Indonesian Mitochondrial DNA and Its Opposition to a Pleistocene Era Origin of Proto-Polynesians in Island Southeast Asia

MURRAY P. COX

Abstract The origin of modern Polynesians, the route of their expansion into the Pacific Ocean, and the timing of their movements all remain contentious topics in modern anthropology. Numerous studies have used molecular data to elucidate settlement patterns in the Indo-Pacific region, but the same evidence is often interpreted in opposing ways by different researchers. Above all, mitochondrial DNA (mtDNA) diversity has been used to discriminate between competing migration models and has narrowed the probable source of proto-Polynesian peoples to southern China and Taiwan or eastern Indonesia. Richards et al. (1998) used a dating method employing the $\rho$ statistic to argue for an origin of Polynesian peoples in eastern Indonesia during the Pleistocene (>10,000 years ago). Here, the time to the most recent common ancestor (TMRCA) is recalculated for a new series of Indonesian mtDNA sequences with Polynesian affinities. These data, which incorporate additional sequences published after 1998, produce dates that cannot rule out the possibility of a common ancestor for these sequences during the Holocene (<10,000 years ago). This implies that previous estimates of TMRCA dates for Indonesian sequences lacked the statistical robustness necessary for replicability. The extant mtDNA evidence can no longer be viewed as favoring a Polynesian origin in eastern Indonesia, but instead remains consistent with an origin of proto-Polynesian peoples in southern China and Taiwan.

Excluding proposals for exclusively American (Heyerdahl 1952) or exclusively Melanesian (Terrell et al. 2001, p. 105) ancestry of proto-Polynesian peoples—both of which lack support from molecular evidence (Bonatto et al. 1996; Merriwether et al. 1999)—only two models dominate contemporary discussion surrounding the origin of modern Polynesians. The first model proposes that Polynesians ultimately derive from populations in southern China (Bellwood
where the development of a farming economy during the early Holocene (<10,000 years ago) stimulated a demographic expansion of local peoples. Radiating outward in search of new agricultural lands, these proto-Polynesians reached Taiwan by about 6,000 years ago and spread successively to the Philippines, eastern Indonesia, and New Guinea about 3,500 years before present (B.P.). They reached Vanuatu and Fiji by about 3,200 years B.P., going on to become the first inhabitants of the remote Pacific Ocean and establishing a direct line of descent to contemporary Polynesian islanders.

The second model proposes that proto-Polynesian peoples originated near the center of island Southeast Asia during the late Pleistocene (>10,000 years B.P.) (Oppenheimer 1999; Oppenheimer and Richards 2001). A series of sea-level rises following the last glacial period flooded sizable regions of the Philippines and Indonesia. This deluge induced an eastward exodus of proto-Polynesians early in the Holocene, which ultimately triggered the settlement of the remote Pacific in the stepwise manner suggested.

East Indonesia plays a key role in discriminating between these two hypotheses. Under the southern China/Taiwan model, proto-Polynesian peoples began passing through eastern Indonesia only about 3,500 years ago. Under the island Southeast Asian model, proto-Polynesian peoples were already living in eastern Indonesia more than 10,000 years before their eastward expansion into the remote Pacific. Therefore these models can be distinguished, at least in theory, by examining genetic lineages with Polynesian affinity in eastern Indonesian populations. A time to the most recent common ancestor (TMRCA) of these lineages that falls within the Holocene would support the southern China/Taiwan model; a TMRCA that excludes the Holocene would support the island Southeast Asian model.

