Hum Genet (2000) 107 : 504–512
DOI 10.1007/s004390000403

ORIGINAL INVESTIGATION

Yong-Gang Yao · W. S. Watkins · Ya-Ping Zhang

Evolutionary history of the mtDNA 9-bp deletion in Chinese populations and its relevance to the peopling of east and southeast Asia

Received: 8 May 2000 / Accepted: 3 August 2000 / Published online: 10 November 2000

Abstract In total, 1218 Chinese from twelve ethnic groups and nine Han geographic groups were screened for the mtDNA 9-bp deletion motif. The frequency of the 9-bp deletion in all samples was 14.7% but ranged from 0% to 32% in the various ethnic groups. Three individuals had a triplication of the 9-bp segment. Phylogenetic and demographic analyses of the mtDNA hypervariable segment 1 (HVS1) sequences suggest that the 9-bp deletion occurred more than once in China. The majority of the Chinese deletion haplotypes (about 90%) have a common origin as a mutational event following an initial expansion of modern humans in eastern Asia. Other deletion haplotypes and the three haplotypes with a 9-bp triplication may have arisen independently in the Chinese, presumably by replication error. HVS1 haplotype analysis suggests two possible migration routes of the 9-bp deletion in east and southeast Asia. Both migrations originated in China with one route leading to the Pacific Islands via Taiwan, the other to southeast Asia and possibly the Nicobar Islands. Along both routes of peopling, a decrease in HVS1 diversity of the mtDNA haplotypes is observed. The "Polynesian motif (16217T/C, 16247A/G, and 16261C/T)" and the 16140T/C, 16266C/A, or C/G polymorphisms appear specific to each migration route.

Introduction

The mtDNA 9-bp deletion is caused by the loss of one copy of the 9-bp tandem repeat sequence (CCCCCTCTA) in the COII/tRNA^{Lys} intergenic region of human mtDNA and was originally thought to be Asian-specific (Wrischnik et al. 1987; Harihara et al. 1992; Passarino et al. 1993). The deletion motif occurs at varying frequencies in Asian, Oceanic, and Amerindian populations (Ballinger et al. 1992; Harihara et al. 1992; Passarino et al. 1993; Redd et al. 1995; Sykes et al. 1995; Betty et al. 1996; Melton et al. 1995, 1998; Lum et al. 1998; Merriwether et al. 1999; Watkins et al. 1999; Shurr et al. 1990; Ward et al. 1991; Shields et al. 1992; Kolman et al. 1995). Additional reports have demonstrated the presence of the 9-bp deletion in African (Soodyall et al. 1996; Watkins et al. 1999) and European (Torroni et al. 1995; Thomas et al. 1998) populations. The coalescent times of the 9-bp deletion haplotypes are estimated to be 50,000–70,000 years before present (YBP; Redd et al. 1995; Soodyall et al. 1996; Thomas et al. 1998). Phylogenetic analyses of these mtDNA haplotypes also suggest multiple independent origins of the 9-bp deletion in Africa (Soodyall et al. 1996; Watkins et al. 1999), Europe (Torroni et al. 1995; Thomas et al. 1998), and Asia (Ballinger et al. 1992; Redd et al. 1995; Watkins et al. 1999). Despite multiple origins of the 9-bp deletion, the deletion motif remains a useful marker for tracing population affinities and migration patterns. The clinal increase of the 9-bp deletion frequencies from west to east, together with a corresponding decrease in deleted mtDNA diversity across contemporary Pacific Islanders are consistent with current ideas regarding the prehistoric peopling of the Pacific (Redd et al. 1995; Melton et al. 1995; Sykes et al. 1995; Lum et al. 1998; Merriwether et al. 1999). In Africa, the distributions of the 9-bp deletion are concordant with the migration route of the recent "Bantu expansion" (Soodyall et al. 1996). Thus, the analysis of 9-bp deletion haplotypes in world populations may demonstrate their genetic contributors and elucidate population evolutionary history (Merriwether et al. 1999; Watkins et al. 1999; Alves-Silva et al. 1999b).

Prehistoric demographic expansion and migration patterns in east and southeast Asia are not well resolved. Recent studies of mtDNA, Y-chromosome biallelic markers, and nuclear variation suggest that the northern Asians and the Pacific Islanders can trace their origins to southeast

Y.-G. Yao · Y.-P. Zhang (✉)
Laboratory of Cellular and Molecular Evolution,
Kunming Institute of Zoology, Chinese Academy of Sciences,
Kunming, Yunnan, 650223 P.R. China
e-mail: zhangyp@public.km.yn.cn,
Tel.: +86 871 519 0761, Fax: +86 871 519 5430

W. S. Watkins
Department of Human Genetics, University of Utah,
Salt Lake City, UT 84112, USA

Asia (Ballinger et al. 1992; Redd et al. 1995; Melton et al. 1998; Su et al. 1999). South China may be the center of Asian mtDNA radiation (Ballinger et al. 1992; Melton et al. 1998; Y.-G. Yao et al., unpublished). It is estimated that the first settlement of modern humans in south Asia occurred 50,000–100,000 years ago, followed by a northward migration into northern China and Siberia (Su et al. 1999) and more recent eastward migrations to the Pacific Islands (Redd et al. 1995; Merriwether et al. 1999). The mtDNA haplotypes with the 9-bp deletion found in the Pacific Islanders and Amerindians trace their ancestral types to China. However, no systematic study of the mtDNA haplotypes with the 9-bp deletion in Chinese populations has been reported.

