

Unexpected NRY Chromosome Variation in Northern Island Melanesia

Laura Scheinfeldt,¹ Françoise Friedlaender,* Jonathan Friedlaender,* Krista Latham,* George Koki,† Tatyana Karafet,‡ Michael Hammer,‡ and Joseph Lorenz*§

*Anthropology Department, Temple University, Philadelphia; †Papua New Guinea Institute for Medical Research, Goroka, Papua New Guinea; ‡Division of Biotechnology, University of Arizona; and §Coriell Institute for Medical Research, Camden

To investigate the paternal population history of populations in Northern Island Melanesia, 685 paternally unrelated males from 36 populations in this region and New Guinea were analyzed at 14 regionally informative binary markers and 7 short tandem repeat (STR) loci from the nonrecombining portion of the Y chromosome. Three newly defined binary markers (K6-P79, K7-P117, and M2-P87) aided in identifying considerable heterozygosity that would have otherwise gone undetected. Judging from their geographic distributions and network analyses of their associated STR profiles, 4 lineages appear to have developed in this region and to be of considerable age: K6-P79, K7-P117, M2-P87, and M2a-P22. The origins of K5-M230 and M-M4 are also confirmed as being located further west, probably in New Guinea. In the 25 adequately sampled populations, the number of different haplogroups ranged from 2 in the single most isolated group (the Aita of Bougainville), to 9, and measures of molecular diversity were generally not particularly low. The resulting pattern contradicts earlier findings that suggested far lower male-mediated diversity and gene exchange rates in the region. However, these earlier studies had not included the newly defined haplogroups. We could only identify a very weak signal of recent male Southeast Asian genetic influence (<10%), which was almost entirely restricted to Austronesian (Oceanic)-speaking groups. This contradicts earlier assumptions on the ancestral composition of these groups and requires a revision of hypotheses concerning the settlement of the islands of the central Pacific, which commenced from this region.

Introduction and Background

Human populations in Northern Island Melanesia are particularly diverse and distinctive because of their very long and relatively isolated settlement history. Yet, prior studies of variation in the nonrecombining portion of the Y chromosome (NRY) that included some populations from this region did not suggest any particular structure or heterozygosity there (Cox 2003; Kayser et al. 2003). Our intensive sampling effort, coupled with an analytic battery that included newly identified single-nucleotide polymorphisms (SNPs), has now revealed considerable NRY variation in this region.

Portions of Northern Island Melanesia were settled by modern humans at least 42 000 years before present (YBP) (Allen 2003; Leavesley and Chappell 2004), apparently not long after they reached New Guinea, which was joined at that time to Australia, forming the ancient Pleistocene continent of Sahul. New Guinea and adjacent Northern Island Melanesia constitute the region of Near Oceania, and the populations there were at the eastern edge of the human species range until 3200 YBP.

The earliest populations of Northern Island Melanesia were small groups of hunter-gatherers. They did not simply roam the coasts and lagoons but intermittently settled the interiors of the large islands (Pavlidis and Gosden 1994; Allen 2003). The intentional introductions of plants and animals from New Guinea over the following millennia indicate continuing outside contacts at a modest level. The colonization of relatively remote Manus Island prior to 20 000 YBP infers the use of steerable watercraft such as canoes (Allen 2003). People had successfully made the

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 obsidian to New Ireland and other nearby islands up to ~7000 YBP (Summerhayes and Allen 1993). The implication of these data 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 Upper Pleistocene. By extension, movements between Near Oceania and Island Southeast Asia also would have been intermittent and small in scale.

The languages dating to this period, referred to as Papuan, are so old that their ties to one another are ambiguous and open to controversy (Wurm 1975; Ross 2001; Dunn et al. 2002, 2005; Reesink 2005; Ross 2005). The most recent analysis suggests an “archipelago-based” relationship of the Island Melanesian Papuan languages (Dunn et al. 2005), with very separate New Britain and Bougainville clusters, and the New Ireland Kuot in between. The Papuan languages of New Guinea are unrelated to those in Island Melanesia. These include the large Trans New Guinea language family, spoken across the mountainous interior, as well as a number of other language families, especially in the north, including the Sepik River region.

It was also in a section of Northern Island Melanesia, the Bismarck Archipelago, that the Lapita cultural complex developed within a short period and subsequently spread to many formerly uninhabited islands of the Pacific by 3200 YBP (Pawley and Green 1973; Shutler and Marck 1975; Green 1997; Spriggs 1997). Lapita is associated with the development of the large Oceanic branch of the Austronesian languages. Although the Austronesian language family has its origin in the vicinity of Taiwan (Gray and Jordan 2000), all Austronesian languages spoken from western New Guinea across the Pacific belong to its Oceanic branch, including all those in our sample. The degree to which the entire Lapita phenomenon represented an intrusion of people and material culture into Northern Island Melanesia from Southeast Asia and Taiwan, a fusion of Southeast Asian and indigenous Island Melanesian components, or an essentially autochthonous Island Melanesian development, is still

¹ Present address: Division of Human Genetics, Children’s Hospital of Philadelphia, Abramson Research Center, Philadelphia.

Key words: Melanesia, NRY, Pacific, phylogeography, Bismarck Archipelago.

E-mail: scheinfeldt@email.chop.edu.

Mol. Biol. Evol. 23(8):1628–1641. 2006

doi:10.1093/molbev/msl028

Advance Access publication June 5, 2006

under hot debate by archaeologists (for a review see Green 2003).

Genetic distributions across the region are consistent with a very long and complex history, with substantial early isolation. Populations in Northern Island Melanesia retain distinctive genetic signatures, best understood until now for mitochondrial DNA (mtDNA) variants (Ingman and Gyllenstein 2003; Friedlaender et al. 2005a, 2005b, forthcoming; Merriwether et al. 2005). Along with aboriginal populations from Australia, the Andaman Islands, the Philippines, and Malaysia, these groups have long been suspected of retaining genetic ties to South Asia and to some of the earliest migrations of modern humans out of Africa (Underhill et al. 2000), although the timing and number of these exoduses remain controversial and unresolved (Cordaux and Stoneking 2003).

However, for the Late Holocene, the reported mtDNA and NRY patterns of variation have been contradictory and controversial. The mtDNA evidence supports an Aboriginal Taiwan origin of the haplotype that formed the background on which the so-called "Polynesian Motif" developed in Near Oceania, where it has extremely high frequencies in many populations, before spreading to Remote Oceania (Trejaut et al. 2005; Friedlaender et al. forthcoming). This has been interpreted as supporting a clear and pervasive ancestral/descendant relationship between Taiwan Aboriginal populations and Austronesian-speaking groups in Near and Remote Oceania. On the other hand, previous NRY studies have found very little evidence for a link between Taiwan Aboriginal or Southeast Asian populations and those of Near or Remote Oceania, suggesting that the Polynesians and Micronesians are primarily derived from Near Oceanic (Melanesian) populations (Kayser et al. 2000; Su et al. 2000; Capelli et al. 2001). Therefore, it is unclear how much of a recent Southeast Asian/Taiwanese connection exists in the contemporary genetic profile of Near and Remote Oceanic populations and how these profiles relate to models of the settlement of Remote Oceania.

Here we describe NRY variation in an intensively studied series of populations in the key region of Northern Island Melanesia. To characterize the paternal lineages in this region, we utilized a battery of SNPs that includes 3 highly informative new markers, in addition to 7 commonly typed short tandem repeats (STRs). The combination of intensive sampling and testing with a finer SNP screen reveals just how remarkable and old the NRY variability across this small region is and relates variation within the region more clearly to that reported elsewhere in the Pacific.

Materials and Methods

The samples were selected from our Northern Island Melanesian collection of over 1500 blood samples from 3 field seasons led by J Friedlaender and G Koki—1998, 2000, and 2003. The collection covers 36 language groups and dialects in the region, from both Oceanic and Papuan language families. The sample design focused on Papuan-speaking groups in New Britain, New Ireland, and north Bougainville and their Oceanic-speaking neighbors.

Information gathered on survey participants included their language, residence, and a short genealogy. A total of

685 males who appeared to be paternally unrelated based on the short genealogy constitute the sample for this particular study. Of the 36 groups, 25 had sample sizes of 10 or more. The samples from New Guinea do not represent a true population because they were men from a number of areas (primarily the Sepik) who had settled in the Northern Island Melanesian region. Also, one of the islands (Mussau) was only represented by a single sample set and was therefore excluded from the core population analysis. Therefore, although we report the results for all 36 groups, only the 23 that were adequately covered and constitute multiple population sets within major island subdivisions were included in most statistical analyses (analysis of the molecular variance [AMOVA] and multidimensional scaling [MDS]).

The samples were collected and analyzed with informed consent protocols approved by the appropriate Human Subjects Ethical Committees of Papua New Guinea and Temple University. DNA from the samples was extracted using the Genra Systems Puregene Genomic DNA Isolation Kit (Genra Systems, Minneapolis, Minnesota).

