Ancient nuclear genomes enable repatriation of Indigenous human remains

Wright, Joanne L.; Wasef, Sally; Heupink, Tim H.; Westaway, Michael C.; Rasmussen, Simon; Pardoe, Colin; Fourmile, Gudju Gudju; Young, Michael; Johnson, Trish; Slade, Joan; Kennedy, Roy; Winch, Patsy; Pappin, Mary; Wales, Tapij; Bates, William “badger”; Hamilton, Sharnie; Whyman, Neville; Van Holst Pellekaan, Sheila; Mcallister, Peter J.; Taçon, Paul S.c.; Curnoe, Darren; Li, Ruiqiang; Millar, Craig; Subramanian, Sankar; Willerslev, Eske; Malaspinas, Anna-sapfo; Sikora, Martin; Lambert, David M.

Published in:
Science Advances

DOI:
10.1126/sciadv.aau5064

Publication date:
2018

Document license:
CC BY

Citation for published version (APA):
Wright, J. L., Wasef, S., Heupink, T. H., Westaway, M. C., Rasmussen, S., Pardoe, C., ... Lambert, D. M. (2018). Ancient nuclear genomes enable repatriation of Indigenous human remains. Science Advances, 4(12), [eaau5064]. https://doi.org/10.1126/sciadv.aau5064
Ancient nuclear genomes enable repatriation of Indigenous human remains

Joanne L. Wright1, Sally Wasef1, Tim H. Heupink1,2, Michael C. Westaway1, Simon Rasmussen3, Colin Pardoe4, Gudju Gudju Fournile5, Michael Young6, Trish Johnson7, Joan Slade8, Roy Kennedy9, Patsy Winch10, Mary Pappin Sr.11, Tapij Wales12, William “Badger” Bates13, Sharnie Hamilton14, Neville Whyman15, Sheila van Holst Pellekaan16,17, Peter J. McAllister17, Paul S.C. Taçon1, Darren Curnoe18, Ruiqiang Li19, Craig Millar20, Sankar Subramanian1,12,11, Eske Willerslev22,23,24, Anna-Sapfo Malaspinas25,26, Martin Sikora22, David M. Lambert16

After European colonization, the ancestral remains of Indigenous people were often collected for scientific research or display in museum collections. For many decades, Indigenous people, including Native Americans and Aboriginal Australians, have fought for their return. However, many of these remains have no recorded provenance, making their repatriation very difficult or impossible. To determine whether DNA-based methods could resolve this important problem, we sequenced 10 nuclear genomes and 27 mitogenomes from ancient pre-European Aboriginal Australians (up to 1540 years before the present) of known provenance and compared them to 100 high-coverage contemporary Aboriginal Australian genomes, also of known provenance. We report substantial ancient population structure showing strong genetic affinities between ancient and contemporary Aboriginal Australian individuals from the same geographic location. Our findings demonstrate the feasibility of successfully identifying the origins of unprovenanced ancestral remains using genomic methods.

INTRODUCTION

European colonization has had widespread effects on Indigenous peoples throughout the world. Dispossessed of their land and forcibly dispersed, their traditional ways of life were disrupted, and the survival of their culture and language was threatened (1). This dispossession has often included ancestral human remains that were taken from their resting places and distributed throughout the world. This is true of Native North Americans and Aboriginal Australians, the latter being one of the oldest continuous cultures outside Africa. Since the arrival of the British First Fleet to Australia in 1788, many ancestral remains of Aboriginal Australians have been collected for scientific research or display in museum collections (1, 2). In addition to the confiscation of ancestral remains, Aboriginal Australian children were often removed from their families and placed into institutions or with European families. These children, many of whom are now adults, are commonly referred to as the “Stolen Generation.” As a consequence, many of these Aboriginal Australians have lost their connections to Place and Country (1).

Over many decades, there has been a concerted effort by Aboriginal Australians to reach agreement with museums that would enable their ancestral remains to be returned to the original communities from which they were taken (3–7). This issue is of particular importance to Aboriginal Australians given their spiritual connection to the place in which they were born and lived (8). Many Aboriginal Australians believe that in order for their ancestor’s spirits to rest, their remains must be returned to their ancestral lands and their kin after death (3). For Aboriginal Australians, this is termed return to “Place and Country.”

Unfortunately, many of these remains have no specific details regarding their geographic origin, tribal affiliation, or language group (9, 10). For many, the only information provided is that these remains are “Aboriginal Australian” in origin. This lack of detailed information regarding their provenance means that many remains are unable to be returned. Museums and other institutions are unable to repatriate remains without first identifying the appropriate communities or custodians (2). Notably, this problem is not limited to Aboriginal Australians but is a worldwide issue affecting almost all Indigenous groups.

Recent advances in ancient DNA methods and bioinformatics suggest that a genomic approach can be successfully used to facilitate, on a large scale, the identification of unprovenanced remains. The success of such an approach has been illustrated by the identification and subsequent repatriation of the 8500-year-old Native American remains of “Kemewick Man” (11). To date, there have been a limited number of genomic studies of ancient Aboriginal Australians with only two focused on pre-European (12, 13) and two focused on post-European remains (14, 15).
The first ancient DNA study of Aboriginal Australians, published in 2001, reported the recovery of 10 short mitochondrial sequences (12). However, a later genomic study of the same ancient remains established that the earlier sequences were likely polymerase chain reaction (PCR) artifacts (13). The latter study reported the recovery of the complete mitogenome of an ancient Aboriginal Australian (labeled WLH4) (13). This was the first example of the recovery of authentic human DNA from an Australian archeological context, proving that it was possible to successfully recover ancient DNA, despite the harsh Australian climatic conditions and resulting poor DNA preservation. However, the recovery of the ancient nuclear genomes of Australia’s First People has been to date elusive.

