Am. J. Hum. Genet. 63:1793–1806, 1998

European Y-Chromosomal Lineages in Polynesians: A Contrast to the Population Structure Revealed by mtDNA

Matthew E. Hurles,[1] Catherine Irven,[2] Jayne Nicholson,[2] Paul G. Taylor,[1] Fabricio R. Santos,[3,*] John Loughlin,[2] Mark A. Jobling,[1] and Bryan C. Sykes[2]

[1]Department of Genetics, University of Leicester, Leicester; and [2]Cellular Genetics Group, Institute of Molecular Medicine, and [3]CRC Chromosome Molecular Biology Group, Department of Biochemistry, University of Oxford, Oxford

Summary

We have used Y-chromosomal polymorphisms to trace paternal lineages in Polynesians by use of samples previously typed for mtDNA variants. A genealogical approach utilizing hierarchical analysis of eight rare-event biallelic polymorphisms, seven microsatellite loci, and internal structural analysis of the hypervariable minisatellite, MSY1, has been used to define three major paternal-lineage clusters in Polynesians. Two of these clusters, both defined by novel MSY1 modular structures and representing 55% of the Polynesians studied, are also found in coastal Papua New Guinea. Reduced Polynesian diversity, relative to that in Melanesians, is illustrated by the presence of several examples of identical MSY1 codes and microsatellite haplotypes within these lineage clusters in Polynesians. The complete lack of Y chromosomes having the M4 base substitution in Polynesians, despite their prevalence (64%) in Melanesians, may also be a result of the multiple bottleneck events during the colonization of this region of the world. The origin of the M4 mutation has been dated by use of two independent methods based on microsatellite-haplotype and minisatellite-code diversity. Because of the wide confidence limits on the mutation rates of these loci, the M4 mutation cannot be conclusively dated relative to the colonization of Polynesia, 3,000 years ago. The other major lineage cluster found in Polynesians, defined by a base substitution at the 92R7 locus, represents 27% of the Polynesians studied and, most probably, originates in Europe. This is the first Y-chromosomal evidence of major European admixture with indigenous Polynesian populations and contrasts sharply with the picture given by mtDNA evidence.

Received May 18, 1998; accepted for publication September 22, 1998; electronically published December 2, 1998.

Address for correspondence and reprints: Dr. Mark A. Jobling, Department of Genetics, University of Leicester, University Road, Leicester, LE1 7RH, UK. E-mail: maj4@le.ac.uk

* Present affiliation: Departimento de Biologia Geral, Instituto Ciências Biologicas/Universidade Federal de Minas Gerais, Minas Gerais, Belo Horizonte, Brazil.

© 1998 by The American Society of Human Genetics. All rights reserved.
0002-9297/98/6306-0026$02.00

Introduction

The Polynesians were skillful ocean-going navigators at a time when the Greeks and Romans were little more than coastal sailors. Prior to 1500 A.D. they constituted the most geographically widespread people on Earth (Bellwood 1987); their remarkable settlement of the remote Pacific islands has been extensively studied from the viewpoint of the archaeological record and cultural and linguistic affiliations (Jennings 1979). The Pacific islands have been traditionally classified into three geographical areas. Melanesia includes Papua New Guinea and islands to the east, including the Solomon Islands and Vanuatu. To the north of Papua New Guinea, the dispersed archipelagos of Micronesia stretch from Palau in the west to Kiribati in the east. The islands to the east of Fiji constitute Polynesia, a vast ocean triangle with sides ~6,500 km in length and apices at Hawaii, Aotearoa (New Zealand), and Rapanui (Easter Island).

Archaeology, anthropology, and linguistics (Bellwood 1987, 1989, 1991) have been used to construct the following picture of the region's prehistory. Approximately 30,000–50,000 years before the present (YBP), australoid hunter-gatherers moved through Southeast Asia into the Sahul landmass, which comprised the present islands of Papua New Guinea and Australia. These peoples are thought to be ancestral to the extant populations of native Australians and Papuan highlanders. The second major prehistoric migration into the region is associated with the expansion of a mongoloid Austronesian-speaking population, which began 5,000–6,000 YBP in Taiwan and coastal Southeast China. Over a relatively short period, their highly developed navigational skills allowed them to settle the islands of Melanesia, beyond the Solomon Islands, that had not previously been colonized by the australoid population. By 3,000 YBP, Fiji in eastern Melanesia and Samoa and Tonga in western Polynesia had been reached. This initial colonization of island Melanesia is associated with the Lapita culture, characterized by its distinctive pottery, and was quickly followed by the migration of other Melanesian populations that had acquired ocean-going skills. Colonization of remote Oceania followed a lengthy period of

adaptation in western Polynesia, during which a distinct Polynesian culture developed. The Marquesas were reached by 300 A.D., followed by Hawaii (by 500 A.D.), Rapanui (by 800 A.D.), and, finally, Aotearoa (by 1200 A.D.).

Linking of the proto-Polynesians to the Lapita culture, whose sites are found in island Melanesia, suggests relatively little genetic contribution of australoid peoples to Polynesians. Rather, it has been thought that the Polynesians had their origins in the relatively rapid maritime expansion from Southeast Asia (Bellwood 1979). An alternative model postulates that the proto-Polynesians evolved within Melanesia, from the resident population that had been there for ⩾30,000 years (Terrell 1986). A third model proposes a substantial colonization of Polynesia from the Americas (Heyerdahl 1950). In addition, during the past 300 years there has been substantial contact with Europeans, who may have contributed genetically to the extant population. Analysis of classic nuclear-encoded markers in Polynesians has weakly supported a Southeast Asian origin with perhaps some Melanesian admixture (Hill and Serjeantson 1989), a contribution that has been further supported by the discovery, in Polynesians, of a specific thalassemia allele of Melanesian origin (Hill et al. 1985). There has been no strong support for an American origin, although data from classic loci were never able to exclude this possibility because Native Americans themselves have an origin in Asia.

In contrast to these uncertainties, studies utilizing mtDNA have been particularly informative (Lum et al. 1994; Melton et al. 1995; Redd et al. 1995; Sykes et al. 1995). A common lineage cluster comprising three or four characteristic base substitutions in the control region and a 9-bp deletion elsewhere in the mitochondrial genome has been found in 94% of all Polynesian mtDNA samples. It has also been found at moderate frequency in coastal Papua New Guinea. Ancestral haplotypes have been traced to Indonesia, the Philippines, and Taiwan (Melton et al. 1995; Sykes et al. 1995). Other maternal lineages identified in Polynesians (Lum et al. 1994) have been confirmed as showing a Melanesian australoid admixture of ~4% (Sykes et al. 1995). A common finding in all genetic studies has been a striking lack of diversity in Polynesians, compared with source populations in Melanesia and Southeast Asia (Flint et al. 1989; Lum et al. 1994; Sykes et al. 1995). Together with a cline of diversity within mitochondrial lineage groups, from high in western Polynesia to low in eastern Polynesia (Sykes et al. 1995), this suggests that there have been, not surprisingly, severe population bottlenecks during the colonization of Polynesia. There is no mtDNA evidence for a substantial input from either the Americas or Europe.

Most of the paternally inherited Y chromosome is haploid and thus escapes recombination. Mutations on this portion of the chromosome represent a simple record of its evolution, which can be used to address questions of human population structure (Jobling and Tyler-Smith 1995). Modern Y chromosomes coalesce back to a common ancestor who has been dated, in independent studies, to 188,000 YBP (Hammer 1995), 37,000–49,000 YBP (Whitfield et al. 1995), or ~170,000 YBP (Underhill et al. 1997). Mating practices, the cultural phenomenon of patrilocality, and the small effective population size of the Y chromosome result in a high degree of geographical differentiation, which has been utilized to investigate prehistoric migration events (Underhill et al. 1997; Zerjal et al. 1997). It is the presence of different Y-chromosomal lineage clusters in Southeast Asia, the Americas, and Europe that allows us to investigate the origins of Polynesian Y chromosomes. The Y chromosome is likely to be more sensitive than other loci to certain kinds of admixture—for example, recent male-dominated admixture between populations normally separated by geographical distance.

