Original Research Article

Y-Chromosome Diversity Is Inversely Associated With Language Affiliation in Paired Austronesian- and Papuan-Speaking Communities from Solomon Islands

MURRAY P. COX* AND MARTA MIRAZÓN LAHR
Leverhulme Centre for Human Evolutionary Studies, Department of Biological Anthropology, University of Cambridge, Downing Street, Cambridge, United Kingdom

ABSTRACT   The Solomon Islands lie in the center of Island Melanesia, bordered to the north by the Bismarck Archipelago and to the south by Vanuatu. The nation's half-million inhabitants speak around 70 languages from two unrelated language groups: Austronesian, a language family widespread in the Pacific and closely related to languages spoken in Island Southeast Asia, and "East Papuan", generally defined as non-Austronesian and distantly related to the extremely diverse Papuan languages of New Guinea. Despite the archipelago's presumed role as a staging post for the settlement of Remote Oceania, genetic research on Solomon Island populations is sparse. We collected paired samples from two regions that have populations speaking Austronesian and Papuan languages, respectively. Here we present Y-chromosome data from these samples, the first from Solomon Islands. We detected five Y-chromosome lineages: M-M106, O-M175, K-M9*, K-M230, and the extremely rare clade, K1-M177. Y-chromosome lineages from Solomon Islands fall within the range of other Island Melanesian populations but display markedly lower haplogroup diversity. From a broad Indo-Pacific perspective, Y-chromosome lineages show partial association with the distribution of language groups: O-M175 is associated spatially with Austronesian-speaking areas, whereas M-M106 broadly correlates with the distribution of Papuan languages. However, no relationship between Y-chromosome lineages and language affiliation was observed on a small scale within Solomon Islands. This pattern may result from a sampling strategy that targeted small communities, where individual Y-chromosome lineages can be fixed or swept to extinction by genetic drift or favored paternal exogamy. Am. J. Hum. Biol. 18:35–50, 2006.   © 2005 Wiley-Liss, Inc.

The Solomon Islands archipelago comprises a string of seven main and innumerable smaller islands lying just south of the equator in the Western Pacific Ocean (Fig. 1). To the northwest, the islands of New Britain and New Ireland reach out toward the major landmass of New Guinea. Together with the Solomons archipelago, these island clusters form a zone of cultural interaction known as Near Oceania. To the southeast, a sizable ocean gap separates the main Solomons chain from Remote Oceania, a region that encompasses the archipelagos of Reefs–Santa Cruz, Vanuatu, New Caledonia, and Fiji, eventually extending as far as the small and widely dispersed islands of Polynesia.

The major islands of Near Oceania were settled during the late Pleistocene (<45,000 BP; O'Connell and Allen, 2004), requiring ocean crossings even at periods of low sea level during glacial maxima. Human popula-

tions reached Buka, an island just north of Bougainville, by 29,000 BP (Wickler and Spriggs, 1988). The earliest radiocarbon date from Solomon Islands is relatively recent by comparison; the cave site of Vatuluma Posovi on Guadalcanal, just west of Honiara, is tenta-
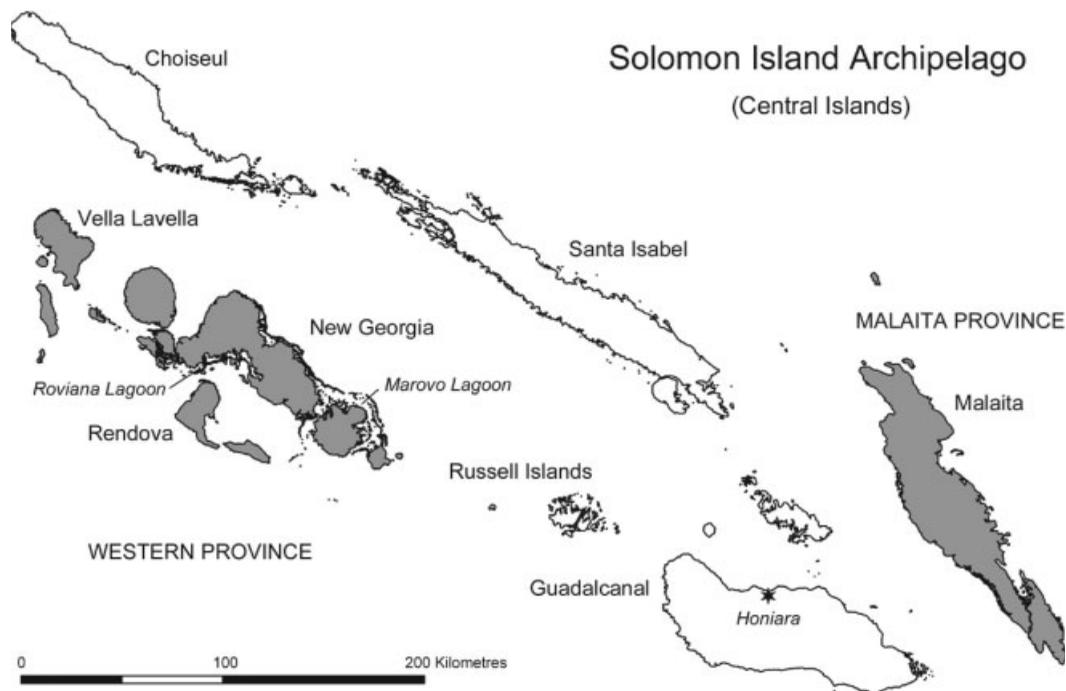
Fig. 1.   Geography of central Solomon Islands identifying major sampling locations.

tively dated to 6,450–5,950 cal. BP (Roe, 1993:176). However, bathymetric readings between islands from Buka to the Russell Islands, just north of Guadalcanal, are suggestive of a contiguous landmass during the last glacial maximum (28,000–18,000 BP). The island groups of New Georgia, Malaita, and San Cristobal were never interconnected with the main Solomon's chain, but sea crossings to these groups were shorter than those overcome previously to reach Bougainville from New Britain or New Ireland. Consequently, the whole of the main Solomons chain may well have been settled, at least in places, since the late Pleistocene.

Conversely, the barrier to Remote Oceania south of the main Solomon Islands chain was breached only around 3,000 BP (Bedford et al., 1998), when new technologies associated with the Lapita cultural complex first appear in the Reefs–Santa Cruz and Vanuatu archipelagos. The Lapita culture is widely regarded as a mark of incursive populations ultimately deriving from Mainland Asia (Bellwood, 1997, 2001, 2005; Cox, 2003; Spriggs, 2000), but modified and enlarged by contact with pre-existing groups during their advance through

Island Southeast Asia and out into the Pacific (cf. the intrusion/innovation/integration model of Green, 1991). Early Lapita pottery has not been found in the Solomons, although late Lapita sites occur intermittently along the coasts of New Georgia, possibly dating to around 2,700 BP (Felgate and Dickinson, 2001; Summerhayes and Scales, 2005). The lack of early Lapita sites is curiously at variance with the vast and rapid spread of the Lapita culture in the Reefs–Santa Cruz group and Vanuatu, just to the south of the Solomon Islands (Bedford, 2003). However, sudden sediment deposition, potentially associated with agricultural land use, has been detected in the Roviana Lagoon region of New Georgia after 3,210–3,330 cal. BP (Grimes, 1998) and may be an alternative marker of early Lapita settlement. Whether the Solomon Islands were ever settled early in the Lapita period remains an open question (Summerhayes and Scales, 2005), but in general, the Lapita impact there seems distorted relative to allied trajectories in the Bismarck and Vanuatu archipelagos.

