Melanesian origin of Polynesian Y chromosomes

Manfred Kayser*, Silke Brauer*, Gunter Weiss*, Peter A. Underhill†, Lutz Roewer‡, Wulf Schiefenhövel§ and Mark Stoneking*

Background: Two competing hypotheses for the origins of Polynesians are the 'express-train' model, which supposes a recent and rapid expansion of Polynesian ancestors from Asia/Taiwan via coastal and island Melanesia, and the 'entangled-bank' model, which supposes a long history of cultural and genetic interactions among Southeast Asians, Melanesians and Polynesians. Most genetic data, especially analyses of mitochondrial DNA (mtDNA) variation, support the express-train model, as does linguistic and archaeological evidence. Here, we used Y-chromosome polymorphisms to investigate the origins of Polynesians.

Results: We analysed eight single nucleotide polymorphisms (SNPs) and seven short tandem repeat (STR) loci on the Y chromosome in 28 Cook Islanders from Polynesia and 583 males from 17 Melanesian, Asian and Australian populations. We found that all Polynesians belong to just three Y-chromosome haplotypes, as defined by unique event polymorphisms. The major Y haplotype in Polynesians (82% frequency) was restricted to Melanesia and eastern Indonesia and most probably arose in Melanesia. Coalescence analysis of associated Y-STR haplotypes showed evidence of a population expansion in Polynesians, beginning about 2,200 years ago. The other two Polynesian Y haplotypes were widespread in Asia but were also found in Melanesia.

Conclusions: All Polynesian Y chromosomes can be traced back to Melanesia, although some of these Y-chromosome types originated in Asia. Together with other genetic and cultural evidence, we propose a new model of Polynesian origins that we call the 'slow-boat' model: Polynesian ancestors did originate from Asia/Taiwan but did not move rapidly through Melanesia; rather, they interacted with and mixed extensively with Melanesians, leaving behind their genes and incorporating many Melanesian genes before colonising the Pacific.

Addresses: *Max Planck Institute for Evolutionary Anthropology, Inselstraße 22, D-04103 Leipzig, Germany. †Department of Genetics, Stanford University, Stanford, California 94305, USA. ‡Institute of Legal Medicine, Humboldt University Berlin, Hannoversche Straße 6, D-10115 Berlin, Germany. §Max Planck Institute for Behavioural Physiology, Von-der-Tann Straße 3, D-82346 Andechs, Germany.

Correspondence: Manfred Kayser
E-mail: kayser@eva.mpg.de

Received: **17 July 2000**
Revised: **18 August 2000**
Accepted: **18 August 2000**

Published: **27 September 2000**

Current Biology 2000, 10:1237–1246

0960-9822/00/$ – see front matter

Background

The origin of the Polynesians — the people living in the area of the Pacific bounded by Fiji to the west, Hawaii to the north, Easter Island to the east and New Zealand to the south — has long drawn the attention of researchers from different fields. Linguists group all the languages spoken in Polynesia, Micronesia, the main part of island Melanesia (excluding the Papuan languages spoken in New Guinea and a few adjacent islands of Melanesia), island Southeast Asia, mainland Malaysia and Madagascar into one Austronesian language family that originated on or near Taiwan [1–3]. Recently, a re-analysis of 1,200 Austronesian languages revealed that nine out of the ten subgroups containing 26 languages are spoken exclusively by Taiwanese aborigines, whereas the other 1,174 Austronesian languages belonging to the tenth subgroup are also spoken outside of Taiwan [2,3]. Archaeological evidence — mostly from pottery remains of the Lapita culture — points to southern China and Taiwan for the origin of Polynesians. Lapita remains, which are widespread throughout Polynesia and island Melanesia dating back to 3,600 to 2,500 years ago, have not been found in New Guinea and Australia, and have been interpreted to be derived from a similar culture established about 6,000 years ago in Taiwan and southern China [4–7].

The archaeological and linguistic evidence have given rise to the 'express-train to Polynesia' hypothesis for the colonisation of the Pacific [4,8,9], according to which Austronesian-speaking people migrated rapidly during the last few thousand years from Asia/Taiwan into the Pacific. Recent genetic data from mitochondrial DNA (mtDNA) have been interpreted to support this hypothesis and a Taiwanese origin of Polynesians [10–12]. Data from human leukocyte antigen (HLA) genes also suggest Polynesian affinities with Asians [13–15].

An alternative 'entangled-bank' hypothesis for the colonisation of Polynesia assumes a long and more complex history of interaction between Polynesia, Melanesia and

Southeast Asia and holds that there was no single discrete migration event or 'express-train' to Polynesia [16]. In this respect, Lapita has also been interpreted as having evolved gradually in island Melanesia and introduced to Polynesia from Melanesia rather than imported from Asia [17,18]. Whereas most of the mtDNA and HLA data have been interpreted as indicating an Asian/Taiwan origin of Polynesians, the haemoglobin genes of Polynesians do indicate affinities with Melanesians [19–22]. A third hypothesis, that Polynesians came from South America [23], receives no support from any genetic data [12,24] and is not generally accepted by either archaeologists or linguists [6].

To date, conclusions from genetic data regarding Polynesian population history are restricted to maternally inherited mtDNA and recombining autosomal DNA. Insights from the male-inherited Y chromosome are not available or are of limited information because of the lack of polymorphic markers in this geographic region [25] or the limited numbers of Melanesian and Southeast Asian populations analysed [26]. To investigate the origin of Polynesian Y chromosomes, we have therefore undertaken an extensive study of eight Y-chromosomal single nucleotide polymorphisms (Y-SNPs) and seven Y-chromosomal short tandem repeat (Y-STR) loci or microsatellites in 18 populations from Polynesia, Melanesia, Asia and Australia.

Results

We initially investigated 28 males from the Cook Islands for eight Y-SNPs (M4, M5, M9, M16, M21, M119, M122 and RPS4Y711). This analysis revealed that every Cook Islander belongs to one of only three Y-chromosome haplotypes (Table 1, Figure 1), defined by polymorphisms at M9, M122 and RPS4Y711. To investigate the origin of these Polynesian Y-chromosome haplotypes, we analysed an additional 583 individuals from 17 populations from Asia, Australia and Melanesia at the three Y-SNPs that were found to be polymorphic in the Cook Islanders. We also analysed seven Y-STRs (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392 and DYS393).

