LETTER

doi:10.1038/nature19844

Ancient genomics and the peopling of the Southwest Pacific

Pontus Skoglund[1,2,3], Cosimo Posth[4,5], Kendra Sirak[6,7], Matthew Spriggs[8,9], Frederique Valentin[10], Stuart Bedford[9,11], Geoffrey A. Clark[11], Christian Reepmeyer[12], Fiona Petchey[13], Daniel Fernandes[6,14], Qiaomei Fu[1,15,16], Eadaoin Harney[1,2], Mark Lipson[1], Swapan Mallick[1,2], Mario Novak[6,17], Nadin Rohland[1], Kristin Stewardson[1,2,18], Syafiq Abdullah[19], Murray P. Cox[20], Françoise R. Friedlaender[21], Jonathan S. Friedlaender[22], Toomas Kivisild[23,24], George Koki[25], Pradiptajati Kusuma[26], D. Andrew Merriwether[27], Francois-X. Ricaut[28], Joseph T. S. Wee[29], Nick Patterson[2], Johannes Krause[5], Ron Pinhasi[6]§ & David Reich[1,2,18]§

The appearance of people associated with the Lapita culture in the South Pacific around 3,000 years ago[1] marked the beginning of the last major human dispersal to unpopulated lands. However, the relationship of these pioneers to the long-established Papuan people of the New Guinea region is unclear. Here we present genome-wide ancient DNA data from three individuals from Vanuatu (about 3,100–2,700 years before present) and one from Tonga (about 2,700–2,300 years before present), and analyse them with data from 778 present-day East Asians and Oceanians. Today, indigenous people of the South Pacific harbour a mixture of ancestry from Papuans and a population of East Asian origin that no longer exists in unmixed form, but is a match to the ancient individuals. Most analyses have interpreted the minimum of twenty-five per cent Papuan ancestry in the region today as evidence that the first humans to reach Remote Oceania, including Polynesia, were derived from population mixtures near New Guinea, before their further expansion into Remote Oceania[2–5]. However, our finding that the ancient individuals had little to no Papuan ancestry implies that later human population movements spread Papuan ancestry through the South Pacific after the first peopling of the islands.

Pacific islanders today derive from a mixture of two highly divergent ancestral populations[3]. The first ancestral modern human population arrived in island southeast Asia more than 40,000 years before present (BP), and contributed to the ancestry of both indigenous Australians and Papuans, and hence to other Pacific islanders[4]. The second ancestral population is more closely related to mainland East Asians[4], and is not found in unadmixed form today. The first humans to reach Remote Oceania—a term we use to refer to the region unoccupied before approximately 3,000 BP beyond the main Solomon Islands and, in this case, excluding Micronesia—were associated with the Lapita culture, which existed between 3,450–3,250 and 2,700–2,500 BP. These people spread into Remote Oceania using the first boats capable of long-distance sea travel and introduced new domesticated animals and plants, and their successors reached the most isolated islands of the eastern and southern Pacific by 1,000–700 BP[6]. Several hypotheses have been proposed to explain why present-day indigenous people of Near Oceania (New Guinea, the Bismarck Islands, and the Solomon Islands area) and Remote Oceania have ancestry both from Papuans and from populations of ultimate East Asian origin. In one set of models that has been favoured by recent genetic studies[3–5,7], the mixture occurred at around 3,000 BP, during the expansion of populations of East Asian origin through the New Guinea region[8]. In the other set of models, the population of ultimate East Asian origin initially mixed little with Papuans[9], and thus later gene exchanges account for the ubiquitous Papuan ancestry today[2,10].

We obtained genome-wide ancient DNA data from three individuals from Teouma, an archaeological site on Efate island, Vanuatu (Supplementary Information section 1), which were all directly radiocarbon dated to between 3,110 and 2,740 BP, an interval that is chronologically part of the Lapita period (Extended Data Table 1). We also obtained genome-wide ancient DNA data from an individual from the Talasiu site on Tongatapu island, Tonga, directly radiocarbon dated to 2,680–2,340 BP, a period spanning the late Lapita and immediately post-Lapita period (Supplementary Information section 2 and Extended Data Table 1). In dedicated clean rooms, we prepared powder from petrous bones[11], extracted DNA[12], and prepared up to four double-stranded libraries from each extract[13]. We enriched the libraries for 1.24 million targeted single nucleotide polymorphisms (SNPs)[14], sequenced the products, and represented each individual by a single randomly drawn sequence for each SNP. This procedure resulted in 139,461–231,944 SNPs that were covered at least once in each of the individuals. The low ratio of sequences aligning to Y chromosome targets compared to targets on other chromosomes[15] reveals that all four individuals are females (Extended Data Table 1). We obtained three mitochondrial DNA sequences from Vanuatu and all were haplogroup B4a1a1a, the classic 'Polynesian motif'[16].

Multiple features of the data suggest that the DNA was authentic and minimally contaminated. First, in all individuals, around 40% of

[1]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [2]Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. [3]Archaeological Research Laboratory, Department of Archaeology and Classical Studies, Stockholm University, 10691 Stockholm, Sweden. [4]Institute for Archaeological Sciences, Archaeo- and Palaeogenetics, University of Tübingen, Tübingen 72070, Germany. [5]Max Planck Institute for the Science of Human History, 07745 Jena, Germany. [6]School of Archaeology and Earth Institute, Belfield, University College Dublin, Dublin 4, Dublin, Ireland. [7]Department of Anthropology, Emory University, Atlanta, Georgia 30322, USA. [8]School of Archaeology and Anthropology, College of Arts and Social Sciences, The Australian National University, Canberra, Australian Capital Territory 2601, Australia. [9]Vanuatu National Museum, Vanuatu Cultural Centre, Port Vila, Vanuatu. [10]Maison de l'Archéologie et de l'Ethnologie, CNRS, UMR 7041, 92023 Nanterre, France. [11]Department of Archaeology and Natural History, College of Asia and the Pacific, The Australian National University, Canberra, Australian Capital Territory 2601, Australia. [12]College of Arts, Society and Education, James Cook University, Queensland 4870, Australia. [13]Radiocarbon Dating Laboratory, University of Waikato, Hamilton 3240, New Zealand. [14]CIAS, Department of Life Sciences, University of Coimbra, 3000-456 Coimbra, Portugal. [15]Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, IVPP, CAS, Beijing 100044, China. [16]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig 04103, Germany. [17]Institute for Anthropological Research, 10000 Zagreb, Croatia. [18]Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts 02115, USA. [19]RIPAS Hospital, Bandar Seri Begawan, Brunei Darussalam. [20]Institute of Fundamental Sciences, Massey University, Palmerston North, 4442, New Zealand. [21]Independent Scientist, Sharon, Connecticut 06069, USA. [22]Department of Anthropology, Temple University, Gladfelter Hall, Philadelphia, Pennsylvania 19122, USA. [23]Estonian Biocentre, Evolutionary Biology group, Tartu, 51010, Estonia. [24]Division of Biological Anthropology, University of Cambridge, Fitzwilliam Street, Cambridge CB2 3QG, UK. [25]Papua New Guinea Institute of Medical Research, Goroka, Eastern Highlands Province 441, Papua New Guinea. [26]Eijkman Institute for Molecular Biology, Jakarta 10430, Indonesia. [27]Department of Anthropology, Binghamton University, Binghamton, New York 13902, USA. [28]Evolutionary Medicine Group, Laboratoire d'Anthropologie Moléculaire et Imagerie de Synthèse UMR 5288 CNRS, Université de Toulouse, Toulouse 31073, France. [29]National Cancer Centre Singapore, Singapore 169610, Singapore. §These authors jointly supervised this work.

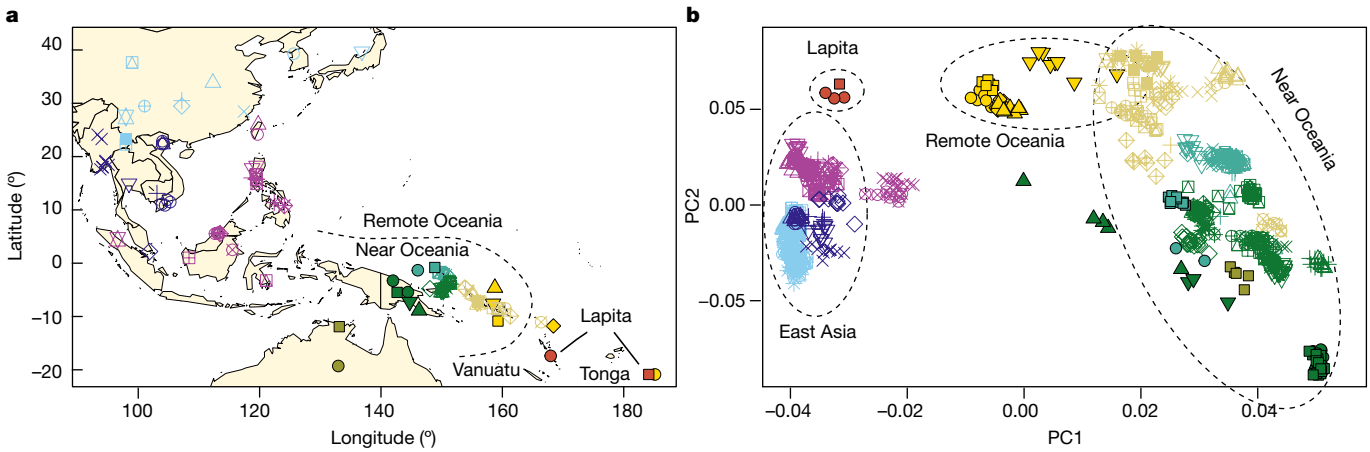


Figure 1 | Data from ancient and present-day populations. a, Locations of 778 present-day individuals genotyped on the Affymetrix Human Origins Array and 4 ancient individuals. **b**, Ancient individuals projected onto principal components (PC) 1 and 2 computed using only present-day samples. Individual population labels are given in Extended Data Fig. 2.

all sites that are cytosines in the human reference sequence appear as thymines in the terminal nucleotide, as expected for genuine ancient DNA (Extended Data Fig. 1a). Second, when we carried out principal component analysis (PCA; Fig. 1) of 778 present-day people from 83 East Asian and Oceanian populations genotyped at 621,799 SNPs (Extended Data Table 2) and projected the ancient individuals, we found that all clustered tightly with each other and with data from the same individuals restricting to sequences with cytosine-to-thymine changes at the terminal nucleotide (these sequences are unlikely to be contaminants[17,18]) (Extended Data Fig. 1b). Third, the cluster of ancient individuals does not overlap with present-day populations, indicating that the data are from a population that is not present in unmixed form today (Fig. 1). The distinctiveness of the ancient individuals is also highlighted by their high differentiation from all present-day groups ($0.05 < F_{ST} < 0.26$; between all modern individuals and the ancient Vanuatu individuals, using the statistic F_{ST}, which is proportional to average squared allele frequency difference) (Extended Data Table 3).