Frequency distributions of the mtDNA "Polynesian motif" trace its origin to Taiwan and support the "express train model" of Polynesian expansion (Melton et al. 1995; Redd et al. 1995). Recently, Richards et al. (1998) have argued that eastern Indonesia might be the origin of the Polynesian expansion based on an almost complete absence of the full motif in western Indonesia and the Philippines (Melton et al. 1995; Sykes et al. 1995). A systematic analysis of the mtDNA deletion motif and HVS1 variation in the Chinese population should help to resolve population affinities and historic migration routes of mainland Asians into southeast Asia and the Pacific Islands.

This study analyzes the 9-bp deletion motif in 12 Chinese ethnic populations, nine Han geographic populations, and 32 Thais and demonstrates (1) the distribution and evolutionary history of the 9-bp deletion in Chinese, (2) its concordance with the geographic distribution and historic migration of the east and southeast Asian populations.

Material and methods

Samples

A total of 1218 individuals from 12 Chinese ethnic populations and nine Han geographic groups were selected for mtDNA 9-bp deletion screening. In addition, we also typed 32 Thais for the presence of the 9-bp deletion. The sample sizes and locations of the populations are shown in Fig. 1 and Table 1. The geographic origin, nationality, and maternal pedigree (unrelated through at least two generations) of each sample were ascertained before sampling.

DNA amplification and sequencing

Genomic DNA was extracted from whole blood by standard phenol/chloroform methods. The intergenic COII/tRNA^{Lys} 9-bp region was amplified and detected with methods described by Harihara et al. (1992) under modified thermal cycling conditions (94°C for 40 s, 57°C for 50 s, and 72°C for 1 min for 30 cycles). The mtDNA HVS1 was amplified and sequenced by using primers L15996 and H16498 following the same methods as those in our previous work (Yao et al. 2000). In total, 25 individuals with the deletion were sequenced with the same primers as those used for the polymerase chain reaction (PCR). These samples included 17 individuals in cluster II and eight individuals in cluster I (see below). Randomly selected individuals lacking the deletion motif were also sequenced. Additionally, the intergenic COII/tRNA^{Lys} 9-bp region was sequenced in all individuals with a triplication.

Data analyses

Sequences were edited, aligned with Dnastar software (DNASTAR), and compared with the published reference sequence (Anderson et al. 1981; Fig. 2). Previously published mtDNA data that was considered here included 78 Africans (32 with the deletion), 336 Amerindians (63 with the deletion), 118 Australian aboriginals (4 with the deletion), 29 Taiwan aboriginals (12 with the deletion), 74 Pacific Islanders with the 9-bp deletion (27 Indonesians, 23 coastal Papua New Guinea, and 24 Samoan), 99 Indians (6 Nicobarese and 22 South Indians with the deletion), 6 Chinese, and 4 Cambodians with the deletion (Ward et al. 1991, 1993; Vigilant et al. 1991; Santos et al. 1994; Kolman et al. 1995; Redd et al. 1995; Mountain et al. 1995; Betty et al. 1996; van Holst Pellekaan et al. 1998; Melton et al. 1998; Krings et al. 1999; Watkins et al. 1999). Except for the 28 Taiwan aboriginal (Melton et al. 1998) and 71 Indians (Mountain et al. 1995) sequences from Dr. T. Melton and Dr. J. L. Mountain respectively, all other sequences were from the mtDNA database (http://www.eva.mpg.de/hvrbase; Handt et al. 1998) or GenBank. Sequences of 325 Chinese and 27 Thais without the 9-bp deletion were also included for comparison (Yao et al. 2000; Y.-G. Yao et al., unpublished). Because some of the published data were only 360 bp (nucleotide positions 16024–16383), we restricted our analyses to that 360-bp segment when global deletion haplotypes were analyzed.

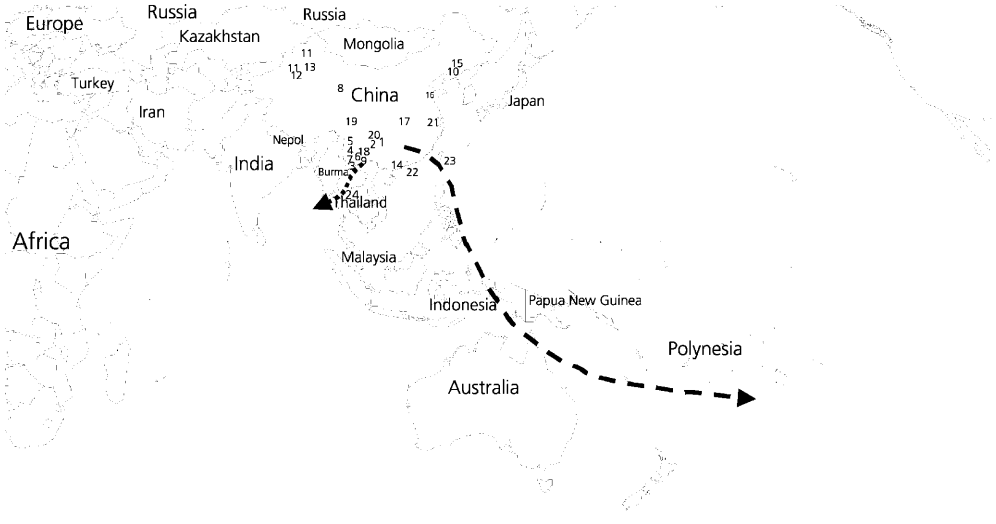

Fig. 1 The localities of the populations considered in this study and the two proposed migration routes of the 9-bp deletion in east and southeast Asia. For the *numbers* of the populations, see Table 1

