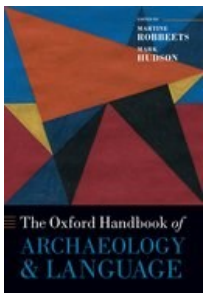

The Oxford Handbook of Archaeology and Language

Martine Robbeets (ed.), Mark Hudson (ed.)

https://doi.org/10.1093/oxfordhb/9780192868350.001.0001

Published online: 22 July 2025 **Published in print:** 26 June 2025 **Online ISBN:** 9780191964299

Print ISBN: 9780192868350

CHAPTER

25 Austronesian archaeolinguistics

Matthew Spriggs, Paul Geraghty, Yue-Chen Liu

https://doi.org/10.1093/oxfordhb/9780192868350.013.25 Pages 466–492

Published: 22 July 2025

Abstract

The Austronesian language family is the second largest in the world in terms of the number of languages, people who speak the languages, and geographic spread. Evidence from linguistics, archaeology, and genetics indicates the Austronesian dispersal began around 5,000 years ago when Taiwan was settled by Neolithic farmers from southeast China. Over the next four thousand years, Austronesian speakers spread to most of the rest of the geographic area where they live today. In this chapter, recent research advances in archaeology, linguistics, and genetics are summarized to provide an overview of the consensus and the controversy on Austronesian history. The current distribution area of the Austronesian speakers can be divided into several regions, including the Philippines, Island and Mainland Southeast Asia, the Pacific Islands, and Madagascar, each discussed in turn. Questions that need further interdisciplinary investigation are pointed to in the concluding section.

Keywords: Austronesian languages, Taiwan, the Philippines, Island Southeast Asia, Mainland Southeast Asia, Pacific Islands, Madagascar, archaeology, genetics, linguistics

Subject: Language Evolution, Linguistics

Series: Oxford Handbooks

Collection: Oxford Handbooks Online

25.1 Introduction

The Austronesian (AN) language family is the second largest in the world, consisting of around 1,250 different languages. There are over 300 million AN speakers. Before the European colonial expansions began, around five centuries ago, the AN language family was the widest in geographical extent, stretching from Madagascar in the west to Rapa Nui (Easter Island) in the east, and Taiwan in the north to Aotearoa/New Zealand in the south (Figure 25.1). The standard work on AN languages and the history of research into them is by Blust (2013; see also 2019a), with some more recent developments in linguistic research covered in this chapter. A summary of the archaeology of the Neolithic settlement of Island Southeast Asia (ISEA) can be found in Bellwood (1997 [2007]) and the later extension of such settlement into the Southwest Pacific is covered most recently in Bedford and Spriggs (2019). At the time of European contact beginning in the sixteenth century, AN speakers were very largely farming populations of Neolithic derivation, although there were some hunter-gatherer groups in ISEA speaking AN languages. In other parts of that region, state-level societies with varying degrees of literacy had developed over the previous millennium, in large part under Hindu-Buddhist influence from South Asia (Bellwood 1997 [2007]).

The wide distribution makes determining the beginning (when) and the geographic cradle (where) of AN languages a very important question. It is a truism that relations between language, culture, and biology are never simple, and we must be careful to separate out the spread of the AN languages, the spread of (arguably) associated cultures, and the biological make-up of the people speaking those languages: there is not a one-to-one correlation (Eisenmann *et al.* 2018). In this chapter, we use AN to refer exclusively to the language family and its distribution in ISEA during the earlier Neolithic and its subsequent eastward extension into the Western Pacific; we use 'Island Southeast Asian (or ISEA) Neolithic' (and related terms) to refer to associated cultures
p. 467 that can be linked—in general—with the initial ↳ spread of AN languages through Island Southeast Asia, and 'Lapita culture' to refer to its eastward extension into the southwestern Pacific Islands, with later cultures and those of Micronesia being named more specifically; and we differentiate several different biological populations that spoke and/or speak AN languages today: southern East Asians, First Remote Oceanians, Negritos, Australo-Melanesians or Papuans, and—in Madagascar and its environs—people of predominantly African descent.

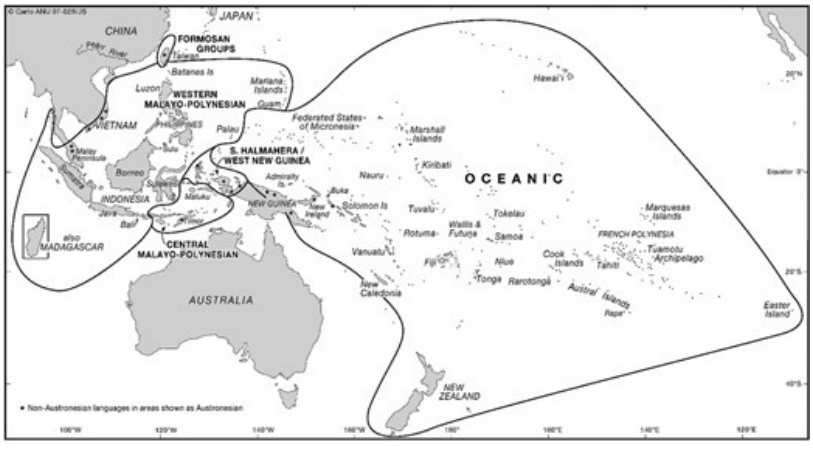

Figure 25.1 The distribution of Austronesian languages and major subgroups

Source: Adapted from Spriggs (2010). Reproduced with the permission of CartoGIS, Scholarly Information Services, The Australian National University.

During their expansion, the original AN speakers of southern East Asian descent assimilated (or were assimilated by) the local populations they encountered, including Negritos and Papuans. These interactions led

to substantial non-East Asian ancestry in most of the AN speakers living today. Thus, to determine genetic admixture patterns of early AN speakers with various groups of distinct genetic ancestry has become a research interest in genetics, complementary to the findings in linguistics and archaeology.

25.2 Disciplinary histories in the Austronesian world

As far back as 1603, Frederick de Houtman had noted the link between Malagasy languages of Madagascar and Malay in Southeast Asia, but the wider extent of the AN language family, albeit not under that name, was first defined by Adrian van Reeland (Hadrianus Relandus) in 1708. He established that the language family included not only Malagasy and ISEA languages but also those in Western Polynesia recorded by the explorers Schouten
p. 468 and ↳ Lemaire in 1616 (de Houtman 1603; van Reeland 1706–1708). Johann Reinhold Forster, a naturalist on Cook's Second Voyage 1772–1775, included all Polynesian languages but not those of Island Melanesia in the family (Forster 1778 [1996]), while von der Gabelentz and Conon (1861–1873) added various languages of Melanesia to the group. The Formosan Indigenous languages of Taiwan were not included in the AN family—in print at least—until the nineteenth century. Meanwhile, the term 'Malayo-Polynesian' had become a popular name for the language family, first formulated by Franz Bopp (1841) as 'malayisch-polynesisch'. In 1906, the Catholic priest and scholar Wilhelm Schmidt coined the term 'Austronesisch' (Austronesian) for the language family as a whole, and by consensus, after the mid-1970s, 'Malayo-Polynesian' is defined as the ancestor of all AN languages outside of Taiwan, with Taiwan generally accepted as the proximate origin of the family, containing nine out of the ten generally accepted primary subgroups (for more details of this history of research and nomenclature see Blust 2013: 20–24 and van Driem 2018: 5–6). The 'centre of gravity' model assumes that the area of greatest diversity of a language family or subgroup is probably its place of origin. There are other methods of homeland detection, however, (see Wichmann, Chapter 6, this volume) and consulting the greatest diversity alone may not be reliable. Recently some linguists have proposed that there is evidence that Proto-Malayo-Polynesian subgroups with East Formosan, implying that it originated from the part of Taiwan where these languages are spoken (V. Chen *et al.* 2022).

Malayo-Polynesian divides into two subgroups, although one of them is really only a residual category (Figure 25.2): Western (WMP) and Central-Eastern (CEMP). WMP comprises the languages of the Philippines, CHamoru —previously Chamorro (Marianas) and Palau in Micronesia, Borneo, Madagascar, Malaysia, Western Indonesia as far as Sulawesi, and the Chamic languages of Mainland Southeast Asia. It is not a well-defined subgroup and Blust (2013: 32) suggested it might be better termed 'Residual Malayo-Polynesian', i.e. all the Malayo-Polynesian languages that are not CEMP. Central-Eastern Malayo-Polynesian is claimed by Blust to be better defined and divides into Central Malayo-Polynesian or CMP (many islands of Eastern Indonesia such as in Maluku/Moluccas) and Eastern Malayo-Polynesian (EMP), although there are growing contrary views on this (Donohue and Grimes 2008; and further discussion in Klamer 2019). CMP is also a residual category rather than a well-defined subgroup. EMP comprises two subgroups: South Halmahera-West New Guinea (SHWNG) and the Oceanic (Oc) subgroup (Blust 2013: 33). These linguistic relationships suggest that, apart from the Marianas and Palau, all Pacific islands where Austronesian languages are spoken were peopled ultimately from the West New Guinea region as they all belong to the Oc subgroup, the sister group to the east of SHWNG and presumably representing an expansion from that area.[1]

p. 469 A link between certain suites of material culture items and the spread of AN languages was first suggested early in the twentieth century and formalized as more archaeological evidence became available and direct ages provided by radiocarbon dating in the mid-1970s (Ellen and Glover 1974; Shutler and Marck 1975; Bellwood 1978; for a more recent example of the genre see Spriggs 2010). Current syntheses of ISEA and Pacific archaeology include Bellwood (2013), Kirch (2017), and Carson (2018). It is notable that no pertinent archaeology has yet been carried out in the putative West New Guinea 'homeland' region of the EMP subgroup.

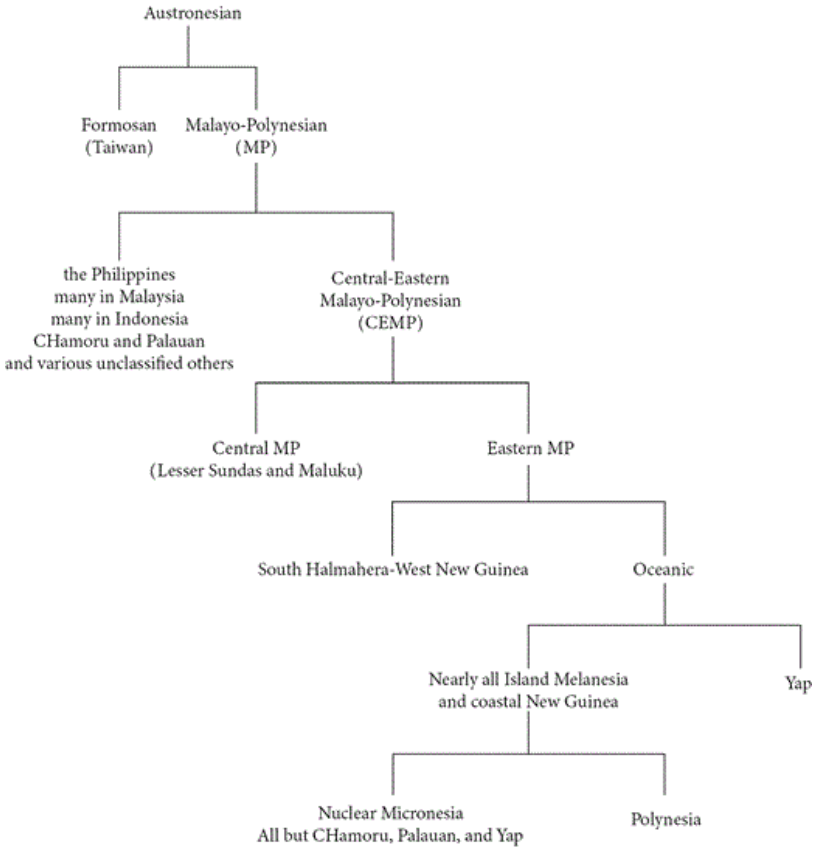

Figure 25.2 Schematic of Austronesian linguistic subgroups

Source: Based on information in Blust (2013) and other literature cited in the text.Image by Yue-Chen Liu.