Fortunately, this expansion can be approached from a different perspective: it

remains the most credible underlying cause of the spread of Austronesian languages across a vast swathe of the Indo-Pacific world (Bellwood, 2005). Austronesian is the sole language group documented in Remote Oceania during historic times (excluding the Reefs–Santa Cruz group), and it accounts for more than three-quarters of the languages spoken along the main Solomons chain today (Grimes, 2005). The close similarity and numerical dominance of Austronesian languages in the Indo-Pacific region are suggestive of a major population expansion during the Holocene (<10,000 BP). Conversely, the diverse and unrelated Papuan languages occur widely on mainland New Guinea but are strewn only intermittently across the islands of Near Oceania, including 23 languages in the Solomon Islands and Reefs–Santa Cruz archipelagos (Grimes, 2005). These populations may represent relict groups ultimately derived from the original Pleistocene settlers of Near Oceania (whether from within the Solomons or elsewhere in eastern Near Oceania), whereas Austronesian-speaking populations would have histories more firmly embedded in a late Holocene expansion from Mainland Asia. However, populations have always interacted in Island Melanesia, and groups speaking Papuan languages can move as easily as those speaking Austronesian tongues (Friedlaender et al., 2002; Spriggs, 1997:174). Evidence of widespread language exchange (Friedlaender, 1975; Ross, 1994) cautions that ethnolinguistic hypotheses very much remain concepts to be tested.

Despite its key geographical role as a staging post for the first settlement of Remote Oceania and despite representing a famous early period of ethnographic research (Rivers, 1914), the Solomon Islands can boast little anthropological investigation today. Genetic research is particularly sparse. Although the subject of limited blood group surveys from the 1950s (Blake et al., 1983; Douglas et al., 1962; Walsh and Kooptzoff, 1955) and during the 1980s the focus of a major research initiative led by Jonathan Friedlaender (1987, 1990), surprisingly little is known about population genetic structuring among Solomon Island peoples. Blood-group studies class Solomon Island populations with other Melanesian groups in close geographic proximity, but long-term and long-range relationships are more obscure. Friedlaender et al. (2002) pre-

sented a preliminary study of mitochondrial DNA variation on Santa Cruz, an island group beyond the Near Oceania boundary, some 300 km southeast of the main Solomons chain. This remains the sole representation of modern molecular genetic research in Solomon Islands. Although Friedlaender and his colleagues postulate a relatively recent spread of both Melanesian mtDNA lineages and Papuan languages into the geographically very isolated Santa Cruz group, the relationship between language and biological affiliation on the main Solomons chain remains an open question.

To address whether languages and genetic markers are associated in the main Solomon Islands archipelago, we sampled a small number of individuals from two island regions during 2004. The first collection represents a group of unrelated individuals from the province of Malaita, where only Austronesian languages of the Malaita–San Cristobal subgroup are currently spoken; the second collection represents a small population that speaks Touo, an isolated Papuan language restricted to the southern end of Rendova Island, Western Province (Dunn et al., 2002). Using a detailed phylogeny of global Y-chromosome lineages (Jobling and Tyler-Smith, 2003), we examined genetic diversity in these paired Solomon Island samples and classified male individuals to Y-chromosome haplogroup level, as defined by the Y-Chromosome Consortium (YCC, 2002). Oceania has proved a fruitful area of Y-chromosome research, with studies of male lineages suggesting a mid-Holocene expansion of Austronesian-speaking peoples who incorporated men from indigenous Papuan populations more frequently than previously believed (Cox, 2003; Hage and Marck, 2003). Yet subsequently, incoming Austronesian and indigenous Papuan groups have had a long shared history of interaction in Island Melanesia. By contrasting genetic variation in populations speaking Austronesian and Papuan languages, we aim to determine whether genetic and linguistic diversity is correlated today in small Island Melanesian communities.

Here we present an analysis of Y-chromosome diversity in two population groups from Solomon Islands. This study presents the first snapshot of genetic variation in Solomon Island men, and we place this variation within a comparative framework of Y-chromosome diversity from Island Melanesia

and the broader Indo-Pacific region. Finally, we compare patterns of Y-chromosome diversity between the paired Solomon Island populations speaking Austronesian and Papuan languages, and determine whether an association between Y-chromosome lineages and language affiliation observed at a broad Indo-Pacific scale holds at a smaller community level within the Solomon Islands archipelago.

MATERIALS AND METHODS

Samples and DNA extraction

Biological samples from men and a few women ($N = 43$) were collected in two provinces of Solomon Islands under research permits from the Government of Solomon Islands (Ministry of Education and Research, Ministry of Health) to MML (as part of the "Pioneers of Island Melanesia" project), with downstream approval from provincial secretaries, village organizers, and local communities. Only volunteers over the age of 18 years who provided fully informed, written consent were included in the study. Samples from the Western Province were collected on southern Rendova Island by MPC; malarial clinic staff in Honiara, Guadalcanal, collected samples representative of Malaita Province from expatriate Malaitans in the capital city. Blood spots produced by single-use lancets (Unistik2, Owen Mumford, Oxford, England) were collected on untreated filter cards (Testkarten, Schleicher and Schuell/Whatman, Brentford, England), air dried, and transported to the United Kingdom. DNA was extracted from blood leukocytes using a standard phenol/chloroform extraction protocol (Aus-ubel et al., 1995) and concentrated with Microcentricon filters (Millipore, Billerica, MA).

Y-SNP screening

The Y-Chromosome Consortium (YCC) has inferred a detailed tree of global Y-chromosome diversity from a set of over 250 phylogenetically informative polymorphisms (Jobling and Tyler-Smith, 2003; YCC, 2002). The YCC currently recognizes nineteen primary clades of Y-chromosome lineages, called haplogroups, which are characterized by one or more unique, monophyletic polymorphisms and are frequently restricted to specific geographical regions. All male samples ($N = 36$) were typed hierarchically according to YCC guidelines (YCC, 2002) with seven haplogroup-defining markers: K-M9, L-M11, M-M106, NO-M214, N-Lly22g, O-M175, and P-92R7. Screening ceased once Y-chromosomes had been assigned to haplogroup level. Additionally, paragroup K-M9* Y-chromosomes were assayed for two sub-haplogroup lineages—K1-M177 and K-M230—which are restricted geographically to Melanesian populations (Kayser et al., 2003; YCC, 2002). Y-SNPs were typed using PCR-RFLP methodologies described elsewhere (Cox, in print), with the exception of protocols for K-M230 (Kayser et al., 2003) and N-Lly22g (kindly provided by C. Tyler-Smith, Wellcome Trust Sanger Institute, personal communication, 2002). The PCR-RFLP protocols employed here were validated against YCC control DNAs with known lineage assignments (kindly provided by M. Hammer, University of Arizona, 2004). Fluorescence-based DNA sequencing of forward and reverse strands confirmed the derived status of a single individual classified by PCR-RFLP to lineage K1-M177.

Population genetic statistics

Y-chromosome haplogroup frequencies of representative Indo-Pacific populations were generated for comparative purposes using data from publicly available sources (Capelli et al., 2001; Cox, 2003; Kayser et al., 2003). Lineage frequencies were collapsed to haplogroup level for full comparability with the data presented here for two Solomon Island populations. Genealogical information collected from Solomon Island donors allowed the exclusion of men who were related paternally within three generations (the resolution limit of the genealogical data). Only Y-chromosome data from men who were unrelated using this criterion (32 of 36) were analyzed further.

Haplogroup diversity ($\hat{H}$) and its variance were ascertained for each sample to determine patterns of population variability across the Indo-Pacific region. Haplogroup diversity values, representing the probability of two different haplotypes being chosen randomly from a population sample, were calculated according to Nei (1987:180) with the freely available software, Arlequin (Excoffier, 2005). Confidence intervals were calculated using the normal approximation. Because differences between groups should not be evaluated by multiple direct compar-

isons of confidence intervals (Sokal and Rohlf, 1995:240), the significance of differences between population pairwise values was determined using the Tukey-Kramer procedure (Sokal and Rohlf, 1995:247). Inverse Studentized range distributions were calculated for each desired significance value ($\alpha = 0.05, 0.01, 0.001$) with the statistical package R (R Project, 2005), and the Tukey–Kramer procedure was executed manually in a Microsoft® Excel spreadsheet developed by one of the authors (M.P.C.).