The Y chromosome contains a wealth of different polymorphic systems with different mutational mechanisms and rates that vary from $\sim 5 \times 10^{-7}$/locus/generation, for base substitutions (Hammer 1995), to a few percent per generation, for the minisatellite MSY1 (Jobling et al. 1998). Initial attempts to use the Y chromosome for regional evolutionary studies were hampered by a lack of well-characterized polymorphisms (Spurdle and Jenkins 1992; Spurdle et al. 1994). Recently, however, a number of studies have successfully addressed questions of regional prehistory by use of the Y chromosome (Underhill et al. 1997; Zerjal et al. 1997). It has been suggested that the best way to utilize the information content from the different polymorphic systems is to adopt a genealogical approach based on a hierarchical analysis of different marker systems (Jobling and Tyler-Smith 1995; Santos and Tyler-Smith 1996; de Knijff et al. 1997). This entails subdivision of sets of chromosomes into distinct lineage clusters defined by compound haplotypes (haplogroups) of rare-event biallelic polymorphisms, followed by assaying of diversity within such haplogroups by use of more variable loci such as mini- and microsatellites. This provides information about the demographic history of haplogroups while minimizing the effect of recurrent mutation in the multiallelic polymorphic systems.

This study complements the mtDNA analysis, by examining the paternal lineages of two populations, one from coastal Papua New Guinea and the other from Polynesia (Cook Islands), by use of Y-chromosomal markers that can be assayed by PCR. Thirty of the 33 Cook Islands samples have the common Polynesian mtDNA haplotype characterized by control-region transitions at positions 16189, 16217, 16247, and 16261.

Seven base substitutions and one insertion/deletion were used to distinguish 10 possible Y-chromosome haplogroups, and diversity within observed haplogroups was assayed by use of seven microsatellite loci and the minisatellite MSY1.

A well-documented and well-dated history suggesting simple population movements in Oceania has been drawn from studies of achaeology, linguistics, and maternal lineages; Oceania thus represents an ideal region for studies of paternal lineages. This is the first Y-chromosomal study to reveal substantial European admixture within the Polynesian Y-chromosomal pool.

Subjects and Methods

Subjects

The DNA samples used in this study were provided by 91 individuals from two locations in the Pacific, all of whom had agreed to take part in a genetic survey. Polynesian samples came from Rarotonga in the Cook Islands in central Polynesia. Melanesian samples were from Port Moresby in Papua New Guinea.

Biallelic-Polymorphism Typing

All of these polymorphisms can be typed by use of similar PCR protocols. Although the cycling programs differ, the reaction volume (10 μl) and composition are the same. Samples (10–20 ng) of genomic DNA were added to a PCR buffer, as described by Jeffreys et al. (1990), with the addition of 1 μM of each primer and 0.5 U of *Taq* polymerase (Advanced Biotechnologies).

All PCR reactions were done in 96-well Thermowell M microtiter plates in an MJR PTC-200 machine. For RFLP analysis, 1 U of the appropriate restriction enzyme in 10 μl of 2 × digestion buffer was added directly to the PCR reaction and incubated at the appropriate temperature for 2 h. All 96 digest/PCR products were run out on a single agarose gel (1%–3%; Seakem agarose [FMC]), in 1 × Tris-borate EDTA (TBE), by means of gel tanks designed and made in house. All pipetting operations were performed with 12-channel pipettes.

A Y *Alu* polymorphism (YAP) was typed according to the procedure of Hammer and Horai (1995); SRY-1532 (identical to SRY 10,831 of Whitfield et al. 1995) was typed according to the procedure of Kwok et al. (1996); and SRY-2627 was typed according to the procedure of Veitia et al. (1997), who referred to it as "SRY-2628." *DYS199* was typed by use of the PCR primers *5*-TAATCAGTCTCCTCCCAGCA-3′ and 5′-AGGTACCAGCTCTTCCCAATT-3′; a cycling program of 94°C for 20 s, 59°C for 20 s, and 72°C for 30 s, repeated 36 times; and the restriction enzyme *Mfe*I. SRY-3225 (identical to SRY 9,138 of Whitfield et al. 1995) was typed by use of the primers 5′-CAACTGTTGAGAAATAGTCATC-3′ and 5′-CCCAGATGCATATATTACAGG-3′; a cycling program of 94°C for 30 s, 58°C for 30 s, and 72°C for 60 s, repeated 34 times; and the restriction enzyme *Hae*III. 92R7 (Mathias et al. 1994) was also typed by use of a PCR-RFLP assay (M. E. Hurles and C. Tyler-Smith, unpublished data). The M4 amplicon (*DYS234*), which has been shown to contain a site polymorphic in Oceanic people (Underhill et al. 1997), was typed by PCR-RFLP analysis, with PCR conditions of 94°C for 30 s, 60°C for 30 s, and 72°C for 45 s, repeated 34 times, and *Nde*I digestion. M9 was amplified by use of the conditions described by Underhill et al. (1997), and the polymorphic site was assayed by digestion with *Hin*fI.

The polymorphism SRY-2627 was typed only on those chromosomes that had been shown to carry the 92R7 (1) allele, since this polymorphism has been shown to occur only on this chromosomal background (M. E. Hurles, unpublished data). Polymorphisms known to occur only on a YAP^+ background were not typed in the sample set, since it contained no YAP^+ chromosomes.

Y chromosomes having the Tat polymorphism (Zerjal et al. 1997) constitute a subset of 50f2/C-deletion chromosomes (Jobling et al. 1996). 50f2/C-deletion chromosomes are identified by the absence of the 203-bp MSY1 homologue, which is coamplified in the MSY1-flanking PCR prior to minisatellite variant repeat–PCR (MVR-PCR) (Jobling et al. 1998). No such deletion chromosomes were detected in the MSY1-flanking reactions, and, therefore, the Tat polymorphism itself was not typed in these samples.

The allelic states of each haplogroup are as follows (in the order *DYS199,* YAP, SRY-3225, SRY-1532, 92R7, M4, SRY-2627, and M9): haplogroup 1 chromosomes have the compound haplotype 00011001, those of haplogroup 2 have 00010000, those of haplogroup 3 have 00001001, those of haplogroup 24 have 00010101, and those of haplogroup 26 have 00010001. Binary nomenclature indicates ancestral (0) and derived (1) forms of the polymorphism: 92R7, YAP (Jobling and Tyler-Smith 1995), SRY-1532, SRY-3225 (Whitfield et al. 1995), *DYS199* (Underhill et al. 1996), M4, M9 (Underhill et al. 1997), and SRY-2627 (Veitia et al. 1997).

Microsatellite Typing

The following seven highly polymorphic Y chromosome–specific microsatellites were analyzed: *DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392,* and *DYS393;* all have tetranucleotide-repeat units, except for *DYS392,* which has a trinucleotide-repeat unit. Primer sequences are as described elsewhere (de Knijff et al. 1997), except for the forward primer of *DYS389II,* which was 5′-TCCAACTCTCATCTGTATTATCTATGTG-3′. PCR reactions were performed in a 15-μl re-

action volume containing 6 pmol of each primer, 200 μM of each dNTP, 1× GeneAmp PCR buffer II, 1.5 mM $MgCl_2$, 0.375 U of Ampli*Taq* Gold (PE Applied Biosystems), and 50 ng of DNA. The forward primer was 5′-labeled with tetrachlorofluoroscein (*DYS19, DYS389I,* and *DYS389II*), carboxyfluoroscein (*DYS390, DYS391,* and *DYS393*), or 4,7,2′,4′,5′,7′-hexachloro-6-carboxyfluoroscein (*DYS392*). The general PCR profile was as follows: 94°C for 14 min, annealing temperature (AT) for 1 min, 72°C for 0.5 min, and 94°C for 1 min, with the latter three steps repeated 35 times; AT for 1 min; and 72°C for 10 min. Annealing temperatures were as follows: 54°C for *DYS19,* 51°C for *DYS389I,* 55°C for *DYS389II,* 56°C for *DYS390,* 55°C for *DYS391,* 53.5°C for *DYS392,* and 56°C for *DYS393.*