RESULTS

Samples, sequencing, and authenticity

In collaboration with Aboriginal Australian Traditional Owners and communities across Australia, we undertook ancient DNA analyses of 27 sets of remains from archeological excavations of known Aboriginal Australian burial sites or previously repatriated remains of known provenance. The distinction between Aboriginal Australian genomes from pre- and post-European contact periods is important for the determination of the provenance of remains. This is because the resulting admixture can confound the determination of provenance (16). The ages of the skeletal remains of the ancient Aboriginal Australians studied here were determined from either archeological or \(^{14}\)C dates (Supplementary Materials). Different regions of Australia were settled by Europeans at different times. However, all of the ages of ancient samples pre-date European contact, thus excluding the possibility of admixture. This distinction is further supported by the presence of morphological and pathological characteristics including the absence of dental caries and distinctive tooth wear patterns typical of hunter-gatherer diets (17). The hair samples used were from individuals born before European settlement of the geographic regions sampled (table S1 and Supplementary Text).

Using DNA in-solution capture methods and second-generation sequencing, we successfully recovered 10 ancient nuclear genomes (0.3× to 6.9× coverage) and 27 mitogenomes (2.3× to 321× coverage) from pre-European contact remains of Aboriginal Australians (dated 95 to 1540 years before the present). In addition, for four of the male ancient Aboriginal Australians (KP1, MH8, PA86, and WLH4), we were able to recover partial or complete Y-chromosome sequences (haplogroups S1a and S1c). Preliminary shotgun sequencing showed that many of the ancient samples had very low levels of endogenous DNA (table S2). We therefore designed and used a modified method that used whole-genome capture baits (myBaits) to enrich targeted

---

**Fig. 1.** Details of the locations and language groups of Aboriginal Australian samples. Light orange shading indicates the distribution and location of Pama-Nyungan language families. Orange shading indicates the distribution of non–Pama-Nyungan language families. Dashed lines show the approximate distribution of accepted major language subgroups as published in (20) with language names in italics. Red symbols indicate previously published mitochondrial or nuclear genomes; blue symbols indicate new unpublished data. Circles indicate contemporary Aboriginal Australian samples, and stars represent ancient individuals. Sample code abbreviations have been included in parentheses.
We found that, by modifying the hybridization temperature to 57°C, the genome coverage obtained was significantly enhanced.

All recovered ancient sequences exhibited damage patterns characteristic of ancient DNA, with elevated levels of cytosine to thymine misincorporations in the 5′ end of fragments and guanine to adenine misincorporations in the 3′ end (Materials and Methods). In addition, contamination estimates for both mitochondrial and genome-wide sequences all displayed low contamination levels (Materials and Methods).

**Establishing comparative contemporary DNA datasets**

We used the recovered ancient Aboriginal Australian mitogenomes and nuclear genomes as proxies for unprovenanced remains to determine whether we could accurately identify their geographic origins using DNA-based methods. While we successfully recovered four partial or complete Y chromosomes from ancient male Aboriginal Australians, previous research on contemporary Aboriginal Australian males (16, 18, 19) has shown considerable levels of Eurasian admixture, with large numbers of research participants carrying non-Indigenous Y-chromosome haplotypes. The level of Eurasian admixture observed in contemporary Aboriginal Australian males varies greatly, with between ~32 and ~70% being observed in different regions of Australia (16, 18, 19). Undeniably, there has been a significant loss of Aboriginal Australian Y-chromosome genetic diversity since European settlement, perhaps with entire lineages lost to the past. This Y-chromosome admixture makes it extremely difficult to obtain a clear picture of the
paternal genomic history of Aboriginal Australians and, as such, would hinder attempts to find possible ancestral connections for repatriation purposes.

We constructed comparative contemporary mitochondrial and nuclear DNA datasets based on self-reported language group affiliations (16, 20), as well as geographic locations (Fig. 1). The contemporary nuclear DNA dataset comprised 100 high-coverage nuclear genomes of Pama-Nyungan language–speaking Aboriginal Australians. A total of 112 mitogenomes showing Aboriginal Australian–specific mitochondrial haplogroups were included in the mitochondrial DNA analyses, including 17 previously published genomes (table S3) (14, 21, 22).

As these contemporary datasets were assembled for the purpose of repatriation, they required a high degree of accuracy. Therefore, only previously published contemporary DNA sequences of known geographic origin and/or language group were used (14, 16, 21, 22). The recent publication of 111 mitogenomes recovered from historic hair samples from locations in Queensland and South Australia was not included, as the deposited sequences lacked precise geographic identifiers (15).

Mitochondrial genetic affinities

Using mitochondrial maximum likelihood phylogenetics (fig. S2), we compared 29 ancient Aboriginal Australian mitogenomes (14) with the 112 contemporary mitogenomes we previously assembled (Fig. 2). We observed 38 distinct mitochondrial haplogroups, with novel subclades discovered within mitochondrial haplotypes M42c *, R12a *, R12b *, and M42a3, while new subtypes were found for most other mitochondrial haplotypes (Supplementary Materials). For 18 ancient Aboriginal Australian individuals (62.1%), the closest contemporary match was an individual from the same geographic region (within 235 km). Within this group, nine ancient individuals could be matched to a contemporary individual within 100 km, and six could be matched to individuals from the exact location from which the ancient remains originated (Fig. 1).