Table 1 Frequencies of the mtDNA 9-bp deletion in Chinese ethnic populations. Populations 13–23 are different Han geographic groups (*capital letters* in *parentheses* are abbreviations for each population). Cantonese, Taiwan Han, and Thai were not included in the total number

Population	Locality	No. of samples	No. with 9-bp deletion	Frequency of 9-bp deletion (%)	No. with deletion sequenced
1 Miao	Kaili, Guizhou	37	12	32.4	12
2 Buyi	Zhenning, Guizhou	26	8	30.8	8
3 Dai	Jinghong, Yunnan	34	8	23.5	8
4 Naxi	Lijiang, Yunnan	18	4	22.2	4
5 Nu	Gongshan, Yunnan	30	3	10.0	3
6 Sali	Yunnan	32	5	15.6	5
7 Wa	Yunnan	36	1	2.78	1
8 Tu	Huzu, Qinghai	42	7	16.7	7
9 Zhuang	Yunnan	30	4	13.3	4
10 Man	Fenchen, Liaoning	30	2	6.67	2
11 Uygur (Wei)	Kashen & Ili, Xinjiang	60	2	3.33	2
12 Kasak	Kashen, Xinjiang	30	0	0.00	0
13 Xinjiang (XJ)	Xinjiang	48	4	8.33	4
14 Guangdong(GD)	Zhanjiang, Guangdong	34	8	23.5	8
15 Liaoning (LN)	Liaoning	57	7	12.3	7
16 Qingdao (QD)	Qingdao, Shandong	30	1	3.33	1
17 Wuhan (WH)	Wuhan, Hubei	57	10	17.5	9
18 Yunnan (YN)	Kunming, Yunnan	375	61	16.3	3
19 Sichuan (SC)	Sichuan	112	18	16.1	0
20 Guizhou (GZ)	Guizhou	70	8	11.4	0
21 Shanghai (SH)	Shanghai	30	6	20.0	6
22 Cantonese[a]	HongKong	20	4	20.0	4
23 Taiwan Han[b]	Taiwan	20	8	40.0	–
24 Thai	Thailand	32	5	15.6	5
Total	21	1218	179	14.7	94

[a]Data from Betty et al. (1996)
[b]Data from Ballinger et al. (1992)

Twenty Han individuals who lacked the 9-bp deletion and who were sequenced in our former work (seven were from the Liaoning Han group, seven from the Guangdong Han group, and six from the Wuhan Han group; Y.-G. Yao et al., unpublished) were also included in the phylogenetic analysis of the deleted haplotypes. After removal of sites with extra C insertions in the 16183–16189 region, a neighbor-joining (N-J) tree of 98 haplotypes identified in 94 individuals with the 9-bp deletion, three with the triplication, and 20 Han individuals without the deletion was constructed by using PHYLIP 3.5C standard packages (Felsenstein 1993). Distances between haplotypes were computed with Kimura's two-parameter model (Kimura 1980), assuming a transition:transversion ratio of 10:1. Phylogenetic analysis was carried out by using the Neighbor program with 500 bootstraps, and the tree was drawn by means of the Consense and Drawtree programs provided in the PHYLIP package. Population phylogenetic analysis was based on the net genetic distances (d_A; Nei 1987) by using the N-J algorithm.

The evolutionary history of the 9-bp deletion in Asian populations was also examined by the pairwise mismatch distribution (Rogers and Harpending 1992) with the Arlequin software package (Schneider et al. 1999). The time of past population expansion was measured in tau (τ) units of mutational time ($\tau=2\,t\mu$, where μ is the mutation rate over all sites in the sequence and t is the time in years). The mutation rate (33% site/million years) from Ward et al. (1991) was used in converting the mutational time into real time. The coalescent dates for the mtDNA clusters were calculated by using the model of Tamura and Nei (1993) with a gamma correction of $a=0.07\pm0.03$, which was estimated directly from the data.

Results

The 9-bp deletion

The frequencies of the 9-bp deletion in the Chinese ethnic populations and Thai population are given in Table 1. The deletion frequencies vary from 0% to 32%. The highest frequency occurs in the Miao (32.4%) from the Guizhou province, whereas no deletion has been found in Kazak from Xinjiang province. The deletion frequencies among the Han geographic populations were moderate (11%–23%) with the exception of those from Qingdao Han, Xinjiang Han, and Taiwan Han. The frequency of the 9-bp deletion in all the Chinese samples was 14.7%, which was slightly lower than that of Thai (15.6%), Vietnamese (17.9%; Ballinger et al. 1992), and Japanese (16%; Harihara et al. 1992), and different from those of Indian Nicobarese (24%; Watkins et al. 1999), the Pacific Islanders (40%–100%; Redd et al. 1995; Sykes et al. 1995; Merriwether et al. 1999), and the Amerindians (0%–45%; Shurr et al. 1990; Shields et al. 1992; Kolman et al. 1995).

The intergenic COII/tRNA^{Lys} 9-bp region of the 25 deletion individuals showed no deviation from the results of the PCR/polyacrylamide gel electrophoresis (PCR-PAGE).

