Ancient genomes from the last three millennia support multiple human dispersals into Wallacea

Sandra Oliveira[1]*, Kathrin Nägele[2]*, Selina Carlhoff[2]*, Irina Pugach[1], Toetik Koesbardiati[3], Alexander Hübner[1,2], Matthias Meyer[1], Adhi Agus Oktaviana[4], Masami Takenaka[5], Chiaki Katagiri[6], Delta Bayu Murti[3], Rizky Sugianto Putri[3], Mahirta[7], Thomas Higham[8,9], Charles F. W. Higham[10], Sue O'Connor[11,12], Stuart Hawkins[11,12], Rebecca Kinaston[13], Peter Bellwood[14], Rintaro Ono[15], Adam Powell[16], Johannes Krause[2]*, Cosimo Posth[2,17,18]*, Mark Stoneking[1,19]*

1 Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
2 Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
3 Department of Anthropology, Faculty of Social Sciences and Political Sciences, Universitay Airlangga, Surabaya, Indonesia
4 The National Research Center for Archaeology, Jakarta, Indonesia
5 Kagoshima Women's College, Kagoshima, Japan
6 Okinawa Prefectural Archaeological Center, Nishihara, Japan
7 Jurusan Arkeologi, Fakultas Ilmu Budaya, Universitas Gadjah Mada, Yogyakarta, Indonesia
8 Department of Evolutionary Anthropology, University of Vienna, Vienna, Austria
9 Oxford Radiocarbon Accelerator Unit, RLAHA, University of Oxford, United Kingdom
10 Department of Anthropology, University of Otago, Dunedin, New Zealand
11 School of Culture, History and Language, College of Asia and the Pacific, Australian National University, Acton, Australia
12 ARC Centre of Excellence for Australian Biodiversity and Heritage, Australian National University, Canberra, ACT, Australia
13 Department of Anatomy, School of Medical Sciences, University of Otago, Dunedin, New Zealand
14 School of Archaeology and Anthropology, College of Arts and Social Sciences, Australian National University, Canberra, Australia
15 Center for Cultural Resource Studies, National Museum of Ethnology, Osaka, Japan
16 Department of Human Behavior, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
17 Institute for Archaeological Sciences, Archaeo- and Palaeogenetics, University of Tübingen, Tübingen, Germany
18 Senckenberg Centre for Human Evolution and Palaeoenvironment, University of Tübingen, Tübingen, Germany
19 Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive, UMR 5558, Villeurbanne, France

*These authors contributed equally

Abstract

Previous research indicates that the human genetic diversity found in Wallacea - islands in present-day Eastern Indonesia and Timor-Leste that were never part of the Sunda or Sahul continental shelves - has been shaped by complex interactions between migrating Austronesian farmers and indigenous hunter-gatherer communities. Here, we provide new insights into this region's demographic history based on genome-wide data from 16 ancient individuals (2600-250 yrs BP) from islands of the North Moluccas, Sulawesi, and East Nusa Tenggara. While the ancestry of individuals from the northern islands fit earlier views of contact between groups related to the Austronesian expansion and the first colonization of Sahul, the ancestry of individuals from the southern islands revealed additional contributions from Mainland Southeast Asia, which seems to predate the Austronesian admixture in the region. Admixture time estimates for the oldest individuals of Wallacea are closer to archaeological estimates for the Austronesian arrival into the region than are admixture time estimates for present-day groups. The decreasing trend in admixture times exhibited by younger individuals supports a scenario of multiple or continuous admixture involving Papuan- and Asian-related groups. Our results clarify previously debated times of admixture and suggest that the Neolithic dispersals into Island Southeast Asia are associated with the spread of multiple genetic ancestries.

Wallacea (Figure 1), a region of deep-sea islands located between the Sunda and Sahul continental shelves[1], has been both a bridge and a barrier for humans migrating from Asia to New Guinea, Australia, and the Pacific Islands. Modern humans presumably first crossed the Wallacean island chains before reaching Sahul, for which the earliest proposed date is ~65 kyr BP, associated with the Madjedbebe rock shelter in northern Australia[2]. However, this date has been questioned[3] and the earliest unequivocal dates for modern humans in Sahul are around 47 kyr BP[4-6]. In Wallacea itself, the archaeological record indicates occupation by AMH starting ~46 kyr BP in the southern Wallacean islands[7-9], ~45.5 kyr BP in Sulawesi[10], and ~36 kyr BP in the northern Wallacean islands (North Moluccas)[11].

After a long period of occupation by hunter-gatherer communities, the region was impacted by the Austronesian expansion. Equipped with novel sailing and farming technologies, Austronesian-speaking groups likely expanded out of Taiwan ~4000-5000 yr BP[12-14] and eventually settled in Island Southeast Asia, Oceania, and Madagascar. Their arrival is generally linked to the earliest appearance of pottery, which dates to ~3500 yr BP in Wallacea[11,15-18].

During the late Neolithic period and the early Metal Age in Wallacea (2300-2000 yr BP) the maritime trade network intensified further, with a movement of spices such as Northern Moluccan cloves (*Syzygium aromaticum*), bronze drums, and glass beads connecting Wallacea to India and Mainland Southeast Asia[11,17,19-24].

The contact between Austronesian-speaking groups and the previously established hunter-gatherer communities is still reflected in the linguistic and biological diversity of Wallacea today. Austronesian languages of the Malayo-Polynesian (MP) subgroup are widespread throughout the region[25], but a few dozen non-Austronesian (i.e., Papuan) languages are also spoken in Wallacea, including in the North Moluccas, Timor, Alor, and Pantar[26]. Additionally, some Austronesian languages spoken in Wallacea show linguistic features acquired from Papuan languages[27].

Analyses of the genomic composition of present-day groups from the region have shown signals of admixture between Papuan-related ancestry and an Asian-related ancestry most similar to that of present-day Austronesian-speaking groups[28-30]. This dual ancestry is geographically distributed as a gradient of increasing Papuan-related ancestry from west to east[28,30]. Previous studies have estimated the time of admixture of the two ancestries using data from present-day groups[28-30], providing the first inferences on the direction and rate of spread of genetic ancestry across the islands[28]. However, the time estimates from different studies show discrepancies of more than 3000 yr (Table S1) that cannot be solely attributed to ascertainment biases in the sampled individuals and analyzed markers but are also affected by the methods used to infer the date of admixture[28,29]. In fact, current methods cannot accurately infer admixture dates for complex scenarios involving continuous gene flow or repeated admixture from closely related sources[31], which could have occurred in the area. As a result, even relatively recent (<4000 yr) demographic movements across Wallacea are not well understood. Resolving the discrepancies between admixture dates has important implications for understanding interactions between Austronesians and indigenous pre-Austronesian populations. Admixture dates close to the archaeological dates proposed for the Austronesian arrival would indicate that admixture occurred soon after contact. In contrast, very recent dates would imply that communities co-existed for a long time before genetically mixing or were mixing for a prolonged time period. Moreover, admixture dates predating the Austronesian arrival would suggest alternative explanations, such as genetic influences from other Asian-related groups in earlier periods[32].

In this study, we leverage the power of ancient DNA to investigate spatio-temporal patterns of variation within Wallacea during the last ~2500 yr. We provide insights into the time of arrival of the Austronesian-related ancestry, the temporal span of admixture, and the relationship between the ancestry of incomers and the ancestry of other modern human groups from Island Southeast Asia and Oceania. Additionally, we explore the impact and timing of an additional migration from Mainland Southeast Asia to Wallacea.

Results

We extracted DNA from skeletal remains from 16 individuals dated to ~2600-250 yr BP from eight archaeological sites spanning the North Moluccas, Sulawesi, and East Nusa Tenggara (Figure 1A, Table 1, Table S2). Sequencing libraries were then constructed and capture-enriched for ~1.2 million genome-wide single nucleotide polymorphisms (SNPs)[33] and the complete mitochondrial genome (mtDNA). The authenticity of ancient DNA was confirmed based on the elevated amounts of deaminated positions at the ends of reads (up to ~45% in non-UDG treated samples) and the short average fragments size (~55 base pairs) (Table S2). Contamination estimates were low (~0.01% for males based on X-chromosomes and ~2% based on mtDNA) (Table S2). The mean mtDNA coverage ranged from ~8X to ~900X (Table S2), while the number of SNPs obtained for each individual ranged from ~123,000 to ~1 million (Table 1).

MtDNA and Y-chromosome haplogroups were determined based on the mitochondrial genomes and SNP data, respectively, and are reported, together with their associated origin, in Table 1. The haplogroup composition shows that both Asian- and Australo-Papuan-related ancestries were already present in the North Moluccas ~2150 yrs BP. Furthermore, two North Moluccas individuals dating to ~2150-2100 yrs BP carry mtDNA and Y chromosome haplogroups associated with different ancestries, indicating that admixture started before that period. In comparison to the individuals from East Nusa Tenggara and Sulawesi, those from North Moluccas show a higher proportion of mtDNA lineages that connect them to Near Oceania, as attested by the Q haplogroups characteristic of Northern Sahul [34] and by the so-called "Polynesian pre-motif" (B4a1a/B4a1a1)[35]. None of the individuals from Sulawesi or East Nusa Tenggara carry Australo-Papuan-related mtDNA haplogroups, which are found there today[36,37], likely reflecting the small number of ancient individuals analyzed.