The PCR products from all seven individual reactions were diluted and pooled together with a size-standard GeneScan-500 TAMRA, so that an individual's seven-locus microsatellite haplotype could be read from a single gel track (Reed et al. 1994). The pooled products were run on denaturing 6% acrylamide gels in 1 × TBE on an Applied Biosystems 373A DNA sequencer. Gels were analyzed by ABI PRISM GeneScan Analysis 2.0.2, and samples were genotyped by Genotyper 1.1 (both from PE Applied Biosystems). By means of GAS (genetic-analysis system [A. Young, personal communication]), the discrete distributions of allele lengths were then put into allele bins, which were assigned numbers according to their sizes, with the smallest allele being denoted "1." To account for the *DYS389I* variable array being contained inside the *DYS389II* amplicon (Cooper et al. 1996), the length of the *DYS389I* product was subtracted from that of *DYS389II,* prior to placement in allele bins. The correspondence between our allele definitions and the repeat-unit number of previously published sources (de Knijff et al. 1997) is as follows: allele 1 corresponds to *DYS19,* repeat-unit number 11; *DYS389I,* repeat-unit number 9; *DYS389II,* repeat-unit number 15; *DYS390,* repeat-unit number 20; *DYS391,* repeat-unit number 9; *DYS392,* repeat-unit number 11; and *DYS393,* repeat-unit number 12.

Minisatellite Coding

Three-state MSY1 MVR-PCR of repeat types 1, 3, and 4 was performed according to the method of Jobling et al. (1998). A code—for example, (1)20(3)35(4)20—represents the minisatellite array as blocks of different repeat-unit variants, in this case 20 type 1 repeats are followed by a block of 35 type 3 repeats followed by a block of 20 type 4 repeats. Modular-structure nomenclature of, for example, the form (1,3,4) refers to a block of type 1 repeats followed by a block of type 3 repeats followed by a block of type 4 repeats.

Sequence Analysis of MSY1 Null Repeats

MSY1 alleles containing null repeats were analyzed by fluorescent automated sequencing of PCR products generated by means of the primers $Y1A^+$ and $Y1B^+$.

Phylogenetic Analysis

The haplogroup tree was based on that of Jobling and Tyler-Smith (1995) and used only those polymorphisms that are presently typable by PCR and incorporated the new polymorphisms *DYS199,* SRY-2627, SRY-3225, M9, and M4, by typing them on a fully representative panel of chromosomes previously typed for all polymorphisms in the original tree. Minimum spanning microsatellite networks were constructed for chromosomes belonging to individual haplogroups, under the assumption of a single-step mutation process, by means of pairwise comparisons between all haplotypes. Initially all haplotypes differing by single mutational steps are linked, and then most-parsimonious steps of greater mutational distance are considered, in turn, until all haplotypes are included in the network. All most-parsimonious steps linking any given haplotype into the network are included (Zerjal et al. 1997).

The root for the maximum-parsimony tree used in dating haplogroup 24 (Bertranpetit and Calafell 1996) was chosen by combining the mode allele length of each individual locus into a seven-locus haplotype. This haplotype is identical to haplotype 44, which is itself represented more than once in the data set. The network gives additional support for this haplotype being a plausible root because the latter occupies a central position within the network and has the greatest number of links to other haplotypes.

Results

A total of 33 unrelated Cook Islander and 58 unrelated Papua New Guinean samples were assayed for all polymorphisms. Of the 10 haplogroups defined by the rare-event biallelic polymorphisms, only 5 (i.e., haplogroups 1-3, 24, and 26) were observed in this sample set. Haplogroup 1 and haplogroup 3 chromosomes are found in Polynesians alone, haplogroup 2 and haplogroup 26 chromosomes are found in both Polynesians and Melanesians, and haplogroup 24 chromosomes are found only in Melanesians. Haplogroup 2 is the most ancestral. These data are summarized in figure 1.

A total of 91 samples gave full biallelic-polymorphism data, and 79 of them also had full seven-locus microsatellite haplotypes; of these, 77 gave MSY1 codes. MSY1 subtypes were defined initially on the basis of modular structures of arrays of different repeat types that are represented more than once in the sample set. Three subtypes were defined: (1), (1,3,4); (2), (3,1,3,4);

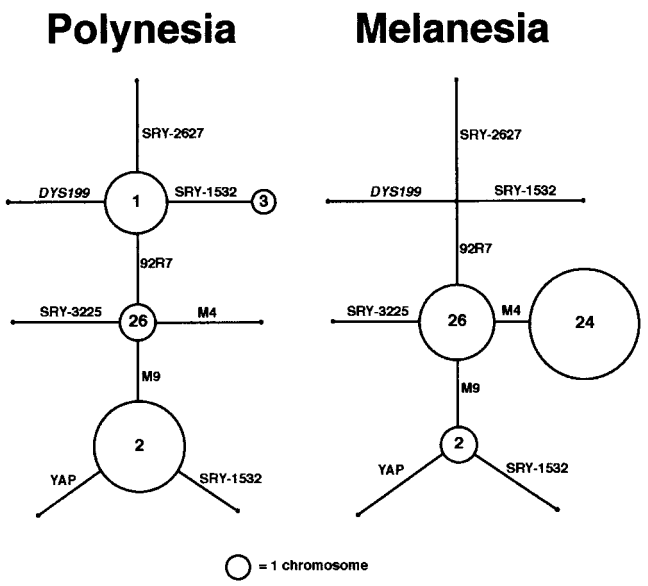


Figure 1 Haplogroup distributions within Polynesians and Melanesians: unrooted trees based on that in the work of Jobling and Tyler-Smith (1995), with new mutations added (see Subjects and Methods). Circles represent haplogroups, and lines represent single mutational steps between them. Lines are labeled with the polymorphisms that they represent. Circle area is proportional to the number of chromosomes within a given haplogroup. Dots at the ends of the lines denote haplogroups not represented in that population. Numbers within circles indicate haplogroup number. All 91 chromosomes with full rare-event biallelic-polymorphism data are included, irrespective of the status of their microsatellite or minisatellite data.

and (3), a group of MSY1 codes that all end with the modular structure (…4,0,4). An additional subtype $(3,1,3^{+},4^{-})$ was defined when nonoverlapping block-size ranges within the (3,1,3,4) chromosomes were shown to be statistically significant ($P < .001$; Student's *t*-test). An example of a code of the former MSY1 subtype is (3)3(1)13(3)61(4)8, and an example of the latter MSY1 subtype is (3)1(1)13(3)39(4)16. These typing data are summarized in figure 2 and are represented diagrammatically in figure 3. Note that the percentages of haplogroup frequencies given in the text refer to all 91 chromosomes studied, whereas the pie charts in figure 3 consider only those chromosomes for which MSY1 codes were available. Sample CI185 is classified as a (…4,0,4) chromosome by virtue of its microsatellite haplotype, the characteristic null at the start of the array, and the similarity between its repeat-block sizes and those in the other (…4,0,4) chromosomes. The null repeats appear to have been converted back to type 4 repeats in this chromosome, perhaps by a repeat-homogenization process (Bouzekri et al. 1998). Complete typing results, including microsatellite haplotypes, are tabulated in the Appendix (table A1).

Minimum spanning microsatellite networks for each haplogroup were constructed with the 79 samples for which there were full microsatellite- and biallelic-polymorphism data. MSY1 subtypes were then mapped onto these networks. The final networks are shown in figure 4.

It is important to distinguish between MSY1 subtypes that represent very rare mutational events and those that may have independently evolved multiple times. Subtypes (3,1,3,4) and (1,3,4) are common to haplogroups 24 and 26. They have also both been observed in other haplogroups, which were not found in the present study (Jobling et al. 1998). Given that a haplogroup defined by a unique mutational event must have been founded by a single chromosome, these MSY1 subtypes have probably arisen independently several times. In addition, there is, in the microsatellite networks, no tight clustering of chromosomes having these MSY1 subtypes. By contrast, the (…4,0,4) and $(3,1,3^{+},4^{-})$ MSY1 subtypes have been found only in haplogroup 2 and haplogroup 26 chromosomes, respectively, and only in this region of the world, suggesting that each has a unique origin. In addition, the tight clustering of these chromosomes within the microsatellite networks further supports this conclusion.