DYS390.3del/RPS4Y711T haplotype

At the Y-STR locus DYS390, alleles of short fragment length (19–21 repeats) were observed in the Cook Islanders, mainland and island Papua New Guineans, eastern Indonesians and Australians, whereas other populations had exclusively alleles of 21–27 repeats in length. To determine the molecular basis of these short alleles, we sequenced DYS390 in all individuals with allele lengths of 18–22 repeats. In addition, from every population in which short alleles were found, we also sequenced DYS390 in a further three or four individuals that had alleles longer than 22 repeats. This was so that any underlying sequence heterogeneity in alleles with the same fragment length would be revealed. In total, we sequenced DYS390 from 319 individuals. DYS390 is a complex STR with four different repetitive segments, designated 390.1–390.4 (Figure 2). All Cook Islanders with alleles 19–21 had a deletion of the 390.3 segment, and all of the length variation was in the combined

Table 1

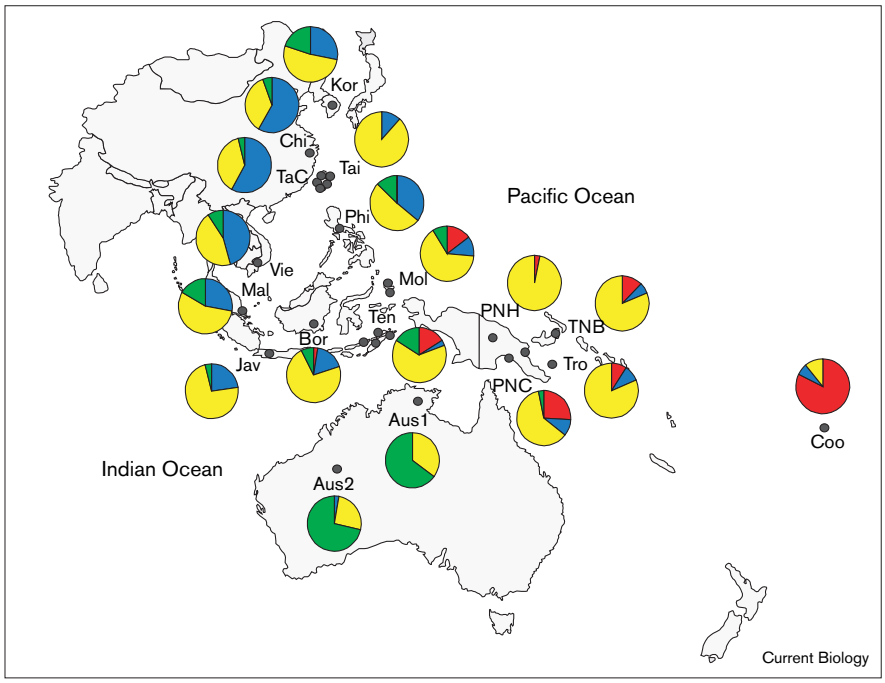
Y-chromosomal haplotypes observed in Polynesia and their frequency distribution in other populations from Melanesia, Asia and Australia.

	Y-chromosomal haplotypes			
Population*	DYS390.3del/RPS4Y711T	M122C/M9G	M9G	Others
Cook Islands (Coo, 28)	82.1	7.1	10.7	0
Papua New Guinea coast (PNC, 31)	25.8	9.7	61.3	3.2
Nusa Tenggara (Ten, 31)	16.1	3.2	64.5	16.1
Moluccas (Mol, 34)	14.7	11.8	64.7	8.8
Tolai New Britain (TNB, 16)	12.5	6.3	81.3	0
Trobriand Islands (Tro, 54)	9.3	9.3	81.5	0
Papua New Guinea highlands (PNH, 31)	3.2	0	96.8	0
Southern Borneo (Bor, 40)	2.5	17.5	72.5	7.5
Korea (Kor, 25)	0	28.0	52.0	20.0
Han Southern China (Chi, 36)	0	58.3	36.1	5.6
Han Chinese Taiwan (TaC, 26)	0	57.7	38.5	3.8
Taiwan aborigines (Tai, 43)	0	11.6	88.4	0
Philippines (Phi, 39)	0	35.9	51.3	12.8
Vietnam (Vie, 11)	0	45.5	45.5	9.1
Malay (Mal, 18)	0	27.8	55.6	16.7
Java (Jav, 53)	0	22.6	73.6	3.8
Australia 1 (Aus1, 60)	0	0	35.0	65.0
Australia 2 (Aus2, 35)	0	2.9	25.7	71.4

*The abbreviations used for each population and the number of individuals sampled are indicated within the brackets.

Figure 1

Y-chromosomal haplotypes observed in Polynesia and their frequency distribution in other populations from Melanesia, Asia and Australia (for abbreviations, see Table 1). Red, DYS390.3del/RPS4Y711T; blue, M122C/M9G; yellow, M9G; green, other haplotypes.

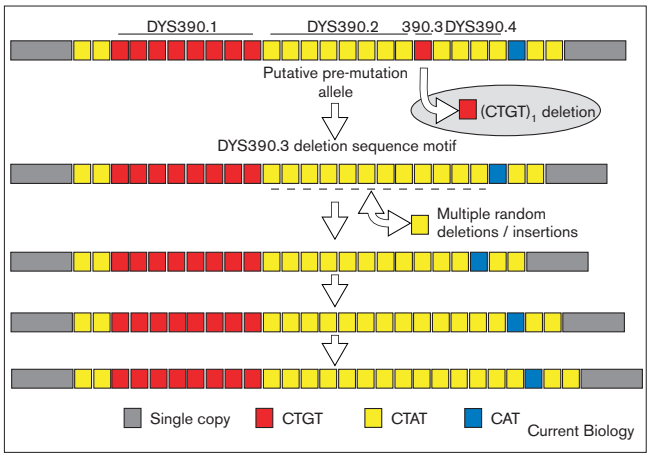


390.2–390.4 segment. All other individuals in our study with alleles 19 and 20 also had the 390.3 deletion, with the exception of Australians who instead had the previously described deletion in 390.1 ([27]; M.K. and M.S., unpublished work). Some individuals from mainland/island Papua New Guinea and Indonesia with alleles 21–23 had the 390.3 deletion and some did not, whereas none of the individuals sequenced with alleles of at least 24 repeats had the 390.3 deletion.

An analysis of eight Y-SNPs revealed that the DYS390.3 deletion is associated with a C→T mutation at position 711 of the RPS4Y gene [28]. The combined DYS390.3 deletion and RPS4Y711T haplotype (DYS390.3del/RPS4Y711T) had a frequency of 82% in the Cook Islands, 26% in coastal Papua New Guinea, 10–15% in the Moluccas and Nusa Tenggara Islands of eastern Indonesia, 9–12% in island Papua New Guinea (New Britain, Trobriand Islands) and was observed once in the Papua New Guinea highlands and southern Borneo, but it was not found in any other Southeast Asian, East Asian or Australian population studied here (Table 1, Figure 1). In the Cook Islanders and mainland/island Papua New Guineans, RPS4Y711T was completely associated with the 390.3 deletion; all individuals with the 390.3 deletion had this mutation, while all individuals lacking the 390.3 deletion carried the ancestral RPS4Y711C. Only in Indonesia, especially eastern Indonesia, some individuals with the RPS4Y711T carried the 390.3 deletion and some did not.