To explore the genome-wide patterns of variation in the ancient individuals from Wallacea, we performed principal component analysis (PCA) based on two different sets of present-day populations from Asia and Oceania (see methods; Figure 1B and Figure S1). The ancient Wallaceans cluster between populations from Papua New Guinea and Asia, together with present-day populations from their respective geographical areas. However, the trajectory outlined by individuals from the northern (North Moluccas) vs. southern (East Nusa Tenggara) islands is slightly different, suggesting they may have distinct genetic histories. Ancient individuals from East Nusa Tenggara cluster on a cline towards mainland Asians and some Western Indonesian groups, while ancient individuals from the North Moluccas align on a trajectory towards present-day Taiwanese and Philippine populations or even towards ancient individuals from Guam 2200 BP, Vanuatu 2900 BP, and Tonga 2500 BP. These ancient individuals were all previously shown to have almost exclusively Austronesian-related ancestry[38,39].

We next used a model-based clustering method (DyStruct) to infer shared ancestry, considering the dating of the individuals used in the analysis[40]. The results for the best supported number of clusters in each of the tested datasets (Figure S2A-B) show that ancient individuals from Wallacea share ancestry with Papuan-speaking

groups from New Guinea (dark blue component) and multiple Asian groups whose ancestry can be partitioned into three main components (Figure 1C; see full results in Figure S2C-D). One component (yellow) is present at high frequencies in Austronesian-speaking groups from Taiwan, Philippines, and Indonesia, as well as ancient individuals from Taiwan; a second component (mango) is maximized in Polynesian-speaking groups from Near Oceania, Tahiti, Tonga, and Samoa, and ancient individuals from the same region; and a third component (dark red) is widespread in present-day and ancient individuals from Southeast Asia. The most striking difference among the ancient Wallaceans is the presence of the Southeast Asian component (dark red) in the ancient East Nusa Tenggara and Sulawesi individuals but the absence of this component in the North Moluccan individuals. A more subtle difference occurs in the relative proportion of the two Austronesian-related components (Figure S3): ancient individuals from Sulawesi and East Nusa Tenggara have a higher relative proportion of the Austronesian-related (yellow) component that predominates in the likely source of the expansion (Taiwan), compared to the ancient individuals from the North Moluccas, who are more similar to groups from the Pacific.

To directly compare the sharing of alleles between ancient individuals from Wallacea and different Asian-related groups, we used f-statistics[41]. First, we computed an f_4-statistic of the form $f_4(Mbuti, ancient Wallacean; Amis, test)$, where the test group includes the ancient and present-day groups from mainland Asia, Island Southeast Asia, and the Pacific that have no discernible Papuan-related ancestry (Figure S4). A significant positive result indicates that the ancient Wallacean shares more drift with the test group than with the Taiwanese Amis, while a significant negative result indicates that the ancient Wallacean shares more drift with the Amis than with the test group. Our results show that ancient individuals from the North Moluccas share more drift with ancient individuals from Vanuatu (2900 BP) and Tonga (2500BP) than with Amis. In contrast, ancient individuals from Sulawesi and East Nusa Tenggara do not share additional drift with any tested groups. Nonetheless, the higher number of f_4-statistics consistent with zero in tests involving ancient individuals from East Nusa Tenggara (Komodo and Liang Bua) indicates that they share as much drift with Amis as with several other groups, not only from Taiwan and Philippines but also from Southeast Asia or Western Indonesia. This result, together with the identification of an additional ancestry component related to Southeast Asia (Figure S2C-D) in the ancient East Nusa Tenggara and Sulawesi individuals, supports a more complex admixture history in these parts of Wallacea.

We next created a series of two-dimensional plots comparing pairs of f_4-statistics designed to capture any differences between the North Moluccas and East Nusa Tenggara individuals that might reflect differential Asian-related ancestries. All pairs had the form $f_4(Mbuti, test; New Guinea Highlanders, ancient Wallacean)$ and always included Amis or ancient Vanuatu (2900 BP) vs. other Asian-related groups as $test$. Since individuals from the North Moluccas lack the Southeast Asian component in the DyStruct analysis, their differences to East Nusa Tenggara individuals regarding the attraction to the Asian-related groups should be maximized by the best proxies for the Southeast Asian ancestry in East Nusa Tenggara. By comparing pairs of present-day groups (Figure S5) and pairs of ancient groups (Figure S6), we conclude that the groups maximizing differences between ancient individuals from Wallacea are the present-day Mlabri or Nicobarese and ancient individuals from Vietnam (Nam Tun 2600BP, Man Bac 3800BP, Mai Da Dieu 4000BP), Cambodia (Vat Komnou 1800BP), Thailand (Ban Chiang 2500BP and 3350BP), Malaysia (Gua Cha Cave 2300BP), and Laos (Tam Pa Ping 3000BP) (Figure 2). These are thus the best proxies among the tested groups and individuals for the Southeast Asian ancestry in East Nusa Tenggara.

We also investigated the relationships between the newly presented ancient Wallaceans and groups associated with the first colonization of Sahul or Wallacea using an f-statistic of the form $f_4(Mbuti, new ancient Wallacean;$

New Guinea Highlanders, test). Since the tested groups are present-day Australo-Papuans with no discernable Asian ancestry and a recently published pre-Neolithic individual from Sulawesi (Leang Panninge)[42], values consistent with zero indicate that ancient Wallaceans are equally related to New Guinea Highlanders and the tested groups, while deviations from zero indicate differences in affinity to the tested group. Most ancient Wallacean individuals show a significantly closer affinity to New Guinea Highlanders than to Australians, the group representing the Bismarck archipelago, or the Leang Panninge individual (Figure S7). The low number of SNPs available for some *f4* tests (e.g., tests involving the Leang Panninge individual) and the low amount of Papuan ancestry carried by the Liang Bua and Topogaro individuals likely account for the non-significant differences in affinity with New Guinea Highlanders vs. Bismarcks or Leang Panninge by broadening the confidence intervals (Figure S7) and reducing the power to distinguish between potential sources of this ancestry (Figure S8), respectively. Nonetheless, tests involving the Leang Panninge individual consistently exhibit the lowest *f4* values. Therefore, despite being from Wallacea, this ancient individual is a worse proxy for the Papuan-related ancestry of the newly reported ancient Wallaceans than any other tested group.

We further investigated potential differences in ancestry among ancient Wallaceans by modeling ancestry sources and proportions with the qpAdm software[41]. Our results indicate that whereas the ancient individuals from the North Moluccas can be modeled as having both Papuan- and Austronesian-related ancestry, ancient individuals from East Nusa Tenggara and Sulawesi are either consistent with or require a three-wave model, with additional Southeast Asian-related ancestry (Figure 3; Table S3). Despite the cases for which we identify more than one fitting model ($P > 0.01$), the estimated proportions under the model with the highest *P*-value correlate with the proportions of Austronesian, Papuan, and Southeast Asian ancestry inferred by DyStruct (mantel statistic r = 0.97, $P < 0.001$). The ancient East Nusa Tenggara individuals display more inter-island variance in their Papuan- and Southeast Asian-related ancestries ($s^2 = 0.026$ and 0.046, respectively) compared to their Austronesian-related ancestry ($s^2 = 0.003$).

A comparison between the ancestry composition of ancient and present-day individuals from the same region (Figure S3; Table S3) suggest that a small part (~8%) of the Austronesian-related ancestry of the ancient North Moluccas individuals was replaced by Southeast Asian ancestry in present-day groups, masking former differences between regions of Wallacea. The present-day groups from Sulawesi and East Nusa Tenggara can be modeled by the same three ancestry components found in the ancient individuals from those regions. However, the ancestry proportions of ancient and present-day groups show some differences at the regional level, which could be explained by ancestry shifts over time, or could also reflect the small sample sizes.

To gain insights into the relative order of admixture events between different ancestries through time in the northern and southern Wallacean islands, we used an approach called Admixture History Graph (AHG)[43] that relies on differences in the covariance between the ancestry components inferred by the DyStruct analysis (Table S4-6). The AHG, applied to both ancient and present-day data from East Nusa Tenggara, suggests that the admixture of Southeast Asian- and Papuan-related ancestries occurred before the arrival of the Austronesian-related ancestry (Table S6). An analogous test based on the three main ancestry components observed among individuals from the North Moluccas (one Papuan and two Austronesian-related) does not provide compelling evidence of backflow from Austronesian-related groups of the Pacific, as the AHG infers that Papuan ancestry was introduced into a population that already had the two Austronesian-related components (Table S6). This result suggests that drift had a more significant role in the occurrence and distribution of the two Austronesian-related components.