Our choice of the NRY SNP markers followed the YCC tree (Jobling and Tyler-Smith 2003) as abstracted in figure 1, and we followed their lineage naming terminology. The 14 SNPs that were regionally informative are shown. P79 and P117 identify new subdivisions within haplogroup K (defined by M9), now named K6-P79 and K7-P117. P87 lies within the M lineage (defined by M4) and is a predecessor of P22, which identifies M2a-P22. Those samples that were positive for P87 but negative for P22 were listed as M2-P87*. We also tested for the 50f2/c deletion (Jobling et al. 1996) that was found only on the K background in our series. We chose 7 STR loci for analysis because they had been reported as polymorphic and informative in the Southwest Pacific in earlier studies: DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 (Forster et al. 1998; Kayser et al. 2001).

The panel of SNP markers employed in this study comes from previously published work: M89, M201, M69, M170, M9, RPS4Y (M130), M175, M4, M230, M74, M38, M119, M122, P22, M16, and M208 (Kayser et al. 2001; Underhill et al. 2001). In each case, we used 25 μ l polymerase chain reactions (PCRs) utilizing reagents purchased from Applied Biosystems (Foster, California): 2 μ l DNA (50 ng/ μ l concentration), 2.5 μ l 10 \times buffer, 2.5 mM MgCl₂, 2 μ l deoxynucleoside triphosphate, 0.2 μ l Taq Gold polymerase, 0.25 μ l of each 25 μ M primer (forward and reverse), and 15.3 μ l distilled H₂O. Previously published primers were utilized (Kayser et al. 2001; Underhill et al. 2001).

The genotyping methods for the 3 novel SNPs are as follows. Two SNPs were sequenced: P79 and P87. The P79 PCR is 1135 bp, the mutation is a T \rightarrow C change at base 258, and the primers are F: ttgtttgctctgctgtg and R: tctga-gaatagtcagatgt. The P87 PCR product is 676 bp, the mutation is an A \rightarrow C change at base 224, and the primers are F: acagcgagcagtaagtaa and R: actaacctcaccaatct. The P117 PCR product is 1081 bp, and the mutation is a G \rightarrow T change at base 68. The T mutation loses an *MnII* restriction site, and the primers are F: ctgattattctttctaccttg, R: gttatgc-caggaaacatgcc.

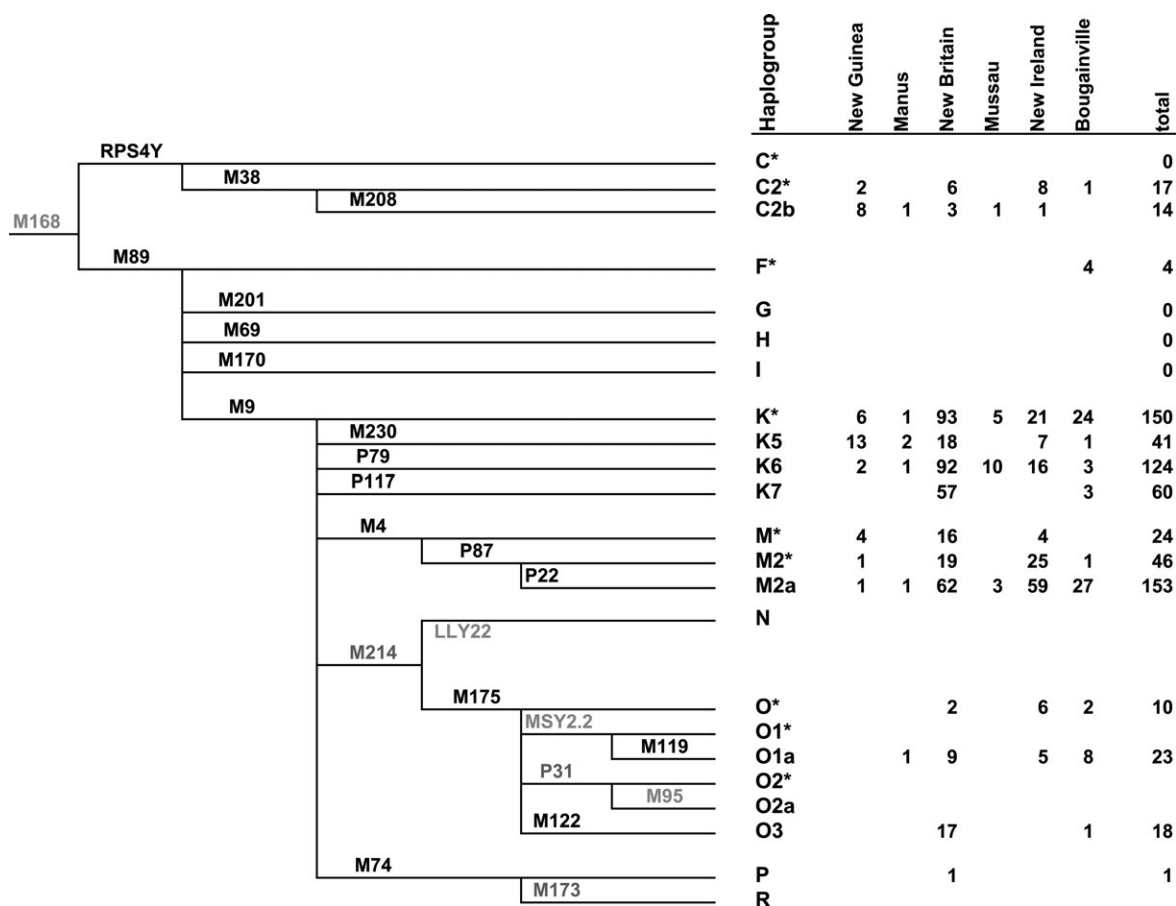


FIG. 1.—The phylogenetic relationship of 14 NRY haplogroups based on 22 SNPs. The results on 18 SNPs (in black) are summarized by primary island group in the right hand column (New Hanover is merged with New Ireland). The SNPs in gray are presented only for context but have not been analyzed in the current study.

For the statistical analyses, we used the program ARLEQUIN 3.000 to calculate several population genetic parameters. These included measures of population diversity and distances on the SNP data (\hat{H} and F_{ST}) and STR data (average π_n and R_{ST}), respectively (Schneider et al. 2000). The NRY haplogroup data were entered restriction fragment length polymorphism data with a haplogroup list. To display the population NRY haplogroup distances, we performed nonparametric MDS on the pairwise SNP F_{ST} values, using the SPSS software program. AMOVA (again using ARLEQUIN 3.000) was run on different formats or population partitioning schemes: 23 core populations (as notated in table 2 with an asterisk), 3 islands (New Ireland, New Britain, and Bougainville), and 2 language groupings (Austronesian and Papuan).

Median-joining networks based on microsatellite variation within each haplogroup were constructed using the NETWORK 4.1.1.2 program (Bandelt et al. 1999). The reduced median output (threshold = 1) was used as input for the median-joining network, which was used for the Steiner output. For all statistical analyses, DYS389I was subtracted from DYS389II because the DYS389II PCR product also contains DYS389I. It was not possible to include each mutational step between nodes because of the complexity of the networks; however, the distance between nodes was

kept to scale relative to the number of steps between each haplotype. NETWORK 4.1.1.2 was also used for generating mismatch distributions from these same files.

To estimate haplogroup divergences, the ancestral node was identified from the network and run in the Ytime program (Behar et al. 2003), using the estimated mutation rate of Zhivotovsky et al. (2004). This method was employed because it is relatively straightforward, easily replicable, and we were particularly interested in the relative dating of the haplogroups rather than the age of each population split.

Results

Table 1 summarizes the incidences of NRY SNPs that have commonly been screened in populations across Southeast Asia and Oceania. We were particularly interested in comparing our data with these 2 regions and were therefore forced to limit table 1 to data containing comparable sets of SNPs. Table 2 presents the extended SNP results for the populations in our Northern Island Melanesian series, as well as the 37 New Guinea men who had settled in the region. It also includes measures of SNP and STR population diversity. Figure 2 shows the geographical frequency of the SNP haplogroups by major population in our series for Northern Island Melanesia. Even at first inspection, the