Physical anthropology using skeletal remains and measurements of living people has also been invoked to track migration patterns (Pietrusewsky and Douglas 2016; Pietrusewsky *et al.* 2016). While indicative of
p. 470 underlying genetic patterns, such studies represent a rather ↳ inexact science compared to current genetic methods, as well as one with a rather sinister past both in relation to its original collection history and often racist assumptions (Douglas and Ballard 2008). This inexactness was also true of the early stages of comparative genetics, focusing initially on parts of the mitochondrial genome (inherited maternally) and later studies conducted with the addition of Y-chromosome data (inherited paternally), but always on samples from living people. Full genome data became available only from the early 2000s but it took some time before these analyses could be reliably extended to ancient DNA from human skeletal remains, starting in 2010 (Reich 2018).

Ancient DNA (aDNA) is DNA from organisms that lived several decades to several hundreds of thousands of years ago, which can be recovered from diverse biological materials (Hofreiter *et al.* 2001). The analysis of ancient DNA from directly dated samples provides a snapshot of genetic variation at a known time and place in the past, which enables the direct evaluation of genetic relationships among past people (or between past and present-day people) and reveals the changes in genetic ancestry through time. The first full genome coverage of aDNA from the AN realm, indeed from anywhere in the Tropics, was published in 2016, based on Lapita-associated skeletal remains from Vanuatu and Tonga in the Pacific (Skoglund *et al.* 2016). The first ISEA aDNA results were published in 2018 (Lipson *et al.* 2018b; McColl *et al.* 2018). Results from Taiwan and Mainland China only became available from 2020 onwards (Yang *et al.* 2020; C-C. Wang *et al.* 2021; Huang *et al.* 2022).

As many studies have shown throughout the world (see recently Gretzinger *et al.* 2022 for Early Medieval England), without ancient samples, it is impossible to assess population change over time through drift and gene flow, processes that have doubtless occurred everywhere. Nor is it possible to assess the pattern of population admixture and to provide anything other than estimates—depending on techniques used which

have been of varying reliability—for when such admixture occurred. Genetic interpretations have see-sawed over the decades since about the 1980s as the discipline rapidly developed, finally reaching maturity and some stability in results only really within the last decade. Thus, earlier studies of contemporary populations in the Pacific and Southeast Asia based only on modern populations and uniparentally inherited markers failed to provide any consensus position from genetics and tended only to confuse historical interpretation. For example, some results from the maternal mitochondrial DNA (mtDNA) data contradicted those from the paternal Y-chromosomal polymorphisms, suggesting either Taiwan or further south in ISEA as the homeland of AN speakers (Melton *et al.* 1995; Sykes *et al.* 1995; Kayser *et al.* 2000, 2008; Su *et al.* 2000; Trejaut *et al.* 2005; Tabbada *et al.* 2010; Ko *et al.* 2014).

Genetic methods, particularly aDNA analyses, have really come into their own in the last decade or so, providing powerful new tools to address questions and issues of anthropological interest that were not available to us previously (Stoneking 2016). In this chapter, we review the most up-to-date discoveries about the population history of AN speakers since we entered the age of genomics and aDNA, and how they have changed our understanding of the origin, the settlement, and the subsequent population dynamics of speakers of this language family. We combine evidence from archaeology, linguistics, and this new, mature phase of genetics, concentrating for the latter on results from aDNA.

p. 471

25.3 Austronesian origins into and out of Taiwan

The origin of the AN language speakers can be considered in two parts: the proximate origins, for which Taiwan is the favoured candidate given its linguistic diversity (Blust 2019a: 418–424) and the ultimate origins, which are more contested, involving, as they do, comparisons with other supposed 'sister' languages as well as postulated 'pre-Austronesian' languages on the Chinese mainland which have left no direct witnesses. The linguistic and archaeological 'out of Taiwan' model was developed particularly by linguist Robert Blust and archaeologist Peter Bellwood in the early 1980s (Bellwood 1984/5; Blust 1984, 1985). Briefly, the 'out of Taiwan' model states that the AN expansion began from Taiwan, somewhat more than a millennium after it was settled by Neolithic rice and millet farmers from southeast China in about 5,500–5,000 BP (BP is years 'before present', defined by convention as 1950 AD). Over the next thousand years, AN speakers spread to most of the rest of the geographic area where they are living today, moving through ISEA around 4,000–3,500 BP (Spriggs 2007), initially bypassing most of New Guinea (except perhaps the western end of its north coast where SHWNG languages are found today) and moving into northern parts of Island Melanesia (the latter two areas designated as 'Near Oceania') around 3,250/3,150 BP (Specht and Gosden 2019). The Lapita culture, identified as the material culture associated with the initial appearance of AN languages in the northern Island Melanesian region, then spread to island groups to the south and east which had never previously been occupied by humans, during the period 3,000–2,800 BP (Petchey *et al.* 2015). This area is designated as 'Remote Oceania' and consists of the Eastern Outer Islands of the Solomons chain, Vanuatu, New Caledonia, Rotuma, Fiji, Micronesia, and Polynesia (see Figure 25.3 for the range of Austronesian dispersal, and Figure 25.4 for distribution of the Lapita culture).

The Lapita culture represents the first pottery-using culture in the South Pacific Islands, including New Guinea. Its elaborate dentate-stamped decorations appear as full-blown in the earliest Lapita sites of the Bismarck Archipelago, just to the east of New Guinea, and have at least generic links to decorated Neolithic pottery in the Philippines and other parts of Southeast Asia. There is also a distinctive non-pottery material culture in early Lapita sites with a range of polished stone adzes and other tools, and shell ornaments of multiple types, utilizing an array of marine shell species. The settlement pattern generally consists of small hamlets of rectangular houses, sometimes on stilts over the reef, and there is evidence of high mobility in the wide distribution of volcanic obsidian flakes and tools originating from sources in New Britain and the Admiralty Islands which are found as far south and east as New Caledonia, Vanuatu, and Fiji and as far west as Sabah in

Borneo. Decorated pottery, too, was moving between island groups, but in quantities that suggest generalized forms of interaction rather than as a trade ware (Spriggs 2021; for recent updates on the Lapita culture see
p. 472 Bedford and Spriggs 2019 and Kirch 2017: 74–106). ↳

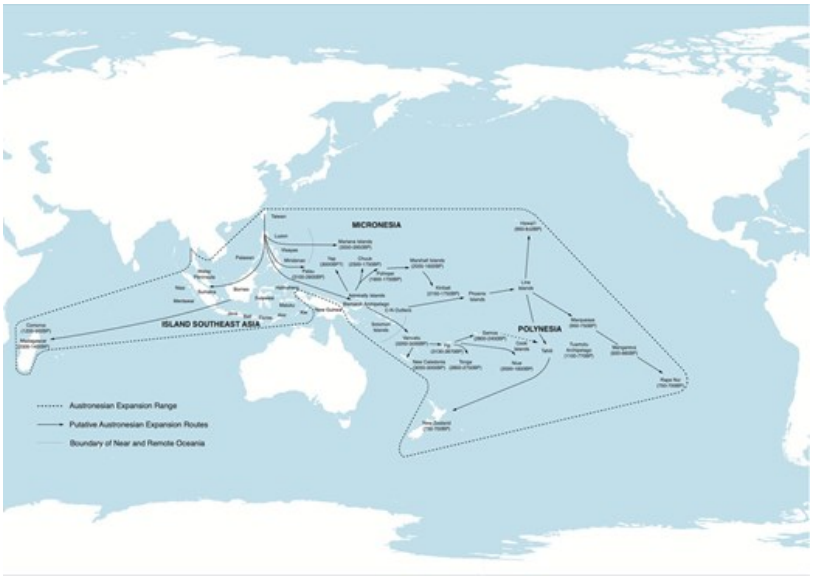

Figure 25.3 The Austronesian dispersal across Southeast Asia, the Pacific and Indian Oceans

The dashed line indicates the boundary of Austronesian expansion. The arrows indicate the putative dispersal routes of Austronesian speakers. Approximate estimates of colonization dates are based on Rieth and Cochrane (2018) and other literature cited in the text. The dashed line from Samoa to Tahiti reflects the pre-Wilson view of settlement from West to East Polynesia. Not indicated are the back-migrations from Western Polynesia to the Polynesian Outliers although the Wilson-proposed dispersal route from the Outliers to Eastern Polynesia is shown. For details see the text. Image by Yue-Chen Liu.

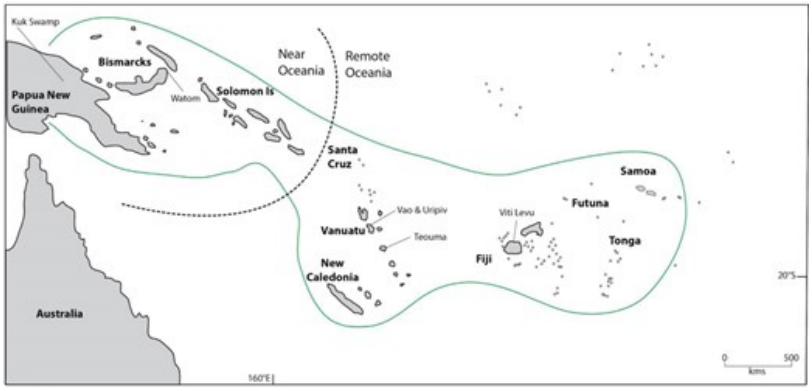

Figure 25.4 Distribution of the Lapita culture in the Western Pacific

Map courtesy of Monica Tromp, used with permission.

p. 473 Settlement of Eastern Polynesia began perhaps about 1,150 BP and was completed within the last thousand years, with the settlement of New Zealand recently dated to 700–675 BP (1250–1275 AD: Bunbury *et al.* 2022). The Micronesian region to the north, also part of 'Remote Oceania', has a more complex history of settlement by Austronesian speakers, which we take up below. The unique situation of Madagascar, initially settled during the first millennium AD, is also discussed below. Computational linguistics has identified two relatively long pauses that occurred during the Austronesian expansion (Gray *et al.* 2009): one was between the settlement of Taiwan and of the northern Philippines, while the other was between the settlement of the Southwest Pacific (as far as Fiji and the western part of Polynesia) and of Eastern Polynesia, the rest of the Polynesian Triangle.

