Ancient Solomon Islands mtDNA: assessing Holocene settlement and the impact of European contact

François-Xavier Ricaut a,*, Timothy Thomas b, Maru Mormina c, Murray P. Cox d, Maggie Bellattic, Robert A. Foley c, Marta Mirazon-Lahr c

a CNRS FRE 2960, Laboratoire d’Anthropologie, Université de Toulouse, Toulouse III Paul Sabatier, 31073 Toulouse cedex 3, France
b Anthropology Department, University of Otago, PO Box 56, Dunedin, New Zealand
c Leverhulme Centre for Human Evolutionary Studies, University of Cambridge, The Henry Wellcome Building, Fitzwilliam Street, Cambridge CB2 1QH, United Kingdom
d The Institute of Molecular Biosciences, the Allan Wilson Centre for Molecular Ecology and Evolution, and the Bio-Protection Centre, Massey University, Palmerston North, New Zealand

A R T I C L E   I N F O

Article history:
Received 6 August 2009
Received in revised form 20 November 2009
Accepted 9 December 2009

Keywords:
Ancient DNA
HVS-I region
Near Oceania
Solomon islands

A B S T R A C T

Archaeologists, linguists and geneticists generally agree that Near Oceania was subject to two major pulses of human dispersal: a Pleistocene occupation around 40,000 BP and a Late-Holocene migration at 3500 BP commonly associated with the Austronesian expansion out of Taiwan. The latter led to the development of the Lapita cultural complex in the Bismark Archipelago which resulted in the settlement of Remote Oceania and there are a variety of competing models (express train, slow boat, entangled bank, etc.) used to explain this. Recent genetic studies have focused on this issue, but none of them have taken into consideration the bias possibly introduced by 19th-century historically reported population decline caused by European contact.

In this paper we present a case study to test the effect of 19th–20th century colonial impact on the mitochondrial DNA diversity of Solomon Islanders and to investigate the complex stratigraphy of settlement in this archipelago during and after the Lapita period. We extracted DNA from hairs and teeth belonging to 21 individuals collected by the Somerville expedition during the late 19th-century, and typed them for mitochondrial DNA (mtDNA) hypervariable region I (HVS-I) and the intergenic COII/tRNALys 9-base pair deletion (9 bp-del). Comparison of these genetic data with those available from the modern Solomon Islanders and Southeast Asian and Oceanic populations conflicts with the hypothesis of drastic changes in Solomon maternal gene pool diversity, indicating that the last century putative bottleneck is not detectable through our genetic data. In addition, the ancient and modern Solomon haplogroup distribution (e.g. M27 haplogroup) suggests, in agreement with some archaeological and linguistic models, that Early Lapita populations expanding out of the Bismarck Archipelago had little or no contact with indigenous non-Austronesian populations in Bougainville and the Solomon Islands. This finding indicates smaller scale analyses reveal a more complex reality of genetic admixture in some parts of Oceania than is often assumed in current debates.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Studies of mitochondrial and non-recombining Y chromosome DNA variation in the modern populations of Oceania are well established and have provided useful insights into the prehistoric colonisation of the Pacific Islands, particularly when analyzed in comparison to archaeological and linguistic data (e.g. Melton et al., 1995; Redd et al., 1995; Lum and Cann, 1998; Cox and Lahr, 2006; Friedlaender et al., 2002, 2005, 2007; Kayser et al., 2006, 2008; Ricaut et al., 2008). In general the genetic ancestry of populations in Melanesia and Polynesia can be traced to the initial Pleistocene era settlement of Papua New Guinea and offshore islands some 40,000 years ago (BP), and subsequent admixture with Southeast Asians in the mid-Holocene. The latter process is thought to be associated with the rapid eastward spread of the Austronesian (AN) language family and a Neolithic economy, culminating in human expansion into Remote Oceania after 3200 BP (Bellwood, 2005).

Within this context a key focus of genetics research has been to define the extent of admixture between the indigenous populations and the AN speaking newcomers, with research favouring either little or no mixing (the ‘fast train’ scenario) or significant...
mixing (the 'slow boat' scenario). Increasingly, archaeological and genetic data lend support to models that posit the integration of Asian (AN speaking) peoples into indigenous Melanesian (non-AN speaking) communities in the vicinity of the Bismarck Archipelago circa 3450–3350 BP, leading to the development of new cultural forms (the 'Lapita cultural complex'), new languages (the Oceanic subgroup of AN) and populations of mixed ancestry, which ultimately went on to colonize Remote Oceania and Polynesia a few hundred years later (Kayser et al., 2008; Green, 1991; Kirch, 1997).

The genetic evidence for this mixing is seen most clearly when comparing the mtDNA and NRY phylogeographies of descendant populations. Polynesian mtDNA types are mostly (94%) of East Asian origin (Kayser et al., 2006), with a particular mtDNA HV1 motif (the 'Polynesian motif') able to be traced back through Island Melanesia to Eastern Indonesia, and arguably towards Taiwan (but see discussions in Friedlaender et al., 2007; Hill et al., 2007; Soares et al., 2008), the linguistic homeland of Austronesian (Redd et al., 1995; Trejaut et al., 2005).

