Austronesian-speaking people.

Originally sourced from mainland Asia, pigs, dogs, and chickens were differentially incorporated into the agricultural practices of ISEA and may have reached ISEA via different translocation pathways. Unlike the Pacific, where they converge and often co-occur in the archaeozoological record, in ISEA pigs and dogs occur patchily, hinting at multiple translocations by different human groups via various routes at potentially different time periods (Denham 2013). In contrast,

chicken archaeological remains are thus far not recorded within ISEA before 3500 ya (Bellwood 2011). In fact, domestic chickens were not mentioned in one of the comprehensive appraisals of avifauna in Indonesian archaeological sites in Liang Bua (Meijer *et al.* 2013). This makes it difficult to reconstruct the translocation history of chickens in ISEA *en route* to the Pacific. It is only within Oceania, particularly in Lapita sites, where chickens appear and often co-occur with pigs and dogs in archaeological contexts. However, pigs and dogs are generally found only in later sites, in contrast to chickens, which are found throughout all archaeological layers. Thus, chickens are crucial to understanding the initial settlement in the Pacific (Matisoo-Smith 2007).

Genetic data from pigs, dogs, and chickens can be used to trace human expansions and interactions in a way that is complimentary to information provided by linguistics, archaeology, and human genetics. The section below focus on genetic research undertaken on these domesticates, particularly their domestication origins and subsequent translocations, and, by extension, how these have elucidated our current understanding of human expansions and migratory trajectories throughout ISEA and into the Pacific.

The Pig

Pigs have a great economic and ritualistic value in the tribal societies of Southeast Asia (Rosman & Rubel 1989) and appear to have undergone domestication multiple times in MSEA (Larson *et al*. 2010). There are wild, domestic, and feral pigs that are widely distributed throughout Southeast Asia and were exploited by hunter-gatherers and farmers alike. However, the co-occurrence of these forms in

ISEA makes morphological studies on pigs from the region challenging (Dobney *et al*. 2008). One of the lineages likely produced by these multiple domestication processes was a mitochondrial lineage called the Pacific Clade, which is relevant to discussions about human-mediated introduction of pigs into ISEA and the Pacific (Larson *et al.* 2005; Larson *et al*. 2007). The Pacific Clade signature is found in pigs from Vietnam, Sumatra, Java, Wallacea, and Melanesia. The geographic distribution of pigs containing the Pacific Clade was used to infer the route taken by human populations translocating pigs into the Pacific. Pigs in Vanuatu are clearly linked archaeologically to the LCC, with their genetics suggesting MSEA as the possible source population (Lum *et al*. 2006).

The Pacific Clade signature has never been documented in pigs from Taiwan or the Philippines (Larson *et al*. 2007). This raises questions regarding the 'Out-of-Taiwan' expansion. However, one pig lineage that is observed in both Taiwan and the Philippines is the Lanyu pig. These pigs stem from an isolated pig population found on the island of Lanyu located southeast of Taiwan and are distantly related to endemic East Asian pigs. The genetic signature of the Lanyu pigs has been documented in modern pigs from the Philippines, potentially indicating the prehistoric translocation of this haplotype from Taiwan to the Philippines (Herrera 2010). The modern extent of the signature has not been fully investigated but may provide support for the 'Out-of-Taiwan' hypothesis.

Clearly the Pacific Clade and the Lanyu pig mitochondrial lineages are relevant to discussions about the prehistoric translocation of pigs into ISEA and the Pacific but appear to be contradictory. To date, it is clear that the Pacific Clade was

transported into the Pacific from MSEA, and that the Lanyu pigs were translocated from Taiwan to the Philippines. Further to this, there is also evidence that pigs might have undergone an independent domestication process in Wallacea where endemic wild pigs (*Sus celebensis*) are also found (Groves 2007). However, this situation also hints at the complexity of the advent of agriculture in ISEA. How this has influenced the genetic makeup of domestic pigs in the Pacific is still to be investigated. In summary the translocation of pigs into the ISEA and Pacific regions appear to be from multiple source populations indicating both a Taiwanese and MSEA link into ISEA.

The Dog

The close and protracted association of dogs with humans means that they are also an important proxy for inferring prehistoric human expansions. Dogs served as a tool for hunting game but also as transportable food source in themselves (Titcomb & Pukui 1969). The dingo (*Canis lupus dingo*) probably represents the earliest canid translocated into ISEA and Australasia from MSEA beginning 4000 ya (Savolainen *et al*. 2004; Oskarsson *et al*. 2011). Although human occupation of Australia dates back to at least 50,000 ya (Mulvaney & Kamminga 1999), it did not involve the movement of domesticates at that time. It was only much later during the advent of agriculture in Southeast Asia that dogs became entrenched in human history (Gollan 1984). Another species of dog translocated into the region is the New Guinea singing dog (*Canis hallstromi*). Like the dingo, they also trace their ancestry to human movement from MSEA (Smith & Litchfield 2009). Studies on the origins of dog domestication indicate that mainland Asia (*i.e.,* South China) was the putative centre (Pang *et al.* 2009; Ding *et al*. 2012). This could have an important implication for

how dogs were transported into ISEA, the Pacific, and Australia in relation to the 'Out-of-Taiwan' hypothesis.

A study estimating the time to most recent common ancestor (TMRCA) from dog Y-chromosomes that included dingoes from Australia, New Guinea singing dogs, and domestic dogs from several localities in ISEA, found that dogs in the Philippines, Bali, and Brunei are more closely related to dogs in Thailand than Taiwan (Sacks *et al.* 2013). It also highlighted translocation histories for dingoes and New Guinea singing dogs being distinct from, and earlier than, that of village dogs in ISEA, indicating two waves of dog translocations. An earlier analysis using dog mtDNA suggested that the ancestry of female lineage in ISEA is traceable to East Asia, with the expansion of the Austronesian speakers from Taiwan as the suggested mechanism for their translocation into ISEA (Savolainen *et al.* 2004). However, a similar study came to a contradictory conclusion suggesting that the source population for village dogs in ISEA and in Australia was MSEA (Oskarsson *et al*. 2011). In summary, the route of introduction of dogs into ISEA and the Pacific is not so dissimilar to that of pigs and may contain multiple origins and translocations.

The domestication and translocation history of chickens

The domestic chicken played a crucial role to humans in the past, much as they do today. They are significant components of the domestic economies of historic and modern societies where they are used for food, sport, and as sacrifices (Crawford 1990). Other than as a dietary source of protein, the bones of chickens were also used prehistorically to make needles, beads, toys, and whistles (Steadman *et al.* 2002) and their feathers as adornment (Carter 1971). The antiquity of chickens

in Oceania is well documented (Storey *et al.* 2008). Archaeological chicken remains recorded in Vanuatu beginning 3000 ya are associated with the LCC, which potentially dates the initial arrival of chickens in Near Oceania (Bedford *et al*. 2006; Storey *et al.* 2008). It has also been proposed that chickens, along with rats, were the only commensals that can be demonstrated to have dispersed in the earliest migrations beyond Vanuatu and the first domesticate to arrive in Remote Oceania (Kennett *et al.* 2006; Anderson 2009). This is in stark contrast to the situation in ISEA and MSEA where little chicken evidence is found in the archaeological record (Glover & Bellwood 2004; Bellwood 2011). Thus, the nature of the arrival of chickens in ISEA is unclear, but their arrival in this region must have been prior to the earliest recorded archaeological chickens remains in the Pacific.

There are many factors influencing the lack of chicken bones in ISEA, which may include attrition by rats and dogs, butchering practices, tool-making, and incomplete identification of bones (Matisoo-Smith & Robins 2004; Storey *et al.* 2008). These factors are in part the reason why almost nothing is known about the population history of chickens in ISEA, specifically the timing and route of their dispersal. Similar to other domesticates, the chicken is of popular interest not only for its economic utility but also for its potential in revealing certain aspects of human movements in the past. However, in comparison to pigs and dogs, chickens are highly transportable and reproduce fast and in high numbers, thus they might have been preferred by ancient mariners over dogs and pigs as a food source during their initial movement into the Pacific.

The domestic chicken belongs to the genus *Gallus*; the four wild *Gallus* species are all native to South and Southeast Asia. Consequently, the wild ancestors of modern domestic chickens must be derived from these regions. The four living *Gallus* species, *Gallus gallus* (Red Jungle Fowl), *Gallus sonneratii* (grey jungle fowl), *Gallus varius* (Green jungle fowl), and *Gallus lafayetii* (Ceylon jungle fowl), vary in morphology and geographic distribution (Delacour 1977; Crawford 1990; Tixier-Boichard *et al.* 2011). The grey jungle fowl is found in the Indian subcontinent while the Ceylon jungle fowl is found only in Sri Lanka. The green jungle fowl is endemic only to the island of Java, and they are morphologically considered as most distant to domestic chickens. The red jungle fowl (RJF) are genetically closest to domestic chickens and are found in South and Southeast Asia (Liu *et al*. 2006). Five subspecies of *Gallus gallus* (RJF) have been defined, which correspond to their geographic location (Johnsgard 1999). The *G. gallus gallus* are found in southern Vietnam, Cambodia, Thailand, and Laos. The *G. gallus spadiceus* are those found in Myanmar, Thailand, Malaysia and southern China. The *G. gallus murghi* are found in India. The *G. gallus jabouillei* are endemic to southern China and northern Vietnam. The RJF subspecies in Indonesia is the *G. gallus bankiva*.

Previous studies indicate that the domestication process for chickens mainly involved the RJF (*Gallus gallus*) (Fumihito *et al.* 1994; Fumihito *et al.* 1996). However, it has more recently been demonstrated that the RJF might not be the sole progenitor of the domestic chicken, as a gene from the grey jungle fowl (*Gallus sonneratii*) coding for yellow skin colour is present in certain domesticated chickens (Eriksson *et al*. 2008). However, our understanding of this apparent hybridisation event is limited. It is unclear whether humans actively mediated the hybridisation of

the two species during domestication or if RJF naturally hybridised with grey jungle fowl in areas where their distributions overlap. Furthermore, it has also been suggested that all *Gallus* species experienced some level of hydidisation except for *Gallus varius* on Java (Nishibori *et al*. 2005).

Several regions in Asia have been identified as putative centres for chicken domestication based on archaeological records. Chicken remains dating to at least 4500-4000 ya have been found in the archaeological ruins at Mohenjo-Daro in the Indus Valley, with morphological characterisation of these chickens suggesting that the domestication process must have started around this period (Zeuner 1963; Crawford 1990). Domestic chicken remains have also been recorded from archaeological sites in China (West & Zhou 1989). The northern Chinese chickens were considered domestic based on morphology and on the fact that north China is not part of the natural range of RJF. Whether they were domesticated *in situ* is unclear. However, an early Holocene (~10,000 ya) domestication of chickens in northern China has recently been reported (Xiang *et al.* 2014), which represents the earliest date reported for chicken domestication. The domestication of the chicken seemingly occurred independently in multiple Asian centres and most likely on various occasions during the Holocene (Liu *et al*. 2006). Based on genetic studies these domestication centres most likely included parts of India (Kanginakudru *et al.* 2008), Thailand (Fumihito *et al*. 1996), and the southern China region (Liu *et al.* 2006) – areas where *Gallus gallus* are found. A multiple origin for the domestic chicken has been corroborated by mtDNA studies where thirteen highly divergent mtDNA clades (haplogroup A-I and W-Z) have been identified (Liu *et al*. 2006; Miao *et al.* 2012). Most of these mitochondrial clades contain both wild RJF and

domestic chickens supporting the multiple South and Southeast Asian origins for domestic chickens.

In the Philippines indigenous chickens abound, but it is not known when they initially appeared and whether they arrived as wild or domestic fowls (Mudar 1997). Indigenous chickens in the Philippines could represent descendant populations of chickens domesticated in MSEA that went feral after their initial introduction into the archipelago. Alternatively, they could be descendants of wild endemic populations that potentially arrived in ISEA during the Pleistocene when lowered sea-levels formed narrow sea gaps and parts of the Philippine archipelago were connected to mainland Asia through the Sunda shelf. This vast dry-land extension of the Asian continent is formed by the Malay Peninsula, Sumatra, Java, and Borneo (Heaney 1985). This likely connected the Philippine island of Palawan to Sunda, but this probably had little effect on other Philippine islands (Robles *et al.* 2014). The exposed land could have potentially been part of the range of ancestral RJF. The presence of the green jungle fowl on the Indonesian island of Java suggests that ancestral jungle fowls must have colonised a larger portion of ISEA during the Pleistocene, and this may have included the Philippines.

Additional evidence also supports an indigenous Philippine jungle fowl with chickens in the Philippines being deeply integrated into the indigenous culture suggesting antiquity (Jocano 1975). Linguistic evidence also suggests that the chicken was probably utilised by the speakers of the proto-MP , a subgroup of the Austronesian language family that developed in the Philippines (Blust 2002). The term for chicken is absent in the Proto-Austronesian language (Formosan) spoken in

Taiwan, suggesting the word developed after this language group arrived in the Philippines from Taiwan.

The majority of published chicken genetic studies undertaken in Indonesia mainly focus on describing the diversity and population structure of indigenous domestic chickens found in the archipelago (Sulandari *et al*. 2008; Katano *et al.* 2011). Other studies focus more on the improvement of production performance and conservation of indigenous domestic chickens (Sartika *et al*. 2005; Nataamijaya 2014). However, no studies have previously attempted to reconstruct the population history of how indigenous chickens got into Indonesia.

Tracking the translocation of chickens using genetics

The variations contained in DNA can be used to track population origins and dispersal. By looking into these variations, researchers have been able to identify population specific genetic markers and depict the evolutionary relationships between individuals containing them (Matisoo-Smith 2008). By examining how variations in DNA occur geographically and temporally, an approach can be used to infer origins, dispersal trajectories, and interactions of populations in the past. To track down prehistoric expansions, scientists usually use haplotypes, which is a genetic unit that contains variations transferred through generations without recombination. The entire mitochondrial DNA molecule and a portion of the Y chromosome are transmitted through the generations as haplotypes.

However, the use of non-recombining genetic markers to infer histories should be used with caution, particularly when using modern DNA. The underlying

assumption of using modern DNA to reconstruct prehistoric dispersal is that modern samples demonstrate continuity with the past populations under investigation. However, the genetic signature of past populations in modern samples can be overwritten by population replacement and obscured by introgression and admixture. Thus, we are mostly unaware of the extent to which modern genetic variation is representative of past genetic diversity (Zeder *et al*. 2006). For domesticates, this can be further confounded by modern day trade and exchange (Storey *et al*. 2012). Thus, inferring histories using modern genetic patterns should be done in tandem with other sources of information such as geography, archaeology, linguistics, or ecology. Ancient DNA (aDNA) and genome level information, if available, should also be used. The genetic information contained within properly provenanced archaeological material can give a window into past genetic histories. Genome level analyses on the other hand can allow the quantification and tracking of hybridisation events. Genome-scale analyses is becoming increasingly important as we realise how modern domestic populations are in fact the result of admixture between widely dispersed populations (Larson & Fuller 2014). Prehistoric populations of domesticates are potentially separated from their modern counterparts by thousands of years of selective breeding, introgression, and hybridisation. Thus, using additional genetic information such as those from genomic regions or genome-wide single nucleotide polymorphisms (SNPs) can potentially circumvent these challenges.

Although the techniques used in aDNA studies are developing rapidly and being constantly optimised (Shapiro & Hofreiter 2014), retrieving DNA from ancient samples is nonetheless still limited by the preservation of the sample. In the Pacific,

chicken skeletal remains have been used in aDNA studies to examine the prehistoric human settlement and interactions (Storey *et al*. 2007; Gongora *et al.* 2008; Storey *et al*. 2010; Thomson *et al*. 2014). However, given that the modern chicken populations in ISEA (Indonesia and the Philippines) have never been extensively investigated, it is paramount to investigate the genetic diversity and historical demographic trends of chickens in ISEA before they can be used to assess their utility for illuminating aspects of prehistoric human migrations in the Pacific.

Aims of the thesis

The research conducted in this thesis is focused on the molecular characterisation of chickens from ISEA and the Pacific to understand their population history and reconstruct the prehistoric expansions of the Austronesian-speaking people. To date, there have been no detailed studies on chickens from ISEA that examine prehistoric human expansions *en route* to the Pacific and also the Indian Ocean. As ISEA is a geographically strategic point in the settlement process of the Pacific and Indian Ocean (as attested by linguistics, archaeology, and human genetics), the study presented here aims to expand on the lack of previous sampling by focusing on indigenous chicken populations from ISEA. In order to overcome the limits of modern DNA in reconstructing prehistoric events, museum and archaeological samples were also included in the study. Genome-wide information was used to further explore the population history of chickens in the region. The overall goal of this project was therefore to reconstruct aspects of chicken translocations in the Indo-Pacific region. In order to achieve this, the research aims were as follows:

1. Create a mitochondrial DNA dataset from indigenous chickens across Island Southeast Asia, characterise their genetic diversity, structure, and relatedness with those in the Pacific and the Indian Ocean rim;
2. Assess the utility of chicken genetics as a biological proxy for inferring human migrations and interactions in the Indo-Pacific region in light of current archaeological and linguistic knowledge; and

3. Evaluate the resolution of human-mediated dispersal histories of chickens inferred from 1) mitochondrial DNA control region, 2) whole mitochondrial genomes, and 3) genome-wide nuclear single nucleotide polymorphisms.

Thesis outline

The thesis has six chapters each describing my research using chickens as a biological proxy for reconstructing human expansions in ISEA, the Pacific, and the Indian Ocean. Each chapter starts by providing a review of the literature to give context, then describes the experiments and analyses used to test hypothesis being examined, presents the results, and discusses these within the broader context.

Chapter 2 assesses the genetic variation within and between chicken populations in ISEA and the Pacific to examine aspects of prehistoric movement and interactions in these regions. The inferences were mainly drawn from a 201 base pair fragment of the mtDNA control region. The chicken mtDNA control region represents the majority of the variation contained in the entire mitochondrial genome, and more importantly the small fragment size allowed direct comparison with the worldwide reference dataset including sequences from archaeological chickens in the Pacific. Here, it was demonstrated that chickens are suitable for investigating aspects of human expansions in the Pacific and that the translocation of domesticates into the region is far more complex than previously envisaged. This chapter has been presented at two international conferences: Indo-Pacific Prehistory Association (IPPA) in Cambodia and International Council for Archaeozoology (ICAZ) in Argentina. This chapter is also in preparation to be submitted to the journal Molecular Biology and Evolution (MBE).

Chapter 3 explores the possibility of chickens having been translocated by Austronesian-speakers from ISEA to Madagascar across the Indian Ocean. The Indian Ocean has been used as a corridor for the translocation of several domesticates including chickens from various geographical points around the Indian Ocean rim from ISEA, South Asia, the Arabian Peninsula, and the east coast of Africa. This chapter suggests an ultimate South Asian origin for Madagascan chicken.

Chapter 4 examines the dispersal history of chickens from Southeast Asia and the Pacific using whole mitochondrial genomes of modern, museum, and archaeological chickens. This study utilises capture-based enrichment and next-generation sequencing techniques to access the whole mitochondrial genome. The additional nucleotide variation outside the mitochondrial control region provides added resolution for inferring the genetic affinities of chickens in the region. The improved resolution provided by sequencing the whole mitochondrial genome was used to investigate the phylogenetic relationship of the most ubiquitous mitochondrial haplogroup (haplogroup D) in ISEA and the Pacific. Scenarios about the origins and dispersal of this chicken lineage are examined.

Chapter 5 uses genome-wide nuclear SNP data to examine the genetic affinities of chickens from mainland Asia, ISEA, and the Pacific. This is an exploratory chapter, as genome-wide SNP data has never been used to look into the population history of chickens before. This can potentially overcome any challenges involved in using non-recombining single molecules, such as the mtDNA to reconstruct the complex processes involved in the domestication of chickens.

Chapter 6 is a general discussion that reviews the significant findings from each data chapter. It summarizes and highlights the contributions that using chicken genetics as a biological proxy can give to tracking human migrations and making narratives about the expansion of the Austronesians in ISEA and the Pacific.

References

Adelaar, KA. 1995. Borneo as a cross-roads for comparative Austronesian linguistics. The Austronesians:81-102.

Adelaar, KA. 1995. Borneo as a cross-roads for comparative Austronesian linguistics. The Austronesians:81-102.

Anderson, A. 2009. The rat and the octopus: initial human colonization and the prehistoric introduction of domestic animals to Remote Oceania. Biological Invasions 11:1503-1519.

Anderson, A, S O'Connor. 2008. Indo-Pacific Migration and Colonization-Introduction. Asian Perspective 47:2-11.

Austin, CC. 1999. Lizards took express train to Polynesia. Nature 397:113-114.

Bedford, S, M Spriggs, R Regenvanu. 2006. The Teouma Lapita site and the early human settlement of the Pacific Islands. Antiquity 80:812-828.

Bellwood, P. 1991. The Austronesian dispersal and the origin of languages. Scientific American 265:88-93.

Bellwood, P. 1995. Austronesian prehistory in Southeast Asia: homeland, expansion and transformation. The Austronesians:96-111.

Bellwood, P. 2007. Prehistory of the Indo-Malaysian Archipelago. Canberra, Australia: Australian National University Press.

Bellwood, P. 2011. Holocene population history in the Pacific region as a model for worldwide food producer dispersals. Current Anthropology 52:S363-S378.

Bellwood, P, C Gamble, SA Le Blanc, M Pluciennik, M Richards, JE Terrell. 2007. First Farmers: the Origins of Agricultural Societies, by Peter Bellwood. Malden (MA): Blackwell, 2005; ISBN 0-631-20565-9 hardback£ 60; ISBN 0-631-20566-7 paperback£ 17.99, xix+ 360 pp., 59 figs., 3 tables. Cambridge archaeological journal 17:87-109.

Bellwood, P, P Hiscock. 2005. Australia and the Austronesians. In: C Scarre, editor. The Human Past. London. p. 264-305.

Bellwood, PS. 1997. Prehistory of the Indo-Malaysian archipelago. Hawaii: University of Hawaii.

Bellwood, PS, JJ Fox, DT Tryon. 1995. The Austronesians: historical and comparative perspectives: Australian National University.

Blust, R. 2002. The History of Fanual Terms in Austronesian Languages. Oceanic linguistics:89-139.

Capelli, C, JF Wilson, M Richards, MP Stumpf, F Gratrix, S Oppenheimer, P Underhill, VL Pascali, T-M Ko, DB Goldstein. 2001. A predominantly indigenous paternal heritage for the Austronesian-speaking peoples of insular Southeast Asia and Oceania. The American Journal of Human Genetics 68:432-443.

Carter, GF. 1971. Pre-Columbian chickens in America. In: CL Riley, JC Kelley, CW Pennington, RL Rands, editors. Man across the Sea: Problems of Pre-Columbian Contacts. Austin, TX: University of Texas Press. p. 178-218.

Chi, Z, H-c Hung. 2010. The emergence of agriculture in southern China. Antiquity 84:11-25.

Cox, MP, TM Karafet, JS Lansing, H Sudoyo, MF Hammer. 2010. Autosomal and X-linked single nucleotide polymorphisms reveal a steep Asian–Melanesian ancestry cline in eastern Indonesia and a sex bias in admixture rates. Proceedings of the Royal Society of London B: Biological Sciences 277:1589-1596.

Cox, MP, AJ Redd, TM Karafet, CA Ponder, JS Lansing, H Sudoyo, MF Hammer. 2007. A Polynesian motif on the Y chromosome: Population structure in remote Oceania. Human biology 79:525-535.

Crawford, RD. 1990. Poultry breeding and genetics: Elsevier.

Delacour, J. 1977. Pheasants of the World.

Denham, T. 2004. The roots of agriculture and arboriculture in New Guinea: looking beyond Austronesian expansion, Neolithic packages and indigenous origins. World Archaeology 36:610-620.

Denham, T. 2013. Early farming in Island Southeast Asia: an alternative hypothesis. Antiquity 87:250-257.

Denham, TP, SG Haberle, C Lentfer, R Fullagar, J Field, M Therin, N Porch, B Winsborough. 2003. Origins of agriculture at Kuk Swamp in the highlands of New Guinea. Science 301:189-193.

Dewar, RE, HT Wright. 1993. The culture history of Madagascar. Journal of World Prehistory 7:417-466.

Diamond, J, P Bellwood. 2003. Farmers and their languages: the first expansions. Science 300:597-603.

Diamond, JM. 2000. Linguistics: Taiwan's gift to the world. Nature 403:709-710.

Ding, Z, M Oskarsson, A Ardalan, H Angleby, L-G Dahlgren, C Tepeli, E Kirkness, P Savolainen, Y Zhang. 2012. Origins of domestic dog in Southern East Asia is supported by analysis of Y-chromosome DNA. Heredity 108:507-514.

Dobney, K, T Cucchi, G Larson. 2008. The pigs of Island Southeast Asia and the Pacific: new evidence for taxonomic status and human-mediated dispersal. Asian Perspectives 47:59-74.

Donohue, M, T Denham. 2010. Farming and language in island southeast Asia. Current Anthropology 51:223-256.

Dunn, M, A Terrill, G Reesink, RA Foley, SC Levinson. 2005. Structural phylogenetics and the reconstruction of ancient language history. Science 309:2072-2075.

Eriksson, J, G Larson, U Gunnarsson, B Bed'hom, M Tixier-Boichard, L Strömstedt, D Wright, A Jungerius, A Vereijken, E Randi. 2008. Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. PLoS genetics 4:e1000010.

Friedlaender, JS, FR Friedlaender, FA Reed, KK Kidd, JR Kidd, GK Chambers, RA Lea, J-H Loo, G Koki, JA Hodgson. 2008. The genetic structure of Pacific Islanders. PLoS genetics 4:e19.

Fumihito, A, T Miyake, S Sumi, M Takada, S Ohno, N Kondo. 1994. One subspecies of the red junglefowl (Gallus gallus gallus) suffices as the matriarchic ancestor of all domestic breeds. Proceedings of the National Academy of Sciences 91:12505-12509.

Fumihito, A, T Miyake, M Takada, R Shingu, T Endo, T Gojobori, N Kondo, S Ohno. 1996. Monophyletic origin and unique dispersal patterns of domestic fowls. Proceedings of the National Academy of Sciences 93:6792-6795.

Glover, I, P Bellwood. 2004. Southeast Asia: from prehistory to history. New York: Routledge Curzon.

Gollan, K. 1984. The Australian dingo: in the shadow of man. Vertebrate zoogeography and evolution in Australasia:921-927.

Gongora, J, NJ Rawlence, VA Mobegi, H Jianlin, JA Alcalde, JT Matus, O Hanotte, C Moran, JJ Austin, S Ulm. 2008. Indo-European and Asian origins for Chilean and Pacific chickens revealed by mtDNA. Proceedings of the National Academy of Sciences 105:10308-10313.

Gray, RD, AJ Drummond, SJ Greenhill. 2009. Language phylogenies reveal expansion pulses and pauses in Pacific settlement. Science 323:479-483.

Grivet, L, C Daniels, J-C Glaszmann, A D'Hont. 2004. A review of recent molecular genetics evidence for sugarcane evolution and domestication.

Groves, C. 2007. Current views on taxonomy and zoogeography of the genus Sus. Pigs and humans 10:15-29.

Hagelberg, E, M Cox, W Schiefenhövel, I Frame. 2008. A genetic perspective on the origins and dispersal of the Austronesians. Past Human Migrations in East Asia: Matching Archaeology, Linguistics and Genetics:356-375.

Heaney, L. 1985. Zoogeographic evidence for middle and late Pleistocene land bridges to the Philippine Islands. Modern Quaternary Research in Southeast Asia 9:127-144.

Herrera, M. 2010. *An archaeogenetic study of the genus Sus in the Philippines and its propositions for understanding human translocation in Island Southeast Asia*, Unpublished Master's Thesis, University of the Philippines.

Higham, C, T Higham. 2009. A new chronological framework for prehistoric Southeast Asia, based on a Bayesian model from Ban Non Wat. Antiquity 83:125-144.

Higham, CF. 2004. Mainland Southeast Asia from the Neolithic to the Iron age. Southeast Asia: from prehistory to history:41-67.

Hurles, ME, E Matisoo-Smith, RD Gray, D Penny. 2003. Untangling Oceanic settlement: the edge of the knowable. Trends in Ecology & Evolution 18:531-540.

Jocano, FL. 1975. The Philippines at the Spanish contact.

Johnsgard, PA. 1999. Pheasants of the World: Smithsonian Institution Press.

Kanginakudru, S, M Metta, R Jakati, J Nagaraju. 2008. Genetic evidence from Indian red jungle fowl corroborates multiple domestication of modern day chicken. BMC evolutionary biology 8:174.

Katano, T, T Shimogiri, K Kawabe, S Okamoto. 2011. Genetic diversity and population structure of Indonesian native chickens based on single nucleotide polymorphism markers. Poultry science 90:2471-2478.

Kayser, M, S Brauer, R Cordaux, A Casto, O Lao, LA Zhivotovsky, C Moyse-Faurie, RB Rutledge, W Schiefenhoevel, D Gil. 2006. Melanesian and Asian origins of Polynesians: mtDNA and Y chromosome gradients across the Pacific. Molecular biology and evolution 23:2234-2244.

Kennett, DJ, A Anderson, B Winterhalder. 2006. The ideal free distribution, food production, and the colonization of Oceania. Behavioral ecology and the transition to agriculture:265-288.

Kirch, PV. 1997. The Lapita peoples : ancestors of the oceanic world. Cambridge, Mass.: Blackwell Publishers.

Larson, G, T Cucchi, M Fujita, E Matisoo-Smith, J Robins, A Anderson, B Rolett, M Spriggs, G Dolman, T-H Kim. 2007. Phylogeny and ancient DNA of Sus provides insights into neolithic expansion in Island Southeast Asia and Oceania. Proceedings of the National Academy of Sciences 104:4834-4839.

Larson, G, K Dobney, U Albarella, M Fang, E Matisoo-Smith, J Robins, S Lowden, H Finlayson, T Brand, E Willerslev. 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. Science 307:1618-1621.

Larson, G, DQ Fuller. 2014. The Evolution of Animal Domestication. Annual Review of Ecology, Evolution, and Systematics 45:115-136.

Larson, G, R Liu, X Zhao, J Yuan, D Fuller, L Barton, K Dobney, Q Fan, Z Gu, X-H Liu. 2010. Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. Proceedings of the National Academy of Sciences 107:7686-7691.

Lebot, V. 2009. Tropical root and tuber crops: cassava, sweet potato, yams and aroids: Cabi.

Lipson, M, P-R Loh, N Patterson, P Moorjani, Y-C Ko, M Stoneking, B Berger, D Reich. 2014. Reconstructing Austronesian population history in Island Southeast Asia. bioRxiv.

Liu, Y-P, G-S Wu, Y-G Yao, Y-W Miao, G Luikart, M Baig, A Beja-Pereira, Z-L Ding, MG Palanichamy, Y-P Zhang. 2006. Multiple maternal origins of chickens: out of the Asian jungles. Molecular phylogenetics and evolution 38:12-19.

Lum, JK, JK McIntyre, DL Greger, KW Huffman, MG Vilar. 2006. Recent Southeast Asian domestication and Lapita dispersal of sacred male pseudohermaphroditic “tuskers” and hairless pigs of Vanuatu. Proceedings of the National Academy of Sciences 103:17190-17195.

Matisoo-Smith, E. 1994. The Human Colonisation of Polynesia. A Novel Approach: Genetic Analyses of the Polynesian Rat (*Rattus exulans*). The Journal of the Polynesian Society 103:75-87.

Matisoo-Smith, E. 2007. Animal translocations, genetic variation and the human settlement of the Pacific. In: JS Friedlaender, editor. Genes, language, and culture history in the southwest Pacific. Oxford, UK: Oxford University Press.

Matisoo-Smith, E. 2008. Using DNA in Landscape Archaeology. In: JT Bruno David, editor. Handbook of Landscape Archaeology. California: Left Coast Press, Inc. p. 521-529.

Matisoo-Smith, E, M Hingston, G Summerhayes, J Robins, HA Ross, M Hendy. 2009. On the Rat Trail in Near Oceania: Applying the Commensal Model to the Question of the Lapita Colonization. Pacific Science 63:465-475.

Matisoo-Smith, E, J Robins. 2009. Mitochondrial DNA evidence for the spread of Pacific rats through Oceania. Biol Invasions 11:1521-1527.

Matisoo-Smith, E, JH Robins. 2004. Origins and dispersals of Pacific peoples: Evidence from mtDNA phylogenies of the Pacific rat. PNAS 101:9167-9172.

Meijer, HJ, T Sutikna, EW Saptomo, RD Awe, Jatmiko, S Wasisto, HF James, MJ Morwood, MW Tocheri. 2013. Late Pleistocene-Holocene non-Passerine Avifauna of Liang Bua (Flores, Indonesia). Journal of Vertebrate Paleontology 33:877-894.

Melton, T, R Peterson, AJ Redd, N Saha, A Sofro, J Martinson, M Stoneking. 1995. Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. American journal of human genetics 57:403.

Miao, Y, M Peng, G Wu, Y Ouyang, Z Yang, N Yu, J Liang, G Pianchou, A Beja-Pereira, B Mitra. 2012. Chicken domestication: an updated perspective based on mitochondrial genomes. Heredity 110:277-282.

Mudar, KM. 1997. Patterns of animal utilization in the Holocene of the Philippines: A comparison of faunal samples from four archaeological sites.

Mulvaney, J, J Kamminga. 1999. Prehistory of Australia. Washington, D.C.: Smithsonian Institution Press.

Nataamijaya, A. 2014. Possibility of Conserving and Improving the Productivity of Native Chicken. JITV 19.

Nishibori, M, T Shimogiri, T Hayashi, H Yasue. 2005. Molecular evidence for hybridization of species in the genus Gallus except for Gallus varius. Animal genetics 36:367-375.

Oskarsson, MC, CF Klütsch, U Boonyaprakob, A Wilton, Y Tanabe, P Savolainen. 2011. Mitochondrial DNA data indicate an introduction through Mainland

Southeast Asia for Australian dingoes and Polynesian domestic dogs. Proceedings of the Royal Society B: Biological Sciences:rspb20111395.

Pang, J-F, C Kluetsch, X-J Zou, A-b Zhang, L-Y Luo, H Angleby, A Ardalan, C Ekström, A Sköllermo, J Lundeberg. 2009. mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. Molecular biology and evolution 26:2849-2864.

Perrier, X, E De Langhe, M Donohue, C Lentfer, L Vrydaghs, F Bakry, F Carreel, I Hippolyte, J-P Horry, C Jenny. 2011. Multidisciplinary perspectives on banana (Musa spp.) domestication. Proceedings of the National Academy of Sciences 108:11311-11318.

Pierron, D, H Razafindrazaka, L Pagani, F-X Ricaut, T Antao, M Capredon, C Sambo, C Radimilahy, J-A Rakotoarisoa, RM Blench. 2014. Genome-wide evidence of Austronesian–Bantu admixture and cultural reversion in a hunter-gatherer group of Madagascar. Proceedings of the National Academy of Sciences 111:936-941.

Robles, E, P Piper, J Ochoa, H Lewis, V Paz, W Ronquillo. 2014. Late Quaternary Sea-Level Changes and the Palaeohistory of Palawan Island, Philippines. The Journal of Island and Coastal Archaeology:1-21.

Rosman, A, PG Rubel. 1989. Stalking the wild pig: hunting and horticulture in Papua New Guinea. In: S Kent, editor. Farmers as hunters: The implication of sedentism. Cambridge: Cambridge University Press. p. 27-36.

Sacks, BN, SK Brown, D Stephens, NC Pedersen, J-T Wu, O Berry. 2013. Y chromosome analysis of dingoes and Southeast Asian village dogs suggests a Neolithic continental expansion from Southeast Asia followed by multiple Austronesian dispersals. Molecular biology and evolution 30:1103-1118.

Sartika, T, RR Noor. 2005. Production performance of some local chicken genotypes in Indonesia: An overview.

Savolainen, P, T Leitner, AN Wilton, E Matisoo-Smith, J Lundeberg. 2004. A detailed picture of the origin of the Australian dingo, obtained from the study of mitochondrial DNA. Proceedings of the National Academy of Sciences of the United States of America 101:12387-12390.

Shapiro, B, M Hofreiter. 2014. A Paleogenomic Perspective on Evolution and Gene Function: New Insights from Ancient DNA. Science 343:1236573.

Smith, BP, CA Litchfield. 2009. A review of the relationship between indigenous Australians, dingoes (Canis dingo) and domestic dogs (Canis familiaris). Anthrozoos: A Multidisciplinary Journal of The Interactions of People & Animals 22:111-128.

Soares, P, T Rito, J Trejaut, M Mormina, C Hill, E Tinkler-Hundal, M Braid, DJ Clarke, J-H Loo, N Thomson. 2011. Ancient voyaging and Polynesian origins. The American Journal of Human Genetics 88:239-247.

Spriggs, M. 1995. The Lapita culture and Austronesian prehistory in Oceania. The Austronesians: historical and comparative perspectives:112-133.

Steadman, DW, GK Pregill, DV Burley. 2002. Rapid prehistoric extinction of iguanas and birds in Polynesia. Proceedings of the National Academy of Sciences 99:3673-3677.

Storey, AA, JS Athens, D Bryant, et al. 2012. Investigating the Global Dispersal of Chickens in Prehistory Using Ancient Mitochondrial DNA Signatures. PloS One 7:1-11.

Storey, AA, T Ladefoged, EA Matisoo-Smith. 2008. Counting Your Chickens: Density and Distribution of Chicken Remains in Archaeological Sites in Oceania. International Journal of Osteoarchaeology 18:240-261.

Storey, AA, JM Ramírez, D Quiroz, DV Burley, DJ Addison, R Walter, AJ Anderson, TL Hunt, JS Athens, L Huynen. 2007. Radiocarbon and DNA evidence for a pre-Columbian introduction of Polynesian chickens to Chile. Proceedings of the National Academy of Sciences 104:10335-10339.

Storey, AA, M Spriggs, S Bedford, SC Hawkins, JH Robins, L Huynen, E Matisoo-Smith. 2010. Mitochondrial DNA from 3000-year old chickens at the Teouma site, Vanuatu. Journal of Archaeological Science 37:2459-2468.

Su, B, L Jin, P Underhill, et al. 2000. Polynesian origins: Insights from the Y chromosome. Proceedings of the National Academy of Sciences 97:8225-8228.

Sulandari, S, MSA Zein, T Sartika. 2008. Molecular characterization of indonesian indigenous chickens based on mitochondrial DNA displacement (D)-loop sequences. HAYATI Journal of Biosciences 15:145.

Tabbada, KA, J Trejaut, J-H Loo, Y-M Chen, M Lin, M Mirazón-Lahr, T Kivisild, MCA De Ungria. 2010. Philippine mitochondrial DNA diversity: a populated viaduct between Taiwan and Indonesia? Molecular biology and evolution 27:21-31.

Thomson, VA, O Lebrasseur, JJ Austin, TL Hunt, DA Burney, T Denham, NJ Rawlence, JR Wood, J Gongora, LG Flink. 2014. Using ancient DNA to study the origins and dispersal of ancestral Polynesian chickens across the Pacific. Proceedings of the National Academy of Sciences 111:4826-4831.

Titcomb, M, MK Pukui. 1969. Dog and Man in the Ancient Pacific, with Special Attention to Hawaii. Star-Bulletin Print. Honolulu.

Tixier-Boichard, M, B Bed'hom, X Rognon. 2011. Chicken domestication: From archeology to genomics. Comptes rendus biologies 334:197-204.

Tofanelli, S, S Bertoncini, L Castrì, D Luiselli, F Calafell, G Donati, G Paoli. 2009. On the origins and admixture of Malagasy: new evidence from high-resolution analyses of paternal and maternal lineages. Molecular biology and evolution 26:2109-2124.

Underhill, PA, T Kivisild. 2007. Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. Annu. Rev. Genet. 41:539-564.

West, B, B-X Zhou. 1989. Did chickens go north? New evidence for domestication. World's Poultry Science Journal 45:205-218.

Wilmshurst, J, T Hunt, Lipo C, A Anderson. 2011.High-precision radiocarbon dating shows recent and rapid initial human colonization of East Polynesia. Proceedings of the National Academy of Sciences 108:1815-1820.

Xiang, H, J Gao, B Yu, H Zhou, D Cai, Y Zhang, X Chen, X Wang, M Hofreiter, X Zhao. 2014. Early Holocene chicken domestication in northern China. Proceedings of the National Academy of Sciences 111:17564-17569.

Xu, S, I Pugach, M Stoneking, M Kayser, L Jin. 2012. Genetic dating indicates that the Asian–Papuan admixture through Eastern Indonesia corresponds to the Austronesian expansion. Proceedings of the National Academy of Sciences 109:4574-4579.

Zeder, MA, E Emshwiller, BD Smith, DG Bradley. 2006. Documenting domestication: the intersection of genetics and archaeology. TRENDS in Genetics 22:139-155.

Zeuner, FE. 1963. A history of domesticated animals. A history of domesticated animals.

CHAPTER 2: Island Southeast Asian origin and the dispersal of Polynesian chickens indicated by mitochondrial DNA

Statement of authorship

Island Southeast Asian origin and the dispersal of Polynesian chickens indicated by mitochondrial DNA

Manuscript in preparation for *Molecular Biology and Evolution*

Michael James Herrera (Candidate)
Collected and liaised for modern and museum samples, performed DNA extractions, PCR amplifications, processed and analysed the data, interpreted the data, created figures and tables, and wrote the paper.

I hereby certifv that the statement of contribution is accurate.