Table 1
Y Chromosome Lineage Frequencies of Island Melanesia and Nearby Regions

		Haplogroup Frequencies (%)																	Other Source ^a
		M89	RPS4Y	M38	M208	M217	390.1d	M9	M230	M4	P22	M175	M119	M95	M122	M74	M173		
		N	F*	C*	C2*	C2b	C3		K*	K5	M*	M2a	O*	O1a	O2a	O3	P	R	
Asia																			
	Korea	25	8	0	0	0	12	0	8	0	0	0	32	4	0	28	4	4	a
	China	36	0	0	0	0	6	0	11	0	0	0	0	22	3	58	0	0	a
Southeast Asia																			
	Vietnam	11	0	0	0	0	9	0	0	0	0	0	9	36	46	0	0	0	a
	Taiwan	26	0	0	0	0	4	0	4	0	0	0	12	23	0	58	0	0	a
	Chinese	43	0	0	0	0	0	0	0	0	0	0	5	79	5	12	0	0	a
	Aborigines	26	0	0	0	0	0	0	0	0	0	0	4	96	0	0	0	0	c
	Paiwan	53	0	0									0	74	24	2	0		c
	Bunun	50	0	0									0	98	0	2	0		c
	Atayal	50	0	0									0	43	6	47	0		c
	Ami	53	0	0									0	70	28	0	0		c
	Yami	40	0	2	?	?	?						0	28	3	39	0	2	a, c, f
	Philippines	115	1	4	0	0	1	0	20	2	0	0	0	8	32	30	0	4	5
	Malaysia	50	6	2	0	0	4	0	8	0	2	0	0	23	42	23	2	4	a, f
	Java	53	2	2	0	0	0	0	2	0	0	0	2	23	42	23	2	4	a
	S. Borneo	40	5	0	3	0	3	0	10	0	0	0	5	15	38	18	3	3	a
	Balinese	551	0	2	0	0	0.2		1	0.4	1		18	59	7	0	4	8	e
	E. Indonesia	55	0	6	27	0	0		20	13	13	?	13		9				f
	Moluccas	34	0	9	15	0	0	0	18	21	21	0	0	6	0	12	0	0	a
	Nusa Tenggara	31	0	7	16	0	0	0	32	10	7	0	0	23	0	3	0	3	a
New Guinea																			
	WNG	94	0	0	0	25	0	0	0	0	75	0	0	0	0	1	0	0	a
	WNG																		
	Highlands	89	0	0	9	0	0	0	11	2	78	0	0	0	0	0	0	0	a
	Lowlands/Coast	31	0	3	13	10	0	0	16	16	29	0	0	0	10	3	0	0	a
	PNG	31	0	0	3	0	0	0	7	52	36	0	0	3	0	0	0	0	a
	PNG	31	0	0	3	0	0	0	7	52	36	0	0	3	0	0	0	0	a
	PNG	7	0	0	0	0	0	0	29	57	14	0	0	0	0	0	0	0	g
	Coast	25	0	0	8	28	0	0	16	32	12	4	0	0	0	0	0	0	g
Island Melanesia																			
	Trobriand	53	0	0	0	9	0	0	23	0	30	0	0	28	0	9	0	0	a
	Manus	7	0	0	0	14	0	0	29	29	0	14	0	14	0	0	0	0	g
	New Britain	East	145	0	0	3	1	0	57	8	11	15	1	1	?	3	0	0	g
	West	245	0	0	1	0.4	0	0	64	3	7	16	0	3	0	5	0.4	?	g
	Mussau	20	0	0	0	5	0	0	75	0	5	15	0	0	0	0	0	0	g
	New Hanover	43	0	0	0	0	0	0	12	2	12	70	2	2	?	0	0	0	g
	New Ireland	109	0	0	7	1	0	0	29	6	21	27	5	4	?	0	0	0	g
	Bougainville	North	54	7	0	2	0	0	48	2	2	18	4	15	?	2	0	0	g
	Central	18	0	0	0	0	0	0	17	0	0	83	0	0	0	0	0	0	g
	Vanuatu	234	0	18	?	?	?	?	41	6	30	?	0	0	0	4	0	2	b
	Fiji	55	0	3	?	?	?	?	41	?	15	?		6		9			26
Polynesia																			
	Tonga	55	0	23	?	?	?		1	?	8	?	0	2	0	58	0		8
	Western Samoa	16	0	69	?	?	?		6	?	0		0	6	6	13	0		c
	Cook Islands	28	0	0	0	82	0	0	4	0	0	0	0	0	0	7	0	7	a
	Atiu	42	0	84	?	?	?		1	?	0		0	0	0	3	3		9
	French																		
	Polynesia	87	0	53	?	?	?		8	?	0		0	2	0	35	2		c
	Maori	54		0	42	?	0		4	?	?					5		33	16
Australia																			
	Arnhem	60	2	10	0	0	0	53	30	0	0	0	0	0	0	0	0	5	a
	Desert	35	3	0	0	0	0	69	17	0	0	0	0	0	0	3	0	9	a

NOTE.—? = untested. Data from current study are shown in bold italics.

^a Data sources— a: Kayser et al. (2003), b: Cox (2003), c: Capelli et al. (2001), d: Underhill et al. (2001), e: Karafet et al. (2005), f: Hammer et al. (2005), and g: this study.

SNP variation across Northern Island Melanesia is remarkable. Not only did New Guinea, New Britain, New Ireland, and Bougainville have very different SNP signatures but SNP frequency variation within islands was also clear, most especially for New Britain, the largest and best-sampled island in the survey.

Lineage C-RPS4Y

C is considered the oldest lineage in Asia and the Southwest Pacific, likely introduced with the first settlers of both the Sunda shelf and the ancient continent of Sahul, with an estimated origin time of perhaps 50 000 YBP (Underhill 2004). The C3 division is Asian, found as far

Table 2
Y Chromosome Lineage Frequencies in Island Melanesia

Island Region	Population ^a	Language ^b	N	Haplogroup Frequencies (individuals)													
				M89 F*	M38 C2*	M208 C2b	M9 K*	M230 K5	P79 K6	P117 K7	M4 M*	P87 M2*	P22 M2a	M175 O*	M119 O1a	M122 O3	M74 P
New Guinea																	
PNG Coast	North Coast		25		2	7	2	8	2		2	1	1				
	Markham		1								1						
	Rigo		2		1			1									
PNG Highlands	Eastern Highlands		4					3			1						
	Morobe Highlands		1				1										
	Western Highlands		2				1	1									
PNG Island	Misima		2				2										
Manus			7			1	1	2	1				1		1		
New Britain																	
West New Britain	Kove*	O	24				8			1	7		2	1		4	1
	Anêm*	P	34			1	5	3	2	16	5		2				
	Mangseng*	O	11				5	1	3	1			1				
	Mamusi*	O	43				10		17			1	11		4		
	Nakanai*	O	36	2			8	1	8	8	1		2		1	5	
	Loso (Nakanai)*	O	15				4		5			1	5				
	Mengen*	O	23				6		9	4			2		2		
	Melamela*	O	14				1		5	5						3	
	Ata*	P	45				13	2	12		2	1	15				
	Kol	P	4				3			1							
East New Britain	Tolai*	O	49	4	2	10	10		8			1	9			5	
	Sulka*	P	33				17		6	4			3	1	2		
	Mali (Baining)*	P	24				1		1	8		6	8				
	Kaket (Baining)*	P	39				2	1	16	9		9	2				
Mussau			20			1	5		10		1		3				
New Hanover	Lavongai*	O	43				5	1				5	30	1	1		
New Ireland																	
New Ireland (O)	Tigak*	O	21				2	3	2		2	1	8	1	2		
	Nalik*	O	17	2			4	1					7	2	1		
	Notsi*	O	14	2			1		4			3	3		1		
	Madak*	O	19				4	1	4			6	3	1			
	Patpatar	O	6	2	1						2		1				
New Ireland (P)	Kuot*	P	32	2			5	1	6			10	7	1			
Bougainville																	
North Bougainville	Saposa*	O	26	4			14		2			1	2		3		
	Teop*	O	18		1		6		1	2			4	1	3		
	Buka*	O	10					1		1			4	1	2	1	
Central Bougainville	Aita*	P	18				3						15				
South Bougainville	Nagovisi		1										1				
	Siwai		2				1						1				
			685	4	17	14	150	41	124	60	24	46	153	10	23	18	1

^a Core populations denoted with *.^b Languages: O-Oceanic, P-Papuan.

to the east as Borneo, and a special branch of C occurs exclusively in Aboriginal Australia. Although C-RPS4Y* (xM38, M217, DYS390.1del) is found in low frequencies across Southeast Asia, Taiwan, Indonesia, and Australia, it has been identified in only a few coastal New Guinea samples and not in our series or further to the east in Remote Oceania. This is an example of a paragroup (designated by an asterisk suffix in the YCC nomenclature), which consists of samples or lineages belonging to a clade, but not assignable to any of its tested subclades.

Lineage C2-M38*(xM208)

Besides the unique Aboriginal Australian lineage shown in table 1, there are separate branches of C that are native to Near Oceania. Although C2-M38* has been identified as far west as Borneo, its major concentration appears to be in eastern Indonesia and coastal New Guinea.

The 17 C2-M38* samples identified in our series were primarily found in New Ireland (and in the Tolai, who migrated from southern New Ireland to East New Britain). One north Bougainville Teop (Oceanic) was also C2-M38*.

Lineage C2b-M208

This subdivision of C has been particularly interesting because it was found in high frequencies in the highlands of West Papua and in 23 of 28 typed Cook Islanders (Kayser et al. 2003). As a result, it was taken as evidence for the Melanesian (New Guinea) origin of most Polynesian males (Kayser et al. 2000, 2003). Surprisingly, only 14 C2b-M208 samples were found in our series, 7 of which were from the Sepik region of New Guinea. All but one of the remainder from our Island Melanesian series was from Oceanic-speaking groups.

