

7-20-2021

Reconstruction of the Austronesian Diaspora in the Era of Genomics

Geoffrey K. Chambers
Victoria University of Wellington

Hisham A. Edinur
Universiti Sains Malaysia

Follow this and additional works at: https://digitalcommons.wayne.edu/humbiol_preprints

Recommended Citation

Chambers, Geoffrey K. and Edinur, Hisham A., "Reconstruction of the Austronesian Diaspora in the Era of Genomics" (2021). *Human Biology Open Access Pre-Prints*. 182.
https://digitalcommons.wayne.edu/humbiol_preprints/182

This Article is brought to you for free and open access by the WSU Press at DigitalCommons@WayneState. It has been accepted for inclusion in Human Biology Open Access Pre-Prints by an authorized administrator of DigitalCommons@WayneState.

Reconstruction of the Austronesian Diaspora in the Era of Genomics

Geoffrey K. Chambers^{1*} and Hisham A. Edinur²

¹School of Biological Sciences, Victoria University of Wellington, Wellington, New Zealand.

²School of Health Sciences, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia.

*Correspondence to: Geoffrey K. Chambers, Victoria University of Wellington, School of Biological Sciences, PO Box 600, Wellington 6140 NZ. E-mail: geoff.chambers@vuw.ac.nz.

Short Title: Genomics of the Austronesian Diaspora

KEY WORDS: AUSTRONESIAN, AUSTRALO-MELANESIAN, AUSTROASIATIC, POLUNESIAN, MELANESIAN, SINGLE NUCLEOTIDE POLYMORPHISMS (SNPS), GENOME-WIDE SNPS, DNA SEQUENCING.

Abstract

The Austronesian Diaspora is a 5000-year account of how a small group of Taiwanese farmers expanded to occupy territories reaching halfway round the world. Reconstructing their detailed history has spawned many academic contests across many disciplines. An outline orthodox version has eventually emerged, but still leaves many unanswered questions. The remarkable power of whole-genome technology has now been applied to people across the entire region. This review gives an account of this era of genetic investigation and discusses its many achievements including revelation in detail of many unexpected patterns of population movement and the significance of this information for medical genetics.

The Austronesian Diaspora is a big story; an account of how a relatively small collection of migrant farmers spread half-way round the world to become a diverse set of populations numbering 380 million people living in 33 separate nation states (Figure 1).

It is even more remarkable that all this has taken place within the last 5,000 years. There are many excellent reviews describing their history including: Bellwood (1991, 2005, 2017); Donohue and Denham (2011); Duggan and Stoneking (2014); Friedlaender (2007); Gosling and Matisoo-Smith (2018); Hung (2019); Kayser (2010); Matisoo-Smith (2015); Pugach and Stoneking (2015); Soares et al. (2011). Most of these are primarily concerned with Island Southeast Asia or Oceania and each set of authors has tended to concentrate on one set of ideas, one particular geographical area or class of evidence. Very few have attempted to build a comprehensive and holistic genes, language and culture picture. Bellwood et al. (2011) may be exception, but even these authors are focused on evidence for and against the 'Out of Taiwan Model' (more on this idea later).

In this review, we contribute to further developing this account and concentrate exclusively on new genetic evidence, but we have tried to take an ultra-wide-angle view encompassing the entire geographical area over which Austronesian people are now distributed (Figure 1). In particular, we ask how, starting from Lipson et al. (2014), Wollstein et al. (2010) and Xu et al. (2012), a virtual avalanche of new data from genomic studies has informed an updated view of Austronesian history. The three early genomic surveys cited above pointed to an origin in Taiwan, showed relatively uniform Austronesian/Australo-Melanesian admixture across Polynesia and introduced Austroasiatic people from mainland southeast Asia as contributors to the genepool of Austronesian people in Indonesia and Peninsula Malaysia. Genome wide SNP surveys and even whole genome work are now quite extensive and routinely conducted at extremely large sampling scales; see for example, Mallick et al. (2016). This approach has been facilitated by developing

effective culturally sensitive ways for scientists to work with indigenous groups: see Resnik (1999) on the wide range problems encountered by the earlier Human Genome Diversity Project. The new phase of work has even included examination of ancient human remains from burial sites. Such ancient DNA studies have been substantially advanced by discoveries that human petrous bones are an excellent source of well-preserved genetic information (Pinhasi et al. 2015), together with improved amplification techniques for trace DNA molecules and next generation DNA sequencing methods. The massively enhanced resolving power inherent in this type of data has significantly impacted on the Austronesian story in three main areas; 1) better resolution of deep ancestry fractions (e.g., around Papua New Guinea), 2) revised settlement narrative (e.g., for Vanuatu) and 3) disclosed otherwise cryptic details of local population movements (e.g., in Samoa and Madagascar).

Before one can begin to attempt a reconstruction that incorporates this new genomics information it is useful to define the participant populations and the ideas that shaped orthodox understanding at the dawn of the genomics era as shown in Figure 1. These follow in order in the two sections below.

A Large Cast of Characters

Although the Austronesian diaspora chiefly concerns the Austronesian people themselves (AN) as typified by the Taiwanese aboriginal Hill Tribes. The account also involves two other major contributors, Australo-Melanesians (AM) and Austroasiatic-speaking people (AA) with extensive admixture interactions over large areas and across large time spans. Further, the story would not be complete without proper consideration of the most ancient people who have interacted with them in the distant past, nor without including human population groups who interacted with them at the periphery of the AN distribution during more recent times.

Homo erectus and Other Ancient Hominins

The first humans to settle Island southeast Asia (specifically in Java) were *Homo erectus*, between 0.75 - 0.4 mya. Their remains are abundant in Java and elsewhere, but do not extend to Wallacea and Sahul. They arrived early and may have persisted in some form or other until more recently, e.g., as represented by the insular dwarf *H. floresiensis* which is arguably a descendant of *H. erectus*: see Tucci et al. (2018) for details and references. A second example may be the less well-known *H. luzonensis* recently described from the Philippines (Détroit et al. 2020). However, these ancient people do seem to have crossed over the ocean passage to the Philippines via what many think was always a fairly wide and deep water channel by around 67 kya. They are also reported to have butchered a rhino there. So, this find prompts the question: ‘How did the rhino make the same sea crossing, or do we really have the right model(s) of past sea level changes for this area?’

Some have suggested that one or more of these ancient hominins or their ancestors may have interbred with other later forms described below including the well-known *H. neanderthalensis* and *H. denisova* (Rogers et al. 2020). The former admixed with early ‘Out of Africa’ *H. sapiens*. Today most anatomically modern humans have around 2% Neanderthal DNA in their individual genomes, including AM Andaman Islanders (Onge); see Jinam et al. (2017). The Denisovans are later entries to the *H. sapiens* expansion story and clearly crossed into Sunda and Sahul. These lineages of *H. denisova* are different from the Altai form found in northern mainland Asia. The southern lineages interacted with AM leaving at least 2-4% introgressed genomic signal (possibly higher in Papua New Guinea). Some authors have suggested that there are two such sequential events, tagged as D1 and D2, with the latter taking place within the last 40 to 50 k years or even later (Jacobs et al., 2019). Others disagree (Bergström et al., 2020). Nonetheless, it is a great advantage to have a selection of genomic sequences from these ancient peoples ready to hand as they are an aid to

sorting out relationships between contemporary human lineages (Skoglund and Mathieson, 2018).

Australo-Melanesians

This term principally describes the inhabitants of Australia and Papua New Guinea, but also applies to other populations with shared ancestry. Older terms for them have been used, including Papuan-speaking Australoids and Melanesians, but are now largely discarded.

These authors prefer the name used here because it seems to have entered popular usage and Melanesian (despite its ambiguity) is reserved principally as the descriptor that the people of Vanuatu and New Caledonia and others use to describe themselves. This group is derived from the earliest 'Round India Coastal Route' migration picking up part Neanderthal ancestry along the way. They are the source of the Onge and the Negrito people of Indonesia and Malaysia (where they form part of the wider Orang Asli group). From here, Bellwood (2018) argues that the data of Lipson et al. (2018a) and McColl et al. (2018) show them to be related to the 'Hòahabinins' of mainland southeast Asia and even the Jōmon of ancient Japan. This is taken by Bellwood to be local support for the 'Two Layer Model' of human settlement in Asia although the original authors would not agree with this suggestion. This proposal will clearly require more extensive investigation. The Indigenous Australians are part of this lineage and their settlement history has now been elaborated in some detail (Malaspinas et al. 2016). Much the same can be said for Papua New Guinea (Bergström et al. 2017) with the exception of the north coast settlements and offshore islands which have extensive AN admixture, as has long been known. Negrito people are also found in Borneo (Yew et al. 2018a,b) and the Philippines (Arenas et al. 2020) where they are particularly well known. There is no concrete evidence that they made it all the way to Taiwan, but it is more than

likely that they did. Today, they continue to exist there only in myth and legend among the Hill Tribes as the ‘little people of the forest’ (Chen et al. 2019).

Austroasiatic-Speaking People

This group emerged within mainland southeast Asia as descendants of the earlier wave(s) of settlement – see discussion above of work by Lipson et al. (2018) and McColl et al. (2018).

Within the last 10 k years they began to move south through Thailand into Indonesia and Peninsula Malaysia, as first elaborated by Lipson et al. (2014).

Austronesians

These people most likely originated in mainland China as an element of the Daic people (Li et al. 2008). Arguably, almost all traces of their former presence have been obscured by expansion of the now dominant Han. A novel new genomics study sequencing whole Y chromosomes provides hypothetical links to northern China (Sun et al. 2020). They moved to Taiwan around 8-10 kya and have since split into numerous sub-tribes, who each speak one of the nine major branches of Austronesian language. The enigmatic tenth branch is Malayo-Polynesian as is spoken elsewhere in one form or another across the whole area of the Austronesian Diaspora. Around 5000 ka Taiwanese AN farmers began to move south via the Batanes Islands and on to the Philippines as supported by a strong combination of genetic and archaeological evidence (Tabbada et al. 2010; Bellwood et al. 2011). Even today, no one single tribal group has been clearly identified as the source of these migrants, but the eastern coastal Amis and Atayal seem to be the leading candidates.

Sources of Recent Admixture

As Austronesians and later their admixed descendants, the Polynesians, settled new lands across the globe they began to come into contact with many new populations. In c20th this process has increased dramatically facilitated by an increasing ease of worldwide travel. Well known examples include contact with the Bantu speaking tribes from Africa in Madagascar and Arabs, Chinese, Europeans and Indians in Malaysia.

It is important that investigators should be familiar with all of these populations and all of their possible interactions, so that they may create well designed studies with appropriate reference comparisons for data analysis. Equally, it is vital to have a strong model of pattern and process elements in the history of the Diaspora. The development of such ideas is described in the next section.

Paradigms Past and Present

Received wisdom relating to the Austronesian Diaspora (see Figure 1 for a sketch with dates) has gone through many changes over the years, has involved many extended debates between ideas and extends all the way back to the Multiregional Continuity Hypothesis vs. the Two Layer Model to explain the replacement of *H. erectus* by *H. sapiens*. Later intellectual contests stem from Forster's observations made during James Cook's voyages of Pacific exploration. Polynesian people all across the various islands spoke their own dialect form of a single language and so must be closely related to one another (see Hung 2019). Captain Cook was rightly lauded for his amazing set of travels. But one might also reflect that he was meeting Pacific people who had covered the same oceans and well before his ships arrived. The Polynesians had done this entirely without the assistance of Western science and metal-based technology, all without the benefit of a written language. They had a wholly organic system of marine architecture supported by indigenous knowledge and navigation systems carried on down through generations by oral tradition, woven maps and dances. Further,

Cook's voyages were made possible in part with help from Tahitian navigators such as Tupaia.