For the remaining 11 ancient individuals (37.9%), the results were either inconclusive due to a lack of contemporary matches or because some mitochondrial haplotypes were geographically widespread. It has been previously documented that some Aboriginal Australian mitochondrial haplotypes have widespread distributions across the continent, while others are regional specific (15, 16, 23–25), reflecting ancient female migration patterns. While this is an interesting anthropological finding and may potentially inform the time depth of these practices, it is less helpful for repatriation. In two instances (6.9%), the closest ancient mitochondrial matches were not from the same geographic locations. In this case, the closest contemporary matches were individuals from opposite sides of Cape York Peninsula, some 635 km away (Fig. 1). As the return to Place and Country of ancestral remains is of paramount importance to many Aboriginal Australian communities, repatriation to an incorrect Country would be problematic. Therefore, the use of mitochondrial DNA alone is not recommended for repatriation.
Nuclear genetic structure of ancient and contemporary populations

Subsequently, we performed a series of analyses on the 10 ancient Aboriginal Australian nuclear genomes recovered. To investigate the overall genetic structure of ancient and contemporary Aboriginal Australian populations, we analyzed the individuals in the context of a reference panel including 2117 modern individuals from worldwide populations genotyped for 593,610 single-nucleotide polymorphisms (SNPs) (26, 27). Principal components analysis (PCA) and supervised model-based clustering (ADMIXTURE) revealed high levels of recent admixture across many Aboriginal groups, particularly those from Bourke (BKM) and Willandra Lakes (WIL) (Fig. 3). While most of the recent admixture is European in origin, we also observed evidence of East Asian gene flow, particularly among individuals from north-eastern Queensland (CAI and WPA). In contrast, individuals from the Western Central Desert (WCD) were almost completely unadmixed and protected ~13% Papuan-related ancestry in the individual from Cairns (PA86). This was also observed for contemporary individuals from New South Wales (WLH4 and KP1) (Fig. 5). Notably, we detected ~13% Papuan-related ancestry in the individual from Cairns (PA86). This was also observed for contemporary individuals from the same region (20).

We next sought to determine whether the ancient Aboriginal Australian individuals were most closely related to the individuals with known traditional connection to the same region, thereby facilitating repatriation. Genetic clustering using PCA or outgroup $f_2$ statistics both suggested a higher genetic affinity of the ancient individuals to local contemporary groups, compared to contemporary Aboriginal Australians from other geographic locations (Fig. 4). We further investigated these patterns using $f_3$ statistics in the form of $f_3$(Mbuti,Ancient; Contemporary,Papuan) on the masked dataset. This measured the amount of excess allele sharing of an ancient Aboriginal Australian individual with a given contemporary group when compared to Papuans. We found that the local contemporary groups consistently showed the highest level of sharing with the respective ancient Aboriginal Australian individual, supporting long-term local population continuity (Fig. 6). This finding is in concordance with previous studies of Aboriginal Australian skeletal remains (28, 29). For the two higher-coverage ancient Aboriginal
Australian individuals (KP1 and MH8), we additionally carried out haplotype sharing analyses. As previously supported with the allele frequency-based results, the largest excess haplotype sharing for both KP1 and MH8 was also with the local contemporary group (fig. S8).

**DISCUSSION**

Our analyses of the first ancient nuclear genomes of Aboriginal Australians reveal substantial past population structure. This result confirms the previous identification of an east versus west genetic divide between the contemporary Australian populations (16) while, at the same time, revealing further major geographic subdivision. When combined with the strong genomic affinities observed among ancient and contemporary populations from the same geographic locations, we showed that we could use these findings to reliably repatriate ancient unprovenanced remains to the correct Place and Country.

Over a long period, mitochondrial and Y-chromosome sequences have proved to be highly informative genetic markers for a diverse array of applications. This includes phylogenetic reconstruction, timing of divergent events, and tracing the spread of humans worldwide (30). Some researchers, as a result of a biogeographic study of mitochondrial DNA diversity (15), have proposed that these genetic marker sequences can be used to facilitate repatriation. However, in contrast, one of the major findings of the current study is that mitochondrial DNA sequence variation performs poorly in this regard. Our results suggest that mitochondrial sequences, if used in repatriation efforts in the Australian context, would result in a significant percentage (~7%) of remains being returned to the wrong Indigenous group.

We show that even under the arid conditions of Australia, low-coverage nuclear genomes can be recovered, and more importantly, these low coverage genomes can be used to precisely and accurately repatriate ancient remains. Furthermore, with advances in DNA capture and recovery methods, as well as with improvements in SNP analyses and the decreasing costs of genome sequencing, this general approach is likely to become more affordable and effective over time. We propose that our approach can be used now and will be used routinely in the future to return remains to their rightful kin. This approach could also allow for the identification of Place for members of the Stolen Generation. Because of the colonial history of removing children from their families, these people have lost not only their links to but also any knowledge of the location of their Country.

Our findings also suggest that a similar approach could be used to facilitate the repatriation of Indigenous remains in other countries.
with a known ancient population history and a contemporary database. This would represent a major scientific and social advance.