The Polynesian motif is characterised by a set of mtDNA polymorphisms in the non-coding part of the mtDNA (nucleotide positions 16189, 16217, 16247 and 16261) that defines subgroup B4a1a1 within haplogroup B, which in turn is defined by the 9 bp-del (Redd et al., 1995). The Polynesian motif is found at highest frequency in Polynesia, with varying frequencies among coastal populations in Island Melanesia. However, the motif has not been found in the New Guinea Highlands, and is relatively rare in Island Southeast Asia although it does occur sporadically in both central and eastern Indonesia (Melton et al., 1995; Redd et al., 1995; Richards et al., 1998; Pierson et al., 2006; Friedlaender et al., 2002, 2005, 2007; Hill et al., 2007; Kayser et al., 2006, 2008; Ricaut et al., 2008). The presence of the Polynesian motif has also been confirmed in Madagascar – the western edge of the Austronesian expansion (Soodyall et al., 1995, 1996; Hurles et al., 2005). It has been argued that this motif developed in eastern Island Southeast Asia or Near Oceania (Trejaut et al., 2005; Friedlaender et al., 2007) during the mid-late-Holocene, between 10 kya and 4 kya (Pierson et al., 2006). Its immediate precursor has been identified through whole mtDNA sequencing in Taiwanese aboriginal groups (Trejaut et al., 2005), seemingly corroborating the Holocene Austronesian expansion out of Taiwan which led to the development of the Lapita cultural complex in the Bismark Archipelago and the subsequent settlement of Polynesia and Micronesia.

Polynesian Y chromosomes on the other hand, are of predominantly (66%) Melanesian origin (Kayser et al., 2006). This disparity suggests sex-biased admixture influenced by matriloc residence and matrilineal kinship reckoning amongst immigrant AN communities that resided in Melanesia for sometime before expansion to Remote Oceania (Hage and Marck, 2003; Cann and Lum, 2004; Kayser et al., 2008). Amongst indigenous non-Austronesian communities in coastal New Guinea and Island Melanesia we also see evidence of this mixing, but here mtDNA suggests a larger contribution of AN women, in keeping with patriloc non-AN (Papuan) social conventions (Ricaut et al., 2008; 363; Kayser et al., 2008).

Consequently, the regional-scale pattern of Holocene population history is generally depicted as involving a single eastward expansion of southeast Asian populations, which either slowed down (Kayser et al., 2008) or paused (Green, 1991; Anderson, 2001; Specht, 2007) in the Bismarck Archipelago, incorporating local peoples before continuing towards Polynesia in one movement. At the sub-regional scale however, there is archaeological and linguistic evidence for a more complex situation. The archaeological distribution of Lapita sites is generally discontinuous, occurring on some offshore islands and coastal locations in the Bismarck Archipelago but not others, and being absent along the north coast of New Guinea (Lilley, 2008). The present-day distribution of Austronesian languages there, happened much later in the post-Lapita period, with islands such as Karkar having a complex linguistic and genetic stratigraphy (Ross, 1988; Ricaut et al., 2008). The earliest Lapita settlements outside the Bismarcks are in the Reefs-Santa Cruz islands at 3200 BP (Green, 2003), perhaps suggesting that the initial Lapita expansion bypassed the entire main Solomons chain where there are no Early Lapita sites (i.e. 3400–2800 BP) (Sheppard and Walter, 2006). Recent linguistic analyses of the Reefs-Santa Cruz languages indicate that they derive from an early branch of the Oceanic subgroup whose proximal homeland is the Bismark Archipelago, in agreement with the archaeological evidence of a 'leapfrog' colonisation (Ross and Naess, 2007; Naess and Boerger, 2008). The Austronesian languages of the main Solomons belong to a later branch of the Oceanic subgroup (Ross, 1989), perhaps having arrived 2800–2600 BP when we first begin to see Late Lapita sites in the archaeological records of Buka, the Western Solomons, and Santa Ana (Sheppard and Walter, 2006). Prior to this period it is likely that the Solomon Islands were occupied solely by non-Austronesian speaking populations as evidenced by the patchy distribution of non-AN languages there today, and the presence of pre-Neolithic archaeological sites, such as Klu cave on Buka, dating to 29,000 BP (Wickler, 2001), and Vatuluma Posovi on Guadalcanal, dating to 6400 BP (Roe, 1993). Similar complexities are considered in recent debates about the grammatically anomalous AN languages of southern Vanuatu, which have been argued by some to indicate non-AN speakers were among the earliest inhabitants of the islands (approx. 3100 BP). In this scenario an early non-AN Lapita population first colonised the island, quickly followed by an AN Lapita population (Blust, 2008; Donohue and Denham, 2008), although this is highly contentious (Pawley, 2006).

It seems evident at this point that Holocene population movements associated with the spread of AN languages and Neolithic economies did not follow a strict ‘wave of advance’ model during expansion through Melanesia to Polynesia (Spriggs, 1997). Thus, more attention to localised sub-regional genetic patterns is warranted. However, whilst population genetics studies have been quite successful at elucidating the broad patterns of settlement in the Pacific, smaller scale analyses are fewer and problematic, with unexpected inversions in language and DNA ancestry remaining difficult to explain (Cox and Lahr, 2006; Ricaut et al., 2008; Friedlaender et al., 2007). Studies using modern DNA to infer demographic processes in the distant past are highly susceptible to the confounding effects of recent population movements. As Hunley et al. (2008) have recently argued localised genetic and linguistic exchanges tend to obscure microevolutionary patterns, yet leaving broader patterns still detectable.