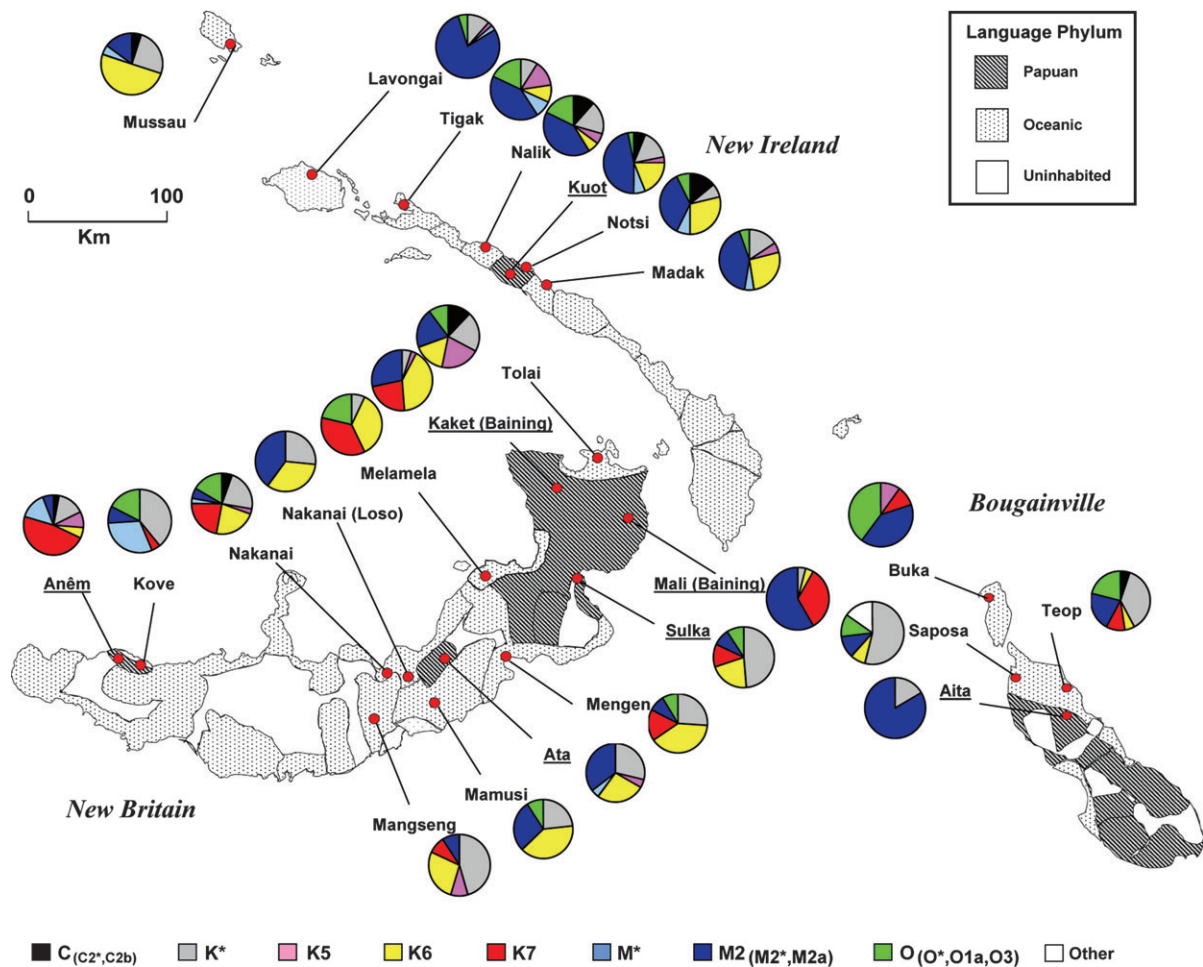


FIG. 2.—NRY haplogroup frequency distributions across 24 Northern Island Melanesian populations. Red nodes represent the location of the sampled populations. Dark shading denotes the Papuan language distributions and lighter shading denotes Oceanic languages.

Lineage C4-P55

This branch was found in one sample in our series from the New Guinea highlands (not shown, but included in diversity calculations).

Lineage K-M9

K-M9* is the residual paragrupo defined by the presence of M9 and the absence of the other haplogroup-defining polymorphisms within it. As shown in figure 1, lineage K was by far the most common macrohaplogroup, but we have been able to detect a number of haplogroups within it.

Most of the populations included in table 1 were tested for K5-M230, but not for the newly defined K6-P79 and K7-P117 haplogroups, which were tested in our series only. Therefore, K-M9* from earlier studies may include representatives of these new haplogroups, but from their very restricted distributions described below, this appears unlikely. In the Southwest Pacific, K-M9* is particularly common in Wallacea, the Trobriand Islands, sections of New Britain and New Ireland, and Fiji.

Lineage K5-M230

This was one of the first polymorphisms found to subclassify K-M9* Y chromosomes in Melanesia (Kayser et al. 2003). It has been found as far west as Bali, where it is rare, in eastern Indonesia (10%–20%), but it reaches its greatest frequency in the highlands of Papua New Guinea (52%), before declining in frequency in Island Melanesia. It has not been found in those Polynesian samples that have been tested (Kayser et al. 2003; Karafet et al. 2005). It therefore appears to be a “Highlands New Guinea” marker with a likely origin there. A large portion of our New Guinea samples (35%) also fall into the K5 lineage (primarily among Sepik men), and it is relatively common among the Tolai and New Ireland populations but almost absent in Bougainville.

Lineage K1-M177

K1-M177 has generally been considered a private polymorphism, only identified in a single Papuan-speaking Nasioi individual collected in a cell line by J Friedlaender from Bougainville (Whitfield et al. 1995). When it was subsequently incorporated in Pacific population surveys, it had

not been found (Cox 2003) until another K1-M177 individual was identified in Malaita, Solomon Islands (Cox and Lahr 2006). We have not been able to identify it in our new Bougainville series, but it should be included in future regional surveys. The 2 cases of K1-M177 are not shown in the tables.

K6-P79

Lineage K6-P79, recently identified, has not been screened widely but was common in our series. It was infrequent along the New Guinea coast but frequent among the Mussau and certain inland New Britain populations: 2 Papuan-speaking (Ata and Kaket Baining) and one inland Oceanic-speaking group (Mamusi). It also had an appreciable frequency in central New Ireland, including the Papuan-speaking Kuot. It was uncommon in Bougainville.

K7-P117

Lineage K7-P117 had a more restricted distribution than K6-P79. Because it was also only recently characterized, it has not been screened outside our series. Except for 3 samples from north Bougainville, it was limited to New Britain populations and undetected in New Ireland. K7-P117 was particularly common in certain Papuan groups. In the Anêm (where K6-P79 was low in frequency), 47% were K7-P117. There was a similar preponderance of K7-P117 to K6-P79 frequencies in the Baining Mali. In contrast, K7-P117 was absent in a number of groups where K6-P79 predominated (the Mamusi, Ata, and Mussau).

50f2/c (data not shown)

We discovered that the 50f2/c deletion occurred in a subset of 3 different haplogroups within the K lineage in our series: in 21% of our K-M9* samples, 14% of our K6-P79 samples, and 5% of the K7-P117 samples. It occurred almost exclusively in New Britain samples. Because of its peculiar distribution and clear recurrence, we have not incorporated the 50f2/c results in further analyses.

Lineage M

The M haplogroup is heavily Near Oceanic in its distribution, found only in a very few samples to the west of the Wallace Line (see table 1). Again, its subdivisions have not been regularly screened, but from our series, its diversity suggests its major expansion (and perhaps origin) in Near Oceania (table 2).

Lineage M-M4(xP87)*

Twenty-four samples were assigned in our series to paralog M-M4*. These were scattered widely from New Guinea (Sepik, Markham Valley, and Eastern Highlands) to New Ireland. Its highest frequency was in the Papuan-speaking Anêm, along with their Kove neighbors in far western New Britain.

Lineage M2-P87(xP22)*

M2-P87* is another recently identified paralog that in this case subdivides the M clade and was found in 46 individuals in our series. It was almost entirely restricted

to New Ireland and east New Britain and was most common in the Papuan-speaking Kuot and both Baining groups there. It was also found in their immediate neighbors in lower frequencies.

Lineage M2a-P22

This lineage was the most frequently occurring haplogroup in our population (153 positive samples). It occurred in particularly high frequencies in the isolated Papuan-speaking Aita of Bougainville (>80%), in New Hanover (70%), and in frequencies ~30% in parts of New Ireland and New Britain (the Papuan-speaking Ata in particular). In contrast, it occurred in only one of the samples from New Guinea.

Lineage O

The O lineage is ubiquitous in East Asia, as table 1 suggests. That O is not more common among some sampled Polynesian populations (i.e., Cook Islands, Western Samoa, Atiu, and Maori) has been interpreted as indicating a relatively small Southeast Asian (and specifically Aboriginal Taiwanese) contribution to their paternal history. However, the sampling in Polynesia has not been particularly comprehensive, and many islands in Remote Oceania have undoubtedly been severely affected by genetic drift. O does have a major presence in Tonga, French Polynesia, and (probably) Western Samoa.

All branches of O were rare in our sample, but when they occurred, they were almost exclusively found in Oceanic-speaking groups. Because there was no great difference in their distributions in our series, the lineages of O will be discussed together.

