An ancient genomic perspective on the human dispersals to tropical islands – implications for the settlement history of the ancient Caribbean and the Pacific.

Dissertation
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Kathrin Nägele
Aus Heilbronn am Neckar

Tübingen
2020

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	09/06/21
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Johannes Krause
2. Berichterstatter:	Prof. Dr. Christopher Miller
3. Berichterstatter:	Prof. Dr. Lluis Quintana-Murci

Index

1. ABBREVIATIONS

3'	three-prime end of the DNA strand
5'	five-prime end of the DNA strand
aDNA	ancient DNA
aDNS	alte DNS
BP	years before present
°C	Degree Celsius
DNA	Deoxyribonucleic Acid
DNS	Desoxyribonucleinsäure
km	kilometers
kns	Knots (1 nautical mile per hour)
mtDNA	mitochondrial DNA
NGS	Next Generation Sequencing
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PNG	Papua New Guinea
RNA	Ribonucleic Acid
SNP	Single Nucleotide Polymorphism
UDG	Uracil DNA Glycosylase

2. SUMMARIES

2.1 Zusammenfassung

Die hier vorgelegte Arbeit nutzt neue Techniken in der Beprobung und Aufbereitung alter DNS (aDNS). Dank fortschrittlicher Sequenziertechnologien ist es heute möglich die genetische Information lange Verstorbener zu analysieren und mit anderen (bio-) archäologischen und genetischen Ergebnissen zu einem detaillierteren Bild der Menschheitsgeschichte zusammenzuführen. Erst zu Beginn der Arbeit an den im Nachfolgenden zusammengefassten Manuskripten wurde es durch die gezielte Beprobung des Pars petrosa, dem das Mittelohr enthaltenden Knochen, möglich, diese Information auch aus tropischen Regionen zu gewinnen, wo hohe Temperaturen und Luftfeuchtigkeit die Zersetzung der Erbmasse beschleunigen.

Die in dieser Arbeit zusammengefassten Studien beschäftigen sich mit der Besiedelungsgeschichte zweier Weltregionen, die zu den spätesten besiedelten Gebieten der Menschheitsgeschichte gehören: Der Karibik und des Pazifiks. In beiden Fällen stellten sich trotz oder gerade durch die Rekonstruktion der Besiedelungsgeschichte mittels anderer bioarchäologischer Methoden Fragen, zu deren Beantwortung die archäogenetischen Analysen beitragen sollen.

Die Karibik war die letzte durch den Menschen besiedelte Region der Amerikas, erfuhr jedoch als erste die Alles umwälzenden Konsequenzen des Kolonialismus. Bereits in den ersten 100 Jahren nach Entdeckung des amerikanischen Kontinentes hatte sich die kulturelle, linguistische und genetische Landschaft der Karibik unwiederbringlich verändert. Analysen moderner DNS, wie häufig in anderen Gebieten unternommen, konnten ebenso wie linguistische Analysen hauptsächlich zur Rekonstruktion der Kolonialgeschichte eingesetzt werden. Die umfassenden Analysen des archäologischen Befunds zeigen zwar einen prä-kolonialen Reichtum an Keramiken und Lebensweisen, sind jedoch vielfach von den Berichten der kolonialen Chronisten beeinflusst.

In **Manuskript A** (Nägele et al. 2020, *Science*) ergänzen wir den Wissensstand durch die Analyse alter DNS von Menschen die zwischen 3200 und 500 Jahren vor heute in der Karibik lebten. In der vergleichenden Analyse der neu generierten Genome mit bereits publizierten Genomen aus den Amerikas konnten wir zeigen, dass zwei vom archäologischen Befund als Gruppen verschiedener Einwanderungsbewegungen

identifizierte Gruppen auch genetisch unterscheidbar sind. Wir bestätigen die Herkunft der jüngeren Einwanderung im nordöstlichen Südamerika und zeigen, dass diese wahrscheinlich über die kleinen Antillen nordwärts geschah. Die ältere Gruppe, vor allem auf Kuba lebende Jäger und Sammler Gesellschaften, warf bislang die meisten Fragen auf. Unklar war woher diese Gruppe kam, und ob es sich um eine genetisch einheitliche, auf einen gemeinsamen Ursprungsort zurückzuführende Gruppe handelte. Unsere Analysen zeigen, dass bereits vor 3000 Jahren die Menschen Kubas die Herkunft zweier verschiedener Regionen zeigen. Zwar können wir den Ursprungsort nicht eingrenzen, jedoch finden wir Verbindungen eines Individuums zu alten Individuen in Kalifornien. Nordamerika sollte demnach nicht als ein möglicher Ursprungsort ausgeschlossen werden. In allen anderen untersuchten Individuen dieses Kontextes findet sich zusätzlich noch eine südamerikanische genetische Komponente, die sich von der späterer Einwanderer unterscheidet. Unsere Analysen zeigen, dass die Karibik mehrfach besiedelt und wiederbesiedelt wurde, bereits durch die frühen Jäger und Sammler.

Die Manuskripte B und C befassen sich mit den Details der Besiedelungsgeschichte des Pazifiks. Sie bauen auf einer vorhergehenden, archäogenetischen, Studie auf welche zeigte, dass die ersten Siedler im entfernten Pazifik asiatischen Ursprungs waren. Diese wurden der archäologisch definierten Lapita Kultur zugeordnet, welche vor ca. 3250 Jahren im Bismarck Archipel entstand und kurz darauf erstmals Fern-Ozeanien besiedelte. Offen blieben die Fragen, wie, woher und wann die heute in Menschen Fern-Ozeaniens zu findende papuanische genetische Komponente in die Region kam. In Manuskript B (Posth, Nägele et al. 2018, *Nature Ecology and Evolution*) analysierten wir hierfür die DNS alter, aber auch heute lebender Menschen aus Vanuatu, der ersten besiedelten Inselkette Fern-Ozeaniens, um die Entwicklung der genetischen Zusammensetzung zu verstehen. Die Ergebnisse zeigen, dass die in heutigen Bewohnern Vanuatus, die die Selbstbezeichnung ni-Vanuatu wählen, dominierende papuanische Herkunft bereits vor 2600 Jahren durch Menschen, vermutlich aus dem Bismarck Archipel, nach Vanuatu kam. Es handelte sich hierbei nicht um eine substanzielle Einwanderung, welche die lokale, genetisch vorwiegend ostasiatische Population, ersetze, sondern um eine über hunderte Jahre anhaltende, graduelle Zuwanderung. Durch diese langsame Änderung der genetischen

Zusammensetzung lässt sich auch die gleichzeitige Erhaltung der von den ersten Siedlern gesprochenen austronesischen Sprache erklären.

Da der Ursprung der in Vanuatu vorherrschenden papuanischen Herkunft im Bismarck-Archipel nur durch moderne Populationen angenähert werden konnte, befasst sich Manuskript C (Nägele et al. *in prep*) mit der genetischen Diversität Nah-Ozeaniens. Anders als Fern-Ozeanien wurde dieses bereits vor 45 000 – 50 000 Jahren erstmals vom Menschen besiedelt. Bis vor kurzem wurde angenommen, dass die Lapita Kultur das Festland Neu Guineas niemals, oder nur weit vorgelagerte Inseln besiedelt hatte. Dies änderte sich mit der Entdeckung einer Fundstelle an der Südküste Papua Neu Guineas, in der sich die für die Lapita Kultur typische Keramik auf 2900 BP datieren ließ. Unklar ist jedoch, ob und wann die Menschen der Lapita Kultur auch genetische Spuren auf dem Festland hinterlassen haben, und ob auch im Bismarck Archipel eine Vermischung von papuanischer und ostasiatischer Herkunft geschah bevor die Inseln Fern-Ozeaniens besiedelt wurden. Die archäogenetische Analyse von 41 Individuen die vor 3700 – 150 Jahren in Papua Neu Guinea und dem Bismarck Archipel lebten, zeigt, dass die untersuchten Individuen an der Süd- und Nordküste eine ostasiatische genetische Komponente zeigen, die derer der ersten Sieder Fern-Ozeaniens am ähnlichsten ist. Die Zusammensetzung unterscheidet sich jedoch zwischen den verschiedenen Fundstellen, was auf unterschiedliche Interaktionen mit Populationen im Inland und den Küsten hindeutet. Die Vermischung der beiden Komponenten lässt sich auf 1500 – 1000 Jahre vor heute datieren, mehr als 1500 Jahre nach dem erstmaligen auftreten der Lapita Kultur in der Region. Die späten Daten der Vermischungsereignisse lassen darauf schließen, dass die verschiedenen Kulturen Jahrtausende nebeneinander gelebt haben, ohne sich genetisch zu vermischen.

Im Bismarck Archipel zeigen die genetischen Daten von 5 Individuen von Watom Island, welche einen großen Zeitintervall von 3000 Jahren abdecken, dass die Einwohner Watoms vor 3700 und 2600 Jahren papuanischer Herkunft waren. Ein vor 2100 Jahren lebendes Individuum zeigt eine Vermischung mit ostasiatischer Herkunft auf, datiert auf 2300 Jahre vor heute. Die geringe Anzahl von Individuen und die schlechte Abdeckung ihrer Genome mahnen zu einer vorsichtigen Interpretation, doch scheint es als sei die in Mauskript B und vorangegangenen Studien getroffene

Aussage, dass die ersten Siedler asiatischen Ursprungs waren und sich erst in Fern-Ozeanien mit Menschen papuanischer Herkunft mischten, vorerst bestätigt.

Zusammenfassend tragen die hier vorgestellten Studien nicht nur zum besseren Verständnis der Besiedelungsgeschichte der jeweiligen Region bei, sondern vermitteln auch ein anderes Bild der Seetauglichkeit vergangener Populationen. Jäger und Sammler Gesellschaften werden gemeinhin nicht für gute Seefahrer gehalten, was die Ergebnisse und Vorhersagen zur Ausbreitung der Menschen auf dem Planeten beeinflusst. Sowohl die Jäger und Sammler Amerikas, als auch die Neu Guineas waren wohl in der Lage große Wasserflächen zu überqueren. Die traditionelle Sichtweise auf Wasserflächen und Ozeane als Barriere sollte im Kontext menschlicher Ausbreitungsereignisse überdacht werden und die hier gezeigte verbindende Funktion in die Fragestellungen zur Menschheitsgeschichte einbezogen werden.

2.2 Summary

This thesis uses new techniques in the sampling and processing of ancient DNA (aDNA). Next-generation sequencing technologies have accelerated the production of ancient genomes and allow analysing the genetic information of people who lived in the distant past. Together with other (bio-) archaeological and modern genetic results, the analysis of ancient genomic sequences allows a more detailed reconstruction of human history. Just at the outset of this thesis, the targeted sampling of the petrous part of the temporal bone, the cranial bone harbouring the inner ear, facilitated the recovery of ancient genomes from tropical regions. There, high temperatures and humidity expedite the decay of DNA molecules.

The studies combined in this thesis focus on the settlement history of two regions. Both the Caribbean and the Pacific are among the last regions to be settled by humans. In both cases, the reconstruction of human history from results in other disciplines has left or even led to open questions. Archaeogenetic methods can add detail to the reconstructions and answer basic questions.

The Caribbean was the last region of the Americas to be settled by humans, yet the first to experience the drastic impact of European colonialism. Within the first century after invasion of the American continents, the linguistic, cultural and genetic landscape had been irrevocably changed. Unlike in other regions, analysis of present-day genomes and linguistic variation was only useful in reconstructing the colonial past.

The comprehensive analysis of archaeological contexts showed a rich pre-colonial variety of ceramics and lifestyles, but the reports of the European chroniclers often influence the interpretations.

Manuscript A (Nägele et al. 2020, *Science*) complements the archaeological evidence with the analysis of ancient genomes of people who lived in the Caribbean between 3200 and 500 years before present (BP). In the comparison of the newly generated sequences with published ancient and present-day sequences from the Americas, a difference in genetic ancestry is revealed, consistent with two distinct archaeological contexts associated with different dispersals into the Caribbean. We support the synthesis that the more recent dispersal originated in northeastern South America, and conclude that they dispersed northwards, through the Lesser Antilles. Most questions,

however, concern the older group, fisher-hunter-gatherer societies inhabiting the Greater Antilles. From (bio-) archaeological analysis, it remained unclear where the dispersal originated and whether it was a single dispersal with one origin. Our analyses show that, already around 3000 BP, two ancestries were present on the island, connected to populations in different regions today. Although we fail to pinpoint the origin of the genetic ancestries, one individual reveals a connection to ancient individuals in California, suggesting not excluding North America as a possible place of origin. All other individuals show an additional, South American component, different from the ancestry of the later dispersal from northeastern South America. We conclude that the Caribbean has been settled and resettled multiple times, already by fisher-hunter-gatherer communities.

Manuscripts B and C focus on details of the settlement history of the Pacific. They build on previous archaeogenetic research, in which the early settlers of Remote Oceania, associated with the archaeological Lapita culture, reveal almost exclusively East Asian-related ancestry. Unanswered remained the question regarding the timing and mode of mixture with Papuan-related ancestry, present in all Pacific Islanders today.

Manuscript B (Posth, Nägele et al. 2018, *Nature Ecology and Evolution*) analyses ancient and present-day human genomes from Vanuatu, the first archipelago settled in western Remote Oceania. Analysing the development of the genetic composition through time, we show that present-day inhabitants of Vanuatu, self-identifying as ni-Vanuatu, have almost exclusively Papuan-related ancestry, which was introduced already 2600 BP by people originating most likely in the Bismarck Archipelago. Our time-transect shows that in contrast to one migration, the shift from exclusively East Asian-related ancestry in the first settlers to the almost exclusively Papuan-related ancestry today happened gradually over centuries. This slow change allows integrating the genetic turnover with linguistic evidence for preservation of Austronesian language, which was reconstructed for the first settlers.

As the Papuan-related ancestry was identified through present-day populations serving as a proxy to the ancient populations, Manuscript C (Nägele et al. *in prep*) aims to understand the ancient genetic diversity in Near Oceania, which had been settled by humans already 45,000-50,000 BP. Until recently it was assumed that the Lapita

culture omitted the mainland of New Guinea, settling only offshore islands. The discovery of a 2900-year-old site with Lapita pottery in Caution Bay, on the south coast of PNG, has changed this view. However, it is still unclear if and when the Lapita culture left genetic traces on the mainland. Additionally, there is doubt regarding the interactions of the Lapita-associated people with Indigenous populations on the Bismarck Archipelago, possibly leading to a mixed population settling western Remote Oceania. The analysis of 41 individuals from Papua New Guinea and the Bismarck Archipelago, dated to 3700 – 500 BP, shows that all individuals from the southern and northern coasts of Papua New Guinea harboured an East Asian-related component, most similar to those of the early Remote Oceanians. The composition and different timing of the admixture events in the different sites suggests complex interactions with inland and coastal populations. The admixture event was inferred to around 1500 BP, 1400 years after the first occurrence of the Lapita culture in the region. This late date implies either a repeated admixture as observed in Vanuatu, or parallel societies for millennia, without genetic exchange.

The five individuals analysed from the Bismarck Archipelago cover a timeframe of 3000 years. The two oldest individuals show exclusively Papuan-related ancestry. One individual dated to 2100 BP shows admixture with Asian-related ancestry around 2300 BP. The low coverage and small amount of individuals ask for cautious interpretation, but it seems that the statement in Manuscript B, regarding the settlement of western Remote Oceania by genetically East-Asian people, can be supported. The admixture event in the Bismarck Archipelago postdates the initial settlement of Vanuatu and Tonga and happens, similar to the one on the mainland, a millennium after the first occurrence of the Lapita cultural complex in the islands.

In summary, the studies presented here add to a better understanding of the settlement history of the respective regions, but also to our understanding of seafaring capabilities in ancient times. Traditionally, hunter-gatherer- communities are not known as great navigators, and models of human dispersal on the planet have favoured land routes for those communities in the past. Hunter-gatherer communities of the Americas, as well as in the Pacific, have been shown to cross large bodies of water to settle islands and interact with island populations. Moving forward, the connecting

nature of bodies of water should be more seriously considered in ideas about the dispersal of humans across the world.

3. LIST OF PUBLICATIONS AND MANUSCRIPTS

3.1 published manuscripts included and discussed in the dissertation

MANUSCRIPT A:

K. Nägele*, C. Posth*, M.I. Orbegozo, Y. Chinique de Armas, S.T. Hernandez Godoy, U.M. González Herrera, M.A. Nieves-Colón, M. Sandoval-Velasco, D. Mylopotamitaki, R. Radzeviciute, J. Laffoon, W. J. Pestle, J. Ramos-Madrigal, T.C. Lamnidis, W.C. Schaffer, R.S. Carr, J.S. Day, C. Arredondo Antúnez, A.R. Rivero, A.J. Martínez-Fuentes, E. Crespo-Torres, I. Roksandic, A.C. Stone, C. Lalueza-Fox, M. Hoogland, M. Roksandic, C.L. Hofman, J. Krause, H. Schroeder (2020) "Genomic insights into the early peopling of the Caribbean" in *Science* (369) pages 456-460

MANUSCRIPT B:

C. Posth*, K. Nägele*, F. Valentin, S. Bedford, K.W. Kami, R. Shing, H. Buckley, R. Kinaston, M. Walworth, G.R. Clark, C. Reepmeyer, J. Flexner, T. Maric, J. Moser, J. Gresky, L. Kiko, K.J. Robson, K. Auckland, S.J. Oppenheimer, A.V.S. Hill, Jana Zech, F. Petchey, P. Roberts, C. Jeong, R.D. Gray, J. Krause, A. Powell (2018): "Language continuity despite population replacement in Remote Oceania" in *Nature Ecology and Evolution* (2), pages 731 – 740

3.2 draft manuscript included in discussed in the dissertation

MANUSCRIPT C:

K. Nägele, R. Kinaston, S. Carlhoff, D. Gaffney, E. Bertolini, M. Tromp, R. Radzeviciute, G. Summerhayes, F. Petchey, D. Anson, P.Petchey, H.Buckley, J. Krause, C. Posth, A. Powell. "Ancient Genetic Diversity in Near Oceania - insights from coastal New Guinea and the Bismarck Archipelago." Draft Manuscript.

4. OWN CONTRIBUTIONS

4.1 MANUSCRIPT A

I performed the lab work for 22 of the published samples from sampling to sequencing. For 40 samples contributed by collaboration partners in Copenhagen, I performed the lab work from building libraries through sequencing. Under the supervision of Cosimo Posth, professor at the University of Tübingen, I performed population genetic analysis and reconstructed the mitochondrial genome. Together with Hannes Schroeder, group leader at the Globe Institute, University of Copenhagen, I wrote the manuscript, with critical input from the other authors involved in the study. I produced the figures and tables for the final publication. I coordinated the Cuban and Puerto Rican collaborations, permit and sample acquisition.

4.2 MANUSCRIPT B

Under the supervision of Cosimo Posth, professor at the University of Tübingen, I performed lab work for all ancient samples included in this study from sampling to sequencing and amplification for capture, for which I also included the 27 modern genomes. I reconstructed the mitochondrial genome and determined the uniparental haplogroups. I wrote parts of the manuscript and supplementary information, generated part of the supplementary tables and edited the manuscript.

4.3 MANUSCRIPT C

Together with Emilie Bertollini, whom I supervised, I performed the sampling of petrous bones from Watom Island. I conducted the population genetic analysis and reconstructed the mitochondrial genomes. I wrote the manuscript with input from Rebecca Kinaston, Postdoc at the University of Otago, Dyllan Gaffney, PhD student at the University of Cambridge, Glenn Summerhayes, professor at the University of Otago, and Cosimo Posth, professor at the University of Tübingen. I produced all figures and supplementary sables and compiled the supplementary information.

5. INTRODUCTION

Paleoanthropologists and archaeologist have investigated questions of human pre-history for centuries and shaped our ideas about and understanding of human evolution. The young field of archaeogenetics utilises genetic analysis of ancient human remains to contribute to answering the questions posed by archaeology. Ancient DNA has proven a useful tool in complementing archaeological methods and answering open questions in human history. Not only has it shown a close relation to and genetic exchange with Neanderthals (*1*), but enriched the human family tree with the Denisovan (*2*), who also contributed to the genetic make-up of present-day humans (*3*). The dating of the split of modern humans from other human forms (*4*) shows that still much can be discovered about the human past, evolution and dispersals through future fossil findings. Through ancient DNA, some essential questions in human history are better understood today, and human history can be tied better to the present. In many cultures, identity is connected to ancestry (*5*). By understanding their roots, people aim to understand the present and "who they are", despite the sometimes clashing genetic ancestry and social identities (*6*). To some, like the descendants of enslaved people, it can be a means of battling a trauma passed down for many generations and seeking the roots so violently pulled out (*7, 8*). To others, the question is connected to the challenging of narratives erasing their culture and history (*9, 10*). Made possible through recent advances in technology, this thesis centres the ancient inhabitants of tropical islands in the Caribbean and the Pacific. Both regions have been highly impacted by European invasion, although in different ways and extents. As modern genetic and linguistic analyses have limited power to help reconstruct the human history of places impacted by colonialism, ancient DNA analysis can provide important perspectives to add to a more detailed reconstruction of the past.

5.1 The history of ancient DNA – overcoming problems and refining methods.

Since the emergence of archaeogenetics in the 1980s (*11, 12*), and through further refinement in the decades after its discovery, the use of ancient DNA (*13*), recovered from archaeological remains, has proven a robust line of evidence in reconstructing the

past. Recovering ancient DNA fragments revolutionised the field of archaeology once again, as did the introduction of radiocarbon dating (*14*). However, like radiocarbon dating, ancient DNA recovery and analysis have pitfalls, impacting the quality of analysis and the certainty of interpretation. Both technologies have undergone a refinement of the methods to overcome the problems mostly affecting data quality, leading to reliable data produced today. Ancient DNA can today successfully be retrieved from a variety of archaeological and palaeontological remains such as skeletal tissue (*11, 12*), hair (*15*), dental calculus (*16*) mummified soft tissue (*17*), plants (*18, 19*), coprolites (*20*), and most recently from sediment (*21*). However, DNA recovery does have limitations. Claims of DNA retrieved from million-year-old fossils (*22-25*) today are regarded as contaminants (*26-29*). To date, a horse, found in the best possible preservation conditions in permafrost, yielded the oldest DNA sequence, 700,000-year-old (*30*).

DNA molecules have limited chemical stability. Their fragmentation starts right after death in the absence of enzymatic repair mechanisms (*31*) and the body's enzymes expedite this process (*32*). During decomposition, DNA is further fragmented through microbial digestive processes (*31, 33*), leading to a strand length of 400 bp shortly after a persons or animals death (*31*). After the decomposition of soft tissue, DNA in bones, hair and the soil degrade further. The DNA continues to destabilise through depurination. In the presence of water, the β-N-glycosidic bond between the DNA backbone and the purine bases is cleaved, leaving the site abasic and more prone to strand breaks, leading to additional fragmentation (*34, 35*). As a result, aDNA molecules have a very short fragment length of on average 40bp. The rate at which DNA fragments varies and depends on several factors. While time is an essential factor (*36, 37*), other environmental conditions such as soil acidity, mineral content and macro-and microclimate (*37*) can significantly influence the preservation, complicating the predictability of DNA preservation. In tropical environments, the two most apparent factors are humidity and temperature, both aiding hydrolytic reactions, and driving the fragmentation and damage of DNA molecules.

Regions with high temperatures and high temperature fluctuations generally show worse DNA preservation (*37, 38*), while in regions with stable cold and dry conditions DNA as old as 700,000 years can be recovered (*30*).

Another typical damage on DNA is depurination. With time, and expedited by high humidity and temperatures, a rise in deaminated sites towards the ends of DNA fragments (*39*) can be observed. In a hydrolytic reaction the base cytosine loses an amino group, resulting in uracil, a base in vivo only observed in RNA. During PCR based sequencing or the stabilisation in libraries for sequencing, a thymine is misincorporated in place of the uracil. This results in a higher CG content in the resulting sequence, potentially leading to erroneous results if not taken into account.

Nevertheless, while DNA damage complicates the recovery and analysis of ancient DNA, it is pivotal for authentication. Ancient DNA is extracted from elements exposed to their environment by digesting the proteins, releasing the DNA molecules (*35*). As this process in unspecific, the extract itself is a mixture of the DNA of different organisms present in and on the sample, mirroring its environment. Ancient human DNA is abundant only in meagre copy numbers, while environmental contaminations such as bacterial, viral, fungal, plant and contemporary human DNA dominate the extract (*40-42*). To overcome low copy-numbers, the initial studies used bacterial cloning (*11, 12*) but were difficult to reproduce, and the targeted DNA amplification through PCR was soon favoured (*43, 44*). However, as PCR is only able to amplify long DNA fragments, while the inherent nature of ancient DNA is short fragments (*45*), this method biased the amplification towards longer molecules, most likely deriving from contamination (*46*). Efforts to reduce said contamination, such as decontamination procedures, separate pre- and post-amplification facilities and the use of negative controls (*47, 48*) improved the quality of the ancient DNA data produced. Today, contamination still poses the biggest threat in the analysis of ancient DNA, despite the efforts to control it before sequencing. Modern DNA does not exhibit the typical damage pattern of ancient DNA, and usually has higher molecule sizes. Bioinformatical tools have made it possible to detect (*49-51*) and remove contamination (*52*), making use of the authentic patterns in ancient DNA.

The advent of next-generation sequencing methods (NGS) in 2005 would prove to revolutionise the field of genetics in general, but even more for ancient DNA studies, only two decades from the first publications of ancient DNA sequences. The power of NGS comes from the parallel sequencing of millions of untargeted molecules (*53, 54*), reducing the costs while significantly increasing the sequencing throughput and

allowing to investigate all molecules in the sample. DNA fragments are unstable for the reasons mentioned above. To stabilise them in preparation for NGS, they are turned into so-called libraries. Universal adaptors are attached to both the molecules' ends in an extract. In addition to stabilising, the adapters function as priming sites for subsequent amplification (*55*). Most commonly this library is double-stranded, consisting of forward 3' and reverse 5' strands of DNA, complementing each other. The independent processing of both strands, using the so-called single-stranded library protocol, results in a more efficient conversion, retrieving shorter fragments (*56, 57*). As NGS allows parallel sequencing of many libraries, unique short sequences, so-called "indexes", are added to the adaptors to differentiate the reads from other libraries and possible laboratory contamination (*58*). Bioinformatical tools for statistical analysis allow quantification of endogenous DNA and authentication through the investigation of DNA damage (*59*). However, the sequencing of entire libraries containing not only the desired ancient human DNA reads but also reads from the environmental contamination is costly, especially when samples show bad DNA preservation. The development of enrichment techniques made it possible to push the boundaries of DNA recovery once again. These assays can be designed for any of the contents, be it of human or microbial origin, and produce sensible data from very old remains or samples from regions with unfavourable conditions such as the tropical regions. By using hybridisation techniques with probes designed to complement known DNA sequences, the target DNA is captured, while environmental contamination can be washed out (*60-62*). The remaining, targeted reads can be sequenced, resulting in lower costs while increasing the chance of covering more of the ancient genome.

The research in this thesis focuses on tropical regions. As described above, these climates pose particular difficulties when attempting to recover ancient DNA. Using the latest technologies paired with targeting skeletal elements with high DNA preservation (*63, 64*), ancient genomes from highly fragmented and damaged molecules can be reconstructed, adding another line of evidence to the understanding of the population history of the Caribbean and the Pacific.

5.2 Human mobility in the past and the role of ancient DNA - shedding new light on old questions.

Human mobility has different categories implying different modes of mobility (*65*). The term migration implies intent and an individual pace of mobility, moving over more significant distances within one generation. The notion of intent comes with certain connotations; genetics alone cannot identify whether a journey happened with intent and an aim. The term dispersal lacks these connotations and leaves the intent and pace open. Expansions can be intentional and over vast distances, but can also be unintentional, moving only a few kilometres within one generation.

Dispersals have been a major source of tension in the archaeological discourse, mirrored in the often-debated question whether cultural change came through the diffusion of new technologies and ideas, limiting the interaction of people to their neighbours and cultural exchange, or whether the interaction continued over larger spatial spheres and extended to genetic exchange. In many cases, big narratives about cultural transitions through the movement of people were based on observations of similarities and changes in the material culture, identifying a source and direction of change (*66-68*). Comparative analysis of archaeological material can be one line of evidence in identifying ancient dispersals or migrations, but when it is the only method used, it leaves to many alternative scenarios and explanations. Another valuable line of evidence can be studying strontium isotopes. The analyses can reveal mobility during the lifetime of a person and has been used on a multitude of archaeological contexts (*54*). The isotopic signature of a person's environment is built into the enamel, hence differences between the isotopic profile of an individual and the environment they were buried in, shows this person moved from their childhood home to the place they were laid to rest. However, isotopic signatures are not unique, and the same profile can be consistent with various geographical regions. Therefore, it is difficult to identify the exact place of origin or the range of mobility.

Although ancient DNA cannot identify the exact geographical place of origin of one person, it has shown to be a useful tool in identifying dispersals through a change in the genetic composition of people within a region. Additionally, the genetic profile of a person can be compared to other ancient and present-day populations and can,

therefore, serve as yet another line of evidence, narrowing down the place or region of origin. Most famously, the long-standing question of whether the Neolithic revolution in Europe was started by cultural diffusion (pots) or by the migration of people from the Fertile Crescent to central Europe (people), was settled through aDNA analysis. Genetically the hunter-gatherer-populations inhabiting Europe in the Mesolithic period were largely genetically replaced by people with Near Eastern ancestry (*69-72*). Apart from contributions to this archaeologically very well studied phenomenon, ancient DNA analysis of 69 European individuals living between 8000 and 4000 BP surprisingly showed a dispersal from the Caspian steppe into Europe. This dispersal introduced the third ancestry component leading to the trifold ancestry of hunter-gatherer, near-eastern farmer and steppe ancestry of the majority of present-day Europeans (*73*), revealing a dispersal not inferred from the archaeological record.

The genetic analysis of an ancient Siberian genome showed the genetic relation of Native Americans to Asians (*74*), excluding ideas of an initial settlement from Europe, as implied by some analysis of stone tools (*75*), or from the Pacific Islands, as a morphological study on a 9000-year-old skull from Washington suggested (*76*).

5.3 The settlement of islands – an evolutionary and a human perspective

The inherent nature of islands (as defined in (*77*)) is to be separated from the mainland by large bodies of water. While isolation is the characteristic trait, islands can form in two fundamentally different ways: They can either arise out of the sea as a result of volcanic activity or a lowering of sea levels. Alternatively, they can form from an existing mainland by incursions, such as continental drift or the rise of sea-levels. In the latter case, flora and fauna of the once united mainland will be found on the island, while in the former case the island will be devoid of terrestrial species. The question of how species with low mobility settle these islands, and how the islands' nature affects speciation processes, have produced an astounding body of research and introduced exciting evolutionary concepts specific to island biogeography (*78*).

The founder principle (*79*) conveys that specimens settling in an island will be a subset of a bigger population, and hence the genetic variation will be a subset of the original

population. This founding effect and the genetic variation of the founding population will be dependent on the number of individuals involved. In a bottleneck event, the founding population experiences a sharp decline in population size, reducing the genetic variability significantly. Although initially smaller, the genetic variation of such a population, isolated from the source population, can increase through mutation and re-sorting. While its importance for island speciation is still under debate (*80, 81*), genetic drift, describing an alteration of allele frequencies by chance (*82*) is one form of genetic re-sorting. It is usually observed under low population sizes, theoretically promoting a rapid shift to a new combination of alleles, while reducing the number of heterozygote sites (*83-85*). Although generally it is assumed that high genetic variation is linked to high fitness, some argue that the narrowing of a gene pool during a bottleneck event will lead to higher adaptive capability by breaking up old adaptive complexes, increasing the speed of speciation (*86-88*). This genetic release can be paired with an ecological release, meaning the absence of predators or competitors leads to phenotypic changes when "liberated" from selective pressures (*89*), as would be for example the loss of defensive traits in the absence of predators. These factors may have led to what has been dubbed the "Island Rule" (*90-92*). In island taxa of mammalian orders, passerine birds, lizards and turtles, unusual sizes have been reported, from gigantism of small (*93, 94*) and dwarfism in large (*95*) mainland species on islands. *Homo floresiensis* (*96, 97*) might be a human example of the island rule, although it is the only example in hominins thus far. A general problem in the comparison and study of island and mainland species is the unclear locational and historical context. Island forms of species can be an evolution from mainland species or represent a more basal form, where the island served as a refuge where selective pressures were absent, while mainland populations were replaced or changed due to selective pressure (*98, 99*).

Of high interest in the biogeographical community is the geographical source from which species settle an island. As mentioned above, island populations often change their morphology drastically compared to the source population, limiting the reliability of morphological studies. Animals and plants arrive on islands by drifting with currents or winds (*100-102*), allowing an approximation of migration routes and inference of origins.

Considering the topic of island biogeography in the context of humans and human evolution is especially interesting. Logical assumptions derived from the studies of plant and animal species would suggest that close islands were colonised before distant, large before small islands and that the populations on the adjacent mainland settled the nearest islands, aided by winds and currents (*103*).

However, unlike animals, humans have demonstrated forethought and intention, allowing them to navigate against prevailing winds and currents (*104, 105*). Additionally, cognitive capabilities allow humans to adapt to their environment, regardless of genetic predispositions (*106, 107*), although the settled environments can affect the genetic make-up by driving genetic adaptation (*108*). Where animal populations cannot be self-sustained on small islands, humans can trade with nearby islands or mainland for essential resources (*109*). Nevertheless, islands present technological and ecological challenges for long-term human settlement, partly caused by the impoverished ecosystems, and require technological and cultural innovation. When investigating the human population histories of islands, the same overall frameworks of island biogeography provide a useful basis, although with a great amount of variation, through space and time (*110, 111*).

The settlement and resettlement of islands by humans is directly connected to questions about their ability to cross bodies of water, teaching us about mobility in ancient times. However, ancient DNA studies of island populations are scarce. The majority of islands on our planet are situated along the equator, where preservation conditions are poor. The few analyses from higher latitudes show a diversity of scenarios, ranging from strong isolation and drift to a high connectedness with the mainland source populations. The ancient Jomon lived isolated in the Japanese Archipelago since 16 000 BP (*112*) until the introduction of rice farming 3000 BP (*113*). Ancient genomic analyses of Jomon-related individuals have shown that the present-day Indigenous population of Japan, the Ainu, are descendants of the Jomon, (*114-116*) confirming the isolation for millennia.

Ancient DNA analysis of populations in the Mediterranean have shown that at least since the Bronze Age, the ocean surrounding them did not isolate the islands, but they were similarly impacted by the Steppe expansion observed in mainland populations (*117*). Similarities in the monumental towers build on Mallorca, Menorca and Sardinia

even suggested cultural connections between the islands (*118*), supported by a genetic similarity of the ancient islanders (*117*). The inhabitants of Crete around 4000 years BP, mystically known as the Minoans, certainly seemed to have been connected to the contemporary mainland population in Greece, known as the Mycenaeans (*119*).

5.4 The settlement of the Caribbean – reconstructing the connections to the mainland

The insular Caribbean was one of the last regions of the Americas to be settled by humans. The dispersal over the almost 3 million km^2 of open water remains as poorly understood as it is a remarkable achievement. Archaeological evidence shows occupational sites dating to the Archaic Age of the Americas as early as 8000-5000 BP in the extreme ends of the archipelago. The oldest site on the Caribbean Islands is the Banwari Trace in Trinidad (*120*), the only island that, at the time, was still connected to the American mainland of what is now Venezuela. The connection to the continents facilitated settlement compared to the other islands, which remained isolated throughout human history. Apart from the Banwari trace, the oldest dates obtained are in the northern Caribbean Islands, such as Cuba, Puerto Rico, St.Martin and Anguilla (*121*). Evidence from the Windward Islands is restricted to proxies of human presence through pollen analysis and supposedly anthropogenic fires (*122*), but the difficulties in linking burned materials to human activity (*123*) ask for a cautious interpretation of this data. European invasion in the late 15^{th} century had dramatically changed the cultural, linguistic and genetic landscape of the region, making reconstructions of the pre-contact Caribbean population history based on those lines of evidence close to impossible. Archaeological analyses have recovered diverse contexts (*124, 125*), reconstructing a cultural diversity on the various islands. However, attempts to connect the early populations of the insular Caribbean to those on the American mainland relied on stylistic comparisons of stone tools and were not able to trace back to a certain point of origin (*126, 127*). Through a high diversity within Archaic Age sites, comparisons showed connections to sites in North America (*128, 129*), South America (*130, 131*) and Central America (*126, 127*). Although simulations of prevailing winds and ocean currents (*132*) and recent analyses of radiocarbon dates (*121*) and cranial morphometry (*133*)

foregrounded South and Central America as the origin of the Archaic Age dispersals into the Caribbean, the origins and routes of Archaic Age dispersals into the region remain unresolved.

More than 2500 years after the first settlement of the islands around 7000 – 5000 BP, new people entered the region, heralding the Ceramic Age of the Caribbean. In the archaeological record, they are recognised by ornate pottery with similarities to the pottery found in the Saladero site in Venezuela (*130*) but show regional variations. This dispersal is well studied, and archaeological, linguistic and genetic evidence is in agreement with an origin in the Orinoco River Delta in the northern Amazon (*134-136*). However, the stylistic differences in the pottery led to the discussion of whether the Ceramic Age dispersals were from different regions and groups. The La Hueca culture (*137*) was identified on Puerto Rico, and the distinct bird shape used in its decorations and ornaments are reminiscent of the Andean condor. The idea of a second "Ceramic" dispersal of people connected to the Andes is not universally accepted in Caribbean archaeology, and more fundamentally the route in which these dispersals happened is debated. The stepping stone model assumes a dispersal from south to north, using the intervisible islands of the Lesser Antilles as stepping stones into the Caribbean (*138*).

In contrast, the Southward expansion model (*121, 139, 140*) proposes a leap from the northeastern South American mainland directly to Puerto Rico, omitting the Lesser Antilles at first, before expanding south and west. Generally, the lack of reliable radiocarbon dates across the Caribbean (*121*) makes it challenging to infer the settlement history from the pattern of colonisation. Additionally, the value of comparisons of ceramic styles has been questioned before (*139, 141*). Undisputed are the observations of an expansion hiatus in the Mona passage bridging Puerto Rico and Hispaniola (*135, 142*) around 2400BP. Only around 1000 years later evidence of people using decorated ceramics can be found on Hispaniola and Cuba. It is not clear why the expansion came to a hold. Based on the reports of Spanish chroniclers such as Velazques de Cuellar and de las Casas (quoted in (*143*)), a popular explanation is that the presence of the Archaic Age populations on the Cuban archipelago prevented the advance of the Ceramic Age settlers east. The archaeological record of Cuba shows the presence of foraging, semi-nomadic horticulturalists throughout the Cuban archipelago and in Hispaniola during the Archaic Age. However, a lack of reliable radiocarbon dates

does not allow assessing whether they persisted until the expansion of the South American settlers into the Caribbean. While it seems likely that the newcomers encountered Indigenous communities in Hispaniola and Cuba, the nature of their interaction remains unclear (*144*). However, based on the abundance of Archaic sites including simple ceramics in Hispaniola (*145-148*), and the east of the Cuban archipelago (*149-151*) it has been suggested that the descendants of the Archaic Age communities played a more critical role in the formation of later Greater Antillean societies than traditionally assumed (*152*).

A last migration into the Caribbean is reported by the colonisers, who were supposedly informed by the Indigenous inhabitants of Hispaniola about people in the east (*153*), who, according to the reports, were described as warlike and practicing cannibalism. Today known as "the Island Caribs", they allegedly inhabited the Lesser Antilles and were connected to Carib speakers from the South American mainland, despite in fact that they were speaking an Arawak language (*154*). The present-day Indigenous Peoples of the Lesser Antilles retrace their ancestry to the historic Island Caribs (*155-158*). However, the existence of ancient, pre-contact Caribs as an ethinc entity is as much debated as is their practice of cannibalism (*159, 160*). Some believe they did exist as a ethnic group with origins in Carib speakers of South America, resolving the language discontinuity with adoption of the local Arawakan language while retaining the cultural practices (*135*). Another theory sees the emergence of Island Caribs as the result of the hostility of colonisers towards the Indigenous People of the Greater Antilles, who fled to the Lesser Antilles and joined a resistance, creating a new ethnicity for themselves (*161*). It is not impossible that the Island Caribs as an entity of warlike, cannibalising people where entirely a construction of the Spanish colonisers. After the harsh decline of the Indigenous population of the Greater Antilles, caused by disease and inhumane labour practices, the colonisers were in need for more people to exploit their workforce (*143, 162, 163*). By creating the hostile, warlike and cannibalistic "Island Caribs" they assured the endorsement of enslavement by the Spanish crown (*164-166*).

In the past, the Cayo pottery style of St. Vincent (*167*) as well as the Suazey ceramic complex (*168-170*) have been interpreted as pre-colonial evidence of the Island Caribs. Despite the similarities of the pottery style to Korabio ceramics from the Guianas (*171*), the discussions of the origins are on-going. Overall, the archaeological discussion has

focused on trying to reconcile the archaeological evidence with the records of the colonisers and has yet been unable to solve the “Island Carib Problem” (*172*).

The Indigenous people of the Caribbean, among them the Kalinago and Garifuna, show the widespread narrative of the extinction of Indigenous Caribbean people is not valid. However, European invasion has drastically altered the linguistic, cultural and genetic landscape of the Caribbean, making it very difficult, if not impossible to add present-day genetics to the lines of evidence in the attempt to understand the population history of the Caribbean and to reconstruct the genetic connections to the American mainland.

5.5 Settling the Pacific region – a first and the last frontier

Traditionally the Pacific region is divided into Micronesia, Melanesia and Polynesia. Coined by a French imperialist, these terms are more a reflection of the racist worldview at the time than a sensible geographical, cultural, linguistic or historic structure. The terms Polynesia and Micronesia refer to the nature of the islands - in Micronesia, islands are small, expressed through the Greek word "mikros" for "small”. With over 1000 islands Polynesia has arguably many islands, reflected in the Greek "polys" for "much". Melanesia, on the other hand, refers to the colour of the peoples' skin, "Mela" meaning "black", singling out this region not based on shared history, language and culture, which are all manifold within this region, but on the colour of people’s skin. While Polynesia has remained a useful category seeing the shared history, language family and beliefs, Micronesia and Melanesia have proven to be artificial categories not useful in the context of the human history of the region (*173*).

A more sensible subdivision more reflective of geography, animal and plant species distribution and settlement history, are into Near and Remote Oceania (*173*). Near Oceania comprises New Guinea, which at the LGM was joint with Australia (*174*), including its offshore islands and the Solomon Islands with the exclusion of the Santa Cruz Islands in the east of the island chain. Remote Oceania comprises all other Islands in the Pacific, including islands previously associated with Poly-, Micro- and Melanesia.

The islands of Remote Oceania are far apart, and also far away from the Asian mainland and the Australian continent. At the same time, the first evidence for occupation of these islands shows the arrival of humans only around 3000 years ago. In contrast, Near

Oceania's intervisible islands (*175*) were settled already more than 45 000 years ago, attributable to its closeness to the Asian mainland (*176, 177*).

The routes early modern humans took to settle Near Oceania and Australia are debated. However, evidence for human occupation is established throughout Indonesia on the Nusa Tenggara (*178*), and Sulawesi (*179*) suggesting they crossed the waters of Wallacea into Near Oceania, requiring skills beyond basic seafaring. At the very minimum by 35 000 years ago, humans had occupied most regions of Australia and New Guinea, among them Island Near Oceania.

It was not until much later, around 3000 years ago, that people navigated beyond the borders of Near Oceania into Vanuatu and Western Polynesia in what would become the last initial settlement of human history. Originating possibly in Taiwan around 5000 years before present (*180*), the Austronesian expansion extended through Island South East Asia to New Guinea and reached the Bismarck Archipelago, bringing with them a cultural landscape of domesticated crops and animals, seafaring and the Austronesian language (*181, 182*).

In the Bismarck Archipelago, the Lapita cultural complex developed around 3250 years ago (*183, 184*). Identified in the archaeological record by dentate stamped ornate pottery, their descendants were the first to extend beyond - possibly even leapfrogging (*185, 186*) - the Solomon Islands into western Remote Oceania. The first evidence for their arrival is from the Teouma site on Vanuatu, dating to 2900 BP (*187, 188*) and around 2800 BP from Fiji and Tonga (*189-192*). By 2700 BP the expansion, as identified through the dentate stamped pottery, had reached its easternmost limit on Samoa (*193-195*).

Despite the well-researched expansion and the consensus on the origin and extent, conflicting theories propose two very different modes for the expansion. The "fast train" model (*196*) assumes the ancestors of Lapita people sailed from Taiwan to the Bismarck Archipelago and out into Western Polynesia without (genetic) contact to people in Near Oceania. By contrast, the "slow boat" hypothesis (*197*), assumes a slower dispersal, involving a fusion of two distinct historical traditions before the navigation into western Remote Oceania. Ancient genomic sequences from individuals associated with the Lapita cultural complex from Tonga and Vanuatu showed exclusively East Asian-related ancestry, supporting the "fast train" model (*190*). Although shedding light on one of the

most debated questions in Pacific archaeology, the origin and timing of admixture of the Papuan–related genetic component remained unclear. In the time following the initial settlement of Western Remote Oceania, the archaeological record shows a shift in material culture.

The present-day similarities of languages, beliefs and cultural practices within the Polynesian triangle, suggest a common origin of the Pacific Islanders (*198, 199*). Complemented by more complex models (*200*), the consensus derived from various lines of evidence (summarised in (*201*)) is that following the initial settlement of Western Polynesia by the Lapita cultural complex, the production shifted towards simpler pottery known as "Polynesian plainware" (*202*). Ensuing was a long period where new developments and innovations were produced in a more local sphere within Western Polynesia, leading to distinctive cultural forms and artefacts, and likely the Proto-Polynesian language (*203*). This period lasted ~800 - 1000 years until the people continued the navigation into the yet further away islands of the Polynesian triangle, spanning from Hawai'i in the north, Aotearoa (New Zealand) in the south and Rapa Nui (Easter Island) in the east, ending the last major expansion of human history with the longest voyage of the preindustrial world, 3000 km from eastern Polynesia to Aotearoa, in the 13th century AD (*204*). Seeing present-day people inhabiting Remote Oceania are genetically a mixture of, apart from the European colonisers contribution, Papuan- and Asian-related ancestry (*205, 206*), and other biological markers trace back to various sources (*200*), the question remains as to where the ancestral populations came from, how and when they mixed to form the genetic make-up of present-day Pacific Islanders.

6. AIM

By utilising the recent technological advances in the field of archaeogenetics, this thesis aims to add to our understanding of the settlement processes of tropical islands, focusing on the Caribbean and the Pacific. While archaeology has recorded anchor points in time and space, providing the when and where of human occupation and differences defined through specific material assemblages, this thesis aims to complement the record by adding insights from ancient genomics.

In the Caribbean, several questions are open regarding the populations that inhabited the islands before the European invasion. Concerning the well studied Saladoid expansion, we aim to add to the discussion regarding the route of this dispersal, explicitly testing the competing models of the Southward Expansion and the Stepping Stone model. Expanding westwards, the newcomers from South America most likely encountered Indigenous populations in Hispaniola and Cuba, where sites from both contexts can be found. After testing whether the two groups differ in their genetic ancestry, we aim to investigate possible interactions between groups through genetic admixture. Finally, we aim to identify the origins of the American mainland for the initial settlement during the Archaic Age populations and explore how they are connected to variations in material culture and subsistence among Archaic Age sites.

The genetic origin of the early settlers of western Remote Oceania was shown to be East Asian populations, with negligible Near Oceanic, Papuan-related contribution. Disregarding the European ancestry component introduced through European invasion, present-day Pacific Islanders show a two-fold ancestry with Papuan-related ancestry besides the East Asian-related ancestry of the early settlers. With a focus on individuals from Vanuatu, the first islands settled in western Remote Oceania, we investigate the changes in genetic composition from the initial settlement to the present-day. We aim to address the questions of timing and nature of admixture events, the origin of subsequent dispersals and corroborate the findings with linguistic observations.

With the analysis of ancient genomes from Papua New Guinea, we aim to assess the genetic diversity of New Guinea in the past and examine how the coastal populations were affected by expansions into the region, such as the Austronesian expansion, leading to the diverse cultural and linguistic mosaic in the region today.

7. RESULTS

7.1 Manuscript A: Genomic insights into the early peopling of the Caribbean

Kathrin Nägele*, Cosimo Posth*, Miren Iraeta Orbegozo, Yadira Chinique de Armas, Silvia Teresita Hernandez Godoy, Ulises M. González Herrera, Maria A. Nieves-Colón, Marcela Sandoval-Velasco, Dorothea Mylopotamitaki, Rita Radzeviciute, Jason Laffoon, William J. Pestle, Jazmin Ramos-Madrigal, Thiseas C. Lamnidis, William C. Schaffer, Robert S. Carr, Jane S. Day, Carlos Arredondo Antúnez, Armando Rangel Rivero, Antonio J. Martínez-Fuentes, Edwin Crespo-Torres, Ivan Roksandic, Anne C. Stone, Carles Lalueza-Fox, Menno Hoogland, Mirjana Roksandic, Corinne L. Hofman, Johannes Krause, Hannes Schroeder in Science (369) pages 456-460, 2020

*Authors contributed equally

The Caribbean Islands were separated from the American continents throughout human history. The origins of the first settlers, their dispersal routes into the islands and their connection to the American mainland are still a matter of debate. The first occupational sites appear in the Archaic Age in the extreme ends of the archipelago. In Cuba, they are dated to 5000 BP, contemporaneous with sites on Barbados in the South. They are represented in the archaeological record by a variety of material cultures, some by lithic tool assemblages, others by tools made from marine resources. Attempts to identify a region of origin have mainly relied on the comparison of artefact assemblages and morphological features and suggested origins in Florida, the Yucatan peninsula and South America. A second migration is well attested archaeologically, identified by ornate pottery with regional variations, marking the beginning of the Ceramic Age in the Caribbean by 2800 BP. The archaeologically identified origins of this dispersal in the upper Amazon region are supported by ancient genomics and linguistics. In this study, we analysed 93 individuals from various islands and found genetic differences consistent with the two contexts. While the groups associated with the Ceramic Age dispersal derive from the same ancestral source connected to

present-day populations in northeastern South America, we find evidence for at least two dispersals into Cuba before 2500 BP. One of the dispersals seems to be connected to radiation events in North America. The genetic diversity supports the results of other bioarchaeological studies reporting differences in morphology, diet, weaning and burial practices. We find a surprising lack of admixture between the early and the Ceramic Age settlers leaving questions as to how the encounters of the different populations shaped the cultural landscape of the Caribbean before the arrival of the European colonisers. Our results shed light on the initial peopling of the Caribbean and the movements of Archaic Age peoples in the Americas.

HUMAN EVOLUTION

Genomic insights into the early peopling of the Caribbean

Kathrin Nägele[1]*†, Cosimo Posth[1,2]†, Miren Iraeta Orbegozo[3], Yadira Chinique de Armas[4], Silvia Teresita Hernández Godoy[5,6], Ulises M. González Herrera[7], Maria A. Nieves-Colón[8], Marcela Sandoval-Velasco[3], Dorothea Mylopotamitaki[3], Rita Radzeviciute[1], Jason Laffoon[9], William J. Pestle[10], Jazmin Ramos-Madrigal[3], Thiseas C. Lamnidis[1], William C. Schaffer[11,12], Robert S. Carr[13], Jane S. Day[14], Carlos Arredondo Antúnez[15], Armando Rangel Rivero[15], Antonio J. Martínez-Fuentes[15]‡, Edwin Crespo-Torres[16]‡, Ivan Roksandic[4], Anne C. Stone[8,12], Carles Lalueza-Fox[17], Menno Hoogland[9,18], Mirjana Roksandic[4], Corinne L. Hofman[9,18], Johannes Krause[1]*, Hannes Schroeder[3,9]*

The Caribbean was one of the last regions of the Americas to be settled by humans, but where they came from and how and when they reached the islands remain unclear. We generated genome-wide data for 93 ancient Caribbean islanders dating between 3200 and 400 calibrated years before the present and found evidence of at least three separate dispersals into the region, including two early dispersals into the Western Caribbean, one of which seems connected to radiation events in North America. This was followed by a later expansion from South America. We also detected genetic differences between the early settlers and the newcomers from South America, with almost no evidence of admixture. Our results add to our understanding of the initial peopling of the Caribbean and the movements of Archaic Age peoples in the Americas.

Archaeological evidence suggests that people first moved into the Caribbean around 8000 calibrated years before the present (cal yr B.P.) (*1*, *2*). Apart from Trinidad, which is located closer to the American mainland, the earliest securely dated archaeological sites in the region date to around 5000 cal yr B.P. and are located in Barbados, Cuba, Curaçao, and St. Martin, followed by sites in Hispaniola and Puerto Rico (*2*). The locations of these sites suggest that the early settlers took long and rapid leaps of exploration across the Caribbean Sea. As a result, there is no gradual wave of advance that would point backward to a single point of origin. In the absence of clear chronological clues, some archaeologists have relied on stylistic comparisons of artifact assemblages to suggest possible links between the Caribbean and surrounding mainland (*3*, *4*), and others have studied the prevailing winds and currents to suggest possible dispersal routes (*5*).

Starting around 2800 cal yr B.P., new people began to enter the islands. Their arrival marks the beginning of the Ceramic Age in the Caribbean as a distinctive new style of pottery starts to appear along with more permanent settlements and agricultural practices (*1*). Archaeological and genetic evidence indicates that the new settlers came from South America (*6*, *7*), but how they reached the islands is debated. Two models have been put forward: The traditional model suggests that people gradually moved northward through the Lesser Antilles until they reached Puerto Rico, and then they eventually moved further west into Hispaniola and Cuba (*6*). Alternatively, it has been suggested that the new settlers first reached Puerto Rico, bypassing the Lesser Antilles, before expanding southward (*8*). Whichever way this expansion took place, it seems likely that the newcomers encountered indigenous communities in the islands, but the nature of their interactions is unclear (*9*).

To shed light on the population history of the Caribbean, we retrieved genome-wide data from 93 ancient Caribbean islanders from 16 archaeological sites dating between 3200 and 400 cal yr B.P. (Fig. 1 and tables S1 to S3) (*10*). The skeletal samples derive from two distinct archaeological contexts, which are referred to as Archaic and Ceramic, respectively (*10*). The 52 Archaic-related individuals come from seven sites in Cuba and date to around 3200 to 700 cal yr B.P., whereas the 41 Ceramic-related individuals stem from nine sites in Cuba, the Bahamas, Puerto Rico, Guadeloupe, and St. Lucia and date to around 1500 to 400 cal yr B.P. (Fig. 1). To overcome the challenges posed by poor DNA preservation, we used a hybridization capture method targeting ~1.2 million genome-wide single-nucleotide polymorphisms (SNPs) (*10*). Additionally, we report mitochondrial DNA (mtDNA) haplogroups for 89 of the 93 individuals and Y chromosome haplogroups for 40 of the 47 males (table S1). Contamination estimates were low (on average <1% on both nuclear and mitochondrial estimates) except for five individuals, who were not included in the final dataset (table S4).

The mtDNA data reveal clear differences in haplogroup frequencies between the individuals from the two contexts (fig. S1). Although most of the individuals from Cuba from 3200 to 700 cal yr B.P. carry haplogroups D1 and C1d (with a frequency of 47 and 30%, respectively), these haplogroups are less common among individuals from Ceramic-related contexts, including those reported in previous studies (*11*, *12*). Overall, mtDNA diversity is higher among Ceramic Age individuals, with haplogroups B2, C1b, and C1c specific to this group (fig. S1).

To explore these differences at a genome-wide level, we performed a principal components analysis (PCA) on the capture data using 12 present-day Native American populations as references (*10*) (Fig. 2A), and we found that the individuals fall into two distinct clusters that are consistent with their archaeological contexts. When plotting the ancient Caribbean individuals with other ancient and modern Native Americans (*7*, *13*–*17*), we find that individuals from Ceramic Age contexts, including those from Cuba, cluster with present-day individuals from South America as well as a published 1000-year-old genome from the Bahamas (*7*). By contrast, individuals from Archaic-related contexts in Cuba from 3200 to 700 cal yr B.P. cluster outside present-day Native American variation (fig. S2).

To assess whether the observed clustering reflects different genetic affinities, we grouped individuals by site and computed f_4 statistics of the form f_4(*Mbuti*, *Test*; *Early San Nicolas*, *Preacher's Cave*), measuring the amount of allele sharing between the tested groups (*Test*) and the 1000-year-old individual from the Bahamas (*Preacher's Cave*) (*7*) versus 4900-year-old individuals from California's Channel Islands (*Early San Nicolas*) (*16*), who represent a branch splitting off the main Native American lineage before the diversification of ancient Central and South Americans (Fig. 2B and table S5) (*15*). As expected, the individuals

[1]Max Planck Institute for the Science of Human History, Jena, Germany. [2]Institute for Archaeological Sciences, Archaeo- and Palaeogenetics, University of Tübingen, Tübingen, Germany. [3]The Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [4]Department of Anthropology, University of Winnipeg, Winnipeg, MB, Canada. [5]Grupo de Investigación y Desarrollo, Dirección Provincial de Cultura, Matanzas, Cuba. [6]Universidad de Matanzas, Matanzas, Cuba. [7]Departamento de Arqueología, Instituto Cubano de Antropología, Havana, Cuba. [8]School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, USA. [9]Faculty of Archaeology, Leiden University, Leiden, Netherlands. [10]Department of Anthropology, University of Miami, Miami, FL, USA. [11]Liberal Arts Department, Phoenix College, Phoenix, AZ, USA. [12]Center for Bioarchaeological Research, Arizona State University, Tempe, AZ, USA. [13]Archaeological and Historical Conservancy Inc., Davie, FL, USA. [14]Research Atlantica Inc., Boca Raton, FL, USA. [15]Museo Antropológico Montané, Facultad de Biología, Universidad de La Habana, Havana, Cuba. [16]Departamento de Sociología y Antropología, Universidad de Puerto Rico Rio Piedras, San Juan, Puerto Rico. [17]Institute of Evolutionary Biology, Spanish National Research Council (CSIC)–Universitat Pompeu Fabra, Barcelona, Spain. [18]Royal Netherlands Institute of Southeast Asian and Caribbean Studies, Leiden, Netherlands.
***Corresponding author. Email: hschroeder@sund.ku.dk (H.S.); naegele@shh.mpg.de (K.N.); krause@shh.mpg.de (J.K.)**
†These authors contributed equally to this work.
‡Deceased.

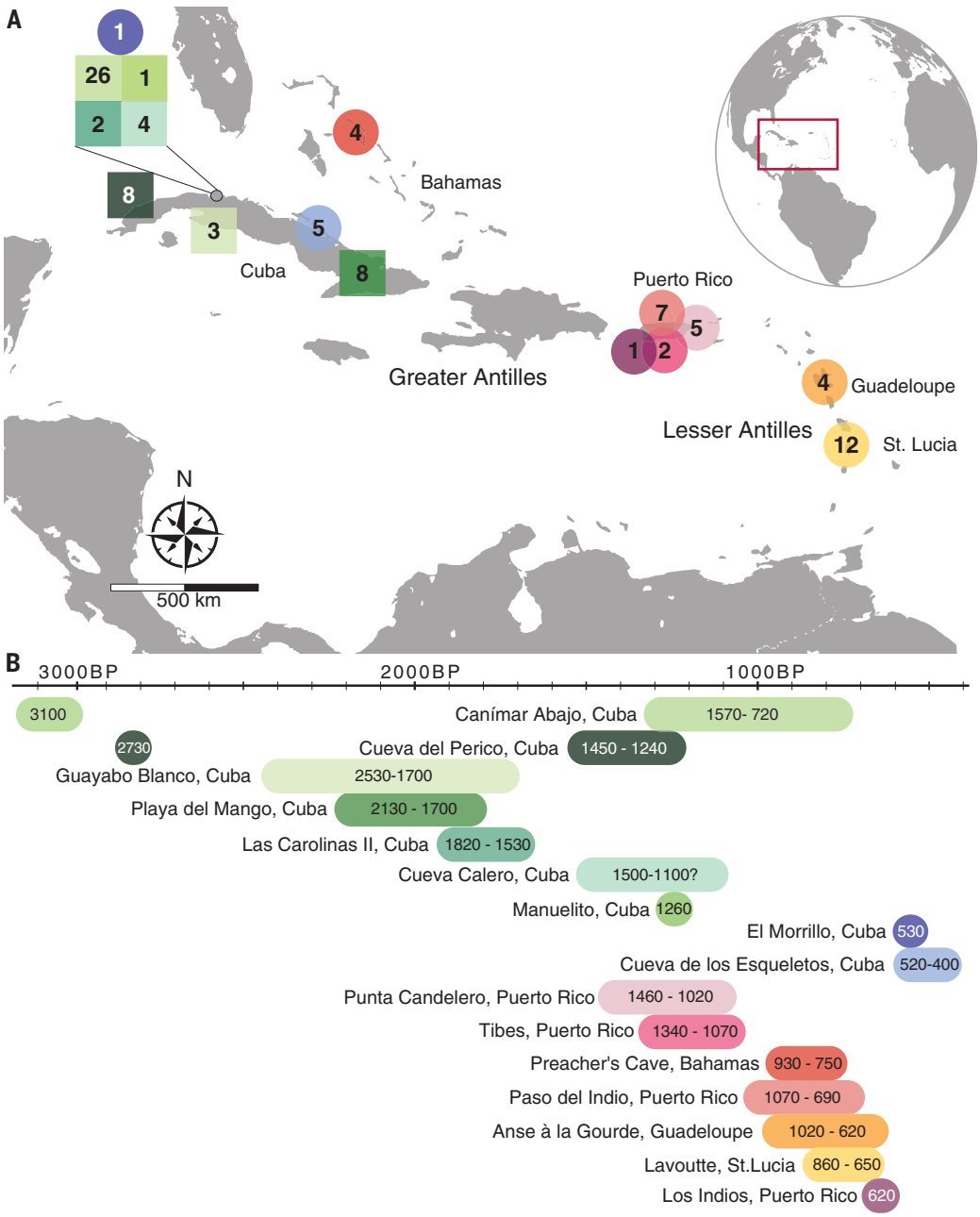


Fig. 1. Sites and samples. (**A**) Map of the Caribbean showing the locations of the sites discussed in the text, including the number of individuals analyzed per site. Squares represent sites with samples from Archaic-related contexts, and circles denote those from Ceramic-related contexts. (**B**) Date ranges for each site are reported in calibrated years before the present (BP). Date ranges derive from directly dated skeletal remains and do not necessarily represent the entire period of occupation of a site. For sites with single individuals, mean point dates are provided. The date ranges for the Cueva Calero individuals are based on archaeological context and indirect radiocarbon dates (*10*).

from Preacher's Cave show the highest affinity to the genome from the same site (*7*), followed by all other Ceramic-related groups. By contrast, all individuals from Cuba from 3200 to 700 cal yr B.P. show less affinity to the Bahamian genome, with one individual from the site of Cueva del Perico (CIP009) being slightly closer to the individuals from California's Channel Islands (*16*). These differences are largely driven by a greater similarity of Ceramic-related groups to present-day populations from northeastern South America (Fig. 2C and figs. S3 and S4) (*7*).

To test whether the two groups derived from the same or distinct ancestral populations, we used *qpWave* (*18*), which estimates the minimum number of sources necessary to explain the genetic composition of an individual or group of individuals (*10*). This analysis was consistent with the groups deriving from at least two separate streams of ancestry (chi-square test, $P = 1.68 \times 10^{-17}$), which demonstrates that the distinction we observe in the PCA cannot be explained by genetic drift alone (table S6). This is also reflected in a supervised clustering analysis, which results in two separate components (fig. S5A) (*10*).

The radiocarbon dates associated with the individuals (Fig. 1B) indicate that both groups were present in the Caribbean at the same time. However, using *qpAdm* (*19*), we do not detect any notable levels of admixture, except for one individual (PDI009) from the Ceramic Age site of Paso del Indio in Puerto Rico, who is dated to 1060 to 910 cal yr B.P. and carries a minor proportion of Archaic-related ancestry (13 ± 7.7%) (table S7). Considering the mounting evidence of the influence of Archaic Age communities on the development of later Caribbean societies (*20*, *21*), it is notable to find so little evidence of admixture between the two groups. However, it is possible that the result is influenced by our limited sampling coverage of the transitional period and of islands such as Hispaniola.

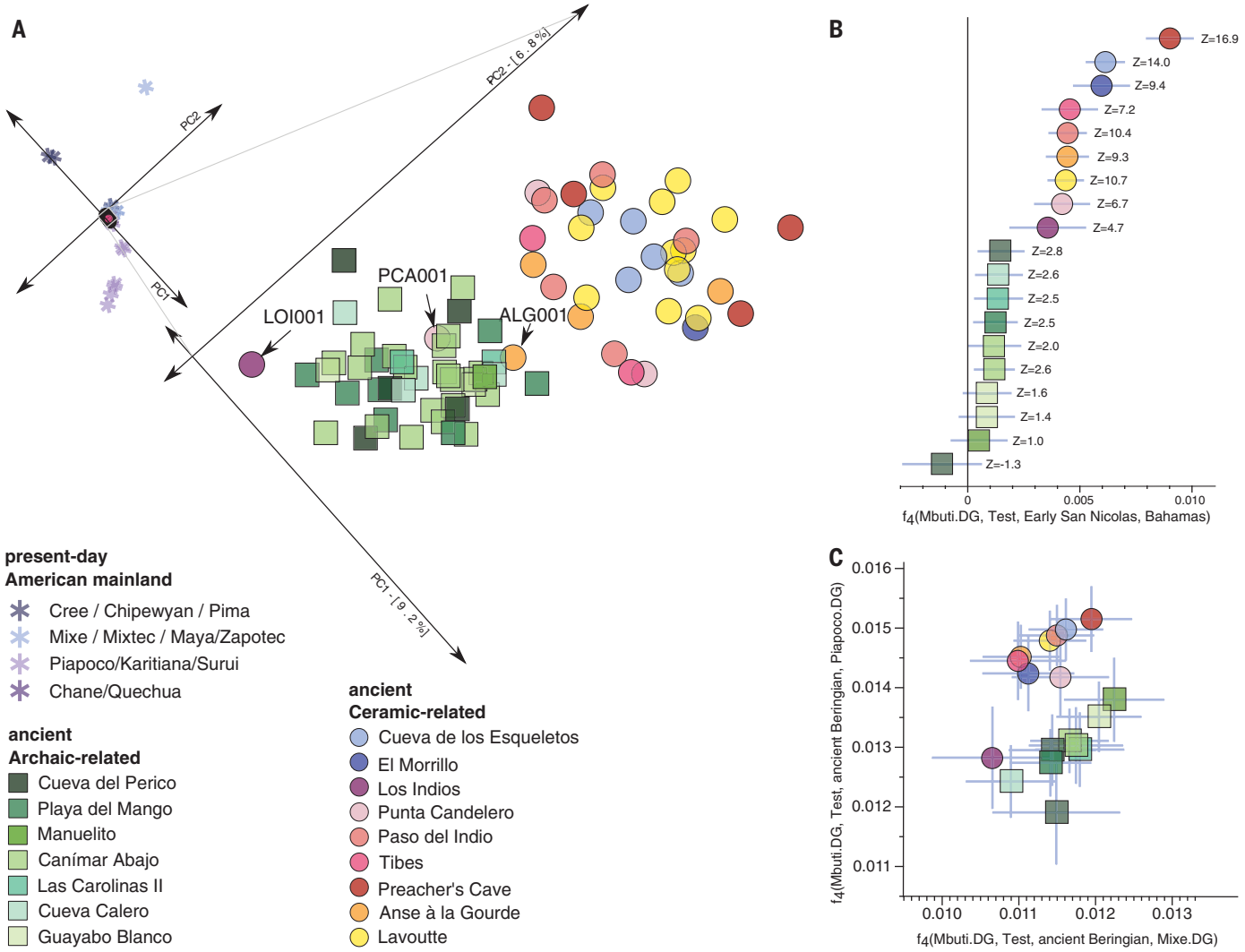


Fig. 2. Population substructure of ancient Caribbean islanders. (**A**) PCA of ancient Caribbean islanders projecting the ancient individuals onto principal components calculated from present-day Native American populations (*10*). Three Ceramic Age individuals (ALG001, LOI001, and PCA001) cluster outside their main grouping, but f_4 statistics indicate that they are more closely related to Ceramic-related than to Archaic-related individuals (table S5). (**B**) f_4 statistics measuring the differential affinities of ancient Caribbean islanders to 4900-year-old individuals from the California Channel Islands (Early San Nicolas) (*16*) and a published 1000-year-old individual from the Bahamas (*7*). The Bahamian genome serves as a proxy for ancient northeastern South American components that are not available from the mainland. (**C**) Differential affinities of ancient Caribbean islanders to present-day Piapoco (y axis) and Mixe (x axis). Light blue lines indicate two standard errors. Squares indicate samples from Archaic-related contexts, and circles denote those from Ceramic-related contexts.

