2005

Mitochondrial Genetic Diversity and its Determinants in Island Melanesia

Jonathan S. Friedlaender

Fred Gentz

Françoise R. Friedlaender

Frederika Kaestle

Theodore G. Schurr
University of Pennsylvania, tgschurr@sas.upenn.edu

See next page for additional authors

Mitochondrial Genetic Diversity and its Determinants in Island Melanesia

Abstract
For a long time, many physical anthropologists and human geneticists considered Island Melanesian populations to be genetically impoverished, dominated by the effects of random genetic drift because of their small sizes, internally very homogeneous, and therefore of little relevance in reconstructing past human migrations. This view is changing. Here we present the developing detailed picture of mitochondrial DNA (mtDNA) variation in eastern New Guinea and Island Melanesia that reflects linguistic distinctions within the region as well as considerable island-by-island isolation. It also appears that the patterns of variation reflect marital migration distinctions between bush and beach populations. We have identified a number of regionally specific mtDNA variants. We also question the widely accepted hypothesis that the mtDNA variant referred to as the 'Polynesian Motif' (or alternatively the 'Austronesian Motif') developed outside this region somewhere to the west. It may well have first appeared among certain non-Austronesian speaking groups in eastern New Guinea or the Bismarcks. Overall, the developing mtDNA pattern appears to be more easily reconciled with that of other genetic and biometric variables.

Keywords
mitochondria, genetic diversity, Melanesia

Author(s)
Jonathan S. Friedlaender, Fred Gentz, Françoise R. Friedlaender, Frederika Kaestle, Theodore G. Schurr, George Koki, Moses Schanfield, John McDonough, Lydia Smith, Sal Cerchio, Charles Mgone, and D. Andrew Merriwether

Papuan pasts:
cultural, linguistic and biological histories of Papuan-speaking peoples

edited by
Andrew Pawley, Robert Attenborough,
Jack Golson and Robin Hide

Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University

Published by Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University
Canberra ACT 0200
Australia

First published 2005

National Library of Australia Cataloguing-in-Publication entry:

Papuan pasts: cultural, linguistic and biological
histories of Papuan-speaking peoples

Bibliography.
ISBN 0 85883 562 2

1. Papuans - History. 2. Papuans - Social life and customs.
3. Papuan languages. I. Pawley, Andrew. II. Australian National University.
Research School of Pacific and Asian Studies. Pacific Linguistics.

995.3

Cover: Bird-shaped pestle from Wonia, Western Province, Papua New Guinea
(AM Reg E065762).
Photographer: John Fields, copyright Australian Museum, Sydney.

Copyedited and typeset by Margaret Forster
Cover design by Addcolour Digital
Printed and bound by Addcolour Digital, Fyshwick, Canberra

23 *Mitochondrial genetic diversity and its determinants in Island Melanesia*

JONATHAN FRIEDLAENDER, FRED GENTZ, FRANÇOISE FRIEDLAENDER, FREDERIKA KAESTLE, THEODORE SCHURR, GEORGE KOKI, MOSES SCHANFIELD, JOHN MCDONOUGH, LYDIA SMITH, SAL CERCHIO, CHARLES MGONE AND D. ANDREW MERRIWETHER

Introduction

For a long time, many physical anthropologists and human geneticists considered Island Melanesian populations to be genetically impoverished, dominated by the effects of random genetic drift because of their small sizes, internally very homogeneous, and therefore of little relevance in reconstructing past human migrations. This view is changing. Here we present the developing detailed picture of mitochondrial DNA (mtDNA) variation in eastern New Guinea and Island Melanesia that reflects linguistic distinctions within the region as well as considerable island-by-island isolation. It also appears that the patterns of variation reflect marital migration distinctions between bush and beach populations. We have identified a number of regionally specific mtDNA variants. We also question the widely accepted hypothesis that the mtDNA variant referred to as the 'Polynesian Motif' (or alternatively the 'Austronesian Motif') developed outside this region somewhere to the west. It may well have first appeared among certain non-Austronesian speaking groups in eastern New Guinea or the Bismarcks. Overall, the developing mtDNA pattern appears to be more easily reconciled with that of other genetic and biometric variables.

Early findings

The remarkable variation we found among Bougainville populations has continued to frame our perspective on Island Melanesian human biological variation (see especially Friedlaender 1975, but also 1971a, 1971b, 1987 and Friedlaender & Steinberg 1970; Friedlaender et al. 1971; Sokal & Friedlaender 1982; Merriwether, Friedlaender et al. 1999 and Merriwether, Kaestle et al. 1999). While people across that rather small island are uniformly black-skinned and distinctive among Papua New Guineans, there remains remarkable diversity from one section of Bougainville to the next. Geographic isolation and short-range migration within Bougainville were assumed to be the major determinants

Andrew Pawley, Robert Attenborough, Jack Golson and Robin Hide, eds, 2005, *Papuan pasts: cultural, linguistic and biological histories of Papuan-speaking peoples*, 693–716.

of any patterning of allele frequencies and other biological variables among human populations there, rather than differential natural selection. In addition, because early work in the Markham Valley of New Guinea had shown at least some genetic patterning followed major linguistic distinctions there (Giles et al. 1965), it was logical to assess this association in other islands in the region.

After identifying the smallest area with the greatest linguistic diversity within Bougainville (the central east coast), a set of adjacent villages along a path transecting those language areas was surveyed away from the coast, with the intention of recovering data from a less disturbed situation (Friedlaender 1975). Along with the biological protocol of a battery of traditional biometrics and blood group and serum protein analyses, we gathered information on contemporary village census sizes, marital migration rates over the sampled villages, and variation in completed fertility rates.

We realised these demographic statistics were probably highly variable over even a few generations and therefore limited in their power to account for patterns of existing biological variation. For example, large villages were a new creation in most areas. They had been encouraged or forced on the populace in the colonial era, and were usually aggregations of formerly dispersed hamlets of extended families. In addition, the Second World War had swept through some regions of Bougainville in the 1940s and displaced many groups. While almost all people in the survey villages who had been so affected claimed they returned to their original gardens and villages afterwards, this had not been independently confirmed.

Even with these changes, the pattern of marital migration in inland Bougainville villages was skewed to the right in this survey, with almost everyone setting up marital residences only a kilometre or two from their birthplaces (Friedlaender 1975:78). This informal 'census' did not account for those individuals who had emigrated from the region, either to towns or entirely off-island (mostly men). The localised endogamy must have characterised earlier periods as well. People were afraid to move far because of pervasive feuding, headhunting, and the fear of malevolent ancestral spirits.

We have reworked these data to distinguish migration by males and females separately. The results are given in Table 1 and Figure 1a. The data suggest very little difference between the male and female pattern of mobility in this rural region, although there is some suggestion of greater female mobility within a range of 10 kilometres. This is an important point because many contemporary geneticists have interpreted distinctions between mitochondrial DNA and Y-chromosome patterns of variation worldwide to be the likely result of differentially larger female than male marital migration rates (Seielstad et al. 1998). Rather than using genetics to hypothesise demographic determinants, we tried to examine directly those variables and their predicted effects.

Table 1: Marital migration rates in 'beach' and 'bush' villages in 1967 and 2003

Bougainville Bush Villages – 1967

Men	343	50	16	3	2	1	415
Women	324	104	20	3	0	4	455
	0	1–5	6–10	11–15	16–20	>20 km	Total

Shore Villages – 2003

Men	69	40	8	4	5	19	145
Women	36	25	4	2	3	20	90
	<1	1–9	10–19	20–29	30–39	≥40 km	Total

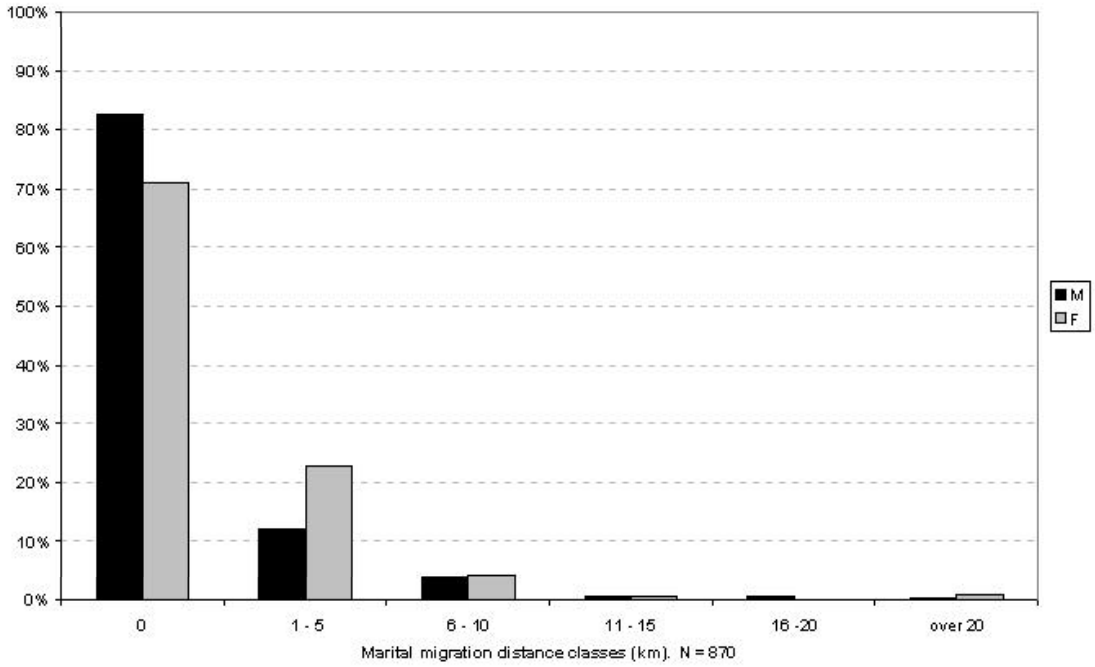

Figure 1a: Marital migration distribution—bush villages (Bougainville, 1967)