Three lineages of O were distinguished—Lineage O-M175*, Lineage O1a-M119, and Lineage O3-M122. O-M175* is generally rare throughout Asia, Aboriginal Taiwan, Indonesia, and Melanesia (table 1). O2 is very common in Southeast Asia, especially Bali, but it is almost absent in Melanesia and Polynesia (it has been found in one Western Samoan). O1a-M119 and O3-M122 have been found in frequent and often equal proportions throughout Island Southeast Asia (Capelli et al. 2001). O3-M122 is the most frequent O haplogroup in Polynesia (over 50% in Tonga and 35% in French Polynesia), whereas the other O branches are either rare or lost there. Interestingly, among Taiwan Aborigines, O3-M122 is common only among the Ami, who have been suggested to be the ancestors of all Austronesians outside Taiwan from mitochondrial evidence (Trejaut et al. 2005 and also refer to table 1).

Both O3-M122 and O1a-M119 are surprisingly rare throughout Island Melanesia and New Guinea, and in our series, no O branches were common. Where they were identified, they almost always occurred together in Oceanic-speaking groups. Among the Papuan-speaking series, 1 Kuot was O-M175* and 3 Sulka were either O-M175* or O1a-M119. The Sulka are known to have been heavily influenced linguistically and otherwise by their Oceanic-speaking neighbors (Reesink 2005). Taken together, although the O lineages had their highest frequencies in Bougainville, O3-M122 was found almost exclusively in New Britain, and there it was concentrated in only a few

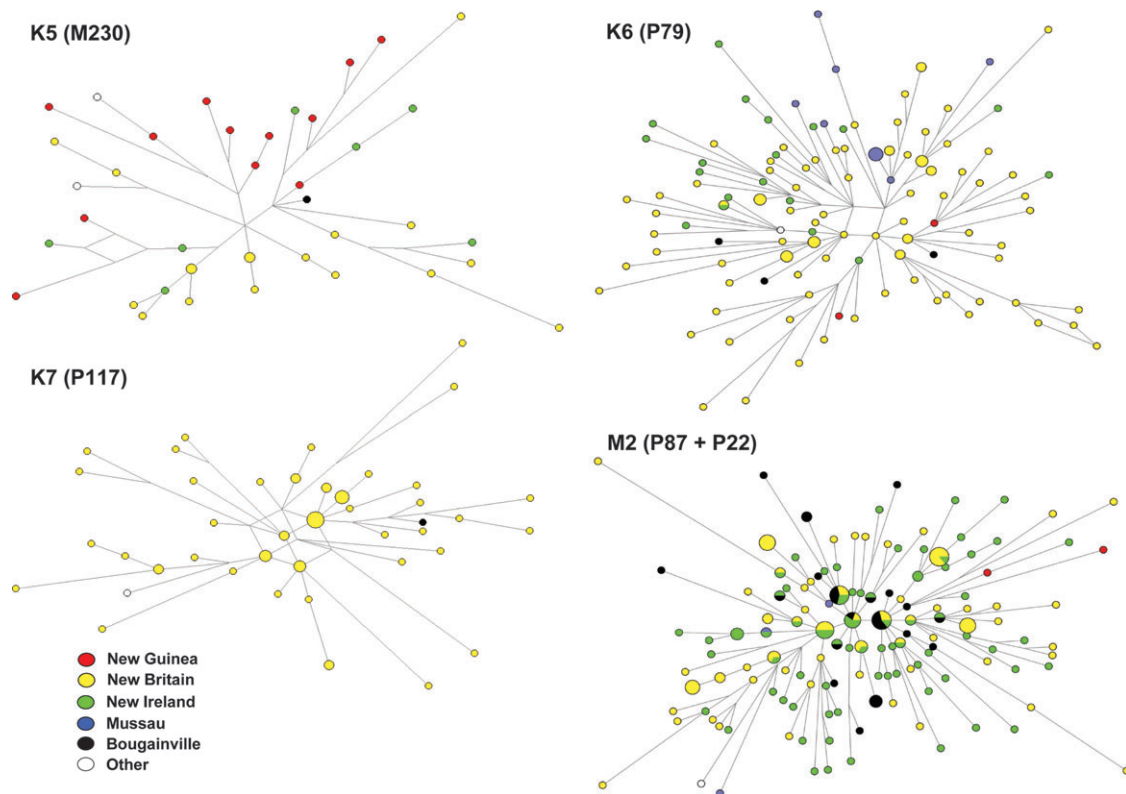


FIG. 3.—Median-joining microsatellite networks for 4 NRY haplogroups in Northern Island Melanesia. Nodes are color coded by island. The nodes are sized according to the number of individuals sharing that haplotype. Nodes consisting of haplotypes not present in the current sample are not shown.

Oceanic-speaking groups along the north coast, which has been thought to be the specific homeland of Lapita.

To summarize the SNP distributions, although some haplogroups were widespread in the general region (K-M9*, K5-M230, M-M4*, and to a lesser extent the C haplogroups), many appeared to be specific to different sections of Northern Island Melanesia, where they appear to have originally developed (K6-P79, K7-P117, M2-P87*, and M2a-P22). This latter set was rare or absent in our (admittedly small) New Guinea series. In contrast, O3-M122 (and secondarily other O lineages) was closely identified with certain Oceanic-speaking groups, but the O lineages were surprisingly scarce, given the preponderance of Oceanic languages in our sample, and also the mtDNA indications of a major presence of the “Polynesian Motif”, B4a1a1 (Friedlaender et al. 2005a, 2005b, forthcoming).

As mentioned, we chose 7 microsatellites for typing because they had been shown to be polymorphic in earlier Pacific studies. Median-joining networks for the microsatellites on each SNP-defined haplogroup background were calculated using NETWORK 4.1.1.2 (Bandelt et al. 1999). Not all networks showed expansions because some haplotypes did not include many individuals, and others clearly included haplogroups that have yet to be identified. Those that were indigenous to Near Oceania, on the other hand, gave clear signals of expansion, and their median networks are shown in figure 3. These are haplogroups K5-M230, K6-P79, K7-P117, and M2-P87 (including M2a). They all show star-like networks associated with population expansions and most have smooth pairwise mismatch dis-

tributions (see Supplemental Figure S1 online). The exception is K7-P117, which is more ragged in its mismatch distribution and may contain sublineages.

Age estimates from microsatellite diversities are problematic as estimates of appropriate mutation rates vary widely from locus to locus, the standard errors are large, and some population split dates used for calibration are open to question, such as the proposed Samoan/Cook/Maori divergence date (i.e., Bonn e-Tamir et al. 2003; Zhivotovsky et al. 2004). One might suspect that O3-M122, for example, should have an age within Island Melanesia of 3200 years ago because it has been closely tied with the Oceanic intrusion along the New Britain north coast. However, higher than expected, microsatellite diversity suggests that the O haplotypes may have already diversified to some degree before the migration from Southeast Asia. Nevertheless, the relative haplogroup divergence age ranking appears clear (see table 3).

All the native Near Oceanic haplogroups (M-M4, M2-P87, M2a-P22, K5-M230, K6-P79, and K7-P117) have essentially equivalent divergence estimates, given their relatively large standard errors. Using the mutation rate estimate of Zhivotovsky et al. (2004), this suggests that they all developed there some time roughly between 45 000 and 30 000 YBP. This compares well with calculations for a number of ancient mtDNA expansions specific to the region (Friedlaender et al. 2005b; Merriwether et al. 2005). It is also consistent with the currently accepted earliest settlement dates for the region, of ~42 000 YBP (Groube 1986; Leavesley et al. 2002).

Table 3
Y Chromosome Haplogroup Age Estimates from STRs

Haplogroup	<i>n</i>	ASD ^a	TMRCAs ^b (years)	Lower Limits ^c (years)	Upper Limits ^c (years)
C2	31	1.369	49 600	42 000	61 000
C2b	14	1.276	46 200	39 000	57 000
K5	41	0.925	33 500	28 000	41 000
K6	124	1.025	37 100	31 000	46 000
K7	60	0.978	35 400	30 000	44 000
M	223	1.052	38 100	32 000	47 000
M2	199	0.904	32 700	28 000	40 000
M2a	150	0.925	33 500	28 000	41 000
O	51	0.831	30 100	25 000	37 000
O3	18	0.532	19 300	16 000	24 000

^a Average square distance (ASD) was calculated using the program Ytime (Behar et al. 2003).

^b Time of the most recent ancestor (TMRCAs) estimates were calculated using ASD with the mutation rate of $(2.8 \pm 0.5) \times 10^{-5}$ per locus per year (Zhivotosky et al. 2004).

^c Limits were calculated based on the standard error from the above mutation rate.

As described, haplogroups C2* and C2b developed earlier outside this region, and their divergence estimates are consistent with this argument. The most recent divergence times belong to haplogroup O which is dated to approximately 30 000 YBP: the small set of O3-M122 microsatellite samples (which are of Southeast Asian origin) yielded an estimated divergence date somewhat younger than 20 000 YBP. Note also that M2 appears to be slightly younger than M2a, an apparent contradiction. The 2 mutations appear to have arisen almost simultaneously, the 2 lineages share a great many haplotypes, and the associated standard errors are very large, in any case.