We also detect two distinct ancestries in Cuba around 2700 to 2500 cal yr B.P., represented by the oldest individuals from Cueva del Perico (CIP009) and Guayabo Blanco (GUY002) (Fig. 3, A and B), which suggests multiple early dispersals into the western Caribbean before the arrival of Ceramic Age groups. Using *qpWave* (*18*), we find that some of the oldest individuals in our dataset (i.e., CIP009 and the individuals from Guayabo Blanco) cannot be modeled as descendants of the same ancestral source (chi-square test, $P = 0.013$) (table S6). When we try to model CIP009 alongside other ancient Native American genomes (*14*–*16*) using *qpGraph* (*18*), a model where CIP009 branches off the main Native American lineage with the individuals from California's Channel Islands (*16*) before the radiation of ancient South and Central Americans fits the data best (Fig. 3A). By contrast, all other Archaic-related individuals, including the 2500-year-old individual from Guayabo Blanco (GUY002), require additional gene flow from ancient South Americans to improve the models (Fig. 3B and fig. S6). Together, these results support multiple dispersals into the western Caribbean before the arrival of Ceramic Age groups. Although it is difficult to determine where these early dispersals originated, it seems that at least one of them was connected to radiation events in North America before the diversification of Central and South Americans (*14*, *15*).

After 2800 cal yr B.P., there was another expansion, which originated in South America and is well supported archaeologically (*1*). When we model this expansion using the Ceramic Age genomes in our dataset, we find that a stepping-stone model with people originating in South America and gradually moving northward through the Lesser Antilles fits the data better than a model assuming a southward expansion from Puerto Rico (Fig. 3C and fig. S7). However, because we do not have any individuals with Ceramic-related ancestry from the earliest phase of the Ceramic Age expansion (around 2800 to 2200 cal yr B.P.), it is difficult to model this process accurately. The expansion of Ceramic Age groups

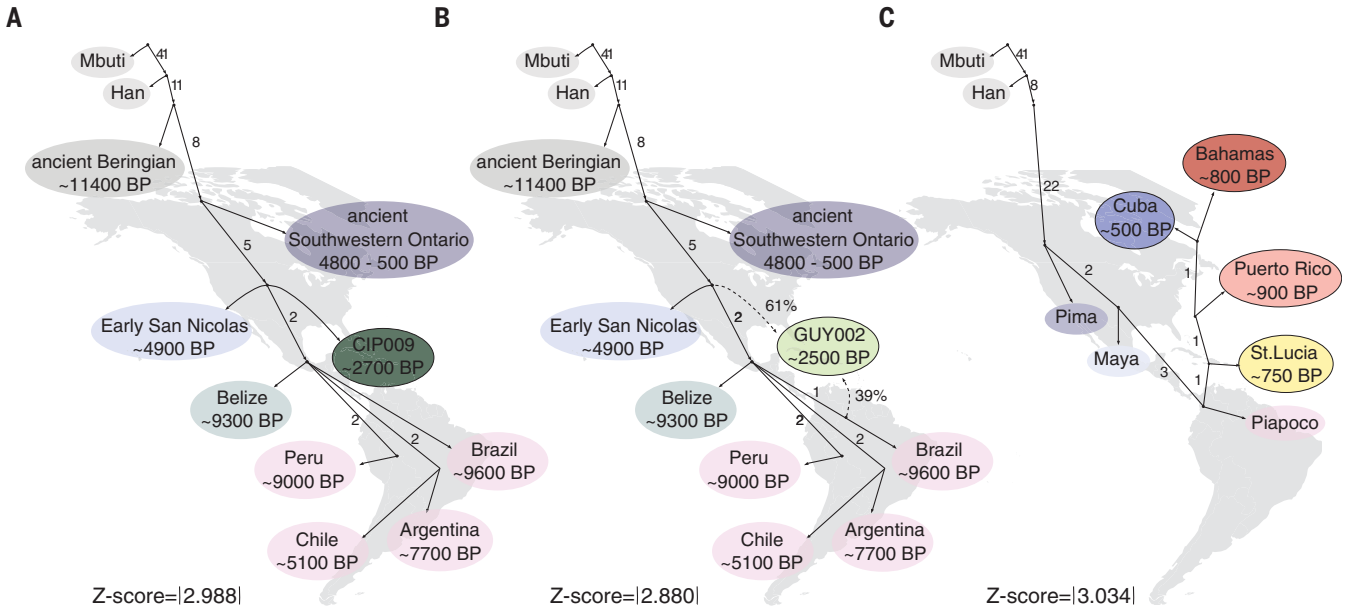


Fig. 3. Admixture graphs modeling the ancestry of ancient Caribbean islanders. (**A** to **C**) We show the best-fitting models for each individual or group as inferred from the final fit score (*10*) for individual CIP009 from the Cueva del Perico (A), individual GUY002 from Guayabo Blanco (B), and several Ceramic Age groups (C). CIP009 (2700 cal yr B.P.) branches off the main Native American lineage along with individuals from the California Channel Islands (*16*) before the diversification of Central and South Americans, whereas GUY002 (2500 cal yr B.P.) requires some South American–related ancestry to make the model fit. The expansion of South American groups after 2800 cal yr B.P. can best be modeled as a stepping-stone process, whereas a southward model results in a worse fit (fig. S7). The geographical positions of ancient groups correspond to their approximate locations. Arrows do not indicate dispersal routes, and node placements do not show the actual geographic regions where the splits took place. Numbers to the right of solid lines are proportional to optimized drift; percentages to the right of dashed lines represent admixture proportions. The label Peru ~9000 BP includes Peru Cunchaicha (9000 cal yr B.P.) and Peru Lauricocha (8600 cal yr B.P.) (*15*). For other groups, see the supplementary materials (*10*).

stalled in Puerto Rico for at least 1000 years before resuming sometime after 1500 cal yr B.P., and it is generally assumed that the advance was halted by the presence of Archaic Age communities in Hispaniola and Cuba (*1*, *6*). Our results are consistent with a temporal gap, as we do not detect any Ceramic-related ancestry in Cuba until 500 cal yr B.P. However, it is still unclear whether we are dealing with a period of genetic turnover (*19*, *22*) or a more-complex history of interaction with intermittent episodes of admixture similar to those that have been observed in other parts of the world (*23*, *24*).

The genetic evidence presented in this work supports the notion that the Caribbean was settled and resettled by successive population dispersals that originated on the American mainland. We find support for at least three separate population dispersals into the region, including two early dispersals, one of which appears to be connected to radiation events in North America. Archaic Age peoples clearly had the seafaring abilities to conquer the Caribbean (*5*). In fact, there is mounting evidence to suggest that, far from being an insuperable barrier, the Caribbean Sea functioned as an aquatic motorway that people crossed frequently, despite its occasional unpredictability (*25*). The initial peopling of the Caribbean was later followed by another expansion from South America. As the newcomers arrived in the islands, they must have encountered descendants of the early settlers, but we find notably little evidence of admixture. This raises questions regarding the nature of their interactions and the role of the early settlers in the development of later Caribbean societies. Additional data and multiple lines of evidence will be needed to explore these questions further and to shed more light on the complex population history of the Caribbean.

REFERENCES AND NOTES

1. W. F. Keegan, C. L. Hofman, *The Caribbean Before Columbus* (Oxford Univ. Press, 2017).
2. M. F. Napolitano *et al.*, *Sci. Adv.* **5**, eaar7806 (2019).
3. W. R. Coe II, *Am. Antiq.* **22**, 280–282 (1957).
4. S. M. Wilson, H. B. Iceland, T. R. Hester, *Lat. Am. Antiq.* **9**, 342–352 (1998).
5. R. T. Callaghan, *Lat. Am. Antiq.* **14**, 323–338 (2003).
6. I. Rouse, *The Tainos: Rise and Decline of the People Who Greeted Columbus* (Yale Univ. Press, 1992).
7. H. Schroeder *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 2341–2346 (2018).
8. S. M. Fitzpatrick, *J. Mari. Arch.* **8**, 101–138 (2013).
9. A. T. Antczak, C. L. Hofman, in *Early Settlers of the Insular Caribbean: Dearchaizing the Archaic*, C. L. Hofman, A. T. Antczak, Eds. (Sidestone Press, 2019), pp. 29–42.
10. See supplementary materials.
11. C. Lalueza-Fox, F. L. Calderón, F. Calafell, B. Morera, J. Bertranpetit, *Ann. Hum. Genet.* **65**, 137–151 (2001).
12. M. A. Nieves-Colón *et al.*, *Mol. Biol. Evol.* **37**, 611–626 (2020).
13. D. Reich *et al.*, *Nature* **488**, 370–374 (2012).
14. J. V. Moreno-Mayar *et al.*, *Science* **362**, eaav2621 (2018).
15. C. Posth *et al.*, *Cell* **175**, 1185–1197.e22 (2018).
16. C. L. Scheib *et al.*, *Science* **360**, 1024–1027 (2018).
17. G. A. Gnecchi-Ruscone *et al.*, *Mol. Biol. Evol.* **36**, 1254–1269 (2019).
18. N. Patterson *et al.*, *Genetics* **192**, 1065–1093 (2012).
19. W. Haak *et al.*, *Nature* **522**, 207–211 (2015).
20. W. F. Keegan, *Caribb. J. Sci.* **42**, 1 (2006).
21. C. L. Hofman, A. T. Antczak, Eds., *Early Settlers of the Insular Caribbean: Dearchaizing the Archaic* (Sidestone Press, 2019).
22. M. E. Allentoft *et al.*, *Nature* **522**, 167–172 (2015).
23. H. McColl *et al.*, *Science* **361**, 88–92 (2018).
24. C. Posth *et al.*, *Nat. Ecol. Evol.* **2**, 731–740 (2018).
25. C. L. Hofman, A. J. Bright, R. Rodríguez Ramos, *J. Caribb. Archaeol.* **3**, 1–18 (2010).

ACKNOWLEDGMENTS

We thank the Comisión Nacional de Monumentos de la República de Cuba; the Consejo para la Protección del Patrimonio Arqueológico Terrestre de Puerto Rico; the Antiquities, Monuments and Museum Corporation of the Bahamas (AMMC); the Direction des affaires culturelles de Guadeloupe; and the St. Lucia Archaeological and Historical Society for supporting this study and providing access to samples. Special thanks go to the staff of the Centro Ceremonial Indígena de Tibes for assistance with sampling and curation at the Tibes site. **Funding:** The research was funded by the Max Planck Society and the European Research Council under the 7th Framework Program (grant agreement no. 319209, ERC Synergy Project NEXUS1492). H.S. was supported by the HERA (Humanities in the European Research Area) Joint Research Program "Uses of the Past" (CitiGen) and the European Union's Horizon 2020 research and innovation program under grant agreement no. 649307. W.J.P. and M.A.N.-C. were supported by the National Science Foundation (BCS-0612727 and BCS-1622479). C.L.-F. was supported by a grant from the Ministry of Science, Innovation and Universities (PGC2018-0955931-B-100, AEI/FEDER, UE). M.R. was supported by the Social Sciences and Humanities Research Council of Canada (435-2016-0529). M.R., Y.C.d.A., U.M.G.H., and S.T.H.G. were supported by the Social Sciences

and Humanities Research Council of Canada (standard research grant SSHRC - 410-2011-1179 and SSHRC postdoctoral fellowship - 756-2016-0180) and several University of Winnipeg internal grants (Major grant 2017, 2018; Partnership Development grant 2017, 2018; and Discretionary grant 2017, 2018). **Author contributions:** H.S., K.N., C.P., and J.K. conceived and led the study. M.R. and Y.C.d.A. funded and coordinated excavations in Cuba with administrative and academic support by A.J.M.-F. U.M.G.H. and Y.C.d.A. led the excavation of the Playa del Mango site. S.T.H.G. and Y.C.d.A. led the excavation at the sites of Canímar Abajo. E.C.-T. excavated and curated individuals from the sites of Punta Candelero, Los Indios, and Paso del Indio. C.L.H. and M.H. led the archaeological excavations at the site of Anse à la Gourde and Lavoutte. W.J.P. sampled and M.A.N.-C. extracted DNA from the individuals excavated at Paso del Indio, Punta Candelero, and Tibes. K.N., M.I.O., R.R., M.S.-V., and D.M. processed the rest of the samples. K.N., C.P., and H.S. analyzed the data with input from T.C.L. and J.R.-M. K.N., H.S., C.P., and J.K. interpreted the data with critical input and contextualization from Y.C.d.A., U.M.G.H., S.T.H.G., C.A.A., A.R.R., C.L.-F., I.R., and M.R. for Cuba; W.J.P., M.A.N.-C., and A.C.S. for Puerto Rico; R.S.C., J.S.D., and W.C.S. for the Preacher's Cave; and C.L.H., J.L., and M.H. for the sites in the Lesser Antilles. K.N. and H.S. wrote the manuscript with critical input from C.P., J.K., M.R., C.L.H., and the remaining authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Alignment files of the nuclear and mtDNA sequences for all analyzed individuals are available at the European Nucleotide Archive (ENA) database under the accession no. PRJEB37518.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6502/456/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S7
Tables S1 to S7
References (*26*–*101*)
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

28 February 2020; accepted 18 May 2020
Published online 4 June 2020
10.1126/science.aba8697

7.2 Manuscript B: Language continuity despite population replacement in Remote Oceania

Cosimo Posth*, Kathrin Nägele*, Frédérique Valentin, Stuart Bedford, Kaitip W. Kami, Richard Shing, Hallie Buckley, Rebecca Kinaston, Mary Walworth, Geoffrey R. Clark, Christian Reepmeyer, James Flexner, Tamara Maric, Johannes Moser, Julia Gresky, Lawrence Kiko, Kathryn J. Robson, Kathryn Auckland, Stephen J. Oppenheimer, Adrian V.S. Hill, Jana Zech, Fiona Petchey, Partick Roberts, Choongwon Jeong, Russel D. Gray, Johannes Krause, Adam Powell in Nature Ecology and Evolution (2), pages 731 – 740 (2018)

*Authors contributed equally

The Lapita cultural complex originated in the Bismarck Archipelago around 3250 years BP. Identified in the archaeological record by their distinct ornate pottery with dentate stamped patterns, they were part of the Austronesian expansion. Bringing with them a cultural landscape of agriculture, seafaring technologies and Austronesian languages, they expanded quickly into Remote Oceania, reaching eastern Remote Oceania around 2900 – 2500 years BP. Ancient DNA analysis of individuals associated with the Lapita cultural complex revealed to have almost exclusively East-Asian ancestry. The new finding raised questions about the formation of the genetic make-up of present-day Pacific Islanders, who in addition to the East Asian component carry Papuan-related ancestry. The amount of Papuan-related ancestry varies across the Pacific region. While Polynesians carry around 40 %, the genetic make-up of present-day inhabitants of Vanuatu was unknown. The cultural aspects show closer connections to Papuan populations, while the languages spoken in Vanuatu are all part of the Austronesian language family associated with the Austronesian expansion. To investigate the present-day genetic make-up of Vanuatu Islanders (ni-Vanuatu), we genotyped the genomes of 27 ni-Vanuatu. To understand the processes leading to the observations from those present-day populations, we analysed a time transect of over 2,500 years with 19 ancient individuals from various islands in Vanuatu, from Tonga and from Malaita in the Solomon Islands. We found evidence for a previously undescribed

dispersal of people with Papuan-related ancestry into Vanuatu, evidenced by an individual of exclusively Papuan-related ancestry dated to 2500 cal. BP excavated on Tanna, sharing the highest affinity with present-day populations on New Britain in the Bismarck Archipelago. The genetic time-transect shows the genetic exchange of the first inhabitants associated with the Lapita cultural complex and individuals of Papuan-related ancestry, commencing immediately after the initial colonisation and leading to a diverse population with varying amounts of Papuan-related ancestry in Vanuatu in the post-Lapita period. Younger individuals in the time transect show higher proportions of Papuan ancestry, maximised in the newly genotyped Ni-Vanuatu, showing a genetic turnover since initial colonisation. At the same time, the first inhabitants preserve the Austronesian language, as all languages spoken in Vanuatu today are part of the Austronesian branch. This peculiarity is rare, if not unprecedented in human history, and assuming a single, substantial migration leading to the genetic turnover would inevitably have resulted in a shift to Papuan languages or Austronesian languages related to those spoken on the Bismarck Archipelago today. When dating the admixture in the individuals along this time-transect, we observe older admixture dates for the more ancient samples, suggesting multiple admixture events. The repeated arrival of people with Papuan-related ancestry could have led to the observed language continuity while resulting in a genetic turnover, leading to the disparity of genetic ancestry and languages spoken in the archipelago.

nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-018-0498-2

Language continuity despite population replacement in Remote Oceania

Cosimo Posth[1,19*], Kathrin Nägele[1,19], Heidi Colleran[2*], Frédérique Valentin[3], Stuart Bedford[2,4], Kaitip W. Kami[2,5], Richard Shing[5], Hallie Buckley[6], Rebecca Kinaston[1,6], Mary Walworth[2], Geoffrey R. Clark[7], Christian Reepmeyer[8], James Flexner[9], Tamara Maric[10], Johannes Moser[11], Julia Gresky[12], Lawrence Kiko[13], Kathryn J. Robson[14], Kathryn Auckland[15], Stephen J. Oppenheimer[16], Adrian V. S. Hill[15], Alexander J. Mentzer[15], Jana Zech[17], Fiona Petchey[18], Patrick Roberts[17], Choongwon Jeong[1], Russell D. Gray[2], Johannes Krause[1*] and Adam Powell[1,2*]

Recent genomic analyses show that the earliest peoples reaching Remote Oceania—associated with Austronesian-speaking Lapita culture—were almost completely East Asian, without detectable Papuan ancestry. However, Papuan-related genetic ancestry is found across present-day Pacific populations, indicating that peoples from Near Oceania have played a significant, but largely unknown, ancestral role. Here, new genome-wide data from 19 ancient South Pacific individuals provide direct evidence of a so-far undescribed Papuan expansion into Remote Oceania starting ~2,500 yr BP, far earlier than previously estimated and supporting a model from historical linguistics. New genome-wide data from 27 contemporary ni-Vanuatu demonstrate a subsequent and almost complete replacement of Lapita-Austronesian by Near Oceanian ancestry. Despite this massive demographic change, incoming Papuan languages did not replace Austronesian languages. Population replacement with language continuity is extremely rare—if not unprecedented—in human history. Our analyses show that rather than one large-scale event, the process was incremental and complex, with repeated migrations and sex-biased admixture with peoples from the Bismarck Archipelago.

Sahul—the continent comprising present-day Australia, Tasmania and New Guinea—was colonized by modern humans during the Pleistocene as early as 65,000 yr BP[1]. However, it took more than 60,000 yr for humans to move east of the Solomon Islands, from Near Oceania out into Remote Oceania[2] (Fig. 1b). These seafaring Neolithic peoples—part of the Austronesian Expansion beginning ~5,500 yr BP probably in present-day Taiwan and the nearby mainland[3–5]—carried farming technology and a major branch of the Austronesian languages[6] into the islands of Southeast Asia, eventually reaching New Guinea and the Bismarck Archipelago and encountering indigenous Papuans. Here, at ~3,300 yr BP, the Lapita cultural complex[3,7] appeared—characterized by distinctive dentate-stamped pottery—and, using the outrigger sailing canoe, Lapita peoples expanded east, leap-frogging beyond the Solomon Islands[8,9]. They transported their landscapes[3] and Oceanic languages out into Remote Oceania, first arriving in the Santa Cruz Islands, Vanuatu[10] and New Caledonia ~3,000 yr BP[11], and rapidly navigated >800 km of open ocean to Fiji, reaching western Polynesia by ~2,850 yr BP[12].

Uncovering the extent of interaction between incoming Austronesian and indigenous Papuan peoples is critical to understanding all subsequent Pacific prehistory. 'Papuan' here refers to both the non-Austronesian languages found across New Guinea and a component of genetic ancestry likely to have diverged from the ancestors of present-day East Asians at least 27,000 yr BP[13]. The linguistic, cultural and genetic diversity in New Guinea is immense, arising through complex histories of differentiation since first arrival[14]. While most Near Oceanians today speak Papuan languages, Remote Oceanians almost exclusively speak Oceanic languages of the Austronesian family[15]. Bayesian phylogenetic analyses of 400 of the >1,200 Austronesian languages[5] broadly support the 'express train' model of the Austronesian expansion, whereby Austronesian-speaking groups had negligible cultural or genetic interaction with indigenous Papuans in Near Oceania before moving further into the Pacific. However, the genetic composition of the present-day South Pacific indicates a more complex history, comprising major East Asian-Austronesian (~79–87%) and minor Papuan (~21–13%)

[1]Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany. [2]Department of Linguistic and Cultural Evolution, Max Planck Institute for the Science of Human History, Jena, Germany. [3]Maison de l'Archéologie et de l'Ethnologie, CNRS, UMR 7041, Nanterre, France. [4]School of Culture, History and Language, College of Asia and the Pacific, Australian National University, Canberra, Australian Capital Territory, Australia. [5]Vanuatu Cultural Centre, Port-Vila, Vanuatu. [6]Department of Anatomy, School of Biomedical Sciences, University of Otago, Dunedin, New Zealand. [7]Archaeology and Natural History, College of Asia and the Pacific, Australian National University, Canberra, Australian Capital Territory, Australia. [8]College of Arts, Society and Education, James Cook University, Cairns, Queensland, Australia. [9]Department of Archaeology, University of Sydney, Sydney, New South Wales, Australia. [10]Service de la Culture et du Patrimoine, Punaauia, Tahiti, French Polynesia. [11]Commission for Archaeology of Non-European Cultures, German Archaeological Institute, Bonn, Germany. [12]Department of Natural Sciences, German Archaeological Institute, Berlin, Germany. [13]Solomon Islands National Museum, Honiara, Solomon Islands. [14]MRC Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, UK. [15]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK. [16]School of Anthropology and Museum Ethnography, University of Oxford, Oxford, UK. [17]Department of Archaeology, Max Planck Institute for the Science of Human History, Jena, Germany. [18]Waikato Radiocarbon Dating Laboratory, The University of Waikato , Hamilton, New Zealand. [19]These authors contributed equally: Cosimo Posth, Kathrin Nägele. *e-mail: posth@shh.mpg.de; colleran@shh.mpg.de; krause@shh.mpg.de; powell@shh.mpg.de

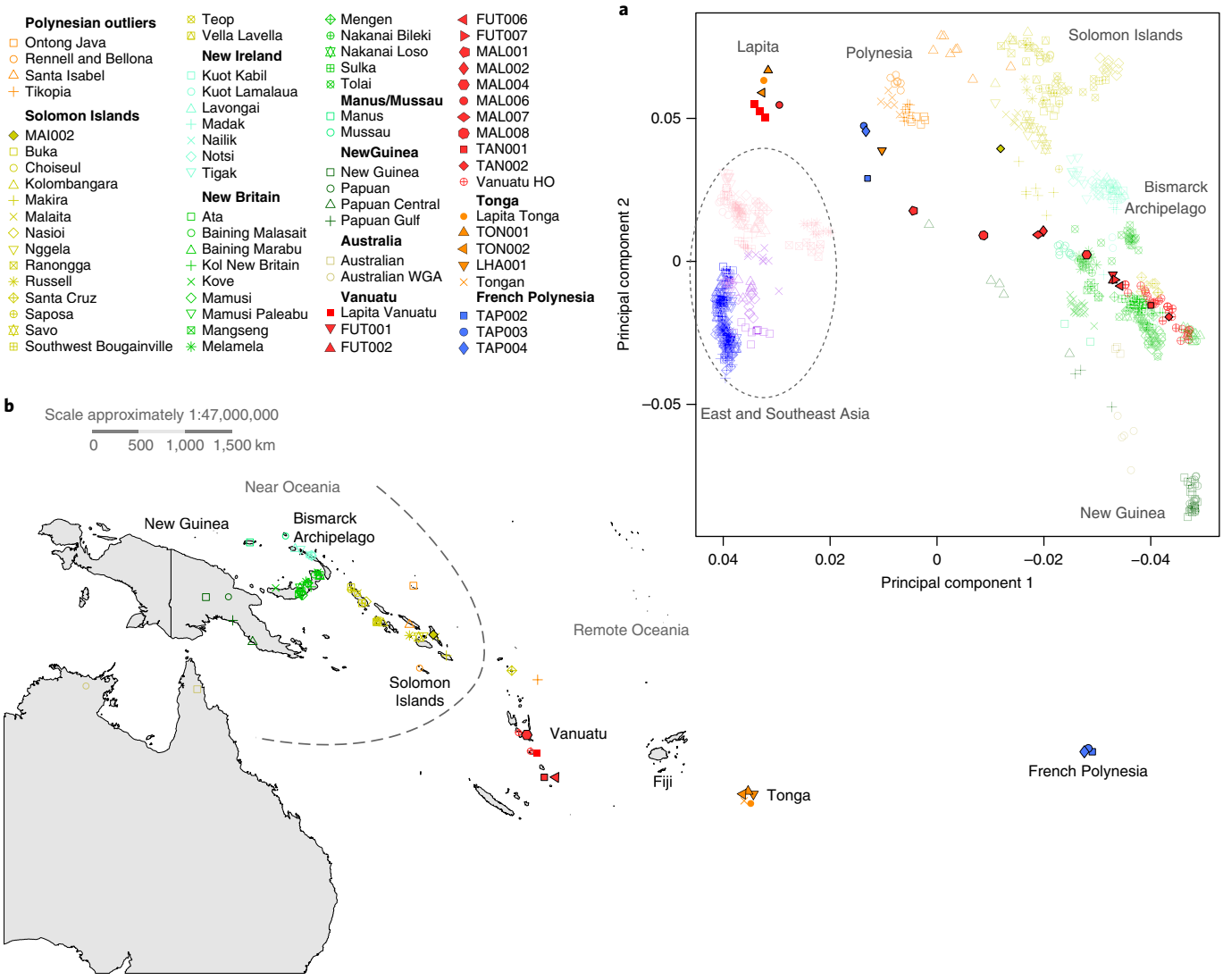


Fig. 1 | Spatial and genetic distribution of ancient and present-day individuals. a, PCA of modern-day East Asian and Near and Remote Oceanian populations genotyped on the Axiom Genome-Wide Human Origins Array, with 23 ancient individuals projected. Ancient samples are indicated by filled symbols and present-day samples are indicated by open symbols (WGA, whole genome assembly). The new data from this study have a black border. **b**, Regional map showing the locations of Near and Remote Oceanian sample populations and ancient individuals.

components of genome-wide ancestry[13,16]. Mitochondrial DNA (mtDNA)[17] and Y-chromosome[18,19] studies show that populations across Polynesia have maternal ancestry largely of Austronesian origin (>96%[20]) while most of their Y-chromosomes derive from Near Oceania (>60%[20]), confirmed in recent X-chromosome analyses[13,21]. This suggests that Oceanic-speaking populations—before or during the formation of the Lapita cultural complex—experienced significantly sex-biased admixture, involving women of Austronesian origin and Papuan men. This model requires that Lapita peoples, while maintaining Oceanic language(s), had admixed ancestry in Near Oceania before their eastward expansion into Remote Oceania. However, the first genome-wide ancient data from the region[21] demonstrate—consistent with craniofacial analyses[22]—that Papuan ancestry is largely absent in individuals from Lapita sites in both Vanuatu and Tonga. The present-day genetic ancestry of Remote Oceania can therefore only be explained by subsequent population expansion, carrying Papuan ancestry into the Pacific.

Vanuatu has been an important hub in the western Pacific[23] from Lapita onwards. Uncovering the detailed demographic processes shaping the genetic and linguistic landscape of Vanuatu is thus crucial to understanding those of the wider Pacific. Here, we provide the earliest direct evidence of Papuan genetic ancestry in Remote Oceania. Our results reveal that peoples from Near Oceania began arriving just a few centuries after the first Lapita settlements in Vanuatu. This was followed by an almost complete, yet incremental, replacement of Lapita-Austronesian by Bismarck Archipelago-like genetic ancestry.

Results

Ancient and modern genome-wide data. We recovered genome-wide and mitochondrial ancient DNA (aDNA) data from the bones and/or teeth of 19 individuals from archaeological sites ^{14}C-dated to ~2,600–200 yr BP across Vanuatu ($n = 12$), Tonga ($n = 3$), French Polynesia ($n = 3$) and the Solomon Islands ($n = 1$) (Table 1 and Supplementary Tables 1 and 2; see Methods). DNA was extracted[24] and converted into double-stranded genetic libraries[25,26] in dedicated cleanroom facilities. Hybridization capture targeted the complete mitochondrial genome and ~1.24 million single nucleotide polymorphisms (SNPs; hereafter referred to as 1,240 K capture)[27,28], followed by next-generation sequencing.

Table 1 | Data description for the newly reported genome-wide data from 19 ancient individuals: radiocarbon dating and ancient DNA summary statistics

Sample name	Country, island	Anatomical element	Cal. BP (CE/BCE) 95.4%	Sex	mtDNA haplogroup	Y-chromosome haplogroup	Damage restricted	Mean coverage	SNPs	Library type
FUT001	Vanuatu, Futuna	Left petrous	1230–980 (720–970 CE)	Female	P1d2a	–	No	1.289	647,595	Non-UDG
FUT002	Vanuatu, Futuna	Right petrous	1240–1000 (710–950 CE)	Female	M28b1	–	No	1.163	626,821	UDG-half
FUT006	Vanuatu, Futuna	Left petrous	1270–1070–880 CE)	Male	P1d2a	K2	No	0.748	453,192	UDG-half
FUT007	Vanuatu, Futuna	Right petrous	1190–970 (760–980 CE)	Male	M28b1	K2b1a3	No	0.596	392,622	UDG-half
LHA001	Tonga, Tongatapu	Molar	780–550 (1170–1400 CE)	Female	B4a1a1	–	Yes	0.048	37,058	UDG-half
MAI002	Solomon Islands, Malaita	Right petrous	540–480 (1410–1470 CE)	Female	B4a1a1a	–	No	5.582	913,583	Non-UDG
MAL001	Vanuatu, Malakula	Left petrous	2330–2100 (380–150 BCE)	Female	B4a1a1	–	No	0.089	78,100	Non-UDG
MAL002	Vanuatu, Malakula	Left petrous	2490–2200 (540–250 BCE)	Female	B4a1a1a	–	No	0.302	220,082	UDG-half
MAL004	Vanuatu, Malakula	Left petrous	2690–2320 (740–370 BCE)	Male	B4a1a1a	M1b	No	1.751	697,939	UDG-half
MAL006	Vanuatu, Malakula	Left petrous	2670–2320 (720–370 BCE)	Female	B4a1a1a11	–	Yes	0.011	10,418	Non-UDG
MAL007	Vanuatu, Malakula	Right petrous	2140–1920 (190–30 BCE)	Female	B4a1a1a	–	No	0.609	394,207	UDG-half
MAL008	Vanuatu, Malakula	Left petrous	2290–1940 (350 BCE–10 CE)	Female	B4a1a1a	–	Yes	0.025	22,381	Non-UDG
TAN001	Vanuatu, Tanna	Left petrous	260–0 (1690–1950 CE)	Male	P1d1	O2a2b2a	No	1.223	629,733	UDG-half
TAN002	Vanuatu, Tanna	Right petrous	2630–2350 (680–400 BCE)	Male	Q2a	K2b1	No	0.241	191,304	UDG-half
TAP002	French Polynesia, Ra'iātea	Molar	270– –10 (1680–1960 CE)	Male	B4a1a1m1	N/A	Yes	0.041	39,897	Non-UDG
TAP003	French Polynesia, Ra'iātea	Molar	270– –10 (1680–1960 CE)	Male	B4a1a1c	CT	No	0.158	137,660	UDG-half
TAP004	French Polynesia, Ra'iātea	Molar	240–10 (1710–1940 CE)	Male	B4a1a1+16126	CT	No	0.072	66,227	Non-UDG
TON001	Tonga, Tongatapu	Right petrous	2670–2320 (720–370 BCE)	Female	B4a1a1a	–	Yes	0.092	82,790	Non-UDG
TON002	Tonga, Tongatapu	Left petrous	2690–2350 (740–400 BCE)	Male	B4a1a1	O1a1a1a	Yes	0.406	285,776	Non-UDG

Cal. BP, calibrated years before present; CE, common era; BCE, before the common era. Y-chromosome haplogroup column is marked with '–' for females and N/A as unavailable for a male individual. Mean coverage and SNP number are calculated on the 1,240 K SNP capture target. UDG-half and non-UDG refer to the type of library protocol applied[25,26] (Methods).

The isolated aDNA was authenticated based on the presence of typical deamination patterns, low levels of mtDNA contamination and X-chromosome contamination in males, and analyses were restricted, if necessary, to the probable endogenous deaminated sequences[29] (Supplementary Tables 3 and 4 and Supplementary Fig. 1; see Methods). The genome-wide aDNA was co-analysed with four published Lapita samples[21], 781 present-day Oceanian and East Asian samples genotyped for ~600 K SNPs on the Axiom Genome-Wide Human Origins Array (hereafter referred to as the HO dataset)[21,30], and 308 high-coverage genomes[31]. We also genotyped 27 ni-Vanuatu samples from the islands of Malakula and Efate (Supplementary Fig. 2; see Methods) on the Axiom Genome-Wide Human Origins Array, with 8 also shotgun sequenced at low coverage (0.6–3-fold) (Supplementary Table 5). All newly generated data were analysed alongside published genome-wide Illumina HumanCore-24 data from 754 individuals across Remote Oceania, including 610 from Vanuatu[32] (Supplementary Table 6).

Demographic history of Vanuatu. While early Lapita people in Vanuatu had largely East Asian-Austronesian ancestry[21], principal component analysis (PCA) shows that, although diverse, the 27 present-day individuals fall instead within the Near Oceanic cline, in close proximity to Santa Cruz and New Britain populations (Fig. 1a,b), demonstrating an almost complete population turnover since initial settlement. Previous analysis based on patterns of linkage disequilibrium (ALDER[33]) estimated the time of Papuan admixture into Remote Oceania at 1,927–1,239 yr BP for Polynesian populations[21], and our analyses on regional populations gave similar estimates of ~2,000–1,500 yr BP (see below). However, the ^{14}C

dates for the ancient samples demonstrate that Papuan ancestry was already in Vanuatu up to 1,000 yr earlier, from ~2,500 yr BP. Both the earliest (TAN002) and latest (TAN001) ancient samples from Tanna (Supplementary Fig. 2) lay inside the distribution of the new present-day HO samples, but it is striking that ancient samples from Malakula and Futuna within this timeframe do not (Fig. 1a). The Malakula time-transect bridges much of the massive genetic distance between initial Lapita inhabitants and contemporary ni-Vanuatu. ADMIXTURE[34] analyses on ancient and modern Vanuatu shotgun-sequenced data support a complex population replacement. With $K=5$ ancestral components—allowing the distinction between Asian-Austronesian (blue) and Near Oceanian-Papuan (green)—Vanuatu demonstrates a general but heterogeneous trend of increasing Papuan ancestry through time (Fig. 2a), from largely Lapita-Austronesian (ref. [21] and MAL006) to predominantly Papuan ni-Vanuatu ancestry.

qpWave analysis[35] determined that ancient Vanuatu could be modelled as a two-way admixture between Papuan and Austronesian populations (Supplementary Table 7), using qpAdm[36] to quantify the relative ancestry proportions (Fig. 2b and Supplementary Table 8). The near-contemporaneous genetic heterogeneity in Malakula is striking. Over the ~500 yr period beginning ~2,500 yr BP, Malakula was home to individuals with between 22 and 46% of their ancestry derived from ancestral Austronesians (Futuna samples ~1,100 yr BP have 11 to 17%). The earliest ancient individual, TAN002, is a male carrying both Papuan mtDNA and Y-chromosome haplogroups (Q2a and K21b, respectively), with autosomes consistent with having no Austronesian ancestry (Fig. 2b and Supplementary Fig. 3). We estimated the excess Austronesian X-chromosome ancestry relative to the autosomes across our time transect, finding diverse levels of maternal ancestry within Malakula (Supplementary Table 8). In particular, MAL004—a male with the typical Papuan Y-chromosome haplogroup *M1b*—carries as much as ~50% Austronesian maternal excess (and Polynesian mtDNA haplogroup B4a1a1a), providing the first direct snapshot of this sex-biased admixture in progress[17–20]. The latest ancient sample, TAN001, shows similar autosomal admixture proportions to contemporary ni-Vanuatu, and carries a Papuan mtDNA haplogroup and Polynesian Y-chromosome haplogroup (P1d1 and O2a2b2a, respectively).

To identify potential source populations of post-Lapita Near Oceanian ancestry, we performed a four-population test for admixture by calculating *D*-statistics[30] on the new ancient Vanuatu data, downsampled to the more geographically extensive HO dataset (Supplementary Table 9). Using the model 'D(Near Oceanian, New Guinea; Vanuatu ancient, Mbuti)', where Near Oceanian is drawn from all potential sources reported in ref. [21], we identified Baining Marabu and Baining Malasait in New Britain, Bismarck Archipelago (Fig. 1b) as the closest present-day proxy sources of Near Oceanian ancestry in the ancient Vanuatu individuals (with standard score Z significantly greater than zero, i.e. $Z \gg 0$). One possible confounding factor is the significant difference in the levels of Austronesian ancestry in Baining populations compared with New Guinea Papuans shown by 'D(Baining Marabu or Baining Malasait, New Guinea; Ami, Mbuti)': $Z=3.7$ or 4.2. However, TAN002 does not show such an attraction to Ami, confirming that its affinity to Baining relative to Papuans is not explained by shared Austronesian ancestry (Supplementary Table 9). Furthermore, although Denisovan admixture levels are observed to decline with increased Austronesian ancestry proportion[37], the best-supported source populations have values consistent with New Guinea Papuans ('D(Baining Marabu or Baining Malasait, New Guinea; Denisovan, Mbuti)': $Z=-0.8$ or -1.9). Thus, *D*-statistics confirm the close relationship observed in PCA between Baining populations and the earliest Vanuatu individual carrying Near Oceanian ancestry (TAN002), despite the immense geographical distance (Fig. 1a,b).

qpGraph[30] analyses (Fig. 3a) showed that TAN002 could be modelled as an unadmixed individual descended from a population ancestral to modern Baining Marabu, before the Baining Marabu receives a 4% Austronesian contribution. In Vanuatu, a population associated with TAN002 would admix with local Lapita people (proxied by Ami) giving rise to ancient Malakula individuals ~2,500–2,000 yr BP. Additional Papuan admixture is needed to account for the lower Austronesian proportion in the ~1,100 yr BP Futuna population (Fig. 2b, Supplementary Table 8 and Supplementary Fig. 3). The most recent ancient individual TAN001 can only be modelled as descended directly from a Baining-related population, suggesting local population replacement. We were unable to fit present-day Vanuatu HO alongside the new ancient samples in a single model (Supplementary Fig. 4), indicating that present-day ni-Vanuatu may carry an additional genetic component not found in ancient populations.

Different genetic trajectory in Polynesia. Analyses of two new Lapita individuals (TON001 and TON002) from the Talasiu site in Tonga[21] confirmed their genetic similarity to early peoples in Vanuatu (Fig. 1a). Notably, TON002 is a male carrying the Y-chromosome haplogroup O1a1a1a, providing direct evidence that this clade—like the 'Polynesian mtDNA motif' haplogroup B4a1a1a—was associated with the Austronesian expansion[38]. After Lapita settlement, the populations of Vanuatu and Tonga appear to follow a considerably different genetic trajectory; PCA analyses indicate that present-day Tongans fall between the East Asian and Near Oceanian clines (Fig. 1a and Supplementary Fig. 5), more specifically between Lapita individuals and Solomon Islanders. A newly sequenced ancient Tongan female sample (LHA001) from 780–550 yr BP, lay relatively close in PCA to modern Tongans, but its lower affinity to Solomon Islanders suggests that modern Tongan ancestry was not yet completely in place by this time ('D(LHA001, Tongan; Savo, Mbuti)': $Z=-3$).

We obtained genome-wide data from three individuals unearthed at the monumental site Taputapuātea (TAP002, TAP003 and TAP004) on the island of Ra'iātea, French Polynesia dated to the time of European contact in the eighteenth century AD[39]. ADMIXTURE[34] analyses (Fig. 2a) show that these individuals have major Austronesian (blue) and minor Papuan (green) ancestry components, and both carry typical Polynesian mtDNA haplogroups (Table 1). In PCA space, they fall in close proximity to the Tongan individual LHA001—slightly more towards the East Asian cline—suggesting that the population expansion to East Polynesia ~900–800 yr BP[40] may have originated in western Polynesia. ADMIXTURE analyses ($K=4$) on a subset of HO data—including 454 present-day and 13 ancient Near and Remote Oceanian individuals (Supplementary Fig. 5)—show that present-day ni-Vanuatu carry a heterogeneous proportion of three major components that are maximized in Near Oceanian populations (Papuan, Baining and Bougainville), with a minor Lapita-related component (Supplementary Fig. 5). In contrast, present-day Tongans have substantial Lapita ancestry, with a minor component of Near Oceanian admixture (with different proportions of Papuan, Baining and Bougainville) (Supplementary Fig. 5). qpAdm analyses further support modelling modern Tongans as a two-way admixture between ancestral Austronesians and a population ancestral to some present-day Solomon Island groups (such as Malaita and Makira) or represented by the ~500 yr BP Malaita individual (MAI002), even when Papuan and Baining Marabu are included as an additional outgroup (Supplementary Table 10). Thus, Solomon Islanders alone can explain the Near Oceanian ancestry found in Tongans, without contribution from New Guinea Papuans. This provides evidence that, post-Lapita, Tonga probably received its Near Oceanian ancestry from a different source than did Vanuatu.

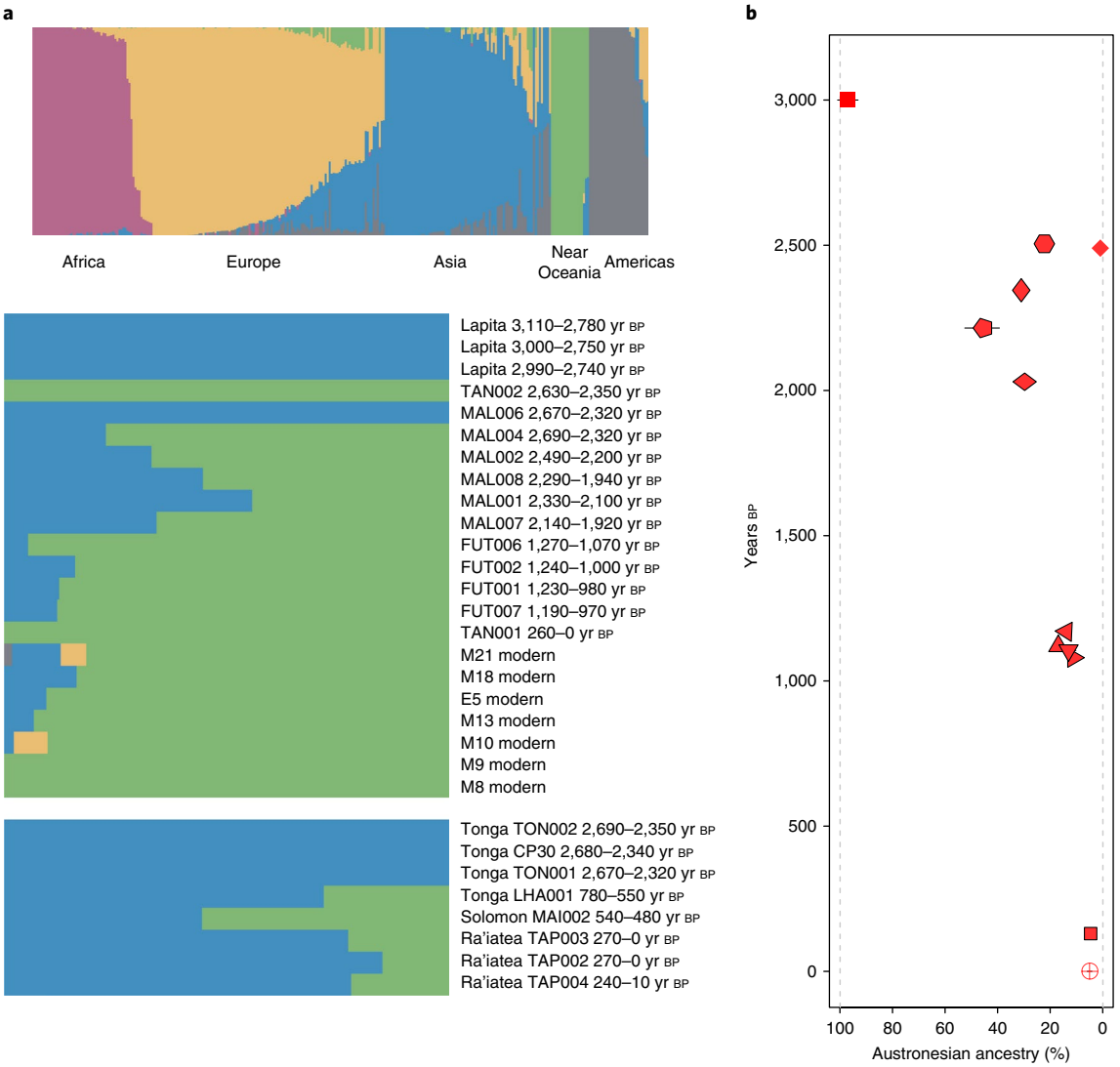


Fig. 2 | Admixture proportions of Papuan- versus Lapita-related ancestry in ancient and present-day populations using 1,240 K genome-wide data. **a**, Unsupervised ADMIXTURE analyses of present-day global populations and ancient Pacific individuals, with five ancestral components. **b**, Austronesian ancestry proportion (modelled by the indigenous Taiwanese population Ami) in ancient and present-day Vanuatu individuals estimated through qpAdm analyses. The symbols are as described in Fig. 1. Black lines indicate s.e. if larger than the symbol (see also Supplementary Table 8).

Genetic cline in present-day Vanuatu. We analysed the new ancient and modern data alongside a dataset from Remote Oceania[32], which includes 754 individuals from New Caledonia, Vanuatu, Fiji and Tonga (Supplementary Table 6), genotyped on the HumanCore-24 BeadChip, with ~160 K and ~50 K SNP overlap with the 1,240 K and HO data, respectively. After removing individuals with genetic evidence of non-autochthonous ancestry, PCA and ADMIXTURE analyses (Supplementary Figs. 6 and 7) demonstrated high genetic diversity in ni-Vanuatu from the islands of Santo and Maewo (north of Malakula; Supplementary Fig. 2), with these individuals lying on a cline running from close to New Britain, through Vanuatu, New Caledonia and Fiji, and towards present-day Tonga. The new Vanuatu HO data from the islands of Malakula and Efate (Supplementary Fig. 2), and the most recent ancient Tanna individual (TAN001), lay overwhelmingly towards the New Britain end of this cline. Downsampled to ~50 K SNPs, the different trajectories for post-Lapita Vanuatu and Tonga populations identified in the HO analyses are less distinguishable. We used *D*-statistics to test whether this cline describes a separate demographic process to that which brought Bismarck-like ancestry to Vanuatu (see Methods), but—at the resolution of currently available regional genotyping data—we are unable to distinguish between the two clines with confidence (Supplementary Fig. 8), suggesting that a Tongan-like ancestry may have played some role in the formation of present-day genetic diversity in Vanuatu. However, the HO analyses demonstrate that present-day Tongan ancestry, forming one end of this cline, was not fully in place before ~780–550 yr BP (LHA001), so this influence may be significantly later than the initial arrival of Bismarck ancestry in Malakula (~2,500 yr BP).

Austronesian-Papuan admixture date estimation. We performed ALDER[33] analyses on both modern and ancient Vanuatu data to gain independent estimates of arrival times for the Papuan ancestry component. We obtained an estimate of 60.7 ± 8.2 generations BP for the 27 HO Vanuatu individuals, which—assuming a 28.1 yr generation time[21]—equates to 1,705 ± 232 yr BP (Fig. 3b; see Methods). Interestingly, admixture time estimates similarly obtained for ancient Vanuatu provided 51.2 ± 17 generations for 3 Futuna individuals (FUT002, FUT006 and FUT007) and 5.6 ± 1.8 generations for 3 ancient Malakula individuals (MAL002, MAL004

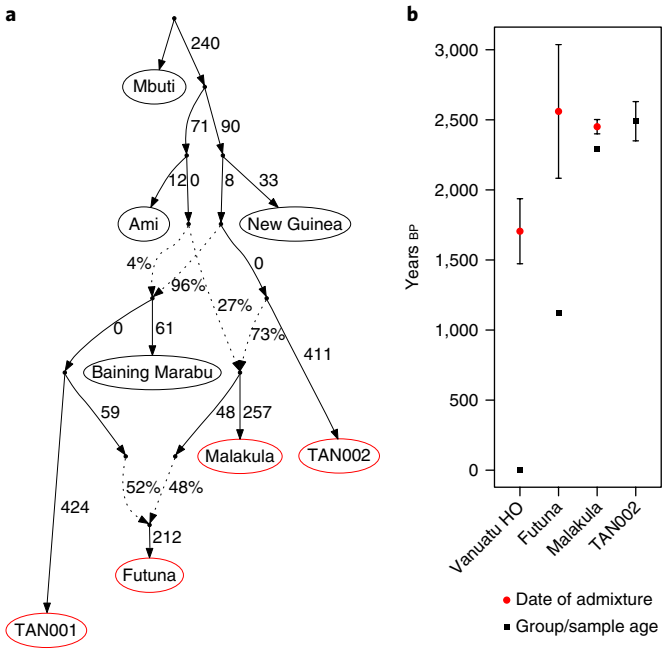


Fig. 3 | Demographic history of ancient Vanuatu individuals. a, qpGraph model that fits the observed allele frequency patterns with branch lengths representing drift in $F_{ST} \times 1{,}000$ units and edge percentages indicating admixture proportions. Ancient samples or groups are indicated with a red border. **b**, ALDER analyses estimating the date of Papuan and East Asian admixture, converted into years with a generation time of 28.1 yr. Bars show s.e. for date estimates. The sample ages for the two ancient groups (Futuna and Malakula) are averaged radiocarbon dating confidence interval midpoints. As the earliest ancient Vanuatu individual with unadmixed Near Oceanian ancestry, TAN002 is included for age comparison, with the error bar indicating the 95.4% radiocarbon dating confidence interval.

and MAL007). Accounting for ancient sample ages, the admixture date is estimated at 2,560 ± 477 yr BP for Futuna and 2,451 ± 51 yr BP for Malakula, coinciding with the latest presence of individuals in the new Vanuatu time transect with unadmixed Papuan (TAN002) or Austronesian (MAL006) ancestry (Fig. 3b). ALDER analyses of the ref. [32] data gave dates ranging from 1,569 ± 79 yr BP (Fiji) to 1,999 ± 101 yr BP (Port Olry, Vanuatu), overlapping the interval proposed by ref. [21], yet still significantly later than the directly dated admixed ancient individuals in Malakula (Supplementary Fig. 9).

Discussion

The population history of Remote Oceania is relatively short, but these early stages appear complex, particularly in Vanuatu. New genome-wide aDNA data directly demonstrate the presence of Papuan peoples in Remote Oceania far earlier than estimated with present-day regional genome-wide data[21] (Supplementary Fig. 9), with unadmixed Bismarck-like individuals apparent in Vanuatu as early as ~2,500 yr BP, possibly contemporaneous with the end of the Lapita horizon. The new HO data from contemporary Malakula and Efate show that while Oceanic-speaking Lapita peoples were genetically replaced by a population closely related to Papuan-speaking Baining people, present-day ni-Vanuatu continue to speak Oceanic languages. The almost complete replacement of a population's genetic ancestry that leaves the original languages in situ is extremely rare—possibly without precedent—in human history and requires explanation. Alongside linguistic and archaeological evidence, our aDNA analyses provide a plausible and compelling model for this language continuity, namely an extended and incremental process of population replacement by peoples from the Bismarck Archipelago (Fig. 3a), rather than a single massive turnover event that would probably have brought a shift from Oceanic to Papuan languages.

The >120 languages spoken today in Vanuatu—per capita, the most linguistically diverse place on Earth—are exclusively Oceanic[14], yet many aberrant, seemingly Papuan, linguistic features are evident[41]. These include quinary numeral systems, rounded labial phonemes, dual exclusion of *p* and *c* phonemes, and serial verb construction[42–45]. These features are heterogeneously distributed across Vanuatu[42–44], extremely rare or absent in other Austronesian languages and shared almost exclusively with Papuan languages (for example, Supplementary Fig. 10). A number of ethnographically attested cultural practices or artefacts also share this near-exclusive distribution, including large nasal piercing ornaments, penis sheaths, head binding and the rearing of full-circle tusker pigs[42,46]. These shared cultural and linguistic features provide further support for the Baining–Papuan genetic connection we identify. While some linguists argue for a single admixed expansion into Vanuatu from Near Oceania[47], or Papuan involvement in initial Lapita settlement[43], others propose a two-wave model[42], where an initial unadmixed proto-Oceanic-speaking population arrives, followed closely by a separate Papuan-speaking expansion. The two-wave model[42] is supported because the putative Papuan linguistic features found in Vanuatu cannot be reconstructed for proto-Oceanic, and their marked deviation from most other Oceanic languages suggests development within Vanuatu[42–44]. Some features can be reconstructed for the proto languages of Vanuatu—rounded labials and the *p*/*c* gap for Proto-North-Central Vanuatu[48], and quinary numeral systems for Proto-Southern Vanuatu[49]—pointing to their early development, and strongly supporting early Papuan influence. An undifferentiated proto-Oceanic operating as a lingua franca for linguistically diverse Papuan migrant groups could explain[42] the continuity of Oceanic languages in the face of secondary Papuan expansion.

Our aDNA analyses lend direct support to this historical linguistic model[42]. Indeed, some archaeologists have argued that the process by which Papuans made their way into Remote Oceania was strikingly different from the initial arrival of Lapita people[23], suggesting a continuing process of long-distance interaction rather than a simple dispersal event. One element of this process—namely, the sex-biased admixture inferred from present-day South Pacific populations (for example, refs [13,21])—is already becoming clearer, with such genetically admixed ancient individuals (for example, MAL004) observed shortly after the earliest arrival of Near Oceanian peoples in Remote Oceania (Fig. 2b and Supplementary Table 8). We show that initially genetically homogeneous Lapita peoples in Vanuatu and Tonga[21] follow strikingly different post-Lapita population trajectories, reflected in the clear cultural separation seen in the archaeological record. As a defined stylistic horizon, Lapita lasted only a few hundred years after settlement; local differentiation in pottery design beginning ~2,700 yr BP suggests significant fragmentation of the previously well-connected Lapita peoples[23]. In central Vanuatu, the appearance of the incised Erueti ceramic complex ~2,550 yr BP[50] seems to parallel a contemporaneous stylistic shift across island Melanesia post-Lapita, including both New Caledonia and the Bismarck Archipelago[3]. It is an intriguing possibility that the early arrival of Bismarck-like people we now directly observe in Vanuatu may have exacerbated—even triggered—the process of Lapita fragmentation[23] and the ongoing long-distance interactions we uncover may also have influenced the convergent processes of stylistic diversification[3,50] found in pottery sequences across the region.

Our analysis of present-day Remote Oceanian data[32] suggests a possible Tongan-like influence on the genetic diversity of present-day eastern Melanesia, with populations in northern Vanuatu, New Caledonia and Fiji lying on a cline towards modern Tonga

(Supplementary Fig. 6). Given the data resolution, we were unable to clearly distinguish this from the other cline formed by the post-Lapita population trajectory in Vanuatu (Fig. 1a), but the ancient Tongan individual LHA001 suggests that it formed later. One possibility is that this genetic structure was influenced by interactions with western Polynesia, leading to the many Polynesian outlier communities—characterized by retention of various Polynesian linguistic features, cultural practices and genetic ancestry[3]—distributed across Micronesia, New Guinea, the Solomon Islands, New Caledonia and Vanuatu. While the timing, scale and impact of this westward Polynesian migration is not yet precisely estimated, it probably coincided with the initial colonization of eastern Polynesia ~900–800 yr BP[40].

In conclusion, our analyses of Vanuatu genome-wide data—both ancient and modern—combined with linguistic and archaeological evidence, strongly support a model of interaction and incremental admixture between Lapita-Austronesian peoples and incoming Bismarck Islanders that led to an eventual population turnover, but left the pre-existing Oceanic languages in place. This multidisciplinary work has begun to uncover the complex, localized demographic processes that drove the initial colonization of the wider South Pacific and formed the enduring cultural and linguistic spheres that continue to shape the Pacific today.

Methods

Ancient and modern-day DNA processing. *Ancient DNA sampling.* All samples were processed in dedicated laboratories at the Max Planck Institute for the Science of Human History in Jena, Germany. Bone powder for DNA extraction was obtained from petrous bones by drilling the densest osseous matter around the cochlea, and from teeth by cutting at the junction between the root and crown and sampling the dental pulp. For detailed information on the analysed samples, their archaeological context and radiocarbon age, see Supplementary Information, Supplementary Tables 1 and 2, Fig. 1 and Supplementary Fig. 2.

Extraction. DNA from the 23 ancient individuals was extracted following established protocols[24]. Negative and cave bear positive controls were included. To release DNA from 50–100 mg of bone powder, a solution of 900 µl EDTA, 75 µl H_2O and 25 µl Proteinase K was added. In a rotator, samples were digested for at least 16 h at 37 °C, followed by an additional hour at 56 °C[51]. The suspension was then centrifuged and transferred into a binding buffer as previously described[24]. To bind DNA, silica columns for high volumes (High Pure Viral Nucleic Acid Large Volume Kit; Roche) were used. After two washing steps using the manufacturer's wash buffer, DNA was eluted in TET (10 mM Tris, 1 mM EDTA and 0.05% Tween) in two steps for a final volume of 100 µl.

Library preparation. For aDNA authentication and contamination estimates, screening DNA libraries were built from 20 µl of DNA extract in the absence of uracil DNA glycosylase (non-UDG libraries), following a double-stranded library preparation protocol[25]. After assessing human DNA contamination levels, one or two additional 25 µl aliquots of DNA extract were transformed into either non-UDG libraries[25] or 'UDG-half' double-stranded libraries with a protocol that makes use of the UDG enzyme to reduce, but not eliminate, the amount of deamination-induced damage towards the ends of aDNA fragments[26]. Negative and positive controls were carried out alongside each experiment. Libraries were quantified using the IS7 and IS8 primers[25] 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[52] in 1–4 parallel 100 µl reactions using PfuTurbo DNA Polymerase (Agilent). 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[25] using the quantification method described above. Some 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 Illumina platforms.

Enrichment. Both UDG-half- and non-UDG-treated libraries were further amplified with IS5/IS6 primers to reach a concentration of 200–400 ng µl^{-1} as measured on a NanoDrop spectrophotometer (Thermo Fisher Scientific). mtDNA capture[27] was performed on screened libraries that, after shotgun sequencing, showed the presence of aDNA, highlighted by the typical CtoT and GtoA substitution pattern towards 5′ and 3′ molecule ends, respectively. Furthermore, samples with a percentage of human DNA in shotgun data around 0.1% or greater were enriched[53] for a list of 1,237,207 targeted SNPs across the human genome (1,240 K capture)[28].

Sequencing. The enriched DNA product was sequenced on an Illumina HiSeq 4000 instrument with 75 single-end-run cycles or 50 pair-end-run cycles (for TAN001 and FUT006) using the manufacturer's protocol. The output was de-multiplexed using bcl2fastq version 2.17.1.14 and dnaclust version 3.0.0.

Modern DNA sampling. Genetic sampling was carried out as part of a long-term linguistic and anthropological fieldwork project, directed by R. Gray and H. Colleran at the Max Planck Institute for the Science of Human History (http://www.shh.mpg.de/456217/vanuatu-languages-lifeways). The saliva samples of 27 present-day ni-Vanuatu from the islands of Malakula and Efate were collected using the Oragene OG-500 saliva collection kit. Ethical approval for this work was granted by the Ethik-Kommission der Friedrich-Schiller-Universität in Jena, Germany, and we obtained research permission from the Vanuatu Kaljoral Senta, the institution that regulates all research in the country. Sampling was carried out in five communities that are already participating in the linguistic and anthropological project, and all participants gave documented informed consent and were provided the means to withdraw from the study if required.

Modern DNA extraction and library preparation. Extraction and library preparation were performed in the molecular biology laboratories of the Max Planck Institute for the Science of Human History in Jena, Germany. Modern-day DNA was extracted from the Oragene kit following the manufacturer's protocols with slight modification (that is, 600 µl of sample volume was taken, and 10 µl of 8 modern-day DNA extracts (Supplementary Table 5) were used to build the double-stranded DNA libraries[25]). They were then indexed in one reaction following the same protocols mentioned above, pooled equimolarly and shotgun sequenced on an Illumina HiSeq 4000 instrument (75 single-end-run cycles).

Genotyping of present-day humans. The company ATLAS Biolabs in Berlin, Germany genotyped 27 modern DNA extracts on the Axiom Genome-Wide Human Origins Array. After checking DNA quality and quantity on both a 1% agarose gel and a NanoDrop, samples were adjusted to 20 ng µl^{-1} using a Qubit high sensitivity kit (Thermo Fisher Scientific), loaded on the Axiom Genome-Wide Human Origins Array (Affymetrix) and genotyped on a GeneTitan. Genotyping was performed using the Affymetrix Genotyping Console, and all individuals had >94% genotyping completeness.

Genomic data processing. Preprocessing of the sequenced reads was performed using EAGER version 1.92.44 (ref. [54]). Reads resulting from the sequencing of modern and ancient DNA libraries were clipped to remove residual adaptor sequences using Clip&Merge[54] and AdapterRemoval version 2 (ref. [55]), respectively. Clipped sequences were then mapped against the human reference genome hg19 using the Burrows–Wheeler Aligner (BWA)[56] turning seeding off and with the *–n* parameter set to 0.01. Duplicates were removed with DeDup[54], which removes reads with identical start and end coordinates. Additionally, a mapping quality filter of 30 was applied using SAMtools[57]. Alignment files were filtered for reads showing the presence of probable deaminated bases as the result of postmortem damage (PMD) using PMDtools version 0.60 (ref. [58]). Both damage-restricted and non-restricted sequences from either non-UDG or UDG-half libraries were trimmed for the first and last three positions to reduce the impact of deamination-induced missincorporations during genotyping. Trimmed reads were genotyped using pileupCaller (https://github.com/stschiff/sequenceTools/tree/master/src-pileupCaller)—a tool that randomly draws one allele at each of the 1,240 K-targeted SNPs covered at least once. The generated pseudo-haploid calls for 19 ancient Pacific individuals (Table 1) were merged to a pulldown of the 1,240 K SNPs from the Simons Genome Diversity Project[31], 8 shotgun-sequenced modern-day individuals from Vanuatu and 4 previously published 1,240 K captured individuals associated with the Lapita culture from Vanuatu and Tonga[21]. Moreover, the newly generated capture data for the ancient individuals as well as 27 genotyped modern-day individuals (Supplementary Table 5) were merged to the ~600 K SNPs of the HO dataset[21,30].

Authentication of ancient DNA. In the field of aDNA, several methods have been developed to assess the authenticity of the retrieved DNA[29]. First, the typical features of aDNA were inspected using DamageProfiler (https://bintray.com/apeltzer/EAGER/DamageProfiler); for example, short average fragment length (~40–70 base pairs) and an increased proportion of miscoding lesions due to deamination at the molecule termini (Supplementary Table 3). Sex determination was performed by comparing the coverage on the targeted X-chromosome SNPs (~50 K positions within the 1,240 K capture) normalized by the coverage on the targeted autosomal SNPs to the coverage on the Y-chromosome SNPs (~30 K), again normalized by the coverage on the autosomal SNPs[59] (Table 1). Individuals falling in an intermediate position between male and female were assigned to undetermined sex and indicate the presence of present-day DNA contamination. For male individuals, ANGSD was run to measure the rate of heterozygosity of polymorphic sites on the X-chromosome after accounting for sequencing errors

in the flanking regions[60]. This provides an estimate of nuclear contamination in males that are expected to have only one allele at each site. For all male samples that exhibited X-chromosome contamination levels below 2% with at least 100 X-chromosome SNPs covered twice, all reads were retained for further analyses (Supplementary Table 4). Otherwise, only PMD fragments that were likely to be of endogenous origin were used[61] (Table 1). For both male and female individuals, mtDNA-captured data were used to jointly reconstruct the mtDNA consensus sequence and estimate contamination levels with schmutzi[62] (Supplementary Table 11). For specimens where a relatively low proportion of mtDNA molecules compared with nuclear DNA was observed (Supplementary Table 11), mtDNA contamination estimates could be used as reliable predictors for nuclear contamination[29]. Population genetic analyses on samples presenting mtDNA levels of contamination above 4% were restricted to PMD fragments. Moreover, for each individual, the positioning in PCA space was compared with the data after restriction to deaminated sequences[21]. Samples that were substantially displaced in PCA space (Supplementary Fig. 1) were restricted to PMD fragments for population genetic analyses.

Population genetic analyses. PCAs were computed with present-day populations from the HO dataset composed of 781 Oceanians and East Asians[21] and 27 modern-day Vanuatu individuals newly genotyped here, for a total of 808 individuals. Ancient individuals were projected onto the two first components using smartpca (version 13050)[63] with the options 'lsqproject: YES' and 'numoutlieriter: 0' (Fig. 1 and Supplementary Fig. 1). Another PCA was computed on the ~50 K SNPs overlapping the HO dataset and a recently published Illumina HumanCore-24 dataset (typed on ~240 K SNPs in total)[32] (Supplementary Fig. 6). The same 808 modern-day Oceanians and East Asians were used to build the principal components on which 669 individuals across Remote Oceania (Supplementary Table 6) and 15 ancient Pacific individuals with more than 6 K SNPs were projected. The software ADMIXTURE version 1.3.0 (ref. [34]) was run in unsupervised mode on the high-coverage genomes of 308 modern-day worldwide individuals[31], 8 shotgun-sequenced present-day Vanuatu individuals and all 23 ancient Pacific individuals. Only transversion sites of the 1,240 K SNPs (~220 K positions) were considered to reduce the impact on the clustering algorithm of residual damage still present in non-UDG-treated libraries. An additional regional ADMIXTURE analysis was carried out on the transversions subset of the HO data (~110 K SNPs), including 13 ancient individuals from Vanuatu and Tonga (more than 15 K SNPs) and 454 modern-day Oceanian individuals (Supplementary Fig. 5). Finally, ADMIXTURE was run on the overlapping SNPs between the HO and ref. [32] datasets for the 27 newly genotyped present-day individuals from Malakula and Efate in Vanuatu (Supplementary Table 5), in addition to 754 present-day individuals from New Caledonia, Vanuatu, Fiji and Tonga (Supplementary Fig. 7). From the dataset of the latter four countries, 85 individuals harbouring more than 2% of non-local ancestry at $K=5$ were removed for a total of 669 individuals retained (Supplementary Table 6). In the following analyses all SNPs were investigated for individuals with UDG-half libraries, whereas only transversion SNPs were used for individuals with non-UDG libraries to avoid spurious results originating from leftover aDNA damage.

D-statistics were calculated using the qpDstats version 711 programme from the ADMIXTOOL suite (https://github.com/DReichLab) in the form 'D(Pop1, Pop2; Pop3, Outgroup)'. A negative value implies that either Pop1 and Outgroup, or Pop2 and Pop3 share more alleles than expected under the null hypothesis of a symmetrical relationship between Pop1 and Pop2 (Supplementary Table 9). To jointly observe the affinity of modern-day Fiji, Tonga, New Caledonia and Vanuatu individuals from the ref. [32] and HO datasets, as well as ancient Vanuatu individuals towards Ami and Tonga populations, we calculated two sets of *D*-statistics in the form (1) 'D(Baining, X; Ami, Mbuti)' and (2) 'D(Baining, X; modern Tongan, Mbuti)', where *X* is drawn from Fiji, Tonga, Maewo (Vanuatu), Port Olry (Vanuatu), Santo (Vanuatu) and New Caledonia from ref. [32], as well as the Vanuatu HO and ancient Malakula, Futuna and Tanna samples. Plotting (1) against (2) (Supplementary Fig. 8) shows that we cannot see a clear deviation between modern and ancient individuals, as all values do not appreciably differ from the straight line expected for no differential ancestry.

qpWave version 400 (ref. [35]) was implemented on the HO dataset to test whether the ancient individuals are consistent with two sources of ancestry represented by modern-day Ami (as the best proxy for ancestral Austronesian) and Papuan individuals, with respect to a set of outgroups (Mbuti, Denisovan, Sardinian, English, Yakut, Chukchi, Mala, Japanese, Ju_hoan_North, Mixe, Onge and Yoruba). This is obtained when rank $n-1$ cannot be rejected ($P>0.05$), as shown for all our ancient Vanuatu individuals, as well as modern Vanuatu HO individuals despite a much lower *P* value (Supplementary Table 7). The same populations for both the HO and 1,240 K datasets were then used in qpAdm version 610 (ref. [36]) to estimate admixture proportions for ancient and modern-day Vanuatu individuals (Supplementary Fig. 3, Fig. 2b and Supplementary Table 8). qpAdm models each individual as a mixture of Ami and Papuan by fitting admixture proportions that match the observed matrix of f4-statistics and computing standard errors with a block jackknife. To evaluate potential sex bias admixture, qpAdm analysis, as described above, was run only on X-chromosome SNPs (option 'chrom:23') of the 1,240 K dataset. Differences in admixture proportions between autosomal and X-chromosome SNPs provide an indication of sex-biased admixture (Supplementary Table 8).