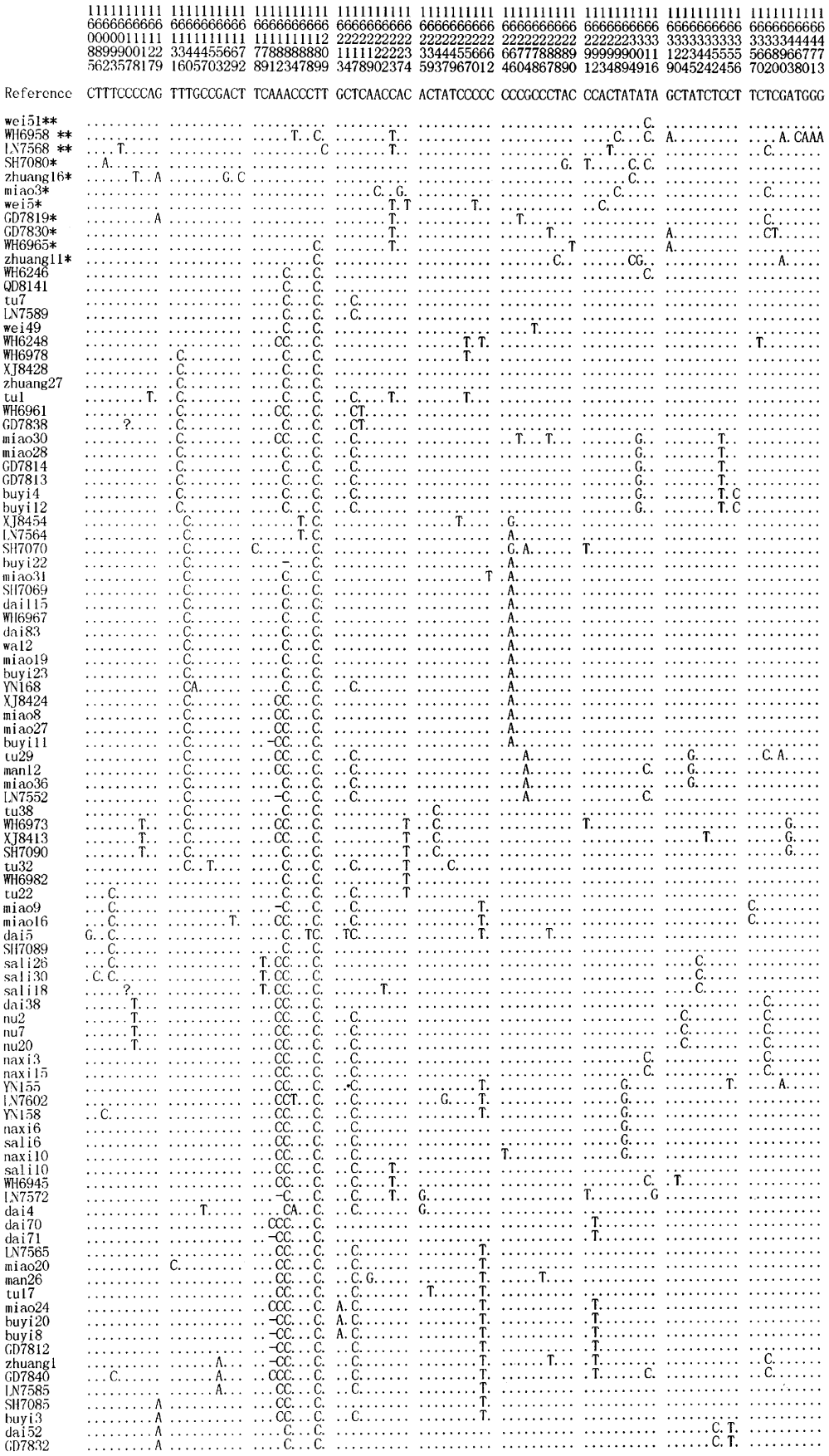

Fig. 2 Variable sites of mtDNA CR segment I sequences in 94 Chinese with the 9-bp deletion and three with the 9-bp triplication (*double asterisks*), with respect to positions 16001–16495 of the reference sequence (Anderson et al. 1981). The length polymorphisms of C in region 16179–16192 because of the 16189T/C substitution were removed from the sequences. *Single asterisks* These eight individuals have divergent haplotypes and may represent multiple 9-bp deletion events in the COII/ $tRNA^{Lys}$ intergenic region

Sequencing of this region indicated that all individuals with the deletion had one 9-bp motif. No expansions or contractions of the poly-C tract in the motif were detected.

Triplication of the 9-bp segment

Three individuals with the 9-bp triplication were detected by PCR-PAGE and confirmed by sequencing. One individual was from Xinjiang Uyghr, the other two were from the Wuhan and Liaoning Han populations, respectively. The Liaoning Han appeared heteroplasmic with two bands being detected by PAGE. Sequencing confirmed that one band contained two 9-bp repeats, whereas the larger band contained three repeats. Thomas et al. (1998) reported similar heteroplasmy in a sample from Glasgow.

The mtDNA HVS1 sequences of the three individuals with the 9-bp triplication did not contain unique polymorphic sites distinguishing them from other Chinese lineages with or without the 9-bp deletion (Fig. 2); however, they differed from each other by at least six mutational events. The polymorphic sites (16192C/T, 16274 G/A, and 16362T/C) reported in a Brazilian and two Portuguese individuals with the triplication (Alves-Silva et al. 1999a) were not found in the three Chinese, suggesting that they may have originated from independent mutational events.