Finally, we investigated the timing of admixture using the software DATES (Figure S9), which can be applied to ancient DNA data from single individuals and has been shown to return estimates that are as accurate as those estimated by other methods based on multiple individuals[44]. Since we estimate admixture dates for individuals from different periods, we expect to gain more insights about admixture through time and reconcile the time estimates from different methods, despite the challenges associated with dating complex scenarios. Using Papuans and a pool of Asian groups as sources, we found that estimates for the oldest individuals from the Northern Moluccas (2150 BP) and East Nusa Tenggara (2600 BP) are very similar (~3000 years BP, adjusting for the archaeological age of each sample; Figure 4), approaching archaeological dates for the arrival of the Austronesians in Wallacea. However, ancient individuals from more recent periods display more recent estimates. Contrary to present-day groups from East Nusa Tenggara, present-day groups from the North Moluccas show even younger admixture dates (~1400 BP) than the ancient individuals from the same region (Figure 4). This pattern probably reflects differential admixture pulses (and/or continuous gene flow) since the initial contact between Austronesians and Papuan-related groups, as suggested by the changes in ancestry composition or ancestry proportions over time (Figure 3).

Discussion

This study greatly increases the amount of ancient genomic data from Island Southeast Asia, a tropical region unsuitable for DNA preservation but particularly important for understanding human population interactions and admixture. Notably, we report the first genomic data from an ancient human from the Liang Bua cave on the island of Flores, the same site that provided the only known remains of *Homo floresiensis*. Although the amount of genomic data recovered precluded detailed investigation into admixture with archaic humans, the fact that authentic ancient DNA could be recovered from 2600 yrs old remains from Liang Bua suggests that such investigations could be possible for other ancient human remains from this site.

Our results reveal striking regional variation in the ancestry of Wallacean individuals and provide insights into the temporal window and nature of events that shaped them. Some of the variation is found among the Austronesian-related ancestry of ancient northern vs. southern Wallaceans. The most remarkable differences are associated with ancestry contributions from Mainland Southeast Asia that were probably already part of the East Nusa Tenggara genomic landscape when Austronesians arrived. The Australo-Papuan ancestry of Wallaceans, on the other hand, supports some shared or parallel histories.

Australo-Papuan ancestry in Wallacea

Based on our analysis, all newly presented ancient Wallacean individuals are genetically closer to present-day Australo-Papuans than to the recently published pre-Neolithic Leang Panninge individual (~7 kyr BP) from Sulawesi[42], suggesting little direct continuity between pre-Neolithic and post-Austronesian Wallaceans. Additionally, the Australo-Papuan ancestry of the newly presented Wallaceans is closer to the ancestry of present-day Papua New Guineans than indigenous Australians. This proximity suggests that either the group that gave rise to Australians split first, or there was contact between Wallacea and Papua New Guinea after their initial settlement. The second scenario is supported by a recent mtDNA study which inferred two major influxes of Papuan ancestry into Wallacea, one following the Last Glacial Maximum ~15 kyr, the other associated with Austronesian contact ~3 kyr[45].

Previous studies have also reported elevated amounts of Denisovan ancestry in present-day Wallaceans, which correlate with the amount of Papuan-related ancestry[46]. We confirmed that the same relationship holds for the ancient Wallaceans included in this study by plotting estimated Papuan vs. Denisovan ancestry (Figure S10), thus supporting the idea that their Denisovan-related ancestry was likely contributed via Papuan-related admixture.

Early Southeast Asian ancestry in East Nusa Tenggara

The area to the east of the Wallace line is generally assumed to have been mostly isolated and shaped by two main streams of ancestry: one related to the first settlement of Sahul and another associated with the Austronesian expansion[28-30]. The results presented here show that the genetic variation of ancient individuals from East Nusa Tenggara also requires ancestry contributions from Mainland Southeast Asia. The Admixture History Graph (AHG) analysis further suggests that this ancestry likely admixed with Papuan-related ancestry before the arrival of Austronesian-related ancestry. The inferred order of events makes it unlikely that the Southeast Asian and Austronesian-related ancestries were introduced together in the south of Wallacea from Western Indonesia, where both ancestries are found. Instead, it seems that human groups from Mainland Southeast Asia crossed into southern Wallacea before the Austronesian-related groups spread into the region.

The broad geographical distribution of ancient and present-day groups best matching the Mainland Southeast Asian ancestry found in the south of Wallacea raises questions about the actual origin(s) of peoples that reached those islands. The best present-day proxies, the Mlabri from Thailand/Laos and the Nicobarese from the Nicobar islands, speak Austroasiatic languages and are genetically similar, with the Nicobarese displaying some additional ancestry related to non-Austroasiatic Onge (from the neighboring Andaman Islands)[47]. Even though the Mlabri are a hunter-gatherer group, their ethnogenesis has been attributed to a strong founder event from an ancestral agricultural population[48]. The fact that both the Mlabri and Nicobarese have been relatively isolated compared to other Southeast Asian groups that recently experienced extensive admixture[49,50], might explain why they are the best present-day proxies for the Mainland Southeast Asian ancestry in southern Wallacea, without being necessarily connected to the inferred migration event. Moreover, ancient individuals from all across Mainland Southeast Asia are also equivalent proxies for this ancestry in southern Wallacea (Figure 2 and S6); thus, there is no clear link between any specific ancient group from Southeast Asia and the actual source that contributed ancestry to East Nusa Tenggara.

The only documented languages in East Nusa Tenggara belong to the Austronesian family (Malayo-Polynesian subgroup) or the Papuan language cluster, and no significant influences from Mainland Southeast Asian language families have been reported in southern Wallacea. While all populations in Island Southeast Asia to the west of Wallacea currently speak Austronesian languages, an Austroasiatic influence has been suggested in some Borneo languages[51], and Borneo individuals carry a relatively high degree of Mainland Southeast Asian genetic ancestry (Figure S2). However, no linguists have suggested a prior presence of Austroasiatic languages in East Nusa Tenggara, and neither are there direct archaeological traces of pre-Austronesian contact between this region and Mainland Southeast Asia.

The earliest evidence of interaction likely comes from the appearance of Dong Son bronze drums, which spread to southern but not northern Wallacea around the early centuries AD, following early maritime trade routes[52]. These drums probably originated in northern Vietnam or adjacent provinces of southern China[52]. Although we cannot rule out some Southeast Asian ancestry contributions from the Dong Son period (or even later) for the younger East Nusa Tenggara individuals, the high amount of Southeast Asian ancestry in the Liang Bua

individual (2600 BP) and our AHG inferences support its earlier presence in southern Wallacea. Ancient DNA from older time periods will help to clarify the time of arrival of this ancestry.

The relationship between the Austronesian ancestry of the North Moluccas and the Pacific

The fine-scale structure observed among Austronesian-related groups from Island Southeast Asia and the Pacific, and the higher genetic proximity of the ancient North Moluccas individuals to the latter, are pertinent for earlier considerations of the role of the North Moluccas in dispersals to the Pacific[20]. When analyzed through sophisticated seafaring and climatic models, the North Moluccas appears to be one of the most likely starting points for settlers that ventured into the Palau or Mariana Islands (Western Micronesia)[53,54]. Their geographical setting also led archeologists to search the region for pottery that could be considered ancestral to the Lapita cultural complex (distributed from the Bismarck archipelago to New Caledonia in the south and Samoa in the east), as well as the Marianas Redware culture[11,18]. However, to date no evidence connects the North Moluccas red-slipped pottery to either of these material cultures, and ancient DNA from Guam supports an origin for the settlement of the Mariana Islands from the Philippines[39]. Overall, the North Moluccas red-slipped Neolithic pottery is closer to that of the Talaud Islands, Northern and Western Sulawesi, North Luzon, Batanes, and Southeastern Taiwan[11].

Our genetic results are compatible with a route through the North Moluccas, leading to the settlement of the Mariana Islands and the Lapita-related dispersals. Under a relatively simple expansion scenario, without significant back migration, the increasing amounts of Austronesian ancestry characteristic of the Pacific (and decrease of ancestry characteristic of Taiwan/Philippines) from the ancient North Moluccas to Guam (~2200 BP), Vanuatu (~2900 BP), and Tonga (~2500 BP) could reflect their relative position along the peopling wave that eventually reached the Eastern parts of the Pacific (Figure S3). Yet, the position here refers to the split order of groups, without any necessary attachment to their geographical location. Therefore, it is possible that the higher proximity between the North Moluccas and groups from the Pacific, compared to East Nusa Tenggara or Sulawesi, simply reflects their more recent ancestry tracing back to a common Austronesian source, regardless of its location. This scenario also implies that the Austronesian-related ancestry found in East Nusa Tenggara or Central Sulawesi is somewhat differentiated from that found in the North Moluccas. Furthermore, we cannot exclude the possibility of more complex migration scenarios (e.g., involving back-migrations).

It is also important to consider that the dates of the oldest individuals from the North Moluccas (2150 BP in Morotai and 1900 BP in Kayoa island) overlap with the start of the Early Metal age ~2300-2000 yr BP in this region[11]. This period is characterized by the appearance of copper, bronze, and iron artifacts and glass beads in the region, as well as the spread of pottery into Morotai island. Thus, these individuals might not be good representatives of the first wave of Austronesians, thought to have reached Kayoa island 3500 yr BP[11,18], but instead might feature additional genetic influences brought by later contacts.