Modern-day Tongans were modelled in qpAdm as resulting from a two-way admixture between Ami (as the best proxy for ancestral Austronesian) and ancient (MAI002) or modern-day Solomon Islanders from the islands of Makira, Malaita and Bougainville (Naisoi and Choiseul populations). When selecting the 12 outgroups listed above, Tongans can successfully be modelled with $P>0.05$, using a block jackknife to calculate standard errors as indicated previously. qpAdm was re-run expanding the outgroup population list with Papuan and Baining Marabu. For present-day individuals from Makira and Malaita and the ancient individual from Malaita (MAI002), rank $n-1$ can still not be rejected, indicating that additional Papuan New Guinea or Baining ancestry is not necessary to model modern-day Tongans (Supplementary Table 10).

Admixture dates were estimated based on linkage disequilibrium using ALDER[33] on the ~160 K overlapping SNPs between the 1,240 K capture and ref. [32] datasets. As source populations, 20 Asian (Ami, Atayal, Igorot, Kinh, Dai, She, Lahu and Han) and 16 Papuan individuals were chosen. The estimated dates of admixture were converted into years assuming a generation time of 28.1 yr[21,64] for the 27 Vanuatu HO individuals (Fig. 3b) and for modern-day New Caledonia, Vanuatu, Fiji and Tonga populations[32] (Supplementary Fig. 9). Admixture dates were also estimated for SNPs overlapping with the 1,240 K capture for three ancient Futuna individuals (FUT002, FUT006 and FUT007) with an average age set to 1,123 yr BP and three ancient Malakula individuals (MAL002, MAL004 and MAL007) with an average age set to 2,293 yr BP (Fig. 3b).

Admixture graphs on the HO dataset were fitted with qpGraph version 5211 (refs [30,65]), which matches a matrix of *f*-statistics testing the relationships between all analysed populations at the same time. An initial backbone graph of modern-day populations without signs of admixture was built into the tree (Mbuti, Ami and New Guinea). The differential proportion of Denisovan ancestry between Mbuti-Ami and New Guinea populations[66] was not modelled here since this is accommodated in the graph by shifting the splitting point of the African Mbuti population. Baining Marabu was then incorporated as admixed between an Ami-related and New Guinea-related lineage, as suggested from *D*-statistics analyses (Supplementary Table 9). Ancient UDG-half individuals from Vanuatu (three Futuna individuals grouped, three Malakula individuals grouped and two Tanna individuals separately) were added chronologically one by one at each possible position of the graph, reporting every time the highest *D*-statistic between the observed and fitted model and calculating the *Z*-score with a block jackknife. The graph reported in Fig. 3a is built with a total of 38,789 SNPs and fits the allele frequency relationships between modern-day and ancient individuals with all empirical *f*-statistics within the 3 s.e. interval and only one significant *D*-statistic ($Z=2.6$). The modern-day Vanuatu HO population can be fitted as admixed between modern-day Baining Marabu and Ami-related populations, but this relatively simple model with only four populations already has the worst *Z*-score, equal to 2.3 (Supplementary Fig. 4a). Moreover, we were unable to fit a modern-day HO Vanuatu population in the graph once ancient individuals are included, neither by replacing the ~200 yr BP TAN001 individual (Supplementary Fig. 4b), nor by modelling Vanuatu HO as deriving part of its ancestry from the ~1,100 yr BP Futuna population (Supplementary Fig. 4c), with worst *Z*-scores of 6 and 5.2, respectively.

Haplogroup assignment for uniparental markers. After enrichment of the libraries for the mitochondrial genome (mtDNA capture), reads were preprocessed in EAGER version 1.92.55 as described above and aligned to the mitochondrial reference genome (rCRS) using CircularMapper, a programme that takes into account the circularity of the mtDNA[54]. Contamination estimation and mitochondrial genome assembly were jointly performed using schmutzi[62] with the parameters '--notusepredC –uselength'. Present-day human contamination estimates were performed using a comparative database of 197 modern-day worldwide mtDNAs provided with the software package. For the resulting sequences, we filtered positions with likelihoods above 20 or 30 (Supplementary Table 11) and used HaploGrep2 (ref. [67]) to assign the corresponding mtDNA haplogroup. For the FUT007 individual, the mtDNA consensus sequence was reconstructed from the mtDNA off-target reads in the combined non-UDG and UDG-half 1,240 K capture data (Table 1 and Supplementary Table 11). Sequenced reads overlapping the Y-chromosome SNPs present in the International Society of Genetic Genealogy database version 11.349 (http://www.isogg.org/tree) were investigated to assign Y-chromosome haplogroups. ANGSD[60] was used to count ancestral and derived allele occurrence and perform a majority call for positions covered at least once. For this analysis, UDG-half and no-UDG data were combined for each sample (Supplementary Table 3). To avoid missassignments due to DNA damage, CtoT and GtoA mutations required a minimum of two consistent nucleotides to be called. Haplogroup assignment was based on the most downstream SNP retrieved after evaluating the presence of upstream mutations along the related haplogroup phylogeny[59].

Life Sciences Reporting Summary. Further information on experimental design is available in the Life Sciences Reporting Summary.

Data availability. All newly reported ancient DNA data, including nuclear DNA and mtDNA alignment sequences, are archived in the European Nucleotide

Archive database (accession number PRJEB24810). Newly reported SNP genotyping and shotgun-sequenced data will be made available on request to H.C. and A.P., subject to a signed agreement to restrict usage to anonymized non-medical studies of population history, as outlined in the ethics and consent documentation.

Received: 9 January 2018; Accepted: 12 February 2018;
Published online: 27 February 2018

References

1. Clarkson, C. et al. Human occupation of northern Australia by 65,000 years ago. *Nature* **547**, 306–310 (2017).
2. Pawley, A. & Green, R. Dating the dispersal of the Oceanic languages. *Ocean. Linguist.* **12**, 1–67 (1973).
3. Kirch, P. V. *On the Road of the Winds: An Archaeological History of the Pacific Islands Before European Contact* (Univ. California Press, Berkeley, 2017).
4. Blust, R. A. The prehistory of the Austronesian-speaking peoples: a view from language. *J. World Prehist.* **9**, 453–510 (1995).
5. Gray, R. D., Drummond, A. J. & Greenhill, S. J. Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* **323**, 479–483 (2009).
6. Blust, R. A. *The Austronesian Languages* (Australian National Univ., Canberra, 2009).
7. Summerhayes, G. R. et al. Tamuarawai (EQS): an early Lapita site on Emirau, New Ireland, PNG. *J. Pac. Archaeol.* **1**, 62–75 (2010).
8. Sheppard, P. J. Lapita colonization across the Near/Remote Oceania boundary. *Curr. Anthropol.* **52**, 799–840 (2011).
9. Pugach, I. et al. The gateway from Near into Remote Oceania: new insights from genome-wide data. *Mol. Biol. Evol.* https://doi.org/10.1093/molbev/msx333 (2018).
10. Petchey, F. J., Spriggs, M., Bedford, S., Valentin, F. & Buckley, H. Radiocarbon dating of burials from the Teouma Lapita cemetery, Efate, Vanuatu. *J. Archaeol. Sci.* **50**, 227–242 (2014).
11. Sand, C. *Lapita Calédonien. Archéologie d'un Premier Peuplement Insulaire Océanien* (Société des Océanistes, Paris, 2010).
12. Burley, D., Weisler, M. I. & Zhao, J.-X. High precision U/Th dating of first Polynesian settlement. *PLoS ONE* **7**, e48769 (2012).
13. Wollstein, A. et al. Demographic history of Oceania inferred from genome-wide data. *Curr. Biol.* **20**, 1983–1992 (2010).
14. Pawley, A., Attenborough, R., Golson, J. & Hide, R. *Papuan Pasts: Cultural, Linguistic and Biological Histories of Papuan-Speaking Peoples* (Australian National Univ., Canberra, 2005).
15. Lynch, J., Ross, M. & Crowley, T. *The Oceanic Languages* (Curzon, Richmond, 2002).
16. Kayser, M. et al. Genome-wide analysis indicates more Asian than Melanesian ancestry of Polynesians. *Am. J. Hum. Genet.* **82**, 194–198 (2008).
17. Melton, T. et al. Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. *Am. J. Hum. Genet.* **57**, 403–414 (1995).
18. Kayser, M. et al. Melanesian origin of Polynesian Y chromosomes. *Curr. Biol.* **10**, 1237–1246 (2000).
19. Hurles, M. E. et al. Y-chromosomal evidence for the origins of Oceanic-speaking peoples. *Genetics* **160**, 289–303 (2002).
20. Kayser, M. et al. Melanesian and Asian origins of Polynesians: mtDNA and Y chromosome gradients across the Pacific. *Mol. Biol. Evol.* **23**, 2234–2244 (2006).
21. Skoglund, P. et al. Origins and genetic legacy of the first people in Remote Oceania. *Nature* **538**, 510–513 (2016).
22. Valentin, F., Détroit, F., Spriggs, M. & Bedford, S. Early Lapita skeletons from Vanuatu show Polynesian craniofacial shape: implications for Remote Oceanic settlement and Lapita origins. *Proc. Natl Acad. Sci. USA* **113**, 292–297 (2015).
23. Bedford, S. & Spriggs, M. in *The Oxford Handbook of Prehistoric Oceania* (eds Cochrane, E. & Hunt, T.) Ch. 8 (Oxford Univ. Press, Oxford, 2014).
24. Dabney, J. et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110**, 15758–15763 (2013).
25. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, pdb.prot5448 (2010).
26. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. B* **370**, 20130624 (2015).
27. Fu, Q. et al. A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr. Biol.* **23**, 553–559 (2013).
28. Fu, Q. et al. An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216–219 (2015).
29. Key, F. M., Posth, C., Krause, J., Herbig, A. & Bos, K. I. Mining metagenomic data sets for ancient DNA: recommended protocols for authentication. *Trends Genet.* **33**, 508–520 (2017).
30. Patterson, N. et al. Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
31. Mallick, S. et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201–206 (2016).
32. Parks, T. et al. Association between a common immunoglobulin heavy chain allele and rheumatic heart disease risk in Oceania. *Nat. Commun.* **8**, 14946 (2017).
33. Loh, P. R. et al. Inferring admixture histories of human populations using linkage disequilibrium. *Genetics* **193**, 1233–1254 (2013).
34. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
35. Reich, D. et al. Reconstructing Native American population history. *Nature* **488**, 370–374 (2012).
36. Haak, W. et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
37. 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).
38. Mirabal, S. et al. Increased Y-chromosome resolution of haplogroup O suggests genetic ties between the Ami aborigines of Taiwan and the Polynesian Islands of Samoa and Tonga. *Gene* **492**, 339–348 (2012).
39. Oliver, D. L. *Ancient Tahitian Society: Ethnography* (Univ. Press of Hawaii, Honolulu, 1974).
40. Wilmshurst, J. M., Hunt, T. L., Lipo, C. P. & Anderson, A. J. High-precision radiocarbon dating shows recent and rapid initial human colonization of East Polynesia. *Proc. Natl Acad. Sci. USA* **108**, 1815–1820 (2011).
41. Blust, R. Review of Lynch, Ross, and Crowley, "*The Oceanic Languages*". *Ocean. Linguist.* **44**, 544–548 (2005).
42. Blust, R. Remote Melanesia: one history or two? An addendum to Donohue and Denham. *Ocean. Linguist.* **47**, 445–459 (2008).
43. Donohue, M. & Denham, T. The language of Lapita: Vanuatu and an early Papuan presence in the Pacific. *Ocean. Linguist.* **47**, 365–376 (2008).
44. Lynch, J. Melanesian diversity and Polynesian homogeneity: the other side of the coin. *Ocean. Linguist.* **20**, 95–129 (1981).
45. Tryon, D. T. in *Melanesia: Beyond Diversity* (eds May, R. J. & Nelson, H.) 241–248 (Australian National Univ., Canberra, 1982).
46. Speiser, F. *Ethnology of Vanuatu. An Early Twentieth Century Study* (Crawford House Press, Bathurst, 1996).
47. Pawley, A. Explaining the aberrant Austronesian languages of Southeast Melanesia: 150 years of debate. *J. Polyn. Soc.* **115**, 215–258 (2006).
48. Clark, R. *Leo Tuai: a Comparative Lexical Study of North and Central Vanuatu Languages* (Australian National Univ., Canberra, 2009).
49. Lynch, J. *The Linguistic History of Southern Vanuatu* (Australian National Univ., Canberra, 2001).
50. Bedford, S. *Pieces of the Vanuatu Puzzle: Archaeology of the North, South and Centre* Vol. 23 157–192 (Australian National Univ., Canberra, 2006).
51. Rohland, N. & Hofreiter, M. Ancient DNA extraction from bones and teeth. *Nat. Protoc.* **2**, 1756–1762 (2007).
52. Kircher, M., Sawyer, S. & Meyer, M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* **40**, e3 (2012).
53. Fu, Q. et al. DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl Acad. Sci. USA* **110**, 2223–2227 (2013).
54. Peltzer, A. et al. EAGER: efficient ancient genome reconstruction. *Genome Biol.* **17**, 60 (2016).
55. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88 (2016).
56. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
57. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
58. Skoglund, P. et al. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc. Natl Acad. Sci. USA* **111**, 2229–2234 (2014).
59. Fu, Q. et al. The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
60. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).
61. Meyer, M. et al. A mitochondrial genome sequence of a hominin from Sima de los Huesos. *Nature* **505**, 403–406 (2014).
62. Renaud, G., Slon, V., Duggan, A. T. & Kelso, J. schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* **16**, 224 (2015).
63. Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
64. Moorjani, P. et al. A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. *Proc. Natl Acad. Sci. USA* **113**, 5652–5657 (2016).
65. Reich, D., Thangaraj, K., Patterson, N., Price, A. L. & Singh, L. Reconstructing Indian population history. *Nature* **461**, 489–494 (2009).

66. Reich, D. et al. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468**, 1053–1060 (2010).
67. Weissensteiner, H. et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–W63 (2016).

Acknowledgements

We thank the communities in Malakula and Efate in Vanuatu who participated in this study, and particularly all sample donors. We are grateful to M. Stoneking, I. Pugach and C.-C. Wang for comments, and to G. Brandt, R. Bianco and technicians at the Max Planck Institute for the Science of Human History for laboratory support. This research was supported by the Max Planck Society. Archaeological investigations on Malakula, Vanuatu were funded by the Sasakawa Pacific Island Nations Fund, the Marsden Fund of the Royal Society of New Zealand (Fast-Start 9011/3602128; 04-U00–007), a National Geographic Scientific Research grant (7738–04) and an Australian Research Council Discovery Project grant (DP0880789). Investigations on Tanna, Vanuatu were supported by an Australian Research Council Discover Project grant (DP160103578). F.V. is funded by CNRS-UMR 7041, H.B. is funded by the Marsden Fund of the Royal Society of New Zealand (Standard Grant UOO0917) and a University of Otago Research Grant, and A.P. is funded by European Research Council Starting Grant 'Waves' (ERC758967).

Author contributions

F.V., S.B., R.S., H.B., R.K., G.R.C., C.R., J.F., T.M., J.M., J.G. and L.K. contributed archaeological material. H.C., K.W.K. and A.P. contributed the 27 present-day Vanuatu samples. J.Z., F.P. and P.R. contributed isotopic data and radiocarbon date calibrations. M.W. and R.D.G. contributed linguistic interpretation. F.V., S.B., J.M., F.P. and P.R. contributed text in the Supplementary Information. K.J.R., K.A., S.J.O., A.V.S.H. and A.J.M. contributed geographical labels for the ref. [32] samples. C.P. and K.N. performed ancient DNA laboratory work. C.P., K.N., C.J. and A.P. performed population genetic analyses. C.P., K.N., H.C. and A.P. wrote the paper with input from F.V., S.B., H.B., M.W., F.P., P.R., C.J., R.D.G. and J.K. C.P. and A.P. created the figures. The study was conceived and coordinated by C.P., K.N., H.C., R.D.G., J.K. and A.P.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-018-0498-2.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to C.P. or H.C. or J.K. or A.P.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims inpublished maps and institutional affiliations.

7.3 Manuscript C: Ancient Genetic Diversity in Near Oceania - insights from coastal New Guinea and the Bismarck Archipelago.

Kathrin Nägele, Rebecca Kinaston, Selina Carlhoff, Dylan Gaffney, Emilie Bertolini, Monica Tromp, Rita Radzeviciute, Glenn Summerhayes, Fiona Petchey, Dimitri Anson, Peter Petchey, Hallie Buckley, Johannes Krause, Cosimo Posth, Adam Powell.

Draft Manuscript.

Despite their essential role in the population history of adjacent regions, the ancient genetic diversity of Near Oceania remains unstudied. The cultural and linguistic diversity of present-day populations in the regions suggest complex interactions in the past. Until recently, the consensus among Pacific archaeologists was that the Lapita cultural complex, as part of the Austronesian expansion, was restricted to smaller, offshore islands and archipelagos. The discovery of Lapita pottery in Caution Bay on the south-western coast of New Guinea has changed this view. Dating to around 3000 BP, evidence of occupation on the mainland of Papua New Guinea is as old as in eastern Near Oceania, where a Lapita burial ground on Vanuatu dates to 2900 BP. The linguistic and cultural diversity, both today and in the archaeological record, implies a complex population history of isolation and exchange. Additionally, previous research has shown the importance of Near Oceanic populations in the formation of the present-day genetic make-up of Near and Remote Oceanic Islanders. However, the genetic variation in the region is solely studied in present-day populations, and no ancient sequences have been analysed to date.

To investigate the ancient genetic diversity, we produced ancient genomes of 41 individuals excavated from four different sites across Near Oceania. The majority of individuals were excavated in Papua New Guinea on the south coast, and are dated between 500 and 150 BP, complemented by two genomes from the north coast dated to ~700 BP. The genetic results show a mixture of highland Papuan- and Asian-related ancestry, however diverse in respect to the proportions between the different sites. The admixture events are dated to 1500 BP and 1000 BP, revealing more differences between the sites. The results suggest different interactions with

inland and island populations, resulting in the mosaic of cultural and linguistic diversity on the south coast today. Five individuals from Watom Island in the Bismarck Archipelago cover a period of 3000 years and show occupation of the island by people with Papuan-related ancestry before the occurence of the Lapita cultural complex on the island. One individual dating to 2100 BP shows admixture of East Asian-related ancestry and local Papuan-related ancestry. The event of the admixture is dated to 2300 BP, which postdates the formation of the Lapita cultural complex on the Bismarck Islands by ~1000 years and the first settlement of eastern Near Oceania by ~600 years. The findings support previous statements in archaeogenetic research, which suggest the arrival of unadmixed East Asian-related populations was followed by a later arrival of and mixture with Papuan-related populations.

DRAFT MANUSCRIPT – do not circulate

Title: Ancient Genetic Diversity in Near Oceania - Insights from coastal New Guinea and the Bismarck Archipelago.

Autors: Kathrin Nägele[1*], Rebecca Kinaston[2], Selina Carlhoff[3], Dylan Gaffney[5,6], Hallie Buckley[3], Emilie Bertolini[3], Monica Tromp[5,6], Rita Radzeviciute[1], Fiona Petchey[9,10], Dimitri Anson[7], Peter Petchey[7], Glenn Summerhayes[7,8], Johannes Krause[1,2], Cosimo Posth*[3,11] , Adam Powell*[1]
*Corresponding authors

Affiliations:
[1] Department of Archaeogenetics, Max-Planck-Institute for Evolutionary Anthropology, Leipzig, Germany
[2] Department of Anatomy, School of Medical Sciences, University of Otago, Dunedin, New Zealand
[3] Department of Archaeogenetics, Max-Planck-Institute for the Science of Human History, Jena, Germany
[4] Department of Archaeology, Max-Planck-Institute for the Science of Human History, Jena, Germany
[5] Department of Archaeology, University of Cambridge, Cambridge, UK
[6] Magdalene College, University of Cambridge, Cambridge, UK
[7] Department of Anthropology and Archaeology, University of Otago, Dunedin, New Zealand
[8] School of Social Science, University of Queensland, St.Lucia, Australia
[9] DARC Centre of Excellence for Australian Biodiversity and Heritage, Canberra Australia
[10] Radiocarbon Dating Laboratory, University of Waikato, Hamilton, New Zealand.
[11]Institute for Archaeological Sciences, Archaeo- and Palaeogenetics, University of Tübingen, Tübingen, Germany

Abstract:

The ancient inhabitants of Near Oceania have played an essential role in the population history of Oceania and the adjacent regions. However, inferences about the genetic population history are solely made from present-day genomes, with no ancient sequences available from the region to date. Present-day coastal populations of Papua New-Guinea (PNG) harbour Asian-related ancestry, raising questions regarding the source of this genetic component and the date of admixture. In this study, we analyse whole genome sequences of 41 individuals from three archaeological sites from mainland PNG and one site from the Bismarck Archipelago (Watom Island), dating between 500 – 50 years before present (BP) and 3800 – 500 BP, respectively. We observe varying amounts of Asian-related ancestry in the coastal groups of mainland PNG ranging from 15% - 60% and genetically infer the admixture date with Papuan-related ancestry to around 1500 – 1000 BP for most individuals. Two geographically close sites in southern PNG show ancestry patterns diverging during the late prehistoric periods, revealing different interaction-spheres with coastal and inland populations.

Our time transect in the Bismarck Archipelago reveals the presence of people with Papuan-related ancestry on Watom Island dating from 3800-2600 BP, and 40% admixture with Early Remote Oceanian-related ancestry inferred around 2300 BP, postdating the earliest occurrence of the Lapita cultural complex by 1000 years. The date of admixture inferred indicates that genetic exchange between Asian-related and Papuan-related populations took place after the initial settlement of western Remote Oceania.

Introduction:

Near Oceania, comprising New Guinea, the Bismarck Archipelago and the Solomon Islands, has been occupied by humans since at least 45,000 years before present (BP) (Groube, Chappell et al. 1986, Allen, Gosden et al. 1988, O'Connell and Allen 2004, Summerhayes, Leavesley et al. 2010). Despite the critical role its occupants have played in the genetic history of the adjacent regions such as Indonesia (Stoneking and Delfin 2010) (Oliveira, Nägele and Carlhoff et al. forthcoming) and Remote Oceania (Kayser, Brauer et al. 2000, Wollstein, Lao et al. 2010, Lipson,

Skoglund et al. 2018, Posth, Nägele et al. 2018), its past genetic diversity remains unstudied. In fact, no ancient genomes from the region are available to date.

Apart from being the stage of one of the earliest maritime dispersals by anatomically modern humans, Near Oceania has also been host to of one of the most recent major dispersals in human history. As part of the Austronesian expansion throughout Island Southeast Asia, starting around 5000 BP (Bellwood and Dizon 2005, Bellwood, Fox et al. 2006, Gray, Drummond et al. 2009), people arrived in the Bismarck Archipelago by 3300 BP (Kirch 2001, Summerhayes, Matisoo-Smith et al. 2010). They are recognised in the archaeological record by distinct ornate pottery and a lifestyle that incorporated horticulture, domestic animals, maritime and terrestrial foraging and seafaring and likely spoke Austronesian languages. This combination of traits is today known as the Lapita cultural complex (Green 1991). The linguistic diversity today includes languages of the Austronesian language family and of various non-Austronesian (i.e Papuan) language families (Hammarström, Forkel et al. 2019). Archaeological analyses suggest this diversity is the result of high mobility and repeated colonization of the coasts by Austronesian speaking people (Dutton 1978, Allen, Holdaway et al. 1997, Allen 2010) and their interactions with local inhabitants, but the nature and extent of their interactions remains unknown (Summerhayes and Allen 2007).

Genetic analyses show the settlement of western Remote Oceania, especially Vanuatu and Tonga, starting ~3250BP in the Bismarck Archipelago (David, McNiven et al. 2011, Petchey, Spriggs et al. 2014), took place without substantial genetic admixture with populations with Papuan-related ancestry, who had resided in Near Oceania for tens of millennia before (Skoglund, Posth et al. 2016). Shortly after initial settlement of Remote Oceania, the first settlers mixed with people of Papuan-related ancestry to form the genetic composition of present-day groups (Lipson, Skoglund et al. 2018, Posth, Nägele et al. 2018). The geographical source of this Papuan-related ancestry was identified in present-day populations from the Bismarck Archipelago. The population most closely resembling the Papuan ancestry reaching Remote Oceania is Baining (Lipson, Skoglund et al. 2018, Posth, Nägele et al. 2018), a population grouped by the use of Baining, a Papuan (i.e. non-Austronesian) language family on New Britain (Friedlaender, Friedlaender et al.

2008). Additionally, Western Polynesia shows additional influences from the Solomon Islands (Lipson, Skoglund et al. 2018) . This interpretation, however, is not undisputed. It has been pointed out (Bedford, Blust et al. 2018) that, without understanding the population history and possibly similarly complex dynamics in Near Oceania, specifically the region where the Lapita cultural complex formed, it is difficult to discern the admixture events and ancestries present thought time in Remote Oceania.

Until recently, it was believed that people associated with the Lapita cultural complex primarily restricted their occupation to smaller offshore islands (Kirch 2001). Archaeological excavations have changed this view for mainland PNG with the discovery of Lapita pottery dated to 2900-2600 cal. BP at the Caution Bay site on the south coast (McNiven, David et al. 2011). From 2200 BP onwards, intensive settlements on the south coast are associated with shell impressed pottery and tools manufactured in an exchange sphere of ~700 km (Summerhayes and Allen 2007, David, McNiven et al. 2012), providing evidence for occupation by groups likely descended from the Lapita cultural complex (Lilley 2008, McNiven, David et al. 2011, McNiven, David et al. 2012). The distribution of Austronesian speaking people along the south coast (Ross 1988) supports this idea, as do the local oral traditions (Swadling 1977, Oram 1981). There is a dearth of archaeological sites on the south coast that date to between 1200 and 700 BP, a period known as the 'Papuan Hiccup' (Allen 2010). After 700 BP, settlement sites reappear across the region but it remains unclear whether this is the return of the previous occupants or sites established by entirely new populations (Bulmer 1971). According to the linguistics, Austronesian languages appear on the north coast of New Guinea and the Vitiaz Strait at the same time as on the south coast. Both regions also sport a small number of Lapita sites (Lilley 1988, Terrell and Schechter 2007, Gaffney, Summerhayes et al. 2015, Gaffney, Summerhayes et al. 2019, Summerhayes 2019), suggesting similar timings and processes on both coasts.

By investigating the genetic diversity of ancient inhabitants of the southern and north-eastern coast of Papua New Guinea and from the Bismarck Archipelago, this study contributes a genetic perspective to questions in Near Oceanic prehistory. It aims to understand the different ancestries involved in the formation of the genetic

make-up of people in the region today, and how they are linked to different dispersal events or migrations in the region.

Results:

Ancient genetic variation.

To add to the understanding of the genetic variation in ancient Near Oceania and the genetic traces of mobility in the region observed from archaeology we generated whole-genome sequences for 41 individuals from three sites on the New Guinea mainland and one in the Bismarck Archipelago (Table 1). To overcome the poor aDNA preservation in the region, we used a targeted enrichment approach to enrich for 1.2 million single nucleotide polymorphisms (SNPs) across the genome (Supplementary Materials). Contamination estimates were low with an average of 2% contamination (Supplementary Table S4). We determined the Y-chromosomal haplogroups for 21 of the 25 genetically male individuals and mitochondrial haplogroups for 21 of the 41 individuals in the dataset (Table 1, Supplementary Figure S1).

Newly produced radiocarbon dates of selected individuals (Supplementary Table S2) place the individuals analysed in an occupational phase for Nebira on the southern coast from around 500 cal. BP to the present day. Largely overlapping in time, individuals from the nearby site of Eriama date to around 400 cal. BP to the present day. An individual from the site of Tilu, situated on the north-eastern coast, dates to around 700 cal. BP. The five individuals excavated from Watom show large time intervals, covering a period from 3800 BP to 500-600 BP (Table 1, Fig.1B).

To investigate the genetic distances between present-day, published ancient and newly produced ancient genomes, we calculated a principal component analysis (Supplementary Materials, Supplementary Table S5). Individuals excavated from the Papuan coast do not cluster with present-day individuals from the New Guinean highlands, but with populations inhabiting the southern coastline today. Illustrating a cline, extending from the present-day individuals from the Papua New Guinean highlands to present-day East Asian and ancient Early Remote Oceanian individuals, the individuals excavated from Eriama on the southern coast cluster on the Papuan end of the cline, together with the two individuals from the site of Tilu,

situated on the northern coast. However, some individuals are removed from the group towards the East Asian end of the cline, suggesting a genetically heterogeneous population in Eriama. Spatially removed from the individuals from Eriama, individuals from the close-by site of Nebira cluster closely together around the midpoint of the same cline. This placement suggests a genetic composition of one half deriving from Papuan-related ancestry maximised in the highland populations of New Guinea (Supplementary Fig. 2) and the other half of East Asian related ancestry.

To identify a genetic grouping of the individuals, we computed f_4-statistics and confirmed the grouping with qpWave (Supplementary Material, Supplementary Table S7). Individuals from Nebira appear homogeneous in the genetic composition, whereas individuals from Eriama form two groups, consistent with their placement in the PCA along the Asian – Papuan cline. Two individuals (ERI006 and ERI004) are shifted towards the Asian end of the cline, while the other individuals cluster closer to the Papuan highlanders forming the second group. The three older individuals excavated from Watom were analysed separately (WAT002, WAT005, WAT006), accounting for the considerable time intervals between them (Fig.1B, Table 1). The exceptions are WAT001 and WAT003, which were grouped based on recent dates and genetic similarity.

Ancestry modelling.

After establishing that the individuals and groups were not consistent with deriving from one stream of ancestry (Supplementary Table S7), we modelled the Austronesian-related ancestry with present-day indigenous Taiwanese (Ami) and the Papuan-related ancestry with New Guinea Highlanders. QpAdm tests (Supplementary Table S7) show that most individuals can indeed be modelled as a mixture of these two ancestries (Fig. 3a). However, ancient individuals from the mainland of Papua New Guinea and the Bismarck Archipelago differ in their affinities to present-day Near Oceanic populations. While the individuals from Eriama, Nebira and Tilu show higher affinity to mainland New Guinean populations, all individuals from Watom show higher affinities to Baining from New Britain (Supplementary Table S6; Supplementary Figure 3a). Additionally, the ancient

individuals differ in their affinities to Asian populations. The Asian ancestry in Nebira and the Eriama individuals with higher Asian ancestry proportion show more affinity to the Early Remote Oceanians from Vanuatu and Tonga, associated with the Lapita cultural complex, than to present-day populations of Ami or Kankanaey, or to ancient individuals from Taiwan (Supplementary Table S6, Supplementary Figure 3b).

The patterns observed through the pc-analysis and f_4-statistics are supported by the qpAdm analysis. It shows higher proportions of East Asian ancestry in individuals from Nebira of 45-60% compared to Eriama and Tilu, where the proportion is around 20%. An exception are the two individuals from Eriama not included in the group, which show higher East Asian ancestry proportions of around 35%.

The individual WAT002, dated to around 2100 BP from Watom island, shows admixture between Papuan-related and East-Asian related populations, at 40% to 60%, respectively. This proportion is reduced to 20% when Early Remote Oceanian related individuals are used as a proxy for the Asian-related ancestry (Fig. 3c)

Timing of the admixture events.

To estimate the time of admixture between the two ancestry components, we performed an admixture dating analysis (Fig. 3d; Supplementary Material). For the ~2100-year-old individual from Watom Island, we inferred an admixture around ten generations ago, resulting in a date of ~2350 BP for the admixture event, matched by the two individuals from the same site, dating to ~500 BP.

We again observe differences in the two sites on the southern PNG coast. Inferred from the dated individuals only, individuals from Eriama have an average admixture date of 1030 BP, while the average date for individuals from Nebira shows an admixture event around 1500 BP. The individual NBR020 is the only exception with a younger admixture date of 650 BP. Coincidentally the same individual shows a non-local isotopic signal (Supplementary Fig. Xb). Because of low coverage, individuals from Tilu from the northern coast of Papua New Guinea were grouped for the analysis, resulting in an inferred admixture event 1030 years before present, similar to the average date for Eriama.

Sex-biased admixture.

While the vast majority of assigned mitochondrial haplogroups shows the "Polynesian motif" B4a1a1 associated with the Austronesian expansion, with only four individuals carrying Near Oceanian-related mitochondrial haplogroups. Y-Chromosome haplogroups show an opposite pattern. Apart from one Y-haplogroup with Asian origin (O2a2b2) all other 18 identified haplogroups are of Near Oceanic origin (Table 1, Supplementary Figure 1). To understand whether the pattern of sex-biased admixture observed in the uniparental markers is detectable at a genome-wide level, we compared the admixture proportions on the X chromosome to those inferred from the autosome. The analysis shows an excess of Austronesian ancestry on the X-Chromosome ranging from 10 to 60 percent, suggesting a sex-biased admixture as previously observed in populations from Vanuatu, where more males carrying Papuan-related ancestry admixed with more females carrying East Asian-related ancestry (Posth, Nägele et al. 2018).

Genetic relationships and burial patterns.

Finally, we investigated the genetic relationships between the individuals for Nebira, the only site where familial relationships could be addressed (Supplementary Fig. 5, Supplementary Table S9). The Nebira cemetery contained primary burials, some of which were used multiple or simultaneous burials. The resulting relatedness-network shows that the first and second-degree relations at the site include both male and female individuals, showing no clear signs of patri- or matrilocality. Genetically related individuals are not interred together but are buried close by, while unrelated individuals can be found in multiple or simultaneous burials. This result suggests the reason for multiple burials or reusing graves was not family-related.

Discussion:

Population History on Papua New Guinea:

All individuals from the coast of New Guinea harbour East Asian ancestry not observed in present-day individuals from the highlands of New Guinea, but present

in coastal New Guinea populations today (Fig. 2a; Supplementary Figure 2). While both the Eriama and Nebira sites on the southern coast are geographically very close together, their genetic composition suggests different population histories for the two sites. On average, the individuals from Eriama show higher Papuan ancestry, albeit with higher variation (Fig 3e) and younger admixture dates (Fig. 3d) compared to the individuals from Nebira. The latter are more homogenous, with higher proportions of East-Asian ancestry and on average older admixture dates (Fig3 e, a, d).

The admixture dates inferred from the dated individuals show the admixture event for people in Nebira occurred at around 1500 BP, much later than in other islands such as Vanuatu (~2600BP) and Watom (~2300 BP). This is possibly the result of a later arrival of the descendants of the Lapita cultural complex in the southern coast. Evidence for the first occurrence of Lapita settlements is dated to 2900 BP (McNiven, David et al. 2011), 400 years after initial Lapita colonisation of the Bismarck Archipelago (Kirch 2001, Summerhayes, Matisoo-Smith et al. 2010) and contemporaneous with evidence from Vanuatu (Petchey, Spriggs et al. 2014). From our analysis it is not clear whether this late admixture date is a result of a long isolation of the first settlers related to Lapita cultural complex, or the result of subsequent admixture of local populations, as this has been observed to lead to similar results (Posth, Nägele et al. 2018). Alternatively, the Lapita pottery found on the south coast could have been introduced by trade, without the extension of the material exchange to genetic exchange. Analysis of older individuals from the region could help specifying the first mixture of the two ancestries.

The individual NBR020 (ACJ-34) shows an admixture date of only 650 years, much younger than the majority of the group. Additionally, analysis of the strontium isotopes shows this individual to be non-local (Shaw, Buckley et al. 2011). Analyses of the archaeological remains in the south coast of PNG suggest abandonment or shift of sites ~1200 to 1000 BP, with a disruption in ceramic production (Frankel and Rhoads 1994, David 2008) and styles (Rhoads 1982, Bickler 1997). A breakdown in marine resources (Allen 1972, Vanderwal 1978, Swadling 1981, Vanderwal 2011, Shaw, Coxe et al. 2020) was possibly connected to climatic changes increasing aridity of the region (Allen 2010, Sutton, Summerhayes et al. 2015, Shaw, Coxe et

al. 2020). The increasingly unfavourable conditions possibly resulted in the “Papuan Hiccup”. Trade supposedly resumed in ~800 - 500 BP, and is associated with new Austronesian speaking groups arriving on the coasts. Similar pottery styles in the south coast and the Massim region in the east (Bulmer 1971) suggest this phase eventually lead to the emergence of the *Hiri* trade networks (Dutton 1982, Hope, Golson et al. 1983, Lilley 2004, Skelly and David 2017). The date for the resumed trade coincides with the admixture date inferred in NBR020 (ACJ-34).

At the same time, admixture dates inferred from the individuals from Eriama (Fig.3d), together with the higher Papuan ancestry (Fig.3a, e) suggest a later arrival of the Papuan-related ancestry. The strong shift within only 100 years from 40% to 15% Asian-related ancestry suggest they were part of an interaction sphere with people with higher Papuan-related ancestry. However, we lack the resolution to identify which populations contributed and where their geographical source was. The different genetic compositions of the two nearby sites show that the southern coast was a genetic, and possibly also a cultural and linguistic mosaic of people, matching the situation today. The ancient genetic diversity is reflected by the languages spoken in the region today: The Motu language is part of a western branch of Central Papuan Austronesian languages (Blust 1990, Ross 1994), and is spoken mostly by people located on the coasts. The Papuan language dominant in the region, Koita (Dutton 2010), is spoken in settlements more inland (Dutton 1969, Swadling 1977, Oram 1981, Swadling 1981).

The two individuals from the site of Tilu, situated on the north-eastern coast of New Guinea, show a genetic pattern and admixture dates similar to Eriama (Fig. 3a,d, Supplementary Table S8). First evidence for a settlement of the site dates to 650 BP (Gaffney, Summerhayes et al. 2018) by people speaking Bel languages, part of the Austronesian language family. The admixture event inferred from the individuals from Tilu predates the first occupation, suggesting an already admixed population established the site. Archaeological research places the site of Tilu in a local trading network extending to the Willaumez peninsula on the north-western coast of New Britain in the Bismarck Archipelago (Gaffney and Summerhayes 2019). Linguistic evidence and oral traditions suggest an origin in the Vitiaz Strait in the Bismarck Sea for Bel, the Austronesian language spoken at Tilu. However, the material

exchange seems not to have extended to genetic exchange, as there is no higher affinity of the individuals of Tilu to the populations of the Bismarck Archipelago, compared to Eriama (Supplementary Table S6, Supplementary Figure 3a). It is possible the oral traditions started recording after the establishment of the site, and the Bel language as adopted later to facilitate trade. As the individuals are dated to the early occupational phase, their descendants might show higher affinity to populations from the Bismarck Archipelago.

Time transect in Watom:

Previous studies have focused on the archipelago of Vanuatu to understand the population genetic events leading to the genetic make-up of present-day people in Remote Oceania, and ni-Vanuatu in particular (Lipson, Skoglund et al. 2018, Posth, Nägele et al. 2018). They revealed repeated admixture events of people with Papuan-related ancestry shortly after the initial settlement by people with almost exclusive East-Asian-related ancestry. A key point in the critiques prompted by this research was the lack of understanding of the population history of the Bismarck Archipelago (Bedford, Blust et al. 2018) as the supposed source of Papuan ancestry. Assuming a "growing Austronesian world", and taking into account frequent natural disasters in the Bismarck Archipelago, it seems plausible that interactions and displacements might have resulted in contact and genetic and cultural mixture before the migration to Vanuatu, as observed in the region (Torrence 2016) and proposed by some archaeologists (Spriggs 1997).

The Reber-Rakival site on Watom island has a well-established sequence, providing evidence for occupation by people of the Lapita cultural complex from 2800 – 2350 BP, the Middle and Late Lapita phases in the Bismarck Archipelago (Petchey and Green 2005, Petchey, Spriggs et al. 2011). The lowermost layer presents evidence for an earlier occupation (Petchey, Buckley et al. 2016). However, ceramics and obsidian are absent, possibly pointing to a different population occupying the island before the arrival of the Lapita cultural complex.

Our analysis shows that the two oldest skeletons analysed, male individuals excavated from the island of Watom, were genetically Papuan, most similar to people from New Britain today (Fig. 2a, Supplementary Figure 3a). The Strontium

352 analysis of WAT006 shows this individual to be local (Shaw, Buckley et al. 2010).
353 The anthropological analysis shows a rare case of cranial deformation in form of
354 trepanation (Pietrusewsky, Buckley et al. 2014) - a practice not known from the
355 Lapita cultural complex, but observed on the Gazelle Peninsula in Southwest New
356 Britain (Blackwood and Danby 1955, Parkinson 2010). The second burial with
357 Papuan-related ancestry and the oldest individual in this time transect, WAT005,
358 does not show cranial deformations. Both individuals show an occupation of
359 Watom by people with Papuan-related ancestry, both before and after the
360 archaeologically attested arrival of the Lapita-cultural complex in the Bismarck
361 Archipelago. Moving forward in time, by 2100 BP, we find an individual that fits the
362 model of mixture between Papuan-related and East-Asian related ancestry, similar
363 in the proportions to the individuals from the Late-Lapita period in Vanuatu. The
364 individual shows higher affinity to present-day Baining from New Britain, again
365 shared with the contemporary individuals from Vanuatu (Supplementary Figure 3a).
366 Dating the admixture between the two ancestries, the resulting date of ~2300 BP for
367 all individuals with Asian-related ancestry suggests admixture with local Papuan
368 people 1000 years after the arrival of the Lapita cultural complex in the Bismarck
369 Archipelago (Figure 3d). We acknowledge the limitations of interpretations deriving
370 from only five individuals covering a wide time transect. However, we would
371 cautiously interpret the results as further support the “express train” model
372 (Diamond 1988) for the expansion into Remote Oceania. A local admixture upon or
373 shortly after arrival in the Archipelago would have resulted in a much earlier
374 admixture date of ~3000 BP. The higher affinity to the Early Remote Oceanian
375 individuals from Vanuatu and Tonga rather than to ancient Taiwanese does not
376 provide support for a subsequent population expansion from East Asia into the
377 region. Rather it is indicative of a period where both groups lived side by side
378 without genetic exchange. This is consistent with the observation that the oldest
379 individuals with Papuan related ancestry in Vanuatu (2600 and 2300BP) are
380 unadmixed. If this is the case, it presents challenging implications for the population
381 size and mobility regarding the maintenance of a maritime colonizing society,
382 without genetic recruitment from established groups.

Lastly, the two most recent individuals show a genetic make-up very similar to people in the region today. However, they are not consistent with deriving from the same population as the Tolai, inhabiting Watom Island today (Supplementary Table S7). Highly affected by natural disasters and colonialism (Spriggs 1997), displacements in the last ~500 years might pose an explanation for this observation.

Conclusion:

The first analysis of ancient Papuan genomes suggests that people on the coasts of Papua New Guinea were involved in diverse and complex interaction spheres with groups of different ancestries, which left genetic marks. The genetic variation observed in the region today had already started forming by 500 BP, where individuals dated to this period from the north-eastern coast, opposite the Bismarck Archipelago, show an East-Asian ancestry component, similar to that observed in coastal populations in Papua New Guinea today. The comparison of two only 10 km apart, and roughly contemporaneous sites on the southern coast shows that repeated admixture events resulted in different ancestry trajectories, possibly extended to culture and language, similar to the diversity observed today. The genetic affinities of the inhabitants of the southern coast of Papua New Guinea show that the occurrence of pottery associated with the Lapita cultural complex can be tied to the arrival of the related ancestry.

The genetic affinities on Watom Island in the Bismarck Archipelago are similar in respect to the Asian-related ancestry but differ in the Papuan-related ancestry. While this ancestry component is most similar to present-day Papuan Highlanders for the individuals from the coasts of Papua New Guinea, the individuals from Watom show higher affinity to present-day populations from the Bismarck Archipelago.

There, we inferred from one admixed individual an admixture event post-dating the first occurrence of the Lapita cultural complex in the Bismarck Archipelago and the admixture events inferred from individuals in Vanuatu (~2500B). This result suggests the admixture on Watom occurred not only after the initial settlement of the Bismarck Archipelago ~3300 BP but also after the continued journey into Remote Oceania ~3000 BP. Although the limited number of individuals advise for a cautious

interpretation, the long interval between the first evidence of occupation by the Lapita cultural complex ~3300 BP on the Bismarck Archipelago and the inferred date of admixture for the first individual harbouring the related ancestry, implies the first, minimally admixed settlers remained isolated from local populations with Papuan-related ancestry. It seems more likely that first an unadmixed East Asian-related and later Papuan-related populations arrived and mixed in Remote Oceania, supporting the statement in previous archaeogenetic research (Lipson, Skoglund et al. 2018, Posth, Nägele et al. 2018, Posth, Nägele et al. 2019).

Author contributions:

RK, HB, DG, GS, MT contributed archaeological material and contextualised. KN, EB and RR performed ancient DNA laboratory work. KN and SC performed population genetic analysis. , KN interpreted the data with critical input from RK, CP, JK and archaeological contextualisation by DG, GS, MT, RK and HB. KN wrote the manuscript with critical input from CP, RK, JK, AP and the remaining authors. KN produced the figures. AP, JK and CP conceived and coordinated the study.

Acknowledgements:

We would like to thank the National Museum and Art Gallery of Papua New Guinea for the permission to analyse the individuals excavated from Nebira, Eriama and Tilu. We would like to acknowledge the pioneering work of the late Dr. Susan Bulmer. She led extensive surveys and excavations on the South Coast and Highlands of Papua New Guinea. We thank the Tilu community at Malmal village for their assistance in the research and Franziska Aron and Luka Papac for support in the wetlab.

This study was funded through the ERC starting grant WAVES granted by the European Research Council (ERC), the Otago Research Grant and the Max Planck Society.

References:

Allen, J. (1972). "Nebira 4: an early Austronesian site in central Papua." Archaeology & Physical Anthropology in Oceania **7**(2): 92-124.

Allen, J. (2010). "Revisiting Papuan ceramic sequence changes: another look at old data." Artefact: the Journal of the Archaeological and Anthropological Society of Victoria, The(33): 4.

Allen, J., C. Gosden, R. Jones and J. P. White (1988). "Pleistocene dates for the human occupation of New Ireland, northern Melanesia." Nature **331**(6158): 707-709.

Allen, J., S. Holdaway and R. Fullagar (1997). "Identifying specialisation, production and exchange in the archaeological record: The case of shell bead manufacture on Motupore Island, Papua." Archaeology in Oceania **32**(1): 13-38.

Bedford, S., R. Blust, D. BURLEY, M. Cox, P. KIRCH, E. MATISOO‐SMITH, Å. Næss, A. Pawley, C. Sand and P. Sheppard (2018). "Ancient DNA and its contribution to understanding the human history of the Pacific Islands." Archaeology in Oceania **53**(3).

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

Bellwood, P., J. J. Fox and D. Tyron (2006). The Austronesians: Historical and comparative perspectives, ANU Press.

Bickler, S. H. (1997). "Early pottery exchange along the south coast of Papua New Guinea." Archaeology in Oceania **32**(2): 151-162.

Blackwood, B. and P. Danby (1955). "A study of artificial cranial deformation in New Britain." The Journal of the Royal Anthropological Institute of Great Britain and Ireland **85**(1/2): 173-191.

Blust, R. (1990). "Three recurrent changes in Oceanic languages." Pacific island languages: Essays in honour of GB Milner: 7-28.

Bulmer, S. (1971). "Prehistoric settlement patterns and pottery in the Port Moresby area." Journal of the Papua and New Guinea Society **5**(2): 29-91.

David, B. (2008). "Rethinking cultural chronologies and past landscape engagement in the Kopi region, Gulf Province, Papua New Guinea." The Holocene **18**(3): 463-479.

David, B., I. J. McNiven, M. Leavesley, B. Barker, H. Mandui, T. Richards and R. Skelly (2012). "A new ceramic assemblage from Caution bay, south coast of mainland PNG: the linear shell edge-impressed tradition from Bogi 1." Journal of Pacific Archaeology **3**(1): 73-89.

David, B., I. J. McNiven, T. Richards, S. P. Connaughton, M. Leavesley, B. Barker and C. Rowe (2011). "Lapita sites in the central province of mainland Papua New Guinea." World Archaeology **43**(4): 576-593.

Diamond, J. M. (1988). "Express train to Polynesia." Nature **336**(6197): 307-308.

Dutton, T. (1978). "Language and trade in central and south-east Papua." The Australian Journal of Anthropology **11**(3): 341.

Dutton, T. E. (1969). The peopling of Central Papua: Some preliminary observations, Pacific Linguistics, Research School of Pacific and Asian Studies, The ….

Dutton, T. E. (1982). The Hiri in History: Further aspects of long distance Motu trade in Central Papua, Better English Language Teaching.

Dutton, T. E. (2010). Reconstructing Proto Koiarian: The history of a Papuan language family, Pacific Linguistics, Research School of Pacific and Asian Studies, The ….

Frankel, D. and J. W. Rhoads (1994). Archaeology of a coastal exchange system: sites and ceramics of the Papuan Gulf, Division of Archaeology and Natural History, Research School of Pacific and ….

Friedlaender, J. S., F. R. Friedlaender, F. A. Reed, K. K. Kidd, J. R. Kidd, G. K. Chambers, R. A. Lea, J.-H. Loo, G. Koki and J. A. Hodgson (2008). "The genetic structure of Pacific Islanders." PLoS Genet **4**(1): e19.

Gaffney, D. and G. R. Summerhayes (2019). "Coastal mobility and lithic supply lines in northeast New Guinea." Archaeological and Anthropological Sciences **11**(6): 2849-2878.

Gaffney, D., G. R. Summerhayes, A. Ford, J. M. Scott, T. Denham, J. Field and W. R. Dickinson (2015). "Earliest pottery on New Guinea mainland reveals Austronesian influences in highland environments 3000 years ago." PloS one **10**(9): e0134497.

Gaffney, D., G. R. Summerhayes and M. Mennis (2019). "A Lapita presence on Arop/Long Island, Vitiaz Strait, Papua New Guinea?" Debating Lapita: Distribution, Chronology, Society and Subsistence **52**: 115.

Gaffney, D., G. R. Summerhayes, M. Mennis, T. Beni, A. Cook, J. Field, G. Jacobsen, F. Allen, H. Buckley and H. Mandui (2018). "Archaeological investigations into the origins of Bel trading groups around the Madang coast, northeast New Guinea." The Journal of Island and Coastal Archaeology **13**(4): 501-530.

Gray, R. D., A. J. Drummond and S. J. Greenhill (2009). "Language phylogenies reveal expansion pulses and pauses in Pacific settlement." science **323**(5913): 479-483.

Green, R. C. (1991). "[DIVERSITY, CONTINUITY AND CHANGE IN THE HISTORY OF THE BISMARCK ARCHIPELAGO] The Lapita Cultural Complex: current evidence and proposed models." Bulletin of the Indo-Pacific Prehistory Association **11**: 295-305.

Groube, L., J. Chappell, J. Muke and D. Price (1986). "A 40,000 year-old human occupation site at Huon Peninsula, Papua New Guinea." Nature **324**(6096): 453-455.

Hammarström, H., R. Forkel and M. Haspelmath (2019). Glottolog 4.0. Jena.

Hope, G., J. Golson and J. Allen (1983). "Palaeoecology and prehistory in New Guinea." Journal of Human Evolution **12**(1): 37-60.

Kayser, M., S. Brauer, G. Weiss, P. Underhill, L. Roewer, W. Schiefenhövel and M. Stoneking (2000). "Melanesian origin of Polynesian Y chromosomes." Current Biology **10**(20): 1237-1246.

Kirch, P. V. (2001). Lapita and Its Transformations in Near Oceania: Archaeological Investigations in the Mussau Islands, Papua New Guinea, 1985-88, Archaeological Research Facility, University of California.

Lilley, I. (1988). "Prehistoric exchange across the Vitiaz Strait, Papua New Guinea." Current Anthropology **29**(3): 513-516.

Lilley, I. (2004). "Trade and culture history across the Vitiaz Strait, Papua New Guinea: the emerging post-Lapita coastal sequence." RECORDS-AUSTRALIAN MUSEUM: 89-96.

Lilley, I. (2008). "Flights of fancy: Fractal geometry, the Lapita dispersal and punctuated colonisation in the Pacific." Islands of inquiry: Colonisation, seafaring and the archaeology of maritime landscapes **29**: 75.

Lipson, M., P. Skoglund, M. Spriggs, F. Valentin, S. Bedford, R. Shing, H. Buckley, I. Phillip, G. K. Ward and S. Mallick (2018). "Population turnover in Remote Oceania shortly after initial settlement." Current Biology **28**(7): 1157-1165. e1157.

McNiven, I. J., B. David, K. Aplin, J. Mialanes, B. Asmussen, S. Ulm, P. Faulkner, C. Rowe and T. Richards (2012). "Terrestrial engagements by terminal Lapita maritime specialists on the southern Papuan coast." Peopled landscapes: Archaeological and biogeographic approaches to landscapes: 121-156.

McNiven, I. J., B. David, T. Richards, K. Aplin, B. Asmussen, J. Mialanes, M. Leavesley, P. Faulkner and S. Ulm (2011). "New direction in human colonisation of the Pacific: Lapita settlement of south coast New Guinea." Australian Archaeology **72**(1): 1-6.

O'Connell, J. F. and J. Allen (2004). "Dating the colonization of Sahul (Pleistocene Australia–New Guinea): a review of recent research." Journal of Archaeological Science **31**(6): 835-853.

Oram, N. (1981). "The history of the Motu-speaking and Koitabu-speaking peoples according to their own traditions." Oral tradition in Melanesia: 207-229.

Parkinson, R. (2010). Thirty years in the South Seas: land and people, customs and traditions in the Bismarck Archipelago and on the German Solomon Islands, Sydney University Press.

Petchey, F. and R. Green (2005). "Use of three isotopes to calibrate human bone radiocarbon determinations from Kainapirina (SAC), Watom Island, Papua New Guinea." Radiocarbon **47**(2): 181-192.

Petchey, F., M. Spriggs, S. Bedford, F. Valentin and H. Buckley (2014). "Radiocarbon dating of burials from the Teouma Lapita cemetery, Efate, Vanuatu." Journal of Archaeological Science **50**: 227-242.

Petchey, F., M. Spriggs, F. Leach, M. Seed, C. Sand, M. Pietrusewsky and K. Anderson (2011). "Testing the human factor: radiocarbon dating the first peoples of the South Pacific." Journal of Archaeological Science **38**(1): 29-44.

Petchey, P., H. Buckley, R. Walter, D. Anson and R. Kinaston (2016). "The 2008-2009 excavations at the SAC locality, Reber-Rakival Lapita site, Watom Island, Papua New Guinea." Journal of Indo-Pacific Archaeology **40**: 12-31.

Pietrusewsky, M., H. Buckley, D. Anson and M. T. Douglas (2014). "Polynesian origins: A biodistance study of mandibles from the Late Lapita site of Reber-Rakival (SAC), Watom Island, Bismarck Archipelago." Journal of Pacific Archaeology **5**(1): 1-20.

Posth, C., K. Nägele, H. Colleran, F. Valentin, S. Bedford, R. D. Gray, J. Krause and A. Powell (2019). "Response to “Ancient DNA and its contribution to understanding

575 the human history of the Pacific Islands"(Bedford et al. 2018)." Archaeology in
576 Oceania **54**: 57-61.

577 Posth, C., K. Nägele, H. Colleran, F. Valentin, S. Bedford, K. W. Kami, R. Shing, H.
578 Buckley, R. Kinaston and M. Walworth (2018). "Language continuity despite
579 population replacement in Remote Oceania." Nature ecology & evolution **2**(4): 731-
580 740.

581 Rhoads, J. W. (1982). "Prehistoric Papuan exchange systems: the hiri and its
582 antecedents." The Hiri in History: Further aspects of long distance Motu trade in
583 Central Papua: 131-151.

584 Ross, M. (1988). Proto Oceanic and the Austronesian languages of western
585 Melanesia, Dept. of Linguistics, Research School of Pacific Studies, The
586 Australian ….

587 Ross, M. D. (1994). "Central Papuan culture history: some lexical evidence."
588 Austronesian terminologies: Continuity and change: 389-479.

589 Shaw, B., H. Buckley, G. Summerhayes, D. Anson, S. Garling, F. Valentin, H.
590 Mandui, C. Stirling and M. Reid (2010). "Migration and mobility at the Late Lapita
591 site of Reber–Rakival (SAC), Watom Island using isotope and trace element
592 analysis: a new insight into Lapita interaction in the Bismarck Archipelago." Journal
593 of Archaeological Science **37**(3): 605-613.

594 Shaw, B., H. Buckley, G. Summerhayes, C. Stirling and M. Reid (2011). "Prehistoric
595 migration at Nebira, South Coast of Papua New Guinea: New insights into
596 interaction using isotope and trace element concentration analyses." Journal of
597 Anthropological Archaeology **30**(3): 344-358.

598 Shaw, B., S. Coxe, V. Kewibu, J. Haro, E. Hull and S. Hawkins (2020). "2500-year
599 cultural sequence in the Massim region of eastern Papua New Guinea reflects
600 adaptive strategies to small islands and changing climate regimes since Lapita
601 settlement." The Holocene: 0959683620908641.

602 Skelly, R. J. and B. David (2017). Hiri: Archaeology of long-distance maritime trade
603 along the south coast of Papua New Guinea, University of Hawai 'i Press.

604 Skoglund, P., C. Posth, K. Sirak, M. Spriggs, F. Valentin, S. Bedford, G. R. Clark, C.
605 Reepmeyer, F. Petchey and D. Fernandes (2016). "Genomic insights into the
606 peopling of the Southwest Pacific." Nature **538**(7626): 510-513.

607 Spriggs, M. (1997). The Island Melanesians, Blackwell.

608 Stoneking, M. and F. Delfin (2010). "The human genetic history of East Asia:
609 weaving a complex tapestry." Current Biology **20**(4): R188-R193.

610 Summerhayes, G., E. Matisoo-Smith, H. Mandui, J. Allen, J. Specht, N. Hogg and S.
611 McPherson (2010). "An early lapita site on Emirau, New Ireland, PNG." Journal of
612 Pacific Archaeology **1**(1): 62-75.

613 Summerhayes, G. R. (2019). "Austronesian expansions and the role of mainland
614 New Guinea: a new perspective." Asian Perspectives **58**(2): 250-260.

615 Summerhayes, G. R. and J. Allen (2007). "Lapita writ small? Revisiting the
616 Austronesian colonisation of the Papuan south coast." Oceanic Explorations.

617 Summerhayes, G. R., M. Leavesley, A. Fairbairn, H. Mandui, J. Field, A. Ford and R.
618 Fullagar (2010). "Human adaptation and plant use in highland New Guinea 49,000 to
619 44,000 years ago." science **330**(6000): 78-81.

620 Sutton, N., G. Summerhayes and A. Ford (2015). Regional interaction networks in
621 southern Papua New Guinea during the late Holocene: Evidence from the chemical
622 characterisation of chert artefacts. Proceedings of the Prehistoric Society,
623 Cambridge University Press.

624 Swadling, P. (1977). "A review of the traditional and archaeological evidence for
625 early Motu, Koita and Koiari settlement along the central south Papuan coast." Oral
626 History **5**(2): 37-43.

627 Swadling, P. (1981). "The settlement history of the Motu and Koita speaking people
628 of the Central Province, Papua New Guinea." Oral Tradition in Melanesia: 240-251.

629 Terrell, J. E. and E. M. Schechter (2007). "Deciphering the Lapita code: the Aitape
630 ceramic sequence and late survival of the'Lapita face'." Cambridge Archaeological
631 Journal **17**(1): 59.

632 Torrence, R. (2016). "Social resilience and long-term adaptation to volcanic
633 disasters: The archaeology of continuity and innovation in the Willaumez Peninsula,
634 Papua New Guinea." Quaternary International **394**: 6-16.

635 Vanderwal, R. (1978). "Exchange in prehistoric coastal Papua." The Australian
636 Journal of Anthropology **11**(3): 416.

637 Vanderwal, R. L. (2011). "Prehistoric studies in central coastal Papua."

638 Wollstein, A., O. Lao, C. Becker, S. Brauer, R. J. Trent, P. Nürnberg, M. Stoneking
639 and M. Kayser (2010). "Demographic history of Oceania inferred from genome-wide
640 data." Current Biology **20**(22): 1983-1992.

Figures and Tables:

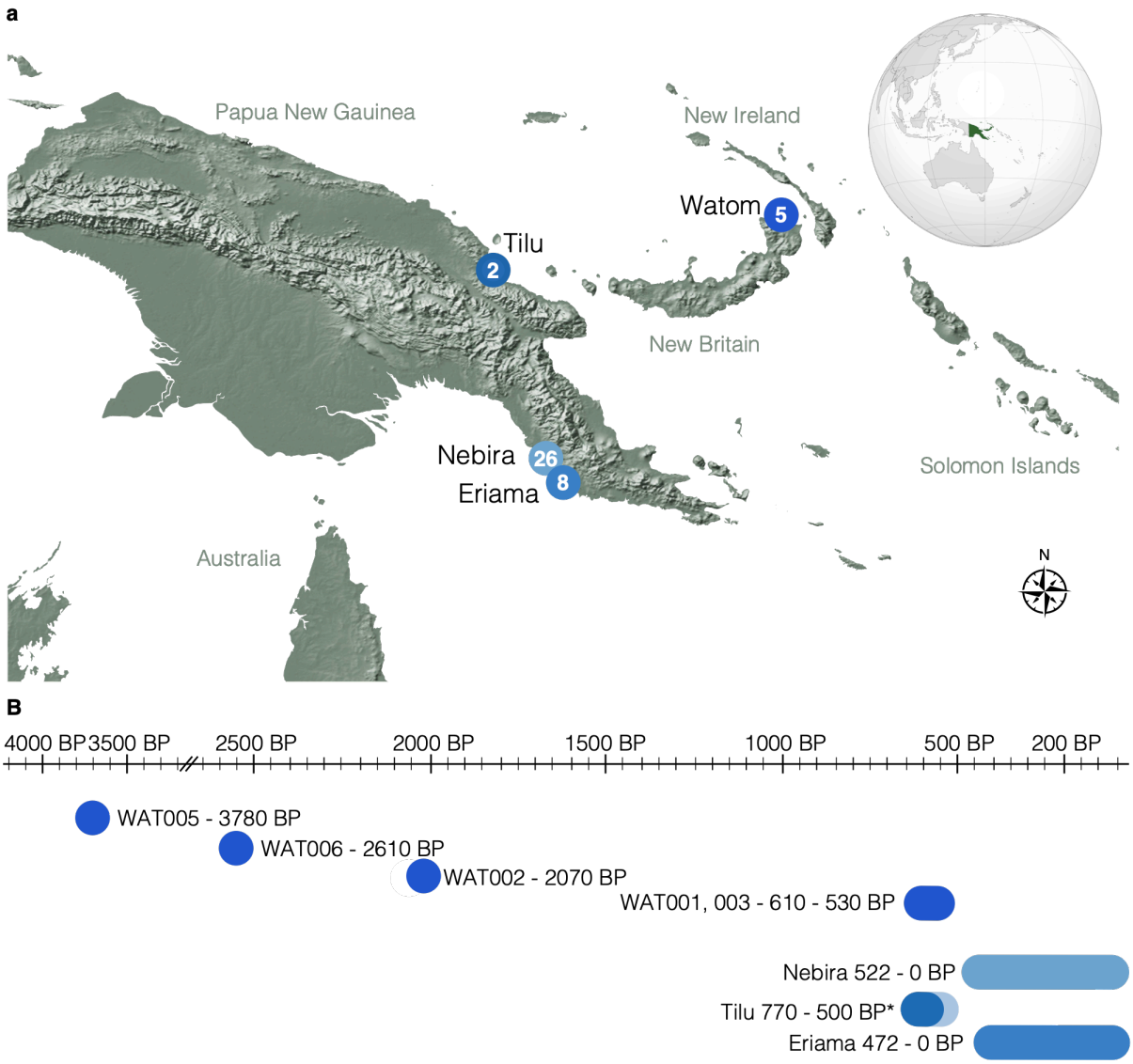


Fig.1: Sites and samples. *Map of Near Oceania showing (**a**) the location of the sites discussed in this study and the number of individuals analysed per site. (**b**) Date ranges for each site/individual in calibrated years BP. Date ranges are based on directly dated skeletal remains and do not necessarily represent the entire occupation of the site. For single individuals, mean point dates are provided. * for Tilu the solid date range is based on direct dating of one individual, transparent range indicates archaeological evidence for occupation of the site.*

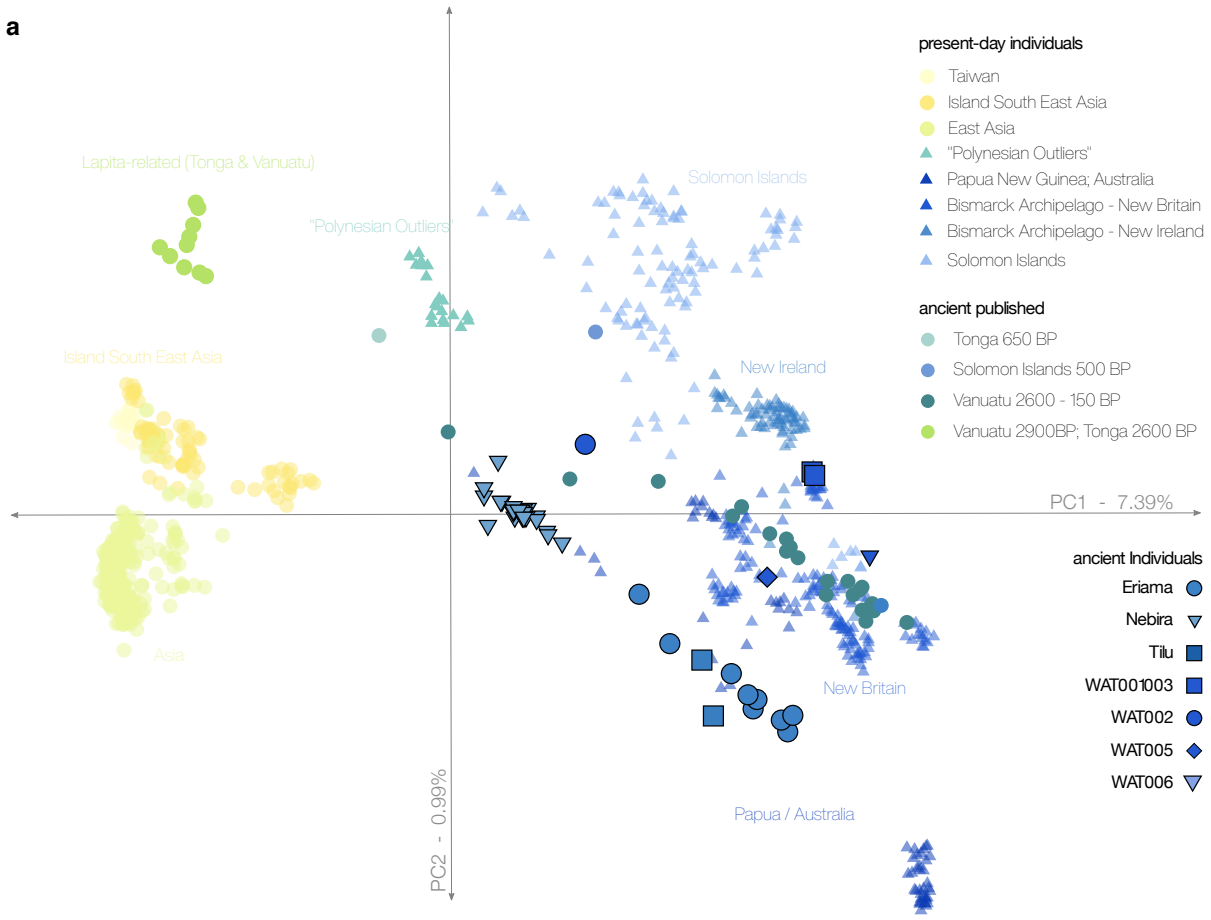


Fig.2: Population substructure in Oceania. *Principal component analysis (PCA) of present-day individuals from Asia, Island South-East Asia, Near and Remote Oceania with ancient samples projected. Outlined individuals are newly reported in this study (**A**).*

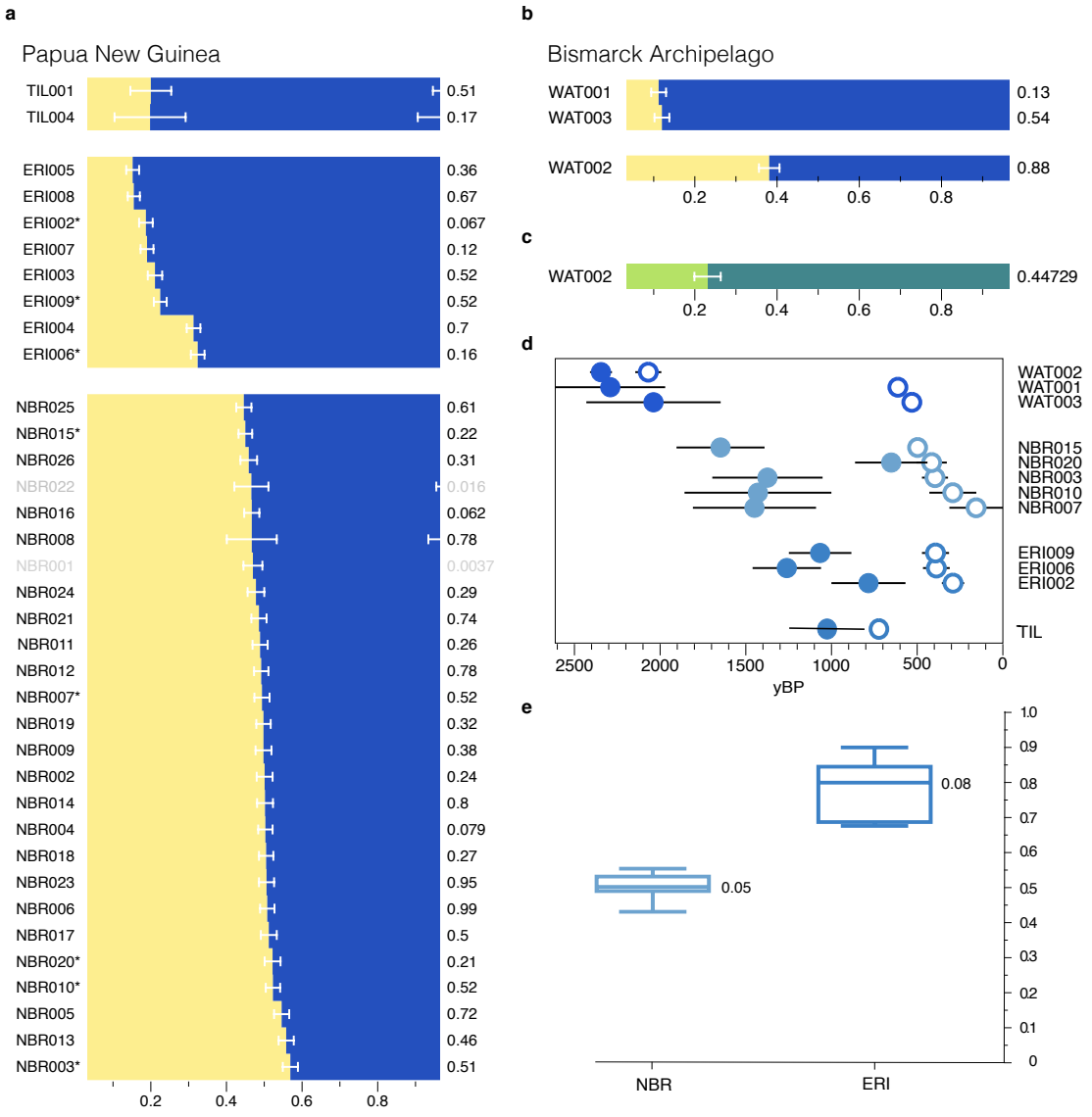


Fig.3: Ancestry modelling and Admixture dating. *Ancestry modelling using Ami (yellow) and New Guinea (blue). (**a**) Representing the ancestral components for all individuals from Papua New Guinea and (**b**)* for Watom Island*. (c) Ancestry modelling for WAT002 using Lapita-related individuals from Vanuatu and Tonga (green) and an individual from Vanuatu with Papuan ancestry closely related to present-day Baining Marabu from the Bismarck Archipelago (TAN002, teal green). White lines indicate the standard error for each component. (d) Inferred dates of admixture (filled circles) and the midpoint of the calibrated C14 date range for each individual (empty circles). Black lines on empty circles indicate the full C14 range and black lines on filled circles indicate the standard errors of the admixture date. (**e**) Variation of the Papuan ancestry proportion in Nebira and Eriama. Numbers show the mean ancestry proportion.*

8. DISCUSSION

8.1 DNA retrieval from tropical environments

As outlined in Part 5.1, DNA decay is expedited by the presence of water and in higher temperatures (*31*), both properties of the tropical environments of the regions focused on the presented studies. Until recently, the prospect of retrieving genome-wide data from tropical regions seemed unrealistic, and the studies summarised in this thesis are the result of improved technologies in the retrieval and reconstruction of ancient genomes.
For this thesis, a total of 351 samples were screened for ancient DNA. While we targeted the petrous portion of the temporal bone, shown to yield most DNA (*63, 64*), they were only available from 108 skeletons, and teeth were therefore sampled for the majority of the skeletons with a total of 239 tooth samples. Additionally, four samples from other bones were included in the screening process but yielded no ancient DNA (Fig.1A). Out of the total samples screened, surprisingly more than half were suitable for whole-genome analysis. The vast majority comes from the petrous part of the temporal bone, which in both regions has a success rate close to 100% (Fig. 1B). Although only ~25-40% of the screened teeth were used in the final analyses, the recovery rate is still surprisingly high. While retrieving DNA from ancient human remains in tropical regions is a success itself, the quality of the analysis depends highly on the quality of the data. Higher coverage of genomes is a necessity for confident statistics and interpretations. The percentage of endogenous DNA represents the proportion of DNA that maps to the human reference genome and shows a typical pattern of deaminated sites towards the end of the molecules (*39*) as a measure of authenticity. This value not only determines whether or not the respective library will be processed further, but it also is a predictor for the number of analysable sites resulting from further processing. The percentage of endogenous DNA in teeth is much lower compared to the percentage contained in petrous portions (Fig 1C). In the targeted enrichment approach used in the presented studies, the library is enriched for molecules containing a set of 1.2 million single nucleotide polymorphisms (SNPs), designed to differentiate between worldwide populations. Genomes produced from teeth result on average in coverage of 200,000

SNPs, while for petrous portion the average is much higher, around 600,000 SNPs. In both elements, the number of markers ranges, with some very poor and some very well covered genomes (Fig. 1D). The successful enrichment, especially in the poorly preserved samples highlights the importance and efficiency of the targeted enrichment approach in tropical regions.