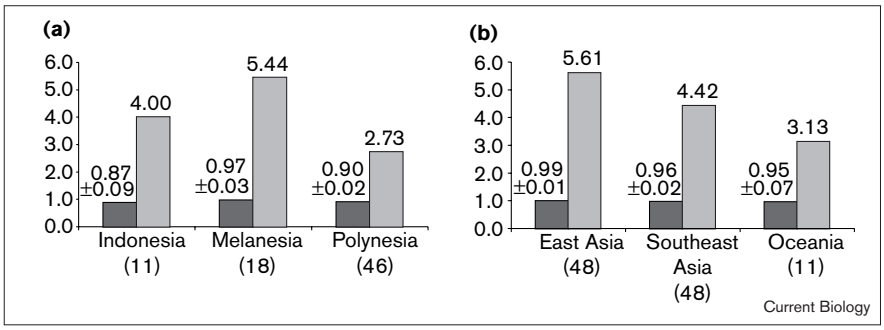
Hurles *et al.* [26] also reported DYS390 alleles 20 and 21 in different samples of Cook Islanders and coastal Papua New Guinea. They found that all chromosomes with these alleles also carried a specific allele at the Y-chromosomal minisatellite locus MSY1. In their study, sequence analysis was not performed to clearly identify the DYS390.3 deletion, nor was the RPS4Y711 marker analysed. Nevertheless, we can infer that the individuals

Figure 2



Schematic structure of the repetitive sequence at the Y-STR locus DYS390 and proposed mutational event of the DYS390.3 deletion. Each square represents one repeat unit.

Figure 3



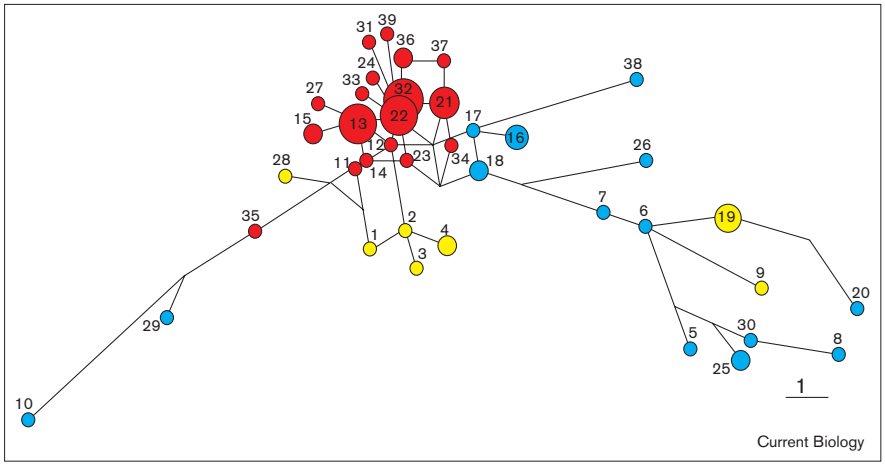
Haplotype diversity (dark grey bars), their standard deviation and the mean number of pairwise differences (light grey bars) of Y-STR haplotypes. **(a)** DYS390.3del/RPS4Y711T haplotype. **(b)** M122C/M9G haplotype. In (a), Indonesia includes five individuals from Mol, five from Ten and one from Bor; Melanesia includes one individual from PNH, eight from PNC, two from Port Moresby, five from Tro, and two from TNB; Polynesia includes 39 from Coo and seven Western Samoans. In (b), East Asia includes 21 individuals from Chi, 15 from TaC, seven from Kor, and five from Vie; Southeast Asia includes 14 individuals from Phi, 12 from Jav, seven from Bor, five Mal, five Tai, four from Mol and one from Ten; and Oceania includes five individuals from Tro, three from PNC, two from Coo and one from TNB (a single Australian was excluded). The units on the vertical axis are the mean number of pairwise differences, for the light grey bars, and the haplotype diversity, for the dark grey bars.

they observed with DYS390 alleles 20–21 belong to the DYS390.3del/RPS4Y711T haplotype that we observed, because they have an A→G mutation at SRY1532 [26], which is associated with the RPS4Y711T mutation [29,30]. In a previous study, the DYS390.3 deletion was also found and characterised by sequence analysis in 7 out of 10 individuals from Western Samoa [27,31], although RPS4Y711 was not analysed.

Combining the present data with these published data revealed a total of 75 individuals from 11 population samples from Polynesia, Papua New Guinea (mainly island and coastal Papua New Guinea), and Indonesia (mainly Moluccas and Nusa Tenggara) that carry the DYS390.3del/RPS4Y711T haplotype (Table 1, Figure 1); 39 seven-locus Y-STR haplotypes were observed within these 75 individuals. Pairwise R_{ST} analysis based on Y-STR haplotypes revealed statistically significant differences ($p < 0.01$) between the following three groups: Melanesia (mainly coastal and island Papua New Guinea), Polynesia (mainly Cook Islands), and Indonesia (mainly Moluccas and Nusa Tenggara Islands). The highest Y-STR haplotype diversity was found in Melanesia and the lowest in Indonesia/Polynesia (Figure 3a). Also, the highest mean number of pairwise differences between haplotypes was found within Melanesia and the lowest was in Polynesia (Figure 3a). Y-STR haplotype sharing was only observed within groups, mainly within Polynesia, but not between any groups.

A median-joining network connecting all 39 haplotypes revealed that all Polynesian haplotypes form a tight cluster and can be connected to each other mostly by single-step mutations, with the exception of one of the West Samoan haplotypes (Figure 4). In contrast, Melanesian and Indonesian haplotypes appeared in different parts of the network and were separated by a large number of mutations. The network shown in Figure 4 was calculated by

Figure 4



Median-joining network of 39 Y-STR haplotypes from 75 individuals belonging to the DYS390.3del/RPS4Y711T haplotype. Circles, Y-STR haplotypes with the area proportional to the number of individuals; lines, mutation steps; parallel lines, identical mutations. Red haplotypes are from Polynesia, blue from Melanesia, and yellow from Indonesia. The scale bar indicates one mutation.

Table 2

Demographic inferences concerning the DYS390.3del/RPS4Y711T and the M122C/M9G haplotype based on associated variation at seven Y-STR loci.

	Posterior probabilities*		
Parameter	Prior probabilities*†	DYS390.3 deletion‡	M122C§
Initial effective population size in 1000 individuals	0.40 (0.06; 2.86)	0.33 (0.11; 1.02)	0.25 (0.07; 0.78)
Population growth rate per generation $\times 10^{-3}$	6.9 (0.3; 36.9)	6.4 (1.0; 22.5)	14.5 (5.9; 33.2)
Time of expansion in 1,000 years	4.9 (0.1; 64.6)	5.0 (1.2; 16.9)	6.0 (2.7; 14.5)
Time back to mrca in 1,000 years	17.1 (1.6; 100.6)	11.5 (5.0; 32.5)	11.1 (5.1; 28.3)

*Median (and 95% equal-tailed intervals). †Prior probabilities used for calculations (see Materials and methods and Figure 5). ‡Based on 75 individuals and 39 Y-STR haplotypes (diversity: 0.958 ± 0.011). §Based on 108 individuals and 81 Y-STR haplotypes (diversity: 0.984 ± 0.006).