MATERIALS AND METHODS

Ancient DNA laboratory methods

All pre-PCR procedures were carried out in a dedicated Ancient DNA facility in the Australian Research Centre for Human Evolution, Griffith University. The facility is sealed, geographically isolated from any modern molecular laboratory, and has one-way airflow under positive pressure, and the air is high-efficiency particulate air (HEPA)–filtered. The skeletal remains and hair samples were processed within an ultraviolet-sterilized ultralow particulate air (ULPA)–filtered vertical laminar flow cabinet (used for this purpose only). Each sample was initially treated with 10% bleach to remove any surface contaminants and then washed with UltraPure DNAse/RNase-Free Distilled Water (Invitrogen) to remove any remaining bleach.

Skeletal material was processed using a Dremel rotary tool with a high-speed diamond cutter head or manually with a sterilized scalpel blade, where the outer surface was discarded. DNA was extracted from ~50 mg of bone or tooth powder, as previously described (13). Extraction blanks were included throughout all procedures. Hair samples were processed in 2 to 4 ml of digestion buffer, as previously described (14). This solution was incubated in a rotating incubator oven for 24 hours at 45°C. After complete digestion, the samples were centrifuged at 9000g for 3 min. The supernatant was combined with 10× volume of a modified binding buffer [500 ml of buffer PB (phosphate buffer; Qiagen), 1:250 pH indicator I, 15 ml of 3 M NaOAC (pH 5.2), and 1.25 ml of NaCl]. Extractions were purified using the MinElute Reaction Cleanup Kit (Qiagen) following the manufacturer’s protocol and eluted using 100 μl of buffer EB (elution buffer; Qiagen) after incubation for 10 min at 37°C.

Ancient DNA library construction, amplification, and screening

Double-stranded Illumina DNA libraries were built according to the methods previously described (13, 14, 31), with some minor modifications in the Taq polymerase used for amplification. Libraries built from different samples were amplified using different polymerases (table S2). All libraries were screened on a Bioanalyzer 2100 (Agilent)
to ensure that the DNA length distributions did not show any significant artifacts from amplification, e.g., artificially long molecules due to serial binding or primer dimers. Where these problems occurred, the number of PCR amplification cycles or primer concentration was adjusted. All PCR and extraction blanks were screened for contaminant library constructs on the Bioanalyzer.

**Whole-genome in-solution target capture**

Between 100 and 500 ng of library, amplified DNA was generated as described above using multiple secondary amplifications, some of which were sent for direct sequencing. DNA libraries were subjected to custom myBaits whole–human genome capture (Arbor Biotechnologies). Target capture enrichment was performed according to the manufacturer’s instructions; however, hybridization was performed for 36 to 42 hours at 55° to 57°C. The bead-binding buffers, initial 30-min incubation, and cleaning steps were also performed at this chosen hybridization temperature. Postcapture libraries were amplified on binding beads using the Kapa HiFi Uracil+ kit (Kapa Biosystems) according to the myBaits manual (version 3) for between 14 and 17 cycles.

**Ancient sequencing**

Ancient samples were sequenced using 100–base pair single-end reads. This sequencing was conducted using either a HiSeq 2500 Sequencing System (Illumina) at the Danish National High-Throughput DNA Sequencing Centre in Copenhagen or on a MiSeq Sequencing System (Illumina) using 150 version 3 kits at the Griffith University DNA Sequencing Facility. Sequences were base called using CASAVA 1.8.2 (Illumina).

**Contemporary sequencing**

DNA library construction and sequencing of contemporary samples were conducted at the Kinghorn Centre for Clinical Genomics at the Garvan Institute in Sydney, Australia or Novogene Bioinformatics Technology Corporation Limited in Beijing, China. Sequencing libraries were generated using the TrueSeq Nano DNA HT Sample Preparation Kit (Illumina, USA) following the manufacturer’s recommendations. Libraries were then 150–base pair paired-end sequenced on an Illumina HiSeq X. Genome coverage of this sequencing averaged between 45–60 x.

**Ancient DNA mapping and consensus calling**

Adapters were trimmed from the sequencing data using fastx_clipper, part of the FASTX-Toolkit version 0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit/) using parameters -Q 33 - 30. Levels of human DNA were determined by mapping reads to the human reference genome (GRCh37/hg19) or the revised Cambridge reference mitochondrial genome (32). Mapping was completed using BWA version 0.6.2 (33) with the following options: seed disabled (34) with terminal low-quality trimming (-q 15), before being aligned using BWA-0.6.2 aln with seed disabled. The mapped reads were sorted, and duplicates were removed and merged using SAMtools version 0.1.18 (35, 36).

**Contemporary data processing**

The paired-end contemporary DNA sequences were mapped to the human reference genome (GRCh37/hg19) or the revised Cambridge reference mitochondrial genome (32) using BWA version 0.6.2 (33). Duplicate sequences were then removed from Alignment/Map (SAM) files using SAMtools (35, 36). Using the mpileup command, with a maximum depth parameter of 1000, variant call format (VCF) files were generated for each chromosome separately. Using an awk command, indel variations were excluded. The VCF files of individual modern genomes were merged using VCFtools (37) after zipping and indexing using tabix.

**Mitochondrial maximum likelihood phylogenetics**

Consensus mitogenomes were generated, and ambiguous bases were checked and manually corrected. Mitochondrial haplotypes were identified using HaploGrep 2.0 software (38), with mitochondrial variations described in PhyloTree mtDNA Build 17 (39). Alternatively, haplotypes were identified manually, and novel ones were named in accordance with recent Aboriginal Australian mitochondrial haplogroup classifications (24).