The measure of SNP and STR diversity calculated for the populations in our series where $n > 9$ is Nei's \hat{H} and average π_n , respectively (Nei 1987, p. 180), calculated using Arlequin 3.0 (Excoffier 2005). These values and their standard deviations are in table 4. The most homogeneous group (the lowest \hat{H} and average π_n values) was the Papuan-speaking Aita of Bougainville, who used to be very isolated and subject to considerable genetic drift, whereas the 2 most heterogeneous populations by both measures were the Notsi and Tolai. Otherwise, there was no apparent pattern or trend of heterogeneity to the distribution of \hat{H} and average π_n within the sample series, especially given their relatively large standard deviations. Comparing values of \hat{H} and average π_n across studies, or simply values of segregating sites, is difficult because SNP and STR panels differ so widely, and in our case, we had added 3 regionally polymorphic SNPs.

In order to see if the pattern of variation clustered according to an island-by-island or language family affiliation, we considered both factors in our AMOVA, the results of which are shown in table 5. For this analysis, we included only those populations with 10 or more individuals and that were well defined (this excluded our heterogeneous samples from New Guinea). Mussau with a single population was also excluded. This left 23 populations for the AMOVA.

With these populations, we found well over 80% of the SNP and STR variance to be within the different populations in the series, but still this left a considerable proportion of among-group variance to be explained. Partitioning the populations into 3 island sets (New Britain, New Ireland,

and Bougainville) produced both significant ($P < 0.01$) among- and within-island variance components for both the SNP and STR results. As for the effect of island land-mass on internal variation, which would be predicted with a simple biogeographical model, New Ireland, which is long and narrow, did have the lowest within-island, among-group, variation with regard to both SNPs and STRs. New Britain, the largest island, only showed a modest increase in within-island, among-group, variation for SNPs but did show a clear increase in STR diversity. The 4 Bougainville populations produced a strikingly smaller SNP within-population variance component than either of the other 2 islands. This was due to the uniformity of the Aita population there (only 2 haplogroups present) and also the relatively small number of other Bougainville populations sampled.

With regard to partitioning the variance by language, there was no simple contrast between the major language divisions represented by the populations. Applying the AMOVA to the populations partitioned by major language affiliation (Oceanic vs. Papuan) produced a nonsignificant variance component ($P > 0.40$) between language groups for both STRs and SNPs. The Papuan populations had a somewhat higher among-group variance than did the Oceanic groups, which is not surprising, because they are generally more isolated and retain more ancient population signatures.

As a corollary to the AMOVA analysis, MDS was used to provide a 2-dimensional representation of the pairwise population F_{ST} values (fig. 4). The plot has a very low stress level (0.01), indicating it is a reliable representation of the pairwise distances. The main features of the MDS display are: New Britain populations generally fall to one side of the plot, whereas New Ireland populations fall to the other; the 4 Bougainville populations are scattered throughout (consistent with the AMOVA results); the Papuan Aita are the most removed group, followed by the Lavongai; none of the Papuan-speaking groups are particularly closely paired, even those considered linguistically closely related (i.e., the Ata and Anêm, and the 2 Baining languages, Kaket and Mali); and geographic proximity, although important overall, does not seem to determine the closest pairwise relationships.

Table 4
Population Diversities in Island Melanesia

Island	Region	Population ^a	Language ^b	N	Standard Gene Diversity (SNP)		Average Gene Diversity (STR)	
					\hat{H}	SD	Average π_n	SD
New Guinea								
	PNG Coast	North Coast		25	0.823	0.050	0.679	0.383
		Markham		1	—	—	—	—
		Rigo		2	—	—	—	—
	PNG Highlands	Eastern Highlands		4	—	—	—	—
		Morobe Highlands		1	—	—	—	—
		Western Highlands		2	—	—	—	—
	PNG Island	Misima		2	—	—	—	—
Manus				7	—	—	—	—
New Britain								
	West New Britain	Kove*	O	24	0.797	0.050	0.553	0.321
		Anêm*	P	34	0.742	0.064	0.485	0.283
		Mangseng*	O	11	0.764	0.107	0.681	0.406
		Mamusi*	O	43	0.732	0.035	0.579	0.327
		Nakanai*	O	36	0.848	0.028	0.621	0.349
		Loso (Nakanai)*	O	15	0.752	0.056	0.652	0.380
		Mengen*	O	23	0.767	0.054	0.585	0.337
		Melamela*	O	14	0.747	0.066	0.502	0.305
		Ata*	P	45	0.747	0.029	0.619	0.346
		Kol	P	4	—	—	—	—
	East New Britain	Tolai*	O	49	0.855	0.018	0.697	0.384
		Sulka*	P	33	0.695	0.071	0.659	0.369
		Mali (Baining)*	P	24	0.743	0.043	0.528	0.308
		Kaket (Baining)*	P	39	0.738	0.042	0.576	0.327
Mussau				20	0.695	0.081	0.607	0.351
New Hanover		Lavongai*	O	43	0.496	0.085	0.543	0.310
New Ireland								
	New Ireland (O)	Tigak*	O	21	0.833	0.065	0.645	0.369
		Nalik*	O	17	0.787	0.075	0.659	0.381
		Notsi*	O	14	0.857	0.056	0.752	0.434
		Madak*	O	19	0.825	0.048	0.634	0.366
		Patpatar	O	6	0.000	0.000	0.000	0.000
	New Ireland (P)	Kuot*	P	32	0.815	0.035	0.671	0.375
Bougainville								
	North Bougainville	Saposa*	O	26	0.686	0.087	0.641	0.363
		Teop*	O	18	0.837	0.057	0.626	0.372
		Buka*	O	10	0.844	0.103	0.523	0.339
	Central Bougainville	Aita*	P	18	0.294	0.119	0.444	0.270
	South Bougainville	Nagovisi		1	—	—	—	—
		Siwai		2	—	—	—	—
				685				

^a Core populations denoted with *.^b Languages: O-Oceanic, P-Papuan.

Discussion

The newly identified regionally specific SNPs have greatly enhanced our ability to identify the extent and pattern of NRY variation in Northern Island Melanesia. Earlier surveys concluded that populations there were comparatively low in NRY SNP heterozygosity but that now appears to have been caused by an ascertainment bias of SNPs (i.e., earlier batteries lacked a number of important regionally informative SNP markers and also there was a general undersampling of these diverse populations). To judge by their microsatellite diversity, all these lineages are of considerable age, with divergence estimates roughly in the 30–45,000 YBP range for those that developed in the region. These estimates are compatible with those from studies of mtDNA and X variation in the region (Friedlaender et al. 2005b; Merriwether et al. 2005; Wilder and Hammer forthcoming) and are accommodated by the

established earliest settlement dates for the region. Even the earliest archaeological dates for Buka/Bougainville, which are the youngest for these 3 islands (29,000 YBP), falls very close to the estimated date for the origin of M2a-P22, which is the most common of the haplogroups in our sample.

Although the AMOVA results showed that the 23 core populations had large within-population SNP and STR variance components (84.1% and 87.1%, respectively), this still left an important among-population variance proportion to explain. Partitioning by major language group produced an insignificant result, and although island-by-island partitioning did indicate some structure along this line, we could not discern a special effect of island size. This result is consistent with limited, but persistent, short-range gene flow and population movement that has blurred distinctions between neighboring groups but has been insufficient to homogenize variation across the region. Similarly, related

Table 5a
AMOVA Based on Y-Chromosomal Haplogroups (F_{ST})

Group	<i>n</i>	No. of Populations	No. of Groups	Variance Components (%) [*]		
				Between Groups	Within Groups	Within Populations
No grouping (23 populations)	608	23	1	—	15.9	84.1
Geography (3 islands) ^{a,b}	608	23	3	8.4	10.7	80.9
Language (2 groups) ^c	608	23	2	<i>-0.6</i>	16.3	84.3
New Ireland	146	6	1	—	9.4	90.6
New Britain	390	13	1	—	9.7	90.3
Bougainville	72	4	1	—	27.5	72.5
Papuans	225	7	1	—	18.6	81.4
Oceanians	383	16	1	—	14.8	85.2

^a Groups for geography: New Ireland, New Britain, and Bougainville.

^b For this AMOVA analysis, Lavongai was included with New Ireland.

^c Linguistic groups: Papuans and Oceanians.

^{*} All *P* values are <0.01 except for the linguistic groups where *P* = 0.48 and the nonsignificant value is shown in italics.

study of the relationship of NRY, mtDNA, autosomal STR, and linguistic variation with geography across this region shows a poor correspondence with either an isolation by distance or a population fissions model of differentiation (Hunley et al. forthcoming).

It appears that each of the autochthonous NRY haplogroups has its highest concentration in a Papuan-speaking region but has diffused from that center across neighboring languages so that the original specific haplogroup—Papuan language association has been blurred. We can still identify likely expansion centers for certain regional variants. New Guinea is the likely origin for K5-M230 and M-M4. Different populations of New Britain were the origins of K6-P79, K7-P117, and M2a-P22. K6-P79 has also dispersed to New Ireland and Mussau, consistent with its lower microsatellite diversity there. M2-P87 likely developed in New Ireland, where the Papuan-speaking Kuot appears to be the epicenter. The Aita of north Bougainville are at the current focus for M2a-P22. It is intriguing that the Lavongai populations just to the north of New Ireland also have very high M2a-P22 frequencies.