Signed _____ Date 6 March 2015

Jeremy Austin
Assisted in the design of the study, gave advice on laboratory work, assisted in the interpretation of results, and edited the manuscript

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 6 March 2015

Vicki Thomson
Assisted in the genetic analysis, interpretation of results, and edited the manuscript

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 5 March 2015

Jessica Wadley
Assisted in the genetic analysis, interpretation of results, and edited the manuscript

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 5 March 2015

Alan Cooper
Assisted in the study design, and commented and edited the manuscript

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 6 March 2015

Philip Piper
Assisted in the interpretation of results, commented and edited the manuscript

I hereby certify that the statement of contribution is accurate.

2 March 2015

Signed _____ Date _____

Jaime Gongora
Liaised the Indonesian samples, commented and edited the manuscript

I hereby certify that the statement of contribution is accurate.

3 March 2015

Signed _____ Date _____

Island Southeast Asian origin and dispersal of Polynesian chickens indicated by mitochondrial DNA

Michael B. Herrera[A], Vicki A. Thomson[A], Jessica J. Wadley[A], Philip J. Piper[b], Alan Cooper[A], Jaime Gongora[C], and Jeremy J. Austin[A]

[A]Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, South Australia 5005, Australia

[B]School of Archaeology and Anthropology, Australian National University, Canberra ACT 0200, Australia

[C]Centre for Advanced Technologies in Animal Genetics and Reproduction, Faculty of Veterinary Science, University of Sydney, Sydney NSW 2006, Australia

Abstract

The colonisation of the Pacific islands was one of the most significant human migrations in prehistory. This process involved the spread of Austronesian languages, genes, and culture, as well as the translocation of several plants and animals from Mainland Southeast Asia (MSEA) through Island Southeast Asia (ISEA) as far as Remote Oceania. The recovery of chicken bones from numerous sites in Near- and Remote Oceania suggests domestic fowl likely played an important economic, social, and ritual role during early colonisation. However, both the origins and routes involved in the translocation of chickens from MSEA and/or ISEA into the Pacific remain uncertain. Within the Pacific most modern day chickens and all archaeological samples possess a group of unique, closely related mitochondrial DNA (mtDNA) haplotypes within haplogroup D. These represent the ancestral mtDNA lineages transported across the Pacific. In order to determine the homeland of Pacific chickens and to explore their human-mediated routes of translocation into the Pacific, we surveyed mtDNA control region variation in native village chickens in ISEA including likely source populations in Indonesia and the Philippines. We compared these with sequences of both modern and ancient chickens from the Pacific, and modern chickens globally and used coalescent simulations to test scenarios that likely explain the distribution of chickens in the region. Village chickens from ISEA predominantly belong to the very diverse haplogroup D, and ancestral Pacific haplotypes (referred to as Polynesian D hereafter) are observed only in chickens from a limited location in the Philippines. These results suggest that Pacific chickens trace their ancestors to the Philippines, but not further north in Taiwan, nor to the south and west in Indonesia. A subset of haplogroup D (*i.e.*, Polynesian D) may have been integrated into the Austronesian expansion in the

Philippines *en route* to the Pacific. This highlights the complex origins of domesticated animals translocated into the Pacific region during prehistoric times.

Keywords

Gallus gallus, Island Southeast Asia, Pacific, control region

Introduction

The colonisation of remote islands and archipelagos in the Pacific is considered one of the greatest global diasporas in prehistory. This migration is thought to have begun around 4200-4000 cal. BP with movements of Austronesian-speaking peoples from Taiwan into the Philippines, south and west through Indonesia, eventually establishing the Lapita Cultural Complex at c. 3300 cal. BP in Near Oceania, before migrating east as far as Remote Oceania (Bellwood 2007). Adaptive strategies for these transoceanic colonisations included the deliberate translocation of several plants and animals (Kirch 1997; Matisoo-Smith 2009). This included three domestic animals that have their origins in Mainland Southeast Asia (MSEA): pigs (*Sus scrofa*), dogs (*Canis* spp.) and chickens (*Gallus gallus*). The early arrival of domesticates in Near Oceania, potentially with the first Austronesian colonists, and subsequent transport across the Pacific, along with rats (*Rattus* spp.), makes them especially informative as proxies for understanding patterns of human migration. For example, genetic studies of pigs (Larson *et al*. 2007a; Dobney *et al*. 2008) and dogs (Oskarsson *et al*. 2011; Sacks *et al*. 2013) have shown that they were translocated from MSEA through ISEA via peninsular Malaysia and the Indonesian islands of Sumatra and Java, then the lesser Sunda chain and into the Pacific. In contrast, studies on Pacific rats (*Rattus exulans*) and Polynesian chickens suggest a homeland

in ISEA (Thomson *et al*. 2014a; Thomson *et al*. 2014b), and their likely routes of translocation potentially vary from those observed for pigs and dogs. Zooarchaeological studies have also begun to demonstrate the unlikely nature of contemporaneous introduction of the three major domesticates to different locales across MSEA and ISEA (Piper *et al*. 2009; Amano *et al*. 2013). The combination of current genetic and zooarchaeological research seems to indicate complex and potentially independent routes of translocation for the three domesticates and the Polynesian rat (*Rattus exulans*) from MSEA or within ISEA and into the Pacific.

Archaeological and genetic studies on chickens have already provided some insights into their domestication and dispersal/translocation histories. For instance, Southeast Asia and the Indian subcontinent have been identified as putative centres for chicken domestication (Liu *et al.* 2006; Kanginakudru *et al.* 2008). Records from Talepakamalai and Etakosarai on Mussau, Watom on New Britain and at Teouma on Vanuatu indicate that chickens had crossed ISEA and reached Near Oceania with the Lapita Culture before 3000 cal. BP (Storey *et al.* 2008; Storey *et al*. 2012) . However, archaeological records of chicken before c.3000 cal. BP in ISEA are non-existent and almost nothing is known about their origins, timings of introduction or routes of translocation even though the genetic composition and diversity of chickens in this important region are essential for reconstructing migration models from MSEA to the Pacific.

Of the several divergent chicken mitochondrial DNA (mtDNA) haplogroups (A-I, control region; W-Z mtDNA genome) (Liu *et al.* 2006; Miao *et al*. 2012), two clades (D and E) are pertinent to understanding Pacific dispersal during prehistory. A

previous study proposed two independent chicken introductions to the Pacific, with haplogroup E representing the initial introduction, and D for a later one (Storey *et al.* 2007a). However, a growing body of evidence from ancient and modern DNA studies refute this finding and indicate haplogroup D is the only authentic lineage prehistorically translocated into the Pacific (Gongora *et al.* 2008; Thomson *et al.* 2014b). Recently, it was demonstrated that all archaeological and most modern chickens in Polynesia possess a unique set of four mtDNA control region single-nucleotide polymorphisms (SNPs) (Thomson *et al*. 2014b). Haplotypes containing this set of SNPs are thought to represent the founding mtDNA lineages dispersed by humans across Remote Oceania. This ancestral SNP motif has only been found in two locations outside the Pacific – Camiguin Island and Luzon in the Philippines – suggesting a homeland for the founding mtDNA lineages that is consistent with other lines of evidence about early Polynesian origins (Anderson 2009). However, sampling of chickens across ISEA was relatively poor until now, preventing definitive conclusions about the homeland of founding mtDNA lineages.

Here, we characterise the mtDNA diversity of contemporary chickens in ISEA, MSEA, and the Pacific. In particular, we aim to determine whether the ancestral Pacific chicken lineages survive in ISEA, and whether the current distribution of D haplotypes can provide clues to the origin of Polynesian chickens, and their introduction routes. To do this, we extensively sampled chickens from ISEA and MSEA and added these to existing mtDNA data from the region in order to provide a framework for examining the origin of chicken lineages present in the Pacific. Finally, we simulated genetic data to explore and test different population

migration scenarios for comparison with the observed patterns within ISEA and the Pacific.

Materials and methods

Sampling

A total of 1112 indigenous village chickens from across MSEA, ISEA, and the Pacific were sampled and sequenced (Figure 2-1). Areas of intensive commercial chicken farming were avoided and only free ranging chickens were selected. Samples consisted of plucked body feathers, or blood samples from the brachial vein stored on FTA cards (Qiagen, Hilden, Germany). Toe pads from historic chicken skins were sampled from the American Museum of Natural History (USA) and the National Museum of Natural History (France). Genomic DNA was extracted from feather samples using the salting-out method (Nicholls *et al*. 2000) and blood was extracted using a phenol-chloroform extraction (Kirby 1956). Museum specimens were extracted using a QIAmp DNA Tissue Kit (Qiagen, Hilden, Germany), following manufacturer's instructions, but with the following exception: the tissue digestion step (buffer ATL + Proteinase K) was conducted overnight at 55 ˚C with the addition of 200nM DTT at a final concentration.

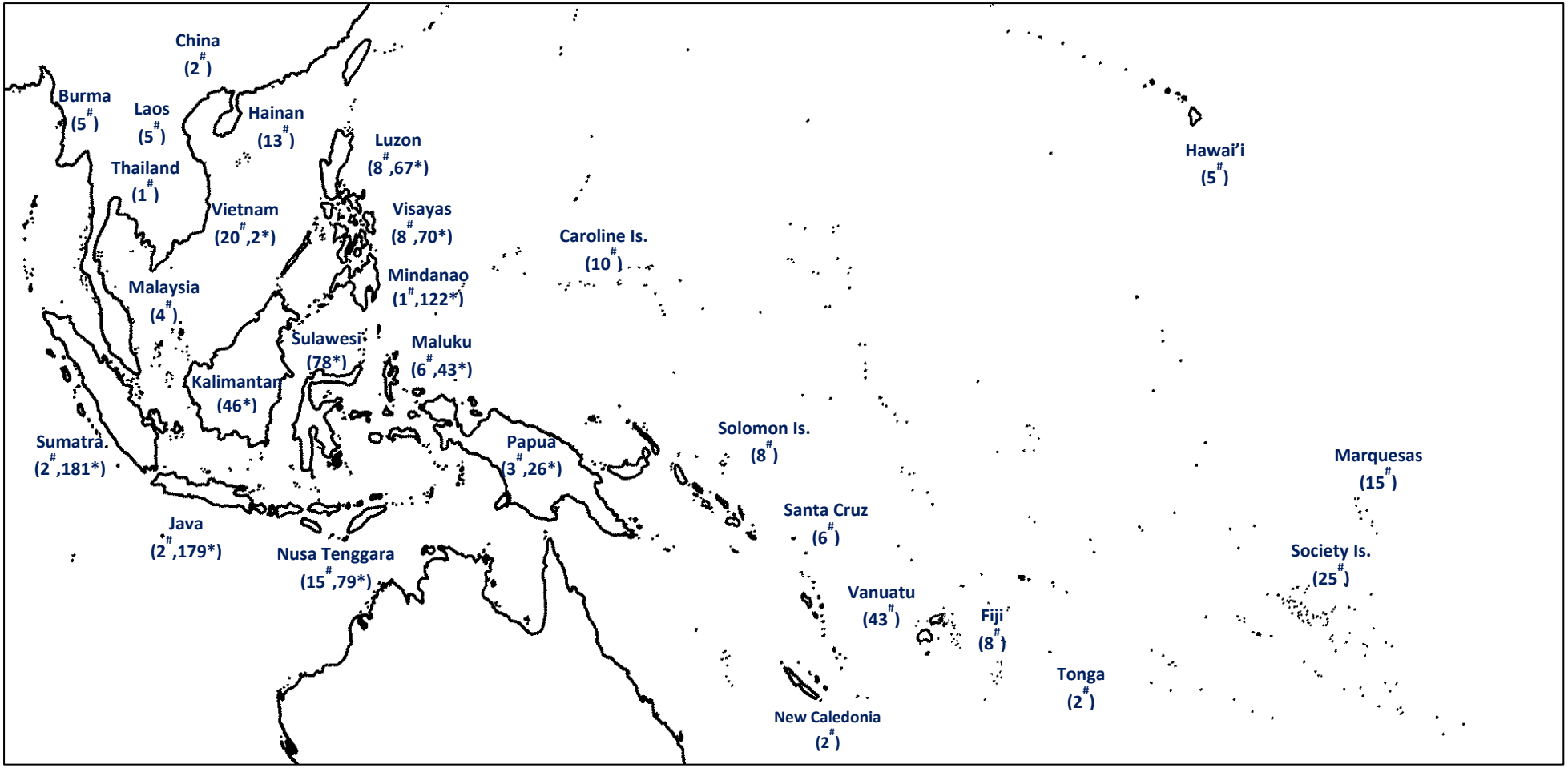


Figure 2-1. Sampling localities of modern and historical chickens collected in this study from Mainland Southeast Asia, Island Southeast Asia, and the Pacific. Name of locality followed by number of samples (# - museum, * - modern). Luzon includes Batanes and the main Luzon Island. Visayas includes islands of Bohol, Cebu, Iloilo, Leyte, Mindoro, Palawan, and Samar.

PCR amplification and sequencing

A 764 base pair (bp) fragment of the mtDNA control region was PCR amplified from modern specimens using primers GallP4F (5'-AACTCCCCTACTAAGTGTACCCCC-3') and GallP4R (5'-TTGACACTGATGCACTTTGGATCG-3') (this study), and a smaller 201bp fragment from the museum specimens using GG144F (5'-ACCCATTATATGTATACGGGCATTAA-3') and GG586R (5'-TCGGTCAGGCACATCCCATGCATAACT-3') (Storey *et al.* 2007a). Each PCR for the modern specimens (25µl final volume) contained 1x Hotmaster buffer (5Prime), 200 µM each dNTP, 0.2 µM of each primer and 0.5U of Hotmaster *Taq* DNA polymerase (5Prime). Thermocycling for modern samples comprised an initial denaturation and enzyme activation at 94°C for 2 min, followed by 30 cycles of denaturing at 94°C for 20 s, primer annealing at 55°C for 10 s and extension at 65°C for 60 s, and a final extension at 65°C for 70 min. PCRs for the museum samples contained 1x HiFi buffer (Invitrogen), 250 µM each dNTP, 2 mM $MgSO_4$, 1mg/mL RSA, 0.4 µM of each primer and 0.5 U of Platinum *Taq* HiFi DNA polymerase (Invitrogen). Thermocycling for museum specimens, included initial denaturation and enzyme activation at 94°C for 1 min, then 50 cycles of denaturing at 94°C for 15 s, primer annealing at 55°C for 15 s and extension at 68°C for 30 s, and a final extension 68°C for 10 m. PCR clean-up, Sanger sequencing and capillary electrophoresis were conducted at the Australian Genome Research Facility Ltd. (Australia).

Sequence chromatograms were assembled, visually inspected, and manually edited using Geneious 6.5.5 (Biomatters) to obtain a consensus sequence of 764 bp

for the modern specimens and 201 bp for the museum specimens. The 764 bp sequences were trimmed to match the 201 bp fragment representing nucleotide positions 166 to 367 of the mtDNA control region, using NC_007235 (Nishibori *et al*. 2005) as a reference sequence. Although short, this 201 bp fragment has the advantage of spanning the most variable region of the chicken mtDNA control region and being comparable to the available ancient DNA sequences and the reference CR dataset from across the world (Table SI2-1). Altogether, 6169 mtDNA control region sequences were available and aligned using the MUSCLE algorithm in Geneious 6.5.5.

Phylogeny and phylogeography

The aligned sequences were collapsed to unique haplotypes using FaBox 1.40 (Villesen 2007). Haplogroup assignment for newly generated sequences was made via identity to existing sequences of known haplogroup using a Neighbour-Joining tree generated with Tamura-Nei substitution model and a *Gallus varius* (haplotype H4, see results) sequence as outgroup. Regional phylogenetic networks were estimated using a median-joining (MJ) algorithm (Bandelt *et al*. 1999) in Network 4.6.1.0 (Fluxus Engineering) for haplotypes in haplogroup D from the Philippines, Indonesia, and the Pacific.

Population variability and structure

Measures of genetic diversity (haplotype diversity and nucleotide differences) and pairwise population F_{ST} values were calculated for each sampling location using Arlequin 3.5 (Excoffier *et al.* 2005). Each island in ISEA and the Pacific were set as the study population, whereas in MSEA this corresponds to countries. To visualise

relationships between populations, a non-parametric multidimensional scaling plot (MDS) on pairwise population F_{ST} scores was performed. These were initially done for all samples (irrespective of haplogroup) from all populations. Given the central role of haplogroup D during the initial colonisation of the Pacific and possibly haplogroup E (Storey *et al*. 2007b; Storey *et al*. 2010; Storey *et al.* 2012), we also carried out MDS on only haplogroup E individuals from all populations, haplogroup D individuals from populations in Indonesia, Philippines and the Pacific and haplogroup E individuals from the same regions. As a corollary to the MDS visualisations, an analysis of molecular variance (AMOVA) was performed to assess mtDNA structure in the ISEA-Pacific region. The AMOVA was done on the overall ISEA-Pacific dataset assuming no groups and with groups between populations found in the Philippines, Indonesia and the Pacific for haplogroups D and E only.

Population demographic history was investigated using mismatch distribution plots (Rogers & Harpending 1992) for all haplogroup D individuals from Indonesia, the Philippines, and the Pacific. Observed and simulated mismatch distributions were tested using the sum of squared differences (SSD) and Harpending's raggedness index (Harpending 1994). The results from the mismatch distribution plots were compared to tests for neutrality, Fu's FS (Fu 1997) and Tajima's D statistics (Tajima 1989), calculated using Arlequin 3.5 (Excoffier *et al.* 2005).

Bayesian coalescent simulation

To explore alternative translocation histories of chickens into the Pacific, we used Bayesian coalescent simulations (BayeSSC v1.0) (Anderson *et al.* 2005) to compare against observed phylogeographic patterns in ISEA and the Pacific. Control

region sequences from ISEA, the Pacific, and mainland Asia were used to model four major scenarios (Figure S2-1) of chicken translocation from Asia to the Pacific. The null model assumes one continuous population across Asia, ISEA, and the Pacific with gene flow everywhere. Model set 1 represents the expansion of the Austronesian-speaking peoples from Taiwan to ISEA and finally in the Pacific. Model set 2 approximates movements either from MSEA directly to the Philippines or initially through Hainan then the Philippines before arriving in the Pacific. Model set 3 allows for gene flow directly from Indonesia into the Pacific with no gene flow from the Philippines or MSEA. MSEA is assumed to be the geographic area (*i.e.*, Thailand, Laos, and Vietnam) where all modern populations currently found in ISEA and the Pacific would have coalesced back to in the past as it is part of the natural biogeographic range of red and grey jungle fowls and it is one of the regions where chicken domestication is believed to have occurred (Fumihito *et al.* 1994). Priors were placed on the timing of migration events. Population genetic statistics were calculated for 1 million simulated genealogies for each model, with population pairwise F_{ST} values selected to compare against those observed for each population. An Approximate Bayesian Framework (ABC) (Beaumont *et al*. 2002) was used to compare the simulated against the observed data, to construct posterior distributions, and to obtain the maximum likelihood estimate (MLE) for the migration events constructed for each set of simulations. A second set of simulations was run using the MLEs in place of the priors for 1000 generations so that goodness of fit for each scenario could be tested using Akaike's Information Criterion (AIC) (Akaike 1974).

Results

Classification of chicken mtDNA control region sequences and haplogroup distribution

We generated 1112 new mtDNA control region sequences from samples across Indonesia and the Philippines (*i.e.*, potential ancestral source populations of Polynesian chickens), plus additional sequences from MSEA and the Pacific Islands (Figure 2-1). When combined with existing Genbank sequences, a total of 6169 *Gallus gallus* sequences were used, which represent 527 haplotypes (termed H10-H536, for the 201 bp mtDNA control region fragment). Phylogenetic analysis shows the major haplogroups (Figure 2-2) previously defined in literature. The geographic distribution and frequency of major haplogroups varies globally. Haplogroups A and B are most common in East Asia and MSEA, with haplogroup C also abundant in East Asia, particularly in Japan. Haplogroup F, G, and K are observed only in mainland Asia, while haplogroup H is detected at low frequencies in ISEA, MSEA, and South Asia (Table 2-1).

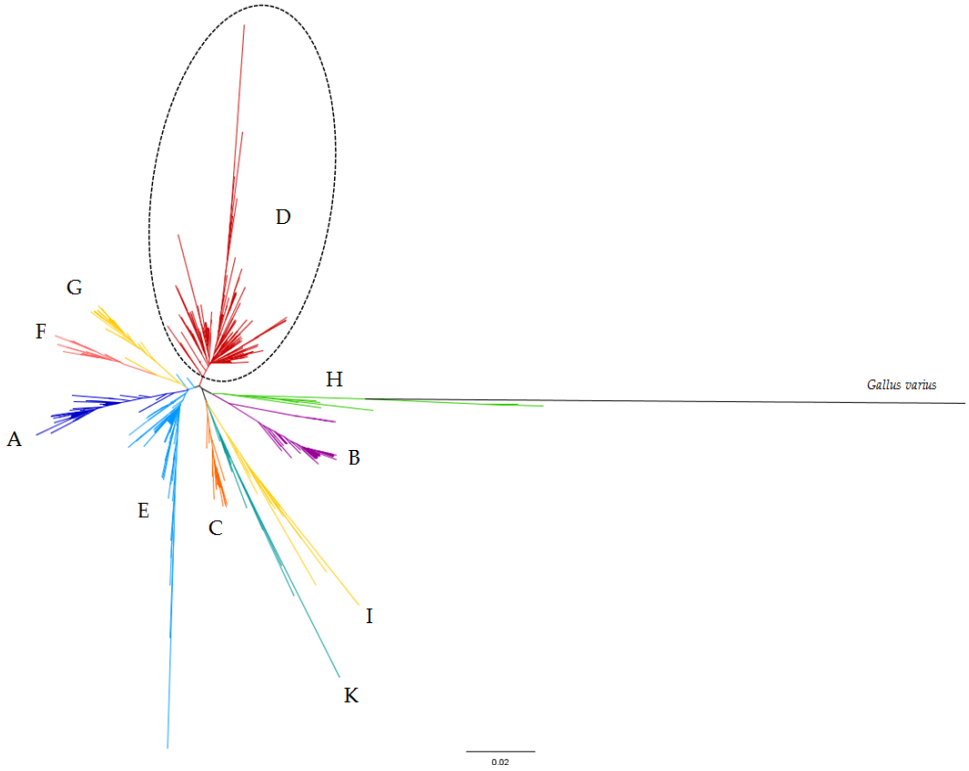


Figure 2-2. Neighbour-joining tree showing the relationships between 527 mtDNA control region haplotypes (201bp) from 6169 worldwide chicken samples. *Gallus varius* was used as an outgroup. Haplotypes are classified into 10 major haplogroups (A-I and K). The majority of haplotypes from Indonesia, Philippines and the Pacific fall within haplogroup D (dotted circle).

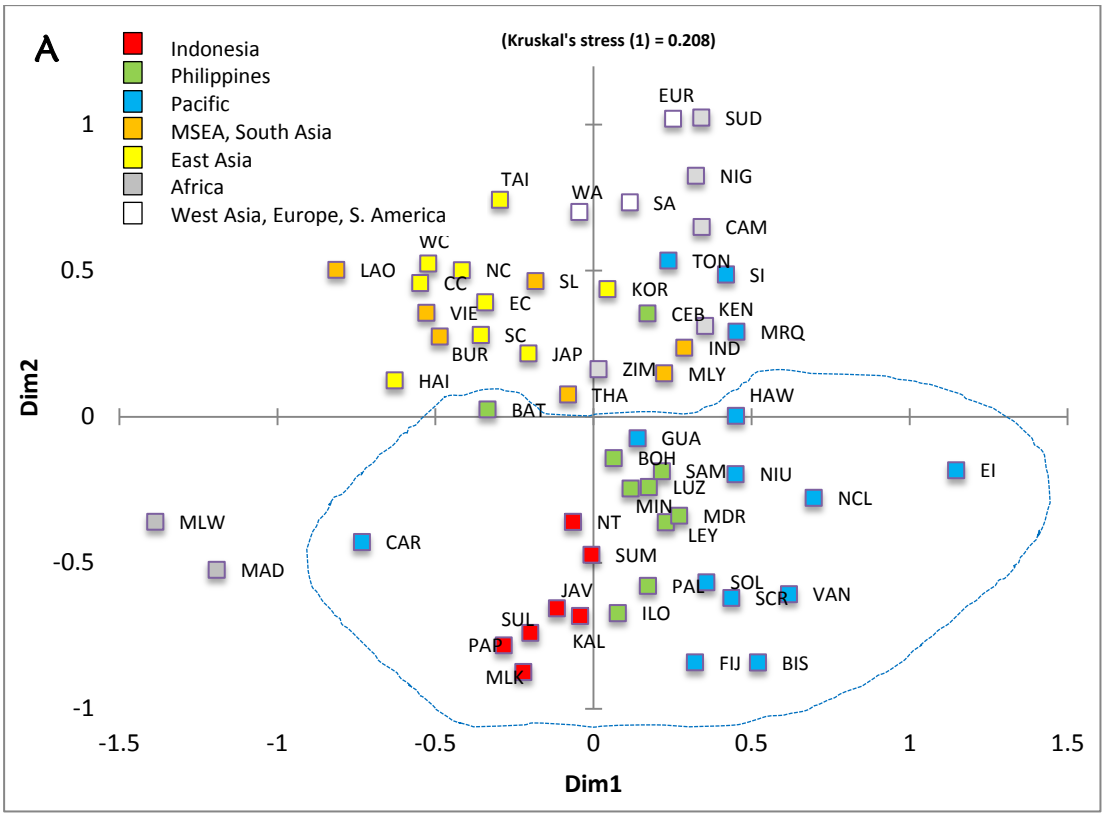
Table 2-1: Sampling locations, sample sizes, haplogroup and haplotype assignments for 6169 worldwide chicken samples.

LOCATION		HAPLOGROUP ASSIGNMENT	NO. OF HAPLOTYPES PER HAPLOGROUP
PHILIPPINES			
Batanes	8	B(4); D(4)	4(B=3; D=1)
Bohol	5	D(3); E(2)	3 (D=2; E=1)
Cebu	6	D(2); E(4)	4 (D=2; E=2)
Iloilo	5	D(5)	3 (D=3)
Leyte	4	D(3); E(1)	2 (D=1; E=1)
Luzon	72	A(2); B(7); C(2); D(47); E(14)	26 (A=2; B=4; C=1; D=15; E=4)
Mindanao	131	A(3); B(10); C(1); D(84); E(33)	34 (A=1; B=3; C=1; D=26; E=3)
Mindoro	19	B(1); C(1); D(14); E(3)	13 (B=1; C=1; D=8; E=3)
Palawan	43	C(1); D(40); E(1); H(1)	17 (C=1; D=14; E=1; H=1)
Samar	10	C(1); D(6); E(3)	9 (C=1; D=5; E=3)
INDONESIA			
Java	189	A(2); B(12); D(155); E(19); H(1)	36 (A=2; B=6; D=20; E=7; H=1)
Kalimantan	50	B(1); D(43); E(5)	14 (B=1; D=11; E=2)
Maluku	57	B(1); D(56)	10 (B=1; D=9)
Nusa Tenggara	93	A(1); B(12); D(74); E(5); H(1)	27 (A=1; B=2; D=21; E=2; H=1)
Papua	34	A(1); D(32); E(1)	6 (A=1; D=4; E=1)
Sulawesi	78	A(1); B(2); D(74); E(1)	22 (A=1; B=2; D=18; E=1)
Sumatra	191	A(4); B(20); D(155); E(6); H(6)	34 (A=2; B=3; D=21; E=4; H=4)
PACIFIC			
Bismarck	13	D(12); E(1)	2 (D=1; E=1)
Caroline Islands	10	D(9); E(1)	3 (D=2; E=1)
Easter Island	16	D (16)	2 (D=2)
Fiji	30	D(30)	4 (D=4)
Guam	5	D(3); E(2)	5 (D=3; E=2)
Hawaii	30	D(14); E (15)	5 (D=2; E=3)
Marquesas	20	D(7); E(13)	3 (D=1; E=2)
New Caledonia	2	D(2)	1 (D=1)
Niue	3	D(2); E(1)	2 (D=1; E=1)
Sta. Cruz	36	D(32); E(4)	7 (D=6; E=1)
Society Islands	25	D(8); E(17)	3 (D=1; E=2)
Solomon	32	B(1); D(28); E(3)	7 (B=1; D=4; E=2)
Tonga	2	E(2)	1 (E=1)
Vanuatu	84	A(1); D(77); E(6)	17 (A=1; D=13; E=3)
MSEA			
Burma	52	A(10); B(20); D(8); E(1); F(11); H(1); I(1)	17 (A=3; B=4; D=3; E=1; F=4; H=1; I=1)
Laos	74	A(17); B(39); D(8); E(1); F(4); G(2); I(3)	19 (A=4; B=3; D=5; E=1; F=2; G=1; I=3)
Malaysia	6	D(1); E(2); H(2); I(1)	5 (D=1; E=1; H=2; I=1)
Thailand	25	A(2); B(4); C(3); D(5); E(1); F(5);H(1); I(4)	13 (A=2; B=1; C=1; D=5; E=1; F=1; H=1)
Vietnam	204	A(37); B(89); C(5); D(11); E(11); F(19); G(14); I(18)	54 (A=10; B=12; C=1; D=8; E=4; F=2; G=4; I=12)
EAST ASIA			
China, Central	472	A(137); B(162); C(85); D(1); (E(84); G(3)	58 (A=16; B=14; C=15; D=1; E=11; G=1)
China, East	264	A(56); B(93); C(31); D(24); E(60)	45 (A=14; B=12; C=6; D=5; E=8)
China, North	108	A(34); B(29); C(21); E(24)	22 (A=5; B=1; C=7; E=9)
China, South	1292	A(277); B(383); C(66); D(41); E(85); F(212); G(228)	134 (A=23; B=28; C=8; D=10; E=13; F=23; G=29)
China, West	302	A(155); B(36); C(28); D(3); E(37); F(2); G(41)	43 (A=15; B=4; C=6; D=2; E=8; F=1; G=7)
Hainan	24	A(3); B(2); C(1); D(18)	14 (A=2; B=1; C=1; D=10)
Japan	400	A(141); B(25); C(107); D(49); E(78)	32 (A=6; B=1; C=8; D=9; E=8)
Korea	31	B(5); C(9); E(17)	6 (B=1; C=1; E=4)
Taiwan	193	A(91); B(1); C(9); E(92)	8 (A=3; B=1; C=1; E=3)
SOUTH ASIA			
India	359	A(2); B(7); C(2); D(88); E(229); F(9); H(4); K(18)	88 (A=1; B=2; C=1; D=28; E=39; F=1; H=1; K=15)
Sri Lanka	131	A(15); B(23); C(2); D(6); E(42); G(43)	26 (A=2; B=3; C=1; D=6; E=10; G=4)
AFRICA			
Cameroon	3	E(3)	1 (E=1)
Kenya	159	D(53); E(106)	21 (D=7; E=14)
Madagascar	79	D(67); E(12)	8 (D=5 ; E=3)
Malawi	19	D(19)	2 (D=2)
Nigeria	38	A(1); D(4); E(33)	31 (A=1; D=4; E=26)
Sudan	20	E(20)	3 (E=3)
Zimbabwe	201	A(36); B(21); D(59); E(85)	28 (A=8; B=1; D=6; E=13)
EUROPE	116	A(7); B(1) E(108)	10 (A=3; B=1; E=6)
SOUTH AMERICA	48	A(3); B(6); D(1); E(38)	11 (A=1; B=2; D=1; E=7)
WEST ASIA	26	A(4); B(4); E(18)	6 (A=2; B=1; E=3)

Haplogroups D and E, considered crucial in understanding the origins and dispersal of chickens in the Pacific, are found at varying frequencies across MSEA, ISEA, and the Pacific. Within the Pacific, the majority of chickens belong to haplogroup D (78%), compared to haplogroup E (21%). Similarly, the adjacent region of ISEA is predominantly composed of haplogroup D lineages (80%), compared to haplogroup E (10%). Haplogroups D and E are both found in high proportion in Africa. However, haplogroup E is also observed at high proportion of chickens in Europe (93%), South America (69%), West Asia (69%), and South Asia (55%).

Population structure

The MDS plot for pairwise F_{ST} values for 57 worldwide populations generally separates island populations, particularly ISEA and the Pacific, from continental populations (below the zero axis vs. above in dimension 2; Figure 2-3a). The plot has a low stress level (0.157), indicating it is a reliable visualisation of the pairwise distances. The distribution of the 57 study populations on the two axes of the MDS appears to be due to their relative haplogroup composition. For instance in dimension 2, the proportion of haplogroup D in populations below the zero axis is substantially higher than those above it (but see Table 1).



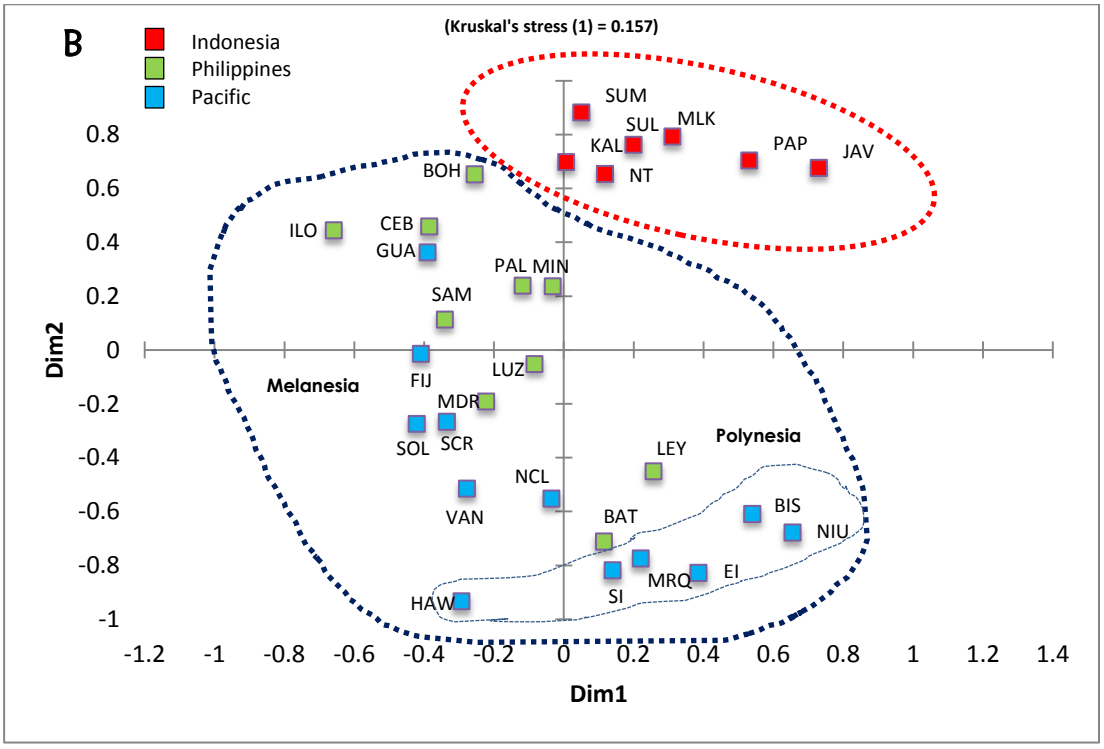


Figure 2-3. Multidimensional scaling plots (MDS) on population pairwise F_{st} for (a) 6169 worldwide chicken samples using all haplogroups (red: Indonesia; green: Philippines; blue: Pacific; orange: MSEA and South Asia; yellow: East Asia; grey: Africa; white: Europe, West Asia, and South America. Populations are abbreviated as follows: BAT, Batanes; BOH, Bohol; CEB, Cebu; ILO, Iloilo; LEY, Leyte; LUZ, Luzon; MIN, Mindanao; PAL, Palawan; SAM, Samar; JAV, Java; KAL, Kalimantan; MLK, Maluku; NT, Nusa Tenggara; PAP, Papua; SUL; Sulawesi; SUM, Sumatra; BIS, Bismarck; CAR, Caroline Islands; EI, Easter Islands; FIJ, Fiji; GUA, Guam; HAW, Hawai'i; MRQ, Marquesas; NCL, New Caledonia; NIU, Niue; SCR, Santa Cruz; SI, Society Islands; SOL, Solomon;

TON, Tonga; VAN, Vanuatu; BUR, Burma; LAO, Laos; MLY, Malaysia; THA, Thailand; VIE, Vietnam; CC, Central China; EC, East China; NC, North China; SC, South China; HAI, Hainan; JAP, Japan; KOR, Korea; TAI, Taiwan; IND, India; SL, Sri Lanka; CAM, Cameroon; KEN, Kenya; MAD, Madagascar; MLW, Malawi; NIG, Nigeria; SUD, Sudan; ZIM, Zimbabwe; EUR, Europe; SA, South America; and WA, West Asia. (b) 1038 haplogroup D chicken samples from Indonesia, Philippines, and Pacific (colours and abbreviations are the same as (a)).

As D haplotypes appear to drive the clustering of island populations in the Philippines, Indonesia, and the Pacific, an MDS plot of just haplogroup D individuals from these regions was generated. This reveals a clear separation between Indonesian populations and those from the Philippines and the Pacific (Figure 2-3b). Forming a tight cluster, Indonesia seems to be only distantly associated with the Philippines and the Pacific. A larger MDS space occupied by Philippine populations also suggests higher diversity compared to Indonesia. Furthermore, a contrast between most Melanesian (Solomon, Santa Cruz, Vanuatu, New Caledonia, and Fiji) and Polynesian (Society Islands, Niue, Marquesas, and Easter Island) islands in dimension one can also be observed. In contrast to the clear phylogeographic patterns seen in haplogroup D, MDS plots just for haplogroup E individuals show no signals of geographic structure (Figure S2-2a, b).

AMOVA shows that a high proportion of genetic variation is found within populations (Table 2-2). The proportion of among group variance (25.59%) for the three regions is significant for haplogroup D, whereas haplogroup E produced a non-significant variance component (2.88%). Furthermore, the highest variance component for D haplogroup is observed in the combined Philippine-Indonesia vs. Pacific groupings (28.36%), whereas genetic structure is not observed in the combined Indonesia-Pacific vs. Philippine groupings (-0.16%).

Table 2-2. Population genetic structure estimated from the analysis of molecular variance (AMOVA) based on mtDNA control region sequences from ISEA and Pacific chickens.

Group	n	Number of population	Number of groups	Variance components (%)		
				Among groups	Among populations within groups	Within Populations
Haplogroup D						
No grouping	1038	29	1	…	**25.02**	**74.98**
Regions*	1038	29	3	**25.59**	**6.52**	**67.90**
Philippines/Pacific vs. Indonesia	1038	29	2	**24.41**	**8.53**	**66.06**
Philippines/Indonesia vs. Pacific	1038	29	2	**28.36**	**9.87**	**61.77**
Indonesia/Pacific vs. Philippines	1038	29	2	**-0.16***	**25.11**	**75.05**
Haplogroup E						
No grouping	164	25	1	…	**29.41**	**70.59**
Regions*	164	25	3	**2.88***	**27.11**	**70.01**
Philippines/Pacific vs. Indonesia	164	25	2	**2.97***	**27.73**	**69.30**
Philippines/ Indonesia vs. Pacific	164	25	2	**2.71***	**27.60**	**69.69**
Indonesia/Pacific vs. Philippines	164	25	2	**0.17***	**29.30**	**70.53**

Distribution and diversity of haplogroup D

Haplogroup D occurs at high frequency in southern Africa, Madagascar, ISEA and the Pacific region (Figure 2-4) and at much lower frequency in continental Asia. However, comparing haplogroup D at a regional level reveals MSEA as the most diverse for haplotype ($h=0.94 \pm 0.02$) and nucleotide ($\pi=0.019 \pm 0.011$) diversities (Table 2-3), despite being observed at a low frequency in the region. Within the ISEA-Pacific region, haplogroup D individuals are most diverse in the Philippines ($h=0.91 \pm 0.01$; $\pi= 0.011 \pm 0.006$), followed by Indonesia ($h=0.75 \pm 0.02$; $\pi=0.007 \pm 0.004$), and then the Pacific ($h=0.73 \pm 0.03$; $\pi=0.008 \pm 0.005$).

The four diagnostic Polynesian SNPs (defined by Thomson *et al.* 2014) are found in D haplotypes from a relatively restricted geographic area in the region: only in the Philippines (*i.e.*, Batanes, Cagayan Valley, Ifugao, Palawan, Mindoro, and Mindanao) and in one sample from Hainan but not further west on continental Asia (Figure 2-4). The Polynesian motif is not detected in the Indonesian archipelago, but it is observed at high frequencies in the Pacific.