The Austronesian (Malayo-Polynesian major subgroup) languages of the Northern Moluccas are part of the South Halmahera–West New Guinea (SHWNG) regional subgroup. According to Blust[55], the SHWNG languages are closer to Oceanic languages - the most clearly defined MP branch - than to any other Western MP subgroup, have a time depth of ~2000 yr, and might represent a replacement movement from the East[55,56]. Additionally, the language phylogeny of Gray et al.[13] places the SHWNG languages as the closest outgroup to all Oceanic languages and further estimates their divergence to around 3500 yr BP, whereas the languages spoken in East Nusa Tenggara are an outgroup to both SHWNG and Oceanic languages. Thus, the closer

genetic proximity of the ancient North Moluccas to Oceanians, compared to East Nusa Tenggarans, is also mirrored in their linguistic relationships.

The timing of Admixture

Besides providing direct evidence of Austronesian-Papuan contact prior to 2150 yr BP in the North Moluccas and 2600 yr BP in East Nusa Tenggara, the oldest individuals gave admixture date estimates close to 3000 years BP. This period is slightly younger than the earliest archaeological traces of the Neolithic (Austronesian) arrival in the North Moluccas (~3500 years BP for Kayoa Island[11,18]) but older than the adoption of pottery on Morotai Island (~2300-2000 years BP), where the oldest North Moluccan individuals in this study were found[11,57]. However, it is similar to some of the earliest secure dates from East Nusa Tenggara (~3000 years BP for eastern Flores[17]. Some of the previous studies conducted on present-day eastern Indonesian populations suggested that this admixture lagged about a millennium after the arrival of Austronesian populations[30]. Our admixture analysis for ancient individuals, and the comparison with present-day data, provides an alternative explanation, and helps to clarify previous debates concerning times of admixture[28-30,32,58]. The decreasing trend in admixture estimates from the oldest individuals until present-day populations is a strong indicator that admixture occurred in multiple pulses, or even continuously, over a considerable time period. Therefore, even our oldest estimates might not correspond to the actual start of admixture but rather to a more recent time due to the bias introduced by continuous admixture.

Evidence of interaction between populations over time comes from multiple sources. In the Metal Age, admixture might have been facilitated by emergent maritime networks and spice trade interactions[11]. In the North Moluccas, this period not only corresponds to a more rapid spread of material culture between regions[11,21,22,57], but also to the inferred period of language leveling or radiation described for both the Austronesian (SHWNG) and Papuan (West Papuan Phylum, Northern Halmahera Stock) languages of the North Moluccas[11]. The historical socio-economic systems of the North Moluccas and western Papua also brought together Papuan-speaking resident populations and a Malay-speaking elite[11], reinforcing that mixing could have occurred until very recent times, and have involved diverse peoples. In contrast to the North Moluccas, the East Nusa Tenggara and Sulawesi ancient and present-day individuals do not show genetic traces of very recent contact. Still, their demographic history was nonetheless characterized by a long-term process of admixture involving at least two Asian-related ancestries.

These findings supporting ongoing contact in Wallacea not only improve our understanding of the history of the region, but also have important implications for efforts that use present-day genomic data to discern the direction and number of human migrations to the Australo-Papuan region (e.g. ref. [59]). In particular, failure to take this ongoing contact into consideration may result in wrongly considering the genetic affinity between Papuans and northern vs. southern Wallaceans to reflect ancestral relationships of these groups rather then differences in the degree of contact. Overall, our findings suggest different histories for northern vs. southern Wallaceans that reflect differences in contact with Southeast Asia, in the temporal span of contact with Papua New Guinea, and perhaps even with different Austronesian-related groups. Future ancient DNA studies involving individuals from earlier periods will help to improve our understanding of the demographic changes occurring both before and after the arrival of farming in Wallacea.

Methods

Sampling. All samples were processed in dedicated ancient DNA laboratories at the Max Planck Institute for the Science of Human History (Jena) and the Max Planck Institute for Evolutionary Anthropology (Leipzig), Germany. In Jena, the petrous bone of samples AMA001, AMA004, and AMA009 was first drilled from the outside, identifying the position of the densest part by orienting on the internal acoustic meter and drilling parallel to it into the target area to avoid damaging the semicircular ducts[60] (protocol: dx.doi.org/10.17504/protocols.io.bqd8ms9w). After that, the petrous bone was cut along the margo superior partis petrosae (crista pyramidis) and 50 − 150 mg bone powder were drilled from the densest part around the cochlea[61]. All other elements processed in Jena (AMA003, AMA005, AMA008, JAB001, KMO001, LIA001, LIA002, LIT001, TOP002, TOP004) were sampled by cutting and drilling the densest part. In Leipzig, the Tanjung Pinang and Uattamdi specimens were sampled by targeting the cochlea from outside. For this, a thin layer of surface (~1 mm) was removed with a sterile dentistry drill. Small holes were then drilled into the cleaned areas, yielding between 42 and 63 mg of bone powder. Detailed information on the analysed samples, radiocarbon dating, and archaeological context are provided in the Supplementary Information and in Tables S2 and S7.

DNA Extraction. DNA extraction in both laboratories was carried out using a silica-based method optimized for the recovery of highly degraded DNA[62], with modifications described elsewhere[63] and large-volume silica spin columns and binding buffer 'D' for DNA binding. In Jena, the second elution of DNA from the spin column was carried out using a fresh aliquot of elution buffer, for a total of 100 µl DNA extract, whereas in Leipzig the same aliquot of elution buffer was loaded twice, for a total of 50 µl DNA extract (protocol: dx.doi.org/10.17504/protocols.io.baksicwe).

Library preparation. In Jena double stranded DNA libraries were built from 25 µl of DNA extract in the presence of uracil DNA glycosylase (UDG-half libraries), following a protocol that uses the UDG enzyme to reduce, but not eliminate, the amount of deamination-induced damage towards the ends of aDNA fragments[64]. Negative and positive controls were carried alongside each experiment (extraction and library preparation) (protocol: dx.doi.org/10.17504/protocols.io.bmh6k39e). Libraries were quantified using the IS7 and IS8 primers[65] in a quantification assay using a DyNAmo SYBP Green qPCR Kit (Thermo Fisher Scientific) on the LightCycler 480 (Roche). Each aDNA library was double indexed[66] in parallel 100 µl reactions using PfuTurbo DNA Polymerase (Agilent) (protocol: dx.doi.org/10.17504/protocols.io.bakticwn). The indexed products for each library were pooled, purified over MinElute columns (Qiagen), eluted in 50 µl TET and again quantified using the IS5 and IS6 primers[65] using the quantification method described above. 4 µl of the purified product were amplified in multiple 100 µl reactions using Herculase II Fusion DNA Polymerase (Agilent) following the manufacturer's specifications with 0.3 µM of the IS5/IS6 primers. After another MinElute purification, the product was quantified using the Agilent 2100 Bioanalyzer DNA 1000 chip. An equimolar pool of all libraries was then prepared for shotgun sequencing on the Illumina HiSeq 4000 platform using a SR75 sequencing kit. Libraries were further amplified with IS5/IS6 primers to reach a concentration of 200-400 ng/µl as measured on a NanoDrop spectrophotometer (Thermo Fisher Scientific). In Leipzig, single-stranded DNA libraries were prepared without UDG treatment using the Bravo NGS workstation B (Agilent Technologies), followed by library quantification, amplification and double-indexing, as described in detail in ref. [67].

Targeted enrichment and high-throughput sequencing. Mitochondrial DNA capture[68] was performed on screened libraries which, after shotgun sequencing, showed the presence of aDNA, highlighted by the typical C to T and G to A substitution pattern towards the 5′ and 3′ molecule ends, respectively. Furthermore, samples with a percentage of human DNA in shotgun data around 0.1% or greater were enriched for a set of 1,237,207 targeted SNPs across the human genome (1,240K capture) as described previously[33]. The enriched DNA product was sequenced on an Illumina HiSeq 4000 instrument with 75 cycles single-reads or 50 cycles pair-end-reads using the manufacturer's protocol (in Jena), or on a HiSeq 2500 with 75 paired-end-reads (in Leipzig). The output was demultiplexed using in-house scripts requiring either a perfect match of the expected and observed index sequences (Leipzig samples) or allowing a single mismatch between the expected and observed index sequences (Jena samples).