If they are incorporated into screens in Remote Oceania (i.e., Southern Island Melanesia, Polynesia, and Micronesia), the newly described NRY haplogroup markers should, along with specific microsatellite associations, help

Table 5b
AMOVA Based on Y-Chromosomal STRs (R_{ST})

Group	Variance Components (%) [*]		
	Between Groups	Within Groups	Within Populations
No grouping (23 populations)	—	12.9	87.1
Geography (3 islands) ^{a,b}	9.7	7.1	83.2
Language (2 groups) ^c	<i>-1.1</i>	13.5	87.6
New Ireland	—	4.3	95.7
New Britain	—	9.3	90.7
Bougainville	—	5.3	94.7
Papuans	—	14.8	85.2
Oceanians	—	12.6	87.4

^a Groups for geography: New Ireland, New Britain, and Bougainville.

^b For this AMOVA analysis, Lavongai was included with New Ireland.

^c Linguistic groups: Papuans and Oceanians.

^{*} All *P* values are <0.01 except for the linguistic groups where *P* = 0.84 and the nonsignificant value is shown in italics.

in identifying particular contributions from sections of Northern Island Melanesia and New Guinea to populations in those regions. We know that M-M4* occurs in Vanuatu in high frequency, as well as Fiji and Tonga (table 1). K-M9* is common in Vanuatu and Fiji, as well as occurring in low frequencies in all Polynesian series to date. Therefore, testing for K6-P79, K7-P117, M2-P87, and M2a-P22 in these series will help to illuminate their population histories.

Any recent Southeast Asian component (associated with the development of Lapita) is most certainly associated with haplogroup O3-M122, which has a remarkably low frequency in this sample series (18 of 685 samples or 2.6%). Even if all O haplogroups might have been introduced just with the immediate Southeast Asian ancestral component of the Lapita peoples, their frequency in our series totals just 8% (table 2), almost entirely restricted to Oceanic-speaking populations. In our sampling strategy, we may have missed higher concentrations of the O haplogroups. For example, many Lapita sites in the Bismarck Archipelago are located on smaller or offshore islands and we did not sample those, except for Mussau (where no O samples were found). It is also possible that most of the “Southeast Asian” male component either moved on to other islands or was lost/obliterated in this region.

The clear deduction from the NRY data is that the Southeast Asian male contribution to contemporary Oceanic-speaking populations in Northern Island Melanesia is unexpectedly small. This contrasts sharply with the accepted interpretation of the mtDNA haplogroup frequency data in this region because of the very high frequency of the so-called “Polynesian Motif” (B4a1a), even in many Papuan-speaking groups, where it approaches a frequency of 0.90 (Merriwether et al. 1999; Friedlaender et al. 2005a, forthcoming). Explanatory scenarios of different male versus female migratory behavior or sexual selection have been proposed to resolve this discrepancy in Remote Oceanic populations (Hage and Marck 2003; Cann and Lum 2004), but the contradictions in the Northern Island Melanesian comparisons only serve to underline the contrast and beg for some other explanation. We suggest that because both the mtDNA and NRY have a fraction of the reproductive sample sizes of autosomal markers, their haplotypes are especially likely to experience pronounced

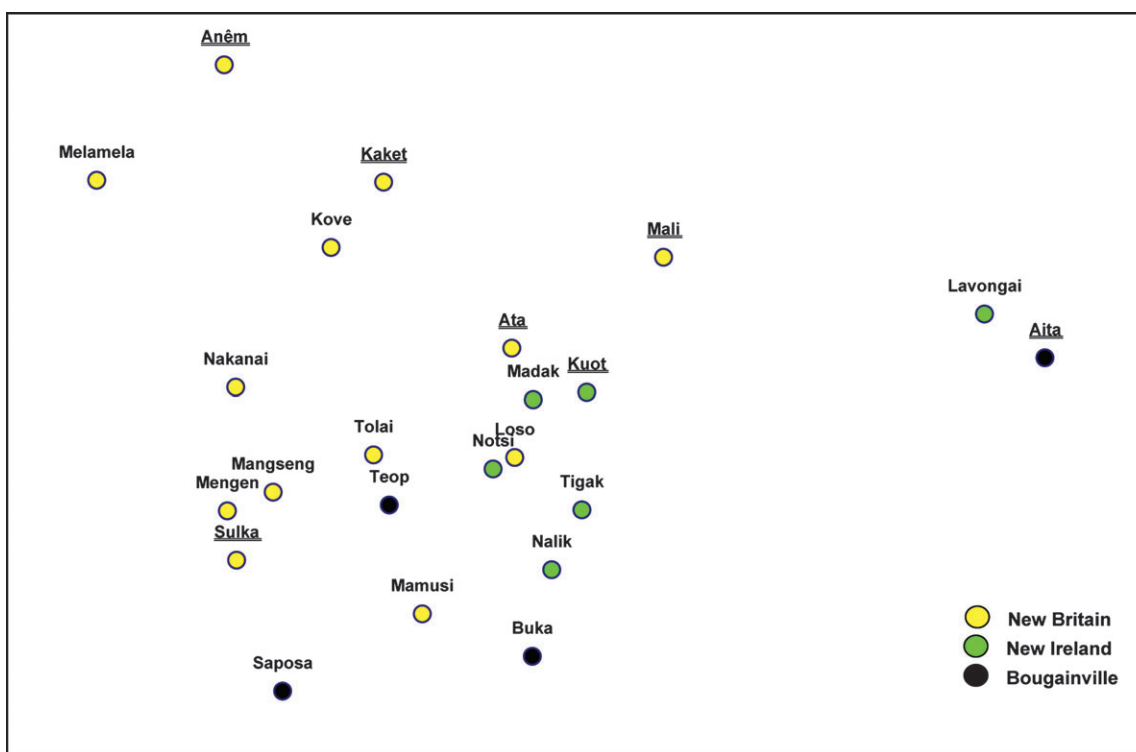


FIG. 4.—Two-dimensional MDS plot generated from pairwise SNP F_{ST} values. Twenty-three core populations are shown. Papuan-speaking populations are underlined.

fluctuations from generation to generation (i.e., extreme genetic drift) in the very small populations that inhabited this entire region. The point is that neither the mtDNA nor the NRY-based interpretations of Aboriginal Taiwanese/Southeast Asian contributions to Oceanic- or to Papuan-speaking populations may be particularly accurate. They are likely to be rather rough (though very informative) approximations of the real male and female genetic contributions to contemporary populations across the Southwest Pacific. They are, after all, only 2 different, though highly variable, “loci.” The analysis of a large battery of independent autosomal loci will undoubtedly help resolve these contradictory interpretations in a more satisfactory manner.

Supplementary Material

Supplementary Figure S1 is available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

Acknowledgments

We are extremely grateful and indebted to the people of the Bismarck Archipelago and Bougainville who willingly participated in this study and often acted as our hosts. We hope that it contributes to a better understanding of their population histories, which is our stated objective. We thank William Beggs, Patrick Bender, Charles Mgone, Heather Norton, Daniel Hrdy, and Andrew Merriwether for their assistance and participation. The laboratory aspect of this research was supported by a Wenner-Gren Founda-

tion predoctoral Grant, whereas the fieldwork was supported by postdoctoral grants from the National Science Foundation, the Wenner-Gren Foundation for Anthropological Research, and the National Geographic Society Research Fund. The materials presented in this paper will appear in similar form as chapter 5 of the forthcoming book, *Genes, Language, and Cultural History in the Southwest Pacific* edited by JS Friedlaender, Oxford University Press, New York.

Literature Cited

- Allen J. 2003. Discovering the Pleistocene in Island Melanesia. In: Sand C, editor. Pacific archaeology: assessments and anniversary of the first Lapita excavation (July 1952) Koné-Nouméa, 2002. Volume 15. Nouméa, New Caledonia: Le Cahiers de l'Archéologie en Nouvelle-Calédonie. p 33–42.
- Bandelt H-J, Forster P, Rohlf A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37–48.
- Behar DM, Thomas MG, Skorecki K, et al. (11 co-authors). 2003. Multiple origins of Ashkenazi Levites: Y chromosome evidence for both Near Eastern and European ancestries. *Am J Hum Genet* 73:768–79.
- Bonné-Tamir B, Korostishevsky M, Redd AJ, Pel-Or Y, Kaplan ME, Hammer MF. 2003. Maternal and paternal lineages of the Samaritan isolate: mutation rates and time to most recent common male ancestor. *Ann Hum Genet* 67:153–64.
- Cann RL, Lum JK. 2004. Dispersal ghosts in Oceania. *Am J Hum Biol* 16:440–51.
- Capelli C, Wilson JF, Richards M, Stumpf MP, Gratrix F, Oppenheimer S, Underhill P, Pascali VL, Ko TM, Goldstein DB. 2001. A predominantly indigenous paternal heritage for