The issue of post-settlement interaction and exchange in Melanesia is often completely ignored in genetic reconstructions. This is surprising since Melanesia is renowned for its numerous exchange networks, fluid social organisation, and mobile populations. Furthermore, it is archaeologically well established that populations went through extreme processes of post-Lapita diversification and regionalisation, involving the breakdown of region-wide exchange networks into smaller circuits of increasing specialisation and complexity (Spriggs, 1997; Kirch, 2000). Compounding this is the fact that Island Melanesian populations were subject to various important disruptions after European contact – introduced diseases caused population bottlenecks (with losses of up to 97% on some islands), pacification removed social boundaries once maintained by warfare, engagement in wage labour on plantations caused new movements and intermarriage, and the rapidity and ease of travel was facilitated by new modes of transportation.
Despite popular images of isolated tribes, Melanesian populations throughout the 20th century became increasingly cosmopolitan (Spriggs, 1997).

Inasmuch as geneticists have considered these confounding effects, simulation studies have been seen as the most viable solution (Cox, 2007). Another possibility is to use ancient DNA methods to recover samples for either the period in question or, ideally, at regular time intervals throughout prehistory. Such a strategy would still suffer from the problems of small sample size, but would at least allow an interrogation of the issue of temporal shifts in gene frequencies.

In this paper we present a case study addressing the above themes. We focus on the genetic ancestry of Solomon Islanders, a group still underrepresented in published studies. Our goal is to investigate the complex stratigraphy of settlement in this archipelago during and after the Lapita period as suggested by the recent archaeological and linguistic accounts presented above – particularly with respect to the possibility of Early Lapita having avoided the Solomons. As an initial step in coping with problems of uncertainty introduced by post-settlement population interactions and bottlenecks, we tested the effect of 20th century colonial impact through a comparative analysis of ancient mtDNA from curated specimens collected in the 1800s. This, of course, does not allow us to investigate sub-regional admixture in prehistory, but it is an important first step in approaching the issue of temporal change.

The samples in question were collected by Lieutenant Somerville (1897) during 1893–1894 whilst conducting a British Royal Navy cartographic survey of the islands of the New Georgia group in the Western Solomon Islands, and have resided in the Duckworth Collection, University of Cambridge, since his return to England from the voyage. Somerville collected hair from living individuals and skeletal remains from abandoned settlement complexes. It is likely that the latter relate to the local practice of keeping the skull of important lineage ancestors in small enclosures for the purpose of ritual activity and veneration (Walter et al., 2004). Many such sites occur in areas depopulated by warfare throughout New Georgia, and it is possible Somerville encountered these during his coastal surveys.

Somerville’s voyage followed the establishment of the British Protectorate of the Solomon Islands in 1893 and marks the beginnings of the colonial period. Throughout the 19th-century European traders and whalers had been active in the islands, but their numbers were few, and their primary impact appears to have been the introduction of metal weapons and firearms perhaps leading to the intensification of warfare. After 1893 however, the New Georgia group was actively pacified by the British. Missionaries arrived in 1902, and conversion to Christianity proceeded after 1910 (Bennett, 1987). The arrival of missionaries coincided with influenza and dysentery epidemics which decimated local populations, and also saw the beginnings of wage labour in the region (Woodford, 1922; Luxton, 1955). ‘Blackbirding’ or the capture of local people for indentured labour on plantations in Fiji and Australia did not affect the New Georgia group, being mostly restricted to Malaita, Guadalcanal and Makira from 1870 to 1903. However, during the 20th century people from these islands, having become accustomed to plantation work, increasingly migrated as workers to New Georgia (Bennett, 1987). It is not uncommon today to find people who have married into New Georgian villages from throughout the Solomons, and indeed men from Vanuatu, Papua New Guinea, Tonga, Malaysia, China, Australia, New Zealand, England and even Africa are also present. The Somerville collection thus presents an opportunity to examine DNA variation in New Georgia immediately prior to the period of greatest social and population disruption.

2. Materials and methods

2.1. Samples

It may be impossible to authenticate human aDNA sequences in certain cases (Cooper, 1997; Pääbo et al., 2004), and even stringent controls can fail to prevent or detect contamination (Handt et al., 1996; Kolman and Tuross, 2000; Gilbert et al., 2005; Sampietro et al., 2006). Consequently, to limit the risk of contamination careful attention was paid to the selection of samples. Because minimal information was available on the number of specialists who handled the specimens, from the time of their collection at the end of the 19th-century to the second half of the 20th century, we chose to use tooth and hair samples as a DNA source rather than bone samples. Indeed, as reported by Gilbert et al. (2006a), hairs are either impermeable to contamination or can be easily decontaminated, and teeth present a relatively impervious outer enamel layer which provides a degree of protection from contaminating DNA sources (Oota et al., 1995; Gilbert et al., 2003, 2006b). Several studies have shown that DNA is preserved better in teeth than in bone samples, and that teeth yield higher amounts of DNA than bone in many environments (Oota et al., 1995; Ricaut et al., 2005).

The tooth and hair samples used for this study were collected by Lt. Somerville during the 1893–1894 voyage of the HMS Penguin to the Solomon Islands. Our samples are from 21 individuals from various locations within the Western Solomons (Fig. 1).