Richards et al. (1998) examined mitochondrial DNA (mtDNA) diversity in a series of samples from east Indonesia. They identified six individuals who carried a suite of hypervariable segment I (HVS-I) polymorphisms that are characteristic of mtDNA lineages in Polynesia (positions 16189, 16217, 16247, and 16261 relative to the Cambridge Reference Sequence) (Andrews et al. 1999). This haplotype reaches such high frequencies in Polynesia that it has been dubbed the Polynesian motif, despite its sporadic occurrence as far west as Indonesia. Richards and his colleagues inferred the TMRCA for these lineages with the ρ statistic—a genetic diversity value that is believed to correlate closely with the time since the divergence of a set of related sequences from their common ancestor (Saillard et al. 2000). Their calculation suggested that the east Indonesian Polynesian motif sequences shared a common ancestor no more recently than the Pleistocene (17,000 years B.P.; 95% credible region 5,500–34,500 years B.P.). This interpretation conflicts with the southern China/Taiwan model, which proposes that proto-Polynesian peoples first reached east Indonesia only during the mid to late Holocene (about 3,500 years B.P.). Because the TMRCA date inferred for the east Indonesian sequences had a 95% confidence interval with a lower bound of 5,500 years B.P., Richards et al. (1998) rejected the southern
China/Taiwan model at a significance level of 5%. The Pleistocene date was interpreted as strong support for an origin of proto-Polynesian peoples within island Southeast Asia. This result produced considerable disquiet in the anthropological community, because the island Southeast Asian model conflicts substantially with linguistic (Bayard 1996) and archeological evidence (Spriggs 2000).

Subsequent writings by proponents of the island Southeast Asia model have presented the Pleistocene TMRCA date for east Indonesian Polynesian motif sequences as a linchpin argument in their model (Oppenheimer and Richards 2001). However, there remains considerable doubt surrounding the statistical robustness of a TMRCA value estimated from only six HVS-I sequences (Hurles et al. 2003). Here, I adduce a new series of Polynesian motif sequences from Indonesia and recalculate the TMRCA date for these lineages using the methods of Richards et al. (1998). I show that the TMRCA date previously inferred changes markedly with the addition of only a few new sequences, and consequently, previous estimates are seen to lack reproducibility and robustness. The recalculated TMRCA values presented here remain compatible with either model and do not exclude the possibility of a proto-Polynesian origin in southern China/Taiwan.

Materials and Methods

This study replicates an analysis described previously by Richards et al. (1998), in which Polynesian motif sequences with transitions at positions 16189, 16217, 16247, and 16261 (Andrews et al. 1999) were assembled from eastern Indonesia, coastal Papua New Guinea, Samoa, and the Cook Islands. Sequences from Taiwan representing the immediate ancestor of the Polynesian motif (lacking only the transition at np 16247) were also collected. Data sets from the original study have been augmented here by sequences published after 1998 (Ingman et al. 2000; Lum and Cann 2000; Ingman and Gyllensten 2003; Tajima et al. 2003). One further data set from Santa Cruz in the southern Solomon Islands was included to fill a geographic lacuna in the original study (Friedlaender et al. 2002).

Furthermore, I present new hypervariable segment I (HVS-I) sequence data from Indonesia. Cox (2003) described mtDNA diversity in individuals \( N/H11505 \) from five localities in Indonesia: southern Borneo \( N = 22 \), Java \( N = 53 \), Lombok \( N = 24 \), northern Sulawesi \( N = 22 \), and Ambon \( N = 23 \). Four individuals from Lombok \( N = 2 \), northern Sulawesi \( N = 1 \), and Ambon \( N = 1 \) were identified as carriers of the Polynesian motif. Their sequences have been deposited in the GenBank Sequence Database (available at http://www.ncbi.nlm.nih.gov/Entrez/; accession numbers AY850732–AY850733).

Phylogenetic relationships between the Polynesian motif sequences were inferred using the computer application Network (available at http://www.fluxus-engineering.com/sharenet.htm), which generates a graphical representation of the
phylogenetic relationships and genetic diversity in each population sample. In this type of representation, circles depict sequence haplotypes and are proportional to haplotype frequency; links represent mutations (less than 16,000) from the hub Polynesian motif sequence.

