Corrections

ANTHROPOLOGY. For the article “Ancient mitochondrial M haplogroups identified in the Southwest Pacific,” by D. Andrew Merriwether, Jason A. Hodgson, Françoise R. Friedlaender, Robin Allaby, Salvatore Cerchio, George Koki, and Jonathan S. Friedlaender, which appeared in issue 37, September 13, 2005, of Proc. Natl. Acad. Sci. USA (102, 13034–13039; first published September 6, 2005; 10.1073/pnas.0506195102), the authors note that, in the third, fourth, and fifth lines from the bottom of Table 3, haplogroup M12 should read M42. The corrected table appears below. In addition, on page 13036, left column, line 7 of the second paragraph, “M28a” should read “M27a.” These errors do not alter the conclusions of the article.

Table 3. Coding region diversity and age estimates for M haplogroups in Near Oceania and Australia

<table>
<thead>
<tr>
<th>Haplogroup</th>
<th>n</th>
<th>(\rho)</th>
<th>(\sigma)</th>
<th>TMRCA, years SD</th>
<th>Founder age, years SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>M27a</td>
<td>2</td>
<td>1.500</td>
<td>0.866</td>
<td>7,700 4500</td>
<td></td>
</tr>
<tr>
<td>M27b</td>
<td>3</td>
<td>0.667</td>
<td>0.667</td>
<td>3,400 3400</td>
<td></td>
</tr>
<tr>
<td>M27c</td>
<td>2</td>
<td>0.500</td>
<td>0.500</td>
<td>2,600 2600</td>
<td></td>
</tr>
<tr>
<td>Total M27</td>
<td>7</td>
<td>13.430</td>
<td>2.176</td>
<td>69,000 11200 84,400 14,300</td>
<td></td>
</tr>
<tr>
<td>M28a</td>
<td>2</td>
<td>3.000</td>
<td>1.225</td>
<td>15,400 6300</td>
<td></td>
</tr>
<tr>
<td>M28b</td>
<td>2</td>
<td>8.000</td>
<td>2.000</td>
<td>41,000 10300</td>
<td></td>
</tr>
<tr>
<td>Total M28</td>
<td>4</td>
<td>7.000</td>
<td>1.458</td>
<td>36,000 7500 61,700 13,700</td>
<td></td>
</tr>
<tr>
<td>M29</td>
<td>3</td>
<td>0.667</td>
<td>0.471</td>
<td>3,400 2400 65,100 18,000</td>
<td></td>
</tr>
<tr>
<td>Q1</td>
<td>6</td>
<td>3.833</td>
<td>0.833</td>
<td>19,700 4300</td>
<td></td>
</tr>
<tr>
<td>Q2</td>
<td>3</td>
<td>4.667</td>
<td>1.247</td>
<td>24,000 6400</td>
<td></td>
</tr>
<tr>
<td>Q3</td>
<td>3</td>
<td>7.333</td>
<td>1.886</td>
<td>37,000 9700</td>
<td></td>
</tr>
<tr>
<td>Total Q</td>
<td>12</td>
<td>8.667</td>
<td>1.434</td>
<td>44,500 7400 70,233 13,700</td>
<td></td>
</tr>
<tr>
<td>Total Near Oceania</td>
<td>26</td>
<td>14.038</td>
<td>1.548</td>
<td>72,100 8,000</td>
<td></td>
</tr>
<tr>
<td>M42a</td>
<td>2</td>
<td>10.500</td>
<td>3.202</td>
<td>54,000 16500</td>
<td></td>
</tr>
<tr>
<td>M42b</td>
<td>1</td>
<td>8.000</td>
<td>2.828</td>
<td>41,000 14500</td>
<td></td>
</tr>
<tr>
<td>Total M42 (Australia)</td>
<td>3</td>
<td>9.667</td>
<td>2.333</td>
<td>49,700 12000 85,600 18,100</td>
<td></td>
</tr>
<tr>
<td>M (Combined)</td>
<td>29</td>
<td>14.310</td>
<td>1.436</td>
<td>73,500 7,400</td>
<td></td>
</tr>
</tbody>
</table>

Haplogroup E1a is omitted as it represents a recent introduction from Southeast Asia. TMRCA, 11,910 (16).

www.pnas.org/cgi/doi/10.1073/pnas.0508800102
**BIOCHEMISTRY.** For the article “Small RNA genes expressed from *Staphylococcus aureus* genomic and pathogenicity islands with specific expression among pathogenic strains,” by Christophe Pichon and Brice Felden, which appeared in issue 40, October 4, 2005, of *Proc. Natl. Acad. Sci. USA* (102, 14249–14254; first published September 23, 2005; 10.1073/pnas.050838102), the authors note that in Table 1, the last set of four sRNA genes appeared incorrectly, due to a printer’s error. The corrected table appears below.