Hence it was natural to ask: 'Where had these people come from and how did they get to their present islands?' This set up a homeland debate between Asia vs. South America. The latter was championed by Thor Heyerdahl (1952) and has since fallen almost entirely out of favour, despite the fact that a number of vegetable resources used by Polynesians do clearly originate from South America (see Matisoo-Smith 2015). These have generally been held to be the fruits of return trading trips by Polynesians. The role of South Americans has recently gained greater prominence following the report by Ioannidis et al. (2020). They discovered introgression of small amounts DNA in eastern Polynesia traced to Columbian Indians (possibly the Zenu or their near neighbours). Contact with Columbia makes good sense as sailing to and from much of the rest of the South American coast is very difficult due to the strong northwards drive of the Humboldt Current. It remains unclear just who was sailing the boats that made contact. Ioannidis et al. (2020) make a claim that this all happened around 640 ya. They present two lines of evidence. First, Columbian genetic markers are fairly widely spread across this section of Polynesia, which takes time. Second, these DNA segments have been reduced to small average size by recombination, much smaller than the segments introduced by the more recent episodes of European contact and admixture. Finally, they argue that it is possible more than one interaction took place on Rapa Nui (Easter Island) maybe as early as 820 ya, just prior to the establishment of their monumental stone structures around 720 - 670 ya. Their study involves 807 individuals from 17 Pacific Islands and 15 South American locations in a survey of genome-wide SNPs.

The Asia Homeland concept is not without its own problems. Regarding appearances, Austronesian residents and visiting Polynesians are rarely confused with mainstream Chinese living in Taiwan, but Polynesians are often taken for Atayal or Amis etc. Also, Melanesia lies

between Remote Oceania and Mainland Asia. Melanesians (i.e. AM people living in Melanesia) are equally well differentiated physically. They are typically described as having dark skin, short fuzzy hair and not all of them speak Austronesian languages. Nonetheless, a new debate was spawned ideas competing Island Southeast Asia against Taiwan (the home of Austronesian languages according to many scholars – see Gray et al. (2009). The Island Southeast Asia Homeland(s) concept arises from the ‘Nusanto Maritime Trade and Communication Network’ idea (see account in Hung, 2019) and was a key element in Oppenheimer’s ‘Eden in the East’ concept which also builds on historical changes in sea level creating land bridges across much of Sundaland (Oppenheimer, 1998). The now dominant ‘Out of Taiwan’ idea was a geographic sub-element of Bellwood’s ‘Farming Language Dispersal Hypothesis’ (Bellwood, 2005). Passage of Austronesian speaking people from Taiwan south through the Batanes Islands and on to the Philippines was well supported by a characteristic pottery horizon, but has lacked strong evidence of widespread rice cultivation and consumption.

The next stage involved contact with areas in northern Papua New Guinea which already had substantial AM populations. This prompted questions about the amount of admixture between residents and incoming AN migrants. This in turn prompted a whole series of metaphoric models featuring ‘boats’ and ‘trains’ travelling in either slow or express mode and with greater or lesser amounts of contact and admixture.

In the following section the authors will argue that both classes of model can claim to capture some aspect of the story and that a pre-existing version of the Lapita Cultural Complex was carried along both. This view is supported by new genomics evidence from both the Bismarck Archipelago and Vanuatu. We will also argue that these exciting new findings alone do not entirely solve the question of the origin of the Polynesians.

An Extended Model as Informed by Genomics

At the outset of the present decade the orthodox received wisdom account (see Figure 1 and Chambers and Edinur 2015 for one recent description) was pretty much as outlined above and as illustrated in Lipson et al. (2014). From around 50 kya AM people began, or continued, their easterly and northerly movements and established large linguistically diverse settlements in Australia and Papua New Guinea (PNG) with smaller Negrito populations in Indonesia, Malaysia and mainland Asia plus another branch in Borneo, the Philippines and possibly Taiwan. Much later AN people moved south through the Philippines and Borneo with one branch going on westwards to Indonesia (with admixture with AA at the western end) and ultimately reaching the Comoros Islands and Madagascar. A second branch sailed east to northern Papua New Guinea and the Bismarck's where they established a partly pre-existing cultural complex known as Lapita (LCC). This spread through the Solomons, New Caledonia, Vanuatu, Tonga and Samoa carried by admixed AN:AM progenitors of the Polynesians (with a characteristic 70:30 genetic mix). Voyaging paused at this point for around two thousand years before setting off once again across the far reaches of Remote Oceania to settle an important cultural centre in the Leeward Society Islands. Starting from around 1000 ya. extensive exploration and settlement voyages filled up what is known as the Polynesian Triangle with corners at Hawaii, Rapa Nui (Easter Island), Aotearoa (known today as New Zealand).

Since then a whole generation of genomics studies collecting either complete whole genomes or large-scale sets of single nucleotide polymorphisms (SNP) has served both to support the general account above and add considerable detail and thereby elaborate many of the episodes. A description of this extensive body of work is given below following the composite AM, AA and AN migration pathways as laid out above.

The early part of this diaspora across Island Southeast Asia is shown in detail in Figure 2 and the major branches are numerically labelled in chronological order to better clarify the structure of the text.

BRANCH 1: Mainland Southeast Asia

There are now many ancestry studies across this region, all with greater or lesser involvement of AN people. Two of the largest are Mallick et al. (2016) reporting the Simons Human Genome Diversity Project including 300 full genomes from 142 populations and Bergström et al. (2020) with 929 genomes from 54 populations. To these one may add Lipson et al. (2018a) and McColl et al. (2018) plus two others, Liu et al. (2020) and Pischedda et al. (2017), both with a special focus on Vietnam. Overall, these publications confirm traces of AM ancestry in southeast Asia, but find no evidence of the continued presence of the direct Daic ancestors of the AN people. There are many AN-speaking groups in modern Vietnam and elsewhere (including the Cham, Chru, Raglai – see Macholdt et al. 2019 for a full list), but very little evidence of their having AN-type genes. This is clearly an example of one or more cultural shifts. Interestingly, there is evidence of both Neanderthal and Denisovan ancestry in Asian populations. Northern groups have Altai Denisovan admixture, whereas southern peoples have received Denisovan DNA which is more characteristic of Oceanic populations and done so possibly on more than one occasion. Interestingly, the Denisovan genomic fractions reported in these studies are around 2 - 4% compared with higher earlier estimates of 4 - 6%.

Taiwan

It is perhaps surprising that no large genomic studies have been carried out with the Aboriginal tribes of Taiwan. They are represented in small numbers in some of the surveys

mentioned above, but there is no shortage of previous conventional style genetic studies: e.g., Ko et al. (2014); Jinam et al. (2012); Mirabel et al. (2013) and Tabbada et al. (2010), sufficient to convince many other scholars that this is the true distal source of the oceanic Austronesian Diaspora, see Bellwood et al. (2011). Nonetheless, it still remains theoretically possible that the first migrants passed directly from Mainland Asia to the Philippines. The enigmatic linguistics data might be thought by some to confirm this scenario. Overall, all one can say for certain that the centrality of Taiwan in the Austronesian expansion is not yet clearly supported by modern genomic evidence.

The Philippines

These authors are not aware that any genomic level study has yet been carried out specifically on AN people resident in the Philippines. However, there is one new and quite extensive mitochondrial DNA study by Arenas et al. (2020). Surprisingly, their work does demonstrate all the conventional affinities for this group as recognised by Pugach et al. (2018), but still concludes that the Philippines were settled from both the north and the south. Further, consideration of this dilemma is plainly merited, but the southern route may reflect a later trading network with AN settlers in Borneo.

It is from this point onwards that we have chosen to divide the AN Diaspora into three streams (Figure 2). The first (BRANCH 2) leads to Micronesia; the second (BRANCH 3) northern Papua New Guinea and out into Near Oceania and then further east to Remote Oceania. The third stream (BRANCH 4) leads to Borneo, Malaysia and Indonesia etc. It is always possible that this split occurred in Borneo and we know of no specific evidence to say clearly that it did not. However, a new study from the Leeward Society Islands (Hudjashov et al., 2018) does point to the Philippines. These interesting developments are examined again in more detail in subsequent sections of this review.

BRANCH 2: Micronesia

There is no doubt that many of the contemporary residents of these widely scattered oceanic islands have AN or part AN ancestry. Further, there have been plenty of opportunities for AN immigration from the east, south and west as shown in Bellwood et al. (2011) and see Figure 1. Equally there were plenty of opportunities for genetic exchange via maritime trading networks. Thus, the human history of Micronesia is complex and to date there is neither a full account of these patterns and processes, nor are there many genetic data that we are aware of from which such accounts might be constructed or tested. Hence, it would generally seem premature to make any attempt include this region in the wider account of the Austronesian Diaspora as presented in this article.

There is one caveat. This is in the form of one very informative new study of ancient DNA from Guam by Pugach et al. (2020). Their genome-wide SNP data track early settlers from around 3500 ya to a home source in the Philippines. These lineages have a close link to early Lapita people from Vanuatu and Tonga (see later) which raises some interesting prospects for ocean-going contact between them but noting that there is no evidence of Lapita pottery horizons anywhere in Micronesia but red slipped pottery are found there and dated to around 3500Y YA does have affinities with forms found in the Philippines and elsewhere in Island Southeast Asia and which may be antecedent forms of Lapita ware (Hung et al., 2011). Given the generally adverse pattern of prevailing winds the settlement of the Marianas must represent a deliberately planned initiative and as such represents the longest ever open ocean voyage made by humans at that time.

BRANCH 3: Papua New Guinea and Its Neighbouring Islands

It has long been known that villages along the north coast of Papua New Guinea (PNG) are home to admixed AN/AM people who speak Austronesian languages (see Figure 3A). They are physically different from typical PNG Highlanders: see Bergström et al. (2017) for more about the internal structure of the AM population. The genetic admixture of these coastal people is now known to be ~ 70AN:30AM (Lipson et al. 2014, Wollstein et al. 2010) with considerable gender bias. Their mtDNA lineages are predominantly AN and their Y chromosomes have an over representation of AM haplotypes arising from matrilineal marriage practices (Hage and Marck, 2003). One sees a similar pattern across the New Britain and New Ireland archipelagos with northern coastal populations having more AN genotypes and the interior and southern settlements more AM genotypes (Friedlaender, 2007). The Reefs Santa Cruz Group is a marked exception. It is an early (around 3000 ya) site of Austronesian Lapita leapfrog expansion into Remote Oceania (see Sheppard, 2019), but is now populated by entirely AM people (more on this later). The adjacent island group, the Solomons, seems to have been settled in a slightly more complex fashion with admixed AN/AM people at the eastern end of the group and a more AM lineage at the western end (see Figure 3B). A recent genome-wide SNP study (Isshiki et al. 2020) show that their admixed genomes have been shaped by natural selection.

The people of this region have one of the highest known representations of Denisovan ancestry in their genomes (see Vernot et al. 2016 for the earliest large-scale genome sequencing data supporting this view). This genetic material must have entered via admixture during the past 30 – 40 k years, i.e. because their ancestors such as the Andaman Islanders have Neanderthal admixture, but not Denisovan. Some claim that this happened in two temporally and geographically separated episodes labelled D1 (around 30 kya) and D2 (around 46 kya), based on complete genome sequences (Jacobs et al., 2019)

Finally, it has long been held that the Lapita Cultural Complex arose as a product of the AN/AM interaction somewhere along the north coast of PNG and or nearby islands in Melanesia. This idea was supported by the fact that locations at the north end of New Britain and New Ireland are the earliest known Lapita sites (Summerhayes et al. 2010). Some commentators claimed that Lapita artefacts were without precedent in archaeological sites along the AN voyaging passage. Others disagreed (Bellwood et al. 2011). The highly characteristic Lapita style pottery has also been found in New Caledonia, Vanuatu, Fiji, Tonga and Samoa. These observations raised some immediate questions, because today the first three island groups are occupied, more or less exclusively, by Melanesians with AM heritage (see next section). The second two are Polynesian as might be expected under the previous set of assumptions.