These pauses echo the archaeological pattern of dating of Neolithic settlement for the same areas, giving confidence to the reality of these linguistically derived pauses.

If we go back to the ultimate origins of AN on the Chinese mainland, we enter a much more uncertain and speculative area, with any AN-related languages on the coast opposite Taiwan having been absorbed by the spread of dominant Sinitic languages associated with the expansion of the Han dynasty south of the Yangzi River from 206 BC to 220 AD (Blust 2013: 424). There is as yet no linguistic consensus as to the possible relations of Austronesian languages to other language families. Blust (2013: 702–721) considers all such models of relationship up to that time and only gives any credence to the Austric hypothesis, originally formulated by Wilhelm Schmidt (1906) as relating the Austronesian and Austroasiatic (both Austric and Austroasiatic being further terms coined by Schmidt) language families, and the Austro-Tai hypothesis linking AN and the Tai-Kadai (or Tai or Daic or Kra-Dai languages), first formulated independently by Benedict (1942) and Wulff (1942).

Austroasiatic languages span a region from central India to Vietnam, including the Mon-Khmer languages of Cambodia, Laos, and Vietnam, the Aslian languages of the Malay Peninsula, Palaung-Wa and Riang of the Myanmar–China border region, the languages of the Nicobar Islands off the coast of Thailand but politically part of India, and the Munda languages as well as Khasi of Assam in India (Blust 2013: 703). Blust considers that morphological agreements between AN and Austroasiatic languages provide 'possible evidence for a distant genetic relationship' (2013: 710). Bellwood (2005) placed the origin of Austroasiatic to the immediate south of the original Tai-Kadai area, that is in Guangxi and northern Vietnam.

The Tai-Kadai languages are spread over an area from southern Thailand through Laos, Vietnam, and into southern China, with Thai for instance being the dominant language in Thailand. They are thought to have arisen in the Guangdong region of southern China (Blench 2018; Guo *et al.*, Chapter 24, this volume). The relationship between Tai-Kadai and Austronesian is seen by some linguists as specifically genetic (i.e. common descent), although Blust notes that it is a point of contention whether the connection is 'due to contact or divergent descent' (2013: 717). Blench (2018) sees the Tai-Kadai languages as having a 'multi-genetic' origin, involving 'pre-Austronesian' languages of the Chinese mainland in the Guangdong area as well as continuing cross-Taiwan Strait interactions between the Formosan languages and an early form of Tai-Kadai, with further connections of at least some Tai-Kadai languages with later Proto-Malayo-Polynesian languages in the Philippines area.

An even more ambitious idea was put forward by Starosta (2005, but submitted just before his death in 2002),
p. 474 who postulated a Proto-East Asian language splitting into ↳ 'Sino-Tibetan-Yangzian' (ST-Yangzian) and 'Proto-Austronesian', with the latter at a later extra-Formosan stage splitting into Tai-Kadai and Malayo-Polynesian. 'ST-Yangzian' split into 'Pre-Sino-Tibetan' and 'Proto-Yangzian' with the latter splitting into Proto-Hmong-Mien and Proto-Austroasiatic. The justification is to explain the spread of Central Chinese millet farming, as practised by his 'Proto-East-Asian' speakers, down into the Yangzi area where rice was domesticated and formed the spur to further migrations of groups that became the AN, Hmong-Mien, Austroasiatic, and (derived from AN in this view, following Sagart (2005b) the Tai-Kadai speakers of southern China and Southeast Asia. Blust (2013) does not comment on this hypothesis, presumably as very little linguistic evidence was offered in support by Starosta.

Early genome-wide surveys of AN speakers can be traced back to the beginning of the 2010s (Wollstein *et al.* 2010; Xu *et al.* 2012; Lipson *et al.* 2014), when modern genome-wide single nucleotide polymorphism (SNP) array data were obtained from Southeast Asia and the Southwest Pacific and pointed to a genetic origin of the AN speakers in Taiwan. A recent genome-wide study on a larger scale (Choin *et al.* 2021), collecting 317 high-coverage genomes from twenty Oceanian populations, also supports the 'out-of-Taiwan' model, inferring that the East Asian ancestors of the Pacific Islanders may have diverged from Taiwanese Indigenous peoples around 5,000 BP. As with many genetically generated dates, this is probably too old—by about 1,000 years in this case—as the archaeological evidence of Neolithic expansion from Taiwan into the Philippines and the rest of ISEA

is now well dated as starting around 4,200/4,000 BP (Hung *et al.* 2022). Given that there is clear evidence of Neolithic cultures in Taiwan with pottery and rice and millet cultivation from about 5,500/5,000 BP, the primacy of Taiwan in terms of ISEA Neolithic origins is clear (Bellwood 1997 [2007]).

Ancient DNA studies also support the expansion of the AN speakers starting from Taiwan. As mentioned earlier, the first genome-wide aDNA study of the Pacific region was published in 2016 (Skoglund *et al.* 2016), with the authors reporting data from three individuals from Vanuatu about 3,000–2,800 BP and one individual from Tonga about 2,700–2,300 BP, all of which were associated with the Lapita culture. More individuals with genome-wide ancient DNA data from Oceania were reported subsequently (Lipson *et al.* 2018b, 2020; Posth *et al.* 2018b; Pugach *et al.* 2021; Liu *et al.* 2022), covering more diverse geographical, temporal, and cultural contexts. Yet the consistent genetic model indicated by these data shows that the First Remote Oceanians (FROs) were of East Asian-related ancestry and related distantly to southeastern Chinese (Yang *et al.* 2020), more closely related to Indigenous people in Taiwan (C.-C. Wang *et al.* 2021), and most closely related to the ancestors of present-day north-central Philippine groups such as the Kankanaey, corresponding to models of population movement along a route from Taiwan to the Philippines to Near Oceania to Remote Oceania derived from linguistics and archaeology.

Recently published aDNA data from northern Maluku and Sulawesi, dating up to ~2,200 BP, are also consistent with being derived from a Taiwan-associated expansion mixed with local Indigenous (Negrito-like) ISEA-associated ancestry (Oliveira *et al.* 2022). Conversely, a ~7,800 BP individual from South Sulawesi had no evidence of ancestry related to early Austronesian speakers (Carlhoff *et al.* 2021). These observations thus provide additional evidence in favour of the 'out-of-Taiwan' model for the spread of Neolithic agricultural populations through ISEA and into the Pacific.

p. 475 Further aDNA data from Neolithic samples from the southern Chinese Mainland and Taiwan Strait islands support an origin on the coast of southern China for Proto-Austronesian speakers (Yang *et al.* 2020; C-C. Wang *et al.* 2021; Huang *et al.* 2022) by showing clear connections with both ancient and present-day individuals with Southern East Asian-related ancestry, like the Lapita-associated Vanuatu and Tongan FRO individuals and present-day Pacific Islanders. This close genetic link is also supported by mitochondrial and Y-chromosomal DNA (Wei *et al.* 2017; Kutanan *et al.* 2018). Mitochondrial and Y-chromosomal lineages found in ancient individuals from Taiwan are shared by modern Indigenous Taiwanese peoples, as well as the Lapita-associated individuals from Vanuatu and Tonga (Skoglund *et al.* 2016; Lipson *et al.* 2018b, 2020; Posth *et al.* 2018b). The 8,060–8,320-year-old Liangdao individual from an island just off the Chinese mainland but under Taiwan administration also carried the mitochondrial haplogroup that is common in AN-speaking populations in Taiwan, the Philippines, and Indonesia today, and is most similar to those found in the Aboriginal Formosans from Taiwan (Ko *et al.* 2014). Moreover, ancient and modern AN-speaking Indigenous Taiwanese share significantly more alleles with the Tai-Kadai speakers of the southern Chinese Mainland and Hainan Island, consistent with the hypothesis that ancient 'pre-Austronesian' populations related to present-day Tai-Kadai speakers introduced agriculture to Taiwan (Bellwood 2011). This supports the proposed connection between Neolithic southern East Asians and early AN-speaking groups.

A competing hypothesis to 'Out of Taiwan' is 'Out of Sundaland' or 'Nusantao' which proposes that the homelands of AN speakers were within ISEA, particularly on the Pleistocene Sundaland landmass which was drowned by rising sea levels during the end of the last glacial period, and claims a south-to-north migration of the Austronesian speakers (Solheim 1984; Solheim *et al.* 2006). Linguistic evidence offers no support at all to this hypothesis.

A recent genomic-wide survey of 1,028 modern individuals representing 115 Indigenous population groups living in the Philippines offers to blend rather than to decide between these two models (Larena *et al.* 2021), as it purports to show that the ancestors of Cordilleran people, who eventually became, more or less, identified with those we have called Austronesian speakers, diverged from Indigenous peoples of Taiwan at least ~8,000

BP, prior to the arrival of rice agriculture in the Philippines with the advent of the Neolithic around 4,000 BP. Some descendants of the ancient Cordilleran people remain the least admixed East Asian groups carrying an ancestry shared by most AN-speaking populations, thereby challenging an exclusive 'out-of-Taiwan' model. Again, the problem may simply be the form of genetic dating used, which as we have seen has a tendency to produce 'dates' too old compared to the archaeological evidence; as yet we have no ancient genomes from the Philippines which can cast definitive light on these hypotheses. Other early linguistically based hypotheses such as Dyen's (1962) hypothesis of a Melanesian homeland for AN have no support today from linguists (cf. Anceaux 1965).

When we talk about AN language history, it always involves other major contributors, such as Australo-Melanesians, Austroasiatic-speaking people, etc., who substantially admixed and interacted with the early AN speakers on a large scale in time and space. Therefore, the story of AN language speakers would not be complete without properly considering the populations originally speaking other languages who interacted with them in the distant past of initial AN spread or during more recent times. In the next part of this ↳ chapter, we provide summaries of the population history of six major areas of AN spread and discuss the prehistoric and recent interactions between AN speakers and other human populations (if any) living there: the Philippines, the rest of ISEA, Mainland Southeast Asia, the South Pacific, Micronesia in the northern Pacific, and Madagascar at the western edge of the Indian Ocean. Given that the Pacific Ocean covers a third of the planet and all three continents would fit within it, we can be excused for giving the Pacific Islands an unequal share of the coverage.

p. 476

25.4 The Philippines

The Philippines is situated at the crossroads of human migrations in the Asia-Pacific region. For millennia, the archipelago has served as a corridor for migration from one continent to another and a link between Southeast Asia, Australia, and the Pacific islands. The oldest modern human remains are from the Tabon Caves of Palawan, dated to around 47,000 BP (Détroit *et al.* 2004). The ancestors of present-day Negrito groups are widely regarded as the first modern human inhabitants of the Philippines (Reid 2013; Jinam *et al.* 2017) and are believed to be descendants of the first human migrations out of Africa. AN speakers reached the Philippines at around 4,200–4,000 BP, and then rapidly spread across the rest of ISEA (Blust 2013, 2019). This population of southern East Asian ancestry mixed with the existing Negritos, resulting in modern Filipino ethnic groups which display various ratios of genetic admixture between East Asian-like Austronesian ancestry and Australo-Melanesian-like Negrito ancestry (Lipson *et al.* 2014). The influence of AN expansion is also evidenced by the fact that all present-day Indigenous Philippine languages, including the languages spoken by Negritos, belong to the AN language family (Blust 2013). There has been significant language-levelling within Philippine languages at least twice in their history, meaning that it is difficult to determine the initial patterns and processes of AN spread there using linguistic methods (Blust 2005a), and whether Proto-Philippines ever really existed is still fiercely debated (Blust 2019a; Liao 2020; Reid 2020; Ross 2020; Zorc 2020). Recent Bayesian phylogenetic analysis of Philippine languages indicates a rapid initial AN expansion followed by northward back-migration (King *et al.* 2021). Genetic signatures also indicate the possibility of more recent migrations of Mainland SEA, Papuan, and South Asian people into the Philippines.