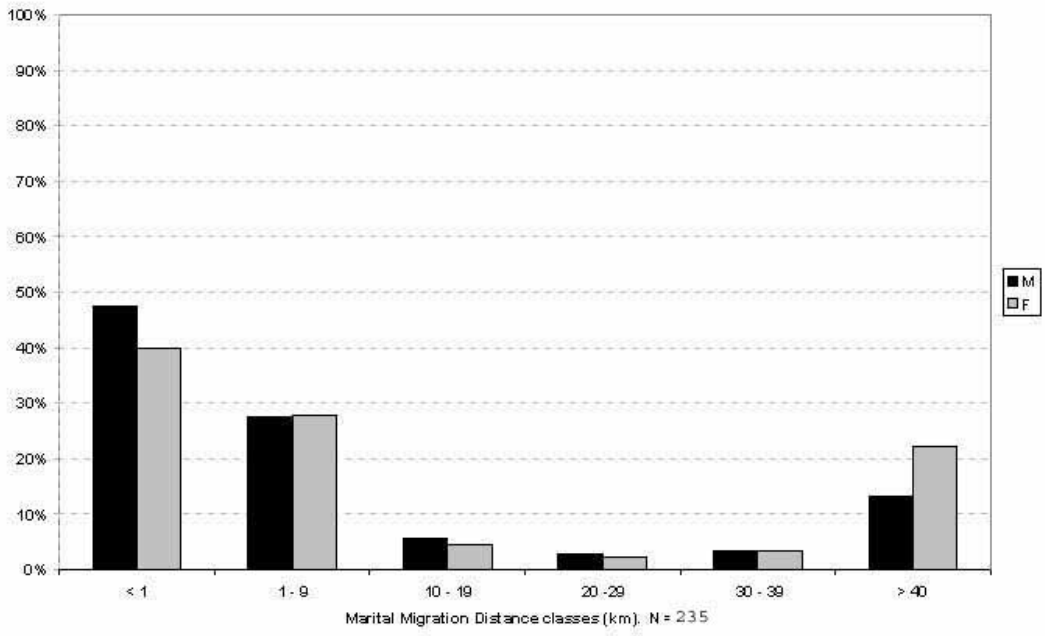

Figure 1b: Marital migration distribution—shore villages (Bougainville, 2003)

Language boundaries per se did not have significant effects in restricting marital migration in this inland region. While only ~5% of marriages in central Bougainville involved movement of one (or both) spouses across language boundaries, the barriers were actually rather porous (Friedlaender 1975:76—especially clear for certain villages). People at the boundaries were often bi- or trilingual. In villages or hamlets near language boundaries, marriages across boundaries happened only slightly less frequently than those at equivalent distances within the language area. The overwhelming tendency was to marry within a sharply circumscribed area around one's birthplace, whether or not that was near a language barrier. In other words, 'language differentiation (itself) probably was rapid and extreme *because of* the lack of movement and intercourse of the people …' although, subsequent to their formation, such barriers might have acted to restrict migration (Friedlaender 1975:72). Language boundaries drawn on maps in this region are over-simplified distinctions.

In 2003, we visited northern beach villages in north Bougainville and collected comparable marital migration data there as well as in a coastal section of West New Britain. The results were very different, with many more individuals moving further away from their birthplaces (Table 1 and Figure 1b). While this may be the result of modernisation, it could also reflect a longer-term distinction in population movements between 'beach' and 'bush' populations. If this were the case, beach groups from different regions would be expected to be genetically similar while bush groups would be expected to develop, or at least to retain, genetic distinctiveness.

Another demographic variable that geneticists have suggested could explain male/female differentials in genetic heterogeneity is their disparity in completed reproductive variance. In his study of South American indigenous groups, Neel (1970) emphasised the magnitude of the 'big man effect', with a few men leaving a disproportionate number of offspring. If only a few men contribute their genes to the next generation, in effect this reduces the effective population size of males, accentuating possibilities of drift and greater allele frequency fluctuations from generation to generation. The Bougainville data on reproductive rates of men and women in the 1960s do not show this distinction. The variance of male reproductive performance was only slightly greater than for women (approximately a half a child). However, it was clear that completed fertility rates had changed by the 1960s. Bougainville women were having very large numbers of children that were surviving to adulthood, and the overall variation in their reproductive performance was increasing (see Table 2).

The biological survey in central Bougainville did show a great deal of variation across the sample. The language affiliation of a village was sometimes a better predictor of its biological similarity to others than was either geographic proximity or current rates of marital exchange among villages. A discriminant function analysis of 12 body and head measurements among villagers showed this most clearly (Friedlaender 1975:147–157). Within this very small area, less than 100 kilometres in extent, villages from the same language group clustered together in anthropometric space. The overall arrangement also reflected the divide between the northern and southern Papuan language family areas within Bougainville.

Table 2: Male vs. female retrospective fertility in central Bougainville—1966
Mean and variance of number of children surviving infancy by current age of a) women and b) men for cohorts older than 40

	Age						
a	40–44	45–49	50–54	55–59	60–64	65+	Total
Women	41	36	29	30	17	15	168
Children	212	126	96	81	41	50	606
Cohort fertility	5.1	3.5	3.3	2.7	2.4	3.3	3.6
Reproductive variance	3.8	4.6	4.6	4.2	4.6	4.9	4.36
b							
Men	62	49	38	14	10	20	193
Children	274	226	155	55	32	58	800
Cohort fertility	4.4	4.6	4.1	3.9	3.2	2.9	4.1
Reproductive variance	4.3	4.6	5.8	5.9	5.3	3.6	4.77

Two other modes of analysis were applied to the Bougainville pattern of biological variability. The first was Malécot's 'isolation by distance' approach, which is appropriate for gene frequency variation (Friedlaender 1971a). The Malécot analysis confirmed a generally similar degree of decline of genetic similarity over very short distances within the Bougainville island sample. The second, 'spatial autocorrelation analysis', highlighted the resulting general lack of consistency of the patterns of variation across all the villages. The conclusion was that such a high degree of generally unstructured variation over distances longer than 30 kilometres was likely to be due to 'a diffusion phenomenon as in migration …,' specifically 'two or more migrations from disparate sources' (Sokal & Friedlaender 1982:221, 224). Only a few biological variables showed a north-south gradient (for example face and body breadths, premolar sizes). The most extreme, and, in retrospect, the most prescient single allele distribution across these villages involved an immunoglobulin marker, $\underline{\text{Km}}^1$ (then called Inv^1), which formed a sharp cline in frequency, from 0.83 to 0.33 (Friedlaender & Steinberg 1970).

To summarise the early studies, the exceptional internal variation on Bougainville appeared to be due to several important influences. These included small population sizes; very limited marital exchange over even small distances of 20 kilometres or less; a tradition of internal movement of lineages and small populations; some major later migration influences from external sources; and a long period of continuous settlement (Friedlaender 1975).

Bougainville mitochondrial DNA variation

More recently, we reanalysed many of the remaining Bougainville plasma samples collected in 1966 and 1985–1986 for their mtDNA variation. Our results were both consistent with, and contradictory to, earlier findings. We found that northern and southern Bougainville Papuan-speaking populations differed dramatically in their frequencies of a special variant (see Figure 2 over page). This reinforced the distinctions between north and south that we had found before. The surprise was that the variant that was so common among the south Bougainville Papuans had been associated with Polynesians and their Austronesian predecessors (Hertzberg et al. 1989; Melton et al.

1995; Redd et al. 1995; Hagelberg & Clegg 1993; Sykes et al. 1995; Soodyall et al. 1996, inter alia). This was especially perplexing to us, since the south Bougainville populations are just as black as north Bougainville groups. There was no other suggestion of overwhelming Polynesian or Austronesian intermixture.