Vanuatu

The next obvious destination on the westward progress of the main Austronesian Diaspora would seem to be the islands that make up Vanuatu. In a sense this is true, but matters begin to get complicated around this location. Certainly, there are Lapita sites in Vanuatu and some of these are associated with human burials. A series of aDNA studies conducted there by a combined team of archaeologists and geneticists produced some very surprising results (Lipson et al. 2018b; Posth et al. 2018; Skoglund et al. 2016). These findings are well described by the two leaders of the team (Spriggs and Reich 2020). In short, Lapita culture is first recorded from 3350 - 3100 ya in the Bismarck Archipelago. It did not immediately reach the Solomons, as far as anyone knows, but seems to have leap-frogged on to Vanuatu from about 3000 ya and then shortly afterwards to Tonga (2850 ya) and Samoa (2750 ya). The bones in the Teouma cemetery near Port Vila were interred in Lapita pots and discovered to have 100% AN ancestry. This stands in marked contrast with the contemporary population

groups which is around 90% AM, but who nonetheless speak a range of modified AN languages. The early Lapita settlements only persisted for a short time and were replaced by the new wave of AM settlers by 2300 ya. The entire genetic history of the three phases of human settlement in Vanuatu is now fully laid out by Lipson et al. (2020).

There are some new fundamental lessons about the AN Diaspora to be learned here. First, the original settlers must have come more or less direct from either Borneo or the Philippines bringing an intact form of the LCC with them or alternately (and perhaps more likely) from the Bismarck Islands before admixture with resident Papuans took place. It is less likely, that Lapita culture arose around the time of AN/AM admixture beginning in PNG, but was carried to the region and developed as a whole by the earliest AN settlers. Either way this, would seem to be a particularly striking form of the old ‘Express Train’ idea (Diamond 1988). Second, the source of the later AM population Vanuatu is unknown even based on the genetic make-up of the present inhabitants (as mentioned earlier), but see Lipson et al. (2020) who favour Reefs-Santa Cruz. Third, for all the elegance of the work described above, Vanuatu may yet turn out to be tangential to the Austronesian Diaspora, because this version of settlement history does not account for the origin of the Polynesians. Their origin as an admixed AN:AM people must now either be seen as northern coastal PNG (as proto-Polynesians) or Samoa (Polynesians). New genetic evidence shows a direct link between these two locations by passing Vanuatu. The AN people living both east and west of Vanuatu have fairly uniform 70:30 admixture with AM (see Kimura et al., 2008 as one early study establishing this idea). Thus, the most parsimonious account sees them leap-frogging Vanuatu as they headed west. This idea was captured (but not promoted) by Skoglund et al. (2016) and shows a minor side branch of admixed people going to Vanuatu. This may well be the case, as there are, for example, contemporary Polynesian settlements on Efate at Mele village and on Ifira Island. Any alternative strictly ‘via Vanuatu’ account would see a steady

diffusion of AM people into pure AN Lapita culture settlements in Tonga and Samoa. True, there is time enough for this to have happened, but it would not be expected to produce such genetic uniformity. Besides, Fiji is now almost wholly AM plus a few Polynesian villages and one needs to explain why this did not happen in Tonga and Samoa.

There are also two important things to learn here about how genetics/genomics are applied to study the history of human populations. In part, this project has been so informative because the genetics team in Vanuatu worked closely together with the archaeologists. Many past (and present) genetics studies use models and data processing algorithms to date past events and/or calculate population sizes. It must always be remembered that such findings are only hypotheses and cannot stand against hard dating evidence associated with cooking fires and pottery horizons etc. (acknowledging that these too are always considered provisional). Another cautionary note is sounded by this work. That is that it is always risky to extrapolate genetic evidence gathered from modern populations as if it applied throughout time. No findings that we are aware of make this point more strongly than the replacement of the Lapita people in Vanuatu by Papuans.

Fiji, Tonga, and Samoa

The settlement histories of these three island groups are interesting for their similarities and differences as discussed above. There are no genomic data for the AM population of Fiji, but it is worthwhile noting again that, like Vanuatu, there are long existing Polynesian villages in Fiji. Both Tonga and Samoa are former Lapita settlements with contemporary populations of Polynesians. Ancient DNA from burial sites on Tonga confirm that they were first settled by AN people and later replaced by Polynesians (Harris et al. 2020; Posth et al. 2018). New genomics data from Harris et al. (2020) date first settlement of Samoa as around 2750 - 2800 ya (shortly after Vanuatu) followed by a large influx of Polynesians between 1500 -1000 ya.

Their study subjects range from 23 to 35% AM and, as might be expected, it is in this AM fraction of their genomes that the Denisovan DNA inserts are located. These data also give a picture of population growth and movement featuring population recent migrations between Savai'i and Upolu and within Upolu.

Leeward Society Islands

During the last 1000 years this island group has played a significant role together with the Marquesas and others as a home base for the settlement of the Polynesian Triangle (see following section) and leading to contact with South America as explained earlier – see Figure 4.

A new genomics study by Hudjasov et al. (2018) favours a model of direct Polynesian settlement from the Solomons via Samoa and Tonga leap-frogging Vanuatu, very much in line with some of the themes developed earlier in this article. They date admixture with AM in two pulses one around 1,200–1,700 ya, and another some time earlier. This settlement scenario fits well with the archaeological history of Samoa and Tonga (see earlier). It is unlikely that this admixture episode occurred in the Philippines themselves as their descendant AN tribes, now living in Borneo, have remained unadmixed (see later), like the Philippine Kankanaey in the Hadjashov et al. (2018) study. Either way, the Leeward Society Islands still developed to become the centre of a reticulated Polynesian trading and cultural exchange network; see Wilmshurst et al. (2011) for some provisional dates and Crowe (2018) for more details.

The Polynesian Triangle

This vast area of the Pacific is bounded by Hawaii, Rapa Nui and Aotearoa. There are many prior genetic studies but few specific genomics reports for these locations except for those in

Ioannidis et al. (2020) as described earlier. Their analysis includes information from many sites in Eastern Polynesia. This whole region was rapidly settled from centres in the Leeward Society and Marquesas Islands. The source quoted above contains a lot of valuable analyses of these movements and these are best read in detail in the original source (Ioannidis et al. 2020). A full description of this era is too extensive for inclusion here, but we will explain how Polynesians reached the remote southernmost vertices of the Polynesian Triangle. In brief, Rapa Nui is now populated by Polynesians with extensive admixture with South American Indians and Europeans. The Polynesian people of Aotearoa are known as the Māori and arrived around 700 ya - see Figure 5.

Remarkably, they set out from there in every conceivable direction to explore the Kermadecs and attempt settlement of the Chatham Islands, Auckland Islands and Norfolk Island (see Crowe 2018). It is only on the first of these that a subtribe called the Moriori persisted through to the time of European discovery. Polynesian settlements on both Norfolk Island and Raoul Island (Kermadecs) did last long enough for garden crop plantations to be established. Well after Norfolk Island was abandoned by Polynesians it was taken over by the Australian colonial administration and converted into a market garden to supply the new immigrants in their Botany Bay enclave with fresh produce. It also served as a very grand Georgian style penal colony. When this institution was finally abandoned, the island was turned over to the booming population of Pitcairn Island. These islanders are the descendants of the nine HMS Bounty mutineers and their six Tahitian partners who accompanied them to their remote sanctuary. The contemporary population of Norfolk Island are the subjects of several large-scale medical genomics studies (e.g., Benton et al. 2019). The significance of this and other similar health-related work will be discussed later.

BRANCH 4: Borneo

Three smaller-scale genomics studies of AM Negrito and AN tribes on Borneo (Deng et al. 2019; Yew et al. 2018a, b) show the expected deep differences between them and confirm Denisovan ancestry in the Negrito peoples. Perhaps surprisingly, the various AN people show little or no AM admixture. This is significant because it suggests that the Austronesian-speaking people who subsequently went on to Malaysia, Indonesia and elsewhere likely arrived with close to 100% AN ancestry – see Figure 6.

Malaysia

It is in northern Malaysia that the three-way AN, AM, AA interactions begin to become a central focus of the present account. Before then, AN people arrived in southern Malaysia (probably from Borneo) and became what are today known as the Proto-Malays and who now form part of the minority Orang Asli (OA) tribes. Nonetheless, they did give rise to today's majority Deutero-Malays, who are not part of the OA. This history and various other complexities including language shifts made it difficult for many previous authors to sort out the intricate Malay genetic 'layer- cake' (Norhalifah et al. 2016). The OA consist of three groups; besides the Proto-Malays there are the Semang (Negritos) and the Senoi who are AA partly admixed with Semang. The Deutero-Malays have a wide range of historical admixture, first with Indian and Arab Malacca Straits traders, later with European colonials and finally with immigrant workers from southern China and India. These last two now form substantial minority groups of their own in Malaysia. In addition, various predominantly AN Malay Sub-ethnic Groups have migrated to Malaysia from surrounding territories.

A surprisingly large number of genomic studies have already been carried out either in Malaysia (Deng et al. 2014; Deng et al. 2015; Deng et al. 2019; Fu et al. 2018; Juhan et al. 2014; Salleh, et al. 2013; Sun et al. 2020; Wong et al. 2013) or including Malay subjects (Wu et al. 2019). Taken together, they provide piecemeal support for the description above. For

instance, Wu et al. (2019) surveyed no less than 4810 Singaporean subjects by whole-genome sequencing at a depth of 13x and were able to show clear separation between Chinese, Malays and Indians.

Indonesia

The three-way interaction between AA, AM (Negrito) and AN across the Indonesian archipelago has been the subject of many conventional genetic studies. In summary, these built up a picture of admixture between AA and AM in the east and between AN and AM in the west. The emergent pattern was not so much clinal, but more as one having a fairly well-defined boundary. These patterns have now been supported by two genome-scale surveys (Jinam et al. 2017 with 516,832 SNPs) and (Hudjashov et al. 2017 with 510,000 SNPs). The consensus of these reports is that AN admixture and farming culture only spread slowly westwards into Java and Sumatra. Set against this view is a single intriguing aDNA report from Gua Harimau in eastern Sumatra where human remains are dated as being early in the Austonesian Diaspora and returned provisional AN type mtDNA haplotype signals (Matsumura et al. 2018). This datum is interesting, but by no means conclusive.

In closing this section, we note that genomic science in Indonesia now seems to have advanced to a fairly sophisticated stage with a recent epigenomic study of ancestral patterns of gene regulation and environmental adaptation (Natri et al. 2020).

Madagascar

Perhaps the most unlikely chapter in the entire Austronesian Diaspora is the settlement of the Indian Ocean islands, The Comoros and Madagascar (Cox et al., 2012). The Comoros AN islanders are now more or less fully admixed with Indian and Arab migrants (Gourjon et al. 2011; Msadie et al. 2010). While those in Madagascar have preserved their identity and

Austronesian language despite extensive contact and admixture with Bantu-speaking Africans who have also settled there. This account has been well supported by numerous genetic studies (e.g., see Chambers and Edinur, 2015 for references and further details). More recently the power of genomic analysis has been brought to bear on this account (Pierron et al. 2017) which adds detail to the general picture. The first settlers arrived around 2000 ya and built villages on the northern and western edges of the island. They have subsequently translocated further south in a series of movements and now occupy a region in the central western part of the island – see Figure 7.