The genetic structure of populations was investigated within an analysis of variance framework. Hierarchical ANOVAs were computed to determine whether language affiliation is a significant factor in the Indo-Pacific region. Variance values among groups (F_{CT}), within groups (F_{SC}), and among populations (F_{ST}) were calculated in Arlequin (Excoffier, 2005), and their significance was evaluated via a permutation approach with 10,000 iterations. Pairwise population variance values (F_{ST}) were adopted as representative genetic distances and compared with a matrix of geographical distances between sampling locations to determine whether Y-chromosome variation in the Indo-Pacific region is genetically isolated by distance. Geographical distances between sampling locations were calculated as geodesic arcs in kilometers using the freely available software, Earth (Byers 2005), and correlated against genetic distance using the Mantel (1967) test implemented in the statistical package R. Geographical distances were ln- transformed to compensate for two-dimensional spatial analysis, and pairwise F_{ST} values were transformed by $F_{ST}/(1 - F_{ST})$ to generate values varying from 0 to ∞ in analogy with geographical distance (Rousset, 1997).

RESULTS

A total of nine Y-chromosome SNPs were screened in 32 men to investigate the association between genetic markers and language affiliation in two provinces of Solomon Islands. Although small, this sample is the first presentation of Y-chromosome variation in Solomon Islands. Lineage frequencies among 1,368 men from 26 additional Indo-Pacific populations were contrasted with the Solomon Island sample to determine major spatial trends across Island Southeast Asia and Oceania (Fig. 2). Only five Y-chromosome lineages were identified in men from Solomon Islands: K-M9*, K-M230, K1-M177, M-M106, and O-M175 (Table 1). These lineages are described in greater detail below.
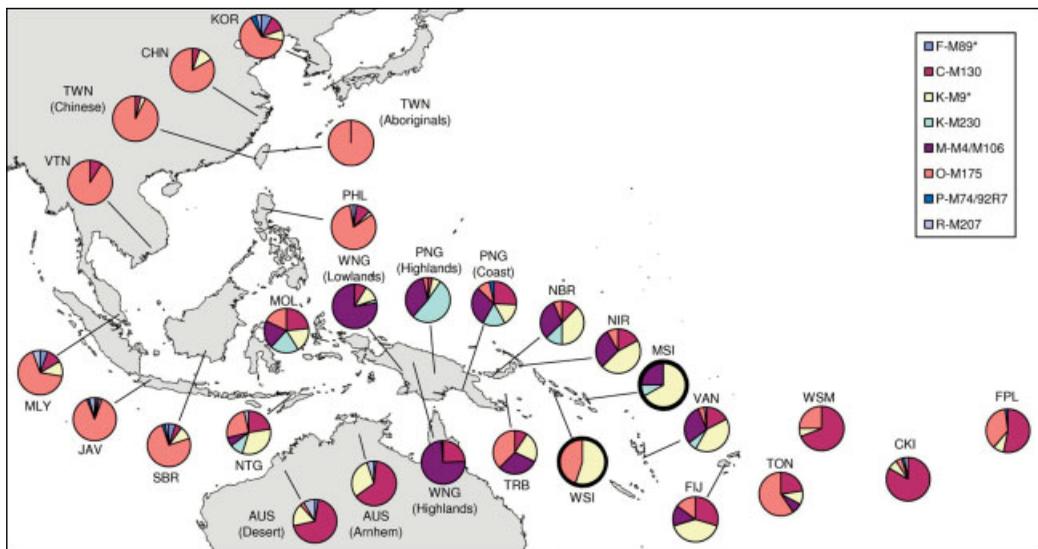

Fig. 2. Frequency distributions of Y-chromosome lineages in Solomon Islands (pie charts in bold outline) and 26 additional populations from the Indo-Pacific region. See Table 1 for population abbreviations.

TABLE 1. Y-chromosome lineage frequencies in Solomon Islands and 26 other populations from the Indo-Pacific region

Location	Code	No.	Lineage frequencies (%)								Haplogroup diversity[a]		Source
			F-M89*	C-M130	K-M9*	K-M230	M-M4/M106	O-M175	P-M74/92R7	R-M207	Ĥ	SD	
1 Korea	KOR	25	8	12.0	8.0	0	0	64.0	4.0	4.0	0.58	0.11	d
2 China	CHN	36	0	5.6	11.1	0	0	83.3	0	0	0.30	0.093	d
3 Vietnam	VNM	11	0	9.1	0	0	0	91.0	0	0	0.18	0.14	d
4 Malaysia	MLY	18	5.6	11.1	11.1	0	0	66.7	0	5.6	0.56	0.13	d
5 Taiwan—Chinese	TWN-C	26	0	3.8	3.8	0	0	92.3	0	0	0.15	0.093	d
6 Taiwan—Aborigines	TWN-A	43	0	0	0	0	0	100.0	0	0	0	0	d
7 Philippines	PHL	39	2.6	10.3	2.6	0	0	82.1	0	2.6	0.32	0.093	d
8 Java	JAV	53	1.9	1.9	1.9	0	0	88.6	1.9	3.8	0.21	0.074	d
9 Southern Borneo	SBR	40	5	5.0	10.0	0	0	75.0	2.5	2.5	0.43	0.094	d
10 Nusa Tenggara	NTG	31	0	22.6	32.3	9.7	6.5	25.8	0	3.2	0.73	0.046	d
11 Moluccas	MOL	34	0	23.5	17.6	20.6	20.6	17.7	0	0	0.75	0.035	d
12 WNG Lowlands/Coast	WNG-L	89	0	9.0	11.2	2.2	77.5	0	0	0	0.38	0.059	d
13 WNG Highlands	WNG-H	94	0	24.5	0	0	74.5	1.1	0	0	0.39	0.047	d
14 PNG Highlands	PNG-H	31	0	3.2	6.5	51.6	35.5	3.2	0	0	0.55	0.060	d
15 PNG Coast	PNG-C	31	0	25.8	16.1	16.1	29.0	9.7	3.2	0	0.76	0.074	d
16 Trobriand Islands	TRB	53	0	9.4	22.6	0	30.2	37.7	0	0	0.72	0.037	d
17 Tolai New British	NBR	16	0	12.5	37.5	12.5	31.2	6.3	0	0	0.68	0.085	d
18 New Ireland	NIR	86	0	17.0	45.0	0	29.0	8.0	0	0	0.68	0.029	b
19 Malaita Province, Solomon Is.	*MSI*	*12*	*0*	*0*	*66.7*	*8.3*	*25.0*	*0*	*0*	*0*	*0.41*	*0.13*	*This paper*
20 Western Province, Solomon Is.	*WSI*	*20*	*0*	*0*	*55.0*	*0*	*0*	*45.0*	*0*	*0*	*0.52*	*0.042*	*This paper*
Composite, Solomon Is.		*32*	*0*	*0*	*59.4*	*3.1*	*9.4*	*28.1*	*0*	*0*	*0.54*	*0.072*	*This paper*
21 Vanuatu	VAN	234	0	17.5	40.6	6.4	29.5	4.3	0	1.7	0.66	0.018	c
22 Fiji	FIJ	55	0	30.0	41.0	0	15.0	15.0	0	0	0.71	0.033	b
23 Tonga	TON	55	0	23.0	10.0	0	8.0	60.0	0	0	0.58	0.059	b
24 Western Samoa	WSM	16	0	69.0	6.0	0	0	25.0	0	0	0.49	0.12	b
25 Cook Islands	CKI	70	0	83.3	7.5	0	0	4.6	1.8	2.84	0.29	0.069	b, d
26 French Polynesia	FPL	87	0	53.0	8.0	0	0	37.0	2.0	0	0.59	0.031	b
27 Australia—Arnhem	AUS-A	60	1.7	63.3	30.0	0	0	0	0	5.0	0.52	0.051	d
28 Australia—Desert	AUS-D	35	2.9	68.7	17.1	0	0	3.0	0	8.6	0.51	0.090	d
TOTAL		1400	0.6	24.3	20.3	3.6	21.3	28.0	0.5	1.4	0.76	0.0023	

[a]Haplogroup diversity values are calculated on combined K-M230/K-M9* and R-M207/P-M74/92R7 proportions due to a lack of data resolution from some population studies. Frequencies taken from publications:
[b]Capelli et al. (2001).
[c]Cox (2003).
[d]Kayser et al. (2003).