DNA was extracted from the teeth of 13 individuals and from locks of hair (without roots) of 8 individuals. Because of the relatively poor preservation of the teeth and the selection criteria used (permanent tooth with closed apex, in a good state of preservation without any lesions (caries etc.), and still fixed to the jaw) only one tooth from each individual was used for DNA analysis. Nevertheless, two extracts were generated from most of the teeth allowing for a replication strategy. The hairs had been kept stored in glass and/or metal cases at room temperature, and are unlikely to have been handled frequently. The number of hairs available from each individual was large enough to allow replication of the extraction and amplification procedures.

2.2. DNA extraction

To extract DNA from teeth we used a modification of a previously published method (Gilbert et al., 2003) which reduces external contamination. One tooth was prepared and extracted at a time. Each whole tooth was washed in 50% bleach for 5 min and rinsed in bi-distilled water, and then exposed to ultraviolet light (254 nm wavelength, 12 W and 5 cm distance) for 30 min on both sides. Teeth were then fully encased in RTV-11 liquid silicone rubber (Tiranti) and were left for 24 h while the matrix hardened. The silicone that encases the top of the root was removed with a sterile blade, and the root tip was removed by sanding with a Dremel tool. The pulp chamber cavity and the surrounding dentine were ground to a fine powder with a Dremel drill bit, and the powder obtained (300–500 mg depending of the size of the tooth) was put in two DNA free tubes in order to perform two independent DNA extractions. When such a replication strategy was not possible (e.g. tooth analyzed was too small) the amplified products were systematically cloned.

Two independent extractions were made for each individual from 4 cm of hair shafts (cut in small pieces of 0.5–1 cm length). Prior to the DNA extraction, the hair shafts were treated to eliminate any sources of contaminating DNA present on their surface. Each hair shaft sample was twice immersed in a solution containing 1% SDS for approximately 30 s, followed by a rinse in bi-distilled water, then sterile ethanol 70%, and air-dried in a sealed sterile chamber.
Extractions were made using a modification of a previously described method (Ricaut et al., 2004). Briefly, each hair shaft and each tooth powder was digested in an extraction buffer containing 2% SDS, 10 mM Tris HCl pH 8.0, 0.3 M sodium acetate, 0.5 M EDTA, 1 mg/ml proteinase K, 40 mM DDT (dithiotreitol) and 0.2 M PTB (N-phenacylthiazolium bromide), and incubated under agitation for 12 h at 56 °C.

Post-digestion, DNA was extracted using phenol/chloroform/isoamyl alcohol (25:24:1, v/v/v) organic extraction performed on the supernatant. The aqueous phase was then purified with the CleanMix Kit (Talent), a method based on the large DNA affinity for silica in the presence of guanidium thiocyanate. Samples were then concentrated to 40 μl employing Microcon-30 filters (Millipore), after an elution step with 400-μl sterile water.

2.3. PCR amplification, cloning and sequencing

Mitochondrial DNA analyses were performed on the hyper-variable region 1 (HVS-I) of the mtDNA control region. The 9-base pair intergenic region V deletion (Cann and Wilson, 1983) was also amplified to confirm affiliation of the mtDNA sequences to haplogroup B. To determine the presence or absence of the 9 bp-deletion we amplified a fragment of approximately 120-bp including the mtDNA region V using primers L8196/H8297 (Handt et al., 1996). Hypervariable region 1 (HVS-I) of the mtDNA control region was amplified using two sets of overlapping primer pairs: L15989 (Gabriel et al., 2001)/H16239 (Ivanov et al., 1996), and L16190/H16410 (Gabriel et al., 2001). We also used the primer H16167 (5’-GGGTTTGATGTGGATTGGG-3’) (Ricaut et al., 2004) to resolve amplification problems linked to the polycytosine region located between nucleotide positions (np) 16184–16193. PCR conditions for these reactions were: pre-denaturation at 94 °C for 10 min, followed by 40 cycles at 94 °C for 30 s, 30 s at 48 °C (L15989/H16239, L15989/H16167 and L8196/H8297) or 51 °C (L16190/H16410), and 72 °C for 45 s; and final extension at 72 °C for 5 min.

We also amplified shorter DNA fragments by using additional overlapping primer sets L15996/H16139, L16131/H16218, L16209/H16356, and L16287/H16410 (from Handt et al. (1996) except for L15996 from Endicott et al. (2003)). Cycling parameters were pre-denaturation at 94 °C for 5 min, followed by 40 cycles at 94 °C for 1 min, 1 min at 53 °C, and 72 °C for 1 min, and final extension at 72 °C for 5 min.

PCR amplifications were carried out in 50 μl of reaction mixture containing 2–6 μl of the ancient DNA extracts, 10 mM Tris HCL pH 8.3, 50 mM KCl, 1.5 mM MgCl2, 1 mg/ml BSA, 200 μM each dNTP, 0.25 μM each primer and 2 U of Taq Gold Star (Eurogentec).

Intergenic region V and HVS-I region amplification products were visualized on a 1% agarose gel and purified with Microcon-PCR filters (Millipore) and systematically sequenced. Sequence reactions were performed on each strand, with the same primers as those employed for PCR amplification, by means of ABI Prism BigDye Terminator Cycle Sequencing Kit (PE Applied Biosystems) according to the manufacturer's specifications. The sequence reaction products were analyzed on an ABI Prism 3100 (PE Applied Biosystems) automated DNA sequencer in the Sequencing Service of the Department of Zoology (Oxford University).