Admixture with their Bantu-speaking African neighbours has been gender-biased with a predominance of introgression into the AN residents coming from Bantu males (reflecting the well-known pattern previously seen in AN/AM admixture in coastal PNG (Hage and Marck 2003 and see Palencia-Madrid et al. 2019). One further surprising detail is that the distal origin of the Madagascar Austronesians appears to be the Banjar of Borneo. In a second study (Brucato et al. 2019) it was revealed that there are relict signals of AN genomes in Somalia and Yemen (Arabian Peninsula). The first of these traces back to Madagascar, but the second leads direct to the Banjar of Borneo.

Most Recent Movements of Austronesian People

One of the lesser known episodes of the Austronesian Diaspora was the translocation of the Indianized colony of Tambraling from northern Malaysia to Sri Lanka around 700 ya, where they became part of the Tamil community. This group increased in numbers as they were joined by slaves exported from Java by the Dutch and later by Malaysian exiles transported by the British Colonial Administration (Hammarström et al. 2017). The Dutch and British also transported people, who were arguably political exiles, to South Africa. Today they form the distinctive Cape Malay community in their brightly painted suburbs of Cape Town.

Finally, even today Austronesians are still on the move as economic migrants mainly from the Pacific Islands to the large cities in eastern Australia and the west coast of the United States of America.

Conclusions

The Austronesian Diaspora is a mystery that has fascinated scholars for a very long time. Indeed, discussion of this topic dates all the way back to the time of first encounter through Cook's voyages. A contested orthodox genes, language and culture account has emerged starting from Taiwan moving south and splitting into three major arms; west into Micronesia, south westward via Island Southeast Asia to Madagascar and eastwards via coastal PNG to occupy the Polynesian Triangle with rise and fall of the Lapita Cultural Complex along the way.

The past decade has seen introduction of new and powerful genomics techniques including the deep time reach of aDNA methodology. This review presents a comprehensive updated account of events. Its novelty lies in what it can contribute to the synthetic picture by including a full survey of new genetic studies and how their emergent evidence has significantly challenged the present orthodox account of this episode of human history. Now, for the first time, scientists have been able to directly examine well-dated ancestral populations and their associated cultures. This is only possible on rare occasions but does give geneticists the opportunity to question the ancestors themselves, rather than having to rely on inferences made by examining their descendants. Consequently, present our understanding of population movements in Near Oceania has become more complex. One branch of the earliest AN people has been shown to have carried a pre-established Lapita Culture direct to Vanuatu, while a second branch admixed with the AM residents of northern PNG. It is this second group that are the ancestors of the Polynesians and who sailed

eastwards settling the Solomons from their eastern end in leapfrog fashion. It is likely that they then made another leapfrog jump by-passing Vanuatu entirely and landing at Tonga and Samoa. It seems that a period of slow population growth then preceded their departure for the Leeward Society Islands. We now have much more complete knowledge of their situation prior to their further dispersal to fill up the islands contained within with Polynesian Triangle.

The westward arm of the diaspora includes widespread settlement of more or less unadmixed AN in Borneo as a source for settling the eastern end of Indonesia and giving rise to the Malays. Interestingly, the study of Sun et al (2020), noted earlier in relation to possible AN ancestors in northern China, not only supports this view of Malay origins but also supports the more general revised account of the Austronesian Diaspora as presented here. It is in Malaysia and Indonesia that the role of Austroasiatic-speaking migrants from mainland Asia becomes important and this too is now far better understood. Remarkably, people from Borneo have now been shown to have voyaged to, and settled in, Madagascar and Yemen (Brucato et al. 2019).

Finally, it has always seemed anomalous that AM should have so much genomic content from ancient hominins, particularly Denisovans. It is now clear that this DNA must have been introduced in Near Oceania because their Onge ancestors (Andaman Islands) have only Neanderthal markers. Hence, the later Denisovan introgression is dated at 60 - 40 kya taking place in, or near, PNG. It is still undecided if this represents one or two episodes, or if it includes some *H. erectus* DNA. Importantly, many of the introgressed regions seem to have been the targets of positive natural selection including immune system markers associated with disease resistance as first recognised by Abi-Rached et al. (2011). Similar patterns have been reported for AN/AM admixture (Edinur et al., 2012) where the HLA-C split is 40:60 but for HLA-B it is 90:10. Isshiki et al. (2020) also recently reported strong signals of natural selection in the HLA regions of these admixed genomes.

It is clear that the advent of this new technology has changed the analytical culture for investigators permanently. The industrial standard for reporting has increased by several orders of magnitude. No longer will it be satisfactory to submit studies with genetic analyses of single targets in just 100-200 subjects. The new surveys described here include hundreds of participants from whom whole genomes were obtained or data from genome-wide SNP panels numbering in hundreds or thousands. The work is typically carried out by big international teams often involving 20 or 30 authors, often from 10 or more different institutions, including a substantial investment in bioinformatics. Consider, for example, the study carried out in Madagascar by Pierron et al. (2017) – this included 257 villages across which 2.5 M SNPs were tested in 700 new genomes and gaining data for 2691 full mtDNA sequences and 1554 Y chromosome haplotypes collected by 36 authors working in 12 different laboratories. Various other examples are quoted in brief in other parts of this review as further illustrations of the new scale of international operation.

This is not to say that smaller scale studies have entirely lost their worth, particularly those concerned with medical genetics (e.g., as stressed by Natri et al. 2020). They confirm ancestry relationships and help to stress AN unity. It is well known that Pacific communities share particularly elevated incidences of diabetes, gout, coronary heart disease and other metabolic disorders (see Matisoo-Smith 2015 and references therein). This will be increasingly important as the long-awaited era of personal pharmacogenomics finally dawns (see also comments by Cox pp. 115-116 in Bellwood 2017). For the present this ethnicity-based approach remains the best way forward despite being considered by some (erroneously in our view) as ‘race-based medicine’. Nonetheless, the value of using ethnic information about patients is widely recognised as being of fundamental value (Borrell et al., 2021). For instance, the drug Plavix (Colpidogrel) carries a black box message on its package warning that poor metabolisers may not receive the full benefit and remain at elevated risk of heart

attack. Such people include up to 75% of Pacific Islanders who lack the *CYP2C19* genetic variant necessary to metabolise the pro-drug into its active form. It is clear that AN people will need to develop their own initiatives or strong strategic relationships if they are to benefit fully from the opportunities provided by these developments. It is unlikely that the data required will come from elsewhere unprompted.

In closing, we note that the mode of conduct for research in this academic field seems to have been changed beyond all expectations. The standard ‘Out of Taiwan’ account has been consolidated and elaborated, but the emerging picture of later migration patterns is more complex and fragmented. It is likely that such changes may continue for some time to come.

Note Added in Proof

Three significant new genomic studies have been published over recent months. The first two (Choin et al. 2021 and Tätte et al. 2021) are more or less in accord with the account given in the main body of our text. The third by Larena et al. (2021) is a somewhat radical departure and offers to blend the otherwise competing ‘Out of Southeast Asia’ and ‘Out of Taiwan’ versions of Austronesian history. It is a particularly large-scale survey of ~2.3 million SNP markers in 1,028 individuals representing 115 indigenous population groups living in the Philippines. As such, it deserves serious detailed consideration.

Much of the Larena et al. (2021) reconstruction (see their Figs 3A - D) hinges on sea level changes which meant that much of the southeast Asia region was dry land up to the end of the Last Glacial Maximum around 11.7 kya. This facilitated population movements starting with Negritos and split into four lineages in mainland southeast Asia, Australia plus PNG and the Philippines (north and south). This all fits well with our standard model. Next in their scheme, the Negrito lineage interacts in some undefined way with ‘Basal East Asians’ (so-named, but otherwise undescribed) which gives rise to ancestral Austroasiatic,

MANOBO-like and Cordilleran peoples. The last lineage eventually becomes, more or less, identified with those we have called Austronesians. The MANOBO-like lineage is a late stage (before 12 kya) migration to the southern Philippines. Larena et al. (2021) extend the early Austroasiatic expansion (Htin/Malabri-related SAMA-like in their terms) beyond Indonesia and Malaysia across Borneo and on to the Philippines (a fairly significant extension of our account). Finally, post sea level rise, but before 7 kya, the Cordilleran people expand across the ocean to settle both Taiwan and the Philippines. In summary, this seems to represent a reasonable blend of previous ideas.

After all due consideration we feel that this account may raise as many questions as it provides answers. In particular, we point to 1) that the postulated relationships between ancestral and derivative groups in mainland Asia is perhaps unconventional and certainly not made fully explicit (as noted earlier), 2) that the model calls for parallel, but non-identical introgression of Denisovan genomes in Philippines and PNG and 3) the authors fails to clearly identify affinities between Cordillerans and Austronesians. Finally, the dates above are derived from molecular trees and coalescent models using genomic data from modern people and not actual physical archaeological evidence. In our view contemporary genetics plus tree models give a dated network of relationships that can only create hypotheses about what went on in the past and when it all took place. Actual dates with real authority come only from buried artefacts and genomic data from ancient humans from the same sites tell us who really lived there and when. Vanuatu history is the object lesson as we have stressed in the main body of our paper.

Acknowledgments

Geoff Chambers thanks Victoria University of Wellington for Alumnus Scholar support.

Work in the Edinur laboratory is supported by internal grants from Universiti Sains Malaysia

and Fundamental Research Grant Scheme: 304/PPSK/6150159 awarded by the Ministry of Higher Education, Malaysia. We are grateful for constructive inputs received from academic colleagues during the preparation of this most particularly to Andrew Crowe and to two anonymous and to two anonymous reviewers for their guidance on revisions. We include a new set of illustrations with extensive help from Norul Hajar Binti Che Ghazali at USM, Kelantan to whom the authors jointly express their sincere gratitude.

Received 10 October 2020; accepted for publication 27 March 2021.

Literature Cited

- Abi-Rached, L., M. J. Jobin, S. Kulkarni et al. 2011. The shaping of modern immune systems by multiregional admixture with archaic humans. *Science* 334:89–94.
- Arenas, M., A. Gorostiza, and J. M. Baquero. 2020. The early peopling of the Philippines based on mtDNA. *Sci. Rep.* 10:4901.
- Bellwood, P. 1991. The Austronesian dispersal and the origin of languages. *Sci. Am.* 265:88–93.
- Bellwood, P. 2005. Examining the farming/language dispersal hypothesis in the East Asian context. In *The Peopling of East Asia*, L. Sagan, R. Blench, and A. Sanchez-Mazas, eds. New York: Routledge Curzon Press, 17–30.
- Bellwood, P. 2017. *First Islanders*. Oxford, UK: John Wiley and Sons.
- Bellwood, P. 2018 The search for ancient DNA heads east. *Science* 361:31–32
- Bellwood, P., G. K. Chambers, M. Ross et al. 2011. Are “cultures” inherited? Multidisciplinary perspectives on the origins of Austronesian-speaking peoples prior to 1000 BC. In *Investigating Archaeological Cultures: Material Culture, Variability and Transmission*, B. W. Roberts and M. V. Linden, eds. Dordrecht, NL: Springer, 321–354.
- Benton, M. C., R. A. Lea, D. McCartney-Coxon et al. 2019. Genome-wide allele-specific methylation is enriched at gene regulatory regions in a multi-generation pedigree from the Norfolk Island isolate. *Epigenetics Chromatin* 12:60.
- Benton, M., D. McCartney-Coxon, D. Eccles et al. 2012. Complete mitochondrial genome sequencing reveals novel haplotypes in a Polynesian population. *PLoS One* 7:e35026.
- Bergström, A., S. A. McCarthy, R. Hui et al. 2020. Insights into human genetic variation and population history from 929 diverse genomes. *Science* 367:eaay5012.