The genomic study by Larena *et al.* (2021) reveals the complexity of the origins, kinship patterns, and genetic diversity of the Philippines, arguing that over the millennia, at least five major streams of migration built up the population: Negrito, Manobo, Sama, Papuan, and Cordillerans (the Indigenous peoples of the Cordillera Mountain Range of northern Luzon, often referred to as Igorots, who have southern East Asian ancestry). Rather than the introduction of farming, Larena and colleagues suggest that climate change, with subsequent sea-level fluctuation, may have played a more important role in driving the mass movement of populations in various directions.

Although Larena *et al.* (2021) provide a new model of the evolutionary history of the Philippines, the dates of population divergence estimated in the study using genomic data from modern people will require support from archaeological evidence. Further work on the dates from archaeological assemblages that can be associated with particular population movements and genomic data from ancient humans from the same sites would tell us ↳ who really lived there and when the split of different lineages happened. Although the model proposed by Larena *et al.* (2021) is complex, there are still some unsettled questions left; for example, the relationship between Philippine groups and ancestral populations in mainland Asia, the relationship between the Cordillerans and the supposedly later arriving AN-speakers, and so on. One of the main values of this study is that it underlines the need for aDNA from directly relevant time periods and regions.

p. 477

25.5 The rest of Island Southeast Asia (ISEA) in relation to Austronesian expansion

Once south of the Philippines, the linguistic pattern becomes if anything more complex. It is well summarized by Blust (2013: 757, references removed):

> Further south the migration history becomes cloudier: southward expanding AN speakers split into at least two major streams, one leading into western Indonesia and the other into eastern Indonesia and the Pacific. However, it is possible that migration streams from the Philippines split three ways in their southward progress, with one leading into Borneo, and then to Sumatra, mainland Southeast Asia and Madagascar, a second into Sulawesi, and the third into eastern Indonesia and the Pacific. What has become increasingly clear in recent years is that the attested distribution of AN languages is due to a complex interaction of primary settlement and differentiation *in situ*, and subsequent language leveling not just in the Philippines, but in a number of areas. Long after the expansion of Proto Philippines precipitated a major episode of language leveling, the expansion of Proto Greater Central Philippines led to another reduction of linguistic diversity throughout most of the central Philippines, and in parts of northern Sulawesi. The recent establishment of a 'Celebic' supergroup that excludes the South Sulawesi languages… leaves the latter group as a clue to probable earlier diversity that was leveled by the expansion of Proto Celebic. Much the same is true of southern Sumatra, where a low level of diversity is found in precisely those areas that would have been settled earliest in almost any plausible scenario of AN migration… Central and Central-Eastern Malayo-Polynesian almost certainly split from one another in the northern Moluccas, and the early diversity of AN languages in this area evidently was reduced by the expansion of South Halmahera-West New Guinea speakers out of Cenderawasih Bay in Irian [now West Papua] into southern Halmahera and the smaller islands between Halmahera and New Guinea.

Genome-wide analyses of AN populations across ISEA reveal four ancestral source components that have contributed to modern human populations (Lipson *et al.* 2014). These source populations are most closely related to contemporary Aboriginal Taiwanese people, Negritos, Papuans, and H'tin (an Austroasiatic-speaking population of Laos and Cambodia in Mainland Southeast Asia). All the present-day AN-speaking populations in ISEA were found to be mixtures of two or three components, one of which is always southern East Asian-related, while the other one or two are from various sources. There seems to be a boundary between western and eastern parts of ISEA, as the second-largest ancestry component is Negrito and/or Papuan in the east (i.e. Philippines, Sulawesi, Lesser Sunda Islands, and Maluku) and 'Austroasiatic' in the west (Sumatra, Java, and Borneo).

The most striking finding of study of the genetic ancestry of ISEA is that it recognizes the genetic contribution of speakers of Austroasiatic languages, especially in Western Indonesia ↳ and Borneo. Today, these languages are confined to Mainland Southeast Asia (MSEA) and mainly spoken in small pockets, except for modern

p. 478

Vietnamese and Cambodian, surrounded by speakers of Tai-Kadai and Sino-Tibetan languages. This suggests that the first farmers who spread south into ISEA from China's agricultural homelands might have spoken Austroasiatic languages, which then became replaced by AN languages carried by subsequent farmer expansions. Current genetic evidence suggests two main possibilities: one is that AN speakers spreading south encountered Austroasiatic farmers already established in Western Indonesia and Borneo and replaced them linguistically; the other is that Austroasiatic farmers spread into Western Indonesia after Austronesian farmers arrived there but at some stage became absorbed linguistically.

The most likely scenario on archaeological grounds, albeit not yet closely dated, would be an initial Neolithic spread by Austroasiatic speakers from MSEA. Early pottery in Sumatra and the western half of Java, as well as in Western Borneo, is a cord-marked or paddle-impressed ware similar to that common in Early Neolithic communities in MSEA such as Kok Phanom Di in Thailand and dating there from 4,190–3,910 BP onwards (Higham and Higham 2022) as well as in Neolithic sites in peninsular Malaysia (Simanjuntak 2017). Adelaar (1995) suggested that there was an Austroasiatic substratum in Land Dayak languages of Western Borneo, and further linguistic discussion of this idea can be found in Blench (2010).

25.6 Mainland Southeast Asia (MSEA) and the later spread of Austronesian languages

Anatomically modern humans reached MSEA at least 65,000 BP (Higham 2013; Oxenham and Buckley 2016; Bae *et al.* 2017), probably the direct ancestors of the Hòabìnhian hunter-gatherers of the early Holocene (Higham 2014). Recent hunter-gatherers in MSEA are considered to be derived from the Hòabìnhian tradition (Aghakhanian *et al.* 2015). Major cultural changes in human society were propelled by the development and expansion of agriculture, and indeed, present-day MSEA populations were substantially influenced by later migrations involving rice and millet farmers from the north (Higham 2014). Rice farming began ~4,500–4,000 BP in the area, accompanied by a relatively uniform and widespread suite of pottery and other material culture, displaying connections to southern China (Higham 2002; Diamond and Bellwood 2003; Chi and Hung 2010; Oxenham *et al.* 2011). It has been hypothesized that this cultural transition was caused by the migration of Austroasiatic-speaking southern East Asians rather than the Indigenous hunter-gatherers of Southeast Asia (Matsumura *et al.* 2008; Oxenham *et al.* 2011; Chi and Hung 2012; Matsumura and Oxenham 2014). People who speak Austroasiatic languages today have a wide but fragmented distribution in the region, the fragmentation reflecting later population movements (Higham 1998; Peiros and Shnirelman 1998; Bellwood *et al.* 2011; Sidwell and Blench 2011).

There was a competing hypothesis on agricultural development that the Hòabìnhian hunter-gatherers adopted agriculture without substantial external gene flow (Hanihara 2006), but there are now many genome-wide present-day and aDNA studies across MSEA (Lipson *et al.* 2018b; McColl *et al.* 2018) that confirm traces of Austroasiatic ancestry in the region. They further suggest that the arrival of AN speakers in Mainland Southeast Asia was a recent event, and the language transition came from recent cultural shifts. For example,
p. 479 the Chamic ↳ language speakers (an AN subgroup), descendants of the people of the Champa kingdoms of central Vietnam, have been shown to be genetically predominantly local to MSEA but with a component of male ancestry derived from ISEA AN-speaking populations, as well as additional South Asian male genetic input presumably linked to trading and religious contacts—the early religion of Champa was a form of Hinduism (Peng *et al.* 2010; He *et al.* 2012).

There were clearly links from perhaps 3,000 BP onwards between MSEA populations and ISEA Austronesian-speaking communities, archaeologically most often associated with the Sa Huynh culture of central Vietnam (2,500–1,850 BP) which immediately preceded the establishment of Champa (Hung *et al.* 2013; Blench 2022). Specific links between the Kalanay culture of the central Philippines and Sa Huynh going back to Solheim

(1957) have recently been disputed and substituted by more convincing material culture links between Kalanay and slightly later MSEA sites such as Hoa Diem in Cam Ranh Bay in central Vietnam and southern peninsular Thailand sites of the early centuries AD (Yamagata and Matsumura 2017). At what point language shift from presumably Austroasiatic to AN languages occurred in central Vietnam is unclear with some estimates that it occurred only about 2,000 BP before substantial loans from Sanskrit entered Southeast Asian languages (Hung *et al.* 2013: 399). The spread of Chamic and Malay languages in parts of MSEA could represent a form of elite dominance with speakers switching from other languages to these apparently more prestigious trade- and religious-associated languages. As to origins within the AN sphere, Blust concludes: 'A single population movement out of southwest Borneo evidently led to the establishment both of Malayic languages in Sumatra and the Malay Peninsula and of Chamic languages in Vietnam and neighbouring areas' (2019: 426).