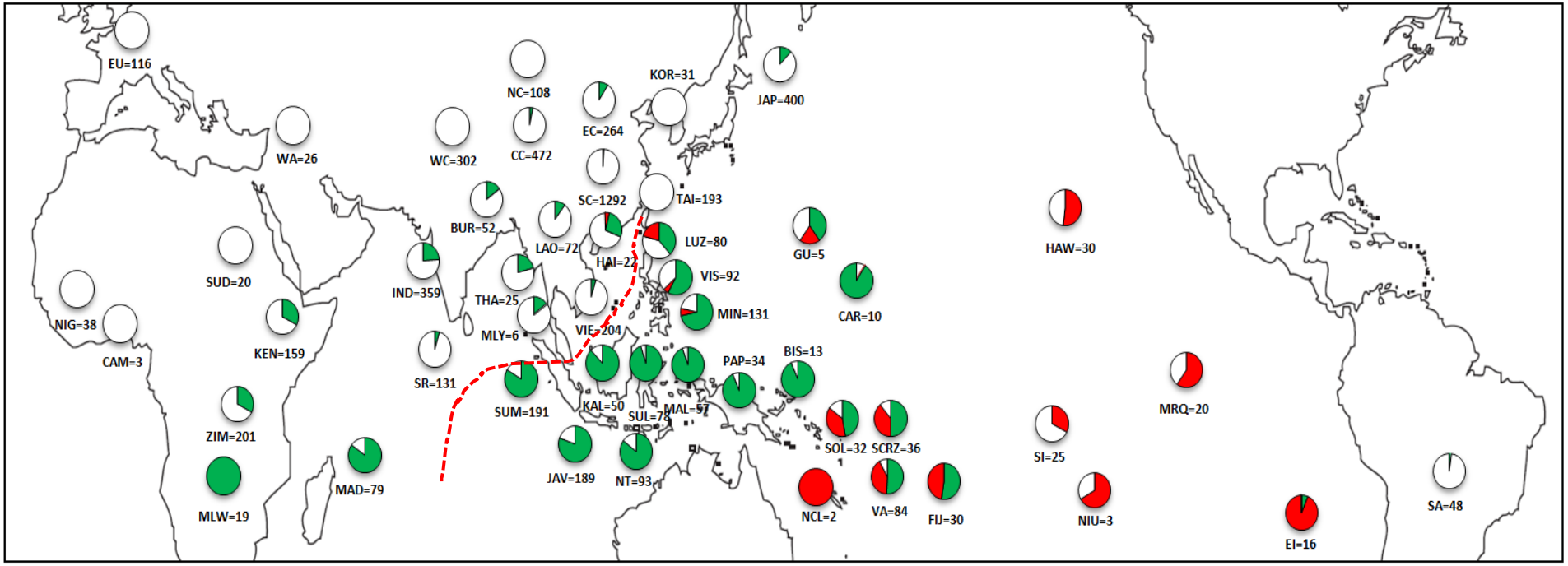


Figure 2-4. Frequency of mtDNA haplogroup D (green), D haplotypes containing the 4-SNP Polynesian motif (red), and other haplogroups (white) in world-wide chicken populations. Sampling locality (abbreviation are the same as Fig. 2-3) is followed by sample size.

Table 2-3: Population genetic summary statistics for regional mtDNA haplogroup D sequences.

Population		Genetic Diversity			Neutrality Tests		Mismatch Distribution Demographic Expansion		Mismatch Distribution Spatial Expansion	
Region	**N (H)**	**HD (SD)**	**ND (SD)**	**Pi**	**Tajima's D**	**Fu's F_S**	**SSD**	**HRI**	**SSD**	**HRI**
Philippines	208 (44)	0.91 (0.01)	0.010764 (0.006350)	2.16	-1.78* (*P*=0.01)	-27.04* (*P*=0.00)	0.0034*	0.0554*	0.0034*	0.05542*
Indonesia	590 (67)	0.75 (0.02)	0.006921 (0.004392)	1.26	-2.21* (*P*=0.00)	-27.68* (*P*=0.00)	0.0031*	0.0784*	0.0031*	0.0784*
Pacific	240 (26)	0.73 (0.03)	0.007718 (0.005124)	1.55	-1.46* (*P*=0.03)	-18.09* (*P*=0.00)	0.0078	0.0367	0.0081*	0.0367
MSEA	33 (19)	0.94 (0.02)	0.019015 (0.010913)	3.82	-1.50 (*P*=0.05)	-9.66 (*P*=0.00)	0.0014	0.0151	0.0012	0.0151
East Asia	136 (27)	0.91 (0.01)	0.020719 (0.011476)	4.16	-0.85 (*P*=0.22)	-7.36 (*P*=0.02)	0.0051	0.0164	0.0038	0.0164
South Asia	94 (32)	0.85 (0.03)	0.018991 (0.010680)	3.82	-1.73 (*P*=0.01)	-17.80 (*P*=0.00)	0.0235	0.0612	0.016	0.0612
Africa	275 (19)	0.44 (0.04)	0.005536 (0.004013)	1.11	-1.64 (*P*=0.02)	-11.99 (*P*=0.00)	0.0273	0.2367	0.0119	0.2367

Phylogenetic relationship of D haplotypes from ISEA and the Pacific

There are 25 D haplotypes observed in the Pacific region (Figure 2-5). Eight haplotypes are shared between Pacific populations and ISEA or mainland Asia while 17 are unique to the Pacific. More than half of D haplotypes (n=15/25, 60%) in the Pacific contain the Polynesian 4 SNP motif. Interestingly, only two Polynesian D haplotypes (H453 and H498) are found in ISEA, and only in the Philippines. Two haplotypes with the Polynesian 4 SNP motif are not yet detected in the Pacific, but are found in ISEA: H440 (n=1) and H446 (n=7) from the Philippines and H410 (n=1) from Hainan. H453 is the most common Polynesian D haplotype both in the Pacific and the Philippines, with a star-shaped cluster of rare haplotypes emanating from H453. The haplotypes that are just one mutation away from H453 are mostly from Melanesia and Polynesia. In contrast, most Micronesian haplotypes are shared with the Philippines and Indonesia (H414, H509, and H536) and do not contain the Polynesian 4 SNP motif, save for one haplotype (H320) from Guam.

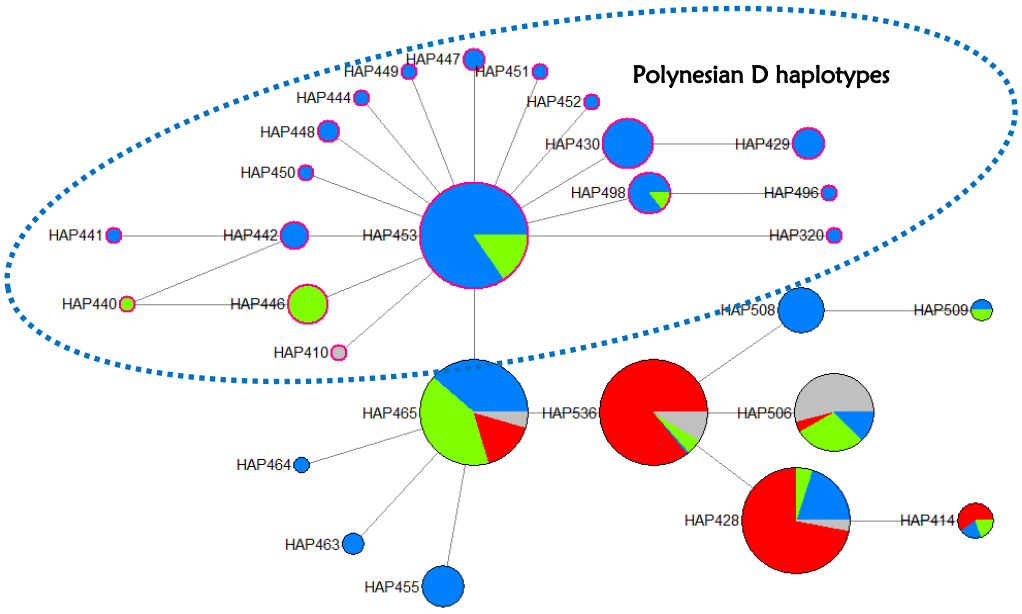


Figure 2-5. Median-joining network of 25 D haplotypes found in the Pacific including Polynesian haplotypes from the Philippines (H440, H446) and Hainan (H410). Pie charts are coloured by region (blue, Pacific; green, Philippines; red, Indonesia; grey, outside of ISEA-Pacific region). The dashed line indicates haplotypes containing the 4 SNP Polynesian motif.

Population dynamics of chickens in ISEA and the Pacific

As the maternal genetic structure of chickens across ISEA and the Pacific appears to be largely driven by haplogroup D, neutrality tests were only done on haplogroup D. Neutrality tests (Tajima's D and Fu's FS) indicate that populations of haplogroup D in ISEA and the Pacific show a significant expansion signal (Table 2-3). Mismatch distribution (MD) patterns in Indonesia and the Philippines are characteristically unimodal (Figure 2-6) and there is support for population expansion under sudden demographic and spatial expansions in both these regions, but only a spatial expansion for the Pacific. As a peak towards the right hand side of the mismatch plot (*i.e.*, a greater number of mismatches) suggests an older expansion event, the position of the peak indicates ancient expansions of D haplogroup in ISEA (peak at 3 and 2 mismatches for the Philippines and Indonesia, respectively), whereas in the Pacific, it shows a more recent expansion (peak at 1 mismatch; Figure 2-6).

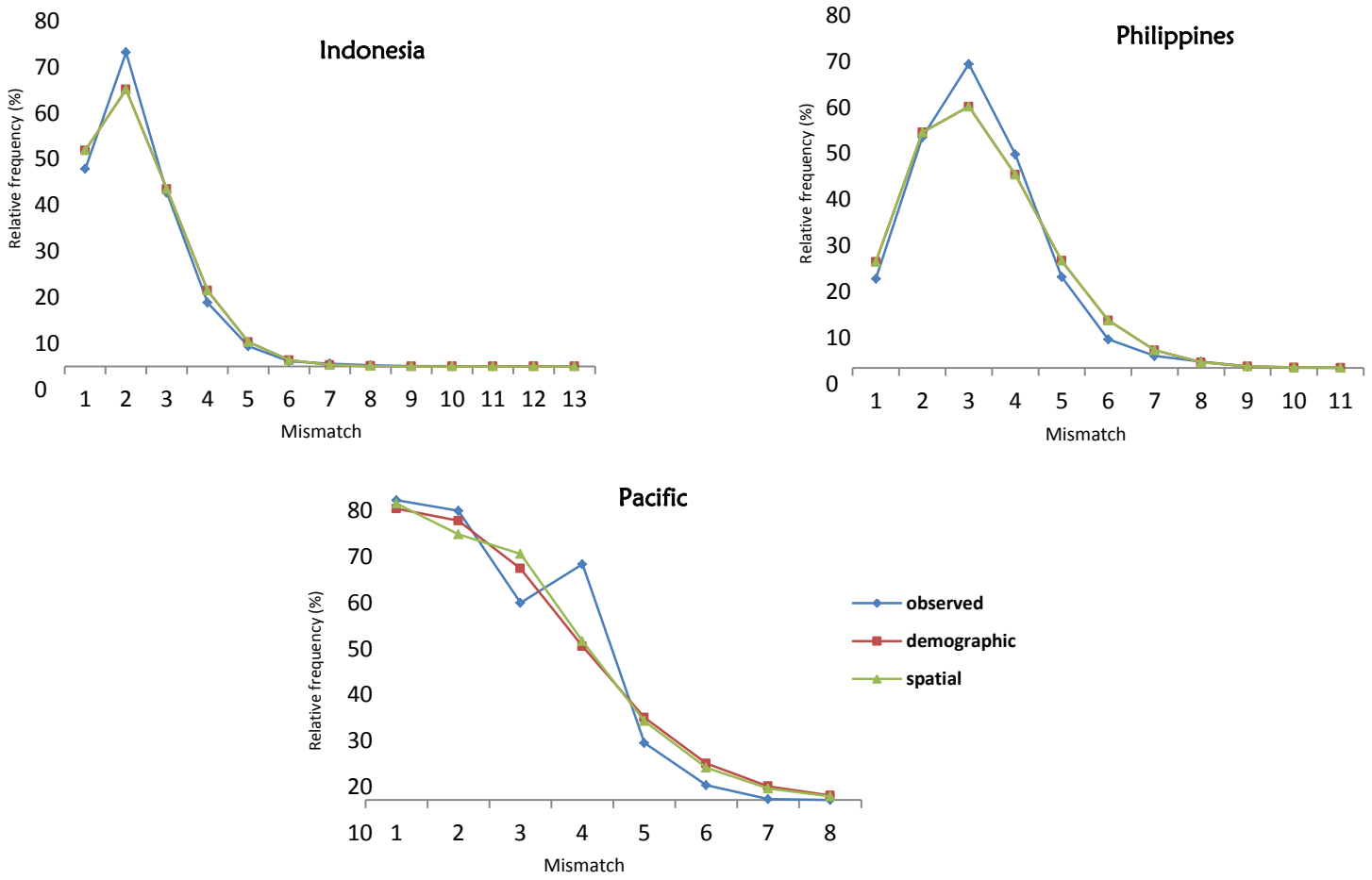


Figure 2-6. Mismatch distribution patterns for mtDNA control region haplogroup D samples from chickens sampled in the Philippines, Indonesia, and the Pacific.

Expansion history testing using BayeSSC

Coalescent simulations identified a movement from mainland Asia through Taiwan and the Philippines then onto the Pacific, but with also some level of gene flow from Indonesia, as the most likely scenario for the translocation history of chickens into the Pacific (Scenario H3-B, Figure SI2-1). This scenario suggests Pacific chickens ultimately arrived from MSEA via two separate routes: 1) from MSEA through Taiwan and the Philippines to the Solomon Island archipelago, and 2) from MSEA through the Malaysian peninsula, Indonesia, Wallacea and PNG to the Solomon Islands.

Discussion

This study demonstrates that ISEA was a regional centre for haplogroup D diversification. It is likely that the initial dispersal of founding haplogroup D lineages into ISEA from mainland Asia and the subsequent founder effects in the archipelagic environment of ISEA led to the current level of diversity observed in the region (ISEA contains 60% of the worldwide haplogroup D diversity, which makes it second only to MSEA). The ubiquity and diversity of haplogroup D chickens in ISEA also suggests its protracted presence in the region. Unfortunately, archaeological chicken remains predating archaeological chickens in the Pacific have yet to be documented in ISEA, so investigating the temporal depth of haplogroup D in ISEA is not possible. Nevertheless, the presence of domestic chickens in Near Oceania before c. 3000 cal. BP (Storey *et al*. 2012) suggests they must have been transported through ISEA before this date.

Distinct geographical patterns of haplogroup D are evident within ISEA and the Pacific. Demographic and spatial expansion in the Philippines and Indonesia appear to have occurred at two different times. First in the Philippines then followed by Indonesia (*i.e.*, the location of the peak in mismatches; Figure 2-6). Spatial expansion in the Pacific clearly occurred in more recent prehistory, as indicated in the mismatch plots and the star-like cluster observed in the phylogenetic network.

The initial spread of domestic chickens into the Pacific appears to have originated from the Philippines. Genetically, this is best exemplified by the distribution of D haplotypes that possess the diagnostic motif of four SNPs that characterises Polynesian archaeological chickens. Our study shows that Polynesian D haplotypes are common in the Philippines and it is not found in Taiwan. This is also not observed in Indonesia despite the extensive sampling regime employed across this archipelago. Thus, the geographic distribution of Polynesian D haplotypes suggests that Indonesia did not contribute to the initial maternal lineage of Pacific chickens east of Near Oceania. It is also possible that Philippine domestic chickens did not spread south and west into Indonesia (but see below).

Although the greatest haplotype D diversity is recorded in MSEA, the Polynesian D haplotypes are yet to be found there. The presence of two haplotypes, H440 and H446, in the Philippines suggests this archipelago as the origin for the Polynesian D group. The lack of the Polynesian D haplotypes in eastern Indonesia also suggests a direct route of human-mediated transfer of chickens from the Philippines to Near Oceania. Most likely, migrating Austronesian-speaking populations transported with them, along with many aspects of material culture, the

founding haplogroup D lineages containing the Polynesian motif that subsequently gave rise to the different Polynesian D haplotypes found throughout much of Remote Oceania.

The high frequency and ubiquity of haplotype H453 in Polynesia suggests that this lineage was transported from the Philippines to the Pacific prehistorically. One hypothesis to explain the diversity of Polynesian D haplotypes in the Pacific today is a combined effect of *in situ* evolution and drift. The presence of Polynesian D haplotype H498 found on both the Philippine island of Mindoro and in Vanuatu is somewhat enigmatic. It is possible that this is a prehistoric introduction that expanded no further than Vanuatu. However, another plausible hypothesis is that the chickens transported into the Pacific contained a mixed group of Polynesian D haplotypes. Neither hypothesis explains the presence of Polynesian D haplotype (H410) from Hainan; however, this one sample could also represent a recent back-migration from Luzon to Hainan.

The mismatch plots also suggest an expansion of domestic chickens through Indonesia, possibly through Sumatra and Java and spreading across eastern Indonesia as far as the Bismarck Archipelago and Solomon Islands as indicated by the simulation results (see H3-B, Figure S2-1). The presence of a second group of lineages from Indonesia in the Pacific (yet in contrast to the absence of Polynesian D haplotypes in Indonesia) implies a second separate introduction of chickens to Near Oceania, unrelated to those carrying the Polynesian D motif. The lack of haplotypes carrying the Polynesian D motif west of the Solomon's means either, that only a subset of chickens from the separate introductions was transported further into the

Pacific, or more likely, initial translocation east of the Solomon's had already occurred prior to the arrival of the Indonesian haplotypes in Near Oceania.

Interestingly, the two-route pattern of chicken introduction to ISEA through the Philippines and Indonesia is remarkably similar to that recorded for pigs (Larson *et al.* 2007b). The 4000 – 3500-year-old status of both these pig introductions has been confirmed by recovery from numerous archaeological sites (Piper *et al.* 2009; Amano *et al*. 2013). However, in the case of pigs, it is the Indonesian 'Pacific Clade', rather than those of Philippine origin, that was eventually transported into Remote Oceania (Larson *et al.* 2007b).

It was initially thought that plants and animals were transported contemporaneously across ISEA and into the Pacific during the initial migration of Austronesian-speaking peoples from Taiwan. The pig data clearly demonstrated that this was not the case, and that there were additional routes of potential human population movements and animal translocation from the mainland across ISEA and into the Pacific. What this current study of chickens does is to further emphasize the complexities involved in human migration into the Pacific, and highlights potential problems in the temporal sequencing of the prehistoric arrivals of translocated domestic animals into Oceania.

Conclusion

Our extensive sampling regime across ISEA allowed us to test for the first time possible origins of the Polynesian chickens. The results indicated two potential routes of translocation, with initial colonisation from the Philippines of chickens

possessing the Polynesian motif followed by a slightly later arrival of domestic fowl of MSEA and Indonesian origin. The distribution of the Polynesian SNP motif demonstrates the potential of chicken genetic studies in testing theories about the prehistoric peopling of the Pacific as hypothesised by archaeological and linguistic studies. It also validates that the highly variable region of the mtDNA control region holds sufficient signal to inform us about the translocation history of chickens. However, studies using nuclear data may allow finer-scale resolution of these theories, particularly if archaeological chicken remains are subsequently found in ISEA with well-preserved ancient DNA.

Acknowledgement

This research was funded by the Australian Research Council (ARC) Discovery Project ID DP110105187. We thank the American Museum of Natural History (USA) and the National Museum of Natural History (France) for providing the historic museum samples. We also thank the Biological Research and Services Laboratory, Natural Science Research Institute of the University of the Philippines for allowing us to perform DNA extractions for the Philippine samples in their facility.

References

Akaike, H. 1974. A new look at the statistical model identification. Automatic Control, IEEE Transactions on 19:716-723.

Amano, N, PJ Piper, H-c Hung, P Bellwood. 2013. Introduced Domestic Animals in the Neolithic and Metal Age of the Philippines: Evidence From Nagsabaran, Northern Luzon. The Journal of Island and Coastal Archaeology 8:317-335.

Anderson, A. 2009. The rat and the octopus: initial human colonization and the prehistoric introduction of domestic animals to Remote Oceania. Biological Invasions 11:1503-1519.

Anderson, CN, U Ramakrishnan, YL Chan, EA Hadly. 2005. Serial SimCoal: a population genetics model for data from multiple populations and points in time. Bioinformatics 21:1733-1734.

Bandelt, H-J, P Forster, A Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular biology and evolution 16:37-48.

Beaumont, MA, W Zhang, DJ Balding. 2002. Approximate Bayesian computation in population genetics. Genetics 162:2025-2035.

Bellwood, P. 2007. Prehistory of the Indo-Malaysian Archipelago. Canberra, Australia: Australian National University Press.

Dobney, K, T Cucchi, G Larson. 2008. The pigs of Island Southeast Asia and the Pacific: new evidence for taxonomic status and human-mediated dispersal. Asian Perspectives 47:59-74.

Excoffier, L, G Laval, S Schneider. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evolutionary bioinformatics online 1:47.

Fu, Y-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915-925.

Fumihito, A, T Miyake, S-I Sumi, M Takada, S Ohno, N Kondo. 1994. One subspecies of the red junglefowl (Gallus gallus gallus) suffices as the matriarchic ancestor of all domestic breeds. PNAS 91:12505-12509.

Gongora, J, NJ Rawlence, VA Mobegi, H Jianlin, JA Alcalde, JT Matus, O Hanotte, C Moran, JJ Austin, S Ulm. 2008. Indo-European and Asian origins for Chilean and Pacific chickens revealed by mtDNA. Proceedings of the National Academy of Sciences 105:10308-10313.

Harpending, H. 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Human biology:591-600.

Kanginakudru, S, M Metta, R Jakati, J Nagaraju. 2008. Genetic evidence from Indian red jungle fowl corroborates multiple domestication of modern day chicken. BMC evolutionary biology 8:174.

Kirby, K. 1956. A new method for the isolation of ribonucleic acids from mammalian tissues. Biochemical Journal 64:405.

Kirch, PV. 1997. The Lapita peoples : ancestors of the oceanic world. Cambridge, Mass.: Blackwell Publishers.

Larson, G, T Cucchi, M Fujita, E Matisoo-Smith, J Robins, A Anderson, B Rolett, M Spriggs, G Dolman, T-H Kim. 2007a. Phylogeny and ancient DNA of Sus provides insights into neolithic expansion in Island Southeast Asia and Oceania. Proceedings of the National Academy of Sciences 104:4834-4839.

Larson, G, R Liuc, X Zhao, et al. 2007b. Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. PNAS 107:7686-7691.

Liu, Y-P, G-S Wu, Y-G Yao, Y-W Miao, G Luikart, M Baig, A Beja-Pereira, Z-L Ding, MG Palanichamy, Y-P Zhang. 2006. Multiple maternal origins of chickens: out of the Asian jungles. Molecular phylogenetics and evolution 38:12-19.

Matisoo-Smith, E. 2009. The Commensal Model for Human Settlement of the Pacific 10 Years on - What Can We Say and Where to Now? Journal of Island and Coastal Archaeology:151-163.

Miao, Y, M Peng, G Wu, Y Ouyang, Z Yang, N Yu, J Liang, G Pianchou, A Beja-Pereira, B Mitra. 2012. Chicken domestication: an updated perspective based on mitochondrial genomes. Heredity 110:277-282.

Nicholls, JA, MC Double, DM Rowell, RD Magrath. 2000. The evolution of cooperative and pair breeding in thornbills Acanthiza (Pardalotidae). Journal of Avian Biology 31:165-176.

Nishibori, M, T Shimogiri, T Hayashi, H Yasue. 2005. Molecular evidence for hybridization of species in the genus *Gallus* except for *Gallus varius*. Animal genetics 36:367-375.

Oskarsson, MC, CF Klütsch, U Boonyaprakob, A Wilton, Y Tanabe, P Savolainen. 2011. Mitochondrial DNA data indicate an introduction through Mainland Southeast Asia for Australian dingoes and Polynesian domestic dogs. Proceedings of the Royal Society B: Biological Sciences:rspb20111395.

Piper, P, H-c Hung, F Campos, P Bellwood, R Santiago. 2009. A 4000 year-old introduction of domestic pigs into the Philippine Archipelago: implications for understanding routes of human migration through Island Southeast Asia~ and Wallacea. Antiquity 83:687-695.

Rogers, AR, H Harpending. 1992. Population growth makes waves in the distribution of pairwise genetic differences. Molecular biology and evolution 9:552-569.

Sacks, BN, SK Brown, D Stephens, NC Pedersen, J-T Wu, O Berry. 2013. Y chromosome analysis of dingoes and Southeast Asian village dogs suggests a Neolithic continental expansion from Southeast Asia followed by multiple Austronesian dispersals. Molecular biology and evolution 30:1103-1118.

Storey, A, T Ladefoged, E Matisoo-Smith. 2008. Counting your chickens: density and distribution of chicken remains in archaeological sites of Oceania. International Journal of Osteoarchaeology 18:240-261.

Storey, AA, JS Athens, D Bryant, et al. 2012. Investigating the Global Dispersal of Chickens in Prehistory Using Ancient Mitochondrial DNA Signatures. PloS One 7:1-11.

Storey, AA, JM Ramírez, D Quiroz, DV Burley, DJ Addison, R Walter, AJ Anderson, TL Hunt, JS Athens, L Huynen. 2007a. Radiocarbon and DNA evidence for a pre-Columbian introduction of Polynesian chickens to Chile. Proceedings of the National Academy of Sciences 104:10335-10339.

Storey, AA, JM Ramirez, D Quiroz, et al. 2007b. Radiocarbon and DNA evidence for a pre-Columbian introduction of Polynesian chickens to Chile. PNAS 104:10335-10339.

Storey, AA, M Spriggs, S Bedford, SC Hawkins, JH Robins, L Huynen, E Matisoo-Smith. 2010. Mitochondrial DNA from 3000-year old chickens at the Teouma site, Vanuatu. Journal of Archaeological Science 37:2459-2468.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.

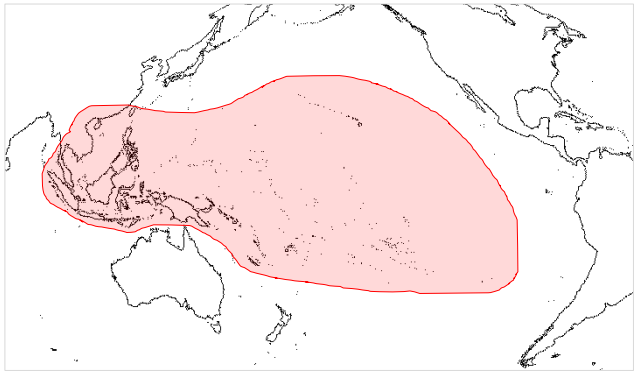
Thomson, V, KP Aplin, A Cooper, S Hisheh, H Suzuki, I Maryanto, G Yap, SC Donnellan. 2014a. Molecular Genetic Evidence for the Place of Origin of the Pacific Rat, Rattus exulans. PloS one 9:e91356.

Thomson, VA, O Lebrasseur, JJ Austin, TL Hunt, DA Burney, T Denham, NJ Rawlence, JR Wood, J Gongora, LG Flink. 2014b. Using ancient DNA to study the origins and dispersal of ancestral Polynesian chickens across the Pacific. Proceedings of the National Academy of Sciences 111:4826-4831.

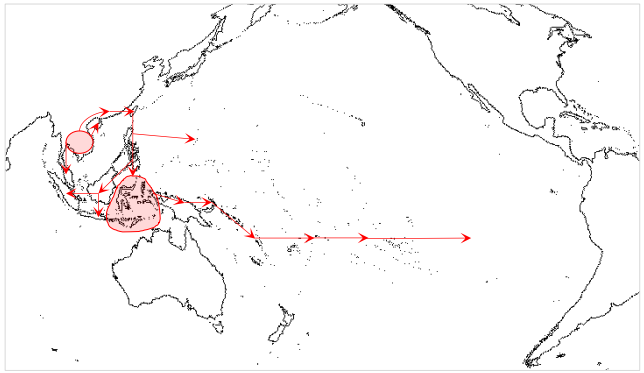
Villesen, P. 2007. FaBox: an online toolbox for fasta sequences. Molecular Ecology Notes 7:965-968.

Supplementary information

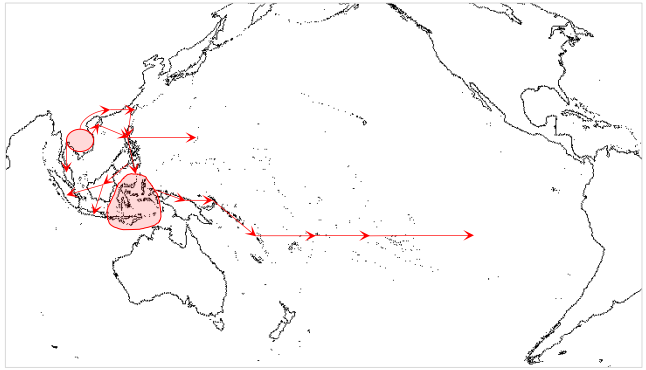
H_0, AIC=487.35



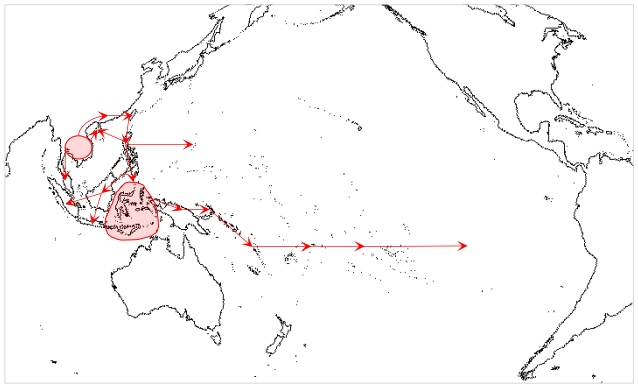
H_1-A, AIC=378.63



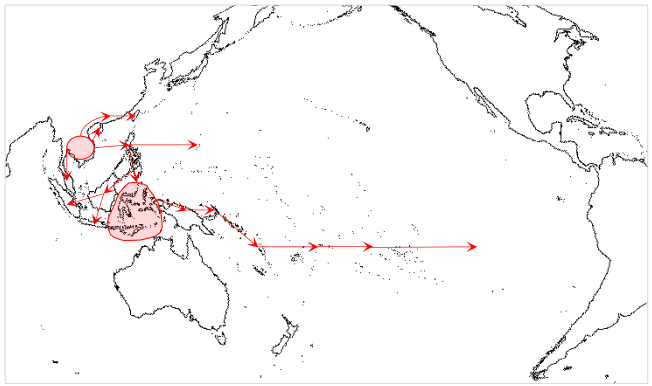
H_1-B, AIC=389.97



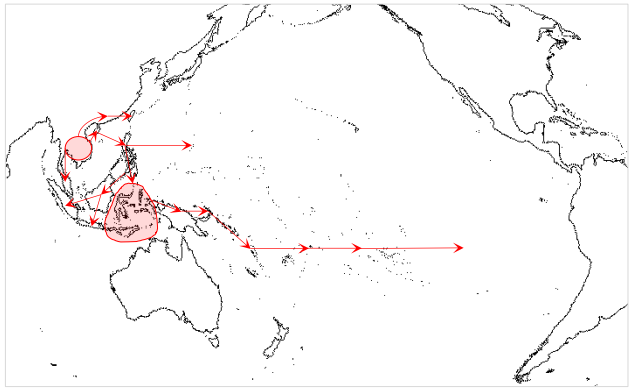
H_1-C, AIC=428.98



H_2-A, AIC=415.63



H_2-B, AIC=411.58



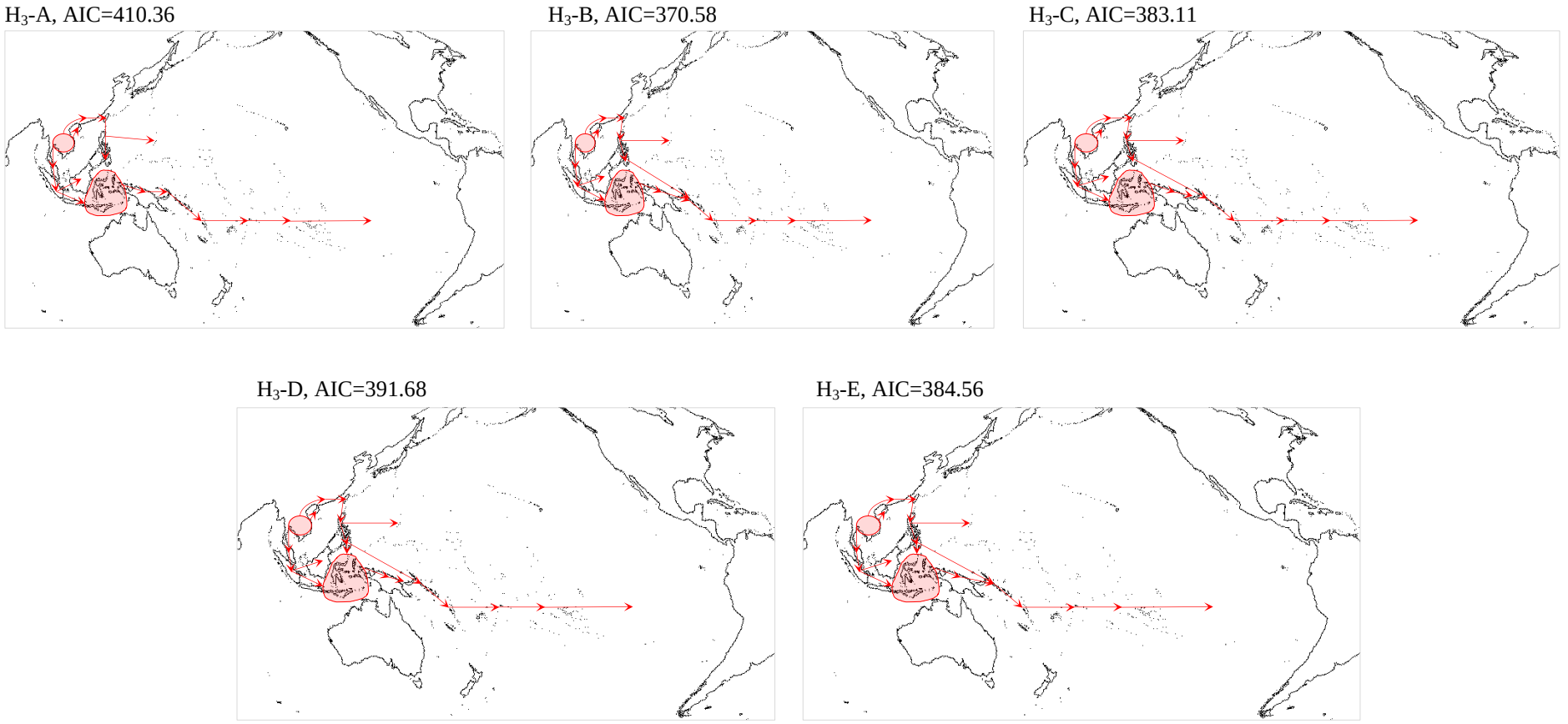
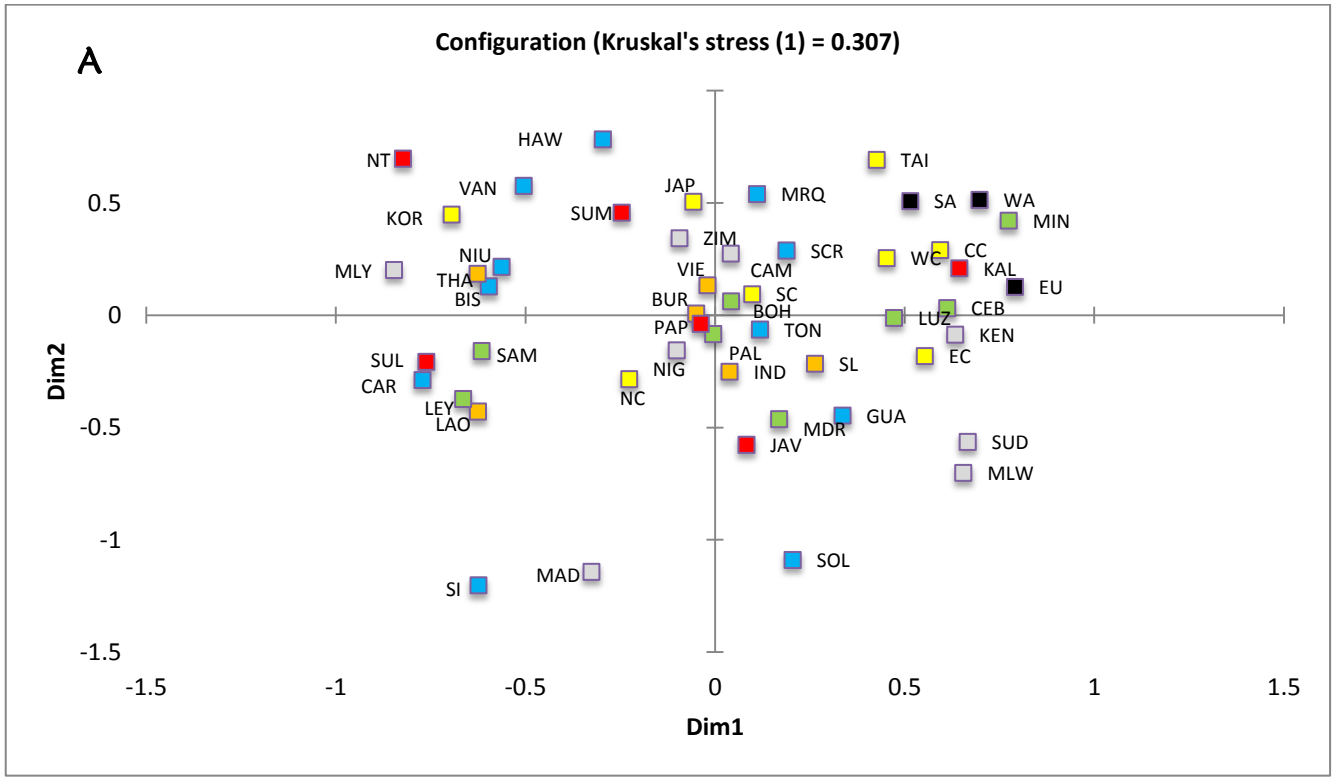


Figure S2-1. Serial Coalescent Simulation and Approximate Bayesian Computation models for the reconstruction chicken translocation history from MSEA through ISEA to the Pacific. Four major scenarios was modelled with the null hypothesis (H_0) as total panmixia (populations in MSEA, South China, ISEA, and the Pacific are considered as continuous), alternative hypotheses (H1) describes three scenarios of translocation of chickens from Asia through Taiwan and the Philippines before arriving the Pacific, (H2) describes two translocation scenarios from Asia directly to the Philippines or initially through Hainan then the Philippines before arriving the Pacific, and (H3) describes translocation scenarios allowing gene flow from Asia through Indonesia then the Pacific.


A
Configuration (Kruskal's stress (1) = 0.307)
HAW
NT
TAI
VAN
JAP
MRQ
SA
WA
MIN
KOR
SUM
ZIM
SCR
WC
CC
NIU
MLY
VIE
CAM
KAL
EU
THA
BIS
BUR
SC
CEB
BOH
LUZ
PAP
TON
KEN
SUL
SAM
PAL
EC
NIG
IND
SL
CAR
NC
LEY
GUA
LAO
MDR
SUD
JAV
MLW
SOL
MAD
SI
0.5
0
-0.5
-1
-1.5
-1.5
-1
-0.5
0
0.5
1
1.5
Dim2
Dim1

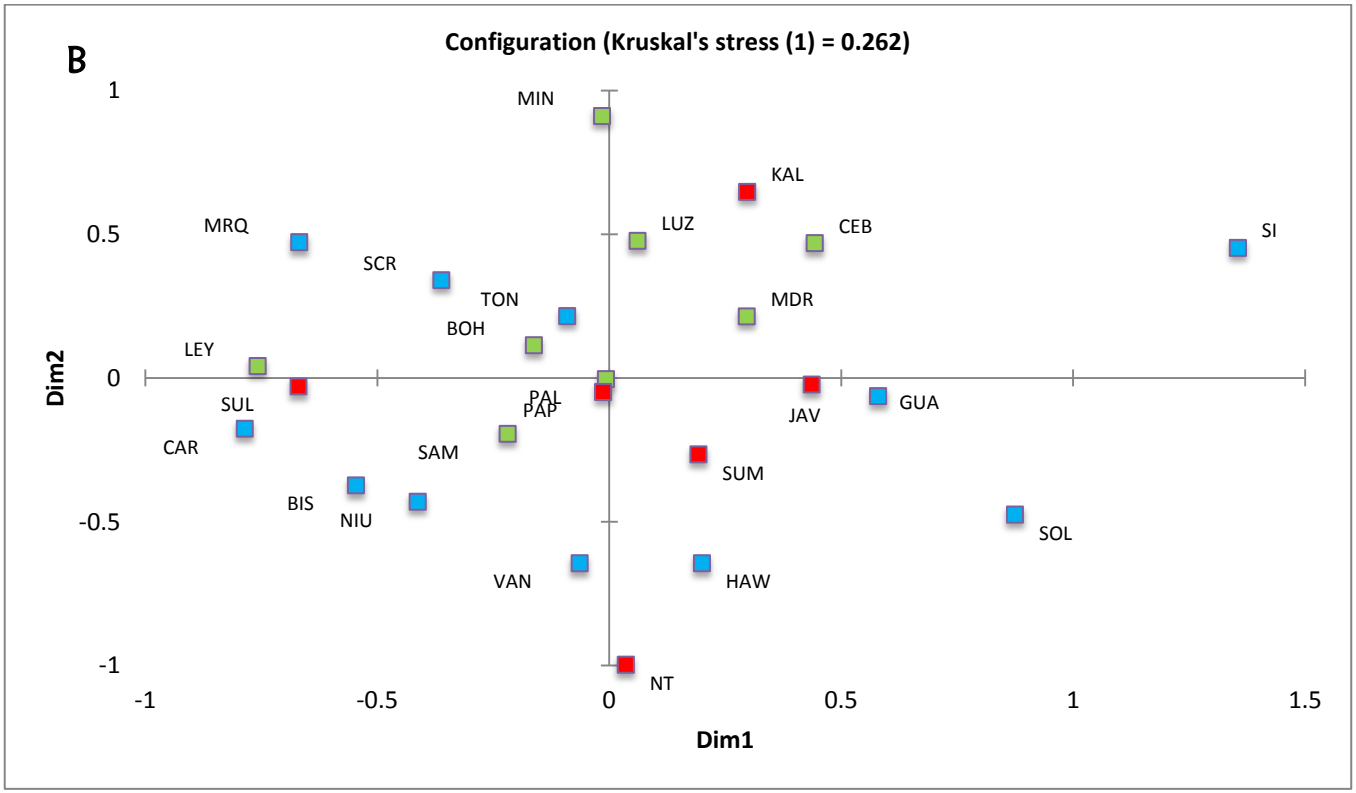


Figure S2-2. Multidimensional scaling plots (MDS) on population pairwise Fst for (a) world-wide chicken samples from haplogroup E only (red: Indonesia; green: Philippines; blue: Pacific; orange: MSEA and South Asia; yellow: East Asia; grey: Africa; white: Europe, West Asia, and South America. Populations are abbreviated as follows: BAT, Batanes; BOH, Bohol; CEB, Cebu; ILO, Iloilo; LEY, Leyte; LUZ, Luzon; MIN, Mindanao; PAL, Palawan; SAM, Samar; JAV, Java; KAL, Kalimantan; MLK, Maluku; NT, Nusa Tenggara; PAP, Papua; SUL; Sulawesi; SUM, Sumatra; BIS, Bismarck; CAR, Caroline Islands; EI, Easter Islands; FIJ, Fiji; GUA, Guam; HAW, Hawai'i; MRQ, Marquesas; NCL, New Caledonia; NUI, Nuie; SCR, Santa Cruz; SI, Society Islands; SOL, Solomon; TON, Tonga; VAN, Vanuatu; BUR, Burma; LAO, Laos; MLY, Malaysia; THA, Thailand; VIE, Vietnam; CC, Central China; EC, East China; NC, North China; SC, South China; HAI, Hainan; JAP, Japan; KOR, Korea; TAI, Taiwan; IND, India; SL, Sri Lanka; CAM, Cameroon; KEN, Kenya; MAD, Madagascar; MAL, Malawi; NIG, Nigeria; SUD, Sudan; ZIM, Zimbabwe; EU, Europe; SA, South America; and WA, West Asia. (b) Chicken samples from Indonesia, Philippines, and Pacific from haplogroup E only.