The samples with sequence electropherograms that indicated heteroplasmic positions or presented a low degree of reproducibility of results (caused by random DNA damage, Taq misincorporation and contamination from exogenous DNA) and the individuals for which only one extraction was possible (e.g. some of the tooth samples), were systematically cloned. For each of these samples, up to 16 colonies from two PCRs/extractions were amplified and sequenced. PCR products were cloned using the TOPO TA Cloning kit (Invitrogen), according to the manufacturer’s instructions. Colonies were used to initiate PCR reamplifications with vector M13R and M13F primers, and purified and sequenced as previously described. The sequences were aligned against the rCRS sequence (Andrews et al., 1999), analyzed for post-mortem damage, induced miscoding lesions and the presence of contaminant DNA sequences. If no evidence of contamination could be observed, the multiple PCR fragments were assembled and the consensus sequence was determined.
2.4. Contamination precautions

As explained above, we chose to work from biological material such as hairs and teeth which have been shown to be relatively impermeable to contamination and/or easily decontaminated.

Pre-extraction, extraction procedures, and mixing of reactions for PCR, were performed in isolated areas at the Palaeogenetics Laboratory (Leverhulme Centre for Human Evolutionary Studies, University of Cambridge, United Kingdom), which is located in a different building from the main laboratory where PCR amplification and subsequent lab work take place. In accordance with suggested ancient DNA procedures (Cooper and Poinar, 2000) the ancient-DNA laboratory is dedicated to the analysis of low copy number samples, and no amplified product is present in this laboratory. The area is protected against contamination with positive air pressure, filtered airflow, routinely cleaned with bleach, DNase Away (Molecular BioProducts) and irradiated with ultraviolet (UV) light (254 nm). All the pre-extraction steps (i.e. cleaning, cutting and drilling of the teeth) took place in a glove-box, within the ancient-DNA laboratory. The glove-box is routinely cleaned with bleach, DNase Away (Molecular BioProducts) and irradiated with ultraviolet (UV) light (254 nm). The extraction step and mixing of reactions for PCR took place in two different hoods, both integrated within the ancient-DNA laboratory, and protected against contamination in the same way (positive air pressure, UV irradiation, bleach, and DNase treatment).

The samples were always handled wearing gloves, breathing masks and complete protective clothing, and all laboratory equipment and consumables (lab coats, pipettes, tubes, filter tips, chemicals and dedicated reagents, etc.) were DNA free or of the highest purity available, and/or sterilized by a long UV exposure.

To detect possible contamination by exogenous modern DNA, extraction and amplification blanks were used as negative controls at a ratio of one control to one sample; and all persons involved in processing samples were genetically typed and compared to the results obtained from the ancient Solomon samples. For each Solomon hair sample and some of the teeth samples when possible, at least two extractions were undertaken at different times, and at least four amplifications by extraction were made to assess the reproducibility of the results. Furthermore, as explained above, PCR products were cloned if results were unable to be reproduced or if only one extraction was performed from an individual. We also attempted to amplify, using the different primer sets described above, mtDNA fragments of different length to test the molecular behavior of amplified products, and to detect an inverse correlation between the amplification efficient and the length of the amplicons.

2.5. Data analysis

The HVS-I sequences obtained from the Solomon Islands samples were compared to the mtDNA sequences of 9464 individuals from East and Southeast Asia, compiled from the DDBJ/EMBL/GenBank international nucleotide sequence database and additional literature (additional Table 1). All these sequences were aligned using the BioEdit 5.0.9 program (Hall, 1999). To determine the frequency of the Solomon Islands HVS-I haplotypes in the populations used for comparative analysis we used the Arlequin package (http://www.lgb.unige.ch/arlequin). Solomon Islands mtDNA sequence haplogroup classification was based on the accepted nomenclature (e.g. Kivisild et al., 2002; Kong et al., 2006; Friedlaender et al., 2005; Trejaut et al., 2005; and Merriwether et al., 2005), and the frequency of these haplogroups was determined in the reference populations. As in previous publications (Merriwether et al., 2005; Friedlaender et al., 2005), coding-region analysis was not performed because the sequencing of hyper-variable region 1 (and 9 bp-deletion for a subset of samples) allowed us to assign all the sequences to a known haplogroup without ambiguity (the full HVS-I haplogroup motifs are represented in all of the ancient Solomon sequences).

Differences in haplogroup frequencies between ancient and modern Solomon samples were determined using an in-house resampling algorithm implemented in R (http://www.r-project.org/) (code available on request).

The spatial frequency distribution of some haplogroups in Island Southeast Asia, Near and Remote Oceania, based on the longitude and latitude for the centre points of each sample location and frequency data of each population, was performed using the ESRI ArcGIS software package.

3. Results

3.1. Ancient DNA authentication

One of the most critical issues in the analysis of ancient DNA concerns the authenticity of the DNA sequences obtained, especially from human remains, because of the possibility of artifacts becoming included in the DNA sequences, and the risk of contamination with modern human DNA molecules. Strict criteria for authenticity have been established (Handt et al., 1996; Cooper and Poinar, 2000; Hofreiter et al., 2001; Pääbo et al., 2004), but even when closely followed these alone cannot prove with certainty the authenticity of human aDNA (Pääbo et al., 2004). A strong logical chain of evidence is required to authenticate aDNA results (Gilbert et al., 2005). In the present study, we can exclude contamination with a high level of confidence and attest to the authenticity of mtDNA results on the following grounds:

(1) The material used for aDNA analysis was either highly impermeable to contamination with modern DNA or easily decontaminates, as explained in the previous section. Moreover, the teeth used were removed from secure sockets in the mandible, and thus their roots had been protected from contaminating DNA during previous handling of the remains.