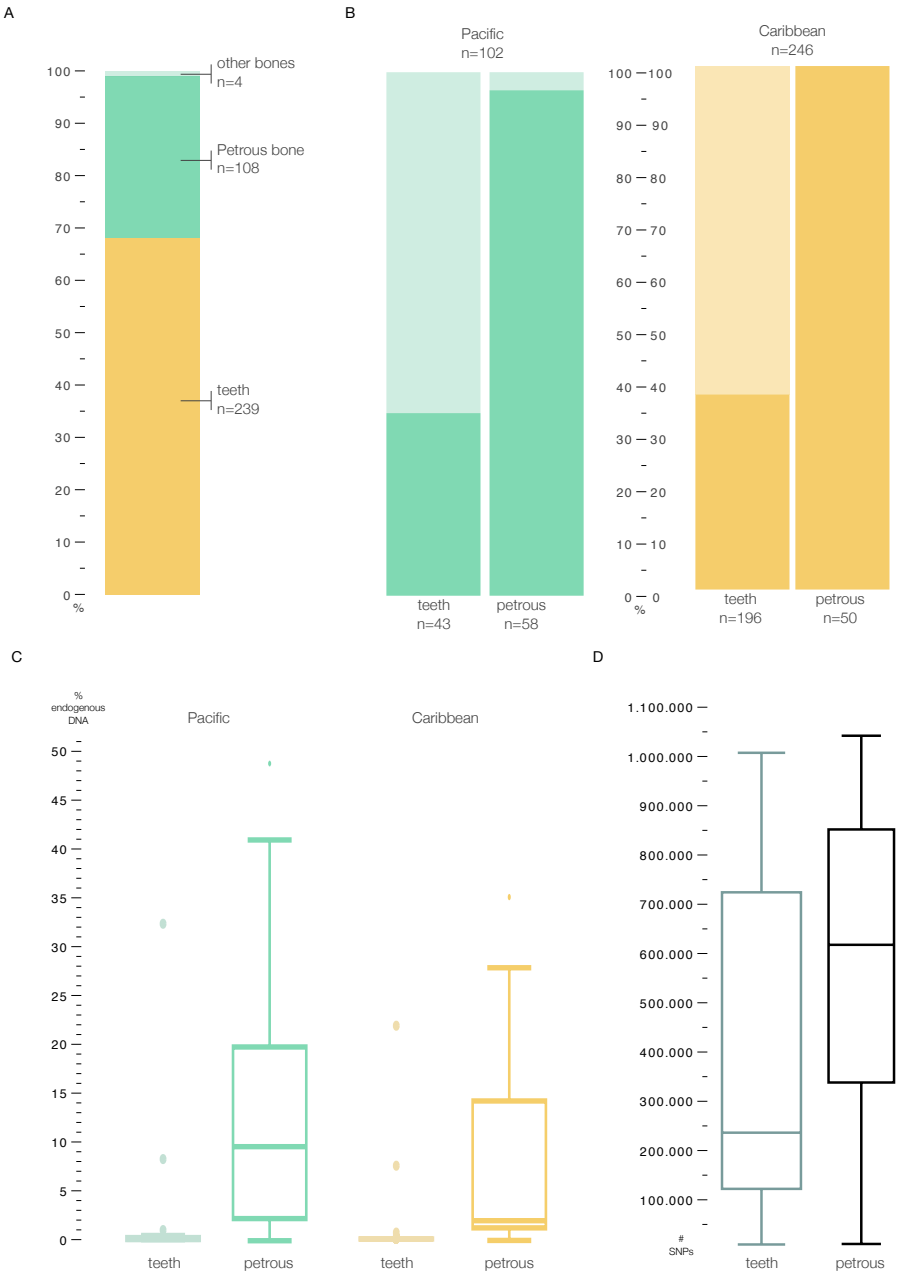


Figure 1: Human DNA preservation in different skeletal elements from tropical regions. Proportions of elements within the dataset of teeth, petrous bone and other bones (A). The proportion of successful and unsuccessful screening efforts as measured by successful whole genome analysis for teeth and petrous, divided by region (B). Per cent of ancient, human DNA (%endogenous DNA) preserved in the samples, dividend by element and region (C). Individuals that showed less than 5% damage were excluded from this analysis. Number of SNPs genotyped on the 1240K SNP panel for genomes recovered from teeth and petrous portions of the temporal bone (D); analysis contains only individuals analysed in the presented studies.

The comparison of the two regions shows that the similar climates lead to similar preservations. The higher average percentages of endogenous DNA in the samples from the Pacific are most likely the result of on average younger radiocarbon dates, and show the effect of age on the preservation, as discussed before (*38*). Comparing roughly contemporaneous individuals from the same climate, the high variation in the preservation demonstrates the considerable effect of the microclimate within the site or the environment of the burial ground. Another property of highly degraded DNA is short fragment lengths. In the established protocols, shorter fragments are lost, affecting the recovery of DNA from samples with high degradation. Most recent advances of the techniques to prepare DNA libraries for sequencing have been applied in this thesis in few occasions, but have produced most promising results. The single-stranded library protocol (*56, 57*) can retrieve shorter fragments and has been used on a few samples in this study that produced insufficient coverage. In those cases, the production of an additional single-stranded library has led to a substantial increase in analysable sites.
In summary, the approach taken in this thesis was successful in the retrieval of DNA from ancient human remains from tropical regions. Combining a careful, targeted sampling of the petrous portion of the temporal bone, where possible, paired with a targeted enrichment approach, has resulted in considerable samples sizes and comprehensive datasets on which future studies can build on. For future projects, the broad application of the single-stranded library protocol promises a higher success rate for poorly preserved samples and elements with lower DNA content such as teeth or other bones.

8.2 Insights into the settlement history of the Caribbean and implications for the American continents.

Manuscript A shows that the Caribbean was settled and resettled multiple times, already during the Archaic Age of the Americas.
The most recent dispersal, heralding the Ceramic Age of the Caribbean around 2800 BP, is very well researched. The origins in the Orinoco River Delta in northeastern

South America are well supported through various disciplines, including archaeogenomics. However, the mode in which this dispersal expanded into the Caribbean was still debated. The dispersal over the Lesser Antilles, known as the Stepping Stone Model, competes with the Southward Expansion Hypothesis, proposing a direct journey from South America to Puerto Rico, before expanding southward into the Lesser Antilles. In this thesis, we analysed data of individuals associated with this context and tested the two models. The data at hand favours a stepping stone mode for the expansion of the Ceramic Age settlers. As the time interval between the first archaeological dates for the occurrence of the newcomers in the Islands is 1,300 years earlier than the oldest genome analysed, the results leave room for future analysis. It is crucial to close the temporal gap in the sampling to model with higher confidence and to exclude the possibility of multiple migrations during the Ceramic Age.

Additionally, analyses within the group of Ceramic Age settlers reveal different interaction spheres. Most profoundly this is the case for Puerto Rico. Situated in the centre of the Caribbean (Fig. 2A), archaeological analysis have long suggested it served as a crossroads in the Caribbean, connecting the western islands, the Bahamas and the Lesser Antilles, possibly even the South American mainland (*207, 208*).

Analysis of the differential affinities (Fig. 2B) shows that the different sites on Puerto Rico differ in their affinities to other Ceramic Age groups in the Caribbean. Paso del Indio, situated in the north, shows higher affinity to the Bahamas, similar to groups in Cuba. Individuals from Punta Candelero in the east of the island have higher affinities to the groups from the Lesser Antilles. The variation in Puerto Rico suggests different interaction spheres and calls for a more detailed investigation of the variation within islands, and the connections to other islands, possibly revealing trade routes and interactions during the Ceramic Age.

Little can be added to the discussion of the “Island Carib Problem” (*172*). The site of Lavoutte has been interpreted as a Carib Ceremonial Centre in the past (*168*). However, genetic affinities of the individuals from the site show a clear connection to the other ceramic-related individuals across the Caribbean and to Arawak-speaking present-day populations on the northeastern South American mainland, rather than Carib speakers. In this study no skeletons from a Cayo-context associated with “Island

Carib" were analysed. Expanding the sampling to individuals associated with such context, dating close to the European invasion, and to islands further south might allow to test specific questions related to the Island Caribs. According to record of the Carib myths of origin by a French missionary, Kalina from the South American mainland entered the islands, killed the men and took the woman as wives (*209*). Such a scenario can, in theory, be tested with genetic data. Available datasets include Carib speakers, enabling us to test higher affinities to those rather than to Arawak-speakers or the ancient ceramic-related individuals. Furthermore, sex-biased admixtures can be detected through already established methods, allowing a corroboration of the French coloniser's reports.

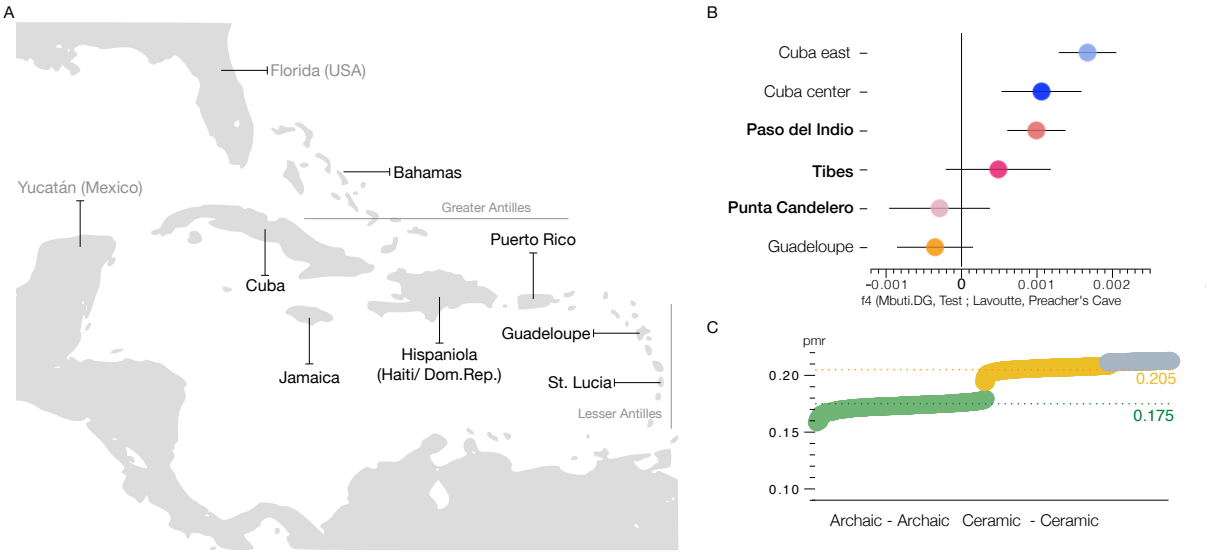


Figure 2: Diversity of Caribbean populations. Map of the Caribbean showing places of interest (**A**). Differential affinities among Ceramic Age groups (**B**). Bold are sites in Puerto Rico. Positive values suggest a higher affinity to the Bahamian genomes; negative values indicate higher affinity to the genomes from St.Lucia. Black lines indicate one standard error. Pairwise mismatch rates calculated from each position covered in pairs of individuals (**C**). Dotted lines indicate the mean value for Archaic Age groups (green) and Ceramic Age groups (yellow).

Surprising is the lack of interaction of Ceramic Age newcomers with Archaic Age populations. The expansion hiatus between the settlement of Puerto Rico and the expansion into Hispaniola has been attributed to the presence of Archaic Age populations on Hispaniola. At the same time, the high abundance of Archaic Age sites, including plain ceramics in Hispaniola and the east of Cuba, have generated ideas of interaction and cultural exchange. In this study, we were not able to detect significant

amounts of admixture in either of the populations. However, data from the island of Hispaniola is missing entirely. Investigating the genetic diversity on Hispaniola should provide insights into nature and extent of possible interactions between the Ceramic Age newcomers and the descendants of Archaic Age settlers.

One objective of the study was to identify the origins of Archaic Age settlers on the American mainland. Morphological studies have proposed several regions of origin for the early settlers of the Western Caribbean; however, the results show that these analyses are unreliable, possibly the result of the genetic bottleneck and a subsequent genetic release (*84, 85*). An analysis of the pairwise mismatches across all sites covered on the 1240K SNP panel shows a reduced variation within all Archaic Age individuals in the study (Fig. 2C), suggesting a bottleneck for the population. It is unclear, however, if this bottleneck occurred during the settlement as a result of a low population size reaching the islands, or if it occurred after colonisation through natural disasters or disease.

The data presented in this study suggests the search for one origin or one route of the Archaic Age settlers into the insular Caribbean is erroneous. The genetic diversity present in the hunter-gatherer populations from Cuba implies the possibility of multiple dispersal events, hence more than one origin and route for the initial settlement of the region. Key to this conclusion is a single individual excavated from the Cueva de Perico (CIP009) in the western tip of Cuba. The radiocarbon date of 2700 cal. BP held many surprises, pushing back the occupation of the site by 1,200 years and revealing an earlier occupational phase. Not much could be inferred from the anthropological analysis and archaeological context, as this skeleton was disturbed, possibly through later burial activity. As the single individual in this dataset, CIP009 showed an ancestry that could be modelled as a descendant of a population connected to individuals from the Californian Channel Islands (*210*) and a branch leading to the already described radiation which gave rise to all Central and South American populations today (*211*) (Fig. 3A, 2). This placement of CIP009 revealed a yet undescribed radiation event, which perhaps occurred on the North American continent (Fig. 3A, 1).

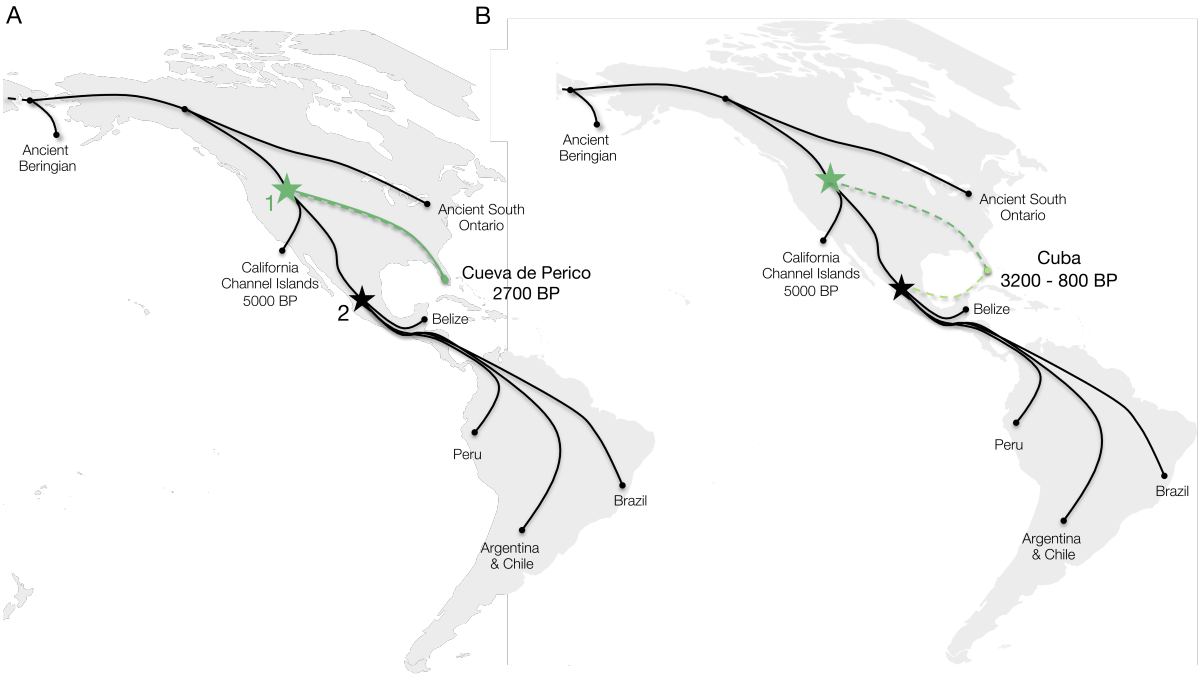


Figure 3 Tree-like representation of the ancestry modeling of Archaic Age individuals from the Western Caribbean. We are modeling the position of CIP009 (A) and all other Archaic-related individuals from Cuba (B). Stars denote radiation events implying a fast divergence of populations. 1 shows the newly revealed radiation, giving rise to individuals from the Californian Channel Islands, the individual CIP009 from the Archaic Age in the Western Caribbean and all ancient individuals and present-day groups in South and Central America. 2 shows the previously identified radiation of Central and South American populations (*211*). The geographical position of ancient groups approximates the site location. Arrows do not indicate dispersal routes and node placements do not show the geographical region where the populations split.

The connection to south-western North American populations does not necessarily mean one of those dispersals originated in North America, entering the Caribbean through the before hypothesised route via Florida. However, recent studies have foregrounded the Yucatan peninsula and South America as likely places of origin. Therefore, the data calls for a reconsideration of the route via Florida. Moreover, all other individuals with an Archaic Age context show multiple, divergent, ancestries. In addition to the component present in CIP009, all other individuals show a South American component. This component is, however, different from the specific one identified in the later dispersals from northeastern South America. The inability to identify the region of origin for either of the components present in the Archaic Age settlers of the Western Caribbean is, of course, unsatisfying, but seeing the few data available from the Americas perhaps not surprising. At the date of publication, this study doubled the available genomes from the Americas, highlighting the lack of suitable references to tie in newly produced genomes and satisfactorily model the events of the colonisation of the continents. The poor preservation in the tropical

climates lead to a paucity of genomes recovered, and models have to be built with few individuals over huge timeframes. Previous studies have shown that the population history of the Americas is complex and that ancient and present-day populations are, in some regions, the result of a mosaic of genetic replacements, and the product of continuity in other regions. Increasing the number of available genomes and closing the temporal and spatial gaps will allow more detailed reconstructions of the settlement and population history of the American continents. American ancestral populations have experienced bottleneck events. As a result of the reduced genetic variability in the Americas (*212*), the 1240K panel offers only little informative markers, as the panel was designed to differentiate worldwide populations. Questions of dispersals within the continents, especially Central and South America, can be answered by analysing genomes within the region. The design of a special SNP panel, targeting variable sites private to American populations, is a necessity to increase the number of expressive markers in American populations.

The analysis here has shown an indication of diversity within Archaic Age contexts in Cuba and no sign for admixture with Ceramic Age groups in Cuba. Targeted sampling of sites with differences in material culture, and in the timeframe of the arrival of the Ceramic Age expansion, will add to the understanding of diversity before, processes during the arrival of Ceramic Age groups, and the influence of Archaic Age communities on the development of the later Ceramic societies. To gain a better understanding of the origins of Archaic Age settlers, a dense sampling, especially in the Circum-Caribbean and targeted towards the timeframe of the first occurrences, is pivotal to identify the ancestral populations. The data presented in this thesis, together with the continued refining of methods for sampling and retrieval of ancient DNA, give a promising prospect of fulfilling this goal in the future.

8.3 Implications of Caribbean settlement for the seafaring abilities of hunter-gatherer- societies in the Americas.

The presence of multiple ancestries suggests the settlement of the Caribbean in the Archaic Age was not a singular event, but took place repeatedly, at least twice,

throughout the occupation of the islands. It raises questions regarding the ability of Archaic Age people to navigate open seascapes, an attribute generally not connected to hunter-gatherer populations.

Based on the earliest dates for human presence, the first islands to be settled in the (then) insular Caribbean were Cuba and Hispaniola (*121*). The points on the mainland closest to the islands are the Yucatan peninsula in Mexico in the west and Florida, separated through the waters of the Florida Strait in the north (Fig. 2A). With a distance from the mainland to the islands of 204 km and 185 km, respectively, the islands are not visible from the mainland and the currents separating them are relatively strong, ranging between 1.5 and 4.5 ktn (*213*). Therefore, crossing the water bodies into the Caribbean would be challenging without knowledge of navigation and seafaring.

There is much evidence recovered for connections of later Ceramic societies with the continents (*135, 136, 214*), but only a few vessels recovered from the islands themselves (*213*). Notably, the single trunk dugout canoes found on the islands do not correspond to the boats depicted in the illustrations of the early colonisers (*215*). Evidence for contact with the mainland from the Archaic Age is scarce. Botanical remains indicate contact between the Antilles and the mainland. In sites dating to the Archaic Age, yellow sapote and avocado seeds were recovered (*216*). Both species, like others found in the Antilles, are native to Central America and Mexico and have modes of dispersal that exclude introduction by drift over water (*217*). As expected by the fast decay of wood in tropical climates, no watercrafts have yet been recovered from Archaic Age contexts. The anthropological analysis of a woman excavated from the Angi site in Nicaragua, dating to the Archaic Age, is the first evidence for the use of watercraft in the Circum-Caribbean (*218*). The strong bones of the upper body and arms indicate this person had been rowing intensively throughout her life. Considering the lower water levels at the time, exposing various small islands bridging into the Greater Antilles, connections to the Caribbean islands seem possible (*219*). However, comparable anthropological analyses of skeletons from Archaic Age burials in the Antilles are pending.

The implications of archaeological, anthropological and genetic analyses for seafaring during the Archaic Age integrate into a broader discussion on the initial settlement of the American continents. The current evidence, including genetic research, supports an

Asian origin for the first Americans (*74, 220-222*). However, several hypotheses regarding events, routes and timing of dispersals into the Americas coexist, leaving the details of the initial settlement of the continents debated.

Geological and environmental research showed the Bering strait had been a land bridge and ice-free during the last glacial period (*223, 224*) and led to ice-free corridors into North America (*225*). Archaeogenetic research provided further support with the analysis of a late Pleistocene individual excavated from the Upward Sun River in Alaska, who contributed to the ancestry of all present-day Native Americans (*226*). A model integrating the pattern of sites showing a high density of the distinct stone tool tradition of the "Clovis- culture" (*227*) in eastern North America, with evidence from linguistic and biological analysis suggested a dispersal through the ice-free corridors (*228*). However, archaeological evidence for human occupation in different regions of the Americas raised the question regarding the dispersal routes. The oldest occupational sites were found in South America, at Monte Verde, dating to roughly 14,000 BP (*229*) conflicting with the theory of an overland dispersal. Recent studies, despite not universally accepted by the American archaeological community, suggest even older occupation at least 20,000 years ago (*230, 231*). There is mounting evidence that the first migrations into the Americas were not over land but by water, along the coast of the American continents and facilitated by the productive ecosystems of the kelp-forests (*232-234*). There are many indications for seafaring abilities in hunter-gatherer and Neolithic societies of the Americas. Mentioned before is the skeleton from Nicaragua, showing signs of extensive rowing through the strong upper body and arm bones. Apart from the Caribbean Islands, evidence for human presence is found on various smaller islands that most likely have never been connected to the continents. The Martha's Vineyard Island off the coast of Massachusetts shows signs of shellfish exploitation by 4000 BP, indicating water crossing by northeastern Native Americans (*235*). In the Californian Channel Islands, evidence for presence can be found starting around 7000 BP (*236, 237*) and analysis of ancient DNA of inhabitants living there 5,000 years ago shows a close relationship to the populations on the mainland (*210*).

Integrating the new findings from archaeology and genomics, it seems reasonable to genuinely integrate models of dispersal over water to reassess the colonisation process and movements in the Americas.

8.4 Towards a more detailed understanding of the settlement history of the Pacific.

Studies of the genetic diversity in the Pacific region (Fig. 4) have revealed a high level of differentiation among people inhabiting Near Oceania, despite lower genetic variation (*212, 238*). This differentiation seems to be not only the result of genetic drift, but is also the result of a complex history, involving periods of isolation, but more importantly several dispersals from other world regions and within the region. A landmark archaeogenetic study found the first settlers of western Remote Oceania from the islands of Vanuatu and Tonga, associated with the Lapita Cultural Complex, were genetically East-Asian, and therefore descendants not of the geographically closer populations of Papua New Guinea, but of the more distant islands of South East Asia (*190*). Despite settling a long-standing debate, the result was perhaps unsurprising, seeing the navigational skills of the early settlers were very well recognised. The movement of goods such as obsidian, pottery and other items between Lapita communities over distances of up to 2000 km is well documented (*239*). Together with the dispersal over 4000km to Vanuatu and Tonga, it attests to the effortless navigation of the seascapes. Computer simulations, taking into account actual conditions of winds and currents in the Pacific, showed that the journey could not have been achieved by drift, but was the result of intentional navigation (*104, 175*). However, analyses of present-day genomes of Pacific Islanders show that they have mixed ancestry, deriving 40 – 60% of ancestry from Papuan-related ancestral populations (*205, 206*).

Figure 4: The Pacific region and places of interest. Map of the Pacific and the adjacent landmasses, Near Oceania, where islands are intervisible, is underplayed in blue and Remote Oceania, where Islands are a minimum of 250 km apart, in green. Dotted line indicates the border of Near and Remote Oceania, with the majority of the Solomon Islands west and the Santa Cruz Islands (Solomon Islands) east.

Manuscript B has shown that the genetic composition of Remote Oceania is the result of subsequent dispersals following the settlement of the region by genetically unadmixed, Asian-related individuals. The study presented a genetically Papuan-related individual in the south of the Vanuatu Archipelago, and presence of the genetic component carried by this individual in others from the north and south. The inference of admixture dates revealed a peculiar pattern. The younger the C14 associated with the individuals in the analyses, the later the admixture event was inferred. The patterns observed led to the conclusion that, after initial settlement through East Asian-related populations, repeated gene flow of Papuan-related populations, likely from the Bismarck Archipelago, shaped the genetic make-up of present-day ni-Vanuatu in a century-long process. Only one day after publication of Manuscript B, a very similar study, focusing on the central islands of the Vanuatu Archipelago was presented (*240*). In this study, the exclusively Papuan-related ancestry of one individual was interpreted as an almost complete genetic turnover shortly after initial colonisation, leading to the conclusion of a substantial, Papuan-related migration into Vanuatu. Both conclusions might be correct, as different events could have shaped the different parts of the archipelago. However, the latter study failed to integrate the linguistic assessment of the islands. Vanuatu is the per capita linguistically most diverse region of the earth, with

125 languages spoken on the Archipelago today (*241*). Those languages exclusively belong to the Austronesian language family, which was reconstructed for the early settlers associated with the Lapita Cultural Complex. Moreover, archaeological evidence for a substantial migration into the archipelago is absent (*242*). Integrating the results of different disciplines allow for a more comprehensive and complex interpretation, very much in line with the conclusions of Manuscript B.

Additionally, results imply that Papuan-related populations in Near Oceania had seafaring abilities to match those of the Lapita Cultural Complex. Not only did they travel to Vanuatu, evidenced by two individuals dating between 2400 and 2300 BP, both in central and southern Vanuatu, but they must have undertaken the journey repeatedly, resulting in subsequent admixtures reflected in the increasingly more recent admixture dates. Archaeological assessment of obsidian artefacts from Near Oceania suggests the Indigenous Papuan-related groups were involved in long-distance trade and inter-island exchanges starting already 18 000 years ago (*243, 244*). The archaeogenetic studies were met with criticism, and a congregation of researchers on Pacific Archaeology pointed out, that, without understanding the genetic variation and possible interactions within Near Oceania, such conclusions were mere suggestions (*242*). Manuscript C investigates the ancient genetic diversity in Near Oceania. In addition to analysing a large time-transect in the Bismarck Archipelago, three sites on the Papua New Guinean mainland are investigated. Investigating the genetic composition of the coastal populations has become increasingly interesting, after the discovery of an extensive Lapita site on the south coast of Papua New Guinea (*245*). Before, it was believed that the settlements were restricted to smaller offshore islands (*246, 247*) and the Lapita Cultural Complex omitted large islands. Predicted by the Asian ancestry component in present-day coastal populations (*248*), we found all but two individuals harboured Asian-related ancestry. The only two individuals that seem to have exclusively Papuan-related ancestry are the two oldest in the dataset; both excavated on Watom Island in the Bismarck Archipelago. Dating to 3700 and 2600 BP, they are evidence that Papuan-related populations inhabited Watom Island before and after the arrival of the Austronesian expansion and the formation of the Lapita Cultural Complex. In a later individual, dated to 2100 BP, we find a twofold genetic ancestry, deriving 60% from Papuan-related populations and 40% from an Asian-related

population, most similar to the early Remote Oceanians from Vanuatu and Tonga. Surprisingly the date of the admixture event was inferred to 2300 BP, a millennium after the first occurrence of Lapita artefacts in the archipelago. The Late admixture date raises questions about the extent and nature of interactions and the exact timing. The pattern of admixture dates observed in Manuscript B calls for a closer examination of the events in the Bismarck Archipelago and a cautious interpretation. However, the result suggests the two divergent ancestries inhabited the Bismarck Archipelago in parallel, without genetic interaction. Surprisingly, the Lapita Cultural Complex appears in the Bismarck Archipelago fully developed, without preceding developmental stages. From the analysis of the ceramic styles and decorations, it is clear that the craftsmanship is different from similar styles in the Philippines, but it remains somewhat mysterious how a fully developed culture appears in the Bismarck Archipelago (*249*). It might be worthwhile to consider other islands as the origin of the Lapita Cultural Complex, perhaps further north, in the Mariana Islands. Expanding the archaeogenetic sampling to the north might elucidate the process of the development of the Lapita Cultural Complex and the formation of the genetic make-up of early Remote Oceanians. Based on the similarities of pottery styles and decorations, a dispersal 'from the Philippines via the Mariana Islands' to the Bismarck Archipelago has been proposed before (*250*). Radiocarbon dates of ceramic artefacts found in the Mariana Islands date the initial settlement of humans to 3500 BP (*249, 251*), according to palaeo-environmental evidence even earlier (*252*), suggesting the settlement occurred at the same time or even earlier to that of the Bismarck Archipelago.

The late admixture date of 2300 BP in an individual from Watom supports the contested statement of previous archaeogenetic research that the first settlers of western Remote Oceania arrived unadmixed. Nevertheless, an extension of the sampling to the larger islands of the Bismarck Archipelago will help evaluate whether the observations on Watom Island can be generalised.

The late admixture dates on Watom are outreached by those inferred on the mainland of Papua New Guinea. Differing between sites, we observe admixture dates ranging between 1500 and 650 BP, postdating the occurrence of Lapita artefacts by at least 1,500 years (*245*). Possibly an indication of long time parallel occupations of East Asian-related and Papuan-related populations, the differences between sites are

echoed in other genetic differences. Two nearby sites on the south coast showed significant differences not only in their proportions but also in the affinity of the Asian component to other East Asian populations. While one site showed more affinity to the ancient individuals from Vanuatu and Tonga, providing 50% of their ancestry, individuals of the other site, together with a site on the northern coast, showed a smaller Asian-related ancestry component ~15 - 35%. Their affinity was higher to Austronesian populations more similar to the ancestors of the Lapita Cultural Complex and those of present-day Taiwanese and Philippine populations. From the differences in the genetic proportions and admixture dates, it is clear that the two nearby sites, despite the temporal overlap, show indications of different interactions with (Austronesian) coastal and island populations and (Papuan) inland populations. In one site, showing higher East Asian-related ancestry, one particularly late admixture date is observed. However, we lack the resolution to identify which populations contributed and where their geographical source was. The different genetic compositions of the two nearby sites show that the southern coast was a genetic, and possibly also a cultural and linguistic mosaic of people, matching the situation today. The ancient genetic diversity is reflected by the languages spoken in the region today: The Motu language is part of a western branch of Central Papuan Austronesian languages (*253, 254*), and is spoken mostly by people located on the coasts. The Papuan language dominant in the region, Koita (*255*), is spoken in settlements more inland (*256-259*). For the site on the northern coast, archaeological evidence dates the establishment of the site to 650 BP (*260*), and places it in a local trade network connected to the Bismarck Archipelago. However, the genetic affinities suggest the material exchange did not extend to genetic exchange. Linguistic evidence and oral traditions suggest an origin in the Vitaiz Strait in the Bismarck Sea for Bel, the Austronesian language spoken in the region. They would imply another case of discontinuity between linguistic and genetic evidence. The oral traditions might have started recording the history after the establishment of the site, and the Bel language was adopted to facilitate trade. As the individuals are dated to the early occupational phase, their descendants might show higher affinity to populations from the Bismarck Archipelago.

As mentioned before, one individual from the south coast shows a particularly recent admixture date of 650 BP. It coincides with the resurgence of trade routes after a

period in which unfavourable conditions (*261-263*) resulted in a dearth of sites, called the 'Papuan Hiccup' (*261*). After 700 BP, settlement sites reappear across the region, but it remains unclear whether this is the return of the previous occupants or sites established by entirely new populations (*264*). The isotopic analysis of this individual suggests it is a non-local (*265*), possibly favouring the re-establishment of the site by new populations. However, we lack the resolution to identify which populations contributed and where their geographical source was. Most groups occupying the coasts show similar ancestry proportions. Therefore, a more detailed analysis is necessary to differentiate different cultural groups. An analysis not feasible with genetic data alone. One challenge for future archaeogenetic work in the Pacific will therefore lie in the refinement of methods, to disentangle the connections between populations shaped by very similar admixture events. After the broad strokes of initial colonisation and formation of gene pools, the highly regionalised interaction spheres call for a better understanding of the history of smaller regions within the Pacific, or even of individual islands. Integration of archaeogenetic and archaeological data, as well as linguistic and cultural studies, might help disentangle the complex and closely intertwined regional histories.

A recent analysis of present-day genomes from central Remote Oceania claimed the arrival of South American genetic component in the Pacific before the settlement of Rapa Nui. It suggested the first settlers of central Remote Oceania could have been Americans, genetically most similar to Indigenous inhabitants of present-day Colombia (*266*). After arrival in the region, they mixed with settlers from the west, resulting in the genetic make-up observed in the analysed individuals. Contact between the Pacific and the American populations is well accepted among researchers. Analyses of domesticates such as the sweet potato and chicken (*267, 268*), show an early contact with South America (*269*), inferred to the time before the colonisation of Rapa Nui, Aotearoa and Hawai'i, as suggested by the authors. However, there is much to be criticised about the study, from details of assumptions for the models to the lack of incorporation of archaeological evidence. The claim of a first settlement of Polynesia by South American populations is currently not supported by all other disciplines, such as archaeology or archaeogenetics. Arguably, there are only a few ancient genomes

recovered from central Remote Oceania. Manuscript B includes only three individuals excavated from Rai'atea, post-dating the alleged mixture. In all three genomes, no American component could be detected (Fig 2A of Manuscript B). In 8.3, I discussed the evidence for seafaring capabilities of pre-contact American populations, and it is not entirely improbable that this journey could have been achieved. Nevertheless, the evidence at hand makes a contact through Polynesians, voyaging to the shores of the Americas, more likely than vice versa.

Further analysis of ancient genomes, together with a reassessment of the oral histories, the archaeological record and new surveys of sites in central Remote Oceania will provide clarity whether the settlement history of Remote Oceania will have to be rewritten entirely, or merely extended by a pre-historic voyage of northern South Americans into the Pacific.

8.5 Conclusion

Western explorers have inclined to assume they alone were capable of intentional exploration over oceans, attributing the colonisation of islands by less or unindustrialised societies to accidental dispersal (*270*). Nevertheless, Pacific Islanders colonised earlier and farther than the first European explorers, and the pattern of colonisation shows that this expansion was intentional rather than accidental (*104, 105*). The mounting evidence for seafaring on the American continents suggests that, also in this region, people have crossed waters with intent, crossing challenging straits to settle the Caribbean Islands (*213*). In addition to providing new insights to questions regarding the initial and subsequent settlements into the Caribbean and the Pacific, this work adds to a body of research that demonstrates how bodies of water, when examined from a human biological and cultural perspective, are not necessarily perceived as barriers but connect populations. This ties into a worldwide comparison of the colonisation history of islands by humans, showing that some quite the opposite observations from the original assumptions are a global phenomenon (*77*). The attractive idea of islands as “ideal laboratories for studying evolution” (*271*) might apply to most animal and plant species, but for humans, more complex models have to be

considered. Apart from the ecological configuration of the land, expansion of subsistence resources on neighbouring islands (*272*) and the surrounding marine resources may have influenced patterns of settlement (*273, 274*). Smaller islands might have been settled as part of growing trade networks (*273, 275*) and cultural effects might have played a role in direction and speed of exploration and colonisation. The discovery of the Azores, Madeira and the Canaries encouraged Columbus to sail west, expecting more islands on the way. The subsequent discovery of the Americas, especially the tales of the wealth, triggered explorations of all oceans. This autocatalytic effect (*276, 277*) might have been similar in pre-historic populations. Especially in the Pacific, where explorations took place against winds and currents facilitating the journey home, tales of the returnees might have triggered additional explorations and technological advances. For populations with the knowledge of decent from voyagers, the oral traditions might have stimulated them to undertake further explorations. The answers to questions about the motivation of peoples to colonise and recolonise close and distant islands are, of course, speculative. However, archaeogenetic results can provide one more line of evidence in the reconstruction of such events and societies. Together with anthropology, archaeology and linguistics, and respecting the oral histories, a more detailed picture can be obtained, and old narratives based on and biased by the observations of colonisers can be challenged successfully.

9. OUTLOOK

The results presented in this thesis would not have been possible without the advancements in technologies in the sampling, processing and bioinformatical analysis of ancient genomes. The possibility to recover DNA from tropical human remains is a landmark in the development of the field. It will allow reconstructing a more wholesome picture of the dispersal of our species across the globe. With the continually improving protocols, we can expect the recovery not only of more, but of older genomes in tropical regions, and advances in data analysis give promising prospects of answering the open questions outlined in the discussion.

10. REFERENCES

1. R. E. Green *et al.*, A draft sequence of the Neandertal genome. *science* **328**, 710-722 (2010).
2. J. Krause *et al.*, The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. *Nature* **464**, 894-897 (2010).
3. D. Reich *et al.*, Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468**, 1053 (2010).
4. M. Meyer *et al.*, Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. *Nature* **531**, 504-507 (2016).
5. R. Tutton, "They want to know where they came from": population genetics, identity, and family genealogy. *New Genetics and Society* **23**, 105-120 (2004).
6. M. D. Shriver, R. A. Kittles, Genetic ancestry and the search for personalized genetic histories. *Nature Reviews Genetics* **5**, 611-618 (2004).
7. E. Clay, Mediated science, genetics and identity in the US African diaspora. *Media, Spiritualities and Social Change*, 25-36 (2011).
8. A. Nelson, The factness of diaspora. *Revisiting race in a genomic age*, 253-268 (2008).
9. J. D. Forbes, *Columbus and other cannibals: The Wetiko disease of exploitation, imperialism, and terrorism*. (Seven Stories Press, 2011).
10. L. Guitar, Documenting the myth of Taino extinction. *Journal of Caribbean Amerindian History and Anthropology*, (2002).
11. R. Higuchi, B. Bowman, M. Freiberger, O. A. Ryder, A. C. Wilson, DNA sequences from the quagga, an extinct member of the horse family. *Nature* **312**, 282-284 (1984).
12. S. Pääbo, Molecular cloning of ancient Egyptian mummy DNA. *nature* **314**, 644-645 (1985).
13. M. Hofreiter, D. Serre, H. N. Poinar, M. Kuch, S. Pääbo, ancient DNA. *Nature Reviews Genetics* **2**, 353-359 (2001).
14. L. C. Renfrew, *Before civilization*. (Random House, 1973).
15. M. Rasmussen *et al.*, Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **463**, 757-762 (2010).
16. C. Warinner *et al.*, Direct evidence of milk consumption from ancient human dental calculus. *Scientific reports* **4**, 1-6 (2014).
17. W. L. Salo, A. C. Aufderheide, J. Buikstra, T. A. Holcomb, Identification of Mycobacterium tuberculosis DNA in a pre-Columbian Peruvian mummy. *Proceedings of the National Academy of Sciences* **91**, 2091-2094 (1994).
18. K. Yoshida *et al.*, The rise and fall of the Phytophthora infestans lineage that triggered the Irish potato famine. *Elife* **2**, e00731 (2013).
19. K. Swarts *et al.*, Genomic estimation of complex traits reveals ancient maize adaptation to temperate North America. *Science* **357**, 512-515 (2017).
20. H. Poinar, M. Kuch, G. McDonald, P. Martin, S. Pääbo, Nuclear gene sequences from a late Pleistocene sloth coprolite. *Current Biology* **13**, 1150-1152 (2003).
21. V. Slon *et al.*, Neandertal and Denisovan DNA from Pleistocene sediments. *Science* **356**, 605-608 (2017).
22. E. M. Golenberg *et al.*, Chloroplast DNA sequence from a Miocene Magnolia species. *Nature* **344**, 656-658 (1990).
23. R. DeSalle, J. Gatesy, W. Wheeler, D. Grimaldi, DNA sequences from a fossil termite in Oligo-Miocene amber and their phylogenetic implications. *Science* **257**, 1933-1936 (1992).

24. R. J. Cano, H. N. Poinar, N. J. Pieniazek, A. Acra, G. O. Poinar, Amplification and sequencing of DNA from a 120–135-million-year-old weevil. *Nature* **363**, 536-538 (1993).
25. N. Weyand, M. Bunnell, DNA sequence from Cretaceous period bone fragments. *Science* **266**, 1229-1232 (1994).
26. I. Marota, C. Basile, M. Ubaldi, F. Rollo, DNA decay rate in papyri and human remains from Egyptian archaeological sites. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* **117**, 310-318 (2002).
27. E. Willerslev *et al.*, Long-term persistence of bacterial DNA. *Current Biology* **14**, R9-R10 (2004).
28. M. T. P. Gilbert *et al.*, Long‐term survival of ancient DNA in Egypt: Response to Zink and Nerlich (2003). *American journal of physical anthropology* **128**, 110-114 (2005).
29. J. Binladen, M. T. P. Gilbert, E. Willerslev, 800 000 year old mammoth DNA, modern elephant DNA or PCR artefact? *Biology Letters* **3**, 55-57 (2007).
30. L. Orlando *et al.*, Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499**, 74-78 (2013).
31. T. Lindahl, Instability and decay of the primary structure of DNA. *nature* **362**, 709-715 (1993).
32. Z. Darzynkiewicz *et al.*, Cytometry in cell necrobiology: analysis of apoptosis and accidental cell death (necrosis). *Cytometry: The Journal of the International Society for Analytical Cytology* **27**, 1-20 (1997).
33. G. Eglinton, G. A. Logan, Molecular preservation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **333**, 315-328 (1991).
34. T. Lindahl, A. Andersson, Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid. *Biochemistry* **11**, 3618-3623 (1972).
35. J. Dabney *et al.*, Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci U S A* **110**, 15758-15763 (2013).
36. C. J. Adler, W. Haak, D. Donlon, A. Cooper, G. Consortium, Survival and recovery of DNA from ancient teeth and bones. *Journal of Archaeological Science* **38**, 956-964 (2011).
37. M. E. Allentoft *et al.*, The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proceedings of the Royal Society B: Biological Sciences* **279**, 4724-4733 (2012).
38. L. Kistler, R. Ware, O. Smith, M. Collins, R. G. Allaby, A new model for ancient DNA decay based on paleogenomic meta-analysis. *Nucleic Acids Research* **45**, 6310-6320 (2017).
39. A. W. Briggs *et al.*, Patterns of damage in genomic DNA sequences from a Neandertal. *Proceedings of the National Academy of Sciences* **104**, 14616-14621 (2007).
40. D. H. Huson, A. F. Auch, J. Qi, S. C. Schuster, MEGAN analysis of metagenomic data. *Genome research* **17**, 377-386 (2007).
41. A. Kazarina *et al.*, Analysis of the bacterial communities in ancient human bones and burial soil samples: Tracing the impact of environmental bacteria. *Journal of Archaeological Science* **109**, 104989 (2019).
42. A. E. Mann *et al.*, Differential preservation of endogenous human and microbial DNA in dental calculus and dentin. *Scientific reports* **8**, 1-15 (2018).
43. H. Erlich, R. Saiki, Specific DNA amplification. *Nature* **331**, 461-462 (1988).

44. R. K. Saiki *et al.*, Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350-1354 (1985).
45. J. Krause *et al.*, A complete mtDNA genome of an early modern human from Kostenki, Russia. *Current Biology* **20**, 231-236 (2010).
46. S. Pääbo, A. C. Wilson, Polymerase chain reaction reveals cloning artefacts. *Nature* **334**, 387-388 (1988).
47. S. Pääbo, Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. *Proceedings of the National Academy of Sciences* **86**, 1939-1943 (1989).
48. S. Pääbo *et al.*, Genetic analyses from ancient DNA. *Annu Rev Genet* **38**, 645-679 (2004).
49. T. S. A. Korneliussen, A; Nielsen, Rasmus, ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 13 (2014).
50. G. Renaud, V. Slon, A. T. Duggan, J. Kelso, Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol* **16**, 224 (2015).
51. N. J. Nakatsuka *et al.*, ContamLD: Estimation of Ancient Nuclear DNA Contamination Using Breakdown of Linkage Disequilibrium. *bioRxiv*, (2020).
52. P. Skoglund *et al.*, Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proceedings of the National Academy of Sciences* **111**, 2229-2234 (2014).
53. M. Margulies *et al.*, Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376-380 (2005).
54. D. R. Bentley *et al.*, Accurate whole human genome sequencing using reversible terminator chemistry. *nature* **456**, 53-59 (2008).
55. M. Meyer, M. Kircher, Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* **2010**, pdb prot5448 (2010).
56. M.-T. Gansauge, A. Aximu-Petri, S. Nagel, M. Meyer, 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. *Nature Protocols*, 1-23 (2020).
57. M.-T. Gansauge *et al.*, Single-stranded DNA library preparation from highly degraded DNA usingT4DNA ligase. *Nucleic Acids Research*, gkx033 (2017).
58. M. Kircher, S. Sawyer, M. Meyer, Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* **40**, e3 (2012).
59. A. Peltzer *et al.*, EAGER: efficient ancient genome reconstruction. *Genome Biol* **17**, 60 (2016).
60. H. A. Burbano *et al.*, Targeted investigation of the Neandertal genome by array-based sequence capture. *science* **328**, 723-725 (2010).
61. Q. Fu *et al.*, A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr Biol* **23**, 553-559 (2013).
62. T. Maricic, M. Whitten, S. Pääbo, Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PloS one* **5**, e14004 (2010).
63. R. Pinhasi *et al.*, Optimal ancient DNA yields from the inner ear part of the human petrous bone. *PloS one* **10**, e0129102 (2015).
64. C. Parker *et al.*, A systematic investigation of human DNA preservation in medieval skeletons. *bioRxiv*, (2020).
65. R. Sanjek, Rethinking migration, ancient to future. *Global networks* **3**, 315-336 (2003).
66. G. Kossina. (Kabitzsch, Würzburg, 1911).

67. V. Heyd, in *Transitional Landscapes? The 3rd Millennium BC in Europe.* (Habelt, 2016), pp. 53-85.
68. K. Kristiansen *et al.*, Re-theorising mobility and the formation of culture and language among the Corded Ware Culture in Europe. *antiquity* **91**, 334-347 (2017).
69. B. Bramanti *et al.*, Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *science* **326**, 137-140 (2009).
70. W. Haak *et al.*, Ancient DNA from European early neolithic farmers reveals their near eastern affinities. *PLoS Biol* **8**, e1000536 (2010).
71. P. Skoglund *et al.*, Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* **336**, 466-469 (2012).
72. I. Lazaridis *et al.*, Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409-413 (2014).
73. W. Haak *et al.*, Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207-211 (2015).
74. M. Raghavan *et al.*, Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87-91 (2014).
75. B. Bradley, D. Stanford, The North Atlantic ice-edge corridor: a possible Palaeolithic route to the New World. *World Archaeology* **36**, 459-478 (2004).
76. J. C. Chatters, The recovery and first analysis of an Early Holocene human skeleton from Kennewick, Washington. *American antiquity*, 291-316 (2000).
77. W. F. Keegan, J. M. Diamond, in *Advances in archaeological method and theory*. (Elsevier, 1987), pp. 49-92.
78. R. J. Whittaker, J. M. Fernández-Palacios, *Island biogeography: ecology, evolution, and conservation*. (Oxford University Press, 2007).
79. E. Mayr, Change of genetic environment and evolution. (1954).
80. N. H. Barton, G. M. Hewitt, Adaptation, speciation and hybrid zones. *Nature* **341**, 497-503 (1989).
81. B. Grant, D. Owen, C. Clarke, Parallel rise and fall of melanic peppered moths in America and Britain. *Journal of Heredity* **87**, 351-357 (1996).
82. R. Berry, A. Berry, T. Anderson, P. Scriven, The house mice of Faray, Orkney. *Journal of Zoology* **228**, 233-246 (1992).
83. M. C. Whitlock, in *Ecology, genetics and evolution of metapopulations*. (Elsevier, 2004), pp. 153-173.
84. R. Chakraborty, M. Nei, Dynamics of gene differentiation between incompletely isolated populations of unequal sizes. *Theoretical Population Biology* **5**, 460-469 (1974).
85. R. Chakraborty, M. Nei, Bottleneck effects on average heterozygosity and genetic distance with the stepwise mutation model. *Evolution*, 347-356 (1977).
86. J. R. Weinberg, V. R. Starczak, D. Jörg, Evidence for rapid speciation following a founder event in the laboratory. *Evolution* **46**, 1214-1220 (1992).
87. H. L. Carson, Genetic change after colonization. *GeoJournal* **28**, 297-302 (1992).
88. R. Woods *et al.*, Rapid size change associated with intra-island evolutionary radiation in extinct Caribbean "island-shrews". *BMC Evolutionary Biology* **20**, 1-12 (2020).
89. M. V. Lomolino, Immigrant selection, predation, and the distributions of Microtus pennsylvanicus and Blarina brevicauda on islands. *The American Naturalist* **123**, 468-483 (1984).
90. M. V. Lomolino, Body size of mammals on islands: the island rule reexamined. *The American Naturalist* **125**, 310-316 (1985).
91. M. V. Lomolino, Body size evolution in insular vertebrates: generality of the island rule. *Journal of Biogeography* **32**, 1683-1699 (2005).

92. Q. Fu *et al.*, Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445-449 (2014).
93. C. Darwin, in *The voyage of the Beagle*. (JM Dent and Sons, London, 1845).
94. E. N. Arnold, Indian Ocean giant tortoises: their systematics and island adaptations. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* **286**, 127-145 (1979).
95. D. Lack, Island birds. *Biotropica*, 29-31 (1970).
96. P. Brown *et al.*, A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia. *Nature* **431**, 1055-1061 (2004).
97. M. J. Morwood *et al.*, Archaeology and age of a new hominin from Flores in eastern Indonesia. *Nature* **431**, 1087-1091 (2004).
98. Q. Cronk, The history of endemic flora of St Helena: a relictual series. *New Phytologist* **105**, 509-520 (1987).
99. A. Engler, *Versuch einer Entwicklungsgeschichte der extratropischen Florengebiete der südlichen Hemisphäre und der tropischen Gebiete: Mit einer pflanzengeographischen Erdkarte*. (Engelmann, 1882).
100. E. J. Censky, K. Hodge, J. Dudley, Over-water dispersal of lizards due to hurricanes. *Nature* **395**, 556-556 (1998).
101. E. Holzapfel, J. Harrell, Transoceanic dispersal studies of insects. *Pacific Insects* **10**, 115-153 (1968).
102. S. Renner, Plant dispersal across the tropical Atlantic by wind and sea currents. *International Journal of Plant Sciences* **165**, S23-S33 (2004).
103. R. H. MacArthur. (Princeton Univ. Press, Princeton, NJ, 1967).
104. M. Levison, *The settlement of Polynesia: a computer simulation*. (Australian National University Press, 1973).
105. B. R. Finney, Anomalous westerlies, El Nino, and the colonization of Polynesia. *American Anthropologist* **87**, 9-26 (1985).
106. J. Smithers, B. Smit, Human adaptation to climatic variability and change. *Global environmental change* **7**, 129-146 (1997).
107. M. Grove, Palaeoclimates, plasticity, and the early dispersal of Homo sapiens. *Quaternary International* **369**, 17-37 (2015).
108. C. Jeong, A. Di Rienzo, Adaptations to local environments in modern human populations. *Current opinion in genetics & development* **29**, 1-8 (2014).
109. J. F. Cherry, Islands out of the stream: isolation and interaction in early East Mediterranean insular prehistory. *Prehistoric production and exchange: the Aegean and eastern Mediterranean*, 12-29 (1985).
110. J. Terrell, *Comparative Study of Human and Lower Animal Biogeography in the Solomon Islands: The Text of a Paper Prepared for the Smithsonian Conference on the Application of Models in Theoretical Biology and Biogeography to Archeology and Anthropology (April 30-May 2, 1974)*. (Field Museum of Natural History, 1974).
111. J. Terrell, Island biogeography and man in Melanesia. *Archaeology & Physical Anthropology in Oceania* **11**, 1-17 (1976).
112. J. Habu, H. Junko, *Ancient Jomon of Japan*. (Cambridge University Press, 2004), vol. 4.
113. S.-M. Ahn, The emergence of rice agriculture in Korea: archaeobotanical perspectives. *Archaeological and anthropological sciences* **2**, 89-98 (2010).
114. T. Gakuhari *et al.*, Ancient Jomon genome sequence analysis sheds light on migration patterns of early East Asian populations. *Communications biology* **3**, 1-10 (2020).
115. H. Kanzawa-Kiriyama *et al.*, A partial nuclear genome of the Jomons who lived 3000 years ago in Fukushima, Japan. *Journal of human genetics* **62**, 213-221 (2017).

116. T. Jinam *et al.*, The history of human populations in the Japanese Archipelago inferred from genome-wide SNP data with a special reference to the Ainu and the Ryukyuan populations. *Journal of human genetics* **57**, 787 (2012).
117. D. M. Fernandes *et al.*, The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean. *Nature Ecology & Evolution* **4**, 334-345 (2020).
118. V. Lull, R. M. Pérez, C. R. Herrada, R. Risch, Los cambios sociales en las islas Baleares a lo largo del II milenio. *Cypsela: revista de prehistòria i protohistòria*, 123-148 (2005).
119. I. Lazaridis *et al.*, Genetic origins of the Minoans and Mycenaeans. *Nature* **548**, 214-218 (2017).
120. K. B. Tankersley, N. P. Dunning, L. A. Owen, J. Sparks, GEOCHRONOLOGY AND PALEOENVIRONMENTAL FRAMEWORK FOR THE OLDEST ARCHAEOLOGICAL SITE (7800–7900 cal BP) IN THE WEST INDIES, BANWARI TRACE, TRINIDAD. *Latin American Antiquity* **29**, 681-695 (2018).
121. M. F. Napolitano *et al.*, Reevaluating human colonization of the Caribbean using chronometric hygiene and Bayesian modeling. *Science Advances* **5**, eaar7806 (2019).
122. P. Siegel *et al.*, Ecosystem engineering during the human occupations of the Lesser Antilles. *Early Settlers of the Insular Caribbean: Dearchaizing the Archaic, CL Hofman, AT Antczak, Eds.(Leiden, Sidestone Press, 2019) pp*, 77-88 (2019).
123. P. Goldberg, C. E. Miller, S. M. Mentzer, Recognizing fire in the Paleolithic archaeological record. *Current Anthropology* **58**, S175-S190 (2017).
124. Y. C. d. Armas, M. Roksandic, R. R. Suarez, D. G. Smith, W. M. Buhay, in *Cuban Archaeology in the Caribbean,* I. Roksandic, Ed. (University of Florida Press, 2016).
125. Y. Chinique de Armas *et al.*, Starch analysis and isotopic evidence of consumption of cultigens among fisher–gatherers in Cuba: the archaeological site of Canímar Abajo, Matanzas. *Journal of Archaeological Science* **58**, 121-132 (2015).
126. W. R. Coe, A distinctive artifact common to Haiti and Central America. *American Antiquity* **22**, 280-282 (1957).
127. S. M. Wilson, H. B. Iceland, T. R. Hester, Preceramic connections between Yucatan and the Caribbean. *Latin American Antiquity*, 342-352 (1998).
128. I. Rouse. (JSTOR, 1994).
129. J. Febles, Estudio comparativo de las industrias de la piedra tallada de Aguas Verdes (Baracoa) y Playitas (Matanzas): Probable relación de estas industrias con otras del SE de los Estados Unidos. *Arqueologia de Cuba y de otras dreas Antillanas, Editorial Academia, La Habana*, 312-371 (1991).
130. J. M. Cruxent, I. Rouse, An Archaeological Chronology of Venezuela. *American Antiquity* **26**, 2 (1961).
131. I. Rouse, The Classification of Artifacts in Archaeology. *American Antiquity* **25**, 313-323 (1960).
132. R. T. Callaghan, Ceramic age seafaring and interaction potential in the Antilles: a computer simulation. *Current Anthropology* **42**, 308-313 (2001).
133. A. H. Ross, W. F. Keegan, M. P. Pateman, C. B. Young, faces Divulge the origins of caribbean prehistoric inhabitants. *Scientific Reports* **10**, 1-9 (2020).
134. H. Schroeder *et al.*, Origins and genetic legacies of the Caribbean Taino. *Proc Natl Acad Sci U S A* **115**, 2341-2346 (2018).
135. I. Rouse, *The Tainos: Rise and decline of the people who greeted Columbus*. (Yale University Press, 1992).
136. J. Granberry, G. Vescelius, *Languages of the pre-Columbian Antilles*. (University of Alabama Press, 2004).
137. L. Chanlatte Baik, Y. Narganes-Sorde, Cultura La Hueca. *Museo de Historia, Antropología y Arte, Universidad de Puerto Rico, Río Piedras*, (2005).

138. I. Rouse, *The Tainos: rise & decline of the people who greeted Columbus*. (Yale University Press, 1992).
139. S. M. Fitzpatrick, M. Kappers, C. M. Giovas, The southward route hypothesis: examining Carriacou's chronological position in Antillean prehistory. (2010).
140. R. T. Callaghan, Crossing the Guadeloupe passage in the Archaic Age. *Island Shores, Distant Pasts: Archaeological and Biological Approaches to the Pre-Columbian Settlement of the Caribbean*, 127-147 (2010).
141. W. F. Keegan, West Indian archaeology. 1. Overview and foragers. *Journal of Archaeological Research* **2**, 255-284 (1994).
142. W. F. Keegan, Modeling dispersal in the prehistoric West Indies. *World Archaeology* **26**, 400-420 (1995).
143. C. O. Sauer, *The early Spanish main*. (Cambridge University Press, 1966).
144. W. F. Keegan, Islands of chaos. *BAR INTERNATIONAL SERIES*, 33-46 (2004).
145. E. Ortega, J. Guerrero, *Estudio de cuatro nuevos sitios paleoarcaicos de la isla de Santo Domingo*. (Museo del hombre dominicano, 1981).
146. R. O. Rimoli, J. E. Nadal, *El horizonte ceramista temprano en Santo Domingo y otras Antillas*. (Editora de la UASD, 1983).
147. M. Veloz, E. Ortega, P. Pina, El Caimito: Un Antiguo Complejo Ceramista de las Antillas Mayores. *Santo Domingo: Fundación García Arévalo*, (1974).
148. M. Veloz Maggiolo, I. Vargas, M. Sanoja, F. L. Calderón, Arqueología de Yuma (Republica Dominicana). *Taller, Santo Domingo*, (1976).
149. P. Godo, Contextos arqueológicos del protoagrícola en el centro-occidente de Cuba. *El Caribe Arqueológico* **5**, 62-75 (2001).
150. P. P. Godo, El problema del protoagricola de Cuba: discusion y perspectivas. *El Caribe Arqueológico* **2**, 19-30 (1997).
151. J. Reyes Carderò, Apropiación y tradición alimentaria en el oriente cubano. *El Caribe Arqueológico* **5**, 42-51 (2001).
152. W. F. Keegan, Archaic influences in the origins and development of Taino societies. *Caribbean Journal of Science* **42**, 1 (2006).
153. M. F. De Navarrete, *Colección de las viajes y descubrimentos que hicieron por mar los españoles desde fines del siglo XV..., 5*. (1837).
154. D. D. Davis, R. C. Goodwin, Island Carib origins: evidence and nonevidence. *American Antiquity*, 37-48 (1990).
155. C. L. Hofman, M. L. Hoogland, Caribbean encounters: rescue excavations at the early colonial Island Carib site of Argyle, St. Vincent. *Analecta Praehistorica Leidensia* **43**, 63-76 (2012).
156. C. L. Hofman *et al.*, Indigenous Caribbean perspectives: archaeologies and legacies of the first colonised region in the New World. *antiquity* **92**, 200-216 (2018).
157. L. Honychurch, *Carib to Creole: A history of contact and culture exchange*. (Dominica Institute, 2000).
158. P. Twinn, Land ownership and the construction of Carib identity in St. Vincent. *Indigenous resurgence in the contemporary Caribbean: Amerindian survival and revival*, 89-106 (2006).
159. R. A. Myers, Island Carib Cannibalism. *New West Indian Guide/Nieuwe West-Indische Gids* **58**, 147-184 (1984).
160. R. B. Moore, Carib" cannibalism": A study in anthropological stereotyping. *Caribbean Studies* **13**, 117-135 (1973).
161. J. Sued-Badillo, *Los caribes: realidad o fábula: ensayo de rectificación histórica*. (Edit. Antillana, 1978).
162. S. F. Cook, W. Borah, The aboriginal population of Hispaniola. *Essays in population history* **1**, 376-410 (1971).

163. D. D. Davis, The Strategy of Early Spanish Ecosystem Management on Cuba. *Journal of Anthropological Research* **30**, 294-314 (1974).
164. I. Altman, S. Cline, J. J. Pescador, *The Early History of Greater Mexico*. (Prentice Hall, 2003).
165. M. Stevenson, Columbus and the war on indigenous peoples. *Race & Class* **33**, 27-45 (1992).
166. N. L. Whitehead, Carib cannibalism. The historical evidence. *Journal de la Société des Américanistes*, 69-87 (1984).
167. A. Boomert, The Cayo complex of St. Vincent: Ethnohistorical and archaeological aspects of the Island Carib problem. *Antropológica* **66**, 3-68 (1986).
168. A. K. Bullen, R. P. Bullen, in *Proceedings of the 3rd International Congress for the Study of Pre-Columbian Cultures of the Lesser Antilles. Grenada: Grenada National Museum*. (1970), pp. 61-86.
169. R. P. Bullen, *The Archaeology of Grenada, West Indies*. (University of Florida, 1964).
170. R. P. Bullen, A. K. Bullen, *Archaeological Investigations on St. Vincent and the Grenadines, West Indies*. (Bryant Foundation, 1972).
171. A. Boomert, Koriabo and the polychrome tradition: the late-prehistoric era between the Orinoco and Amazon mouths. *BAR International Series*, 251-268 (2004).
172. L. Allaire, Ph. D. dissertation, yale University. University Microfilms, Ann Arbor, Mich, (1977).
173. A. Pawley, R. Green, Dating the dispersal of the Oceanic languages. *Oceanic Linguistics* **12**, 1-67 (1973).
174. K. Lambeck, H. Rouby, A. Purcell, Y. Sun, M. Sambridge, Sea level and global ice volumes from the Last Glacial Maximum to the Holocene. *Proceedings of the National Academy of Sciences* **111**, 15296-15303 (2014).
175. G. Irwin, S. Bickler, P. Quirke, Voyaging by canoe and computer: experiments in the settlement of the Pacific Ocean. *Antiquity* **64**, 34-50 (1990).
176. C. Clarkson *et al.*, Human occupation of northern Australia by 65,000 years ago. *Nature* **547**, 306-310 (2017).
177. G. R. Summerhayes, J. H. Field, B. Shaw, D. Gaffney, The archaeology of forest exploitation and change in the tropics during the Pleistocene: The case of Northern Sahul (Pleistocene New Guinea). *Quaternary International* **448**, 14-30 (2017).
178. S. Hawkins *et al.*, Oldest human occupation of Wallacea at Laili Cave, Timor-Leste, shows broad-spectrum foraging responses to late Pleistocene environments. *Quaternary Science Reviews* **171**, 58-72 (2017).
179. M. Aubert *et al.*, Pleistocene cave art from Sulawesi, Indonesia. *Nature* **514**, 223-227 (2014).
180. J. Sun *et al.*, Paternal gene pool of Malays in Southeast Asia and its applications for the early expansion of Austronesians. *American Journal of Human Biology*, e23486 (2020).
181. R. D. Gray, A. J. Drummond, S. J. Greenhill, Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *science* **323**, 479-483 (2009).
182. P. Bellwood, *Austronesian prehistory in Southeast Asia: homeland, expansion and transformation*. (Canberra: ANU E Press, 1995).
183. C. Gosden *et al.*, Lapita sites of the Bismarck Archipelago. *Antiquity* **63**, 561-586 (1989).
184. J. Specht, C. Gosden, Dating Lapita pottery in the Bismarck Archipelago, Papua New Guinea. *Asian Perspectives*, 175-199 (1997).
185. I. Pugach *et al.*, The gateway from Near into Remote Oceania: new insights from genome-wide data. *Molecular biology and evolution* **35**, 871-886 (2018).
186. R. Walter, P. Sheppard, *Archaeology of the Solomon Islands*. (2017).