Lineage K-M9*

K-M9* is a paragroup defined by the presence of M9, a marker found widely across Eurasia and the Americas, in the absence of the haplogroup-defining polymorphisms L-M11, M-M106, NO-M214, and P-92R7. In this study, K-M9* individuals also lack the sublineage polymorphisms K-M230 and K1-M177. The K-M9* paragroup is effectively a catch-all classification of Y-chromosomes that cannot be defined more accurately with the set of Y-SNP markers available today. K-M9* is particularly common in parts of Melanesia, sporadically exceeding 50% across Island Melanesia. When phylogenetically informative Y-SNPs are screened more thoroughly in these populations, this paragroup may be found to encompass one or several monophyletic lineages. However, K-M9* is treated in this study as a single group for statistical purposes. Although the distribution of K-M9* Y-chromosomes is highly suggestive of Australo-Melanesian affinity, low K-M9* frequencies on Mainland Asia cannot exclude the possibility of an Asian connection for this paragroup. The majority of Y-chromosomes from Malaita (67%) and the Western Province (55%) fall within the K-M9* paragroup, which is at higher frequency in Solomon Islands than to the north in New Ireland (45%) and New Britain (38%), and to the south in Vanuatu (41%). The prevalence of unclassifiable K-M9* Y-chromosomes places Solomon Islands firmly within an Island Melanesian context, but adds little to our understanding of the region's male biological history.

Lineage K-M230

K-M230 was the first polymorphism found to sub-classify K-M9* Y-chromosomes in Melanesia (Kayser et al., 2003). The marker occurs in eastern Indonesia (10–20%) and Island Melanesia ($\sim$ 10%), but reaches greatest frequency in the highlands of Papua New Guinea (52%). Genetic dating places the origin of the M230 T→A mutation early in the Holocene, followed by demographic expansion around 3,500 BP (Kayser et al. 2003). Given its prevalence in Island Melanesia, K-M230 may represent a subset of indigenous Melanesian lineages adjoined to highly variable, incursive populations in the vicinity of New Guinea, which spread rapidly in concert through Near Oceania during the late

Holocene. However, given uncertainty surrounding the accuracy of genetic dates, a dispersal of K-M230 lineages either earlier or during modern times cannot be discounted.

Lineage K1-M177

K1-M177 was first identified in the YCC cell-line repository, from which most of our phylogenetically informative Y-SNPs were initially developed. Until now, a single Papuan-speaking Nasioi individual from the eastern coast of Bougainville represented the entire dataset of K1-M177 carriers (YCC, 2002). Although this marker is not screened regularly in population surveys (e.g., Capelli et al., 2001; Kayser et al., 2003), K1-M177 was not detected in over 200 men from Vanuatu (Cox, 2003). K1-M177 is often considered a "private" lineage—a rare mutation event that occurred recently in the immediate paternal genealogy of the single Nasioi carrier. However, this study identified the M177 polymorphism using a PCR-RFLP methodology in one Solomon Islander from Malaita Province. This unexpected discovery was confirmed by forward- and reverse-strand DNA sequencing, and compared against known ancestral and derived control DNAs. Modern times have seen moderate levels of contact between Bougainville and Malaita, but the identification of K1-M177 carriers at opposite ends of the main Solomons chain may hint at an earlier spread of the polymorphism. Perhaps K1-M177 demarcates a subset—albeit small—of unresolved K-M9* lineages in Island Melanesia. If so, wider population screening of K1-M177 may prove helpful for reconstructing population substructure in the region.

Lineage M-M106

The distribution of M-M106 is centered firmly on Melanesia. Frequencies hover around 75% across West New Guinea, and the marker reaches frequencies of 20% in the Moluccas and 29–36% across Island Melanesia from New Britain to Vanuatu (Capelli et al., 2001; Cox, 2003; Kayser et al., 2003). Three Papuan-speaking Nasioi individuals from eastern Bougainville also carried M-M106 (Su et al., 2000). This lineage has not been detected north or west of a line drawn from the Moluccas to Nusa Tenggera, and its absence from Mainland

Asia is especially notable (Su et al., 1999). Furthermore, M-M106 does not occur in Polynesian populations, with the exception of Tonga (8%). M-M106 is largely restricted within the geographical distribution of Papuan languages. Although modern Polynesians carry high proportions of one indigenous Melanesian Y-chromosome lineage (C2b-M208; Kayser et al., 2003), the absence of M-M106 (and K-M230) argues forcefully against a simple admixture scenario. The underlying processes of population contact during the Austronesian expansion must have been more complex, or they involved small numbers of men and considerable genetic drift. Adding to this more complex picture of Austronesian and Papuan admixture in the late Holocene, M-M106 is not associated with Papuan languages in Solomon Islands. A quarter of men from Malaita Province are M-M106 carriers, although only Austronesian languages are currently spoken in this province. No M-M106 lineages were detected in the Western Province sample, even though the individuals sampled speak an isolated Papuan language. This distribution of lineage M-M106—the inverse of that expected from broad-scale linguistic associations—stresses the importance of viewing Indo-Pacific genetic diversity from diverse perspectives: associations that are perceptible at broad scales do not necessarily hold at the level of community surveys.

Lineage O-M175

In comparison, O-M175 has a center of gravity seated firmly in East Asia. The lineage varies from 64% to 100% among Mainland and Island Southeast Asian populations (and approaches 35% in the Austronesian-speaking population of Madagascar: Cox, 2003; Hurles et al., 2005), but O-M175 is always carried by a majority of individuals in Southeast Asia. In a sample of Taiwanese aborigines, descendants of a presumptive homeland population under one model of the Austronesian expansion (Bellwood, 2005), O-M175 has reached fixation (Kayser et al., 2003). Contrary to the distribution of M-M106, O-M175 frequencies drop rapidly south and east of a line drawn from the Moluccas to the Nusa Tenggeras. The lineage is all but absent from West New Guinea and the Papua New Guinea highlands, and Y-STR analysis of the few indivi-

duals who carry it suggests recent admixture from surrounding regions (Kayser et al., 2003:289). East of New Guinea, O-M175 frequencies become more variable in Island Melanesia (6–37%), but they still indicate moderate levels of Asian contribution. Frequencies are even more variable in Polynesia (5–60%). Some subgroups of lineage O-M175 (particularly O3-M122) have probable associations with the expansion of Austronesian languages into the Pacific (Capelli et al., 2001; Kayser et al., 2001). However, in analogy to the inverse distribution observed for lineage M-M106, the O-M175 association is also switched in Solomon Islands. O-M175 was not identified in the sample from Malaita Province, where only Austronesian languages are currently spoken. Conversely, nearly half the men sampled from the Western Province are O-M175 carriers (45%), despite speaking a Papuan language. It is unclear whether this effect is an artefact of small sample size or an accurate representation of Y-chromosome diversity in Solomon Islands, but the finding again emphasizes that community studies can discern levels of variability that may be missed at higher scales of analysis.

Patterns of genetic diversity and population substructuring

The genetic structure of Y-chromosome diversity in the Indo-Pacific region was investigated with an analysis of variance approach. Haplogroup variance values (F_{ST}) were calculated for each population pair, and correlated against the geographical distance between sampling locations. Y-chromosome variation in the Indo-Pacific shows marked genetic isolation by distance; populations that are close geographically are also significantly more similar genetically ($r = 0.30$, $P < 0.001$). However, only a small proportion of the variation in F_{ST} values ($r^2 = 0.09$) can be attributed to the effects of increasing geographical distance. Additional factors—some probably cultural, some merely stochastic—are necessary to account for Y-chromosome haplogroup distributions in the Indo-Pacific.