(2) The analyses were undertaken in a dedicated aDNA laboratory, under strictly controlled conditions. Extraction and PCR blanks showed no signs of contamination. Three teeth samples, one of which (Sol-18) did not yield positive amplifications, and two (Sol-4421 and Sol-4429) which showed irreproducible results, were discarded.

(3) The DNA sequences obtained from each individual were either fully reproducible from multiple extractions and amplifications. Alternatively, when only a single extraction was possible or clear heteroplasmic polymorphisms observed, nucleotide changes were confirmed by cloning (Appendix B). The sequences were confirmed on both strands and reconstructed from overlapping HVS-I fragments obtained with the different primers set (reproducibility of substitutions from different overlapping fragments/clones). This approach also enables us to exclude mosaics patterns or phantom mutations (Bandelt et al., 2002; Bandelt, 2005).

(4) Post-mortem DNA damage typical in ancient samples was observed in all samples (Hofreiter et al., 2001), but could not be reproduced. It has been shown that post-mortem damage may also accumulate in old (>10 years) contaminant DNA to similar
Mitochondrial HV-I sequences (between positions 16018 and 16384) from ancient Solomon islanders and their haplogroup attributions.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Haplotypes</th>
<th>Location</th>
<th>Polymorphic positions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sol-AS21.12 (H)</td>
<td>CCCC X.I B4</td>
<td>Solomon Isl.</td>
<td>1166</td>
</tr>
<tr>
<td>Sol-147 (H)</td>
<td>CCCC T X.I B4a</td>
<td>Western Solomon Islands</td>
<td></td>
</tr>
<tr>
<td>Sol-4424 (T)</td>
<td>CCCC CCCC T X.I B4a</td>
<td>New Georgia</td>
<td></td>
</tr>
<tr>
<td>Sol-4426 (T)</td>
<td>CCCC CCCC T X.I B4a</td>
<td>New Georgia</td>
<td></td>
</tr>
<tr>
<td>Sol-C1 (H)</td>
<td>A CCCC CCCC G T X.I B4a1a1</td>
<td>Solomon Isl.</td>
<td></td>
</tr>
<tr>
<td>Sol-S1 (H)</td>
<td>CCC CCCC G T X.I B4a1a1</td>
<td>Simbo</td>
<td></td>
</tr>
<tr>
<td>Sol-OC14.2 (H)</td>
<td>A CCCC CCCC G T X.I B4a1a1</td>
<td>Western Solomon Islands</td>
<td></td>
</tr>
<tr>
<td>Sol-117 (H)</td>
<td>CCC CCCC G T B4a1a1</td>
<td>Simbo</td>
<td></td>
</tr>
<tr>
<td>Sol-41 (T)</td>
<td>CCC CCCC G T B4a1a1</td>
<td>Western Solomon Islands</td>
<td></td>
</tr>
<tr>
<td>Sol-4428 (T)</td>
<td>A T CCC T T G C T XX M27</td>
<td>New Georgia</td>
<td></td>
</tr>
<tr>
<td>Sol-4430 (T)</td>
<td>A T CCC T T G C T XX M27</td>
<td>New Georgia</td>
<td></td>
</tr>
</tbody>
</table>

Note: Dash (-) indicates that the nucleotide positions could not be determined.

a. Numbered according to the revised Cambridge Reference Sequence (Andrews et al., 1999).

b. Region V length-polymorphisms variant according to Lum and Cann (1998) classification.

c. Region one 9-bp repeat (ACCCCCTCT); XX = two 9-bp repeats (ACCCCCTCT ACCCCCTCT).

d. ND, not determined.

These findings allow us to consider it highly unlikely that the data arise from artifacts (specific contamination, post-mortem damage, heteroplasmy, or sequencing artifacts), and validate and authenticate the molecular results obtained in this study. A 367 base pair (bp) segment of the mitochondrial DNA HVS-I region was sequenced and confirmed (nucleotide position (np) 16018–16384 of the rCRS, revised Cambridge Reference Sequence) (Andrews et al., 1999) for 17 Solomon teeth/hair samples out of the 21 initial samples (Table 1). DNA sequence data have been deposited to Genbank (Accession Numbers: GQ275330–GQ275347).

### 3.2. Phylogeographic analyses

All the retrieved ancient Solomon Islands sequences belong to East Asian, southeast Asian (B4/B4a, B4a1a1) and ancient Near Oceanic (Q1, M27) mtDNA haplogroups, common amongst populations from Island Melanesia (Appendix D and E; Table S2 of Friedlaender et al., 2007). Despite the small size of our 19th-century Western Solomon Islands sample, the mtDNA gene pool composition is relatively similar to that of modern Solomon Islanders (Appendix D; Table S2 of Friedlaender et al., 2007).