Table S2-1. Samples used in the study. (Excel File/CD-ROM)

Table S2-2. Population genetic summary statistics for the study populations worldwide.

LOCATIONS	*n*	H	HD (SD)	ND (SD)	Pi	Ts	Tv	Tajima 's D	p-value	Fu's FS	p-value
PHILIPPINES											
Batanes	8	4	0.750 (0.140)	0.04 (0.02)	7.50	15	0	1.52	0.96	3.31	0.94
Bohol	5	3	0.800 (0.164)	0.01 (0.01)	2.80	5	0	1.12	0.82	1.22	0.73
Cebu	6	4	0.800 (0.172)	0.013 (0.009)	2.60	6	0	-0.06	0.49	-0.02	0.40
Iloilo	5	3	0.700 (0.218)	0.006 (0.005)	1.20	3	0	-1.05	0.17	-0.19	0.25
Leyte	4	2	0.500 (0.265)	0.009 (0.008)	2.00	4	0	-0.78	0.20	2.20	0.82
Luzon	72	26	0.938 (0.013)	0.024 (0.013)	4.84	40	0	-1.34	0.07	-8.84	0.01
Mindanao	131	34	0.890 (0.017)	0.021 (0.011)	4.15	35	0	-1.06	0.12	-15.41	0.0
Mindoro	19	13	0.953 (0.031)	0.022 (0.012)	4.33	20	0	-0.94	0.17	-4.93	0.01
Palawan	43	17	0.907 (0.023)	0.017 (0.01)	3.51	21	7	-1.55	0.04	-5.72	0.02
Samar	10	9	0.978 (0.054)	0.028 (0.016)	5.53	19	1	-1.02	0.16	-3.26	0.04
INDONESIA											
Java	189	36	0.780 (0.022)	0.014 (0.008)	2.78	28	1	-1.26	0.06	-23.3	0.00
Kalimantan	50	14	0.857 (0.038)	0.011 (0.007)	2.22	19	0	-1.51	0.03	-5.14	0.02
Maluku	57	10	0.694 (0.054)	0.007 (0.005)	1.35	14	0	-1.64	0.02	-3.48	0.03
Nusa Tenggara	93	27	0.875 (0.026)	0.020 (0.011)	4.01	30	8	-1.45	0.06	-10.41	0.01
Papua	34	6	0.613 (0.078)	0.006 (0.004)	1.30	11	0	-1.62	0.03	-0.70	0.35
Sulawesi	78	22	0.792 (0.044)	0.010 (0.006)	2.04	18	1	-1.40	0.05	-14.5	0.0
Sumatra	191	34	0.811 (0.022)	0.018 (0.010)	3.69	33	7	-1.35	0.05	-14.32	0.002
PACIFIC											
Bismarck	13	2	0.154 (0.126)	0.004 (0.003)	0.77	5	0	-1.86	0.01	1.91	0.79
Caroline Islands	10	3	0.378 (0.181)	0.006 (0.005)	1.20	6	0	-1.8	0.01	0.84	0.68
Easter Island	16	2	0.125 (0.106)	x	x	x	x	x	x	x	x
Fiji	30	4	0.614 (0.061)	0.008 (0.006)	1.66	4	0	1.65	0.95	1.82	0.82
Guam	5	5	1.00 (0.127)	0.023 (0.016)	4.60	10	0	-0.30	0.48	-1.48	0.09
Hawaii	30	5	0.729 (0.053)	0.016 (0.009)	3.29	7	1	1.90	0.98	3.21	0.91
Marquesas	20	3	0.542 (0.076)	0.012 (0.008)	2.39	5	0	2.14	0.99	3.9	0.95
New Caledonia	2	1	x	x	x	x	x	x	x	x	x
Niue	3	2	0.667 (0.314)	0.023 (0.019)	4.67	7	0	0	0.74	2.88	0.86
Sta. Cruz	36	7	0.818 (0.036)	0.011 (0.007)	2.31	9	0	0.19	0.63	0.31	0.58
Society Islands	25	3	0.627 (0.061)	0.014 (0.008)	2.79	6	0	2.25	0.99	5.09	0.98
Solomon	32	7	0.778 (0.052)	0.014 (0.008)	2.81	15	0	-0.81	0.24	0.77	0.67
Tonga	2	1	x	x	x	x	x	x	x	x	x
Vanuatu	84	17	0.841 (0.028)	0.012 (0.007)	2.36	22	1	-1.47	0.05	-5.85	0.02
MSEA											
Burma	52	17	0.923 (0.015)	0.042 (0.022)	8.52	35	4	0.04	0.6	0.31	0.59
Laos	74	19	0.821 (0.034)	0.029 (0.015)	5.79	32	3	-0.54	0.32	-1.52	0.35
Malaysia	6	5	0.933 (0.122)	0.045 (0.028)	9.07	22	1	-0.63	0.36	0.66	0.54
Thailand	25	13	0.917 (0.031)	0.052 (0.027)	10.51	35	8	-0.29	0.44	0.45	0.61
Vietnam	204	54	0.892 (0.017)	0.036 (0.019)	7.19	56	8	-0.92	0.18	-21.41	0.00
EAST ASIA											
China, Central	472	58	0.879 (0.008)	0.031 (0.016)	6.14	40	3	0.02	0.59	-19.91	0.01
China, East	264	45	0.898 (0.012)	0.032 (0.017)	6.44	35	3	0.20	0.65	-11.73	0.01
China, North	108	22	0.868 (0.020)	0.032 (0.017)	6.53	27	1	0.67	0.81	-1.09	0.41
China, South	1292	134	0.926 (0.004)	0.038 (0.019)	7.59	67	9	-0.46	0.37	-23.85	0.004
China, West	302	43	0.893 (0.013)	0.032 (0.017)	6.37	38	4	-0.07	0.58	-9.20	0.06
Hainan	24	14	0.888 (0.057)	0.023 (0.132)	4.71	25	0	-1.10	0.13	-4.00	0.04
Japan	400	32	0.916 (0.006)	0.034 (0.018)	6.85	34	2	0.68	0.80	-1.09	0.47
Korea	31	6	0.783 (0.036)	0.029 (0.016)	5.91	16	0	1.60	0.97	5.21	0.96
Taiwan	193	8	0.762 (0.016)	0.022 (0.012)	4.51	17	0	1.44	0.93	7.10	0.95
SOUTH ASIA											
India	359	88	0.869 (0.015)	0.026 (0.014)	5.19	56	33	-1.77	0.006	-24.84	0.00
Sri Lanka	131	26	0.917 (0.010)	0.033 (0.017)	6.59	34	2	0.08	0.611	-2.29	0.28
AFRICA											
Cameroon	3	1	x	x	x	x	x	x	x	x	x
Kenya	159	21	0.795 (0.022)	0.021 (0.011)	4.19	21	0	0.35	0.71	-2.35	0.26
Madagascar	79	8	0.412 (0.067)	0.008 (0.005)	1.6	11	0	-0.76	0.25	-0.58	0.46
Malawi	19	2	0.651 (0.044)	0.017 (0.010)	3.47	7	0	2.92	1.00	7.13	0.99
Nigeria	38	31	0.981 (0.013)	0.012 (0.007)	2.41	25	5	-2.25	0.00	-26.7	0.0
Sudan	20	3	0.232 (0.064)	0.007 (0.005)	1.46	7	0	0.01	0.57	3.85	0.93
Zimbabwe	201	28	0.907 (0.005)	0.030 (0.016)	5.96	29	2	0.76	0.82	-2.17	0.34
EUROPE	116	10	0.565 (0.052)	0.008 (0.005)	1.60	15	0	-1.18	0.11	-1.54	0.26
S. AMERICA	48	11	0.692 (0.070)	0.016 (0.009)	3.31	19	1	-0.85	0.22	-0.67	0.43
WEST ASIA	26	6	0.602 (0.101)	0.019 (0.011)	3.85	12	0	0.75	0.82	2.39	0.88

H – number of haplotypes, HD – haplotype diversity, ND – nucleotide diversity, Ts – transitions, Tv – transversions, * - statistically sig nificant p-values ($p<0.05$ for Tajima's D, $p<0.02$ for Fu's FS)

CHAPTER 3: East African origins for Malagasy chickens as indicated by mitochondrial DNA

Statement of authorship

East African origins for Malagasy chickens as indicated by mitochondrial DNA

Michael James Bannister Herrera (Candidate)
Conceptualised and designed the study, performed the data processing, analysis, and interpretation, created figures and tables, and wrote the paper.

I hereby certify that the statement of contribution is accurate.

Signed _____ _____ Date 6 March 2015

Jeremy Austin
Assisted in the study design, gave advice on data interpretation, and edited the manuscript.

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 6 March 2015

Vicki Thomson
Gave advice on data interpretation, commented and edited the manuscript.

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 5 March 2015

Jessica Wadley
Gave advice on data interpretation, commented and edited the manuscript.

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 5 March 2015

Philip Piper
Assisted in the interpretation of results, commented and edited the manuscript

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 2 March 2015

Jaime Gongora
Liaised the Indonesian samples, commented and edited the manuscript

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 3 March 2015

East African origins for Malagasy chickens as indicated by mitochondrial DNA

Michael B. Herrera[A], Vicki A. Thomson[A], Philip J. Piper[B], Jessica J. Wadley[A], Jaime Gongora[C] and Jeremy A. Austin[A]

[A]Australian Centre for Ancient DNA, School Biological Sciences, University of Adelaide, South Australia 5005, Australia.

[B]School of Archaeology and Anthropology, Faculty of Arts, The Australian National University, Canberra ACT 0200, Australia.

[C]Centre for Advanced Technologies in Animal Genetics and Reproduction, Faculty of Veterinary Science, University of Sydney, Sydney NSW 2006, Australia

Abstract

The colonization of Madagascar by Austronesian-speaking populations between A.D. 50–500 represents the western most extension of the greatest global diaspora in prehistory that also resulted in human population movements from Taiwan into Island Southeast Asia and the Pacific Ocean. These human territorial expansions also entailed the translocation of a range of economically important plants and animals that have been useful in tracing the origins, timing and routes of migration of these populations. The, three major domesticates that were translocated were the pig (*Sus scrofa*), dog (*Canis lupus familiaris*) and chicken (*Gallus gallus*). Recent studies have shown that of these three, the chicken, with an origin in Island Southeast Asia, was the most widely transported across Oceania arriving in Melanesia with the early Lapita. What is still unclear though is whether the initial Austronesian-speaking populations that reached Madagascar via the Indian Ocean also transported chickens, which could be useful in tracking the westward migration of these Austronesian-speaking peoples that colonised Madagascar. To address this question, we investigated the mitochondrial control region diversity and connectivity of chicken populations around the Indian Ocean rim (Southeast Asia, South Asia, the Arabian Peninsula, East Africa, and Madagascar). In contrast to the linguistic and human genetic evidence indicating dual African and Southeast Asian ancestry, we observe that chickens on Madagascar share a common ancestor with East Africa, which together are genetically closer to South Asian chickens than to those in Southeast Asia. This suggests the earliest expansion of Austronesian-speaking people in the Indian Ocean is unlikely to have included chickens. In contrast with the linguistic evidence of a link between Madagascar and Island Southeast Asia, chickens appear to have arrived in Madagascar from East Africa, with the East African populations

derived from maritime exchange with South Asia. Our results further demonstrate the complexity of the translocation history that created the current biodiversity in Madagascar.

Keywords

Austronesian, Madagascar, Indian Ocean, control region, migration

Introduction

Beginning in the first few centuries A.D. Austronesian speakers from Island Southeast Asia (ISEA) established trade links with India and eventually colonised Madagascar between ca 50 - 500 A.D. (Dewar & Wright 1993; Burney *et al.* 2004; Bellwood 2007). This makes Madagascar the most westerly point of the great Austronesian expansion. Linguistic (Beaujard 2003) and genetic (Hurles *et al*. 2005) evidence suggests a dual ancestry for the indigenous people of Madagascar, involving African and South-east Asian origins. For example, Malagasy, the language spoken in Madagascar, is a member of the Austronesian language family related to the Barito and Dayak languages spoken in south-east Kalimantan (Dahl 1991). However, recent genetic studies indicate that Malagasy populations are derived from genetic admixture involving Indonesian and African ancestors (*i.e.*, Bantu) (Tofanelli *et al.* 2009; Pierron *et al*. 2014). In addition to genetic and linguistic evidence, transfers of material culture are also evident in the Austronesian inherited traditions connecting Madagascar to Indonesia (Blench 2010). However, Madagascar had numerous maritime connections with regions around the Indian Ocean, the legacy of which included the translocation of domestic and commensal animals, as well as plants (Blench & Dendo 2006; Fuller *et al.* 2011).

Prehistoric exchanges have influenced the current day distribution of domestic plants and animals around the Indian Ocean rim (Boivin *et al*. 2013). The most notable Madagascan domesticates that have originated from Island Southeast Asia (ISEA) include taro (*Colocasia esculenta*), Asian yam (*Dioscorea alata*), and banana (*Musa sapientum*) (Fuller *et al.* 2011). Their absence in the intervening regions of north India and the Arabian Peninsula makes a translocation via a central Indian Ocean maritime corridor (Fuller *et al.* 2011) more likely than a coastal route (Murdock 1959). However, genetic studies of ship-borne commensals found in Madagascar, such as rats (*Rattus sp*.), mouse (*Mus musculus*), and shrew (*Suncus murinus*), suggests India as the point of origin (Hingston *et al.* 2005; Kurachi *et al.* 2007; Tollenaere *et al.* 2010). An investigation of diversity in domesticate/commensal species and their translocation patterns around the Indian Ocean rim suggests a deeply entrenched trade and contact network (Boivin *et al.* 2013). For instance, exchanges in the Arabian Sea led to the movement of certain cereal crops, such as sorghum (*Sorghum bicolor*), pearl millet (*Pennisetum galucum*), and finger millet (*Eleusine coracana*), from Africa to the Arabian Peninsula and India (Fuller & Boivin 2009), while Zebu cattle were translocated from India to the Arabian Peninsula and Africa (Hanotte *et al*. 2002). Translocations across the Bengal Sea included the movements of mung bean (*Vigna raidata*) and horsegram (*Macrotyloma uniflorum*) from India to Southeast Asia (Castillo & Fuller 2010) and in reverse mango (*Mangifera indica*) and citron (*Citrus medica*) from Southeast Asia to India (Asouti 2008). The protracted connection between India and the Austronesian-speakers in Indonesia probably culminated in the development of the Srivijayan Empire in Indonesia and the Malay Peninsula (Munoz 2006). The

Srivijayan religion, culture, and even language contains a mixture of influences not only from India but from Persia as well (Hall 1977). These instances demonstrate that there has been a long history of contact, trade, and exchange between geographically distant groups of people around the Indian Ocean rim.

As chickens are not native to Africa and Madagascar, they must ultimately derive from South and Southeast Asia, which are the natural biogeographic distribution of jungle fowls. Similar to banana, taro and yam, chickens are deeply integrated into the subsistence culture of Africa indicating a certain level of antiquity (Williamson 2000), with chickens first appearing in Madagascar around the late 8th-mid 9th century A.D. (Mudida & Horton 1996). However, the routes and timing of chicken introductions to Madagascar are unclear. Archaeological evidence suggests that chickens were introduced into Africa via trade links between East Africa and Southeast Asia (MacDonald 1992). Furthermore, on Madagascar, the Malagasy term for chicken is borrowed from Bantu languages (from the east coast of Africa), not Austronesian (Blench 2010), suggesting that chickens were initially established on the east coast of Africa before being introduced into Madagascar.

Previous studies have characterised partial mtDNA control region sequences in African village chickens, which fell into two major mitochondrial lineages (referred to as haplogroup D and E) and suggest two origins: 1) in Southeast Asia; and 2) the Indian subcontinent (Muchadeyi *et al.* 2008; Mtileni *et al.* 2011; Mwacharo *et al*. 2011). Chickens from Madagascar have been reported to belong to two major mitochondrial lineages (Razafindraibe *et al.* 2008). Although Razafindraibe *et al.* (2008) suggested that this was evidence for a dual geographic

origin for Malagasy chickens, from continental Africa and Indonesia; they provided no analysis to support their conclusions.

Thus, the geographic origin of Madagascan chickens remains uncertain, with Africa, South Asia and/or ISEA as possible sources. Here, we assess the genetic relationships and diversity of chicken populations found around the Indian Ocean rim. We aim to address the question of whether indigenous chickens in Madagascar trace their ancestry to Indonesia as reflected in the expansion of the Austronesian culture, or they resulted from a more complex connection with other areas around the Indian Ocean.

Materials and Methods

Samples analysed in this study were previously described elsewhere (Table S3-1). In total, 3115 chickens from Madagascar, Africa, the Arabian Peninsula, West Asia, East Asia, South Asia, ISEA, Mainland SEA (MSEA), and the Pacific were used for analyses. DNA sequences consisted of variable lengths of the mtDNA control region. Sequences were aligned using the MUSCLE algorithm in Geneious v.6.5.5 and truncated to 349 bp, the longest sequence common to all samples. The number of haplotypes and the number of samples per haplotype were determined by collapsing the sequences to unique haplotypes using FaBox 1.40 (Villesen 2007). Identifying the mitochondrial haplogroup of each unique sequence was done by ordering them into a neighbour-joining (NJ) tree generated using the Tamura-Nei substitution model and comparing the assignments to those from previously published papers (Liu *et al.* 2006; Thomson *et al.* 2014).

To assess population genetic differentiation and gene flow among chicken populations in East Africa, Madagascar, the Arabian Peninsula, East Asia, South Asia, Southeast Asia, and the Pacific, population pairwise F_{ST} scores were computed using Arlequin v.3.1 (Excoffier *et al.* 2005). To visualise the relationships between populations, a non-parametric multidimensional scaling plot (MDS) was performed using the pairwise F_{ST} scores. To further explore geographic structure, an analysis of molecular variance (AMOVA) was also calculated in Arlequin (Excoffier 2005) using Indonesia, South Asia, East Africa, and Madagascar as groups. These regions were selected for AMOVA because they are the most likely regions involved in the translocation of chickens to Madagascar. Furthermore, the extensive linguistic and human genetic studies mostly involve these regions. A substantial proportion of haplotypes found on Madagascar and ISEA belong to mitochondrial haplogroup D. Thus, a separate set of analyses were performed on a dataset containing only haplogroup D lineages.

The evolutionary relationships among haplotypes was estimated via a median-joining (MJ) network (Bandelt *et al.* 1999) using NETWORK v.4.6.1 (fluxus-engineering.com) for haplogroup D and E haplotypes. Pairwise genetic distances between haplotypes was calculated using GENALEX (Peakall & Smouse 2006) and ordered into a principal coordinate analysis (PCoA) plot to further explore the relationships of the lineages and to see how the haplotypes around the Indian Ocean rim are distributed. Intra-population genetic variation (*i.e.,* haplotype and nucleotide diversity) for populations at a regional level were calculated using Arlequin 3.1 (Excoffier *et al.* 2005). To understand the historical demography of

populations, expansion statistics were also calculated (*i.e.*, Tajima's D and Fu's FS) using Arlequin 3.1 (Excoffier *et al.* 2005).

Results

Mitochondrial haplogroup distribution patterns

Chickens on Madagascar belong to only two haplogroups - with the majority (85%) of samples belonging to haplogroup D and the rest belonging to haplogroup E (Figure 3-1). Similarly, in east African chickens only haplogroup D and E are observed with increasing haplogroup D frequency with increasing latitude southwards. In contrast, haplogroup D is not observed in chickens from the Arabian Peninsula and western Asia: the haplogroup composition in these regions is dominated by haplogroup E. Within India, haplogroup E chickens are also observed at a high level (67%) while haplogroup D is next highest (22%) and all other haplogroups make up the balance (11%). To the east in MSEA, all haplogroups (A to I) are observed, but the frequency of haplogroup D and E are dramatically lower in comparison to other haplogroups (4.8% and 2.4% respectively). However, chickens from islands further east (in the Pacific Ocean) have a high proportion of haplogroup D (84%).

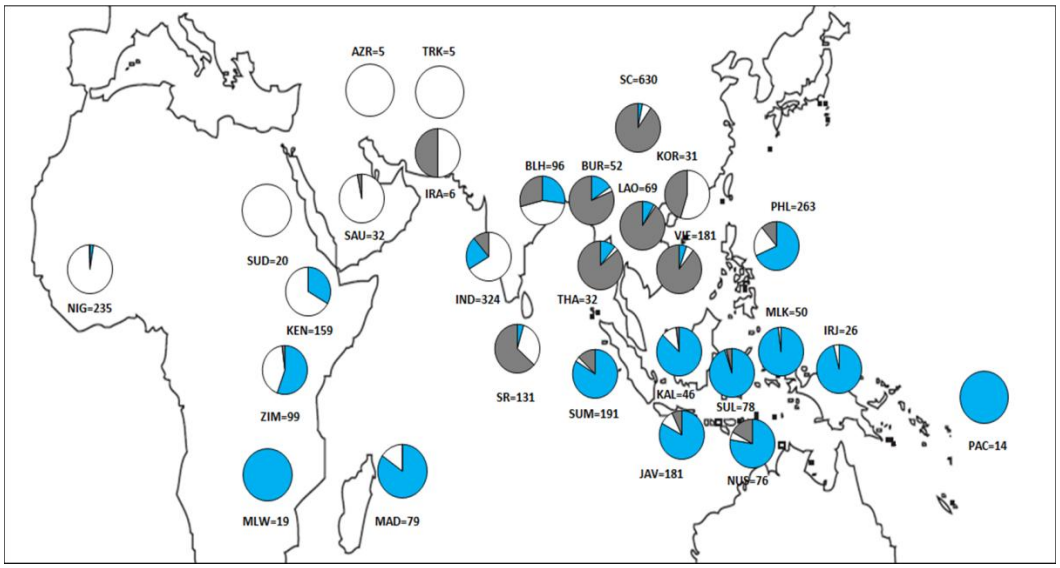


Figure 3-1. Frequency distribution of chicken mitochondrial DNA haplogroup (*blue* – haplogroup D, *white* – haplogroup E, and grey – other haplogroups) by geographic location followed by sample size. Sample localities are Azerbaijan (AZR), Bangladesh (BLH), Burma (BUR), India (IND), Iran (IRA), Irian Jaya (IRJ), Java (JAV), Kalimantan (KAL), Kenya (KEN), Korea (KOR), Laos (LAO), Madagascar (MAD), Malawi (MLW), Maluku (MLK), Nigeria (NIG), Nusa Tenggara (NUS), Pacific (PAC; Fiji, Solomon, and Vanuatu), Philippines (PHL; Luzon, Visayas, and Mindanao), Saudi Arabia (SAU), South China (SC), Sri Lanka (SRI), Sudan (SUD), Sulawesi (SUL), Sumatra (SUM), Thailand (THA), Turkmenistan (TRK), Vietnam (VIE), and Zimbabwe (ZIM).

Population genetic structure

The overall genetic structure in the MDS plot using all haplogroups reveals structuring at a broad geographic scale (Figure 3-2A). Distinctive clustering of populations from ISEA (shown in green) and from the Pacific (in blue) can be seen. African populations, including Madagascar, sit closer to South and West Asian populations, than ISEA populations. Madagascar falls closest to Malawi, which is the closest African population to Madagascar. The other continental African populations show broad geographic clines (northern-southern cline follows high-low 'Dim 2' values). When only haplogroup D is used the Madagascan samples form a distinct cluster with the geographically closest East African populations, Malawi and Zimbabwe (Figure 3-2B).

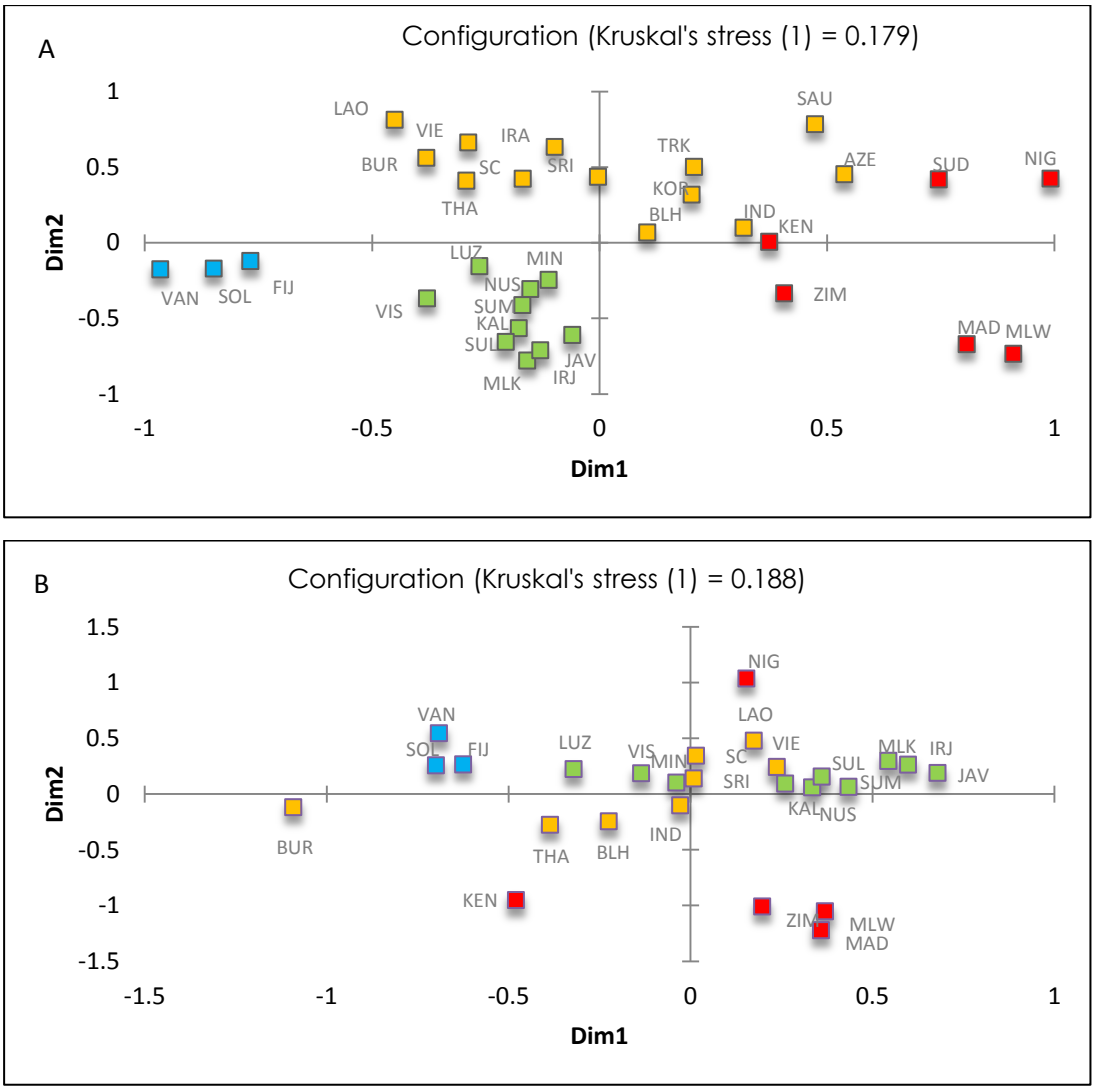


Figure 3-2. Multidimensional scaling plots (MDS) on population pairwise F_{ST} scores for (A) 3115 chickens from Asia (*Orange*), Africa (*Red*), Island Southeast Asia (*Green*), and the Pacific (*Blue*) using all haplogroups. (B) 1081 haplogroup D chickens from the same regions. Azerbaijan, Iran, Korea, Saudi Arabia, Sudan, and Turkmenistan are not present in plot B because they do not contain haplogroup D lineages. See Figure 1 for abbreviations.

The AMOVA performed on chickens from Indonesia, South Asia, East Africa, and Madagascar show population structure (Table 3-1). However, the among group variance components are high (*i.e.*, >50 %) only when haplogroup D chickens are used. For haplogroup D samples, the among-group variance component (-1.98 %) is only not significant when Indonesia, East Africa, and Madagascar are grouped together *versus* South Asia. All other comparisons show a high among group

variance. The variance components of the regional grouping are generally low when using only haplogroup E.

Table 3-1. Population genetic structure estimated from the analysis of molecular variance (AMOVA) based on mtDNA control region sequences from relevant regions in the Indian Ocean rim: (A) Indonesia, (B) South Asia, (C) East Africa, and (D) Madagascar.

Group	n	No. of population	No. of groups	Variance components (%)		
				Among groups	Among populations within groups	Within Populations
Haplogroup D and E combined						
No grouping	1347	14	1	…	31.07	68.93
Group 1 (A vs. B vs. C vs. D)	1347	14	4	30.38	5.75	63.87
Group 2 (A, C, & D vs. B)	1347	14	2	17.87	19.66	62.46
Group 3 (A vs. B, C, & D)	1347	14	2	23.58	14.77	61.64
Group 4 (A, B, vs. C, D)	1347	14	2	11.38	24.11	64.51
Haplogroup D only						
No grouping	845	14	1	…	48.00	52.00
Group 1 (A vs. B vs. C vs. D)	845	14	4	51.98	6.95	41.07
Group 2 (A, C, & D vs. B)	845	14	2	*-1.98*	49.19	52.79
Group 3 (A vs. B, C, & D)	845	14	2	44.25	15.07	40.68
Group 4 (A, B, vs. C, D)	845	14	2	56.42	9.30	34.28
Haplogroup E only						
No grouping	502	10	1	…	9.86	90.14
Group 1 (A vs. B vs. C vs. D)	502	10	4	4.88	6.26	88.86
Group 2 (A, C, & D vs. B)	502	10	2	1.80	8.65	89.55
Group 3 (A vs. B, C, & D)	502	10	2	4.78	8.72	86.50
Group 4 (A, B, vs. C, D)	502	10	2	0.49	9.56	89.96

Population genetic variability and dynamics

The genetic differentiation observed from the population pairwise F_{ST} values (Table 3-2) using both haplogroup D and E samples show that the highest level of divergence is between Indonesia and Madagascar (1.33034), whereas it is the lowest between South Asia and Africa (0.14823). When using only haplogroup D samples, the highest level of divergence is still between Indonesia and Madagascar (0.69607)

and the lowest is between Africa and Madagascar (0.13317). All population pairwise comparisons were significant at the 5 % level.

Table 3-2. Population pairwise (F_{ST}) between chicken samples from Indonesia, South Asia, East Africa, and Madagascar based on mitochondrial control region.

Population	Abbreviation	SA	INDO	AFR	MAD
		Haplogroup D & E combined			
South Asia	**SA**	0			
Indonesia	**INDO**	0.60650*	0		
Africa	**AFR**	0.14823*	0.50893*	0	
Madagascar	**MAD**	0.79735*	1.33034*	0.29177*	0
		***Haplogroup* D**			
South Asia	**SA**	0			
Indonesia	**INDO**	0.28224*	0		
Africa	**AFR**	0.38310*	0.66755*	0	
Madagascar	**MAD**	0.38160*	0.69607*	0.13317*	0

*Significant differences at P<0.05

Both Tajima's D and Fu's FS neutrality statistics indicate that chickens from South Asia, Indonesia, East Africa, and Madagascar deviate from neutrality when using only haplogroup D (Table 3-3). However, when using both haplogroup D and E all consistently deviate from neutrality except Africa. These results support a model of demographic expansion of haplogroup D for each of the four regions.

Table 3-3. Genetic diversity measures and historical demographic patterns of chickens from Indonesia, South Asia, East Africa, and Madagascar.

Region	N(H)	Molecular Diversity Indices			Neutrality Test	
		HD (SD)	ND (SD)	π	Tajima's D	Fu's F_S
		Haplogroup D & E combined				
South Asia	409 (114)	0.9141 ± 0.0109	0.0124(0.0068)	4.31	-1.97*	-25.10*
Indonesia	583 (115)	0.9519 ± 0.0036	0.0047 (0.0031)	1.63	-2.01*	-27.13*
Africa	276 (42)	0.9111 ± 0.0075	0.0123 (0.0068)	4.30	-0.03	-17.36*
Madagascar	79 (10)	0.4340 ± 0.0680	0.0048 (0.0032)	1.68	-1.15	-1.88
		D haplogroup only				
South Asia	100 (40)	0.9166 (0.0186)	0.0118 (0.0066)	4.13	-1.67*	-25.74*
Indonesia	551 (102)	0.9464 (0.0039)	0.0038 (0.0026)	1.32	-2.14*	-28.85*
Africa	127 (19)	0.7768 (0.0242)	0.0041 (0.0028)	1.41	-1.50*	-11.94*
Madagascar	67 (6)	0.2243 (0.0675)	0.0008 (0.0009)	0.27	-1.90*	-5.32*

N(H) – size (haplotypes #), HD(SD) – haplotype diversity (standard deviation), ND – nucleotide diversity, π – mean # of pairwise difference , SSD – sum of squared differences, * - statistically significant p-values ($p<0.05$ for Tajima's D, $p<0.02$ for Fu's FS)

Phylogenetic relationships of Madagascan mtDNA haplotypes in East Africa, South Asia and Indonesia

The median-joining phylogenetic network created using only haplogroup D samples show that Madagascar and East Africa share a closely related set of haplotypes and together these two regions have a closer phylogenetic relationship with South Asia than with Indonesia (Figure 3-3A). Altogether, there are six Madagascan D haplotypes: H45 and H36 are shared with East Africa and the rest are unique to Madagascar (H16, H40, H41, and H42) (Figure 3-3B). H45 is the most common D haplotype in East Africa and Madagascar and forms the central node from which the rest of the Madagascan and East African haplotypes radiate. Two predominantly Indonesian haplotypes (H65 and H74, Fig 3B) are observed at very low frequencies in continental Africa, but not at all in Madagascar and in fact they are phylogenetically distant to the other Madagascan D haplotypes. The PCoA plot (Figure S3-1) that used the genetic distances between all 112 D haplotypes from Madagascar, East Africa, South Asia, and Indonesia also supports the broad separation of regions indicated in the network. Additionally, there are four haplotypes belonging to haplogroup E found on Madagascar. One haplotype is

unique to Madagascar, one is shared with East Africa and two are shared with East Africa, South Asia and Indonesia (Figure S3-2).

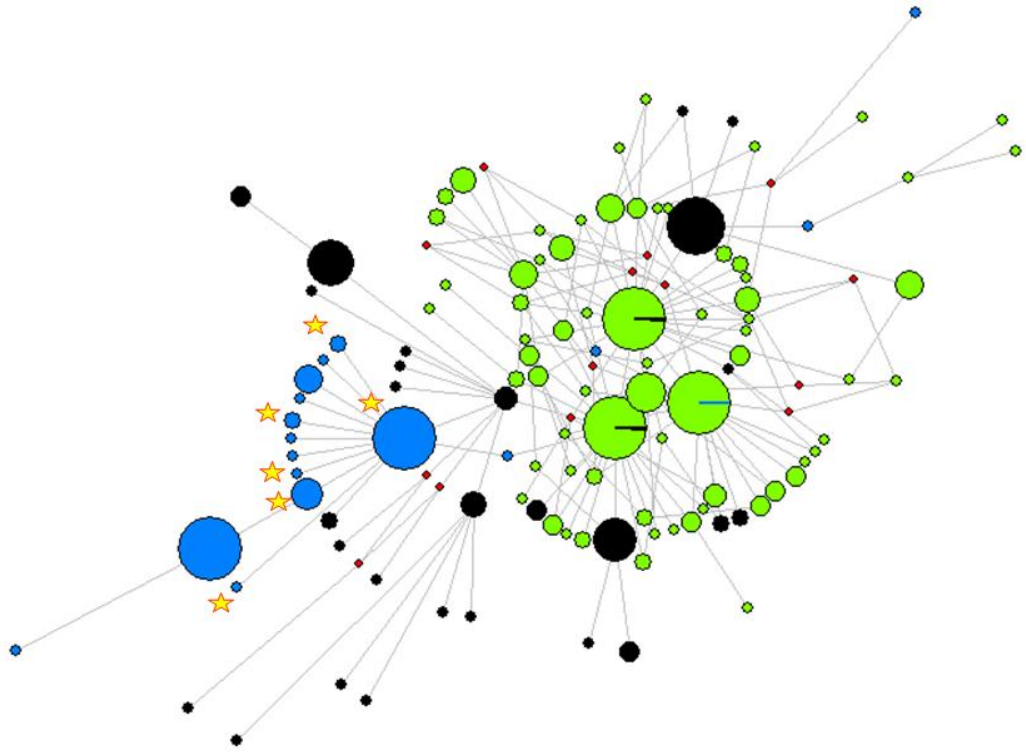


Figure 3-3A. Median-joining network depicting the relationship of D haplotypes of chickens from East Africa and Madagascar (*blue*), South Asia (*black*) and Indonesia (*green*) using all haplotypes regardless of frequency. Stars mark Madagascan samples. Inferred haplotypes are indicated by small red dots.

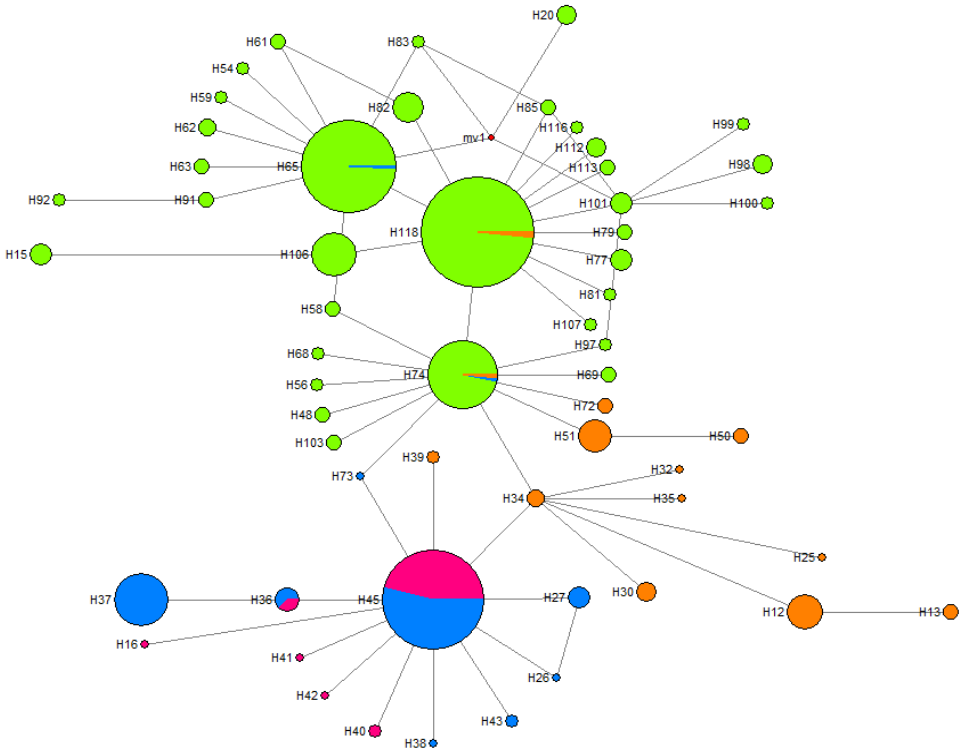


Figure 3-3B. Median-joining (MJ) network of mtDNA-CR D haplotypes observed in Africa (*blue*), South Asia (*brown*), Indonesia (*green*), and Madagascar (*purple*) excluding most haplotypes represented by 1 sample. The circle sizes are proportional to the haplotype frequencies, and the length of the lines corresponds to the number of mutations connecting haplotypes.