The ancient Vanuatu and Tongan individuals are not shifted in the PCA in the direction of Papuan ancestry, in contrast to all present-day Remote Oceanians. In this respect, they are similar to indigenous Taiwanese populations such as the Ami and Atayal as well as to populations from the Philippines such as the Kankanaey, who have no detectable Papuan ancestry (Fig. 1). To test whether the ancient individuals had any evidence of Papuan ancestry, we used the *qpWave*/*qpAdm* software (Methods) to analyse allele frequency correlation statistics[19]. The results were consistent with the ancient individuals and the Taiwanese Ami having descended from a common ancestral population to the exclusion of 14 worldwide outgroups ($P > 0.05$ for the ancient individuals from both Vanuatu and Tonga). We estimate the possible range of Papuan ancestry in the Vanuatu individuals to be 0–11% and in the Tongan individual to be 0–17% (99% confidence intervals truncated at zero), which does not overlap the point estimates of at least 25% Papuan ancestry in all present-day Oceanians (Fig. 2a). To test the hypothesis that the ancient Remote Oceanian individuals might be from the source population of the non-Papuan ancestry in Oceanians today, we computed the statistic f_4(Africa, *Test*; Australian, Polynesian), which evaluates the degree of allele sharing of a candidate *Test* population with Polynesians (at sites where Polynesians differ

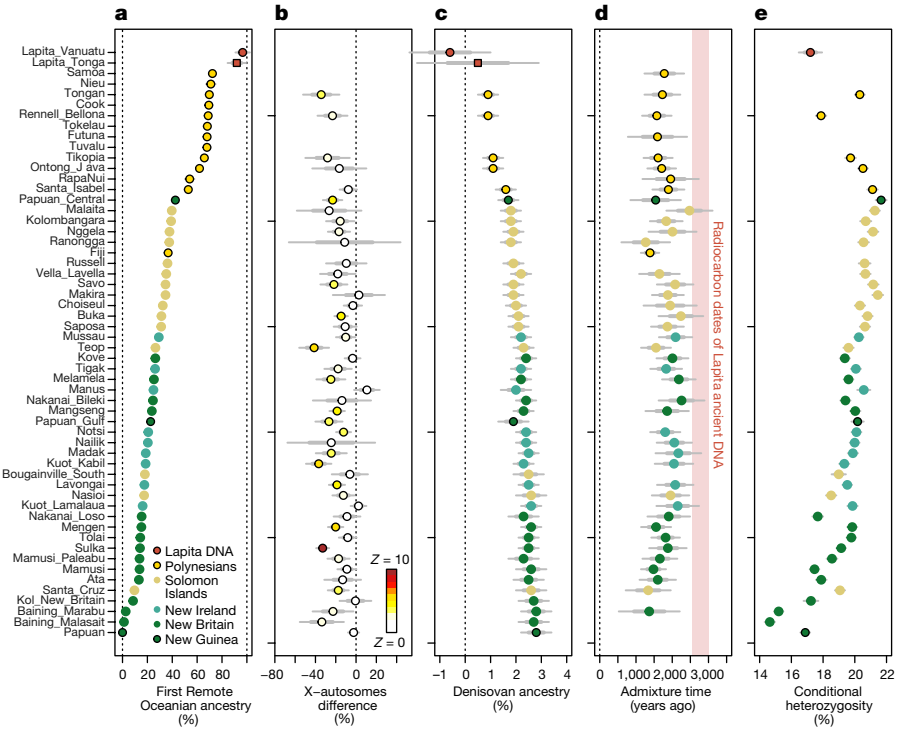


Figure 2 | Genetic characteristics of the Oceanian ancestry cline. a, Estimated proportion of First Remote Oceanian ancestry. The Papuan ancestry can be estimated as 100% minus the estimate of First Remote Oceanian ancestry. **b**, Difference between First Remote Oceanian ancestry estimates on chromosome X and the autosomes. **c**, Denisovan ancestry estimates are inversely related to First Remote Oceanian ancestry estimates. **d**, Estimated date of admixture in all populations with at least four individuals and significant evidence of decay of weighted admixture linkage disequilibrium as measured in ALDER. We used Han and New Guinean Highlanders as surrogates for the ancestral populations. We assumed a generation interval of 28.1 years, and show 95% confidence intervals (thin whiskers) incorporating uncertainty both in the ALDER date and the value of the human generation interval. We show the range of radiocarbon dates for the ancient individuals. **e**, Conditional heterozygosity (genetic diversity) estimated by drawing two random chromosomes from different individuals at each locus, using only SNPs ascertained in a single Yoruba, and restricting to transversion SNPs to avoid any concerns about inflated heterozygosity due to ancient DNA degradation. Thick and thin error bars in all five panels correspond to 1 and 1.96 standard error of the estimate, respectively.

from Australians), and found that it was maximized when *Test* was the ancient Vanuatu or Tonga individuals (Extended Data Fig. 2b), as expected if a population related to them was the true source. We conclude that the non-Papuan ancestry that is ubiquitous in Oceania is derived from a population related to the ancient individuals we analysed, and that this ancestry reached uninhabited islands in Remote Oceania with little or possibly no mixture with Papuans. We call the population of which both the ancient Vanuatu and Tongan individuals were a part the 'First Remote Oceanians' and find that the ancestry fraction from this population is the single most important factor shaping genetic variation among Pacific islanders, accounting for most variation in measurements including genetic diversity (Pearson's $R = 0.86$, $P = 2 \times 10^{-12}$ for 42 non-Polynesian groups; Extended Data Fig. 2) and the proportion of archaic Denisovan ancestry ($R = -0.96$, $P < 10^{-16}$ for all 56 Oceanian groups; Fig. 2).

Our evidence that early and geographically diverse Remote Oceanian individuals had little or no Papuan ancestry contradicts models in which there were significant Papuan contributions to Lapita people before their dispersal into Remote Oceania[3–5]. Instead, our results show that the Papuan genetic signature appeared in many Remote Oceanian populations only subsequent to initial settlement. To gain further insight into when the Papuan ancestry may have become ubiquitous in Remote Oceanians, we leveraged the fact that chromosome segments from ancestral populations break up at a known rate due to recombination and that the length distribution of these segments translates to a date of mixture[20]. We estimate dates of approximately 50–80 generations ago using ALDER[21], or 1,500–2,300 BP assuming 28.1 years (see Methods) per generation[22] (Fig. 2d and Extended Data Fig. 3). We combined the statistical error of the genetic estimate and the uncertainty about the generation interval (Methods), and obtained a 95% confidence interval of 1,239–1,927 BP for a pool of Polynesians, all of whom have similar Papuan ancestry proportions. This finding that Papuan–First Remote Oceanian mixture occurred long after the end of the Lapita period implies that the Polynesian ancestral population was not fully formed at that time, although we note that alternative methods for dating Papuan admixture in Remote Oceanians arrived at older dates[4,23–25]. However, our ALDER dates are supported by direct ancient DNA evidence, as the Tongan individual at 2,680–2,340 BP carried little or no Papuan ancestry, providing unambiguous confirmation that the ancestral population of Polynesians was not fully formed by the end of the Lapita period.

We used *qpGraph* to explore models of population separation and mixture that might accommodate the ancient DNA data[26] (Supplementary Information section 3). We obtained fits using models in which Polynesians today are mixtures of First Remote Oceanians and a Papuan population related to Highland New Guineans (Fig. 3a). We also obtained consistent findings using *TreeMix*[27] (Extended Data Fig. 4). In Fig. 3 we show the best fitting model, which suggests that the ancient individuals from Vanuatu and Tonga descended from an ancestral (presumably Lapita) population that separated earlier from the population that is the primary component in present-day Polynesians. This implies that not just Papuan ancestry but also deeply branching First Remote Oceanian ancestry was introduced to Remote Oceania through movement of people after the time of the ancient individuals. Thus, the minimum 25% Papuan ancestry seen in present-day Remote Oceanians is a conservative underestimate of the later population displacement. It is unlikely that there was 100% replacement, however, as we observed weak excess affinity of present-day Tongans to the ancient Tongan individuals in symmetry tests (see Methods). More deeply in time, our modelling indicates that Philippine populations (Kankaney) are the closest outgroup to the First Remote Oceanians, indigenous Taiwanese (Atayal) second closest, and mainland southeast Asians such as the Dai most remote, consistent with models of population movement along a route from Taiwan to the Philippines to Near Oceania to Remote Oceania[28]. We were surprised that we could not fit Australians as outgroups to New Guinean Highlanders and the

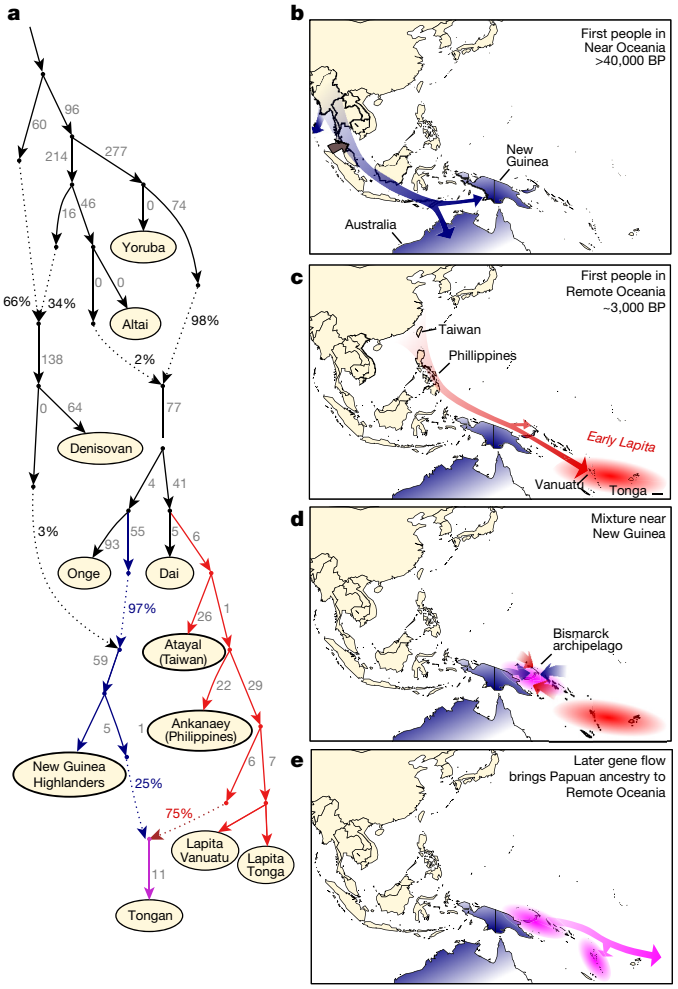


Figure 3 | A model of population history. a, A model of population relationships that fits allele frequency patterns (all empirical *f*-statistics within 3 standard errors of expectation). Branch lengths are shown in units of $F_{ST} \times 1{,}000$. Admixture edges show mixture proportions. **b**, A model of population movements more than 40,000 years ago in which modern humans arrived in the Australia–New Guinea region (blue shading) and mixed with archaic Denisovans (brown shading). **c**, A model of events before 3,000 years ago, in which the First Remote Oceanian population formed by spread of a population of ultimate East Asian origin to a region including Vanuatu and Tonga, and experienced little or no mixture with the Papuans they encountered along the journey (red shading). Note that geographic routes are speculative. **d**, A model of populations of mixed Papuan–First Remote Oceanian ancestry in Near Oceania less than 3,000 years ago in a patchwork of islands with different proportions of First Remote Oceanian ancestry (pink shading). **e**, A model of secondary expansion of admixed populations bringing Papuan ancestry into Remote Oceania, which was still not complete in Tonga by the date of the Talasiu individual at 2,680–2,340 BP.

Papuan ancestry in Polynesians (Extended Data Fig. 5). However, we could fit Australians as deriving from a mixture of an ancient Australian lineage and a Papuan lineage from the same group that expanded into Polynesia. This is plausible if there was continuing gene flow between New Guinea and Australia. Another parsimonious model is that the ancestry in present-day Polynesians is not all Papuan, but a Papuan–Australian mix.

Previous studies of mitochondrial DNA and Y chromosomes suggested that present-day people of the South Pacific harbour more East Asian ancestry from female than from male ancestors[3]. Our genome-wide analyses confirm a significant excess of First Remote Oceanian ancestry on the X chromosome compared to the autosomes (*Z* scores up to 10) (Fig. 2b). Females carry two-thirds of the X chromosomes in

a population but only half of the autosomes (Extended Data Fig. 6), and we compared the ancestry estimates in these two parts of the genome to obtain the most accurate estimates of sex-biased admixture in diverse Oceanians to date (Extended Data Fig. 6 and Extended Data Table 4). It has been suggested that matrilocal social structure in the primarily First Remote Oceanian ancestry populations of the region is one likely factor to explain these patterns[29,30]. However, it is also possible that some of these patterns reflect a scenario in which the later movement of Papuan ancestry into Remote Oceania was largely mediated by males who then mixed with resident females.

Our study has shown that many of the first humans in Remote Oceania had little, if any, Papuan ancestry, in stark contrast to the situation today. While our findings cannot rule out the possibility that multiple groups—some of which carried substantial amounts of Papuan ancestry—settled Remote Oceania early on, the lack of such ancestry in both Vanuatu and Tonga can be more parsimoniously explained by later population movements bringing the Papuan ancestry. The scenario emerging from ancient DNA analysis is thus radically different from that suggested by previous genetic studies, which have generally posited that the first people in Remote Oceania and Polynesia[2–5] had substantial Papuan ancestry. Our finding of major post-Lapita movements of Papuan ancestry into Remote Oceania also cannot be related to the later arrival of Papuan ancestry that has been suggested for Fiji, which is estimated to have occurred at least a millennium later at 500 BP[4] or 1,100 BP[24] (Fig. 2). Systematic study of ancient DNA from throughout Remote Oceania should make it possible to provide a detailed chronicle of the population movements and sex-biased population mixtures that shaped the ancestry of present-day Oceanians.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 20 April; accepted 13 September 2016.
Published online 3 October 2016.