All mitogenomes were aligned using SeaView version 4.6.1 (40). The mitochondrial evolutionary history of Aboriginal Australians was inferred using the maximum likelihood method based on the Tamura-Nei model (41) with 1000 bootstrap replications, as implemented in MEGA7 (42). The tree with the highest log likelihood (~19648.2315) was used (Fig. 2). Initial tree(s) for the heuristic search were obtained by applying Neighbor-Joining and BioNJ algorithms to a matrix of pairwise distances estimated using the maximum composite likelihood approach and then selecting the topology with a superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (categories +G, parameter =). This rate variation model allowed for some sites to be evolutionarily invariable ([+I], % sites).

The final tree was drawn to scale, with branch lengths measured in the number of substitutions per site and involved 141 mitochondrial sequences. All positions containing gaps and missing data were eliminated, and a total of 11,042 nucleotide positions were used in the final dataset. Subsequent annotation and presentation of the tree were completed using Interactive Tree of Life version 3.4.3 software (43).

**Analysis panel**

Genotyping of newly sequenced as well as previously reported modern individuals was carried out using SAMtools/bcf tools (35), followed by filtering, as previously described (44). Briefly, per-individual diploid genotypes were called using SAMtools mpileup (-C50 option) and bcftools call with the consensus caller (-c option). Calls from each genome were then filtered, excluding calls with low (six reads or one-third of the average sequencing depth, whichever was higher) or high (>2 times the average depth) coverage. We further filtered variants called within 5 base pairs of each other, with a Phred posterior probability of < 30 or a strand bias or end distance bias P value of <10, or with deviations from Hardy-Weinberg equilibrium with a P value of <10 across all samples.

For population genetic analyses, those genotypes were merged with a reference panel of 2286 modern individuals from worldwide populations, genotyped at 593,610 SNPs using the Affymetrix Human Origins array (26, 27). All ancient individuals were represented by pseudo-haploid genotypes obtained by sampling a random allele at each SNP position. For the two ancient samples with higher coverage, KP1 (6.9x) and MH8 (6.8x), we also carried out diploid genotyping, as described above, to be used in the haplotype sharing analyses.

**Population structure and admixture modeling**

PCA was carried out using smartpca (45) by projecting ancient individuals onto the components inferred from modern individuals using the “lsqproject” option. Genetic affinities of ancient and modern
individuals were investigated using the \( f \)-statistic framework (46). We used "outgroup \( f_s \)" statistics to determine the amount of shared genetic drift between pairs of individuals and/or groups, as well as \( f_s \) statistics for allele sharing symmetry tests (tables S5 and S6). Model-based clustering implemented in ADMIXTURE was used to investigate patterns of recent admixture, in supervised mode using European (French), East Asian (Han), Oceanian (Papuan), and Australian (WCD) individuals as putative source populations.

**Local ancestry inference**

Local ancestry deconvolution of the modern individuals was carried out using RFMix (47) and a panel of four reference populations: European (French), East Asian (Han), Oceanian (Papuan), and Australian (WCD). Before this analysis, we subsampled each reference population to the smallest population (WCD; 12 individuals) to avoid potential bias due to unbalanced panel sizes. A "masked" dataset was then obtained by restricting the analysis to SNPs for individuals who were homozygous for Australian ancestry (WCD).

**Haplotype sharing analyses**

Haplotype sharing analyses were performed using Shape-IT (49). We then performed chromosome painting for the two ancient individuals as recipients, using all modern Australians, as well as selected outgroups (French, Han, Papuan, and Bougainville) as potential donors. Differential sharing for the pair of ancient individuals was quantified using the symmetry statistic (50)

\[
S(A, B) = \frac{\text{Donor } A - \text{Donor } B}{\text{Donor } A + \text{Donor } B}
\]

SEs were obtained using a block jackknife across the 22 autosomes.

**Ancient DNA authentication**

Recovered ancient DNA sequences were authenticated using a number of methods. First, DNA damage patterns were estimated for each sample using mapDamage software (51). Samples showed a mean fragment length of 49.2 to 97.4 base pairs, with higher fragment lengths observed in the better-preserved hair samples. All samples exhibited damage patterns characteristic of ancient DNA, with elevated levels of cytosine to thymine misincorporations in the 5' end of fragments and guanine to adenine misincorporations in the 3' end (52) (table S2).

**Ancient DNA contamination estimates**

Mitochondrial contamination estimates were obtained using the contDeam and Schmutzi command in the Schmutzi software package (53). The estimates obtained from contDeam are based on the assumption that only endogenous DNA has deamination patterns typical of ancient DNA and that contaminant fragments are not deaminated and will therefore only reduce overall deamination rates. The Schmutzi command iteratively refines contamination estimates and produces consensus calls (results are presented in table S2). We excluded all ancient DNA libraries that showed contamination estimates higher than 3% from further merging and analyses.

It has been previously shown that deamination patterns typical of ancient DNA are rare in remains younger than 100 years in age (54), and this resulted in higher contamination estimates for the historical hair samples tested here. To verify the results obtained, we checked both the endogenous and contaminant consensus sequences generated by Schmutzi using the SAMtools tview command (35) and HaploGrep 2 (38). We showed that both endogenous and contaminant consensus sequences generated by Schmutzi carry the characteristic SNPs for the same mitochondrial haplotype.