187. F. Petchey, M. Spriggs, S. Bedford, F. Valentin, H. Buckley, Radiocarbon dating of burials from the Teouma Lapita cemetery, Efate, Vanuatu. *Journal of Archaeological Science* **50**, 227-242 (2014).
188. F. Valentin, S. Bedford, H. R. Buckley, M. Spriggs, Lapita burial practices: evidence for complex body and bone treatment at the Teouma cemetery, Vanuatu, Southwest Pacific. *Journal of Island & Coastal Archaeology* **5**, 212-235 (2010).
189. F. Petchey, P. V. Kirch, The importance of shell: Redating of the To'aga site (Ofu Island, Manu'a) and a revised chronology for the Lapita to Polynesian Plainware transition in Tonga and Sāmoa. *PloS one* **14**, e0211990 (2019).
190. P. Skoglund *et al.*, Genomic insights into the peopling of the Southwest Pacific. *Nature* **538**, 510-513 (2016).
191. A. Anderson, G. Clark, The age of Lapita settlement in Fiji. *Archaeology in Oceania* **34**, 31-39 (1999).
192. D. V. Burley, W. R. Dickinson, Origin and significance of a founding settlement in Polynesia. *Proceedings of the National Academy of Sciences* **98**, 11829-11831 (2001).
193. D. J. Addison, A. Morrison, in *The Gotland Papers: Selected Papers from the VII International Conference on Easter Island and the Pacific: Migration, Identity, and Cultural Heritage*. (Gotland University Press Gotland, 2010), pp. 359-376.
194. F. J. Petchey, Radiocarbon determinations from the Mulifanua Lapita site, Upolu, western Samoa. *Radiocarbon* **43**, 63-68 (2001).
195. T. M. Rieth, A. E. Morrison, D. J. Addison, The temporal and spatial patterning of the initial settlement of Sāmoa. *The Journal of Island and Coastal Archaeology* **3**, 214-239 (2008).
196. J. M. Diamond, Express train to Polynesia. *Nature* **336**, 307-308 (1988).
197. S. Oppenheimer, M. Richards, Fast trains, slow boats, and the ancestry of the Polynesian islanders. *Science Progress* **84**, 157-181 (2001).
198. R. Green, Near and Remote Oceania: disestablishing" Melanesia" in culture history. *Man and a half: Essays in Pacific Anthropology and Ethnobotany in honour of Ralph Bulmer*, (1991).
199. R. Green, the initial identification of a people as Polynesian in race, language and culture. *Ethnicity and Culture. R. Auger, ME Glass, S. McEachern and Ph McCartney, eds*, 175-180 (1987).
200. D. J. Addison, E. Matisoo‐Smith, Rethinking Polynesians origins: A West‐Polynesia Triple‐I model. *Archaeology in Oceania* **45**, 1-12 (2010).
201. P. V. Kirch, R. C. Green, *Hawaiki, ancestral Polynesia: an essay in historical anthropology*. (Cambridge University Press, 2001).
202. R. C. Green, Revision of the Tongan sequence. *The Journal of the Polynesian Society* **81**, 79-86 (1972).
203. A. Pawley, Polynesian languages: A subgrouping based on shared innovations in morphology. *The Journal of the Polynesian Society* **75**, 39-64 (1966).
204. R. Walter, H. Buckley, C. Jacomb, E. Matisoo-Smith, Mass migration and the Polynesian settlement of New Zealand. *Journal of World Prehistory* **30**, 351-376 (2017).
205. M. Kayser, The human genetic history of Oceania: near and remote views of dispersal. *Current Biology* **20**, R194-R201 (2010).
206. A. Wollstein *et al.*, Demographic history of Oceania inferred from genome-wide data. *Current Biology* **20**, 1983-1992 (2010).
207. C. L. Hofman, A. J. Bright, R. R. Ramos, R. de Utuado, P. de Ciencias Sociales, Crossing the Caribbean Sea: Towards a holistic view of pre-colonial mobility and exchange. *Journal of Caribbean Archaeology* **3**, 1-18 (2010).

208. R. R. Ramos, *Rethinking puerto rican precolonial history*. (University of Alabama Press, 2010).
209. C. Gullick, in *Proceedings of the Eighth International Congress for the Study of the Pre-Columbian Cultures of the Lesser Antilles. Arizona State University Anthropological Research Papers*. (1980), vol. 22, pp. 464-472.
210. C. Scheib *et al.*, Ancient human parallel lineages within North America contributed to a coastal expansion. *Science* **360**, 1024-1027 (2018).
211. C. Posth *et al.*, Reconstructing the Deep Population History of Central and South America. *Cell* **175**, 1185-1197 e1122 (2018).
212. A. Bergström *et al.*, Insights into human genetic variation and population history from 929 diverse genomes. *Science* **367**, (2020).
213. R. T. Callaghan, in *The Oxford Handbook of Caribbean Archaeology*. (2013).
214. R. R. Ramos, in *The Oxford handbook of Caribbean archaeology*. (Oxford University Press Oxford, 2013), pp. 155-170.
215. R. Callaghan, S. J. Schwabe, in *Proceedings of the XVIII congress of the international association for Caribbean archaeology. IACA Martinique, St. Georges, Grenada*. (2001), vol. 1999, pp. 231-242.
216. I. Rouse, R. Alegría, *Excavations at Maria de la Cruz Cave and Hacienda Grande Village Site, Loiza, Puerto Rico*. (Department of Anthropology, Yale University, 1990).
217. L. A. Newsom, E. S. Wing, *On land and sea: Native American uses of biological resources in the West Indies*. (University of Alabama Press, 2004).
218. M. Roksandic *et al.*, The earliest burial from the Caribbean coast of Nicaragua: the Angi shell-matrix site. *antiquity* **92**, 1478-1491 (2018).
219. M. Roksandic, I. Roksandic, The role of the Nicaraguan Rise in the early peopling of the Greater Antilles. *Cuban archaeology in the Caribbean*, 7-16 (2016).
220. M. H. Crawford, *The origins of Native Americans: evidence from anthropological genetics*. (Cambridge University Press, 2001).
221. D. H. O'Rourke, Blood groups, immunoglobulins, and genetic variation. *Handbook of North American Indians* **3**, 762-776 (2006).
222. H. Yu *et al.*, Paleolithic to Bronze Age Siberians Reveal Connections with First Americans and across Eurasia. *Cell*, (2020).
223. S. A. Elias, S. K. Short, R. L. Phillips, Paleoecology of late-glacial peats from the Bering Land Bridge, Chukchi Sea shelf region, northwestern Alaska. *Quaternary Research* **38**, 371-378 (1992).
224. D. M. Hopkins, Cenozoic history of the Bering land bridge. *Science* **129**, 1519-1528 (1959).
225. M. W. Pedersen *et al.*, Postglacial viability and colonization in North America's ice-free corridor. *Nature* **537**, 45-49 (2016).
226. J. V. Moreno-Mayar *et al.*, Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature* **553**, 203-207 (2018).
227. M. B. Collins, M. Kay, *Clovis blade technology: a comparative study of the Keven Davis cache, Texas*. (University of Texas Press, 2002).
228. J. H. Greenberg *et al.*, The settlement of the Americas: A comparison of the linguistic, dental, and genetic evidence [and comments and reply]. *Current Anthropology* **27**, 477-497 (1986).
229. T. D. Dillehay *et al.*, Monte Verde: seaweed, food, medicine, and the peopling of South America. *science* **320**, 784-786 (2008).
230. C. F. Ardelean *et al.*, Evidence of human occupation in Mexico around the Last Glacial Maximum. *Nature* **584**, 87-92 (2020).
231. L. Becerra-Valdivia, T. Higham, The timing and effect of the earliest human arrivals in North America. *Nature*, 1-5 (2020).

232. T. J. Braje, T. D. Dillehay, J. M. Erlandson, R. G. Klein, T. C. Rick, Finding the first Americans. *Science* **358**, 592-594 (2017).
233. J. M. Erlandson, T. J. Braje, From Asia to the Americas by boat? Paleogeography, paleoecology, and stemmed points of the northwest Pacific. *Quaternary International* **239**, 28-37 (2011).
234. B. A. Potter *et al.*, Current evidence allows multiple models for the peopling of the Americas. *Science Advances* **4**, eaat5473 (2018).
235. W. A. Ritchie, *The archaeology of Martha's vineyard: a framework for the prehistory of southern New England*. (Natural History Press, Garden City, New York, 1969).
236. M. A. Glassow, Changes in subsistence on marine resources through 7,000 years of prehistory on Santa Cruz Island. *Archaeology on the northern Channel Islands: Studies of subsistence, economics, and social organization. Coyote Press Archives of California Prehistory* **34**, 75-94 (1993).
237. P. C. Martz, in *Proceedings of the Sixth California Islands Symposium (Santa Barbara Museum of Natural History, 2008)*. (2003), pp. 65-82.
238. J. S. Friedlaender *et al.*, The genetic structure of Pacific Islanders. *PLoS Genet* **4**, e19 (2008).
239. P. J. Sheppard, Lapita lithics: trade/exchange and technology. A view from the Reefs/Santa Cruz. *Archaeology in Oceania* **28**, 121-137 (1993).
240. M. Lipson *et al.*, Population turnover in Remote Oceania shortly after initial settlement. *Current Biology* **28**, 1157-1165. e1157 (2018).
241. A. François, M. Franjieh, S. Lacrampe, S. Schnell. (2015).
242. S. Bedford *et al.*, Ancient DNA and its contribution to understanding the human history of the Pacific Islands. *Archaeology in Oceania* **53**, (2018).
243. J. Allen, Notions of the Pleistocene in greater Australia. *A Community of Culture: the People and Prehistory of the Pacific*, 139-151 (1993).
244. G. R. Summerhayes, J. Allen, The transport of Mopir obsidian to late Pleistocene New Ireland. *Archaeology in Oceania* **28**, 144-148 (1993).
245. I. J. McNiven *et al.*, New direction in human colonisation of the Pacific: Lapita settlement of south coast New Guinea. *Australian Archaeology* **72**, 1-6 (2011).
246. P. V. Kirch, *Lapita and Its Transformations in Near Oceania: Archaeological Investigations in the Mussau Islands, Papua New Guinea, 1985-88*. (Archaeological Research Facility, University of California, 2001).
247. D. Lepofsky, The environmental context of Lapita settlement locations. *Archaeology of the Lapita Cultural Complex: A critical review*, 33-47 (1988).
248. A. Bergström *et al.*, A Neolithic expansion, but strong genetic structure, in the independent history of New Guinea. *Science* **357**, 1160-1163 (2017).
249. M. T. Carson, H. Kurashina, Re-envisioning long-distance Oceanic migration: Early dates in the Mariana Islands. *World Archaeology* **44**, 409-435 (2012).
250. P. Bellwood, in *Lost maritime cultures: China and the Pacific*. (Bishop Museum Press, 2007).
251. H.-c. Hung *et al.*, The first settlement of Remote Oceania: the Philippines to the Marianas. *Antiquity* **85**, 909-926 (2011).
252. J. S. Athens, J. V. Ward, Holocene vegetation, savanna origins and human settlement of Guam. *Records-Australian Museum*, 15-30 (2004).
253. R. Blust, Three recurrent changes in Oceanic languages. *Pacific island languages: Essays in honour of GB Milner*, 7-28 (1990).
254. M. D. Ross, Central Papuan culture history: some lexical evidence. *Austronesian terminologies: Continuity and change*, 389-479 (1994).
255. T. E. Dutton, *Reconstructing Proto Koiarian: The history of a Papuan language family*. (Pacific Linguistics, Research School of Pacific and Asian Studies, The …, 2010).

256. T. E. Dutton, *The peopling of Central Papua: Some preliminary observations*. (Pacific Linguistics, Research School of Pacific and Asian Studies, The …, 1969).
257. N. Oram, The history of the Motu-speaking and Koitabu-speaking peoples according to their own traditions. *Oral tradition in Melanesia*, 207-229 (1981).
258. P. Swadling, A review of the traditional and archaeological evidence for early Motu, Koita and Koiari settlement along the central south Papuan coast. *Oral History* **5**, 37-43 (1977).
259. P. Swadling, The settlement history of the Motu and Koita speaking people of the Central Province, Papua New Guinea. *Oral Tradition in Melanesia*, 240-251 (1981).
260. D. Gaffney *et al.*, Archaeological investigations into the origins of Bel trading groups around the Madang coast, northeast New Guinea. *The Journal of Island and Coastal Archaeology* **13**, 501-530 (2018).
261. J. Allen, Revisiting Papuan ceramic sequence changes: another look at old data. *Artefact: the Journal of the Archaeological and Anthropological Society of Victoria, The*, 4 (2010).
262. B. Shaw *et al.*, 2500-year cultural sequence in the Massim region of eastern Papua New Guinea reflects adaptive strategies to small islands and changing climate regimes since Lapita settlement. *The Holocene*, 0959683620908641 (2020).
263. N. Sutton, G. Summerhayes, A. Ford, in *Proceedings of the Prehistoric Society*. (Cambridge University Press, 2015), vol. 81, pp. 343-359.
264. S. Bulmer, Prehistoric settlement patterns and pottery in the Port Moresby area. *Journal of the Papua and New Guinea Society* **5**, 29-91 (1971).
265. B. Shaw, H. Buckley, G. Summerhayes, C. Stirling, M. Reid, Prehistoric migration at Nebira, South Coast of Papua New Guinea: New insights into interaction using isotope and trace element concentration analyses. *Journal of Anthropological Archaeology* **30**, 344-358 (2011).
266. A. G. Ioannidis *et al.*, Native American gene flow into Polynesia predating Easter Island settlement. *Nature* **583**, 572-577 (2020).
267. C. Ballard, P. Brown, R. M. Bourke, T. Harwood, *The sweet potato in Oceania: a reappraisal*. (2005).
268. S. M. Fitzpatrick, R. Callaghan, Examining dispersal mechanisms for the translocation of chicken (Gallus gallus) from Polynesia to South America. *Journal of Archaeological Science* **36**, 214-223 (2009).
269. J. Hather, P. V. Kirch, Prehistoric sweet potato (Ipomoea batatas) from Mangaia Island, central Polynesia. *Antiquity* **65**, 887-893 (1991).
270. A. Sharp, *Ancient voyagers in the Pacific*. (Penguin Books, 1957).
271. D. R. Drake, C. P. Mulder, D. R. Towns, C. H. Daugherty, The biology of insularity: an introduction. *Journal of biogeography* **29**, 563-569 (2002).
272. J. DaVanzo, Microeconomic approaches to studying migration decisions. (1980).
273. P. V. Kirch, D. E. Yen, Tikopia. The prehistory and ecology of a Polynesian outlier. *Bernice P. Bishop Museum Bulletin Honolulu*, 1-396 (1982).
274. E. S. Wing, E. Reitz, Prehistoric fishing communities of the Caribbean. *Journal of New World Archaeology* **5**, 13-32 (1982).
275. W. Keegan, University of California, Los Angeles--Anthropology, (1985).
276. J. Diamond, The Riddle of the Ancient Mariners. *The Sciences* **25**, 42-46 (1985).
277. J. M. Diamond, W. F. Keegan, Astropology: Supertramps at sea. *Nature* **311**, 704-705 (1984).

10. ACKNOWLEDGEMENTS

I have many people to thank for their support in this long journey and would like to express my gratitude.

First and foremost, to my wonderful partner David. Without you none of this would have been possible. You have served as a discussion partner, practice audience, proof reader and councillor. Not only have you given me strength throughout the last four years, but also are a wonderful father to our beautiful children. Enduring my never-ending series of deadlines, this is as much your accomplishment as it is mine.

To my children Paul and Henri, ihr seid mein größtes Glück. Danke für eure Geduld mit mir. Ich hoffe dass ihr auf mich so stolz sein könnt wie ich es auf euch bin. Ich bin dankbar eure Mama sein zu dürfen.

To my supervisors Johannes and Cosimo. You both have given me a great opportunity and never failed to support me. I believe it is rare to find supervisors who place so much trust in their PhD student. Your trust has allowed me to (almost) fearlessly leave my comfort zone, present our research all over the world and through this allowed me to grow as a person and a researcher.

To Maria Spyrou. With your patience and professionalism you have taught me everything I know in the lab.

To Michal Feldman. Your kind advice, both in professional and private matters, has always been extremely helpful. Your support and friendship is invaluable to me.

To Tiago Ferraz da Silva. With your open ear and open heart you have been a wonderful companion. Thank you for gifting me with your friendship and a new passion.

To Betsy, thank you for your invaluable support as writing buddy and the inspiring discussions about ethics. You are a role model and an inspiration, I'm happy to call you part of my modern family.

To the entire wet lab team of the MPI, but especially Rita Radzeviciute for taking the processing off my hands in the last two years, enabling me to learn and grow in the dry lab and to Antje Wissgott for her flexibility in squeezing in the "high priority" samples.

Last, but not least the great DAG family, especially my fellow PhD-students who were always very supportive. In particular to Thiseas for your patience with my never-ending questions.

11. APPENDIX

11.1 Supplementary Material for Manuscript A

science.sciencemag.org/cgi/content/full/science.aba8697/DC1

Supplementary Materials for

Genomic insights into the early peopling of the Caribbean

Kathrin Nägele*†, Cosimo Posth†, Miren Iraeta Orbegozo, Yadira Chinique de Armas, Silvia Teresita Hernández Godoy, Ulises M. González Herrera, Maria A. Nieves-Colón, Marcela Sandoval-Velasco, Dorothea Mylopotamitaki, Rita Radzeviciute, Jason Laffoon, William J. Pestle, Jazmin Ramos-Madrigal, Thiseas C. Lamnidis, William C. Schaffer, Robert S. Carr, Jane S. Day, Carlos Arredondo Antúnez, Armando Rangel Rivero, Antonio J. Martínez-Fuentes, Edwin Crespo-Torres, Ivan Roksandic, Anne C. Stone, Carles Lalueza-Fox, Menno Hoogland, Mirjana Roksandic, Corinne L. Hofman, Johannes Krause*, Hannes Schroeder*

*Corresponding author. Email: hschroeder@sund.ku.dk (H.S.); naegele@shh.mpg.de (K.N.); krause@shh.mpg.de (J.K.)
†These authors contributed equally to this work.

Published 4 June 2020 on *Science* First Release
DOI: 10.1126/science.aba8697

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at science.sciencemag.org/cgi/content/full/science.aba8697/DC1)

Table of Contents

Supplementary Text

Ethics statement

The human remains analyzed in this study are considered the cultural heritage of the originating countries. Permissions for the export and destructive analysis was obtained through the *Consejo Nacional de Patrimonio Cultural*, as part of the projects "*Poblamiento temprano de la cuenca hidrográfica de río Canímar: Estudio arqueológico, antropológico y paisajístico*" and "*Arqueología de prácticas mortuorias en sociedades aborígenes de bajos niveles productivos de Cuba*" under the principal investigation of Silvia Hernández Godoy and Ulises M. González Herrera, respectively. Samples are covered by the permits PEA 11/17, PEA 9/19, PEA17/17, PEA9/18, PEA17/16. For Puerto Rico, permissions were granted to William J. Pestle and Antonio Curet from the *Consejo para la Protección del Patrimonio Arqueológico Terrestre de Puerto Rico*, the *Autoridad de Carreteras y Transportación* and the *Secretaría de Cultura y Turismo del Municipio Autónomo de Ponce*. They were obtained as part of the project NSF BCF-0612727 and are extended to this project. For The Bahamas, permissions were granted to Robert S. Carr through *The National Museum of the Bahamas, the Antiquities, Monuments and Museum Corporation*. For Guadeloupe, permissions were granted to Hannes Schroeder by the *Direction des affaires culturelles de Guadeloupe, Service régional de l'archéologie*. For St.Lucia, permission was granted to Hannes Schroeder by the *Saint Lucia Archaeological and Historical Society*.

Terminology

The individuals analyzed in this study were excavated from different islands, regions, archaeological horizons, contexts and ages. This poses certain difficulties when referring to them in a more generalized manner. The Caribbean is a place of culturally rich and contextually diverse pre-contact populations, which are only poorly reflected in the terms previously used to describe them (*6*). In absence of written records, we cannot know how people in ancient times self-identified or how they referred to themselves and each other, and using established terminology that was primarily developed to describe archaeological horizons or material culture to refer to people in the past is highly problematic as it implies a connection between material culture, identity and biology that does not necessarily exist. The communities encountered by the colonizers during the first voyages to the Caribbean, were eventually, and perhaps erroneously, referred to as "Taíno", assuming the shared Arawakan language of the encountered indigenous groups reflected a nowadays not verifiable cultural unity (*26*). Indeed,

the diversity in the material culture uncovered by archaeologists in the region suggests otherwise (*1*). We therefore avoid using the term "Taíno", although we acknowledge its use by the neo-Taíno community as a form of social affirmation.

The terms "Ciboney" and "Guanahatabey" to refer to the early settlers of the Western Caribbean are similarly problematic (*1*, *27*). These communities are represented in the archaeological record by a variety of lithic material cultures, some resembling more those of North American (*28*), others those of Central and South American hunter-gatherers (*6*, *29*) and some even featuring ceramics (*20*). First occurrences of those sites are almost contemporaneous on the both extreme ends in the Caribbean, in Cuba (*30*) and Trinidad (*31*). The differences in material culture and the occurrence in two far apart regions suggest a cultural diversity and our genetic analysis shows a diversity also on the genetic level even within only one island. Regardless, sites and individuals excavated from such contexts were traditionally referred to as "Archaic", reflecting the timing of first occurrences in the Caribbean during the Archaic Age in the Americas. When talking about those groups in a genetic context, the term "Archaic" appears especially inadequate. Not only does it imply a less complex society, discriminating against people living this lifestyle today, but it is also confusing, since in ancient genomics the term is used for archaic hominins, such as Neandertals or Denisovans. Other terms previously used reflected the subsistence strategy as fisher-hunter-gatherers or proto-agriculturalists, or, seeing they occupied the Caribbean before the arrival of groups using ornate ceramic pottery, as pre-ceramic. Those terms disregard findings showing a certain form of horticulture, suggesting the subsistence was complemented by plant cultivation (*32*) and ignores the sites with abundance of ceramics (*20*). The latter implies a succession of Archaic Age groups by Ceramic Age groups, while we see contemporaneous occurrence of "Archaic" and "Ceramic" contexts. We acknowledge our inability to reflect the cultural diversity in one term and want to avoid cultural terms for genetic groups. However, we have to refer to the various groups on a broader scale. The debate about the terminologies applied to said groups has been ongoing for decades, and we will be unable to solve this debate in a study about the genetics of the pre-colonial Caribbean. Only together with other scholars knowledgeable of the region and considering all available results produced by different disciplines through a variety of methods, a sensible, non-colonial and unambiguous terminology can be established. Reluctantly, we therefore revert to using a variation of the established terms and refer to the individuals and genetic groups from the Lesser Antilles, Puerto Rico, Bahamas and the individuals excavated on Cuba from Ceramic contexts, exhibiting the same genetic signal with more affinity to north-eastern South American populations, as "Ceramic-related". We will refer to the other groups

analyzed, genetically distinct from the individuals from the “Ceramic-related” group, from Cuba, dating to 3200-700 cal. BP as “Archaic-related”.

Materials and Methods

Sampling

All samples were processed in dedicated ancient DNA laboratories at the Max Planck Institute for the Science of Human History (MPI) in Jena, Germany, the Globe Institute, University of Copenhagen (UCPH), Copenhagen, Denmark, and Arizona State University (ASU), Tempe, AZ, USA. Petrous bones were sampled by isolating the densest part of the cochlea as described in (*33*). Before sampling, the teeth were cleaned by removing any dirt from the surface and wiping them with a cloth dipped in 1% sodium hypochlorite solution. The teeth were then sampled by separating the root using a Dremel diamond-coated cutting disc as described in (*34*). The roots were then UV irradiated for 5 mins on each side in a UVP CL-1000 Ultraviolet Crosslinker and crushed into a coarse powder using a pestle and mortar.

Radiocarbon dating

Radiocarbon dating was carried out at the Curt-Engelhorn-Zentrum Archäometrie gGmbH in Mannheim, Germany. Collagen from bone and dentin was extracted using a modified Longin method (*35*) and long molecules removed with ultrafiltration before freeze-drying the product (*36*). Where no bone material was available, radiocarbon dating was carried out on tooth enamel following hydrochloric acid pretreatment (*37*). After the catalytic reduction to graphite the ^{14}C content was measured with an AMS-System type MICADAS. The isotopic ratios of $^{14}C/^{12}C$ and $^{13}C/^{12}C$ of samples, standards (Oxalic acid II) and controls were measured simultaneously. The resulting ^{14}C dates were normed with $\delta^{13}C$=-25‰ (*38*) and calibrated using the software SwissCal 1.0 (L.Wacker, ETH-Zürich) and the INTCAL13 calibration curve (*39*). The radiocarbon dates and quality collagen indicators (collagen yields, C/N ratios, %C and %N) are reported in Table S2.

DNA Extraction

DNA extraction was carried out following established protocols (*40*). Negative and positive controls were included. To release DNA from 50-100 mg of bone powder, a solution of 900 µl EDTA, 75 µl H2O and 25 µl Proteinase K was added. In a rotator, samples were digested for at least 16 h at 37°C, followed by an additional hour at 56°C (*41*). The suspension was then centrifuged and transferred into a binding buffer as previously described (*40*). To bind DNA, silica columns for high volumes (High Pure Viral Nucleic Acid Large Volume Kit; Roche) were used. After two washing steps using the manufacturer's wash buffer, DNA was eluted in

TET (10 mM Tris, 1 mM EDTA and 0.05% Tween) in two steps for a final volume of 100 µl. The samples processed in Copenhagen were extracted following a modified silica-in-solution protocol (*22*, *34*) using between 50-100 mg of starting material and eluted in 64 µl of EB. Samples processed at ASU were extracted using c. 100 mg of dentine powder following Dabney et al. (*40*) as described in Nieves-Colón et al. (*42*).

Library preparation

Double stranded DNA libraries were built from 25 µl of DNA extract in the presence of uracil DNA glycosylase (non-UDG libraries), following a double-stranded 'UDG-half' library preparation with a protocol using the UDG enzyme to reduce, but not eliminate, the amount of deamination-induced damage towards the ends of aDNA fragments (*43*). Negative and positive controls were carried alongside each experiment. Libraries were quantified using the IS7 and IS8 primers (*44*) 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 (*45*) in 1-4 parallel 100 µl reactions using PfuTurbo DNA Polymerase (Agilent). 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 (*44*) 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 Illumina platforms. The libraries for the four individuals from Guadeloupe were produced in Copenhagen without UDG treatment. In this case, a double-stranded library was generated from 32 µl of extract following the BEST protocol, using adapters compatible with Illumina sequencing (*46*). Quantitative real-time PCR (qPCR) was performed using SYBR green and Amplitaq Gold (Thermo Fisher) in order to estimate the required number of cycles for library index amplification. Each library was then amplified and indexed using a dual indexing protocol (*45*) and purified using SPRI beads as described in Rohland and Reich (*47*). The amplified libraries were quantified using the High-Sensitivity DNA Assay on an Agilent 2200 TapeStation (Agilent Technologies, Palo Alto, CA, USA). The libraries were then pooled in equimolar amounts and sequenced at the Danish National High-throughput DNA Sequencing Centre using an Illumina 2500 run in SR80 mode.

Targeted enrichment and high-throughput sequencing

Both UDG-half- and non-UDG-treated 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). Mitochondrial DNA capture (*48*) was performed on screened libraries which, after shotgun sequencing, showed the presence of aDNA, highlighted by the typical CtoT and GtoA substitution pattern towards 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 in (*49*). The enriched DNA product was sequenced on an Illumina HiSeq 4000 instrument with 75 single-end-run cycles or 50 pair-end-run cycles using the manufacturer's protocol. The output was de-multiplexed using bcl2fastq version 2.17.1.14 (Illumina conversion Software) and dnaclust version 3.0.0 (*50*).

Genomic data processing

Pre-processing of the sequenced reads was performed using EAGER version 1.92.55 (*51*). The resulting reads were clipped to remove residual adaptor sequences using *Clip&Merge* (*51*) and *AdapterRemoval* version 2 (*52*). Clipped sequences were then mapped against the human reference genome hg19 using the Burrows–Wheeler Aligner (BWA) version 0.7.12 (*53*) disabling seeding (-l 16500, –n 0.01). Duplicates were removed with DeDup version 0.12.2 (*51*), which removes reads with identical start and end coordinates. Additionally, a mapping quality filter of 30 was applied using SAMtools version 1.3 (*54*). Reads obtained from libraries without UDG treatment were trimmed for 10 base pairs on both ends according to the observed damage patterns and 2 base pairs for UDG half treated libraries to reduce the impact of deamination induced misincorporations during genotyping. Different sequencing runs and libraries from the same individuals were merged, duplicates removed and sorted again using SAMtools (*54*). Trimmed and untrimmed reads were genotyped separately using pileupCaller v. 8.6.5 (https://github.com/stschiff/sequenceTools/tree/master/srcpileupCaller), a tool that randomly draws one allele at each of the 1,240 K-targeted SNPs covered at least once. We combined the genotypes keeping all transversions from the untrimmed genotypes and transitions only from the trimmed genotypes to eliminate problematic, damage-related transitions on the ends. A second genotype was produced drawing only transversions to eliminate any residual damage for the non-UDG libraries from the individuals from Anse à la Gourde. The generated pseudo-haploid calls for all ancient individuals (Table S1) were merged

to a pulldown of the 1,240 K SNPs from the Simons Genome Diversity Project (*55*)and previously published ancient American individuals (*7*, *14–16*, *56–63*), as well as a published individual from the Bahamas (*7*). For some symmetry tests and the PCA shown in Fig. S2C, the resulting dataset was merged to previously published sets of present-day American individuals (*30, 31*).

Quality control

The typical features of ancient DNA were inspected with *DamageProfiler* version 0.3.1 (http://bintray.com/apeltzer/EAGER/DamageProfiler) (*51*) (Table S3). Sex determination was performed by comparing the coverage on the targeted X-chromosome SNPs (~50 K positions within the 1,240 K capture) normalized by the coverage on the targeted autosomal SNPs to the coverage on the Y-chromosome SNPs (~30 K), again normalized by the coverage on the autosomal SNPs (*64*) (Table S4) and individuals where sex could not be determined were excluded from the analysis (CAO004). For male individuals, ANGSD version 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 (*65*). This provides an estimate of nuclear contamination in males that are expected to have only one allele at each site. For all male samples that exhibited X-chromosome contamination levels below 7% with at least 100 X-chromosome SNPs covered twice, all reads were retained for further analyses (Table S4). For both male and female individuals, mtDNA-captured data were used to jointly reconstruct the mtDNA consensus sequence and estimate contamination levels with schmutzi (*66*) (Table S4). For specimens where a relatively low proportion of mtDNA molecules compared with nuclear DNA was observed (Table S3), mtDNA contamination estimates are used as reliable predictors for nuclear contamination (*67*, *68*). The software ADMIXTURE version 1.3.0 (*69*) was used in a supervised mode to allow for genetic clustering with preset clusters for African (Mbuti.DG, Yoruba.DG), European (French.DG, Sardinian.DG, English.DG) , Asian (Han.DG, Ami.DG, Atayal.DG), Oceanian (Papuan.DG), Siberian (Itelmen.DG, Chukchi.DG, Ulchi.DG) and American (Karitiana.DG, Surui.DG, Mixe.DG, Pima.DG) individuals, resulting in a mosaic of those for the newly produced individuals. Individuals with substantial proportions of African, European, Asian or Oceanian components were excluded from further analysis (CAO004, PCA009, PDI003) (Fig. S5A) A principle component analysis was computed with worldwide present-day populations from the Simons Genome Diversity Project (*55*) (Fig. S5B). Individuals visibly shifted from the American cluster, suggesting non-

American admixture or contamination, were removed from further analyses (CAO004, PDI003, PCA010).

Principal components analysis

Principal components analyses were carried out using *smartpca* version 13050 (*70*) with worldwide present-day populations from the Simons Genome Diversity Project (*55*) (Fig. S5B) and a regional one calculated on the unadmixed American individuals represented in the dataset (Fig. S2A). Ancient individuals as well as the present-day individuals with lower SNP overlap were projected onto the calculated components using the options 'lsqproject: YES', 'shrinkmode: YES' and 'numoutlieriter: 0'. Individuals with less than 20,000 SNPs were not projected. Three individuals (PCA001, ALG003, LOI001) from the sites of Punta Candelero and Los Indios in Puerto Rico and Anse à la Gourde on Guadeloupe, who derive from a Ceramic context, cluster with individuals from Cuba 3000-800 cal. BP. However, an f_4-statistic of the form *f_4(Mbuti, PCA001/ALG003/LOI001; Ceramic, Archaic-related)* reveals that, contrary to the PCA, these individuals have greater affinity to other Ceramic-related individuals than Archaic-related individuals (Table S5).

F-statistics

To identify the differences on an individual basis and to identify a sensible grouping in the subsequent analysis we used *qp3Pop* version 5.0 (*70*) and computed an f_3-outgroup statistics comparing all individuals to each other with Mbuti.DG serving as an outgroup. We used *qpDstat* version 5.0 to run f_4-statistics of the form *f4*(Mbuti, Piapoco.DG Individual 1 site X; Individual 2 site X). This test expects values close to zero for Individual 1 and Individual 2 if they are more related to each other then to Piapoco. This is the case for all Archaic-related and Ceramic-related individuals in position of Individual 1 and Individual 2, respectively. Values indicate much higher allele sharing between Piapoco and individuals from Ceramic-related contexts when tested with Archaic-related contexts. A similar test of the form f_4(Mbuti.DG, Piapoco.DG, grouped site X, grouped site Y) was performed to support a grouping of all sites with the same archaeological context for subsequent analyses. To test the affinities of the different sites to the already published individual from the The Bahamas against ancient Californian Channel Island individuals, we computed an f_4-statistic of the form f_4(Mbuti.DG, Caribbean sites; Early_San_Nicholas.SG, Bahamas_Taino.SG) (Fig. 2B, Table S5). To identify the driving component for the differences in sites from Cuba 3000-800 cal. BP versus

sites with a Ceramic-related context we produced a 2-dimensional plot using the following tests: f_4(Mbuti.DG, Caribbean sites; USR1, Piapoco.DG), and f_4(Mbuti.DG, Caribbean sites; USR1, Mixe.DG), using the individual excavated from the Upward Sun River (USR1), identified as an representing a group basal to all present-day Native Americans, to normalize the shared drift in the Americas (Fig. 2C, Table S5). To elucidate genetic similarities with a set of present-day groups from the Americas, which were represented only by individuals with minimal non-Native American ancestry (*13*), we computed an f_3-outgroup statistic comparing the two archaeologically different groups with the unadmixed American groups and Mbuti.DG serving as an outgroup (Fig. S4). For the four individuals from the site of Anse à la Gourde where libraries were produced in the absence of USER enzyme we restricted the genotype to the transversions only. To understand the differential affinities of anciewith both present-day and ancient groups in the Americas we tested f_4(Mbuti, Test; Archaic-related, Ceramic-related), using in "Test" all published, uncontaminated and well-covered ancient individuals and all present-day groups from the Americas present in the SGDP dataset (Fig. S3, Table S5).

Ancestry modelling (*qpWave*)

We used *qpWave* version 410 (*13*) to test whether the two groups from a different context are consistent with deriving from two distinct ancestral sources relative to a set of reference groups (Mbuti.DG, Onge.DG, Papuan.DG, Han.DG, Russia_MA1_HG.SG, USA_Ancient_Beringian.SG; USA_Anzick.SG, Mixe.DG, Mexico_Zapotec.DG, Belize_MayahakCabPek_9300, Karitiana.DG, Piapoco.DG.). Reference Groups were chosen to represent branches that are considered basal to the populations under investigation keeping the amount of populations at minimum as suggested in the software documentation. After establishing that the sites grouped by archaeological context could be clearly distinguished with the given set of reference groups (Table S6) we used *qpAdm* version 5.0 (*18*) to model all sites and the individuals in each site covered by more than 50,000 SNPs as a two-way admixture between the two groups, while excluding the tested site from the analysis (Table S7).

Admixture graph modeling (*qpGraph*)

To model the ancestry of the ancient individuals in a tree-like form allowing for admixture events we used *qpGraph* version 5.0 (*18*), which compares the fitted tree model to the observed allele frequency correlation inferred from f_2, f_3 and f_4 statistics. The groups chosen to be grafted on the tree had to fulfill the following criteria. 1) The library preparation had to include UDG-

treatment to reduce the number of residual deaminated sites in the analysis. 2) We only fitted individuals with a coverage of at least 100,000 SNPs and grouped them based on pairwise *qpWave* analysis if they were consistent with deriving from one ancestral source (Table S6), with the exception of CIP009 and GUY002 who were fitted individually due to their older date and distinct behavior in various f_4-statistics and *qpWave* analyses. After removing the transitions on CpG sites we used the standard settings of *qpGraph* with the exception of "outpop: NULL" and "allsnps: YES". Starting from an established scaffold (*15*) we simplified the basal portion of the graph (using only Mbuti.DG and Han.DG as non-American populations) and explored the placement of the subsequent groups testing all placements along the scaffold both as a direct branch and as two-way admixtures. When the position of the group or individual was modeled as one- and two-ways with an equally good fit we preferred the one-way placement in the tree. When multiple trees resulted in the same Z-score we chose the one with less outliers and in cases where two trees had both the same Z-score and number of outliers, we chose the tree with the highest likelihood. Due to their different age, we were unable to model all groups on a single tree with an acceptable fit and, therefore, decided to model Archaic- and Ceramic-related individuals using different scaffolds (Fig. S6 and Fig. S7). For the Archaic-related individuals we built a scaffold using ancient Native American genomes (*14–16*), while for the younger Ceramic-related individuals we used present-day Native American groups (*55*).

Site descriptions

Preacher's Cave, The Bahamas

The archaeological site of Preacher's cave is located on the northern part of the Bahamian Island of Eleuthera (25°33'25"N 76°41'41"W). Archaeological surveys revealed multiple phases of occupation dating to both pre- and post-colonial periods. Radiocarbon dates of the site indicate a prehistoric occupation from 1640 cal. BP through European contact (*71*, *72*). The inhabitants of the Bahamas in pre-colonial times were dubbed "Lucayans" (*71*) and the analysis of ceramics placed them in an interaction sphere with the inhabitants of both Cuba and Hispaniola (*73*). Out of a total of six primary burials three were well preserved and disarticulated teeth were found next to the burials, which are well described (*72*) and a 12.4× genome was previously obtained from one of these individuals (*7*). New direct dates for the individuals used in this study were produced by the Klaus-Tschira Archaeometriezentrum, Mannheim, Germany (Table S1 and S2) and range from 880 - 820 cal. BP, placing them in the range of the site.

Anse à la Gourde, Guadeloupe

The multi-component settlement site of Anse à la Gourde, Guadeloupe, is located on the eastern shore of Grande-Terre on the Pointe de Chateaux peninsula (16°15"0' N and 61°13"12' W). The site was excavated by a team of researchers from Leiden University and the Archaeological Service of the Direction Régionale des Affaires Culturelles of Guadeloupe (DRAC) in the 1990s (*74*). The site comprises a habitation area and a plaza surrounded by thick midden deposits reflecting the disposal of garbage over many centuries (1550-550 cal. BP). An initial Early Ceramic Age occupation (dated to 1550-1250 cal. BP) is documented by the presence of Cedrosan Saladoid ceramics but much of this component of the site was lost due to the retreat of the coastline (*75*). The Ceramic Age occupation partly overlays the previous Saladoid settlement and suggests that the site was permanently reoccupied around 1050 cal. BP and remained so until roughly 550 cal. BP. This occupation phase is characterized by a number of round and oval houses of various sizes with human burials under the floors and outside the structures (*76*). The ceramic assemblages include Mamoran/Troumassan Troumassoid to early and late Suazan Troumassoid materials and overall display a high diversity of influences and also include stylistic traits seen in Morne Cybèle and Morne sites on La Désirade and in Cayo sites in the Windward Islands (*25*). Distribution patterns of non-perishable objects and materials suggest that local and micro-regional spheres of interaction were created through

monopolizing and manipulating the manufacture and/or exchange of goods and marriage partners (*25*, *77*, *78*) possibly out of the need to establish elaborate alliance networks among neighbors in order to form larger local socio-political units. A minimum of 99 individuals has been reported from approximately 86 burials, as many of the graves contain the remains of more than one individual. The burial population is composed primarily of adult individuals with relatively few juveniles (*74*). The burials at Anse à la Gourde generally occur in clusters of three to ten burials, are closely associated with house structures, and appear to be exclusively associated with the late Troumassoid occupation of the site (950-600 cal. BP). A wide variety of mortuary practices have been identified including both primary and secondary burials, and single and composite burials (*74*). Detailed analysis of taphonomic processes, anatomical positioning, and burial contexts indicate that many of the interred may have been wrapped (possibly in a hammock) prior to interment and in some cases there is also evidence for desiccation of the corpse and for post-burial manipulation of the corpse in an open grave (*74*, *76*). Newly produced direct radiocarbon dates from the skeletons used in this study (Table S2) were generated by the Klaus-Tschira Archaeometriezentrum, Mannheim, Germany and ranges from 620-1020 cal. BP, placing them within the range of the site.

Lavoutte, St. Lucia

The Lavoutte site is located along the shore of a bay (Cas-en-Bas) on the northeastern coast of the island of St. Lucia in the Lesser Antilles. (14°5'29.3" N and 60°55'32.5" W). The midden area of the site has been dated to roughly 1150-500 cal. BP (*79*). Portions of a large, distinct, ceramic figurine of a seated female figure, known as the 'Lavoutte statue', was discovered at this site in addition to a number of other unique, highly decorated figurines and a guaíza-like (carved face) artifact. The discovery of these objects and materials, believed to be of Greater Antillean origins, in addition to the size and location of the site led to the interpretation that the settlement represented a Carib ceremonial center (*79*). The presence of such ceremonial paraphernalia and Taíno-derived iconography in the Lesser Antilles may also indicate attempts by local leaders to acquire and project power and influence (*80*, *81*). Large-scale rescue excavations were carried out in 2009 and 2010 by an international team from Leiden University, the University of Florida, and the St. Lucia Archaeological and Historical Society (*82*). This most recent fieldwork focused on the portions of the site that were most vulnerable to impending damage from human and natural processes. In total, 48 burials containing 53 individuals were recovered within a relatively small portion of the site (*83*). The vast majority

of the burials date to the last phase of the Ceramic-related Age (800-450 cal. BP). Burial practices were variable and include primary, secondary, and composite internments, with most of the primary burials in a flexed or semi-flexed position. Although grave goods were uncommon, they included flaked stone, shell ornaments, and bone artifacts (*83*). Newly produced direct radiocarbon dates for some of the skeletons used in this study (Table S2) were generated by the Klaus-Tschira Archaeometriezentrum, Mannheim, Germany placing them within the range of the site.

Paso del Indio, Puerto Rico

Paso del Indio is located in the *municipio* of Vega Baja in north-central Puerto Rico, on the west bank of the Río Indio, 6 km south of the Atlantic coast (18°25'51.32"N and 66°22'54.45"W) (Fig. 1A). Based on the excavations and the following analysis, Paso del Indio would appear to have consisted of complex arrangement of domestic, production, and possible ritual/ceremonial contexts (*84*). Assessment of numerous (n=44) radiocarbon dates suggests that the site's occupational history may have spanned some 3500 years (*85*). Skeletal remains from nearly 150 burials were recovered from the site, and Pestle subsequently studied ninety-eight of those individuals isotopically (*86*). Direct AMS dates of these individuals range from roughly 1130-590 cal. BP (median probability). Irrespective of the exact dates, all of these individuals would appear to have lived during the Ceramic Age, in a period generally associated with agricultural subsistence economies.

Punta Candelero, Puerto Rico

Punta Candelero is located in the municipality of Humacao in south-eastern Puerto Rico. The site is on a sandy coastal peninsula on the grounds of the private Palmas del Mar Resort (18°05'37.4" N and 65°47'21.6" W). It was excavated between 1986 and 1989 by Miguel Rodríguez López (*87*). Punta Candelero has two well-defined and successive cultural components. The first corresponds to the "La Hueca" cultural complex (2300-1740 cal. BP) and the second corresponds to the "Cuevas" cultural complex (1290-940 cal. BP). Each component is associated with distinct ceramic assemblages. Multiple household structures and a central plaza were built at Punta Candelero during the later occupation of the site (*87*). Skeletal remains from 106 human burials were recovered from the second cultural component. Direct radiocarbon dates of these individuals range between 1690-510 cal. BP (*86*). Newly produced direct radiocarbon dates for one individual (PCA001) (Table S2) were generated by

the Klaus-Tschira Archaeometriezentrum, Mannheim, Germany, placing it in the range of this site.

Tibes, Puerto Rico

Tibes is located in the municipality of Ponce in southern Puerto Rico, on the alluvial terrace of the Rio Portugués, 8 km north of the Caribbean Sea coast (18°02’28.9" N and 66°37'13.2" W). The site was first excavated in 1975 by the *Sociedad Guaynía de Arqueología e Historia* and the *Socieded Arqueológica del Sur-Oeste de Puerto Rico.* Afterwards the site was acquired by the municipality of Ponce and an archaeological park was opened. Active research continues today at Tibes through the *Proyecto Arqueológico del Centro Ceremonial de Tibes* led by L. Antonio Curet *(88)*. Tibes was occupied continuously from 1650-750 cal. BP. The site has 12 monumental stone structures, including two plazas and seven ball-courts, which suggests the site was used as a civic-ceremonial center (*88*). Over 130 human burials have been found at Tibes, 126 of which have been the focus of bioarcheological research (*88*). Most human skeletal remains from Tibes have been directly dated between 1350-1150 cal. BP (*89*).

Los Indios, Puerto Rico

Los Indios is a multi-component site in the *municipio* of Santa Isabel in Southern Puerto Rico (18°0'14.5" N and 66°48'14.3" W). The domestic occupation of the site is well documented through radiocarbon dates spanning from 1130-580 cal. BP. The 130 excavated individuals of which only one genomic data point was produced, were predominantly entombed in primary single burials. The vast majority of individuals are adult skeletons buried in a flexed position similar to other sites in Puerto Rico such as Paso del Indio and Punta Candelero. The material culture of the site was described as Ostionoid with predominantly western, rather than eastern, Puerto Rican pottery styles. The radiocarbon date for the individual presented in this study (Table S2) was generated by the Klaus-Tschira Archaeometriezentrum, Mannheim, Germany, placing it within the range of the site.

Cueva de los Esqueletos, Cuba

The site Cueva de los Esqueletos 1 is located in the Sierra de Cubitas in the province of Camagüey. It was investigated during the 1970s and 1980s when several skeletons were found which were variously classified as “Taíno” or “sub-Taíno” due to the fact that they showed

signs of artificial cranial modification. The skeletons analysed in this study were directly dated to 520-400 cal. BP (Table S1).

El Morrillo, Cuba

El Morrillo is an archaeological site located at the embouchure of the Canímar River, north coast of Matanzas, Cuba (23°02'46.0" N 81°30'13.1" W). Dates for the site were obtained from a shell and human bone from the excavated individuals suggesting an occupation from 590-490 cal. BP. The site is identified in the National Archaeological Survey with the code 25009 and was first reported in 1964. The site was subject to repeated excavations, in which a rich variety of cultural artifacts were recovered. Some of the ceramic fragments recovered were decorated with basket impressions and fixation technique. Among the fragments recovered were zoomorphic handles. Other artifacts included spoons, plates, hammers, gouges and shell peaks, grinding stones and small mortars as well as pendants and lithic beads, but also flakes and other lithic artifacts. Two human burials have been exhumed at the site resulting from rescue excavations on the coastal shore. The first individual was excavated in 1979 and determined to be a male individual of about 45 years of age. The skeleton was oriented North to South, buried with one hand crossed on the back and the other on the forehead. The fragmented state of the cranium made identification of cranial deformation impossible. The second individual was excavated in 2009 and determined to be around 20-24 years old and displayed a cranial deformation. It was buried with the head facing southeast, an extended right arm and slightly bent left arm placed under the body and legs. The strontium isotope analysis of MO_2009 showed a mixed diet dominated by marine resources (*90*). New direct dates of 550-510 cal. BP for the individual presented in this study were produced at the Klaus-Tschira Archaeometriezentrum, Mannheim, Germany (Table S1 and S2).

Las Carolinas II, Cuba

Las Carolinas is an archaeological site located on the eastern bank of the Canímar River, north coast of Matanzas, Cuba (22°59'6" N, 81°27'43" W). It is identified in the National Archaeological Survey with the code 25118. The site was reported in 2007, when amateur archaeologists of the Matanzas Speleological Committee located several human remains on the surface in a collapsing area near the river channel. The pieces were delivered to the patrimonial authorities of the territory. The site is located on the edge of a cliff at 24 m altitude and the preliminary report revealed the existence of four burials in a disturbance context resulting from

the drag of the slope. The chronology of the site is directly dated on the excavated human remains by dates produced for this study to 1820-1530 cal. BP (Table S1 and S2) and places them along other groups of low food production ("Archaic"context) contemporaneous with other occupations in the Canímar river basin such as the younger cemetery in Canímar Abajo site and the Playita site.

Canímar Abajo, Cuba

Canímar Abajo is a shell-matrix site located near Matanzas City (Cuba) at the estuary on the western bank of the Canímar River (23°2'16.29" N and 81°29'48.25" W). The site consists of five stratigraphic levels and includes two cemeteries: the Old Cemetery (3080 ±110 cal. BP) and the Young Cemetery (1370 ±120 cal. BP) separated by an approximately 1 m thick shell-midden layer (*30*). Since the site was first discovered in 1984, at least 213 individuals have been excavated from the two cemeteries, including 130 juveniles. While Canímar Abajo was first described as an exclusively fisher-gatherer population, more recent palaeodietary studies showed that the use of terrestrial resources, including cultigens, formed a significant part of their diet (*91*). New direct dates for the individuals used in this study were produced by the Klaus-Tschira Archaeometriezentrum, Mannheim, Germany (Table S1 and S2) and range from 3200-720 cal. BP.

Playa del Mango, Cuba

The archaeological site of Playa del Mango is located in the Rio Cauto Basin in the province of Granma in eastern Cuba (20°33'14.38" N, 76°59'08.97" W). The site is adjacent to the 'Laguna El Mango', approximately 3.5 km east of the lagoon system of Las Playas and 14 km inland north from the Gulf of Guacanayabo. The site covers approximately 6 km^2 and includes three mounds. The two largest mounds, mound 1 and mound 2 are approximately 60 and 40 meters in diameter and 7 and 3 meters in height, respectively. According to the traditional Cuban classification systems, the site occupants were either "Ciboney" (*92*), "Preagroalfareros" (*93*), or "Guacanayabo Fisher-Gatherers" (*94*). Playa del Mango was first excavated in 1941 by Dr. Bernardo Utset Masía who uncovered approximately 40 skeletons at a depth of ~21 cm below the surface of mound 1. The skeletons were in anatomical position, laying on their backs (dorsal decubitus) in primary burials. Some individuals had beads made of vertebral fish and sharks around their necks, wrists and ankles. Unfortunately, the osteological remains are no longer available and only limited contextual information regarding

these burials was recorded in Utset. Between 1980 and 1986 the Academy of Sciences of Cuba excavated two new areas in M1 and M2, recovering additional faunal remains and artifacts (*95*). Between 2014 and 2018, a joint project by the Cuban Institute of Anthropology and The University of Winnipeg (led by Dr. Mirjana Roksandic) resumed excavations at the site. As a result of the first surveys in 2014 and 2015, some isolated human bones and beads made of fish/shark vertebrae were found. In 2016 and 2018, Block 3 and 7 on the mound, and Block 4 and 6 in the marginal area of the mound were excavated under the direction of Dr. Ulises Gonzalez Herrera (Cuban Institute of Anthropology) and Dr. Yadira Chinique de Armas (University of Winnipeg). In the marginal area of mound 2, 20 skeletons in anatomical position were found, some of which were partially or completely altered due to disturbance by modern agricultural practices. The radiocarbon dates for the skeletons excavated from mound 2 that were produced for this and another study (*96*) and forthcoming dates range from 2160-1515 cal. BP (Table S1 and S2).

Manuelito, Cuba

Manuelito (or Morejón 01) is an archaeological site located in the center of the province of Matanzas, Cuba, very close to the southern part of the watershed of the Canímar River (22°49'24.144" N, 81°23'25.979" W). It is identified in the National Archaeological Survey with the code 25166. The site was reported in 2002, when amateur archaeologists located hundreds of lithic pieces (*majadores, percutores, lajas* grinders, beads, rings), shell (gouges, vessels), flint, faunal remains and human remains on the surface of extensively plowed land. The preliminary report revealed the existence of two burials in a context of disturbance resulting from anthropogenic impacts. The remains were delivered to the patrimonial authorities of the territory. Site dates were obtained by directly dating the human remains to 1290-1180 cal. BP (Table S1 and S2) showing them to be roughly contemporaneous to the younger cemetery of the Canímar Abajo site, raising the question whether the fluvial connection facilitated the connection of the communities in the Canímar river delta. The newly produced direct radiocarbon dates for the skeletons used in this study (Table S2) were generated by the Klaus-Tschira Archaeometriezentrum, Mannheim, Germany.

Cueva del Perico, Cuba

The Cueva del Perico I cemetery is located inside a cave in the northern part of Artemisa province, Cuba (83°16'29.9" W and 22°57'29.4" N). It is one kilometer from the Mani-Mani

River and 25 km from Bahia Honda (*97*)*97(97)*. Excavations conducted between 1970 and 1997 at Cueva del Perico I recovered evidence for a minimum of 162 individuals entombed at the cave site (*98*)*98(98)*. The population was classified as “fisher-gatherers” or “Ciboney” (*97*)*97(97)*. Some artifacts, such as gouges and grinding stones, were found at the site, animal remains were found outside of the cave. The radiocarbon dates obtained indicate that the site was used as a cemetery at least since 2050-1730 cal. BP and until 1350-1150 cal. BP (*97*, *99*)*97(97, 99)99(97, 99)*. A recent AMS ^{14}C date on human bone collagen of one individual indicated that the cemetery was in use at 1560-1370 cal. BP. Newly produced direct radiocarbon dates for the skeletons used in this study (Table S1, Fig. 1B) were generated by the Klaus-Tschira Archaeometriezentrum, Mannheim, Germany and range from 2750-1190 cal. BP, showing an occupation 700 years earlier than previously assumed.

Guayabo Blanco, Cuba

The site of Guayabo Blanco is located 30 kilometers from the nearest coast in the wetland region Ciénaga de Zapata in the south-coast of the Matanzas province (80°55' 25.7" W and 22°18'2.9" N). Seven human skeletons were found in a 1.5 meter high elliptic mound, intentionally formed by adding soil to bury the human bodies. This site was used as a type site for the “Ciboney Guayabo Blanco type” (*92*, *100*)*100(92, 100)*, which is associated with a material culture and industry based on marine resources, including shell gouges from marine mollusks. Newly produced direct radiocarbon dates for the skeletons used in this study (Table S2) were generated by the Klaus-Tschira Archaeometriezentrum, Mannheim, Germany and show occupation of the site from 2530-1700 cal. BP.

Cueva Calero, Cuba

The Cueva Calero is a cave site more than 5 km inland from the nearest coast in the municipality of Cardenas in the Matanzas Province (81°19'1.6" W and 23°2' 58.7" N). The site is situated in the Camarioca river basin, which is adjoining the Canímar river delta, possibly connecting this site with other sites analyzed in this study from the Canímar River delta. Excavation in 1989 uncovered 55 individuals from a 150 sqm burial. A lack of preserved collagen allowed no direct dating of the individuals studied in this work, but a direct date for one skeleton excavated from the Area 2, Trinchera 1, Sección C is available with an age of 1380 ±50 cal. BP (Chinique de Armas, forthcoming). New direct dates for the individuals used

in this study were attempted by the Klaus-Tschira Archaeometriezentrum, Mannheim, Germany but no collagen was preserved to date the individuals (Table S2).

Supplementary Figures

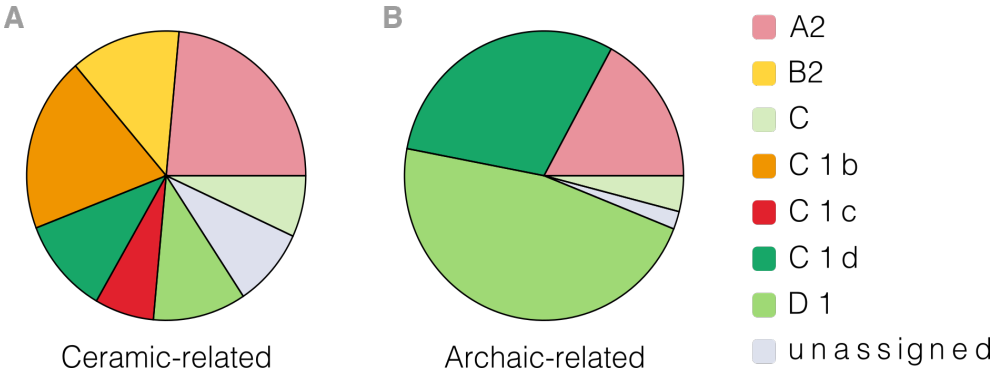


Figure S1. Mitochondrial Haplogroup Frequencies. Frequencies of mitochondrial haplogroups as determined by Haplogrep 2.0 (*101*)*101(101)* (**A**) for Ceramic-related individuals and (**B**) for Archaic-related individuals.

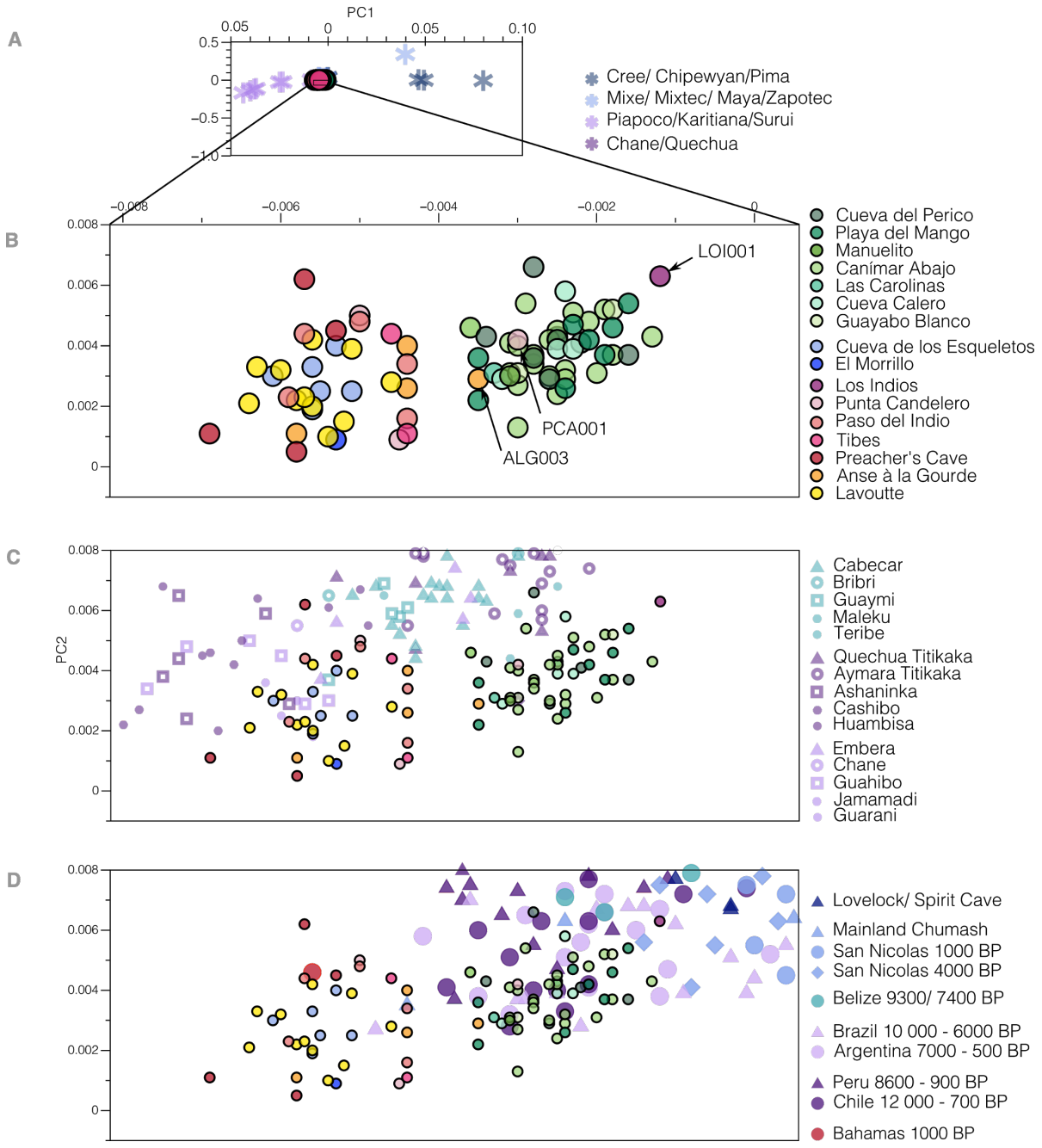


Figure S2: Principal component analysis (PCA) of ancient Caribbean islanders. (**A**) PCA showing ancient Caribbean islanders projected onto PCs calculated based on present-day Native American populations (*55*). (**B**) Focusing on the projected ancient Caribbean individuals. Labeled individuals who cluster outside their main grouping are most similar to the assigned group based on f_4-statistics (Table S5). (**C**) Caribbean individuals and present-day Native Americans, also projected (*13*, *17*). (**D**) Caribbean individuals and previously published projected ancient Native Americans (*14–16*).

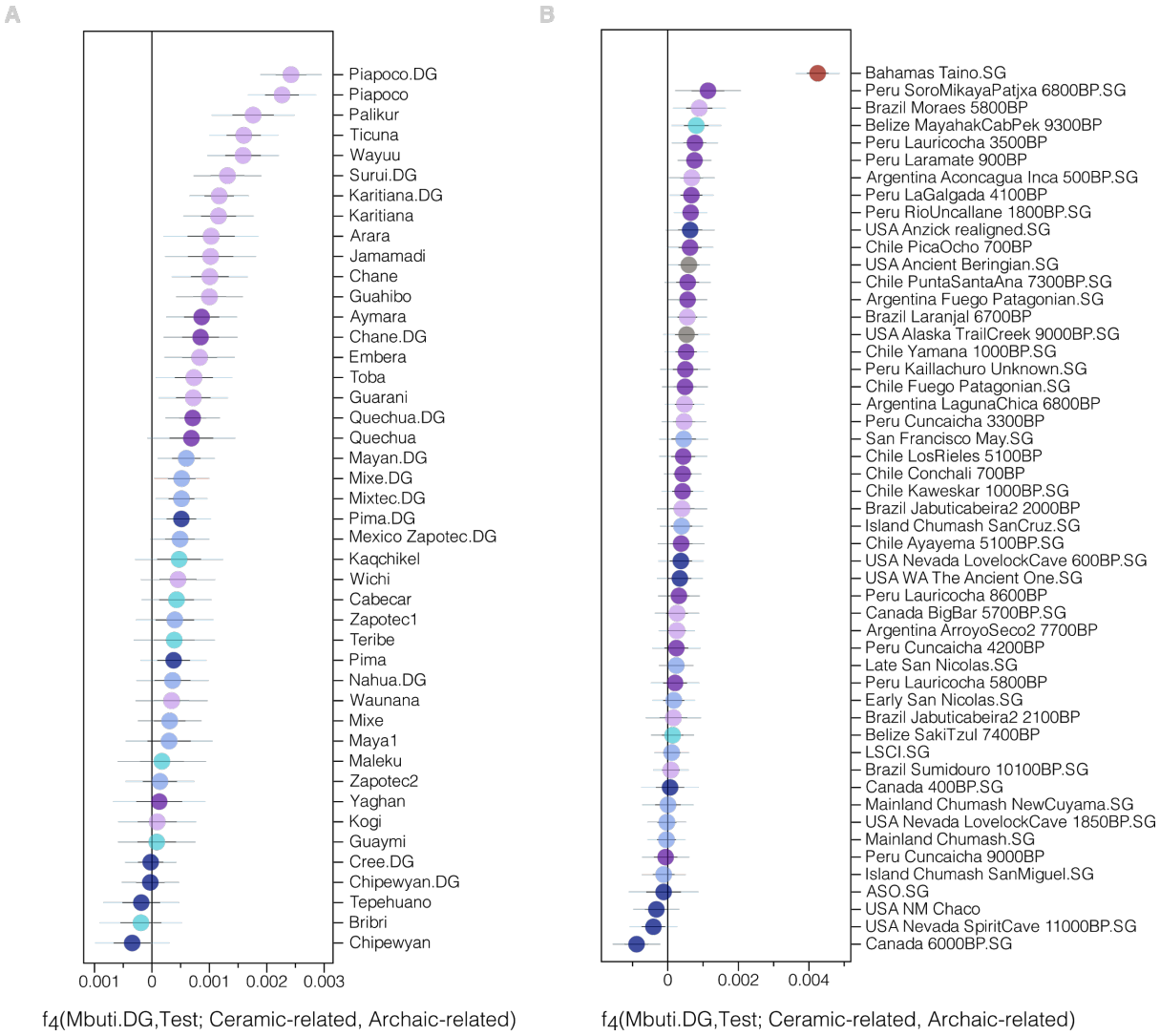


Figure S3. Differential affinities of the two groups to ancient groups and individuals from the Americas. *f_4(Mbuti.DG, Test; Archaic-related, Ceramic-related)* (**A**) with present-day individuals (*13*, *55*), and (**B**) with published ancient genomes (*7*, *14–16*, *56–59*). Negative values indicating more affinity of the tested group to the group with Ceramic-related contexts while positive values show greater affinities to the group from Cuba 3000-800 cal. BP (Archaic-related).

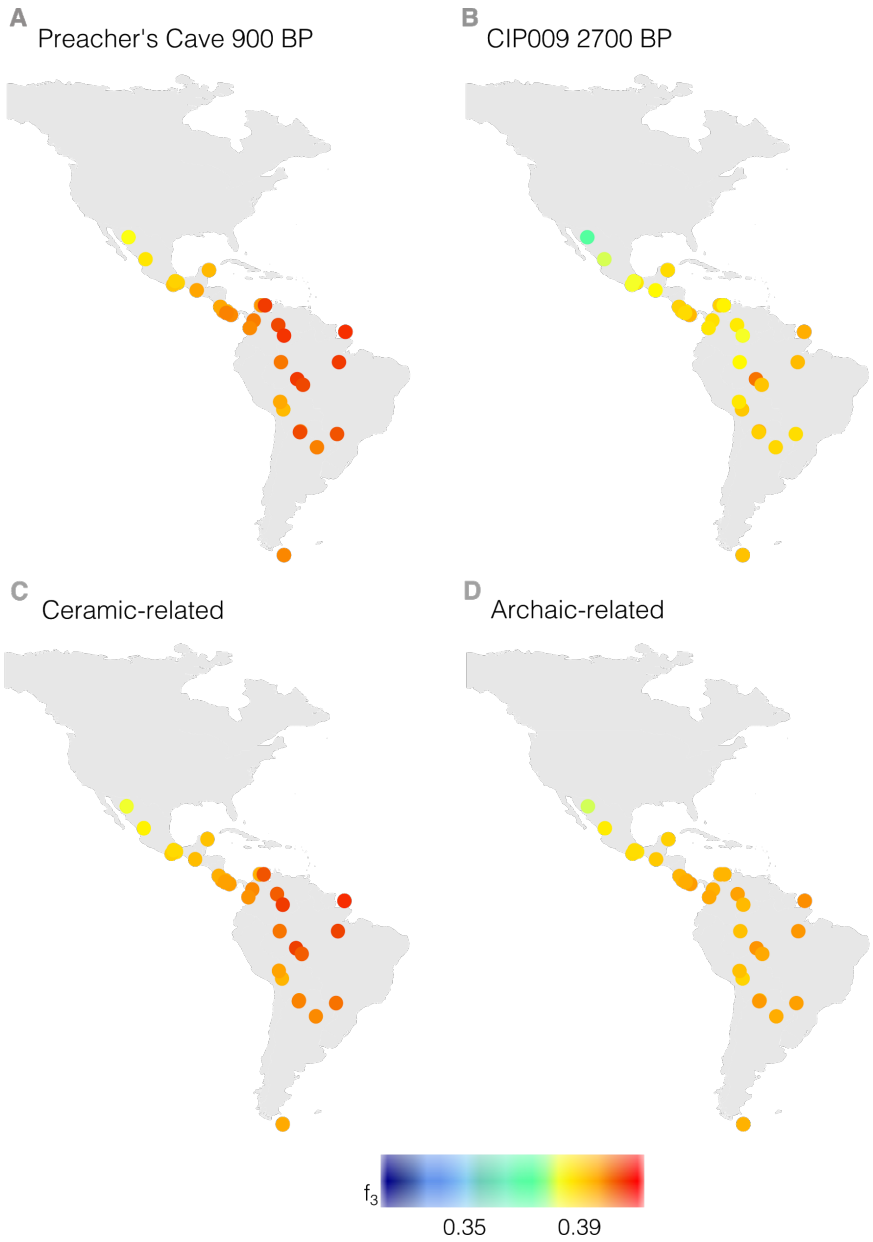


Figure S4. Shared genetic drift. Heat maps showing the shared genetic drift between 29 present-day Native American groups and four newly sequenced individuals from Preacher's Cave in the Bahamas (**A**), the oldest individual from the Cueva del Perico (CIP009, 2700 cal. BP) (**B**), all Ceramic-related individuals (**C**), and all Archaic-related individuals (**D**), as measured by the *f*-statistic f_3(Mbuti.DG; Test, Caribbean individuals/group).