1. Sheppard, P. J., Chiu, S. & Walter, R. Re-dating Lapita movement into Remote Oceania. *J. Pacific Archaeol.* **6,** 26–36 (2015).
2. Kayser, M. *et al.* Genome-wide analysis indicates more Asian than Melanesian ancestry of Polynesians. *Am. J. Hum. Genet.* **82,** 194–198 (2008).
3. Kayser, M. The human genetic history of Oceania: near and remote views of dispersal. *Curr. Biol.* **20,** R194–R201 (2010).
4. Wollstein, A. *et al.* Demographic history of Oceania inferred from genome-wide data. *Curr. Biol.* **20,** 1983–1992 (2010).
5. Matisoo-Smith, E. Ancient DNA and the human settlement of the Pacific: a review. *J. Hum. Evol.* **79,** 93–104 (2015).
6. Bellwood, P. S. *First Farmers: the Origins of Agricultural Societies* (Blackwell Publishing, 2005).
7. Duggan, A. T. *et al.* Maternal history of Oceania from complete mtDNA genomes: contrasting ancient diversity with recent homogenization due to the Austronesian expansion. *Am. J. Hum. Genet.* **94,** 721–733 (2014).
8. Kayser, M. *et al.* Melanesian origin of Polynesian Y chromosomes. *Curr. Biol.* **10,** 1237–1246 (2000).
9. Blust, R. Remote Melanesia: one history or two? An addendum to Donohue and Denham. *Oceanic Linguistics* **47,** 445–459 (2008).
10. Friedlaender, J. S. *et al.* The genetic structure of Pacific Islanders. *PLoS Genet.* **4,** e19 (2008).
11. Pinhasi, R. *et al.* Optimal ancient DNA yields from the inner ear part of the human petrous bone. *PLoS One* **10,** e0129102 (2015).
12. 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).
13. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. Lond. B* **370,** 20130624 (2015).
14. Fu, Q. *et al.* An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524,** 216–219 (2015).
15. Skoglund, P., Storå, J., Götherström, A. & Jakobsson, M. Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J. Archaeol. Sci.* **40,** 4477–4482 (2013).
16. Melton, T. *et al.* Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. *Am. J. Hum. Genet.* **57,** 403–414 (1995).
17. Skoglund, P. *et al.* Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* **336,** 466–469 (2012).
18. 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).
19. Reich, D. *et al.* Reconstructing Native American population history. *Nature* **488,** 370–374 (2012).
20. Moorjani, P. *et al.* The history of African gene flow into Southern Europeans, Levantines, and Jews. *PLoS Genet.* **7,** e1001373 (2011).
21. Loh, P.-R. *et al.* Inference of admixture parameters in human populations using weighted linkage disequilibrium. (2012).
22. Fenner, J. N. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am. J. Phys. Anthropol.* **128,** 415–423 (2005).
23. Xu, S., Pugach, I., Stoneking, M., Kayser, M. & Jin, L. Genetic dating indicates that the Asian-Papuan admixture through Eastern Indonesia corresponds to the Austronesian expansion. *Proc. Natl Acad. Sci. USA* **109,** 4574–4579 (2012).
24. Pugach, I., Matveyev, R., Wollstein, A., Kayser, M. & Stoneking, M. Dating the age of admixture via wavelet transform analysis of genome-wide data. *Genome Biol.* **12,** R19 (2011).
25. Lipson, M. *et al.* Reconstructing Austronesian population history in Island Southeast Asia. *Nat. Commun.* **5,** 4689 (2014).
26. Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192,** 1065–1093 (2012).
27. Pickrell, J. K. & Pritchard, J. K. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8,** e1002967 (2012).
28. Bellwood, P. Holocene population history in the Pacific region as a model for worldwide food producer dispersals. *Curr. Anthropol.* **52,** S363–S378 (2011).
29. Jordan, F. M., Gray, R. D., Greenhill, S. J. & Mace, R. Matrilocal residence is ancestral in Austronesian societies. *Proc. R. Soc. Lond. B* **276,** 1957–1964 (2009).
30. Stephen Lansing, J. *et al.* An ongoing Austronesian expansion in Island Southeast Asia. *J. Anthropol. Archaeol.* **30,** 262–272 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the 356 volunteers who donated samples for genome-wide analysis; M. Stoneking for co-funding genotyping of the Bismarck samples; M. Brilliant, H. Norton, and L. Scheinfeldt, for help in the preparation of the Bismarck samples and establishment of a repository for them at the Marshfield Foundation; A. Kim, I. Pugach, and M. Stoneking for comments, and I. Mathieson for critiques and advice on estimating sex-specific ancestral contributions. The maps in Figs 1a and 3b–e maps were plotted in R using the world() map of the 'fields' and 'maps' packages (using public domain data from the CIA World Data Bank II). P.S. was supported by the Wenner-Gren foundation, SciLifeLab, and the Swedish Research Council (VR grant 2014-453). The Teouma research by M.S. and S.B. was supported by the Australian Research Council (Discovery Grants DP0880789 and DP110101415), the National Geographic Society, and the Australia-Pacific Science Foundation. F.V. was supported by CNRS-UMR 7041. M.N. was supported by an Irish Research Council grant (GOIPD/2013/1). D.F. was supported by an Irish Research Council grant (GOIPG/2013/36). Q.F. was funded by the National Natural Science Foundation of China (L1524016), the Chinese Academy of Sciences Discipline Development Strategy Project (2015-DX-C-03) and the Bureau of International Cooperation of the Chinese Academy of Sciences. T.K. was supported by ERC starting grant FP7-261213. C.P. and J.K. were supported by the Baden Wuerttemberg Foundation. J.K was supported by the DFG grant KR 4015/1-1 and the Max Planck Society. R.P. was supported by ERC starting grant ADNABIOARC (263441). D.R. was supported by NIH grant GM100233, by NSF HOMINID BCS-1032255, and is a Howard Hughes Medical Institute investigator.

Author Contributions N.P., J.K., R.P. and D.R. supervised the study. M.S., F.V., S.B., G.A.C., and C.R. assembled archaeological material and information. P.S., C.P., Q.F., M.L., S.M., N.R. and D.R. analysed genetic data. C.P., K.Si., F.P., D.F., E.H., M.N., N.R, and K.St. performed laboratory work. S.A., M.P.C., F.R.F., J.S.F., T.K., G.K., P.K., D.A.M., F-X.R., and J.T.S.W. assembled the sample collection from present-day populations. P.S. and D.R. wrote the manuscript with major input from C.P., M.S., F.V., G.A.C., M.P.C., J.S.F, J.K. and R.P. and additional input from all other co-authors.

Author Information The aligned sequences are available through the European Nucleotide Archive under accession number PRJEB14728. The newly reported SNP genotyping data for the subset of individuals who provided informed consent consistent with fully public distribution are available at http://genetics.med.harvard.edu/reichlab/Reich_Lab/Datasets.html. To access data for the remaining samples, researchers should send a signed letter to D.R. containing the following text: "(a) I will not distribute the data outside my collaboration; (b) I will not post the data publicly; (c) I will make no attempt to connect the genetic data to personal identifiers for the samples; (d) I will use the data only for studies of population history; (e) I will not use the data for any selection studies; (f) I will not use the data for medical or disease-related analyses; (g) I will not use the data for commercial purposes." Extended Data Table 2 specifies which samples are consistent with which type of data distribution. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.S. (skoglund@genetics.med.harvard.edu), R.P. (ron.pinhasi@ucd.ie) or D.R. (reich@genetics.med.harvard.edu).

Reviewer Information *Nature* thanks P. Bellwood, C. Capelli and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Ancient DNA sampling, extraction, library preparation, enrichment and sequencing. The Vanuatu skeletal samples B30A, B10B, B17 were analysed with permission from the Vanuatu National Museum and the excavators of the Teouma site. The Tonga skeletal sample SK10 was analysed with permission from the excavators of the Talasiu site.

All preparation of skeletal samples, DNA extraction, and library preparation was carried out in dedicated ancient DNA laboratories at University College Dublin, Ireland (sample preparation of the three Vanuatu individuals), at Harvard Medical School in Boston, USA (DNA extraction and library preparation of the three Vanuatu individuals), and at the Max Planck Institute for the Science of Human History in Jena, Germany (sample preparation, DNA extraction and library preparation of the Tonga individual). Each of these facilities is spatially separated from other molecular biology laboratories, and measures are taken to protect ancient individuals from contamination including HEPA filtered air, head-to-toe suits, face masks with visors, multiple layers of gloves, bleaching of all surfaces, ultraviolet light (UVC) decontamination of (non-sensitive) consumables and chemicals, and UVC decontamination of the facility when researchers are not in the room[31]. The final step of the library preparation (amplification) was performed outside the ancient DNA laboratory.

We prepared powder from the cochlea of petrous bones, extracted DNA[12], and prepared libraries with standard protocols (ref. 13 for the Vanuatu individuals and ref. 32 for the Tonga individual). For the three Vanuatu individuals, the first library was prepared in the presence of uracil DNA glycosylase (UDG) to cut out errors due to ancient DNA damage, whereas the remaining three libraries as well as the Tonga library were prepared without UDG as this preserves more DNA for any given sample. We performed in-solution enrichment using previously reported protocols[13,14,33,34] for a targeted set of 1,237,207 SNPs that comprises two previously separately reported sets of 394,577 SNPs[34] and 842,630 SNPs[14]. We sequenced the product on an Illumina NextSeq500 instrument for 2×75 cycles. Following demultiplexing, and, for the Vanuatu samples, removal of both oligonucleotide barcodes that were used to identify the libraries and trailing adaptor sequences, we merged the forward and reverse reads of each read pair requiring a 15-base pair overlap (allowing one mismatch). We then aligned merged sequences to the human genome *hg19* using *BWA 0.6.1* (ref. 35). We removed sequences aligned to identical outer coordinates, choosing the highest quality sequence for each duplication cluster. We merged the data from the four libraries for each Vanuatu individual.

Genomic analysis. We determined sex by comparing the number of X and Y chromosome alignments[15]. We estimated damage patterns using *PMDtools* v0.60[18], separating damage patterns observed inside and outside a CpG context. Since all four individuals were female, we could not estimate contamination using X chromosome data. We investigated whether there was evidence of excess relatedness between any pair of individuals among the Vanuatu individuals, but found that the pairwise mismatch rate using panel 5 of the Affymetrix Human Origins array (see below) was $19.8\% \pm 0.4\%$ for I1368/I1369, $19.7\% \pm 0.6\%$ for I1368/I1370, and $20.5\% \pm 0.4\%$ for I1369/I1370. This suggests no atypical pair of individuals and a similar within-population mismatch rate (heterozygosity) as some present-day Polynesian populations (Fig. 2).

Genotyping of present-day humans. We genotyped 356 individuals from 38 southeast Asian and Oceanian populations on the Affymetrix Human Origins array (Extended Data Table 2). The individuals all contributed DNA samples voluntarily and provided informed consent consistent with studies of human genetic variation and history. Ethical approval of the component studies was provided by the Singapore Health IRB, the Research Ethics Committee at the Facultés de Médecine de Toulouse, the Brunei Medical and Health Research Ethics Committee, the University of Cambridge Biology Research Ethics Committee, the Government of Papua New Guinea Medical Research Advisory Committee, and the Temple University IRB. The collection of genome-wide variation data on de-identified samples was approved by the Harvard Human Research Protection Program (Protocol 11681), re-reviewed on 12 July 2016.

We restricted analysis to samples that had >95% genotyping completeness and that were not visual outliers in PCA with respect to the main cluster of samples in the group. We merged with previously reported Affymetrix Human Origins SNP array data[26,36–39]. We also co-analysed our data with samples genotyped on the Affymetrix 6.0 platform where we removed three previously published[39] Rapa Nui individuals (*5s5j*, *XB3B*, and *3p3p*), and two previously published[40] Samoan individuals (*PLY_07* and *PLY_11*), all of which appeared to have recent European ancestry based on clustering analyses. We finally compared our data to high-coverage genomes from an archaic Neanderthal and an archaic Denisovan, both from Denisova Cave in the Altai Mountains of Siberia[41–43].

Population genetic analysis. When overlapping with the Affymetrix Human Origins SNP array data set of present-day human populations, we have between 74,000 and 126,000 SNPs covered at least once for each of the four individuals (Extended Data Table 1). This is more than the minimum coverage required for high-resolution analysis using allele frequency correlation statistics, e.g. is 10,000 SNPs per individual according to Supplementary Information section 6.2 of ref. 44, a study that had the same median coverage (0.19×) as ours (the range in the present study is 0.14–0.26×). For all analyses, we called genotypes by randomly sampling a single non-duplicate sequence read at each position[45]. This procedure is standard for analysis of low-coverage ancient DNA data and is also often used for higher-coverage data to minimize reference genome biases that can be introduced when determining diploid genotypes[14,17,34,36,41,44–50]. For the *qpAdm, qpWave* and *qpGraph* analyses we excluded transition SNPs to avoid potential biases from postmortem damage (see below).

We performed PCA using *smartpca*[51], with the option *inbreed: YES* in order to sample a single genotype from each individual randomly to match the pseudo-haploid nature of the ancient DNA genotypes from the ancient individuals[52]. We computed f_3-, f_4- and D-statistics as in ref. 26, and F_{ST} using the Hudson estimator and randomly sampled a single haploid sequence to represent each individual at each SNP position, using *popstats*[38]. We estimated the date of admixture using ALDER[21]. We tested the consistency of a matrix of f_4-statistics with one or more sources of ancestry with respect to a set of outgroups (New_Guinea, Denisova, Sardinian, English, Yakut, Chukchi, Mala, Japanese, Ju_hoan_North, Mixe, Onge, Yoruba, and Mbuti) using *qpWave*[19,34].