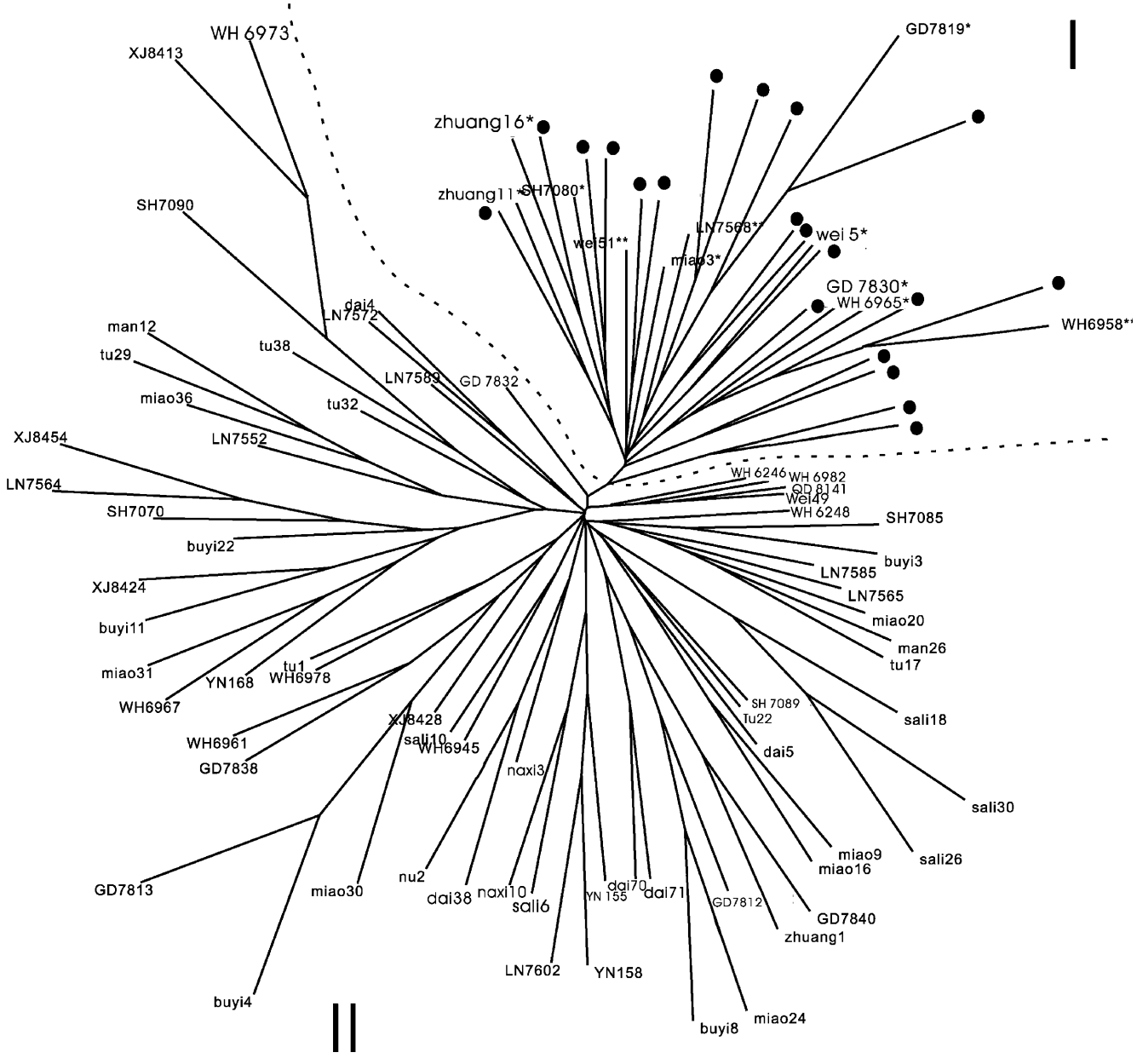

Fig. 3 Phylogenetic tree of 75 Chinese haplotypes with the 9-bp deletion, 20 non-deleted haplotypes (*filled circles* in the tree), and three haplotypes with the 9-bp triplication (abbreviations as in Table 1). The eight deleted haplotypes labeled in Fig. 2, together with the non-deleted haplotypes and the 9-bp triplication individuals converge into cluster I (*top right*), whereas other deleted haplotypes converge mainly in cluster II (*bottom* and *left*)