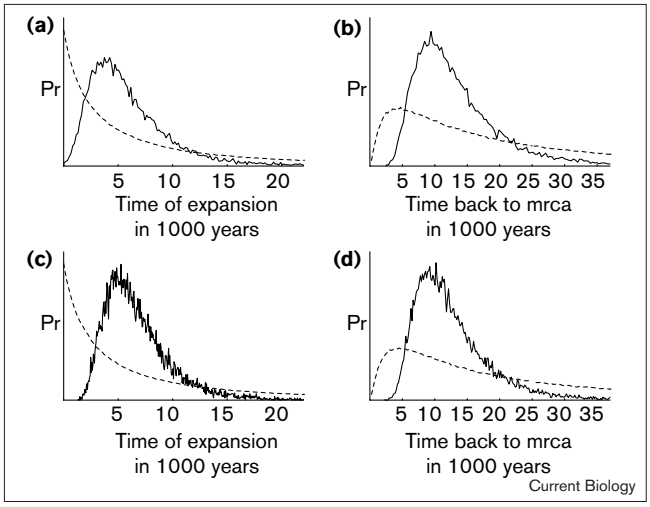
weighting the loci according to mutation rate; calculating the network with equal weights for all loci resulted in increased reticulation within the Polynesian cluster while the rest of the network remained virtually unchanged (data not shown). A minimum-spanning network based on the sum of size differences between the distinct haplotypes also gave a very similar picture of haplotype relationships (data not shown).

To infer demographic data from Y-STR haplotypes of individuals carrying the DYS390.3del/RPS4Y711T haplotype, a Bayesian-based coalescence approach was used. The time back to the most recent common ancestor (mrca) of all 75 individuals carrying the DYS390.3 deletion on the RPS4Y711T chromosome background was estimated to be 11,500 years (Table 2, Figure 5). A signal of slight population growth dating back to the start of a population expansion ~5,000 years ago was detected. When the analysis was restricted to Polynesians, a much stronger signal of population growth was detected, indicating a population expansion starting about 2,200 years ago. These results did not change significantly when different prior probability distributions were applied (data not shown), indicating that these results do reflect the data and not the prior distributions.

Additional major Y-chromosome haplotypes identified in Polynesia

In addition to the DYS390.3del/RPS4Y711T haplotype only two other Y-SNP haplotypes were observed in the Cook Islanders (Figure 1). The M122C/M9G haplotype occurred at a frequency of 7.1% whereas M9G had a frequency of 10.7% (Table 1). Outside Polynesia, the M122C/M9G haplotype was frequent in East Asian and Southeast Asian populations, but was also present in low frequency in all Melanesian populations analysed; it was completely absent in highland Papua New Guinea and nearly absent (1 out of 95) in Australia. Classifying all individuals into three groups consisting of East Asia, Southeast Asia, and Oceania revealed statistically significant differences for R_{ST} based on Y-STR haplotypes for all pairwise comparisons ($p < 0.05$) except Southeast Asia with Oceania. Y-STR haplotype diversity and mean number of pairwise differences were highest in East Asia, lower in Southeast Asia and lowest in Oceania (Figure 3b). Shared Y-STR haplotypes between groups were observed for East versus Southeast Asia (four haplotypes) and Southeast Asia versus Oceania (one haplotype). The time back to the mrca of the 108 individuals carrying the M122C mutation was estimated to be 11,100 years (Table 2, Figure 5). For all individuals carrying the M122C/M9G haplotype, a signal of moderate population growth was detected, with the start of population expansion about 6,000 years ago. The M9G haplotype was

Figure 5



Bayesian-based demographic data inferred from Y-STR variation associated with the **(a,b)** DYS390.3del/RPS4Y711T ($n = 75$) and **(c,d)** M122C/M9G ($n = 108$) haplotypes. Each panel shows the prior (dashed lines) and posterior (continuous lines) probability distribution of the (a,c) time of population expansion and (b,d) time back to the mrca. Pr indicates probability.

present in all populations analysed and at higher frequency than in the Cook Islanders (Table 1). As M9G also occurs outside Asia, especially in Europe, it does not provide any additional information concerning Polynesian origins.

Discussion

Three Y-chromosomal haplotypes, as defined by three Y-SNPs and a specific deletion at the Y-STR locus DYS390, were identified in a Polynesian population sample from the Cook Islands. One of these, the DYS390.3del/RPS4Y711T haplotype, was found at a markedly high frequency of 82%. Hurles *et al.* [26] observed a lower frequency of this haplotype (48%) in an independent sample from the Cook Islands. However, they observed significant European admixture in their sample, whereas for our sample Cook Island paternal ancestry was documented for at least two generations and, thus, was more likely to reflect native Polynesian ancestry. Moreover, we found the DYS390.3del/RPS4Y711T haplotype at a frequency of 70% in Western Samoa, and Deka *et al.* [32] reported a frequency of 60% of alleles 17–22 at DYS390 in American Samoa (although they did not characterise these alleles by DNA sequencing or SNP analysis, they are likely to belong to the DYS390.3del/RPS4Y711T haplotype). Thus, we conclude that this is the major Y-chromosome haplotype in Polynesia. Outside Polynesia, the DYS390.3del/RPS4Y711T haplotype was found to be present only in Melanesia and eastern Indonesia. As the DYS390.3 deletion in Polynesia was always observed on RPS4Y711T chromosomes, whereas RPS4Y711T chromosomes without the DYS390.3 deletion were observed in nearly all of the populations analysed here, with the exception of Melanesia and Polynesia (data not shown), we conclude that the DYS390.3 deletion occurred on a RPS4Y711T chromosome. RPS4Y711T chromosomes are widespread among East and Southeast Asian, Indian, Australian and native American populations, but are not found in Europe or Africa ([28,30,33]; M.K. and M.S., unpublished data). Also, the DYS390.3 deletion has not been identified anywhere else other than in the populations mentioned above when more than 600 individuals from 35 populations from Asia, Africa, Europe and America have been sequenced ([27,34]; this study; M.K., unpublished data), with the exception of a single Turk and a single Ovambo from Namibia [34]. This gives rise to the conclusion that the DYS390.3del/RPS4Y711T haplotype is indeed restricted to Melanesia, Polynesia and eastern Indonesia, and thus can serve as a suitable marker for investigating population history in this region of the world.

We have used Y-STR haplotypes to infer the time and place of origin of the DYS390.3 deletion on the RPS4Y711T background and, thus, the direction of gene flow. Haplotype diversity and the mean number of pairwise differences were higher in Melanesia than in Polynesia or Indonesia and a coalescence-based approach indicated that the deletion arose about 11,500 years ago. These results therefore indicate that the major Y-chromosome haplotype in Polynesians originated in Melanesia.

A second hypothesis is that the DYS390.3del/RPS4Y711T haplotype arose in Proto-Polynesians (presumably Austronesian speakers) in Melanesia as they were expanding on their way to Polynesia. Support for this hypothesis comes from the frequency distribution of the DYS390.3del/RPS4Y711T haplotype in Melanesia; it was present in coastal Papua New Guinea and island Melanesia but nearly absent from highland Papua New Guinea. However, this is contradicted by the amount of Y-STR diversity in Melanesia and the related age of the haplotype (about 11,500 years), as Proto-Polynesians did not reach Melanesia until about 3,500 years ago [9].