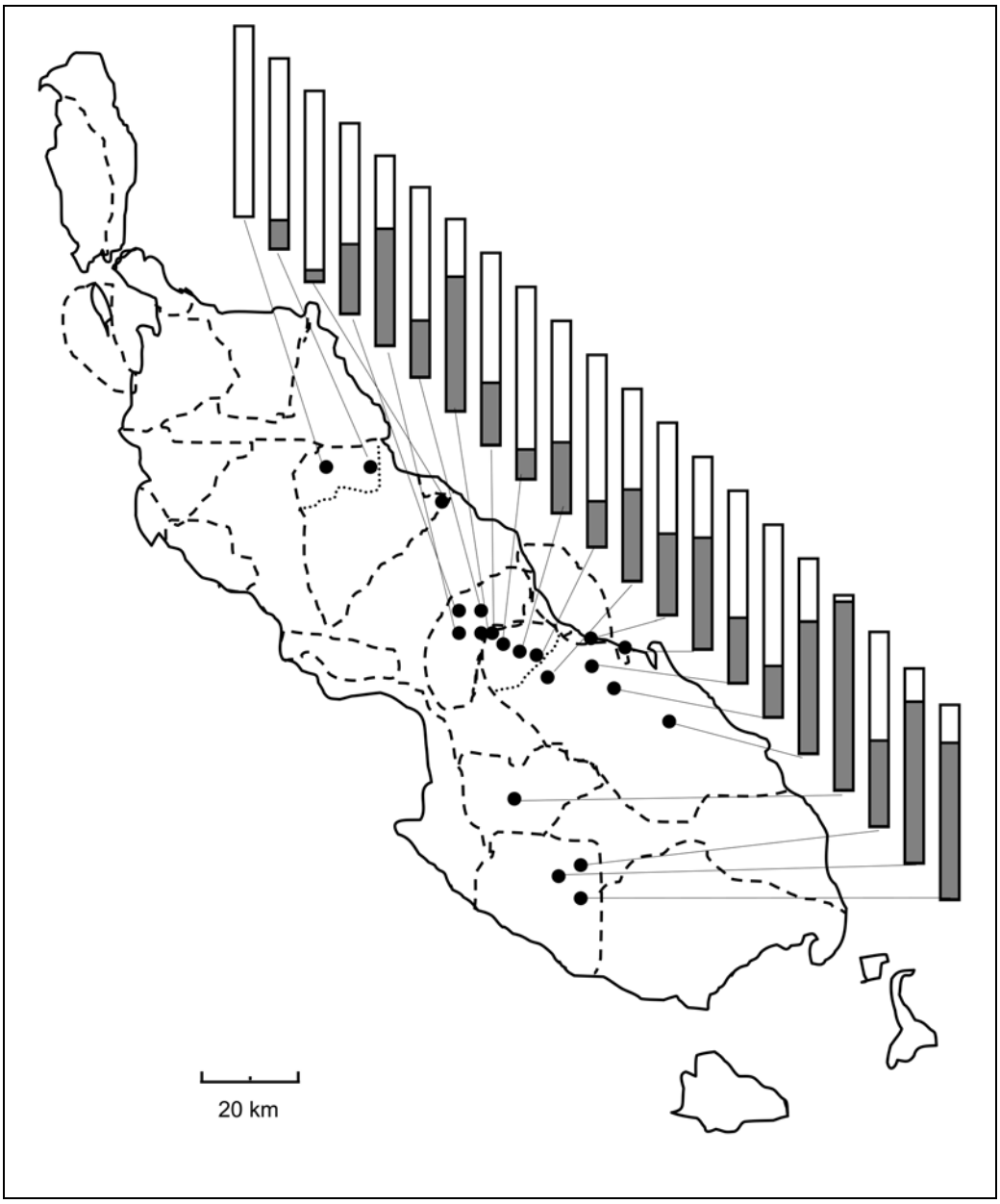

Figure 2: 9 base-pair deletion frequencies across surveyed Bougainville villages. Language boundaries denoted by dashed lines, relevant dialect boundaries by dotted lines. (For details, see Merriwether, Friedlaender et al. 1999).

This variant set, called the 'Polynesian Motif', was supposed to have developed in the following way. It began with a deletion of a nine-DNA base pair block (9-bp) in one section of the circular mtDNA molecule (the intergenic region V) of a woman living somewhere in Asia tens of thousands of years ago. She likely already had one other variant in the control region of her mtDNA at position 16189, and this, in combination with the 9-bp deletion, would have been passed on to all her children as a unit, or haplotype. This combination, and the variants that subsequently developed from it, are all part of Haplogroup B, which is found throughout Asia, the New World, and the Pacific. As an aside, the same 9-bp deletion has been detected in Australian Aborigines (Betty et al. 1996) and Africans (Soodyall et al. 1996), but these cases probably represent independent mutational events, because they occur on different variant backgrounds. A second mutation occurred on the developing 'Polynesian Motif' at position 16217 (this combination is called Haplogroup B4). A sub-variety of that, B4a, adds a mutation at 16261, and this was thought to be limited to Austronesian-speaking groups (found from Madagascar to Remote Oceania). Finally, a fourth mutation occurred on B4a at site 16247, resulting in the full-blown 'Polynesian Motif', which was not found further west than eastern Indonesia, but reached frequencies of 100% or slightly less in the central and eastern Pacific. The 'Motif' was also found in Austronesian-speaking coastal New Guinea, but it was not detected in early studies in the New Guinea Highlands or among Australian Aborigines. This pattern of mutations was interpreted to mean that, as Austronesian-speaking populations moved east from an Asian homeland through the Pacific, the final 16247 mutation in the 'Motif' occurred along the migration route, likely in eastern Indonesia, and became more frequent as this migration moved eastward into Remote Oceania (Diamond 1988; Hertzberg et al. 1989; Stoneking & Wilson 1989; Lum et al. 1994; Lum & Cann 2000; Melton et al. 1995). Later on, others suggested the variant should be called the 'Austronesian Motif' (Richards et al. 1998). They argued its high diversity in the six samples found with the Motif from eastern Indonesia indicated it developed there approximately 17,000 years ago, which was much earlier than most authorities dated any Austronesian presence in that region (Bellwood et al. 1995).

How, then, could we explain the high frequency of the 'Polynesian Motif' in south Bougainville Papuan speakers? The most likely explanation seemed to be heavy female Austronesian influence in the southern groups. Ethnographic accounts had indicated at least some past marital exchange from Austronesian-speaking populations in the Shortland Islands to south Bougainville (Oliver 1954), but the degree of intermarriage was not well established. As explained below, we have now come to a different explanation as a result of our recent survey in the Bismarcks.

Island Melanesian mitochondrial DNA variation: recent developments

In 1998 and 2000, we extended the intensive sampling scheme of covering a number of villages across adjacent language groups to New Britain, New Ireland, Lavongai (New Hanover) and Mussau in the Bismarcks. New Britain and New Ireland were likeliest to replicate the pattern of heterogeneity we had found in Bougainville because of their large size. They each contained populations speaking Papuan and Austronesian languages—at least five separate Papuan languages in New Britain and one in New Ireland. At least one other language in New Ireland (Madak) retains evidence of Papuan heritage (Ross 1988).

We supplemented this core set with a collection of old plasmas and urine samples from a variety of expeditions dating as far back as the 1950s. Additional sequences from New Guinea were added from the literature for reference. These included eight sequences described as 'New Guineans from various parts of coastal and highland Papua New Guinea' by Vigilant et al. (1991:1503–1504), as well as 23 (primarily north) Coastal New Guinea sequences from Redd et al. (1995). The resulting coverage extends from west New Guinea to Tonga in the east, straddles the boundary between Near and Remote Oceania (Green 1991), and includes close to 900 analysed samples to date.

The geographic locations and sample sizes for the larger dataset are given in Table 3 and those for the core sample from the Bismarcks and Bougainville in Table 4. In the Bismarcks and Bougainville, we focused on sampling the Papuan-speaking groups and their Austronesian-speaking immediate neighbours. Most of the 'Motif' samples analysed to date are from non-Austronesian speaking groups (including groups in New Guinea as well as Island Melanesia). In the Bismarcks, we have good coverage from the Papuan-speaking Baining, Sulka, and Ata (Wasi) in New Britain, as well as the Kuot in New Ireland. The analysis of the Austronesian samples is incomplete. If known, we limited analysis to males from unrelated matrilines, so as to avoid obvious bias in the calculated haplotype frequencies. The methods used to extract and amplify the mtDNA segments, and to align sequences and define haplotypes and haplogroups, are detailed elsewhere (Friedlaender et al. 2002; Merriwether, Friedlaender et al. 1999).[1]

The first task was to relate our findings to those of other studies. Different researchers have sequenced different lengths of the mitochondrial DNA molecule or used different restriction fragment length polymorphisms (RFLPs) to identify haplotypes, and the resulting classifications make for difficult comparisons. The most common analysis employs a standard set of 14 restriction enzymes to cut the mtDNA at various points (high resolution RFLP analysis), along with sequencing 200 to 400 bases of the first segment of the hyper variable control region (HVS 1). The protocol for our results discussed here included sequencing a contiguous set of 1000 bases of the control region (sites 15960–00429) encompassing both segments 1 and 2 (HVS 1– and HVS 2), along with RFLP analysis and sequencing in approximately half the coding region on select representative samples. More extended sequencing is becoming the norm, and we will be presenting whole mtDNA genome results on representative samples in our collection shortly. These will enable the construction of more firm phylogenies of the variant haplotypes we have found, but not alter the frequency distributions. A remarkable picture of patterned mtDNA diversity has already emerged as a consequence of the intense sampling of the genetically diverse populations in this geographically limited region.

The mtDNA findings from these analyses amplify the early Bougainville results. The entire region is distinctive and diverse in terms of mtDNA. Besides finding most of the haplogroups known from earlier less intensive studies in the region, we have identified a number of new haplogroups or distinct branches of already defined ones. These are often restricted in their distributions to sections of New Guinea or Island Melanesia.

[1] The protocol for the sample collection and analysis was approved by the Human Subjects Committee Internal Review Boards of Temple University and the University of Michigan, and by the Medical Research Advisory Committee of Papua New Guinea.