- Bergström, A., S. J. Oppenheimer, A. J. Mentzer et al. 2017. A Neolithic expansion, but strong genetic structure, in the independent history of New Guinea. *Science* 357:1,160–1,163.
- Borrell, L. N., J. R. Elhawary, E. Fuentes-Afflick et al. 2021. Race and genetic ancestry in medicine — A time for reckoning with racism. *N. Engl. J. Med.* 384:474–480.
- Brucato, N., V. Fernandes, P. Kusuma et al. 2019. Evidence of Austronesian genetic lineages in East Africa and South Arabia: Complex dispersal from Madagascar and Southeast Asia. *Genome Biol. Evol.* 11:748–758.
- Chambers, G. K., and H. A. Edinur. 2015. The Austronesian Diaspora: A synthetic total evidence model. *Glob. J. Anthropol. Res.* 2:53–65.
- Chen, L.-R., J. A. Trejaut, Y.-H. Lai et al. 2019. Mitochondrial DNA polymorphisms of the Saisiyat indigenous group of Taiwan, search for a Negrito signature. *Edelweiss J. Biomed. Res. Rev.* 1:12–18.
- Choin, J., J. Mendoza-Revilla, L. R. Arauna et al. 2021. Genomic insights into population history and biological adaptation in Oceania. *Nature* 592:583–589.
- Cox, M. P., M. G. Nelson, M. K. Tumonggor et al. 2012. A small cohort of Island Southeast Asian women founded Madagascar. *Proc. Biol. Sci.* 279:2,761–2,768.
- Crowe, A. 2018. *Pathway of the Birds*. Auckland, NZ: Bateman.
- Deng, L., B. P. Hoh, D. Lu et al. 2014. The population genomic landscape of human genetic structure, admixture history and local adaptation in Peninsular Malaysia. *Hum. Genet.* 133:1,169–1,185.
- Deng, L., B. P. Hoh, D. Lu et al. 2015. Dissecting the genetic structure and admixture of four geographical Malay populations. *Sci. Rep.* 5:14375.

- Deng, L., H. Lou, X. Zhang et al. 2019. Analysis of five deep-sequenced trio-genomes of the Peninsula Malaysia Orang Asli and North Borneo populations. *BMC Genomics* 20:842.
- Détroit, F., A. S. Mijares, J. Corny et al. 2019. A new species of *Homo* from the late Pleistocene of the Philippines. *Nature* 568:181–186.
- Diamond, J. M. 1988. Express train to Polynesia. *Nature* 336:307–308.
- Donohue, M., and T. Denham. 2011. Languages and genes attest different histories in island Southeast Asia. *Ocean. Linguist.* 50:536–542.
- Duggan, A. T., and M. Stoneking. 2014. Recent developments in the genetic history of East Asia and Oceania. *Curr. Opin. Genet. Dev.* 29:9–14.
- Edinur, H. A., P. P. J. Dunn, L. Hammond et al. 2012. Using HLA loci to inform ancestry and health in Polynesian and Maori populations. *Tissue Antigens* 80:509–522.
- Friedlaender, J., ed. 2007. *Genes, Language, and Culture History in the Southwest Pacific*. New York: Oxford University Press.
- Fu, R., S. S. Mokhtar, M. E. Phipps et al. 2018. A genome-wide characterization of copy number variants in native populations of Peninsular Malaysia. *J. Hum. Genet.* 26:886–897.
- Gosling, A. L., and E. A. Matisoo-Smith. 2018. The evolutionary history of human settlement of Australia and the Pacific. *Curr. Opin. Genet. Dev.* 53:53–59.
- Gourjon, G., G. Boëtsch, and A. Deggioanni. 2011. Gender and population history: Sex bias revealed by studying genetic admixture of Ngazidja population (Comoro Archipelago). *Am. J. Phys. Anthropol.* 144:653–660.
- Gray, R. D., A. J. Drummond, and S. J. Greenhill. 2009. Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* 323:479–483.

- Hage, P., and J. Marck. 2003. Matrilineality and the Melanesian origin of Polynesian Y chromosomes. *Curr. Anthropol.* 44:S121–S127.
- Hammarström, H., R. Forkel, and M. Haspelmath, eds. 2017. *Sri Lanka Malay. Glottolog 3.0.* Jena, DE: Max Planck Institute for the Science of Human History.
- Harris, D., M. D. Kessler, and A. C. Shetty. 2020. Evolutionary history of modern Samoans. *Proc. Natl. Acad. Sci. U. S. A.* 117:9,458–9,465.
- Heyerdahl, T. 1952. *American Indians in the Pacific: The Theory behind the Kon-Tiki Expedition.* London: Allen and Unwin.
- Hudjashov, G., P. Endicott, H. Post et al. 2018. Investigating the origins of eastern Polynesians using genome-wide data from the Leeward Society Isles. *Sci. Rep.* 8:1823.
- Hudjashov, G., T. M. Karafet, D. J. Lawson et al. 2017. Complex patterns of admixture across the Indonesian Archipelago. *Mol. Biol. Evol.* 34:2,439–2,452.
- HUGO Pan-Asian SNP Consortium. 2009. Mapping human genetic diversity in Asia. *Science* 326:1,541–1,545.
- Hung, H.-C. 2019. History and current debates of archaeology in Island Southeast Asia. In *Encyclopedia of Global Archaeology*, C. Smith, ed. Switzerland: Springer Nature, 1–22.
- Hung, H.-C. 2011. The first settlement of Remote Oceania: The Philippines to the Marianas. *Antiquity* 85:909–926.
- Ioannidis, A. G., J. Blanco-Portillo, K. Sandoval et al. 2020. Native American gene flow into Polynesia predating Easter Island settlement. *Nature* 583:572–577.
- Isshiki, M., I. Naka, Y. Watanabe et al. 2020. Admixture and natural selection shaped genomes of an Austronesian-speaking population in the Solomon Islands. *Sci. Rep.* 10:6872.

- Jacobs, G. S., G. Hudjasov, L. Saag et al. 2019. Multiple deeply divergent Denisovan ancestries in Papuans. *Cell* 177:1–12.
- Jinam, T. A., L. C. Hong, M. E. Phipps et al. 2012. Evolutionary history of continental Southeast Asians: “Early train” hypothesis based on genetic analysis of mitochondrial and autosomal DNA data. *Mol. Biol. Evol.* 29:3,513–3,527.
- Jinam, T. A., M. E. Phipps, F. Aghakhanian et al. 2017. Discerning the origins of the Negritos, First Sundaland People: Deep divergence and archaic admixture. *Genome Biol. Evol.* 9:2,013–2,022.
- Juhan, W. K. W., N. A. M. Tamin, M. H. R. M. Daud et al. 2014. A whole genome analyses of genetic variants in two Kelantan Malay individuals. *Hugo J.* 8:4.
- Kayser, M. 2010. The human genetic history of Oceania: Near and remote views of dispersal. *Curr. Biol.* 20:R192–R201.
- Kimura, R., J. Ohashi, Y. Matsumura et al. 2008. Gene flow and natural selection in oceanic human populations inferred from genome-wide SNP typing. *Mol. Biol. Evol.* 25:1,750–1,761.
- Ko, A. M., C. Y. Chen, Q. Fu et al. 2014. Early Austronesians: Into and out of Taiwan. *Am. J. Hum. Genet.* 94:426–436.
- Larena, M., F. Sanchez-Quinto, P. Sjödin et al. 2021. Multiple migrations to the Philippines during the last 50,000 years. *Proc. Natl. Acad. Sci. U. S. A.* 118:e2026132118.
- Li, H., B. Wen, S. J. Chen et al. 2008. Paternal genetic affinities between western Austronesians and Daic populations. *BMC Evol. Biol.* 8:146.
- Lipson, M., O. Cheronet, S. Mallick et al. 2018a. Ancient genomes document multiple waves of migration in Southeast Asian prehistory. *Science* 361:92–95.
- Lipson, M., P.-R. Loh, N. Patterson et al. 2014. Reconstructing Austronesian population history in Island Southeast Asia. *Nat. Commun.* 5:4689.

- Lipson, M., P. Skoglund, M. Spriggs et al. 2018b. Population turnover in Remote Oceania shortly after initial settlement. *Curr. Biol.* 28:1157–1165.e7.
- Lipson, M., M. Spriggs, F. Valentin et al. 2020. Three phases of ancient migration shaped the ancestry on human populations in Vanuatu. *Curr. Biol.* 30:4,846–4,856.
- Liu, D., N. Duong, N. Ton et al. 2020. Extensive ethnolinguistic diversity in Vietnam reflects multiple sources of genetic diversity. *Mol. Biol. Evol.* 37:2,503–2,519.
- Macholdt, E., L. Arias, N. T. Duong et al. 2019. The paternal and maternal genetic history of Vietnamese populations. *Eur. J. Hum. Genet.* 28:636–645.
- Malaspinas, A.-S., M. C. Westaway, C. Muller et al. 2016. A genomic history of Aboriginal Australia. *Nature* 538:207–214.
- Mallick, S., H. Li, M. Lipson et al. 2016. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* 538:701–706.
- Matisoo-Smith, E. 2015. Ancient DNA and the settlement of the Pacific. *J. Hum. Evol.* 79:93–104.
- Matsumura, H., K. Shinoda, T. Shimanjuntak et al. 2018. Cranio-morphometric and aDNA corroboration of the Austronesian dispersal model in ancient Island Southeast Asia: Support from Gua Harimau, Indonesia. *PLoS One* 13:e0198689.
- McColl, H., F. Racimo, L. Vinner et al. 2018. The prehistoric peopling of Southeast Asia. *Science* 261:88–92.
- Mirabel, S., A. M. Cadenas, R. Garcia-Bertrand et al. 2013. Ascertaining Taiwan as a source of the Austronesian expansion. *Am. J. Phys. Anthropol.* 150:551–564.
- Msadie, S., A. Ducourneau, G. Boetsch et al. 2010. Genetic diversity on the Comoros Islands shows seafaring as major determinant of human bicultural evolution in the Western Indian Ocean. *Eur. J. Hum. Genet.* 19:89–94.

- Natri, H. M., K. S. Bobwik, P. Kusuma et al. 2020. Genome-wide DNA methylation and gene expression patterns reflect genetic ancestry and environmental differences across the Indonesian archipelago. *PLoS Genet.* 16:e1008749.
- Norhalifa, H. K., F. H. Syaza, G. K. Chambers et al. 2016. The genetic history of Peninsular Malaysia. *Gene* 586:129–135.
- Oppenheimer, S. 1998. *Eden in the East: The Drowned Continent of Southeast Asia*. London: Phoenix Press.
- Palencia-Madrid, L., M. Baeta, P. Villaescusa et al. 2019. The Marquesans at the fringes of the Austronesian expansion. *Eur. J. Hum. Genet.* 72:801–810.
- Pierron, D., M. Heiske, H. Razfindrazaka et al. 2017. Genomic landscape of human diversity across Madagascar. *Proc. Natl. Acad. Sci. U. S. A.* 114:E6498–E6506.
- Pinhasi, R., D. Fernandes, K. Sirak et al. 2015. Optimal ancient DNA yields from the inner ear part of the human petrous bone. *PLoS One* 10:e0129102.
- Pischedda, S., R. Barral-Arca, A. Gómez-Carballa et al. 2017. Phylogeographic and genome-wide investigations of Vietnam ethnic groups reveal complex historical and demographic movements. *Sci. Rep.* 7:12630.
- Posth, C., K. Nägele, H. Colleran et al. 2018. Language continuity despite population replacement in Remote Oceania. *Nat. Ecol. Evol.* 2:731–740.
- Pugach, I., A. T. Duggan, A. D. Merriweather et al. 2018. The gateway from Near into Remote Oceania: New insights from genome-wide data. *Mol. Biol. Evol.* 35:871–886.
- Pugach, I., A. Hübner, and H.-C. Hung. 2020. Ancient DNA from Guam and the peopling of the Pacific. *Proc. Natl. Acad. Sci. U. S. A.* 118:e2022112118.
- Pugach, I., and M. Stoneking. 2015. Genome-wide insights into the genetic history of human populations. *Investig. Genet.* 6:6.