For nuclear sequences, contamination was estimated using ANGSD for all ancient males (55). All contamination estimates generated are presented in table S2.

Supervised admixture of the ancient nuclear genomes was undertaken using five putative ancestry sources (fig. S9). Low levels of contamination from a European source were observed in some of the low-coverage ancient samples such as PA109, although it has been reported that using ADMIXTURE on low-coverage samples can result in statistical uncertainty associated with SNP and genotype calling, resulting in high error rates due to sampling, alignment, and sequencing errors (56). Our additional analyses using PCA confirmed that all ancient individuals cluster tightly with previously described Aboriginal Australians without recent admixture (WCD).

**Ancient DNA sex determination**

Sex determination was carried out using the method previously described (57), comparing the morphological and archeological information provided with each set of remains. In all instances, the assigned sex was as expected, which also rules out contemporary contamination from members of the opposite sex.

**SUPPLEMENTARY MATERIALS**

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/4/12/eaau5064/DC1

**Supplementary Text**

Fig. S1. Novel mitochondrial clades M42c1b and M42c1b1.
Fig. S2. Novel mitochondrial clades R12a, R12a1, and R12b.
Fig. S3. Novel mitochondrial haplotype M42a3.
Fig. S4. Novel mitochondrial haplotype P5a1a.
Fig. S5. Novel mitochondrial haplotypes P5a1a and P5a1a1.
Fig. S6. Novel mitochondrial haplotypes P12a and P12a1.
Fig. S7. Novel mitochondrial haplotype P12b.
Fig. S8. Local continuity between ancient and contemporary groups.
Fig. S9. Supervised admixture of ancient Aboriginal Australians using five putative ancestry sources.
Table S1. Weipa and Mapoon Hair samples background.
Table S2. Summary statistics for ancient Aboriginal Australians samples.
Table S3. Supervised admixture.
Table S4. Contemporary nuclear dataset.
Table S5. Masked outgroup \( f_s \) statistics.
Table S6. Masked outgroup \( f_s \) statistics.

**REFERENCES AND NOTES**

1. I. J. McNiven, L. Russell, Appropriated Pasts: Indigenous Peoples and the Colonial Culture of Archaeology (Altamira Press, 2005).
2. M. G. Simpson, Making Representations: Museums in the Post-Colonial Era (Routledge, 2001).
3. M. Truscott, Repatriation of Indigenous cultural property, paper prepared for the 2006 Australian State of the Environment Committee (Department of the Environment and Heritage, Canberra, Australia, 2006).
4. R. F. Langford, Our heritage—Your playground. Aust. Archaeol. 16, 1–6 (1983).
5. C. Pardoe, The eye of the storm: The study of Aboriginal human remains in Australia. J. Indig. Stud. 2, 16–23 (1991).
6. C. Pardoe, in Archaeological sites: Conservation and Management, S. Sullivan, R. Mackay, Eds. (Getty Conservation Institute, 2012), pp. 615–625.
8. V. Grieves, Aboriginal spirituality: Aboriginal philosophy, the basis of Aboriginal social and emotional wellbeing. Discussion paper no. 9, (Cooperative Research Centre for Aboriginal Health, 2009).

9. C. Pardee, In The Oxford Handbook of the Archaeology of Death and Burial, S. Tarlof, L. Nilsson Stutz, Eds. (Oxford Univ. Press, 2013), pp. 733–761.

10. D. Hanchart, in The Dead and Their Possessions: Archaeology and the Reburial Issue, C. Fforde, J. Hubert, P. Turnbull, Eds. (Gerald Duckworth & Co. Ltd., 2004).

11. J. Littleton, Hunter–gatherer dental pathology: Do historic accounts of Aboriginal Australians correspond to the archeological record of dental disease? Am. J. Hum. Biol. 30(7), e23076 (2018).

12. A. S. Al-Rasheid, F. Kronenberg, A. Salas, S. Schönherr, HaploGrep 2: Mitochondrial haplogroup classification in the era of high-throughput sequencing. Mol. Biol. Evol. 33, 1497–1507 (2016).

13. A. R. Jha, A. Tarekegn, D. Toncheva, S. Turdikulova, I. Uktveryte, O. Utevska, R. Vasquez, M. Villena, A. C. Windker, Y. Lepiskoposyan, P. Zalloua, T. Zemunik, A. Cooper, C. Capelli, M. G. Thomas, A. Ruiz-Linares, S. A. Tishkoff, L. Singh, K. Thangaraj, R. Villens, D. M. Das, R. Sukernik, M. Metspalu, E. Eichler, J. Burger, M. Slatkin, S. Pålbo, J. Kelso, D. Reich, J. Krause, Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature 513, 409–413 (2014).

14. C. Pardee, Global human variation: Polariated positions and alternative perspectives, in Before Farming (Liver University Press Online, 2010), pp. 1–21.

15. C. Pardee, Bioscapes: The evolutionary landscape of Australia. Archaeol. Ocean. 29, 182–190 (1994).

16. R. Nielsen, J. M. Akey, J. Pritchard, S. A. Tishkoff, E. Willerslev, Tracing the peopling of the world through genomics. Nature 541, 302–310 (2017).

17. M. Meyer, M. Kircher, Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. 6, pdb.prot5448 (2010).

18. R. Andrews, I. Kubacka, P. F. Chinnery, R. N. Lightowlers, D. M. Turnbull, N. Howell, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat. Genet. 23, 147 (1999).

