Title

Genetic Ancestry of Rapanui before and after European Contact

Journal

Current Biology, 27(20)

ISSN

0960-9822

Authors

Fehren-Schmitz, Lars
Jarman, Catrine L
Harkins, Kelly M
et al.

Publication Date

2017-10-01

DOI

10.1016/j.cub.2017.09.029

Peer reviewed

Report

Current Biology

Genetic Ancestry of Rapanui before and after European Contact

Highlights

- We sequenced the first low-coverage genomes from pre-European-contact Rapanui
- The individuals fall within Polynesian genetic diversity
- We found no Native American admixture in pre- and post-European-contact individuals

Authors

Lars Fehren-Schmitz, Catrine L. Jarman, Kelly M. Harkins, Manfred Kayser, Brian N. Popp, Pontus Skoglund

Correspondence

lfehrens@ucsc.edu

In Brief

Fehren-Schmitz et al. sequence low-coverage genomes of three pre-European-contact and two post-European-contact individuals from Rapa Nui (Easter Island, Chile). They find that the pre-European-contact Rapanui individuals fall well into the genetic variability of prehistoric and modern Polynesian populations and show no admixture with Native Americans.

Fehren-Schmitz et al., 2017, Current Biology *27*, 3209–3215
October 23, 2017
http://dx.doi.org/10.1016/j.cub.2017.09.029

CellPress

Current Biology
Report

Genetic Ancestry of Rapanui before and after European Contact

Lars Fehren-Schmitz,[1,2,8,9,*] Catrine L. Jarman,[3,8] Kelly M. Harkins,[1,8] Manfred Kayser,[4] Brian N. Popp,[5] and Pontus Skoglund[6,7,8]

[1]UCSC Paleogenomics Lab, Department of Anthropology, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA
[2]UCSC Genomics Institute, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA
[3]Department of Archaeology and Anthropology, University of Bristol, 43 Woodland Road, Bristol BS8 1UU, UK
[4]Department of Genetic Identification, Erasmus MC University Medical Center Rotterdam, Wytemaweg 80, 3015 CN Rotterdam, the Netherlands
[5]Department of Geology & Geophysics, University of Hawaii, 1680 East-West Road, Honolulu, HI 96822, USA
[6]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA
[7]The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK
[8]These authors contributed equally
[9]Lead Contact
*Correspondence: lfehrens@ucsc.edu
http://dx.doi.org/10.1016/j.cub.2017.09.029

SUMMARY

The origins and lifeways of the inhabitants of Rapa Nui (Easter Island), a remote island in the southeast Pacific Ocean, have been debated for generations. Archaeological evidence substantiates the widely accepted view that the island was first settled by people of Polynesian origin, as late as 1200 CE [1–4]. What remains controversial, however, is the nature of events in the island's population history prior to the first historic contact with Europeans in 1722 CE. Purported contact between Rapa Nui and South America is particularly contentious, and recent studies have reported genetic evidence for Native American admixture in present-day indigenous inhabitants of Rapa Nui [5–8]. Statistical modeling has suggested that this genetic contribution might have occurred prior to European contact [6]. Here we directly test the hypothesis that the Native American admixture of the current Rapa Nui population predates the arrival of Europeans with a paleogenomic analysis of five individual samples excavated from Ahu Nau Nau, Anakena, dating to pre- and post-European contact, respectively. Complete mitochondrial genomes and low-coverage autosomal genomes show that the analyzed individuals fall within the genetic diversity of present-day and ancient Polynesians, and we can reject the hypothesis that any of these individuals had substantial Native American ancestry. Our data thus suggest that the Native American ancestry in contemporary Easter Islanders was not present on the island prior to European contact and may thus be due to events in more recent history.

RESULTS

Ancient Samples and Sequencing

Rapa Nui (Easter Island) plays a central role in the debate about pre-European contact between Polynesians and Native Americans due to its comparatively close geographic proximity to South America [3, 6, 7, 9]. The feasibility of early sea voyages from South America to Polynesia [9] and from Polynesia to South America [10] has been firmly demonstrated. However, discovery of sweet potatoes—a South American crop—in Polynesia circa 1000 CE long remained the key evidence for this possible contact [11, 12]. Early genetic studies of contemporary [2, 13, 14] and ancient Polynesians [4, 15] did not detect Native American ancestry, but more recent genomic investigations have shown that contemporary indigenous Easter Islanders share human leukocyte antigen (HLA) complex alleles with Native South American populations [5, 8]. However, because of their functional nature, HLA markers are not suitable to examine population history. A recent study analyzing genome-wide single-nucleotide polymorphisms (SNPs) in contemporary Rapanui identified an average of 6% Native American ancestry [6]. Statistical modeling of the length of chromosomal segments exhibiting Native American ancestry as well as variance between individuals signals an admixture event most likely occurring between 1280 and 1425 CE [6]—hence prior to European contact in Polynesia. Although two previous mitochondrial DNA (mtDNA) studies of pre-colonial individuals from Rapa Nui have identified only Polynesian haplotypes [4, 16], a small proportion of Native American ancestry is not necessarily detectable with mtDNA analysis depending on the scale and timing of the admixture. Moreover, mtDNA analysis would generally not allow detection of male-mediated genetic admixture.

To explore the genomic history of Rapa Nui, and to explicitly test the hypothesis of pre-European gene flow from South America, we conducted an ancient DNA analysis based on genome-wide autosomal and mitochondrial DNA markers from ribs of seven candidate human individuals pre- and post-dating the historic first European contact with the Rapanui in 1722 CE. These

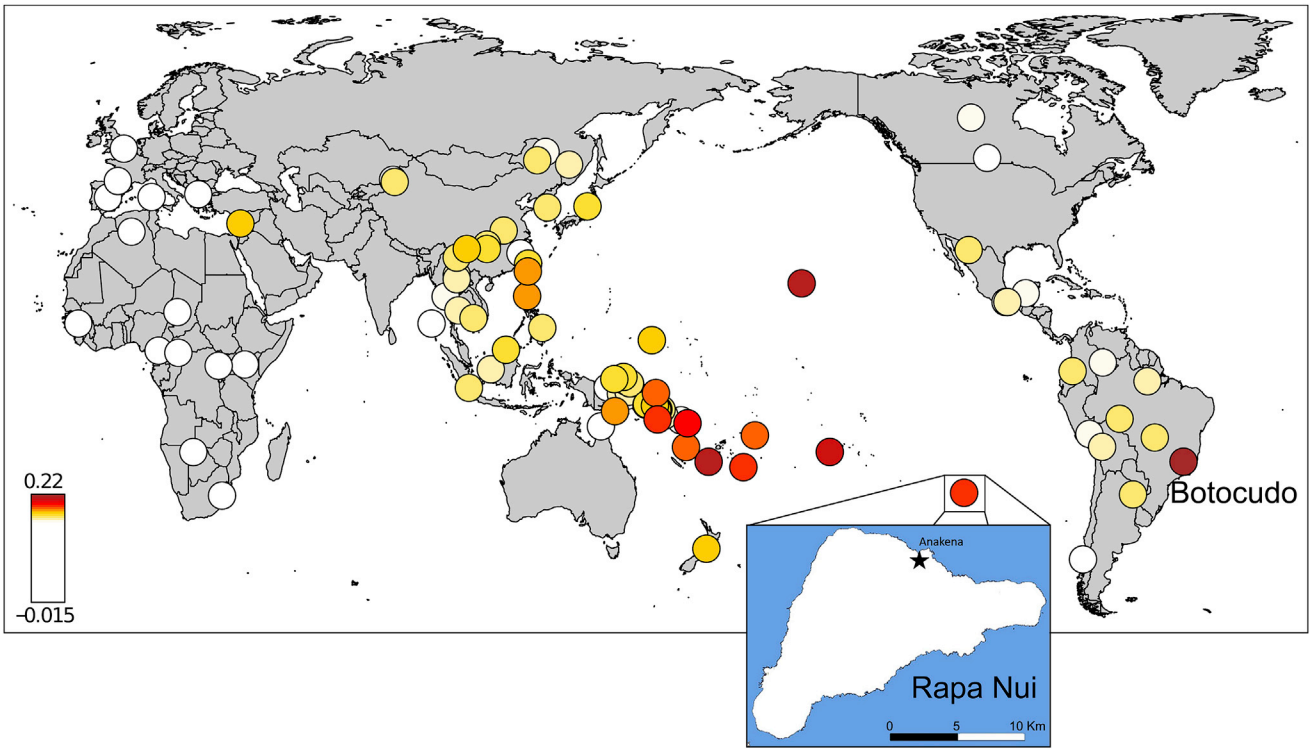

Figure 1. Map of Easter Island Showing the Archaeological Site from Which Samples Derive and Heatmap Illustrating the Amount of Shared Genetic Drift between the Ancient Rapanui Individuals and Global Populations