Haplogroup 1 Chromosomes

Haplogroup 1 chromosomes were found only in Polynesians, in whom they comprised 9 (27%) of 33 chromosomes in the sample. The nine haplogroup 1 chromosomes have eight different microsatellite haplotypes, and all belong to the same MSY1 subtype, (1,3,4).

Haplogroup 2 Chromosomes

Haplogroup 2 chromosomes are found in both Melanesians and Polynesians, in whom they comprise 3 (5%) of 58 and 19 (58%) of 33 of the Y chromosomes, respectively. Although this haplogroup is found in both populations, it can also be found in most areas of the world (C. Tyler-Smith and A. Pandya, personal communication) and thus, on its own, cannot be taken as evidence of recent common ancestry for the two populations. There are 11 different microsatellite haplotypes in the 22 chromosomes analyzed here, and two different MSY1 subtypes—(…4,0,4) and (3,1,3,4)—are represented more than once within this haplogroup. Only one of these MSY1 subtypes, the (…4,0,4) subtype, is found in both Polynesians and Melanesians.

The (…4,0,4) chromosomes found in Melanesians have modular structures different than those found in Polynesian (…4,0,4) chromosomes. The Polynesian chromosomes all have the modular structure (0,1,3,4,0,4), whereas the two (…4,0,4) chromosomes in Melanesians have the modular structures (3,1,3,4,0,4) and (0,3,1,3,4,0,4,0,4). To determine whether these

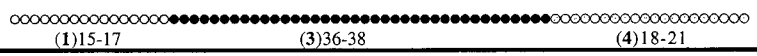
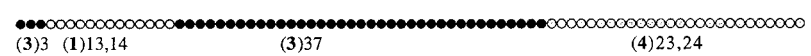
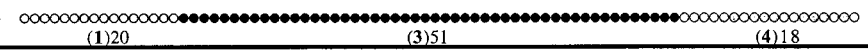
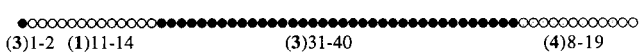
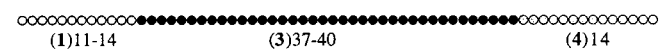
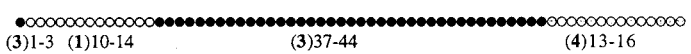
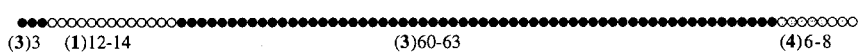
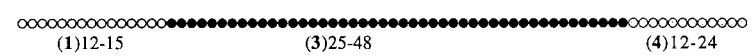
Haplogroup	MSY1 subtype	MSY1 code range	PNG	CI	Different MSY1 codes	Different microsat haplotypes
1	(1,3,4)		0	9	7	8
2	(...4,0,4)	Variable modular structure	2	15	11	7
	(3,1,3,4)		0	2	2	2
	Others (singletons)	Variable	1	2	3	3
3	(1,3,4) singleton		0	1	1	1
24	(3,1,3,4)		21	0	21	18
	(1,3,4)		4	0	3	3
	Others (singletons)	Variable	3	0	3	3
26	(3,1,3,4)		5	0	5	5
	(3,1,3^{+},4^{-})		4	3	5	4
	(1,3,4)		4	0	4	4
	Others (singletons)	Variable	0	1	1	1
		Total	**44**	**33**	**66**	**59**

Figure 2 MSY1 structural diversity within haplogroups defined by rare-event biallelic polymorphisms. Each circle represents a single repeat unit: black circles represent type 3 repeats, white circles represent type 1 repeats, and gray circles represent type 4 repeats. Only the 77 chromosomes with full rare-event, biallelic-polymorphism, microsatellite, and minisatellite data are included. (PNG = Papua New Guineans; CI = Cook Islanders).

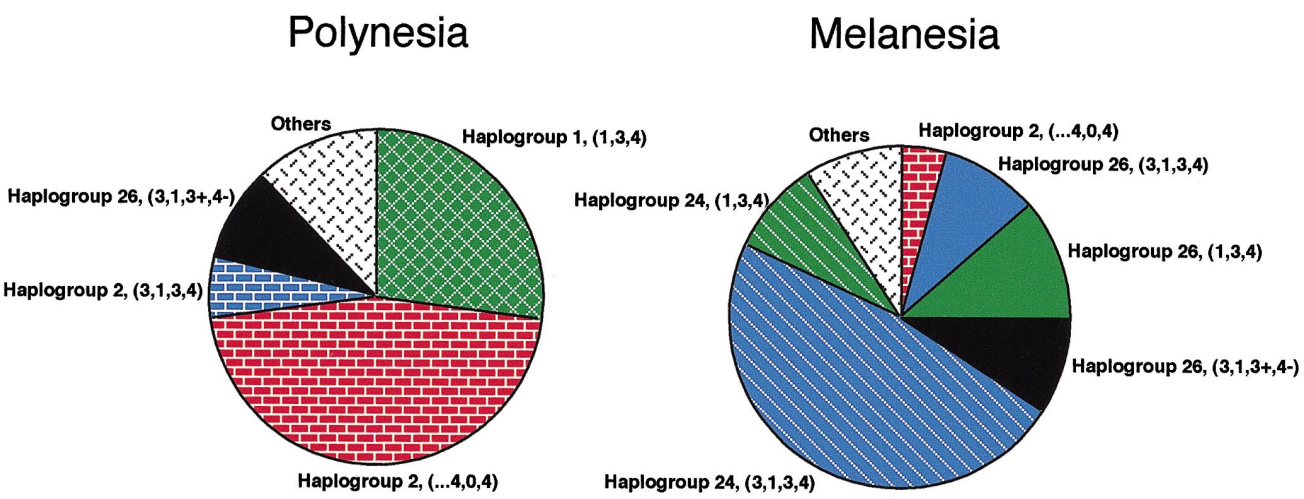


Figure 3 Distribution of MSY1 subtypes within Polynesians and Melanesians. Only the 77 chromosomes for which there are full biallelic-polymorphism and minisatellite-typing data are included. The background pattern represents the haplogroup, and color represents MSY1 subtype. Only those MSY1 subtypes that are represented more than once are displayed. "Others" includes all singleton MSY1 modular structures.

chromosomes do, indeed, belong to the same lineage cluster, the null repeats from a Cook Islander chromosome (CI140) and from a Papua New Guinean chromosome (7092) were sequenced and shown to be identical. The null repeats at the start of these arrays were both shown to be type 3 repeats modified by a T→A transversion at position 21 within the 25-bp repeat unit. The null repeats near the end of the array were found to be type 2 repeats, which are characterized by having a C→G transversion at position 13 within the repeat, relative to type 4 repeats (Jobling et al. 1998). Additional support for the common ancestry of these chromosomes comes from their association with a rare short *DYS390* microsatellite allele that is found almost exclusively in this region of the world (Forster et al. 1998).

The other MSY1 subtype found more than once within haplogroup 2 chromosomes is the (3,1,3,4) subtype, which is found in only two chromosomes (6%) in Polynesians. These two chromosomes, along with a third Polynesian haplogroup 2 chromosome that has neither the (3,1,3,4) nor the (…4,0,4) MSY1 modular structure, cluster together in the microsatellite network, perhaps suggesting a recent common ancestry.

Haplogroup 3 Chromosomes

Only a single haplogroup 3 chromosome has been found in this study—in Polynesians, in whom it is found in 1 (3%) of 33 chromosomes in the sample. Haplogroup 3 chromosomes are found in both Europe and Asia (as haplotype "1D" in the work of Hammer et al. [1998]) and have also been found at low frequency in Indonesians (M. E. Hurles, unpublished data).