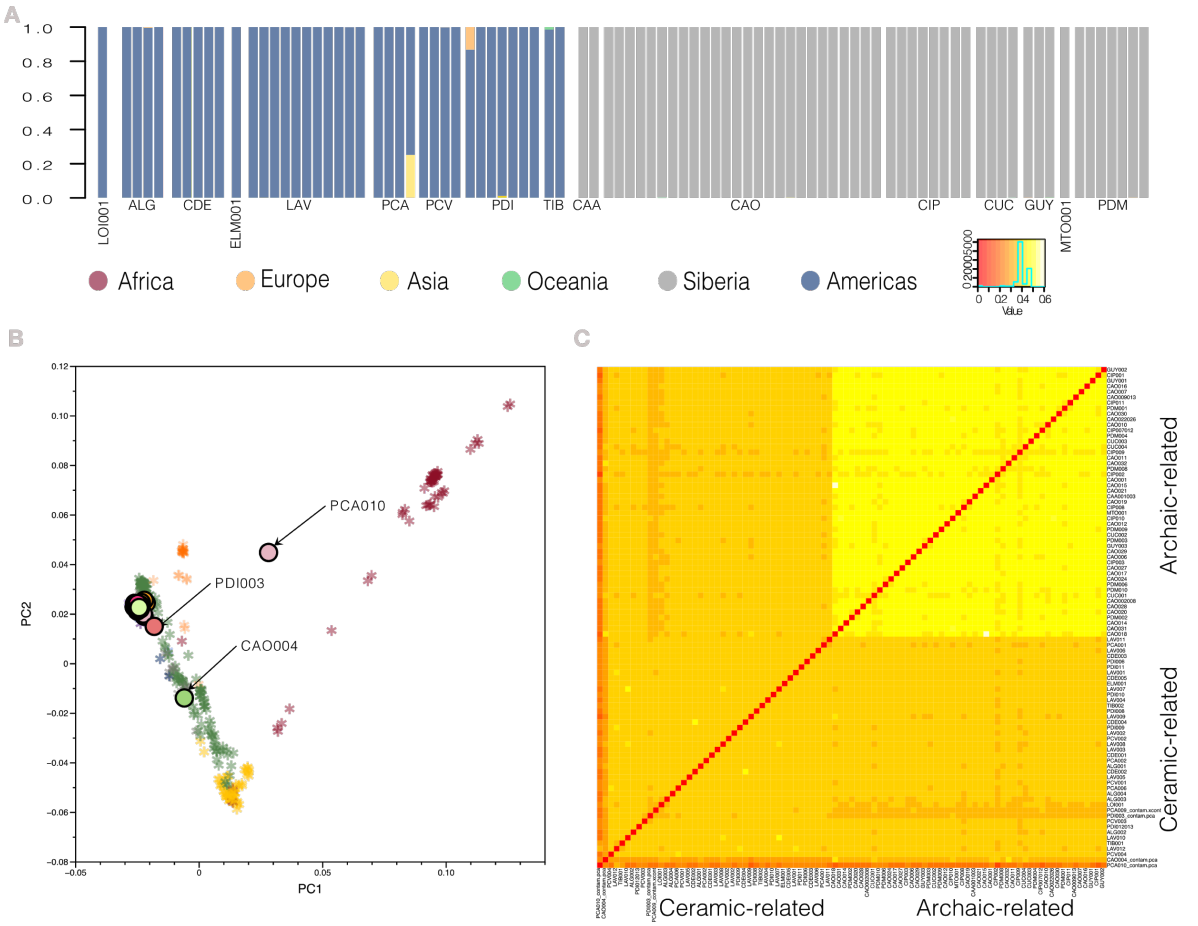


Figure S5. Quality control of newly sequenced individuals. (A) Model based clustering run in supervised mode. The six selected ancestry components are African (Mbuti.DG, Yoruba.DG), European (French.DG, Sardinian.DG, English.DG), Asian (Han.DG, Ami.DG, Atayal.DG), Oceanian (Papuan.DG), Siberian (Itelmen.DG, Chukchi.DG, Ulchi.DG) and Native American (Karitiana.DG, Surui.DG, Mixe.DG, Pima.DG). (**B**) Worldwide PCA calculated on present-day individuals from the Simons Genome Diversity Project (*55*). Ancient samples are projected and the three indicated samples are shifted towards Europe and Africa indicating possible DNA contamination. (**C**) Heatmap of pairwise f_3-ougroups statistics for all individuals compared where light colours indicate high genetic similarity, as is the case for all Archaic-related individuals. Darker colors indicate less genetic similarity as is the case for the individuals already identified in a worldwide PCA to potentially carry contamination.

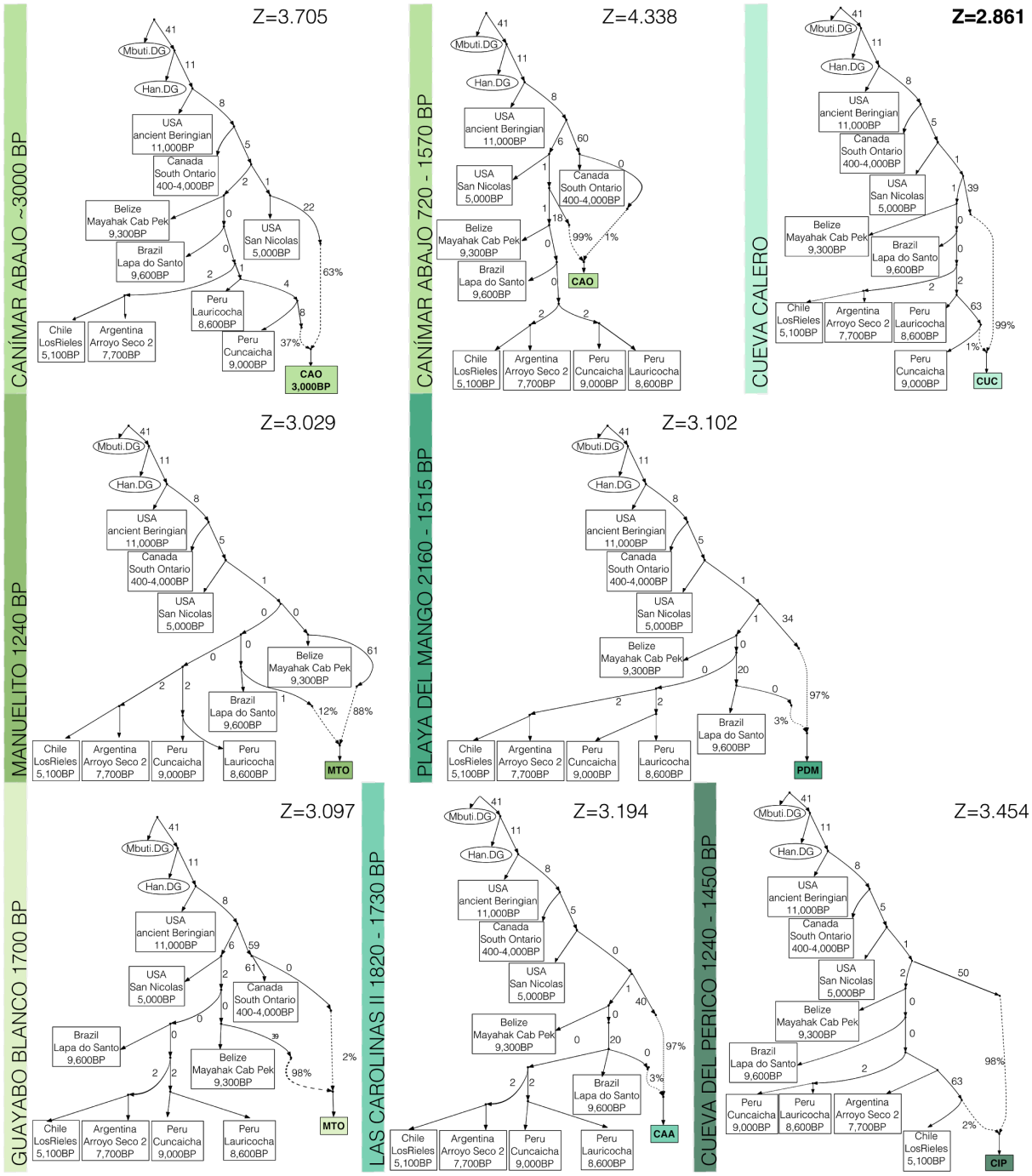


Figure S6. Admixture Graph modeling for Archaic-related groups. Admixture graphs modeling the ancestry of Archaic-related Caribbean islanders. We show the best-fitting model for each genome (or group of genomes) as inferred from the associated worst *Z*-score. Numbers to the right of solid edges are proportional to optimized drift while percentages next to the dashed edges represent admixture proportions.

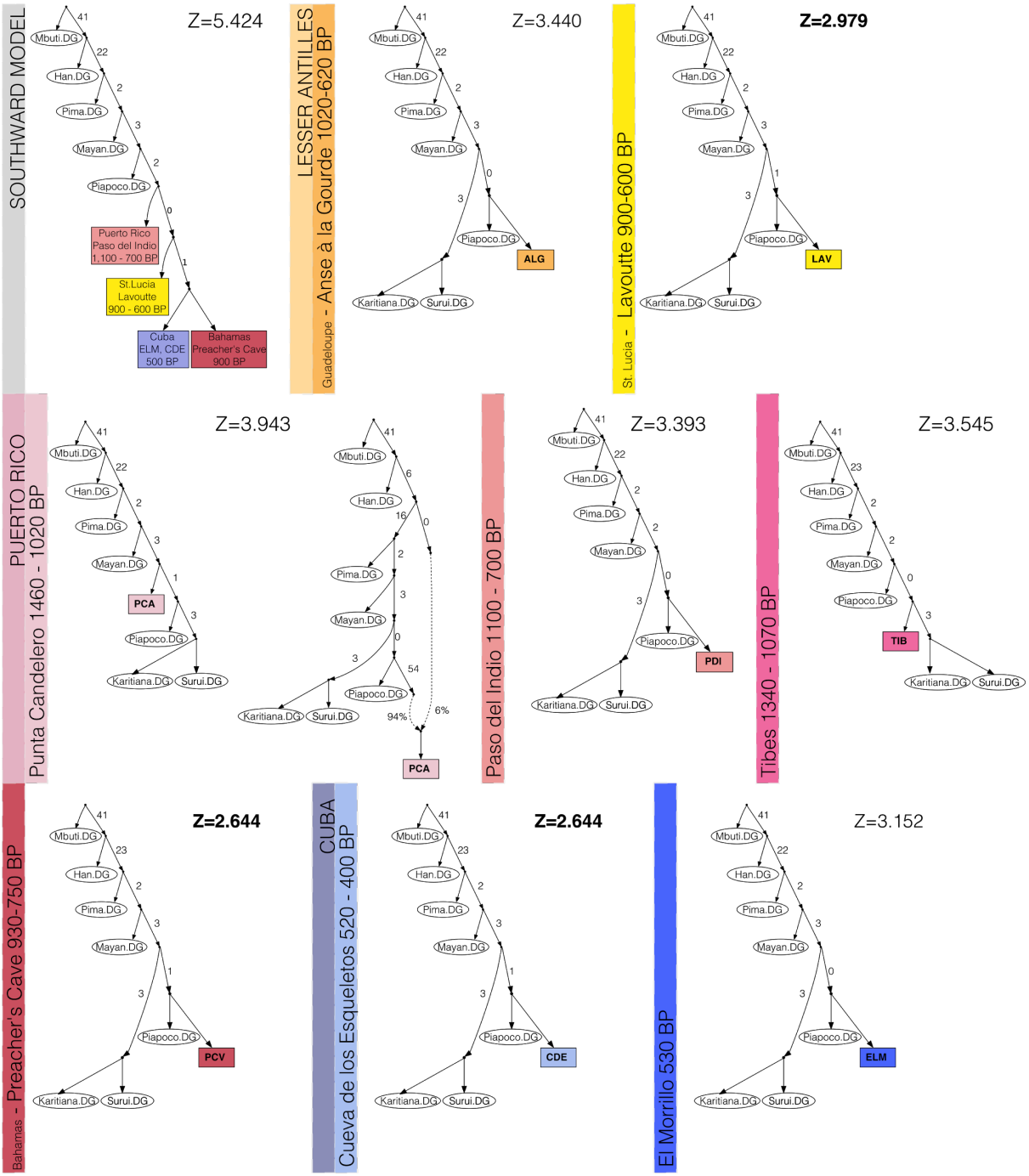


Figure S7. Admixture graphs modeling the ancestry of Ceramic-related Caribbean islanders. We show the best-fitting model for each genome (or group of genomes) as indicated in Figure S6 with the exception of the first panel where we test a specific model of an expansion from the South American mainland to Puerto Rico first (Paso del Indio) and then back to the Lesser Antilles (Lavoutte) and further north and west to Cuba and the Bahamas (Preacher's Cave) (Southward expansion).

Captions for Tables S1-S7 (available as external data sheets)

Table S1. Samples. Overview of the samples analysed in this study listing country of origin, site name, latitude and longitude, Museum ID, Lab ID, sampled element, C14 date cal. BP 2-sigma, genetic sex, mtDNA haplogroup, ChrY-haplogroup, and SNPs covered on the 1240K SNP panel.

Table S2. Radiocarbon dates. Uncalibrated dates are reported as C14 ages BP and calibrated dates as cal. BP. CN ratios, %C and % collagen are also reported.

Table S3. Summary statistics for the sequenced libraries. Summaries obtained through EAGER showing depth of sequencing (# of raw reads), library complexity (# of reads after removed duplicates; Cluster factor), DNA proportion (Endogenous DNA QF %), percentage of damaged sites on the 3' and 5' end of the reads, average fragment length and CG content.

Table S4. Contamination estimates. Genome-wide estimates are based on ANGSD and mtDNA estimates are based on Schmutzi. Samples with higher contamination estimates are highlighted in red and were excluded from downstream analyses.

Table S5. *f*-statistics. Various sets of outgroup f_3- and f_4-statistics testing the amount of shared genetic drift between the ancient Caribbean islanders and other ancient and present-day Native American groups/individuals. Standard errors were calculated by dividing the test statistics by the resulting Z-score.

Table S6. qpWave results. Results of the qpWave analysis relative to the set of outgroups: Mbuti.DG, Onge.DG, Papuan.DG, Han.DG, Russia_MA1_HG.SG, USA_Ancient_Beringian.SG; USA_Anzick.SG, Mixe.DG, Mexico_Zapotec.DG, Belize_MayahakCabPek_9300, Karitiana.DG, Piapoco.DG. The assumption of the two tested groups deriving from the same ancestral source is rejected if the p-value is lower than 0.05.

Table S7. qpAdm results. Results of the qpAdm analysis using the following outgroups: Mbuti.DG, Onge.DG, Papuan.DG, Han.DG, Russia_MA1_HG.SG, USA_Ancient_Beringian.SG; USA_Anzick.SG, Mixe.DG, Mexico_Zapotec.DG, Belize_MayahakCabPek_9300, Karitiana.DG, Piapoco.DG, CIP009. The target is modelled as a mixture of Source 1 and Source 2. P-values below 0.05 indicate a poor fit of the model.

References and Notes

1. W. F. Keegan, C. L. Hofman, *The Caribbean Before Columbus* (Oxford Univ. Press, 2017).

2. M. F. Napolitano, R. J. DiNapoli, J. H. Stone, M. J. Levin, N. P. Jew, B. G. Lane, J. T. O'Connor, S. M. Fitzpatrick, Reevaluating human colonization of the Caribbean using chronometric hygiene and Bayesian modeling. *Sci. Adv.* **5**, eaar7806 (2019). doi:10.1126/sciadv.aar7806 Medline

3. W. R. Coe II, A Distinctive Artifact Common to Haiti and Central America. *Am. Antiq.* **22**, 280–282 (1957). doi:10.2307/276564

4. S. M. Wilson, H. B. Iceland, T. R. Hester, Preceramic Connections between Yucatan and the Caribbean. *Lat. Am. Antiq.* **9**, 342–352 (1998). doi:10.2307/3537032

5. R. T. Callaghan, Comments on the Mainland Origins of the Preceramic Cultures of the Greater Antilles. *Lat. Am. Antiq.* **14**, 323–338 (2003). doi:10.2307/3557564

6. I. Rouse, *The Tainos: Rise and Decline of the People Who Greeted Columbus* (Yale Univ. Press, 1992).

7. H. Schroeder, M. Sikora, S. Gopalakrishnan, L. M. Cassidy, P. Maisano Delser, M. Sandoval Velasco, J. G. Schraiber, S. Rasmussen, J. R. Homburger, M. C. Ávila-Arcos, M. E. Allentoft, J. V. Moreno-Mayar, G. Renaud, A. Gómez-Carballa, J. E. Laffoon, R. J. A. Hopkins, T. F. G. Higham, R. S. Carr, W. C. Schaffer, J. S. Day, M. Hoogland, A. Salas, C. D. Bustamante, R. Nielsen, D. G. Bradley, C. L. Hofman, E. Willerslev, Origins and genetic legacies of the Caribbean Taino. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 2341–2346 (2018). doi:10.1073/pnas.1716839115 Medline

8. S. M. Fitzpatrick, Seafaring Capabilities in the Pre-Columbian Caribbean. *J. Mari. Arch.* **8**, 101–138 (2013). doi:10.1007/s11457-013-9110-8

9. A. T. Antczak, C. L. Hofman, in *Early Settlers of the Insular Caribbean: Dearchaizing the Archaic*, C. L. Hofman, A. T. Antczak, Eds. (Sidestone Press, 2019), pp. 29–42.

10. See supplementary materials.

11. C. Lalueza-Fox, F. L. Calderón, F. Calafell, B. Morera, J. Bertranpetit, MtDNA from extinct Tainos and the peopling of the Caribbean. *Ann. Hum. Genet.* **65**, 137–151 (2001). doi:10.1046/j.1469-1809.2001.6520137.x Medline

12. M. A. Nieves-Colón, W. J. Pestle, A. W. Reynolds, B. Llamas, C. de la Fuente, K. Fowler, K. M. Skerry, E. Crespo-Torres, C. D. Bustamante, A. C. Stone, Ancient DNA Reconstructs the Genetic Legacies of Precontact Puerto Rico Communities. *Mol. Biol. Evol.* **37**, 611–626 (2020). doi:10.1093/molbev/msz267 Medline

13. D. Reich, N. Patterson, D. Campbell, A. Tandon, S. Mazieres, N. Ray, M. V. Parra, W. Rojas, C. Duque, N. Mesa, L. F. García, O. Triana, S. Blair, A. Maestre, J. C. Dib, C. M. Bravi, G. Bailliet, D. Corach, T. Hünemeier, M. C. Bortolini, F. M. Salzano, M. L. Petzl-Erler, V. Acuña-Alonzo, C. Aguilar-Salinas, S. Canizales-Quinteros, T. Tusié-Luna, L. Riba, M. Rodríguez-Cruz, M. Lopez-Alarcón, R. Coral-Vazquez, T. Canto-Cetina, I. Silva-Zolezzi, J. C. Fernandez-Lopez, A. V. Contreras, G. Jimenez-Sanchez, M. J. Gómez-Vázquez, J. Molina, A. Carracedo, A. Salas, C. Gallo, G. Poletti, D. B. Witonsky, G. Alkorta-Aranburu, R. I. Sukernik, L. Osipova, S. A. Fedorova, R. Vasquez, M. Villena, C. Moreau, R. Barrantes, D. Pauls, L. Excoffier, G. Bedoya, F. Rothhammer, J.-M. Dugoujon, G. Larrouy, W. Klitz, D. Labuda, J. Kidd, K. Kidd, A. Di Rienzo, N. B.

Freimer, A. L. Price, A. Ruiz-Linares, Reconstructing Native American population history. *Nature* **488**, 370–374 (2012). doi:10.1038/nature11258 Medline

14. J. V. Moreno-Mayar, L. Vinner, P. de Barros Damgaard, C. de la Fuente, J. Chan, J. P. Spence, M. E. Allentoft, T. Vimala, F. Racimo, T. Pinotti, S. Rasmussen, A. Margaryan, M. Iraeta Orbegozo, D. Mylopotamitaki, M. Wooller, C. Bataille, L. Becerra-Valdivia, D. Chivall, D. Comeskey, T. Devièse, D. K. Grayson, L. George, H. Harry, V. Alexandersen, C. Primeau, J. Erlandson, C. Rodrigues-Carvalho, S. Reis, M. Q. R. Bastos, J. Cybulski, C. Vullo, F. Morello, M. Vilar, S. Wells, K. Gregersen, K. L. Hansen, N. Lynnerup, M. Mirazón Lahr, K. Kjær, A. Strauss, M. Alfonso-Durruty, A. Salas, H. Schroeder, T. Higham, R. S. Malhi, J. T. Rasic, L. Souza, F. R. Santos, A.-S. Malaspinas, M. Sikora, R. Nielsen, Y. S. Song, D. J. Meltzer, E. Willerslev, Early human dispersals within the Americas. *Science* **362**, eaav2621 (2018). doi:10.1126/science.aav2621 Medline

15. C. Posth, N. Nakatsuka, I. Lazaridis, P. Skoglund, S. Mallick, T. C. Lamnidis, N. Rohland, K. Nägele, N. Adamski, E. Bertolini, N. Broomandkhoshbacht, A. Cooper, B. J. Culleton, T. Ferraz, M. Ferry, A. Furtwängler, W. Haak, K. Harkins, T. K. Harper, T. Hünemeier, A. M. Lawson, B. Llamas, M. Michel, E. Nelson, J. Oppenheimer, N. Patterson, S. Schiffels, J. Sedig, K. Stewardson, S. Talamo, C.-C. Wang, J.-J. Hublin, M. Hubbe, K. Harvati, A. Nuevo Delaunay, J. Beier, M. Francken, P. Kaulicke, H. Reyes-Centeno, K. Rademaker, W. R. Trask, M. Robinson, S. M. Gutierrez, K. M. Prufer, D. C. Salazar-García, E. N. Chim, L. Müller Plumm Gomes, M. L. Alves, A. Liryo, M. Inglez, R. E. Oliveira, D. V. Bernardo, A. Barioni, V. Wesolowski, N. A. Scheifler, M. A. Rivera, C. R. Plens, P. G. Messineo, L. Figuti, D. Corach, C. Scabuzzo, S. Eggers, P. DeBlasis, M. Reindel, C. Méndez, G. Politis, E. Tomasto-Cagigao, D. J. Kennett, A. Strauss, L. Fehren-Schmitz, J. Krause, D. Reich, Reconstructing the Deep Population History of Central and South America. *Cell* **175**, 1185–1197.e22 (2018). doi:10.1016/j.cell.2018.10.027 Medline

16. C. L. Scheib, H. Li, T. Desai, V. Link, C. Kendall, G. Dewar, P. W. Griffith, A. Mörseburg, J. R. Johnson, A. Potter, S. L. Kerr, P. Endicott, J. Lindo, M. Haber, Y. Xue, C. Tyler-Smith, M. S. Sandhu, J. G. Lorenz, T. D. Randall, Z. Faltyskova, L. Pagani, P. Danecek, T. C. O'Connell, P. Martz, A. S. Boraas, B. F. Byrd, A. Leventhal, R. Cambra, R. Williamson, L. Lesage, B. Holguin, E. Ygnacio-De Soto, J. Rosas, M. Metspalu, J. T. Stock, A. Manica, A. Scally, D. Wegmann, R. S. Malhi, T. Kivisild, Ancient human parallel lineages within North America contributed to a coastal expansion. *Science* **360**, 1024–1027 (2018). doi:10.1126/science.aar6851 Medline

17. G. A. Gnecchi-Ruscone, S. Sarno, S. De Fanti, L. Gianvincenzo, C. Giuliani, A. Boattini, E. Bortolini, T. Di Corcia, C. Sanchez Mellado, T. J. Dàvila Francia, D. Gentilini, A. M. Di Blasio, P. Di Cosimo, E. Cilli, A. Gonzalez-Martin, C. Franceschi, Z. A. Franceschi, O. Rickards, M. Sazzini, D. Luiselli, D. Pettener, Dissecting the Pre-Columbian Genomic Ancestry of Native Americans along the Andes-Amazonia Divide. *Mol. Biol. Evol.* **36**, 1254–1269 (2019). doi:10.1093/molbev/msz066 Medline

18. N. Patterson, P. Moorjani, Y. Luo, S. Mallick, N. Rohland, Y. Zhan, T. Genschoreck, T. Webster, D. Reich, Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012). doi:10.1534/genetics.112.145037 Medline

19. W. Haak, I. Lazaridis, N. Patterson, N. Rohland, S. Mallick, B. Llamas, G. Brandt, S. Nordenfelt, E. Harney, K. Stewardson, Q. Fu, A. Mittnik, E. Bánffy, C. Economou, M.

Francken, S. Friederich, R. G. Pena, F. Hallgren, V. Khartanovich, A. Khokhlov, M. Kunst, P. Kuznetsov, H. Meller, O. Mochalov, V. Moiseyev, N. Nicklisch, S. L. Pichler, R. Risch, M. A. Rojo Guerra, C. Roth, A. Szécsényi-Nagy, J. Wahl, M. Meyer, J. Krause, D. Brown, D. Anthony, A. Cooper, K. W. Alt, D. Reich, Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015). doi:10.1038/nature14317 Medline

20. W. F. Keegan, Archaic Influences in the Origins and Development of Taino Societies. *Caribb. J. Sci.* **42**, 1 (2006).

21. C. L. Hofman, A. T. Antczak, Eds., *Early Settlers of the Insular Caribbean: Dearchaizing the Archaic* (Sidestone Press, 2019).

22. M. E. Allentoft, M. Sikora, K.-G. Sjögren, S. Rasmussen, M. Rasmussen, J. Stenderup, P. B. Damgaard, H. Schroeder, T. Ahlström, L. Vinner, A.-S. Malaspinas, A. Margaryan, T. Higham, D. Chivall, N. Lynnerup, L. Harvig, J. Baron, P. Della Casa, P. Dąbrowski, P. R. Duffy, A. V. Ebel, A. Epimakhov, K. Frei, M. Furmanek, T. Gralak, A. Gromov, S. Gronkiewicz, G. Grupe, T. Hajdu, R. Jarysz, V. Khartanovich, A. Khokhlov, V. Kiss, J. Kolář, A. Kriiska, I. Lasak, C. Longhi, G. McGlynn, A. Merkevicius, I. Merkyte, M. Metspalu, R. Mkrtchyan, V. Moiseyev, L. Paja, G. Pálfi, D. Pokutta, Ł. Pospieszny, T. D. Price, L. Saag, M. Sablin, N. Shishlina, V. Smrčka, V. I. Soenov, V. Szeverényi, G. Tóth, S. V. Trifanova, L. Varul, M. Vicze, L. Yepiskoposyan, V. Zhitenev, L. Orlando, T. Sicheritz-Pontén, S. Brunak, R. Nielsen, K. Kristiansen, E. Willerslev, Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015). doi:10.1038/nature14507 Medline

23. H. McColl, F. Racimo, L. Vinner, F. Demeter, T. Gakuhari, J. V. Moreno-Mayar, G. van Driem, U. Gram Wilken, A. Seguin-Orlando, C. de la Fuente Castro, S. Wasef, R. Shoocongdej, V. Souksavatdy, T. Sayavongkhamdy, M. M. Saidin, M. E. Allentoft, T. Sato, A.-S. Malaspinas, F. A. Aghakhanian, T. Korneliussen, A. Prohaska, A. Margaryan, P. de Barros Damgaard, S. Kaewsutthi, P. Lertrit, T. M. H. Nguyen, H. C. Hung, T. Minh Tran, H. Nghia Truong, G. H. Nguyen, S. Shahidan, K. Wiradnyana, H. Matsumae, N. Shigehara, M. Yoneda, H. Ishida, T. Masuyama, Y. Yamada, A. Tajima, H. Shibata, A. Toyoda, T. Hanihara, S. Nakagome, T. Deviese, A.-M. Bacon, P. Duringer, J.-L. Ponche, L. Shackelford, E. Patole-Edoumba, A. T. Nguyen, B. Bellina-Pryce, J.-C. Galipaud, R. Kinaston, H. Buckley, C. Pottier, S. Rasmussen, T. Higham, R. A. Foley, M. M. Lahr, L. Orlando, M. Sikora, M. E. Phipps, H. Oota, C. Higham, D. M. Lambert, E. Willerslev, The prehistoric peopling of Southeast Asia. *Science* **361**, 88–92 (2018). doi:10.1126/science.aat3628 Medline

24. C. Posth, K. Nägele, H. Colleran, F. Valentin, S. Bedford, K. W. Kami, R. Shing, H. Buckley, R. Kinaston, M. Walworth, G. R. Clark, C. Reepmeyer, J. Flexner, T. Maric, J. Moser, J. Gresky, L. Kiko, K. J. Robson, K. Auckland, S. J. Oppenheimer, A. V. S. Hill, A. J. Mentzer, J. Zech, F. Petchey, P. Roberts, C. Jeong, R. D. Gray, J. Krause, A. Powell, Language continuity despite population replacement in Remote Oceania. *Nat. Ecol. Evol.* **2**, 731–740 (2018). doi:10.1038/s41559-018-0498-2 Medline

25. C. L. Hofman, A. J. Bright, R. Rodríguez Ramos, Crossing the Caribbean Sea: Towards a holistic view of Pre-Colonial mobility and exchange. *J. Caribb. Archaeol.* **3**, 1–18 (2010).

26. L. A. Curet, The Taino: Phenomena, Concepts, and Terms. *Ethnohistory* **61**, 467–495 (2014). doi:10.1215/00141801-2681759

27. W. F. Keegan, Creating the Guanahatabey (Ciboney): The modern genesis of an extinct culture. *Antiquity* **63**, 373–379 (1989). doi:10.1017/S0003598X00076109

28. I. Rouse, The Classification of Artifacts in Archaeology. *Am. Antiq.* **25**, 313–323 (1960). doi:10.2307/277514

29. J. M. Cruxent, I. Rouse, *An Archaeological Chronology of Venezuela* (Pan American Union, 1958).

30. M. Roksandic, W. Mark Buhay, Y. Chinique de Armas, R. Rodríguez Suárez, M. C. Peros, I. Roksandic, S. Mowat, L. M. Viera, C. Arredondo, A. Martínez Fuentes, D. Gray Smith, Radiocarbon and Stratigraphic Chronology of Canímar Abajo, Matanzas, Cuba. *Radiocarbon* **57**, 755–763 (2015). doi:10.2458/azu_rc.57.18313

31. A. Boomert, *Trinidad, Tobago, and the Lower Orinoco Interaction Sphere: An Archaeological/Ethnohistorical Study* (Cairi Publications, 2000).

32. Y. Chinique de Armas, M. Roksandic, R. R. Suárez, D. G. Smith, W. M. Buhay, in *Cuban Archaeology in the Caribbean*, I. Roksandic, Ed. (Univ. of Florida Press, 2016).

33. C. Gamba, E. R. Jones, M. D. Teasdale, R. L. McLaughlin, G. Gonzalez-Fortes, V. Mattiangeli, L. Domboróczki, I. Kővári, I. Pap, A. Anders, A. Whittle, J. Dani, P. Raczky, T. F. G. Higham, M. Hofreiter, D. G. Bradley, R. Pinhasi, Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5**, 5257 (2014). doi:10.1038/ncomms6257 Medline

34. P. B. Damgaard, A. Margaryan, H. Schroeder, L. Orlando, E. Willerslev, M. E. Allentoft, Improving access to endogenous DNA in ancient bones and teeth. *Sci. Rep.* **5**, 11184 (2015). doi:10.1038/srep11184 Medline

35. R. Longin, New method of collagen extraction for radiocarbon dating. *Nature* **230**, 241–242 (1971). doi:10.1038/230241a0 Medline

36. T. F. G. Higham, R. M. Jacobi, C. B. Ramsey, AMS Radiocarbon Dating of Ancient Bone Using Ultrafiltration. *Radiocarbon* **48**, 179–195 (2006). doi:10.1017/S0033822200066388

37. R. J. A. Hopkins, C. Snoeck, T. F. G. Higham, When Dental Enamel is Put to the Acid Test: Pretreatment Effects and Radiocarbon Dating. *Radiocarbon* **58**, 893–904 (2016). doi:10.1017/RDC.2016.52

38. M. Stuiver, H. A. Polach, Discussion Reporting of ^{14}C Data. *Radiocarbon* **19**, 355–363 (1977). doi:10.1017/S0033822200003672

39. P. J. Reimer, E. Bard, A. Bayliss, J. W. Beck, P. G. Blackwell, C. B. Ramsey, C. E. Buck, H. Cheng, R. L. Edwards, M. Friedrich, P. M. Grootes, T. P. Guilderson, H. Haflidason, I. Hajdas, C. Hatté, T. J. Heaton, D. L. Hoffmann, A. G. Hogg, K. A. Hughen, K. F. Kaiser, B. Kromer, S. W. Manning, M. Niu, R. W. Reimer, D. A. Richards, E. M. Scott, J. R. Southon, R. A. Staff, C. S. M. Turney, J. van der Plicht, IntCal13 and Marine13 Radiocarbon Age Calibration Curves 0–50,000 Years cal BP. *Radiocarbon* **55**, 1869–1887 (2013). doi:10.2458/azu_js_rc.55.16947

40. J. Dabney, M. Knapp, I. Glocke, M.-T. Gansauge, A. Weihmann, B. Nickel, C. Valdiosera, N. García, S. Pääbo, J.-L. Arsuaga, M. Meyer, Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc.*

Natl. Acad. Sci. U.S.A. **110**, 15758–15763 (2013). doi:10.1073/pnas.1314445110 Medline

41. N. Rohland, M. Hofreiter, Ancient DNA extraction from bones and teeth. *Nat. Protoc.* **2**, 1756–1762 (2007). doi:10.1038/nprot.2007.247 Medline

42. M. A. Nieves-Colón, A. T. Ozga, W. J. Pestle, A. Cucina, V. Tiesler, T. W. Stanton, A. C. Stone, Comparison of two ancient DNA extraction protocols for skeletal remains from tropical environments. *Am. J. Phys. Anthropol.* **166**, 824–836 (2018). doi:10.1002/ajpa.23472 Medline

43. N. Rohland, E. Harney, S. Mallick, S. Nordenfelt, D. Reich, Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. B* **370**, 20130624 (2015). doi:10.1098/rstb.2013.0624 Medline

44. M. Meyer, M. Kircher, Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. *Cold Spring Harb. Protoc.* **2010**, pdb.prot5448 (2010). doi:10.1101/pdb.prot5448

45. M. Kircher, S. Sawyer, M. Meyer, Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* **40**, e3 (2012). doi:10.1093/nar/gkr771 Medline

46. C. Carøe, S. Gopalakrishnan, L. Vinner, S. S. T. Mak, M. H. S. Sinding, J. A. Samaniego, N. Wales, T. Sicheritz-Pontén, M. T. P. Gilbert, Single-tube library preparation for degraded DNA. *Methods Ecol. Evol.* **9**, 410–419 (2018). doi:10.1111/2041-210X.12871

47. N. Rohland, D. Reich, Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* **22**, 939–946 (2012). doi:10.1101/gr.128124.111 Medline

48. Q. Fu, A. Mittnik, P. L. F. Johnson, K. Bos, M. Lari, R. Bollongino, C. Sun, L. Giemsch, R. Schmitz, J. Burger, A. M. Ronchitelli, F. Martini, R. G. Cremonesi, J. Svoboda, P. Bauer, D. Caramelli, S. Castellano, D. Reich, S. Pääbo, J. Krause, A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr. Biol.* **23**, 553–559 (2013). doi:10.1016/j.cub.2013.02.044 Medline

49. Q. Fu, M. Hajdinjak, O. T. Moldovan, S. Constantin, S. Mallick, P. Skoglund, N. Patterson, N. Rohland, I. Lazaridis, B. Nickel, B. Viola, K. Prüfer, M. Meyer, J. Kelso, D. Reich, S. Pääbo, An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216–219 (2015). doi:10.1038/nature14558 Medline

50. M. Ghodsi, B. Liu, M. Pop, DNACLUST: Accurate and efficient clustering of phylogenetic marker genes. *BMC Bioinformatics* **12**, 271 (2011). doi:10.1186/1471-2105-12-271 Medline

51. A. Peltzer, G. Jäger, A. Herbig, A. Seitz, C. Kniep, J. Krause, K. Nieselt, EAGER: Efficient ancient genome reconstruction. *Genome Biol.* **17**, 60 (2016). doi:10.1186/s13059-016-0918-z Medline

52. M. Schubert, S. Lindgreen, L. Orlando, AdapterRemoval v2: Rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88 (2016). doi:10.1186/s13104-016-1900-2 Medline

53. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009). doi:10.1093/bioinformatics/btp324 Medline

54. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; 1000 Genome Project Data Processing Subgroup, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009). doi:10.1093/bioinformatics/btp352 Medline

55. S. Mallick, H. Li, M. Lipson, I. Mathieson, M. Gymrek, F. Racimo, M. Zhao, N. Chennagiri, S. Nordenfelt, A. Tandon, P. Skoglund, I. Lazaridis, S. Sankararaman, Q. Fu, N. Rohland, G. Renaud, Y. Erlich, T. Willems, C. Gallo, J. P. Spence, Y. S. Song, G. Poletti, F. Balloux, G. van Driem, P. de Knijff, I. G. Romero, A. R. Jha, D. M. Behar, C. M. Bravi, C. Capelli, T. Hervig, A. Moreno-Estrada, O. L. Posukh, E. Balanovska, O. Balanovsky, S. Karachanak-Yankova, H. Sahakyan, D. Toncheva, L. Yepiskoposyan, C. Tyler-Smith, Y. Xue, M. S. Abdullah, A. Ruiz-Linares, C. M. Beall, A. Di Rienzo, C. Jeong, E. B. Starikovskaya, E. Metspalu, J. Parik, R. Villems, B. M. Henn, U. Hodoglugil, R. Mahley, A. Sajantila, G. Stamatoyannopoulos, J. T. S. Wee, R. Khusainova, E. Khusnutdinova, S. Litvinov, G. Ayodo, D. Comas, M. F. Hammer, T. Kivisild, W. Klitz, C. A. Winkler, D. Labuda, M. Bamshad, L. B. Jorde, S. A. Tishkoff, W. S. Watkins, M. Metspalu, S. Dryomov, R. Sukernik, L. Singh, K. Thangaraj, S. Pääbo, J. Kelso, N. Patterson, D. Reich, The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201–206 (2016). doi:10.1038/nature18964 Medline

56. A.-S. Malaspinas, O. Lao, H. Schroeder, M. Rasmussen, M. Raghavan, I. Moltke, P. F. Campos, F. S. Sagredo, S. Rasmussen, V. F. Gonçalves, A. Albrechtsen, M. E. Allentoft, P. L. F. Johnson, M. Li, S. Reis, D. V. Bernardo, M. DeGiorgio, A. T. Duggan, M. Bastos, Y. Wang, J. Stenderup, J. V. Moreno-Mayar, S. Brunak, T. Sicheritz-Ponten, E. Hodges, G. J. Hannon, L. Orlando, T. D. Price, J. D. Jensen, R. Nielsen, J. Heinemeier, J. Olsen, C. Rodrigues-Carvalho, M. M. Lahr, W. A. Neves, M. Kayser, T. Higham, M. Stoneking, S. D. J. Pena, E. Willerslev, Two ancient human genomes reveal Polynesian ancestry among the indigenous Botocudos of Brazil. *Curr. Biol.* **24**, R1035–R1037 (2014). doi:10.1016/j.cub.2014.09.078 Medline

57. M. Rasmussen, M. Sikora, A. Albrechtsen, T. S. Korneliussen, J. V. Moreno-Mayar, G. D. Poznik, C. P. E. Zollikofer, M. P. de León, M. E. Allentoft, I. Moltke, H. Jónsson, C. Valdiosera, R. S. Malhi, L. Orlando, C. D. Bustamante, T. W. Stafford Jr., D. J. Meltzer, R. Nielsen, E. Willerslev, The ancestry and affiliations of Kennewick Man. *Nature* **523**, 455–458 (2015). doi:10.1038/nature14625 Medline

58. M. Raghavan, P. Skoglund, K. E. Graf, M. Metspalu, A. Albrechtsen, I. Moltke, S. Rasmussen, T. W. Stafford Jr., L. Orlando, E. Metspalu, M. Karmin, K. Tambets, S. Rootsi, R. Mägi, P. F. Campos, E. Balanovska, O. Balanovsky, E. Khusnutdinova, S. Litvinov, L. P. Osipova, S. A. Fedorova, M. I. Voevoda, M. DeGiorgio, T. Sicheritz-Ponten, S. Brunak, S. Demeshchenko, T. Kivisild, R. Villems, R. Nielsen, M. Jakobsson, E. Willerslev, Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87–91 (2014). doi:10.1038/nature12736 Medline

59. J. V. Moreno-Mayar, B. A. Potter, L. Vinner, M. Steinrücken, S. Rasmussen, J. Terhorst, J. A. Kamm, A. Albrechtsen, A.-S. Malaspinas, M. Sikora, J. D. Reuther, J. D. Irish, R. S. Malhi, L. Orlando, Y. S. Song, R. Nielsen, D. J. Meltzer, E. Willerslev, Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature* **553**, 203–207 (2018). doi:10.1038/nature25173 Medline

60. M. Rasmussen, S. L. Anzick, M. R. Waters, P. Skoglund, M. DeGiorgio, T. W. Stafford Jr., S. Rasmussen, I. Moltke, A. Albrechtsen, S. M. Doyle, G. D. Poznik, V. Gudmundsdottir, R. Yadav, A.-S. Malaspinas, S. S. White 5th, M. E. Allentoft, O. E. Cornejo, K. Tambets, A. Eriksson, P. D. Heintzman, M. Karmin, T. S. Korneliussen, D. J. Meltzer, T. L. Pierre, J. Stenderup, L. Saag, V. M. Warmuth, M. C. Lopes, R. S. Malhi, S. Brunak, T. Sicheritz-Ponten, I. Barnes, M. Collins, L. Orlando, F. Balloux, A. Manica, R. Gupta, M. Metspalu, C. D. Bustamante, M. Jakobsson, R. Nielsen, E. Willerslev, The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* **506**, 225–229 (2014). doi:10.1038/nature13025 Medline

61. J. Lindo, R. Haas, C. Hofman, M. Apata, M. Moraga, R. A. Verdugo, J. T. Watson, C. Viviano Llave, D. Witonsky, C. Beall, C. Warinner, J. Novembre, M. Aldenderfer, A. Di Rienzo, The genetic prehistory of the Andean highlands 7000 years BP though European contact. *Sci. Adv.* **4**, eaau4921 (2018). doi:10.1126/sciadv.aau4921 Medline

62. C. de la Fuente, M. C. Ávila-Arcos, J. Galimany, M. L. Carpenter, J. R. Homburger, A. Blanco, P. Contreras, D. Cruz Dávalos, O. Reyes, M. San Roman, A. Moreno-Estrada, P. F. Campos, C. Eng, S. Huntsman, E. G. Burchard, A.-S. Malaspinas, C. D. Bustamante, E. Willerslev, E. Llop, R. A. Verdugo, M. Moraga, Genomic insights into the origin and diversification of late maritime hunter-gatherers from the Chilean Patagonia. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E4006–E4012 (2018). doi:10.1073/pnas.1715688115 Medline

63. D. J. Kennett, S. Plog, R. J. George, B. J. Culleton, A. S. Watson, P. Skoglund, N. Rohland, S. Mallick, K. Stewardson, L. Kistler, S. A. LeBlanc, P. M. Whiteley, D. Reich, G. H. Perry, Archaeogenomic evidence reveals prehistoric matrilineal dynasty. *Nat. Commun.* **8**, 14115 (2017). doi:10.1038/ncomms14115 Medline

64. Q. Fu, C. Posth, M. Hajdinjak, M. Petr, S. Mallick, D. Fernandes, A. Furtwängler, W. Haak, M. Meyer, A. Mittnik, B. Nickel, A. Peltzer, N. Rohland, V. Slon, S. Talamo, I. Lazaridis, M. Lipson, I. Mathieson, S. Schiffels, P. Skoglund, A. P. Derevianko, N. Drozdov, V. Slavinsky, A. Tsybankov, R. G. Cremonesi, F. Mallegni, B. Gély, E. Vacca, M. R. G. Morales, L. G. Straus, C. Neugebauer-Maresch, M. Teschler-Nicola, S. Constantin, O. T. Moldovan, S. Benazzi, M. Peresani, D. Coppola, M. Lari, S. Ricci, A. Ronchitelli, F. Valentin, C. Thevenet, K. Wehrberger, D. Grigorescu, H. Rougier, I. Crevecoeur, D. Flas, P. Semal, M. A. Mannino, C. Cupillard, H. Bocherens, N. J. Conard, K. Harvati, V. Moiseyev, D. G. Drucker, J. Svoboda, M. P. Richards, D. Caramelli, R. Pinhasi, J. Kelso, N. Patterson, J. Krause, S. Pääbo, D. Reich, The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016). doi:10.1038/nature17993 Medline

65. T. S. Korneliussen, A. Albrechtsen, R. Nielsen, ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356 (2014). doi:10.1186/s12859-014-0356-4 Medline

66. G. Renaud, V. Slon, A. T. Duggan, J. Kelso, Schmutzi: Estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* **16**, 224 (2015). doi:10.1186/s13059-015-0776-0 Medline

67. F. M. Key, C. Posth, J. Krause, A. Herbig, K. I. Bos, Mining Metagenomic Data Sets for Ancient DNA: Recommended Protocols for Authentication. *Trends Genet.* **33**, 508–520 (2017). doi:10.1016/j.tig.2017.05.005 Medline

68. A. Furtwängler, E. Reiter, G. U. Neumann, I. Siebke, N. Steuri, A. Hafner, S. Lösch, N. Anthes, V. J. Schuenemann, J. Krause, Ratio of mitochondrial to nuclear DNA affects contamination estimates in ancient DNA analysis. *Sci. Rep.* **8**, 14075 (2018). doi:10.1038/s41598-018-32083-0 Medline

69. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009). doi:10.1101/gr.094052.109 Medline

70. N. Patterson, A. L. Price, D. Reich, Population structure and eigenanalysis. *PLOS Genet.* **2**, e190 (2006). doi:10.1371/journal.pgen.0020190 Medline

71. W. F. Keegan, *The People Who Discovered Columbus: The Prehistory of the Bahamas* (Univ. Press of Florida, 1992).

72. W. C. Schaffer, R. S. Carr, J. S. Day, M. P. Pateman, Lucayan-Taíno burials from Preacher's cave, Eleuthera, Bahamas. *Int. J. Osteoarchaeol.* **22**, 45–69 (2010). doi:10.1002/oa.1180

73. M. J. Berman, P. L. Gnivecki, M. P. Pateman, in *The Oxford Handbook of Caribbean Archaeology*, W. F. Keegan, C. L. Hofman, R. Rodríguez Ramos, Eds. (Oxford Univ. Press, 2013), pp. 264–280.

74. M. L. P. Hoogland, C. L. Hofman, R. Panhuysen, in *Island Shores, Distant Pasts: Archaeological and Biological Approaches to the Pre-Columbian Settlement of the Caribbean*, S. M. Fitzpatrick, A. H. Ross, Eds. (Univ. Press of Florida, 2010), pp. 148–162.

75. C. J. Beets, S. R. Troelstra, P. M. Grootes, M.-J. Nadeau, K. Borg, A. F. M. Jong, C. L. Hofman, M. L. P. Hoogland, Climate and pre-Columbian settlement at Anse à la Gourde, Guadeloupe, Northeastern Caribbean. *Geoarchaeol.* **21**, 271–280 (2006). doi:10.1002/gea.20096

76. C. L. Hofman, in *The Oxford Handbook of Caribbean Archaeology*, W. F. Keegan, C. L. Hofman, R. Rodríguez Ramos, Eds. (Oxford Univ. Press, 2013), pp. 205–220.

77. C. L. Hofman, A. J. Bright, A. Boomert, S. Knippenberg, Island Rhythms: The Web of Social Relationships and Interaction Networks in the Lesser Antillean Archipelago between 400 B.C. and A.D. 1492. *Lat. Am. Antiq.* **18**, 243–268 (2007). doi:10.2307/25478180

78. C. L. Hofman, M. L. P. Hoogland, in *Communities in Contact: Essays in Archaeology, Ethnohistory and Ethnography of the Amerindian Circum-Caribbean*, C. L. Hofman, A. van Duijvenbode, Eds. (Sidestone Press, 2011), pp. 14–44.

79. A. K. Bullen, R. P. Bullen, "The Lavoutte site : A Carib Ceremonial Center" in *Proceedings of the Third International Congress for the Study of Pre-Columbian Cultures of the Lesser Antilles* (Grenada National Museum, 1970), pp. 61–86.

80. M. L. P. Hoogland, C. L. Hofman, Expansion of the Taino cacicazgos towards the Lesser Antilles. *J. Soc. Am.* **85**, 93–113 (1999). doi:10.3406/jsa.1999.1731

81. C. L. Hofman, A. J. Bright, M. L. P. Hoogland, W. F. Keegan, Attractive Ideas, Desirable Goods: Examining the Late Ceramic Age Relationships between Greater and Lesser Antillean Societies. *J. Island Coast. Archaeol.* **3**, 17–34 (2008). doi:10.1080/15564890801984097

82. C. L. Hofman, M. L. P. Hoogland, Connecting Stakeholders: Collaborative preventive archaeology projects at sites affected by natural and/or human impacts. *Caribbean Connections* **5**, 1–31 (2016).

83. C. L. Hofman, M. L. P. Hoogland, H. L. Mickleburgh, J. E. Laffoon, D. A. Weston, M. H. Field, Life and death at precolumbian Lavoutte, Saint Lucia, Lesser Antilles. *J. Field Archaeol.* **37**, 209–225 (2012). doi:10.1179/0093469012Z.00000000020

84. P. E. Siegel, Ed., *Ancient Borinquen: Archaeology and Ethnohistory of Native Puerto Rico* (Univ. of Alabama Press, 2005).

85. R. Rodríguez Ramos, in *Ancient Borinquen: Archaeology and Ethnohistory of Native Puerto Rico*, P. E. Siegel, Ed. (Univ. of Alabama Press, 2005), pp. 1–54.

86. W. J. Pestle, "Diet and society in prehistoric Puerto Rico an isotopic approach," thesis, University of Illinois at Chicago (2010).

87. M. Rodríguez López, "Arqueología de Punta Candelero, Puerto Rico" in *Proceedings of the XIII International Congress of the International Association for Caribbean Archaeology*, E. N. Ayubi, J. B. Haviser, Eds. (IACA, 1991), pp. 605–627.

88. A. L. Curet, L. M. Stringer, Eds., *Tibes: People, Power, and Ritual at the Center of the Cosmos* (Univ. of Alabama Press, 2010).

89. W. J. Pestle, M. Colvard, Bone collagen preservation in the tropics: A case study from ancient Puerto Rico. *J. Archaeol. Sci.* **39**, 2079–2090 (2012). doi:10.1016/j.jas.2012.03.008

90. J. Orihuela León, R. A. Viera Muñoz, L. Pérez Orozco, Contribución a la cronología y la paleodieta de un individuo aborigen excavado en el sitio arqueológico El Morrillo (Matanzas, Cuba). *Cuba Arqueológica* **10**, 16–31 (2017).

91. Y. Chinique de Armas, W. M. Buhay, R. Rodríguez Suárez, S. Bestel, D. Smith, S. D. Mowat, M. Roksandic, Starch analysis and isotopic evidence of consumption of cultigens among fisher–gatherers in Cuba: The archaeological site of Canímar Abajo, Matanzas. *J. Archaeol. Sci.* **58**, 121–132 (2015). doi:10.1016/j.jas.2015.03.003

92. I. Rouse, *Archaology of the Maniabon Hills, Cuba* (Yale Univ. Press, 1942).

93. E. Tabío, E. Rey, Nueva periodización para el estudio de las comunidades aborígenes de Cuba. *Islas Universidad Central de Las Villas* **78**, 34–35 (1984).

94. J. M. Guarch Delmonte, *Estructura para las Comunidades Aborígenes de Cuba* (Ediciones Holguín, 1990).

95. A. P. Cordova, O. Arredondo, "Análisis de restos dietarios del sitio arqueológico El Mango, Río Cauto, Granma" in *Anuario de Arqueología, Centro de Arqueología y Etnología* (Editorial Academia, 1988), pp. 111–132.

96. Y. Chinique de Armas, U. M. González Herrera, W. M. Buhay, J. M. Yero Masdeu, L. M. Viera Sanfiel, M. Burchell, C. Crann, E. R. Grau González-Quevedo, M. Roksandic, Chronology of the archaeological site of Playa Del Mango, Rio Cauto, Granma, Cuba. *Radiocarbon* 10.1017/RDC.2020.52 (2020). doi:10.1017/RDC.2020.52

97. M. Pino, E. Alonso, Excavaciones en la Cueva del Perico I. *Serie Espeleológica y Carsológica* **45**, 5–32 (1973).

98. R. T. Ruíz, D. R. Hernández, M. R. de la Calle, S. M. Jaca, Estudio de los restos óseos humanos aborígenes encontrados en la Cueva del Perico I, Pinar del Río, Cuba. *Biología* **13** (1999).

99. J. Cooper, Modelling mobility and exchange in pre-Columbian Cuba: GIS led approaches to identifying pathways and reconstructing journeys from the archaeological record. *J. Caribb. Archaeol.* **3**, 122–137 (2010).

100. E. E. Tabío, E. Rey, *Prehistoria de Cuba* (Academia de Ciencias de Cuba, 1966).

101. H. Weissensteiner, D. Pacher, A. Kloss-Brandstätter, L. Forer, G. Specht, H.-J. Bandelt, F. Kronenberg, A. Salas, S. Schönherr, HaploGrep 2: Mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–W63 (2016). doi:10.1093/nar/gkw233 Medline

11.2 Supplementary Material for Manuscript B

nature ecology & evolution

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41559-018-0498-2

In the format provided by the authors and unedited.

Language continuity despite population replacement in Remote Oceania

Cosimo Posth[1,19]*, Kathrin Nägele[1,19], Heidi Colleran[2]*, Frédérique Valentin[3], Stuart Bedford[2,4], Kaitip W. Kami[2,5], Richard Shing[5], Hallie Buckley[6], Rebecca Kinaston[1,6], Mary Walworth[2], Geoffrey R. Clark[7], Christian Reepmeyer[8], James Flexner[9], Tamara Maric[10], Johannes Moser[11], Julia Gresky[12], Lawrence Kiko[13], Kathryn J. Robson[14], Kathryn Auckland[15], Stephen J. Oppenheimer[16], Adrian V. S. Hill[15], Alexander J. Mentzer[15], Jana Zech[17], Fiona Petchey[18], Patrick Roberts[17], Choongwon Jeong[1], Russell D. Gray[2], Johannes Krause[1]* and Adam Powell[1,2]*

[1]Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany. [2]Department of Linguistic and Cultural Evolution, Max Planck Institute for the Science of Human History, Jena, Germany. [3]Maison de l'Archéologie et de l'Ethnologie, CNRS, UMR 7041, Nanterre, France. [4]School of Culture, History and Language, College of Asia and the Pacific, Australian National University, Canberra, Australian Capital Territory, Australia. [5]Vanuatu Cultural Centre, Port-Vila, Vanuatu. [6]Department of Anatomy, School of Biomedical Sciences, University of Otago, Dunedin, New Zealand. [7]Archaeology and Natural History, College of Asia and the Pacific, Australian National University, Canberra, Australian Capital Territory, Australia. [8]College of Arts, Society and Education, James Cook University, Cairns, Queensland, Australia. [9]Department of Archaeology, University of Sydney, Sydney, New South Wales, Australia. [10]Service de la Culture et du Patrimoine, Punaauia, Tahiti, French Polynesia. [11]Commission for Archaeology of Non-European Cultures, German Archaeological Institute, Bonn, Germany. [12]Department of Natural Sciences, German Archaeological Institute, Berlin, Germany. [13]Solomon Islands National Museum, Honiara, Solomon Islands. [14]MRC Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, UK. [15]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK. [16]School of Anthropology and Museum Ethnography, University of Oxford, Oxford, UK. [17]Department of Archaeology, Max Planck Institute for the Science of Human History, Jena, Germany. [18]Waikato Radiocarbon Dating Laboratory, The University of Waikato , Hamilton, New Zealand. [19]These authors contributed equally: Cosimo Posth, Kathrin Nägele. *e-mail: posth@shh.mpg.de; colleran@shh.mpg.de; krause@shh.mpg.de; powell@shh.mpg.de

Archaeological Information

Talasiu, Tonga (TON001, TON002, TON004/CP30). The Talasiu site (TO-Mu-2), Tongatapu, Kingdom of Tonga, is located on the shoreline of the Fanga 'Uta Lagoon, ~2.5km south of the Nukuleka site which is regarded as the place of initial human landfall in Tonga[1]. Talasiu contains a dense shell midden deposit ~90cm thick covering some 450m^2 that includes fire features and burials[2]. In 2008, a concentration of burned and partially burned human bone eroding from a road cut was excavated, revealing a mortuary context with partially heated and incomplete skeletal remains of four individuals[3]. In 2011, new adult inhumations were again found eroding from the road cut. As the area was about to be intensively gardened a rescue archaeology project to recover human remains was directed by Frederique Valentin (CNRS) and Geoffrey Clark (ANU) in 2013-2014 and 2016 with the support of the Ministry of Internal Affairs (Kingdom of Tonga) and funded by the French Government (MEAE, Commission des fouilles à l'étranger). In total the excavations identified 19 burial contexts holding early human remains of one or more individuals.

Radiocarbon ages were obtained on human bone from articulated burials (*n*=6), coconut endocarp (*n*=5), unidentified charcoal (*n*=2) and worked shell grave goods (*n*=3). All calibrated results fall between 2,750 and 2,150 calibrated years before present (y BP, 95% probability range) with charcoal and bone results between 2,600 and 2,300y BP influenced by curve flattening resulting in wide age ranges (Hallstatt Plateau). A high-resolution chronology based on U-Th dating of coral files and AMS determinations on charcoal - a material with minimal inbuilt age - demonstrates that the Lapita period on Tongatapu spanned 2,860-2,680y BP[1]. As Lapita ceramics occur throughout the Talasiu deposits it is probable that the midden and burials at Talasiu date to ~2,700-2,600y BP[2] and are of late Lapita age. This is supported by a new U-Th result on a coral file from the Talasiu deposits as well as intact lenses of shell midden and fire features that sealed several burial contexts, which demonstrate that interments were made as the midden was accumulating.

The Talasiu burials represent the oldest human remains found so far in Polynesia and provide the first opportunity to understand the origins, health and mortuary practices of the first people to colonize the eastern islands of Remote Oceania[3-5]. Ancient DNA had previously been obtained from the right petrous bone from a single primary interment of an adult female SK10[5]. The results indicated that this female, like three other Lapita (~2,900y BP) individuals from Teouma site in Vanuatu derived from an East Asian population that no longer exists in unmixed form. The initial aDNA result suggested that later population movements must have spread Papuan ancestry in the South Pacific region after the period of Lapita colonization[5]. Ancient DNA employed in this paper was successfully obtained from two more Talasiu burials. One contained the remains of two individuals who were buried simultaneously (context SK3) in which SK3.1, an old female was sampled (TON001), and the second sample (TON002) was from a male skull (SK6) that had been reburied in an abandoned oven. Finally we report new mtDNA data of individual SK10 (from a molar TON004 and a petrous bone CP30) from whom genome-wide data was previously published[5] and assigned to haplogroup B4a1a1a (Supplementary table 11).

J09, Tongatapu, Tonga (LHA001). The J09 site is a royal tomb (*langi*) called 'Tauatonga' located in Lapaha village on Tongatapu Island just south of the Talasiu (TO-Mu-2) site[6]. In 2012, an excavation through the fill of J09 to recover charcoal to radiocarbon date a tomb, identified a burial in pre-tomb sediments. A fragment of a distal humerus was AMS dated at the Waikato Radiocarbon Dating Laboratory in New Zealand and returned an age of 955 ± 25 ^{14}C years BP (Wk-36401). The bone sample was well-preserved with a C:N ratio of 3.26 and a ^{13}C value of -15.63 indicating a diet with a significant marine contribution (marine contribution estimated as 54%). The calibrated age result of 780-550y BP (Table 1) and the burial location indicates that the individual lived during the inception of the ancient Tongan state when the Tu'i Tonga lineage began to rule the Tonga Islands – an event which was manifested by the construction of a monumental centre at Lapaha and an extensive set of maritime networks[7,8]. A tooth from the skeleton beneath J09 was sampled for aDNA (LHA001).

Rockshelter excavations of Tanna and Futuna, Vanuatu. Skeletal material from the islands of Tanna was excavated in 1963-1964 by Richard and Mary Shutler. The Shutler's were sent to the New Hebrides (it became Vanuatu at independence in 1980) as part of an initiative of the Pacific Science Congress of 1961, under the auspices of the Bishop Museum. During fieldwork on Tanna and Futuna, the Shutlers excavated a number of rockshelter sites, which contained human burials and other materials, as well as some open sites[9].

After the Shutlers' pioneering work, there was a hiatus of nearly 50 years with little or no additional archaeological fieldwork carried out on Futuna and Tanna until very recently. Archaeologists supported by

the Australian Research Council and the French Ministry of Foreign Affairs (MEAE, commission des fouilles) are currently revisiting the materials and sites excavated by the Shutlers, in addition to excavating new sites on Futuna, Tanna, and the neighbouring Polynesian Outlier Aniwa[10]. Part of this current study focuses on the long-term history of human interactions in the area of southern Vanuatu, including the skeletal, genetic, and isotopic signatures of human migration. One of the key dynamics for the two Polynesian Outliers, Futuna and Aniwa, is the timing and nature of the Polynesian settlement, presumably some time between 1,000-500y BP[11].

TaRS and Lowenpakal , Tanna, Vanuatu (TAN001 and TAN002). The skeletal material reported here from Tanna was excavated from two cave sites. One was from TaRS3 located on the west coast of Tanna Island, near the present-day village of Bethel. Tanna is a volcanic island that has been tilting to the south and east due to the active volcano Iasur. As a result, much of the west coast of the island is composed of upraised limestone reef terraces containing rockshelters[12]. TaRS3 is located on an uplifted reef terrace, has an opening approximately 27.5m wide and encompasses a cave that is 12 x 6m in area. The site was excavated completely by the Shutlers down to bedrock. A full skeleton (TAN001) was excavated from the cave. It was an extended burial in prone position, located 1.5m below the surface, with the skull facing to southwest. The Shutlers excavated a further burial from the nearby cave site of TaRS1 but an attempt to extract collagen from this skeleton was not successful..

The other Tanna sample (TAN002) comes from a 1 x 1m testpit excavation carried out in 2016 in a cave site located at Lowenpakel, at the very north coast of Tanna. The excavation was carried out as part of the new South Vanuatu Archaeological Survey program in a location seen as having high potential for early settlement. Excavation revealed deeply stratified deposits. Charcoal from a hearth feature at 1.13mbd returned a date of 900-720y BP and two dates from a lower layer (1.27-1.50mbd) returned 970-830y BP and 1,230-1,010y BP. Scattered human bone including the petrous bone investigated here were found in these lower cultural levels of the testpit. The dating of the petrous bone is much older (2,630-2,350y BP) than other dates from the site but it may originate from earlier deposits that were disturbed by later occupations. Pottery was also found at these levels which tends to lend support to the earlier date for the human bone, since it has been firmly established that on other islands in the south of Vanuatu pottery disappeared around 2,000y BP.

FuRS, Futuna, Vanuatu (FUT001, FUT002, FUT006, FUT007, FUT008). Futuna is a small island roughly 5km long that rises steeply to the highest point 666m above sea level. The island is a *makatea* (raised coral) island and presents an extensive system of rockshelters on former reef terraces. The Shutlers recorded a large number of rockshelters on Futuna, and excavated several of them[9]. The skeletal samples analyzed in this study came from rockshelters FuRS1A and FuRS12. These rockshelters are located on the limestone slopes of the northeastern Ipau district of Futuna. FuRS12 is 13.7m long by 3.6m wide. Excavations uncovered 15 inhumations buried in various positions close to the bedrock towards the back of the rockshelter, many of which were rock-lined or covered with rocks and included grave goods[13]. Samples from four adult burials 1, 7, 8-9 and 12 were investigated in this study (FUT001, FUT002, FUT007 and FUT008). FuRS1A is a roughly 12 x 6m area. It contained the buried remains of two partial individuals, one of which was analyzed here (FUT006), as well as a variety of artifacts including an adze fragment and a sandstone abrader. All five individuals are radiocarbon dated to an interval between to 970 and 1,270y BP, a time period corresponding roughly to the first major Polynesian dispersals to the east to Eastern Polynesia and to the west to Melanesia that occurred around 1,000 BP[11]. In this study we obtained genome-wide and mtDNA data from four individuals (FUT001, FUT002, FUT006, FUT007) and only mtDNA data from the upper incisor of one individual (FUT008) from FuRS12 in burial 12 and newly radiocarbon dated here to 1,376 ± 29 ^{14}C years BP (MAMS-29689).

Urupiv and Vao, Malakula, Vanuatu (MAL001, MAL002, MAL004, MAL006, MAL007, MAL008). The samples from Malakula, northern Vanuatu, come from excavations undertaken on two small islands (c. 2km^2), Uripiv and Vao, located on the north-east coast[14,15]. Excavations on these islands began in 2001 and continued intermittently until 2011 (2002-2004 on Vao; 2001-2002, 2005, 2009-2011 on Uripiv). The sites comprised deeply stratified deposits that encompassed the entire period of human occupation on the respective islands, 3,000 years from Lapita through to the Historic period. During those excavations a total of 7 burials were identified on Vao and 38 on Uripiv. This series of burials offers the rare opportunity to explore changes over time in mortuary behavior, health, diet and migration through different markers including burial features, morphological characteristics, palaeopathological indicators, isotopic data as well as ancient DNA[16,17].

Only accessible petrous bones were selected for this study preferentially sampling skeletons that were previously directly dated[17]. Together with two additional radiocarbon dates (MAL001 and MAL002) the

burials included here correspond well with their archaeological contexts. The sample MAL001 is from Vao Island while all others, MAL002, MAL004, MAL006, MAL007 and MAL008 come from Uripiv Island. MAL001, of Post-Lapita age, was retrieved from a dispersed collection of human bones in a single 1 x 1m test-pit. All of the samples analyzed from Uripiv came from in-situ burials dating to a ~500 years interval, ranging from circa 2,500 to 2,000y BP, which were uncovered during large aerial excavations. These burials represent a variety of mortuary situations. MAL002 is a near complete 18 month old infant (Burial 1) lying on its left side and back. It is of Late Lapita age, buried in the natural beach sand, 1.5m below the current ground surface. MAL004 (Burial 8) is a Late Lapita age young child, buried on its back again in the beach sand. MAL006 (Burial 15) is a Lapita burial of an infant who died in the perinatal period. The body was placed on its left side with the upper limbs extended, hands close to the face and covered with white beach sand. MAL007 (Burial 18) is a post-Lapita burial of a female adult lying in a semi-seated position, with the lower limbs tightly flexed, feet against the pelvis, in a pit dug into a black sediment rich in charcoal and coral gravel. MAL008 (Burial 23) is again a post-Lapita burial of a male adult placed in a seated position in a small sepulchral pit dug into the sand and filled up with dark sediment.

Taputapuātea site, Ra'iātea, French Polynesia (TAP001, TAP002, TAP003, TAP004). The Taputapuātea ceremonial complex on Ra'iātea island (Society Islands, French Polynesia) is of central importance in Polynesian cosmology of the Society Islands. This significance for Polynesian identity justified its inscription in 2017 to the UNESCO World Heritage List. The site is located on the east coast of the island, east of Opoa village, on the flat wide point named *Matahiratera'i*. Extending over a surface of five hectares, this ceremonial complex is dedicated to the cult of several Polynesian deities, and comprises a number of monuments and periods of construction. The great *marae* Taputapuātea dated to the 17th century[18], is surrounded by five other *marae*, including *marae* Hauviri and *marae* Hititai, along with other constructions and enclosures.