25.7 The South Pacific region

Modern humans arrived in Near Oceania at least 47,000 BP and possibly as far back as 65,000 BP, and spread in subsequent millennia through Australia, New Guinea, the Bismarck Archipelago, and the Solomon Islands (Wickler and Spriggs 1988; O'Connell and Allen 2015; Allen and O'Connell 2020). They would have been the ancestors of current Non-Austronesian or 'Papuan' speakers (see Schapper *et al.*, Chapter 26, this volume for coverage of Papuan languages) who represent the majority of the population on the island of New Guinea (Foley 1986; Pawley *et al.* 2005) and maintain a more scattered presence on islands to the immediate west and east of it, probably to the west representing a back-migration from the New Guinea mainland to as far west as the islands of Timor, Alor, and Pantar (about two dozen languages) and on Halmahera and adjacent islands (about one dozen languages) further north in Eastern Indonesia (Schapper 2018 and Schapper *et al.*, Chapter 26, this volume). An extinct language on Sumbawa in Nusa Tengara Timur province of Indonesia may well have been non-Austronesian (Donohue 2007). The suggestion is that the current non-Austronesian languages of at least the Timor-Alor-Pantar group represent a westward migration from New Guinea after the independent development of agriculture there by 6,000 BP, and presumably prior to AN expansion into the area (Pawley 2005). To the east of New Guinea in the Bismarck Archipelago and Solomon Islands, non-Austronesian languages have a pattern that is taken to be a relic of their former wider distribution down to the eastern end of the main Solomons chain (Dunn 2009).

p. 480 After 3,100 BP, humans expanded into Remote Oceania, apparently previously unoccupied. In the southwest Pacific beyond the main Solomon Islands, the first evidence of human presence comes from artefacts of the Lapita archaeological complex (Figure 25.5), which appeared in the Bismarck Archipelago by ~3,350–3,150 BP, reached the Remote Oceanic islands of the Reefs-Santa Cruz Group, Vanuatu, New Caledonia, and Fiji by ~3,000 BP, and spread into geographical Polynesia thereafter (including Tonga by ~2,850 BP and Samoa by ~2,750 BP) (Bedford and Spriggs 2019; Rieth and Athens 2019). As alluded to earlier, aDNA from eleven individuals from Vanuatu and Tonga 3,000–2,500 BP indicates that these pioneers were related distantly to Neolithic southern East Asians (Yang *et al.* 2020), more closely related to Neolithic and Iron Age people of Taiwan (C. C. Wang *et al.* 2021), and most closely related to the ancestors of present-day north-central Philippine groups such as Kankanaey (Skoglund *et al.* 2016; Lipson *et al.* 2018b, 2020; Posth *et al.* 2018b). Using the latest terminology (Liu *et al.* 2022) they are designated First Remote Oceanians—Southwest Pacific ($FRO_{SouthwestPacific}$).

These patterns are consistent with a scenario in which groups from ISEA entered uninhabited Remote Oceania, probably from a staging point along the northern coast of Western New Guinea, bypassing the eastern part of mainland New Guinea and through other Near Oceanian islands without significant genetic admixture with long-established local groups (at least initially). The genetic pattern reported earlier of coastal admixture and the spread of Western Oceanic AN languages through the Bismarcks and parts of the Solomons, with related spreads westwards along both the north and south coasts of New Guinea, represent slightly later expansions of groups with substantial FRO inheritance (Ross 1998).

Figure 25.5 Decorated dentate-stamped Lapita pottery, a unique (for the Pacific) decorative technique and design system

Photograph by Matthew Spriggs.

p. 481 Such late Lapita or immediately post-Lapita period population spreads are supported by dated archaeological assemblages of Lapita affiliation in these areas (David *et al.* 2019; Lilley 1999, 2000; Walter and Sheppard 2017). Clearly such dispersals involved interaction with Papuan speakers already in place, and today some New Guinea AN languages show considerable influence from Papuan languages, and vice versa. Blust (2013: 700–701) describes languages such as Maisin and Magori and its related languages of southeast New Guinea as 'Papuanized' AN languages and Mailu in the same region as an 'Austronesianized' Papuan language. Language levelling may have also taken place in parts of the Western Oceanic area. In the Massim (islands immediately off the eastern tip of New Guinea) are recently discovered assemblages on Brooker Island contemporary with the earliest phase of Lapita and showing interaction of already-resident Papuan populations with presumed Lapita communities nearby, a pattern seen also on Nissan Island between New Ireland and Bougainville in the northern Solomons (Shaw *et al.* 2022). Current languages in both areas are Western Oceanic AN and, thus, presumably represent a somewhat later AN spread and language levelling. The occasional finds of dentate-stamped sherds along the north coast of Eastern New Guinea may represent, at least in part, early Lapita forays along this coast for which we have no surviving linguistic correlates.

There are no people with only Lapita-related FRO ancestry anywhere in southern Remote Oceania today, since later migrations resulted in the arrival of Papuan ancestry from Near Oceania, more specifically the Bismarck Archipelago. The earliest Lapita migrants to Vanuatu showed only 3% Papuan admixture at most, while most Ni-Vanuatu today range from 80 to 88% Papuan contribution. By ~2,400 BP, in some individuals in central and southern Vanuatu where the Lapita culture had ended by 2,800 BP, the proportion of Papuan ancestry most closely linked to New Britain in the Bismarck Archipelago, designated $PAPUAN_{NewBritain}$ (Liu *et al.* 2022), was 96% and 99% respectively. In the small northeastern islands off Malakula, a recognizably Lapita culture possibly continued for a further 400 years or so to about 2,400 BP. There, Late Lapita and immediately Post-Lapita individuals present only 54 to 78% $PAPUAN_{NewBritain}$ genetic inheritance suggesting some spatial unevenness in the immediate impact of these New Britain migrants (figures extracted from Lipson *et al.* 2018b and Posth *et al.* 2018b). Papuan ancestry from this secondary expansion eventually spread further east into Remote Oceania—Fiji and thereafter the whole of Polynesia—at substantial levels (minimum of ~25% $PAPUAN_{NewBritain}$) (Skoglund *et al.* 2016; Lipson *et al.* 2020).

Many linguists, including Blust (2005b, 2008), have pointed to features of languages in some areas of Remote Oceania—including Reefs-Santa Cruz in the Solomons, Santo, and Malakula in northern Vanuatu, southern Vanuatu, and New Caledonia—that suggest former contact with non-Austronesian languages, as summarized in Geraghty (2017).

The final pre-European population expansion into Island Melanesia occurred from about 1,150 BP, with the back-migration of Polynesian communities of Western Polynesian origin who succeeded in linguistically taking over a series of small islands in the Autonomous Region of Bougainville, the Solomons, Vanuatu, and the Loyalty Islands of New Caledonia and two—Kapingamarangi and Nukuoro—in geographical Micronesia that are today part of the Federated States of Micronesia (Spriggs 1997: 187–222; Feinberg and Scaglion 2012). Evidence for this takeover, in some cases accomplished through genocide or ethnic expulsion of other groups —comes from the presence of Polynesian languages on these islands in Island Melanesia and Micronesia which derive from the Samoan area (Figure 25.6), oral traditions of these communities having come from particular
p. 482 Polynesian islands and, at times more ambiguously, from the archaeological record; the latter shows a ↳ range of expressions from very Polynesian-like material culture in the Central-Northern Outliers of the Solomons, which may or may not have been previously inhabited, to a blend of Polynesian and pre-existing Island Melanesian assemblages in the Outliers in Vanuatu and the Eastern Outer Islands of the Solomons which show a 3,000-year occupation history going back to Lapita (Flexner *et al.* 2019; Zinger *et al.* 2020; see also Kirch and Swift 2017 and Kirch and Yen 1982 for Tikopia and Leach and Davidson 2008 for Taumako, both in the Eastern Outer Islands of the Solomons).

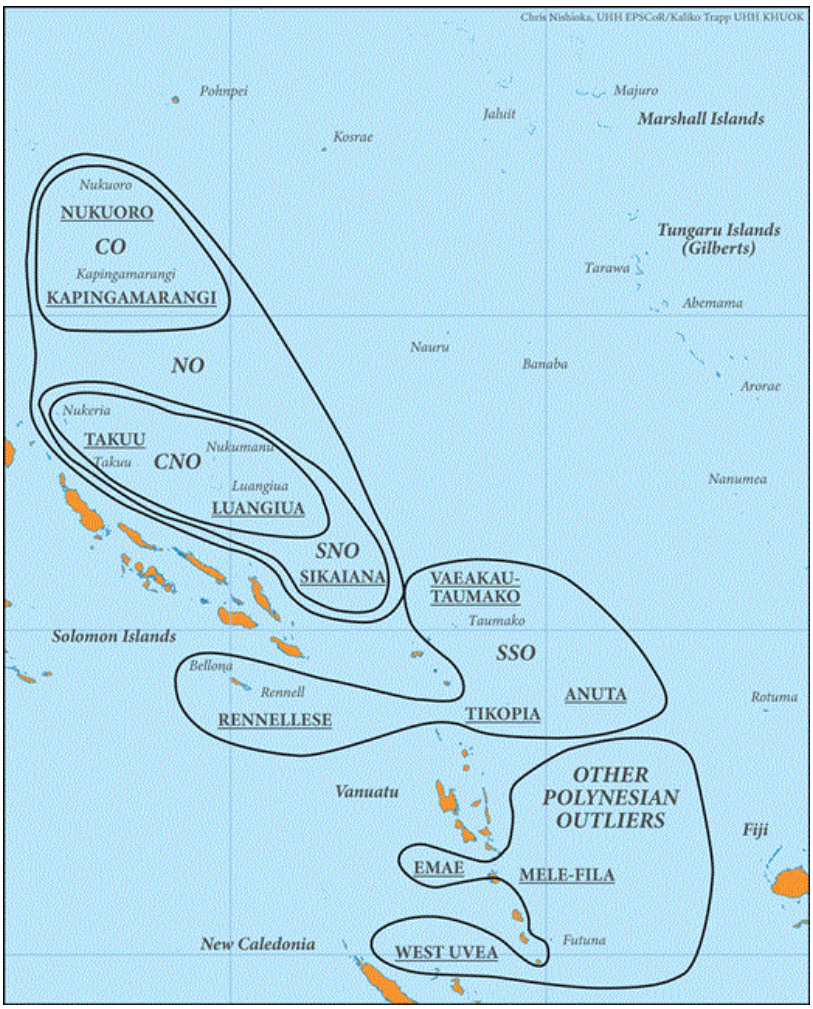

Figure 25.6 Distribution of the Polynesian Outlier islands in Melanesia and Micronesia

Source: Map from Wilson (2018), used with permission.

p. 483 Polynesian genetic and linguistic presence extends beyond the individual Outliers in central and southern Vanuatu, affecting much of the population of central Vanuatu and Aneityum and Tanna to the south, with extensive Polynesian loans into these languages (Clark 1994; Lynch and Fakamuria 1994; Lynch 1996). Some of the people buried with Chief Roi Mata in a mass grave on Eretok islet off the northeast coast of Efate in Vanuatu in ca. 400 BP displayed a major Polynesian genomic component, which went along with the Polynesian-influenced material culture accompanying the burials (for the genomics, see Lipson *et al.* 2020; for the archaeology see Garanger 1972).

Modern Hawaii, French Polynesia, Cook Islands, Rapa Nui, New Zealand, and a few other isolated islands, many of which had been abandoned by European contact (most of the Phoenix and Line Islands, Henderson and Pitcairn, for example) comprise Eastern Polynesia. There was a major pause in AN settlement in Western Polynesia post-Lapita before these more eastern areas were settled, reflected in the many innovations in Proto-Polynesian that developed in Tonga and Samoa, as they progressively separated from Fijian languages and dialects of the ancestral language stage Proto-Central Pacific (Geraghty 1983, 2021). After the major pause before the development of Proto-Malayo-Polynesian, ca. 5,500–4,000 BP, and the move out of Taiwan, the other major AN expansion pause was between Proto-Polynesian and the development of Eastern Polynesian languages, ca. 2,800–1,200 BP. The conventional view is that around 1,150 BP, groups of migrants from the area where Samoan and closely related languages were spoken, presumably aided by developments in sailing technology, were able to move against the prevailing winds to discover the islands of Eastern Polynesia, Tahiti

being often mentioned as the initial landfall on grounds of oral traditions (Kirch 2017). This was the view from the initial subgrouping of Polynesian languages (Green 1966; Pawley 1966; restated in Walworth 2014) until the last decade. Linguist William 'Pila' Wilson started to challenge this model in 1985, and his linguistic challenge has really picked up in the last decade (Wilson 1985, 2012, 2014, 2018, 2021, 2022; see also Thomson *et al.* 2020).