A third hypothesis is that the DYS390.3del/RPS4Y711T haplotype arose in Polynesia, and therefore the presence of this haplotype in Melanesia represents a substantial back-migration from Polynesia. The high frequency of this haplotype in Polynesia, and the central position of Polynesian types in the Y-STR haplotype network, provide some support for this explanation. However, the diversity analyses suggest a Melanesian origin, and the date for the origin of this haplotype (11,500 years) substantially predates any other evidence for human occupation of Polynesia. Bottleneck effects provide a likely alternative explanation for the elevated frequency in Polynesia and for the central position of Polynesian types in the Y-STR network. Moreover, there is no other genetic evidence and little archaeological evidence to suggest a substantial back-migration from Polynesia to Melanesia; although one study [31] does suggest back-migration as an explanation for the high frequency of the mtDNA 9 bp deletion marker in island and coastal Papua New Guinea, others do not favour this interpretation [10–12]. A Melanesian origin thus remains the most likely explanation for the DYS390.3del/RPS4Y711T haplotype in Polynesians.

What about the other Y-SNP haplotypes found in Polynesia? The M122C/M9G haplotype was observed in 7.1% of Cook Islanders, but to date the M122 marker has not been analysed in other Polynesian populations. This haplotype has a high frequency in East and Southeast Asia, as observed here and elsewhere [35,36], but has not been found in Africa, America or Europe [35]. Thus, the M122C mutation probably arose in Asia, on an M9G Y chromosome, about 11,100 years ago as indicated by the coalescence approach. The M122C/M9G haplotype is found in Melanesia, but only in coastal Papua New Guinea and island Melanesia; it is absent from highland Papua New Guinea and nearly absent from Australia (1 in 95). This distribution corresponds closely to that of the mtDNA 9 bp deletion marker [10–12], which is thought to reflect the presumed Austronesian expansion from Asia/Taiwan through coastal and island Melanesia into the

Pacific [4,6,9]. The M122T/M9G haplotype also shows a reduction in Y-STR haplotype diversity and in the mean number of pairwise differences from mainland East Asia to Southeast Asia to Oceania (coastal and island Papua New Guinea, Polynesia). Furthermore, the detected moderate population growth and the estimated start of population expansion at about 6,000 years ago is in perfect agreement with archaeological data, which suggest that the Austronesian expansion started about 6,000 years ago from Asia/Taiwan [4,6,9]. Unfortunately, the frequency of the M122C/M9G haplotype in our Melanesian and Polynesian samples is too low to permit accurate analysis for signs of regional population expansion. Although the M122C/M9G haplotype in Polynesia could reflect an Austronesian expansion, it could also reflect simply a direct Melanesian contribution, as with the DYS390.3del/RPS4Y711T haplotype.

The third Y-chromosome haplotype observed in the Cook Islanders was M9G, at a frequency of 10.7%. This haplotype is not useful to investigate population relationships in Polynesia as it is the common ancestor of the majority of non-African haplotypes [35,37]. Thus, M9G alone is not suitable to characterise a distinct Y-chromosomal haplotype, as many Y-SNPs have occurred on an M9G background [37], including, for example, the M122C mutation. In particular, M9G could reflect recent European admixture; in fact, three Cook Islanders that were typed but not included in our study because family records indicated paternal European ancestry all had the M9G haplotype. Additional markers on the M9G background are needed to address the possible Melanesian origin of M9G in Polynesia.

Our study showed reduced Y-STR haplotype diversity within the DYS390.3del/RPS4Y711T haplotype in Polynesians when compared with Melanesians and Indonesians. Furthermore, the overall Y-SNP and Y-STR haplotype diversity was found to be lowest in Polynesia when compared with the other 17 populations analysed here (M.K. and M.S., unpublished data). Reduced genetic diversity in Polynesians has also been reported for many other genetic markers, indicating a Polynesian bottleneck [11,12,38,39]. Moreover, when dividing the total sample set of individuals carrying the DYS390.3del/RPS4Y711T haplotype into Polynesians and non-Polynesians, the population growth rate of Polynesians was estimated to be four times larger than for non-Polynesians, with a population expansion starting 2,200 years ago. This is in agreement with the hypothesis of a bottleneck in the colonisation of Polynesia, which would result in a stronger signal of population growth coming out of the bottleneck. This date is also in remarkably good agreement with archaeological data suggesting that the Cooks and surrounding islands were settled about 2,200 years ago [9].

We conclude that the major Y-chromosome haplotype in Polynesians has an origin from Melanesia, and in fact this may be the case for all Y-chromosome haplotypes in Polynesia, as all of the haplotypes in Polynesia are also found in Melanesia. Recently, Su *et al.* [40] found no evidence for a Melanesian origin of Polynesian Y chromosomes, because their major Melanesian Y-chromosomal haplotype (H17, characterised by mutations at M4, M5 and M9 [37]) was not found in Polynesia. We also found this haplotype in high frequency in Melanesia (M.K. and M.S., unpublished data) and concur that it is absent from Polynesia; however, the absence of this Melanesian haplotype can be explained by founder effects and genetic drift. Su *et al.* [40] did not analyse the markers RPS4Y711 and DYS390; the DYS390.3del/RPS4Y711T haplotype would be included in their complete ancestral (with respect to the analysed markers) haplotype H1, which they reported were found in frequencies of 15–33% in Melanesians and 30–48% in Polynesians. Therefore, their results may actually be compatible with our conclusions.

Other genetic studies have also provided evidence for Melanesian gene flow into Polynesia. For example, the α-haemoglobin $-\alpha^{3.7}$ III deletion is restricted to (mainly coastal) Papua New Guinea, island Melanesia, Micronesia and Polynesia, and probably originated in Melanesia [19–22]. Although some HLA genes in Polynesia are claimed to show Asian rather than Melanesian origin [13–15], the HLA data are equivocal. A recent phylogenetic analysis of DRB1–DQB1 haplotypes grouped Polynesians with Melanesians [41]. Also, a particular allele, HLA DRB1-0901, that was observed at high frequency (26–45%) in Polynesians [41,42] and at moderate frequency (10–15%) in mainland Asia [43–45] was originally reported to be absent or rare in Melanesian populations [41,44,46] but, recently, this allele has been observed at a frequency of 18% in the Trobriands from island Melanesia and 14% in the Roro from the coast of Papua New Guinea [42]. Thus, for the HLA DRB1-0901 allele, a Melanesian contribution to Polynesia is not ruled out, especially as the DYS390.3del/RPS4Y711T haplotype is also found in these same populations.

Studies of the mtDNA 9 bp deletion marker, and the associated 'Polynesian' sequence motif in hypervariable region I of the mtDNA control region, have suggested a Taiwanese origin for Polynesian mtDNAs [10–12]. This is in agreement with the express-train to Polynesia hypothesis [4,6,9], but seems to disagree with the Y-chromosome data presented here and elsewhere [40]. One explanation might be sex-specific migration patterns [47]. However, although most Polynesian mtDNAs have the 9 bp deletion, about 3.5–10% do not [12,48]. Significantly, these Polynesian mtDNA types that lack the 9 bp deletion are found in Papua New Guinea but not elsewhere [12,48]; moreover, all Polynesian mtDNA types are also found in Melanesia, as is the case with Polynesian Y chromosomes. This would suggest that Polynesian mtDNA ancestors did

not simply migrate through New Guinea, but rather that they interbred with the local Melanesian populations, leaving behind mtDNA types with the 9 bp deletion and incorporating Melanesian mtDNA types. This scenario is compatible with the Y chromosome results.