Using the same methods and assumptions described by Richards et al. (1998), I calculated the \( \rho \) statistic and its variance for each population data set. For samples sequenced over different extents of HVS-I, I used a weighted mutation rate:

\[
\mu = \frac{N_1 \mu_1 + N_2 \mu_2}{N_1 + N_2},
\]

where \( N_1 \) and \( N_2 \) are the numbers of samples sequenced over the two ranges and \( \mu_1 \) and \( \mu_2 \) are the rates appropriate to those ranges, as Richards et al. (1998) defined them. Credible regions encompassing a central 95% confidence interval for the \( \rho \) statistic were calculated using the normal approximation. From these values the time elapsed since the most recent common ancestor (TMRCA) of the sequences could be inferred (Saillard et al. 2000). HVS-I sequences were assigned the mutation rates used by Richards et al. (1998): 1 mutation in 20,180 years for positions 16024–16380, and 1 in 26,600 years for positions 16189–16380. Credible regions (\( \alpha = 0.05 \)) were calculated as described earlier. As in all date estimates based on the \( \rho \) statistic, the quoted credible regions do not account for uncertainties in the mutation rate.

**Results and Discussion**

Phylogeographic analysis of mtDNA variation has commonly been used to clarify the prehistory of Polynesian peoples. Hertzberg et al. (1989) identified an mtDNA lineage in Polynesians that is characterized by a COII/tRNALys intergenic 9-bp deletion. This lineage, which approaches fixation in Polynesia, is often called the Polynesian motif. Melton et al. (1995) and Redd et al. (1995) reconstructed the evolutionary history of this motif lineage by determining the order in which four diagnostic HVS-I transitions occurred (positions 16189, 16217, 16247, and 16261). They concluded that the Polynesian motif probably assumed its final form in central or eastern Indonesia, thereby implying that it was somewhat erroneously named. The motif’s immediate ancestor (lacking only the transition at np 16247) is dispersed throughout Asia as far as India (Melton et al. 1995) and Mongolia (Kolman et al. 1996). Because this ancestor motif also occurs in Taiwan (Sykes et al. 1995) and southern China (Kivisild et al. 2002), it is generally presented as evidence for an ultimate mainland Asian origin of modern Polynesians.

Because successive studies have failed to detect the Polynesian motif in western Indonesia, the Philippines, Taiwan, or mainland Asia (Sykes et al. 1995;
Handoko et al. 2001; Kivisild et al. 2002; Tajima et al. 2003), the Polynesian motif evidently evolved within central or eastern Indonesia. The main point of contention is when. The island Southeast Asian model presumes that proto-Polynesian peoples were living in eastern Indonesia during the Pleistocene; the southern China/Taiwan model presumes that proto-Polynesian peoples only began passing through eastern Indonesia during the mid Holocene. Therefore the age of Polynesian motif lineages in eastern Indonesia immediately presents one way to discriminate between these two hypotheses.

Richards et al. (1998) assembled Polynesian motif sequences from across the Indo-Pacific region to address this question. Using a genetic dating method based on the $\rho$ statistic, Richards and his colleagues inferred the time elapsed since the most recent common ancestor of the Polynesian motif sequences. They also measured the genetic diversity found within the Polynesian motif’s ancestor haplotype in Taiwanese aboriginals and inferred that this sample of sequences diverged from a common ancestor about 30,000 years B.P. (95% credible region, 17,500–47,000 years B.P.). This date fits well with the ancestor motif’s widespread distribution across mainland Asia.

Since this study was performed, 30 new Polynesian motif sequences carried by Taiwanese aboriginals have been published (Melton et al. 1998; Tajima et al. 2003). These sequences were combined with those used during the 1998 study, and new estimates of the $\rho$ statistic were inferred along with its associated TMRCA (Table 1). The recalculated TMRCA estimate is similar to that estimated previously (about 26,500 years B.P.; 95% credible region, 0–52,500 years B.P.). Although it encompasses a conspicuously larger confidence interval, this TMRCA date remains suggestive of a Pleistocene origin for the ancestor motif sequence.