In Wilson's view, the settlement of Eastern Polynesia did not proceed directly from the Samoan area to Eastern Polynesia but instead proceeded initially to the west to the Polynesian Outlier islands offshore from the Central and Northern Solomon Islands. It was from here that Polynesian settlers then made long-distance forays to the east, via the Phoenix and Line Islands, to the Marquesas Islands. Independent of Wilson's ideas, voyaging simulations were also pointing out just how difficult it was given the sailing abilities of the time to travel directly east from the Samoan area. Settlement could be made far to the east, particularly during El Niño years, much more easily along a west–east 'corridor' to the north, with favourable winds and currents (Montenegro *et al.* 2014, 2016).

This 'corridor' begins in exactly the area of the Central-Northern Outliers (Ontong Java, Nukumanu, Takuu, and Nuguria as well as the ex-Outlier of the Carteret Group). Wilson established a Proto-Central Northern Outlier-East Polynesian subgroup (PCNO-EPn), with over 200 linguistic innovations uniquely shared between the C-N Outliers and East Polynesian languages, and not found in Samoan or other closely related non-Outlier languages (Wilson 2018). A connection between the Solomon Islands and islands way to the east, particularly the Marquesas, was posited long ago on the basis of material culture similarities (Balfour 1917; von den Steinen 1925–1928; Reichard 1933; Suggs 1961) but there was no context in which to place it; Wilson has now provided such a context and lists several additional similarities in material culture uniquely shared between the C-N
p. 484 Outliers and ↳ Eastern Polynesia, such as whalebone clubs (called *paraamoa* in Takuu and Nukumanu in the Outliers and *paraaoa* in New Zealand Maori), stone or wooden food pounders, *Ruvettus* fishhooks and upturned canoe ends, as well as details of some spiritual beliefs (Wilson 2018: 414–416). More recently, it has been argued that most, if not all, the varieties of *fēʔī*, the upright-fruiting banana found in profusion in Tahiti and the Marquesas, originated not from Samoa or Western Polynesia, but from the Solomon Islands (Thomson *et al.* 2022).

Linguists have been slow to react to Wilson's ideas, although Blust (2019a: 429) recently accepted them, and archaeologists have been even slower to take them on board for testing against archaeological evidence (the exception being Sheppard 2022). In part this is because the archaeology of the intermediate areas, the C-N Outliers of the Solomons, and the Phoenix and Line Islands (mostly administered by Kiribati but with some islands under United States control) is very little known except for the efforts in the latter archipelagos of di Piazza and Pearthree (2001; Pearthree and di Piazza 2003; di Piazza and Pearthree 2004) and Anderson *et al.* (2002). Much more directed archaeological consideration needs to be given to these remote archipelagoes as they are crucial to archaeological consideration of Wilson's arguments.

Wilson's revised subgrouping of East Polynesian languages postulates that Proto-East Polynesian (EPn) developed in an area bounded by the Phoenix and Line Islands and the Marquesas and had two main branches, labelled Proximal and Distal, that developed in the Line Islands (and possibly also the Phoenix Islands) and the Marquesas respectively (Wilson 2021). Proximal East Polynesian (EPnP) has been reconstructed with two sub-dialects, a northern one that spread to Hawaii and a southern one ancestral to Tahitian, Tuamotuan, Cook Islands, and Maori. Distal East Polynesian (EPnD) in the Marquesas then split, as settlers left the Marquesas for Mangareva and Rapa Nui, forming Marquesan and Far Eastern Polynesian (FEPn). Wilson's subgrouping is contradictory to the other recent model of East Polynesian subgrouping by Walworth (2014), a much flatter model based on archaeological ideas of very rapid settlement of the whole of Eastern Polynesia. Both contradict the earlier standard model of Pawley (1966) and clearly the last word on the matter has not yet been spoken (see Figure 25.7 for schematics of all three models).

Wilson's view has implications for the development of the archaeologically defined Archaic East Polynesian Culture, which presumably developed in the Phoenix and Line Islands from Northern Outlier roots and spread over much of the rest of Eastern Polynesia, with the possible exception of Hawaii (cf. Pearthree and di Piazza 2003). The modelling of radiocarbon dates presented by Wilmshurst *et al.* (2011) which informed Walworth's (2014) linguistic subgrouping and which suggested it was all accomplished in the window of 925–830 BP for Tahiti and possibly the Mangarevan group, with 760–657 BP for the outer islands including the Marquesas is by no mean universally accepted (Dye 2011; Mulrooney 2013; Athens *et al.* 2014). It could just as easily have taken place as a slightly longer and earlier phase from ca. 1,050–750 BP (Spriggs 2014) with a minor pause before the settlement of New Zealand, now convincingly modelled by Bunbury *et al.* (2022) as occurring in 700–675 BP.

Hudjashov *et al.* (2018) combined modern and aDNA to examine the genetic inheritance of East Polynesians, a study amplified by Tätte *et al.* (2022), who concentrated on modern Marquesan samples. Combining the results, we can conclude that East Polynesians are a genetically homogeneous group, although some differences between Tahitians and Marquesans can be detected. Hudjashov *et al.* (2018) gave tentative support to Wilson's
p. 485 model of East Polynesian origins from the C-N Outliers of the Solomons with their inclusion of samples ↳ from Ontong Java showing uniquely shared gene flow, but Tätte *et al.* (2022) suggest the gene flow was from East Polynesia to Ontong Java rather than the other way round. Another recent study of modern Polynesian DNA (Ioannidis *et al.* 2021) is marred by its lack of consideration of Outlier populations as a source for East Polynesian settlement. Again, we need ancient samples from the Outliers, the Phoenix and Line Islands, and Eastern Polynesia to unravel the ancestral relationships among these areas.

p. 486 The role of South Americans has recently gained greater prominence following Ioannidis *et al.*'s (2020) report of introgression of small amounts of DNA in eastern Polynesia traced to Colombian Indians (possibly the Zenu or their neighbours). They claimed that the contact, a one-off admixture event, happened around 800 BP in one of the islands from which Rapa Nui was ultimately settled, possibly the Marquesas. A rather 'romantic' suggestion going way beyond their data was made that South Americans might have reached the Marquesas slightly BEFORE the Polynesians—a tip of the genetic hat to the ideas of Thor Heyerdahl (1952), an overliteral reliance on genetic dating results as well as in apparent ignorance of the ongoing debate over the dating of East Polynesian settlement! How the putative contacts were made, whether by Polynesians continuing to voyage farther to the east and making landfall in the Americas or by raft voyages from South America arriving coincidentally just as the Polynesians reached the Marquesas, is, not surprisingly, a subject of ongoing debate (Jett 2017). Again, aDNA samples from the relevant time period would be definitive in examining the issue.

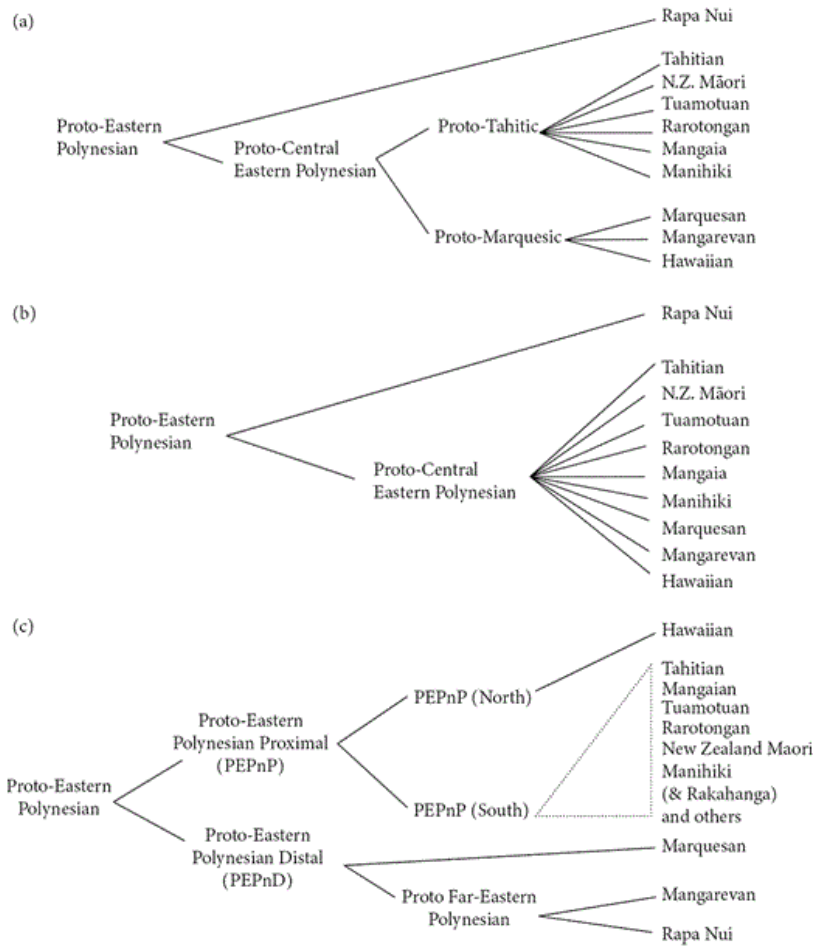

Figure 25.7 The competing subgrouping hypotheses for Eastern Polynesian

A: The conventional subgrouping of East Polynesian based on Pawley (1966). B: 'Flat' subgrouping of Walworth (2014). C: Subgrouping based on Wilson (2021). A and B Schematics taken from Walworth (2014) and C adapted from Wilson (2021), used with permission. The dashed triangle in C denotes that the relationships between these Proximal East Polynesian languages have yet to be worked out.

25.8 The northern Pacific: Micronesia

In the northern part of the Pacific, geographical Micronesia (Figure 25.8, Plate section 2), the first people to reach the Mariana Archipelago arrived at some point in the period 3,500–3,200 BP (Petchey and Clark 2010, 2021; Carson 2014; Petchey *et al.* 2018). Their material culture differed somewhat from the near contemporary Lapita assemblages in the southwest Pacific, with Marianas Redware ceramics being more like those found in the Philippines and the northern tip of Sulawesi (Spoehr 1957; Clark and Winter 2019). The oldest evidence of human occupation in Palau in Western Micronesia dates to ~3,000 BP (Clark 2005; Fitzpatrick and Jew 2018). The oldest evidence in the rest of Micronesia is currently only ~2,000 BP; shell artefacts and ceramics at these sites are similar to late Lapita pottery assemblages, and thus could reflect roots in earlier Lapita cultures in either the Bismarck Archipelago or the southwest Pacific (Athens 1990; Rainbird 2004: 94–96). The last major archaeological synthesis of Micronesia was published about twenty years ago (Rainbird 2004) but there is useful coverage in recent wider-ranging syntheses such as Kirch (2017) and Carson (2018).