- Resnik, D. B. 1999. The Human Genome Diversity Project: Ethical problems and solutions. *Politics Life Sci.* 18:15e23.
- Rogers, A. R., N. S. Harris, A. A. Achenbach et al. 2020. Neanderthal-Denisovan ancestors interbred with a distantly related hominin. *Sci. Adv.* 6:eaay5483.
- Sallah, M. Z., L. K. Teh, L. S. Lee et al. 2013. Systematic pharmacogenomics analysis of a Malay whole genome: Proof of concept for personalized medicine. *PLoS One* 8:e71554.
- Sheppard, P. 2019. Early Lapita colonisation of Remote Oceania: An update on the leapfrog hypothesis. In *Debating Lapita*, S. Bedford and M. Spriggs, eds. Australia: ANU Press, 135–153.
- Skoglund, P., and I. Mathieson. 2018. Ancient DNA genomics: The first decade. *Annu. Rev. Genomics Hum. Genet.* 19:8.1–8.24.
- Skoglund, P., C. Posth, K. Sirak et al. 2016. Genomic insights into the peopling of the southwest Pacific. *Nature* 538:510–513.
- Soares, P., T. Rito, J. A. Trejaut et al. 2011. Ancient voyaging and Polynesian origins. *Am. J. Hum. Genet.* 88:239–247.
- Spriggs, M., and D. Reich. 2020. An ancient DNA Pacific journey: A case study of collaboration between archaeologists and geneticists. *World Archaeol.* 51:620–639.
- Summerhayes, G., E. Matisoo-Smith, H. Mandui et al. 2010. Tamuarawai (EQS): An early Lapita site on Emirau, New Ireland, PNG. *J. Pac. Archaeol.* 1:62–75.
- Sun, J., L.-H. Wei, L.-X. Wang et al. 2020. Paternal gene pool of Malays in Southeast Asia and its applications for the early expansion of the Austronesians. *Am. J. Hum. Biol.* 33:e23486.
- Tabbada, K. A., J. Trejaut, J. H. Loo et al. 2010. Philippine mitochondrial DNA diversity: A populated viaduct between Taiwan and Indonesia. *Mol. Biol. Evol.* 27:21–31.

- Tätte, K., E. Metspalu, L. Palencia-Madrid et al. 2021 (online ahead of print). The Ami and Yami Aborigines of Taiwan and their genetic relationships to East Asian and Pacific populations. *Eur. J. Hum. Genet.*
- Tucci, S., S. H. Vohr, R. C. McCoy et al. 2018. Evolutionary history and adaptation of a human pygmy population on Flores Island, Indonesia. *Science* 361:511–516.
- Vernot, B., S. Tucci, J. Kelso et al. 2016. Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals. *Science* 352:235–239.
- Wallin, P. 2020. Native South Americans were early inhabitants of Polynesia. *Nature* 583:524–525.
- Wilmshurst, J. M., T. L. Hunt, C. P. Lipo et al. 2011. High-precision radiocarbon dating shows recent and rapid initial human colonization of East Polynesia. *Proc. Natl. Acad. Sci. U. S. A.* 108:1,815–1,820.
- Wollstein, A., O. Lao, and C. Becker. 2010. Demographic history of Oceania inferred from genome-wide data. *Curr. Biol.* 20:1,983–1,992.
- Wong, L. P., R. T. Ong, W. T. Poh et al. 2013. Deep whole-genome sequencing of 100 southeast Asian Malays. *Am. J. Hum. Genet.* 92:52–66.
- Wu, D., J. Dou, X. Chai et al. 2019. Large-scale whole-genome sequencing of three diverse Asian populations in Singapore. *Cell* 179:736–749.
- Xu, S., I. Pugach, M. Stoneking et al. 2012. Genetic dating indicates that the Asian-Papuan admixture through Eastern Indonesia corresponds to the Austronesian expansion. *Proc. Natl. Acad. Sci. U. S. A.* 109:4,574–4,579.
- Yew, C.-W., M. Z. Hoque, J. Pugh-Kitingan et al. 2018a. Genetic relatedness of Indigenous ethnic groups in northern Borneo to neighboring populations from Southeast Asia as inferred from genome-wide SNP data. *Ann. Hum. Genet.* 82:216–226.

Yew, C.-W., D. Lu, L. Deng et al. 2018b. Genomic structure of the native inhabitants of Peninsular Malaysia and North Borneo suggests complex human history in Southeast Asia. *Hum. Genet.* 137:161–173.

Figure Captions

Figure 1. A general overview of Austronesian population movements. This global map was developed from one first shown in Bellwood et al. (2011) is taken from Benton et al. (2012). It shows the global spread of Austronesian people with routes and dates as known at that time. Its purpose is to present a picture of the 5000-year expansion of Austronesian-speaking people from a focal origin in mainland south east Asia to Madagascar in the west to Columbia in the east. The figure represents received wisdom at the dawn of the genomics era. Subsequent detailed revisions to this general picture are shown in the following maps (Figures 2 – 7).

Figure 2. A close up view of early migration through southeast Asia. This diagram shows the overall movements of Australo-Melanesians (Blue) into the Asia Pacific region starting from early settlement of Indonesia, Australia and Papua New Guinea. This is followed by their later spread into near Oceania and through mainland Asia. Southerly movements of Austroasiatic-speaking people from mainland Asia (Green) reach Indonesia and Malaysia. Finally, Austronesian people (Red) exited Asia and moved into Island Southeast Asia via Taiwan and spread south via the Philippines splitting into three branches (BRANCH 1-4 as named on the diagram) and giving rise to Proto-Polynesians (Purple) by admixture in northern coastal Papua New Guinea. The later migrations from Borneo to Madagascar and elsewhere (BRANCH 5) are not shown. Both Austronesians and Proto-Polynesians went on to settle in Vanuatu and nearby islands in Remote Oceania as did Australo-Melanesians – shown in greater detail in Figures 3A and 3B. Migration routes across the Pacific including to South America are illustrated in Figure 4 and later voyages by Austronesian from Borneo to Indian Ocean destinations including Madagascar are shown in Figure 7. NB: In this and all

other figures well established migration routes are shown as solid lines and hypothetical pathways are shown with dotted lines and circles

Figures 3A and 3B. Dispersal from northern PNG to western Polynesia. A: Early AN voyagers from the Borneo/Philippines area settle northern coastal PNG and offshore islands and move directly or indirectly on to take Lapita culture to Vanuatu, New Caledonia, Fiji Tonga and Samoa. The resident PNG Australo-Melanesians (Blue) settle offshore Islands in the Bismarcks and Solomons. B: The AM who have occupied PNG offshore islands move on to displace AN from Vanuatu and New Caledonia and also occupy Fiji. Proto-Polynesians (Purple) move from northern coastal PNG and its near offshore islands to occupy the Solomons from the eastern end and develop minority settlements in Vanuatu and Fiji, but majority populations in Samoa, Tonga and Niue.

Figure 4. Settlement of eastern Polynesia. This figure shows how late stage movements from Western Polynesia (Purple) reached to Leeward Society Islands and Marquesas Islands which were a centre for dispersal across the Polynesian Triangle. Contact was also made with the Columbian (Zenu) tribe (Orange) resulting in admixture (Brown) which spread south as shown. Wider Polynesian voyages across the region to Hawaii, Aotearoa and Rapa Nui (via Nuku Hiva) are not shown in detail. This figure is taken from Wallin (2020) with modification

Figure 5. Settlement of Aotearoa (New Zealand) and neighbouring islands. The map shows movement of Polynesian people into New Zealand from Raiatea Lagoon direct and via Rarotonga in the Cook Island group (still debated). New Zealand is known to be the starting point for later voyages to neighbouring island groups.

Figure 6. Austronesian settlement of Island Southeast Asia. This map shows movement from the Philippines through Borneo to Malaysia (Red) an eastern branch leads to Indonesia where admixture begins with pre-established Australo-Melanesians (Blue) a western side branch represents trading voyages to Vietnam and Thailand with cultural exchange. In Malaysia there is a three-way interaction now including early Austroasiatic-speaking people (Green), Negrito (Blue) and Proto-Malay (Red) who form an original Orang Asli community giving rise to the later majority Deutero-Malay (Austronesian) modern population of Malaysia.

Figure 7. Indian Ocean voyages from Borneo to Madagascar and other locations. This figure shows the remarkable early Austronesian voyages (Red) from Borneo to Madagascar and the Comoros Islands. Later movements include those internal to Madagascar and external to coastal Africa including Somalia. There is present admixture with Bantu immigrants (Green) from Mainland Africa. There is also evidence (see text) for a recent more northerly voyage from Borneo reaching Yemen.

Figure 1.

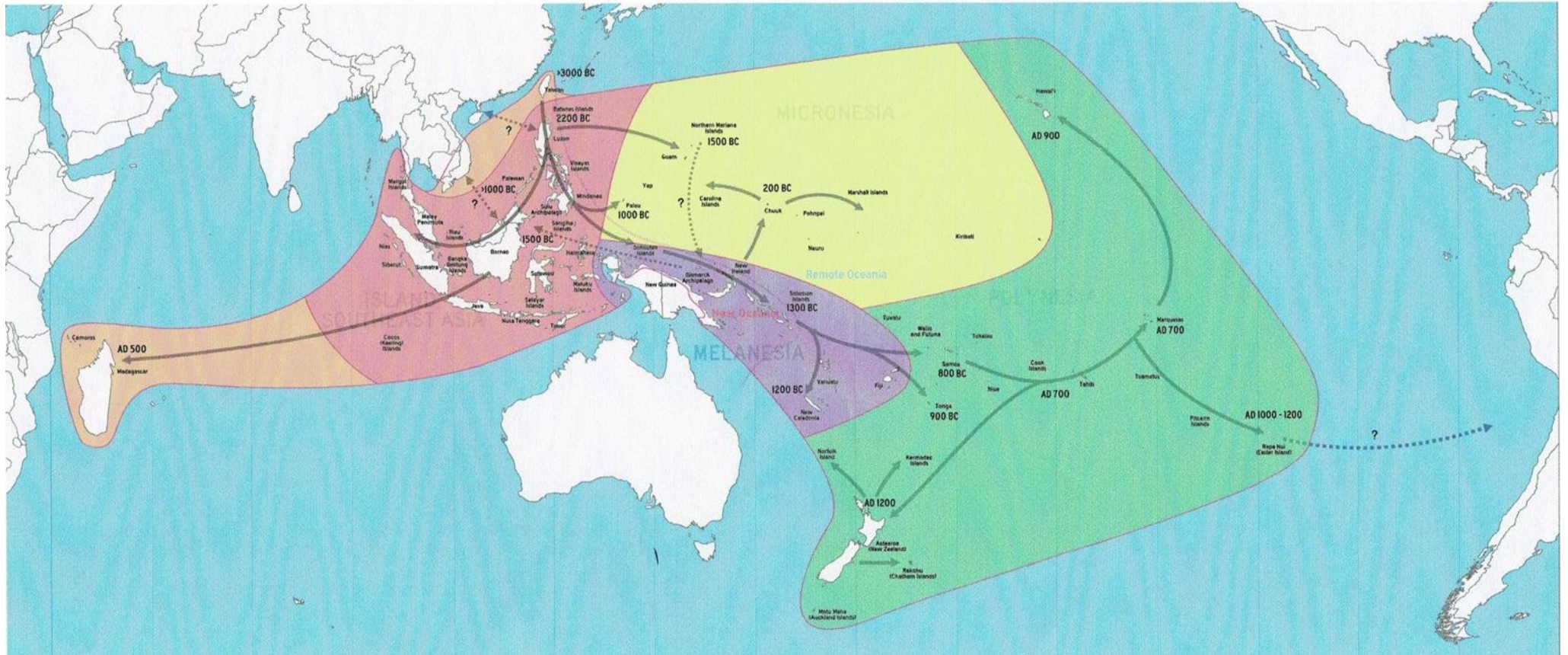


Figure 2.