Haplogroup 24 Chromosomes

This group of Y chromosomes was found in 37 (64%) of 58 chromosomes in the Papua New Guinean sample but not at all in the Polynesian sample. Previously, haplogroup 24 chromosomes had been found in Papua New Guineans and at low frequency in Indonesians (Underhill et al. 1997; M. E. Hurles, unpublished data). There are 24 different microsatellite haplotypes in the 30 chromosomes that have full microsatellite haplotypes. Two different MSY1 subtypes are found in these chromosomes. By far the most frequent MSY1 subtype is (3,1,3,4), found in 21 of 28 chromosomes with identified MSY1 structure. The (1,3,4) MSY1 subtype is found in four of the remaining seven chromosomes.

The question arises as to whether the mutation defining haplogroup 24 occurred prior to or after the colonization of Polynesia. The absence of such chromosomes in Polynesians might constitute significant evidence of reduced diversity. Consequently, microsatellite diversity was used to date the origin of these chromosomes, so as to compare this date with that of the initial migration into Polynesia, which, on the basis of archaeological evidence, has been deduced to be ~3,000 YBP. Two different dating methods were used.

In the first approach (Bertranpetit and Calafell 1996), a rooted maximum-parsimony tree (not shown) was constructed from the pairwise comparisons between compound microsatellite haplotypes. The root of the tree was placed at haplotype 44, on the basis of the mode-allele length at each locus (see Subjects and Methods section). Haplogroup 26 is ancestral, but it could not be used as an outgroup for the purpose of rooting, because

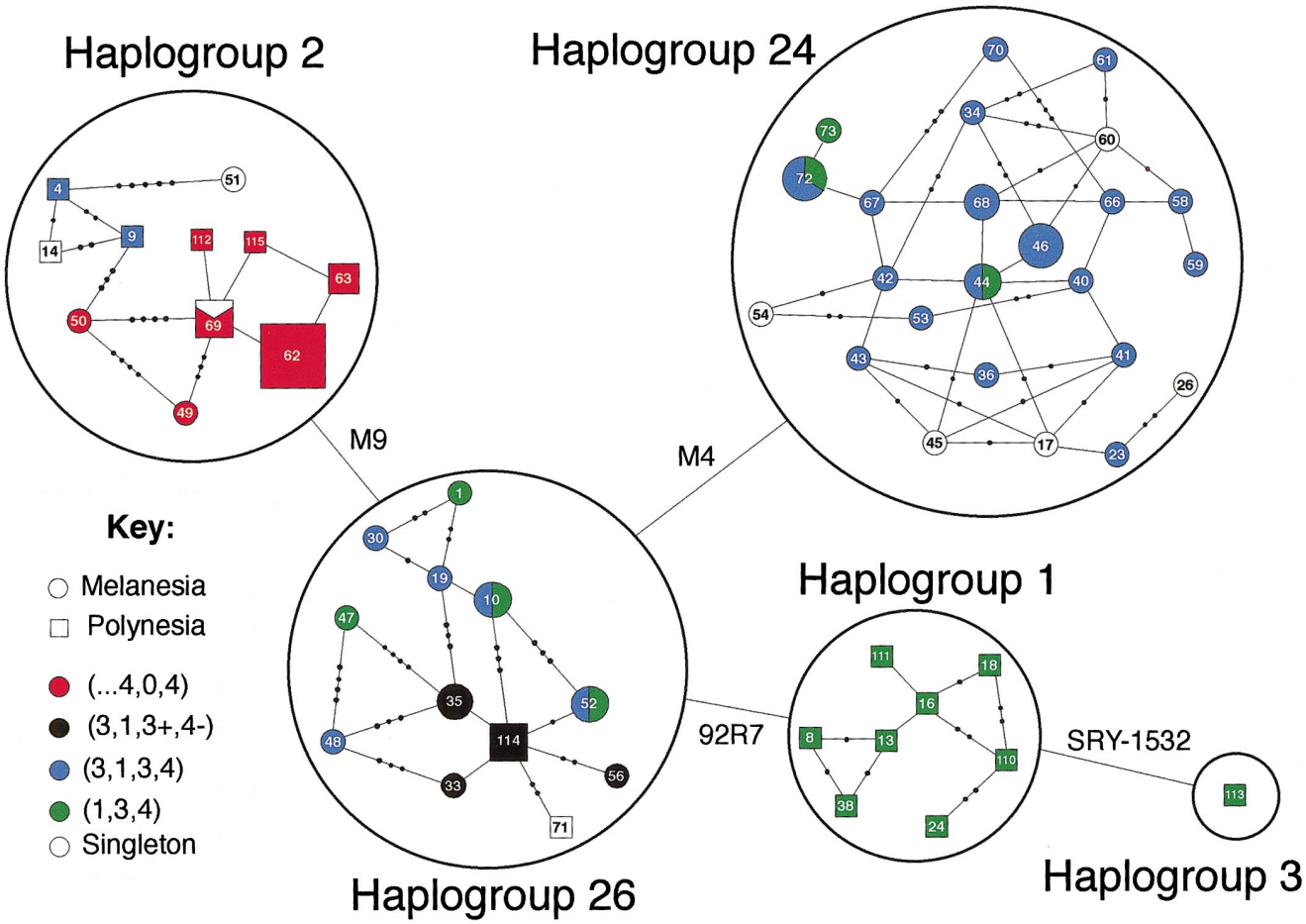


Figure 4 Microsatellite networks for individual haplogroups. Numbered circles and squares represent compound microsatellite haplotypes; small black circles represent intermediate haplotypes not observed in this sample set. Color indicates MSY1 subtype. All 79 chromosomes for which there are complete microsatellite and biallelic-polymorphism data are included.

there are multiple equally parsimonious links between the networks of the two haplogroups. The average number of mutations per locus is calculated from the root to every haplotype in the tree, with consideration of the number of times that that haplotype is represented in the sample. A generation time of 25 years (Thomas et al. 1998) and a mutation rate of 2.1×10^{-3}/locus/generation for Y-chromosome microsatellites (Heyer et al. 1997) was used to convert the calculated figure of 0.38 mutations/locus/generation to a time of origin of ~4,400 YBP for this base substitution. Factoring in the 95% confidence limits, $0.6–4.9 \times 10^{-3}$, of the microsatellite mutation rate gives an interval of 1,900–15,300 YBP (Heyer et al. 1997).

The second approach (Goldstein et al. 1996) relies on the variance of linked microsatellite loci and assumes a single-step–mutation model. As long as the observed variance can be shown to be significantly different from that expected at mutation/drift equilibrium, the time taken for the variance to accumulate since its initial value of zero can be calculated. Given a mutation rate of 2.1×10^{-3}/locus/generation and an effective population size of 4,500 (Goldstein et al. 1996), the variance expected at mutation/drift equilibrium is 9.4, with 95% confidence limits of 4.2–18.4. The variance observed in our sample is 0.45, with 95% confidence limits of 0.36–0.56. This gives a time to the origin of this haplogroup (i.e., when variance was zero) of ~5,500 YBP, with 95% confidence limits of 3,500–5,450 YBP. If the 95% confidence limits for the mutation rate are factored into the equations, an expanded interval of 2,000–30,400 years is obtained. Despite the difference in the various estimates of haplogroup age, it seems likely, on the basis of microsatellite haplotype diversity, that this mutation predates the migration into Polynesia, 3,000 YBP.

Preliminary attempts were also made to use MSY1 diversity for dating, by treating each block of repeat

types as an independent repeat locus, analogous to a microsatellite, with the mutation rate of each block proportional to the number of repeats within the block and with a mutation rate range of 2%–11% for the entire array (Jobling et al. 1998). By means of the first method described above, an age range of 1,200–12,700 YBP is obtained. The latter method gives an age range of 2,300–54,400 YBP. The midpoint mutation rate within the aforementioned range (6.5%) gives absolute ages for haplogroup 24 that agree very well with those from microsatellite dating (data not shown).