One such factor is language affiliation. Membership of a community that speaks an Austronesian, Papuan, or Mainland Asian language is a significant factor affecting the distribution of Y-chromosome variation across the Indo-Pacific (Table 2). For comparative purposes, samples from Asian popu-

TABLE 2. *Y-chromosome haplogroup ANOVA variance components factored by language affiliation and geographical location (excluding Australian samples); group assignments are described in the text*

Groups	N	Variance components (%)		
		Between groups	Within groups	Within populations
No Subdivision by language	1	—	33.7***	66.3***
Austronesian, Papuan	2	13.7*	24.6***	61.7***
Austronesian, Mainland Asia	2	15.5*	23.0***	61.5***
Papuan, Mainland Asia	2	47.5*	13.1***	39.5***
Austronesian, Papuan, Mainland Asia	3	17.0**	23.1***	59.9***
Mainland Asia, Southeast Asia	2	0^{ns}	4.1**	97.2**
Southeast Asia, Greater Australian Outliers	2	39.2***	1.9***	58.9***
Greater Australian Outliers, Polynesia	2	19.9**	6.1***	74.1***
Mainland Asia, Greater Australian Outliers	2	34.5**	2.0***	63.5***
Mainland Asia, Polynesia	2	29.2*	12.3***	58.5***
Papuan, Mainland Asia	2	47.5*	13.1***	39.5***
Papuan, Southeast Asia	2	53.5**	10.3***	36.3***
Papuan, Greater Australian Outliers	2	11.6*	8.7***	79.8***
Papuan, Polynesia	2	31.3*	17.7***	51.1***

[a]Significance values: *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ns, not significant.

lations that speak languages from a range of language families (e.g., Austro-Asiatic, Sino-Tibetan) were placed in a residual "Mainland Asian" category, and because this study focuses on the genetic relationship between Austronesian- and Papuan-speaking populations (and their potential links to Mainland Asia), the two Australian samples included in the Indo-Pacific isolation by distance analysis were excluded from the following ANOVA analyses. A substantial proportion of Y-chromosome variance occurs between the two groups speaking non-Austronesian languages, the Papuan and Mainland Asian populations. These are discriminated by nearly half of their total genetic variation ($F_{CT} = 0.48$, $P < 0.05$). Conversely, Austronesian-speaking groups are equally distant genetically from Mainland Asians ($F_{CT} = 0.16$, $P < 0.05$) and Papuans ($F_{CT} = 0.14$, $P < 0.05$). There is significant statistical support for the view that Y-chromosome lineage frequencies vary between Austronesian, Papuan, and Mainland Asian groups, *from a broad Indo-Pacific perspective*.

However, Austronesian-speaking populations are clearly not a simple subset of either the Mainland Asian or Papuan Y-chromosome profiles. Instead, they carry lineages that have connections with Mainland Asia (e.g., haplogroup O-M175), as well as with Papuan-speaking groups (e.g., haplogroups C-M130 and M-M106). The Papuan connection is especially pronounced in Austronesian-speaking populations from the

vicinity of New Guinea, a region that may be termed the "Greater Australian Outliers" (see Fig. 1 in Terrell, 2004:604). To clarify the historical process behind these statistical patterns, we subdivided the Austronesian-speaking populations into three geographical groups (abbreviations as in Table 1): Southeast Asia (JAV, MLY, PHL, SBR, TWN-A), the Greater Australian Outliers (FIJ, MOL, MSI, NBR, NIR, NTG, PNG-C, TRB, VAN), and Polynesia (CKI, FPL, TON, WSM). We compared the Y-chromosome variance of these groups with those of the Mainland Asian (CHN, KOR, TWN-C, VTN) and Papuan (PNG-H, WNG-L, WNG-H, WSI) samples to determine how Y-chromosome lineage frequencies vary across the Indo-Pacific region (Fig. 3, Table 2). While the Austronesian-speaking populations of Southeast Asia are indistinguishable from the non-Austronesian speaking groups of Mainland Asia ($F_{CT} = 0$, ns), the Austronesian-speaking populations in Southeast Asia and the Greater Australian Outliers are significantly different ($F_{CT} = 0.39$, $P < 0.001$). At the same time, the Greater Australian Outliers are more similar to Papuan-speaking groups ($F_{CT} = 0.12$, $P < 0.05$) than are the Austronesian-speaking populations of Southeast Asia ($F_{CT} = 0.54$, $P < 0.01$). Further east in Oceania, Polynesian populations show greater difference from the Papuan-speaking groups ($F_{CT} = 0.31$, $P < 0.05$) than do the Greater Australian Outliers. However, Polynesians are more similar to Mainland Asian popula-
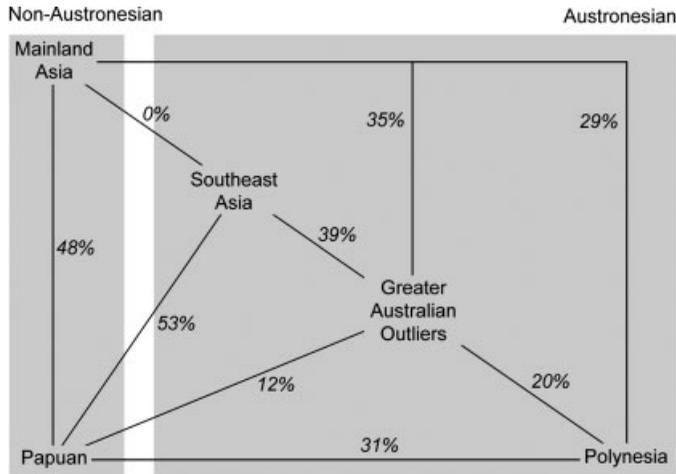
Fig. 3. Percentage dissimilarity variance values (F_{CT}) among geographical linguistic groups. Mainland Asia and Papuan groups represent populations that speak non-Austronesian languages; Southeast Asia, the Greater Australian Outliers, and Polynesia represent Austronesian-speaking populations. Group assignments are described in the text.

tions ($F_{CT} = 0.29$, $P < 0.05$) than are the Greater Australian Outliers ($F_{CT} = 0.35$, $P < 0.01$). This complex patterning is consistent with an expansion of Austronesian-speaking populations ultimately from Mainland Asia, followed by sizable admixture with Papuan-speaking populations as they interacted in the vicinity of New Guinea (Bellwood, 2005; Cox, 2003). The greater Papuan characteristic of the Greater Australian Outliers relative to Polynesian populations probably signifies major population admixture in the vicinity of New Guinea during the 3,000–4,000 years since incursive Asian and indigenous Papuan groups first came into contact. An alternative model (Oppenheimer and Richards, 2002) postulating an expansion of Austronesian-speaking groups from Wallacea (the western part of the Greater Australian Outliers) into northern and western Island Southeast Asia and Oceania cannot explain these patterns. Under the Wallacean model of a northern spread of Austronesian-speaking peoples, Austronesian-speaking populations in the Greater Australian Outliers and Southeast Asia should not be particularly dissimilar (which they are), and the non-Austronesian populations of Mainland Asia should be quite different from Austronesian-speaking populations in Southeast Asia (which they are not). A southern Chinese/Taiwanese origin of Austronesian-speakers (Bellwood, 2005) seems best supported by

this analysis of Y-chromosome data from the Indo-Pacific region.

Consequently, from a statistical perspective, an association seems to hold in the Indo-Pacific between Papuan-speaking groups (and Y-chromosome haplogroups C-M130 and M-M106) on the one hand, and Austronesian-speaking groups (and Y-chromosome haplogroup O-M175) on the other. Of course, such broad-scale associations need not hold for *all* population samples. Much of the total Y-chromosome variation occurs within populations (e.g., 60% for the entire Indo-Pacific sample), and as with genetic isolation by distance, the effect of language affiliation on genetic variation is relatively minor in comparison (e.g., 17% for the entire Indo-Pacific sample). Although there are significant statistical associations between Y-chromosome lineages and language groups in the Indo-Pacific region, non-linguistic factors clearly play dominant roles.