Table 3: mtDNA haplogroup counts in the current survey, from New Guinea, Island Melanesia, as well as a small set from Micronesia and Polynesia

		New Guinea										Island Melanesia																
		West New Guinea		PNG	PNG Highlands				PNG Coastal			PNG island		New Britain		New Ireland		Bougainville		Solomon Islands		Vanuatu	New Caledonia	Fiji	Micronesia	Polynesia				
Haplogroup		Southwest Riverine	Lowland Riverine	PNG	Western Highlands	Fringe Highlands	Eastern Highlands	Morobe Highlands	Sepik	Markham	South Coast	Misima	Rossel	West New Britain	East New Britain	Mussau	New Ireland	North Bougainville	South Bougainville	Solomon Islands	Santa Cruz	Vanuatu	New Caledonia	Fiji	Kiribati	Ontong Java	Samoa	Tonga	Wallis	
B	B4a			2				1		1	4				2		13		2	6	1	2	1	1	1					37
	B4aPM	1		5				6	4	21	6			17	9	1	39	15	15	11	22	6	5	8	2	16	4	3	2	218
	B4*																20	2	7											29
	B4b1														1					1					1					3
P	P1	7	3	2	10	3	2	5	5	9	1	4	3		3		1	4		1	5	4	2							74
	P2	6			1			1		5	6	2											8	1						30
F										2												1	1							4
N*																							3							3
R*	R(V)			1				6																						7
	R(XIV)		1			8	1			7											1									18
Q	Q1	19	22	4	6	2	1	1	4	12				11	9			38			12			2						143
	Q2									10				14	21	1	5				7	2	1	1						62
	Q3			1		2																								3
	Q4								2																					2
E				2							2		2	20	1		8			2			1							38
M*	M*														9			2						1						12
	M(VII)																	12	4	2										18
	M(VIII)											1		40	73		4		1	2	21	7	4	2						155
	M(IX)			2											14					1		1								18
M7																							2		2	8				12
Total		33	26	19	17	15	4	20	15	67	19	7	5	102	142	2	90	73	29	26	69	23	28	16	6	24	4	3	2	886

Table 4: MtDNA haplogroup count in the Bismarcks and Bougainville

		New Britain												New Ireland						Bougainville								
		West New Britain				East New Britain								Mussau	New Ireland					North Bougainville				South Bougainville				
Haplogroup		Uasilau (Ata)	Lugei (Ata)	Kol	Pomio (Mengen)	Marabu	Malasait	Ganai (Sulka)	Watwat (Sulka)	Kabakada	Vunairoto	Matupit (Tolai)	Tolai	Mussau	Kavieng	Lamalaua (Kuot)	Kabil (Kuot)	Kuot	Madak	Aita	Rotokas	Eivo	Simeku	Nasioi	Torau	Nagovisi	Siwai	
B	B4a								1				1			3	1		9					1		1		17
	B4*																15	3	2			2				7		29
	B4b1								1																			1
	B4aPM			14	3			2		1	4	1	1	1	1	6	22	6	4	1		11	3		4	7	4	96
P	P1					2						1			1					4								8
Q	Q1		1	7	3	1		3	2	3										31	5	2						58
	Q2	5	3	6		21								1					5									41
E		8	4	7	1			1											8									29
M*	M*									6		1	2									1	1					11
	M(VII)																			2	6	2	2	4				16
	M(VIII)	4	8	28		34	19	5	6	3	3	2	1				3		1						1			118
	M(IX)									5		1	8															14
Total		17	16	62	7	58	19	11	10	18	7	6	13	2	2	9	41	9	29	38	11	18	6	5	5	15	4	438

Haplogroup definitions (also discussed in Attenborough, this volume). Table 5 lists the mtDNA haplogroups identified in New Guinea and Island Melanesia with the most characteristic mutations that distinguish each from the Cambridge Reference Sequence (CRS) (Anderson et al. 1981). There are 389 different haplotypes distributed among the haplogroups thus far defined from the sequencing of 886 samples. However, about 10% of these involve only minor distinctions due to hypervariable poly 'c' length polymorphisms which are of no phylogenetic importance.

More than 50 other sequences are not included here, as their phylogenetic relationships still remain unresolved. Similarly unresolved haplotypes are being reported elsewhere in isolated populations, as in Australia and South Asia (for example Thangaraj et al. 2003; Huoponen et al. 2001). Hopefully, their exact phylogenetic relations will become clear with complete mtDNA genome sequencing.

The haplogroups are arranged by similarity. Variants of the macro haplogroups M and N are ubiquitous outside Africa, both having developed from a single African clade, L3. Those haplogroups that do not belong to the M macro haplogroup lack the defining mutation T at 10400 and G at 10398, and often lack a T at position 16223 (see Table 5 over page). They are presumed to be part of the very large N macro haplogroup and its major subclade R, which contains Haplogroups B4a (and its 'Polynesian Motif' subdivision), B4b1, F, and P.

Haplotype networks of these sequences from the control region were generated using the Median Joining Network program (Bandelt et al. 1995; Bandelt, Forster & Rohl 1999). Figure 3 (over page) presents a network of our defined primary haplotypes, excluding singletons in the heaviest populated haplogroups for simplicity (that is within Haplogroups B, Q, P, and M(VIII)). This arrangement gives a sense of both the comparative internal diversity of the haplogroups and the phylogenetic relationships among them. The nodes represent specific haplotypes, and their sizes are scaled by the number of constituent samples. The distances between nodes reflect the number of mutation differences between haplotypes, weighted according to the relative frequency of their occurrences (recurrent mutations are weighted less). The black star denotes the approximate placement of the African origin L3. The branches of N are to the left of the star, while the branches of M are to the right.

1. Haplogroup B4a. Of the N clusters, Haplogroup B4a is the most important in our survey, since it has the largest number of constituent haplotypes (90) and samples (284). Most of its haplotypes are very tightly clustered around the core 'Motif' node as shown in Figure 4 (page following Figure 3). This indicates that most haplotypes within this haplogroup are only a single mutational step or two apart from each other, and suggests that the 'Motif' originated recently compared with the other haplogroups in the Figure 3 network. Twenty-nine samples within Haplogroup B4a lack the usual 16261 transition. This is likely because of back mutations from the full 'Motif', which apparently occurred in the Kuot of New Ireland and the Nagovisi of south Bougainville.

Such star-like haplogroup patterns are attributed to rapid population growth, and in certain circumstances their times of expansion can be reasonably estimated (Saillard et al. 2000; Forster et al. 2001; Richards et al. 1998). This is a comparatively diverse star array for any reported 'Motif' network. Its expansion time estimate is 8500 years ago. For eastern Indonesia, Richards et al. estimated an expansion time for the six samples found

Table 5: Defining mutations for mtDNA haplotypes in the Pacific

Numbers in HVS 1 and HVS 2 (the two segments of the control region) refer to transitions from the reference sequence. Transversions are denoted by letters; bold denotes most important; parentheses denotes inferred. Coding region polymorphisms are indicated by a combination of sequence distinctions plus RFLPs that infer key variants as follows: +10394c : +10394*Dde*I - infers (10398G); +10397a : +10397*Alu*I, - infers (10400T)

Haplogroup Designation	HVS 1	HVS 2	Coding Region	N
B4a	16189, **16217, 16261**	**146**	**9bp del**	**37**
B4aPM	16189, **16217, 16261, 16247**	**146**	**9bp del**	**218**
B4*	16189, **16217, 16247**	**146**	**9bp del**	**29**
B4b1	16189, **16217, 16136**	**207**	(9bp del), -10394c, -10397a	**3**
P1	**16357, 16176, 16266**	**212**	-10394c, -10397a, **15607, 6077**	**74**
P2	**16278, 16497**	**143**	-10394c, -10397a, **15607**	**30**
F	**16304, 16129, 16172**	**249 del**	-10394c, -10397a	**4**
R (V)	**16184, 16256**		-10394c, -10397a	**7**
R (XIV)	**16319**	**35, 36, 146, 152**	-10394c, -10397a	**18**
Q1	16223, 16129, **16241,** 16311**, 16265C, 16144, 16148, 16343**	**89, 146, 92**	+10394c, +10397a, **12940**	**143**
Q2	16223, 16129, **16241, 16066**	**228T, 195**	+10394c, +10397a, **12940**	**62**
Q3	16223, 16129, **16241,** 16311	**143**	+10394c, +10397a, **12940**	**3**
Q4	16223, 16129, **16241,** 16311	152		**2**
E	16223, **16362, 16390**	152	+10394c, +10397a**, 4491, 7598**	**38**
M (VII)	16223, **16048, 16077, 16172, 16311, 16320**	**195, 234**	+10394c, +10397a	**18**
M (VIII)	16223**, 16148, 16468,** 16362	**195**	+10394c, +10397a	**155**
M (IX)	16223, 16189	**211,** 151,152	+10394c, +10397a	**18**
M7b	16223, **16297**	**199**	+10394c, +10397a	**1**
M7c	16223, **16295**	**199, 146**	+10394c, +10397a	**11**

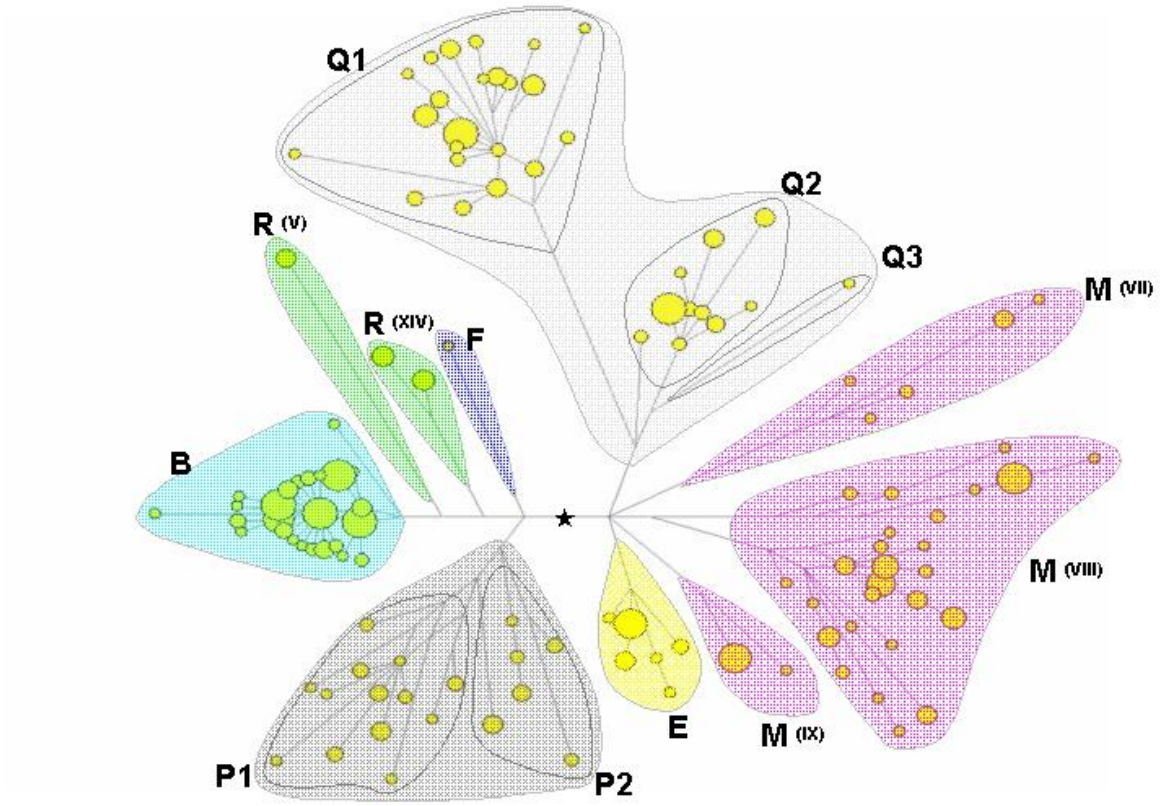

Figure 3: Bandelt median-joining network of haplotypes in the current study, primarily from New Guinea and Island Melanesia