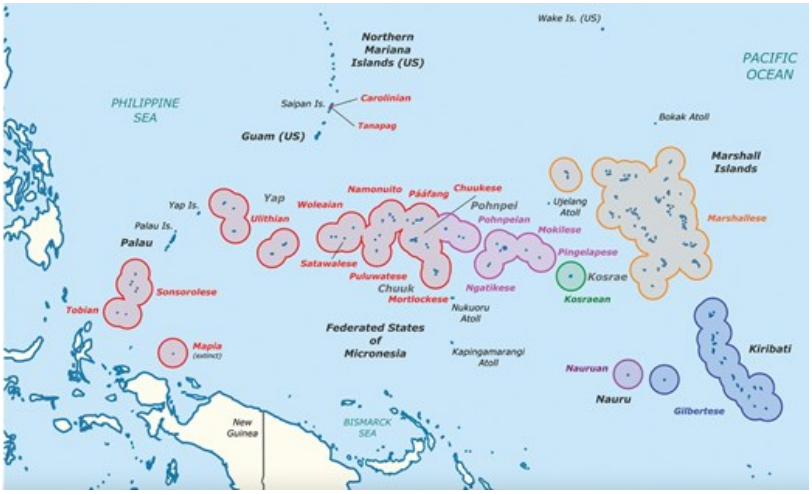

Figure 25.8 Map of Micronesia, showing linguistic groupings and incorporating Marck's (1986) 100-mile sailing radius around the islands

Source: From https://commons.wikimedia.org/wiki/File:Micronesian_languages.en.svg, by user Tyk, CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0, via Wikimedia Commons.

The linguistic situation demonstrates a series of migrations from different directions over the course of Micronesian prehistory. CHamoru of the Mariana Islands and Palauan represent two separate migrations of 'Western Malayo-Polynesian' speakers from ISEA, but linguistic levelling in the Philippines and some areas further south means that their precise origins cannot now be reconstructed linguistically (Blust 2019a: 427). Yapese, spoken on Yap and some nearby atolls, represents a very early split from Proto-Oceanic AN ca. 3,000 BP and thus provides evidence for a Bismarck Archipelago origin, perhaps more specifically the Admiralty Islands (Ross 1996). We can thus expect to find traces of Lapita-related settlement there at some point. The rest of the Micronesian languages belong to a larger grouping, Nuclear Micronesian, whose origins have been argued to be found in particular Oceanic language subgroups of an area stretching all the way from the Admiralties to North-Central Vanuatu (Song 2009). Blust has argued consistently for an origin in the Southeast Solomons,
p. 487 positing a Malaita-Micronesian subgroup (Blust 1984, 2010). The area where ↳ Longgu/Malaita/Makira languages are spoken that represent the 'Malaitan' component of this subgroup is currently very little known archaeologically and what the timing and nature of settlement from there would look like is at present unknown (Walter and Sheppard 2017; Sheppard 2022). Two Polynesian Outlier languages, spoken on Kapingamarangi and Nukuoro, complete the complex linguistic picture, with a likely settlement date of around 1,150–950 BP if Polynesians were in fact their first inhabitants.

By reporting aDNA data from 164 ancient Micronesians ranging from 2,500–500 BP and the genomic data from 112 present-day Micronesians, Liu *et al.* (2022) document five migrations that reveal a complex series of events, as we might expect from the picture given by the linguistic data, and to some extent by the currently known archaeology (Figure 25.9, Plate section 2). Micronesia was impacted by two East Asian migrations not affecting other parts of the Pacific: $FRO_{Marianas}$ and FRO_{Palau}. Present-day Indigenous people of the Marianas, including Guam, consistently have all their pre-European contact ancestry derived from these two East Asian-associated migrations and are thus the only people of the Pacific Islands who lack admixture from Papuans. Ancient people from Guam dating between 2,800 and 2,200 BP have exclusively $FRO_{Marianas}$ ancestry. The $FRO_{Marianas}$ population was impacted by a second wave of FRO East Asian ancestry similar to FRO_{Palau} so that Latte-period populations in the Marianas of 700 BP and later derive 15% of their ancestry from that source. Admixture is estimated to have occurred between 2,400 and 1,700 BP (Liu *et al.* 2022). Central Micronesia, including Pohnpei and Chuuk, derived its East Asian genetic component from $FRO_{SouthwestPacific}$, like the East

Asian-related ancestry found in Island Melanesia and Polynesia and brought into the Pacific initially by Lapita settlers.

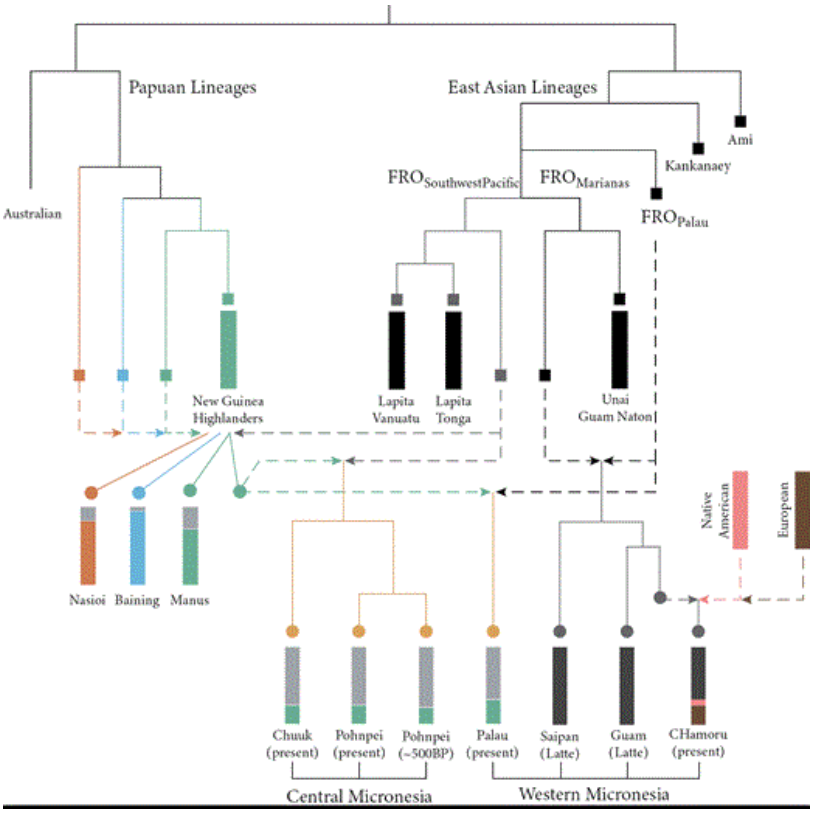

Figure 25.9 Admixture graphs for Micronesia based on ancient and modern DNA. Arrow pairs (head-to-head) denote admixture events.

The heights of the coloured bars give mixture proportions. Source: From Liu *et al.* (2022).

Papuan ancestry in Micronesia has a different source from that in Island Melanesia beyond the main Solomons and Polynesia. The Papuan ancestry in Palau and in Central Micronesia comes from the northern fringe of New Guinea, most likely the Admiralty Islands, rather than from New Britain, which is its source in the southwest Pacific, highlighting an entirely different episode or episodes of secondary Papuan spread: $PAPUAN_{NewGuinea}$ as distinct from $PAPUAN_{NewBritain}$. Modern Palauans, for instance, derive 38% of their ancestry from $PAPUAN_{NewGuinea}$, an admixture estimated as occurring 2,500–2,200 BP, while modern and ancient Pohnpeians (550–150 BP) as well as modern populations from Chuuk derive 27% of their ancestry from the same source. In Central Micronesia, the admixture is estimated to have occurred 2,100–1,800 BP, consonant with the currently accepted archaeological dates for earliest settlement there (Liu *et al.* 2022).

Given that Palau shows no traces of $FRO_{SouthwestPacific}$ settlement but has a similar source for its Papuan ancestry to Central Micronesia where the East Asian component is $FRO_{SouthwestPacific}$, we may well be seeing a situation similar to that in Vanuatu where a secondary migration of almost completely-unadmixed Papuans impacted a previously settled FRO population genetically but do not seem to have had caused significant linguistic changes to the AN languages already spoken there (Lipson *et al.* 2018b; Posth *et al.* 2018b). This is clearly the case in Palau and would seem to be the case too in the Nuclear Micronesian area of central and more eastern parts of Micronesia if Blust's (1984, 2010) source for Nuclear Micronesian in the southeast Solomons proves to be the case. We need pertinent aDNA samples from the Nuclear Micronesian area to allow us to judge. As in the Vanuatu example, such a secondary migration stream in Central Micronesia may have followed

p. 488 shortly after initial settlement by FRO populations. We have no genetic data from Yap, which may present ↳ a somewhat different genetic history given that on linguistic grounds it would have been settled a millennium earlier than other parts of Central Micronesia (Liu *et al.* 2022).

The recent genomic research also confirms that the first seafarers of the open Pacific were matrilocal. By comparing data from early Micronesian seafarers to early southwest Pacific seafarers, extreme female-line DNA genetic drift in both can only be explained by a female household-centred social system where men migrated to find their mates whereas females rarely moved to join men (Liu *et al.* 2022; Palencia-Madrid *et al.* 2019). Just such a social system was postulated for early AN seafarers by Hage (1998) and Hage and Marck (2002, 2003) on linguistic and ethnographic grounds.

25.9 Madagascar

Perhaps one of the most incredible events in AN expansion history is their reaching the Indian Ocean islands off the coast of East Africa: Madagascar and the Comoro Islands (Cox *et al.* 2012). Frequent trade activities along the Swahili Corridor have led to the present-day Comoros society displaying linguistic, cultural, and religious features that were influenced by African, Middle Eastern, and Southeast Asian populations (Gourjon *et al.* 2011; Msaidie *et al.* 2011), while the contemporary Malagasy peoples of Madagascar have preserved their identity and AN language despite extensive contact and admixture with their East African neighbours who have also settled there. Malagasy, spoken in a variety of dialects on the island, derives from the South East Barito languages mainly spoken along the eastern bank of the Barito River in South Kalimantan/Borneo in Indonesia. The language also contains loanwords from Javanese, Malay, and South Sulawesi and indirectly (via Malay and Javanese) from Sanskrit of South Asia (Adelaar 2017). These loans and a knowledge of the sociopolitical history of Island Southeast Asia strongly suggest that AN migration to Madagascar began ca. 1,450–1,350 BP. Given that it derived from a very defined area of ISEA, it was probably a planned migration of a distinct language community organized by a state-level polity, one of the Indianized kingdoms of the time, such as Banjar in southern Borneo or one of the Javanese kingdoms. This is considered in detail by Adelaar (2017), who believes there was a symbiotic relationship between those who would migrate to Madagascar and the Malay traders who would take them there; the settlers would thus have been willing voyagers with a degree of assimilation to Malay culture, rather than slaves to the Malays whose migration would have been forced.