**MEDICAL SCIENCES.** For the article “Roles of PSF protein and VL30 RNA in reversible gene regulation,” by Xu Song, Ying Sun, and Alan Garen, which appeared in issue 34, August 23, 2005, of *Proc. Natl. Acad. Sci. USA* (102, 12189–12193; first published August 3, 2005; 10.1073/pnas.0505179102) and was discussed in “Normal and pathological functions of mammalian retroelements,” a commentary by Albert Deisseroth in issue 35, August 30, 2005 (102, 12292–12293; first published August 23, 2005; 10.1073/pnas.0505866102), the authors note that the human genome contains a sequence on chromosome 11 (clone RP11–419K3) almost identical to a mouse retroelement VL30. This sequence was registered in GenBank as AC019351 on September 9, 2000, but was removed from the database after publication of this article because it is now considered to be a contaminating mouse sequence. The authors also referenced a 742-bp EST sequence (CX757918) that has extensive sequence identity to a region of a mouse VL30 RNA. The EST was cloned from a pleuripotent cell line derived from a human blastocyst inner cell mass, but the apparent lack of a coding sequence in the human genome suggests that the EST could be a mouse contaminant. The only conclusion in the papers that has changed concerns the presence of a VL30 gene in the human genome. All other conclusions remain valid.

### Table 1. The sRNAs expressed in *S. aureus*

<table>
<thead>
<tr>
<th>sRNA gene</th>
<th>Strand*</th>
<th>IGR location†</th>
<th>Lengths‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5S RNA</td>
<td>&gt;</td>
<td>501358/502001</td>
<td>~270</td>
</tr>
<tr>
<td>tmRNA</td>
<td>&gt;</td>
<td>843706/844543</td>
<td>~360</td>
</tr>
<tr>
<td>RNAse P</td>
<td>&lt;</td>
<td>1483578/1484012</td>
<td>~380</td>
</tr>
<tr>
<td>RNAIII</td>
<td>&lt;</td>
<td>2078477/2079445</td>
<td>514</td>
</tr>
<tr>
<td>65 RNA</td>
<td>&lt;</td>
<td>1660650/1660796</td>
<td>~230 (180)</td>
</tr>
</tbody>
</table>

Pathogenicity islands§ SaPln3 (vSa)

<table>
<thead>
<tr>
<th></th>
<th>Strand*</th>
<th>IGR location†</th>
<th>Lengths‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>sprA</td>
<td>&gt;</td>
<td>1856223/1856978</td>
<td>~210</td>
</tr>
<tr>
<td>sprB</td>
<td>&lt;</td>
<td>1866661/1867134</td>
<td>~110</td>
</tr>
<tr>
<td>sprC</td>
<td>&lt;</td>
<td>1871167/1872531</td>
<td>~170</td>
</tr>
</tbody>
</table>

Pathogenicity islands§ N315 island (dSa3)

<table>
<thead>
<tr>
<th></th>
<th>Strand*</th>
<th>IGR location†</th>
<th>Lengths‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>sprD</td>
<td>&lt;</td>
<td>2006878/2007561</td>
<td>~150</td>
</tr>
<tr>
<td>sprE</td>
<td>&gt;</td>
<td>2008571/2009086</td>
<td>~280</td>
</tr>
<tr>
<td>sprF</td>
<td>&gt;</td>
<td>2010789/2011001</td>
<td>~140, 180</td>
</tr>
<tr>
<td>sprG</td>
<td>&lt;</td>
<td>2010789/2011001</td>
<td>~90, 300</td>
</tr>
</tbody>
</table>

*The orientation of the adjacent genes, from the replication origin clockwise.
†From the *S. aureus* N315 European Molecular Biology Laboratory genome database revision 67.
‡Estimations of the nucleotide lengths by Northern blots; the exact size of RNAIII is known (17) and its expression is only detected in strains 502A and 252. 65 RNA has two forms, with lower amounts of the shorter RNA (in parentheses). SprF and sprG have multiple RNA forms.
§The location of the additional copies is available as supporting information.