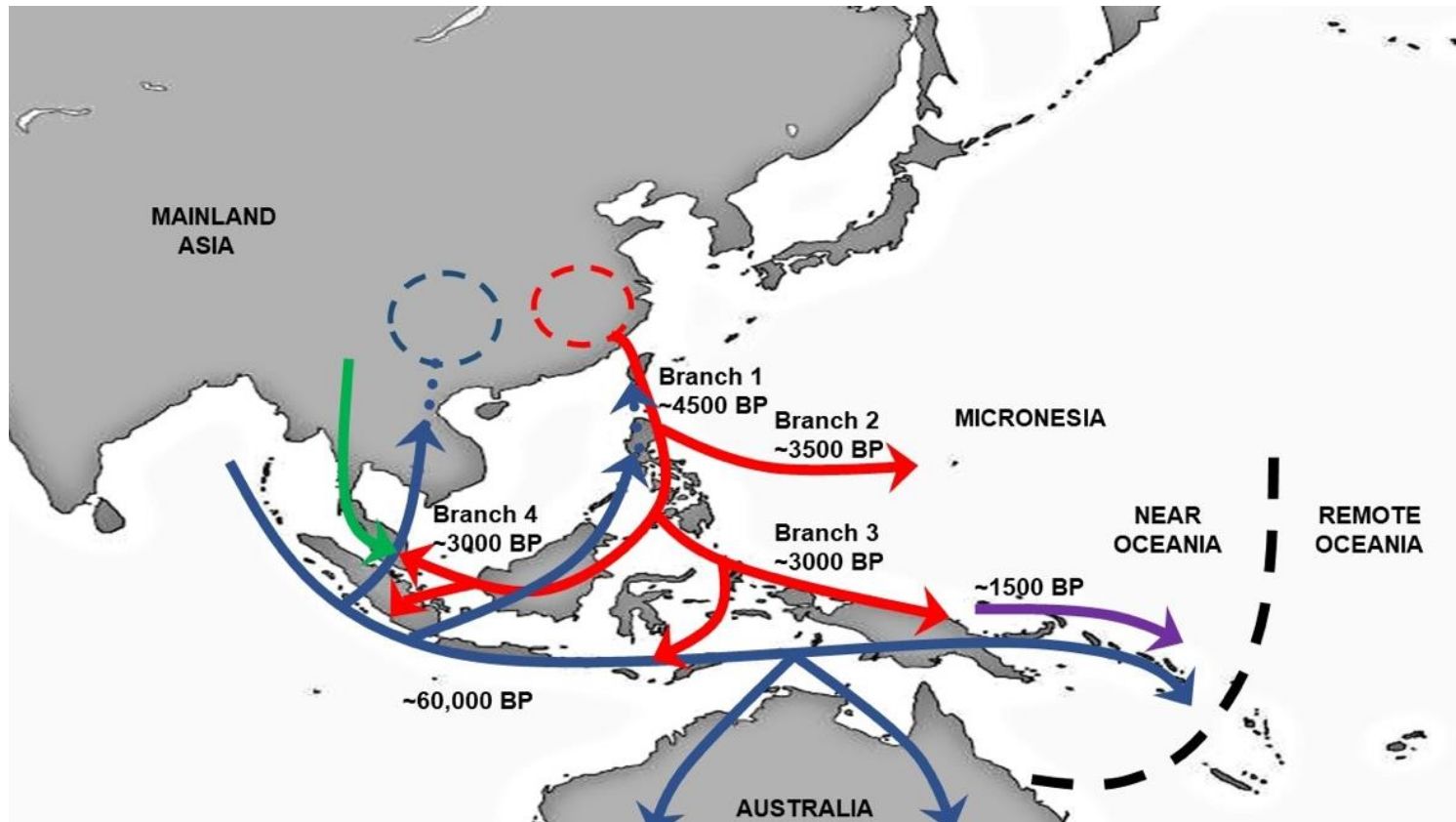


Figure 3A.

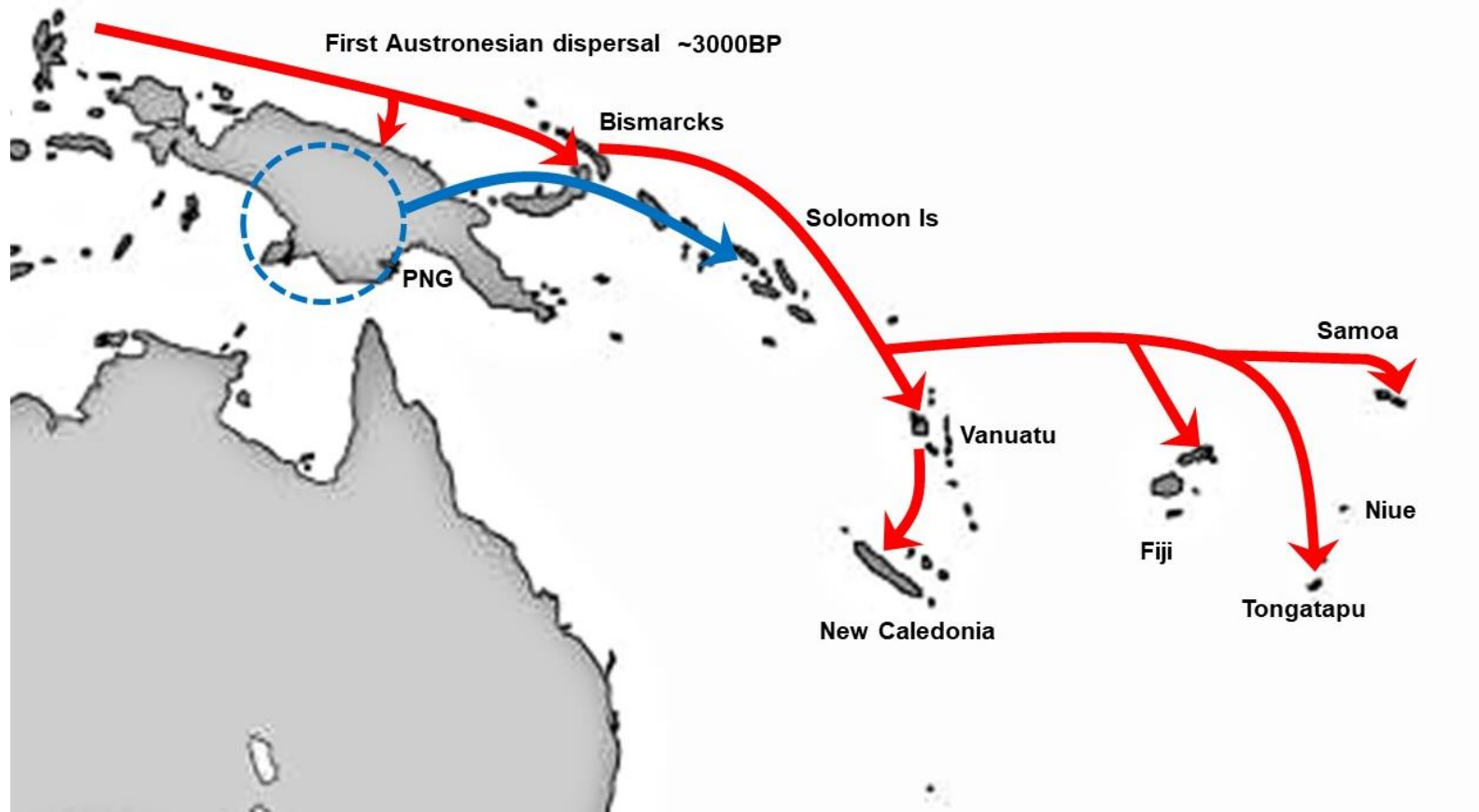


Figure 3B.

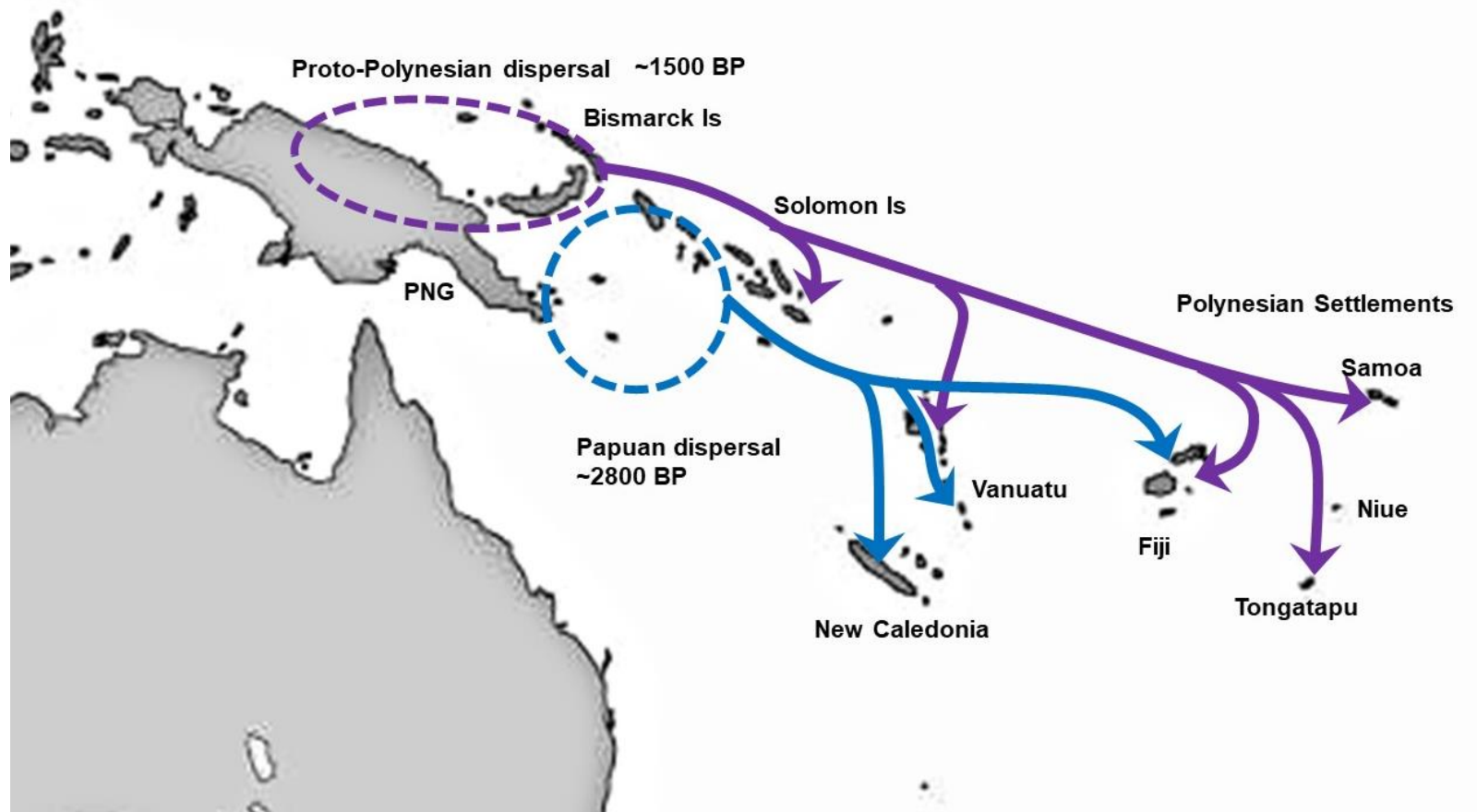


Figure 4.