In similar fashion, genetic diversity also varies significantly across the Indo-Pacific. Haplogroup diversity was calculated for each population sample (Table 1), and ranked from lowest to highest diversity (Fig. 4). Haplogroup diversity is lowest in Mainland and Island Southeast Asian populations, such as China and Taiwan; the sample of Taiwanese aboriginals has no diversity at all. The Highlands and Lowlands of West
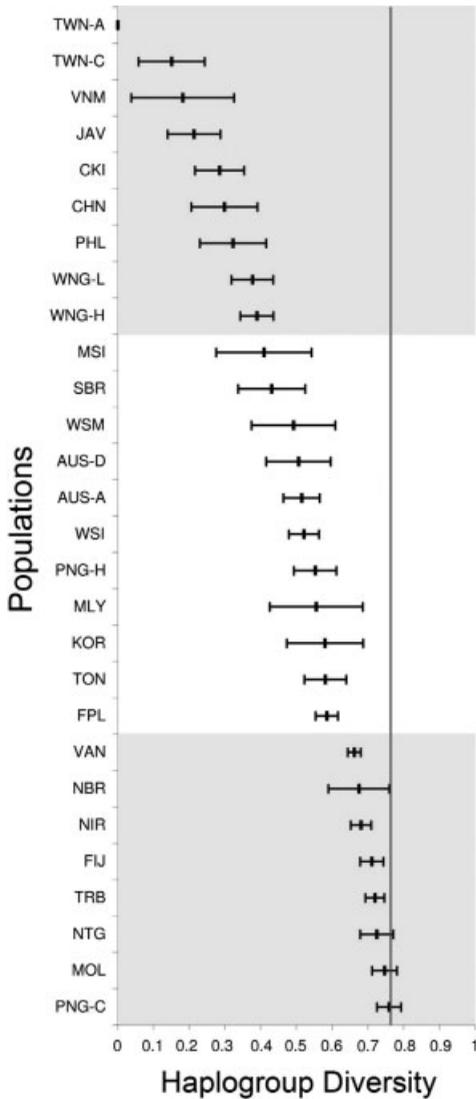
Fig. 4. Y-chromosome haplogroup diversity (Ĥ) in 28 populations from the Indo-Pacific region, including Solomon Islands. Haplogroup diversity is ranked from lowest to highest; error bars indicate one standard deviation from the mean. Composite haplogroup diversity for the entire dataset is designated by a vertical bar, and approximates a maximal trend for Ĥ values in the Indo-Pacific region. Any two populations from the paired shaded boxes have haplogroup diversity values that are highly significantly different from one another ($\alpha = 0.001$).

Outliers"; Terrell, 1989, 2004): that is, Nusa Tenggera and Moluccas to the west, and Island Melanesia from New Britain to Fiji. Coastal Papua New Guinea has the highest haplogroup diversity of any population sampled from the Indo-Pacific region (0.76; 95% confidence interval [0.69, 0.83]). These values trend toward the haplogroup diversity value calculated for a theoretically completely admixed Indo-Pacific population (Ĥ = 0.76; 95% CI [0.76, 0.76]); haplogroup diversity along the Papua New Guinea coast may be approaching the maximum possible within this region today. Such high haplogroup diversity is suggestive of admixture from two or more relatively homogeneous parental sources. In corroboration, Y-chromosome lineages found in the Greater Australian Outliers tend to be composite profiles of Y-chromosome lineages found at greatest frequency in populations with low haplogroup diversity; for instance, O-M175 in Asia (China, Taiwanese Aborigines), and C-M130, M-M106, and K-M230 in Melanesia (West New Guinea Highlands, West New Guinea Lowlands, Papua New Guinea Highlands). This pattern is reminiscent of the population admixture expected during a recent movement of Asian-derived populations into Oceania with subsequent integration of individuals from pre-existing groups in the region (Bellwood, 2001; Cox, 2003).

Patterns of haplogroup diversity incidentally shed light on the relationship between populations in Melanesia and Polynesia. Polynesian assemblages are characterized in particular by high Y-chromosome variability; genetic diversity ranges from low in the Cook Islands (0.29; 95% CI [0.15, 0.42]) to moderately high in French Polynesia (0.59; 95% CI [0.52, 0.65]). High genetic variability among Polynesian populations may result from repeated founder events with associated stochastic genetic drift, as small, kin-related groups stepped out across the uninhabited islands of the greater Pacific (Fix, 2004). As a result of these processes, the haplogroup diversity of French Polynesia and Tonga is significantly higher than that of the Cook Islands (P< 0.001). Consequently, no one Polynesian sample can be treated as representative of Polynesia as a whole, and a range of Polynesian populations should be adopted if used for comparative purposes with other geographical regions; for instance, as here with Island Melanesia.

New Guinea have similarly low diversity. Conversely, the highest haplogroup diversity is found in populations adjacent to New Guinea (cf. the "Greater Australian

DISCUSSION

Here we present the first snapshot of Y-chromosome diversity in Solomon Islands. Although limited by sample size, we place these data within an Island Melanesian framework, and determine whether Y-chromosome variation is apportioned relative to language affiliation within the Solomon Island archipelago. Finally, we address questions regarding variability in small populations, and the effect of nonlanguage factors, such as geographical distance, on population substructure in Island Melanesia in general, and Solomon Islands in particular.

Y-chromosomes were analyzed from unrelated men ($N = 32$) from two regions of Solomon Islands: Malaita Province and the Western Province. Both regions are dominated by the K-M9* paragroup, which represents an assemblage of unclassifiable Y-chromosome lineages. K-M9* frequencies are higher in these Solomon Island samples than anywhere else in the Indo-Pacific (59% in the combined Solomon Island dataset), and while the prevalence of K-M9* places these samples within a Melanesian context, high unresolved K-M9* frequencies across the entirety of Island Melanesia hampers reconstruction of the region's prehistory. Identification of additional phylogenetically informative Y-SNP markers that subdivide the K-M9* paragroup is a necessary prerequisite for teasing apart the population substructure in Island Melanesia. Indo-Pacific K-M9* lineages have their center of gravity in Near Oceania (although they are distributed from Mainland Asia to Australia), so new monophyletic lineages may well prove informative for reconstructing Pleistocene settlement across a large geographical region, including Australia, New Guinea, and Near Oceania.

In comparison, the distribution of Y-chromosome haplogroups M-M106 and O-M175 already reveals broad spatial patterning relative to major language groups. M-M106 is distributed in close approximation to the geographical extent of Papuan languages. The lineage reaches greatest frequency in the Highlands of West and Papua New Guinea, and in the Lowlands of West New Guinea, regions where Papuan languages are spoken exclusively today. Consequently, M-M106 has been proposed as a marker of populations speaking Papuan languages that ultimately derive from Pleistocene groups in and around New Guinea (Capelli et al., 2001:440; Kayser et al., 2003:297). Conversely, O-M175 occurs most frequently across Mainland and Island Southeast Asia with moderate, but variable, frequencies in Polynesian populations. O-M175 follows the current distribution of Austronesian languages across the Indo-Pacific region, and the predominant Pacific sublineage O3-M122 (Capelli et al., 2001) has been proposed as a marker of the spread of Austronesian-speaking populations through Near Oceania and into the greater Pacific (Kayser et al., 2003:299).

This pattern is reflected by an analysis of haplogroup variance, which suggests that Y-chromosome diversity across the Indo-Pacific is statistically associated with language-group affiliation. This association is underpinned by the prevalence of haplogroups C-M130 and M-M106 among Papuan-speaking groups, and haplogroup O-M175 among Austronesian-speaking populations. However, the broad Indo-Pacific association between language affiliation and Y-chromosome variation has been blurred by post-contact interaction in the Greater Australian Outliers, and it does not hold on a smaller scale within Solomon Islands. A quarter of Y-chromosomes from Malaita Province are M-M106 carriers, although only languages of the Austronesian family are spoken there. No O-M175 carriers were detected. Conversely, the Western Province assemblage represents people speaking an isolated Papuan language at the southern end of Rendova Island. Nearly half of the men in this sample are O-M175 derived, but no men were detected with lineage M-M106. This pattern is the inverse of that expected from broad-scale genetic and linguistic associations. Although sample sizes are limited, there may well be no correlation between language affiliation and Y-chromosome diversity within the Solomon Islands.

The underlying process behind this effect is unclear, yet four reasons suggest themselves: an invalid assumption of correlation between genetic markers and language, small sample size leading to inaccurate frequency estimates (sampling error), language shift, and frequency sweeps resulting from limited migration with subsequent genetic drift.

Firstly, the association between language groups and Y-chromosome lineages observed from a broad perspective of the Indo-Pacific region may be erroneous. However, to support this hypothesis, it would have to be argued that O-M175 frequencies increased randomly among Austronesian-speaking populations in Island Southeast Asia and Oceania, while M-M106 frequencies increased randomly in regions where Papuan languages are found today. If M175 and M-M106 frequencies fluctuated randomly among populations, they should not display broad-scale spatial trends. Consequently, patterns of genetic diversity that can be discerned at a regional linguistic scale probably represent the downstream effects of real population processes.