The Polynesian motif itself arose from its ancestor haplotype by a single mutation event at position 16247. For each of the remaining populations in the original study, new data sets were produced that included all the Polynesian motif sequences published for these populations since 1998. The $\rho$ statistic was recalculated for each data set, and new TMRCA were values inferred (Table 1). Although few sequences are available from each sampling location, the TMRCA values inferred for Polynesian motif sequences in coastal New Guinea (about 5,000 years B.P.), the southern Solomon Islands (about 3,000 years B.P.), Samoa (about 5,000 years B.P.), and the Cook Islands (about 1,000 years B.P.) are broadly consistent with the estimates of the 1998 study. Furthermore, graphical representations of the phylogenetic relationships between these Polynesian motif sequences are starlike (Figure 1), a pattern that is highly suggestive of a recent population expansion (Slatkin and Hudson 1991). This pattern is generally considered to represent the genetic signal of a rapid proto-Polynesian dispersal. Consequently, the inferred dates and phylogenetic patterns can be viewed as corresponding well with a mid to late Holocene spread of proto-Polynesian peoples into the Pacific Ocean, as predicted by both the island Southeast Asia and southern China/Taiwan models of the Polynesian dispersal.
Table 1. Divergence Time Estimates for the Polynesian Motif in Eastern Indonesia, Coastal Papua New Guinea, the Southern Solomon Islands, Samoa, and the Cook Islands and Its Ancestor Haplo-type in Taiwan

<table>
<thead>
<tr>
<th>Ancestral Sequence</th>
<th>Sampling Location</th>
<th>N</th>
<th>ρ</th>
<th>Mean Divergence Time t (Years)</th>
<th>Central 95% Credible Region (Years)</th>
<th>Richards et al. (1998) Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>16189-16217-16261</td>
<td>Taiwan</td>
<td>44</td>
<td>1.23</td>
<td>26,500</td>
<td>0–52,500</td>
<td>30,500</td>
</tr>
<tr>
<td>16189-16217-16261</td>
<td>Eastern Indonesia</td>
<td>10</td>
<td>0.60</td>
<td>12,000</td>
<td>1,000–23,500</td>
<td>17,000</td>
</tr>
<tr>
<td>16189-16217-16247</td>
<td>Coastal Papua New Guinea</td>
<td>22</td>
<td>0.23</td>
<td>5,000</td>
<td>500–9,000</td>
<td>5,000</td>
</tr>
<tr>
<td>16189-16217-16247</td>
<td>Southern Solomon Islands</td>
<td>22</td>
<td>0.14</td>
<td>3,000</td>
<td>0–7,000</td>
<td>1,500–10,000</td>
</tr>
<tr>
<td>16189-16217-16247</td>
<td>Samoa</td>
<td>56</td>
<td>0.21</td>
<td>5,000</td>
<td>2,000–8,500</td>
<td>1,000–6,000</td>
</tr>
<tr>
<td>16189-16217-16247</td>
<td>Cook Islands</td>
<td>49</td>
<td>0.04</td>
<td>1,000</td>
<td>0–2,000</td>
<td>0–3,000</td>
</tr>
</tbody>
</table>

a. Rounded to the nearest 500 years.
b. Data are from (a) Lum et al. (1994); (b) Redd et al. (1995); (c) Sykes et al. (1995); (d) Lum et al. (1998); (e) Melton et al. (1998); (f) Ingman et al. (2000); (g) Lum and Cann (2000); (h) Friedlaender et al. (2002); (i) Cox (2003); (j) Ingman and Gyllensten (2003); and (k) Tajima et al. (2003). HV-S1 sequences were assigned the mutation rates used by Richards et al. (1998). Sequences from publications b, c, f, h, i, j, and k were used with a mutation rate of 1 in 20,180 years for positions 16024–16380; sequences from publications a, c, d, and g were used with a mutation rate of 1 in 26,600 years for positions 16189–16380.
mtDNA Opposes Polynesian Origin in Indonesia

Figure 1. Phylogenetic network of mtDNA sequence haplotypes containing the Polynesian motif in (a) eastern Indonesia, (b) coastal Papua New Guinea, (c) the southern Solomon Islands, and (d) Samoa. Values beneath the network diagrams show the number of Polynesian motif sequences available from each population. Data encompass the mtDNA control region’s first hypervariable segment (circa nucleotide positions 16024–16380) from publications listed in Table 1. Circles represent sequence haplotypes and are proportional to haplotype frequency. Links represent mutations (less 16,000) from the hub Polynesian motif sequence, which deviates from the Cambridge Reference Sequence (Andrews et al. 1999) by transitions at positions 16189, 16217, 16247, and 16261.