Conclusions

Most, if not all, Polynesian genes examined to date are also found in Melanesia, and thus a Melanesian origin is not only demonstrable for some genes (such as the Y-chromosomal DYS390.3del/RPS4Y711T haplotype), it cannot be ruled out for any gene. Although this would appear to argue against the express-train hypothesis and in favour of the alternative entangled-bank hypothesis of Polynesian origins, it is also true that some of the genes in Melanesia do reflect an expansion out of Southeast Asia that appears to be associated with the expansion of the Austronesian language family. This has been demonstrated for mtDNA [10–12] and may also be true for the Y-chromosomal M122C/M9G haplotype. The overall picture appears to be more complicated than a simple express-train hypothesis of Polynesian ancestors moving quickly through Melanesia with little genetic impact; a more apt metaphor might be a 'slow-boat' model, in which Polynesian ancestors moved gradually across Melanesia, mixing extensively with local Melanesian populations, and thereby not only left behind their genes, but also incorporated many Melanesian genes. This scenario is compatible with cultural indications of interactions between Polynesian ancestors and Melanesians, and with linguistic evidence for 'pauses' during the spread of Austronesian languages through the Pacific [2,49,50].

Materials and methods

DNA samples

DNA samples from 611 male individuals from 18 populations of the following geographic locations were analysed. Polynesia: 28 Cook Islanders (13 from Rarotonga, 4 from Mauke, 3 from Manihiki, 1 from Penrhyn, 2 from Mitiaro, 3 from Aitutaki, 1 from Atiu and 1 from Mangaia); Papua New Guinea: 31 highland New Guineans and 31 coastal New Guineans, both described elsewhere [51], 16 Tolai from Vunapope New Britain, and 54 Trobriand Islanders from Tauwema Kaileuna Island; Australia: 60 aborigines from Arnhem Land (Aus1) and 35 from Great Sandy Desert (Aus2), both described elsewhere [52]; Southeast Asia: 34 eastern Indonesians from the Moluccan Islands (20 from Hiri and 14 from Ternate), 31 from the Nusa Tenggara Islands (8 from Alor, 2 from Flores, 11 from Roti and 10 from Timor, all described elsewhere [11]), 53 Javanese from a rural area near Jakarta, 40 individuals from southern Borneo and 18 Malay, both described elsewhere [10]; Eastern Asia: 11 southern Vietnamese, 26 Han Chinese from Taiwan, 25 southern Koreans (the latter three groups are first generation immigrants to the USA), 36 Han Chinese from Hanzhou (province Zhejiang, south coast of China), 43 Taiwan aborigines (10 Ami, 10 Atayal, 10 Bunum, 13 Paiwan) and 39 Philippinos; the latter were described elsewhere [10]. Care was taken to include unrelated males only for all populations and to assure native ancestry for at least one or two generations for the following areas: Cook Islands, Trobriand Islands, New Britain, coastal and highland Papua New Guinea, eastern Indonesia. For statistical analyses, published data for 33 Cook Islanders from Rarotonga and 58 Papua New Guineans from Port Moresby [26] and 10 Western Samoans [27,31] were included.

Genotyping and DNA sequencing

The Y-STRs DYS19 (or DYS394), DYS389I, DYS389II, DYS390, DYS391, DYS392 and DYS393 were analysed by PCR using published primers, allele nomenclature and protocols [53,54]. Fragment length analysis was performed using either an A.L.F. express (Pharmacia) or an ABI PRISM 377 DNA Sequencer (PE Biosystems). DNA sequencing of DYS390 was performed on both strands, either directly from the PCR product or after reamplification, using published primers [53]. Prior to sequence analysis, PCR products were purified using the QUIquick PCR purification kit (Qiagen) or, before and after reamplification, using a 3% NuSieve (FMC Bioproducts) agarose gel and the Wizard Plus Minipreps DNA purification system (Promega). Sequencing reactions were prepared with the Big Dye Reaction Terminator Cycle Sequencing Kit (PE Biosystems) and purified by isopropanol precipitation prior to analysis on an ABI PRISM 377 DNA Sequencer.

The Y-SNPs analysed include M4, M5, M9, M16, M21, M119, M122 and RPS4Y711 [28,37]. M9, M122 and RPS4Y711 were analysed by PCR using the following standard conditions: 0.4 μM of each primer, 1× GeneAmp PCR buffer (PE Biosystems), 1 U AmpliTaq Gold DNA polymerase (PE Biosystems), 0.2 μM dNTPs (Pharmacia Biotech), 147 μM bovine serum albumin (Sigma), 10–100 ng DNA and a hot-start PCR of 11 min at 95°C (initial denaturation), followed by 30–50 cycles of 30 sec at 94°C, 30 sec at the locus-specific annealing temperature, and 45 sec at 72°C, followed by a final step of 10 min at 72°C. For M9, the PCR primers 5′-GCAGCATATAAAACTTTCAGG-3′ and 5′-GAAATGCATAATGAAGTAAGCG-3′ were used with an annealing temperature of 54°C. The M9 C→G mutation was detected by single-strand oligonucleotide (SSO) hybridisation as described elsewhere [10] using the 5′-biotin-labelled probes: 5′-GATGGTTGAATCCTCTTTAT-3′ for the ancestral C allele, and 5′-ATAAAGAGCATTCAACCATC-3′ for the mutated G allele, with two stringent washes at 55°C for 10 min for each of the probes. Alternatively, the M9 C→G mutation was screened by PCR–RFLP using 10–20 μl PCR product, 1× buffer 2 (New England Biolabs), 10 U *Hinf*I (New England Biolabs) at 37°C overnight, resulting in one undigested fragment (164 bp) for the mutant G allele or two digested fragments (100 bp, 64 bp) for the ancestral C allele. For M122, the PCR primers 5′-GTTGCCTTTTGGAAATGAATAAATC-3′ and 5′-CACTTGCTCTGTGTTAGAAAAGATAGC-3′ were used with an annealing temperature of 58°C. The M122 T→C mutation was detected by PCR–RFLP using 10–20 μl PCR product, 1× buffer 2 (Promega), 1.1 μM bovine serum albumin and 10 U *Hsp*92II (Promega) at 37°C overnight. Fragments were resolved on a 3% NuSieve agarose gel, resulting in one undigested fragment (109 bp) for the mutant C allele, or two digested fragments (58 bp, 51 bp) for the ancestral T allele. For RPS4Y711, the PCR-primers 5′-CTGTACTTACTTTTATCTCCTC-3′ and 5′-CAGCAACAGTAAGTCGAATG-3′ were used with an annealing temperature of 55°C. The RPS4Y711 C→T mutation was detected by PCR–RFLP using 10–20 μl PCR product, 1× buffer 2 *Bsl*I (New England Biolabs), 5 U *Bsl*I (New England Biolabs) at 55°C overnight. Fragments were resolved on a 3% NuSieve agarose gel, resulting in one undigested fragment (91 bp) for the mutant T allele or two digested fragments (34 bp, 57 bp) for the ancestral C allele. For the Y-SNPs M4, M5, M16, M21 and M119, primers, PCR conditions, and genotyping methods will be described in detail elsewhere (M.K. and M.S., unpublished work) and are available from the authors.