Discussion

Our population genetic and phylogenetic analyses of >3000 chicken mtDNA sequences from around the Indian Ocean rim strongly support an African origin for Madagascan chickens. This contrasts with a previous study of Madagascan chickens that suggested a combined African and Southeast Asian origins (Razafindraibe *et al.* 2008), likely due to the increased number of samples available to our study.

The presence of only two haplogroups, D and E, in Madagascar and East Africa suggests that these were the only lineages translocated to these regions prehistorically, with haplogroup D the dominant lineage in both Madagascar and East Africa. Despite a long and complex history of maritime exchange around the

Indian Ocean rim a strong phylogeographic signal (Figure 3-2A, B) remains. The phylogeography indicates a shared common ancestry for Madagascan and East African chickens. The star-like radiation stemming from the most common Madagascar haplotype (H45, which is also the most common east African haplotype; Figure 3-3B) and the fact that the Malagasy term for chicken comes from the African Bantu language suggests that the initial chicken populations arrived in Madagascar via the East African coast. What is not known, however, is whether H45 represents the founding lineage, with *in-situ* evolution responsible for the one and two base pair derivations radiating from H45. As H45 haplotype is currently not observed outside of East Africa/Madagascar and there are no archaeological chicken remains representing the earliest chickens on Madagascar, we cannot tell where H45 originated.

The strong phylogeographic signal within haplogroup D suggests a South Asian, rather than Southeast Asian, source for East African/Madagascan chickens (Figure 3-3A, B). Therefore, despite the clear linguistic and human genetic associations between Madagascar and Indonesia, the Austronesians do not appear to have successfully translocated chickens to Madagascar. If they have, the translocated chickens may have died out quickly and not left a genetic signal in modern chicken population. Thus, modern chickens in Madagascar are more likely to be South Asian in origin brought either by other cultures, or by Austronesians that came via the Indian subcontinent rather than directly across the Indian Ocean. Support for the latter theory can be inferred from the presence of Austronesian-speakers in South Asia during the first millennium A.D. (Hall 1977; Mahdi 1999).

It is interesting that haplogroup D chickens are not observed in the Arabian Peninsula and occur at low frequencies in northeast Africa, suggesting that chickens might have been transported via a direct sea link from India across the Arabian Sea to eastern Africa/Madagascar. Alternatively, Madagascan chickens could also have been transported along a coastal route through the Arabian Peninsula and northeast Africa but with the signal overwritten by subsequent and repeated translocation of haplogroup E chickens.

Haplogroup E is less common in Madagascar compared to haplogroup D. As haplogroup E lacks phylogeographic signature most likely due to modern day translocation (Chapter 2, but see Figure S3-3), it is difficult to derive fine-scale inferences based on haplogroup E other than establishing that it is also most likely South Asian in origin. Furthermore, it is difficult to ascertain whether the arrival of haplogroups D and E was contemporaneous. However, the most parsimonious explanation is that a mixed population of both haplogroup D and E chickens were transported from India to Madagascar via east Africa. Testing this hypothesis is difficult using existing samples. A more deliberate sampling regime, combined with nuclear genetic data and ancient DNA from archaeological samples, would help establish the full history of chicken translocation around the Indian Ocean rim.

Conclusion

Mitochondrial DNA data suggests chickens were introduced into Madagascar from South Asia via east Africa. The translocation of chickens from South Asia to the east coast of Africa and Madagascar might have occurred through direct sea links between the regions. A scenario whereby chickens arrived in Madagascar along with

the expansion of the Austronesian-speaking people directly across the India Ocean is not supported. However, it remains a possibility that Austronesian traders and mariners integrated South Asian chickens during their voyages *en route* to east Africa and Madagascar. The hypothesis presented here can be further tested by establishing an extensive and intensive sampling regime, by genetic investigation of securely dated archaeological chickens, or by using genome-scale studies of modern chickens along the Indian Ocean rim.

Acknowledgement

This research was funded by the Australian Research Council (ARC) Discovery Project ID DP110105187. We thanked Hanta Razafindraibe for making the Madagascan chicken sequences available to this study.

References

Asouti, E. 2008. Trees and woodlands of South India: Archaeological perspectives: Left Coast Press.

Bandelt, H-J, P Forster, A Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular biology and evolution 16:37-48.

Beaujard, P. 2003. Les arrivées austronésiennes à madagascar: Vagues ou continuum?(partie I). Études Océan Indien:59-147.

Bellwood, P. 2007. Prehistory of the Indo-Malaysian Archipelago. Canberra, Australia: Australian National University Press.

Blench, R. 2010. Evidence for the Austronesian voyages in the Indian Ocean. The global origins and development of seafaring:239-248.

Blench, R, M Dendo. 2006. The Austronesians in Madagascar and on the East African coast: surveying the linguistic evidence for domestic and translocated animals. International Conference on Austronesian Languages Palawan: 17-20 January 2006.

Boivin, N, A Crowther, R Helm, DQ Fuller. 2013. East Africa and Madagascar in the Indian Ocean world. Journal of world prehistory 26:213-281.

Burney, DA, LP Burney, LR Godfrey, WL Jungers, SM Goodman, HT Wright, A Jull. 2004. A chronology for late prehistoric Madagascar. Journal of Human Evolution 47:25-63.

Castillo, C, D Fuller. 2010. Still too fragmentary and dependent upon chance? Advances in the study of early Southeast Asian archaeobotany.

Dahl, OC. 1991. Migration from Kalimantan to Madagascar: Norwegian University Press Oslo.

Excoffier, L, G Laval, S Schneider. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evolutionary bioinformatics online 1:47.

Fuller, DQ, N Boivin. 2009. Crops, cattle and commensals across the Indian Ocean. Current and Potential Archaeobiological Evidence. Études Océan Indien:13-46.

Fuller, DQ, N Boivin, T Hoogervorst, RG Allaby. 2011. Across the Indian Ocean: the prehistoric movement of plants and animals. Antiquity 85:544-558.

Hall, K. 1977. The coming of Islam to the Archipelago. Economic Change and Social Interaction in Southeast Asia, Ann Arbor, Center for South and Southeast Asian Studies, University of Michigan.

Hanotte, O, DG Bradley, JW Ochieng, Y Verjee, EW Hill, JEO Rege. 2002. African pastoralism: genetic imprints of origins and migrations. Science 296:336-339.

Hingston, M, SM Goodman, JU Ganzhorn, S Sommer. 2005. Reconstruction of the colonization of southern Madagascar by introduced Rattus rattus. Journal of Biogeography 32:1549-1559.

Hurles, ME, BC Sykes, MA Jobling, P Forster. 2005. The dual origin of the Malagasy in Island Southeast Asia and East Africa: evidence from maternal and paternal lineages. The American Journal of Human Genetics 76:894-901.

Kurachi, M, B-L Chau, V-B Dang, T Dorji, Y Yamamoto, MM Nyunt, Y Maeda, L Chhum-Phith, T Namikawa, T Yamagata. 2007. Population structure of wild musk shrews (Suncus murinus) in Asia based on mitochondrial DNA variation, with research in Cambodia and Bhutan. Biochemical genetics 45:165-183.

Liu, Y-P, G-S Wu, Y-G Yao, Y-W Miao, G Liukart, M Baig, A Beja-Pereira, Z-L Ding, MG Palanichamy, Y-P Zhang. 2006. Multiple maternal origins of chickens: Out of the Asian jungles. Molecular phylogenetics and evolution 38:12-19.

MacDonald, KC. 1992. The domestic chicken (*Gallus gallus*) in sub-Saharan Africa: A background to its introduction and its osteological differentiation from indigenous fowls (*Numidinae* and *Francolinus sp*.). Journal of Archaeological Science 19:303-318.

Mahdi, W. 1999. The dispersal of Austronesian boat forms in the Indian Ocean. Archaeology and Language 3:144-179.

Mtileni, B, F Muchadeyi, A Maiwashe, E Groeneveld, L Groeneveld, K Dzama, S Weigend. 2011. Genetic diversity and conservation of South African indigenous chicken populations. Journal of Animal Breeding and Genetics 128:209-218.

Muchadeyi, F, H Eding, H Simianer, C Wollny, E Groeneveld, S Weigend. 2008. Mitochondrial DNA D-loop sequences suggest a Southeast Asian and Indian origin of Zimbabwean village chickens. Animal genetics 39:615-622.

Mudida, N, M Horton. 1996. Subsistence at Shanga: the faunal record. Shanga: The archaeology of a muslim trading community on the coast of East Africa:378-393.

Munoz, PM. 2006. Early kingdoms of the Indonesian Archipelago and the Malay peninsula: Editions Didier Millet.

Murdock, GP. 1959. Africa: its peoples and their culture history.

Mwacharo, J, G Bjørnstad, V Mobegi, K Nomura, H Hanada, T Amano, H Jianlin, O Hanotte. 2011. Mitochondrial DNA reveals multiple introductions of domestic chicken in East Africa. Molecular phylogenetics and evolution 58:374-382.

Peakall, R, PE Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6:288-295.

Pierron, D, H Razafindrazaka, L Pagani, F-X Ricaut, T Antao, M Capredon, C Sambo, C Radimilahy, J-A Rakotoarisoa, RM Blench. 2014. Genome-wide evidence of Austronesian–Bantu admixture and cultural reversion in a hunter-gatherer group of Madagascar. Proceedings of the National Academy of Sciences 111:936-941.

Razafindraibe, H, VA Mobegi, SC Ommeh, G Bjørnstad, O Hanotte, H Jianlin. 2008. Mitochondrial DNA origin of indigenous malagasy chicken. Annals of the New York Academy of Sciences 1149:77-79.

Thomson, VA, O Lebrasseur, JJ Austin, TL Hunt, DA Burney, T Denham, NJ Rawlence, JR Wood, J Gongora, LG Flink. 2014. Using ancient DNA to study the origins and dispersal of ancestral Polynesian chickens across the Pacific. Proceedings of the National Academy of Sciences 111:4826-4831.

Tofanelli, S, S Bertoncini, L Castrì, D Luiselli, F Calafell, G Donati, G Paoli. 2009. On the origins and admixture of Malagasy: new evidence from high-resolution analyses of paternal and maternal lineages. Molecular biology and evolution 26:2109-2124.

Tollenaere, C, C Brouat, JM Duplantier, L Rahalison, S Rahelinirina, M Pascal, H Moné, G Mouahid, H Leirs, JF Cosson. 2010. Phylogeography of the introduced species Rattus rattus in the western Indian Ocean, with special emphasis on the colonization history of Madagascar. Journal of Biogeography 37:398-410.

Villesen, P. 2007. FaBox: an online toolbox for fasta sequences. Molecular Ecology Notes 7:965-968.

Williamson, K. 2000. Did chickens go west. The origins and development of African livestock:368-448.

Supplementary information

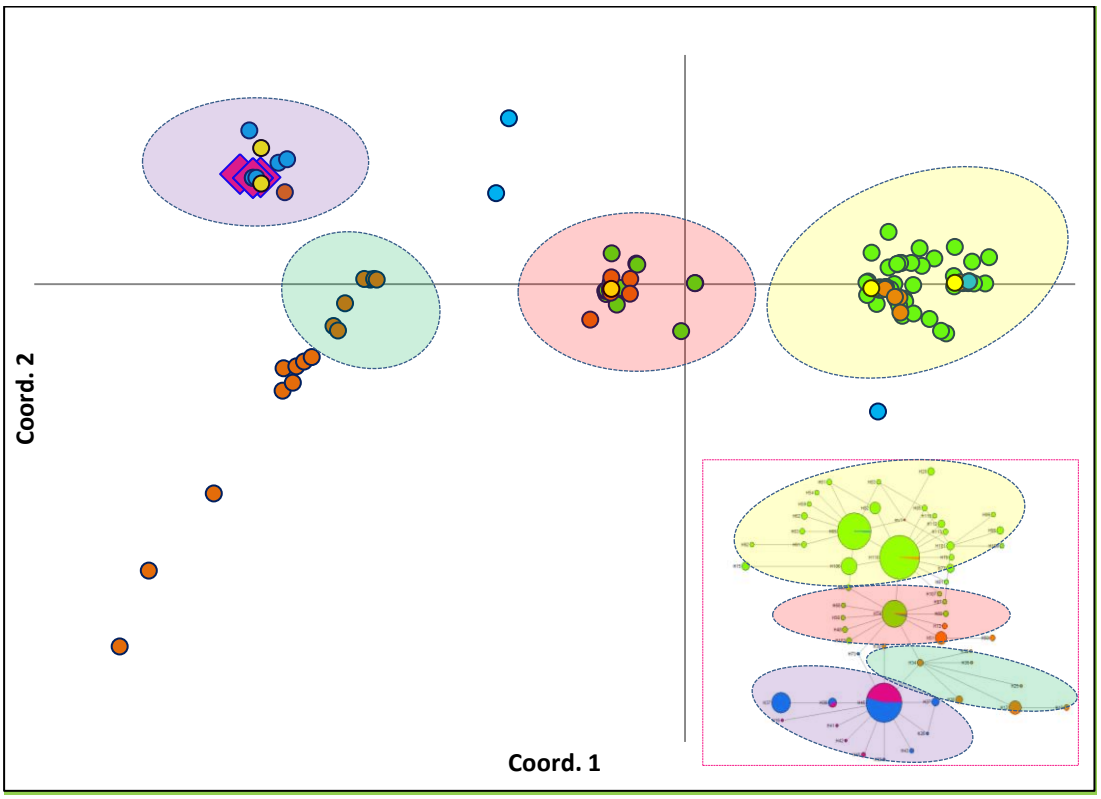


Figure S3-1. Principal Coordinate Analysis (PCoA) via covariance matrix of pairwise genetic distances of D haplotypes observed in Africa (*blue*), South Asia (*brown*), Indonesia (*green*), and Madagascar (*purple*). Haplotypes found in more than one geographic region are in yellow. Shaded haplotypes in the PCoA corresponds to the haplotypes in the MJ network (inset). Unshaded haplotypes corresponds to haplotypes not shown in the network.

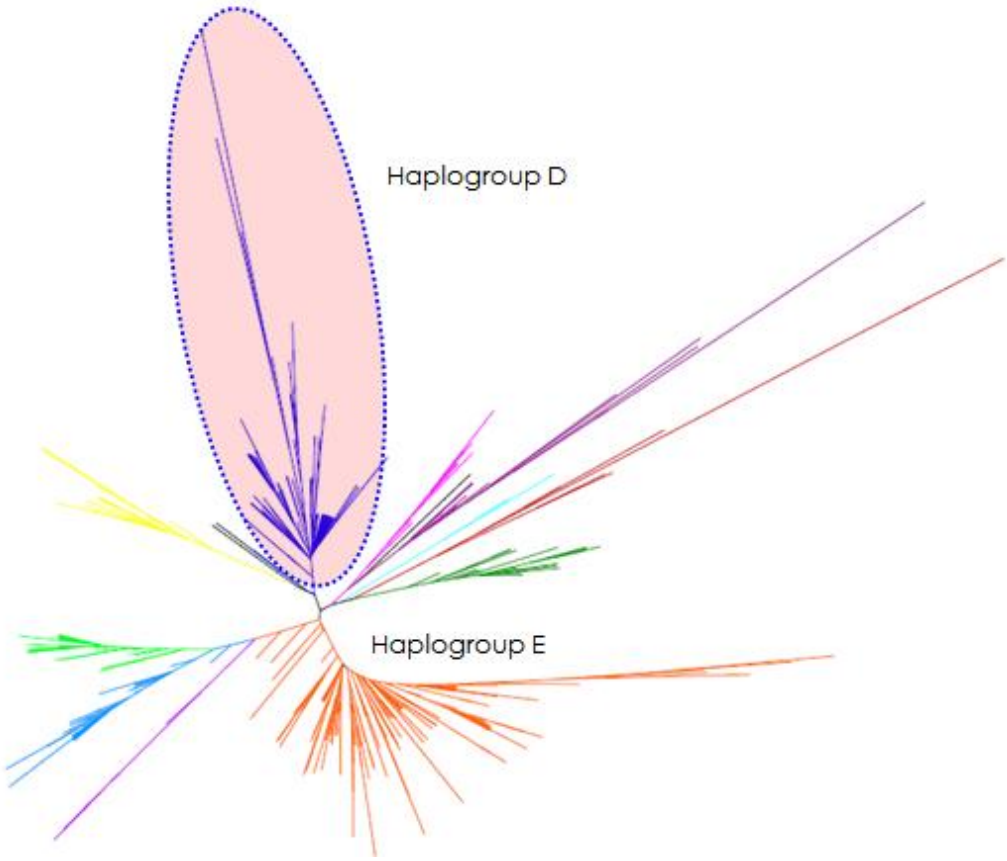


Figure S3-2. Neighbour-joining tree of the 429 haplotypes generated by the 349 bp dataset used in the study. Haplogroup D and E are depicted.

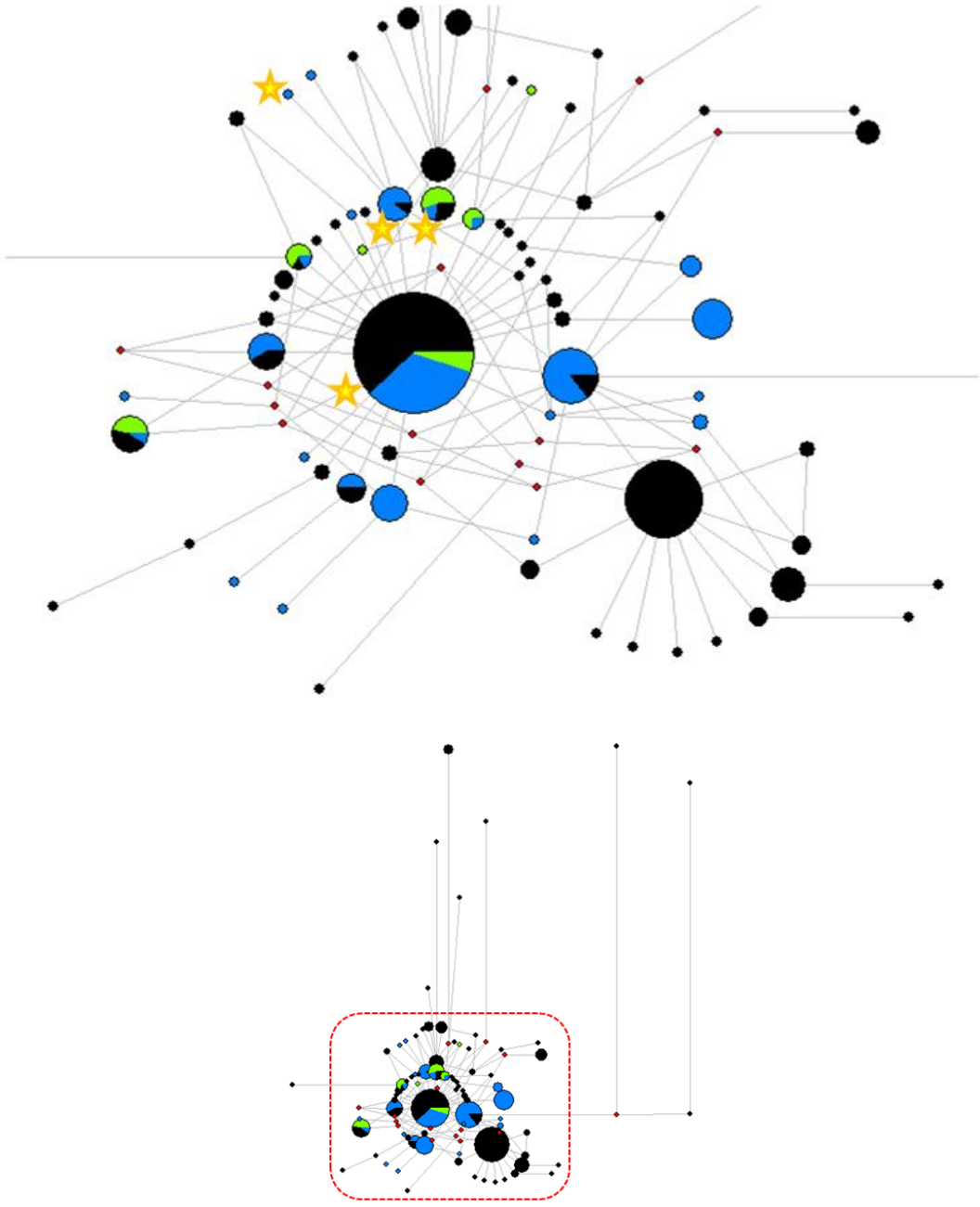


Figure S3-3: Median-joining network depicting the relationship of the E haplotypes observed in East Africa and Madagascar (*blue*), South Asia (*black*) and Indonesia (*green*). Stars mark the positions of Madagascan samples. Inferred haplotypes are indicated by small red dots. The inset shows all 80 observed haplotypes.

Table S3-1. Samples used in the study. (Excel Table, CD-ROM)

Table S3-2. Genetic diversity measures for each chicken population from Indonesia, South Asia, Continental Africa, and Madagascar.

Region	N(H)	HD (SD)	ND (SD)	π
		All haplogroups		
South Asia	552 (153)	0.9494 (0.0063)	0.0192 (0.0100)	6.71
India	324 (90)	0.8693 (0.0164)	0.0153 (0.0082)	5.35
Sri Lanka	132 (38)	0.9392 (0.0085)	0.0208 (0.0108)	7.24
Bangladesh	96 (25)	0.9276 (0.0115)	0.0202 (0.0106)	7.06
Indonesia	648 (142)	0.9602 (0.0030)	0.0089 (0.0051)	3.11
Sumatra	191 (38)	0.8224 (0.0207)	0.0115 (0.0064)	4.01
Java	181 (35)	0.7861 (0.0224)	0.0081 (0.0048)	2.84
Kalimantan	46 (13)	0.8493 (0.0405)	0.0070 (0.0043)	2.46
Sulawesi	78 (23)	0.8315 (0.0377)	0.0061 (0.0038)	2.14
Nusa Tenggara	76 (21)	0.8702 (0.0286)	0.0111 (0.0063)	3.89
Maluku	50 (7)	0.6188 (0.0641)	0.0032 (0.0023)	1.10
Irian Jaya	26 (5)	0.5662 (0.0862)	0.0038 (0.0027)	1.32
Africa	277 (43)	0.9118 (0.0075)	0.0123 (0.0068)	4.32
Kenya	159 (27)	0.8505 (0.0171)	0.0127 (0.0069)	4.45
Zimbabwe	99 (13)	0.7170 (0.0356)	0.0104 (0.0059)	3.61
Malawi	19 (3)	0.2924 (0.1274)	0.0009 (0.0011)	0.30
Madagascar	79 (10)	0.4340 (0.0680)	0.0048 (0.0032)	1.68
		D haplogroup only		
South Asia	100 (40)	0.9166 (0.0186)	0.0118 (0.0066)	4.13
India	71 (31)	0.8769 (0.0350)	0.0128 (0.0071)	4.56
Sri Lanka	4 (4)	1.00 (0.1768)	0.0105 (0.0079)	3.67
Bangladesh	25 (5)	0.6433 (0.0710)	0.0064 (0.0041)	2.25
Indonesia	551 (102)	0.9464 (0.0039)	0.0038 (0.0026)	1.32
Sumatra	154 (25)	0.7385 (0.0276)	0.0034 (0.0024)	1.17
Java	150 (21)	0.6915 (0.0259)	0.0032 (0.0023)	1.11
Kalimantan	40 (10)	0.8077 (0.0507)	0.0049 (0.0032)	1.71
Sulawesi	74 (19)	0.8127 (0.0407)	0.0048 (0.0032)	1.68
Nusa Tenggara	59 (17)	0.8101 (0.0452)	0.0044 (0.0029)	1.52
Maluku	49 (6)	0.6029 (0.0649)	0.0022 (0.0018)	0.76
Irian Jaya	25 (4)	0.5300 (0.0861)	0.0021 (0.0018)	0.74
Africa	127 (19)	0.7768 (0.0242)	0.0041 (0.0028)	1.41
Kenya	53 (9)	0.5457 (0.0759)	0.0038 (0.0027)	1.32
Zimbabwe	55 (7)	0.3003 (0.0802)	0.0020 (0.0017)	0.71
Malawi	19 (3)	0.2924 (0.1274)	0.0009 (0.0011)	0.30
Madagascar	67 (6)	0.2243 (0.0675)	0.0008 (0.0009)	0.27

N(H) – size (haplotypes #), HD(SD) – haplotype diversity (standard deviation), ND – nucleotide diversity, π – mean # of pairwise difference , SSD – sum of squared differences, * - statistically significant p-values ($p<0.05$ for Tajima's D, $p<0.02$ for Fu's FS)

CHAPTER 4: Mitochondrial DNA genomes resolve the genetic origins of Polynesian chickens

Statement of authorship

Mitochondrial DNA genomes resolve the genetic origins of Polynesian chickens

Michael James Bannister Herrera (Candidate)
Conceptualised and designed the study, performed the data processing, analysis, and interpretation, created figures and tables, and wrote the paper.

I hereby certify that the statement of contribution is accurate.

Signed _____ _____ Date 6 March 2015

Jeremy Austin
Assisted in the study design, gave advice on laboratory work and in data interpretation, and edited the manuscript.

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 6 March 2015

Vicki Thomson
Gave advice on data interpretation, commented and edited the manuscript.

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 5 March 2015

Jessica Wadley
Gave advice on laboratory work, assisted in the preparation of the sequencing libraries, assisted in the data processing, gave advice on data interpretation, commented and edited the manuscript.

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 5 March 2015

Alan Cooper
Assisted in the design of the study.

I hereby certify that the statement of contribution is accurate.

5 March 2015

Signed _____ Date _____

Philip Piper
Assisted in the interpretation of results, commented and edited the manuscript

I hereby certify that the statement of contribution is accurate.

2 March 2015

Signed _____ Date _____

Jaime Gongora
Liaised the Indonesian samples.

I hereby certify that the statement of contribution is accurate.

3 March 2015

Signed _____ Date _____

Mitochondrial DNA genomes resolve the genetic origins of Polynesian chickens

Michael B. Herrera[A], Jessica J. Wadley[A], Vicki A. Thomson[A], Philip J. Piper[B], Jaime Gongora[C], Alan Cooper[A] and Jeremy J. Austin[A]

[A]Australian Centre for Ancient DNA, School Biological Sciences, University of Adelaide, South Australia 5005, Australia.

[B]School of Archaeology and Anthropology, Faculty of Arts, The Australian National University, Canberra ACT 0200, Australia.

[C]Centre for Advanced Technologies in Animal Genetics and Reproduction, Faculty of Veterinary Science, University of Sydney, Sydney NSW 2006, Australia

Abstract

Indigenous chickens (*Gallus gallus domesticus*) from Island Southeast Asia predominantly comprise one mitochondrial DNA (mtDNA) haplogroup (haplogroup D) that also likely represents the founding lineage transported prehistorically to the Pacific Islands. Currently, inferences about human-mediated transport of chickens into Island Southeast Asia and the Pacific are based on short mtDNA control region sequences, which lack the ability to resolve relationships among individuals within closely related lineages. Here we investigate the maternal population history of chickens in Island Southeast Asia and the Pacific using 57 complete mtDNA genomes from archaeological, historic, and modern chickens. We combine these new sequences with existing mtDNA genomes from mainland Asia to infer the translocation history of chickens into Island Southeast Asia and the Pacific. We also examine how genome-scale based inferences compare with current hypotheses on human-mediated movement of chickens in Island Southeast Asia and the Pacific based on short control region sequences. Our results reveal that haplogroup D lineages in the ISEA and Pacific regions are divergent from mainland Asian haplogroup D chickens, and the Polynesian chickens are ultimately genetically derived from Philippine chickens. Furthermore, there are no significant contradictions between inferences based on mitochondrial control region sequences and whole mitochondrial genomes.

Keywords

Whole mitochondrial genomes, Island Southeast Asia, Pacific, chickens, migration

Introduction

The domestic chickens (*Gallus gallus domesticus*) appears to have experienced multiple domestication events in distinct areas across South and Southeast Asia, with the genetic relationship and distribution of these chickens suggesting the red jungle fowl (*Gallus gallus*; Figure 4-1) was the primary ancestor (Liu *et al*. 2006). Thirteen mitochondrial haplogroups (A-I; W-Z) have been identified from control region sequences (Liu *et al*. 2006; Miao *et al.* 2012) and all can be found in mainland Asia. However, in contrast to mainland Asia, where all thirteen haplogroups occur, chickens in Island Southeast Asia (ISEA) are restricted to predominantly haplogroup D. This hints to a single or a limited dispersal of chickens from mainland Asia onto the islands and archipelagos of ISEA. The level of current diversity of haplogroup D in ISEA also suggests it has had a protracted history in the region (Chapter 2). In the Pacific, the situation is similar, with the haplogroup composition being represented by a subset of lineages from haplogroup D.

Figure 4-1. Photographs of red jungle fowl (left) and green jungle fowl (right).

The origin of current domestic chickens in ISEA, as either an introduction of domestic chickens from elsewhere or as a separate domestication event within ISEA, remains an enigma. The red jungle fowl (RJF) are indigenous to mainland Southeast Asia (MSEA), southern China, and ISEA. Certain islands in Indonesia (*i.e.*, Java and Nusa Tenggara) are also home to another indigenous jungle fowl - the green jungle fowl (*Gallus varius*; Figure 4-1) whose contribution to domestic chickens is unknown.

In the Philippines, the timing of the arrival of red junglefowl (locally known as "labuyo") and their influence on the domestication process of chickens are uncertain. Two possibilities can be proposed about their arrival in the Philippines. First, the wild "labuyo" populations in the Philippines represent the descendants of domestic chickens prehistorically translocated into the archipelago from MSEA that subsequently became feral. Second, the prehistoric range of RJF included certain islands in the Philippines (*e.g.* Palawan) that were joined to MSEA during periods of lowered sea level, from where they subsequently spread onto other islands via human movements and hybridised with incoming domestic chickens during the recent past. Archaeological evidence for chickens in ISEA prior to European contact is almost nil, thus the inferred history of chickens in the region is unknown. However, as chickens appear in the archaeological record alongside the expansion of the Lapita people in Near Oceania beginning c.a. 3000 before present (BP) (Storey *et al.* 2008), chickens are likely to have been in ISEA prior to this period.

Mitochondrial DNA studies of contemporary chickens in ISEA (Indonesia and the Philippines) reveal that the majority belong to haplogroup D and the high

level of haplogroup D diversity there suggests their protracted antiquity in the region (Chapter 2). The majority of modern and archaeological chickens in the Pacific also belong to haplogroup D (Chapter 2). However, a small subset of closely related haplogroup D lineages, called the "Polynesian D", appear to have been prehistorically translocated into the Pacific (Thomson *et al.* 2014). To date, studies have mainly relied on short stretches of the mtDNA control region to infer the translocation history of chickens in ISEA and the Pacific regions (Storey *et al*. 2007; Gongora *et al.* 2008; Storey *et al*. 2010; Dancause *et al*. 2011; Thomson *et al*. 2014). While the control region is useful in establishing the broad geographic distributions of chicken lineages, it is unable to fully resolve the phylogenetic relationships of haplogroup D chicken lineages in the Asia-Pacific region. Thus, it is difficult to differentiate the founding lineages from the more derived lineages particularly in ISEA where prehistoric chicken remains are unavailable for ancient DNA research. Whole mitochondrial genome (WMG) datasets have been shown to yield more accurate phylogenetic trees than analyses based solely on control region sequences or a few mitochondrial genes, and they also reduce stochastic errors and minimise the effect of homoplasy (Campbell & Lapointe 2011; Havird & Santos 2014). Therefore, the additional nucleotide polymorphisms found outside the control region can potentially aid in resolving the relationships of haplogroup D lineages within ISEA, and could shed light on the evolutionary and translocation history of chickens across ISEA and the Pacific.

To investigate the relationship of haplogroup D chickens in the Asia-Pacific region, we sequenced 57 WMGs from archaeological, historical museum specimens, and modern chickens belonging to haplogroup D from the Pacific, and historic and

modern chickens from ISEA and MSEA. The 57 new WMGs represent all known D haplotypes found in the Pacific and major D haplotypes found in the Philippines and Indonesia previously identified using the control region (Chapter 2). We chose to focus on haplogroup D because it likely represents the sole or predominant lineage involved in the translocation history of chickens in the ISEA-Pacific region. A continuous genomic haplogroup D record across the Asia-Pacific region could provide an opportunity to explore the translocation history of chickens in these regions at a resolution that has never been done previously. We also incorporated 63 previously published WMGs representing all thirteen haplogroups (Nishibori *et al.* 2003; Wada *et al.* 2004; Froman & Kirby 2005; Nishibori *et al.* 2005; Tong *et al*. 2006; Miao *et al.* 2012) in our analyses to examine the relationship of lineages found in mainland Asia and in the ISEA-Pacific region.

Materials and methods

Sample collection, DNA extraction, library preparation, hybridisation-enrichment and sequencing

In order to select haplogroup D individuals for WMG sequencing, a phylogenetic network was created using control region sequences (Chapter 2). The selection process ensured that all known D haplotypes from the Pacific were represented along with the major D haplotypes from ISEA (Table S4-1).

Modern and historic DNA extracts were extracted as per Chapter 2. The DNA was then sheared to an average 200 base pair (bp) sized fragment using a Covaris S220 machine (Woburn, MA). Archaeological (ancient) samples did not require DNA fragmentation due to the degraded nature of the DNA and the short fragment sizes typical of ancient DNA samples. DNA polishing, phosphorylation, adapter

ligation and polymerase ‘fill-in’ reactions were done successively to create fully double-stranded adaptor-tagged DNA libraries for each of the samples (Meyer & Kircher 2010). For modern and historic samples, DNA polishing and phosphorylation reactions were done in a 40 µL final volume with 20 µL of DNA extract added to a 20 µL reaction containing 1 mM ATP, 0.2 mg/mL rabbit serum albumin (RSA; Sigma), 0.1 mM of each dNTP (Invitrogen), 0.5 U/µL T4 polynucleotide kinase (PNK; New England Biolabs, NEB), 1 X NEB Buffer 2, and 0.1125 U/µL T4 DNA Polymerase (NEB). The mixture was heated for 25 min at 25˚C using a thermal block and then the repaired DNA was purified using AMPure magnetic beads (Agencourt) following manufacturer’s instructions. Purified DNA was eluted with 30 µL elution buffer (Qiagen) containing 0.05 % Tween 20. After purification, PNK was denatured at 46˚C for 15 min. For ancient DNA samples, repair reactions were done at 40 µL final volume with 20 µL of DNA extract added to 20 µL reaction containing 10 mM ATP, 0.2 mg/mL rabbit serum albumin (RSA; Sigma), 0.1 mM of each dNTP (Invitrogen), 0.5 U/µL T4 polynucleotide kinase (PNK; New England Biolabs, NEB), 1 X NEB Buffer 2, and 0.1125 U/µL T4 DNA Polymerase (NEB). The mixture was heated for 15 min at 25˚C, 5 min at 12˚C, and 20 min at 75˚C using a thermal block and then the repaired DNA was purified using Minelute spin columns (Qiagen) as per manufacturer’s instructions. Purified DNA was eluted with 22.5 µL elution buffer (Qiagen) containing 0.05 % Tween 20.

Short adapters were ligated to each sample, with one adapter containing a unique 5 bp index, which facilitated sample identification and the exclusion of contaminating DNA (Table S4-2). For all sample types adapter ligation were conducted in 40 µL final volume comprising 1 X T4 Ligase buffer (NEB), 4 %

polyethylene glycol (PEG-400), 0.125 U/µL T4 DNA ligase (NEB), 25 mM barcoded P5 adaptor, 25 mM generic P7 adaptor and 20 µL purified DNA. The ligation mixture was heated for 60 min at 22 ˚C, and then purified using AMPure. Polymerase 'fill-in' reactions, to remove nicks and to create fully double-stranded adaptor-tagged DNA, were done in 40 µL final volume containing 1 X Thermopol buffer, 0.25 mM each dNTP, and 0.3 U/µL Bst-DNA polymerase Large Fragment (NEB) and 20 µL double-stranded adaptor-tagged DNA. The thermal cycling condition was 20 min at 37 ˚C, then 10 min at 80 ˚C.

To immortalise the DNA libraries, PCR amplification reactions were performed in 6 X 25 µL volumes containing 5 µL eluted DNA per tube, 2.5 U AmpliTaq Gold (Applied Biosystems), 250 µM of each dNTP (Invitrogen), 2.5 mM $MgCl_2$, 1 X AmpliTaq Gold buffer, and 0.5 µM short amplification P5 and P7 primers (Table S4-2). The thermocycling condition consisted of 94 ˚C for 6 min, followed by 12 cycles of 30 sec at 94 ˚C, 30 sec at 60 ˚C and 45 sec at 72 ˚C, followed by a final 10 min at 72 ˚C. The amplifications were separated into six separate PCRs to minimise amplification bias. The 6 X 25 µL products were pooled and purified using AMPure magnetic beads (Agencourt) and eluted in 30 µL EB buffer containing 0.05 % Tween 20. A second round of amplification following the conditions above was then performed. The DNA libraries were checked via gel electrophoresis against quantified size markers (HyperLadder^{TM} V, Bioline) and quantified using a Qubit (Life Technologies).

Capture-based enrichment was performed after the creation of the barcoded libraries via hybridisation to biotinylated RNA baits synthesised by MYcroarray (MI,

USA). The RNA baits were designed using published chicken WMG sequence NC_007236 (Miao et al. 2013). Between 4 and 8 barcoded libraries were pooled in equimolar amounts ensuring a total volume of 5.9 µL with total DNA between 100-500ng. Subsequent steps involved hybridisation by incubation for 36 hours and recovery of the captured targets involved immobilisation of the library DNA on magnetic streptavidin beads following the MYbaits Manual Version 2 (http://www.mycroarray.com). Post-capture "on-bead" amplifications (MYbaits Manual v2) were then performed in 6 X 25 µL volumes containing 4 µL of pooled library on beads, 2.5 U AmpliTaq Gold (Applied Biosystems), 250 µM of each dNTP (Invitrogen), 2.5 mM $MgCl_2$, 1 X AmpliTaq Gold buffer, and 0.5 µM short amplification P5 and P7 primers. The thermocycling conditions consisted of 94 ˚C for 12 min, followed by 12 cycles of 30 sec at 94 ˚C, 30 sec at 60 ˚C and 45 sec at 72 ˚C, followed by a final 10 min at 72 ˚C. The 6 X 25 µL products were pooled and purified using AMPURE (Agencourt). After elution and quantification of the DNA libraries using Qubit (Life Technologies) a final amplification step was performed to add the Illumina full-length adapters. After pooling the enriched DNA libraries in equimolar amounts the final concentration of the library was determined on the Agilent 2100 Bioanalyzer using the High Sensitivity DNA kit. The post-enrichment amplified products for modern samples were submitted to South Australia Pathology for Illumina HiSeq (paired-end) sequencing. Ancient and museum samples were sent to the Australian Genome Research Facility for Illumina MiSeq sequencing.

Following sequencing, the read were processed and quality filtered. This involved de-multiplexing the sequence reads according to the 5' index sequence used, allowing for no mismatches for the modern samples and one mismatch for the

ancient and museum samples. The adapters were then trimmed from the de-multiplexed sequence reads using Adapter Removal v1.5 (Lindgreen 2012). After the removal of the 5' index (*i.e.,* first five base pairs), the reads were then used for mapping and assembling to a consensus using the reference WMG sequence used to design the baits. To check for authenticity, particularly for the ancient samples, patterns of damage across the mapped reads were performed using MapDamage (Ginolhac *et al*. 2011).

Network analysis & phylogenetic construction

In addition to the 57 WMGs (haplogroup D, n=56; haplogroup Z, n=1) generated in this study, 61 previously published WMGs composing of haplogroups A-I (n=4, 5, 8, 8, 16, 8, 5, 1, 1 respectively) and W-Z (n=1, 1, 1, 2 respectively) and 2 *Gallus sonneratii* (grey jungle fowl) WMGs (Table S4-1) were also used to reconstruct the evolutionary history of chickens in the Asia-Pacific region. The combined 120 WMGs were aligned using the MUSCLE algorithm in GENEIOUS v.6.0 (Drummond *et al.* 2011).

A phylogenetic network was constructed using a median-joining (MJ) algorithm (Bandelt *et al.* 1999) implemented in NETWORK 4.6 (www.fluxus-engineering.com). Additional MJ networks were estimated using only the control region (1231bp) and another using a short, highly variable 201 bp fragment of the control region (Chapter 2) to examine concordance among phylogenetic networks using different data matrices. Lastly, a phylogenetic network for all haplogroup D WMGs (n=64) was also generated to investigate the relationship between haplogroup D genomes from the Asia-Pacific region. To obtain support values for the clades, a

phylogenetic tree of the haplogroup D WMGs was reconstructed using maximum likelihood implemented in RaxML v7.04 (Stamatakis 2006). The ML tree was performed with bootstrapping via 500 iterations followed by an optimised maximum likelihood search. The tree was rooted to AP003319 – a haplogroup E mitochondrial genome.

Population genetic statistics and structure

Arlequin v3.5 (Excoffier *et al.* 2005) was used to calculate haplotype diversity, nucleotide diversity, number of variable sites, transitions, transversions, and the number of haplotypes. Principal coordinate analysis (PCoA) was used to describe and visualise the genetic relationships of the 64 haplogroup D WMGs from mainland Asia, ISEA, and the Pacific. The pairwise genetic distances between individual haplogroup D WMGs used for the PCoA were calculated using GenAlEX v6.5 (Peakall & Smouse 2012).