For the ancient individuals and all present-day populations genotyped on the Human Origins array, we used *qpAdm*[34], which estimates ancestry proportions from two or more proxy source populations assuming that the proxies are more closely related to the real source populations than they are to a set of outgroups (*qpAdm* also provides a formal statistical test for whether this is the case, which passes in the context that we use it here). We estimated First Remote Oceanian and Papuan ancestry using Denisova, Sardinian, English, Yakut, Chukchi, Mala, Japanese, Ju_hoan_North, Mixe, Onge, Yoruba, and Mbuti as outgroups and New_Guinea and Ami as proxies for the Papuan and First Remote Oceanian source populations, respectively. For the ancient individuals, we excluded all transition SNPs to avoid possible biases due to post-mortem damage, resulting in 35,194 transversion SNPs for Vanuatu (covered by at least one of the individuals) and 22,030 for Tonga. For estimating *qpAdm* ancestry proportions in the Affymetrix 6.0 Polynesian data, we used whole-genome sequences from the same populations as outgroups[53]. We estimated Denisovan ancestry using the Denisovan genome and Japanese as the two sources, and chimpanzee (*Pan troglodytes*), Ju_hoan_North, Mbuti, Yoruba, Dinka and the Altai Neanderthal genome as outgroups.

We computed conditional heterozygosity using panel 5 of the Affymetrix Human Origins array, which contains SNPs ascertained as heterozygous in a single West African Yoruba individual. This provides an unbiased estimate of relative heterozygosity since the Yoruba individual is approximately symmetrically related to all Oceanians (Denisovan ancestry violates this assumption but is not expected to change the ranking of populations). We estimated heterozygosity as the average pairwise mismatch rate when sampling 2 chromosomes from two different individuals using *popstats*[38], restricting to transversion SNPs for all populations and computing standard errors using a Weighted Block Jackknife.

For authentication, we used *PMDtools*[18] to extract sequences with clear evidence of postmortem damage patterns (PMD score of at least 3), disregarding individual bases with phred-scaled base quality <30. We randomly sampled new pseudo-haploid genotypes from the resulting set of sequences and projected the ancient individuals onto the principal components inferred from the present-day populations as above. After this filtering, we retained 68,450 SNPs for I1368; 98,722 SNPs for I1369; 83,024 SNPs for I1370; and 117,023 SNPs for CP30. The ninety-nine per cent confidence intervals for *qpAdm* estimates of Papuan ancestry (see above) using the PMD score-restricted data were 0–21% for the ancient Vanuatu individuals and 0–24% for the ancient Tonga individual, consistent with the confidence intervals obtained from the full data.

To test whether the ancient Vanuatu and the ancient Tonga individuals form a clade, we used *qpWave* to test whether a model of Dai, Ami, Kankanaey and a fourth population were consistent with being outgroups to the two ancient sample groups (we used Dai, Ami and Kankanaey as these span present-day Mainland East Asia, Taiwan, and the Philippines, and lack Papuan ancestry to the limits of our resolution). The analysis used the ~12,000 SNPs that remained after excluding transition SNPs and SNPs missing in one of the two ancient sample groups. We found that the model was consistent with the data for all tested Oceanian and Asian populations shown in Fig. 1, but that the lowest P value was observed for present-day Tongans ($P = 0.09$). We also found that f_4(*Ami, Present day Tongan; Ancient Vanuatu,Ancient Tonga*) $= 0.006$, $Z = 3.2$, when using all SNPs. This suggests a possible affinity between present-day Tongans and the ancient Tongan individual, consistent with the hypothesis that the ancient population of Tonga with little or no Papuan ancestry may have contributed some of the ancestry of present-day Tongans.

Admixture date estimation. To estimate the date of historical admixture between First Remote Oceanians and Papuans, we used ALDER[21,25] on the full Human Origins array data, with New Guinean Highlanders and Han Chinese as the two sources. We use Han Chinese for this analysis owing to their substantial sample size compared to populations more closely related to the ancestral First Remote Oceanian population such as the ancient individuals we analysed, indigenous Taiwanese, and indigenous Philippine groups. ALDER estimates are known to be robust even when using imperfect surrogates for the ancestral populations in this way[26]. We estimate an admixture date for a pool of Polynesian populations by combining data from Tongan, Tikopia, Russell and Bellona populations, all genotyped on the Affymetrix Human Origins SNP array.

ALDER and other methods based on admixture linkage disequilibrium estimate dates in units of generations, which need to be converted to years. For this purpose, we require an estimate of the generation interval—the average age of a parent at the time their gametes were formed—weighted by the fraction of recombination events that occur in each sex (62.3% of all autosomal crossovers are estimated to occur in females, based on Table 1 of ref. 54.). Using estimates from the anthropological literature, this quantity is 27.8 years for hunter-gathering societies, 28.6 years for developed nation states, and 29.6 years for less developed nation states[22]. These numbers are in the range of the point estimate we use of 28.1 years based on breakdown of admixture linkage disequilibrium in radiocarbon-dated ancient genomes[55]. To account for the substantial variability in generation intervals across human societies, we use the sample standard error of 2.15 years measured across eleven diverse hunter-gatherer groups based on Table 4 of ref. 22. The date estimates in Fig. 2 and Extended Data Fig. 4 thus use a generation interval of 28.1 years, and combine the standard error from ALDER (a) with the uncertainty in generation time, that is, $\sqrt{a^2 \times 2.15^2 + A^2 \times 2.15^2 + 28.1^2 \times a^2}$, where A is the ALDER point estimate in number of generations.

We do not subtract 66 years from the dates produced by ALDER to obtain BP dates (conventionally the date before 1950 AD, 66 years ago), because what ALDER is estimating is a number that is close to the BP date. To see this, note that ALDER estimates the date between when chromosomes of the two ancestries began crossing over (one generation after mixing began), and the date of the last crossover (when the germ cells that mixed to produce the present-day samples in our study were formed, likely one or two generations before 2016 CE). Accounting for these corrections means that ALDER is estimating a date of mixture that is likely to be within a generation of the true BP date.

Fitting models of population history. We used *qpGraph*[26,56] to assess the fit of admixture graph models to allele frequency correlation patterns as measured by f_2, f_3-, and f_4-statistics. We started with a skeleton phylogenetic tree consisting of Yoruba, New_Guinea, Dai, Atayal, Kankanaey and the pool of ancient Vanuatu individuals. We added Tongan, Mamanwa (a Philippine Negrito group), Nasioi and Kolombangara, respectively, to all possible edges in the tree, and retained only the graph solutions that provided no individual f_4 statistics with $|Z| > 3$ between empirical and predicted statistics. For the extended version of the admixture graph, we also added Australians to all possible edges of the graph that included these populations. Finally, we modelled the previously documented admixture history relating Denisovans and the Altai Neanderthal genome to the outgroup chimpanzee and the anatomically modern human populations, to which we added the Andamanese Onge and the ancient Tongan individual. The final graph visualized in Fig. 3 used 10,893 SNPs after restricting to transversion SNPs to avoid complications due to ancient DNA damage and also SNPs with coverage in all groups. For more information on the admixture graph inference procedure, see Supplementary Information section 3.

As an alternative inference method, we used *Treemix* v1.12 (ref. 27) to test models for Yoruba, Dai, Atayal, Kankanaey, Tongan, New Guinean Highlanders, the ancient Vanuatu individual and the ancient Tongan individual. The total number of SNPs after excluding transitions, SNPs with minor allele count of less than 4 in the selected data, and SNPs where one population had missing data, was 10,119, which we divided into 337 blocks of 30 consecutive SNPs each to estimate the covariance matrix. We first fitted a maximum likelihood tree of all populations, but found that several of the fitted allele frequency covariances deviated from those empirically observed by up to 16.4 standard errors. We then used the automated heuristic optimization in *Treemix* to infer a graph model with one admixture event using the same populations, and found that the optimal fit was for a model with an admixture event in the history of Tongans, where one portion of their ancestry diverged before the split of the ancestors of the ancient Vanuatu and Tonga individuals, and the other (25% ± 3%) derived from the New Guinean lineage. This maximum deviation between empirical and model covariances observed for the graph with one admixture edge was 1.6, indicating a good fit, consistent with our investigation of models using *qpGraph*.

Female and male ancestral contributions. To estimate the proportion of female ancestors (F) and male ancestors (M) for a given population, we used two different methods both based on the estimates of ancestry for the X chromosome and autosomes. Both used the same underlying model, in which the observed admixture proportion estimates that $\hat{H}_{auto}$ and $\hat{H}_X$ for the autosomes and X chromosome, respectively, depend on M and F such that:

$$\hat{H}_{auto} = \left(\frac{M}{2} + \frac{F}{2}\right) \quad (1)$$

$$\hat{H}_X = \left(\frac{M}{3} + \frac{2F}{3}\right) \quad (2)$$

The first approach obtains unbounded point estimates of M and F by rearranging equations:

$$M = 4 \times \hat{H}_{auto} - 3 \times \hat{H}_X \quad (3)$$

$$F = 3 \times \hat{H}_X - 2 \times \hat{H}_{auto} \quad (4)$$

Similarly, we obtained standard errors for M and F using the weighted block jackknife standard errors for $\hat{H}_{auto}$ and $\hat{H}_X$, SE_{auto} and SE_X, as

$$\mathrm{SE}_M = \sqrt{(9 \times \mathrm{SE}_X{}^2) + (16 \times \mathrm{SE}_{auto}{}^2)} \quad (5)$$

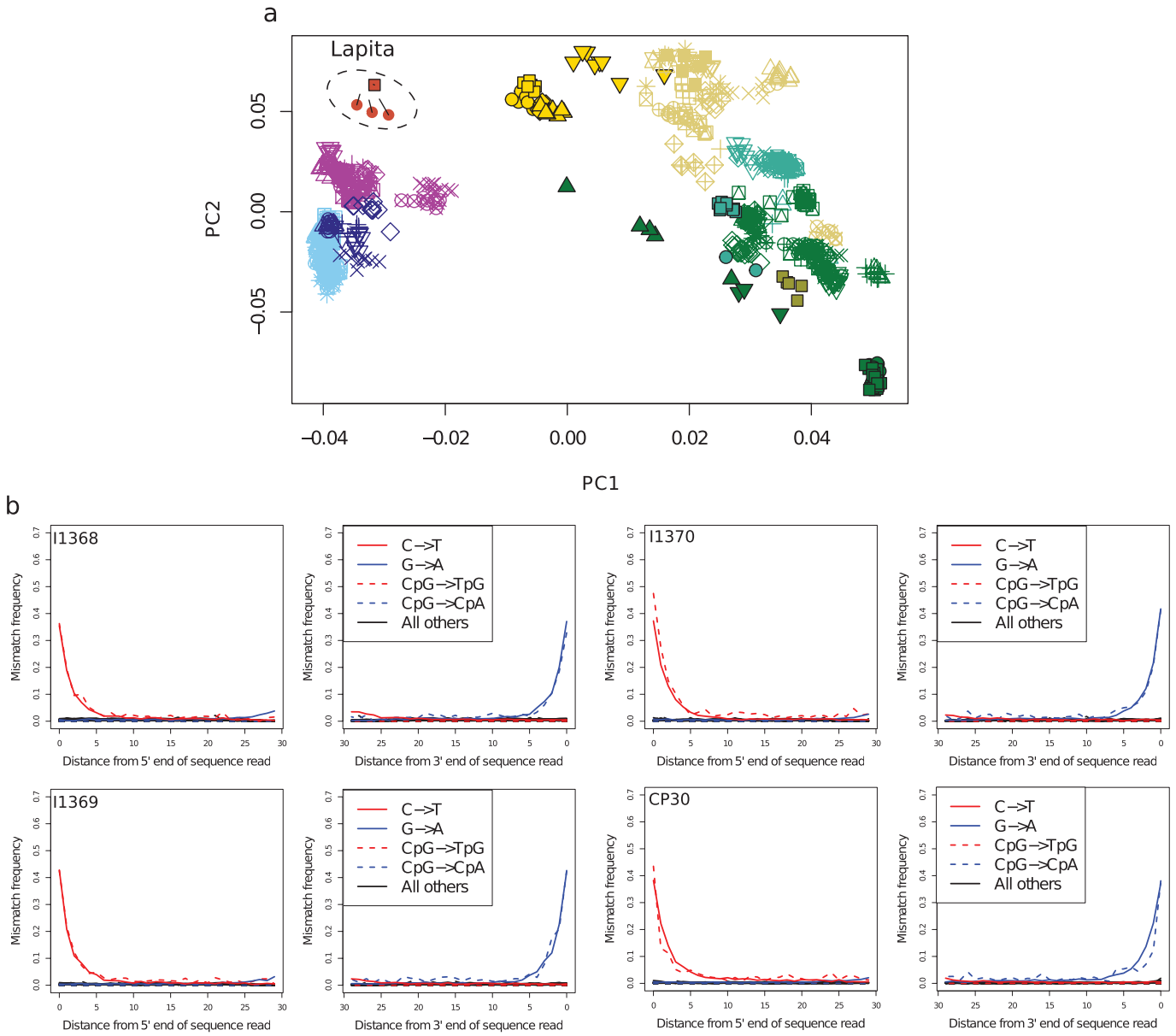
$$SE_F = \sqrt{(9 \times \mathrm{SE}_X{}^2) + (4 \times \mathrm{SE}_{auto}{}^2)} \quad (6)$$

As an alternative to estimating M and F, we took an approximate Bayesian approach by performing 1 million simulations in which M and F were sampled from a uniform prior distribution (0, 1). We then simulated ancestry estimates specifying normal distributions with means and standard errors matching the empirical values (equations 1 and 2). We used the *abc* R package[57] to run a rejection algorithm retaining the 1% of all simulation replicates with the closest Euclidean distances to the empirical $\hat{H}_{auto}$ and $\hat{H}_X$, and performed local linear regression on log-transformed summary statistics to obtain a posterior distribution. The results of the two methods are qualitatively similar. In Extended Data Fig. 6, we plot the posterior intervals of these distributions for selected populations.