Genomic data processing. Pre-processing of the sequenced reads was performed using EAGER v.1.92.55[69]. The resulting reads were clipped to remove residual adaptor sequences using *Clip&Merge*[70] and *AdapterRemoval* v.2[71]. Clipped sequences were then mapped against the human reference genome hg19 using the Burrows–Wheeler Aligner (BWA) v.0.7.12[72], disabling seeding (-l 16500) and allowing for 2 mismatches (–n 0.01). Duplicates were removed with DeDup v. 0.12.2[69]. Additionally, a mapping quality filter of 30 was applied using SAMtools v.1.3[73]. Different sequencing runs and libraries from the same individuals were merged, duplicates removed, and sorted again using SAMtools v.1.3[73]. Genotype calling was performed separately for trimmed and untrimmed reads using pileupCaller v.8.6.5 (https://github.com/stschiff/sequenceTools), a tool that randomly draws one allele at each of the targeted SNPs covered at least once. For the UDG-treated libraries produced in Jena, 2 bases were trimmed on both ends of the reads. For libraries produced in Leipzig (without UDG treatment) the damage-plots were inspected to determine the number of bases to trim from each read. For all libraries, the residual damage extended 8bp into the read, after which it was below 0.05%, and trimmed accordingly. We combined the genotypes keeping all transversions from the untrimmed genotypes and transitions only from the trimmed genotypes to eliminate problematic, damage-related transitions overrepresented at the ends of reads. The generated pseudo-haploid calls were merged with previously published ancient data[38,39,42,47,74-78] present-day genomes from the Simons Genome Diversity Project[79] and worldwide populations genotyped on the Affymetrix Human Origins array[38,41,49,74,75,80-85]. For PCA and DyStruct analyses we additionally merged the data with populations from Island Southeast Asia genotyped on the Affymetrix 6.0[46,86] (dataset 1) or the Affymetrix Axiom Genome-Wide Human Array[30] (dataset 2), filtering out SNPs with a missing rate higher than 10%. Related individuals were excluded if they exhibited a proportion of identity by descent (IBD) higher than 0.3, computed in PLINK v.1.9[87] as P(IBD=2) + 0.5*P(IBD=1). We additionally pruned datasets 1 and 2 for linkage disequilibrium with PLINK v.1.9, removing SNPs with $r^2 > 0.4$ in 200 kb windows, shifted at 25 SNP intervals. After pruning, a total of 89,597 and 65,880 SNPs remained in datasets 1 and 2, respectively.

Y-chromosomal haplogroups were identified by calling the SNPs covered on the Y-chromosome of all male individuals using the pileup from the Rsamtools[88] package and by analysing the SNPs overlapping with the ISOGG SNP index v.14.07 (https://isogg.org/tree) as described in ref. [89].

Authentication of ancient DNA. The typical features of ancient DNA were inspected with *DamageProfiler* v.0.3.1 (http://bintray.com/apeltzer/EAGER/DamageProfiler)[69]. Sex determination was performed by comparing the coverage on the targeted X-chromosome SNPs to the coverage on the Y-chromosome SNPs, both normalized by the coverage on the autosomal SNPs[70] (Table S2). For male individuals, ANGSD v.0.919 was run

to measure the rate of heterozygosity of polymorphic sites on the X-chromosome after accounting for sequencing errors in the flanking regions[90]. This provides an estimate of nuclear DNA contamination in males, as they are expected to have only one allele at each site. For both male and female individuals, mtDNA-captured data were used to jointly reconstruct the mtDNA consensus sequence and estimate contamination levels with contamMix[68] (Table S2) using an in-house pipeline (https://github.com/alexhbnr-mitoBench-ancientMT[39]).

Statistical analyses. Principal component analyses were carried out using *smartpca* v.10210[91] based on present-day Asian and Oceanian populations from datasets 1 and 2. Ancient individuals were projected onto the calculated components using the options 'lsqproject: YES' and 'numoutlieriter: 0'. We used DyStruct v.1.1.0[40] to infer shared genetic ancestry from time-series genotype data. The uncalibrated radiocarbon dates of each ancient sample were converted to generations, assuming a generation time of 29 years[92]. For each dataset (1 and 2), we performed 25 independent runs, using 2 to 15 ancestral populations (K). To compare runs for different values of K, a subset of loci (5%) was held out during training and the conditional log likelihood was subsequently evaluated (Figure S2A-B). Within the best K, the run with the highest objective function was selected (Figure S2C-D).

To formally test population relationships we used f_4 statistics, implemented in the ADMIXTOOLS software[41]. This analysis was carried out using the admixr R package[93]. We used *qpWave* v.410[94] and *qpAdm* v.650[41] to test two- and three-wave admixture models, following a "rotating" strategy[95]. A reference set of populations was chosen to represent diverse human groups and include potential source populations for the ancient Wallacean individuals: Mbuti, English, Brahui, Onge, Yakut, Oroqen, Lahu, Miao, Dai, Khomu, Denisova, Papuan, Kankanaey, Mlabri. We rejected models if their *P*-values were lower than 0.01, if there were negative admixture proportions, or if the standard errors were larger than the corresponding admixture proportion. When more than one model was accepted (Table S3), the estimated admixture proportions under the model with the highest *P*-value was preferred and used in subsequent analyses (Figure S8 and S10) because the results better matched the DyStruct ancestry proportions, and the ability to reject models might be affected by several factors (e.g., the ancestry proportion, the quality of the target sample, the combination of ancient and present-day samples in the same analysis, etc). The correlation between ancestry proportions inferred with *qpAdm* and DyStruct was assessed using a Mantel test with 10,000 permutations of the distance matrix to determine significance.

The relative order of the mixing of different ancestries was inferred using the Admixture History Graphs (AHG) approach described previously[39]. We used the DyStruct proportions for each ancient and present-day Wallacean individual included in dataset 1 and 2 (Table S4-5) to calculate the covariances between ancestry components as indicated in Table S6. The time since admixture was estimated based on the decay of ancestry covariance using the software DATES v.753[44], with parameters: binsize: 0.001; maxdis: 1.0; jackknife: YES; qbin: 10; runfit: YES; afffit: YES; lovalfit: 0.45; mincount: 1. To maximize the number of SNPs included in the analysis and have equal sample sizes, we used as sources 16 Papuan individuals and 16 Asian-related individuas (2 Ami, 1 Atayal, 2 Kankanaey, 5 Dai, 2 Dusun, 2 She, 2 Kinh) with data covering the ~1,240,000 SNPs captured in the ancient samples.

References

1   Dickerson, R. E. Distribution of life in the Philippines. *Monographs of the Bureau of Science* **21**, 1-322 (1928).

2   Clarkson, C. *et al.* Human occupation of northern Australia by 65,000 years ago. *Nature* **547**, 306-310 (2017).

3   O'Connell, J. F. *et al.* When did Homo sapiens first reach Southeast Asia and Sahul? *Proc Natl Acad Sci* **115**, 8482-8490 (2018).

4   Allen, J. & O'Connell, J. Both half right: updating the evidence for dating first human arrivals in Sahul. *Aust Archaeol* **79**, 86-108 (2014).

5   Veth, P. Breaking through the radiocarbon barrier: Madjedbebe and the new chronology for Aboriginal occupation of Australia. *Aust Archaeol* **83**, 165-167 (2017).

6   Allen, J. & O'Connell, J. F. A different paradigm for the initial colonisation of Sahul. *Archaeol Oceania* **55**, 1-14 (2020).

7   Hawkins, S. *et al.* Oldest human occupation of Wallacea at Laili Cave, Timor-Leste, shows broad-spectrum foraging responses to late Pleistocene environments. *Quat Sci Rev* **171**, 58-72 (2017).

8   Shipton, C., O'Connor, S., Reepmeyer, C., Kealy, S. & Jankowski, N. Shell Adzes, exotic obsidian, and inter-island voyaging in the early and middle Holocene of Wallacea. *J Island Coast Archaeol* **15**, 525-546 (2020).

9   Sutikna, T. *et al.* The spatio-temporal distribution of archaeological and faunal finds at Liang Bua (Flores, Indonesia) in light of the revised chronology for Homo floresiensis. *J Hum Evol* **124**, 52-74 (2018).

10   Brumm, A. *et al.* Oldest cave art found in Sulawesi. *Sci Adv* **7**, eabd4648 (2021).

11   Bellwood, P. *The Spice Islands in Prehistory: Archaeology in the Northern Moluccas, Indonesia*. (ANU Press, 2019).

12   Bellwood, P. & Dizon, E. The Batanes archaeological project and the "Out of Taiwan" hypothesis for Austronesian dispersal. *Journal of Austronesian Studies* **1**, 1-33 (2005).

13   Gray, R. D., Drummond, A. J. & Greenhill, S. J. Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* **323**, 479-483 (2009).

14   Ko, A. M.-S. *et al.* Early Austronesians: into and out of Taiwan. *Am J Hum Genet* **94**, 426-436 (2014).

15   Anggraeni, S. T., Bellwood, P. & Piper, P. Neolithic foundations in the Karama valley, West Sulawesi, Indonesia. *Antiquity* **88**, 740-756 (2014).

16   O'Connor, S. Rethinking the Neolithic in Island Southeast Asia, with Particular Reference to the Archaeology of Timor-Leste and Sulawesi. *Archipel. Études interdisciplinaires sur le monde insulindien*, 15-47 (2015).

17   Galipaud, J.-C. *et al.* The Pain Haka burial ground on Flores: Indonesian evidence for a shared Neolithic belief system in Southeast Asia. *Antiquity* **90**, 1505 (2016).