Secondly, estimates of Y-chromosome lineage frequencies from small samples may be incorrect. Although further sampling from Solomon Islands would undoubtedly improve our picture of Y-chromosome variability in the archipelago, some Indo-Pacific samples used for comparative purposes are also small (e.g., those from Vietnam, Malaysia, New Britain, Western Samoa). However, frequency estimates for these populations are invariably consistent with frequency estimates for adjacent populations with larger sample sizes. For instance, the Y-chromosome lineage profile for New Britain falls within the range of neighboring groups. Limited sampling must have played some role in the variability observed in Y-chromosome lineage distributions in Solomon Islands, but simple sampling error is unlikely to provide a complete account for the inverted language/genetic lineage association.

Thirdly, the lack of association between Y-chromosome lineages and language may result from language shifts. The prehistory of the Solomon Islands is understood only poorly, and instances of language adoption and language extinction are almost completely unidentified (Pawley, 2002:267). Groups speaking Papuan languages are usually perceived as tending to adopt Austronesian languages (for instance, as probably occurred in New Ireland; Ross, 1994), primarily in order to explain the predominance of Austronesian languages across the Pacific (Bellwood, 2005). But Austronesian-speaking groups can adopt Papuan languages as easily as the reverse, as was observed among Uruava speakers on Bougainville during the early twentieth century (Friedlaender, 1975). The Papuan language spoken in southern Rendova (Touo) is currently surrounded by Austronesian-speaking populations, and isolated from related Papuan languages on Vella Lavella and the Russell Islands (Dunn et al., 2002). Yet, there is no compelling reason to assume that populations on southern Rendova have always spoken a Papuan language. Perhaps Touo was adopted from the Roviana Lagoon region of New Georgia, where populations spoke Kazukuru, a now extinct but possibly related Papuan language (Grimes, 2005). However, the only clear evidence of language shift in this region today is the encroachment of neighboring Austronesian languages (e.g., Ughele, Roviana, Marovo) on Touo-speaking individuals in the form of multilingualism (Terrill and Dunn, 2003). There are no compelling reasons either to support or refute a prehistoric language shift on southern Rendova.

Fourthly, sweeps in the frequency of certain genetic lineages may result from limited immigration and/or stochastic genetic drift. A key characteristic of genealogies collected from the Western Province is the predominance of large extended family groups. Individuals related within three generations along their paternal lineage were not analyzed in this study, so as not to bias Y-chromosome lineage frequencies unduly. However, the Western Province collection probably incorporates individuals who shared a recent paternal ancestor just four or more generations ago (prior to approximately AD 1850–1900). Genealogical records and community history indicate immigrants from Vella Lavella, Choiseul, and the Marovo Lagoon region of New Georgia, and men from elsewhere in the Solomon Islands (or even further afield) may well have settled in southern Rendova intermittently over an unknown period. If the descendants of such men had more children than average—either through favorable exogamous marriage practices or simply by chance—their Y-chromosome lineages could easily have swept through these small village populations (Fix, 2004; Wright, 1955). Additionally, there may well have been some level of genetic drift during the massive depopulation immediately following initial European contact. Consequently, haplogroup frequencies should be interpreted cautiously; they may be unstable in regions

where populations periodically fluctuate in size or are constantly small.

Limited immigration (perhaps kin-structured) with differential fecundity driving sweeps in the frequency of Y-chromosome lineages is a likely explanation for the genetic variability observed in the two Solomon Island samples. Similarly, high variability of Y-chromosome lineages among Polynesian samples—for instance, O-M175 frequencies of 5–60% and C-M130 frequencies of 23–83%—are highly suggestive of founder events and bottlenecks with stochastic genetic drift in small island populations. The paucity of O-M175 (and corresponding prevalence of C-M130 sublineages) in some Polynesian populations has been taken as support for a major biological contribution from Melanesian rather than Asian sources (Capelli et al., 2001; Kayser et al., 2000). However, while a substantial Melanesian contribution is not in doubt, the variability of these populations is a more interesting phenomenon. It suggests that simple admixture models may not be sufficient to explain the diversity of Pacific populations, and it intimates that more sophisticated models of population genetic processes are probably necessary to explain the extreme variability of many Oceanic samples.

Although seldom documented, the variability of Y-chromosome lineages in Solomon Islands also emphasizes the effects of different sampling strategies. Studies at a community level can discern patterns of variability that may be missed at higher scales of analysis. However, broad population surveys are pre-requisite for observing large-scale spatial trends. For instance, the coastal Papua New Guinea sample adopted here for comparative purposes is a composite assemblage; it represents individuals living at several locations along the northern and southern coasts of Papua New Guinea (Kayser el al., 2000; Stoneking et al., 1990). This sample is representative of a typical broad-scale population survey, but it does not represent a population in any normal biological sense. A breakdown of Y-chromosome data from specific locations along Papua New Guinea is not available, but the genetic diversity at each sampling location would almost certainly be less than that of the composite sample. The collection taken in the Western Province of Solomon Islands represents a more meaningful biological population: a largely panmictic group of individuals with an extended history of subdivision from neighboring groups. The genetic diversity at *individual* locations along coastal Papua New Guinea may well mimic the variable Y-chromosome haplogroup profiles observed here in Solomon Islands.

Importantly, no one sampling strategy is better than another. Composite samples from Coastal and Highland Papua New Guinea emphasize large-scale differences between these two broad geographical regions. Composite samples are better suited to addressing questions relating to large geographical areas; for instance, whether Austronesian and Papuan languages correlate with particular genetic markers across the Indo-Pacific as a whole. Samples from small populations, such as that from the Western Province used here, are less easily applied to such large-scale questions, because the effects of founder events and genetic drift cannot be averaged out among a number of individual communities, the effect of which is to leave behind residual distributions that reflect only major demographic trends.

Community-level surveys address different questions. In particular, they can elucidate patterns of population interaction within small geographical regions. Although such research is exceptional today, it is often interesting in its own right and it can produce novel insights; for instance, detailed blood group typing in the New Guinea Highlands helped solve the etiology of Kuru, a neurodegenerative disorder (Simmons et al., 1961). Additionally, patterns of genetic variability among small populations help elucidate large-scale prehistoric processes. Composite samples— taken now or hypothetically at any stage throughout prehistory—cannot accurately reflect genetic variability among small populations; yet small populations are the currency of widespread demographic processes. Reconstruction of the key processes underlying Indo-Pacific settlement requires more research into the genetic variability of small habitation groups in order to determine realistic values for the parameters determining genetic variation in real-world populations. This groundwork is vital if future simulation studies are to reconstruct the vagaries of prehistoric settlement processes across the greater Pacific region.

Here we present the first snapshot of Y-chromosome variation in Solomon Islands.

We find that Y-chromosome diversity varies between the Western Province and Malaita Province, but the distribution of Y-chromosome lineages does not match expectations based on genetic associations with language that are observable in broad-scale Indo-Pacific studies. Although a linguistic/genetic association is suggested by analyses at large geographical scales, we find no association between language affiliation and Y-chromosome variation in the paired population samples examined here from Solomon Islands. Finally, analyses of between-group genetic variance in the Indo-Pacific region strongly support a complex admixture model, in which the Austronesian-speaking peoples of Melanesia and Polynesia have a substantial biological contribution from Mainland Asia, followed subsequently by sizable contact with Papuan groups within Island Melanesia.

ACKNOWLEDGMENTS

LITERATURE CITED

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. 1995. Short protocols in molecular biology. Indianapolis, IN: Wiley.

Bedford S. 2003. New Zealand Archaeological Association 2003 Conference Abstract—From gaps to galore: Lapita sites in Northern Vanuatu. Archaeol N Z 46:62–63.

Bedford S, Spriggs M, Wilson M. 1998. The Australian National University–National Museum of Vanuatu Archaeological Project: a preliminary report on the establishment of cultural sequences and rock art research. Asian Perspect 37:165–193.

Bellwood P. 1997. Prehistory of the Indo-Malaysian archipelago. Honolulu: University of Hawai'i Press.

Bellwood P. 2001. Foregone conclusions? Current anthropology comment. Curr Anthropol 42:107–108.

Bellwood P. 2005. The first farmers: the origins of agricultural societies. Oxford: Blackwell Publishers.