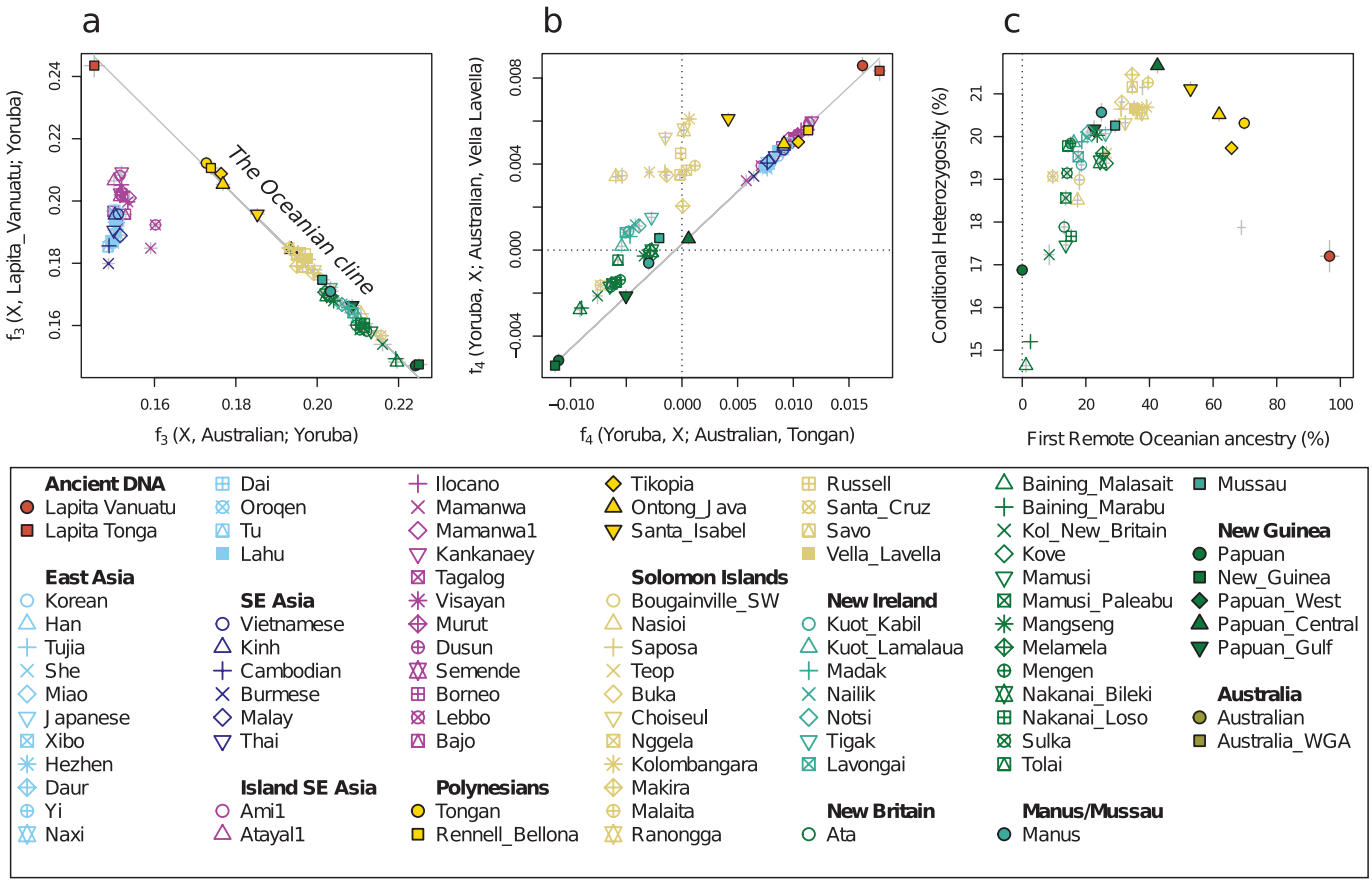
Sample size. No statistical methods were used to predetermine sample size.

31. Knapp, M., Clarke, A. C., Horsburgh, K. A. & Matisoo-Smith, E. A. Setting the stage – Building and working in an ancient DNA laboratory. *Ann. Anat. Anat. Anz.* **194,** 3–6 (2012).
32. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Prot.* http://dx.doi.org/10.1101/pdb.prot5448 (2010).
33. Fu, Q. *et al.* DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl Acad. Sci. USA* **110,** 2223–2227 (2013).
34. Haak, W. *et al.* Massive migration from the steppe is a source for Indo-European languages in Europe. *Nature* **522,** 207–211 (2015).
35. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).
36. Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513,** 409–413 (2014).
37. Qin, P. & Stoneking, M. Denisovan ancestry in East Eurasian and Native American populations. *Mol. Biol. Evol.* **32,** 2665–2674 (2015).
38. Skoglund, P. *et al.* Genetic evidence for two founding populations of the Americas. *Nature* **525,** 104–108 (2015).
39. Moreno-Mayar, J. V. *et al.* Genome-wide ancestry patterns in Rapanui suggest pre-European admixture with Native Americans. *Curr. Biol.* **24,** 2518–2525 (2014).
40. 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).
41. Reich, D. *et al.* Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468,** 1053–1060 (2010).
42. Meyer, M. *et al.* A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338,** 222–226 (2012).
43. Prüfer, K. *et al.* The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505,** 43–49 (2014).
44. Allentoft, M. E. *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522,** 167–172 (2015).
45. Green, R. E. *et al.* A draft sequence of the Neandertal genome. *Science* **328,** 710–722 (2010).
46. Raghavan, M. *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505,** 87–91 (2014).
47. Raghavan, M. *et al.* The genetic prehistory of the New World Arctic. *Science* **345,** 1255832 (2014).

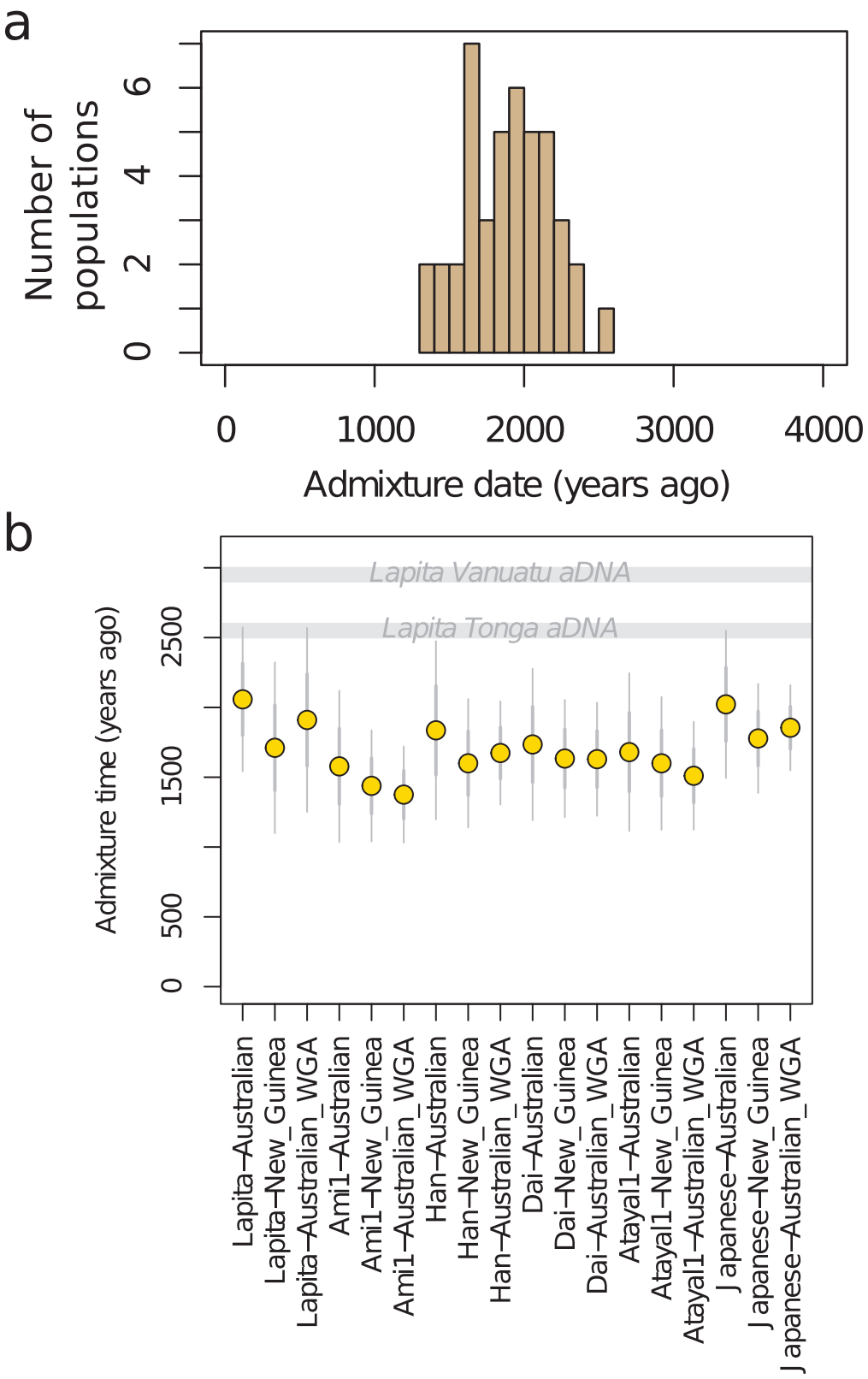
48. Rasmussen, M. *et al.* The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* **506,** 225–229 (2014).
49. Rasmussen, M. *et al.* The ancestry and affiliations of Kennewick Man. *Nature* **523,** 455–458 (2015).
50. Skoglund, P. *et al.* Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science* **344,** 747–750 (2014).
51. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2,** e190 (2006).
52. Skoglund, P. & Jakobsson, M. Archaic human ancestry in East Asia. *Proc. Natl Acad. Sci. USA* **45,** 18301–18306 (2011).
53. Sudmant, P. H. *et al.* Global diversity, population stratification, and selection of human copy-number variation. *Science* **349,** 1174–1181 (2015).
54. Kong, A. *et al.* A high-resolution recombination map of the human genome. *Nat. Genet.* **31,** 241–247 (2002).
55. 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).
56. Reich, D., Thangaraj, K., Patterson, N., Price, A. L. & Singh, L. Reconstructing Indian population history. *Nature* **461,** 489–494 (2009).
57. Csilléry, K., Blum, M. G. B., Gaggiotti, O. E. & François, O. Approximate Bayesian Computation (ABC) in practice. *Trends Ecol. Evol.* **25,** 410–418 (2010).
58. Bronk Ramsey, C. *OxCal Program v4.2.4.* (Radiocarbon Accelerator Unit, Univ. Oxford, 2016).
59. Reimer, P. J. *et al.* IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 Years cal BP. *Radiocarbon* http://dx.doi.org/10.2458/azu_js_rc.55.16947 (2013).
60. Ambrose, S. H. Isotopic analysis of paleodiets: methodological and interpretive considerations in *Food and Nutrition in History and Anthropology (USA)* (1993).
61. Petchey, F., 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).
62. Petchey, F., Anderson, A., Zondervan, A., Ulm, S. & Hogg, A. New marineΔ R values for the South Pacific subtropical gyre region. *Radiocarbon* **50,** 373–397 (2008).