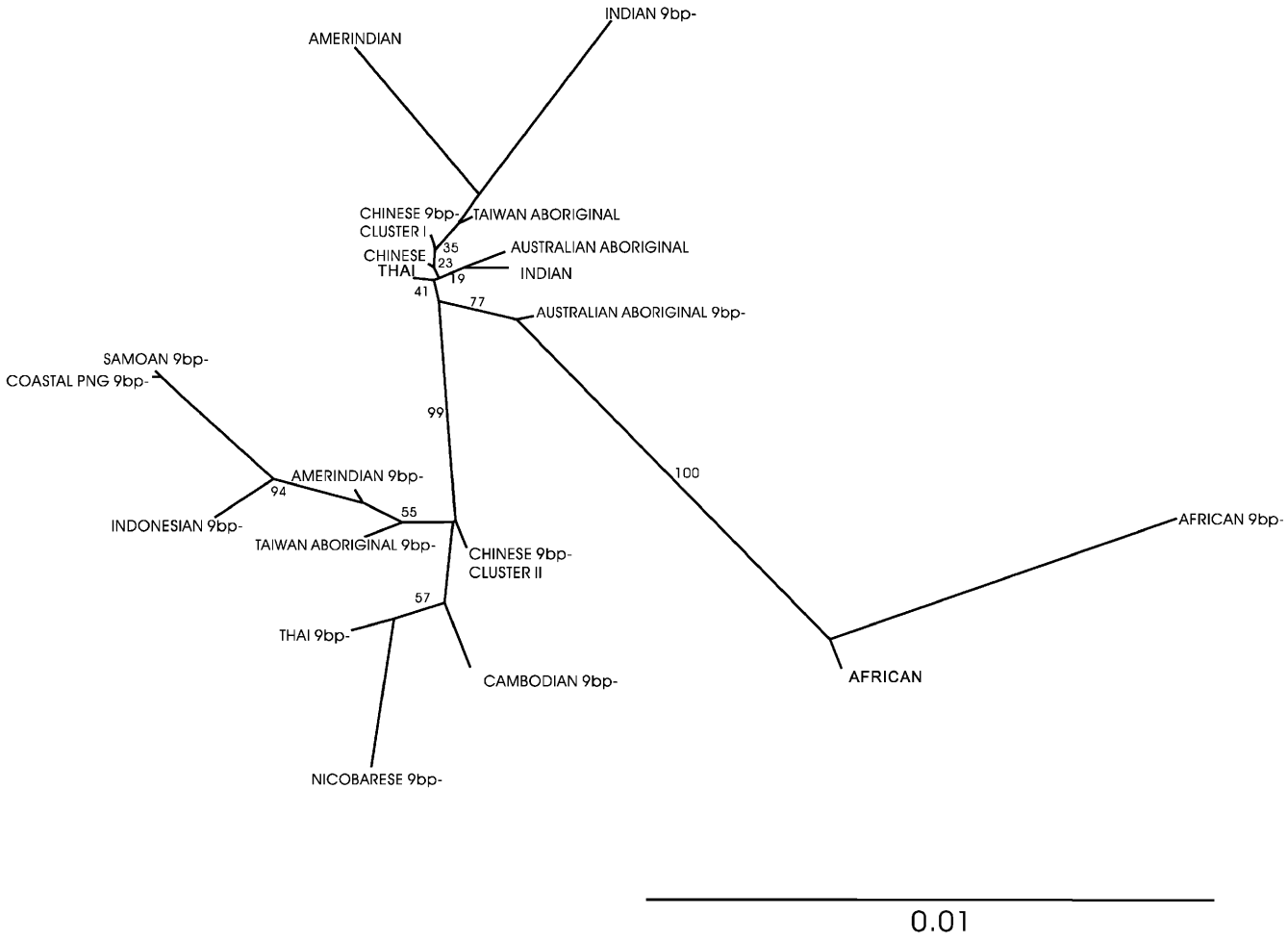

Fig. 4 Unrooted neighbor-joining tree of world populations with or without the 9-bp deletion based on the genetic distances D_A calculated according to the formula: $D_A=D_{XY}-(D_X+D_Y)/2$, where D_{XY} is the pairwise mean between population X and Y, and D_X and D_Y are the pairwise means within population X or Y (Nei 1987). The *numbers* on the *branches* are bootstrap values (percentages) based on 500 bootstrap replications (*9bp*– deleted populations)

Sequence variation in Chinese with the 9-bp deletion

In all, 75 haplotypes were observed in 94 Chinese with the 9-bp deletion (Fig. 2). When six previously reported Chinese were included (Betty et al. 1996; Watkins et al. 1999), 79 unique haplotypes could be identified. The "Polynesian motif" (16217T/C, 16247A/G, 16261C/T) was only detected in one Chinese; however, the 16217T/C and 16261C/T polymorphisms were relatively common, with frequencies of 50% and 20.5%, respectively. The unique polymorphic sites found in Indian (16017T/C, 16126T/C, 16256C/G, 16296C/T; Watkins et al. 1999), in European (16306C/G, 16332C/G; Thomas et al. 1998), and in African (16148C/T, 16187C/T, 16188C/A or G, 16230A/G, Soodyall et al. 1996; Watkins et al. 1999) HVS1 haplotypes were not found in the Chinese with the 9-bp deletion. This finding is consistent with multiple continent-specific deletion events. The 16189T/C substitution was found in the Chinese containing the 9-bp deletion at a frequency of 93.5%. Other sites (16136T/C, 12%; 16140T/C, 26%; 16223C/T, 8.7%; and 16266C/A or C/G, 17%) were also found at a high frequency in Chinese with the 9-bp deletion. The 16292C/T, 16299A/G, and 16309A/G polymorphisms that were observed in the deleted Chinese (6%–9%) were not present in other populations with the 9-bp deletion. Based on the distribution of the variable HVS1 sites (see Fig. 2), the group of eight deletion haplotypes and the other remaining 67 haplotypes could not be derived from an ancestral type by one or two mutation events. This suggests that the 9-bp deletion motif may have occurred more than once in the Chinese population.

Phylogenetic analysis

In the unrooted N-J tree of the 75 Chinese deletion haplotypes, three haplotypes with the triplication, and 20 non-deleted haplotypes, two main clusters (I and II) can be discerned (Fig. 3). The deletion haplotypes marked in Fig. 2 and the three haplotypes with the 9-bp triplication were clustered with those non-deleted haplotypes, whereas other deleted haplotypes were mainly joined in cluster II. The phylogenetic branching pattern suggests that some 9-bp deletion and insertion events have occurred several times

in the Chinese; however, cluster II deletion haplotypes probably originate from a common ancestor.

In the tree of world populations with or without the 9-bp deletion (Fig. 4), deleted Africans cluster with non-deleted Africans and are well separated from other world populations. In the non-African populations, two main clusters (deleted and non-deleted) can be discerned. The deleted Indian and aboriginal Australian populations, which have been reported to have independent origins (Watkins et al. 1999; Betty et al. 1996), are also distinct in the tree and are separated from other deleted populations. The Chinese deletion cluster I presents closer affinity with the non-deleted Chinese population, further arguing for its independent origin. The deleted Thai, Cambodian, Nicobarese, Amerindian, Taiwanese aboriginal, and the Pacific Islander populations converge into two small clades that diverge from the deleted Chinese cluster II population.