Haplotype and haplogroup (Appendices D and E) assignments for our samples can be grouped into three different clusters, each having different geographic and temporal origins. Haplogroup Q1 (6% of the sample), with a coalescent time of approximately 22 kya and a geographic distribution centred on New Guinea and Island Melanesia (Friedlaender et al., 2007), is thought to derive from the population which first settled New Guinea/Island Melanesia prior to 40 kya BP (Groube et al., 1986; Leavesley et al., 2002). Haplogroup M27 (17.6% of the ancient Solomons) probably emerged in Island Melanesia approximately 40 kya ago (Merriwether et al., 2005; Friedlaender et al., 2007) and has a geographic distribution restricted to this area (Appendices D and E), with its highest frequency centred on Bougainville and the Solomon Islands (Appendix D). Haplogroup B4/B4a (29.4% of the ancient Solomons) is widely reported throughout Southeast Asia and Oceania (Appendix D), while its sub-haplogroup B4a1a1 (47% of our sample), which includes the “Polynesian motif”, is restricted to Near and Remote Oceania, Indonesia and Madagascar, and probably originated either in Indonesia or in Near Oceania sometime during the Holocene (Pierson et al., 2006; Oppenheimer, 2004).
4. Discussion

4.1. Colonial bias

Recent genetic studies focusing on the population history of Oceania have not taken into consideration the potential bias introduced by population movements and population decline through introduced diseases during the colonial era. European contact has been reported to have led to the loss of more than 97% of the population on some Melanesian islands (Spriggs, 1997). Such declines could potentially decrease genetic diversity, with minor haplogroups perhaps experiencing extinction. Our current analysis, through a comparison of the mtDNA genepool of 19th-century Western Solomon Islands within the modern Solomon Islands population, conflicts with the hypothesis of drastic changes in genetic diversity, and indicates that a 19th-century bottleneck is not easily detectable for most haplogroups, at least within the Solomon Islands maternal genepool. The presence and frequency of haplogroups are similar in both the modern and 19th-century Solomon samples, with the key exceptions of haplogroups B4 and B4a. Indeed, statistical analysis shows that the frequencies of these two lineages is higher in the 19th-century sample compared to Solomon Islands samples, with the key exceptions of haplogroups B4 and B4a. Indeed, statistical analysis shows that the frequencies of these two lineages is higher in the 19th-century sample compared to Solomon Islands samples, with the key exceptions of haplogroups B4 and B4a. This analysis also indicates that (i) except for B4 and B4a, none of the ancient haplogroup frequencies are statistically different from modern frequencies, and (ii) although haplogroups P, Q2, M27b, M28 and M29 are not observed in the 19th-century Solomon sample, their absence is not statistically unlikely given the relatively low ancient sample size (n = 17) and should not be attributed to increasing diversity due to colonial migration in recent times. Except for haplogroups B4 and B4a (whose elevated ancient frequencies may reflect recent demographic events or possible sampling bias), these results suggest relative continuity and stability in the Solomon mtDNA genepool since the time these samples were collected. This therefore shows that the mixed Asian and indigenous Melanesian composition of the population is not of recent origin and has implications for the study of population history in this area. Admixture is most likely to have occurred in prehistory, due to numerous small and large scale population movements during the Holocene (Allen, 2003; Ricaut et al., 2008; Friedlaender et al., 2007). This result consequently increases our confidence in the representativeness of samples from the region, since these do not appear to have been significantly distorted by periods of unusual population decline or mobility. It also decreases the degree of uncertainty when addressing population relationships at the sub-regional level through genetic studies of modern populations.

4.2. The Solomon Islands in the context of Lapita

The question of how the Solomon Islands fit into the Lapita expansion is a matter of some debate amongst archaeologists. The Lapita cultural complex developed in the Bismarck Archipelago 3450–3350 BP and led to the settlement of the Reefs–Santa Cruz, Vanuatu, New Caledonia, Fiji, Tonga and Samoa by Austronesian speakers 3200–2800 BP (Spriggs, 1997; Specht, 2007). Lapita sites in the Solomon Islands are anomalous within this trajectory, being of very low concentration and dating exclusively to the Late Lapita period, post 2800 BP (Summerhayes and Scales, 2005; Felgate, 2007). Explanations for this pattern vary but can be grouped into three schools of thought: 1) Lapita expansion was geographically continuous, but archaeological surveys have missed Early Lapita sites in the Solomons due to patchy sampling or poor site preservation (Green, 1978; Spriggs, 1997; Kirch and Hunt, 1988: 18, Felgate, 2007); 2) Lapita expansion was discontinuous, with Early Lapita communities ‘leapfrogging’ the Solomon Islands (Gorecki, 1992; Roe, 1993; Sheppard and Walter, 2006); 3) Early Lapita expansion was demographically weak in the Solomons, ‘limping’ due to already present populations and the effects of Malaria and other natural predators. Thus sites from the period are far rarer than in Remote Oceania, where expansion was explosive (Felgate, 2007).