there with the 'Motif' of 17,000 years ago.[2] Their argument for an earliest expansion date and origin for the Motif in eastern Indonesia is not convincing, given the very small Indonesian sample. The available evidence is more compatible with an origin of the 'Motif' in eastern New Guinea (this includes their own more reliable estimates, as well as Forster's, on larger samples from New Guinea) or even Island Melanesia (our data). From the Bismarcks or New Guinea it could well have spread east and west (Richards et al. did mention that because 'Melanesia' had been so poorly sampled, data from that region could change their conclusion). When we create 'Motif' networks for specific islands, New Ireland yields the highest value for mutational diversity in our series, with an associated estimated expansion time of over 12,200 years—and that is based on a sample of 60 'Motifs'.

[2] Richards et al. provide an associated 95% confidence interval of 5500 years to 34,500 years for their small Indonesian sample of Motifs. We have not provided standard errors with our time estimates for Motif expansions because the underlying distribution of mutation rates is not symmetrical at lower values, and therefore standard errors are inappropriate (Peter Forster, pers. comm.). The expansion age estimates in this range should be used as a very rough relative measure.

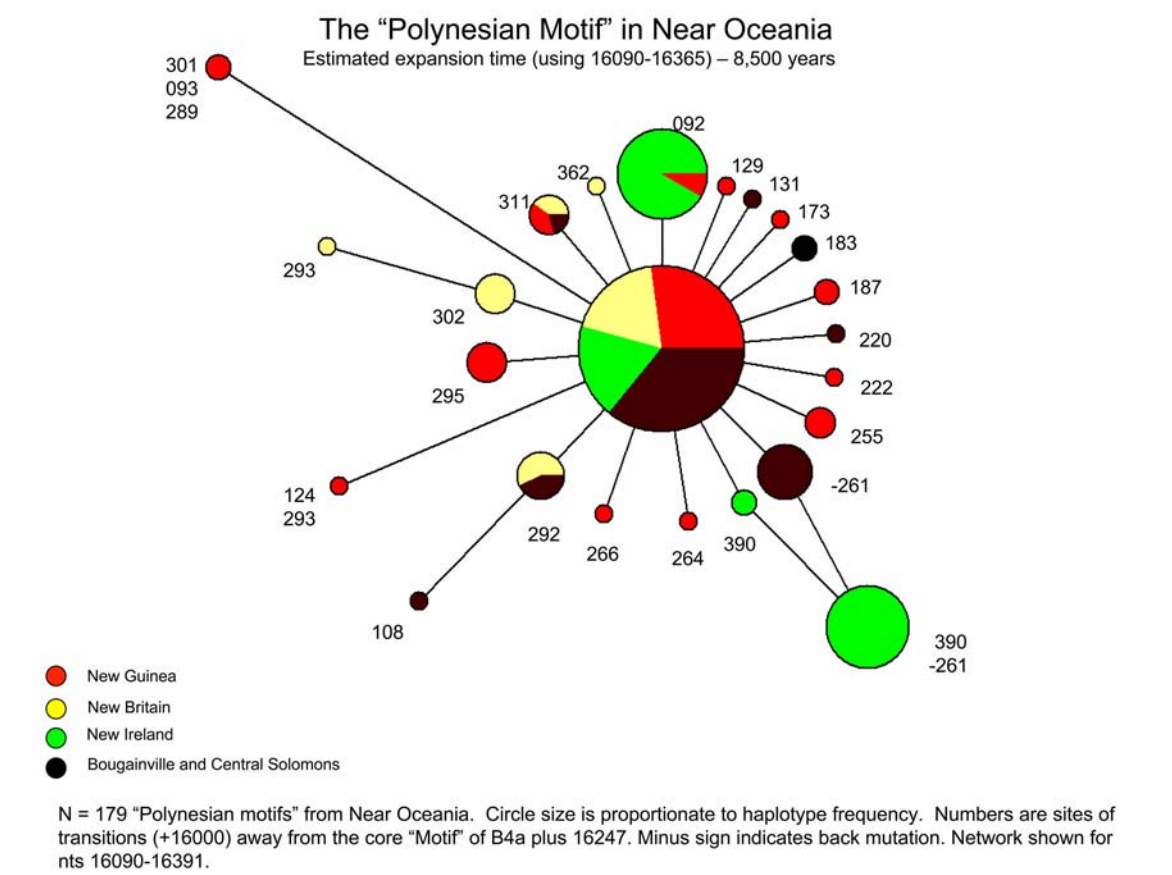

Figure 4: 'Polynesian motif' network

The configurations of all the other mtDNA haplogroup networks shown in Figure 3 are far less star-like or dense than B4a. The comparatively diffuse appearances of the other networks suggest they diversified considerably earlier than the 'Motif'.

2. Haplogroup P. Haplogroup P is the other common and very widespread haplogroup within the N clade in the region. As indicated in Table 3, P is a very diverse haplogroup, but its branches share an important common mutation in the coding region (15607). It has sequences that Forster et al. (2001) believe are ancestral to all haplotypes coming out of Africa in the hypothetical 'southern route'. Forster estimates its expansion date at either 65,400 or 33,300 years, depending on how narrowly defined the haplogroup is. The second date is for the main P1 branch alone. As is common in these estimations, the standard error is very large—about 23,000 years for the older date and 8100 for the younger. Our data give similar results for the respective definitions of P. It does not seem particularly star-like in our network because its most informative mutations lie outside the control region, which was the focus of our analysis (contra the RFLP based analysis of Forster et al. 2001). It is a candidate for being introduced to New Guinea/Australia (or Sahul) with the first arriving women, and appears to have distant branches in Australian Aborigines (Ingman & Gyllensten 2003; Huopenen et al. 2001).

3. Haplogroup F. Figure 3 and Table 5 show that Haplogroup F, which is very common in Southeast Asia, is rare in our sample. Among other characteristics, F is distinguished by a deleted base at 00249 (Kivisild et al. 2002) and the 16304C transition (Schurr & Wallace 2002). It is infrequent in the Pacific.

4. Haplogroups R(V) and R(XIV). The two remaining small branches in the macro N side of Figure 3, provisionally denoted as R(V) and R(XIV),[4] are novel haplotypes we found in different highlands New Guinea areas. We have not yet been able to link them specifically to a particular region of the R branch within the N clade. Although Sykes et al. (1995) found the two HVS1 variants we highlight for R(V) in Borneo and the Philippines, those belong to Haplogroup F, which R(V) does not seem to be a part of.

The other side of the network depicts the branches of macro haplogroup M. M is pervasive in Asia and extends westward as far as East Africa (Quintana Murci et al. 1999). Of those haplogroups belonging to M in this study, Haplogroup Q has been previously defined (Forster et al. 2001; Stoneking et al. 1990; Ballinger et al. 1992; Lum et al. 2000; Sykes et al. 1995; Tommaseo-Ponzetta et al. 2001). Refer to Table 5 for the defining sites. Combining our results with published data (especially Ingman & Gyllensten 2003), we can refer to branches Q1, Q2, Q3 and Q4, all sharing 16241 and probably 12940. Unlike P, the sub-branches of Q are often neatly regionally restricted, while the base haplogroups are more widespread. Although earlier studies suggested Q was less diverse and hence younger than P (with an expansion date of about 15,000 years with a very large standard error), the additional branches we have identified suggest an overall age possibly comparable with that of P. The pattern of Q branching is greatest in eastern New Guinea and Island Melanesia, suggesting the primary expansion of Q occurred there.

5. Haplogroup M(VIII). A number of the other haplogroups that belong to M, all newly defined here, are regionally important. Haplogroup M(VIII) has 67 constituent haplotypes that occur in a total 155 individuals, making it third in frequency in our series behind Haplogroups B and Q. It has a number of very long internal branches but is not star-like. We have sequenced half of the genome of two M(VIII) samples, and its specific affiliations are unresolved other than its situation within the M clade. It is apparently unique in its transition at 16468. There are hints of very distant relationships with mtDNAs from certain Veddic groups in India.

6. Haplogroups M(VII) and M(IX). M(VII) and M(IX) are two other variant haplotypes that remain to be specifically linked to the world-wide mtDNA tree. As with M(VIII), sequencing half the mtDNA genome has not made their connections clearer. Sequencing the rest of their mtDNA genomes should help. Both M(VIII) and M(IX) are regionally restricted and are far less diverse in their branching than M(VIII). From the network results, as well as an inspection of control region sequence similarities in the mtDNA phylogeny in Kivisild et al. (2002), M(VII) may turn out to be an especially deep additional branch of Q.