Mismatch analyses and coalescent time estimation

Mismatch analyses of the 100 deleted Chinese (including the reported six Chinese) presented a unimodal distribution with a τ value (6.8) slightly lower than that of the 325 non-deletion individuals (7.11). Assuming the mutation rate to be 33% site/million years (Ward et al. 1991), the expansion time of the 9-bp deletion is approximately 57,000 YBP. This result is consistent with the previous estimation (Redd et al. 1995) and later than the settlement of modern humans in east Asia based on an initial population expansion estimated by Su et al. (1999).

The average nucleotide diversities of the deleted individuals in the two clusters in Fig. 3 are 0.02133±0.01268 and 0.01686±0.00898, suggesting a coalescent time of approximately 65,000 and 51,000 YBP, respectively, and indicating different origins. The coalescent time of cluster II, which might have a common origin and be the ancestors of those deleted haplotypes found in the Pacific and southeast Asia, is in good agreement with the previous estimation (Redd et al. 1995).

Discussion

Previous studies have shown that the COII-tRNA^{Lys} intergenic region is unstable and may produce multiple deletions and insertions by slipped-strand mispairing during DNA replication (Shields et al. 1992; Passarino et al. 1993; Redd et al. 1995; Thomas et al. 1998; Lum et al. 1998; Watkins et al. 1999). Our results support this opinion. The phylogenetic relationship of the Chinese 9-bp deletion haplotypes intermixing with those of non-deleted Chinese and separating them from the other deleted haplotypes suggests that they did not have a single recent ancestral type. The different coalescent times of the Chinese deleted haplotype clusters and their different phylogenetic positions in the tree of world populations further support this finding.

The triplication of the 9-bp segment in this region has been observed in a few Europeans (Thomas et al. 1998; Alves-Silva et al. 1999a) and three Asians (Shields et al. 1992; Passarino et al. 1993; Lum et al. 1998). Alves-Silva et al. (1999a) have suggested that the two Portuguese with the 9-bp triplication might be from an ancestral type as they differ from each other by only one transition in the mtDNA control region. However, the three Chinese with the triplication are different from each other by at least six mutational events in HVS1, indicating that they have been diverging for a long time and do not have a recent common origin. Triplication of the 9-bp motif appears to be a rare and independent event in the Chinese based on the low observed frequency of the triplication motif.

The evidence presented here suggests multiple origins of the 9-bp deletion in Chinese. However, this motif is still useful as a population-specific marker for tracing the affinity and migration patterns of Chinese ethnic populations, as most of the deletion haplotypes (about 90%) have a common origin. A frequency cline from south to north and from coastal to inland, together with the more intermixed phylogenetic pattern of the deletion haplotypes in the south, suggests that the southern populations are more ancient (Ballinger et al. 1992; Su et al. 1999; Y.-G. Yao et al., unpublished). The various frequencies of the 9-bp deletion in the nine Han geographic populations support previous data suggesting that the Han is a heterogeneous population (Du et al. 1993).

Other studies have suggested that Taiwan is the origin for a proto-Polynesian expansion (Melton et al. 1995, 1998; Sykes et al. 1995). Our results support this proposal but suggest another migration route into southeast Asia. As shown in Fig. 4, two clusters with the 9-bp deletion are rooted in China and expanded via two migration routes. The route from China through a Taiwanese aboriginal bottleneck and then to the Pacific Island has been discussed in previous studies (Sykes et al. 1995; Melton et al. 1995, 1998). The second route, also originating in China, expanded via Cambodia and Thailand to the Nicobar Islands and has not been previously reported (see Fig. 1). Evidence for the two migration routes is also shown by an increasing or decreasing trend of the frequencies of some polymorphic sites along the route. The 16140T/C, 16266C/A, or C/G polymorphisms present increasing frequencies in southeast Asia and are nearly fixed in the Nicobarese, whereas these frequencies decrease along the route to the Pacific Islands. In contrast, the "Polynesian motif" sites (16217 T/C, 16261C/T, and 16247A/G) show an opposing distribution to those of the 16140 and 16266 sites in the two migration routes (Table 2). The deleted haplotypes shared among or between the populations can trace their roots to the Chinese and thus give indirect evidence for the migration routes. One haplotype shared in 13 Chinese was found in one Indonesian, one Thai, and one Nicobarese. The same sequence in two Chinese was seen in 29 Amerindians, and a haplotype shared by five Indonesian, two coastal PNG, two Samoan and two Taiwan aboriginals was also presented in one Chinese.

The expansion times of the 9-bp deletion populations estimated by mismatch analysis further indicate successive peopling events from China. Along the migration

Table 2 Frequencies of polymorphic sites in populations with the 9-bp deletion

Population	*n*	Pi	τ	Polymorphic site (%)								
				16129G/A	16136T/C	16140T/C	16189T/C	16217T/C	16247A/G	16261C/T	16266C/A or G	16362T/C
Chinese[a]	100	0.0179	6.46	6.5	13	26	93.5	52	1.1	22	16.3	13
Cambodian	4	0.0142	–	25	0.0	50	100	50	0.0	25	50	0.0
Thai	5	0.0114	–	40	0.0	60	100	20	0.0	20	80	0.0
Nicobarese	6	0.0112	–	0.0	0.0	100	100	0.0	0.0	0.0	83	0.0
Southeast Asian[b]	15	0.0152	5.77	20	0.0	73.3	100	20	0.0	13.3	73.3	0.0
Taiwan aboriginal	12	0.0137	5.34	16.7	16.7	25	91.7	66.7	0.0	50	25	16.7
Indonesian	27	0.0112	4.34	7.4	14.8	18.6	100	88.9	22.2	55.6	7.4	0.0
Coastal PNG	23	0.0036	1.33	0.0	0.0	0.0	100	100	73.9	100	0.0	0.0
Samoan	24	0.0022	1.48	0.0	0.0	0.0	100	100	79.2	100	0.0	0.0
Amerindian	63	0.0048	2.32	0.0	0.0	0.0	100	96.8	0.0	6.6	1.6	1.6