The evidence from historical linguistics adds to this picture but is not able to resolve the matter. The Austronesian languages distributed from the Bismarck Archipelago to Santa Isabel in the Solomons are classified as Western Oceanic, with New Ireland, Bougainville and the Western Solomons forming a lower order linkage known as Meso-Melanesian. To the east of Santa Isabel the remaining islands of Melanesia belong to Central/Eastern Oceanic, with the eastern Solomons and the Reefs–Santa Cruz forming a subgroup known as Southeast Solomonic. Explaining this boundary has proved difficult. Ross (1989) has argued that it is the result of successive Oceanic expansions, with Southeast Solomonic speakers initially colonising the entire region (i.e. Lapita), only to be replaced at a later date by Meso-Melanesian speakers moving down the archipelago as far as Santa Isabel. This is problematic because there is no evidence of Meso–Melanesian having replaced an earlier Oceanic language. Moreover the Western and Northern Solomonic languages have a significant number of ‘non-autochthonous’ (i.e. non-Melanesian) languages that are considerably older (Dunn et al., 2005) yet survive, despite on-going language replacement (Terrill, 2003). Sheppard and Walter (2006) present an alternative hypothesis, arguing that the distribution of Central/Eastern Oceanic reflects a leapfrog migration of Early Lapita to the Reefs–Santa Cruz, and that Meso-Melanesian was introduced via a Late Lapita migration through the Northern and Western Solomons in accord with archaeological evidence. The latter scenario has received recent support from linguists, who after re-examining the Reefs–Santa Cruz languages, have argued that these broke off from proto-Oceanic very early, prior to later Austronesian expansion through the Solomons chain, and are closest historically to the St. Matthias languages of the northern Bismarcks (Ross and Näss, 2007; Näss and Boerger, 2008).

We would expect each of these scenarios presented by archaeologists and linguists to potentially leave slightly different genetic traces. Key to the argument is the diverse indigenous Melanesian haplogroups which have deep ancestry in the region, and present some geographic stratification. The distribution and frequency of Asian derived haplogroups, on the other hand, is largely irrelevant since all scenarios accept an eventual Austronesian spread to all locales. If Lapita expansion from the Bismarck Archipelago was continuous, with intrusive southeast Asian peoples ‘picking up’ local Melanesians during a gradual eastward movement, we would expect to find the genetic makeup of Remote Oceanic populations to include a broad cross-section of indigenous DNA types present in Near Oceania. On the other hand, if the Lapita expansion was discontinuous we would expect to find Remote Oceanic populations clustering more closely with Bismarck Archipelago populations than groups in the Solomon Islands.

Our aDNA data, in addition to modern sampling by previous researchers, indicates the second of these possibilities is currently to be favoured. The M27 haplogroup has been argued to have developed in Bougainville and the Solomon Islands due to its frequency and diversity there (Fig. 2; Merriwether et al., 2005; Friedlaender et al., 2007). M27a is restricted solely to Bougainville and the Solomons; M27b is mostly found in the Solomons but has rare occurrences in New Britain (14 individuals), Buka (1 individual) and New Caledonia (1 individual); and M27c is found in South Bougainville and the Solomons in low numbers, and even more rarely in New Ireland (Appendix D and E). Thus, with the exception of the single individual from New Caledonia, M27 is not
found in Remote Oceania, as would be expected had Lapita populations moved slowly through the Solomon Islands. Instead, amongst populations in the Reefs-Santa Cruz, Vanuatu, New Caledonia, and Fiji we find indigenous Melanesian haplogroups P, Q, M28 and M29 to be present, most variants of which are rare or absent in the Solomons but very frequent in the Bismarck Archipelago (Appendix D; Friedlaender et al., 2007). The Reefs-Santa Cruz islands particularly, group closely with populations in New Britain in terms of haplogroup frequencies (Friedlaender et al., 2002).

On the basis of this evidence there appears to be congruence between archaeological, linguistic and genetic data with regard to the possibility that Early Lapita populations expanding out of the Bismarck Archipelago had little or no contact with indigenous non-Austronesian populations in Bougainville and the Solomon Islands. However, our evidence cannot distinguish between Sheppard and Walter’s (2006) ‘leapfrog’ and Felgate’s (2007) ‘limping’ archaeological models, in the sense that the observed genetic patterns could be accounted for by either (a) an initial long-range colonisation of the Reefs-Santa Cruz by a mixed Asian/Bismarck Melanesian population, followed by a secondary more gradual Austronesian colonisation of the Solomons chain; or (b) a simultaneous coverage of the Solomons and the Reefs/Santa Cruz by a mixed Asian/Bismarck Melanesian population, which was highly successful in the unoccupied Remote Pacific, but was demographically very weak in the Solomon Islands. Option (b), however, would also require a secondary Austronesian colonisation of the Solomons chain in order to account for the high proportion of Asian lineages. The linguistic evidence perhaps currently favours option (a), but there is still room for uncertainty here. This is particularly the case since our analysis is based on mitochondrial DNA alone, reflecting only maternal population history and being susceptible to genetic drift, especially in contexts with relatively small population sizes.

Whatever the smaller scale details, the results of this analysis suggest that it may be time to move beyond the rather simple models of Lapita expansion commonly examined by population genetics studies. Beyond the two commonly presented options of either (i) little or no mixing of rapidly migrating Asian populations (the ‘fast train’ hypothesis) or (ii) significant admixture of Asian and indigenous Melanesian populations during expansion (the ‘slow boat’ hypothesis), we should consider a third possible scenario based on a more complex reality of genetic admixture in some areas, but not in others.

Acknowledgements

This research was supported by NERC as part of the EFCHED Thematic Programme ‘Searching for traces of the Southern Dispersal’ Project, and by the AHRC through the European Science Foundation EUROCORES Origins of Man, Language, and Languages (OMLL) programme “Pioneers of Island Melanesia”. Further support was also provided by the Isaac Newton Trust of Trinity College, Cambridge, and the Leverhulme Trust.

Appendix. Supplementary material

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.jas.2009.12.014.

References