Blake NM, Hawkins BR, Kirk RL, Bahtia K, Brown P, Garruto RM, Gajdusek DC. 1983. A population genetic study of the Banks and Torres islands (Vanuatu) and of the Santa Cruz islands and Polynesian Outliers (Solomon Islands). Am J Phys Anthropol 62:343–361.

Byers, JA. 2005. Surface distance between two points of latitude and longitude. http://www.wcrl.ars.usda.gov/cec/java/lat-long.htm.

Capelli C, Wilson JF, Richards M, Stumpf MPH, Gratrix F, Oppenheimer S, Underhill P, Pascali VL, Ko T-M, Goldstein DB. 2001. A predominantly indigenous paternal heritage of the Austronesian-speaking peoples of Insular Southeast Asia and Oceania. Am J Hum Genet 68: 432–443.

Cox MP. 2003. Genetic patterning at Austronesian contact zones. Unpublished Ph.D. thesis. Dunedin, New Zealand: University of Otago.

Cox MP. 2005. Minimal hierarchical analysis of global human Y-chromosome SNP diversity by PCR-RFLP. Anthropological Science (In print)

Douglas R, Jacobs J, Hoult GE, Staveley JM. 1962. Blood groups, serum genetic factors and hemoglobin in Western Solomon Islanders. Transfusion 2:413–418.

Dunn M, Reesink G, Terrill A. 2002. The East Papuan languages: a preliminary typological appraisal. Ocean Linguist 41:28–62.

Excoffier L. 2005. Arlequin: software for population genetics data analysis. http://lgb.unige.ch/arlequin/

Felgate MW, Dickinson WR. 2001. Late-Lapita and post-Lapita pottery transfers: evidence from intertidal-zone sites of Roviana Lagoon, Western Province, Solomon Islands. In: Jones M, Sheppard P, editors. Australasian connections and new directions. Proceedings of the 7th Australasian Archaeometry Conference. Research in anthropology and linguistics no. 5. Auckland: Department of Anthropology, University of Auckland. p 105–122.

Fix AG. 2004. Kin-structured migration: causes and consequences. Am J Hum Biol 16:387–394.

Friedlaender J. 1975. The demography, genetics, and phenetics of Bougainville Islanders. Cambridge, MA: Harvard University Press.

Friedlaender JS. 1987. The Solomon Islands Project: a long-term study of health, human biology, and culture change. Oxford: Clarendon Press.

Friedlaender JS. 1990. The Solomon Island Project: an Introduction. Am J Phys Anthropol 81:459–464.

Friedlaender JS, Gentz F, Green K, Merriwether DA. 2002. A cautionary tale on ancient migration detection: mitochondrial DNA variation in Santa Cruz Islands, Solomon Islands. Hum Biol 74:453–471.

Green RC. 1991. The Lapita cultural complex: current evidence and proposed models. In: Bellwood PS, editor. Indo-Pacific prehistory 1990, Vol 2. Bulletin of the Indo-Pacific Prehistory Association 11. Canberra and Jakarta: IPPA and Asosiasi Prehistorisi Indonesia. p 295–305.

Grimes BF. 2005. Ethnologue: languages of the world. http://www.ethnologue.com/

Grimes S. 1998. Landscape history of the Solomon Islands: Palynological study of Roviana Lagoon, Western Province, New Georgia. http://www.arts.auckland.ac.nz/ant/NGAS/report98.htm

Hage P, Marck J. 2003. Matrilineality and the Melanesian origin of Polynesian Y chromosomes. Curr Anthropol 44:S121–S127.

Hurles ME, Sykes BC, Jobling MA, Forster P. 2005. The dual origin of the Malagasy in Island Southeast Asia and East Africa: evidence from maternal and paternal lineages. Am J Hum Genet 76:894–901.

Jobling MA, Tyler-Smith C. 2003. The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet 4:598–612.

Kayser M, Brauer S, Weiss G, Underhill PA, Roewer L, Schiefenhövel W, Stoneking M. 2000. Melanesian origin of Polynesian Y chromosomes. Curr Biol 10:1237–1246.

Kayser M, Brauer S, Weiss G, Schiefenhövel W, Underhill PA, Stoneking M. 2001. Independent histories of human Y chromosomes from Melanesia and Australia. Am J Hum Genet 68:173–190.

Kayser M, Brauer S, Weiss G, Schiefenhövel 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.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220.

Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.

O'Connell JF, Allen J. 2004. Dating the colonization of Sahul (Pleistocene Australia–New Guinea): a review of recent research. J Archaeol Sci 31:835–853.

Oppenheimer S, Richards M. 2002. Polynesians: devolved Taiwanese rice farmers or Wallacean maritime traders with fishing, foraging and horticultural skills? In: Bellwood P, Renfrew C, editors. Examining the farming/language dispersal hypothesis. Cambridge, England: McDonald Institute for Archaeological Research. p 287–297.

Pawley A. 2002. The Austronesian dispersal: languages, technologies and people. In: Bellwood P, Renfrew C, editors. Examining the farming/language dispersal hypothesis. Cambridge, England: McDonald Institute for Archaeological Research. p 251–273.

R Project. 2005. The R project for statistical computing. http://www.R-project.org/

Rivers WHR. 1914. The history of Melanesian society. Cambridge, England: Cambridge University Press.

Roe D. 1993. Prehistory without pots: prehistoric settlement and economy of North-West Guadalcanal, Solomon Islands. Unpublished Ph.D. thesis. Canberra, Australia: Australian National University.

Ross M. 1994. Areal phonological features in North Central New Ireland. In: Dutton T, Tryon D, editors. Language contact and change in the Austronesian world. Berlin: Mouton de Gruyter. p 551–572.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145:1219–1228.

Simmons RT, Graydon JJ, Zigas V, Baker LL, Gajdusek DC. 1961. Studies on Kuru: V. A Blood group genetical survey of the Kuru region and other parts of Papua-New Guinea. Am J Trop Med Hyg 10:639–664.

Sokal R, Rohlf F. 1995. Biometry. San Francisco: W.H. Freeman and Company.

Spriggs M. 1997. The island Melanesians. Oxford: Blackwell Publishers.

Spriggs M. 2000. Out of Asia: The spread of Southeast Asian Pleistocene and Neolithic maritime cultures in island Southeast Asia and the western Pacific. Mod Quaternary Res SE Asia 16:51–75.

Stoneking M, Jorde LB, Bhatia K, Wilson AC. 1990. Geographic variation in human mitochondrial DNA from Papua New Guinea. Genetics 124:717–733.

Su B, Xiao J, Underhill P, Deka R, Zhang W, Akey J, Huang W, Shen D, Lu D, Luo J, Chu J, Tan J, Shen P, Davis R, Cavalli-Sforza L, Chakraborty R, Xiong M, Du R, Oefner P, Chen Z, Jin L. 1999. Y-chromosome evidence for a northward migration of modern humans into Eastern Asia during the last Ice Age. Am J Hum Genet 65:1718–1724.

Su B, Jin L, Underhill P, Martinson J, Saha N, McGarvey S, Shriver M, Chu J, Oefner P, Chakraborty R, Deka R. 2000. Polynesian origins: insights from the Y chromosome. Proc Natl Acad Sci U S A 97:8225–8228.

Summerhayes GR, Scales I. 2005. New Lapita pottery finds from Kolombangara, western Solomon Islands. Archaeol Oceania 40:14–20.

Terrell J. 1989. Commentary: what Lapita is and what Lapita isn't. Antiquity 63: 623–626.

Terrell J. 2004. The 'sleeping giant' hypothesis and New Guinea's place in the prehistory of Greater Near Oceania. World Archaeol 36:601–609.

Terrill A, Dunn M. 2003. Orthographic design in Solomon Islands: the social, historical, and linguistic situation of Touo (Baniata). Written Lang Lit 6:177–192.

Walsh RJ, Kooptzoff O. 1955. The Blood groups of some native inhabitants of the British Solomon Islands. Oceania 25:188–193.

Wickler S, Spriggs M. 1988. Pleistocene human occupation of the Solomon Islands, Melanesia. Antiquity 62:703–706.

Wright S. 1955. Classification of the factors of evolution. Cold Spring Harbor Symp Quant Biol 20:16–24.

Y Chromosome Consortium. 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome Res 12:339–348.