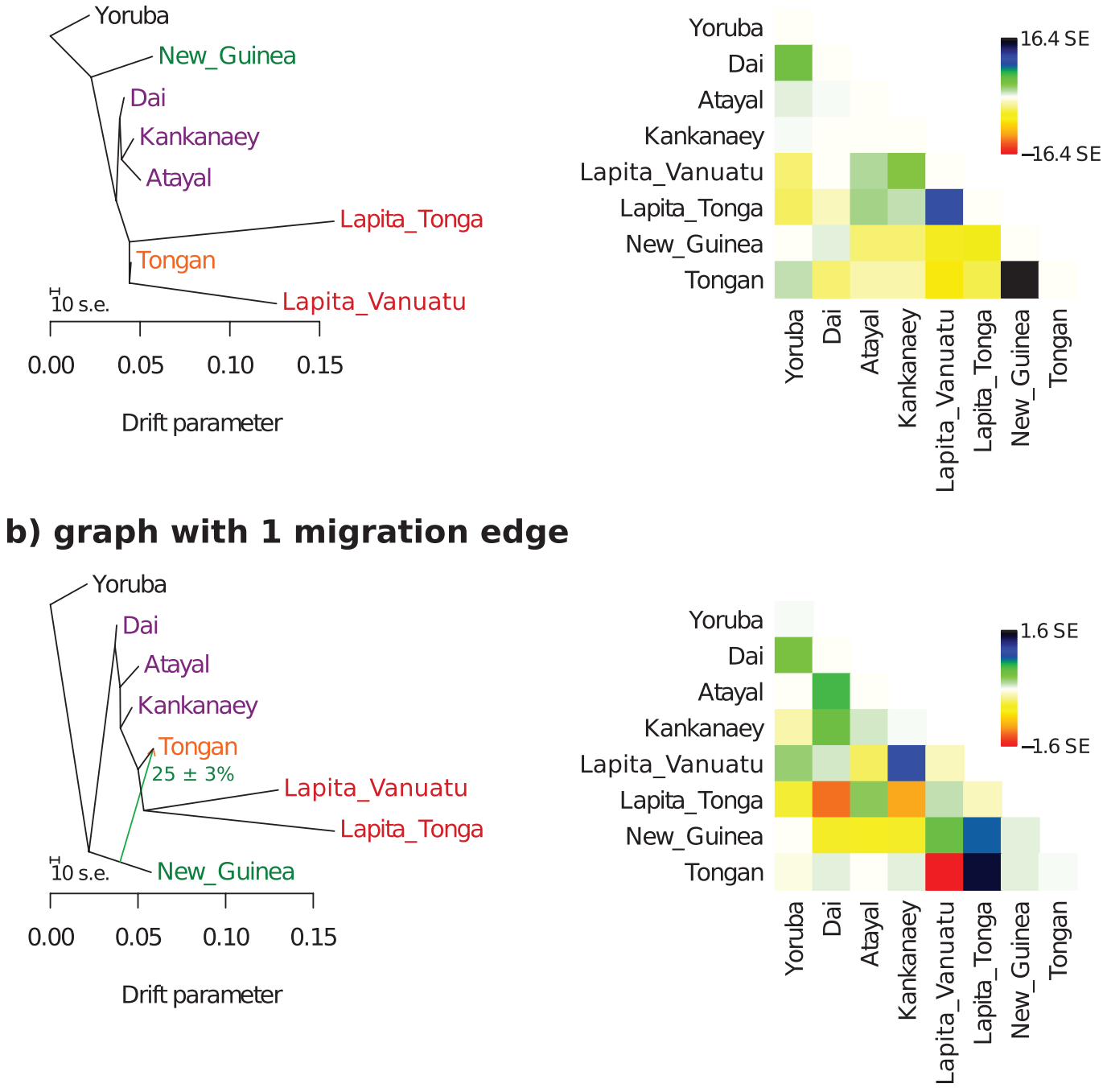
Extended Data Figure 1 | Ancient DNA authenticity. a, PCA performed as for Fig. 1, but with the four ancient individuals represented only by sequences that show clear evidence of postmortem damage (PMD score of at least 3) to remove contaminating sequences that might be present[17,18]. The numbers of SNPs remaining after restriction to damaged sequences is 68,450 SNPs for I1368; 98,722 SNPs for I1369; 83,024 SNPs for I1370; and 117,023 SNPs for CP30. The lines indicate the projection of the samples when no damage-restriction is performed. The large number of SNPs retained, and the fact that the ancient individuals cluster tightly and have the same qualitative positioning in the plot as Fig. 1, indicates that contamination did not contribute to our findings. We also find that estimates of Papuan ancestry using PMD score restricted data are consistent with those obtained using the full data (see Methods). **b**, Postmortem damage patterns for genome-wide in-solution enrichment data from four ancient individuals.



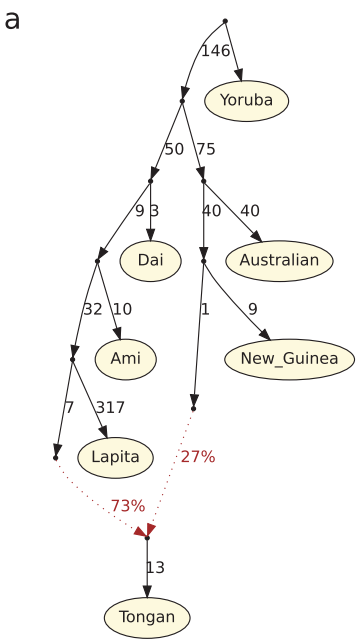
Extended Data Figure 2 | *f*-statistics document the Oceanian ancestry cline. a, Shared genetic drift with the ancient Vanuatu individuals is negatively related to shared drift with Australians. Except for the ancient Tongan individual, populations from Taiwan, the Philippines and Polynesia share the most genetic drift with the ancient Vanuatu individuals, who are not shown in the plot because they are used as reference in the computation. The trend line was fitted without the East Asian populations in the off-cline cluster. The absence of off-cline Oceanian individuals suggests the possibility that present-day Oceanians may largely be derived from a mixture of two source populations. **b**, The ancient Vanuatu individuals and the ancient Tongan individual maximize statistics of the form f_4(Yoruba, *Test*; Australian, Oceanian), suggesting that they are the most closely related to the East Asian ancestry in Oceanians of any sampled population. The trend line was fitted using populations >0.005 on the *x*-axis, together with the two populations with the lowest values on the *x*-axis (Papuan and New_Guinea). **c**, Biplot of First Remote Oceanian ancestry proportions against conditional heterozygosity. Populations with intermediate admixture proportion show the greatest genetic diversity. Thick and thin error bars in all panels are 1 and 1.96 standard errors of the estimate, respectively.

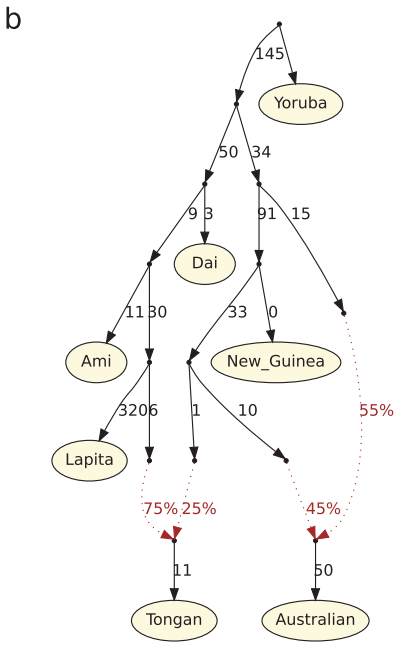


Extended Data Figure 3 | Admixture date estimates. a, Histogram of the point estimate dates in Fig. 2d. **b**, Admixture date estimates for Tongans using different pairs of source populations ('Lapita' in this figure refers to the pool of ancient Vanuatu individuals). Error bars show 1 (thick whiskers) and 1.96 (thin whiskers) standard errors. WGA, whole-genome amplified DNA.

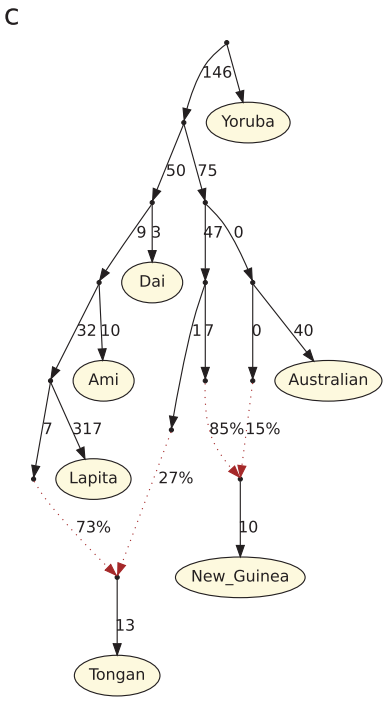


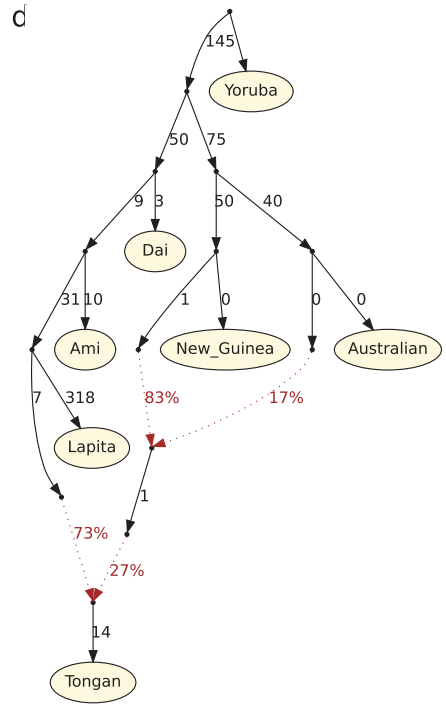
Extended Data Figure 4 | Admixture graph inferred using *Treemix*. **a**, A simple tree-like model without admixture fits the data poorly, as can be seen from the matrix of residuals between empirical and modelled allele frequency covariance on the right. **b**, The optimal placement of a single 25% admixture event is from the lineage related to New Guinean Highlanders into the lineage leading to Tongans. Tongans derive the other portion of their ancestry from the lineage leading to the two ancient groups of individuals. This graph has no significant deviations between empirical and modelled allele frequency covariances.





Outliers	fitted	empirical	z
f4(New_Guinea, Australian; Dai, Tongan)	-0.010860	-0.006083	4.768
f4(New_Guinea, Australian; Tongan, Ami)	0.010860	0.006205	-4.431
f4(New_Guinea, Australian; Tongan, Lapita)	0.010860	0.005384	-3.113
f4(Dai, Tongan; Tongan, Australian)	-0.017124	-0.012147	3.653
f4(Tongan, Ami; Tongan, Australian)	0.010509	0.005834	-3.395