19. H. Li, R. Durbin, Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595 (2010).

20. M. Schubert, A. Ginolhac, S. Lindgreen, J. F. Thomas, K. A. S. Al-Rashied, E. Willerslev, A. Krogh, L. Orlando, Improving ancient dna read mapping against modern reference genomes. BMC Genomics 13, 178 (2012).

21. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; 1000 Genome Project Data Processing Subgroup, The Sequence alignment/ map format and SAMtools. Bioinformatics 25, 2078–2079 (2009).

22. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27, 2987–2993 (2011).

23. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin; 1000 Genomes Project Analysis Group, The variant call format and VCFtools. Bioinformatics 27, 2156–2158 (2011).

24. H. Weissstein, D. Facher, A. Kloss-Brandstädtler, L. Forer, G. Specht, H. J. Bandelt, F. Kronenberg, A. Sala, S. Schönher, HaploGrep 2: Mitochondrial haplogroup classification in the era of high-throughput sequencing. Nucleic Acids Res. 44, W58–W63 (2016).

25. M. van Oyen, M. Kayser; Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum. Mutat. 30, E386–E394 (2009).

26. M. Gouy, S. Guindon, O. Gascuel, SeaView version 4: A multipurpose graphical interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27, 221–224 (2010).

27. K. G. H. Carder, D. M. Nei, Estimation of the number of nucleotide substitutions in the control region of mitochondrial dna in humans and chimpanzees. Mol. Biol. Evol. 10, 512–526 (1993).

28. S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874 (2016).
43. I. Letunic, P. Bork, Interactive tree of life (itol) v3: An online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 44, W242–W245 (2016).

44. M. Sikora, A. Seguin-Orlando, V. C. Sousa, A. Albrechtsen, T. Korneliusen, A. Ko, S. Rasmussen, I. Dupanloup, P. R. Nigst, M. D. Bosch, G. Renaud, M. E. Allentoft, A. Margaryan, S. V. Vasilyev, E. V. Veselovskaya, S. B. Borutskaya, T. Deviase, D. Comeskey, T. Higham, A. Manica, R. Foley, D. J. Meltzer, R. Nielsen, E. Excoffier, M. M. Lahr, L. Orlando, E. Willerslev, Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. Science eaao1807 (2017).

45. N. Patterson, A. L. Price, D. Reich, Population structure and eigenanalysis. PLOS Genet. 2, e110 (2006).

46. N. Patterson, P. Moorjany, Y. Luo, S. Mallick, N. Rohland, Y. Zhan, T. Genschoreck, T. Webster, D. Reich, Ancient admixture in human history. Genetics 192, 1065–1093 (2012).

47. B. K. Maples, S. Gravel, E. E. Kenny, C. D. Bustamante, RFMix: A discriminative modeling approach for rapid and robust local-ancestry inference. Am. J. Hum. Genet. 93, 278–288 (2013).

48. D. J. Lawson, G. Hellenthal, S. Myers, D. Falush, Inference of population structure using dense haplotype data. PLOS Genet. 8, e1002453 (2012).

49. O. Delaneau, J.-F. Zagury, J. Marchini, Improved whole-genome phasing for disease and population genetic studies. Nat. Methods 10, 5–6 (2013).

50. P. Skoglund, S. Mallick, M. C. Bortolini, N. Chennagiri, T. Hünemeier, M. L. Petzl-Erler, F. M. Salzano, N. Patterson, D. Reich, Genetic evidence for two founding populations of the Americas. Nature 525, 104–108 (2015).

51. H. Jónsson, A. Ginolhac, M. Schubert, P. L. F. Johnson, L. Orlando, mapDamage2.0: Fast and accurate detection of damage in ancient DNA. Bioinformatics 29, 1682–1684 (2013).

52. J. Dabney, M. Meyer, S. Pääbo, Ancient DNA damage. Cold Spring Harb. Perspect. Biol. 5, a012567 (2013).

53. G. Renaud, V. Slon, A. T. Duggan, J. Kelso, Schmutzi: Estimation of contamination and error rates in ancient DNA. Bioinformatics 30, 2928–2930 (2014).

54. S. Sawyer, J. Krause, K. Guschanski, V. Savolainen, S. Pääbo, Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. PLOS ONE 7, e34131 (2012).

55. T. S. Korneliussen, A. Albrechtsen, R. Nielsen, ANGSD: Analysis of next generation sequencing data. BMC Bioinformatics 15, 356 (2014).

56. R. Nielsen, T. Korneliusen, A. Albrechtsen, Y. Li, J. Wang, SNP calling, genotype calling, and sample allele frequency estimation from new-generation sequencing data. PLOS ONE 7, e37558 (2012).

57. P. Skoglund, J. Storå, A. Götherström, M. Jakobsson, Accurate sex identification of ancient human remains using DNA shotgun sequencing. J. Archaeol. Sci. 40, 4477–4482 (2013).

58. C. C. Macknight, Macassans and the Aboriginal past. Archaeol. Ocean. 21, 69–75 (1986).

59. C. R. c. heritage ALLIANCE, Department of Environment & Resource Management, A Thematic History of the City of Cairns and Its Regional Towns: Prepared for the Department of Environment & Resource Management, Queensland and the Cairns Regional Council (North Melbourne, Australia, 2011).