Haplogroup 2 is defined by the polymorphism SRY-1532, which, on the basis of both its wide geographic distribution (Jobling and Tyler-Smith 1995; C. Tyler-Smith and A. Pandya, unpublished data) and MSY1 diversity (Jobling et al. 1998), is thought to be a relatively old haplogroup and was not dated by means of the microsatellite data, because mutation/drift equilibrium renders such dating methods unable to resolve suitably far back in time. The sample size for haplogroup 1 is too small for a reasonable attempt at dating. We would also like to date the origin of the (…4,0,4) MSY1 subtype chromosomes; however, the small sample size and the obvious occurrence of a bottleneck in the Polynesian chromosomes would render any such dating inaccurate.

Haplogroup 26 Chromosomes

Haplogroup 26 chromosomes have been found in 3 (9%) of 33 chromosomes in Polynesians and in 13 (22%) of 58 chromosomes in Melanesians. Within this haplogroup, there are three main subgroups—(3,1,3,4), (1,3,4), and (3,1,3^+,4^-)—defined by different MSY1 subtypes. Of these subgroups, only the last is shared between the Melanesian and Polynesian populations; the others are found solely in Melanesians. As discussed above, the (3,1,3^+,4^-) subtype defines a unique Y-chromosomal lineage. There are three different microsatellite haplotypes among the four (3,1,3^+,4^-) chromosomes in Melanesians, whereas all three (3,1,3^+,4^-) chromosomes in Polynesians share the same microsatellite haplotype.

Discussion

The present study illustrates the informative capacity of a genealogical approach to Y-chromosomal analysis and is also the first application of MVR-PCR coding of the minisatellite MSY1 to address a specific issue in human evolution. If we consider all 77 chromosomes for which we have obtained complete typing data, there are 59 different microsatellite haplotypes and 66 different MSY1 codes. Although MSY1 modular structures can be good predictors of lineage clusters (Jobling et al. 1998; P. G. Taylor, unpublished data), no single polymorphic system can distinguish between all the lineage clusters observed in our sample set, and this emphasizes the advantage in the use of different polymorphic systems in conjunction. When the data from MSY1 and the seven-locus microsatellite haplotype are combined, 70 different compound haplotypes of 77 chromosomes are differentiated. Thus, almost all chromosomes can be distinguished from one another.

What can we say about Papua New Guinean Y-chromosomal diversity? Haplogroup 26 is well represented in coastal Papua New Guinea, and, to judge on the basis of its abundance in eastern Asia (M. E. Hurles and C. Tyler-Smith, unpublished data), it is probably derived from the Southeast Asian migration that contributed the majority of Polynesian mtDNA samples. Haplogroup 24 chromosomes, which represent the majority of Melanesian Y chromosomes, are derived from haplogroup 26, and their geographic distribution (Underhill et al. 1997) suggests that they originated in this region of the world, most probably in Papua New Guinea itself.

Haplogroup 24 is not found in the Polynesian sample, but, although the microsatellite diversity potentially indicates an origin for haplogroup 24 chromosomes that predates the colonization of Polynesia, 3,000 YBP, the confidence limits are such that we cannot be certain about this. In addition, even though the origin of haplogroup 24 chromosomes may predate the colonization of Polynesia, at the time of the colonization the frequency of haplogroup 24 is likely to have been substantially less than it is at the present time. Consequently, the absence of haplogroup 24 chromosomes from Polynesians cannot be taken as evidence of bottleneck events. However, there is a reduction of diversity, within both the (…4,0,4) and (3,1,3^+,4^-) MSY1 subtypes, between Papua New Guineans and Polynesians, and this is most probably due to the multiple population bottleneck events that accompanied the colonization of Polynesia. This picture of reduced diversity in Polynesians compared with Melanesians is common to nearly all genetic studies of the region, independent of the locus studied (Flint et al. 1989; Sykes et al. 1995).

If the blocks of MSY1 repeat-unit variants are considered to be independent loci with mutation rates proportional to their sizes, the use of MSY1 diversity to date haplogroup 24 generates ranges of ages that agree well with those derived from microsatellite data. In the future, greater empirical knowledge about the mutation dynamics of this locus will allow more-sophisticated dating analyses.

What are the possible origins of the three major lineage clusters found in Polynesians? Clearly, the haplogroup 2 chromosomes with the (…4,0,4) MSY1 subtype and the haplogroup 26 chromosomes with the (3,1,3^+,4^-) MSY1 subtype found in Polynesians share a recent common origin with those found in Melanesians.

Together, these account for 55% of the Polynesian Y chromosomes in this study.

Haplogroup 1 chromosomes, which comprise 27% of the Polynesian Y chromosomes in this study, are not found in our Melanesian sample. From where have these Y chromosomes come? Haplogroup 1 chromosomes are found at high frequency in Europeans and at low frequency in Asians and peoples of the Americas (Santos and Tyler-Smith 1996; C. Tyler-Smith and A. Pandya, personal communication). Within haplogroup 1, the MSY1 subtype (1,3,4) is found at appreciable frequencies only in Europeans and peoples of the Americas, with Asian chromosomes belonging almost exclusively to the (3,1,3,4) subtype (Jobling et al. 1998). In addition, in a microsatellite network of haplogroup 1 Y chromosomes, the Polynesian Y chromosomes cluster closely with the European and not with the Asian Y chromosomes (M. E. Hurles, unpublished data). Thus, we can discount an Asian origin for these haplogroup 1 chromosomes. The majority (90%) of indigenous Y chromosomes in South Americans have the derived form of the *DYS199* polymorphism (Underhill et al. 1996). Haplogroup 1 chromosomes do not carry this base substitution and, consequently, represent only a small minority of South American Y chromosomes. In contrast, haplogroup 1 chromosomes of the (1,3,4) MSY1 subtype are the most common type of Y chromosome in western Europeans (Jobling et al. 1998), representing approximately two-thirds of all Y chromosomes within this region (Santos and Tyler-Smith 1996; M. E. Hurles, unpublished data). Within our Polynesian sample, we do not observe any Y chromosomes that have the derived form of *DYS199*. Thus, in this study there is no evidence for a Native American contribution to the Polynesian Y-chromosomal pool. It therefore seems likely that the haplogroup 1 chromosomes found in the Polynesians have a recent European origin.

If the haplogroup 1 chromosomes are indeed European in origin, then we should also expect to see in Polynesians some other haplogroups that are found at appreciable frequencies in western European populations. Haplogroup 2 chromosomes belonging to the (3,1,3,4) subtype are the other major type of Y chromosome found in western Europeans, representing approximately one-quarter of all Y chromosomes from this population (Jobling et al. 1998). The Polynesian sample contains two of these chromosomes, and the Melanesian sample contains none.

Thus, the Polynesian sample contains two types of chromosomes that are neither found within the Melanesian sample nor known to be common in Southeast Asians in general (M. E. Hurles, C. Tyler-Smith, and A. Pandya, unpublished data). These two types of Y chromosome represent the most common types found in western Europe. Indeed, in Polynesians these chromosomes are found in approximately the same ratio at which they are present in western European populations, providing additional evidence for recent European admixture. In summary, 55% of Polynesian Y chromosomes can be traced to Melanesians and have Southeast Asian origins, 33% (i.e., haplogroup 1 and [3,1,3,4] haplogroup 2 chromosomes]) appear to be European in origin, and 12% remain of indeterminate origin.

This study illustrates the power of the phylogeographic approach to population-structure analysis using the Y chromosome. The use of microsatellite data alone in Y-chromosome studies of this region of the world will not differentiate between European and genuine Polynesian Y chromosomes. Attempts that investigate paternal relationships within this region of the world (Lum et al. 1998) but that do not take into account this substantial European contribution run the risk of obscuring, with the "noise" of recent admixture, the real patterns of prehistoric population movements and, thus, of potentially drawing spurious conclusions.