When Madagascar was first settled by humans, and where they came from, are much contested in the archaeological literature with claimed dates ranging from 10,000 BP to almost 1,000 BP (summarized in Anderson *et al.* 2018; Anderson 2019; see also Douglass *et al.* 2019; Mitchell 2020). Despite continuing vigorous debate, some of the key protagonists seem recently to have come to a consensus position of around 2,000–1,600 BP (Hixon *et al.* 2021), but this is still a fast-changing picture. Whether earlier settlement than the attested AN one was also initially by AN-speaking traders or by African populations remains in question.

A series of genetic studies have demonstrated the presence of both African and East Asian ancestry components in all regions of Madagascar (Heiske *et al.* 2021). On average, two-thirds genetic ancestry of the present-day Malagasy people is closely related to Bantu-speaking populations, while one-third of their
p. 489 ancestry is closely related to AN speakers who ↳ originated in the Banjar region of Borneo (Hurles *et al.* 2005; Tofanelli *et al.* 2009; Pierron *et al.* 2017). Recent genetic evidence has also revealed that admixture between the AN- and Bantu-speaking people has been sex-biased, with a genetic contribution to the AN-speaking population coming from Bantu males, reflecting the well-known pattern previously seen in FRO-Papuan admixture in Oceania (see discussion above in relation to Micronesia and the southwest Pacific).

Relatedness to Bantu-speaking and AN-speaking populations supports a recent arrival of the Malagasy ancestors on the island, especially when considering no African hunter-gatherer ancestry is found in the Malagasy population. Note that the present-day hunter-gatherer groups living in Madagascar share the same

ancestries as other Malagasy people, suggesting a recent shift to a hunting and gathering specialization (Pierron *et al.* 2014). Through the analyses based on uniparental markers and whole genome analyses, the admixture between the two lineages possibly happened within the last two millennia (Pierron *et al.* 2014, 2017). The arrival of their African ancestors related to East African coast Bantu-speaking populations is suggested as being a more recent phenomenon compared to that of the Asian ancestors of the Malagasy (Semo *et al.* 2020). AN-related genomic signals have also recently been detected in Somalia in Africa and Yemen in the Arabian Peninsula, originating either from the AN-speaking populations in Madagascar and/or more directly from the Banjar of Borneo (Brucato *et al.* 2019).

25.10 The 'why' of the Austronesian spread

The most common explanation for the expansion of Austronesian languages is that it was associated in large part with the spread of agricultural ('Neolithic') societies through an area that was previously inhabited by hunting and gathering groups living at low population densities the so-called 'language/farming hypothesis' (Bellwood 2005, 2017). In support of this can be noted the retention of non-Austronesian or 'Papuan' languages across most of the island of New Guinea, a centre of independent agricultural development (Golson *et al.* 2017) with its own evidence of language expansion of the Trans-New Guinea phylum languages suggested as being linked to agricultural spread about 6,000 years ago (Pawley 2005). Later-arriving AN speakers were very largely confined to coastal and island enclaves, perhaps in competition with relatively dense already agricultural populations speaking non-Austronesian languages. The language/farming model has considerable explanatory power but is not without its critics (see recently Klamer 2019). Others have stressed more the material and non-material cultural package associated with the ISEA Neolithic and its eastward extension in the Lapita culture, including pottery, and new forms of bodily decoration and food preparation rather than the practice of agriculture *per se*. A new material culture, doubtless linked to new ritual and social practices, would have formed an attraction to already in-place Indigenous groups, leading to intermarriage, language shift, and perhaps a more gradual 'neolithization' process in already inhabited areas than envisaged in the farming/language model (Spriggs 2003).

More recently, ideas of social fission, caused by stresses in both the demographic and social sphere, have been invoked to explain the AN spread, providing a safety valve for younger sons of chiefly lines who would never
p. 490 inherit power ordinarily to establish their apex ↳ position in new settlements, either on the same island or by colonizing new lands beyond the horizon. The agency of commoner families 'voting with their feet' cannot be underestimated in such situations of increasing chiefly exactions in food and labour, providing the canoe crews and colonizing propagules necessary to establish new colonies (Furholt *et al.* 2020; see also Earle and Spriggs 2015). The vessels (large voyaging canoes and rafts) needed for such colonizing ventures could most likely only be constructed by powerful families able to mobilize labour and resources, and perhaps happy to see potentially disruptive younger sons move on elsewhere. Thus the bottom-up interests of the commoners seeking to free themselves from oppressive chiefly power and the top-down interests of junior chiefly lines hoping to establish their own primacy coincided, at least for the course of the voyage and establishment of new settlements. This is a pattern that may have been repeated across the AN range, explaining both its rapid spread and occasional long pauses, as chiefly power successively waxed and waned and new social systems developed (Spriggs 2023).

25.11 Still open questions

Among the topics that have not yet been fully addressed is the geographic origin of the three earliest FRO lineages ($FRO_{Marianas}$, FRO_{Palau}, and $FRO_{SouthwestPacific}$). The only aDNA data to date from relevant parts of Eastern Indonesia are from northern Maluku and Sulawesi. From northern Maluku the data come from sites dating to ~2,200–900 BP which show a relatively stable population of mixed Philippine/Taiwan-related and Papuan-related ancestry (Oliveira *et al.* 2022). Thus, this ancient population cannot be the source of the FRO lineages as they did not have Papuan ancestry. Of course, the movement of people into the Marianas, Palau, and Vanuatu occurred a millennium earlier than the earliest of these samples, so it is certainly possible that the ancestry in the region changed in that period and a putative FRO source population that might have lived in northern Maluku, or a broader region of Eastern Indonesia that included northern Maluku, has not been sampled thus far. From Sulawesi, there is an aDNA datapoint from ~250 BP, and another from ~7800 BP. Both of these have Papuan-like or Papuan/Negrito-like ancestry, which is not present in FRO populations, so they cannot be proxies for the source population. Again, there is no current aDNA sampling from the right time period (ca. 4,000–3,000 BP), but the main point here is that there is no evidence at present to support any Eastern Indonesian population, either ancient or modern, as a good genetic proxy for the FRO lineages.

From the Philippines, there are no aDNA data yet, and obtaining such data should be a priority. Putting this all together, there is suggestive genetic evidence supporting a Philippines origin for the FRO lineages. This evidence consists of the fact that the Kankanaey and other Philippine groups look like they might be directly descended without mixture from the FRO source population, and furthermore the FROs are more closely related genetically to modern Philippine groups like Kankanaey than to modern aboriginal Taiwan groups like Ami or Iron Age ancient Taiwanese. There is also suggestive genetic evidence against a Sulawesi or northern Maluku or other Eastern Indonesian origin for the FROs. The evidence is that no modern or ancient populations sampled from these regions to date are consistent with being a source for the FROs or descended from them without admixture. However, it is possible that our knowledge might change with more aDNA
p. 491 sampling from the appropriate ↳ time period in Eastern Indonesia and the Philippines, and it is very important to carry out this work.

We still lack estimations of the population split date among the FRO lineages and the population split date between the FROs and Iron Age and present-day Indigenous Taiwan lineages and the present-day AN-related Philippine lineages, such as Kankanaey. Determining how the FRO lineages and Kankanaey lineages are related to the Taiwan-associated lineages present in admixed form in different parts of Southeast Asia, the Pacific, and Madagascar would be also very informative.

Future studies of disease patterns, such as when yaws, sexually transmitted diseases, and particular kinds of malaria spread across the region will illuminate further details of the lives of AN speakers previously invisible to us. This would help to explain why the Papuan genetic component is dominant in Island Melanesia today, with the possibility that they had a selective advantage through greater resistance to malaria than the $FRO_{SouthwestPacific}$ initial migrants. When malaria was introduced to the Reefs-Santa Cruz Islands and Vanuatu is also an open question.

Unresolved questions also include, for example: how did Austronesian speakers pick up Austroasiatic ancestry in Southeast Asia? How and when did Sinitic languages replace putative Austronesian languages in coastal South China? Why do almost all present-day Indonesian and Filipino people, having both AN and non-Austronesian ancestry, speak AN languages? The colonization of Madagascar was around 1,450–1,350 BP by Austronesian farmers whose closest relatives today are in Borneo. What motivated this extremely long-distance migration? Will future genetic studies illuminate any intermediate steps for this settlement from Borneo to Madagascar?

As well as genetic and linguistic questions, there are many unresolved archaeological ones. They are generally questions of sampling; we simply do not have enough archaeological information for many key areas. No significant archaeological research, for instance, has been carried out in the West New Guinea region believed to be the homeland of the SHWNG languages (see Wright *et al.* 2013). It could well turn out to be the proximate Lapita 'homeland'! Nor do we have any early Lapita sites from the Admiralty Islands, which, on linguistic evidence, might represent one of the earliest areas of Lapita settlement (Blust 2019a: 408). We might expect, on similar grounds, Lapita settlement on Yap but no sites that old have yet been found. Whether and how the Islands of Makira and Malaita in the Solomon Islands relate to the settlement of the Nuclear Micronesian area are unknown in the absence of any sites in these islands of the right age. Questions remain, too, relating to the settlement of Eastern Polynesia, whether from the Samoan area, as in the conventional narrative, or via the C-N Outliers off the Solomon Islands.

25.12 Conclusions

Austronesian language distribution matches the archaeological and genomic evidence well. The latter shows distinct patterns of admixture between Indigenous groups and incoming AN speakers, dependent on the demographic and socio-economic milieux within which such contact occurred. Future aDNA research, particularly from early cemetery sites, has the potential to illuminate kinship patterns that until very recently,
p. 492 archaeologists could only ↳ speculate upon, usually based on modern ethnographic and/or linguistic reconstructions and thus previously difficult to date. Genomics has clearly added new and exciting vistas of future research unimaginable little over a decade ago but we should not forget the 300 years of linguistic scholarship and the more than 150 years of archaeological studies that have provided not only a firm basis for future advances, but which continue to provide both questions and interpretive power in relation to genomic studies.

Acknowledgements

MS thanks his wife Rosemary Leona, a native Raga speaker (North/Central Vanuatu Subgroup of Oceanic), for her patience while he struggled with his contribution to this chapter over the Christmas 2022 and New Year period. Jennifer Sheehan of CartoGIS, Scholarly Services, ANU is thanked for timely assistance with sourcing and adapting the figures. MS thanks Roger Blench and Andy Pawley for continuing discussions over AN language prehistory and late past masters of AN linguistics who have influenced his thinking over many years and who were always willing to engage with the archaeological evidence: Bob Blust, George Grace, and John Lynch. PG also thanks Frank Thomas and Pila Wilson for valuable assistance. YCL thanks David Reich and Mark Lipson for helpful discussions.

Notes

1 After this text was submitted, linguist Malcolm Ross announced a major challenge to the validity of C-EMP and EMP as innovation defined subgroups, contending that SHWNG is not a sister-group to Oc and is more closely related to some of the CMP languages in a group he calls 'Wallacean' (Ross *et al.* 2023: 40–41). If this turns out to be the case, Oc is either a sister-group to Wallacean in a larger grouping comparable to current C-EMP, or comes off the Malayo-Polynesian tree at a higher level, equivalent to WMP/Residual MP and Wallacean groups (Ross et al. 2023: 45–46). The latter subgrouping could suggest a more direct movement from an area nearer the northern Philippines to the Oc and Lapita 'homeland' in the Bismarck Archipelago.