7. Minor haplogroups. Not shown in Figure 3 because of space considerations is a minor haplogroup ancillary to B4a – Haplogroup B4b1, which also has the 9-bp deletion and two of the associated mutations (16189 and 16217—see Table 5). It also shows consistent differences at a number of positions, lacking 16247, 16261, and 146, and adding the distinguishing 16136 and 207 mutations (Kivisild et al. 2002; Yao et al. 2002). Also

[4] The bracketed roman numerals refer to prior nomenclature used by our group, which will ultimately be replaced once more definitive phylogenetic associations are established.

not shown in Figure 3 is Haplogroup M7, which is very common in Asia generally and rare in our sample, and 12 other samples within M (M*) that are unresolved.

Interpreting the distances between pairs of haplogroups in Figure 3 is more problematic. One thing is clear, however. Most of these haplogroups are separated by considerable mutational distinctions. For example, within M it appears that the split between Haplogroups Q and M(VIII) is especially deep and therefore old.

Haplogroup geographical distributions. Figures 5 and 6 show the geographic distributions of the major haplogroup data listed in Tables 4 and 5, using the same colour code as in Figure 3. Many haplogroup distributions do not overlap while others tend to cluster together. Taken together, the overall pattern suggests considerable island-by-island distinctions along with a very high degree of within-island variation. Certain aspects require discussion.

The Polynesian Motif. As discussed, the 'Polynesian Motif' is the most common haplogroup throughout Island Melanesia, coastal New Guinea, and Remote Oceania. Besides the dating problems discussed, we cannot reconcile its geographical distribution within Island Melanesia with a neat Austronesian/Papuan split. Besides its very high frequency among south Bougainville Papuan speakers, we have now found the 'Motif' to be very common among some other inland Papuan-speaking populations in New Ireland, New Britain, and in some non-coastal eastern New Guinea regions. As shown in Figures 5 and 6 and Tables 4 and 5, the new Papuan-speakers with high 'Motif' frequencies are the Kuot (New Ireland), the Kol and Sulka (New Britain), and, among Trans New Guinea speakers in New Guinea, the Garaina in the Morobe Highlands and Gadio Enga of the upper Karawari River area in East Sepik Province. On the other hand, the Baining and Ata in central New Britain join the north Bougainville Papuans in their lack of the full 'Motif' or its antecedent B haplotypes. Perhaps there were ancient ties between the Baining, Ata, and Rotokas/Aita (lacking the 'Motif'), as opposed to the Kol, Sulka, Kuot, and south Bougainville Papuan speakers, where the 'Motif' is common and diversified.

New Guinea. Together, haplogroups P and Q predominate across New Guinea, with Q being more common in the west (see also Tommaseo-Ponzetta et al. 2001 for compatible and more extensive results from west New Guinea). In the Bismarcks and Solomon Islands, P occurs at very low frequency in the Baining and the Aita. It also has been detected in Vanuatu, Santa Cruz, New Caledonia, Central Micronesia, Palau (that variant resembles West New Guinea versions) as well as in four Indonesian samples. Q is found in Bougainville, Santa Cruz, New Caledonia, Micronesia, and at least in parts of Polynesia.

The green pie in Figure 5 represents the frequencies of two provisional minor haplogroups: R(V) and R(XIV). Haplogroup R(XIV) is more widely distributed than R(V), in the Fringe Highlands of New Guinea, and has also been detected in Santa Cruz and Vanuatu. R(V) has been found only among the Garaina in the Morobe Highlands.

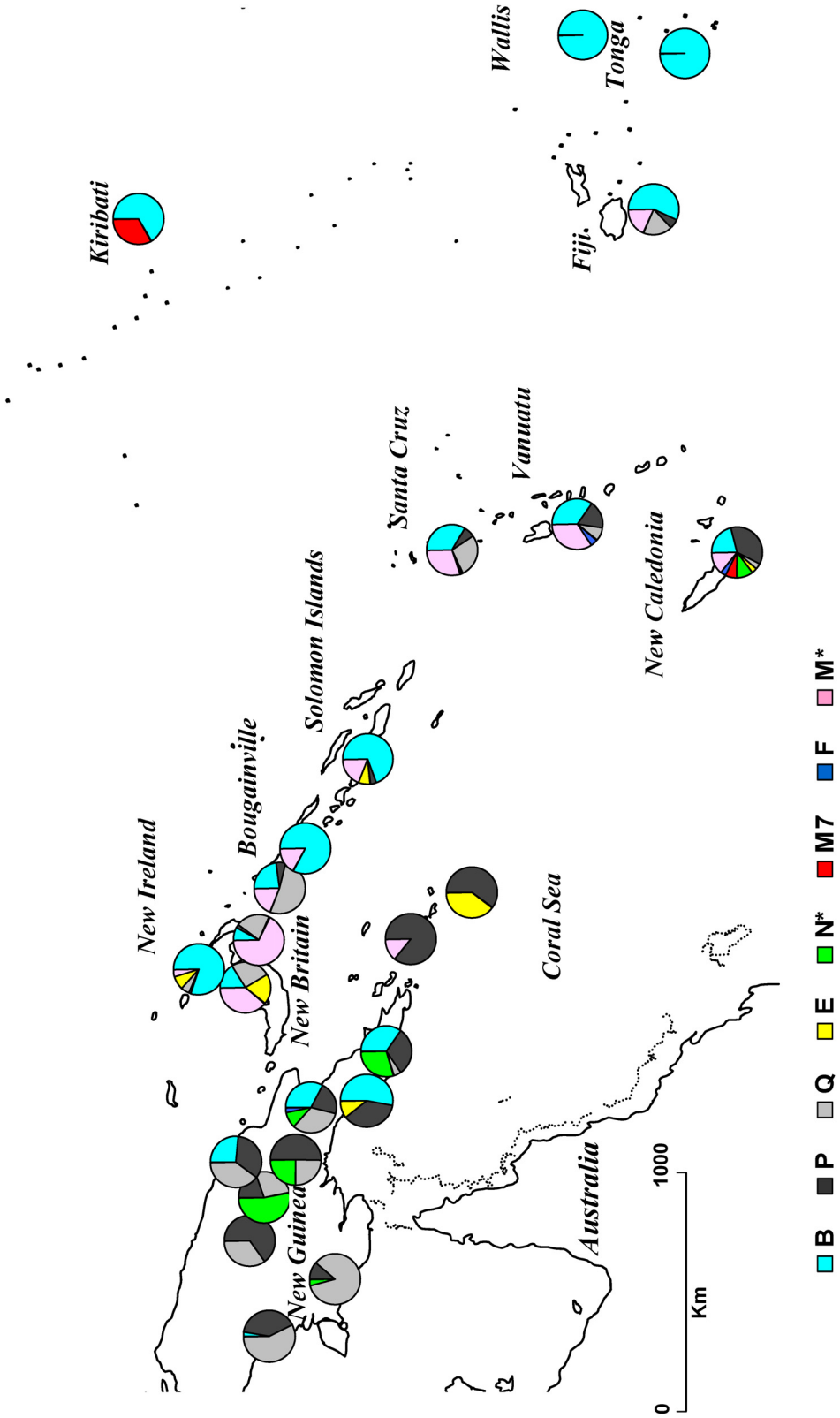

Figure 5: Southwest Pacific mtDNA haplogroup distributions

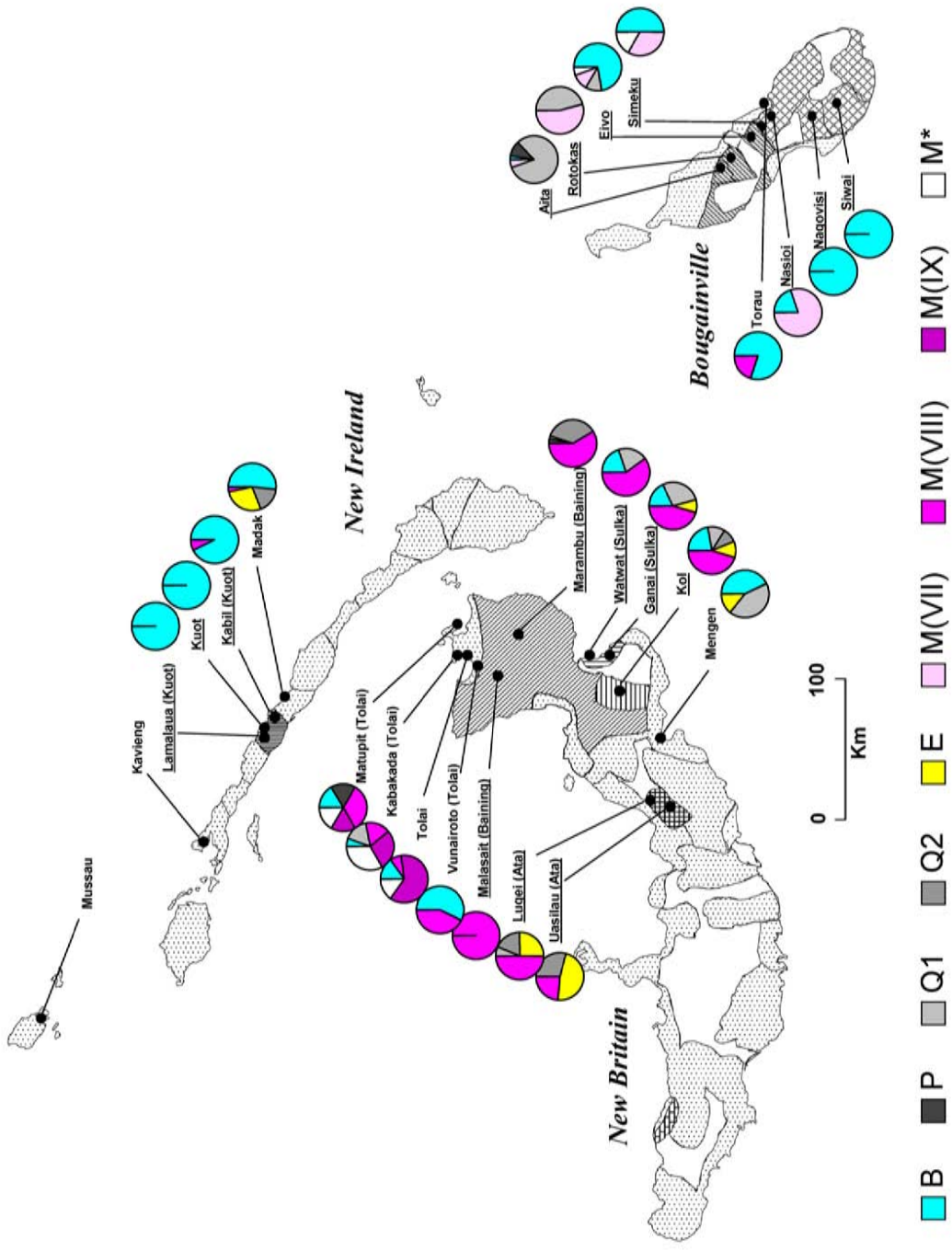

Figure 6: mtDNA haplogroup distributions in the Bismarcks and Bougainville. Within islands, dark hatching indicates Papuan-speaking areas, light stippling Austronesian areas, and uninhabited regions are blank.