18   Ono, R, Oktaviana AA, Sriwigati, Aziz N. In *The Archaeology of Island Colonization* (ed Matthew F. Napolitano, Robert J. DiNapoli) 293-326 (University Press of Florida, 2021).

19   Bellwood, P., Waluyo, A., Nitihaminoto, G. & Irwin, G. Archaeological research in the northern Moluccas; interim results, 1991 field season. *Bulletin of the Indo-Pacific Prehistory Association* **13**, 20-33 (1993).

20   Bellwood, P. *First islanders: prehistory and human migration in Island Southeast Asia*. (John Wiley & Sons, 2017).

21   Ono, R. *et al.* Development of regional maritime networks during the Early Metal Age in northern Maluku Islands: A view from excavated glass ornaments and pottery variation. *The J Island Coast Archaeol* **13**, 90-108 (2018).

22   Ono, R. *et al.* Development of pottery making tradition and maritime networks during the Early Metal Ages in Northern Maluku Islands. *AMERTA* 35 **2**, 109-122 (2017).

23   Lape, P. V. Political dynamics and religious change in the late pre-colonial Banda Islands, Eastern Indonesia. *World Archaeol.* **32**, 138-155 (2000).

24   Solheim, W. G. The University of Hawaiʻi archaeological programme in Eastern Indonesia. *Southeast Asian Archaeology* **1996**, 61-73 (1998).

25   Blust, R. The Austronesian homeland and dispersal. *Annu Rev Linguist* **5**, 417-434 (2019).

26   Holton, G. & Klamer, M. The Papuan languages of East Nusantara and the Bird's Head. In *The languages and linguistics of the New Guinea Area,* 569-640 (De Gruyter Mouton, 2017).

27   Klamer, M. The dispersal of Austronesian languages in Island South East Asia: Current findings and debates. *Lang Linguist Compass* **13**, e12325 (2019).

28   Xu, S. *et al.* Genetic dating indicates that the Asian–Papuan admixture through Eastern Indonesia corresponds to the Austronesian expansion. *Proc Natl Acad Sci* **109**, 4574-4579 (2012).

29   Lipson, M. *et al.* Reconstructing Austronesian population history in island Southeast Asia. *Nat Commun* **5**, 1-7 (2014).

30   Hudjashov, G. *et al.* Complex patterns of admixture across the Indonesian archipelago. *Mol Biol Evol* **34**, 2439-2452 (2017).

31   Pugach, I. *et al.* The gateway from Near into Remote Oceania: new insights from genome-wide data. *Mol Biol Evol* **35**, 871-886 (2018).

32   Denham, T. & Donohue, M. Lack of correspondence between Asian-Papuan genetic admixture and Austronesian language dispersal in eastern Indonesia. *Proc Natl Acad Sci* **109**, E2577-E2577 (2012).

33   Fu, Q. *et al.* An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216-219 (2015).

34   Hudjashov, G. *et al.* Revealing the prehistoric settlement of Australia by Y chromosome and mtDNA analysis. *Proc Natl Acad Sci* **104**, 8726-8730 (2007).

35   Soares, P. *et al.* Ancient voyaging and Polynesian origins. *Am J Hum Genet* **88**, 239-247 (2011).

36   Mona, S. *et al.* Genetic admixture history of Eastern Indonesia as revealed by Y-chromosome and mitochondrial DNA analysis. *Mol Biol Evol* **26**, 1865-1877 (2009).

37   Tumonggor, M. K. *et al.* The Indonesian archipelago: an ancient genetic highway linking Asia and the Pacific. *J Hum Genet* **58**, 165-173 (2013).

38   Skoglund, P. *et al.* Genomic insights into the peopling of the Southwest Pacific. *Nature* **538**, 510-513 (2016).

39   Pugach, I. *et al.* Ancient DNA from Guam and the Peopling of the Pacific. *Proc Natl Acad Sci* **118** (2021).

40   Joseph, T. A. & Pe'er, I. Inference of population structure from time-series genotype data. *Am J Hum Genet* **105**, 317-333 (2019).

41   Patterson, N. *et al.* Ancient Admixture in Human History. *Genetics* **192**, 1065-1093 (2012).

42   Carlhoff, S. *et al.* Genome of a middle Holocene hunter-gatherer from Wallacea. *Nature* **596**, 543-547 (2021).

43   Pugach, I. *et al.* The complex admixture history and recent southern origins of Siberian populations. *Mol Biol Evol* **33**, 1777-1795 (2016).

44   Narasimhan, V. M. *et al.* The formation of human populations in South and Central Asia. *Science* **365**, 6457 (2018).

45   Purnomo, G. A. *et al.* Mitogenomes reveal two major influxes of Papuan ancestry across Wallacea following the last glacial maximum and Austronesian contact. *Genes* **12**, 965 (2021).

46   Reich, D. *et al.* Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania. *Am J Hum Genet* **89**, 516-528 (2011).

47   Lipson, M. *et al.* Ancient genomes document multiple waves of migration in Southeast Asian prehistory. *Science* **361**, 92-95 (2018).

48   Oota, H. *et al.* Recent origin and cultural reversion of a hunter–gatherer group. *PLoS Biol* **3**, e71 (2005).

49   Liu, D. *et al.* Extensive ethnolinguistic diversity in Vietnam reflects multiple sources of genetic diversity. *Mol Biol Evol* **37**, 2503-2519 (2020).

50   Kutanan, W. *et al.* Reconstructing the human genetic history of mainland Southeast Asia: insights from genome-wide data from Thailand and Laos. *Mol Biol Evol* **38**, 3459-3477 (2021).

51   Blench, R. Was there an Austroasiatic presence in Island Southeast Asia prior to the Austronesian expansion? *Bulletin of the Indo-Pacific Prehistory Association* **30**, 133-144 (2010).

52   Calo, A. *Trails of bronze drums across early Southeast Asia: exchange routes and connected cultural spheres* (Institute of Southeast Asian Studies, 2013).

53   Montenegro, Á., Callaghan, R. T. & Fitzpatrick, S. M. Using seafaring simulations and shortest-hop trajectories to model the prehistoric colonization of Remote Oceania. *Proc Natl Acad Sci* **113**, 12685-12690 (2016).

54   Fitzpatrick, S. M. & Callaghan, R. T. Estimating trajectories of colonisation to the Mariana Islands, western Pacific. *Antiquity* **87**, 840-853 (2013).

55   Blust, R. *Eastern Malayo-Polynesian: a subgrouping argument*. In Wurm, S.A. and Carrington, L. (eds) Second International Conference on Austronesian Linguistics 1, 181–234 (Pacific Linguistics Series C 61, 1978)

56   Kamholz, D. C. *Austronesians in Papua: Diversification and change in South Halmahera–West New Guinea*. (University of California, Berkeley, 2014).

57   Ono, R. *et al.* Early Metal Age interactions in Island Southeast Asia and Oceania: jar burials from Aru Manara, northern Moluccas. *Antiquity* **92**, 1023-1039 (2018).

58   Xu, S. & Stoneking, M. Reply to Denham and Donohue: Asian-Papuan genetic admixture is in excellent agreement with Austronesian dispersal in eastern Indonesia. *Proc Natl Acad Sci* **109**, E2578-E2578 (2012).

59   Brucato, N. *et al.* Papua New Guinean genomes reveal the complex settlement of north Sahul. *Mol Biol Evol* (2021).

60   Himmel, M. *Etablierung und Evaluierung einer CT-Scan-informierten minimalinvasiven Methode zur Maximierung der DNA-Gewinnung aus prähistorischem Skelettmaterial vorgelegt von* Bachelor thesis, Ernst-Abbe-Hochschule Jena, (2017).

61   Pinhasi, R. *et al.* Optimal ancient DNA yields from the inner ear part of the human petrous bone. *PloS one* **10**, e0129102 (2015).

62   Dabney, J. *et al.* Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci* **110**, 15758-15763 (2013).

63   Rohland, N., Glocke, I., Aximu-Petri, A. & Meyer, M. Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nat Protoc* **13**, 2447-2461 (2018).

64   Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Philos Trans R Soc Lond, B, Biol Sci* **370**, 20130624 (2015).

65   Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* **2010**, pdb prot5448 (2010).

66   Kircher, M., Sawyer, S. & Meyer, M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* **40**, e3 (2012).

67   Gansauge, M.-T., Aximu-Petri, A., Nagel, S. & Meyer, M. Manual and automated preparation of single-stranded DNA libraries for the sequencing of DNA from ancient biological remains and other sources of highly degraded DNA. *Nat Protoc* **15,** 1-23 (2020).

68   Fu, Q. *et al.* A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr Biol* **23**, 553-559 (2013).

69   Peltzer, A. *et al.* EAGER: efficient ancient genome reconstruction. *Genome Biol* **17**, 60 (2016).

70   Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* **534**, 200-205 (2016).

71   Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes* **9**, 88 (2016).

72   Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).

73   Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).

74   Lipson, M. *et al.* Population turnover in Remote Oceania shortly after initial settlement. *Curr Biol* **28**, 1157-1165. e1157 (2018).