- the Austronesian-speaking peoples of insular Southeast Asia and Oceania. *Am J Hum Genet* 68:432–43.
- Cordaux R, Stoneking M. 2003. South Asia, the Andamanese, and the genetic evidence for an “early” human dispersal out of Africa. *Am J Hum Genet* 72:1586–90.
- Cox MP, Lahr MM. 2006. Y-chromosome diversity is inversely associated with language affiliation in paired Austronesian- and Papuan-speaking communities from Solomon Islands. *Am J Hum Biol* 18:35–50.
- Cox MP. 2003. Genetic patterning at Austronesian contact zones [doctoral dissertation]. Dunedin, New Zealand: University of Otago.
- Dunn M, Terrill A, Reesink G, Foley R, Levinson SC. 2005. Structural phylogenetics and the reconstruction of ancient language history. *Science* 309:2072–5.
- Dunn M, Terrill A, Reesink G. 2002. The East Papuan languages: a preliminary typological appraisal. *Ocean Ling* 41:28–62.
- Excoffier L, Schneider S. 2005. Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47–50.
- Forster P, Kayser M, Meyer E, Roewer L, Pfeiffer H, Benkmann H, Brinkmann B. 1998. Phylogenetic resolution of complex mutational features at Y-STR DYS390 in aboriginal Australians and Papuans. *Mol Biol Evol* 15:1108–14.
- Friedlaender J, Schurr T, Gentz F, et al. (13 co-authors). 2005a. Expanding Southwest Pacific mitochondrial haplogroups P and Q. *Mol Biol Evol* 22:1506–17.
- Friedlaender JS, Friedlaender FR, Hodgson J, McGrath JS, Stoltz M, Koki G, Schurr TG, Merriwether DA. Mitochondrial DNA variation in Northern Island Melanesia. In: Friedlaender JS, editor. *Genes, language, and culture in the Southwest Pacific*. New York: Oxford University Press. Forthcoming.
- Friedlaender JS, Gentz F, Friedlaender F, Kaestle F, Schurr TG, Koki G, Mgone C, McDonough J, Smith L, Merriwether DA. 2005b. Mitochondrial genetic diversity and its determinants in Island Melanesia. In: Pawley A, Attenborough R, Golson J, Hyde R, editors. *Papuan pasts: studies in the cultural, linguistic and biological history of the Papuan speaking peoples*. Canberra: Pacific Linguistics. p 693–716.
- Gray RD, Jordan FM. 2000. Language trees support the express-train sequence of Austronesian expansion. *Nature* 405:1052–5.
- Green RC. 1997. Linguistic, biological, and cultural origins of the initial inhabitants of Remote Oceania. *N Z J Archaeol* 17:5–27.
- Green RC. 2003. The Lapita horizon and traditions—signature for one set of oceanic migrations. In: Sand C, editor. *Pacific archaeology: assessments and anniversary of the first Lapita excavation (July 1952) Koné-Nouméa, 2002*. Volume 15. Nouméa, New Caledonia: Le Cahiers de l’Archéologie en Nouvelle-Calédonie. p 95–120.
- Groube L. 1986. Waisted axes of Asia, Melanesia, and Australia. In: Ward GK, editor. *Archaeology at ANZAAS*. Canberra: Canberra Archaeological Society. p 168–77.
- Hage P, Marck JC. 2003. Matrilineality and the Melanesian origin of Polynesian Y chromosomes. *Curr Anthropol* 44:S121–7.
- Hunley K, Dunn M, Lindström E, et al. (11 co-authors). Inferring prehistory from genetic, linguistic, and geographic variation. In: Friedlaender JS, editor. *Genes, language, and culture in the Southwest Pacific*. New York: Oxford University Press. Forthcoming.
- Ingman M, Gyllensten U. 2003. Mitochondrial genome variation and evolutionary history of Australian and New Guinean aborigines. *Genome Res* 13:1600–6.
- Jobling M, Samara V, Pandya A, et al. (16 co-authors). 1996. Recurrent duplication and deletion polymorphisms on the long arm of the Y chromosome in normal males. *Hum Mol Genet* 5:1767–75.
- Jobling MA, Tyler-Smith C. 2003. The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet* 4:598–612.
- Karafet TM, Lansing JS, Redd AJ, et al. (11 co-authors). 2005. Balinese Y-chromosome perspective on the peopling of Indonesia: genetic contributions from pre-neolithic hunter-gatherers, Austronesian farmers, and Indian traders. *Hum Biol* 77:93–114.
- Kayser M, Brauer S, Weiss G, Schiefenhover W, Underhill P, Shen P, Oefner P, Tommaseo-Ponzetta M, Stoneking M. 2003. Reduced Y-chromosome, but not mitochondrial DNA, diversity in human populations from West New Guinea. *Am J Hum Genet* 72:281–302.
- Kayser M, Brauer S, Weiss G, Underhill PA, Roewer L, Schiefenhover W, Stoneking M. 2000. Melanesian origin of Polynesian Y chromosomes. *Curr Biol* 10:1237–46.
- Kayser M, Krawczak M, Excoffier L, et al. (12 co-authors). 2001. An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations. *Am J Hum Genet* 68:990–1018.
- Leavesley M, Chappell J. 2004. Buang Merabak: additional early radiocarbon evidence of the colonisation of the Bismarck Archipelago, Papua New Guinea. *Antiquity* 78(301): Project gallery.
- Leavesley MG, Bird MI, Fifield LK, Hausladen PA, Santos GM, di Tada ML. 2002. Buang Merabak: early evidence for human occupation in the Bismarck Archipelago, Papua New Guinea. *Aust Archaeol* 54:55–7.
- Merriwether DA, Friedlaender JS, Mediavilla J, Mgone C, Gentz F, Ferrell RE. 1999. Mitochondrial DNA variation is an indicator of Austronesian influence in Island Melanesia. *Am J Phys Anthropol* 110:243–70.
- Merriwether DA, Hodgson JA, Friedlaender FR, Allaby R, Cerchio S, Koki G, Friedlaender JS. 2005. Ancient mitochondrial M haplogroups identified in the Southwest Pacific. *Proc Natl Acad Sci USA* 102:13034–9.
- Nei M. 1987. *Molecular evolutionary genetics*. New York: Columbia University Press.
- Pavlidis C, Gosden C. 1994. 35,000 year-old sites in the rainforests of West New Britain, Papua New Guinea. *Antiquity* 69:604–10.
- Pawley A, Green RC. 1973. Dating the dispersal of Oceanic languages. *Ocean Ling* 12:1–67.
- Reesink G. 2005. Sulka of East New Britain: a mixture of Oceanic and Papuan traits. *Ocean Ling* 44:145–93.
- Ross M. 2001. Is there an East Papuan Phylum? Evidence from pronouns. In: Pawley A, Ross M, Tryon D, editors. *The boy from Bundaberg: studies in Melanesian linguistics in honour of Tom Dutton*. Canberra: Pacific Linguistics. p 301–21.
- Ross M. 2005. Pronouns as a preliminary diagnostic for grouping Papuan languages. In: Pawley A, Attenborough R, Golson J, Hyde R, editors. *Papuan pasts: investigations into the cultural, linguistic and biological history of the Papuan speaking peoples*. Canberra: Pacific Linguistics.
- Schneider S, Roessli D, Excoffier L. 2000. Arlequin, version 2000; a software for population genetics data analysis. Geneva: University of Geneva Press.
- Shutler RJ, Marck JC. 1975. On the dispersal of the Austronesian horticulturalists. *Archaeol Phys Anthropol Ocean* 10:81–113.
- Spriggs MJT. 1997. *The Island Melanesians*. Cambridge, MA: Blackwell Scientific.
- Su B, Jin L, Underhill P, et al. (11 co-authors). 2000. Polynesian origins: insights from the Y chromosome. *Proc Natl Acad Sci USA*. 97:8225–8.

- Summerhayes GR, Allen J. 1993. The transport of Mopir obsidian to Late Pleistocene New Ireland. *Archaeol Ocean* 28:145–9.
- Trejaut JA, Kivisild T, Loo JH, Lee CL, He CL, Hsu CJ, Li ZY, Lin M. 2005. Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations. *PLoS Biol* 3:e247.
- Underhill PA. 2004. A synopsis of extant Y chromosome diversity in East Asia and Oceania. In: Sagart L, Blench R, Sanchez-Mazas A, editors. *The peopling of East Asia: putting together archaeology, linguistics and genetics*. London: Routledge Curzon. p 301–19.
- Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley RA, Oefner PJ, Cavalli-Sforza LL. 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet* 65:43–62.
- Underhill PA, Shen P, Lin AA, et al. (21 co-authors). 2000. Y chromosome sequence variation and the history of human populations. *Nat Genet* 26:358–61.
- Whitfield LS, Sulston JE, Goodfellow PN. 1995. Sequence variation of the human Y chromosome. *Nature* 378:379–80.
- Wilder JA, Hammer MF. Extraordinary population structure among the Baining of New Britain. In: Friedlaender JS, editor. *Genes, language, and culture in the Southwest Pacific*. New York: Oxford University Press. Forthcoming.
- Wurm S. 1975. The East Papuan Phylum in general. In: Wurm S, editor. *Papuan languages and the new Guinea linguistic scene*. Research School of Pacific Studies, Department of Linguistics, Canberra: Australian National University. p 783–804.
- Zhivotovsky LA, Underhill PA, Cinnioglu C, et al. (17 co-authors). 2004. The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *Am J Hum Genet* 74:50–61.

Hope Hollocher, Associate Editor

Accepted May 26, 2006