The Bismarcks and Bougainville. M(VIII) is primarily Island Melanesian in its distribution, particularly the interior of New Britain. It has not been found in New Guinea—perhaps we should call it the 'Island Melanesian Motif'? It is most common among the Baining, being present in 100% of our samples from Mali (Malasait) dialect, and occurs in intermediate frequencies elsewhere on the island, except for the Mengen. It is low in frequency in New Ireland and is all but absent in Bougainville. It does occur sporadically elsewhere in Island Melanesia—in Misima, Vanuatu, New Caledonia, and Fiji. This secondary distribution of M(VIII) would seem to be best explained by recent population movements (Friedlaender et al. 2002).

Provisional haplogroups M(IX) and M(VII) have even more restricted distributions within Island Melanesia. M(IX) has been found primarily among the Tolai of East New Britain (who migrated from southern New Ireland a few generations ago), along with some sporadic cases in eastern Melanesia. Haplogroup M(VII) (Gentz et al. 2000) has been found in north and central Bougainville populations and in one Solomon Islander. Its highest frequency in Bougainville is in the Rotokas.

Minor haplogroups with an Asian affiliation. The mtDNA haplogroups which are most clearly associated with Asian or Southeast Asian populations are rare or absent in our sample.

We detected Haplogroup E in low but significant frequencies in New Britain among the Ata and Kol, and among the Madak of New Ireland, but scattered elsewhere in our sample. In Southeast Asia, E is found primarily among Sabah aboriginal groups and in the Philippines.

Haplogroup F was found in three widely divergent haplotypes and only four individuals, all in coastal situations—the Markham Valley in New Guinea, Vanuatu, and New Caledonia. F is one of the most common Asian haplogroups, from Japan, Korea, Mongolia, through China to western Indonesia.

We found M7 in our small Micronesian sample from Kiribati and in Ontong Java, a Polynesian Outlier with some Micronesian influence. In Asia, M7 is common across Mongolia, Japan, China, the Philippines, aboriginal Taiwan, to Indonesia.

B4B1 also occurs in Taiwan aborigines, the Philippines, Indonesia, Vietnam and China. We detected it in only three samples—from the Sulka region, Solomon Islands, and Kiribati.

If we take the 'Motif' as not being directly Asian-derived, the clear indicators of Asian influence are very rare in our large sample, excepting Haplogoup E in New Britain.

Generalisations

A great deal has been written about the relationship between distributions of language families and genetic variants in human populations, with the most prominent proponent of a direct connection being Luca Cavalli-Sforza. Our conclusion concerning this relationship in Island Melanesia argues for a more complex relationship. At particular moments, there may be a relatively neat bundle of a language, culture, and set of some alleles or gene variants that stay together for a time. But inevitably, this association will decay and become diluted. The rate of dilution will be influenced by the relative isolation of populations, measured most directly by their marital migration rates. Language affiliations may be informative of longer-term historical relations in particular situations.

In Island Melanesia, the distribution of Papuan languages still tracks, to a degree, with some old mtDNA haplogroups. The relation is not always strong. The more remote regions of the largest Melanesian islands retain the oldest genetic signatures. This goes for New Britain, Bougainville, and also for the eastern half of New Guinea. Major distinctions have developed among these three centres affecting a number of mtDNA haplogroups—not just one at a time. The more widespread haplogroups have diverse branching in this region, and some minor haplogroups occur in only one island section here. It may be, as more analysis is completed, that we can see clearer suggestions of multiple population movements within the region extending over the 40,000 years of its habitation.

The Bismarcks and Bougainville are distinguishable from the western part of Remote Oceania (that is Santa Cruz, Vanuatu, New Caledonia, and Fiji). Besides a high frequency of the B4a 'Motif', we found a mixture of various haplotypes at low frequencies with no clear geographic structuring there (refer again to Table 4, and see also Friedlaender et al. 2002). The variety in New Caledonia is puzzling and could be the result of poor sampling or recent migration. The ocean was clearly more of a highway than an obstacle to travel in this region.

As anthropologists, we are inclined to assume that no one gene will provide comprehensive information for population relationships and histories. Many hoped that with the power of mtDNA (and more recently Y-chromosome) analysis, the very confusing patterns of earlier genetic and biometric studies in this region could be essentially ignored and discarded. It now appears that regimens of intensive sampling of key geographic regions and the analysis of longer and longer segments of the mtDNA and Y chromosome are beginning to produce more compatible conclusions with one another. In turn, these are making more sense of other gene and biological trait variation. This was what we found to be the case in Bougainville, with a number of coinciding genetic distinctions between north and south.

A group of variants that appear to have their origins in New Guinea and Island Melanesia have been identified in the Y- and on other nuclear chromosomes (Martinson, Boyce et al. 1994; Kayser et al. 2001). For mtDNA, haplotypes Q and many branches of P all apparently had their origins in the region east of Wallace's Line and west of the Solomons. Almost no Ps or Qs have been found further west—about as many as the 6 'Motifs' from eastern Indonesia. In addition, M(VIII) is a mtDNA haplogroup that is common in Island Melanesia but nowhere else. We believe the 'Polynesian Motif' also belongs in this category. Judging from their relative diversities, P and Q are the oldest haplogroups in the region, followed by M(VIII), and with the 'Motif' developing relatively recently, but still prior to the Neolithic. This all suggests significant population movements in the region over a long time period.

To the west of Wallace's Line, a very different set of haplogroups predominates. These include haplogroups F, B4, M7, and D (which is frequent in Taiwan aborigines), and less commonly, E. Except for E, the others are very rare in our sample, especially in the Bismarcks, Bougainville, or New Guinea. These may well be the best indicators of the (low) degree of recent Southeast Asian (female) intermixture.

The Polynesians have a variety of allele frequencies that are a subset of Melanesian genetic variety, again with limited clear Asiatic influence. Again this is similar to the Y-chromosome findings (Kayser et al. 2001). It is hardly inevitable that all population expansions in this region spread out of Southeast Asia from west to east. Biologically, the initial settlers of Remote Oceania must have been an admixed lot, heavily influenced by

genetic drift, whatever their linguistic and cultural backgrounds. There now seem to be indicators of inter-island contact subsequent to settlement as well.

Our analysis of mtDNA variation in this sample set is far from complete. Intriguing areas within Island Melanesia should be covered (that is southern New Ireland, the ancient home of the distinctive Tolai; north Bougainville and Buka; the region including the Anêm of West New Britain; and the Western Solomons). Broader geographic comparisons and more extensive and complete mtDNA sequencing of representative samples will follow, including the many ambiguous haplotypes. Our group is also examining genetic variation in other non-recombining DNA segments in our large series, including the Y chromosome. The mitochondrial DNA is only one 'gene', in the genetic scheme of things, although, once again, it has proven to be a most informative one.

Acknowledgments

The late Ryk Ward played a special role in the development of this manuscript, and we would like to think he would have enjoyed the finished version. Thanks also go to an anonymous reviewer. This project has been supported by the US National Science Foundation, Temple University, the University of Michigan, the National Geographic Society Exploration Fund, and the Wenner-Gren Foundation. We especially wish to thank the sponsorship of the Papua New Guinea Institute for Medical Research, Mahonia Na Dari Nature Conservancy in Kimbe, and the aid of Noah Lurang, Father Joe Manteras, as well as the many people in Bougainville, New Ireland, and New Britain who were active and interested participants in this project. We hope our efforts will help develop more knowledge about their ancestry.

References

Anderson, Stephen, A.T. Bankier, B.G. Barrel, M.H.L. DeBulin, A.R. Coulson, J. Drouin, I.C. Eperon, D.P. Nierlich, B.A. Roe, F. Sanger, P.H. Schreier, A.J.H. Smith, R. Staden and I.G. Young, 1981, Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465.

Ballinger, Stephen W., T.G. Schurr, A. Torroni, Y.Y. Gan, J.A. Hodge, K. Hassan, K.-H. Chen and D.C. Wallace, 1992, Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient Mongoloid migrations. *Genetics* 130 (January):139–152.

Bandelt, Hans-Jurgen, P. Forster and A. Rohl, 1999, Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16(1):37–48.

Bandelt, Hans-Jurgen, P. Forster, B.C. Sykes and M.B. Richards, 1995, Mitochondrial portraits of human populations using median networks. *Genetics* 141(2):743–753.

Bellwood, Peter, J.J. Fox and D. Tryon, 1995, The Austronesians in history: common origins and diverse transformations. In Peter Bellwood, James J. Fox and Daryl Tryon, eds *The Austronesians: historical and comparative perspectives*, 1-16. Canberra: Department of Anthropology, Australian National University.