However, the addition of newly identified Polynesian motif carriers from eastern Indonesia leads to a reduction in the corresponding TMRCA estimate, and the recalculated value suggests that the most recent common ancestor of the Indonesian Polynesian motif carriers lived more recently than inferred by Richards et al. (1998). A recent study of mtDNA diversity among Indonesians (N = 144) identified four individuals who carry Polynesian motif sequences (Cox 2003). The Polynesian motif itself was detected in three individuals from Lombok (N = 2) and northern Sulawesi (N = 1), and a Polynesian motif haplotype with one further transition at position 16274 was found in a single individual from Ambon. To date, Lombok represents the most westerly location in island Southeast Asia identified with Polynesian motif carriers. These new sequences were combined with those from the 1998 study, and the ρ statistic and TMRCA values were recalculated. The mean of the new TMRCA date still falls within the Pleistocene (about 12,000 years b.p.) and therefore superficially favors the island Southeast Asian model for proto-Polynesian origins. However, the 95% credible region for the date has widened considerably (1,000–23,500 years b.p.) and now encompasses most of the Holocene era as well. Therefore these mtDNA data can no longer discriminate with any statistical confidence between a Pleistocene and a Holocene era date of origin for the Polynesian motif carriers in eastern Indonesia.

As future studies uncover additional Polynesian motif sequences in eastern Indonesia, the ρ statistic and TMRCA values inferred here may well be refined further. Nevertheless, it now seems clear that one cannot confidently exclude either the island Southeast Asia or southern China/Taiwan models of proto-Polynesian origins on this evidence alone. The breadths of the 95% credible
regions themselves indicate the paucity of discrimination power available from extant sequence data from eastern Indonesia. In pragmatic terms, reproducible estimates of TMRCA dates cannot be approximated accurately from just 6 (and probably even 10) mtDNA sequences—assuming that they can be estimated accurately at all. Unfortunately, no research currently exists on the effect that sample size has on \( \rho \) value calculations, and we urgently require further study on how sampling sizes coupled with demographic factors affect dating estimates using the \( \rho \) statistic.

An island Southeast Asian source of modern Polynesian peoples has been criticized on the grounds of linguistic and archeological improbability. Indeed, the model is largely argued from genetic evidence, within which a Pleistocene origin of the Polynesian motif in eastern Indonesia plays a pivotal role (Oppenheimer and Richards 2001). Here, a new series of Polynesian motif carriers have been identified in eastern Indonesia, and a new value for the TMRCA of these sequences has been inferred. These recalculated values suggest that previous age estimates for the Polynesian motif in eastern Indonesia lack statistical rigor and reproducibility. The mtDNA evidence as it stands today can no longer be viewed as credibly supporting a Polynesian expansion that “did not originate in either Taiwan or southern China but within tropical island Southeast Asia” (Richards et al. 1998, p. 1236). Extant mtDNA data from Indonesia do not clearly favor an origin of proto-Polynesian peoples within the Indonesian archipelago. Rather, the recalculated TMRCA date for the Polynesian motif in eastern Indonesia remains consistent with this model as well as with a rapid expansion of proto-Polynesian peoples from southern China/Taiwan, through eastern Indonesia, and into the Oceanic world during the mid Holocene.

Acknowledgments

I extend my sincerest appreciation to the Indo-Pacific individuals and communities that participated in this study. I also thank John Clegg (Weatherall Institute of Molecular Medicine, University of Oxford, United Kingdom) for kindly supplying Indonesian DNA samples, and Erika Hagelberg (Zoology Department, University of Oslo, Norway) for research support. The New Zealand Foundation for Research, Science, and Technology (FRST), the Kon-Tiki Museum (Oslo, Norway), the University of Otago (New Zealand), and the University of Oslo (Norway) sponsored this research.

Received 26 July 2004; revision received 15 December 2004.

Literature Cited

**mtDNA Opposes Polynesian Origin in Indonesia**