Since its first visits by early voyagers such as Joseph Banks in 1769[19], the site has been described by several archaeologists such as Emory and Sinoto[18,20] who mentioned the presence of human remains at various points of the monuments surfaces. Restorations were engaged twice[21,22]. In 1994-1995 the Centre Polynésien des Sciences Humaines[22] recovered burials and concentrations of human remains on *marae* Taputapuātea, Hauviri and Hititai. The studied remains (TAP001, TAP002, TAP003, TAP004) represent individuals deposited at the monuments during funerary ceremonies. Inhumations with the head placed in the vicinity of the main upraised stone is a main mortuary feature at *marae* Hauviri (TAP003) while skulls secondary deposited seems to distinguish *marae* Hititai (TAP004). These events, based on direct dating of human remains, occurred at the earliest in the beginning of the eighteenth century (1710-1730 AD / 1800-1950 AD, Wk-40993) at *marae* Hauviri and more certainly during the nineteenth century (1810-1950 AD, Wk-40995). Three individuals (TAP002, TAP003, TAP004) provided mtDNA and genome-wide data (Table 1) while only mtDNA was obtained from individual TAP001 *(marae* Hauviri), whose mtDNA sequence was assigned to haplogroup B4a1a1 (Supplementary table 11).

Ria-rockshelter, Malaita, Solomon Islands (MAI002, MAI003). The archaeological investigations at the dwelling site and burial place 'Ria-rockshelter' within the research project "Settlement History of Melanesia – Prehistory of the Solomon Islands" are conducted in close cooperation with the National Museum Honiara and the Ministry of Culture and Tourism, Solomon Islands. The rock overhang 'Ria' is located in the province East Are in southern Malaita and was formed by an isolated natural limestone cliff. The overhang could have served as a shelter for one to two families. The archaeological potential of the site was suspected during a survey in the region in 2011 and finally confirmed through archaeological excavations between 2013 and 2017.

The 'Ria-rockshelter' shows evidence of human presence in prehistoric times. The excavations under the shelter disclosed cultural deposits and features and a large collection of knapped stone tools, shells and faunal remains. In the upper layers besides several fire places a pavement made from accurately placed pebbles (*hau poro*) – all affected by heat - was unearthed, possibly indicating an earth oven (*umu*). The set of lithic consists of a great variety of flake adzes, serrated and denticulated pieces, unmodified flakes and cores. As ornaments diverse shell pectoral pendants were found. In the shelter's rear two extended supine burials (Individuals I and II) were discovered under the pebble pavement. Individuals I (MAI001) is an adult of around 25-30 years old assigned to female sex while Individual II (MAI002) is a child of around 11-13 years old. During the excavation in 2015 the remains of a third individual (Individual III, MAI003) came to light, an infant circa 4-5 yeas old. Radiocarbon dating was performed for all three remains providing the following results: Individual I (MAI001): 502 ± 37 ^{14}C years BP (Erl-20179), Individual II (MAI002): 460 ± 30 ^{14}C years BP (Beta-433422) and Individual III (MAI003): 640 ± 30 ^{14}C years BP (Beta-451930). Beside mtDNA and nuclear DNA data from MAI002 (Table 1), an mtDNA sequence assigned to haplogroup B4a1a1a was retrieved for individual MAI003 (Supplementary table 11).

Radiocarbon Dating and Isotopic Analyses

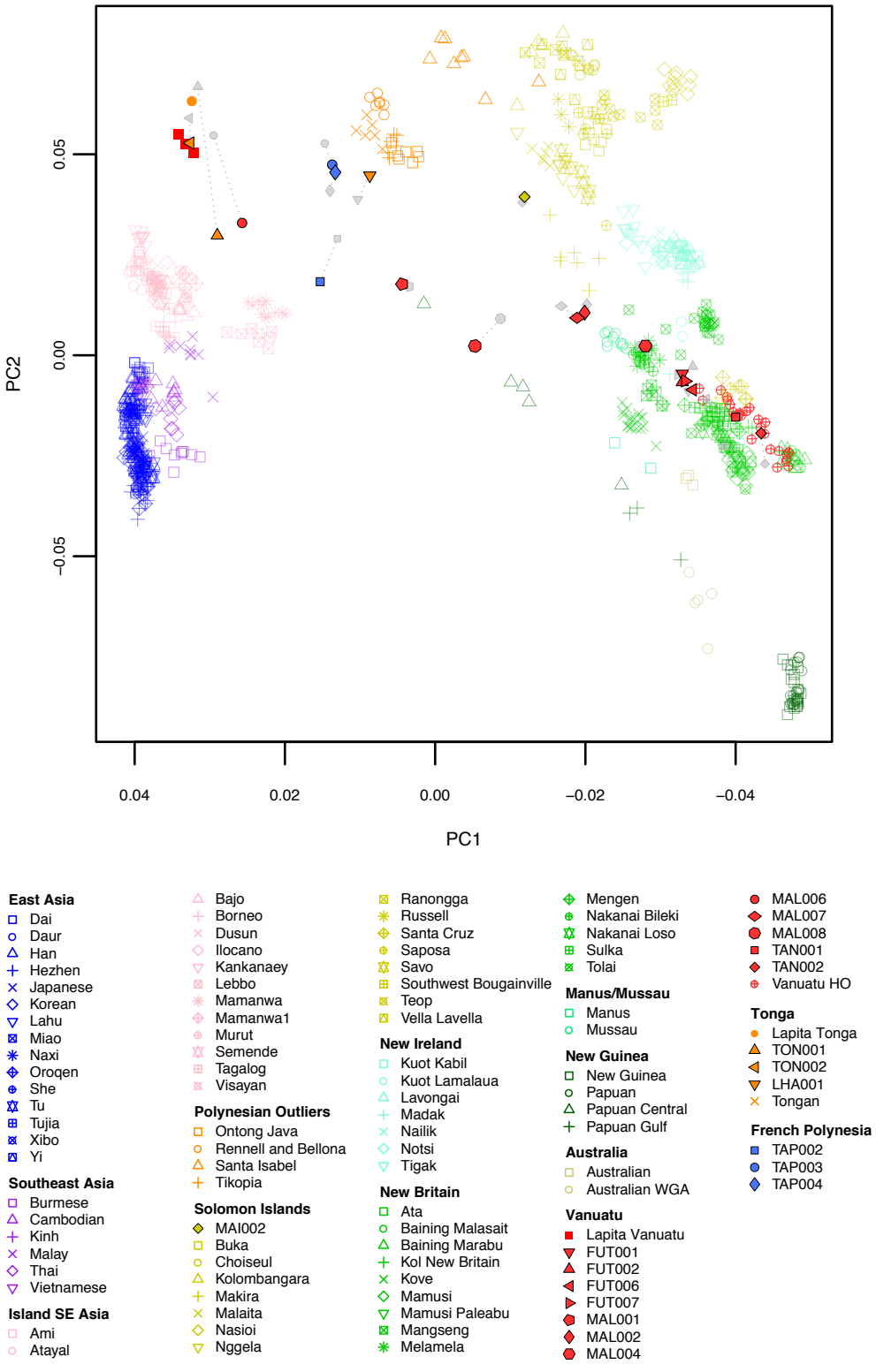
Dates on sampled individuals were undertaken at three different laboratories (University of Heidelberg [MAMS-], University of Waikato [Wk-]) and Beta Analytic [Beta-]), either newly generated (13 dates) or previously published (5 dates) as indicated in Supplementary table 2 for individuals who provided genome-wide data. For individuals who provided only mtDNA data 2 new non-calibrated radiocarbon dates (FU008 and MAI003) are reported within the Supplementary Text section of each site. Supplementary table 2 lists the skeletal elements subjected to stable isotope and dating analyses as well as the protocol followed by each dating laboratory. Prior to conversion to a calendar age it was important to determine whether there were any dietary offsets that could influence the calibrated ages. To enable comparison across the different individuals and results from different dating laboratories stable carbon and nitrogen isotopes from the sampled individuals were measured (Supplementary table 12). Samples of bone or dentine were collected from the skeletal element available for each individual. Bone samples were subsequently cleaned using an air abrasive system with 5 μm aluminium oxide powder and then crushed. Dentine was obtained as a powder from the crown of the tooth using a diamond-tipped drill. Collagen was then extracted following standard procedures[23]. Approximately 500mg of pre-cleaned bone was demineralized in 10ml aliquots of 0.5M HCL at 4°C, with changes of acid until CO_2 stopped evolving. The residue was then rinsed three times in deionized water before being gelatinized in pH3 HCl at 75°C for 48 hours. The resulting solution was filtered, with the supernatant then being lyophilized over a period of 24 hours.

Purified collagen samples (1mg) were analysed at the Department of Archaeology, Max Planck Institute for the Science of Human History in duplicate by EA-IRMS on a ThermoFisher Elemental Analyser coupled to a ThermoFisher Delta V Advantage Mass Spectrometer via a ConFloIV system. Accuracy was determined by measurements of international standard reference materials within each analytical run. These were USGS40 $\delta^{13}C_{raw}$ = -26.4 ± 0.1, $\delta^{13}C_{true}$ = -26.4 ± 0.0, $\delta^{15}N_{raw}$ = -4.4 ± 0.1, $\delta^{15}N_{true}$ = -4.5 ± 0.2; IAEA N2 $\delta^{15}N_{raw}$ = 20.2 ± 0.1, $\delta^{15}N_{true}$ = 20.3 ± 0.2; IAEA C6 $\delta^{13}C_{raw}$ = -10.9 ± 0.1, $\delta^{13}C_{true}$ = -10.8 ± 0.0. In addition, a homogenised bovid bone extracted and analysed within the same batch as the samples produced the following values; $\delta^{13}C$ = -20.1 ± 0.1; $\delta^{15}N$ = 6.8 ± 0.2. The overall mean value among 30 separate extracts of this bone sample produced values of $\delta^{13}C$ = -20.2 ± 0.1; $\delta^{15}N$ = 6.8 ± 0.2.
In all cases, stable isotope ratios are expressed as 'per mil' or parts per thousand (‰). The difference in the $^{13}C/^{12}C$ ratio between the sample and the internationally defined standard AIR (atmospheric air) in ‰ units is referred to as $\delta^{13}C$, and $\delta^{15}N$ refers to the difference in $^{15}N/^{14}N$ ratio between the sample and the internationally defined standard, VPDB (Vienna Peedee Belemnite Limestone). The reported ratios are calculated using the equation: $\delta X = ((R_{sample} - R_{standard})/ R_{standard}) \times 1000$. The full $\delta^{13}C$ and $\delta^{15}N$ results for all samples analysed can be found in Supplementary table 12. All of the samples have C:N ratios within the acceptable range (2.9-3.6)[24] and collagen yields above 1%[25] (Supplementary table 12).
Based on $\delta^{13}C$ and $\delta^{15}N$ measurements of modern flora and fauna from the Pacific[26-28] a percent marine carbon (%MarineC) contribution to the diet was estimated for the human bone (this ranged between 3 and 67%). All radiocarbon dates were calibrated using OxCal v4.2[29] with a mixture of the Marine13 and Intcal13 curves[30] as determined by the calculated %MarineC. A marine reservoir correction (ΔR) value was applied based on pre-AD 1950 shell values for each island group[31].

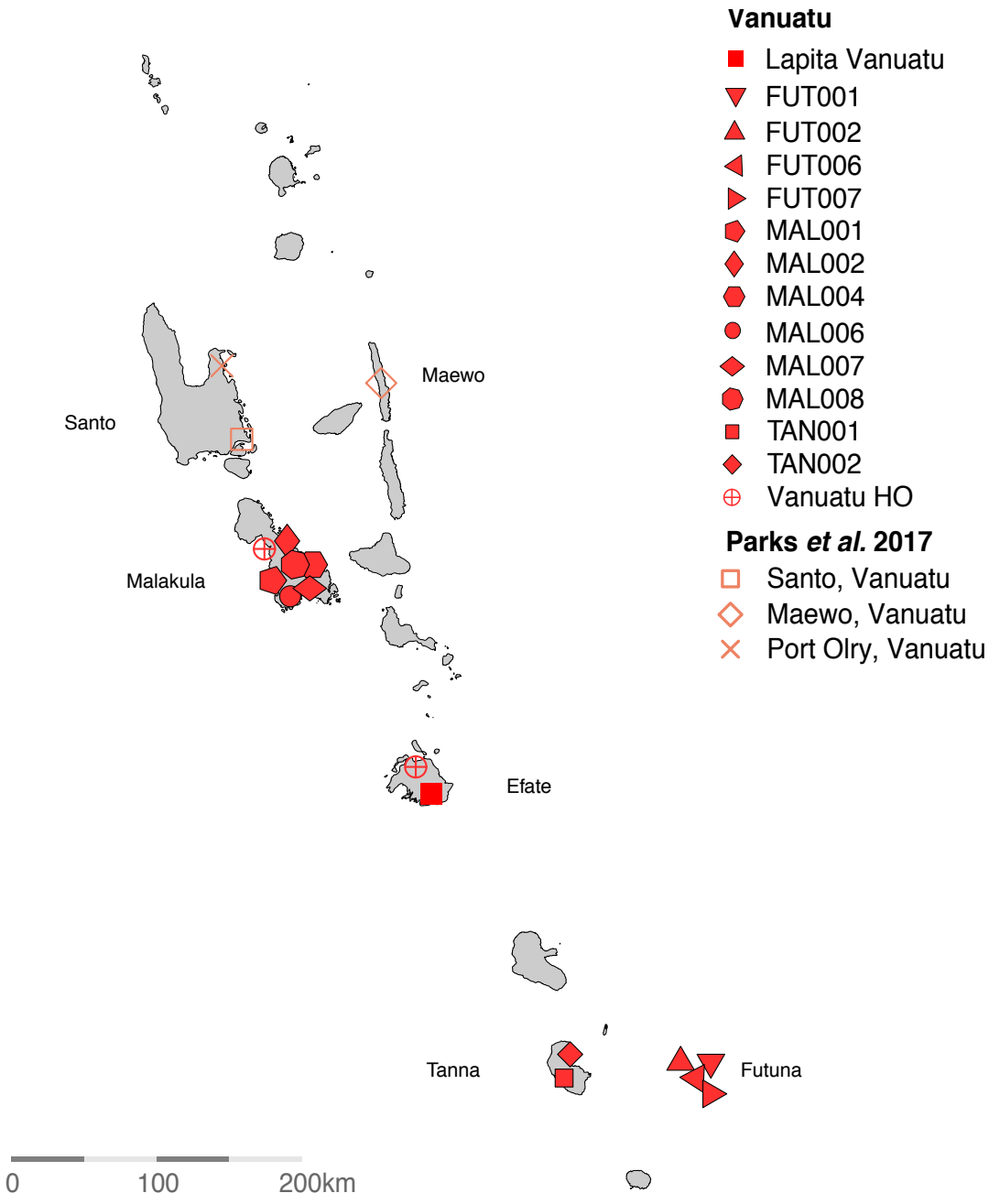
The radiocarbon determinations, and their calibrations before and after the application of the resulting reservoir and dietary effects can be found in Supplementary table 2 and Table 1, respectively. For the majority of the samples this correction does not make a substantial different. However, in all cases the corrected calibrated dates with the reservoir correction applied during calibration have been used in further data interpretation within the paper.

Supplementary Figures

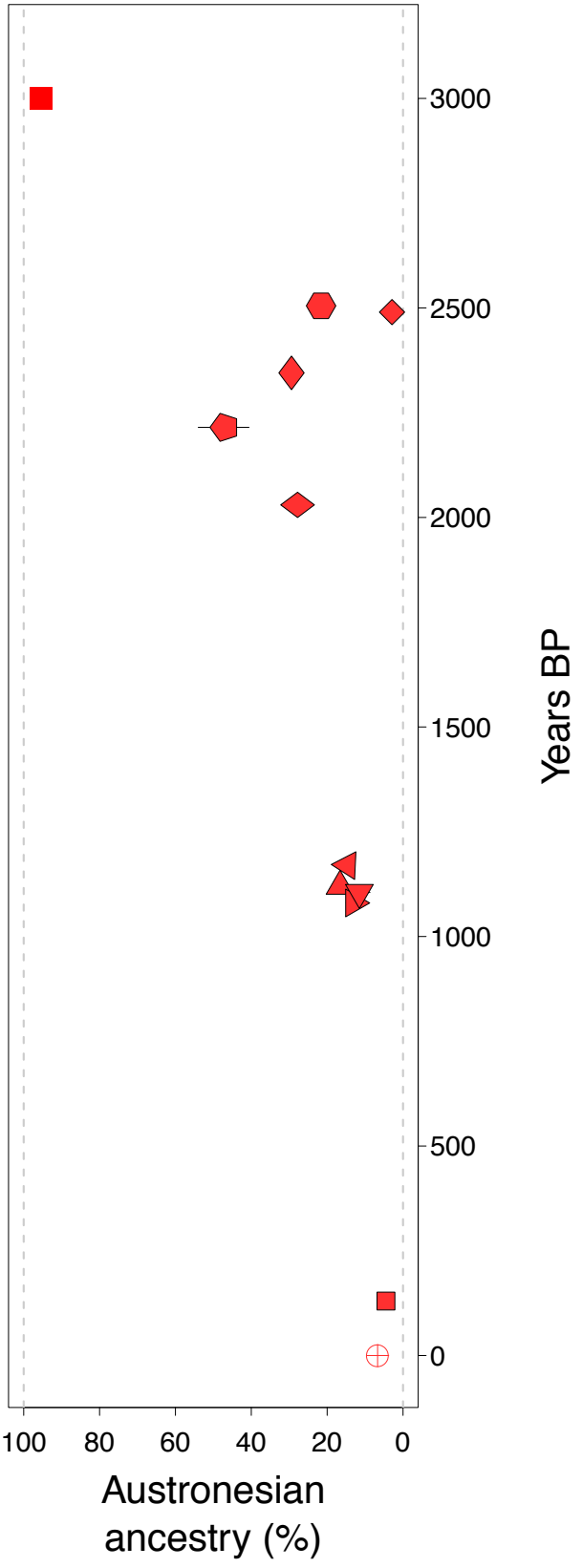
Supplementary figure 1. Principal component analyses of modern-day East Asian and Oceanian populations genotyped on the *Affymetrix Human Origins Array*, with ancient individuals projected before (color filled symbols) and after (grey filled symbols) the restriction to damaged DNA fragments, supposedly of ancient origin. Each individual's pre- and post-filtering symbol is connected with a grey dotted line.



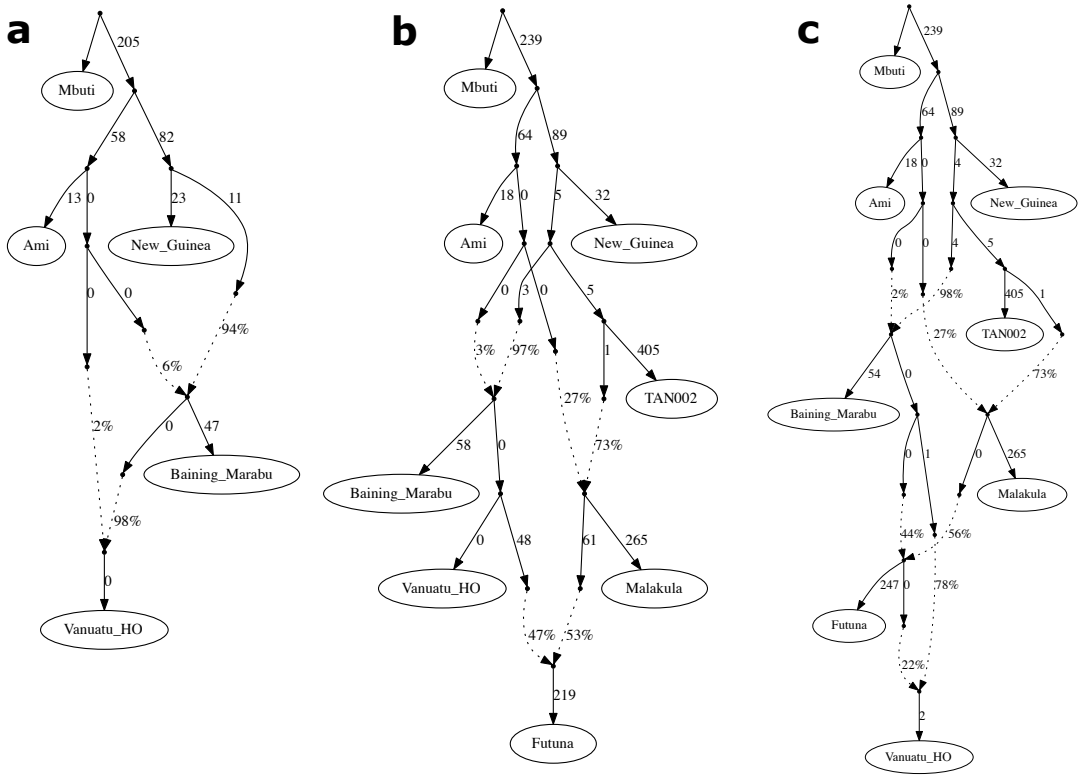
Supplementary figure 2. Map of Vanuatu, showing approximate locations of ancient individuals and modern sampling locations from this study and Parks *et al.*[32].



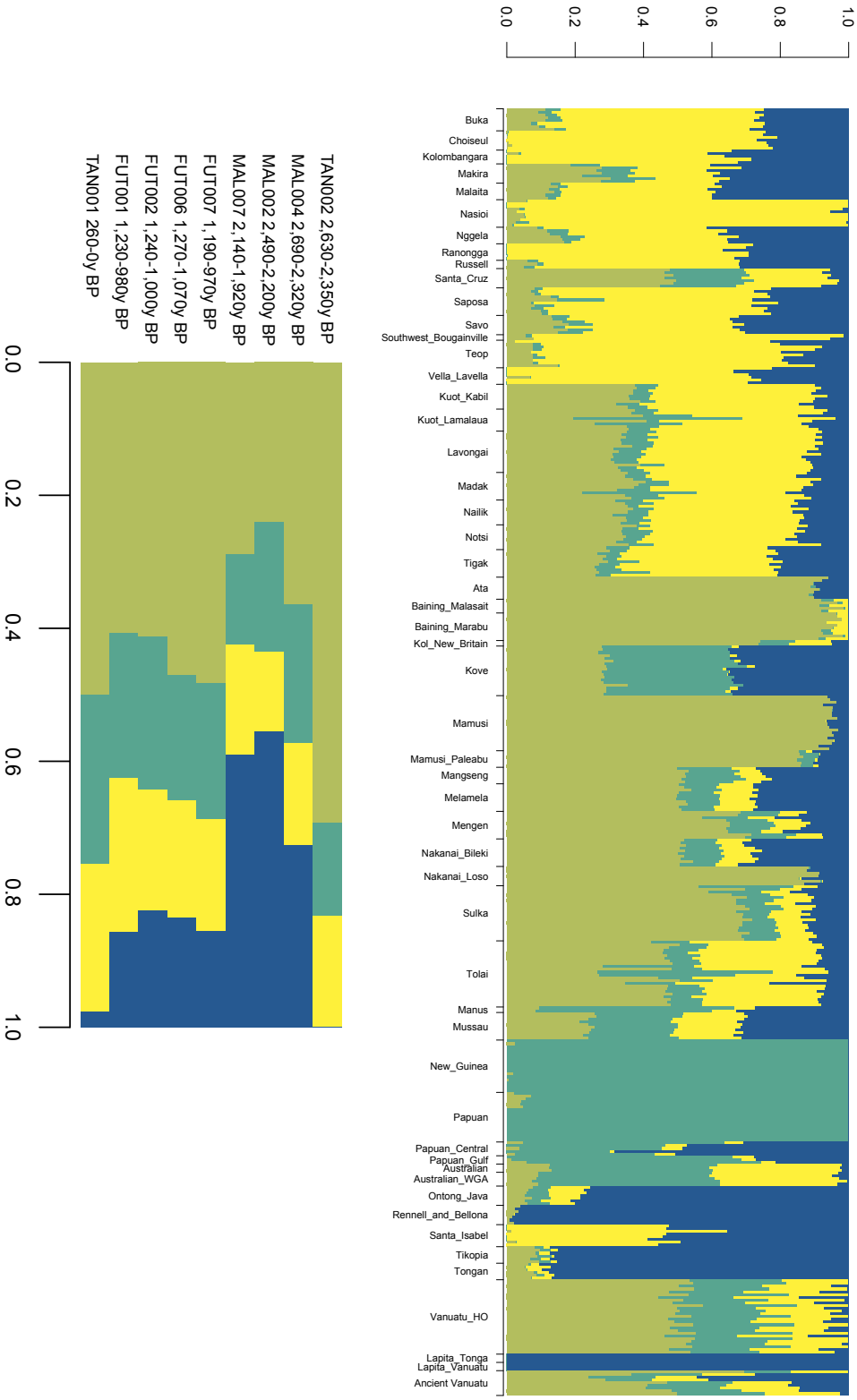
Supplementary figure 3. *qpAdm* analyses on the down-sampled SNPs of the HO dataset modeling the Austronesian ancestry proportion (represented by Ami population) in ancient individuals (filled symbols) and 27 present-day individuals grouped together (unfilled symbol) from Vanuatu. Standard errors are shown as black lines if larger than the sample symbol (legend in Fig. 1).



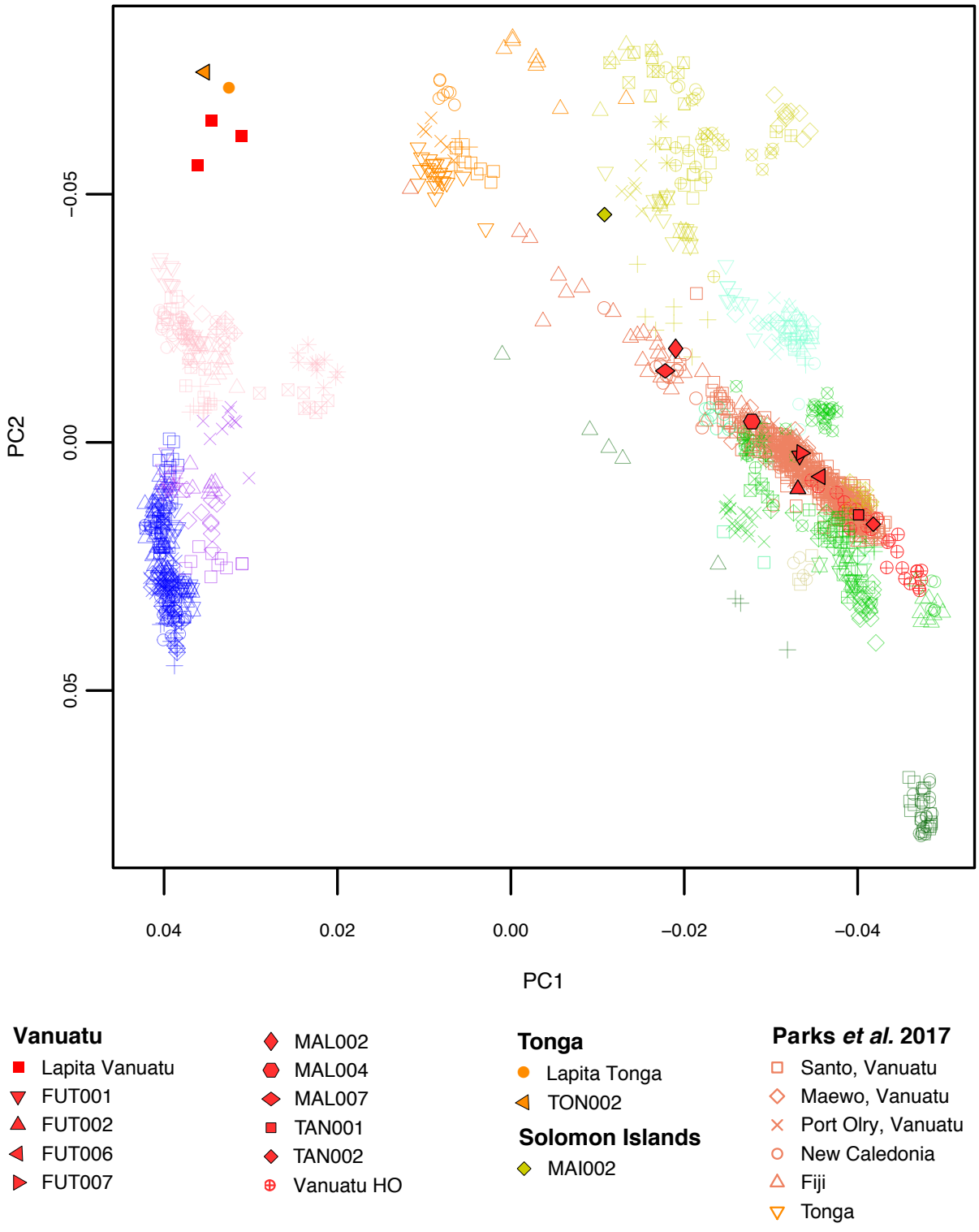
Supplementary figure 4. *qpGraph* analyses modelling population relationships for present-day HO Vanuatu individuals. Vanuatu HO is modelled as **(a)** admixed between modern populations related to Ami and Baining Marabu (Z: 2.3); **(b)** a sister group of Baining Marabu (Z: 6.0); and **(c)** admixed between Baining Marabu and an ancient Futuna-related lineage (Z: 5.2).



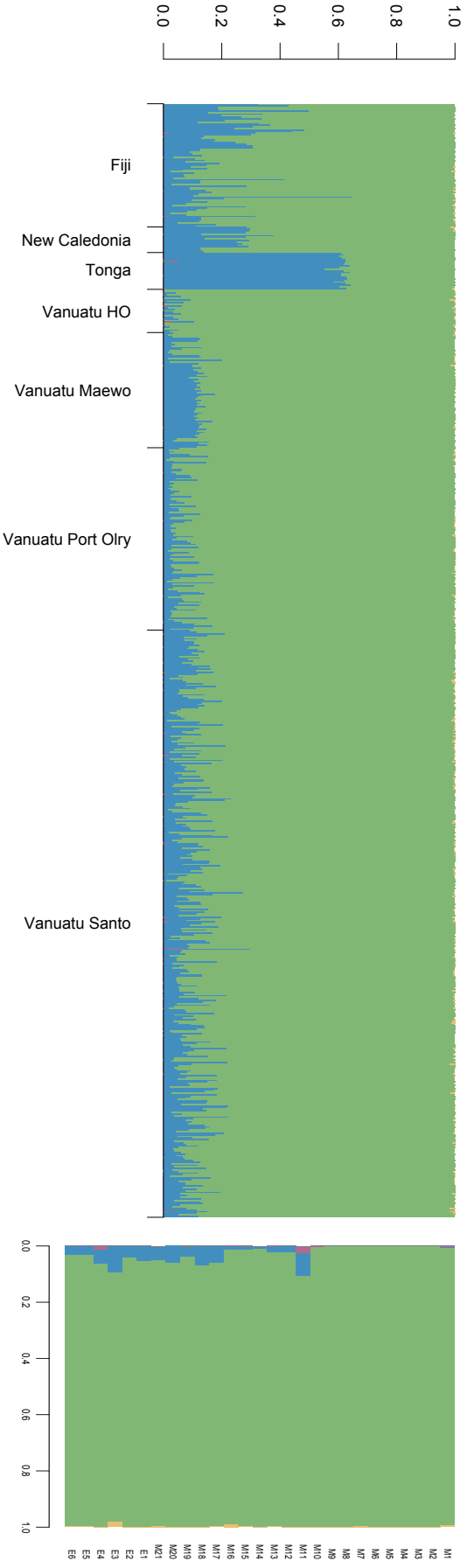
Supplementary figure 5. Unsupervised *ADMIXTURE* analyses (*K*=4) on a regional selection of HO data comprising 454 modern-day Near and Remote Oceanian individuals, 4 previously published Lapita-associated individuals[5] and 9 ancient individuals from Vanuatu (both in the right figure and enlarged on the left).



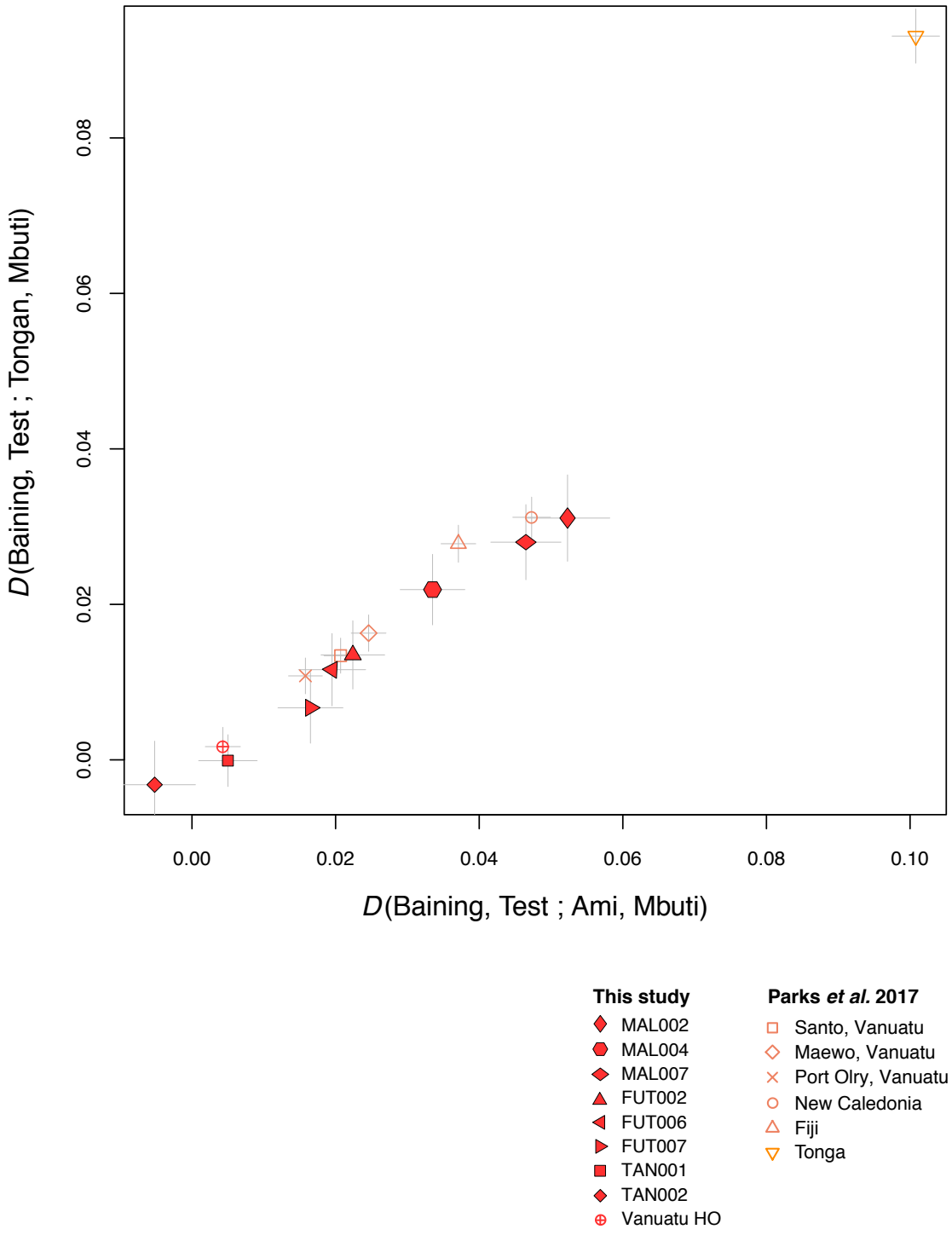
Supplementary figure 6. 15 ancient and 669 modern-day individuals from New Caledonia, Vanuatu, Fiji and Tonga from Parks *et al.*[32] projected onto principal components 1 and 2, computed using the overlapping ~50k SNPs of the HO populations reported in Fig. 1 and Supplementary figure 1.



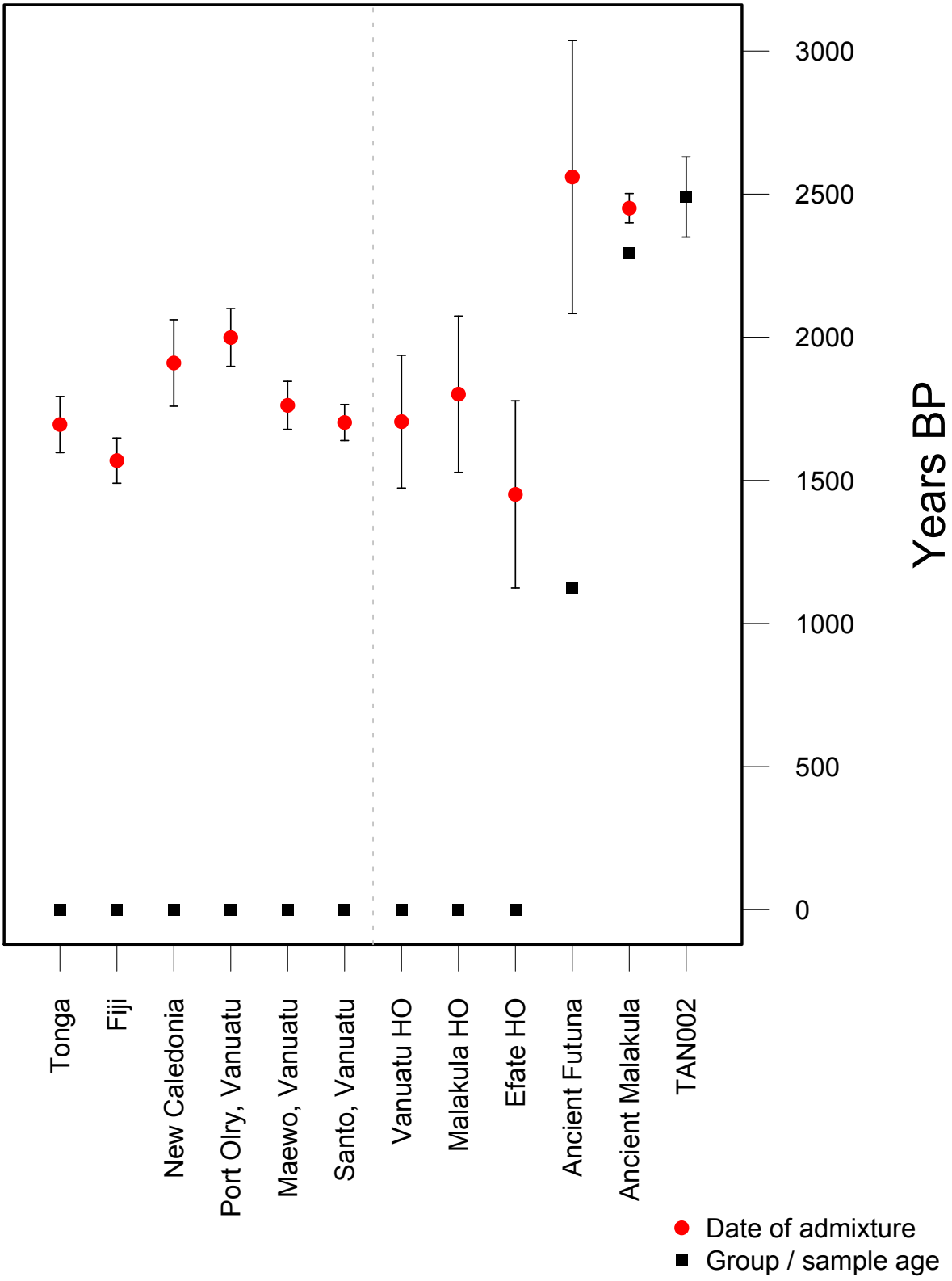
Supplementary figure 7. Unsupervised *ADMIXTURE* analyses on the ~50k SNPs overlapping between 669 individuals from Parks *et al.*[32] and 27 Vanuatu (Malakula and Efate) individuals genotyped on the HO array (both above and enlarged at the bottom).



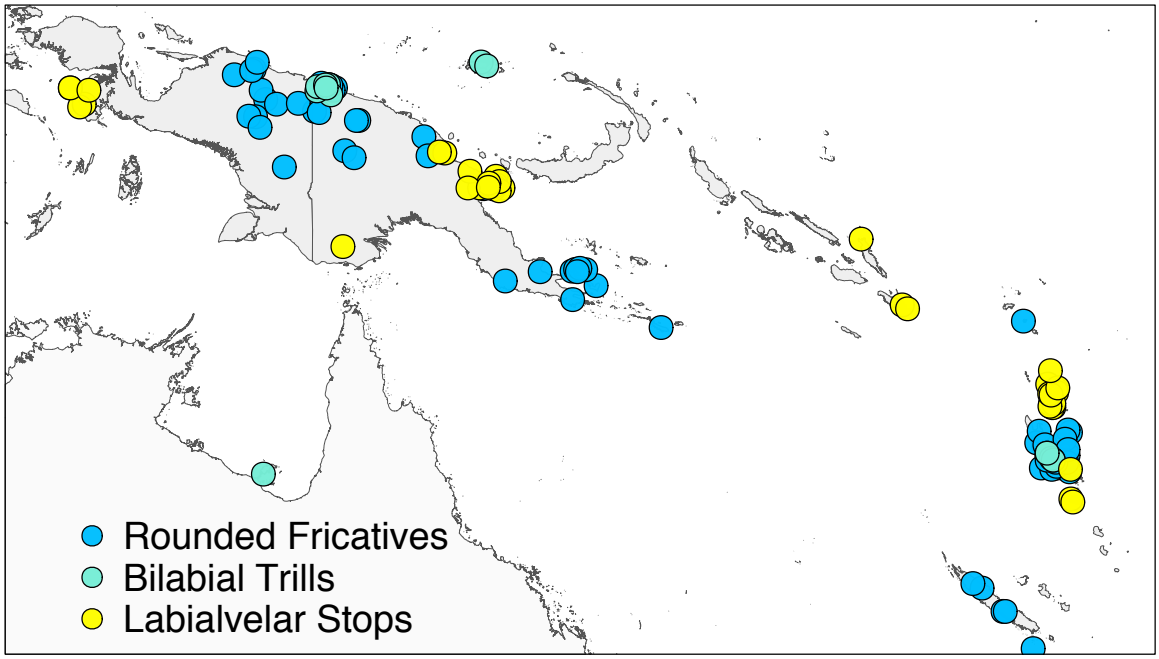
Supplementary figure 8. *D*-statistics in the form *D(Baining, Test ; Ami, Mbuti)* and *D(Baining, Test ; modern Tongan, Mbuti)* plotted against each other, where *Test* is *Fiji*, *Tonga*, *Maewo (Vanuatu)*, *Port Olry (Vanuatu)*, *Santo (Vanuatu)* and *New Caledonia* populations from Parks *et al.*[32], modern-day Vanuatu HO individuals and ancient Malakula, Futuna and Tanna individuals from this study. Standard errors for each point are shown in both dimensions as gray lines.



Supplementary figure 9. *ALDER* analyses estimating the date of Papuan and East Asian admixture, converted into years with a generation time of 28.1 years. Populations investigated derive from Parks *et al.*[32] (left of the dashed gray line), the newly HO genotyped Vanuatu individuals, either grouped together Vanuatu HO (n=27) or divided as Malakula HO (n=21) and Efate HO (n=6), and ancient Futuna (n=3) and Malakula (n=3). Standard error bars are shown for date estimates, while sample ages for the two ancient groups (Futuna and Malakula) are averaged radiocarbon dating confidence interval (CI) midpoints. As the earliest ancient Vanuatu individual with unadmixed Near Oceanian ancestry, *TAN002* is included for age comparison, with error bar indicating the 95.4% radiocarbon dating CI.



316 **Supplementary figure 10.** Rare linguistic features shared between Papuan languages of Near Oceania and
317 the languages of Vanuatu.
318
319



Supplementary Tables

Supplementary table 1. Archaeological information of 19 individuals providing genome-wide data reported in this study.

Sample Name	Country, Island	Site	Burial	Archaeological assignment	Latitude	Longitude
FUT001	Vanuatu, Futuna	FURS 12	Burial 7	n/a	19°31'10.80"S	170°13'33.98"E
FUT002	Vanuatu, Futuna	FURS 12	Burial 8-9	n/a	19°31'10.80"S	170°13'33.98"E
FUT006	Vanuatu, Futuna	FURS1A	Burial 1	n/a	19°31'15.01"S	170°13'48.23"E
FUT007	Vanuatu, Futuna	FURS 12	Burial 4	n/a	19°31'10.80"S	170°13'33.98"E
LHA001	Tonga, Tongatapu	Lapaha	J09	n/a	21°10'35.67"S	175°06'55.75"W
MAI002	Solomons, Malaita	Ria Cave	Individual II, RS1	n/a	9°15'15.5"S	161°13'21.7"E
MAL001	Vanuatu, Malakula	Vao	Burial 7	Post-Lapita	15°54'3.00"S	167°18'16.71"E
MAL002	Vanuatu, Malakula	Uripiv	Burial 1	Late Lapita	16°04'25.97"S	167°26'52.03"E
MAL004	Vanuatu, Malakula	Uripiv	Burial 8	Late Lapita	16°04'25.97"S	167°26'52.03"E
MAL006	Vanuatu, Malakula	Uripiv	Burial 15	Lapita	16°04'25.97"S	167°26'52.03"E
MAL007	Vanuatu, Malakula	Uripiv	Burial 18	Post-Lapita	16°04'25.97"S	167°26'52.03"E
MAL008	Vanuatu, Malakula	Uripiv	Burial 23	Post-Lapita	16°04'25.97"S	167°26'52.03"E
TAN001	Vanuatu, Tanna	TaRS 3	Burial I	n/a	19°33'22.36"S	169°16'56.51"E
TAN002	Vanuatu, Tanna	Lowenpakal	TP5	Late Lapita	19° 19' 59"S	169°20'37"E
TAP002	French Polynesia, Ra'iatea	Taputapuatea	Taputapuatea complex	n/a	16°50'10.50"S	151°21'30.86"W
TAP003	French Polynesia, Ra'iatea	Taputapuatea	Marae Hauviri	n/a	16°50'10.50"S	151°21'30.86"W
TAP004	French Polynesia, Ra'iatea	Taputapuatea	Marae Hititai	n/a	16°50'10.50"S	151°21'30.86"W
TON001	Tonga, Tongatapu	Talasiu	Sk3.1	Late Lapita	21°10'37.63"S	175°06'32.68"W
TON002	Tonga, Tongatapu	Talasiu	Sk6	Late Lapita	21°10'37.63"S	175°06'32.68"W

Supplementary table 2. Radiocarbon dating details for the 19 individuals with nuclear DNA data analyzed here. * indicates that date is not direct but from associated archaeological layer.

Sample Name	absolute dating AMS (uncal)	Lab code-number date	C:N	C [%]	Collagen [%]	Material dated	Protocol	Publication
FUT001	1284 ± 20	MAMS-29775	2.9	36.2	3.7	L petrous	Ultrafiltered gelatin	New
FUT002	1377 ± 20	MAMS-29686	3	35.9	3.1	R petrous	Ultrafiltered gelatin	New
FUT006	1306 ± 20	Wk-44199	3.35	42.57	0.91	L scapula	Ultrafiltered gelatin	New
FUT007	1303 ± 20	MAMS-29688	3	35.5	1.1	R petrous	Ultrafiltered gelatin	New
LHA001	965 ± 25	Wk-36401	3.26	43.61	0.67	R Humerus	Ultrafiltered gelatin	New
MAI002	460 ± 30	Beta-433422	n/a	n/a	n/a	L humerus	Alkali Collagen extraction	New
MAL001	2320 ± 23	MAMS-29692	n/a	n/a	1.7	L petrous	Ultrafiltered gelatin	New
MAL002	2482 ± 26	MAMS-29693	n/a	n/a	0.5	L petrous	Ultrafiltered gelatin	New
MAL004	2515 ± 28	Wk-30882	3.4	42.71	0.9	Rib	Ultrafiltered gelatin	Kinaston et al.2014
MAL006	2608 ± 30	Wk-27489	3.4	43.9	0.4	Rib	Ultrafiltered gelatin	Kinaston et al.2014
MAL007	2111 ± 30	Wk-30883	3.3	42.88	0.8	Foot phalanx	Ultrafiltered gelatin	Kinaston et al.2014
MAL008	2310 ± 33	Wk-30885	3.3	42.93	1.7	Rib	Ultrafiltered gelatin	Kinaston et al.2014
TAN001	228 ± 20	MAMS-29690	3.2	39.8	0.5	L petrous	Ultrafiltered gelatin	New
TAN002	2610 ± 17, 2471 ± 17	MAMS-31124, Wk-46423	3,4	16,5	0,4	R petrous	Ultrafiltered gelatin	New
TAP002	236 ± 18	MAMS-30075	4,1	45,1	4,1	Molar	Ultrafiltered gelatin	New
TAP003	318 ± 18	MAMS-30076	3,8	43,3	3,7	Molar	Ultrafiltered gelatin	New
TAP004	257 ± 19	MAMS-30077	3,8	40,7	1,6	Molar	Ultrafiltered gelatin	New
TON001	2594 ± 20*	WK-41883*	3.4	45.3	0.4	Fibula SK10	Ultrafiltered gelatin	Skoglund et al. 2016
TON002	2594 ± 20*	WK-41883*	3.4	45.3	0.4	Fibula SK10	Ultrafiltered gelatin	Skoglund et al. 2016

Supplementary table 3. Ancient DNA statistics for *1240K* capture libraries. Library type column refers to UDGhalf (UDGh) and non-UDG (nUDG) treatments. Endogenous DNA is reported before capture as the percentage of DNA fragments mapping against the human reference sequence *hg19* (End. DNA Shotgun). All subsequent values refer to the *1240K* capture libraries. End. DNA Capture indicates the percentage of reads overlapping the targeted *1240K* capture SNPs.

Sample Name	Country, Island	Library type	End. DNA Shotgun (%)	# of Raw Reads	Dedup Mapped Reads	Duplication Factor	End. DNA Capture (%)	DMG 1st Base 5' (%)	Average length (bp)
FUT001	Vanuatu, Futuna	nUDG	6.12	41,392,080	4,668,165	3.96	29.2	17.4	57.4
FUT002	Vanuatu, Futuna	nUDG+UDGh	9.36	70,233,017	10,666,919	3.18	29.6	11.3	60.4
FUT006	Vanuatu, Futuna	UDGhalf	8.56	91,420,714	3,487,664	5.94	28.8	13.4	51.5
FUT007	Vanuatu, Futuna	nUDG+UDGh	8.12	46,308,005	4,525,691	4.35	27.0	21.4	57.0
LHA001	Tonga, Tongatapu	UDGh	0.24	24,149,883	967,978	3.48	8.7	26.9	46.0
MAI002	Solomons, Malaita	nUDG	13.96	26,305,390	11,947,705	1.25	32.7	26.9	67.7
MAL001	Vanuatu, Malakula	nUDG	0.07	66,233,021	285,473	4.18	1.1	51.5	51.5
MAL002	Vanuatu, Malakula	nUDG+UDGh	1.54	42,073,018	2,241,947	3.98	13.0	37.9	53.2
MAL004	Vanuatu, Malakula	nUDG+UDGh	3.23	66,970,084	8,752,963	2.44	20.9	37.9	55.0
MAL006	Vanuatu, Malakula	nUDG	0.04	39,240,200	114,561	7.07	1.4	34.9	56.7
MAL007	Vanuatu, Malakula	nUDG+UDGh	1.74	41,715,211	2,453,036	4.33	17.6	35.1	51.5
MAL008	Vanuatu, Malakula	nUDG	0.11	35,130,215	242,876	5.30	2.4	45.2	49.5
TAN001	Vanuatu, Tanna	UDGh	4.19	74,290,450	4,300,891	3.62	28.9	3.44	53.6
TAN002	Vanuatu, Tanna	nUDG+UDGh	0.62	50,924,133	2,079,462	4.09	11.6	39.5	55.8
TAP002	Tahiti, Raiatea	noUDG	0.07	11,069,616	244,934	3.57	5.5	29.2	67.6
TAP003	Tahiti, Raiatea	nUDG+UDGh	0.31	3,745,562	534,951	1.59	16.1	29.3	67.9
TAP004	Tahiti, Raiatea	nUDG	0.05	17,335,921	167,739	3.93	2.7	30.8	63.0
TON001	Tonga, Tongatapu	nUDG	0.49	15,071,679	829,569	3.01	9.9	31.9	57.4
TON002	Tonga, Tongatapu	nUDG	2.97	37,681,096	3,378,089	3.52	20.0	41.0	52.0

Supplementary table 4. Nuclear contamination estimate for the X-chromosome of male individuals performed only on UDGhalf libraries when both library types were available.

Sample Name	Country, Island	Contamination X chromosomes	Std. err.	X-chr SNPs covered twice
FUT006	Vanuatu, Futuna	0.013	6.50E-03	902
FUT007	Vanuatu, Futuna	0.008	6.62E-03	548
MAL004	Vanuatu, Malakula	0.006	3.35E-03	2071
TAN001	Vanuatu, Tanna	0.012	4.73E-03	1671
TAN002	Vanuatu, Tanna	0.019	2.05E-02	138
TAP002	Tahiti, Raiatea	0.091	5.60E-02	84
TAP003	Tahiti, Raiatea	0.021	2.59E-02	99
TAP004	Tahiti, Raiatea	0.060	6.91E-02	37
TON002	Tonga, Tongatapu	0.080	1.38E-02	1220

Supplementary table 5. Details on present-day individuals from Malakula (M#) and Efate (E#) genotyped for the *Human Origins* array (*n*=27) and shotgun sequenced (*n*=8).

Sample Name	Country	Island	Sex	SNPs HO	End. DNA Shotgun (%)	Mean coverage 1240K	SNPs 1240K
M1	Vanuatu	Malakula	Male	594,008			
M2	Vanuatu	Malakula	Male	597,009			
M3	Vanuatu	Malakula	Male	596,726			
M4	Vanuatu	Malakula	Male	596,673			
M5	Vanuatu	Malakula	Male	597,093			
M6	Vanuatu	Malakula	Male	596,165			
M7	Vanuatu	Malakula	Male	596,300			
M8	Vanuatu	Malakula	Male	596,455	43.7	2.64	1,115,485
M9	Vanuatu	Malakula	Male	595,743	24.8	1.05	756,253
M10	Vanuatu	Malakula	Male	596,725	14.9	0.72	531,519
M11	Vanuatu	Malakula	Male	596,699	32.3	2.03	985,442
M12	Vanuatu	Malakula	Female	596,345			
M13	Vanuatu	Malakula	Male	596,197	47.6	1.92	950,843
M14	Vanuatu	Malakula	Female	596,300			
M15	Vanuatu	Malakula	Male	596,276			
M16	Vanuatu	Malakula	Male	596,955			
M17	Vanuatu	Malakula	Male	596,821			
M18	Vanuatu	Malakula	Male	596,532	22.6	1.43	869,874
M19	Vanuatu	Malakula	Female	596,404			
M20	Vanuatu	Malakula	Male	595,483			
M21	Vanuatu	Malakula	Female	596,265	12.7	0.62	535,583
E1	Vanuatu	Efate	Male	596,506			
E2	Vanuatu	Efate	Female	596,302			
E3	Vanuatu	Efate	Male	596,610			
E4	Vanuatu	Efate	Female	591,273			
E5	Vanuatu	Efate	Male	596,073	52.3	3.04	1,146,888
E6	Vanuatu	Efate	Female	596,160			

345 **Supplementary table 6.** Number of individuals retained and removed for each population in the Parks *et*
346 *al.* dataset[32] based on a threshold of non-local ancestry above 2% estimated in *ADMIXTURE* analyses
347 (Supplementary figure 7).
348

Population	Number individuals retained	Number individuals removed
Vanuatu_Port_Olry	114	3
Vanuatu_Maewo	72	11
Vanuatu_Santo	366	44
Fiji	78	3
New_Caledonia	16	15
Tonga	23	9
Total	669	85

349

Supplementary table 7. *qpWave* analyses on the HO dataset, to test whether the ancient (*n=8*) and modern (*n*=27) Vanuatu individuals are consistent with deriving from two streams of admixture represented by Papuan and Ami, using the following populations as outgroup: Mbuti, Denisovan, Sardinian, English, Yakut, Chukchi, Mala, Japanese, Ju_hoan_North, Mixe, Onge, Yoruba. In all cases rank *n-1* cannot be rejected (p>0.05).

Vanuatu individuals	**f4rank: 1**
TAN002	0.40
MAL002	0.92
MAL004	0.47
MAL007	0.44
FUT002	0.43
FUT007	0.09
FUT006	0.57
TAN001	0.36
Modern Vanuatu HO	0.08

Supplementary table 8. *qpAdm* analyses on the *1240K* capture dataset, for the three Vanuatu Lapita individuals from Skoglund *et al.*[5], ancient Vanuatu individuals from this study (*n*=10), and the shotgun sequenced present-day individuals from Malakula and Efate grouped together (*n*=8).

Ancient Vanuatu	Austronesian ancestry autosome	STD error	Austronesian ancestry Xchr	STD error	Xchr-autosome Austronesian ancestry
Lapita_Vanuatu	0.970	0.039	-	-	-
TAN002	0.009	0.024	0.101	0.129	0.092
MAL004	0.222	0.022	0.719	0.072	0.497
MAL002	0.31	0.026	0.407	0.369	0.097
MAL001	0.459	0.066	-	-	-
MAL007	0.297	0.023	0.505	0.07	0.208
FUT006	0.139	0.021	0.595	0.134	0.456
FUT002	0.169	0.022	0.01	0.047	-0.159
FUT001	0.130	0.028	-	-	-
FUT007	0.112	0.021	0.134	0.131	0.022
TAN001	0.046	0.019	0.039	0.12	-0.007
Modern Vanuatu shotgun	0.05	0.009	0.24	0.037	0.19

Supplementary table 9. *D-statistics* in the form *D*(*Pop1*, *Pop2* ; *Pop3*, *Outgroup*). Three individuals are grouped together for the analyses of Futuna (*FUT002*, *FUT006*, *FUT007*) and Malakula (*MAL002*, *MAL002*, *MAL007*). *TAN002* affinity to Baining compared to Papuan New Guinea populations can be caused neither by shared Austronesian (Ami) ancestry nor by Denisovan differential admixture.

Pop1	Pop2	Pop3	Outgroup	Dstat	Zscore	SNPs	Interpretation
Baining_Marabu	New_Guinea	Futuna	Mbuti	0.0122	4.335	454736	Ancient Vanuatu individuals have higher affinity to Baining_Marabu and Baining_Malasait than New_Guinea
Baining_Marabu	New_Guinea	Malakula	Mbuti	0.011	3.876	415565	
Baining_Marabu	New_Guinea	TAN002	Mbuti	0.0136	2.923	94557	
Baining_Marabu	New_Guinea	Lapita_Vanuatu	Mbuti	0.0071	2.432	247385	
Baining_Marabu	New_Guinea	TAN001	Mbuti	0.0069	1.87	336854	
Baining_Malasait	New_Guinea	Malakula	Mbuti	0.0158	4.937	412532	
Baining_Malasait	New_Guinea	Futuna	Mbuti	0.0144	4.623	451250	
Baining_Malasait	New_Guinea	TAN002	Mbuti	0.0192	3.796	93940	
Baining_Malasait	New_Guinea	TAN001	Mbuti	0.0139	3.509	334551	
Baining_Malasait	New_Guinea	Lapita_Vanuatu	Mbuti	0.0055	1.615	245945	

Pop1	Pop2	Pop3	Outgroup	Dstat	Zscore	SNPs	Interpretation
TAN002	New_Guinea	Baining_Malasait	Mbuti	0.0141	2.125	93940	Baining_Malasait and Baining_Marabu are genetically closer to TAN002 than New_Guinea with more than 2 standard deviations
TAN002	New_Guinea	Baining_Marabu	Mbuti	0.0121	1.966	94557	
TAN002	Baining_Marabu	New_Guinea	Mbuti	-0.0015	-0.251	94557	
TAN002	Baining_Malasait	New_Guinea	Mbuti	-0.0051	-0.816	93940	

Pop1	Pop2	Pop3	Outgroup	Dstat	Zscore	SNPs	Interpretation
Lapita_Vanuatu	New_Guinea	Ami	Mbuti	0.1384	35.165	247494	TAN002 forms a clade with New_Guinea in comparisons to Ami. All later Vanuatu individuals show higher affinity to Ami.
Malakula	New_Guinea	Ami	Mbuti	0.0504	15.353	415781	
Futuna	New_Guinea	Ami	Mbuti	0.0306	9.845	455021	
Baining_Marabu	New_Guinea	Ami	Mbuti	0.0098	4.252	592591	
TAN001	New_Guinea	Ami	Mbuti	0.0144	3.809	337036	
Baining_Malasait	New_Guinea	Ami	Mbuti	0.0099	3.733	586950	
TAN002	New_Guinea	Ami	Mbuti	0.0039	0.704	94608	

Pop1	Pop2	Pop3	Outgroup	Dstat	Zscore	SNPs	Interpretation
TAN001	New_Guinea	Denisovan	Mbuti	0.0041	0.779	336968	Equal Denisovan admixture levels in TAN002 and New_Guinea
TAN002	New_Guinea	Denisovan	Mbuti	-0.0038	-0.494	94589	
Baining_Malasait	New_Guinea	Denisovan	Mbuti	-0.0027	-0.839	586824	
Baining_Marabu	New_Guinea	Denisovan	Mbuti	-0.0056	-1.905	592465	
Malakula	New_Guinea	Denisovan	Mbuti	-0.0106	-2.582	415692	
Futuna	New_Guinea	Denisovan	Mbuti	-0.0107	-2.635	454929	
Lapita_Vanuatu	New_Guinea	Denisovan	Mbuti	-0.0454	-7.641	247446	

Pop1	Pop2	Pop3	Outgroup	Dstat	Zscore	SNPs	Intepretation
LHA001pmd	Tongan	Choiseul	Mbuti	-0.0307	-2.937	19290	Modern-day Tongans genetically closer to some modern-day Solomon populations than a 780-550y BP Tongan individual
LHA001pmd	Tongan	Savo	Mbuti	-0.0302	-3.038	19290	

Supplementary table 10. *qpAdm* analyses modeling present-day Tongan individuals as a two-way admixture between Ami and present-day and ancient Solomon Islanders. Mbuti, Denisovan, Sardinian, English, Yakut, Chukchi, Mala, Japanese, Ju_hoan_North, Mixe, Onge, Yoruba are used as outgroup populations with the addition in subsequent runs of Papuan and Papuan plus Baining Marabu, respectively. Models given in red are rejected at rank *n-1* ($p<0.05$).

In-group: Ami +	Outgroups	Outgroups + Papuan	Outgroups + Papuan + Baining_Marabu
Makira	0.5181	0.2742	0.3211
Malaita	0.5429	0.0703	0.0957
MAI002	0.9068	0.7190	0.6037
Nasioi	0.1680	0.0004	0.0005
Choiseul	0.2384	0.0001	0.0001

Supplementary table 11. Ancient DNA statistics for mtDNA capture libraries. mt/nuclear DNA ratio refers to the relative proportion of mtDNA molecules compared to nuclear DNA in shotgun sequencing. Library type column refers to UDGhalf (UDGh) and non-UDG (nUDG) treatments. All subsequent values refer to mtDNA capture libraries. "End. DNA Capture" indicates the percentage of reads overlapping the targeted mtDNA reference sequence (rCRS). Consensus sequence assembly with different likelihood filters (q) and contamination estimate was performed with *schmutzi* and haplogroup assignment with *HaploGrep2*.

Sample name	mt/nuclear DNA ratio	Library type	# of Raw Reads	Dedup Mapped Reads	End. DNA Capture (%)	Mean coverage	DMG 1st Base 5' (%)	Average length (bp)	Contamination estimate	Haplogroup	q filter
FUT001.A0101	55	nUDG	2,442,852	35,441	51.8	115.6	22.3	54.0	0.01 (0-0.02)	P1d2a	q30
FUT002.A0102	45	UDGh	3,525,324	36,623	49.1	132.6	6.2	60.0	0.01 (0-0.02)	M28b1	q30
FUT006.A0101	65	nUDG	727,518	31,897	58.5	105.0	32.0	54.5	0.02 (0.01-0.03)	P1d2a	q20
FUT007.A0101,2	57	nUDGh	46,308,005	3630	0.011	11.8	23.6	53.9	0.01 (0-0.02)	M28b1	q0
FUT008.A0101	986	nUDG	629,122	17,252	52.0	51.2	36.4	49.2	0.02 (0.01-0.03)	P1d2a	q30
LHA001.A0101	58	nUDG	192,550	5,519	11.3	17.2	19.9	51.5	0.03 (0.02-0.04)	B4a1a1	q30
MAI002.A0101	74	nUDG	1,828,034	166,900	43.3	628.2	32.2	62.4	0.01 (0-0.02)	B4a1a1a	q30
MAI003.A0101	n/a	nUDG	507,558	6,524	9.2	24.5	31.0	62.3	0.04 (0.03-0.05)	B4a1a1a	q30
MAL001.A0101	180	nUDG	463,490	4,359	26.8	13.0	53.0	49.3	0.03 (0.02-0.04)	B4a1a1	q30
MAL001.A0102		nUDG	401,588	3,519	21.3	10.4	55.7	48.9	0.03 (0.02-0.04)		
MAL001.A0103		nUDG	461,714	3,254	18.4	9.5	55.9	48.4	0.03 (0.02-0.04)		
MAL002.A0102	67	UDGh	632,870	15,061	45.6	46.3	22.5	50.9	0.01 (0-0.02)	B4a1a1a	q30
MAL004.A0101	31	nUDG	759,638	51,837	50.8	160.4	50.8	51.3	0.01 (0-0.02)	B4a1a1a	q30
MAL006.A0101	247	nUDG	703,402	3,279	24.1	10.5	45.7	53.3	0.04 (0.03-0.05)	B4a1a1a11b	q30
MAL006.A0102		nUDG	2,285,598	3,837	21.7	12.4	47.5	53.6	0.04 (0.03-0.05)		
MAL006.A0103		nUDG	1,982,940	4,002	24.6	13.2	41.3	54.5	0.04 (0.03-0.05)		
MAL007.A0102	99	UDGh	1,083,436	33,703	48.7	102.7	24.9	50.5	0.01 (0-0.02)	B4a1a1a	q30
MAL008.A0101	n/a	nUDG	590,796	2,353	21.7	6.6	51.9	46.7	0.09 (0.07-0.11)	B4a1a1a	q30
MAL008.A0102		nUDG	511,046	1,715	24.0	4.9	44.7	46.9	0.09 (0.07-0.11)		
MAL008.A0103		nUDG	1,686,192	1,628	14.9	4.6	41.6	47.3	0.09 (0.07-0.11)		
TAN001.A0101	63	nUDG	801,460	62,114	49.1	213.0	22.3	56.8	0.01 (0-0.02)	P1d1	q30
TAN002.A0102	36	UDGh	1,034,876	12,490	36.9	40.3	20.7	53.5	0.01 (0-0.02)	Q2a	q30
TAP001.A0101	34,880	nUDG	487,136	37,437	31.6	152.1	25.5	67.3	0.02 (0.01-0.03)	B4a1a1	q30
TAP002.A0101	4,545	nUDG	628,286	54,537	40.6	210.6	34.7	64.0	0.02 (0.01-0.03)	B4a1a1m1	q30
TAP003.A0102	2,596	UDGh	657,580	51,692	40.2	200.4	8.3	64.2	0.01 (0-0.02)	B4a1a1c	q30
TAP004.A0101	16,114	nUDG	3,130,100	171,911	33.7	675.8	22.9	65.1	0.01 (0-0.02)	B4A1a1+16126	q30
TAP004.A0102		UDGh	1,626,690	126,500	47.9	485.7	6.7	63.6	0.01 (0-0.02)		
TON001.A0101	48	nUDG	439,574	7,305	17.9	21.8	46.0	49.4	0.06 (0.04-0.08)	B4a1a1a	q30
TON001.A0102		UDGh	812,702	4,566	18.4	13.6	47.1	49.4	0.06 (0.04-0.08)		
TON002.A0101	59	nUDG	537,478	29,963	43.2	88.3	46.3	48.8	0.02 (0.01-0.03)	B4a1a1	q30
TON002.A0102		UDGh	3,321,464	44,330	55.1	132.3	46.3	49.4	0.02 (0.01-0.03)		
TON004.A0101	368	nUDG	455,234	851	13.5	2.5	40.4	47.9	0.11 (0.08-0.14)	B4a1a1a	q30
CP30	n/a	nUDG	1,357,664	22,386	36.2	63.1	46.9	46.7	0.02 (0.01-0.03)		

Supplementary table 12. Stable $\delta^{13}C$ and $\delta^{15}N$ measurements on collagen from samples analyzed in this study including duplicate analysis and averages of each sample, C:N ratios, collagen yields, and % of marine protein in an individual's diet based on equations developed by Petchey *et al* [33].