Statistical analyses

Y-STRs were analysed with respect to haplotype diversity and the associated standard deviation, mean number of pairwise differences between haplotypes, pairwise R_{ST} values and associated p values based on 10,000 permutations, using the software package Arlequin version 2.000 [55]. A median-joining network [56] based on Y-STR haplotypes was calculated using the Network 2.0b software (http://www.fluxus-engineering.com/sharenet.htm). For network calculation, locus-specific weights were given according to the recently observed mutation rates for the Y-STRs used here [57] so that loci with the highest mutation rates were given the lowest weights (ratio of DYS393:DYS392:DYS19:DYS389I:DYS389II:DYS391:DYS390 = 10:10:5:5:2:2:1). As the DYS389II PCR product also contains

DYS389I, for all statistical analyses, a simple subtraction of the DYS389I repeat length from that of DYS389II was done.

Bayesian-based coalescence analyses of Y-STR haplotypes were performed using the software BATWING (http://www.maths.abdn.ac.uk/~ijw/batwing). The principles of this Markov chain Monte Carlo based inference method were described elsewhere [58]. We chose a two-phase population model, where in the past the population size was of constant size, N, followed by a period of exponential growth until present. We assigned gamma-distributed prior distributions to the mutation rates of the seven loci adjusted to the corresponding estimates in Kayser *et al.* [57]. For the initial effective population size, we used a lognormal prior distribution with mode 148, median 403 and mean 665, corresponding to a small initial founder population. The population growth rate prior (per N generations) was an exponential distribution with mean 1, which covers the simple constant population size model as well as reasonable growth rates for human population. The length of the growth period (in units of N times generation time) also had an exponential prior with mean 1.

Acknowledgements

We thank the original donors of samples, and E. Hagelberg, S. Ulijaszek, K. Katayama, N. Saha, A.S.M. Sofro, K. Bhatia, J. Kuhl, N. Kretchmer, D. Bugawan, J. Martinson, B. Budowle and C. Tyler-Smith for providing DNA or blood samples. C. Tyler-Smith is additionally acknowledged for sharing data before publication. We thank S. Pääbo and M. Nagy for helpful comments. This work was funded by a grant from the National Science Foundation to M.S. and by the Max Planck Society.

References

1. Blust R: **The Austronesian homeland: a linguistic perspective.** *Asian Perspect* 1984/85, **26:**45-67.
2. Blust R: **Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics.** *Symp Ser Inst Linguist Acad Sinica* 1999, **1:**31-94.
3. Diamond JM: **Taiwan's gift to the world.** *Nature* 2000, **403:**709-710.
4. Bellwood PS: *Man's Conquest of the Pacific: The Prehistory of Southeast Asia and Oceania.* Oxford: Oxford University Press; 1978.
5. Spriggs M: **The Lapita cultural complex: origins, distribution, contemporaries and successors.** In *Out of Asia.* Edited by Kirk R, Szathmary E. Canberra: Journal of Pacific History; 1985:185-206.
6. Bellwood PS: *The Polynesians.* London: Thames and Hudson; 1987.
7. Kirch PV: *The Lapita Peoples.* Cambridge, Massachusetts: Blackwell; 1995.
8. Diamond JM: Express train to Polynesia. *Nature* 1988, **336:**307-308.
9. Bellwood PS: **The colonization of the Pacific: some current hypotheses.** In *The Colonization of the Pacific: A Genetic Trail.* Edited by Hill AVS, Serjeantson SW. Oxford: Oxford University Press; 1989:1-59.
10. Melton T, Peterson R, Redd AJ, Saha N, Sofro ASM, Martinson J, *et al.*: **Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis.** *Am J Hum Genet* 1995, **57:**403-414.
11. Redd AJ, Takezaki N, Sherry ST, McGarvey ST, Sofro ASM, Stoneking M: **Evolutionary history of the COII/tRNA-lys intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific.** *Mol Biol Evol* 1995, **12:**604-615.
12. Sykes B, Leiboff A, Low-Beer J, Tetzner S, Richards M: **The origins of the Polynesians: an interpretation from mitochondrial lineage analysis.** *Am J Hum Genet* 1995, **57:**1463-1475.
13. Serjeantson SW, Ryan DP, Thompson AR: **The colonization of the Pacific: the story according to human leukocyte antigens.** *Am J Hum Genet* 1982, **34:**904-918.
14. Serjeantson SW: **Migration and admixture in the Pacific.** *J Pac Hist* 1984, **19:**123-146.
15. Serjeantson SW: **HLA genes and antigens.** In *The Colonization of the Pacific: A Genetic Trail.* Edited by Hill AVS, Serjeantson SW. Oxford: Oxford University Press; 1989:120-173.
16. Terrell J: *Prehistory in the Pacific Islands.* Cambridge: Cambridge University Press; 1986.
17. Anson D: **Lapita pottery of the Bismarck Archipelago and its affinities.** *Archeol Oceania* 1986, **21:**157-165.
18. Terrell J: **What Lapita is and what Lapita isn't.** *Antiquity* 1989, **63:**623-626.
19. Hill AVS, Bowden DK, Trent RJ, Higgs DR, Oppenheimer SJ, Thein SL, *et al.*: **Melanesians and Polynesians share a unique α-thalassemia mutation.** *Am J Hum Genet* 1985, **37:**571-580.
20. Flint J, Hill AVS, Bowden DK, Oppenheimer SJ, Sill PR, Serjeantson SW, *et al.*: **High frequencies of α-thalassemia are the result of natural selection by malaria.** *Nature* 1986, **321:**744-749.
21. Hill AVS, Gentile B, Bonnardot JM, Roux J, Weatherall DJ, Clegg JB: **Polynesian origins and affinities: globin gene variants in eastern Polynesia.** *Am J Hum Genet* 1987, **40:**453-463.
22. O'Shaughnessy DR, Hill AVS, Bowden DK, Weatherall DJ, Clegg JB: **Globin gene in Micronesia: origin and affinities of Pacific Island peoples.** *Am J Hum Genet* 1990, **46:**144-155.
23. Heyerdahl T: *Kontiki: Across the Pacific by Raft.* Chicago: Rand McNally; 1950.
24. Bonatto SL, Redd AJ, Salzano FM, Stoneking M: **Lack of ancient Polynesian-Amerindian contact.** *Am J Hum Genet* 1996, **59:**253-256.
25. Spurdle AB, Woodfield DG, Hammer MF, Jenkins T: **The genetic affinity of Polynesians: evidence from Y chromosome polymorphisms.** *Ann Hum Genet* 1994, **58:**251-263.
26. Hurles ME, Irven C, Nicholson J, Taylor PG, Santos FR, Loughlin J, *et al.*: **European Y-chromosomal lineages in Polynesia: a contrast to the population structure revealed by mtDNA.** *Am J Hum Genet* 1998, **63:**1793-1806.
27. Forster P, Kayser M, Meyer E, Roewer L, Pfeiffer H, Benkmann H, *et al.*: **Phylogenetic resolution of complex mutational features at Y-STR DYS390 in aboriginal Australians and Papuans.** *Mol Biol Evol* 1998, **15:**1108-1114.
28. Bergen AW, Wang C-Y, Tsai J, Jefferson K, Dey C, Smith KD, *et al.*: **An Asian-Native American paternal lineage identified by *RPS4Y* resequencing and by microsatellite haplotyping.** *Ann Hum Genet* 1999, **63:**63-80.
29. Pandya A: Human Y-chromosomal DNA variation. DPhil thesis. Oxford, Oxford University; 1998.
30. Karafet TM, Zegura SL, Posukh O, Osipova L, Bergen A, Long J, *et al.*: **Ancestral Asian source(s) of New World Y-chromosome founder haplotypes.** *Am J Hum Genet* 1999, **64:**817-831.
31. Hagelberg E, Kayser M, Nagy M, Roewer L, Zimdahl H, Krawczak M, *et al.*: **Molecular genetic evidence for the human settlement of the Pacific: analysis of mitochondrial DNA, Y chromosome and HLA markers.** *Phil Trans R Soc Lond B* 1999, **354:**141-152.
32. Deka R, Jin L, Shriver MD, Yu LM, Saha N, Barrantes R, *et al.*: **Dispersion of human Y chromosome haplotypes based on five microsatellites in global populations.** *Genome Res* 1996, **6:**1177-1184.
33. Santos FR, Pandya A, Tyler-Smith C, Pena SD, Schanfield M, Leonard WR, *et al.*: **The central Siberian origin for native American Y chromosomes.** *Am J Hum Genet* 1999, **64:**619-628.
34. Forster P, Rohl A, Lunnemann P, Brinkmann C, Zerjal T, Tyler-Smith C, *et al.*: **A short tandem repeat-based phylogeny for the human Y chromosome.** *Am J Hum Genet* 2000, **67:**182-196.
35. Su B, Xiao J, Underhill P, Deka R, Zhang W, Akey J, *et al.*: **Y-chromosome evidence for a northward migration of modern humans into Eastern Asia during the last Ice Age.** *Am J Hum Genet* 1999, **65:**1718-1724.
36. Qian Y, Qian B, Su B, Yu J, Ke Y, Chu Z, *et al.*: **Multiple origins of Tibetan Y chromosomes.** *Hum Genet* 2000, **106:**453-454.
37. Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, *et al.*: **Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography.** *Genome Res* 1997, **7:**996-1005.
38. Flint J, Boyce AJ, Martinson JJ, Clegg JB: **Population bottlenecks in Polynesia revealed by minisatellites.** *Hum Genet* 1989, **83:**257-263.
39. Lum JK, Cann RL, Martinson JJ, Jorde LB: **Mitochondrial and nuclear genetic relationships among Pacific Island and Asian populations.** *Am J Hum Genet* 1998, **63:**613-624.
40. Su B, Jin L, Underhill P, Martinson J, Saha N, McGarvey ST, *et al.*: **Polynesian origins: insights from the Y chromosome.** *Proc Natl Acad Sci USA* 2000, **97:**8225-8228.
41. Mack SJ, Bugawan TL, Moonsamy PV, Erlich JA, Trachtenberg EA, Paik YK, *et al.*: **Evolution of Pacific/Asian populations inferred from HLA class II allele frequency distributions.** *Tissue Antigens* 2000, **55:**383-400.
42. Zimdahl H, Schiefenhovel W, Kayser M, Roewer L, Nagy M: **Towards understanding the origin and dispersal of Austronesians in the Solomon Sea: HLA class II polymorphism in eight distinct populations of Asia-Oceania.** *Eur J Immunogenet* 1999, **26:**405-416.