Global population data is from the Human Origins reference population set, including the Polynesian Botocudo and modern Rapanui individuals, as calculated with outgroup f_3 statistics.

samples originate from 1986–1988 archaeological investigations at Ahu Nau Nau, Anakena (Figure 1), including three radiocarbon-dated rib samples [17]. Taking seafood consumption into account to correct for marine reservoir effects, samples RN035 and RN036 yielded calibrated radiocarbon dates ranging from 1445 to 1624 CE (95.4% confidence; Table 1) [17]. The archaeological context places RN041 in the same age range, while the third dated sample, RN037, and associated individual RN039 date to a later, post-European period (1815–1945 CE, 95.4% confidence), most likely representing a more recent disturbance [17].

DNA was extracted from sternal rib fragments. Double-stranded partially uracil-DNA-glycosylase (UDG)-treated sequencing libraries [18] generated from the extracts yielded low amounts of endogenous DNA, ranging from 0.02% to 0.5%, for samples RN035, RN036, RN037, RN039, and RN041. To increase the endogenous DNA yield, we conducted targeted whole-genome enrichment (WGE) through hybridization capture (STAR Methods). The yield increased by more than 10× for some samples (1.5%–11% endogenous), but the molecular complexity of the libraries decreased significantly (Table S1). All libraries were sequenced to between 0.001- and 0.004-fold coverage using Illumina sequencing technology (Table 1), largely exhausting the preserved unique endogenous molecules in the libraries (Table S1). We were able to obtain complete mitochondrial genomes from five of seven individuals at 16× to 33× average coverage, as well as up to ~36,000 autosomal SNPs overlapping with population genomic reference datasets (Table 1). All samples had degradation patterns at terminal CpG dinucleotides consistent with ancient DNA ~100–600 years old from a tropical environment [19], with rates of C-to-T transitions at the terminus exceeding 11% for all individuals (Tables 1 and S3). The samples also exhibited low mitochondrial contamination estimates of 0.004%–0.78% (Table 1; STAR Methods), except for RN037 (~9%). Molecular sex assessment revealed that all tested samples are male, except for RN037, for which we could not obtain enough gonosomal sequence reads (Table 1).

Mitochondrial DNA

All five recovered mitochondrial genomes exhibit polymorphisms that characterize the so-called "Polynesian motif" (A14022G, T16217C, A16247G, and C16261T) found widespread across contemporary Polynesians [13, 14, 20]. More precisely, following the Phylotree Build 17 nomenclature [21], RN035 and RN039 can be classified as belonging to haplogroup B4a1a1, wheras RN036, RN037, and RN041 can be determined as B4a1a1m1, characterized by the A1692G polymorphism [22] (see Table S2). The latter haplogroup has been described for eastern Polynesia [13, 14].

Autosomal DNA

To investigate population affinities of the Rapanui samples with modern and ancient Oceanian and global populations, we first

Table 1. Radiocarbon Dates and DNA Sequencing Results for Five Ancient Rapanui Samples

Sample	Date[a]	Mapped Reads	Genome Coverage	Covered SNPs		mtHG	Sex	mtContamination (Average)	Damage[c]
				Human Origins	SGDP[b]				
RN035	1445–1620 CE	223,312	0.0041	2,402	35,629	B4a1a1	XY	0.57%	14%
RN036	1458–1624 CE	45,212	0.001	459	8,368	B4a1a1m1	XY	0.04%	16%
RN041	–	101,754	0.002	746	14,622	B4a1a1m1	XY	0.78%	19%
RN037	1815–1945 CE	24,162	0.0004	104	4,060	B4a1a1m1	?	9.46%	12%
RN039	–	36,518	0.001	369	6,658	B4a1a1	XY	0.35%	11%

The following abbreviations are used: SNP, single-nucleotide polymorphism; SGDP, Simons Genome Diversity Project; mtHG, mitochondrial haplogroup.

[a]The ^{14}C calibrations have been corrected for marine radiocarbon reservoir effect, considering a mixed marine/terrestrial diet consumed by the tested individuals. 95.4% confidence intervals are shown.

[b]Filtered reference set; transition variants excluded.

[c]For full damage estimates and standard error, see Table S3.

compared them to the full Human Origins (HO) population reference dataset [15, 21, 23], including the genomes of two Botocudo individuals of un-admixed Polynesian ancestry [24, 25]. We obtained a maximum overlap of ~2,400 SNPs between the five Rapanui samples and the ~600,000 SNPs of the HO reference dataset (Table 1). Principal-component analysis (PCA) shows that all five ancient Rapanui individuals fall within the genomic variation found in contemporary Polynesians and the two Botocudos (Figure S1). Outgroup f_3 statistics of the form f_3(Yoruba; RapaNui_preContact, *X*), which quantify the amount of shared genetic drift between the pre-European-contact Rapanui individuals (RN035, RN036, and RN041 together) and *X*—a modern or ancient population—since they diverged from an African outgroup, confirm the affinity of the ancient Rapanui individuals with Polynesian populations (Figure 1). The highest amount of shared drift with the pre-European-contact Rapanui is observed for the Botocudo individuals of Polynesian ancestry (f_3 = 0.22, SE = 0.0158, Z = 13.921; Data S1) and populations from Polynesia (Figure 1).

Aiming to increase statistical power, we merged our ancient Rapanui data with (1) the 300 high-coverage genomes of the Simons Genome Diversity Project (SGDP) [26], (2) the Affymetrix Human SNP Array 6.0 data of eight contemporary Rapanui individuals [6], and (3) the genomes of the Botocudo individuals [24]. We used the default filters for the SGDP dataset [26] including only dinucleotide transversion SNPs to avoid potential biases from post-mortem damage that affect transitions (STAR Methods). For 318,950 such sites, we had data for at least one ancient or contemporary Rapanui individual, with overlap for the ancient individuals ranging from ~35,000 SNPs (RN035) to ~4,000 SNPs (RN037), respectively (Table 1).

We performed model-based clustering analysis using the maximum-likelihood approach implemented in ADMIXTURE [27], assuming 2–10 ancestral populations (*K* = 2–10). The lowest cross-validation (*CV*) error was found for the model with 6 modeled clusters or ancestral populations (*K* = 6). This revealed ~80% Asian and ~20% Papuan ancestry for all three pre-European Rapanui individuals, similar to contemporary Polynesians and the Botocudo individuals [2, 15] (Figures 2 and S3). Despite the low SNP coverage, the root-mean-squared error (RMSE) per cluster and *k* value was low for the individuals with more than 10,000 overlapping SNPs (RN035, RN041; RMSE $\leq$ 0.05), consistent with other reports [27, 28], and RMSE for RN036 with ~8,300 SNPs remained in the acceptable range (Table1; Figure S2A).

One of the two post-European-contact samples (RN039) exhibited an ~20% European ancestry signal (SE = 0.0996), similar to European admixture proportions previously reported in contemporary Easter Islanders [6]. The contamination estimates for this sample were low, and damage patterns reflected expectations for ancient DNA (Table 1). Notably, the other post-European-contact sample (RN037) did not show any European genomic admixture.