Results

Sequence and phylogenetic network of *Gallus gallus*

In total, 57 new chicken WMGs were generated consisting of 9 archaeological, 10 historical museum, and 38 modern samples. All archaeological and historic samples exhibited patterns of DNA damage, consistent with being endogenous ancient DNA – elevated levels of C>T and A>G transition at the 5' and 3' ends of sequence reads (Figure S4-1). The combined analysis of 57 new and 63 published chicken WMGs from the Asia-Pacific region yielded an aligned data set of 16,785 bp. The WMG analysis revealed 118 haplotypes defined by 456 polymorphic sites, 417 of these sites are transitions and 39 are transversions. These variable sites combine to yield an

overall mean pairwise uncorrected distance of 0.00169, this translates to approximately 30.65 differences between any two genomes. A high proportion of polymorphic sites (22%, 102 of 456) were found in the control region, and all control region SNPs in the new WMGs matched with the previous control region data (Chapter 2) produced via direct PCR and Sanger sequencing.

The new 57 WMGs from ISEA and the Pacific did not identify any new haplogroups, which has previously been defined as haplogroups A-I and W-Z (Liu *et al.* 2006; Miao *et al*. 2012). Fifty-six of the new WMGs belong to haplogroup D, with one sample belonging to haplogroup Z (Figure 4-2). When we combine all 120 WMGs, it is possible to observe a degree of geographic structuring for some of the haplogroups. Haplogroup D is ubiquitously found in ISEA, and a distinct set of lineages within haplogroup D (called Polynesian D) are found predominantly in the Pacific (Figure 4-3). The haplogroup Z sample we had was located in Hainan, confirming the limited range of this haplogroup from Miao *et al*. (2012). Lastly, WMGs of grey jungle fowls (*Gallus sonneratii*) fall within the diversity of haplogroup E (black circles in Figure 4-2).

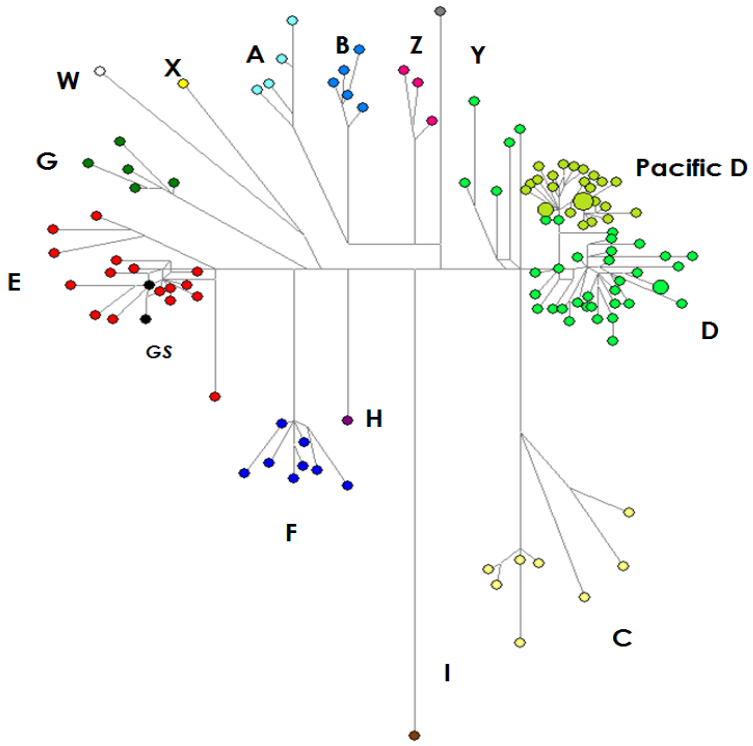


Figure 4-2. Median-joining network of 120 chicken mitochondrial genomes (117 haplotypes). Colours represent 13 known haplogroups (A-I, W-Z, and GS = *Gallus sonneratii*). Length of branch corresponds to number of mutations.

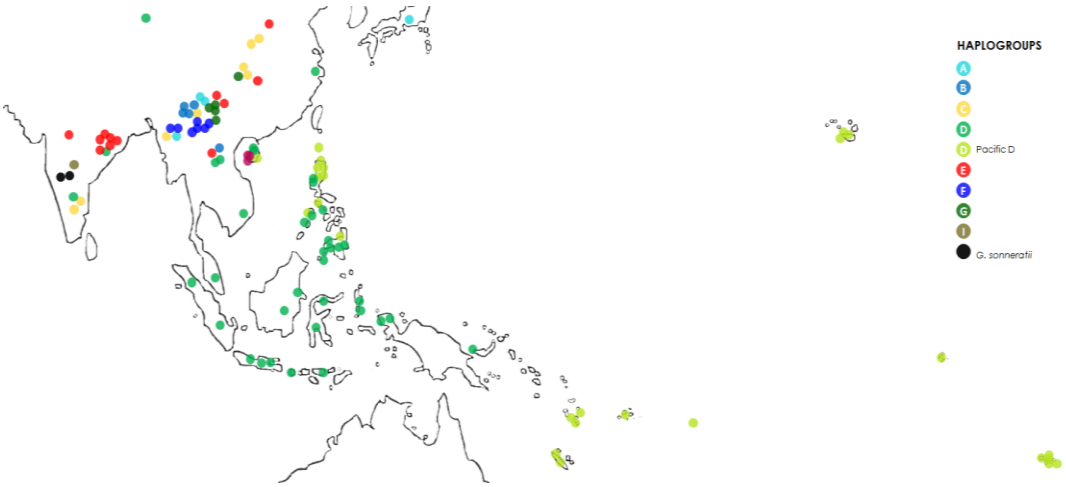


Figure 4-3. Map of the Asia-Pacific region showing the geographic distribution of 120 whole mitochondrial genomes representing the major mtDNA haplogroups. Haplogroups H, W, X, and Y are all found in Yunnan, China (not shown in the map). The location of the *Gallus sonneratii* samples on the map is only indicative, as the samples were not wild specimens. However, *G. sonneratii* are indigenous only to India, we are confident it is approximately correct.

Haplogroup D in the Asia-Pacific

Haplogroup D WMGs from the Asia-Pacific region (n=64) yielded 62 haplotypes defined by 156 polymorphic sites. The haplogroup D WMGs have a haplotype diversity of 0.999 +/- 0.003 and nucleotide diversity of 0.00066 +/- 0.0003. The phylogenetic network shows a clear phylogenetic discontinuity between haplotypes that contain the Polynesian motif (defined in Thomson *et al.* 2014 as nucleotide changes from A to G at base 281, C to T at base 296, T to C at base 306, A to G at base 342 compared to NC_007235) and those that do not (Figure 4-4).

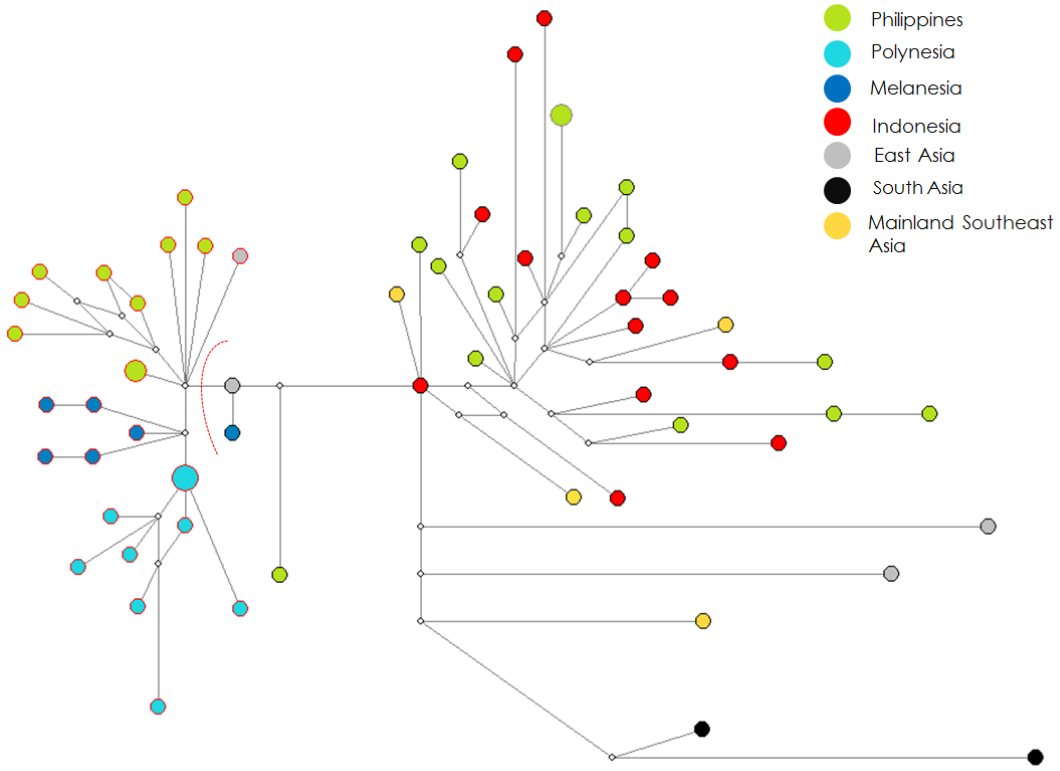


Figure 4-4. Median-joining network of 64 haplogroup D mitochondrial genomes (62 haplotypes). Colours represent the geographic location of the haplotype. Haplotypes with red outline are Polynesian D haplotypes, which is also demarcated by a red line from haplotypes that do not contain the signature in the right. Polynesian haplotypes found in Melanesia are in dark blue and Polynesia in light blue. White circles correspond to inferred haplotype and length of branch corresponds to number of mutations.

Haplotypes containing the Polynesian-motif are predominantly found in the Philippines and Pacific, with one museum sample from Hainan also containing the motif (Figure 4-4). For samples containing the Polynesian motif, genomes from the Philippines are clearly distinguishable from the Pacific, as are the genomes from Melanesia vs. Polynesia within the Pacific. None of the haplogroup D genomes found in Indonesia and mainland Asia contain the Polynesian motif, and in fact are quite distinct from those found in the rest of the ISEA-Pacific region. Of the haplogroup D genomes without the Polynesian motif, those found in Indonesia and the Philippines appear to be more phylogenetically related to each other than to any of those containing the motif.

The observed phylogeographic affinities within haplogroup D, as illustrated by the network analysis (Figure 4-4), are corroborated by principal component analysis (PCoA) using the genetic distances between the haplogroup D genomes (Figure 4-5). The PCoA plot shows three genetic clusters which correspond to geographic regions in the Asia-Pacific region. Again, haplogroup D chickens in continental Asia are divergent to those found in ISEA and the Pacific. It also shows a clear affinity of samples from the Philippines, Melanesia, and Polynesia that contains the Polynesian D motif.

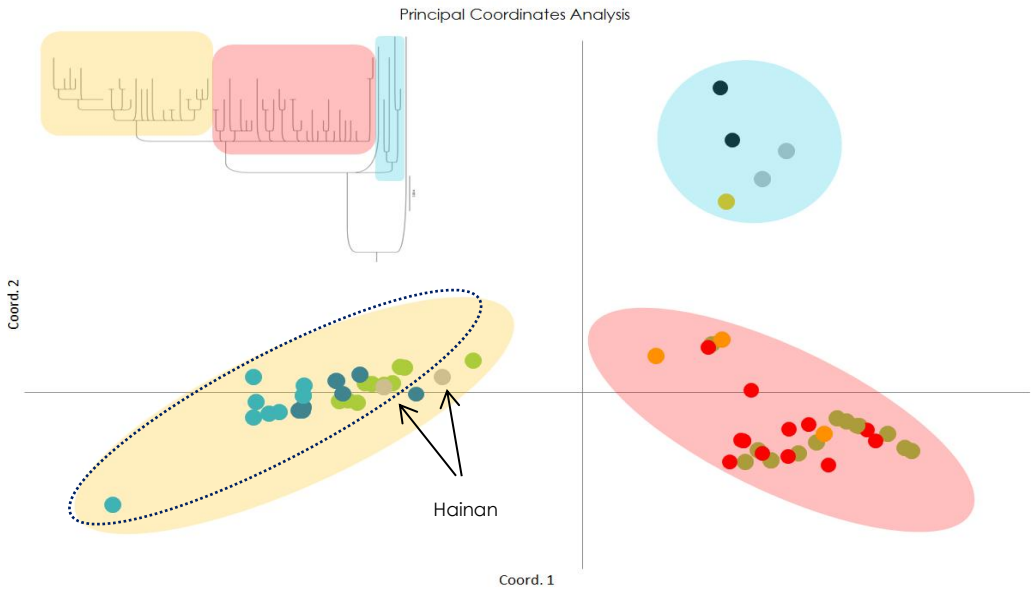


Figure 4-5. PCoA based on the genetic distances of haplogroup D mitochondrial genomes of chickens from the Asia-Pacific region. Geographic locations of the samples are indicated as colour points: light blue – Polynesia, dark blue – Melanesia, green – Philippines, grey – east Asia, red – Indonesia, orange – mainland Southeast Asia, and black – South Asia. Samples containing the Polynesian D motif are within broken line. Colour shading corresponds to the position in the inset ML tree (see next section).

Phylogenetic reconstructions of haplogroup D mitochondrial genomes

The phylogenetic tree of 64 haplogroup D samples shows a polytomy consisting of samples mostly from Indonesia and the Philippines (within red dashed box, Figure 4-6). The inclusion of two samples from Laos within the haplogroup D diversity in ISEA could either mean that chickens in Indonesia and the Philippines are ultimately derived from MSEA or it indicates modern day movement of chickens. Furthermore, continental Asian chickens (*i.e.*, India, China, and Vietnam) are basal in relation to chickens from ISEA and the Pacific. Although with low bootstrap supports, the structure of the tree also shows a step-wise regional clustering of a subset of haplogroup D samples that proceeds from mostly Philippine haplotypes (within green box in Figure 4-6) then to Melanesia and Polynesia. However, a Polynesian sample from Niue clusters with Melanesia rather than the other Polynesian samples. It appears that genomes have an increasingly derived status as they proceed east into

the Pacific, with samples from Polynesia (*i.e.*, Hawaii and Easter Islands) being the most derived. Interestingly, this occurs even though the Polynesian genomes are mostly archaeological samples (*i.e.*, their ancestral status does not result in them falling basal within the tree). The tree also shows that haplogroup D genomes from mainland Asia are distinct from haplogroup D chickens found in ISEA and the Pacific.

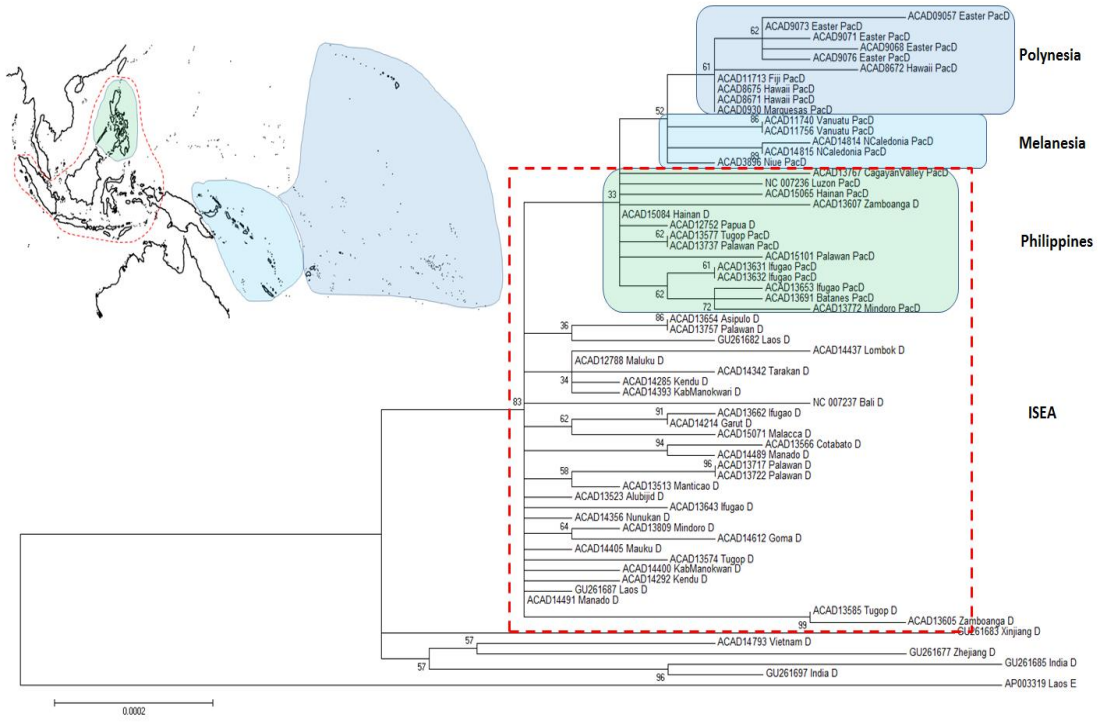


Figure 4-6. Maximum likelihood tree of 64 haplogroup D chicken mitochondrial genomes constructed using a GTR+G model of nucleotide evolution and rooted by a haplogroup E chicken from Laos. The percentage of trees in which the associated samples clustered together is shown next to the branches. Clustering on the tree correspond to the general regions on the inset map.

Comparisons between evolutionary inferences based on WMG and mtDNA-CR

There are no contradictions between the results from phylogenetic analyses using the WMG dataset compared to those using the whole control region or even the shorter 201bp fragment (Figure 4-7). The 13 chicken mitochondrial haplogroups are clearly distinct when using WMGs, however, when using only control region the haplogroups are still identifiable, although haplogroups C, H, Y, and Z are falling close to the diversity of haplogroup D (Figure 4-7). The same observation is true when using only the 201 bp fragment usually used for ancient DNA studies (e.g., Storey *et al.* 2007, Gongora *et al.* 2008; Thomson *et al*. 2014). Furthermore, Polynesian D lineages are most distinguishable when using WMGs.

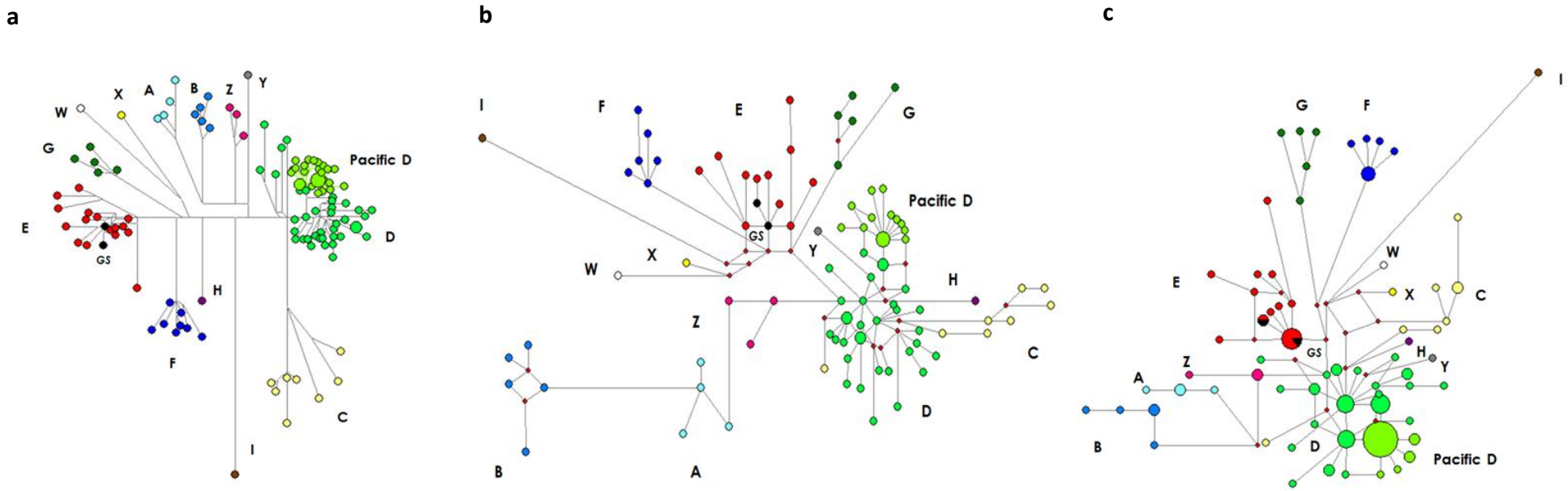


Figure 4-7. Phylogenetic networks of 120 chickens based on (a) whole mitochondrial genome, (b) control region, and (c) 201 bp fragment of the control region. Colour corresponds to haplogroup assignment.

Discussion

The mtDNA phylogeny for domestic chickens, with thirteen clearly defined haplogroups (Figure 4-2) suggests that the domestication process for chickens occurred multiple times involving potentially divergent source populations and jungle fowl species from several centres in Asia (Liu *et al.* 2006). Each haplogroup could potentially represent separate domestications of wild chickens from localised areas in South and Southeast Asia. However, these putative domestication centres are not distinctly reflected in the current geographic distribution of the haplogroups, which may be due to a long history of inter-regional interaction. These interactions most likely included the translocation of chickens through exchange and trade during modern times. Thus, chicken WMGs show a strong phylogenetic signature but with overlapping geographic distributions. To pinpoint exact geographic locations where jungle fowls experienced domestication events can only be resolved through ancient DNA analysis from archaeologically provenance and securely dated chicken bones. Ancient chicken remains in MSEA and ISEA are scarce due to the combined effects of a hot wet climate and other taphonomic processes, which degrade highly porous chicken bones and the DNA inside them. Furthermore, it was not until recently that archaeologists began recognising small animal bones as valuable sources of information. Previously archaeological endeavours were biased towards big cultural artefacts, with ecofacts (*e.g.* animal bones, plant remains, pollen, etc.) often discarded.

Although most chicken mtDNA haplogroups show overlapping geographic distributions, haplogroup D is unique in that it shows reasonable levels of both phylogenetic and geographic separation. In the same way that discreet haplogroups

can be used to infer domestication, a similar signature can be observed within haplogroup D in the ISEA-Pacific region. For example, a sub-group of haplogroup D – the Polynesian D - has been used to define patterns in the Pacific from a short fragment of the control region (Figure 4-7). Haplotypes containing the Polynesian motif may represent an independent domestication process in the Philippines. However, it is difficult to distinguish whether the 'Polynesian D' group represents: 1) an *in situ* domestication process in the Philippines from 'labuyo' red jungle fowl; 2) a re-domestication of chickens that were initially domesticated in mainland Southeast Asia and later introduced into the Philippines; or 3) a naturally isolated population on one of the Philippines islands that then happened to be the only lineage prehistorically translocated by humans into the Pacific.

The long branch separating the closely related Polynesian D haplotypes from the rest of haplogroup D in ISEA and the star-like clustering of Polynesian D chickens (Figure 4-4) certainly suggest the third possibility, where the Polynesian D cluster could represent domestic descendants of a localised progenitor of wild RJF in the Philippines. Domestication may also explain why the Pacific D is also a very specific lineage within haplogroup D diversity, as the domestication process itself would necessarily select a small portion of available diversity (in this case from haplogroup D) in the population. The possibility of an *in situ* domestication of chickens in the Philippines has previously been suggested based on linguistic data (Blust 2002), as there is no known Proto-Austronesian (Formosan) term for chicken. The Austronesian term for chicken only appears in the Proto Malayo-Polynesian branch, a language group that developed in the Philippines and is ancestral to all Austronesian languages outside of Taiwan.

Polynesian D's represent a specific set of haplotypes and seem to suggest that movement of chickens into the Pacific may have involved a small number of birds. Our analyses suggest that chickens with haplogroup D WMGs found in archaeological sites in Polynesia are descendants of domestic chickens still currently found in the Philippines. The strong phylogeographic signature of Polynesian D chickens therefore suggests that chickens endemic to the Philippines played a part in the Polynesian dispersal. Furthermore, the restricted distribution of Polynesian D chickens from the Philippines, Melanesia, and Polynesia readily support the existing models on Austronesian dispersal. This means that the Polynesian chicken is the first animal domesticate that is supportive of part of the linguistic and archaeological narrative about the Austronesian expansion out of Taiwan. However, the most likely scenario from the chicken data is that they were integrated into the expansion of the Austronesian-speakers while in the Philippines *en route* to the Pacific rather than chickens being transported directly from Taiwan. Thus, the WMG data presented here support the previous inferences made about the Philippine homeland for Pacific chickens based on short control region data (Chapter 2; Thomson *et al.* 2014). Altogether, the modern Polynesian D chickens retain at least the maternal ancestral genetic pattern seen in the ancient chicken samples and therefore appear to represent the chickens prehistorically transported into the Pacific.

Conclusion

This study shows that chicken mitochondrial DNA can be used to infer prehistoric processes, particularly when genetic analysis is anchored to archaeological and linguistic narratives. The data presented here demonstrates that haplogroup D chicken lineages in mainland Asia are distinct from those found in ISEA. Chickens representing the Pacific D cluster likely originated in the Philippines and potentially represent a discrete episode of animal domestication on island environments. However, chicken archaeological remains from relevant regions in ISEA that predate the colonial period are yet to be documented. We also highlight that, although the mitochondrial control region is valuable in revealing general phylogeographic patterns, the added utility of whole mitochondrial data extensively sampled across relevant geographic regions can reveal genetic patterns associated with putative episodes of domestication and prehistoric translocation of chickens.

Acknowledgement

This research was funded by the Australian Research Council (ARC) Discovery Project ID DP110105187. We thank the American Museum of Natural History (USA) and the National Museum of Natural History (France) for providing the historic museum samples.

References

Bandelt, H-J, P Forster, A Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular biology and evolution 16:37-48.

Blust, R. 2002. The History of Fanual Terms in Austronesian Languages. Oceanic linguistics:89-139.

Campbell, V, F-J Lapointe. 2011. Retrieving a mitogenomic mammal tree using composite taxa. Molecular phylogenetics and evolution 58:149-156.

Dancause, KN, MG Vilar, R Steffy, JK Lum. 2011. Characterizing genetic diversity of contemporary pacific chickens using mitochondrial DNA analyses. PloS one 6:e16843.

Drummond, A, B Ashton, S Buxton, M Cheung, A Cooper, C Duran, M Field, J Heled, M Kearse, S Markowitz. 2011. Geneious, version 5.4. Geneious, Auckland, New Zealand.

Excoffier, L, G Laval, S Schneider. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evolutionary bioinformatics online 1:47.

Froman, D, J Kirby. 2005. Sperm mobility: phenotype in roosters (Gallus domesticus) determined by mitochondrial function. Biology of reproduction 72:562-567.

Ginolhac, A, M Rasmussen, MTP Gilbert, E Willerslev, L Orlando. 2011. mapDamage: testing for damage patterns in ancient DNA sequences. Bioinformatics 27:2153-2155.

Gongora, J, NJ Rawlence, VA Mobegi, H Jianlin, JA Alcalde, JT Matus, O Hanotte, C Moran, JJ Austin, S Ulm. 2008. Indo-European and Asian origins for Chilean and Pacific chickens revealed by mtDNA. Proceedings of the National Academy of Sciences 105:10308-10313.

Havird, JC, SR Santos. 2014. Performance of Single and Concatenated Sets of Mitochondrial Genes at Inferring Metazoan Relationships Relative to Full Mitogenome Data. PloS one 9:e84080.

Lindgreen, S. 2012. AdapterRemoval: easy cleaning of next-generation sequencing reads. BMC research notes 5:337.

Liu, Y-P, G-S Wu, Y-G Yao, Y-W Miao, G Luikart, M Baig, A Beja-Pereira, Z-L Ding, MG Palanichamy, Y-P Zhang. 2006. Multiple maternal origins of chickens: out of the Asian jungles. Molecular phylogenetics and evolution 38:12-19.

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

Miao, Y, M Peng, G Wu, Y Ouyang, Z Yang, N Yu, J Liang, G Pianchou, A Beja-Pereira, B Mitra. 2012. Chicken domestication: an updated perspective based on mitochondrial genomes. Heredity 110:277-282.

Nishibori, M, M Hanazono, Y Yamamoto, M Tsudzuki, H Yasue. 2003. Complete nucleotide sequence of mitochondrial DNA in White Leghorn and White Plymouth Rock chickens. Animal Science Journal 74:437-439.

Nishibori, M, T Shimogiri, T Hayashi, H Yasue. 2005. Molecular evidence for hybridization of species in the genus Gallus except for Gallus varius. Animal genetics 36:367-375.

Peakall, R, PE Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28:2537-2539.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688-2690.

Storey, A, T Ladefoged, E Matisoo-Smith. 2008. Counting your chickens: density and distribution of chicken remains in archaeological sites of Oceania. International Journal of Osteoarchaeology 18:240-261.

Storey, AA, JM Ramírez, D Quiroz, DV Burley, DJ Addison, R Walter, AJ Anderson, TL Hunt, JS Athens, L Huynen. 2007. Radiocarbon and DNA evidence for a pre-Columbian introduction of Polynesian chickens to Chile. Proceedings of the National Academy of Sciences 104:10335-10339.

Storey, AA, M Spriggs, S Bedford, SC Hawkins, JH Robins, L Huynen, E Matisoo-Smith. 2010. Mitochondrial DNA from 3000-year old chickens at the Teouma site, Vanuatu. Journal of Archaeological Science 37:2459-2468.

Thomson, VA, O Lebrasseur, JJ Austin, TL Hunt, DA Burney, T Denham, NJ Rawlence, JR Wood, J Gongora, LG Flink. 2014. Using ancient DNA to study the origins and dispersal of ancestral Polynesian chickens across the Pacific. Proceedings of the National Academy of Sciences 111:4826-4831.

Tong, X, Y Liang, W Wang, S Xu, X Zheng, J Wang, J Yu. 2006. [Complete sequence and gene organization of the Tibetan chicken mitochondrial genome]. Yi chuan= Hereditas/Zhongguo yi chuan xue hui bian ji 28:769-777.

Wada, Y, Y Yamada, M Nishibori, H Yasue. 2004. Complete nucleotide sequence of mitochondrial genome in silkie fowl (Gallus gallus var. domesticus). The Journal of Poultry Science 41:76-82.

Supplementary information

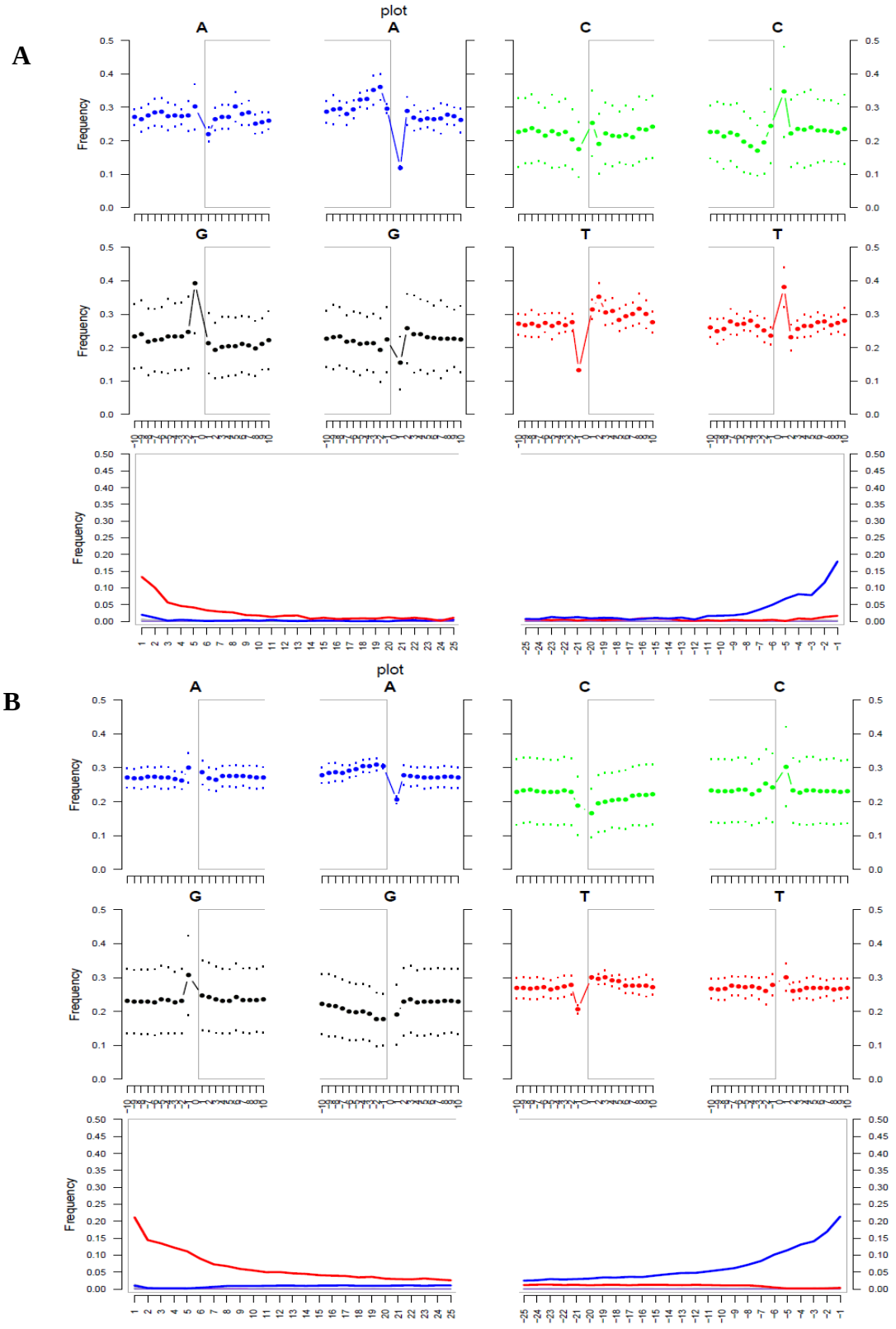


Figure S4-1: Examples of DNA map damage profiles for (A) ancient (ACAD 3896) and (B) museum (ACAD 14814) samples. The four panels above illustrate the high frequency of purines immediately before the reads. The two lower panes illustrate the accumulation of 5' C-to-T (red) and 3' G-to-A (blue) misincorporations characteristic of ancient DNA.

Table S4-1: Samples used in the study, reference number, description, and haplogroup.

No.	Provenance	Sample ID	Description	Reference	Haplogroup
1	Hoihow, Hainan, China	ACAD15065	Museum toe pad	This study	D
2	Isle de Pins, New Caledonia	ACAD14814	Museum toe pad	This study	D
3	Isle de Pins, New Caledonia	ACAD14815	Museum toe pad	This study	D
4	Hoihow, Hainan, China	ACAD15063	Museum toe pad	This study	Z
5	Aborlan, Palawan, Philippines	ACAD15101	Museum toe pad	This study	D
6	You-Boi, Hainan, China	ACAD15084	Museum toe pad	This study	D
7	Maluku, Wetar Islands	ACAD12788	Museum toe pad	This study	D
8	Long Island, New Guinea	ACAD12752	Museum toe pad	This study	D
9	North Annan, Vietnam	ACAD14793	Museum toe pad	This study	D
10	Malacca, Malay Peninsula	ACAD15071	Museum toe pad	This study	D
11	Batanes Archipelago, Philippines	ACAD13691	Modern feather	This study	D
12	Ifugao, Philippines	ACAD13631	Modern feather	This study	D
13	Ifugao, Philippines	ACAD13632	Modern feather	This study	D
14	Mindoro, Philippines	ACAD13772	Modern feather	This study	D
15	Tugop, Philippines	ACAD13577	Modern feather	This study	D
16	Palawan, Philippines	ACAD13737	Modern feather	This study	D
17	Eton, Efate Island, Vanuatu	ACAD11740	Modern feather	This study	D
18	Zamboanga, Philippines	ACAD13605	Modern feather	This study	D
19	Ifugao, Philippines	ACAD13662	Modern feather	This study	D
20	Palawan, Philippines	ACAD13722	Modern feather	This study	D
21	Cotabato, Philippines	ACAD13566	Modern feather	This study	D
22	Manticao, Philippines	ACAD13513	Modern feather	This study	D
23	Ifugao, Philippines	ACAD13653	Modern feather	This study	D
24	Alubijid, Philippines	ACAD13523	Modern feather	This study	D
25	Tugop, Philippines	ACAD13585	Modern feather	This study	D
26	Cagayan Valley, Philippines	ACAD13767	Modern feather	This study	D
27	Palawan, Philippines	ACAD13717	Modern feather	This study	D
28	Marquesas Islands, French Polynesia	ACAD9030	Modern feather	This study	D
29	Tugop, Philippines	ACAD13574	Modern feather	This study	D
30	Palawan, Philippines	ACAD13757	Modern feather	This study	D
31	Asipulo, Philippines	ACAD13654	Modern feather	This study	D
32	Mindoro, Philippines	ACAD13809	Modern feather	This study	D
33	Lombok, Indonesia	ACAD14437	Blood	This study	D
34	Zamboanga, Philippines	ACAD13607	Modern feather	This study	D
35	Ifugao, Philippines	ACAD13643	Modern feather	This study	D
36	Goma, Sulawesi, Indonesia	ACAD14612	Blood	This study	D
37	Kendu, Java, Indonesia	ACAD14292	Blood	This study	D
38	Manado, Sulawesi, Indonesia	ACAD14491	Blood	This study	D
39	Tarakan, Kalimantan, Indonesia	ACAD14342	Blood	This study	D
40	Ponangisu, Efate, Vanuatu	ACAD11756	Modern feather	This study	D
41	Maluku, Indonesia	ACAD14405	Blood	This study	D
42	Manado, Sulawesi, Indonesia	ACAD14489	Blood	This study	D
43	Kab Manokwari, Irian Jaya, Indonesia	ACAD14393	Blood	This study	D
44	Nunukan, Kalimantan, Indonesia	ACAD14356	Blood	This study	D
45	Kab Manokwari, Irian Jaya, Indonesia	ACAD14400	Blood	This study	D
46	Namatakula, Viti Levu, Fiji	ACAD11713	Modern feather	This study	D
47	Garut, Java, Indonesia	ACAD14214	Blood	This study	D
48	Kendu, Java, Indonesia	ACAD14285	Blood	This study	D
49	Anatoloa, Niue	ACAD3896	Archaeological bone	This study	D
50	Makauwahi Cave, Kauai, Hawaii	ACAD8671	Archaeological bone	This study	D
51	Makauwahi Cave, Kauai, Hawaii	ACAD8672	Archaeological bone	This study	D
52	Makauwahi Cave, Kauai, Hawaii	ACAD8675	Archaeological bone	This study	D
53	Anakena, Easter Island	ACAD9057	Archaeological bone	This study	D
54	Anakena, Easter Island	ACAD9071	Archaeological bone	This study	D
55	Anakena, Easter Island	ACAD9068	Archaeological bone	This study	D
56	Anakena, Easter Island	ACAD9073	Archaeological bone	This study	D
57	Anakena, Easter Island	ACAD9076	Archaeological bone	This study	D
58	Japan: Hiroshima	AB086102	domestic chicken	Wada et al. 2004	A
59	China: Yunnan	GU261684	domestic chicken	Miao et al. 2013	A
60	China: Yunnan	GU261695	wild fowl	Miao et al. 2013	A
61	Myanmar	GU261700	wild fowl	Miao et al. 2013	A
62	Laos: Vientiane	NC_007235	wild fowl	Nishibori et al. 2005	B
63	China: Yunnan	GU261704	wild fowl	Miao et al. 2013	B
64	China: Yunnan	GU261705	domestic chicken	Miao et al. 2013	B

65	China: Yunnan	GU261714	domestic chicken	Miao et al. 2013	B
66	China: Yunnan	GU261699	domestic chicken	Miao et al. 2013	B
67	China: Hainan	GU261674	wild fowl	Miao et al. 2013	Z
68	China: Hainan	GU261696	wild fowl	Miao et al. 2013	Z
69	China: Yunnan	GU261693	wild fowl	Miao et al. 2013	Y
70	China: Henan	GU261701	domestic chicken	Miao et al. 2013	C1
71	China: Hunan	GU261675	domestic chicken	Miao et al. 2013	C1
72	China: Hunan	GU261681	domestic chicken	Miao et al. 2013	C1
73	China: Yunnan	GU261718	domestic chicken	Miao et al. 2013	C1
74	China: Henan	GU261679	domestic chicken	Miao et al. 2013	C1
75	Southern India	GU261680	domestic chicken	Miao et al. 2013	C2
76	Myanmar	GU261716	wild fowl	Miao et al. 2013	C3
77	India	GU261707	wild fowl	Miao et al. 2013	C3
78	Philippine: Manila	NC_007236	wild fowl	Nishibori et al. 2005	D1
79	Indonesia: Bali	NC_007237	wild fowl	Nishibori et al. 2005	D1
80	Laos	GU261687	domestic chicken	Miao et al. 2013	D1
81	Laos	GU261682	domestic chicken	Miao et al. 2013	D1
82	China: Xinjiang	GU261683	domestic chicken	Miao et al. 2013	D2
83	China: Zhejiang	GU261677	domestic chicken	Miao et al. 2013	D3
84	Southern India	GU261697	domestic chicken	Miao et al. 2013	D3
85	Northeast India	GU261685	domestic chicken	Miao et al. 2013	D3
86	China: Henan	GU261686	domestic chicken	Miao et al. 2013	E1
87	China: Yunnan	GU261713	domestic chicken	Miao et al. 2013	E1
88	Commercial Line	AP003317	domestic chicken	Nishibori et al. 2003	E1
89	Commercial Lines	AY235571	domestic chicken	Froman & Kirby 2005	E1
90	Commercial Line	AP003318	domestic chicken	Nishibori et al. 2003	E1
91	China: Yunnan	GU261712	domestic chicken	Miao et al. 2013	E1
92	India	GU261709	domestic chicken	Miao et al. 2013	E1
93	Commercial Line	AY235570	domestic chicken	Froman & Kirby 2005	E1
94	Commercial Line	AP003580	domestic chicken	Nishibori et al. 2003	E1
95	China: Hebei	GU261694	domestic chicken	Miao et al. 2013	E1
96	Laos: Vientiane	AP003319	domestic chicken	Nishibori et al. 2005	E1
97	Northeast India	HQ857210	domestic chicken	Miao et al. 2013	E1
98	Northeast India	HQ857209	domestic chicken	Miao et al. 2013	E2
99	India	GU261708	wild fowl	Miao et al. 2013	E3
100	Northeast India	HQ857212	domestic chicken	Miao et al. 2013	E3
101	Northeast India	HQ857211	domestic chicken	Miao et al. 2013	E3
102	Myanmar	GU261691	wild fowl	Miao et al. 2013	F
103	China: Yunnan	GU261702	wild fowl	Miao et al. 2013	F
104	China: Yunnan	GU261688	domestic chicken	Miao et al. 2013	F
105	China: Yunnan	GU261711	domestic chicken	Miao et al. 2013	F
106	China: Yunnan	GU261689	domestic chicken	Miao et al. 2013	F
107	Myanmar	GU261703	wild fowl	Miao et al. 2013	F
108	China: Yunnan	GU261717	domestic chicken	Miao et al. 2013	F
109	China: Yunnan	DQ648776	domestic chicken	Tong et al. 2006	F
110	China: Henan	GU261678	domestic chicken	Miao et al. 2013	G
111	China: Yunnan	GU261710	domestic chicken	Miao et al. 2013	G
112	China: Yunnan	GU261676	domestic chicken	Miao et al. 2013	G
113	China: Yunnan	GU261719	domestic chicken	Miao et al. 2013	G
114	China: Yunnan	GU261690	wild fowl	Miao et al. 2013	G
115	China: Yunnan	GU261715	domestic chicken	Miao et al. 2013	H
116	China: Yunnan	GU261706	wild fowl	Miao et al. 2013	W
117	China: Yunnan	GU261692	wild fowl	Miao et al. 2013	X
118	Northeast India	GU261698	domestic chicken	Miao et al. 2013	I
119	Unknown	AP003320	*Gallus sonneratii*	Nishibori et al. 2005	?
120	Tama Zoological Park, Tokyo	AP006746	*Gallus sonneratii*	Nishibori et al. 2005	?