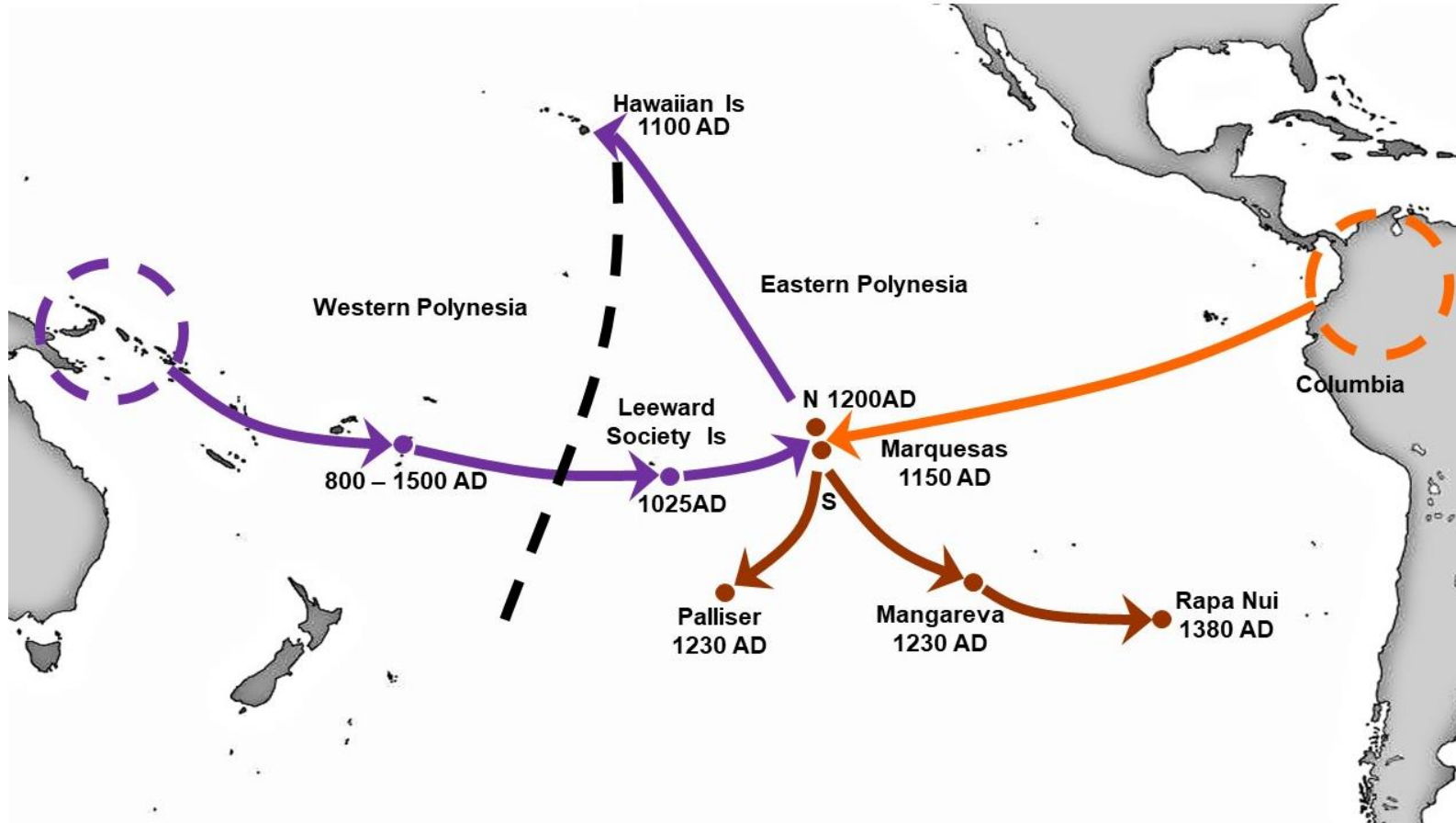


Figure 5.

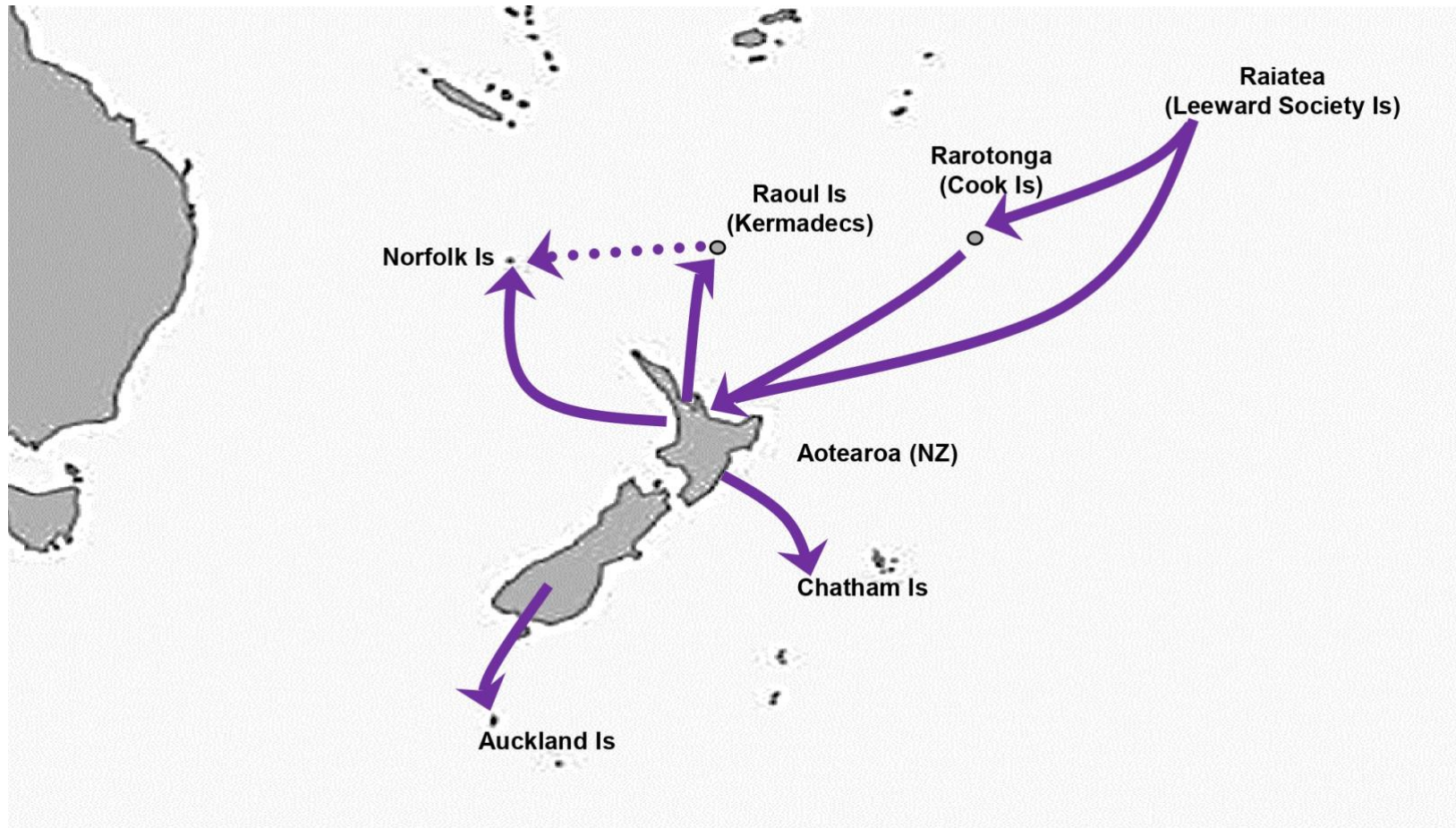


Figure 6.

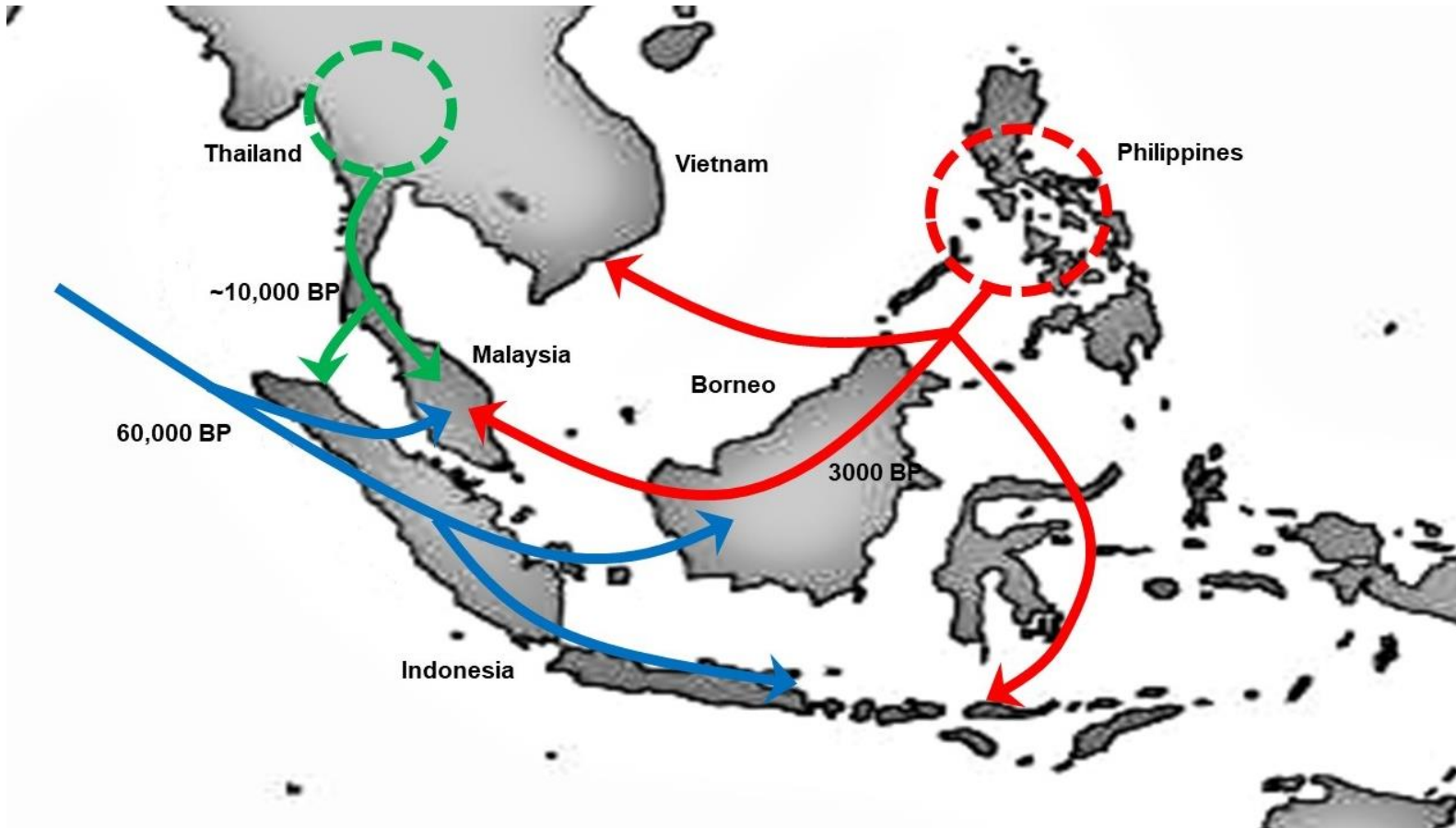


Figure 7.