Testing for Native American Admixture

We used the ancestry proportions inferred from our ADMIXTURE analyses, and the standard errors (SEs) obtained by running 1,000 bootstrap replicates, to calculate the 95% confidence interval (CI), and we used its upper limit to assess the maximum Native American ancestry that is consistent with the data. For all *K* above 4 (when the distinct Native American cluster appears) and below 8 (when the distinct Polynesian cluster appears), the point estimates of Native American ancestry ranged between 0.0 and 0.00001 for all five ancient Rapanui samples. The SEs ranged between 0.007 and 0.02, suggesting that the 95% CI for Native American ancestry is 0%–1% for the individual with the highest statistical power and 0%–3.3% for the individuals with the lowest power (Table 2). In contrast, for contemporary Rapanui individuals described elsewhere [6], our analysis revealed an average Native American ancestry component of ~5% (95% CI upper limit range: 5.6%–9.2% across individuals) for *K* = 6 (Table 2), replicating the ~6% reported by Moreno-Mayar et al. [6]. Thus, the maximum possible Native American ancestry consistent with the data from the ancient Rapanui individuals is significantly lower than the Native American ancestry observed in the contemporary Rapanui samples. Our data therefore allow rejection of the previously suggested hypothesis that Native American genetic ancestry was homogeneous in Rapa Nui prior to European contact. In addition, in an experiment randomly drawing European individuals from France, England, and Norway from the SGDP dataset, we obtained upper 95% confidence limits of Native American ancestry up to 4% (Table 2). This supports the notion that the upper 95% confidence limit of 3.3% Native American ancestry we estimated for the ancient

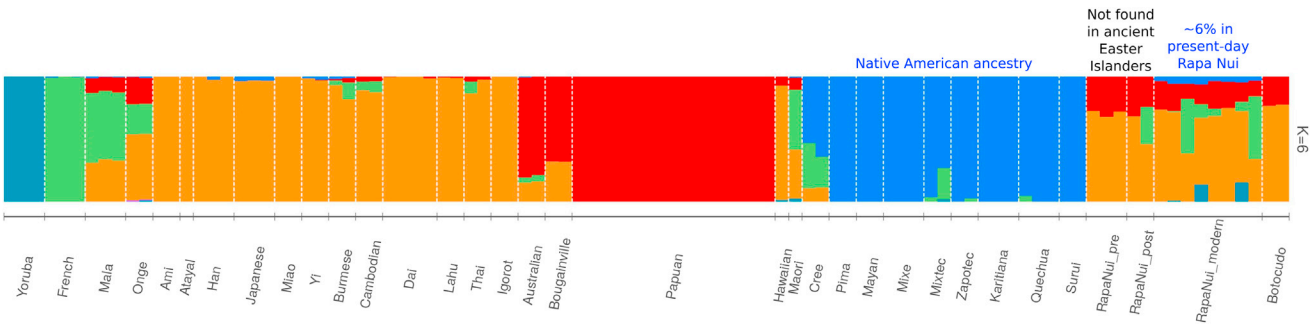

Figure 2. Ancestry Proportions of the Ancient Rapanui Samples When Compared to the Reference Population Dataset Assuming $K = 6$ Ancestral Components

Proportions are in comparison to the SGDP reference population dataset, Polynesian Botocudo, and modern Rapanui assuming $K = 6$ ancestral components, as inferred by ADMIXTURE analysis (see also Figure S3 for all K).

Rapanui samples reflects statistical uncertainty common in this type of analysis.

To further test whether the low number of SNPs available for the ancient Rapanui samples impacts the ability to detect Native American ancestry, we randomly downsampled the contemporary Rapa Nui data to SNP numbers ranging from 25,000 to 500 by randomly drawing one allele for heterozygous SNPs to mimic the data quality of our ancient DNA analysis. Rerunning ADMIXTURE for $K = 6$ for each downsampled dataset, we were able to detect the 6% Native American ancestry observed by Moreno-Mayar et al. [6] for these samples with as few as 500 SNPs (Figure S2B), with a 95% CI ranging from 3%–7% (25,000 SNPs; Table 2) to 0%–17% (500 SNPs). Although the ascertainment of the Affymetrix 6.0 SNPs is different from the SNPs obtained between the 300 SGDP individuals, this test suggests that the number of SNPs available in the ancient Rapanui samples would allow detection of Native American admixture if present.

Additionally, we used D statistics [29]—providing a formal test of a tree-like population history that can detect deviations caused by admixture—of the form D(Atayal, X; Botocudo, RapaNui_pre) and D(Han, X; Botocudo, RapaNui_pre), with X being one of 14 North to South American populations represented in the SGDP dataset (Data S1, Sheets 2–5). None of the comparisons produces a Z score > 2, indicating no evidence for admixture between Native American and the ancient Rapanui [30]. Using the probabilistic model of population splits and divergence implemented in TreeMix [31], we detected a signal of Native American gene flow into the contemporary Rapanui individuals, but not into either the ancient Rapanui or the Botocudo individuals when allowing three or more migration events (Figure S4).

DISCUSSION

Here we present, to our knowledge, the first genome-wide data from ancient Rapanui that allow direct testing of the timing of Native American admixture observed in Rapanui today. While some previous genetic studies of contemporary indigenous inhabitants of Rapa Nui concluded that Native American admixture occurred prior to European contact [5, 6, 8], none of our autosomal DNA (aDNA) analyses of pre- or post-European-contact Rapanui individuals detected Native American ancestry. Although the genomic data we obtained are limited due to unfavorable DNA preservation, the data are statistically well-powered to confirm that all ancient Rapanui samples analyzed here fall into the genomic variation found within ancient and modern Polynesians [2, 15]. Our direct aDNA evidence from three pre-European-contact and two post-European-contact Rapanui suggests that any European-mediated genetic contact between Native Americans and Easter Islanders had not spread to the entire population earlier than the 19th century. However, by use of contemporary Rapanui samples, Moreno-Mayar et al. [6] estimated the Native American admixture event at ~22 generations ago, which translates to 1310–1420 CE if assuming a generation time of 25–30 years [32]. Although it is theoretically possible that pre-European-contact admixture with Native Americans occurred only within some spatiopolitical groups on Rapa Nui [3], to which the five individuals investigated here did not belong, we do not regard this as plausible. Rapa Nui is a very small island, and Anakena is one of the few beaches that are suitable for landing boats [3], meaning that cross-island interactions would be likely at this site. Furthermore, given the wide time span represented by our samples as well as the intrusive burial of the post-European-contact individuals, it is unlikely that all of our individuals represent the same group.

In principle, Native American and European ancestry components found among modern indigenous Rapa Nui inhabitants [6, 33] could have been introduced through European and post-European-contact American involvement at any time after the initial European contact. Between 1722 and 1862, at least 53 documented ships from the Americas and Europe reached Rapa Nui and other Polynesian islands, along with numerous unrecorded whaling ships [1, 32–34]. These ships often carried crews of mixed Native American and European ancestry [1, 32, 33] who could have remained on Rapa Nui, as is documented for at least one case in 1863 [35]. When Peru was pressured to repatriate its few remaining enslaved Polynesians in 1884, the only 15 individuals who survived the voyage were returned to Rapa Nui, irrespective of their previous origin [35]. Following this repatriation, an epidemic is thought to have inflicted mass mortality on an already weakened population, leaving only ~100 survivors [33–35]. Many of these survivors left for Mangareva and Tahiti in the 1870s and 1880s before eventually returning. It is likely that these demographic disruptions

Table 2. Native American Ancestry Components as Determined by ADMIXTURE at *K* = 6 and Upper and Lower Limit 95% Confidence Intervals

	n	Average	Range (across Individuals)	95% CI Upper Limit (Average)	95% CI Upper Limit (Range across Individuals)	95% CI Lower Limit (Range across Individuals)
RapaNui_Modern	8	4.8%	3.6%–6.7%	7.1%	5.6%–9.2%	1.7%–4.3%
RapaNui_Modern (downsampled 25K SNPs)	8	4.0%	3.2%–5.6%	6.7%	5.8%–8.3%	1.3%–4.1%
RapaNui_Pre	3	0.0%	0.0%–0.0%	1.6%	0.4%–3.3%	0.0%–0.0%
RapaNui_Pre (only >10K SNPs)	2	0.0%	0.0%–0.0%	0.7%	0.4%–1%	0.0%–0.0%
RapaNui_Post	2	0.0%	0.0%–0.0%	2.3%	1.7%–2.9%	0.0%–0.0%
Botocudo	2	0.0%	0.0%–0.0%	0.0%	0%–0%	0.0%–0.0%
French	3	0.6%	0.3%–0.8%	1.4%	0.9%–2%	0.0%–0.7%
English	3	1.9%	1.3%–2.5%	3.5%	2.8%–4.2%	0.0%–0.8%
Norwegian	1	3.7%	3.7%	5%	5%	2.4%

See also Table S1.

introduced ancestry on the island that was not present in the pre-European-contact population. The change in the island's ethnic composition, with an influx of foreign Polynesians, was noted by William J. Thomson in 1886 [36]. This is supported by higher than expected diversity of Polynesian Y chromosome lineages in the Rapa Nui gene pool when compared to other Eastern Polynesian islands [33].