This study is also a dramatic example of the advantages of combining the Y-chromosomal results with mtDNA data from the same samples. Comparisons of Y-chromosomal and mtDNA data have previously been used to characterize ethnic introgression within Native American (Bianchi et al. 1997) and African-derived (Bravi et al. 1997) populations. In the Polynesian population studied here, although almost all maternal lineages are derived from native Polynesian ancestors, at least one-third of Y chromosomes are probably of recent European origin. Studies of human leukocyte antigen (HLA) haplotypes in Oceania have identified 5%–10% as being of recent caucasoid admixture (Serjeantson 1989). Analysis of autosomal loci, such as HLA, will always reflect an average of maternal and paternal contributions. By contrast, the current study, in combination with the earlier mtDNA analysis, vividly demonstrates the differential input from males and females. In the case of Polynesians, the predominantly paternal route for European admixture can be explained by the exclusively male composition of the postcontact groups, which included sailors, traders, whalers, and missionaries.

Acknowledgments

We thank Chris Tyler-Smith and Arpita Pandya, for unpublished data on biallelic polymorphisms; Manfred Kayser, for useful discussions; Stuart Bayliss, for automated DNA sequencing; and the MRC Molecular Haematology Unit and John Mitchell, for access to DNA collections. We are grateful to the Prime Minister's Office (Tere Tangiiti) and Health Department (Dr. George Koteka and Vaevae Pare) of the Government of the Cook Islands. This work was supported by the Wellcome Trust and the Medical Research Council. M.A.J. is a Wellcome Trust Career Development Fellow supported by grant 044910.

Appendix

Table A1

Rare-Event Biallelic-, Microsatellite-, and MSY1-Polymorphism Data for All Samples, Classified by Haplogroup and MSY1 Modular Structure

Population and Sample[a]	MSY1 Code	Microsatellite Haplotype	Haplotype Number
Haplogroup 1 (1,3,4):			
CI75A	(1)17(3)36(4)21	2236334	24
CI115	(1)15(3)37(4)20	2225332	16
CI183	(1)16(3)38(4)20	2214232	8
CI188	(1)17(3)38(4)18	2226342	18
CI194	(1)15(3)36(4)21	2224332	13
CI196	(1)16(3)38(4)20	3224232	38
CI120	(1)15(3)39(4)21	2236232	110
CI139	(1)17(3)38(4)18	2226342	18
CI149	(1)15(3)39(4)18	2225331	111
Haplogroup 2 (...4,0,4):			
CI140	(0)1(1)13(3)33(4)6(0)3(4)16	4131223	62
CI147	(0)1(1)13(3)33(4)5(0)4(4)15	4131323	63
CI156	(0)1(1)12(3)35(4)5(0)3(4)17	4231223	69
CI175	(0)1(1)13(3)33(4)6(0)3(4)16	4131223	62
CI180	(0)1(1)13(3)32(4)5(0)4(4)14	4131223	62
CI181	(0)1(1)13(3)33(4)6(0)3(4)16	4131223	62
CI198	(0)1(1)13(3)34(4)5(0)3(4)14	4131223	62
CI206	(0)1(1)13(3)33(4)5(0)4(4)15	4131323	63
CI153	(0)1(1)13(3)33(4)5(0)4(4)16	4131223	62
CI155	(0)1(1)12(3)34(4)5(0)3(4)16	4231223	69
CI186	(0)1(1)13(3)33(4)6(0)3(4)16	4131223	62
CI123	(0)1(1)13(3)33(4)6(0)3(4)16	4131223	62
CI138	(0)1(1)14(3)32(4)6(0)3(4)16	4231323	115
CI167	(0)1(1)13(3) >11 (4)5(0)3(4)16	4232223	112
CI185	(0)1.(1)13.(3)33.(4)25	4131223	62
6591	(3)2(1)12(3)29(4)4(0)5(4)24	3251213	49
7092	(0)1(3)1(1)11(3)29(4)7(0)3(4)9(0)3(4)11	3321212	50
Haplogroup 2 (3,1,3,4):			
CI142	(3)3(1)13(3)37(4)24	2125212	4
CI191	(3)3(1)14(3)37(4)23	2223212	9
Haplogroup 2 (others):			
CI145	(0)1(1)12(3)38(4)11	4231223	69
CI192	(3)3(1)1(1/3?)1(1)5(1/3?)2(1)4(3)28(4)24	2225211	14
14791	(1)14(3/4)21(0)27	3324331	51
Haplogroup 3 (1,3,4):			
CI151	(1)20(3)51(4)18	3234213	113
Haplogroup 24 (3,1,3,4):			
2592	(3)1(1)13(3)36(4)12	4225332	66
2792	(3)1(1)12(3)34(4)15	3335322	53
2892	(3)1(1)13(3)34(4)17	3128232	34
2992	(3)1(1)13(3)40(4)12	4226232	67
3091	(3)1(1)12(3)31(4)16	2236333	23
3092	(3)1(1)14(3)36(4)8	4125332	58
3392	(3)2(1)13(3)33(4)15	3226233	43
3792	(3)2(1)11(3)39(4)15	3215233	36
4292	(3)1(1)13(3)32(4)12	3225333	41
4992	(3)1(1)13(3)35(4)17	4129332	61
5092	(3)1(1)13(3)36(4)19	4226332	68
5492	(3)1(1)12(3)34(4)13	3227332	46
6092	(3)1(1)12(3)37(4)16	3227332	46
8092	(3)1(1)13(3)37(4)12	4125432	59
8392	(3)1(1)13(3)38(4)13	3226232	42

(*continued*)

Table A1 (continued)

Population and Sample[a]	MSY1 Code	Microsatellite Haplotype	Haplotype Number
9391	(3)1(1)11(3)36(4)13	5226232	72
09921	(3)1(1)12(3)34(4)14	4226332	68
12191	(3)2(1)11(3)37(4)15	3226332	44
12591	(3)1(1)12(3)36(4)16	3225332	40
12991	(3)2(1)11(3)35(4)16	4233232	70
13191	(3)1(1)12(3)33(4)13	3227332	46
Haplogroup 24 (1,3,4):			
14991	(1)14(3)37(4)14	3226332	44
6991	(1)12(3)38(4)14	6226232	73
7292	(1)11(3)40(4)14	5226232	72
9192	(1)12(3)38(4)14	5226232	72
Haplogroup 24 (others):			
6192	(3)1(1)13(3)35(4)1(3)1(4)15	4127332	60
8692	(3)1(1)11(0)1(3)32(4)14	2226333	17
14591	(3)1(1)12(3/1)2(3)34(4)16	3336232	54
Haplogroup 24 (MSY1 undetermined):			
9092	x	2245233	26
3691	x	3226334	45
Haplogroup 24 (microsatellites incomplete):			
1892	(3)1(1)12(3)>30(4)11	33x6333	103
4092	(3)1(1)9(3)34(4)16	42x6332	106
4492	(3)1(1)13(3)34(4)16	42x5342	105
4692	(3)1(1)12(3)>27(4)14	42x6343	107
5292	(3)1(1)9(3)36(4)14	32x6232	99
6292	(3)1(0)1(1)11(0/3)1(3)38(4)15	32x5233	96
7992	(3)1(1)13(3)34(4)17	41x9332	104
Haplogroup 26 (1,3,4):			
1992	(1)12(3)25(4)24	3235162	47
5192	(1)15(3)47(4)12	3325232	52
9791	(1)14(3)48(4)14	1333332	1
13591	(1)14(3)46(4)15	2223232	10
Haplogroup 26 (3,1,3^+,4^-):			
CI128	(3)2(1)13(3)63(4)6	3125232	114
CI135	(3)3(1)14(3)59(4)8	3125232	114
CI166	(3)2(1)13(3)63(4)6	3125232	114
2692	(3)3(1)13(3)61(4)8	3135232	35
9191	(3)3(1)13(3)61(4)8	4115132	56
5692	(3)3(1)12(3)63(4)6	3126232	33
13991	(3)3(1)14(3)60(4)8	3135232	35
Haplogroup 26 (3,1,3,4):			
4591	(3)1(1)13(3)39(4)14	2233232	19
5791	(3)1(1)12(3)39(4)14	2223232	10
7792	(3)3(1)10(3)37(4)16	3246242	48
8792	(3)1(1)14(3)44(4)13	3325232	52
14391	(3)1(1)13(3)39(4)15	2433232	30
Haplogroup 26 (others):			
CI190	(3)1(1)11(0)1(1)1(3)51(4)15	5125231	71
Haplogroup 26 (microsatellites incomplete):			
3192	(3)1(1)14(3)>25(4)15	43x5232	109
5392	(3)1(1)13(3)>25(4)12	23x5253	86
5592	(3)1(1)14(3)42(4)14	33x5232	101
6692	(3)3(1)13(3)>33(4)8	31x6232	93
8492	(3)5(1)11(3)>10(4)14	22x5242	85

[a] Sample numbers with the prefix "CI" are from the Cook Islands; all others are from Papua New Guinea.