www.pnas.org/cgi/doi/10.1073/pnas.0508805102
Ancient mitochondrial M haplogroups identified in the Southwest Pacific

D. Andrew Merriwether1,†, Jason A. Hodgson1, Françoise R. Friedlaender3, Robin Allaby1, Salvatore Cerchio1, George Koki1, and Jonathan S. Friedlaender3

1Department of Anthropology, Binghamton University, Binghamton, NY 13902-6000; 2Department of Anthropology, Temple University, Philadelphia, PA 19122; and 3Papua New Guinea Institute of Medical Research, Goroka, Papua New Guinea

Communicated by R. C. Green, University of Auckland, Auckland, New Zealand, July 21, 2005 (received for review June 6, 2005)

Based on whole mtDNA sequencing of 14 samples from Northern Island Melanesia, we characterize three formerly unresolved branches of macrohaplogroup M that we call haplogroups M27, M28, and M29. Our 1,399 mtDNA control region sequences and a literature search indicate these haplogroups have extremely limited geographical distributions. Their coding region variation suggests diversification times older than the estimated date for the initial settlement of Northern Island Melanesia. This finding indicates that they were among the earliest mtDNA variants to appear in these islands or in the ancient continent of Sahul. These haplogroups from Northern Island Melanesia extend the existing schema for macrohaplogroup M, with many independent branches distributed across Asia, East Africa, Australia, and Near Oceania.

Portions of Northern Island Melanesia were settled by at least 42,000 years before present (YBP) (1, 2), apparently not long after New Guinea, which was joined at that time to Australia as the ancient Pleistocene continent of Sahul. Non-14C dates from after New Guinea, which was joined at that time to Australia as there was a detectable and repeated trickle of New Britain inville from New Ireland by 29,000 YBP, and after 20,000 YBP had successfully made the longer windward crossing to Bougainville. Short voyages between islands are inferred (2, 6) because people were not “strandlopers” there were at the eastern edge of the human species range until 130,000 YBP did people from Island Southeast Asia, with northern populations in Near Oceania, as shown in Fig. 1, and the populations there were at the eastern edge of the human species range until 3,200 YBP.

The earliest populations of Northern Island Melanesia were small groups of hunter-gatherers. They were not “strandlopers” restricted to the coasts and lagoons, but settled the interiors of the large islands intermittently (2, 5). The intentional introductions of plants and animals from New Guinea over the following millennia indicate continuing outside contacts at a modest level. Short voyages between islands are inferred (2, 6) because people had successfully made the longer windward crossing to Bougainville from New Ireland by 29,000 YBP, and after 20,000 YBP there was a detectable and repeated trickle of New Britain obisidan to New Ireland and Nissan up to ~7,000 YBP. The implication is that isolation of these small island populations was an incomplete but persistent condition across the region for tens of thousands of years during the Pleistocene. By extension, movements between Near Oceania and Island Southeast Asia also would have been intermittent and small in scale.

Only ~3,200 YBP did people from Island Southeast Asia, with advanced sailing and agricultural skills, make a major impact on this region (7, 8). This led, in turn, to the rapid colonization of the formerly uninhabited islands of Remote Oceania as far as Western Polynesia by the “Lapita People” and also to a following period with variable contacts among those populations and coastal groups back in Near Oceania. As a result of this history, Near Oceanic populations are linguistically and biologically extraordinarily diverse (9–11).

mtDNA continues to be a particularly powerful source for reconstructions of early human demographics because of the effective absence of recombination and its comparatively high mutation rate. In the Southwest Pacific, the haplogroups P and Q have been recognized as old and specific to the general region (12–15). They have long branches, and analyses of their control region variation indicate population expansions in the Pleistocene (12, 13). P is more widespread and heterogeneous than Q, and may therefore be the older of the two. With one clear exception, different branches of P occur either in Australia or New Guinea, but not both. Q is absent in Australia but very common in New Guinea and Island Melanesia (12). This pattern suggests substantial isolation of Australia from New Guinea/Island Melanesia since around the time of first settlement. Haplogroup B4a, which is very common in Remote Oceania and coastal Near Oceania but absent in aboriginal Australia, the New Guinea highlands, and interior sections of Bougainville and New Britain, was introduced to Near Oceania in the terminal Pleistocene (16–18).