The arrival of already admixed individuals with partial Native American ancestry in the late 18th and 19th century could be consistent with the relatively old admixture data estimated by Moreno-Mayar et al. [6]. Additionally, the bottleneck experienced by the island's population could have resulted in a homogeneous proportion of 5%–6% Native American ancestry today. The post-European-contact ancient DNA sample RN039 dating to 1815–1945 CE, which did not exhibit any Native American ancestry but exhibited significant European ancestry, could support this postulated post-European-contact population structure. It is also possible that Native American ancestry was already widespread on Rapa Nui by 1815–1945 CE but that this is not seen in RN039 due to this individual being a recent migrant from another Polynesian island. The other unadmixed and putatively post-European-contact sample, RN037, is not directly dated. Further resolution of the timing of admixture between Native Americans and the ancestors of present-day Easter Islanders will require future studies using dense time series of ancient genomes analyzed from Rapa Nui and South America.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Ancient Rapanui samples
- METHOD DETAILS
  - DNA extraction
  - Library preparation and whole-genome enrichment
- QUANTIFICATION AND STATISTIC ANALYSES
  - Processing of the NGS sequencing data
  - DNA authenticity
  - Molecular sex determination
  - Mitochondrial DNA analyses
  - Population genetic analyses
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, three tables, and one dataset and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2017.09.029.

AUTHOR CONTRIBUTIONS

Conceptualization: L.F.-S., C.L.J., K.M.H., and P.S.; Resources: L.F.-S., C.L.J., M.K., BP, and P.S.; Investigation: L.F.-S., C.L.J., and K.M.H.; Formal Analyses: L.F.-S., P.S., and B.N.P.; Writing: L.F.-S., C.L.J., K.M.H., M.K., B.N.P., and P.S.

ACKNOWLEDGMENTS

We would like to thank David Reich for sharing the full HO and SGDP datasets and Victor Moreno-Mayar, Anna-Sapfo Malaspinas, and Eske Willerslev for sharing the data of modern Rapanui individuals. We also thank Reidar Solsvik at the Kon-Tiki Museum for providing access to the samples. This research was funded by a University of California Presidential Research Catalyst Award (CA-17-445724). P.S. was supported by the Swedish Research Council (VR2014-453). K.M.H. was supported by the National Science Foundation (NSF 1513501). This is SOEST contribution number 10229.

Received: July 7, 2017
Revised: August 21, 2017
Accepted: September 13, 2017
Published: October 12, 2017

REFERENCES

1. Hunt, T.L., and Lipo, C.P. (2006). Late colonization of Easter Island. Science *311*, 1603–1606.
2. Wollstein, A., Lao, O., Becker, C., Brauer, S., Trent, R.J., Nürnberg, P., Stoneking, M., and Kayser, M. (2010). Demographic history of Oceania inferred from genome-wide data. Curr. Biol. *20*, 1983–1992.

3. Gill, G.W., and Stefan, V.H. (2016). Rapanui origins, relationships, and warfare: A summary in theoretical context. In Skeletal Biology of the Ancient Rapanui (Easter Islanders) (Cambridge Studies in Biological and Evolutionary Anthropology), V.H. Stefan, and G.W. Gill, eds. (Cambridge: Cambridge University Press), pp. 286–302.
4. Hagelberg, E., Quevedo, S., Turbon, D., and Clegg, J.B. (1994). DNA from ancient Easter Islanders. Nature *369*, 25–26.
5. Lie, B.A., Dupuy, B.M., Spurkland, A., Fernández-Viña, M.A., Hagelberg, E., and Thorsby, E. (2007). Molecular genetic studies of natives on Easter Island: evidence of an early European and Amerindian contribution to the Polynesian gene pool. Tissue Antigens *69*, 10–18.
6. Moreno-Mayar, J.V., Rasmussen, S., Seguin-Orlando, A., Rasmussen, M., Liang, M., Flåm, S.T., Lie, B.A., Gilfillan, G.D., Nielsen, R., Thorsby, E., et al. (2014). Genome-wide ancestry patterns in Rapanui suggest pre-European admixture with Native Americans. Curr. Biol. *24*, 2518–2525.
7. Thorsby, E. (2012). The Polynesian gene pool: an early contribution by Amerindians to Easter Island. Philos. Trans. R. Soc. Lond. B Biol. Sci. *367*, 812–819.
8. Thorsby, E., Flåm, S.T., Woldseth, B., Dupuy, B.M., Sanchez-Mazas, A., and Fernandez-Vina, M.A. (2009). Further evidence of an Amerindian contribution to the Polynesian gene pool on Easter Island. Tissue Antigens *73*, 582–585.
9. Heyerdahl, T. (1950). Kon-Tiki: Across the Pacific by Raft (Chicago: Rand McNally).
10. Irwin, G. (2011). Sailing from Polynesia to the Americas. In Polynesians in America: Pre-Columbian Contacts with the New World, T.L. Jones, A.A. Storey, E.A. Matisoo-Smith, and J.M. Ramírez-Aliaga, eds. (Plymouth: Rowman Altamira), pp. 247–262.
11. Hather, J., and Kirch, P.V. (1991). Prehistoric sweet potato from Mangaia Island, Central Polynesia. Antiquity *65*, 887–893.
12. Roullier, C., Benoit, L., McKey, D.B., and Lebot, V. (2013). Historical collections reveal patterns of diffusion of sweet potato in Oceania obscured by modern plant movements and recombination. Proc. Natl. Acad. Sci. USA *110*, 2205–2210.
13. Kayser, M., Brauer, S., Cordaux, R., Casto, A., Lao, O., Zhivotovsky, L.A., Moyse-Faurie, C., Rutledge, R.B., Schiefenhoevel, W., Gil, D., et al. (2006). Melanesian and Asian origins of Polynesians: mtDNA and Y chromosome gradients across the Pacific. Mol. Biol. Evol. *23*, 2234–2244.
14. Duggan, A.T., Evans, B., Friedlaender, F.R., Friedlaender, J.S., Koki, G., Merriwether, D.A., Kayser, M., and Stoneking, M. (2014). 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.
15. Skoglund, P., Posth, C., Sirak, K., Spriggs, M., Valentin, F., Bedford, S., Clark, G.R., Reepmeyer, C., Petchey, F., Fernandes, D., et al. (2016). Genomic insights into the peopling of the Southwest Pacific. Nature *538*, 510–513.
16. Dudgeon, J.V., Commendador, A.S., and Tromp, M. (2016). Archaeogenetics and paleodemographic estimation of founding populations: Features of residential geography on Rapa Nui. In Skeletal Biology of the Ancient Rapanui (Easter Islanders) (Cambridge Studies in Biological and Evolutionary Anthropology), V.H. Stefan, and G.W. Gill, eds. (Cambridge: Cambridge University Press), pp. 202–221.
17. Jarman, C.L., Larsen, T., Hunt, T., Lipo, C., Solsvik, R., Wallsgrove, N., Ka'apu-Lyons, C., Close, H.G., and Popp, B.N. (2017). Diet of the prehistoric population of Rapa Nui (Easter Island, Chile) shows environmental adaptation and resilience. Am. J. Phys. Anthropol. Published online June 30, 2017. https://doi.org/10.1002/ajpa.23273.
18. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., and Reich, D. (2015). Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. Philos. Trans. R. Soc. Lond. B Biol. Sci. *370*, 20130624.
19. Briggs, A.W., Stenzel, U., Johnson, P.L.F., Green, R.E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., and Pääbo, S. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. Proc. Natl. Acad. Sci. USA *104*, 14616–14621.
20. Lum, J.K., Rickards, O., Ching, C., and Cann, R.L. (1994). Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. Hum. Biol. *66*, 567–590.
21. Lazaridis, I., Nadel, D., Rollefson, G., Merrett, D.C., Rohland, N., Mallick, S., Fernandes, D., Novak, M., Gamarra, B., Sirak, K., et al. (2016). Genomic insights into the origin of farming in the ancient Near East. Nature *536*, 419–424.
22. van Oven, M., and Kayser, M. (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum. Mutat. *30*, E386–E394.
23. Reich, D., Patterson, N., Campbell, D., Tandon, A., Mazieres, S., Ray, N., Parra, M.V., Rojas, W., Duque, C., Mesa, N., et al. (2012). Reconstructing Native American population history. Nature *488*, 370–374.
24. Malaspinas, A.-S., Lao, O., Schroeder, H., Rasmussen, M., Raghavan, M., Moltke, I., Campos, P.F., Sagredo, F.S., Rasmussen, S., Gonçalves, V.F., et al. (2014). Two ancient human genomes reveal Polynesian ancestry among the indigenous Botocudos of Brazil. Curr. Biol. *24*, R1035–R1037.
25. Gonçalves, V.F., Stenderup, J., Rodrigues-Carvalho, C., Silva, H.P., Gonçalves-Dornelas, H., Líryo, A., Kivisild, T., Malaspinas, A.-S., Campos, P.F., Rasmussen, M., et al. (2013). Identification of Polynesian mtDNA haplogroups in remains of Botocudo Amerindians from Brazil. Proc. Natl. Acad. Sci. USA *110*, 6465–6469.
26. Mallick, S., Li, H., Lipson, M., Mathieson, I., Gymrek, M., Racimo, F., Zhao, M., Chennagiri, N., Nordenfelt, S., Tandon, A., et al. (2016). The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. Nature *538*, 201–206.
27. Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*, 1655–1664.
28. Sikora, M., Carpenter, M.L., Moreno-Estrada, A., Henn, B.M., Underhill, P.A., Sánchez-Quinto, F., Zara, I., Pitzalis, M., Sidore, C., Busonero, F., et al. (2014). Population genomic analysis of ancient and modern genomes yields new insights into the genetic ancestry of the Tyrolean Iceman and the genetic structure of Europe. PLoS Genet. *10*, e1004353.
29. Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.Y., et al. (2010). A draft sequence of the Neandertal genome. Science *328*, 710–722.
30. Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics *192*, 1065–1093.
31. Pickrell, J.K., and Pritchard, J.K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet. *8*, e1002967.
32. Dolin, E.J. (2007). Leviathan: The History of Whaling in America (New York: W.W. Norton & Company).
33. Hurles, M.E., Maund, E., Nicholson, J., Bosch, E., Renfrew, C., Sykes, B.C., and Jobling, M.A. (2003). Native American Y chromosomes in Polynesia: the genetic impact of the Polynesian slave trade. Am. J. Hum. Genet. *72*, 1282–1287.
34. Peiser, B. (2005). From genocide to ecocide: the rape of Rapa Nui. Energy Environ. *16*, 513–539.
35. Maude, H.E. (1982). Slavers in Paradise: The Peruvian Slave Trade in Polynesia, 1862-1864 (Stanford: Stanford University Press).
36. Thomson, W.J. (1891). Te Pito Te Henua, or Easter Island. In Report of the National Museum, 1888–1889 (Washington, DC: Smithsonian Institution), pp. 447–552.
37. Allentoft, M.E., Sikora, M., Sjögren, K.-G., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P.B., Schroeder, H., Ahlström, T., Vinner, L., et al. (2015). Population genomics of Bronze Age Eurasia. Nature *522*, 167–172.
38. Schmieder, R., and Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. Bioinformatics *27*, 863–864.

39. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., and Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics *29*, 1682–1684.
40. Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., and Howell, N. (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat. Genet. *23*, 147.
41. Gilbert, M.T., Willerslev, E., Hansen, A.J., Barnes, I., Rudbeck, L., Lynnerup, N., and Cooper, A. (2003). Distribution patterns of postmortem damage in human mitochondrial DNA. Am. J. Hum. Genet. *72*, 32–47.
42. van Oven, M. (2015). PhyloTree Build 17: Growing the human mitochondrial DNA tree. Forensic Sci. Int. Genet. Supplemental Ser. *5*, e392–e394.
43. Weissensteiner, H., Pacher, D., Kloss-Brandstätter, A., Forer, L., Specht, G., Bandelt, H.-J., Kronenberg, F., Salas, A., and Schönherr, S. (2016). HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. Nucleic Acids Res. *44* (W1), W58–W63.
44. Commendador, A.S., Dudgeon, J.V., Fuller, B.T., and Finney, B.P. (2013). Radiocarbon dating human skeletal material on Rapa Nui: evaluating the effect of uncertainty in marine-derived carbon. Radiocarbon *56*, 277–294.
45. Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., and Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. USA *110*, 15758–15763.
46. Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. *2010*, t5448.
47. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.
48. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.
49. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.
50. Skoglund, P., Northoff, B.H., Shunkov, M.V., Derevianko, A.P., Pääbo, S., Krause, J., and Jakobsson, M. (2014). Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. Proc. Natl. Acad. Sci. USA *111*, 2229–2234.
51. Renaud, G., Slon, V., Duggan, A.T., and Kelso, J. (2015). Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. Genome Biol. *16*, 224.
52. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics *15*, 356.
53. Skoglund, P., Storå, J., Götherström, A., and Jakobsson, M. (2013). Accurate sex identification of ancient human remains using DNA shotgun sequencing. J. Archaeol. Sci. *40*, 4477–4482.
54. Llamas, B., Fehren-Schmitz, L., Valverde, G., Soubrier, J., Mallick, S., Rohland, N., Nordenfelt, S., Valdiosera, C., Richards, S.M., Rohrlach, A., et al. (2016). Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. Sci. Adv. *2*, e1501385.
55. Patterson, N., Price, A.L., and Reich, D. (2006). Population structure and eigenanalysis. PLoS Genet. *2*, e190.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
RN035	Ahu Nau Nau, Anakena, Rapa Nui, Kon-Tiki Museum (Oslo), and Museo Antropologico P. Sebastian Englert (Rapa Nui)	RN035
RN036	Ahu Nau Nau, Anakena, Rapa Nui, Kon-Tiki Museum (Oslo), and Museo Antropologico P. Sebastian Englert (Rapa Nui)	RN036
RN037	Ahu Nau Nau, Anakena, Rapa Nui, Kon-Tiki Museum (Oslo), and Museo Antropologico P. Sebastian Englert (Rapa Nui)	RN037
RN039	Ahu Nau Nau, Anakena, Rapa Nui, Kon-Tiki Museum (Oslo), and Museo Antropologico P. Sebastian Englert (Rapa Nui)	RN039
RN041	Ahu Nau Nau, Anakena, Rapa Nui, Kon-Tiki Museum (Oslo), and Museo Antropologico P. Sebastian Englert (Rapa Nui)	RN041
Chemicals, Peptides, and Recombinant Proteins		
Sodium acetate	Sigma Aldrich	Cat#127-09-3
Silicon dioxide 500 G	SLS	Cat#7631-86-9
Isopropanol	Sigma-Aldrich	Cat#67-63-0
Ultra Pure 0.5 M EDTA (pH 8.0)	Life Technologies	Cat#15575020
Tween-20	Sigma-Aldrich	Cat#9005-64-5
Critical Commercial Assays		
High Sensitivity D1000 ScreenTape	Agilent	Cat#5067-5584
Deposited Data		
RN035	This paper	ENA: PRJEB22217 ERS1876352
RN036	This paper	ENA: PRJEB22217 ERS1876353
RN041	This paper	ENA: PRJEB22217 ERS1876354
RN037	This paper	ENA: PRJEB22217 ERS1876355
RN039	This paper	ENA: PRJEB22217 ERS1876356
Oligonucleotides		
P5-adaptor: CTTTCCCTACACGACGCTCTTCCGA TCTxxxxxxx	[19]	custom generated by Integrated DNA Technologies
P7-adaptor: GTGACTGGAGTTCAGACGTGTGCTCT TCCGATCTxxxxxxx	[19]	custom generated by Integrated DNA Technologies
P5/P7comp: xxxxxxxAGATCG	[19]	custom generated by Integrated DNA Technologies
IS4: 5′-AATGATACGGCGACCACCGAGATCTACACT CTTTCCCTACACGACGCTCTT-3′	[37]	custom generated by Integrated DNA Tech-nologies
P7 indexing: 5′-CAAGCAGAAGACGGCATACGAGAT xxxxxxxxGTGACTGGAGTTCAGACGTGT-3′	[37]	custom generated by Integrated DNA Tech-nologies