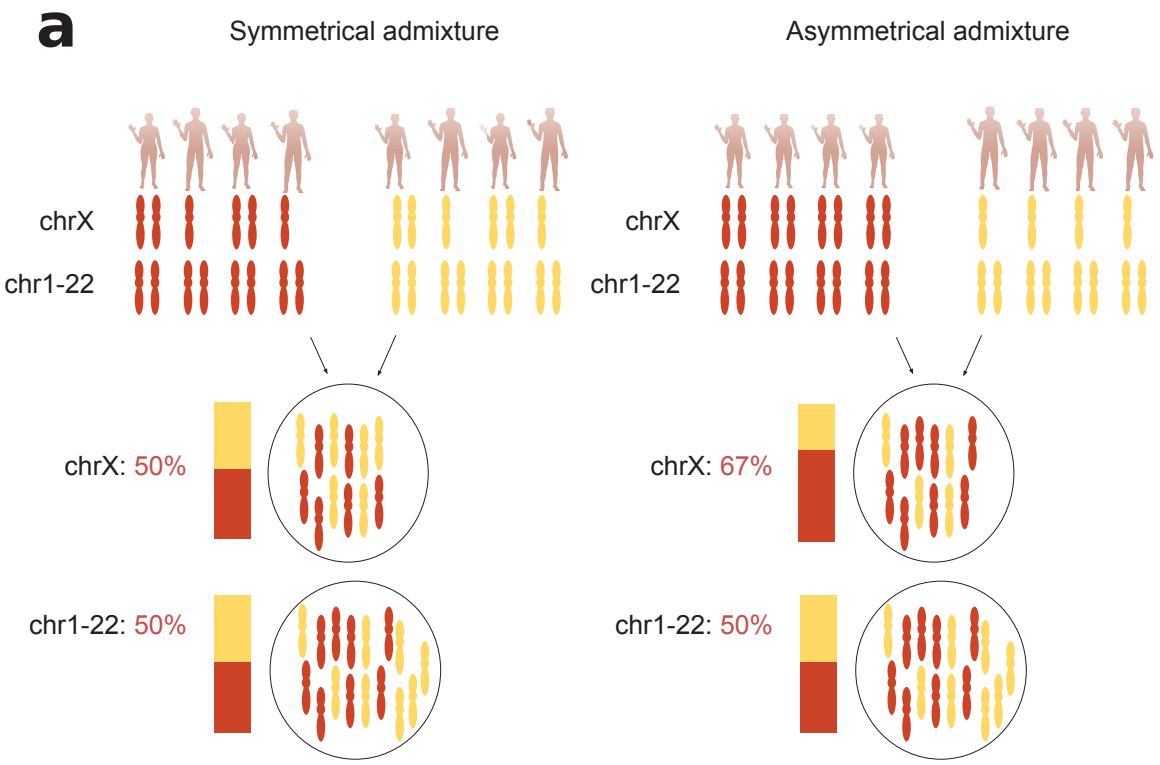


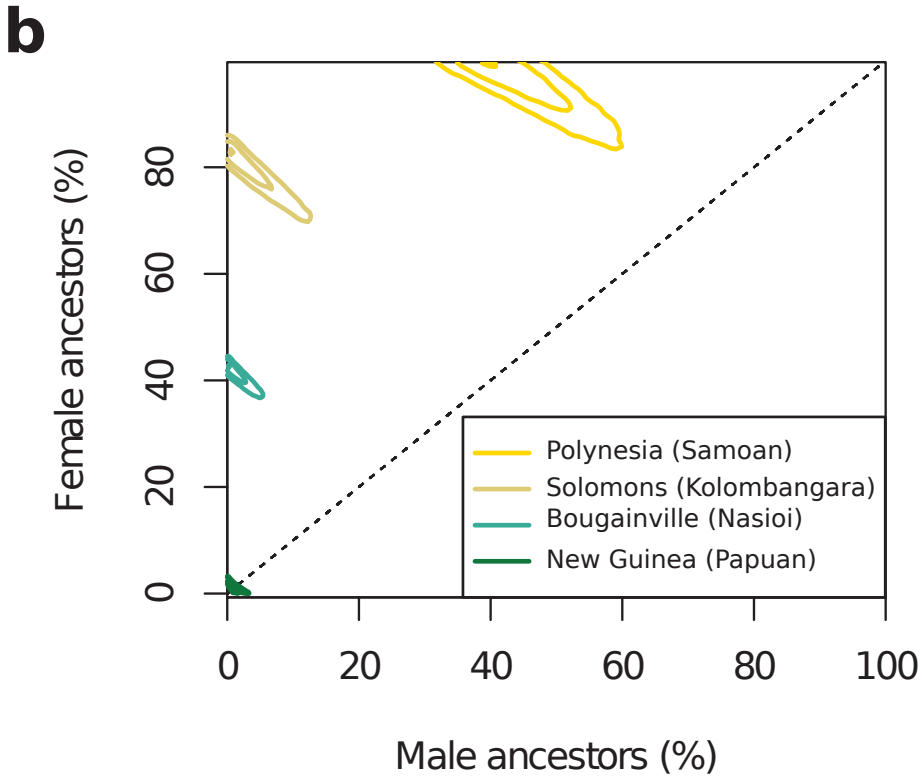


Outliers	fitted	empirical	z
f4(New_Guinea, Australian; Dai, Tongan)	-0.010850	-0.006083	4.758
f4(New_Guinea, Australian; Tongan, Ami)	0.010850	0.006205	-4.422
f4(New_Guinea, Australian; Tongan, Lapita)	0.010850	0.005384	-3.107
f4(Dai, Tongan; Tongan, Australian)	-0.017123	-0.012147	3.652
f4(Tongan, Ami; Tongan, Australian)	0.010509	0.005834	-3.395

Outlier	fitted	empirical	z
f4(New_Guinea, Australian; Dai, Tongan)	-0.009261	-0.006083	3.172

Extended Data Figure 5 | Admixture graphs modelling the population history of Australians. Outlier f_4-statistics are shown ($|Z| > 3$). **a**, A model with a single admixture edge positing that Australians are an outgroup to the Papuan ancestry in Tongans does not fit the data (5 outlier statistics). **b**, An alternative model with 2 admixture edges in which the Papuan ancestry in Tongans also contributed to Australians fits the data (no outliers). **c**, A model with 2 admixture edges in which New Guinean Highlanders are admixed from an Australian source after the divergence of the Papuan source in Tongans does not fit the data (5 outliers). **d**, A model with 2 admixture edges in which the Papuan ancestry in Tongans is intermediate between the New Guinean Highlander lineage and the Australian lineage. Branch lengths are in units of $F_{ST} \times 1{,}000$. Lapita in this figure refers only to Vanuatu, which is the only group for which we have multiple individuals (needed to compute F_{ST}).





Extended Data Figure 6 | First Remote Oceanian ancestry today comes primarily from females. a, Illustration of the rationale for using the X chromosome to study asymmetrical admixture between males and females. The example on the left illustrates admixture with equal proportion of males and females in both the red and the yellow ancestral population. The example on the right illustrates an extreme case of asymmetrical admixture where the red ancestral population only contributes females and the yellow ancestral population only contributes males to the admixed generation, demonstrating the disproportional contribution of X chromosomes by females to the admixed population. **b**, Female and male ancestral contributions based on an admixture model fitted to estimated ancestry proportions on the autosomes and X chromosome. We show the 95%, 70%, and 5% highest posterior intervals for four selected populations from Polynesia (Samoans), the Solomon Islands (Kolombangara), Bougainville (Nasioi), and mainland New Guinea (Papuans).

Extended Data Table 1 | In-solution DNA enrichment and sequencing of ancient individuals

Sample information						Coverage on chromosomes 1-22			Sex determination		
Location	ID1	ID2	Bone for aDNA	Bone for dating	^{14}C Date: Calibrated 95.4% Conf. Int. (Uncalibrated date, Lab number)	Mean depth	All SNPs	SNPs overlapping array	Y SNPs	X SNPs	Sex
Vanuatu	I1368	B30A	Petrous	Skull	2990-2740 BP (2983±32 BP, Wk-22657)	0.26	139,461	74,631	321	18,231	F
Vanuatu	I1369	B10B	Petrous	Petrous	3000-2750 BP (3045±30 BP, Poz-81126)	0.14	199,500	107,523	341	24,255	F
Vanuatu	I1370	B17	Petrous	Skull	3110-2780 BP (3083±26 BP, Wk-21026)	0.21	167,311	90,402	231	19,303	F
Tonga	CP30	SK10	Petrous	Fibula	2680-2340 BP (2594±20 BP, Wk-41883)	0.16	231,994	125,908	75	25,943	F

All dates are calibrated using OxCal v4.2.4[58] with a mixture of the Marine13 and Intcal13 curves[59] as determined by linear interpolation between dietary terrestrial and marine δ^{13}C isotopic endpoints (−21‰/−12‰) with an uncertainty of ±10% on the per cent marine carbon result following previous recommendations[60]. Two of the dates have been previously reported (for I1368/B30A and I1370/B17)[61], and in this study we add two new dates: for I1369/B10B from Tonga (on the same petrous bone used for ancient DNA analysis) and on CP30/SK10 from Tonga (on a fibula). Measured ^{13}C and ^{15}N values for I1369/B10B are −14.5‰ and 13.7‰ respectively, and for SK10 −16.44‰ and 10.48‰. As justified in ref. 61, we also applied a location-specific reservoir correction (ΔR) of 40 ± 44 ^{14}C years to the marine curve to adjust for regional oceanic variation in ^{14}C around Vanuatu, and 11 ± 83 ^{14}C years for Tongatapu[62].

Extended Data Table 2 | 356 individuals newly genotyped on the Human Origins Array

Population	N	Country of origin	Land mass	Language	Lat.	Long.	Co-authors for samples	Protocol Numbers for informed consent	Data distribution
Ata	8	Papua New Guinea	New Britain	Papuan	-5.7	150.9	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Baining_Malasait	5	Papua New Guinea	New Britain	Papuan	-4.47	151.9	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Baining_Marabu	10	Papua New Guinea	New Britain	Papuan	-4.63	152.3	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Bajo	10	Indonesia	Sulawesi	Austronesian	-3.97	122.59	M.P.C, P.K., F.-X.R.	4.13.2013 approval AMIS-UPS Ethics Committee	Signed Letter
Buka	8	Papua New Guinea	Bougainville	Austronesian	-5.42	154.67	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Burmese	10	Myanmar	Asia	Sino-Tibetan	16.41	95.89	T.K., J.T.S.W.	2010/820/A and HBREC.2011.01	Fully public
Dusun	10	Brunei	Borneo	Austronesian	4.71	114.67	T.K., S.A.	MHREC/EDU/2012/3(1) and HBREC.2011.01	Fully public
Ilocano	2	Philippines	Luzon	Austronesian	14.6	120.98	T.K., J.T.S.W.	2010/820/A and HBREC.2011.01	Fully public
Kankanaey	10	Philippines	Luzon	Austronesian	17.07	121.03	T.K., J.T.S.W.	2010/820/A and HBREC.2011.01	Fully public
Kol_New_Britain	2	Papua New Guinea	New Britain	Papuan	-5.38	151.63	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Kove	18	Papua New Guinea	New Britain	Austronesian	-5.47	148.95	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Kuot_Kabil	9	Papua New Guinea	New Ireland	Papuan	-3.07	151.7	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Kuot_Lamalaua	8	Papua New Guinea	New Ireland	Papuan	-3	151.5	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Lavongai	15	Papua New Guinea	New [illegible]	Austronesian	-2.57	150.43	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Lebbo	8	Indonesia	Borneo	Austronesian	1.66	117.16	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Madak	10	Papua New Guinea	New Ireland	Austronesian	-3.1	151.7	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Malay	9	Singapore	Asia	Austronesian	1.35	103.82	T.K., J.T.S.W.	2010/820/A and HBREC.2011.01	Fully public
Mamusi	20	Papua New Guinea	New Britain	Austronesian	-6	151	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Mamusi_Paleabu	6	Papua New Guinea	New Britain	Austronesian	-5.95	150.9	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Mangseng	6	Papua New Guinea	New Britain	Austronesian	-5.93	150.7	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Manus	2	Papua New Guinea	Manus	Austronesian	-2.08	147	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Melamela	10	Papua New Guinea	New Britain	Austronesian	-5	151.25	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Mengen	10	Papua New Guinea	New Britain	Austronesian	-5.1	151.4	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Murut	10	Brunei	Borneo	Austronesian	4.62	115.14	T.K., J.T.S.W.	MHREC/EDU/2012/3(1) and HBREC.2011.01	Fully public
Mussau	10	Papua New Guinea	St. Matthias	Austronesian	-1.58	149.73	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Nakanai_Bileki	10	Papua New Guinea	New Britain	Austronesian	-5.75	150.8	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Nakanai_Loso	7	Papua New Guinea	New Britain	Austronesian	-5.48	150.8	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Nailik	9	Papua New Guinea	New Ireland	Austronesian	-2.98	151.52	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Notsi	9	Papua New Guinea	New Ireland	Austronesian	-3.05	151.65	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Saposa	10	Papua New Guinea	Bougainville	Austronesian	-5.58	154.67	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Southwest_Bougainville	2	Papua New Guinea	Bougainville	Papuan	-6.6	155.5	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Sulka	20	Papua New Guinea	New Britain	Papuan	-4.5	152.3	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Tagalog	5	Philippines	Luzon	Austronesian	14.6	120.98	T.K., J.T.S.W.	2010/820/A and HBREC.2011.01	Fully public
Teop	10	Papua New Guinea	Bougainville	Austronesian	-5.85	155.18	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Tigak	10	Papua New Guinea	New Ireland	Austronesian	-2.57	150.83	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Tolai	24	Papua New Guinea	New Britain	Papuan	-4.31	152.14	F.R.F., J.S.F., G.K., D.A.M.	MRAC.1998.2000.2010 and 99-226.4320	Signed Letter
Vietnamese	10	Vietnam	Asia	Austroasiatic	10.82	106.64	T.K., J.T.S.W.	2010/820/A and HBREC.2011.01	Fully public
Visayan	4	Philippines	Mindanao	Austronesian	9.76	125.51	T.K., J.T.S.W.	2010/820/A and HBREC.2011.01	Fully public