75   McColl, H. *et al.* The prehistoric peopling of Southeast Asia. *Science* **361**, 88-92 (2018).

76   Posth, C. *et al.* Language continuity despite population replacement in Remote Oceania. *Nat Ecol Evol* **2**, 731-740 (2018).

77   Yang, M. A. *et al.* Ancient DNA indicates human population shifts and admixture in northern and southern China. *Science,* **369**, 282-288 (2020).

78   Yang, M. A. *et al.* 40,000-year-old individual from Asia provides insight into early population structure in Eurasia. *Curr Biol* **27**, 3202-3208. e3209 (2017).

79   Mallick, S. *et al.* The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201-206 (2016).

80   Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409-413 (2014).

81   Skoglund, P. *et al.* Genetic evidence for two founding populations of the Americas. *Nature* **525**, 104-108 (2015).

82   Meyer, M. *et al.* A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222-226 (2012).

83   Nakatsuka, N. *et al.* The promise of discovering population-specific disease-associated genes in South Asia. *Nat Genet* **49**, 1403-1407 (2017).

84   Prüfer, K. *et al.* The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**, 43-49 (2014).

85   Qin, P. & Stoneking, M. Denisovan ancestry in East Eurasian and native American populations. *Mol Biol Evol* **32**, 2665-2674 (2015).

86   Jinam, T. A. *et al.* Discerning the origins of the Negritos, first Sundaland people: deep divergence and archaic admixture. *Genome Biol Evol* **9**, 2013-2022 (2017).

87   Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575 (2007).

88   Rsamtools: Binary alignment (BAM), FASTA, variant call (BCF), and tabix file import, R package version 1.34.1. (2019).

89   Rohrlach, A. B. *et al.* Using Y-chromosome capture enrichment to resolve haplogroup H2 shows new evidence for a two-Path Neolithic expansion to Western Europe. *Sci Rep* **11**, 15005 (2021).

90   Korneliussen, T. S. A., A; Nielsen, Rasmus. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinform* **15**, 13 (2014).

91   Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet* **2**, e190 (2006).

92   Fenner, J. N. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am J Phys Anthropol* **128**, 415-423 (2005).

93   Petr, M., Vernot, B. & Kelso, J. admixr—R package for reproducible analyses using ADMIXTOOLS. *Bioinformatics* **35**, 3194-3195 (2019).

94   Reich, D. *et al.* Reconstructing Native American population history. *Nature* **488**, 370-374 (2012).

95   Harney, É., Patterson, N., Reich, D. & Wakeley, J. Assessing the Performance of qpAdm: A Statistical Tool for Studying Population Admixture. *Genetics* **217**(4), iyaa045 (2020).

Acknowledgements

Author contributions

RO, PB, RK, TK, and SH contributed archaeological material, collected with the critical support of AAO, MT, CK, DBM, RSP, M, and SOC. TH, CFWH, RK, and PB contributed radiocarbon data. KN, SC and MM conducted the ancient DNA laboratory work. SO, KN, SC, IP, and AH performed the genetic analysis. SO and KN wrote the manuscript with input from all authors. MS, CP, and JK concieved and coordinated the study.

Competing interests

The authors declare no competing interests.

| Sample name	Island, region	C14 date ± SD (yrs BP)	Assigned group	Sex	mtDNA haplogroup	Y chromosome haplogroup	Number of SNPs
AMA001	Morotai, North Moluccas	2258 ± 30	Aru Manara 2150 BP	M	B4a1a1	C1b1a2b	255,458
AMA003008	Morotai, North Moluccas	2130 ± 24		F	Q1d		576,009
AMA004	Morotai, North Moluccas	2009 ± 24		F	M73a		933,715
AMA005	Morotai, North Moluccas			F	B4a1a1		209,933
AMA009	Morotai, North Moluccas	968 ± 20	Aru Manara 950 BP	F	Q1d		935,157
TanjungPinang1	Morotai, North Moluccas	2090 ± 180	Tanjung Pinang 2100 BP	M	Q	S1d1~	870,652
TanjungPinang2	Morotai, North Moluccas			M	Q1	O2a2b2a2b2	939,665
TanjungPinang4	Morotai, North Moluccas			M	Q	S1a1b1d2b~	897,982
TanjungPinang6	Morotai, North Moluccas			M	B4a1a	O2a2b2a2b2	1,028,190
Uattamdi1	Kayoa, North Moluccas	1915 ± 27	Uattamdi 1900 BP	M	E1a1a1	O1a2a1	895,300
TOP002	Sulawesi, Central Sulawesi	211 ± 24	Topogaro 250 BP	M	E2a	O2a2a1a2a2	697,028
TOP004	Sulawesi, Central Sulawesi	324 ± 24		M	E2a	M1a	249,209
KMO001	Komodo, East Nusa Tenggara	726 ± 19	Komodo 750 BP	F	B4a1a1		122,610
LIA001002	Flores, East Nusa Tenggara	2588 ± 23	Liang Bua 2600 BP	F	M17a		873,614
LIT001	Flores, East Nusa Tenggara	861 ± 20	Liang Toge 850 BP	F	E1a2		623,960
JAB001	Pantar, East Nusa Tenggara	457 ± 19	Jareng Bori 450 BP	F	M7b1a2a1		848,849