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Ilumina Pipeline v1.4	Illumina	https://support.illumina.com/downloads.html
SeqPrep	John St. John	https://github.com/jstjohn/SeqPrep
Burrows-Wheeler Aligner (BWA) 0.7.5a-r405	[38]	http://bio-bwa.sourceforge.net/
Picards-tools-1.98	Broad Institute	https://sourceforge.net/projects/picard/files/picard-tools/1.98/
sequenceTools	Stephan Schiffels	https://github.com/stschiff/sequenceTools
PMDtools	[39]	https://code.google.com/archive/p/pmdtools/
SAMtools-0.1.19	[40]	https://sourceforge.net/projects/samtools/files/samtools/0.1.19/
mapDamage	[41]	https://ginolhac.github.io/mapDamage/
Haplogrep	[42]	http://haplogrep.uibk.ac.at/
EIGENSOFT 5.0.1	[43]	https://github.com/DReichLab/EIG
TreeMix	[32]	https://bitbucket.org/nygcresearch/treemix/wiki/Home
ADMIXTOOLS	[30]	https://github.com/DReichLab/AdmixTools
Other		
Proteinase K	Promega	Cat#MC5005
MinElute PCR Purification Kit	QIAGEN	Cat#28004
Zymo-spin V column extension reservoir	Zymo Research	Cat#C1016-25
T4 polymerase	Fermentas/ThermoFisher	Cat#EP0062
T4 Polynucleotide kinase	Fermentas/ThermoFisher	Cat#EK0031
Buffer Tango	Fermentas/ThermoFisher	Cat#BY5
ATP	ThermoFisher	Cat#R0441
PEG-4000	Sigma Aldrich	Cat#1546569
T4-ligase	Fermentas/ThermoFisher	Cat#EL0011
Bst-polymerase, large fragment (supplied with 10X ThermoPol reaction buffer)	New England BioLabs	Cat#M0275 S
Pfu Turbo Cx Hotstart DNA Polymerase	Agilent	Cat#600412
USER enzyme	New England BioLabs	Cat#M5505L
UGI (Uracil Glycosylase Inhibitor)	New England BioLabs	Cat#M0281L

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for reagents may be directed to and will be fulfilled by the Lead Contact, Lars Fehren-Schmitz (lfehrens@ucsc.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Ancient Rapanui samples

We obtained bone samples from seven individuals buried at Ahu Nau Nau originating from the 1986-88 archaeological investigations at Anakena (Figure 1), carried out as a collaboration between the Kon-Tiki Museum (Oslo) and the Museo Antropologico P. Sebastian Englert (Rapa Nui). All human bone was sampled as rib fragments from adult individuals, with no further information about age or sex available. These securely provenanced samples were collected by C.L.J. as part of a stable isotope study (University of Bristol and University of Hawai'i) [17]. Small pieces of sternal rib that remained from these investigations were transferred to the University of California Paleogenomics Lab, Santa Cruz (UC_PL) to avoid unnecessary sampling of further material and to maximize the information available from each sample.

Permission for destructive analysis of the human remains was granted by the Kon-Tiki Museum, consistent with original permissions to export and analyze excavated material in collaboration with the Museo Antropologico P. Sebastian Englert on Rapa Nui and granted by Consejo de Monumentos Nacionales, Chile. Further ethical approval for the study was granted by the University of Bristol Committee for Research Ethics.

Three samples of human bone from the current study (RN035, RN036, and RN037) were submitted for ^{14}C AMS dating at the Center for Applied Isotope Studies at the University of Georgia as described in Jarman et al. [17].

The three dated bone samples, RN035, RN036, and RN037, yielded uncalibrated dates of 490 ± 25BP, 440 ± 25BP, and 110 ± 25BP respectively. As described in Jarman et al. [17] the radiocarbon dates were calibrated after correcting for marine reservoir effects on the basis of estimated fractions of seafood in each individual's diet as determined from nitrogen and carbon isotope analyses of collagen and individual amino acids isolated from the collagen [17]. A local reservoir correction (ΔD) value of −83 ± 34 was used, following Commendador et al. [44] and the values available on the CHRONO database (http://calib.qub.ac.uk/marine). Results in Table 1 show the mixed marine/terrestrial calibrations. Samples RN035 and RN036 yielded calibrated radiocarbon dates consistent with charcoal ^{14}C dates for Ahu Nau Nau (Anakena). Contextual information places RN041 into the same age range, while the third dated sample RN037 and associated individual RN039, date to a considerably later period (1815-1945 CE, 95.4% confidence), most likely representing a more recent disturbance [17].

METHOD DETAILS

DNA extraction

All pre-amplification DNA procedures were carried out in dedicated clean lab ancient DNA facilities at the University of California, Santa Cruz Paleogenomics Lab (UCSC-PL, USA). All laboratory tools used to process the samples were either sterile and disposable, or decontaminated with full strength bleach (6%) and exposed to UV light for 1h before use. To decontaminate the samples the small pieces of rib were emerged in full strength bleach for two minutes and then rinsed with ddH20 and 70% Ethanol. Each side of the sample was then exposed to UV light for 10 min, and left to dry. Subsequently the samples were pulverized using a ball mill (Retsch 400, Germany).

The samples were extracted following the protocol described in [45], which optimizes the recovery of short DNA fragments from small quantities of bone powder. Because the samples derive from museum collections, we extended the protocol by including an initial buffer replacement as suggested in [37] as an additional prevention from potential external contaminations. From each sample, 50 mg of bone powder were incubated with 1 mL of an extraction buffer consisting of 0.45 M EDTA (pH 8.0) and 0.25 mg/ml of proteinase K for 30 minutes at 37°C. The samples were then centrifuged for three minutes and all supernatant discarded. After adding another 1 mL of extraction buffer, we continued lysis for 18h at 37°C under continuous rotation. After 18h, 1 mL of lysate was added to 13 mL of binding buffer (5M Guanidine hydrochloride (MW 95.53), 40% Isopropanol, 0.05% Tween-20, 9 mM Sodium Acetate) and transferred into an apparatus consisting of an extension reservoir (Zymo Research) fitted to a MinElute silica spin column (QIAGEN). The binding apparatus was placed into a 50 mL falcon tube and centrifuged at 1000-1500 rpm. As the sample-containing binding buffer passes through the MinElute column during centrifugation, its silica-based membrane will retain the DNA molecules. This filter is washed twice by adding 700 μL of PE buffer (QIAGEN). Finally, the filters are dried by centrifugation and the DNA molecules are eluted in a total of 50 μL of TET buffer (1 M Tris-HCL, pH 8.0, 0.5 M EDTA, 10% Tween-20). The elution buffer was heated to 56°C for the elution process. The extraction experiments included two negative controls.