Besides these common mtDNA haplogroups, a number of other mtDNA variants in Island Melanesia have not been possible to characterize without full mtDNA genome sequencing (17). We have now sequenced 14 whole mtDNA genomes of these and can define three additional branches of macrohaplogroup M. They have not been found anywhere to the west of Island Melanesia, including New Guinea. These haplogroups (M27, M28, and M29) are not closely related to each other, although M29 might be distantly related to Q.

Materials and Methods

The samples analyzed were selected from our Southwest Pacific collection. Its core consisted of blood samples collected in three recent field seasons in the Bismarck Archipelago. This primary set was augmented with plasmas and urines from older collections, described elsewhere (12). Information on survey subjects included their language, a short genealogy, current residence, and birthplace (used to assign location), although such details were not available from some of the other collections. One sample from each identified matriline was included in the initial mtDNA control region analysis. The primary samples were collected, and all selected samples were analyzed, with informed consent protocols approved by the appropriate Human Subjects Ethical Committees of Papua New Guinea, the University of Michigan, Binghamton University, and Temple University.

The mtDNA analysis of the selected samples occurred in three phases: (i) sequencing of hypervariable segments 1 and 2 (HVS1 and HVS2); (ii) for those samples not definitely assigned to a known haplogroup on this basis, restriction fragment-length polymorphism (RFLP) screening for the two mutations defining macrohaplogroup M (DdeI 10394, AluI 10397), and depending on the presence or absence of these, additional RFLPs known to

Abbreviation: YBP, years before present.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. DQ137398–DQ137411).

†To whom correspondence should be addressed. E-mail: andym@binghamton.edu.

© 2005 by The National Academy of Sciences of the USA
identify other haplogroups in the Southwest Pacific (19–21); and (iii) sequencing of the coding region on 14 representative samples from each of the major haplogroups that could not be assigned to currently published M sublineages.

DNA was extracted from 100–200 μl of buffy coat, plasma, or urine (depending on the source of the sample) by using either the guanidine-silica based IsoQuick extraction kit (Orca Scientific, Bothell, WA) or the column-based Qiagen extraction kit (Qiagen, Valencia, CA). In preparation for sequencing, DNA was PCR amplified following standard protocols, employing Platinum Taq Polymerase (Invitrogen). The control region was amplified by using primers spanning nucleotide positions 15938–200429. The coding region was amplified by using the PCR primers and conditions of Rieder et al. (22). Successful amplification was verified by electrophoresis on 1% ethidium bromide-stained agarose gels. Samples were prepared for sequencing by an ExoI digest followed by filtration through a Millipore 96-well filter plate (Millipore, Billerica MA) to remove single-stranded DNA and unincorporated nucleotides. PCR product was sequenced by using various versions of the BigDye Terminator Sequencing kits from Applied Biosystems on an Applied Biosystems 377XL automated sequencer using described conditions (23). Custom designed internal sequencing primers were used for all large PCR fragments to increase double-fold coverage. Contig assemblage and sequence alignment was accomplished with SEQUENCE (Forensic Version, GeneCodes, Ann Arbor, MI). The phylogenetic tree was inferred from median-joining networks rooted to L3. The tree was hand-checked to resolve several homoplasies. A few ambiguities remained, and we tended to be conservative in interpreting those cases.

Results

The control region mutations that initially identified the M variants are listed in Table 1 with key mutations in bold. The mtDNA haplogroup incidences in our 1,399 control region results are presented in Table 2 by island and the frequencies of the three haplogroups are indicated in Fig. 1. Besides the commonly occurring haplogroups P, Q, and B described else-