References

Bellwood PS (1979) Man's conquest of the Pacific: the prehistory of Southeast Asia and Oceania. Oxford University Press, Oxford

——— (1987) The Polynesians. Thames & Hudson, London

——— (1989) The colonization of the Pacific: some current hypotheses. In: Hill AVS, Serjeantson SW (eds) The colonization of the Pacific: a genetic trail. Clarendon Press, Oxford, pp 1–60

——— (1991) The Austronesian dispersal and the origin of languages. Sci Am (July): 70–75

Bertranpetit J, Calafell F (1996) Genetic and geographic variability in cystic fibrosis: evolutionary considerations. In: Chadwick D, Cardew G (eds) Variation in the human genome. Wiley & Sons, Chichester, pp 97–118

Bianchi NO, Bailliet G, Bravi CM, Carnese RF, Rothhammer F, Martínez-Marignac VL, Pena SDJ (1997) Origin of Amerindian Y-chromosomes as inferred by the analysis of six polymorphic markers. Am J Phys Anthropol 102:79–89

Bouzekri N, Taylor PG, Hammer MF, Jobling MA (1998) Novel mutation processes in the evolution of a haploid minisatellite, MSY1: array homogenization without homogenization. Hum Mol Genet 7:655–659

Bravi CM, Sans M, Bailliet G, Martinez-Marignac VL, Portas M, Barreto I, Bonilla C, et al (1997) Characterization of mitochondrial DNA and Y-chromosome haplotypes in a Uruguayan population of African ancestry. Hum Biol 69: 641–652

Cooper G, Amos W, Hoffman D, Rubinsztein DC (1996) Network analysis of human Y microsatellite haplotypes. Hum Mol Genet 5:1759–1766

de Knijff P, Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, et al (1997) Chromosome Y microsatellites: population genetics and evolutionary aspects. Int J Legal Med 110:134–140

Flint J, Boyce AJ, Martinson JJ, Clegg JB (1989) Population bottlenecks in Polynesia revealed by minisatellites. Hum Genet 83:257–263

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 15:1108–1114

Goldstein DB, Zhivotovsky LA, Nayar K, Linares AR, Cavalli-Sforza LL, Feldman MW (1996) Statistical properties of the variation at linked microsatellite loci: implications for the history of human Y chromosomes. Mol Biol Evol 13: 1213–1218

Hammer MF (1995) A recent common ancestry for human Y chromosomes. Nature 378:376–378

Hammer MF, Horai S (1995) Y chromosomal DNA variation and the peopling of Japan. Am J Hum Genet 56:951–962

Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, Griffiths RC, et al (1998) Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. Mol Biol Evol 15:427–441

Heyer E, Puymirat J, Dieltjes P, Bakker E, de Knijff P (1997) Estimating Y-chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. Hum Mol Genet 6:799–803

Heyerdahl T (1950) Kontiki: across the Pacific by raft. Rand McNally, Chicago

Hill AVS, Gentile B, Bonnardot JM, Roux J, Weatherall DJ, Clegg JB (1987) Polynesian origins and affinities: globin gene variants in eastern Polynesia. Am J Hum Genet 40:453–463

Hill AVS, Serjeantson SW (1989) The colonization of the Pacific: a genetic trail. Clarendon Press, Oxford

Jennings JD (1979) The prehistory of Polynesia. Harvard University Press, Cambridge, MA

Jeffreys AJ, Neumann R, Wilson V (1990) Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. Cell 60:473–485

Jobling MA, Bouzekri N, Taylor PG (1998) Hypervariable digital DNA codes for human paternal lineages: MVR-PCR at the Y-specific minisatellite, MSY1 (*DYF155S1*). Hum Mol Genet 7:643–653

Jobling MA, Samara V, Pandya A, Fretwell N, Bernasconi B, Mitchell RJ, Gerelsaikhan T, et al (1996) Recurrent duplication and deletion polymorphisms on the long arm of the Y chromosome in normal males. Hum Mol Genet 5: 1767–1775

Jobling MA, Tyler-Smith C (1995) Fathers and sons: the Y chromosome and human evolution. Trends Genet 11: 449–456

Kwok C, Tyler-Smith C, Medonca BB, Hughes I, Berkovitz GD, Goodfellow PN, Hawkins JR (1996) Mutation analysis of 2kb 5′ to SRY in XY females and XY intersex subjects. J Med Genet 33:465–468

Lum JK, Cann RL, Martinson JJ, Jorde LB (1998) Mitochondrial and nuclear genetic relationships among Pacific Island and Asian populations. Am J Hum Genet 63:613–624

Lum JK, Richards O, Ching C, Cann RL (1994) Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. Hum Biol 66:567–590

Mathias N, Bayés M, Tyler-Smith C (1994) Highly informative compound haplotypes for the human Y chromosome. Hum Mol Genet 3:115–123

Melton T, Peterson R, Redd AJ, Saha N, Sofro ASM, Martinson J, Stoneking M (1995) Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. Am J Hum Genet 57:403–414

Redd AJ, Takezaki N, Sherry ST, McGarvey ST, Sofro ASM, Stoneking M (1995) Evolutionary history of the COII/tRNAlys intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. Mol Biol Evol 12:604–615

Reed PW, Davies JL, Copeman JB, Bennett ST, Palmer SM, Pritchard LE, Gough SCL, et al (1994) Chromosome-specific microsatellite sets for fluorescence-based, semi-automated genome mapping. Nat Genet 7:390–395

Santos FR, Tyler-Smith C (1996) Reading the human Y chromosome: the emerging DNA markers and human genetic history. Braz J Genet 19:665–670

Serjeantson SW (1989) HLA genes and antigens. In: Hill AVS, Serjeantson SW (eds) The colonization of the Pacific: a genetic trail. Clarendon Press, Oxford, pp 120–173

Spurdle AB, Jenkins T (1992) The search for Y-chromosome polymorphism is extended to negroids. Hum Mol Genet 1: 169–170

Spurdle AB, Woodfield DG, Hammer MF, Jenkins T (1994)

The genetic affinity of Polynesians: evidence from Y-chromosome polymorphisms. Ann Hum Genet 58:251–263

Sykes B, Leiboff A, Low-Beer J, Tetzner S, Richards M (1995) The origins of the Polynesians: an interpretation from mitochondrial lineage analysis. Am J Hum Genet 57: 1463–1475

Terrell J (1986) Prehistory in the Pacific islands. Cambridge University Press, Cambridge

Thomas MG, Skorecki K, Ben-Ami H, Parfitt T, Bradman N, Goldstein DB (1998) Origins of Old Testament priests. Nature 394:138–139

Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, et al (1997) Detection of numerous Y-chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. Genome Res 7:996–1005

Underhill PA, Jin L, Zemans R, Oefner PJ, Cavalli-Sforza LL (1996) A pre-Columbian Y-chromosome–specific transition and its implications for human evolutionary history. Proc Natl Acad Sci USA 93:196–200

Veitia R, Ion A, Barbaux S, Jobling MA, Souleyreau N, Ennis K, Ostrer H, et al (1997) Mutations and sequence variants in the testis-determining region of the Y chromosome in individuals with a 46, XY female phenotype. Hum Genet 99:648–652

Whitfield LS, Sulston JE, Goodfellow PN (1995) Sequence variation of the human Y chromosome. Nature 378:379–380

Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhövel W, et al (1997) Genetic relationships of Asians and northern Europeans, revealed by Y-chromosomal DNA analysis. Am J Hum Genet 60:1174–1183