Table 1 – Ancient samples from Wallacea included in this study. MtDNA and Y-chromosome haplogroups are colored according to their most likely origin based on previous publications: red – Asian, blue – Australo-Papuan.
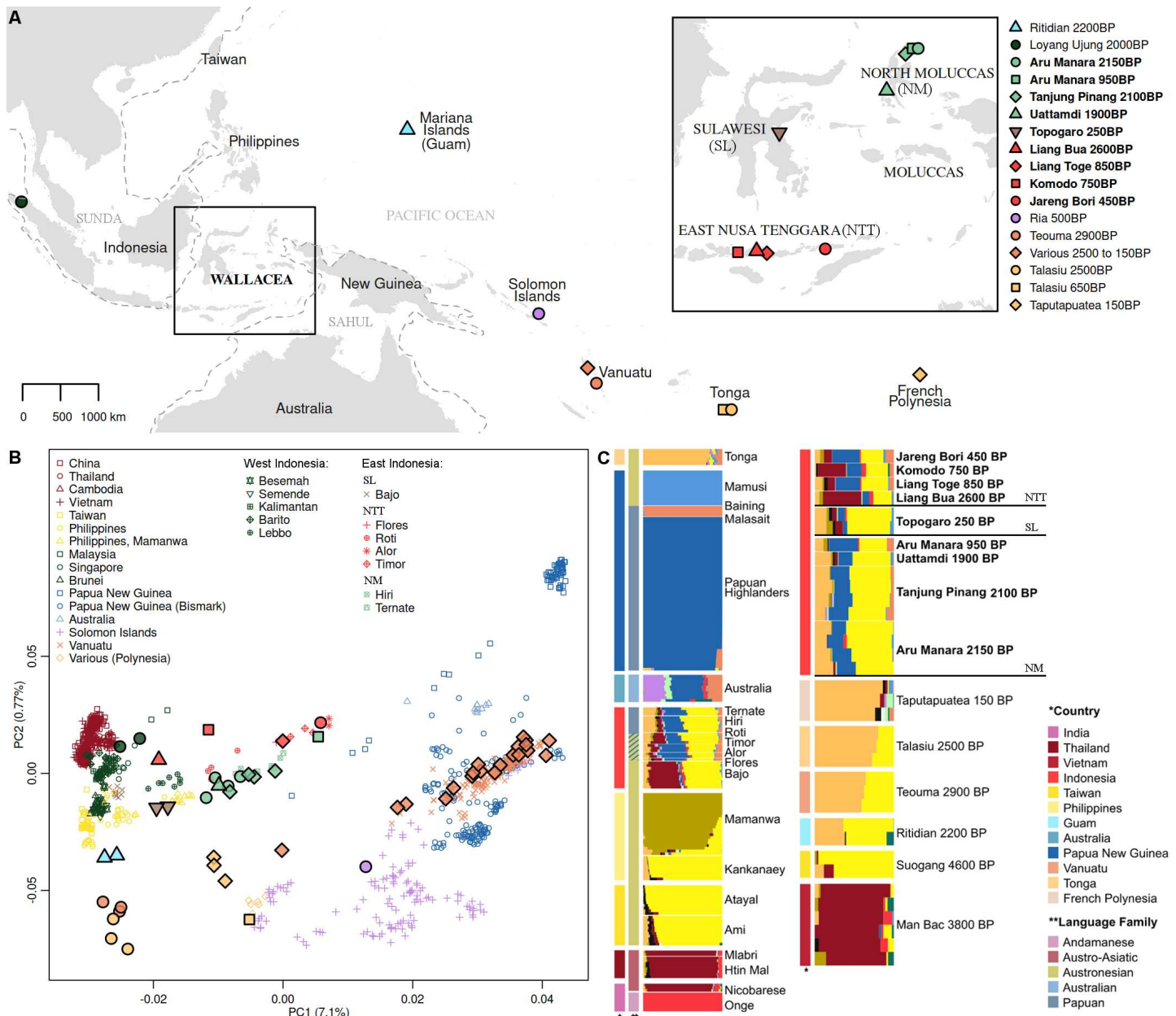
Figure 1 – Sample provenience, PC and DyStruct analyses. A) Map showing the location of ancient individuals. B) PCA of publicly available whole genome data merged with Human Origins and Affymetrix 6.0 genotype data. Ancient individuals (shown with a black contour) are projected and their fill color matches the color of present-day individuals from the same geographic area. C) DyStruct results for the same merged dataset as B) displaying only a subset of the individuals included in the full analysis (Figure S2C). Newly generated individuals are highlighted in bold in the legend. Country and Language information are displayed as color bars to the left of the inferred ancestry components.
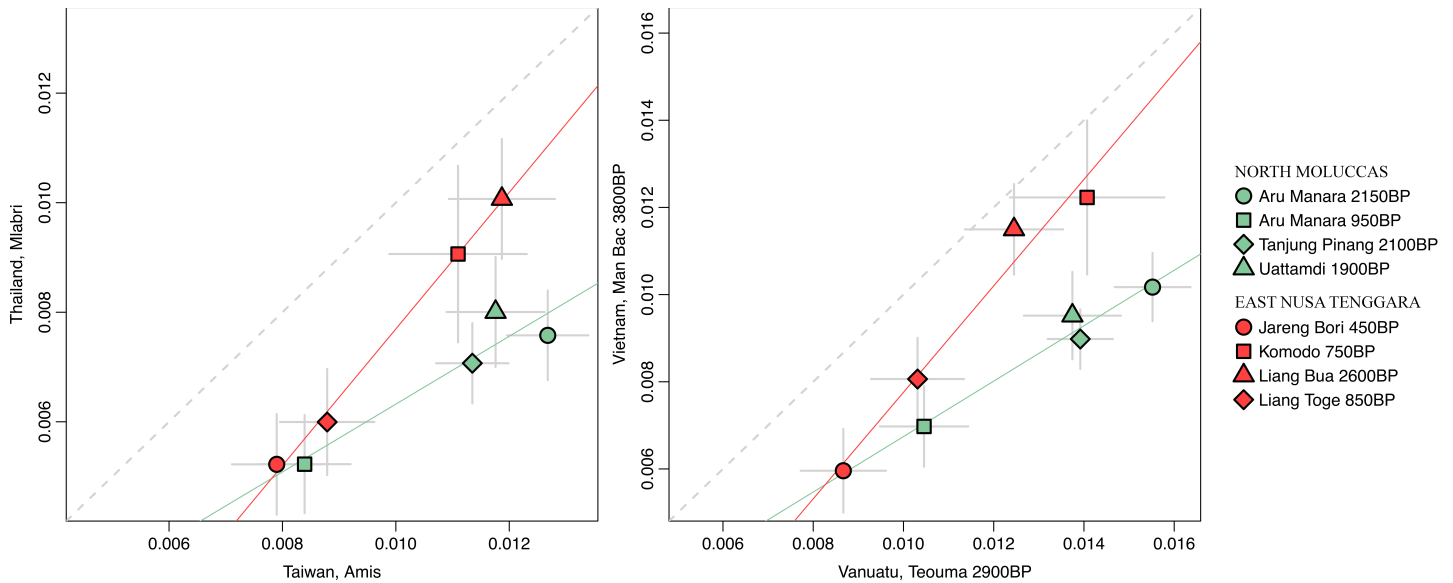
Figure 2 – Biplots showing the results of two pairs of f_4-statistics of the form: $f_4(Mbuti, test; New Guinea Highlanders, ancient Wallacea)$. The test groups are shown on the x-y axis label. Grey bars show two standard errors in each direction. Linear regression lines for the North Moluccas and East Nusa Tenggara individuals are shown in green and red, respectively. The results for all tested pairs are shown in Figures S5 and S6.
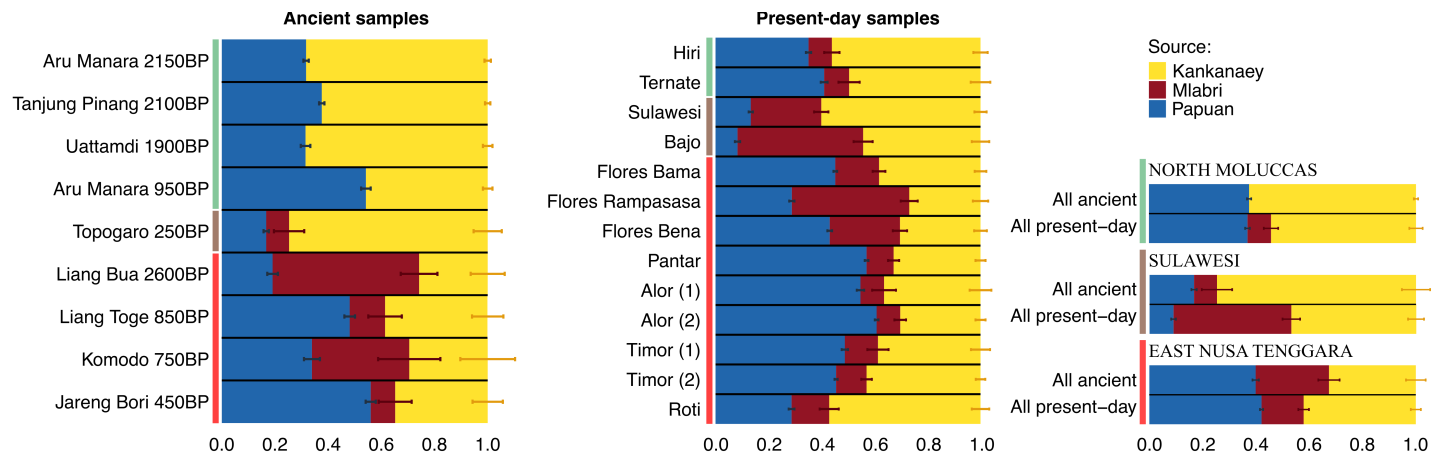
Figure 3 – Ancestry proportions for the model with the highest *P*-value in each group. Individuals from the North Moluccas, Sulawesi, and East Nusa Tenggara are marked with green, brown, and red vertical bars, respectively. Horizontal bars show standard errors. The number following the name of the present-day individuals indicates the genotyping array used: (1) Affymetrix 6.0 and (2) Affymetrix Axiom Genome-Wide Human.
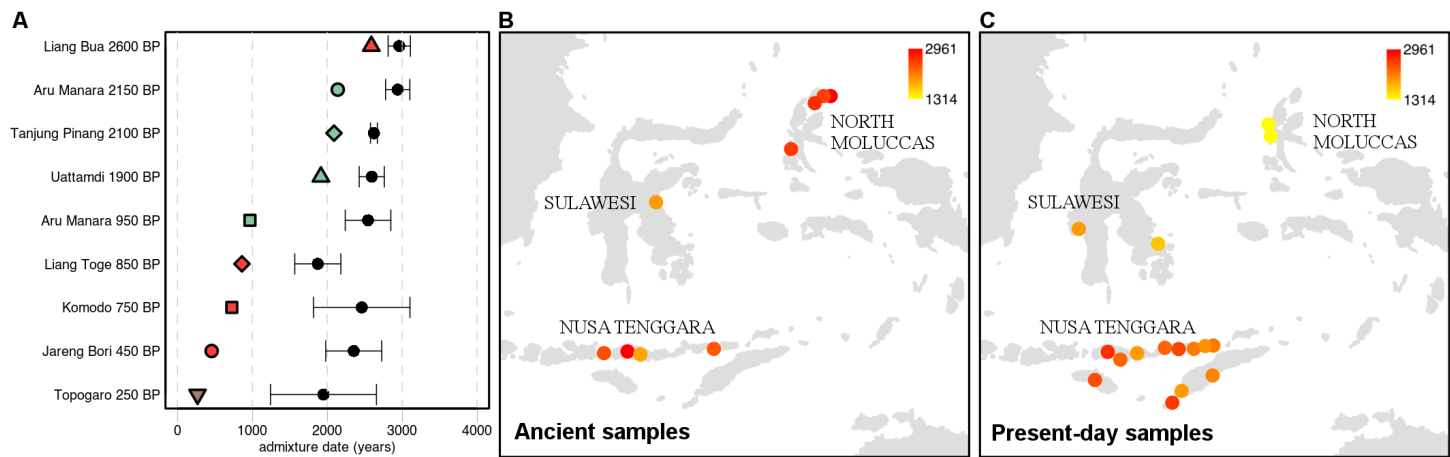
Figure 4 – Admixture estimates. A) Admixture date point estimates and standard errors are shown in black. The individuals age is indicated with filled symbols: green – North Moluccas, brown – Sulawesi, red – East Nusa Tenggara. B-C) Spatial distribution of admixture date estimates for ancient and present-day individuals, with admixture dates depicted according to the heat plot.