Table 1. Defining control region mutations for M27, M28, and M29

<table>
<thead>
<tr>
<th>Haplogroup</th>
<th>HVS 1 (add 16000)</th>
<th>HVS 2</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>M27a</td>
<td>223 048 077 172 311 320 136 189</td>
<td>195 234 150 204 228</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>223 048 077 172 311 320</td>
<td>195 234</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>223 048 077 172 311 320 264</td>
<td>195 234</td>
<td>9</td>
</tr>
<tr>
<td>M27b</td>
<td>209 299 390 145</td>
<td>152 64 199 236</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>209 299 390 086 223</td>
<td>152</td>
<td>13</td>
</tr>
<tr>
<td>M27c</td>
<td>223 301 304</td>
<td>146 186</td>
<td>8</td>
</tr>
<tr>
<td>M28a</td>
<td>223 148 468 362 086 129 320</td>
<td>152 195</td>
<td>47</td>
</tr>
<tr>
<td>2</td>
<td>223 148 468 362 086 129 429</td>
<td>152 195</td>
<td>104</td>
</tr>
<tr>
<td>3</td>
<td>223 148 468 362 086 129 189 209</td>
<td>152 195 279</td>
<td>27</td>
</tr>
<tr>
<td>4</td>
<td>224 148 468 362 086 129 51</td>
<td>152 195 198</td>
<td>10</td>
</tr>
<tr>
<td>M29b</td>
<td>223 148 468 362 318C</td>
<td>152 94</td>
<td>59</td>
</tr>
<tr>
<td>2</td>
<td>223 148 468 362 318T</td>
<td>152 94</td>
<td>14</td>
</tr>
<tr>
<td>M29a</td>
<td>223 189 311</td>
<td>211 310</td>
<td>24</td>
</tr>
<tr>
<td>M29b</td>
<td>223 189 294</td>
<td>211 200</td>
<td>14</td>
</tr>
</tbody>
</table>

Numbers are transitions from the Cambridge Reference Sequence, except where suffix letters denote transversions. Bold indicates a haplogroup defining mutation.
where (12), the newly defined branches of M were found in
certain circumscribed regions in intermediate to high frequen-
cies. The phylogenetic M tree for the old Near Oceanic lineages
is presented in Fig. 2 with Australian Aborigine M sequences
included for comparison.

Haplogroup M27, as shown in Table 1 and Fig. 2, had three
branches with no key shared control region mutations, but they
were ultimately linked in the coding region sequencing by the
shared transitions at nucleotide positions 5375, 9201, and 12358.
We found no convincing associations of M27 with other M
branches, although 5177 was also found in Q3 and in M21b from
the aboriginal Semang of Malaysia (24). Position 5585 in M28a
has been found in Indian M6b, but this also appears to be a
recurrence. M27 was most frequent in central Bougainville,
especially branch M27a, with the other M27 branches detected
in Bougainville and sporadically in New Britain, New Ireland, the
Solomon Islands, Santa Cruz (17), and Vanuatu (25).

Haplogroup M28 was relatively common. It had 71 constituent
haplotypes that were identified in 261 individuals in our series.
Its defining control region variants were the transitions at
nucleotide positions 16468 and 16148, and each of its major
subbranches had other control region distinctions, as shown in
Table 1 and Fig. 2. Coding region sequencing on a representa-
tive subset reinforced these relationships. Within M, haplogroup
M28 shared the transition at nucleotide position 16362 with
haplogroups D, G, and M9/E, so we interpreted this as a
recurrence.

M28 had its greatest prevalence and diversity in the interior of
East New Britain, where it was associated with an ancient
population expansion (17). In some sample sets, it had a
frequency of 100%. M28 was rare in our New Ireland sample set
and almost absent in Bougainville, a distribution contrasting with
M27. We found M28 at variable frequencies elsewhere in Island
Melanesia (i.e., Santa Cruz, Vanuatu, New Caledonia, and Fiji),
reported in studies that relied on short sequencing within HVS1
with different terminologies (refs. 25–27, and possibly ref. 28).
Because the most characteristic single mutation of M28 (i.e.,
16468) lies outside the region of HVS1 most commonly se-
quenced in early mtDNA sequence analyses, it is possible its
distribution is still underestimated. One transversion in a branch
of M28b is 16318T, which is a defining mutation for Indian M18.
However, we are assuming it is recurrent as there are no other
shared mutations.

Haplogroup M29 was initially distinguished by the transition
at nucleotide position 211, subsequently complemented by a long
series of coding region transitions (an average of 13). It was
relatively rare in our series, but, as with M28, its distribution
beyond our series could have been underestimated because of
the general early reliance on short HVS1 sequencing. In our
series, haplogroup M29 was found most frequently in East New
Britain and especially among the Tolai (who originally resided in
southern New Ireland). It was also identified in single samples
from the Solomon Islands and Vanuatu. The M29 sequences
share the 13500 transition with Q, which could indicate a remote
connection. However, it may also be a recurrence, because it has
also been found in a variety of other haplogroups (D, M7a, R2,
F, U7, and V). Additional M29 sequencing should help resolve
this issue.