Extended Data Table 3 | *f*-statistics for populations on the Oceanian cline

	F_{ST}(Lapita_Vanuatu, Test)		f_3(Lapita_Vanuatu, Australian; Test)	
Test population	**Estimate**	**SE**	**Estimate**	**Z-score**
Baining_Malasait	0.263	0.005	0.143	38.4
Baining_Marabu	0.249	0.004	0.115	40.1
Papuan	0.225	0.004	0.066	33.3
Kol_New_Britain	0.216	0.006	0.086	16.7
Mamusi	0.204	0.004	0.059	26.5
Ata	0.197	0.004	0.050	21.6
Nakanai_Loso	0.194	0.004	0.055	21.9
Mamusi_Paleabu	0.187	0.004	0.040	17.8
Santa_Cruz	0.185	0.004	0.021	10.7
Nasioi	0.178	0.004	0.028	14.7
Bougainville_South	0.176	0.006	0.034	8.1
Sulka	0.174	0.004	0.022	13.1
Mengen	0.168	0.004	0.017	9.1
Tolai	0.163	0.004	0.009	5.6
Kuot_Kabil	0.162	0.004	0.014	7.7
Lavongai	0.160	0.004	0.009	5.2
Kuot_Lamalaua	0.160	0.004	0.008	4.3
Nakanai_Bileki	0.153	0.004	0.013	7.0
Melamela	0.152	0.004	0.010	5.1
Madak	0.151	0.004	0.003	1.8
Papuan_Gulf	0.151	0.005	0.002	0.7
Kove	0.150	0.004	0.013	7.4
Mangseng	0.145	0.004	0.003	1.7
Nailik	0.145	0.004	0.000	-0.1
Teop	0.144	0.004	0.005	2.8
Notsi	0.144	0.004	-0.002	-1.4
Manus	0.141	0.006	-0.005	-1.4
Tigak	0.141	0.004	0.002	1.2
Mussau	0.132	0.004	-0.005	-3.3
Choiseul	0.123	0.004	-0.007	-3.7
Saposa	0.122	0.004	-0.011	-6.5
Buka	0.118	0.004	-0.016	-9.9
Vella_Lavella	0.112	0.004	-0.016	-8.6
Ranongga	0.110	0.004	-0.015	-8.2
Savo	0.109	0.004	-0.020	-12.1
Russell	0.108	0.005	-0.017	-6.8
Kolombangara	0.108	0.004	-0.015	-8.0
RenBel	0.106	0.004	0.035	14.0
Gela	0.103	0.004	-0.024	-14.3
Makira	0.101	0.004	-0.024	-16.1
Malaita	0.096	0.004	-0.025	-15.5
Papuan_Central	0.094	0.004	-0.031	-18.2
Bajo	0.082	0.004	0.022	12.8
Isabel	0.079	0.004	-0.024	-15.9
Tikopia	0.077	0.004	-0.003	-1.3
Ontong_Java	0.069	0.004	-0.018	-10.4
Tongan	0.053	0.004	-0.018	-9.9

Standard error (SE) is shown for F_{ST} between each *Test* population and the pool of ancient Vanuatu individuals. The *Z* score is given for the statistic f_3(Ancient Vanuatu, Australian; *Test*), where $Z < 3$ provides significant evidence that the *Test* is admixed between sources related to the ancient Vanuatu and Australians.

Extended Data Table 4 | Ancestry estimates for populations on the Oceanian cline

	First Remote Oceanian ancestry estimate							Anc. contrib. (method of moments)				Two source model
Test	Auto.	SE	chrX	SE	Diff.	SE	Z	Male	SE	Female	SE	*P*-value
Tongan	69.8%	1.0%	104.1%	8.9%	-34.3%	9.0%	-3.8	-33.1%	27.0%	172.7%	22.0%	0.22
Rennel & Bellona	68.9%	1.2%	92.1%	7.3%	-23.2%	7.4%	-3.1	-0.7%	22.4%	138.5%	32.8%	0.89
Tikopia	65.8%	1.0%	93.9%	10.9%	-28.1%	10.9%	-2.6	-18.5%	32.9%	150.1%	39.9%	0.58
Ontong_Java	61.9%	0.9%	78.4%	13.3%	-16.5%	13.3%	-1.2	12.4%	40.1%	111.4%	12.7%	0.83
Santa Isabel	52.9%	0.9%	60.5%	4.2%	-7.6%	4.3%	-1.8	30.1%	13.1%	75.7%	14.2%	0.87
Papuan_Central	42.5%	0.9%	65.6%	4.7%	-23.1%	4.8%	-4.8	-26.8%	14.6%	111.8%	47.7%	0.14
Malaita	39.6%	0.9%	66.1%	15.9%	-26.5%	15.9%	-1.7	-39.9%	47.8%	119.1%	20.8%	0.25
Kolombangara	39.1%	1.0%	54.7%	6.9%	-15.6%	7.0%	-2.2	-7.7%	21.1%	85.9%	16.3%	0.87
Nggela	37.8%	0.9%	54.6%	5.4%	-16.8%	5.5%	-3.1	-12.6%	16.6%	88.2%	82.5%	0.59
Ranongga	37.6%	1.0%	48.8%	27.5%	-11.2%	27.5%	-0.4	4.0%	82.6%	71.2%	29.8%	0.15
Russell	36.2%	1.2%	45.8%	9.9%	-9.6%	10.0%	-1.0	7.4%	30.1%	65.0%	26.2%	0.19
Vella_Lavella	35.2%	1.0%	53.2%	8.7%	-18.0%	8.8%	-2.1	-18.8%	26.4%	89.2%	19.9%	0.10
Savo	34.6%	0.9%	56.4%	6.6%	-21.8%	6.7%	-3.3	-30.8%	20.1%	100.0%	38.4%	0.18
Makira	34.6%	0.9%	31.8%	12.8%	2.8%	12.8%	0.2	43.0%	38.6%	26.2%	13.4%	0.08
Choiseul	32.4%	1.0%	35.5%	4.4%	-3.1%	4.5%	-0.7	23.1%	13.8%	41.7%	9.8%	0.20
Buka	31.3%	0.9%	46.0%	3.2%	-14.7%	3.3%	-4.4	-12.8%	10.3%	75.4%	16.9%	0.52
Saposa	31.1%	0.9%	41.8%	5.6%	-10.7%	5.7%	-1.9	-1.0%	17.2%	63.2%	14.2%	0.25
Mussau	29.2%	0.9%	39.4%	4.7%	-10.2%	4.8%	-2.1	-1.4%	14.6%	59.8%	22.0%	0.54
Teop	26.5%	0.9%	67.8%	7.3%	-41.3%	7.4%	-5.6	-97.4%	22.2%	150.4%	11.5%	0.86
Kove	26.4%	0.9%	29.8%	3.8%	-3.4%	3.9%	-0.9	16.2%	12.0%	36.6%	20.2%	0.01
Tigak	26.2%	0.9%	43.9%	6.7%	-17.7%	6.8%	-2.6	-26.9%	20.4%	79.3%	21.4%	0.87
Melamela	25.3%	0.9%	50.1%	7.1%	-24.8%	7.2%	-3.5	-49.1%	21.6%	99.7%	18.8%	0.94
Manus	24.9%	1.2%	14.2%	6.2%	10.7%	6.3%	1.7	57.0%	19.2%	-7.2%	42.9%	0.87
Nakanai_Bileki	24.5%	0.9%	38.4%	14.3%	-13.9%	14.3%	-1.0	-17.2%	43.1%	66.2%	11.2%	0.28
Mangseng	23.6%	0.9%	42.2%	3.7%	-18.6%	3.8%	-4.9	-32.2%	11.7%	79.4%	19.9%	0.34
Papuan_Gulf	22.6%	1.1%	49.3%	6.6%	-26.7%	6.7%	-4.0	-57.5%	20.3%	102.7%	10.9%	0.07
Notsi	20.7%	0.8%	33.0%	3.6%	-12.3%	3.7%	-3.3	-16.2%	11.3%	57.6%	64.8%	0.13
Nailik	20.4%	0.8%	44.8%	21.6%	-24.4%	21.6%	-1.1	-52.8%	64.9%	93.6%	22.9%	0.08
Madak	18.7%	0.8%	43.1%	7.6%	-24.4%	7.6%	-3.2	-54.5%	23.0%	91.9%	18.7%	0.11
Kuot_Kabil	18.6%	0.9%	55.4%	6.2%	-36.8%	6.3%	-5.9	-91.8%	18.9%	129.0%	26.5%	0.04
Bougainville_Sout	18.0%	1.4%	24.1%	8.8%	-6.1%	8.9%	-0.7	-0.3%	27.0%	36.3%	11.8%	0.26
Lavongai	17.6%	0.8%	36.4%	3.9%	-18.8%	4.0%	-4.7	-38.8%	12.1%	74.0%	16.0%	0.07
Nasioi	17.4%	1.0%	29.9%	5.3%	-12.5%	5.4%	-2.3	-20.1%	16.4%	54.9%	11.5%	0.40
Kuot_Lamalaua	16.2%	0.8%	13.8%	3.8%	2.4%	3.9%	0.6	23.4%	11.8%	9.0%	19.6%	0.07
Nakanai_Loso	15.4%	1.1%	24.3%	6.5%	-8.9%	6.6%	-1.4	-11.3%	20.0%	42.1%	11.2%	0.37
Mengen	15.2%	0.8%	35.3%	3.7%	-20.1%	3.8%	-5.3	-45.1%	11.6%	75.5%	13.0%	0.41
Tolai	14.3%	0.8%	22.5%	4.3%	-8.2%	4.4%	-1.9	-10.3%	13.3%	38.9%	9.4%	0.01
Sulka	14.1%	0.8%	47.0%	3.1%	-32.9%	3.2%	-	-84.6%	9.8%	112.8%	16.3%	0.48
Mamusi_Paleabu	13.7%	1.0%	30.9%	5.4%	-17.2%	5.5%	-3.1	-37.9%	16.7%	65.3%	13.9%	0.21
Mamusi	13.7%	0.9%	22.7%	4.6%	-9.0%	4.7%	-1.9	-13.3%	14.3%	40.7%	27.1%	0.35
Ata	13.2%	1.0%	26.4%	9.0%	-13.2%	9.1%	-1.5	-26.4%	27.3%	52.8%	15.4%	0.24
Santa_Cruz	9.6%	0.9%	27.1%	5.1%	-17.5%	5.2%	-3.4	-42.9%	15.7%	62.1%	23.3%	0.17
Kol_New_Britain	8.5%	1.4%	9.1%	7.7%	-0.6%	7.8%	-0.1	6.7%	23.8%	10.3%	30.1%	0.66
Baining_Marabu	2.6%	1.0%	25.3%	10.0%	-22.7%	10.0%	-2.3	-65.5%	30.3%	70.7%	32.5%	0.24
Baining_Malasait	1.2%	1.2%	34.9%	10.8%	-33.7%	10.9%	-3.1	-99.9%	32.8%	102.3%	8.2%	0.12
Papuan	0.0%	0.5%	2.4%	2.7%	-2.4%	2.7%	-0.9	-7.2%	8.3%	7.2%	0.0%	0.58

Auto., estimate on the autosomes (chromosomes 1–22). Diff., difference between the autosome and X chromosome estimates.

RESEARCH LETTER

Author Queries

Journal: **Nature**
Paper: **nature19844**
Title: **Ancient genomics and the peopling of the Southwest Pacific**

Query Reference	Query
1	AUTHOR: A PDF proof will be produced on the basis of your corrections to this preproof and will contain the main-text figures edited by us and the Extended Data items supplied by you (which may have been resized but will not have been edited otherwise by us). When you receive the PDF proof, please check that the display items are as follows (doi:10.1038/nature19844): Figs none (black & white); 1–3 (colour); Tables: None; Boxes: None; Extended Data display items: Figs 1–6, Tables 1–4. Please check the edits to all main-text figures (and tables, if any) very carefully, and ensure that any error bars in the figures are defined in the figure legends. If you wish to revise the Extended Data items for consistency with main-text figures and tables, please copy the style shown in the PDF proof (such as italicising variables and gene symbols, and using initial capitals for labels) and return the revised Extended Data items to us along with your proof corrections.
2	AUTHOR: Should ref. 21 be Loh, P.-R. *et al.* Inferring admixture histories of human populations using linkage disequilibrium. *Genetics* 193, 1233–1254 (2013)? If not please provide full citation details.
3	AUTHOR: Please provide publisher of ref. 60.
Web summary	Analysis of ancient DNA from five individuals who lived in Vanuatu and Tonga between 2,300 and 3,100 years ago suggests that the Papuan ancestry seen in present-day occupants of this region was introduced at a later date.

For Nature office use only:

Layout	☐	Figures/Tables/Boxes	☐	References	☐
DOI	☐	Error bars	☐	Supp info	☐
Title	☐	Colour	☐	Acknowledgements	☐
Authors	☐	Text	☐	Author contribs	☐
Addresses	☐	Methods	☐	COI	☐
First para	☐	Received/Accepted	☐	Correspondence	☐
		AOP	☐	Author corrx	☐
		Extended Data	☐	Web summary	☐
				Accession codes link	☐
				Referee accreditation	☐

SUBJECT WORDS
Biological sciences/Genetics/Population genetics/Genetic variation [URI /631/208/457/649];Biological sciences/Genetics/Evolutionary biology [URI /631/208/182].

TECHNIQUE TERMS
Not Applicable.