43. Tsuji K, Aizawa M, Sazazuki T: *HLA 1991: Proceedings of the 11th International Histocompatibility Workshop and Conference.* Oxford: Oxford University Press; 1992.
44. Gao X, Serjeantson SW: **Analysis of 2600 HLA-DR, DQ haplotypes in Asia-Oceania.** In *HLA 1991: Proceedings of the 11th International Histocompatibility Workshop and Conference.* Edited by Tsuji K, Aizawan M, Sasazuki T. Oxford: Oxford University Press; 1992:229-232.
45. Hashimoto M, Kinoshita T, Yamasaki M, Tanaka H, Imanishi T, Ihara H, *et al.*: **Gene frequencies and haplotypic associations within the HLA region in 916 unrelated Japanese individuals.** *Tissue Antigens* 1994, **44:**166-173.
46. Gao X, Bhatia K, Trent RJ, Serjeantson SW: **HLA-DR, DQ nucleotide sequence polymorphisms in five Melanesian populations.** *Tissue Antigens* 1992, **40:**31-37.
47. Perez-Lezaun A, Calafell F, Comas D, Mateu E, Bosch E, Martinez-Arias R, *et al.*: **Sex-specific migration patterns in central Asian populations, revealed by analysis of Y-chromosome short tandem repeats and mtDNA.** *Am J Hum Genet* 1999, **65:**208-219.
48. Lum JK, Rickards O, Ching C, Cann RL: **Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters.** *Hum Biol* 1994, **66:**567-590.
49. Cann RL: **Talking trees tell tales.** *Nature* 2000, **405:**1008-1009.
50. Gray RD, Jordan FM: **Language trees support the express-train sequence of Austronesian expansion.** *Nature* 2000, **405:**1052-1055.
51. Stoneking M, Jorde LB, Bhatia K, Wilson AC: **Geographic variation in human mitochondrial DNA from Papua New Guinea.** *Genetics* 1990, **124:**717-733.
52. Redd AJ, Stoneking M: **Peopling of Sahul: mtDNA variation in aboriginal Australian and Papua New Guinea populations.** *Am J Hum Genet* 1999, **65:**808-828.
53. Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, *et al.*: **Evaluation of Y chromosomal STRs: a multicenter study.** *Int J Legal Med* 1997, **110:**125-133,141-149.
54. Redd AJ, Clifford SL, Stoneking M: **Multiplex DNA typing of short-tandem-repeat loci on the Y chromosome.** *Biol Chem* 1997, **378:**923-927.
55. Schneider S, Roessli D, Excoffier L: *Arlequin ver 2.000: a Software for Population Genetics Data Analysis.* University of Geneva, Switzerland: Genetics and Biometry Laboratory; 2000.
56. Bandelt H-J, Forster P, Rohl A: **Median-joining networks for inferring intraspecific phylogenies.** *Mol Biol Evol* 1999, **16:**37-48.
57. Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S, *et al.*: **Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs.** *Am J Hum Genet* 2000, **66:**1580-1588.
58. Wilson IJ, Balding DJ: **Genealogical inference from microsatellite data.** *Genetics* 1998, **150:**499-510.