Haplogroup E, an Asian M variant with a more recent
coevolution (16), was also found in our series, most commonly
in West New Britain (Table 1).

Discussion

Northern Island Melanesia is clearly a relict area, retaining a
remarkable number of ancient population genetic signatures.
Four old lineages of macrohaplogroup M have now been iden-
tified in Near Oceania (Fig. 2). Besides haplogroup O, which
had been shown to have two clearly “star-like” subdivisions (12),
the three M haplogroups described here are centered in different
locations within Northern Island Melanesia: M27 is most com-
mon and diverse in Bougainville, M28 in the interior of east New
Britain, and M29 in southern New Ireland and east New Britain.
The remarkably limited and different distributions of each of
these M branches are the likely result of restricted marital
migration rates (refs. 29 and 30, pp. 71–75) within the region
from the times of their appearances, possibly coupled with their
separate introductions.

Fig. 2 includes the Australian M42 haplogroup (T. Kivisild,
personal communication). The Australian and Near Oceanic M
branches share no special relationships (1598 is clearly recur-
rent) and do not overlap in their distributions. This finding only
reinforces the remarkable distinctiveness of the different South-
west Pacific populations.

The old Northern Island Melanesian M haplogroups have
distributions that tie their original associations to different
Papuan speaking populations in the region (17). However, links
to specific languages or language families have decayed some-

what over time because of language shift and intermarriage so that correlations of gene and language distributions will inevitably be weak.

On present evidence, the spread of M27, M28, and M29 beyond their Northern Island Melanesian centers was limited to the adjacent island chains in Southern Island Melanesia (also reported in refs. 17, 25, and 26). A selection of particular P and Q haplotypes are also scattered across this region from New Guinea to Fiji, specifically Q2*, Q2b, P1, P1e, Plf, and P2 (see tables s2 and s3 of ref. 12). Because none of these are restricted to Southern Island Melanesia, they indicate population interactions right across the region. They must have all spread to Southern Island Melanesia after its colonization by the Lapita People, who would already have been predominantly haplogroup B4a, like all their direct descendants across the rest of Remote Oceania. A more complex scenario would have all of these haplotypes carried with Lapita groups as far as Fiji, and then all have lost in separate founder events in every subsequent voyage into Remote Oceania, a most unlikely alternative. Also, archaeological evidence does indicate considerable influence from Near Oceania into Southern Island Melanesia after the Lapita period, compatible with the preferred first scenario (7, 31–35).

Table 3 gives lineage diversity (ρ and σ) and age estimates for old M branches from the Southwest Pacific employing the commonly used technique of Saillard et al. (36), in conjunction with the estimated mutation calibration rate of Mishmar et al. (37), because these have been recently applied, with an adjustment, to a number of mtDNA sequences from other relict populations in Southeast Asia, including Andaman Islanders and aboriginal Malays (24, 38). The overall (average) age of M had been estimated with the Saillard and Mishmar approach at 64,800 years with a standard error of 7,100 years, taken to conform to proposed dates for a “Southern Route” expansion between 55,000 and 85,000 YBP (39). With our data, the combined estimate for Near Oceanic Ms becomes 72,100/8,000, with the estimate for the M27 branch, taken alone, as 84,400/14,300 years (Table 3). Our calculation for the Australian Aboriginal M42 haplogroup age estimate is in the same range, as shown. However, questions on this approach include the lack of true independence of the branch lengths because they are time-constrained (also commented on in ref. 24), and because of uncertainties on the proper rate calibration (40). A final issue is that the coalescence times of these and other molecular estimates under 1 million years may contain a significant rate artifact (41, 42), which, if verified, would make a variety of absolute date estimates considerably younger.

The length of the branches between the ancestral haplotype M and the M27, M28 and M29 branches vary considerably (see Fig.
approximated to $2N_f$. Assuming a female generation time of $4Ne$ generations (43), the comparable value for mtDNA can be estimated by fixing the ancestral M haplotype was close to fixation in the Near Oceanic lineages branched off sequentially, thereby accounting for the different numbers of mutations accumulated to the present. This scenario would imply that the ancestral M haplotype was fixed in the ancestral population, perhaps reflecting an initial expansion into unpopulated territory by a small group. Because the time of allele loss from fixation in nuclear genes can be approximated by $4N_e$ generations (43), the comparable value for mtDNA can be approximated to $2N_f$. Assuming a female generation time of 20 years, this approximation yields a long-term effective population and Australia.