Betty, David J., A.N. Chin-Atkins, L. Croft, M. Sraml and S. Easteal, 1996, Multiple independent origins of the COII/tRNA(Lys) intergenic 9-bp mtDNA deletion in aboriginal Australians. *American Journal of Human Genetics* 58(2):428–433.

Diamond, Jared M., 1988, Express train to Polynesia. *Nature* 326:307–308.

Forster, Peter, A. Torroni, C. Renfrew and A. Rohl, 2001, Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution. *Molecular Biology and Evolution* 18:1864–1881.

Friedlaender, Jonathan S., 1971a, Isolation by distance in Bougainville. *Proceedings of the National Academy of Sciences USA* 68(4):704–707.

1971b, The population structure of south-central Bougainville. *American Journal of Physical Anthropology* 35(1):13–25.

1975, *Patterns of human variation: the demography, genetics, and phenetics of Bougainville islanders*. Cambridge, MA: Harvard University Press.

1987, Conclusion. In Jonathan Friedlaender, ed. *The Solomon Islands Project: a long term study of human biology and culture change*, 351–364. Oxford: Oxford Science Publications.

Friedlaender, Jonathan S., F. Gentz, K. Green and D.A. Merriwether, 2002, A cautionary tale on ancient migration detection: mitochondrial DNA variation in Santa Cruz Islands, Solomon Islands. *Human Biology* 74:453–471.

Friedlaender, Jonathan S., L.A. Sgaramella-Zonta, K.K. Kidd, L.Y. Lai, P. Clark and R.J. Walsh, 1971, Biological divergences in south-central Bougainville: an analysis of blood polymorphism gene frequencies and anthropometric measurements utilizing tree models, and a comparison of these variables with linguistic, geographic, and migrational "distances". *American Journal of Human Genetics* 23(3):253–270.

Friedlaender, Jonathan S. and A.G. Steinberg, 1970, Anthropological significance of gamma globulin (Gm and Inv) antigens in Bougainville Island, Melanesia. *Nature* 228:59–61.

Gentz, Frederick, J.S. Friedlaender and D.A. Merriwether, 2000, New mtDNA haplogroup variation in Bougainville Island. American Association of Physical Anthropology, San Antonio, Texas, *American Journal of Physical Anthropology* - Abstracts.

Giles, Eugene, E. Ogan and A.G. Steinberg, 1965, Gamma-globulin factors (Gm and Inv) in New Guinea: anthropological significance. *Science* 150:1158–1160.

Green, Roger C., 1991, Near and Remote Oceania - disestablishing "Melanesia" in culture history. In Andrew Pawley, ed. *Man and a half: essays in Pacific anthropology and ethnobiology in honour of Ralph Bulmer,* 491–502. Auckland: The Polynesian Society.

Hagelberg, Erika and J.B. Clegg, 1993, Genetic polymorphisms in prehistoric Pacific islanders determined by analysis of ancient bone DNA. *Proceedings of the Royal Society of London,* Series B. *Biological Sciences* 252*:*163–170.

Hertzberg, Mark, K.N.P. Mickleson, S.W. Serjeantson, J.F. Prior and R.J. Trent, 1989, An Asian specific 9-bp deletion of mitochondrial DNA is frequently found in Polynesians. *American Journal of Human Genetics* 44:504–510.

Huoponen, Kirsi, T.G. Schurr, Y. Chen and D.C. Wallace, 2001, Mitochondrial DNA variation in an aboriginal Australian population: evidence for genetic isolation and regional differentiation. *Human Immunology* 62:954–969.

Ingman, Max and U. Gyllensten, 2003, Mitochondrial genome variation and evolutionary history of Australian and New Guinean aborigines. *Genome Research* 13:1600–1606.

Kayser, Manfred, S. Brauer, G. Weiss, P.A. Underhill, L. Roewer, W. Schiefenhovel and M. Stoneking, 2001, Melanesian origin of Polynesian Y-chromosomes. *Current Biology* 11(2):1–2.

Kivisild, Toomas, H.V. Tolk, J. Parik, Y. Wang, S.S. Papiha, H.-J. Bandelt and R. Villems, 2002, The emerging limbs and twigs of the East Asian mtDNA tree. *Molecular Biology and Evolution* 19:1737–1751.

Lum, Ko-Ji and R.L. Cann, 2000, mtDNA lineage analyses: origins and migrations of Micronesians and Polynesians. *American Journal of Physical Anthropology* 113(2):151–168.

Lum, Ko-Ji, M.O. Richards, C. Ching and R.L. Cann, 1994, Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. *Human Biology* 66:567–590.

Martinson, Jeremy J., A.J. Boyce and J.B.Clegg, 1994, VNTR alleles associated with the alpha-globin locus are haplotype and population related. *American Journal of Human Genetics* 55(3):513–525.

Melton, Terry, R. Peterson, A.J. Redd, N. Saha, A.S.M. Sofro, J. Martinson and M. Stoneking, 1995, Polynesian genetic affinities with Southeast Asian populations identified by mtDNA analysis. *American Journal of Human Genetics* 57:403–414.

Merriwether, D. Andrew, J.S. Friedlaender, J. Mediavilla, C.F. Mgone, F. Gentz and R.E. Ferrell, 1999, Mitochondrial DNA variation is an indicator of Austronesian influence in Island Melanesia. *American Journal of Physical Anthropology* 110(3):243–270.

Merriwether, D. Andrew, F.A. Kaestle, B. Zemel, G. Koki, C.F. Mgone, M.P. Alpers and J.S. Friedlaender, 1999, Mitochondrial DNA variation in the southwest Pacific. In Ranjan Deka, Ranajit Chakraborty and S.S. Papia, eds *Genomic diversity: applications in human population genetics*, 153–181. New York: Kluwer Academic/Plenum Publishers.

Neel, James V., 1970, Lessons from a "primitive" people. *Science* 170(960):815–822.

Oliver, Douglas L., 1954, *A Solomon Island society: kinship and leadership among the Siuai of Bougainville*. Cambridge MA: Harvard University Press.

Quintana-Murci, L., O. Semino, H.-J. Bandelt, G. Passarino, K. McElreavey and A.S. Santachiara-Benerecetti, 1999, Genetic evidence of an early exit of *Homo sapiens sapiens* from Africa through eastern Africa. *Nature Genetics* 23*:*437–441.

Redd, Alan J., N. Takezaki, S.T. Sherry, S.T. McGarvey, A.S.M. Sofro and M. Stoneking, 1995, Evolutionary history of the COII/tRNALys intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. *Molecular Biology and Evolution* 12(4):604–615.

Richards, Martin, S. Oppenheimer and B. Sykes, 1998, mtDNA suggests Polynesian origins in Eastern Indonesia. *American Journal of Human Genetics* 63*:*1234–1236.

Ross, Malcolm, 1988, *Proto Oceanic and the Austronesian languages of Western Melanesia*. Canberra: Pacific Linguistics.

Saillard, Juliette, P. Forster, N. Lynnerup, H.-J. Bandelt and S. Norby, 2000, mtDNA variation among Greenland Eskimos: the edge of the Beringian expansion. *American Journal of Human Genetics* 67(3):718–726.

Schurr, Theodore G. and D.C. Wallace, 2002, Mitochondrial DNA diversity in Southeast Asian populations. *Human Biology* 74:431–452.

Seielstad, Mark T., E. Minch and L.L. Cavalli-Sforza, 1998, Genetic evidence for a higher female migration rate in humans. *Nature Genetics* 20:278–280.

Sokal, Robert and J.S. Friedlaender, 1982, Spatial autocorrelation analysis of biological variation on Bougainville Island. In M.H. Crawford and J.H. Mileke, eds *Current developments in anthropological genetics*, vol. 2: *Ecology and population structure,* 205–227. New York: Plenum Press.

Soodyall, Himla, L. Vigilant, A.V. Hill, M. Stoneking and T. Jenkins, 1996, mtDNA control region sequence variation suggests multiple independent origins of an "Asian - specific" 9-bp deletion in sub-Saharan Africans. *American Journal of Human Genetics* 58:595–608.

Stoneking, Mark, L.B. Jorde, K. Bhatia and A.C. Wilson, 1990, Geographic variation in human mitochondrial DNA from Papua New Guinea. *Genetics* 124(March):717–733.

Stoneking, Mark and A.C. Wilson, 1989, Mitochondrial DNA. In A.V.S. Hill and S.W. Serjeantson, eds *The colonization of the Pacific: a genetic trail,* 215–245. Oxford: Oxford Scientific Publications.

Sykes, Brian, A. Lieboff, J. Low-Beer, S. Tetzner and M.O. Richards, 1995, The origins of the Polynesians: an interpretation from mitochondrial lineage analysis. *American Journal of Human Genetics* 57:1463–1475.

Thangaraj, Krishna, L. Singh, A.G. Reddy, V.R. Rao, S.C. Sehgal, P.A. Underhill, M. Pierson, I.G. Frame and E. Hagelberg, 2003, Genetic affinities of the Andaman Islanders, a vanishing human population. *Current Biology* 13:86–93.

Tommaseo-Ponzetta, Mila, M. Attimonelli, M. De Robertis, F. Tanzariello and C. Saccone, 2001, Mitochondrial DNA variability of West New Guinea populations. *American Journal of Physical Anthropology* 117:49–67.

Vigilant, Linda, M. Stoneking, H. Harpending, K. Hawkes and A.C. Wilson, 1991, African populations and the evolution of human mitochondrial DNA. *Science* 253(September 27):1503–1507.

Yao, Yong-Gang, Q.-P. Kong, H.-J. Bandelt, T. Kivisild and Y.-P. Zhang, 2002, Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *American Journal of Human Genetics* 70:635–651.