Library preparation and whole-genome enrichment

We prepared sequencing libraries for all extracts, including the extraction blanks, employing a protocol that partially treats the DNA with UDG to drive down the rate of ancient DNA errors while preserving a damage signal at the terminal bases of the molecules, as described by [18], thereby allowing us to identify characteristic traits of ancient DNA for authentication purposes.

For each sample we used 20 μL of DNA extract in 60 μL blunt-end reactions with an initial USER enzyme treatment (NEB). The first part of the reaction (the partial UDG treatment, 52.2 μL total) consisted of 6 μL 10 × buffer Tango, 0.24 μL 25 mM dNTP mix, 0.6 μL 100 mM ATP (all reagents from Thermo Scientific Fermentas Molecular Biology Solutions), and 3.6 μL USER enzyme mix (1 U μl−1, NEB). During the 30 min incubation period at 37°C, most deaminated cytosines were excised by UDG and abasic sites were cut by Endo VIII. Next, 3.6 μL of UGI (2 U μl, UDG inhibitor, NEB) was added and incubated for 30 min at 37°C, after which 3 μL T4 polynucleotide kinase and 1.2 μL T4 DNA polymerase were added, and the final 60 μL was incubated for 15 min at 25°C, followed by 5 min at 12°C. Purification of the blunt-end reactions followed on MinElute columns (QIAGEN), with a final elution in 16 μL of EB buffer.

Subsequently, 1 μL of each of a barcoded (7bp unique for each sample), partially double-stranded P5 and P7 Illumina adapters (10 μM) were added to the blunted DNA and mixed before adding the ligation mix to bring adapters and DNA into close proximity (final concentration 0.25 μM for each barcoded adaptor). The concentrations in the 40 μL final ligation reactions were as follows: 1 × T4 DNA ligase buffer, 5% PEG-4000, 5 U T4 DNA ligase (all reagents from Thermo Scientific Fermentas Molecular Biology Solutions, Waltham, MA). Incubation followed for 30 min at room temperature. MinElute purification was performed by adding 200 μL buffer PB to each finished ligation reaction, washing with PE buffer twice, and eluting in 20 μL 10 mM Tris–HCl.

The fill-in reaction was performed in a final volume of 40 μL with 1 × ThermoPol buffer (New England Biolabs (NEB), Ipswich, MA), 250 μM each dNTP (Thermo Scientific Fermentas Molecular Biology Solutions) and 16 U Bst polymerase, large fragment (NEB) and incubated at 37°C for 20 min followed by a heat-inactivation at 80°C for 20 min.The entire 40 μL heat-inactivated fill-in reaction was

used in the PCR that finished the library preparation. A total of 200 μL PCR mix (divided into four reactions) per sample was prepared with the following final concentration: 1 × Pfu Turbo Cx reaction buffer, 20U Pfu Turbo Cx Hotstart DNA Polymerase (both Agilent Technologies, Santa Clara, CA), 200 μM each dNTP (Thermo Scientific Fermentas Molecular Biology Solutions) and 400 nM of each of the two primers (PreHyb-F, PreHyb-R, see STARS table) that do not extend the adaptor sites but keep them truncated (short). After an initial denaturation and activation of the polymerase at 95°C for 2 min, 30 cycles at 95°C for 30 s, 55°C for 30 s and 72°C for 1 min, we performed a final extension at 72°C for 5 min. Following PCR, the product was cleaned up with the MinElute PCR purification kit and eluted in 50 μL 1 × TTE buffer. The amplified libraries were quantified on a TapeStation 2200 using a High Sensitivity D1000 tape (Agilent).

To finalize the libraries with truncated adapters we performed an indexing PCR using 10 μL of library as template for each reaction and the primers IS4 (5′AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTT-3′) and P7 indexing (5′-CAAGCAGAAGACGGCATACGAGATxxxxxxxGTGACTGGAGTTCAGACGTGT-3′) where the x represent an unique 7 bp barcode described in [46]. Amplifications were performed with 1 μL Herculase II Fusion DNA Polymerase (Agilent Technologies) in 50 μL reactions consisting of 1 × Herculase II reaction buffer, 400 nM each primer and 250 μM each dNTP using 1 μL of a 20-fold dilution of the 'short' library for 20 cycles (95°C for 2 min, 20 cycles at 95°C for 30 s, 58°C for 30 s, 72°C for 30 s, final extension 72°C for 10 min). MinElute purification was performed by adding 250 μL PB, followed by two PE-washes, and DNA was eluted in 15 μL 1 × TE (with 0.05% Tween-20). Subsequently, libraries were quantified on the TapeStation again. The preservation and endogenous DNA content of the libraries was then tested by sequencing on an Illumina MiSeq platform at UC_PL (UCSC Paleogenomics Labs), using 2x75 bp paired end sequencing.

Because we observed overall low endogenous content of DNA in our samples, we performed whole genome in-solution enrichment for samples RN035, 036, 037, 039, and 041 using the MYbait Human Whole Genome Capture Kit from MYcroarray (Ann Arbor, MI) using the pre-indexed, barcoded libraries. The libraries were captured following the manufacturer's instructions (http://www.mycroarray.com/pdf/MYbaits-manual-v3.pdf). The captured libraries were amplified for 20 cycles with IS4 and indexed P7 primers as described above. Subsequently, libraries were purified with AMPure XP beads and quantified by 2200 TapeStation (Agilent Technologies).

Both the shotgun and enriched libraries were then sequenced on two lanes of a Illumina HiSeq 4000 sequencer at the UC San Diego sequencing facility. With two lanes of paired end (2x75 bp) data, we exhausted the unique molecules preserved in the libraries. Sequencing statistics can be found in Tables 1 and S1.

QUANTIFICATION AND STATISTIC ANALYSES

Processing of the NGS sequencing data

The Illumina software Casava 1.8.2 processes the MiSeq/HiSeq sequencing images (.bcl files) by performing image analysis, base calling, .bcl conversion and demultiplexing –whereby the samples are identified by the unique index sequences (default = 0 mismatches) that were ligated to the template during library preparation. Using an in-house script, forward (R1) and reverse (R2) reads were further sorted by identifying the correct internal barcodes and discarding those without, allowing up to one mismatch. The software SeqPrep2 (https://github.com/jstjohn/SeqPrep) trimmed the adapters and merged the R1and R2 reads. Default parameters were used for minimum base pair overlap to merge the forward and reverse reads (-o 15) with a quality score cutoff (-q 13) for mismatches to be counted in an overlap. The minimum length allowed for a merged read was set to 30 bp. Reads containing low-complexity regions, defined as stretches of nucleotides with little information content, were filtered by the program PrinSeq [38], using the DUST algorithm (t −12). The low-complexity-filtered, merged reads that passed all quality filters were mapped to the human reference genome using the software Burrows-Wheeler Aligner (BWA) version 0.7.5a-r405 [47], with default parameters and seed option disabled (-l 1024). We used hg19 (GRCh37 build) as reference genome, excluding the mitochondrial contigs. Mitochodrial reads were mapped to the revised Cambridge Reference Sequence (rCRS, NC_012920; [40]) employing the same BWA parameters.

Clonal sequences were removed using the rmDup function implemented in SAMtools [48] and indels were realigned using the tools RealignerTargetCreator and IndelRealigner from GATK-3.0-0 [49]. The resulting bam files were filtered for a minimum mapping quality of 30 using SAMtools −0.1.19. Finally, the tool GenomeAnalysisTK.jar from GATK-3.0-0 was used to calculate the depth of coverage of the bam files. All sequencing statistics are summarized in Table S3.

DNA authenticity

We included negative controls during the wet lab stages and sequenced them together with the samples in the MiSeq screening run. A total of four blanks – two extraction blanks and two library blanks – were sequenced. Furthermore, for each sample we analyzed patterns of molecular damage and presence of contaminants as explained below.