### Table 3. Coding region diversity and age estimates for M haplogroups in Near Oceania and Australia

<table>
<thead>
<tr>
<th>Haplogroup</th>
<th>(n)</th>
<th>(\rho)</th>
<th>(\sigma)</th>
<th>TMRCA, years</th>
<th>SD</th>
<th>Founder age, years</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>M27a</td>
<td>2</td>
<td>1.500</td>
<td>0.866</td>
<td>7,700</td>
<td>4500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M27b</td>
<td>3</td>
<td>0.667</td>
<td>0.667</td>
<td>3,400</td>
<td>3400</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M27c</td>
<td>2</td>
<td>0.500</td>
<td>0.500</td>
<td>2,600</td>
<td>2600</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total M27</td>
<td>7</td>
<td>13.430</td>
<td>2.176</td>
<td>69,000</td>
<td>11200</td>
<td>84,400</td>
<td>14,300</td>
</tr>
<tr>
<td>M28a</td>
<td>2</td>
<td>3.000</td>
<td>1.225</td>
<td>15,400</td>
<td>6300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M28b</td>
<td>2</td>
<td>8.000</td>
<td>2.000</td>
<td>41,000</td>
<td>10300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total M28</td>
<td>4</td>
<td>7.000</td>
<td>1.458</td>
<td>36,000</td>
<td>7500</td>
<td>61,700</td>
<td>13,700</td>
</tr>
<tr>
<td>M29</td>
<td>3</td>
<td>0.667</td>
<td>0.471</td>
<td>3,400</td>
<td>2400</td>
<td>65,100</td>
<td>18,000</td>
</tr>
<tr>
<td>Q1</td>
<td>6</td>
<td>3.833</td>
<td>0.833</td>
<td>19,700</td>
<td>4300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q2</td>
<td>3</td>
<td>4.667</td>
<td>1.247</td>
<td>24,000</td>
<td>6400</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q3</td>
<td>3</td>
<td>7.333</td>
<td>1.886</td>
<td>37,000</td>
<td>9700</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Q</td>
<td>12</td>
<td>8.667</td>
<td>1.434</td>
<td>44,500</td>
<td>7400</td>
<td>70,233</td>
<td>13,700</td>
</tr>
<tr>
<td>Total Near Oceania</td>
<td>26</td>
<td>14.038</td>
<td>1.548</td>
<td>72,100</td>
<td>8,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M12a</td>
<td>2</td>
<td>10.500</td>
<td>3.202</td>
<td>54,000</td>
<td>16500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M12b</td>
<td>1</td>
<td>8.000</td>
<td>2.828</td>
<td>41,000</td>
<td>14500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total M12 (Australia)</td>
<td>3</td>
<td>9.667</td>
<td>2.333</td>
<td>49,700</td>
<td>12000</td>
<td>85,600</td>
<td>18,100</td>
</tr>
<tr>
<td>M (Combined)</td>
<td>29</td>
<td>14.310</td>
<td>1.436</td>
<td>73,500</td>
<td>7,400</td>
<td></td>
<td>4,700</td>
</tr>
</tbody>
</table>

Haplogroup E1a is omitted as it represents a recent introduction from Southeast Asia. TMRCA, 11,910 (16).

We thank Heather Norton, Dan Hrdy, Charles Mgone, and the people of Bougainville, New Ireland, New Hanover, and New Britain who participated in the study, Theodore Schurr, John McDonough, Stacy McGrath, Fred Gentz, and our other collaborators at the Papua New Guinea Institute for Medical Research. We also thank Ger Reesink, Eva Lindström, and Gisela Horvat for advice, and particularly Roger Green. Toomas Kivisild has kindly permitted us to include the M42 Australian Aboriginal sequences, and has provided a number of constructive criticisms. The research was supported by grants from the National Geographic Society Exploration Fund, the Wenner-Gren Foundation for Anthropological Research, and National Science Foundation Grants BSR-9796054 and BSR-9601020.