Sample	$\delta^{13}C$ i	$\delta^{13}C$ ii	$\delta^{13}C$ av	SD	$\delta^{15}N$ i	$\delta^{15}N$ ii	$\delta^{15}N$ av	SD	C:N i	C:N ii	C:N av	Collagen yield	Diet 1 % marine	Diet 2 % marine
FUT001	-18.3	-18.4	**-18.4**	0.1	9.0	8.9	**9.0**	0.1	3.1	3.1	**3.1**	**3.7%**	**16.0**	**28.9**
FUT002	-16.5	-16.5	**-16.5**	0.0	11.9	12.0	**12.0**	0.1	3.1	3.1	**3.1**	**3.1%**	**24.0**	**37.8**
FUT006	-18.6	-18.5	**-18.6**	0.1	9.8	9.6	**9.7**	0.1	3.5	3.4	**3.5**	**10.0%**	**35.0**	**50.0**
FUT007	-17.6	-17.6	**-17.6**	0.0	10.5	10.4	**10.4**	0.1	3.1	3.1	**3.1**	**1.1%**	**33.3**	**48.1**
MAI002	-20.4	-20.4	**-20.4**	0.0	9.6	9.6	**9.6**	0.0	3.1	3.3	**3.2**	**3.1%**	**44.6**	**60.7**
MAL001	-17.4	-17.5	**-17.5**	0.1	11.4	11.2	**11.3**	0.1	3.4	3.4	**3.4**	**1.3%**	**-3.9**	**6.8**
MAL002	-16.8	-16.7	**-16.8**	0.0	11.6	11.6	**11.6**	0.0	3.4	3.5	**3.5**	**2.1%**	**12.7**	**25.2**
MAL004	-17.2	-17.3	**-17.3**	0.1	9.8	10.0	**9.9**	0.2	3.5	3.7	**3.6**	**3.0%**	**27.4**	**41.6**
MAL006	-15.3	-14.9	**-15.1**	0.3	11.8	11.9	**11.8**	0.1	3.4	3.5	**3.5**	**2.3%**	**31.0**	**45.6**
MAL007	-17.0	-17.0	**-17.0**	0.0	9.0	8.8	**8.9**	0.1	3.1	3.2	**3.2**	**6.6%**	**-2.5**	**8.3**
MAL008	-15.3	-15.5	**-15.4**	0.1	10.1	9.8	**10.0**	0.2	3.5	3.3	**3.4**	**1.9%**	**32.3**	**47.0**
TAN001	-18.7	-18.8	**-18.7**	0.0	8.8	8.8	**8.8**	0.0	3.4	3.4	**3.4**	**2.3%**	**48.9**	**65.4**
TAN002	-18.4	-18.4	**-18.4**	0.0	7.0	7.0	**7.0**	0.0	3.5	3.6	**3.6**	**3.1%**	**12.5**	**25.0**
TAP002	-16.4	-16.5	**-16.4**	0.1	12.6	12.6	**12.6**	0.0	3.0	3.0	**3.0**	**4.1%**	**16.0**	**28.9**
TAP003	-15.1	-15.0	**-15.0**	0.1	14.6	14.6	**14.6**	0.0	3.0	3.0	**3.0**	**3.7%**	**-7.5**	**2.8**
TAP004	-15.9	-16.0	**-16.0**	0.1	14.3	14.2	**14.2**	0.1	3.0	3.0	**3.0**	**1.6%**	**30.0**	**44.4**
TON001	-15.4	-15.7	**-15.5**	0.2	12.2	12.2	**12.2**	0.0	3.2	3.3	**3.3**	**6.8%**	**46.0**	**62.2**
TON002	-16.6	-16.7	**-16.7**	0.1	11.1	11.1	**11.1**	0.0	3.5	3.5	**3.5**	**4.2%**	**21.5**	**35.0**

Supplementary References

1. Burley, D., Barton, A., Dickinson, W. R., Connaughton, S. P. & Taché, K. Nukuleka as a founder colony for West Polynesian settlement: New insights from recent excavations. *Journal of Pacific Archaeology* **1**, 128-144 (2010).
2. Clark, G., Grono, E., Ussher, E. & Reepmeyer, C. Early settlement and subsistence on Tongatapu, Kingdom of Tonga: Insights from a 2700–2650calBP midden deposit. *Journal of Archaeological Science: Reports* **3**, 513-524 (2015).
3. Valentin, F. & Clark, G. Early Polynesian mortuary behaviour at the Talasiu site, Kingdom of Tonga. *Journal of Pacific Archaeology* **4**, 1-14 (2013).
4. Valentin, F., Detroit, F., Spriggs, M. J. & Bedford, S. Early Lapita skeletons from Vanuatu show Polynesian craniofacial shape: Implications for Remote Oceanic settlement and Lapita origins. *Proc Natl Acad Sci U S A* **113**, 292-297, doi:10.1073/pnas.1516186113 (2016).
5. Skoglund, P. *et al.* Genomic insights into the peopling of the Southwest Pacific. *Nature* **538**, 510-513, doi:10.1038/nature19844 (2016).
6. Clark, G., Burley, D. & Murray, T. Monumentality and the development of the Tongan maritime chiefdom. *Antiquity* **82**, 994-1008 (2008).
7. Clark, G. Chiefly Tombs, Lineage History, and the Ancient Tongan State. *The Journal of Island and Coastal Archaeology* **11**, 326-343 (2016).
8. Clark, G. *et al.* Stone tools from the ancient Tongan state reveal prehistoric interaction centers in the Central Pacific. *Proceedings of the National Academy of Sciences* **111**, 10491-10496 (2014).
9. Shutler, M. E., Shutler Jr, R. & Bedford, S. Further detail on the archaeological explorations in the Southern New Hebrides, 1963–1964. *Fifty years in the field: Essays in honor and celebration of Richard Shutler Jrs archaeological career*, 189-206 (2002).
10. Flexner, J. L. *et al.* Preliminary results of the South Vanuatu Archaeological Survey: Cultural landscapes, excavation and radiocarbon dating. *Asian Perspectives* (in press).
11. Horsburgh, K. & McCoy, M. D. Dispersal, Isolation, and Interaction in the Islands of Polynesia: A Critical Review of Archaeological and Genetic Evidence. *Diversity* **9**, 37 (2017).
12. Carney, J. & Macfarlane, A. *Geology of Tanna, Aneityum, Futuna and Aniwa.* (New Hebrides Government, 1979).
13. Valentin, F., Spriggs, M., Bedford, S. & Buckley, H. Vanuatu mortuary practices over three millennia: Lapita to the early European contact period. *Journal of Pacific Archaeology* **2**, 49-65 (2011).
14. Bedford, S. The timing and nature of Lapita colonisation in Vanuatu: the haze begins to clear. *Pacific Archaeology: assessments and prospects*, 147-158 (2003).
15. Bedford, S. in *From Southeast Asia to the Pacific: archaeological perspectives on the Austronesian expansion and the Lapita cultural complex* (Academia Sinica, 2007).
16. Bedford, S., Buckley, H., Valentin, F., Tayles, N. & Longga, N. F. Lapita burials, a new Lapita cemetery and post-Lapita burials from Malakula, northern Vanuatu, Southwest Pacific. *Journal of Pacific Archaeology* **2**, 26-48 (2011).
17. Kinaston, R. *et al.* Diet and human mobility from the Lapita to the early historic period on Uripiv Island, Northeast Malakula, Vanuatu. *PloS one* **9**, e104071 (2014).
18. Emory, K. P. & Sinoto, Y. H. *Preliminary Report on the Archeological Investigations in Polynesia: Field Work in the Society and Tuamotu Islands, French Polynesia, and American Samoa in 1962, 1963, 1964*. (Bernice P. Bishop Museum, 1965)
19. Banks, J. & Beaglehole, J. C. *The Endeavour Journal of Joseph Banks, 1768-1771*. Vol. 1 (Trustees of the Public Library of New South Wales in association with Angus and Robertson, 1962).
20. Emory, K. Stone remains in the Society Islands. *Bernice Puahi Bishop Museum* **Bulletin 116**, 182 (1933).
21. Sinoto, Y. H. Restauration de Marae aux Iles de la Société. *Bulletins de la Société des Etudes Océaniennes* 14, 7-8 (1969).
22. Navarro, M., Tchong, J. & L., B. Mise en valeur du site de Taputapuatea, Opoa-Raiatea. *Centre polynésien des Sciences Humaines «Te Anacaharau»* (1995).
23. Richards, M. P. & Hedges, R. E. Stable isotope evidence for similarities in the types of marine foods used by Late Mesolithic humans at sites along the Atlantic coast of Europe. *Journal of Archaeological Science* **26**, 717-722 (1999).
24. Deniro, M. J. Postmortem Preservation and Alteration of Invivo Bone-Collagen Isotope Ratios in Relation to Paleodietary Reconstruction. *Nature* **317**, 806-809, doi:DOI 10.1038/317806a0 (1985).
25. Ambrose, S. H. Isotopic analysis of paleodiets: methodological and interpretive considerations. *Food and nutrition in history and anthropology (USA)* (1993).
26. Leach, F., Quinn, C., Morrison, J. & Lyon, G. The use of multiple isotope signatures in reconstructing prehistoric human diet from archaeological bone from the Pacific and New Zealand. *New Zealand Journal of Archaeology* (2003).

27. Petchey, F. *et al.* Testing the human factor: radiocarbon dating the first peoples of the South Pacific. *Journal of Archaeological Science* **38**, 29-44 (2011).
28. Kinaston, R.L., *et al.* Lapita diet in Remote Oceania: new stable isotope evidence from the 3000-year-old Teouma site, Efate Island, Vanuatu. *PLoS ONE* 9(3):e9037 (2014).
29. Bronk Ramsey, C. OxCal Program v4. 2.3. *Radiocarbon Accelerator Unit, University of Oxford. Available at:* http://c14. *arch. ox. ac. uk/oxcalhelp/hlp_contents. html (accessed 11 June 2013)* (2013).
30. Reimer, P. J. *et al.* Intcal13 and Marine13 Radiocarbon Age Calibration Curves 0-50,000 Years Cal Bp. *Radiocarbon* **55**, 1869-1887 (2013).
31. Petchey, F., Anderson, A., Zondervan, A., Ulm, S. & Hogg, A. New marine ΔR values for the South Pacific subtropical gyre region. *Radiocarbon* **50**, 373-397 (2008).
32. Parks, T. *et al.* Association between a common immunoglobulin heavy chain allele and rheumatic heart disease risk in Oceania. *Nat Commun* **8**, 14946, doi:10.1038/ncomms14946 (2017).
33. Petchey, F., Clark, G., Winter, O., O'Day, P. & Litster, M. Colonisation of Remote Oceania: New dates for the Bapot-1 site in the Mariana Islands. *Archaeology in Oceania* **52**, 108-126 (2017).

11.3 Supplementary Material for Manuscript C

Supplementary Information
for the manuscript

Ancient Genetic Diversity in Near Oceania - Insights from coastal New Guinea and the Bismarck Archipelago.

Kathrin Nägele [1*], Rebecca Kinaston[2], Selina Carlhoff [3], Dylan Gaffney [5,6], Hallie Buckley[3], Emilie Bertolini [3], Monica Tromp [5,6], Rita Radzeviciute[1], Fiona Petchey[9,10], Dimitri Anson[7], Peter Petchey[7], Glenn Summerhayes [7,8], Johannes Krause [1,2], Cosimo Posth* [3,11], Adam Powell* [1]

Ethics statement

Permissions for the samples included in this study were provided by the National Museum and Art Gallery of Papua New Guinea to Rebecca Kinaston for Nebira and Eriama, Hallie Buckley for Watom and Glenn Summerhayes and Dylan Gaffney for Tilu. The permits extend to this study.

Terminology

The Pacific region has a long and complex cultural and biological history. Is was shaped through multiple dispersal events, followed by genetic admixture and cultural evolution, but also greatly impacted by European colonization starting with Vasco Núñez de Balboas voyages to the Pacific in 1530. As observed in many places of the world, European invasion had a lasting impact on the cultural and biological landscape of the Pacific, imposing their culture, religion and languages on the people of the region, renaming and naming people and places from their Eurocentric and racist view. This led to a subdivision that only at first glance seems reasonable. The region was divided into Micronesia (“the small islands”) Polynesia (“the many islands”) and Melanesia (“the black islands”). While Poly- and Micronesia are based on a topographical description

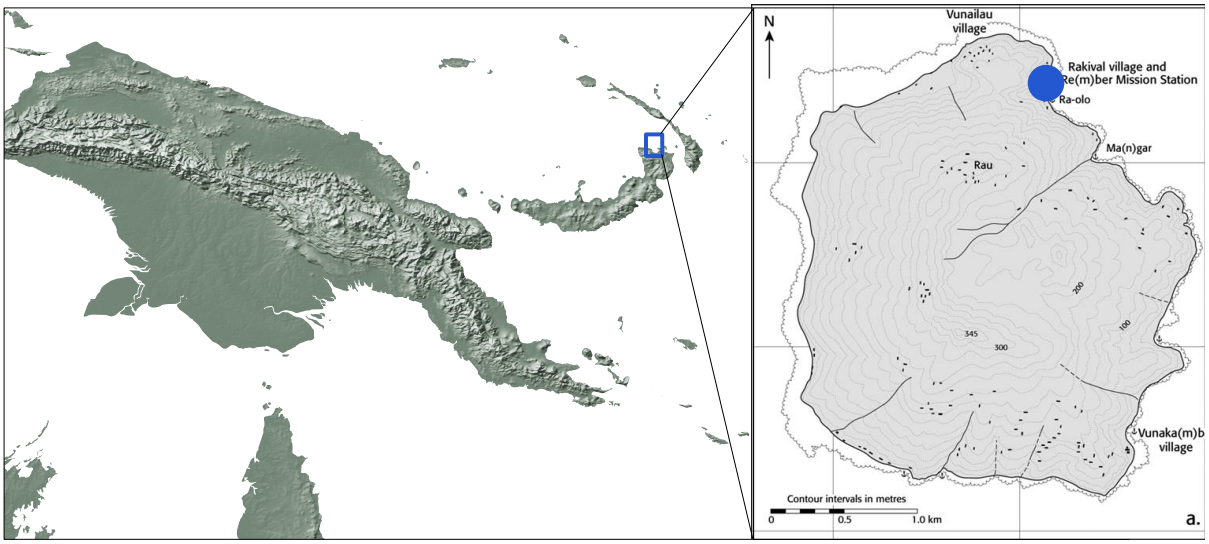
of the islands, Melanesia, comprising New Guinea, the Bismarck Archipelago and the island groups of Fiji, New Caledonia, Vanuatu and Solomon Islands, referred to the colour of peoples skin. Additionally, the concepts of Micronesia and Melanesia have proven inadequate especially in archaeological contexts (*1-3*), as the divisions are not rooted in prehistory (*2*). People inhabiting this region today partly have reclaimed the word (*4*) and are challenging negative representation perpetuated since colonial times. However, we are unable to infer how people in ancient times self identified and have to assume the term introduced by the colonizers homogenized the different peoples based on the colour of their skin, ignoring diverse cultures and identities.
Taking into consideration the biogeography, linguistics and ethnography of the region there is no useful application of the term Melanesia (*2*), hence a perpetuation of this racially inspired term can be avoided. Genetically some populations included in the Melanesian geographical sphere derive their main ancestry from a population that settled in the region ~45,000 years ago. This ancestry signal is maximised in populations inhabiting the highlands of Papua New Guinea. However, a substructure can be observed spreading along a cline without clear grouping of the different islands or archipelagos. We therefore chose to refer to a genetic component that is linked to people from New Guinea, the Bismarck Archipelago or the Solomon Islands as “Papuan-related” if a more general term is needed. We use the names of islands or archipelagos in which a certain genetic signal is maximized in a more regional analysis, and if the signal is specific to a certain group on said islands, e.g. Baining from New Britain, we refer to the genetic ancestry as specifically related to those groups, e.g. “Baining-related”.

The people associated with the Lapita Cultural Complex were the first ones to settle on Vanuatu and Tonga. Their genetic ancestry has been shown to derive from an east-asian ancestral population most similar to ancient and present-day Indigenous individuals from Taiwan and the Philippines (*5*), but show a distinct genetic profile. Although the distinctness of material culture and its great similarity across the islands implies also a cultural unity, it is impossible to reconstruct how people in the past identified, and the vast distance between Vanuatu and Tonga might even suggest different identities despite the same

origins and material culture. However, seeing the genetic similarity of the first settlers of both Vanuatu and Tonga, we are in need of a term to refer to them in a more general way. Often times, studies revert to using the term for the archaeological culture associated with a certain genetic signal. Referring to the genetic composition of the first inhabitants of Vanuatu and Tonga as “Lapita” assumes a connection of material culture, identity and genetic composition we do not want to imply. Other genetic studies have attempted to resolve this by use of the term “First Remote Oceanians” (*5, 6*). However, Remote Oceania does not only cover the islands east of the Solomon Islands, but includes also the islands of Micronesia, north of New Guinea. Based on the similarities of pottery styles and decorations, a dispersal 'from the Philippines via the Mariana Islands' to the Bismarck Archipelago has been proposed before (*7*). Radiocarbon dates of ceramic artefacts found in the Mariana Islands date the initial settlement of humans to 3500 BP (*8, 9*), according to palaeo-environmental evidence even earlier (*10*), suggesting the settlement occurred at the same time or even earlier to that of the Bismarck Archipelago. A recent genetic study investigated the genetic make-up of ancient inhabitants of Guam, dated to 2200 BP (*11*). Ancestry modelling of this individual together with other populations in the Pacific have shown that this individual derives from a lineage ancestral to that of the individuals from Vanuatu and Tonga. Based on the archaeological and genetic evidence available, it is likely that the Mariana Islands were settled before Vanuatu and Tonga, possibly makingg the ancestors of the individual from Guam the “First Remote Oceanians”. As future studies will include more ancient individuals not only from the southern and eastern parts, but also northern Remote Oceania, terminology describing genetic signals has to be reconsidered. To not confuse the geographical anchors involved in the settlement of Remote Oceania, we refer to the genetic profile as observed in the skeletons dating to the Lapita period in Vanuatu and Toga, as “Early Remote Oceanians”.

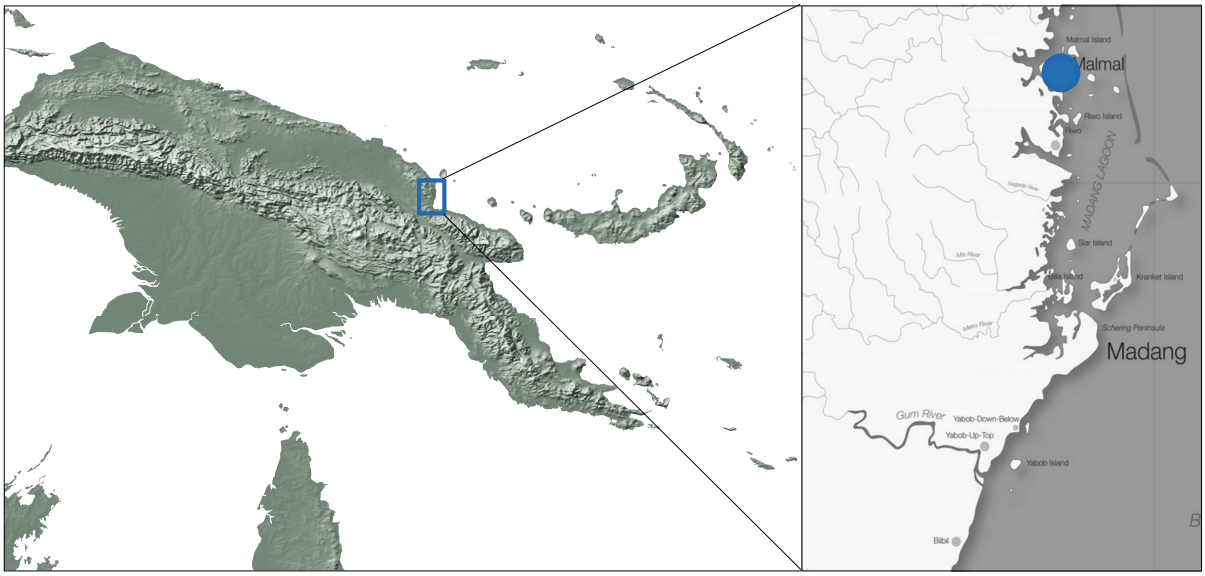
Site descriptions

Reber-Rakival.



The Reber-Rakival site is located on Watom Island, north of East New Britain, 9 km off the Gazelle Peninsula in the Bismarck Sea. A number of sites containing Lapita pottery were excavated on Watom between 1965 and 2009, including the SAC and SAD sites from Reber-Rakival represented in this study (*12-16*). Out of a total of 14 individuals recovered, five are analysed here. Multiple readiocarbon dates, previously produced from human bone from the Watom burials (ca. 2800–2350 cal BP) (*17-19*), placed them in the Middle/Late Lapita phases in the Bismarck Archipelago (ca. 3000–2800/2700–2200 BP)(*20*). To date, the Lapita settlement of site SAC is the earliest evidence of human occupation on the island of Watom, although Lapita populations had been settled on other islands in the Bismarck Archipelago since ~3400 BP (*20*). Newly produced dates of burials B1 and B12 (WAT001, WAT003) represent a later phase of occupation, and burial B10 yields and intermediate date of 2100 BP, showing continued use of the burial site. Additionally, a number of radiocarbon dates from floral and faunal remains, in addition to the burials, have been compiled (*16*). An analysis of the pottery designs from the Lapita layers on Watom further supports the occupation during the Middle and Late Lapita phases in Near Oceania (*21*). A total of five samples were destructively sampled, which all yielded ancient DNA.

Tilu.

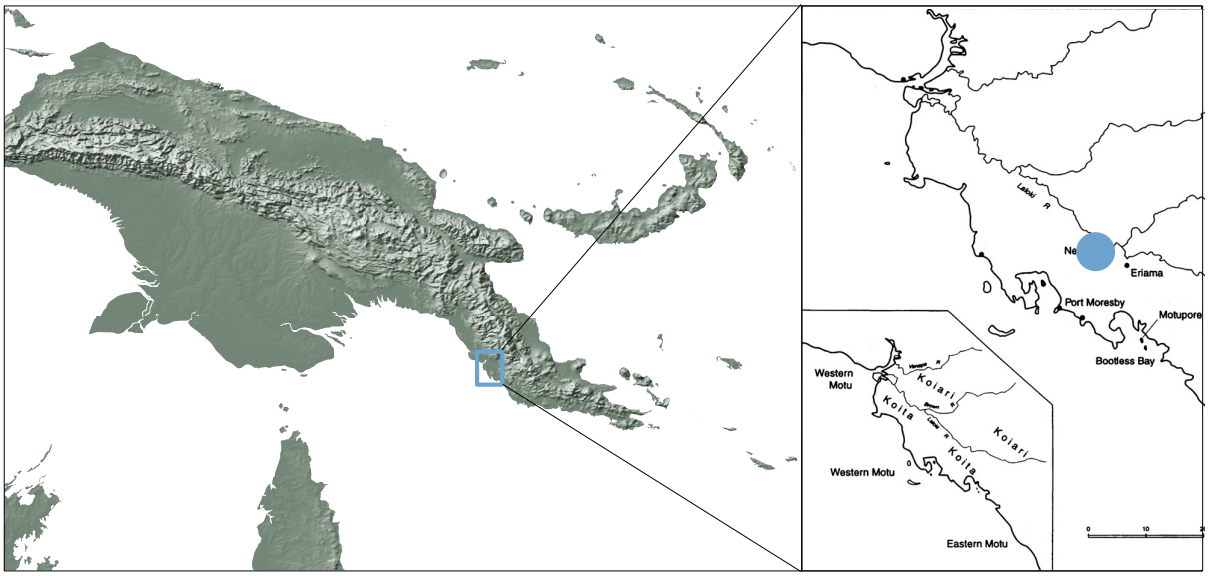


Malmal village is situated 12 km north of Madang Township. Tilu (JCA) clan area (*22*), consists of two elongate mounds about 20 m long. Given around 3 m higher sea levels in the late Holocene (*23*), Tilu would have probably been an island during occupation, starting about 650 BP (*22*). Seven wood charcoal samples were dated and suggest a single phase of site occupation about 550–650 cal BP, consistent with published dates for the site, but somewhat younger than the radiocarbon date retrieved from one individual produced for this study (Table S2). However, corrections for the marine reservoir effect are pending.

Artifacts at the site comprise pottery, obsidian, animal bones of pigs and dogs, and shell artifacts. The pottery sherds show the typical styles for the Madang region. They are red-slipped, produced using paddle and anvil with hand molding. Decorated by appliqué, incision, paddle impression, and impression (*24*). Additionally, 10 sherds from a different, unidentified, ceramic tradition were excavated.

Human remains, mostly mandibular fragments and teeth, were recovered of which two have been analysed in this study. Individual T1702 (TIL001) shows signs of staining indicative for Bethel nut chewing. A total of four samples were destructively samples, of which 2 yielded ancient DNA.

Nebira.

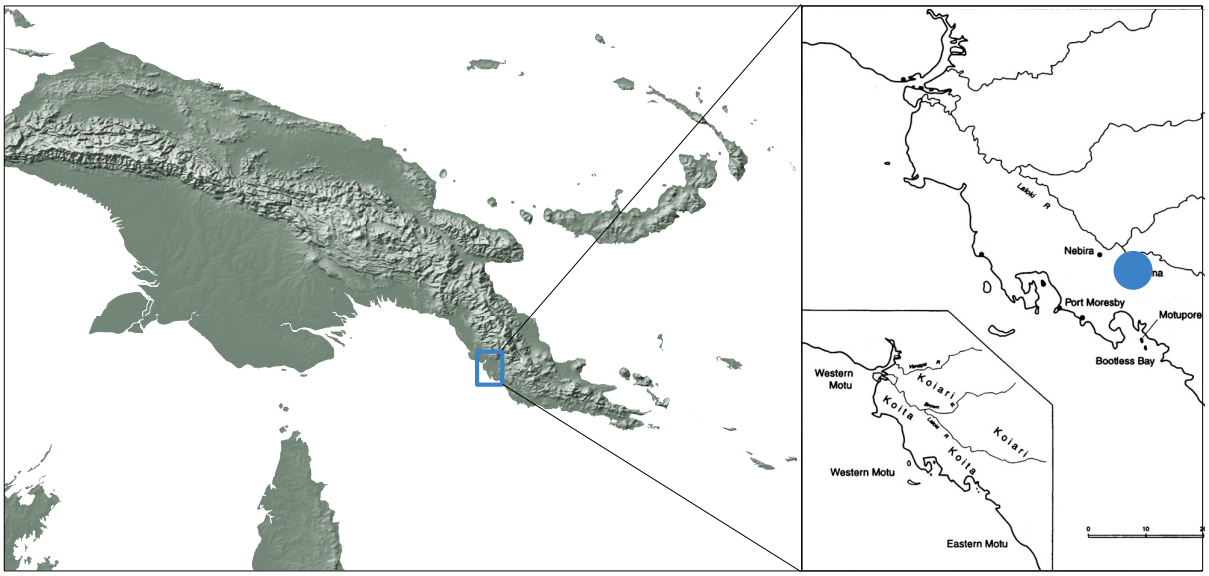


Nebira is located 20 kilometers from the ocean on the South Coast of Papua New Guinea. The first excavations at Nebira were conducted by Susan Bulmer as rescue excavations prompted by the impending destruction of the site by quarrying (*25*). The burials analysed in this study were excavated from central saddle of the two peaked hill Nebira was on, a site with the code ACJ. Radiocarbon dates obtained from the site (*25*) suggest occupation from 950 – 0 BP. More recent radiocarbon dates (*26*) and radiocarbon dates produced for this study (Table 1) range from 540 and 0 BP, indicating the burials were interred during latter part of the Middle Period of South Coast prehistory (A.D.1000 - ~A.D.1500) into the Protohistoric Period (~A.D.1500[1]- A.D.1875), when it was abandoned.

The earlier phases of occupation show a marine subsistence (*27*), with a shift to more terrestrial resources during the later occupations. A total of 38 individuals were found at the site, including five individuals with a non-local Strontium-isotope signature (Supplementary Fig 5a,b,c) (*28, 29*), with varying burial practices and gravegoods (Supplementary Fig. 5a). Based on the stratigraphy the burials were divided into an earlier and later burial phase (Supplementary Fig.5a). Besides single burials, graves were reused for further interments and, due to disarticulation of bones, it has been proposed that the graves had been left open for a period of time, a practice known from historical documents of the local Koita people (*30*). Additionally some bones were likely intentionally removed from the graves, a practice also documented for both groups inhabiting the region today, Motu and Koita (*25, 30*). The pottery traditions indicate that the settlers arrived around 2000 BP and were descendants of the Lapita Cultural Complex (*30, 31*).

A total of 26 samples were destructively sampled of which all yielded ancient DNA.

Eriama.



Eriama ridge is located southeast of Nebira approximately 5 km up the Laloki river and approximately 7.2 km from the nearest point on the coast. From the 24 identified sites at Eriama, individuals analysed in this study are from the site with the code ACV (*25*). 48-50 individuals were excavated. Both adults and non-adults were found at the site, together with shell and animal bone remains, and stone artefacts including a small piece of obsidian possibly from Fergusson Island (*25*). Additionally over 1500 pottery sherds were excavated. The style of pottery associated with one burial was identified as the Waigani Style, which is a late style of bowl decoration found on Motupore Island. One cranium had evidence of painting in the form large oval red spots. Due to the comingled nature of the burial ground association of grave goods with individuals was impossible. It was suggested that the site ACV was only used as a place for interring secondary burials as very few hand and foot bones were found and none of the skeletons were articulated (*25*).

Inferred from the stratigraphy, the site was first inhabited from around 2000-1000 BP, and the use as burial ground commenced after that. Radiocarbon dates from charcoal and human bone indicated a use of the site beginning from 1930 ± 230 (GaK-2670) (*32*). However, the Gakashuin (GaK) lab dates are now thought to be suspect. Direct radiocarbon dates from petrous bones produced for this study provided dates between 154 and 472 cal BP (MAMS 45448, MAMS 45449, MAMS 45450) (Table 1). These dates indicate that the ACV site was used for burial purposes at a contemporaneous time period as Nebira site ACJ, during the latter part of the Middle Period (A.D. 1000 - ~A.D. 1500) and into the Protohistoric Period (~A.D. 1500- A.D. 1875). Unfortunately, all the human remains from the site were mixed during the post-excavation processing of the material in Papua New Guinea. A total of nine samples were destructively sampled, of which all yielded ancient DNA, but one individual was excluded based on a recent radiocarbon date.

Material and Methods

Ancient DNA processing.

Sampling. All samples were processed in dedicated ancient DNA laboratories at the Max Planck Institute for the Science of Human History in Jena, Germany. Bone powder from the petrous part of the temporal bone was obtained through cutting along the margo superior partis petrosae (crista pyramidis) and drilling 50 – 150 mg bone powder from the densest part around the cochlea (*33*). Teeth were sampled by cutting along the junction of the root and the crown and drilling ~50mg from the pulp chamber. In total 46 samples were destructively sampled of which 41 could be included in the analysis.

Radiocarbon dating. The 14 new radiocarbon dates for this study were produced at the Curt-Engelhorn-Zentrum Archäometrie gGmbH in Mannheim, Germany. Collagen from bone and dentin was extracted using a modified Longin method (*34*) and long molecules removed with ultrafiltration before freeze-drying the product (*35*). After the catalytic reduction to graphite the ^{14}C content was measured with an AMS-System type MICADAS. The isotopic ratios of $^{14}C/^{12}C$ and $^{13}C/^{12}C$ of samples, standards (Oxalic acid II) and controls were measured simultaneously. The resulting ^{14}C dates were normed with $\delta^{13}C$=-25‰ (*36*) and calibrated using the software SwissCal 1.0 (L.Wacker, ETH-Zürich) and the INTCAL13 calibration curve (*37*). The radiocarbon dates and quality collagen indicators (collagen yields, C/N ratios, %C and %N) are reported in Table S2.

DNA Extraction. DNA extraction was carried out following established protocols (*38*). Negative and positive controls were included. To release DNA from 50-100 mg of bone powder, a solution of 900 µl EDTA, 75 µl H2O and 25 µl Proteinase K was added. In a rotator, samples were digested for at least 16 h at 37°C, followed by an additional hour at 56°C (*39*). The suspension was then centrifuged and transferred into a binding buffer as previously described (*38*). To bind DNA, silica columns for high volumes (High Pure Viral Nucleic Acid Large Volume Kit; Roche) were used. After two washing steps using the

manufacturer's wash buffer, DNA was eluted in TET (10 mM Tris, 1 mM EDTA and 0.05% Tween) in two steps for a final volume of 100 µl.

Library preparation. Double stranded DNA libraries were built from 25 µl of DNA extract in the presence of uracil DNA glycosylase (UDGhalf libraries), following a double-stranded 'UDG-half' library preparation with a protocol using the UDG enzyme to reduce, but not eliminate, the amount of deamination-induced damage towards the ends of aDNA fragments (*40*). Negative and positive controls were carried alongside each experiment. Libraries were quantified using the IS7 and IS8 primers (*41*) 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 (*42*) in 1-4 parallel 100 µl reactions using PfuTurbo DNA Polymerase (Agilent). 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 (*41*) 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 Illumina platforms in 75 base pair single-end-run cycles using the manufacturer's protocol. To increase the yield for nine individuals with low DNA content in double stranded libraries, we produced a second, single-stranded library (*43*) for those samples in an automated protocol as detailed in (*44*).

Targeted enrichment and high-throughput sequencing. 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). Mitochondrial DNA capture (*45*) was performed on screened libraries which, after shotgun sequencing, showed the presence of aDNA, highlighted by the typical CtoT and GtoA substitution pattern towards 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 in (*46*). The enriched DNA product was sequenced on an Illumina HiSeq 4000 instrument with 75 single-end-run cycles using the manufacturer's protocol. The output was de-multiplexed using bcl2fastq version 2.17.1.14 (Illumina conversion Software) and dnaclust version 3.0.0 (*47*).

***Genomic data processing*.**

Pre-processing of the sequenced reads was performed using EAGER version 1.92.55 (*48*). The resulting reads were clipped to remove residual adaptor sequences using *Clip&Merge (48)* and *AdapterRemoval* version 2 (*49*). Clipped sequences were then mapped against the human reference genome hg19 using the Burrows–Wheeler Aligner (BWA) version 0.7.12 (*50*) disabling seeding (-l 16500, –n 0.01). Duplicates were removed with DeDup version 0.12.2 (*48*), which removes reads with identical start and end coordinates. Additionally, a mapping quality filter of 30 was applied using SAMtools version 1.3 (*51*). In double stranded libraries, reads were trimmed for 2 base pairs to reduce the impact of deamination induced misincorporations during genotyping. Different sequencing runs and libraries from the same individuals were merged, duplicates removed and sorted again using SAMtools (*51*). Trimmed and untrimmed reads were genotyped separately using pileupCaller version 8.6.5 (https://github.com/stschiff/sequenceTools/tree/master/srcpileupCaller), a tool that randomly draws one allele at each of the 1,240 K-targeted SNPs covered at least once. We combined the genotypes keeping all transversions from the untrimmed genotypes and transitions only from the trimmed genotypes to eliminate problematic, damage-related transitions on the ends. Single stranded libraries were genotyped based on the untrimmed reads using the --singleStrandMode. The generated pseudo-haploid calls from both single- and double-stranded libraries were merged using a custom python script, which keeps all identical positions across the two genotypes, as well as the sites covered only in one of the two. For sites covered in both libraries, but with different base calls, the state of the genotype was randomely picked from one of the libraries. The final genotypes of all ancient individuals (Table S1) were merged to a pulldown of the 1,240 K SNPs from the Simons Genome Diversity

Project (*52*), a set of individuals from Asia and the Pacific as reported in Skoglund et al. 2016 (*5*) genotyped on the Human Origins array and previously published ancient Asian and Oceanian individuals (*5, 6, 53, 54*).

Quality control. The typical features of ancient DNA were inspected with *DamageProfiler* version 0.3.1 (http://bintray.com/apeltzer/EAGER/DamageProfiler) (*48*) (Table S3). Sex determination was performed by comparing the coverage on the targeted X-chromosome SNPs (~50 K positions within the 1,240 K capture) normalized by the coverage on the targeted autosomal SNPs to the coverage on the Y-chromosome SNPs (~30 K), again normalized by the coverage on the autosomal SNPs (*55*) (Table S1). For male individuals, ANGSD version 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 (*56*). This provides an estimate of nuclear contamination in males that are expected to have only one allele at each site. All male samples exhibited X-chromosome contamination levels below 7% with at least 100 X-chromosome SNPs covered twice, hence all reads were retained for further analyses (Table S1, Table S4). For both male and female individuals, mtDNA-captured data were used to jointly reconstruct the mtDNA consensus sequence and estimate contamination levels with schmutzi (*57*) (Table S1). For specimens where a relatively low proportion of mtDNA molecules compared with nuclear DNA was observed (Table S3), mtDNA contamination estimates are used as reliable predictors for nuclear contamination (*58, 59*). The software ADMIXTURE version 1.3.0 (*69*) was used in unsupervised mode to allow for free genetic clustering with a worldwide set of individuals (Fig. S2)

Population genomic analysis.

Principal component analysis. Principal component analyses were performed using *smartpca* version 13050 (*60*) with a set of populations from East Asia and the Pacific (Figure 2, Supplementary Data Table S5). Ancient individuals were projected onto the calculated components using the options 'lsqproject: YES', 'shrinkmode: YES' and 'numoutlieriter: 0'. Individuals with less than 20,000 SNPs were not projected with the exception of WAT006.

f-statistics. To identify the differences on an individual basis and to identify a sensible grouping in the subsequent analysis we used *qp3Pop* version 5.0 (*70*) and computed an f_3-outgroup statistics comparing all individuals to each other with Mbuti.DG serving as an outgroup. We used *qpDstat* version 5.0 (*61*) to run f_4-statistics of the form *f_4(Mbuti, Ami.DG Individual 1 site X; Individual 2 site X)* (Table S6). This test expects values close to zero for Individual 1 and Individual 2 if they share more alleles between each other rather than with Ami. In this test all individuals excavated from Nebira showed no Z-score above |3|, suggesting a grouping of the individuals by site was sensible for certain analysis. For individuals excavated from Eriama the two individuals ERI004 and ERI006 produced values above |3| with all other individuals from the site and hence were kept as a separate group. Both individuals excavated from the site Tilu and the two younger individuals excavated on Watom (WAT001 and WAT003) were also grouped on this basis, whereas WAT002, WAT005 and WAT006 were kept separate, also accounting for the long time intervals between them. To test which present-day populations represented best the Papuan and East-Asian related ancestries in the individuals we computed f4-statistics of the form *f_4(Mbuti.DG, Test, X, New_Guinea)* and *f_4(Mbuti.DG, Test, X, Ami),* respectively, testing in X all other populations from the region (Supplementary Table S6). To understand wether the Asian ancestry component was more similar to the Early Remote Oceanians (ERO) from Vanuatu and Tonga (*5*) compared to ancient Austronesians from Taiwan (Suogang) (*53*), predating the expansion to Near and Remote Oceania, we calculated and *f_4(Mbuti.DG, Test; Suogang, ERO)*, expecting positive test scores for a higher affinity to Early Remote Oceanians (Supplementary Table S6, Supplementary Figure S3b). To understand the differential affinities in respect to Near Oceanian Populations, disregarding the differences in Asian ancestry, we produced a biplot (Supplementry Fig. S3a) based on the *f_4(Mbuti. DG, Ami, Test, Baining_Marabu/ New_Guinea)*.

Ancestry modelling. We used *qpWave* version 410 (*62*) to test whether individuals were consistent with deriving from the same group as other individuals from the same site, relative to a set of reference groups (Mbuti, Onge, New_Guinea, Baining_Marabu, Ami, Han, Enlglish, Chukchi, Nasioi, Denisova_published.DG). To test whether some of the individuals could be modeled as consisting of a single ancestry component, we modeled the respective individual and Ami, to test for exclusively Asian ancestry, and New_Guinea, to test for exclusively Near Oceanian ancestry (Supplementary table S.7). For this we used the same references as detailed above, excluding the respective populations used in the test (*63*). After identifying the individuals not consistent with deriving from one respective ancestry, we used *qpAdm* version 5.0 (*61*) to model all groups and the individuals in each group covered by more than 50,000 SNPs as a two-way admixture between New_Guinea and Ami (Supplementary Table S7), and the grouped individuals as a mixture of Early Remote Oceanians and TAN002, a previously published individual with exclusively Near Oceanic, Baining like, ancestry (*64*). As reference groups, we used Mbuti, Onge, Han, Chukchi, English, and Denisova_published.DG. To test for sex-biased admixture we calculated the excess ancestry on the X chromosome, by repeating the admixture modeling on the grouped data with the above settings, but restricted to the X-chromosome (Supplementary Table S7). We subtracted the value obtained from all chromosomes from that obtained from the X Chromosome alone (Supplementary Table S7).

The date of admixture between Papuan related ancestry and East Asian related ancestry was estimated based on linkage disequilibrium. To estimate the admixture in single individuals we used DATES (*65*). As sources we used Papuan.DG for the Papuan related ancestry and, to increase power through a higher number of individuals, a combination of East Asian populations (Han.DG, Ami.DG, Atayal.DG, Igorot.DG, Kinh.DG, She.DG, Dai.DG) to represent the East Asian related ancestry (Figure 3d). The analysis was run with the settings binsize: 0.001, maxdis: 0.5, mincount: 1, lovalfit: 0.45. LD decay curves were inspected to support the soundness of the analysis (Supplementary Figure 4).

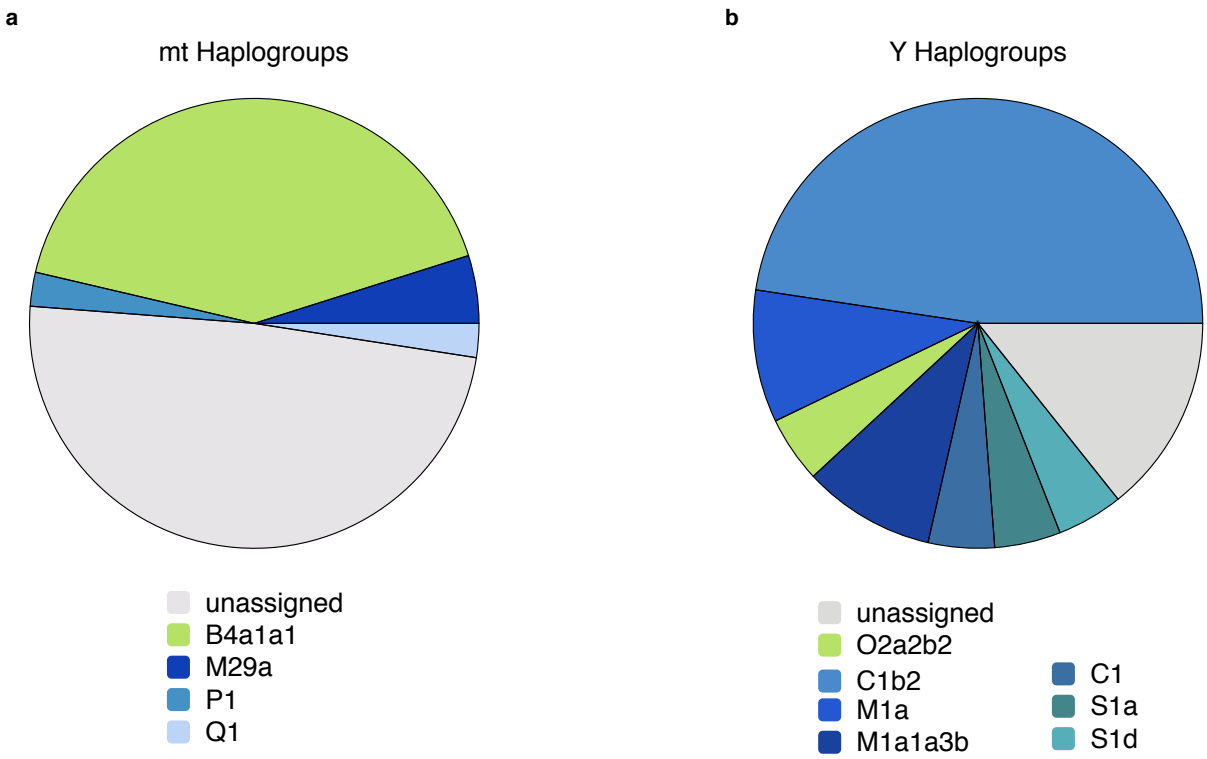
References

1. N. Thomas *et al.*, The force of ethnology: Origins and significance of the Melanesia/Polynesia division [and comments and replies]. *Current Anthropology* **30**, 27-41 (1989).
2. R. Green, Near and Remote Oceania: disestablishing" Melanesia" in culture history. *Man and a half: Essays in Pacific Anthropology and Ethnobotany in honour of Ralph Bulmer*, (1991).
3. J. Terrell, Causal pathways and causal processes: Studying the evolutionary prehistory of human diversity in language, customs, and biology. *Journal of Anthropological Archaeology* **5**, 187-198 (1986).
4. T. Kabutaulaka, Re-Presenting Melanesia: Ignoble Savages and Melanesian Alter-Natives. *EDITED BY STEWART FIRTH*, 193 (2015).
5. P. Skoglund *et al.*, Genomic insights into the peopling of the Southwest Pacific. *Nature* **538**, 510-513 (2016).
6. M. Lipson *et al.*, Population turnover in Remote Oceania shortly after initial settlement. *Current Biology* **28**, 1157-1165. e1157 (2018).
7. P. Bellwood, in *Lost maritime cultures: China and the Pacific*. (Bishop Museum Press, 2007).
8. M. T. Carson, H. Kurashina, Re-envisioning long-distance Oceanic migration: Early dates in the Mariana Islands. *World Archaeology* **44**, 409-435 (2012).
9. H.-c. Hung *et al.*, The first settlement of Remote Oceania: the Philippines to the Marianas. *Antiquity* **85**, 909-926 (2011).
10. J. S. Athens, J. V. Ward, Holocene vegetation, savanna origins and human settlement of Guam. *Records-Australian Museum*, 15-30 (2004).
11. I. Pugach *et al.*, Ancient DNA from Guam and the Peopling of the Pacific. *bioRxiv*, (2020).
12. J. Allen, In search of the Lapita Homeland: Reconsructing the prehistory of the Bismarck Archipelag. *The Journal of Pacific History* **19**, 186-201 (1984).
13. D. Anson, Reber-Rakival dentate-stamped motifs: documentation and comparative implications. *New Zealand Journal of Archaeology* **20**, 119-135 (2000).
14. R. Green, D. Anson, Archaeological investigations on Watom Island: early work, outcomes of recent investigations and future prospects. *New Zealand Journal of Archaeology* **20**, 183-197 (2000).
15. R. Green, D. Anson, J. Specht, The SAC burial ground, Watom Island, Papua New Guinea. *Records of the Australian Museum* **41**, 215-221 (1989).

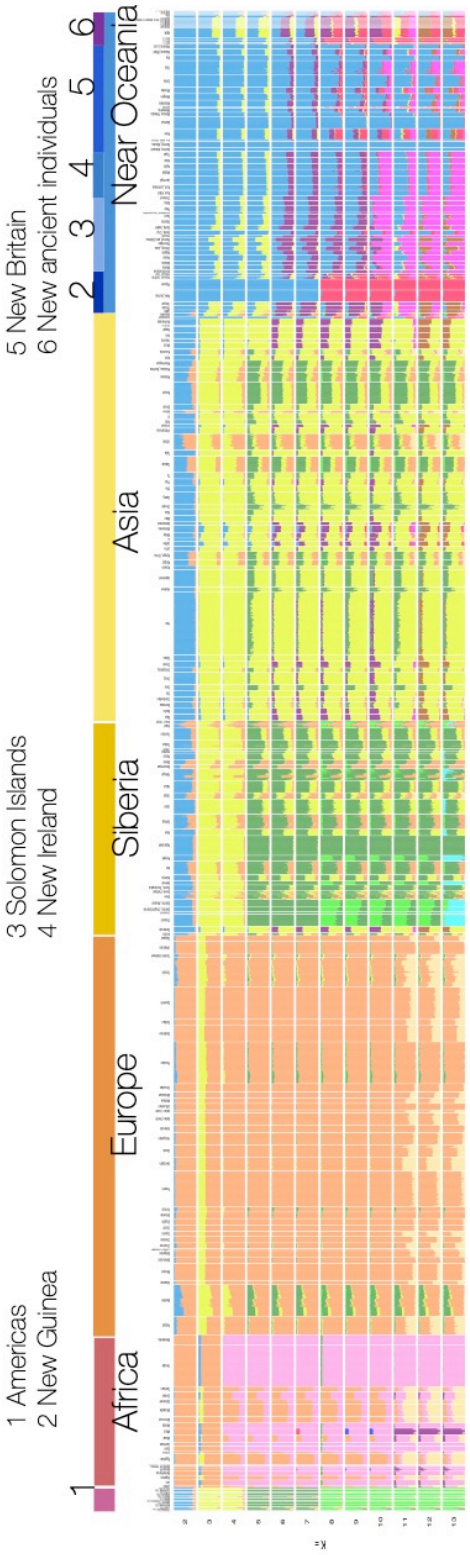
16. D. Anson, R. C. Green, R. K. Walter, *A revised and redated event phase sequence for the Reber-Rakival Lapita site, Watom Island, East New Britain Province, Papua New Guinea*. (Department of Anthropology, University of Otago, 2005).
17. F. Petchey, R. Green, Use of three isotopes to calibrate human bone radiocarbon determinations from Kainapirina (SAC), Watom Island, Papua New Guinea. *Radiocarbon* **47**, 181-192 (2005).
18. F. Petchey *et al.*, Testing the human factor: radiocarbon dating the first peoples of the South Pacific. *Journal of Archaeological Science* **38**, 29-44 (2011).
19. N. B. Athfield, R. C. Green, J. Craig, B. McFadgen, S. Bickler, Influence of marine sources on 14C ages: isotopic data from Watom Island, Papua New Guinea inhumations and pig teeth in light of new dietary standards. *Journal of the Royal Society of New Zealand* **38**, 1-23 (2008).
20. G. R. Summerhayes, Defining the chronology of Lapita in the Bismarck Archipelago. *The chronology of Lapita dispersal in Oceania*, (2001).
21. D. Anson, Compositional analyses of dentate-stamped Lapita and nail-incised and applied relief pottery from Watom Island. *The Pacific from* **5000**, 85-103 (1999).
22. B. Egloff, *Archaeological investigations in the coastal Madang area and on Eloaue Island of the St. Matthias Group*. (Trustees of the Papua New Guinea Public Museum and Art Gallery, 1975).
23. A. W. Tudhope *et al.*, Alternating seismic uplift and subsidence in the late Holocene at Madang, Papua New Guinea: evidence from raised reefs. *Journal of Geophysical Research: Solid Earth* **105**, 13797-13807 (2000).
24. D. Gaffney, University of Otago, (2017).
25. S. Bulmer, University of Papua New Guinea, (1978).
26. R. Kinaston, University of Otago, (2010).
27. J. Allen, Nebira 4: an early Austronesian site in central Papua. *Archaeology & Physical Anthropology in Oceania* **7**, 92-124 (1972).
28. R. Kinaston, H. Buckley, A. Gray, B. Shaw, H. Mandui, Exploring subsistence and cultural complexes on the south coast of Papua New Guinea using palaeodietary analyses. *Journal of Archaeological Science* **40**, 904-913 (2013).
29. B. Shaw, H. Buckley, G. Summerhayes, C. Stirling, M. Reid, Prehistoric migration at Nebira, South Coast of Papua New Guinea: New insights into interaction using isotope and trace element concentration analyses. *Journal of Anthropological Archaeology* **30**, 344-358 (2011).
30. J. Allen, Management of resources in prehistoric coastal Papua. *The Melanesian Environment*, 35-44 (1977).
31. G. R. Summerhayes, J. Allen, Lapita writ small? Revisiting the Austronesian colonisation of the Papuan south coast. *Oceanic Explorations*, (2007).
32. S. Bulmer, Settlement and economy in prehistoric Papua New Guinea: a review of the archeological evidence. *Journal de la Société des Océanistes* **31**, 7-75 (1975).
33. R. Pinhasi *et al.*, Optimal ancient DNA yields from the inner ear part of the human petrous bone. *PloS one* **10**, e0129102 (2015).

34. R. Longin, New method of collagen extraction for radiocarbon dating. *Nature* **230**, 241-242 (1971).
35. T. F. Higham, R. M. Jacobi, C. B. Ramsey, AMS radiocarbon dating of ancient bone using ultrafiltration. *Radiocarbon* **48**, 179-195 (2006).
36. M. Stuiver, H. A. Polach, Discussion reporting of 14 C data. *Radiocarbon* **19**, 355-363 (1977).
37. P. J. Reimer *et al.*, IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. *Radiocarbon* **55**, 1869-1887 (2013).
38. J. Dabney *et al.*, Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci U S A* **110**, 15758-15763 (2013).
39. N. Rohland, M. Hofreiter, Ancient DNA extraction from bones and teeth. *Nat Protoc* **2**, 1756-1762 (2007).
40. N. Rohland, E. Harney, S. Mallick, S. Nordenfelt, D. Reich, Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Philos Trans R Soc Lond B Biol Sci* **370**, 20130624 (2015).
41. M. Meyer, M. Kircher, Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* **2010**, pdb prot5448 (2010).
42. M. Kircher, S. Sawyer, M. Meyer, Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* **40**, e3 (2012).
43. M.-T. Gansauge *et al.*, Single-stranded DNA library preparation from highly degraded DNA usingT4DNA ligase. *Nucleic Acids Research*, gkx033 (2017).
44. M.-T. Gansauge, A. Aximu-Petri, S. Nagel, M. Meyer, 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. *Nature Protocols*, 1-23 (2020).
45. Q. Fu *et al.*, A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr Biol* **23**, 553-559 (2013).
46. Q. Fu *et al.*, An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216-219 (2015).
47. M. Ghodsi, B. Liu, M. Pop, DNACLUST: accurate and efficient clustering of phylogenetic marker genes. *BMC bioinformatics* **12**, 1-11 (2011).
48. A. Peltzer *et al.*, EAGER: efficient ancient genome reconstruction. *Genome Biol* **17**, 60 (2016).
49. M. Schubert, S. Lindgreen, L. Orlando, AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes* **9**, 88 (2016).
50. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
51. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
52. S. Mallick *et al.*, The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201-206 (2016).

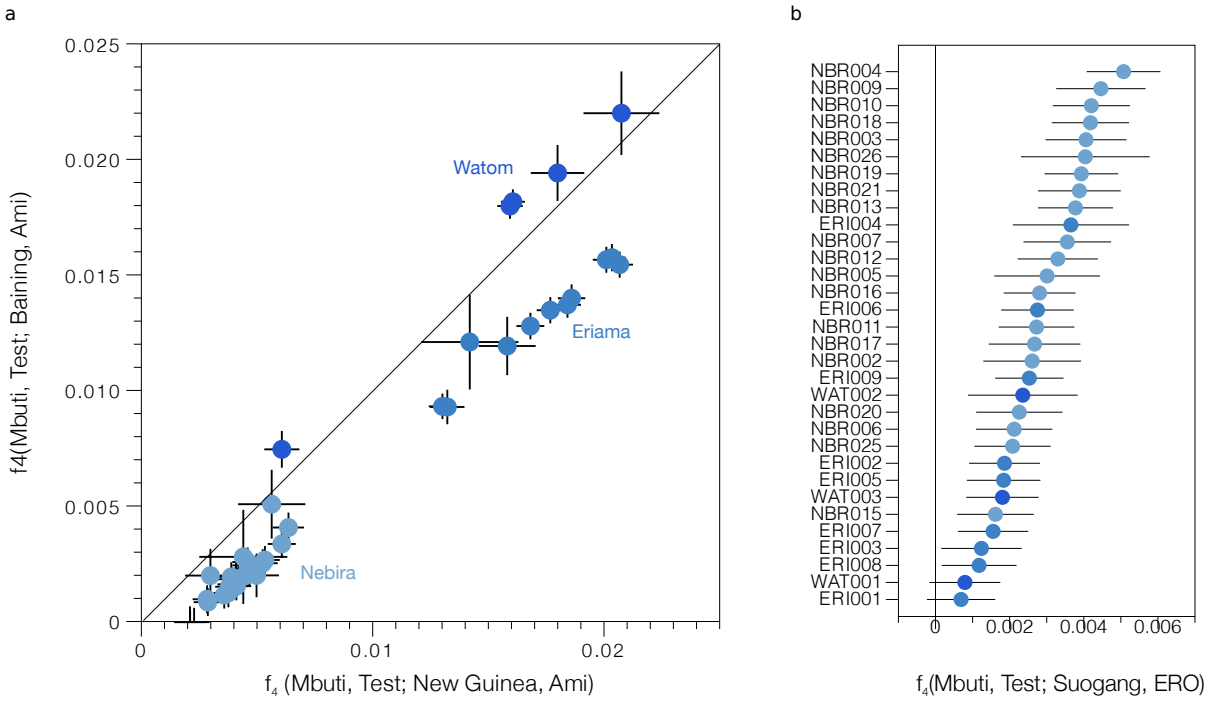
53. M. A. Yang *et al.*, Ancient DNA indicates human population shifts and admixture in northern and southern China. *Science*, (2020).
54. C. Posth *et al.*, Reconstructing the Deep Population History of Central and South America. *Cell* **175**, 1185-1197 e1122 (2018).
55. Q. Fu *et al.*, The genetic history of Ice Age Europe. *Nature* **534**, 200-205 (2016).
56. T. S. A. Korneliussen, A; Nielsen, Rasmus, ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 13 (2014).
57. G. Renaud, V. Slon, A. T. Duggan, J. Kelso, Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol* **16**, 224 (2015).
58. F. M. Key, C. Posth, J. Krause, A. Herbig, K. I. Bos, Mining Metagenomic Data Sets for Ancient DNA: Recommended Protocols for Authentication. *Trends Genet* **33**, 508-520 (2017).
59. A. Furtwangler *et al.*, Ratio of mitochondrial to nuclear DNA affects contamination estimates in ancient DNA analysis. *Sci Rep* **8**, 14075 (2018).
60. N. Patterson, A. L. Price, D. Reich, Population structure and eigenanalysis. *PLoS Genet* **2**, e190 (2006).
61. N. Patterson *et al.*, Ancient Admixture in Human History. *Genetics* **192**, 1065-1093 (2012).
62. D. Reich *et al.*, Reconstructing Native American population history. *Nature* **488**, 370-374 (2012).
63. É. Harney, N. Patterson, D. Reich, J. Wakeley, Assessing the Performance of qpAdm: A Statistical Tool for Studying Population Admixture. *bioRxiv*, (2020).
64. C. Posth *et al.*, Language continuity despite population replacement in Remote Oceania. *Nature ecology & evolution* **2**, 731-740 (2018).
65. V. M. Narasimhan *et al.*, The formation of human populations in South and Central Asia. (2018).



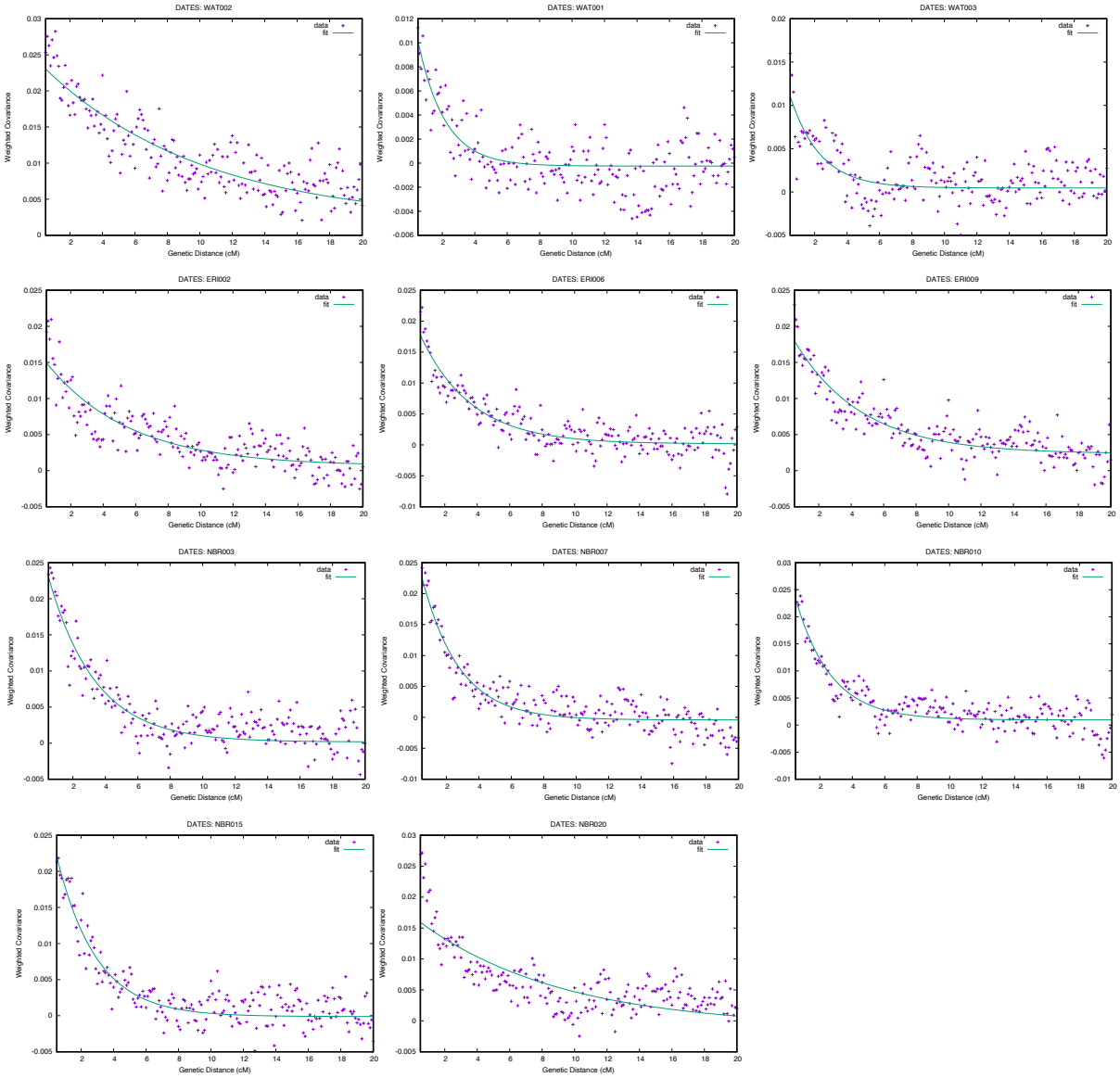
Supplemetary Fig. S1: **Uniparental Markers**. Haplogroup frequencies for mt-Haplogroups for all individuals (a) and Y-Haplogroups (b) for male individuals. Green halplogroups have an Asian origin, blue and teal haplogroups have a Near Oceanian origin.



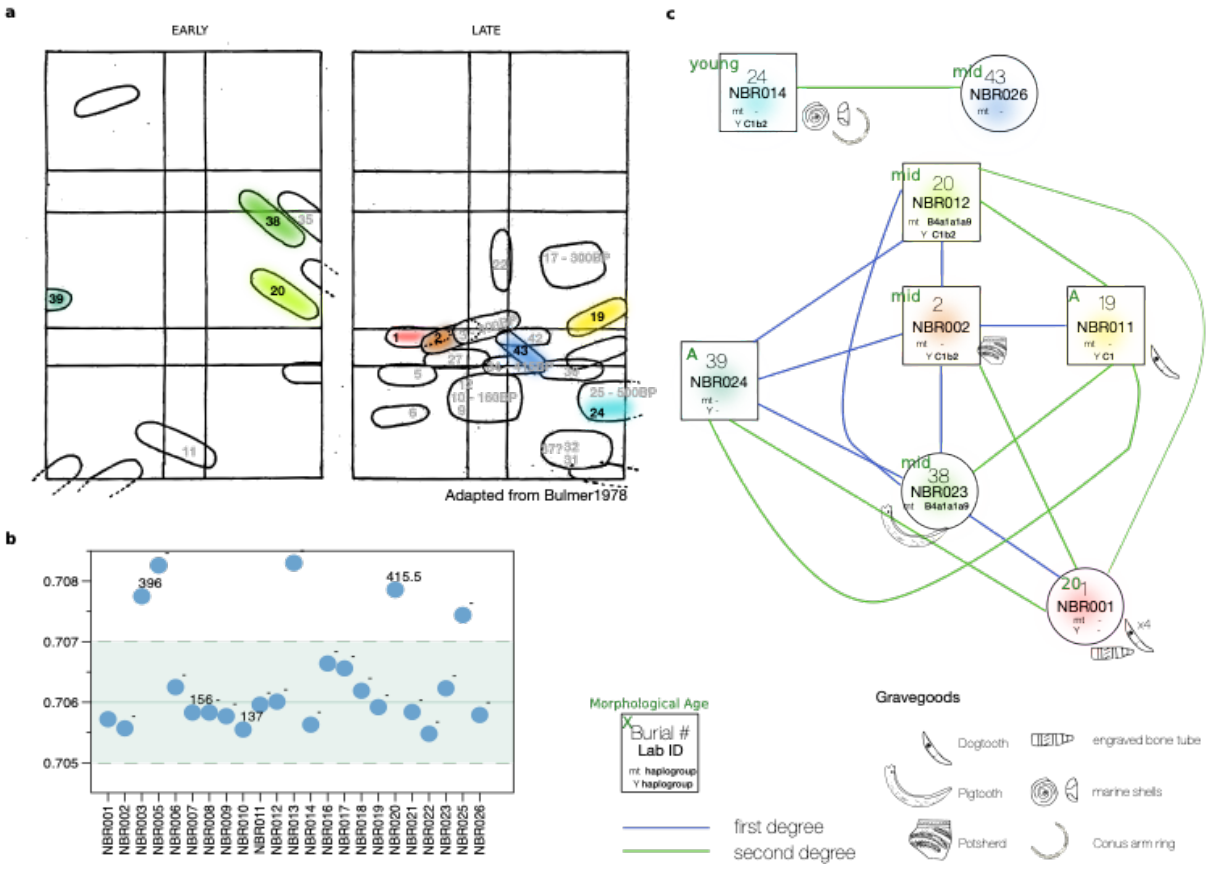
Supplementary Figure S2: Genetic clustering analysis using ADMIXTURE. Clusters K=2-13 are computed based on the available individuals genotyped of the Human Origins array and individuals from the Americas from the SGDP dataset.



Supplementary Figure 3: f_4-statistics. (a) f_4-biplot investigating the affinities to New Guinean highlanders (x-axis) and Baining from New Britain in the Bismarck Archipelago (y-axis). (b) f_4-statistic showing affinity of the Asian component to ancient Taiwanese (Suogang, negative test scores) or Early Remote Oceanians (ERO) from Vanuatu and Tonga (positive test scores). Black lines indicate one standard error.



Supplementray Figure 4: LD dekay curves for the Admixture dating models, showing the LD decay between segments of Asian (Han.DG, Ami.DG, Atayal.DG, Igorot.DG, Kinh.DG, She.DG, Dai.DG) and Near Oceanic (Papuan.DG) ancestry in the individuals included in the analysis (Fig. 3d).



Supplementrary Figure S5: ***Burial patterns and genetic relationships at the site of Nebira.*** *Position and orientation of graves from the early and late burial phase at Nebira (adapted from (25)) (**a**). Coloured graves indicate individuals with genetic relations. Individuals with black numbers show a genetic relationship; individuals with grey numbers are included in the genetic analysis but do not show generic relations**.** Numbers correspond to burial numbers in (25). Strontium isotope analysis for individuals excavated from Nebira (**b**). Turquoise background indicates the local variation at the site; numbers are the midpoint of the calibrated C14 dates.*

*Genetic relationship network (**c**), showing first (blue lines) and second degree (green lines) relationships, mitochondrial and Y-chromosomal haplogroups where available, morphological age and grave goods found with the individuals.*

Supplementary Tables

Available as separate data sheets

Supplementary Table 1: Summary of the individuals analysed in this study. Details on site, location, Archaeological and lab IDs, C14-dating, Sr-values published in Shaw et al 2015, Libraries, genetic sex, contamination estimates, uniparental haplogroups and coverage on the 1240K SNP panel.

Supplementary Table 2: Newly Produced Radiocarbon Dates produced in the Kurt-Engelhorn Zentrum für Archäometrie, Mannheim. Detailed are the C14Lab IDs, Jena Lab IDs, unclaibraed C14 age and standard error, dc13 value, calibrations for 1 sigma and 2 sigma, C:N ratio, concentration and quality of Collagen and the tissue type used for dating.

Supplementary Table 3: Sequencing summaries for initiall screening (SG), 1240K capture (TF) and mitochondrial capture (mt), detailing Sequencing depth, endogenous DNA content on average and on target, coverage statistics, miochondrial to nuclear ratio, damage observed, fragment lenghths and GC content.

Supplementary Table S3: Sex determination and contamination estimation through ANGSD for nuclear contamination in males, and through schmutzi assessing the contamination based on mitochondrial data.

Supplementary Table S5: principal component analysis, present-day populations used to calculate the priciple components, results for the principle components 1-10.

Supplementary Table 6: *f*-statistics. Investigation of differential affinities between Early Remote Oceanians (ERO) and ancient Austronesians (Suogang); New Guinea and populations from the Bismarck Archipelago, and details on the Papuan ancestry in Tilu.

Supplementary Table 7: Ancestry modeling with qpWave/qpAdm. Details on the test scores of individuals against their respective group, the model of Asian (Ami) and Near Oceanic (New_Guinea) ancestry and testing continuity with ancient individuals from Watom and the present-day inhabitants (Tilu).

Supplementary Table 8: Admixture Dating. Dating of admixture between the Asian related and Papuan-related ancestry components. Papuan component modeled through Papuan.SG, Asian component through a combination of Han.DG, Ami.DG, Atayal.DG, Igorot.DG, Kinh.DG, Dai.DG, She.DG

Supplementary Table 9: Genetic relatedness analysis. Degree of genetic relatedness calculated with READ. Determination of first (parent-offspring; siblings) and second degree (Aunt/Uncle - Nephew/Niece; Grandparent - Grandchildren).