Table S4-2. Sequencing library structure and PCR primers used for the construction of the sequencing libraries.

Sequencing library structure	**Sequence (5'-3')**
5' truncated Illumina adapter	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
5' barcode	Variable
3' truncated Illumina adapter	AGATCGGAAGAG
Primers used for library amplification	**Sequence (5'-3')**
Library amplification Forward	ACACTCTTTCCCTACACGAC
Library amplification Reverse	GTGACTGGAGTTCAGACGTGT
Illumina Forward	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTT
Illumina Reverse	CAAGCAGAAGACGGCATACGAGATACCTAGGGTGACTGGAGTTCAGACGTGT

CHAPTER 5: Exploring the population history of chickens in the Asia-Pacific using genome-wide single nucleotide polymorphism

Statement of authorship

Exploring the population history of chickens in the Asia-Pacific region using genome-wide single nucleotide polymorphism

Michael James Herrera (Candidate)

Collected samples, performed DNA extractions, liaised for the genotyping of the samples, and processed, analysed, and interpreted the data, created figures and tables, and wrote the paper.

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 6 March 2015

Jeremy Austin

Helped design the study, liaised for the genotyping of the samples, assisted in the interpretation of results, edited and commented on the manuscript

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 6 March 2015

Vicki Thomson

Assisted in the genetic analysis, interpretation of results, and edited the figures and the manuscript

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 5 March 2015

Jessica Wadley

Assisted in the genetic analysis, interpretation of results, and edited the manuscript

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 5 March 2015

Alan Cooper

Assisted in designing and planning the study.

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 5 March 2015

Philip Piper

Assisted in the interpretation of results, commented and edited the manuscript

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 2 March 2015

Jaime Gongora

Gave advice in the study design.

I hereby certify that the statement of contribution is accurate.

Signed _____ Date 3 March 2015

Exploring the population history of chickens in the Asia-Pacific using genome-wide single nucleotide polymorphisms

Michael B. Herrera[A], Vicki A. Thomson[A], Jessica J. Wadley[A], Philip J. Piper[b], Alan Cooper[A], Jaime Gongora[C], and Jeremy J. Austin[A]

[A]Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, South Australia 5005, Australia

[B]School of Archaeology and Anthropology, Australian National University, Canberra ACT 0200, Australia

[C]Centre for Advanced Technologies in Animal Genetics and Reproduction, Faculty of Veterinary Science, University of Sydney, Sydney NSW 2006, Australia

Abstract

Our current understanding about the population and translocation history of indigenous chickens in Island Southeast Asia and the Pacific is based mostly on the geographic distribution of genetic variations found in the mitochondrial DNA. It has previously been revealed that Island Southeast Asia is a region where a specific chicken mitochondrial lineage, called haplogroup D, diversified. Chickens from this lineage were subsequently transported by prehistoric cultures from the Philippine archipelago to the islands further east in the Pacific. However, the routes of this prehistoric translocation have not been tested using nuclear DNA. Here we use high-density nuclear data composed of > 500,000 single-nucleotide polymorphic markers in 51chickens of indigenous, wild, or commercial origin to explore the population and dispersal history of chickens in the Asia-Pacific region. Our limited analysis of this data reveals that chickens from Island Southeast Asia are distinct from chickens in mainland Asia. It also shows that red jungle fowl in the Philippines are genetically distant from the local domestic chickens, which potentially confirms the natural or endemic status of red jungle fowls, called ‘labuyo’ in the Philippines.

Keywords

Chickens, Asia-Pacific, SNPs, translocation, domestication

Introduction

Chickens in Island Southeast Asia (ISEA) are an important component of local societies as they serve numerous utilitarian and religious purposes (*e.g.* Crawford 1990; Tixier-Boichard *et al*. 2011). Chickens in ISEA consist of both indigenous populations of free ranging domestic chickens as well as a wild endemic version of the red jungle fowl (RJF) (*Gallus gallus*), which has been involved in multiple domestication processes across South and Southeast Asia (Liu *et al*. 2006; Chapter 2). It does not appear that the domestic chicken (*Gallus gallus domesticus*) is purely RJF, as there are three other species of the genus *Gallus* that may have played a role in the domestication process: Sri Lankan jungle fowl (*Gallus lafayetti*), grey jungle fowl (*Gallus sonneratii*), and green jungle fowl (*Gallus varius*). Hybridisation between these species and the RJF does not produce fertile offspring (*e.g*. Bekisar in Indonesia) (Artawan 1996). However, recent research has revealed the absence of the yellow skin gene in wild RJF that is found in both domestic birds and grey jungle fowl, which suggests hybridisation with the grey jungle fowl during or after the domestication of *G. g. domesticus* (Eriksson *et al*. 2008).

In addition to hybridisation with other congeneric species, the domestication process that resulted in *G. g. domesticus* may also have involved all four RJF subspecies, the *G. g. gallus* (Indochina RJF), *G. g. spadiceus* (Burmese RJF), *G. g. jabouillei* (Vietnamese RJF), and *G. g. bankiva* (domestic Java RJF). However, other than *G. g. bankiva* almost nothing is known about the population history of RJF in

ISEA. In particular, which subspecies does the Philippine RJF, locally known as 'labuyo', belong to? Although labuyo are abundant in the jungles across the Philippines it is unclear whether these chickens are a wild endemic species of RJF or a feral population of *G. g. domesticus*. However, speculation exists that an independent domestication of labuyo might have occurred in the Philippines sometime during the Holocene (Mudar 1997).

The indigenous domestic chickens in ISEA, particularly those found in the Philippines, are important because they represent likely descendants of chickens that were translocated into the Pacific during prehistory (Thomson *et al.* 2014; Chapter 2). In the colonisation process of the Pacific, chickens, along with a range of other organisms, served as an important food resource during long-distance sea voyages and on newly colonised islands (Kirch 1997). Several scholars have previously looked at this settlement process using the geographic distribution of genetic lineages of chickens in the Pacific (Storey *et al.* 2007; Gongora *et al.* 2008; Storey *et al*. 2010; Dancause *et al.* 2011; Thomson *et al*. 2014), and this has improved our understanding of human dispersal. However, these studies were based on short fragments of mitochondrial DNA (mtDNA) and, the population and dispersal history of chickens has not yet been examined using nuclear DNA before.

Advances in genome technology, allowing rapid and inexpensive access to numerous loci or genomic regions, can improve inference of population and evolutionary history (Luikart *et al*. 2003). One of the most promising approaches in population genomics is the use of genome-wide single nucleotide polymorphisms (SNPs). These unlinked nuclear genetic markers are considered the most widespread

sequence variation in the genome and they are believed to estimate population genetic parameters well (Brumfield *et al*. 2003; Morin *et al*. 2004). SNPs are thus better able to reveal a genome-wide picture of population history (Nielsen 2000; Hare 2001). This has fuelled the use of genome-wide SNP data for studies investigating the population and evolutionary history of domestic (e.g., Pollinger *et al.* 2010; McCue *et al.* 2012) and non-model organisms (*e.g.* Moura *et al*. 2014; Spinks *et al*. 2014; Moura *et al*. 2015). Similarly, genome-wide SNP analyses have also been used to study admixture events in human populations (*e.g*. Lipson *et al.* 2014; Pierron *et al.* 2014) where it has revealed the rich complexities of past human history.

Although the application of genome-wide SNPs is promising there are certain challenges including: 1) issues with ascertainment bias; 2) inclusion of non-neutral loci; and 3) the sheer number of loci generated in population genomic analysis (Helyar *et al*. 2011). Firstly, SNP genotyping arrays, which are one method of generating nuclear SNP datasets, may contain a bias due to the way the SNPs were 'ascertained' that can distort inferences made about the demographic history of a population (Lachance & Tishoff 2013). Ascertainment bias stems from the initial discovery of SNP markers used in designing the array, where often a small panel of individuals is used with rare alleles not usually represented (Gravel *et al.* 2011). This mainly happens when SNPs that are discovered in one population are then used to genotype another (Nielsen 2004). Ascertainment bias can result in errors in estimating population genetic parameters such as measures of population differentiation, linkage disequilibrium, and selection scans (Lachance & Tishkoff 2013). Secondly, the genotyping of thousands of genetic markers results in loci that

are evolutionary neutral as well as those under selection, and behaviour of each type can potentially vary from the other (Luikart *et al.* 2003). Neutral loci usually reflect demography and the evolutionary history of a population or species, whereas selected loci often display outlier patterns of variation due to the selection they experience (Luikart *et al.* 2003). Thus, it is essential that outlier loci are detected and removed in order to infer a reliable demographic and evolutionary history. Lastly, population genetic analysis based on SNP data is theoretically straightforward, but the number of loci (100,000's) being genotyped can make calculation of basic descriptive statistics difficult. However, population genetic software programs are increasingly improving their capacity to handle large data sets (Helyar *et al.* 2011).

With these considerations on the power of genome-wide SNPs in reconstructing evolutionary history, we explore the population history of chickens in MSEA, the Philippines, and the Pacific. We aim to address the question of whether, and to what extent, chickens in Asia are related to chickens further east in ISEA and the Pacific. We also examine how different the inferred history based on nuclear markers is from that based on mitochondrial DNA. To do this we performed a genome-wide analysis of >500,000 SNP markers on indigenous chickens from mainland Asia, the Philippines, the Pacific, and on Asian domestic breeds and examined their genetic differentiation and population structure.

Materials and methods

Samples and SNP genotyping array

DNA samples from previously extracted chicken feathers from the Philippines and Pacific Islands were evaluated for quality and DNA concentration (Chapter 2). The samples selected for SNP genotyping had ratios of absorbance that did not deviate from 1.8 using the NanoDrop-1000 (NanoDrop Technologies, Inc.), however, the concentrations of the samples varied widely between 30 and 500 ng/µL. Based on mitochondrial haplogroup results, most of the samples belong to chicken mitochondrial haplogroup D. This mitochondrial lineage was chosen because it is ubiquitous in ISEA and the Pacific (Chapter 2). We compared thirty-eight samples selected from our previous mitochondrial work (thirty-four samples from five Philippine islands including five labuyo chickens, and four samples from the Pacific), with eleven domestic Asian breeds, and two red jungle fowl from Asia. Details on all fifty-one individuals used in the analyses are shown in Table 5-1.

Table 5-1. List of chicken samples successfully genotyped and used in the study.

Population/Region	Sampling Locality/Breeds*	Sample Reference
Vanuatu, Pacific	Takara	A11751
	Eton	A11742
	Malakula	A11837
	Pangpang	A11746
Batanes, Philippines	Batanes	A13693
	Batanes	A13692
	Batanes	A13693
Palawan, Philippines	Palawan[+]	A13714
	Palawan	A13736
	Palawan	A13743
	Palawan	A13747
	Palawan	A13760
	Palawan[+]	A13724
	Palawan	A13717
Mindoro, Philippines	Mindoro[+]	A13768
	Mindoro	A13773
	Mindoro	A13807
	Mindoro	A13776
	Mindoro[+]	A13770
	Mindoro	A13777
Mindanao, Philippines	Bukidnon, Mindanao	A13834

	Cotabato, Mindanao	A13562
	Talakag, Mindanao	A13528
	Lurugan, Mindanao	A13561
	Aya-Aya, Mindanao	A13502
	Lugait, Mindanao	A13507
	Zamboanga, Mindanao	A13590
	Davao, Mindanao	A13676
	Libertad, Mindanao	A13519
	Naawan, Mindanao	A13518
	Zamboanga, Mindanao	A13614
	Ifugao, Luzon	A13641
	Cagayan Valley, Luzon+	A13767
	Kalinga, Luzon	A13705
Luzon, Philippines	Ifugao, Luzon	A13664
	Quirino, Luzon	A13688
	Ifugao, Luzon	A13657
	Ifugao, Luzon	A13665
	Silkies black	1688
	Silkies white	2405
	Brahma gold	360
	Brahma light	383
	Cochin black	580
	Marans copper	1287
Mainland Asia domestic breeds	Orphington	1381
	Sundheimer light	1769
	Wyandotte white	1880
	Wyandotte silver	2166
	Toutenkou	2509
	Gallus gallus spadiceus+	2621
	Gallus gallus gallus+	2760

+Red jungle fowl samples

The genotyping of DNA samples of indigenous domestic chickens from the Philippines and Pacific was done by the Federal Research Institute for Animal Health (SYNBREED Project), Germany using a 600 K Affymetrix® Axiom® HD genotyping array to generate 580, 961 SNPs. This is the first high-density genotyping array developed for chickens. A detailed description for this genotyping array from SNP discovery to final array validation is reported in Kranis *et al*. (2013). The domestic Asian chickens and the two mainland RJF genotypes were provided by the SYNBREED Project from previous work. Additional details about this research group are available in http://www.synbreed.tum.de.

Detection of loci under selection & population structure analysis

Altogether 580,961 SNPs were used in the analyses. Initially, the entire dataset was used to explore population structure and relationships between samples using principal component analysis (PCA) and maximum likelihood (ML) inference. The PCA analyses were performed using the smartPCA function of the EIGENSOFT 4.0 software (Patterson *et al.* 2006) and the ML analysis was done using RaXML (Stamatakis 2006). The ML tree was performed using the MULTIGAMMA model with bootstrapping via 500 iterations followed by an optimised likelihood search.

Considering that loci under selection can potentially hinder the accuracy of in inferred population histories (Luikart *et al*. 2003), these loci were identified and removed to create a neutral loci dataset. To remove loci under selection, the outlier F_{ST} approach (Beaumont & Nichols 1996) implemented in LOSITAN (Antao *et al*. 2008) with the selection detection workbench for co-dominant markers was used. This was used to detect putative signs of selection for each of the SNP loci using a bisection algorithm simulated for 50,000 iterations at a significance *P* value of 0.005 and a false discovery rate of 0.1. Loci with unusually high F_{ST} are putatively under positive selection; in contrast loci with very low F_{ST} are under stabilising selection. Simulations were done on batches of 10,000 SNPs as LOSITAN cannot test large numbers of SNPs in one simulation run. Both loci under positive and stabilising selection were identified and removed. To examine the clustering of populations based only on the neutral SNPs, we used the same smartPCA method as above (Patterson *et a*l. 2006).

Results

Phylogenetic inference

The ML tree clearly illustrates that Asian domestic breeds are distinct from indigenous domestic chickens found in the Philippines and the Pacific (Figure 5-1). Chickens from the Batanes archipelago (orange stars in Figure 5-1) appear to have grouped together and share a lot of common alleles as illustrated by the long branch leading to the Batanes nodes. The red jungle fowl (labuyo) samples from the Philippines, although sampled from different islands, cluster together into a distinct clade (red stars in Figure 5-1). However, for the domestic Philippine samples there is no apparent clustering of individuals based on island. Even from the few non-D mitochondrial haplogroups present, we can see that clustering of individual chickens does not correspond to their mitochondrial haplogroup assignment. We used samples that are predominantly a subset of haplogroup D that contains a combination of 4 mitochondrial SNPs, the 'ancestral Polynesian D motif' (Thomson *et al.* 2014), with several other samples belonging to haplogroup B and E. The mitochondrial haplogroups for the Asian breeds and the Asian red jungle fowls are not known.

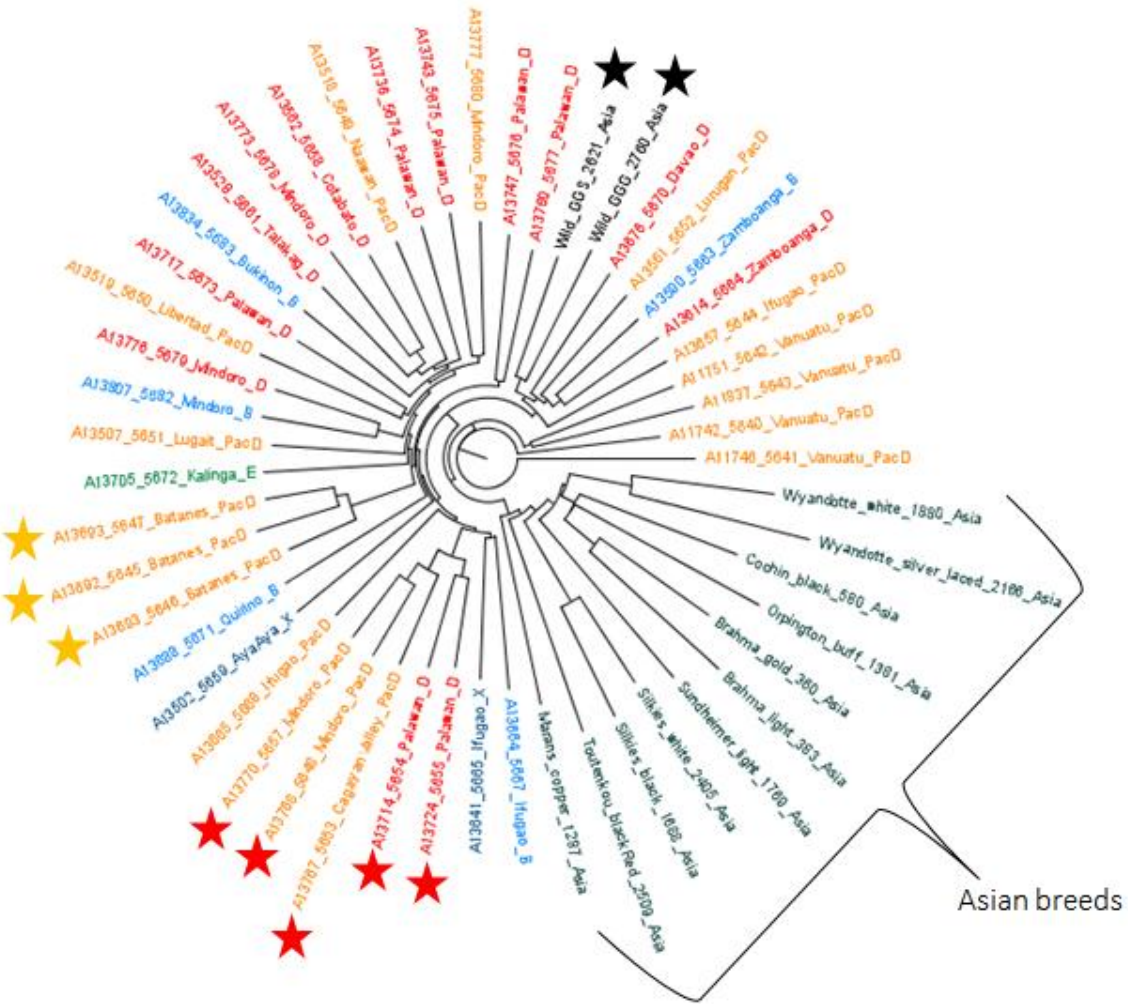


Figure 5-1. Maximum likelihood tree of 51 chickens based on genome-wide SNPs constructed using MULTIGAMMA model. Labels indicate the locality of the samples with colour matching mitochondrial haplogroup assignment. Green – haplogroup E, Light blue – Haplogroup B, Red – Haplogroup D (excluding Polynesian Ds), Orange – Polynesian Ds, Black – *Gallus gallus* subspecies, Navy – Asian domestic breeds. Haplogroup assignment for sample A13503and Asian domestic breeds are not known. Coloured stars represent the following: red – Philippine red jungle fowls, orange – Batanes archipelago chickens, and black – Asian red jungle fowl. For bootstrap support values see Supplementary Figure S1.

Outlier loci detection & neutral population structure

Using LOSITAN (Antao *et al.* 2008), a total of 42,220 SNP loci were identified as outliers and therefore potentially under selection. Specifically 25,469 are under positive selection and 16,751 are under balancing selection (Figure 5-2).

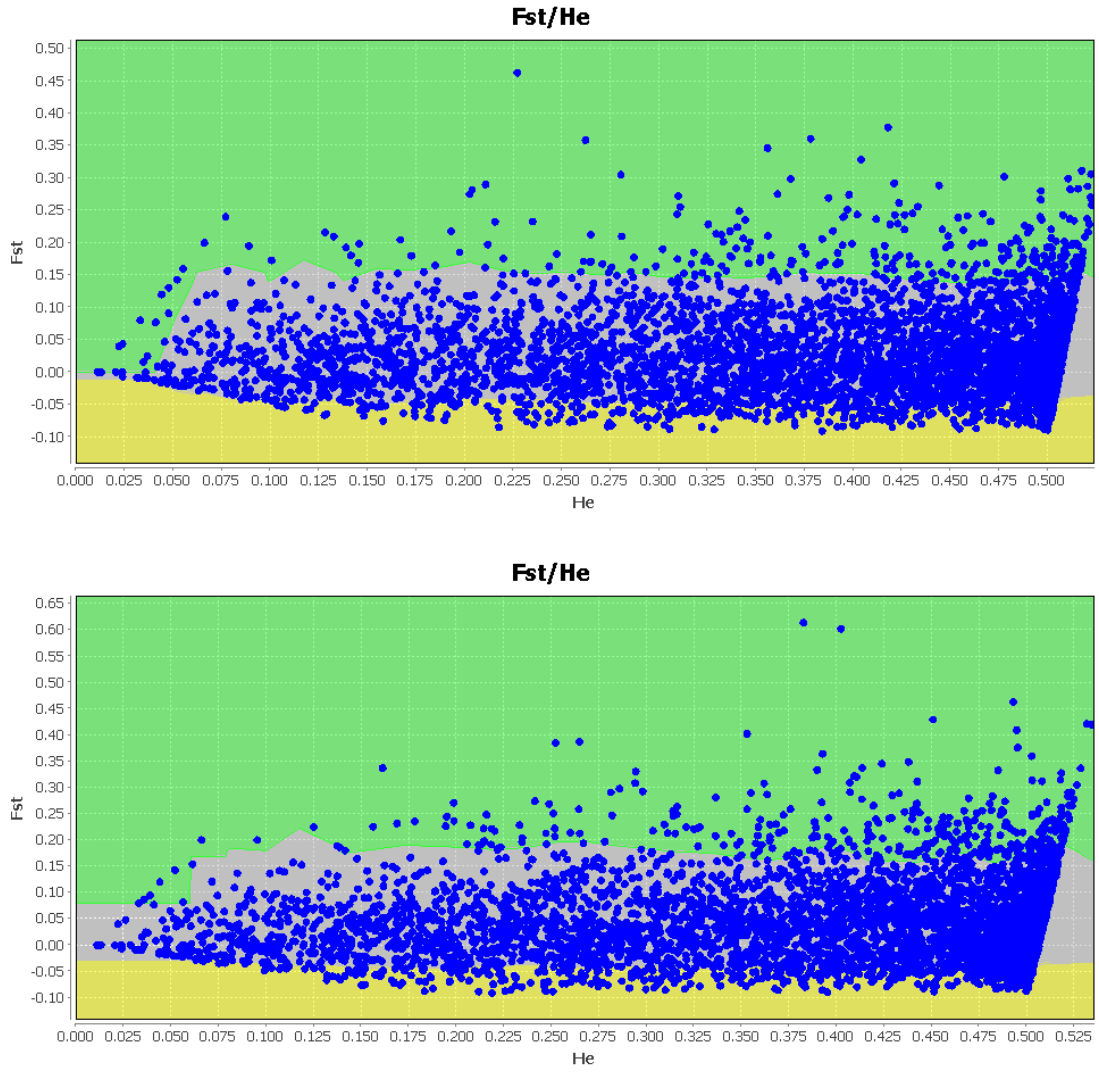


Figure 5-2. Two examples of LOSITAN plots identifying outlier SNP loci, with blue dots representing SNPs. Loci in the green field are potentially under positive selection, loci in the yellow field are under balancing selection, and in the grey field are neutral SNPs. 'He' stands for heterozygosity.

The PCA implemented in EIGENSOFT (Patterson *et al*. 2006) using only the neutral SNPs seems to clearly distinguish between Philippine, Asian, and Pacific populations (Figure 5-3). The first two components clearly cluster three groups: the labuyo (Philippine RJF, outlined in red), indigenous domestic chickens found in the Philippines (outlined in green), and Pacific (from Vanuatu, outlined in blue) with the Asian domestic breed scattered across the plot (Figure 5-3A). There is some geographic clustering of individuals from certain localities in the Philippines, namely chickens from the Ifugao and Batanes when the Asian domestic breeds are removed (Figure 5-3B). However, only Batanes retains that separate clustering when eigenvectors 3 and 4 are used (Figure 5-4). Ifugao falls close to other localities found in the Philippines. It is also notable that the labuyo (Philippine RJF) appear to be genetically different from the Philippine indigenous domestic chicken. Furthermore, mainland RJF (*G. g. gallus* and *G. g. spadiceus*) appear to be genetically closer to Philippine indigenous domestic chickens than to the Philippine RJF. Additionally, there is an absence of any observable clustering among mainland Asian chicken breed in the PCA plot. The PCA plot using both selected and neutral SNPs did not differ significantly to the plot that used only neutral SNPs (Supplementary Figure S5-2).

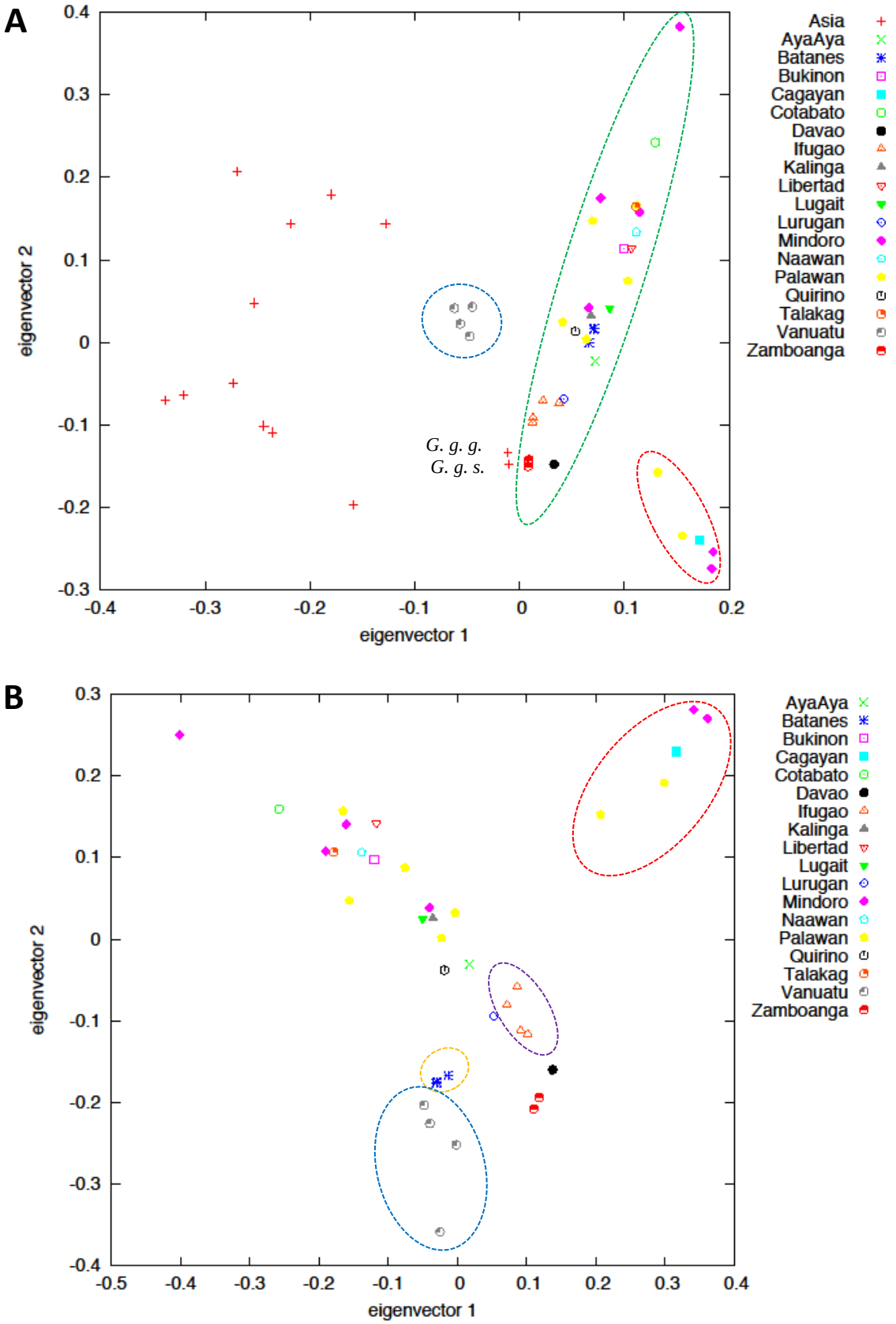


Figure 5-3. Principal coordinate analysis performed using Eigensoft on genome-wide neutral SNPs of chickens from (A) the Asia-Pacific region. Dashed outlines represent the following groups: green – Philippine indigenous chickens, red – Philippine red jungle fowl, blue – Vanuatu. The Mainland Asia RJF samples are labelled as: *G. g. g.* – *Gallus gallus gallus* (Indochina) or *G. g. s.* – *Gallus gallus spadiceus* (Burmese RJF); and (B) the Philippines and Pacific only. Dashed outlines represent the

following groups: red – Philippine RJF, blue – Vanuatu, orange – Batanes, purple – Ifugao populations.

The PCA plot using the components 3 and 4 and without the Asian breeds creates three distinct clusters, namely the Batanes chickens, Philippine indigenous chickens, and Philippine RJF (Figure 5-4). Batanes chickens form a distinct and distant cluster from the rest of the indigenous chickens in the Philippines.

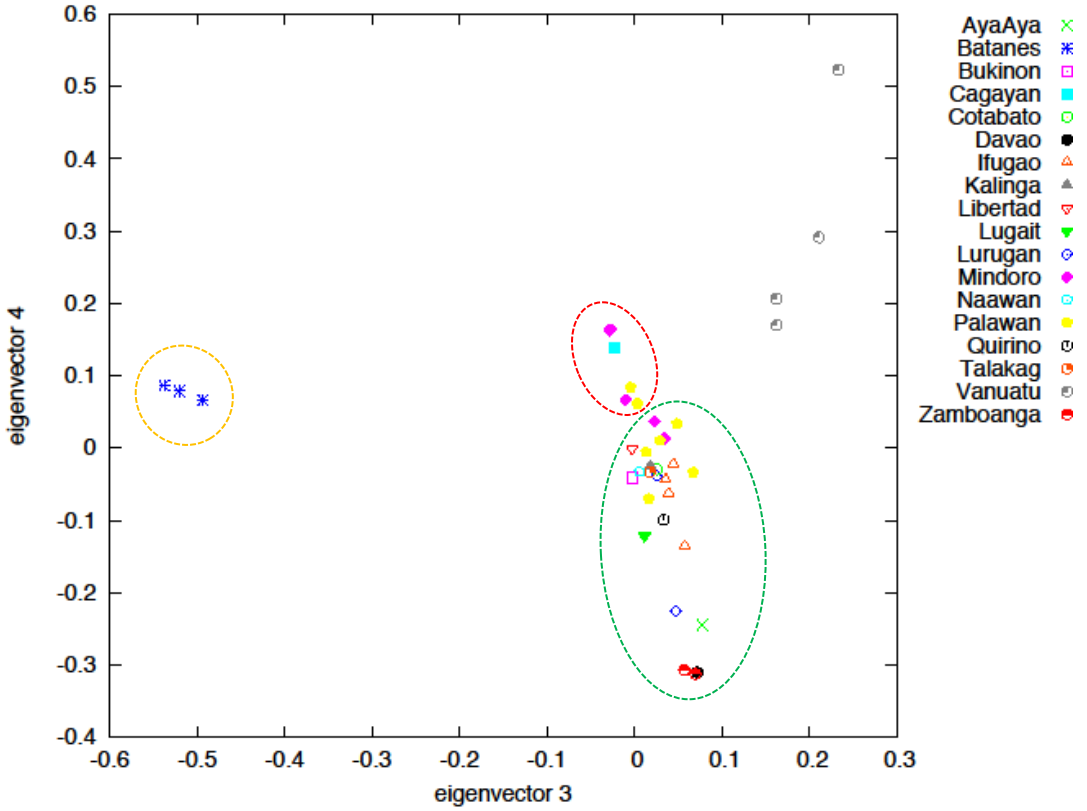


Figure 5-4. Principal coordinate analysis performed using Eigensoft on genome-wide neutral SNPs of chickens from the Philippines and the Pacific using eigenvectors 3 and 4. Dashed outlines represent the following groups: red – Philippine RJF, orange – Batanes, green – Philippine indigenous chickens.

Discussion

The stark separation between chickens found in Asia and the Philippines as revealed by the phylogenetic tree (Figure 5-1) and PCA plot (Figure 5-3A) is concordant with observation based on mitochondrial data (Chapter 2). This suggests that Asian and Philippine domestic chickens experienced separate population histories, probably owing to vicariant events. There is, however, a close positioning of the two Asian RJF sub-species (*Gallus gallus gallus* and *Gallus gallus spadiceus*) in relation to Philippine indigenous domestic chickens. The range of these two Asian RJF subspecies include south China, Myanmar, Thailand, Laos, Cambodia, Vietnam, and down to the Malay peninsula (Johnsgard 1999). Therefore, the ancestors of the Philippine indigenous domestic chickens were most likely domesticated in mainland Southeast Asia from Asian RJF before being introduced into the Philippines. However, in the absence of corroborating zooarchaeological evidence, the timing and processes surrounding their introduction into the Philippines cannot be ascertained. The position of the Vanuatu chickens indicates a closer relationship to the Philippine indigenous domestic chickens than to the Philippine RJF or Asian domestic breeds. This is consistent, with previous mitochondrial studies indicating that the chickens translocated into the Pacific were derived from the Philippine haplotypes (Thomson *et al.* 2014; Chapter 2). The tight clustering of Vanuatu chickens and the separation seen in both the phylogenetic tree and PCA may be caused by genetic drift or a bottleneck during the migration process.

The separation of Philippine RJF from the rest of the indigenous domestic chickens found in the Philippines seems to suggest a different population history for these wild endemic ‘labuyo’ chickens. It is not known whether they represent a

natural dispersal of RJF into the Philippines during the Pleistocene when sea level was significantly lower or an earlier episode of human mediated introduction of chickens from mainland Asia into the Philippines. Due to the genetic distance from the Philippine indigenous domestic chickens and their morphological differences (Delacour 1977), it is more likely that the labuyo (Philippine RJF) represent a subspecies of RJF isolated in the Philippines when sea levels rose. Additionally, these wild 'labuyo' chicken populations are ubiquitous in tropical jungles of the Philippines. If the samples from Palawan, Mindoro, and Luzon used in this study are representative of the Philippine RJF, then it appears that they experienced very little genetic mixing with indigenous chickens.

The genetic distance of Batanes indigenous chickens from the rest of the Philippine indigenous domestic chickens is interesting (Figure 5-3B). The Batanes archipelago is located in between Taiwan and the Philippine island of Luzon. The islands here were crucial during the initial spread of the Austronesian speakers from Taiwan into the Philippines (Bellwood *et al.* 1995). The distinctness of the chickens on the island appears to reflect their protracted presence on the archipelago. Furthermore, the long branch leading to the Batanes samples suggests the presence on this island of a constant population that shares a lot of common alleles (Figure 5-1). This could indicate long-term management of chickens leading to the development of a distinct chicken lineage. The genetic isolation of chickens in Batanes is also reflected in another island endemic on the island of Lanyu, southeast of Taiwan - the Lanyu pigs (Wu *et al.* 2007; Larson *et al.* 2010). Distinct from East Asian and European pigs (Chang *et al*. 2009), this miniature pig breed is believed to have been locally developed on Lanyu. Although genetic drift due to lack of gene

flow cannot be ruled out altogether, the distinct lineages of these two domesticates on adjacent islands could potentially indicate that the Austronesians engaged in the act of domestication with livestock on island environments. Additionally, a 4000-year-old domestic pig recovered from Nagsabaran, Cagayan Valley on the northern Philippine island of Luzon (Piper *et al.* 2009) has a similar mitochondrial DNA signature to the Lanyu pigs (pers. comm. Piper). This potentially indicates the movement of domesticates from Taiwan into northern Philippines. Altogether, the population history of chickens in the Philippines appears complex, although it is clear that the genetic lineage brought into the Pacific was derived from the Philippines. Multiple episodes of arrivals of chickens into the Philippines archipelago from Asia, then the protracted isolation, and subsequent introductions of chickens likely gave rise to this complex history.

One limitation of not having nuclear SNPs from Indonesian samples is that we cannot compare these results to those gathered from mtDNA data, which suggests a lack of Indonesian contributions for the initial translocation of chickens into the Pacific (Chapter 2). We also cannot determine the extent of shared population history for Philippine chickens with those from Indonesia. When available, this will certainly allow us to understand the population history of chickens in the entirety of ISEA, and thus provide a better foundation for discussing how past people in ISEA managed and translocated chickens within the region and into the Pacific.

Conclusion

This study is the first to use genome-wide SNPs to explore the population history of chickens in the Asia-Pacific region. Although, the study could further benefit from adding genotypes from indigenous chickens in Asia, Indonesia, and the Pacific we have clearly shown the link between Asia, the Philippines and the Pacific. This study has demonstrated the separate histories of chickens on mainland Asia and the Philippines, at the same time also showing how chickens in the ISEA-Pacific region are ultimately derived from mainland Asia. Furthermore, we have suggested that the Philippine RJF (labuyo) could represent an endemic wild chicken population that naturally dispersed into the Philippines, specifically on the islands of Palawan and Mindoro during the Pleistocene. This study also hints at the possibility that the Austronesian-speaking people might have engaged more in livestock management than previously envisaged. Although crucial to farming societies, domestic animals could also have precipitated cultural expansion and trade, which is crucial in understanding the processes behind the expansion of the Austronesian-speaking people. This emerging pattern warrants further investigation and integration with current archaeological data in Luzon and Taiwan.

Acknowledgment

This research was funded by the Australian Research Council (ARC) Discovery Project ID DP110105187. We thank Steffen Weigand of the SYNBREED project for genotyping chicken samples from the Philippines and the Pacific, and for providing us with additional genotypes from Asia. We also acknowledge and extend our gratitude to Julien Soubrier for assisting us in the data processing and analyses.

References

Antao, T, A Lopes, RJ Lopes, A Beja-Pereira, G Luikart. 2008. LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. BMC bioinformatics 9:323.

Artawan, IK. 1996. Studi tentang fertilitas dan karyotipe bekisar (Hibrid interspesifik Gallus): Universitas Gadjah Mada.

Beaumont, MA, RA Nichols. 1996. Evaluating loci for use in the genetic analysis of population structure. Proceedings of the Royal Society of London. Series B: Biological Sciences 263:1619-1626.

Bellwood, PS, JJ Fox, DT Tryon. 1995. The Austronesians: historical and comparative perspectives: Australian National University.

Brumfield, RT, P Beerli, DA Nickerson, SV Edwards. 2003. The utility of single nucleotide polymorphisms in inferences of population history. Trends in Ecology & Evolution 18:249-256.

Chang, W, H Chu, Y Jiang, S Li, Y Wang, C Chen, K Chen, C Lin, Y Ju. 2009. Genetic variation and phylogenetics of Lanyu and exotic pig breeds in Taiwan analyzed by nineteen microsatellite markers. Journal of animal science 87:1-8.

Crawford, RD. 1990. Poultry breeding and genetics: Elsevier.

Dancause, KN, MG Vilar, R Steffy, JK Lum. 2011. Characterizing genetic diversity of contemporary pacific chickens using mitochondrial DNA analyses. PloS one 6:e16843.

Delacour, J. 1977. Pheasants of the World.

Eriksson, J, G Larson, U Gunnarsson, B Bed'hom, M Tixier-Boichard, L Strömstedt, D Wright, A Jungerius, A Vereijken, E Randi. 2008. Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. PLoS genetics 4:e1000010.