60. I. S. McIntosh, The birrinydji legacy: Aboriginals, macassans and mining in north-east Queensland. Aboriginal History 21, 70–89 (1997).

61. M. J. Morwood, D. R. Hobbs, The Asian connection: Preliminary report on Indonesian human remains using DNA shotgun sequencing. PLOS ONE 7, e1002453 (2012).

62. D. F. Branagan, Then look not coldly on science: Joseph Campbell, M. A journeyman cleric. J. Proc. R. Soc. N. S. W. 133, 245–293 (1993).

63. J. M. Bowler, H. Johnston, J. M. Olley, J. R. Prescott, R. G. Roberts, W. Shawcross, N. A. Spooner, New site for human occupation and climatic change in Lake Mungo, Australia. Nature 412, 837–840 (2001).

64. K. Douglas, Pictures of Time beneath Science, Heritage and the Uses of the Deep Past (CSIRO Publishing, 2010).

65. J. M. Bowler, H. Johnston, J. M. Olley, J. R. Prescott, R. G. Roberts, W. Shawcross, N. A. Spooner, New site for human occupation and climatic change in Lake Mungo, Australia. Nature 412, 837–840 (2001).

66. J. Allen, J. F. O’Connell, Both half right: Updating the evidence for dating first human arrivals in Sahul. Aust. Archaeol. 79, 86–108 (2014).

67. J. M. Bowler, H. Johnston, J. M. Olley, J. R. Prescott, R. G. Roberts, W. Shawcross, N. A. Spooner, New site for human occupation and climatic change in Lake Mungo, Australia. Nature 412, 837–840 (2001).

68. J. Allen, J. F. O’Connell, Both half right: Updating the evidence for dating first human arrivals in Sahul. Aust. Archaeol. 79, 86–108 (2014).

69. J. M. Bowler, H. Johnston, J. M. Olley, J. R. Prescott, R. G. Roberts, W. Shawcross, N. A. Spooner, New site for human occupation and climatic change in Lake Mungo, Australia. Nature 412, 837–840 (2001).
Author contributions: Project conceptualization: J.L.W., C.M., E.W., and D.M.L. Methodology: J.L.W., S.W., T.H.H., C.M., E.W., A.-S.M., and D.M.L. Validation of results: J.L.W., S.W., T.H.H., S.S., A.-S.M., and M.S. Formal analysis: J.L.W., S.W., T.H.H. Resources (samples): J.L.W., M.C.W., C.P., G.G.F., M.Y., T.J., J.S., R.K., P.W., M.P.s., T.W., W.B.B., S.H., N.W., S.v.H.P., P.J.M., and D.M.L. Resources (historical context and Indigenous knowledge): M.C.W., C.P., G.G.F., M.Y., T.J., J.S., R.K., P.W., M.P.s., T.W., W.B.B., S.H., N.W., S.v.H.P., P.J.M., P.S.C.T., and D.C. Software: S.R., R.L., S.S., A.-S.M., and M.S. Writing: J.L.W., S.W., M.S., and D.M.L., with critical input from all other authors.

Competing interests: The authors declare that they have no competing interests. Data and materials availability: Aboriginal Australian contemporary nuclear genome sequence data generated in this study have been deposited at the European Genome-phenome Archive (EGA; www.ebi.ac.uk/ega/), which is hosted by the EBI, under the accession number EGAS00001003359. Ancient Aboriginal Australian nuclear sequence data have been deposited at the European Nucleotide Archive (ENA; www.ebi.ac.uk/ena), which is hosted by the EBI, under the accession number PRJEB29663. Ancient Aboriginal mitochondrial data have been deposited in the GenBank database (http://ncbi.nlm.nih.gov/genbank), which is hosted by the National Center for Biotechnology Information, under accession numbers MK165665 to MK165690. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 17 June 2018
Accepted 20 November 2018
Published 19 December 2018
10.1126/sciadv.aau5064

Citation: J. L. Wright, S. Wasef, T. H. Heupink, M. C. Westaway, S. Rasmussen, C. Pardoe, G. G. Fourmile, M. Young, T. Johnson, J. Slade, R. Kennedy, P. Winch, M. Pappin Sr., T. Wales, W. B. Bates, S. Hamilton, N. Whyman, S. van Holst Pellekaan, P. J. McAllister, P. SC. Taçon, D. Curnoe, R. Li, C. Millar, S. Subramanian, E. Willerslev, A.-S. Malaspinas, M. Sikora, D. M. Lambert, Ancient nuclear genomes enable repatriation of Indigenous human remains. Sci. Adv. 4, eaau5064 (2018).
Ancient nuclear genomes enable repatriation of Indigenous human remains

Joanne L. Wright, Sally Wasef, Tim H. Heupink, Michael C. Westaway, Simon Rasmussen, Colin Pardoe, Gudju Gudju Fourmile, Michael Young, Trish Johnson, Joan Slade, Roy Kennedy, Patsy Winch, Mary Pappin, Sr., Tapij Wales, William “Badger” Bates, Sharnie Hamilton, Neville Whyman, Sheila van Holst Pellekaan, Peter J. McAllister, Paul S.C. Taçon, Darren Curnoe, Ruiqiang Li, Craig Millar, Sankar Subramanian, Eske Willerslev, Anna-Sapfo Malaspinas, Martin Sikora and David M. Lambert

Sci Adv 4 (12), eaau5064.
DOI: 10.1126/sciadv.aau5064