Ancient DNA molecules exhibit specific characteristics like short read lengths and specific nucleotide misincorporations due to post-mortem damage that can be used in part to verify the authenticity of ancient DNA molecules [19, 39, 41]. Because our libraries were partially UDG treated, most of damage accumulating at the ends of the molecules, except at the terminal bases, was repaired. However, terminal CpG dinucleotides are unaffected by the UDG treatment when methylated [18, 50]. We estimated patterns of DNA damage using PMDtools [50] and observed that damage for all samples exceeded 11% as to be expected for ancient DNA samples [18]. The distribution of read lengths, as calculated by MapDamage2 [39], show that the shotgun libraries range between 39-49 bp, while the fragments sequenced from the WGS captured libraries are slightly longer ranging from 61-69 bp (see Table S3).

We estimated mitochondrial contamination rates employing the modules contDeam and mtCont implemented in the software tool SCHMUZTI using the recommended parameters [51]. Samples RN035, 36, 39, and 41 exhibit low average contamination rates between 0.04%–0.8%. Only RN037 exhibits an elevated contamination rate of average 9.46% We have to caution that the partial UDG treatment of the sequencing libraries might lead to overestimation of contamination rates due to the reduced number of damaged sites [51]. The obtained number of gonosomal reads was not sufficient to perform X chromosome based contamination assessment as implemented in ANGSD for any sample [52].

Molecular sex determination

Sex for all samples was determined by evaluating the ratio (Ry) of reads aligning to the Y chromosome (nY) compared to the total number of reads aligning to the sex chromosomes (nX + nY), i.e., Ry = (nY/nY+nX), as described in [53]. All samples except RN037 could be assigned as male with confidence higher than 95%. For RN037 it was not possible to assign sex.

Mitochondrial DNA analyses

To identify the mitochondrial haplotypes of the individuals we manually analyzed each variant as described in [54] rather than relying on completely automated procedures. All mitochondrial reads mapped to the rCRS using BWA were visualized in Geneious v7.1.3 (Biomatters; available from https://www.geneious.com/) for each sample. Initially, SNPs were called in Geneious for all polymorphisms with minimum coverage 2 and a minimum variant frequency 0.7. The assembly and the resulting list of SNPs were verified manually and compared to SNPs reported at phylotree.org (mtDNA tree Build 17 [18 Feb 2016]) [42]. Following recommendations in van Oven and Kayser 2009 [22] we excluded common indels and mutation hotspots at nucleotide positions 309.1C(C), 315.1C, AC indels at 515–522, 16182C, 16183C, 16193.1C(C), and C16519T.

We embedded the consensus mitochondrial genomes in the existing mitochondrial tree (mtDNA tree Build 17 [18 Feb 2016]) using the online tool HaploGrep2 [43]. All ancient sequences represent unique linages falling into the Polynesian haplogroups B4a1a1 and B4a1a1m1. The sequence polymorphisms for each sample are listed in Table S2. Consensus sequence are available at the National Center for Biotechnology Information (accession numbers MF681702 to MF681706).

Population genetic analyses

The samples were first compared to the full Affymetrix Human Origins (HO) public dataset as reported in [15, 21] containing data from 2846 present-day people genotyped at 621,799 SNPs. The reference data from [15] contains 778 individuals from 83 East Asian and Oceanian populations and is the most comprehensive reference set available for the Oceanian samples studied here and includes the ancient early Oceanian Lapita samples. We also realigned BAM files from the two ''Botocudo'' samples of Polynesian ancestry [24], as described above, and added these samples to our dataset. Genotypes for the ancient samples, which overlapped with this dataset, were called using SAMtools mpileup with the –B flag to disables base alignment quality recalibration [48] and pileupCaller provided by Stephan Schiffels (https://github.com/stschiff/sequenceTools). Triallelic SNPs were discarded and bases were required to have quality $\geq$ 30. Alleles were not called within the first and last 2 bp of reads. For positions with more than one base call, one allele was randomly chosen with a probability equal to the frequency of the base at that position employing the RandomCalling function implemented in pileupCaller.

Principal components were defined using the East Asian – Oceanian subset published in [15] using EIGENSOFT 5.0.1 smartpca [55]. Results of the PCA analyses are visualized in Figure S1.

We computed outgroup f_3 statistics for the whole HO population set using the qp3Pop program in the ADMIXTOOLS package [30] in order to evaluate the amount of shared drift between our ancient samples with other global populations since their divergence from an African (Yoruba) outgroup. Outgroup f_3 statistics for the Rapa Nui genomes are reported in Figure 1 mapping a subset of the tested HO populations (see also Data S1, Sheet 1).

Because of the relatively low SNP overlap of our Rapa Nui genomes with the HO reference set we also compared them to the 300 high coverage genomes from 142 populations published by the Simons Genome Diversity Project (SGDP) [26]. Again we added the two 'Botocudo' individuals, the ancient Rapanui, and also the modern the Affymetrix Human SNP Array 6.0 data of the eight modern Rapanui individuals published by Moreno-Mayar et al. [6], this time sampling a single sequence read at each position, requiring base quality of 30, mapping quality of 30, and a minimum fragment length of 35. We used the default filter set for the SGDP dataset following Mallick et al. including only dinucleotide transversion SNPs to avoid potential biases from post-mortem damage affecting transition SNPs in the ancient Rapanui. There were 318,950 such transversion SNPs where we had data for at least one ancient or modern Rapanui individual, with overlap ranging from ~35,000 SNPs (RN035) to ~4,000 SNPs (RN037) for each of the historic individual samples (Table 1).

We performed model-based clustering analysis for the population set using the maximum-likelihood approach implemented in ADMIXTURE [27] assuming 2-10 ancestral populations (K = 2-10). ADMIXTURE was run with the cross validation (– cv) flag, the random seed flag (-s time), and 2000 block bootstrap replicates (-B 1000) considering K = 2 to K = 10, with 25 replicates for each value of k. We observe the lowest cross-validation (CV) error for the model assuming 6 ancestral populations (K = 6) (Figures 2 and S2A). We further assessed the potential clustering bias due to the low SNP coverage by calculating the root mean squared error (RMSE) per cluster and k value finding that accuracy remains sufficiently high for the individuals with more than 10,000 overlapping SNPs (RMSE < = 0.05) as reported by others [27, 28] (Table 1; Figure S1). We used the ancestry proportions inferred from our ADMIXTURE analyses and the standard error obtained by running 1000 block bootstrap replicates (the default behavior of

ADMIXTURE is to perform the block bootstrap over approximately 10 cM blocks [27]) to calculate the 95% confidence interval, as 2 times the standard error, as an estimate of the maximum undetected Native American ancestry.

To further test if our observed lack of Native American ancestry in the historic samples is affected by the low-coverage we randomly downsampled the modern Rapa Nui SNP data in several increments ranging from to 25,000-500 SNPs to mimic the low number of reference dataset sites overlapped by each ancient DNA sample, and the per-site single coverage (randomly select one allele per SNP). We then repeated the model-based clustering analysis for $K = 6$ using ADMIXTURE for each down sampled dataset (stepwise from 25K to 500 SNPs) and obtained standard errors to calculate the 95% confidence interval using the same parameters described above.

Using the SGDP reference dataset we calculated *D* statistics [29] of the form *D*(Yoruba, X; Ami, RapaNui_pre), *D*(Yoruba, X; Han, RapaNui_pre), as well as *D*(Atayal, X; Botocudo, RapaNui_pre), and *D*(Han, X; Botocudo, RapaNui_pre) using the qpDstat function of ADMIXTOOLS [30], X being one of 14 North to South American populations represented in the SGDP set. Results can be found in Data S1, sheets 2-5.

We applied TreeMix [31] to a reduced set of the SGDP reference data including the 'Botocudo' and modern Rapanui individuals to infer maximum likelihood trees and admixture graphs. We root the graphs with Yoruba, disabled sample size correction (-noss), performed a round of global realignments of the graph (-global), and computed standard error of migration weights (-se). We considered up to five migration edges and kept for each edge the graph with the highest log-likelihood among ten replicate runs (Figure S4).

DATA AND SOFTWARE AVAILABILITY

BAM files for the ancient Rapanui individuals have been deposited in European Nucleotide Archive under the study accession number ENA: PRJEB22217. Mitochondrial consensus sequences are available at the National Center for Biotechnology Information under the accession numbers NCBI: MF681702, MF681703, MF681704, MF681705, MF681706.