Gongora, J, NJ Rawlence, VA Mobegi, H Jianlin, JA Alcalde, JT Matus, O Hanotte, C Moran, JJ Austin, S Ulm. 2008. Indo-European and Asian origins for Chilean and Pacific chickens revealed by mtDNA. Proceedings of the National Academy of Sciences 105:10308-10313.

Gravel, S, BM Henn, RN Gutenkunst, AR Indap, GT Marth, AG Clark, F Yu, RA Gibbs, CD Bustamante, DL Altshuler. 2011. Demographic history and rare allele sharing among human populations. Proceedings of the National Academy of Sciences 108:11983-11988.

Hare, MP. 2001. Prospects for nuclear gene phylogeography. Trends in Ecology & Evolution 16:700-706.

Helyar, S, J Hemmer-Hansen, D Bekkevold, M Taylor, R Ogden, M Limborg, A Cariani, G Maes, E Diopere, G Carvalho. 2011. Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Molecular Ecology Resources 11:123-136.

Johnsgard, PA. 1999. Pheasants of the World: Smithsonian Institution Press:USA.

Kirch, PV. 1997. The Lapita peoples : ancestors of the oceanic world. Cambridge, Mass.: Blackwell Publishers.

Kranis, A, AA Gheyas, C Boschiero, F Turner, L Yu, S Smith, R Talbot, A Pirani, F Brew, P Kaiser. 2013. Development of a high density 600K SNP genotyping array for chicken. BMC genomics 14:59.

Lachance, J, SA Tishkoff. 2013. SNP ascertainment bias in population genetic analyses: Why it is important, and how to correct it. Bioessays 35:780-786.

Larson, G, R Liu, X Zhao, J Yuan, D Fuller, L Barton, K Dobney, Q Fan, Z Gu, X-H Liu. 2010. Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. Proceedings of the National Academy of Sciences 107:7686-7691.

Lipson, M, P-R Loh, N Patterson, P Moorjani, Y-C Ko, M Stoneking, B Berger, D Reich. 2014. Reconstructing Austronesian population history in Island Southeast Asia. bioRxiv.

Liu, Y-P, G-S Wu, Y-G Yao, Y-W Miao, G Luikart, M Baig, A Beja-Pereira, Z-L Ding, MG Palanichamy, Y-P Zhang. 2006. Multiple maternal origins of chickens: out of the Asian jungles. Molecular phylogenetics and evolution 38:12-19.

Luikart, G, PR England, D Tallmon, S Jordan, P Taberlet. 2003. The power and promise of population genomics: from genotyping to genome typing. Nature Reviews Genetics 4:981-994.

McCue, ME, DL Bannasch, JL Petersen, J Gurr, E Bailey, MM Binns, O Distl, G Guérin, T Hasegawa, EW Hill. 2012. A high density SNP array for the domestic horse and extant Perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. PLoS genetics 8:e1002451.

Morin, PA, G Luikart, RK Wayne. 2004. SNPs in ecology, evolution and conservation. Trends in Ecology & Evolution 19:208-216.

Moura, A, J Kenny, R Chaudhuri, M Hughes, R Reisinger, P de Bruyn, M Dahlheim, N Hall, A Hoelzel. 2015. Phylogenomics of the killer whale indicates ecotype divergence in sympatry. Heredity 114:48-55.

Moura, AE, JG Kenny, R Chaudhuri, MA Hughes, A J Welch, RR Reisinger, P Bruyn, ME Dahlheim, N Hall, AR Hoelzel. 2014. Population genomics of the

killer whale indicates ecotype evolution in sympatry involving both selection and drift. Molecular ecology 23:5179-5192.

Mudar, KM. 1997. Patterns of animal utilization in the Holocene of the Philippines: A comparison of faunal samples from four archaeological sites.

Nielsen, R. 2000. Estimation of population parameters and recombination rates from single nucleotide polymorphisms. Genetics 154:931-942.

Nielsen, R. 2004. Population genetic analysis of ascertained SNP data. Human genomics 1:218-224.

Patterson, N, AL Price, D Reich. 2006. Population structure and eigenanalysis. PLoS genetics 2:e190.

Pierron, D, H Razafindrazaka, L Pagani, F-X Ricaut, T Antao, M Capredon, C Sambo, C Radimilahy, J-A Rakotoarisoa, RM Blench. 2014. Genome-wide evidence of Austronesian–Bantu admixture and cultural reversion in a hunter-gatherer group of Madagascar. Proceedings of the National Academy of Sciences 111:936-941.

Piper, PJ, H-c Hung, FZ Campos, P Bellwood, R Santiago. 2009. A 4000 year-old introduction of domestic pigs into the Philippine Archipelago: implications for understanding routes of human migration through Island Southeast Asia~ and Wallacea. Antiquity 83:687-695.

Pollinger, JP, KE Lohmueller, E Han, HG Parker, P Quignon, JD Degenhardt, AR Boyko, DA Earl, A Auton, A Reynolds. 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature 464:898-902.

Spinks, PQ, RC Thomson, HB Shaffer. 2014. The advantages of going large: genome-wide SNPs clarify the complex population history and systematics of the threatened western pond turtle. Molecular ecology 23:2228-2241.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688-2690.

Storey, AA, JM Ramírez, D Quiroz, DV Burley, DJ Addison, R Walter, AJ Anderson, TL Hunt, JS Athens, L Huynen. 2007. Radiocarbon and DNA evidence for a pre-Columbian introduction of Polynesian chickens to Chile. Proceedings of the National Academy of Sciences 104:10335-10339.

Storey, AA, M Spriggs, S Bedford, SC Hawkins, JH Robins, L Huynen, E Matisoo-Smith. 2010. Mitochondrial DNA from 3000-year old chickens at the Teouma site, Vanuatu. Journal of Archaeological Science 37:2459-2468.

Thomson, VA, O Lebrasseur, JJ Austin, TL Hunt, DA Burney, T Denham, NJ Rawlence, JR Wood, J Gongora, LG Flink. 2014. Using ancient DNA to

study the origins and dispersal of ancestral Polynesian chickens across the Pacific. Proceedings of the National Academy of Sciences 111:4826-4831.

Tixier-Boichard, M, B Bed'hom, X Rognon. 2011. Chicken domestication: From archeology to genomics. Comptes rendus biologies 334:197-204.

Wu, C, Y Jiang, H Chu, S Li, Y Wang, Y Li, Y Chang, Y Ju. 2007. The type I Lanyu pig has a maternal genetic lineage distinct from Asian and European pigs. Animal genetics 38:499-505.

Supplementary information

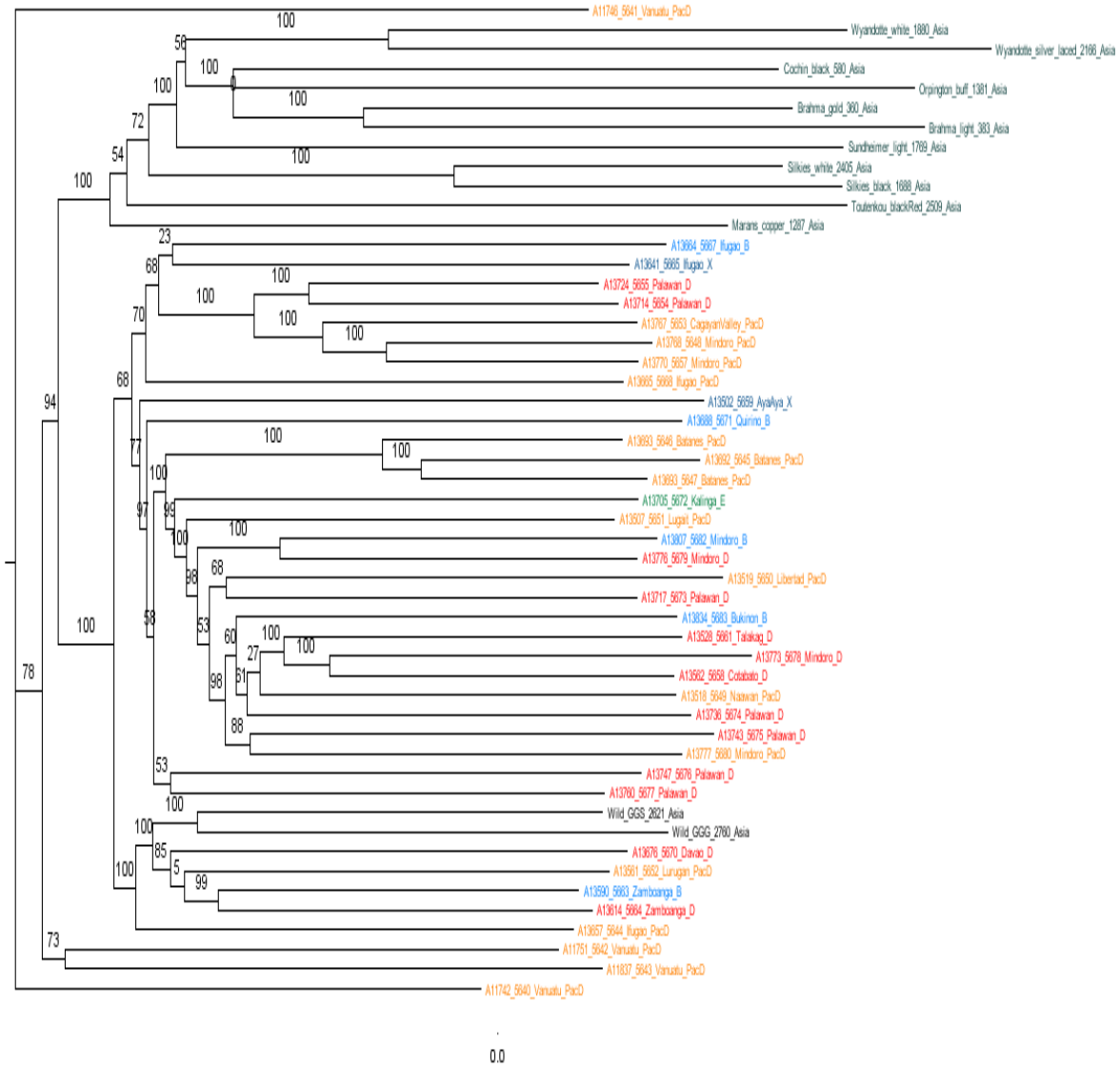


Figure S5-1. Maximum likelihood tree of 51 chickens based on genome-wide SNPs constructed using MULTIGAMMA model showing bootstrap supports.

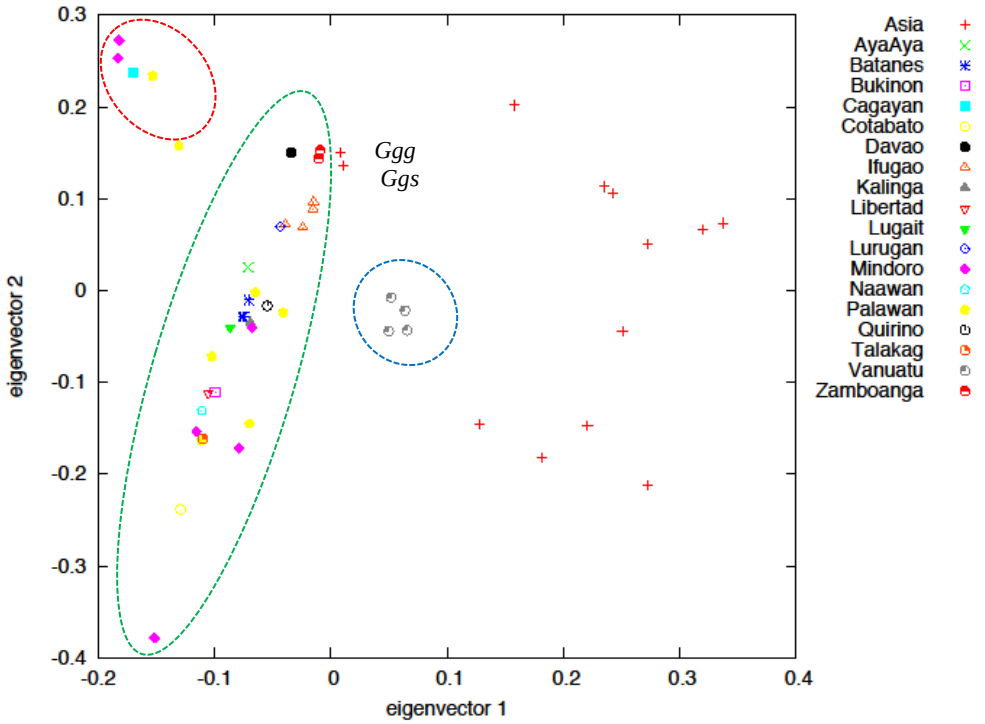


Figure S5-2. Principal coordinate analysis performed using Eigensoft on genome-wide of the combined neutral and selected SNPs of chickens from the Asia-Pacific region. Dashed outlines represent the following groups: green – Philippine indigenous chickens, red – Philippine red jungle fowl, blue – Vanuatu. The Mainland Asia RJF samples are labelled as: *Ggg* – *Gallus gallus gallus* (Indochina) or *Ggs* – *Gallus gallus spadiceus* (Burmese RJF).

CHAPTER 6: Concluding discussion

Concluding discussion

Here, I (1) review the overall aims of the thesis and give an overview about the significance of the findings based on the current understanding of prehistoric translocations of chickens in the Indo-Pacific region, and its inferences about the Austronesian expansion; and (2) outline the limitations of the work and some recommendations are forwarded for future directions on the use of parallel histories for understanding the human past. The overall goal of the thesis was to use modern and ancient DNA to understand the population history of domesticated chickens in ISEA, reconstruct the translocation history of domesticated chickens from ISEA into the Pacific and Indian Oceans, and thus provide new data on the prehistoric expansion of the Austronesian-speaking people. The specific aims of the thesis were to:

(1) characterise the genetic diversity, structure, and relatedness of indigenous chickens across ISEA and compare with those in the Pacific and Indian Oceans,

(2) reconstruct the Austronesian expansion and interactions in the Indo-Pacific region using chicken as a biological proxy in light of current archaeological and linguistic knowledge, and

(3) assess the resolution of human-mediated dispersal histories of chickens based on mitochondrial DNA control region, whole mitochondrial genomes, and genome-wide nuclear SNPs.

Aim 1: Genetic characterisation of Indo-Pacific chickens

Summary of findings and significance: The molecular characterization of indigenous domestic chickens from ISEA has never been done with such a large number of samples spread across the entire region. When combined with the available published data from adjacent regions in MSEA and the Pacific, the population history of chickens in ISEA and their geographic context can be explored more fully than ever before. The data assembled from indigenous chickens in ISEA has filled significant gaps in our understanding of the origin of chickens translocated into the Pacific and Indian Oceans. Furthermore, the level of molecular resolution created in this study far surpasses anything previously achieved. This resolution is demonstrated in the progression of this thesis from firstly using mitochondrial control region data (Chapter 2 and 3), before progressing to whole mitochondrial genomes (Chapter 4), and finally genome-wide nuclear SNPs (Chapter 5) for important lineages in Oceania. These studies reveal that chickens in ISEA have a unique evolutionary history that is closely tied to the geography of the region, the expansion of human groups particularly the Austronesian-speakers, and the complexities involved in the development of agricultural life-ways in the region.

Using the mitochondrial control region, Chapters 2 and 3 demonstrate that modern chickens in Indonesia, the Philippines, the Pacific, and Madagascar mainly belong to a specific mitochondrial lineage called haplogroup D. This is in contrast to MSEA and South Asia where all known mitochondrial haplogroups are observed, although at different frequencies (Liu *et al.* 2006; Miao *et al.* 2012). A few other haplogroups are also observed in ISEA (haplogroups A, B, C, E and H) but they are a minority. As haplogroup D is the dominant lineage in ISEA, the Pacific, and

Madagascar, its history has important implications for the population and translocation history of chickens in the region. Haplogroup D's importance is also suggested by the level of haplogroup D diversity observed in ISEA today, indicating an early introduction/expansion several millennia ago. The insular geography of ISEA must have played a crucial role in the level of diversity observed within haplogroup D today. Another major finding from Chapter 2 supports the recently reported Philippine homeland for the Polynesian D chickens, which contains a unique group of haplotypes, called the Polynesian D haplotypes (Thomson *et al.* 2014b). This distinct group of haplotypes contains a unique set of mutations that characterise contextually secure archaeological chickens in the Pacific. This haplotype was also found in samples from the Philippines and in one sample from Hainan. Chapter 3 similarly shows that chickens in Madagascar mainly belong to haplogroup D. The lack of diversity and the star-like pattern of diversification are also seen from these samples hinting at the limited translocation history of chickens on the island.

The whole mitochondrial genome analysis that I performed was predominantly focused on Polynesian Ds from different ISEA and Pacific islands, and revealed interesting aspects about the population history of chickens in the region. In Chapter 4, the additional nucleotide polymorphisms outside the mitochondrial control region demonstrate that although the Polynesian D lineages were defined based only on control region data, it is still phylogenetically distinct when whole mitochondrial genomes are examined. The distinctness of the Polynesian D lineages suggest that they were isolated from the rest of haplogroup D lineages either before or during translocation. One possibility is that the creation of

the Polynesian D lineages was the result of animal husbandry practices that involved the initial selection of chickens containing the Polynesian D signature, followed by subsequent breeding, and dispersal of this bottlenecked population by humans to remote islands. On the other hand, the Polynesian D signature could have occurred after chickens were actively integrated into the development of agriculture in ISEA and therefore reflect the origin of the transported Polynesian chickens. However, we can be certain that the chickens that were prehistorically transported into the Pacific carried this mitochondrial signature.

Chapter 5 uses genome-wide information that could better reflect the population history of chickens and to date, this work is the first of its kind looking into the population history of chickens in the Asia-Pacific regions from a nuclear genome point-of-view. Although Chapter 5 is only an exploratory investigation of the nuclear SNP dataset, the large number of SNPs used in this study was able to reveal that indigenous chickens found in the Philippines trace their genetic heritage to MSEA. This result is inferred from the fact that Philippine indigenous chickens show a higher affinity with red jungle fowl in MSEA than jungle fowl found in the Philippines. This indicates that the involvement of Philippine red jungle fowls in the animal husbandry practices of early farmers in the region might have been minimal or non-existent.

Furthermore, the genetic distance between red jungle fowl and indigenous domestic chickens in the Philippines could also indicate the arrival of the former predated the arrival of the latter, which might explain why they are very distinct populations. Altogether, this shows that there were several and most likely

temporally distinct episodes of chicken arrivals in ISEA. Several of the findings in this research have applications outside of understanding prehistoric processes in ISEA. Although conservation and livestock improvement is not a priority for this research, my work is nonetheless important for researchers who are aiming to improve the poultry industry, as incorporating Philippine RJF into their breeding programmes could increase the genetic diversity of poultry. This study suggests that the Philippine RJF are a good candidate to explore for population conservation programmes because they are genetically distinct.

Altogether, molecular characterisation of indigenous chickens from ISEA has contributed to our understanding of the population history of chickens in the region and it has revealed five major points. First, the prehistoric translocation of chickens into ISEA mainly involved mitochondrial haplogroup D chickens. Second, while within the Philippines, the Polynesian D signature developed through a mechanism probably involving isolation making it phylogenetically distinct from other haplogroup D lineages. Third, Polynesian D chickens were translocated into the Pacific eventually reaching Remote Oceania. Fourth, nuclear SNP data suggests that indigenous chickens in the Philippines are ultimately derived from red jungle fowl found in MSEA. Lastly, Philippine RJF are a distinct population. These findings highlight the complexities associated with human-mediated animal translocations in the Indo-Pacific region.

Limitations and future directions: This study has added to our understanding of chicken diversity, origins and dispersal in ISEA through dense sampling across Indonesia and the Philippines, However, given the lack of archaeological samples in

the region, these new modern samples can only represent the history of populations surviving to the present day. Although, I argue that the level of haplogroup D diversity in ISEA suggests its long-term presence in the region and thus shows population continuity, the characterisation of archaeological chickens from ISEA would allow for a more direct reconstruction of prehistoric chicken translocations. Archaeological chickens from ISEA before 3000 ya, if found, would be critical for improving current interpretations about the translocation of chickens into and out of the region. If available these chicken remains should be subjected to ancient DNA (aDNA) analysis, preferably using Next Generation Sequencing technologies, as the inclusion of aDNA analysis in reconstructing prehistoric processes is paramount (Storey *et al*. 2012). Ancient DNA from samples in ISEA could suggest when red jungle fowls and/or indigenous chickens first entered the region. The generation of aDNA datasets from ISEA, MSEA, and the Indian Ocean rim would also help improve our understanding of translocation histories in the region. This is the only way to resolve the complicating effects of processes such as trade, exchange networks, and modern day movements of chickens and their population histories.

Key areas, such as Taiwan, would also benefit from further sampling. In this study, only 193 published sequences from Taiwan were used and haplogroup D is not observed on the island. Thus, it cannot be ascertained whether the sequences from Taiwan represents commercial or indigenous chickens or whether haplogroup D was absent prehistorically or they experienced population replacement. An improved sampling regime (both modern and archaeological) from New Guinea, Madagascar, and key areas along the Indian Ocean rim is also desirable.

Nuclear SNP genotypes from Indonesia are also absent in the study (Chapter 5). Generating and assembling SNP genotypes from Indonesia is paramount for investigating the genetic affinity of chickens from the area with those in the Philippines and the potential subsequent movement into the Pacific. Another interesting point that could be investigated with Indonesian SNP data would be whether the red jungle fowl found in Indonesia shows a similar isolated signature to that of the Philippine red jungle fowl. Additional red jungle fowl genotypes from MSEA and ISEA will also be useful in discussing the regional contributions to the domestication process of chickens. Once a more comprehensive dataset is assembled involving chickens from across Austronesia, more advanced analysis can be performed such as quantifying the levels of admixture between distinct populations. Additionally, functional analyses on potentially selected SNP loci should also be investigated to further understand the domestication process of chickens. It would also be interesting to find out whether the chickens containing the Polynesian D signature had actually undergone independent domestication processes in ISEA or the Pacific.

Aim 2: Reconstruct the Austronesian expansion using a proxy species

***Summary of findings and significance*:** The distribution of genetic lineages across the landscape and the relationship between them is very informative for understanding the mechanisms on how a species arrived in a particular geographic area. This is the main aim of phylogeography (Avise *et al*. 1987; Avise 2000). In studies involving translocated animals, it becomes advantageous to anchor the distribution found to relevant information coming from other sources such as linguistics, archaeology, and human genetics. This is particularly true for

domesticates in ISEA and the Pacific where considerable sea-crossings are necessary for these species to get from one island to another. As chickens lack any natural ability to cross the sea, the distribution of chicken genetic lineages is thus only possible via human movements. In the Pacific, the dispersal of animal domesticates (including chickens) from ISEA is likely linked to the expansion of the Austronesian-speaking people. However, it is not known how domestic chickens initially got into ISEA. It is possible that domestic chickens arrived via a similar mechanism to that seen in pigs (Larson *et al.* 2007) and therefore potentially by Austroasiatic people from MSEA (Lipson *et al.* 2014), as Austroasiatic people were likely present in ISEA prior to the spread of the Austronesian (Oppenheimer & Richards 2001; Kayser *et al.* 2006; Blench 2011; Lipson *et al.* 2014).

Chapters 2 and 3 illustrate the distribution of haplogroup D lineages in the ISEA-Pacific region and in the Indian Ocean rim, respectively, and discuss the implications for understanding the Austronesian expansion of agricultural components from ISEA. Several plants and animals are suggested to derive their ancestral origins in ISEA (Denham *et al.* 2003; Denham 2004; Fuller and Boivin 2009; Donohue & Denham 2010). Considering that the expansion of the Austronesian language is dominant and widespread it is logical to think that this was also the mechanism responsible for the translocation of domesticates. This is also true for haplogroup D chickens, as they are ubiquitous in areas corresponding to where the Austronesian-speaking people initially spread.

In Chapter 2 and 4, I show that it is likely chickens from the Philippines were translocated by Austronesian-speakers into Melanesia, and were eventually dispersed

along with the Lapita Culture. The Polynesian D signature illustrates that the ancestral origin of the Polynesian chickens is in the Philippines. However, unlike linguistics where the Austronesian language clearly originated in Taiwan, no haplogroup D chickens are found in Taiwan. Thus, the likely scenario is that chickens were integrated into the Austronesian expansion in the Philippines while *en route* to the Pacific. The spatial distribution and the increasingly derived state of Polynesian D chicken WMGs as they proceeds east to Remote Oceania suggests that they were associated with the first explorers who arrived in this region. These first explorers were identified as the Lapita culture based on inferences from linguistic and archaeological evidence. However, the phylogenetically based Philippine origin of the Polynesian chicken is demonstrated using purely modern Philippine chickens that retain the Polynesian motif, not from archaeological Philippine chickens, due to a lack of these remains. Although possibly biased by later admixture events, the large sampling effort of modern chickens in ISEA demonstrated here shows for the first time the level of diversity of haplogroup D in ISEA and its protracted history in the region.

The advent of agriculture in ISEA is often portrayed as being synonymous with the expansion of the Austronesians (Bellwood 1991; Diamond & Bellwood 2003; Bellwood *et al.* 2007). However, the solely 'South-China-through-Taiwan' source of agriculture in ISEA has been challenged recently with a theory about more indigenous and multiple origins within ISEA as demonstrated in Denham (2004) and Denham (2013). Also, it seems that the agricultural system of early farmers in ISEA did not include animal domesticates (Denham 2013). Animal domesticates such as pigs, dogs and chickens all seem to be imports from MSEA that arrived at, and were

integrated into ISEA agricultural practices at, different times (Denham 2013). Furthermore, it is worthwhile to explore the contribution of the Austroasiatic expansion on the advent of agriculture in ISEA and their contribution to the dispersal of domestic chickens into the region. For pigs two introduction pathways into ISEA are known: one via the Malay Peninsula and the other via Taiwan (Larson *et al.* 2007; Dobney *et al.* 2008). The dispersal of pigs via the Malay Peninsula might have been instigated by the Austroasiatic expansion, which could also be hypothesised to be the case for the initial arrival of domesticated chickens in ISEA from MSEA. Support for this theory is seen in the nuclear data from Philippine domestic chickens, which shows a high affinity with red jungle fowl from MSEA (Chapter 5). Genome-wide analysis of humans also indicates a significant Austroasiatic contribution to human populations in ISEA (Lipson *et al.* 2014). Therefore, although Austronesians are thought to be the cause of the chicken dispersal into the Pacific, the Austroasiatic speakers are more likely to have initially translocated domestic pigs, dogs, and chickens into ISEA.

Chapter 3 explores whether chickens were transported in association with the expansion of the Austronesians in the Indian Ocean. Madagascar, the westernmost range of the Austronesian also experienced agricultural inputs from South Asia and ISEA, such as the yam, taro, and banana (Blench & Dendo 2006; Fuller & Boivin 2009). Human genetic studies also confirm the Austronesian component in the genetic heritage of the Malagasy people (Razafindrazaka *et al.* 2009; Pierron *et al.* 2014). Thus, there is a clear east to west expansion in the Indian Ocean. However, Chapter 3 shows that Malagasy chickens are most closely associated with East African chickens that are derived from Indian origins. This is not altogether

surprising because India is the primary source of chickens that moved west into the Middle-East, Europe, and Africa (Crawford 1990; Hanotte *et al.* 2002; Mwacharo *et al.* 2011). Furthermore, this is also partially supported linguistically because the Malagasy word for chicken has an African origin (*i.e.*, Bantu) rather than Austronesian origins for the rest of their language (Blench & Dendo 2006).

Limitations and future directions: This thesis clearly adds to our current understanding of the relationship between the Austronesian expansion and the dispersal of chickens from ISEA through the phylogeographic analyses conducted in Chapter 2 and 3 using the mitochondrial control region and in Chapter 4 using whole mitochondrial genomes. The inclusion of additional whole mitochondrial genomic data from Austronesian chickens, including both modern and ancient samples (where available) would be further beneficial. This additional data would strengthen the inferences developed in this thesis. In instances where ancient samples are unavailable, genome-wide nuclear information is the next best direction.

Knowledge about the dispersal of chickens in the Indian Ocean will definitely benefit from more deliberate sampling regimes from South Asia, the Arabian Peninsula, East Africa and Madagascar. Assembling a whole mitochondrial genomic dataset, similar to the one done in the ISEA-Pacific (Chapter 4) should reveal a more resolved relationship between chicken populations around the Indian Ocean rim. This can improve inferences about the dispersal of chickens and their association with the Austronesian expansion in the Indian Ocean.

Aim 3: Assessing the resolution of human-mediated dispersal histories of chickens using control region, whole mitochondrial genome, and genome-wide SNP based histories

***Summary of findings and significance*:** Each standalone chapter in this study characterises, with increasing molecular resolution, the chicken populations in the Indo-Pacific region. This information is then used to investigate the population history of chickens and the dispersal of chickens in the Indo-Pacific-region. A comparison of the 201 bp control region fragment, the entire control region, and the whole mitochondrial genome (Chapter 4) illustrates that molecular characterisations made on mitochondrial evidence do not differ considerably, regardless of fragment size. Therefore, inferences derived from the geographic distribution of the different haplotypes are consistent across all levels of data analysis. However, utilising the whole mitochondrial genome dataset allows the relationship between the D haplotypes to be revealed more clearly. In particular, it shows that Polynesian D chickens in the Philippines are more basal in comparison to the ones in the Pacific. This supports the phylogenetic based theory of a Philippine homeland for the Polynesian chicken as hypothesised in Thomson *et al.* (2014). The basal position of Philippine Polynesian D chickens is not apparent when using the shorter control region dataset (Chapter 2). However, the control region dataset still allows for the same conclusion based on the geographic distribution of the Polynesian D haplotype in the Philippines and the Pacific, and its absence in Indonesia and MSEA.

Any inferences relating to domestication are only apparent when using whole mitochondrial genomes, as Polynesian D haplotypes appear to have experienced a

separate history in comparison to the remaining haplogroup D chicken haplotypes without the Polynesian signature. The domestication of chickens is known to have occurred multiple times in Mainland Asia (Fumihito *et al*. 1996; Liu *et al.* 2006; Kanginakudru *et al.* 2008; Miao *et al*. 2012), but given the expansion of the Polynesian D lineages from the Philippines it is possible that this lineage may be the result of another separate domestication event in the Philippines. Other scholars have also hinted at the possibility of a separate domestication episode of chickens in the Philippines (Mudar 1997; Blust 2002), so the results presented here identify a potentially unknown domestication process for chickens in an island environment.

It is only with the nuclear-wide genomic information in Chapter 5 that the distant relationship between the Philippine RJF and the indigenous domestic chicken is revealed. Philippine RJF may represent relic populations that predate the introduction of domestic chickens into ISEA. The close clustering of Philippine indigenous domestic chickens and Asian RJF also suggest an ultimate MSEA origin for chickens in ISEA today. Also, it is demonstrated using nuclear SNPs that mainland Asian chicken breeds are distant from chickens in ISEA and the Pacific. This has potential implications for genetic rescue of highly inbred commercial chicken breeds.

The chicken mitochondrial control region contains enough information to define the haplogroups and other major sub-haplogroup divisions (*i.e.,* Polynesian D; Thomson *et al.* 2014). This has an added advantage in terms of sampling and the number of published sequences available for comparison. However, as demonstrated in Chapter 4, finer-scale relationships and thus novel inferences are only apparent

using genome-wide information, such as the putative *in-situ* domestication of chickens in the Philippines. In the end, the decision on molecular scale will depend on the questions that a researcher sets out to answer.

Limitation and future directions: I recommend that all available aDNA material from the Pacific needs to be subjected to whole mitochondrial genomic analysis, particularly for CR haplotypes where whole mitochondrial genomes have not been generated previously. This can potentially inform us about fine-scale inter-island interactions not revealed using the short control region fragment.

The technology used to assemble the genome-wide nuclear SNP dataset in this study (Chapter 5) required large amounts of good quality DNA, therefore it currently cannot be used for ancient chicken material. Thus, a technology that can accommodate the analysis of nuclear loci, preferably genome-wide, from archaeological material is ideal. Capture-based enrichment of relevant nuclear genomic regions is one possible future direction. The use of genome-wide information is critical for investigating complex population histories, such as the chickens in ISEA that show signs of multiple translocation events from different types of molecular data (mtDNA vs. nuclear). Thus it would be ideal to genotype nuclear SNPs from more samples used in the control region analysis (Chapter 2 & 3).

In relation to other domesticates prehistorically transported across ISEA and the Pacific, it would be interesting to find out what genome-scale information (mitochondrial and nuclear) can further reveal about their histories. To date, no reports have been made using genomic information of commensals to understand

their translocation histories in ISEA and the Pacific other than this study. Most of the studies to date rely on uni-parental markers, such as mitochondrial control region fragments and the non-recombination region of the Y-chromosome (*e.g.* Larson *et al*. 2007; Oskarsson *et al.* 2011; Ding *et al*. 2012; Thomson *et al.* 2014a). Developing nuclear SNP genotypes might reveal other observations not apparent when only using fragments of the mitochondrial control region.

Concluding remarks

The results of this study demonstrated the utility of chicken genetics in improving our understanding and testing our current knowledge about the prehistoric movements of human groups both in the Pacific Ocean and Indian Ocean. It also genetically characterised the population history of indigenous chickens in ISEA, which in previous studies is limited only to a small number of samples. This examination revealed that the chicken population in ISEA is complex, and intimately linked to the population history of chickens in mainland Asian and crucial to the dispersal history of chickens in the Pacific Ocean and the Indian Ocean. The study also provides support to the Philippine homeland of the Polynesian chicken. This illustrates the complex nature of animal translocations in the region, which does not privilege a single source location, but one that occurred multiple times and via different routes, not only for chickens but for pigs and dogs as well.

The benefit of using the proxy approach is in its ability to provide an independent line of information about human expansions during prehistory. The approach of using domesticates, such as chickens, and other domesticates in examining past human movements through phylogeography is coming to an exciting

period due to the advances in molecular technology. This study shows how genome-wide information can allow us to infer original insights about the past which previously is not apparent using the short fragments of the mitochondrial control region. Thus, this study highlights the added value of genome-wide information in elucidating patterns and population histories. This is seen in the inference made on the possible domestication process for chickens in the Philippines.

Altogether, the study provided added novel insights about human-mediated translocation of chickens in the Indo-Pacific region. This is will hopefully add interest in the examination and lively discussion on prehistoric translocation and human movements, not only in the Indo-Pacific, but is other regions of the world as well.

References

Avise, JC. 2000. Phylogeography: the history and formation of species: Harvard university press.

Avise, JC, J Arnold, RM Ball, E Bermingham, T Lamb, JE Neigel, CA Reeb, NC Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annual review of ecology and systematics:489-522.

Bellwood, P. 1991. The Austronesian dispersal and the origin of languages. Scientific American 265:88-93.

Bellwood, P, C Gamble, SA Le Blanc, M Pluciennik, M Richards, JE Terrell. 2007. First Farmers: the Origins of Agricultural Societies, by Peter Bellwood. Malden (MA): Blackwell, 2005, Cambridge archaeological journal 17:87-109.

Blench, R. 2011. Was there an Austroasiatic presence in Island Southeast Asia prior to the Austronesian expansion? Bulletin of the Indo-Pacific Prehistory Association 30:133-144.

Blench, R, M Dendo. 2006. The Austronesians in Madagascar and on the East African coast: surveying the linguistic evidence for domestic and translocated animals. International Conference on Austronesian Languages Palawan: 17-20 January 2006.

Blust, R. 2002. The History of Fanual Terms in Austronesian Languages. Oceanic linguistics:89-139.

Crawford, RD. 1990. Poultry breeding and genetics: Elsevier.

Denham, T. 2004. The roots of agriculture and arboriculture in New Guinea: looking beyond Austronesian expansion, Neolithic packages and indigenous origins. World Archaeology 36:610-620.

Denham, T. 2013. Early farming in Island Southeast Asia: an alternative hypothesis. Antiquity 87:250-257.

Denham, TP, SG Haberle, C Lentfer, R Fullagar, J Field, M Therin, N Porch, B Winsborough. 2003. Origins of agriculture at Kuk Swamp in the highlands of New Guinea. Science 301:189-193.

Diamond, J, P Bellwood. 2003. Farmers and their languages: the first expansions. Science 300:597-603.

Ding, Z, M Oskarsson, A Ardalan, H Angleby, L-G Dahlgren, C Tepeli, E Kirkness, P Savolainen, Y Zhang. 2012. Origins of domestic dog in Southern East Asia is supported by analysis of Y-chromosome DNA. Heredity 108:507-514.

Dobney, K, T Cucchi, G Larson. 2008. The pigs of Island Southeast Asia and the Pacific: new evidence for taxonomic status and human-mediated dispersal. Asian Perspectives 47:59-74.

Donohue, M, T Denham. 2010. Farming and language in island southeast Asia. Current Anthropology 51:223-256.

Fuller, DQ, N Boivin. 2009. Crops, cattle and commensals across the Indian Ocean. Current and Potential Archaeobiological Evidence. Études Océan Indien:13-46.

Fumihito, A, T Miyake, M Takada, R Shingu, T Endo, T Gojobori, N Kondo, S Ohno. 1996. Monophyletic origin and unique dispersal patterns of domestic fowls. Proceedings of the National Academy of Sciences 93:6792-6795.

Hanotte, O, DG Bradley, JW Ochieng, Y Verjee, EW Hill, JEO Rege. 2002. African pastoralism: genetic imprints of origins and migrations. Science 296:336-339.

Kanginakudru, S, M Metta, R Jakati, J Nagaraju. 2008. Genetic evidence from Indian red jungle fowl corroborates multiple domestication of modern day chicken. BMC evolutionary biology 8:174.

Kayser, M, S Brauer, R Cordaux, A Casto, O Lao, LA Zhivotovsky, C Moyse-Faurie, RB Rutledge, W Schiefenhoevel, D Gil. 2006. Melanesian and Asian origins of Polynesians: mtDNA and Y chromosome gradients across the Pacific. Molecular biology and evolution 23:2234-2244.

Larson, G, T Cucchi, M Fujita, E Matisoo-Smith, J Robins, A Anderson, B Rolett, M Spriggs, G Dolman, T-H Kim. 2007. Phylogeny and ancient DNA of Sus provides insights into neolithic expansion in Island Southeast Asia and Oceania. Proceedings of the National Academy of Sciences 104:4834-4839.

Lipson, M, P-R Loh, N Patterson, P Moorjani, Y-C Ko, M Stoneking, B Berger, D Reich. 2014. Reconstructing Austronesian population history in Island Southeast Asia. bioRxiv.

Liu, Y-P, G-S Wu, Y-G Yao, Y-W Miao, G Luikart, M Baig, A Beja-Pereira, Z-L Ding, MG Palanichamy, Y-P Zhang. 2006. Multiple maternal origins of chickens: out of the Asian jungles. Molecular phylogenetics and evolution 38:12-19.

Miao, Y, M Peng, G Wu, Y Ouyang, Z Yang, N Yu, J Liang, G Pianchou, A Beja-Pereira, B Mitra. 2012. Chicken domestication: an updated perspective based on mitochondrial genomes. Heredity 110:277-282.

Mudar, KM. 1997. Patterns of animal utilization in the Holocene of the Philippines: A comparison of faunal samples from four archaeological sites.

Mwacharo, J, G Bjørnstad, V Mobegi, K Nomura, H Hanada, T Amano, H Jianlin, O Hanotte. 2011. Mitochondrial DNA reveals multiple introductions of domestic chicken in East Africa. Molecular phylogenetics and evolution 58:374-382.

Oppenheimer, SJ, M Richards. 2001. Polynesian origins: slow boat to Melanesia? Nature 410:166-167.

Oskarsson, MC, CF Klütsch, U Boonyaprakob, A Wilton, Y Tanabe, P Savolainen. 2011. Mitochondrial DNA data indicate an introduction through Mainland Southeast Asia for Australian dingoes and Polynesian domestic dogs. Proceedings of the Royal Society B: Biological Sciences:rspb20111395.

Pierron, D, H Razafindrazaka, L Pagani, F-X Ricaut, T Antao, M Capredon, C Sambo, C Radimilahy, J-A Rakotoarisoa, RM Blench. 2014. Genome-wide evidence of Austronesian–Bantu admixture and cultural reversion in a hunter-gatherer group of Madagascar. Proceedings of the National Academy of Sciences 111:936-941.

Razafindrazaka, H, F-X Ricaut, MP Cox, M Mormina, J-M Dugoujon, LP Randriamarolaza, E Guitard, L Tonasso, B Ludes, E Crubézy. 2009. Complete mitochondrial DNA sequences provide new insights into the Polynesian motif and the peopling of Madagascar. European Journal of Human Genetics 18:575-581.

Storey, AA, JS Athens, D Bryant, M Carson, K Emery, C Higham, L Huynen, M Intoh, S Jones, PV Kirch. 2012. Investigating the global dispersal of chickens in prehistory using ancient mitochondrial DNA signatures. PloS one 7:e39171.

Thomson, V, KP Aplin, A Cooper, S Hisheh, H Suzuki, I Maryanto, G Yap, SC Donnellan. 2014a. Molecular Genetic Evidence for the Place of Origin of the Pacific Rat, Rattus exulans. PloS one 9:e91356.

Thomson, VA, O Lebrasseur, JJ Austin, TL Hunt, DA Burney, T Denham, NJ Rawlence, JR Wood, J Gongora, LG Flink. 2014b. Using ancient DNA to study the origins and dispersal of ancestral Polynesian chickens across the Pacific. Proceedings of the National Academy of Sciences 111:4826-4831.