[a]The deleted Chinese include 94 individuals sequenced in the current study and six reported individuals by Betty et al. (1996) and Watkins et al. (1999). The values of nucleotide diversity (*Pi*) and expansion time (τ) were estimated from the 86 individuals belong to cluster II in Fig. 3 and the six published individuals

[b]Values were computed in total samples of Thai, Cambodian, and Indian Nicobarese

routes shown in Fig. 1 and Table 2, the oldest expansion time are represented in the Chinese populations. We have not estimated the respective expansion times of the deleted Thai, Cambodian, and Nicobarese, because of their small sample size; however, the sequence divergence (τ=5.77) of these 15 deleted individuals is lower than that of the Chinese with the deletion.

Our results show that the 16261C/T substitution probably arose in south or central China based on a high frequency (22%) in Chinese and did not occur uniquely in Taiwan as suggested by Melton et al. (1995). The following scenario for the evolutionary history of the 9-bp deletion in southeast Asia and the Pacific Islands could be compatible with our results. The 9-bp deletion arose about ~60,000 YBP (Redd et al. 1995) in China, followed by the nucleotide substitutions at position 16217 and then at position 16261. The mtDNAs carrying these markers spread extensively throughout east and southeast Asia and expanded northward to Siberia and America, and eastward into Taiwan and then to the Pacific Islands as shown in other studies (Redd et al. 1995; Melton et al. 1995, 1998; Su et al. 1999). During the expansion, the populations underwent strong population bottlenecks (Melton et al. 1995, 1998; Bonatto and Salzona 1997), producing a general decrease in HVS1 diversity among the deleted populations along the migration route (Redd et al. 1995). We have observed this trend along both migration routes into southeast Asia and the Pacific Island as shown in Fig. 1 and Table 2. Analysis of additional individuals throughout Asia should continue to improve our understanding of the expansions, migrations, and diversity represented in the populations of the Asian mainland and southeast Asia.

Acknowledgements We thank the two anonymous reviewers for helpful comments on our earlier version of the manuscript. This work was supported by the Natural Sciences Foundation of China, the Chinese Academy of Sciences, the Natural Sciences Foundation of Yunnan Province, and National Science Foundation grants SBR 9514733 and 9512178.

References

Alves-Silva J, Guimaraes PEM, Rocha J, Pena SDJ, Prado VF (1999a) Identification in Portugai and Brazil of a mtDNA lineage containing a 9-bp triplication of the intergenic COII/tRNA^{Lys} region. Hum Hered 49:56–58

Alves-Silva J, Santos MS, Pena SD, Prado VF (1999b) Multiple geographic sources of region V 9-bp deletion haplotypes in Brazilians. Hum Biol 71:245–249

Anderson S, Bankier AT, Barrell BG, Bruijn MHL de, Coulson AR, Drouin J, Eperon IC, et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465

Ballinger SW, Schurr TG, Torroni A, Gan YY, Hodge JA, Hassan K, Chen K-H, et al (1992) Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient Mongoloid migrations. Genetics 130:139–152

Betty DJ, Chin-Atkins AN, Croft L, Sraml M, Easteal S (1996) Multiple independent origins of the COII/tRNA^{Lys} intergenic 9-bp mtDNA deletion in aboriginal Australians. Am J Hum Genet 58:428–433

Bonatto SL, Salzona FM (1997) Diversity and age of the four major mtDNA haplogroups, and their implications for the peopling of the New World. Am J Hum Genet 61:1413–1423

Du R, Yip VF (1993) Ethnic groups in China. Science Press, Beijing

Felsenstein J (1993) PHYLIP-Phylogeny inference package (version 3.5C). University of Washington, Seattle

Handt O, Meyer S, Haeseler A von (1998) Compilation of human mtDNA control region sequences. Nucleic Acids Res 26:126–9

Harihara S, Hirai M, Suutou Y, Shimizu K, Omoto K (1992) Frequency of a 9-bp deletion in the mitochondrial DNA among Asian populations. Hum Biol 64:161–166

van Holst Pellekaan S, Frommer M, Sved J, Boettcher B (1998) Mitochondrial control-region sequence variation in aboriginal Australians. Am J Hum Genet 62:435–449

Kimura M (1980) A simple method for estimating evolutionary rates of base substitution through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Kolman CJ, Bermingham E, Cooke R, Ward RH, Arias TD, Guionneau-Sinclair F (1995) Reduced mtDNA diversity in the NgÖbè Amerinds of Pannamà. Genetics 140:275–283

Krings M, Salem Ae, Bauer K, Geisert H, Malek AK, Chaix L, Simon C, et al (1999) mtDNA analysis of Nile river valley populations: a genetic corridor or a barrier to migration? Am J Hum Genet 64:1166–1176

Lum KJ, Cann RL (1998) mtDNA and language support a common origin of Micronesians and Polynesians in Island Southeast Asia. Am J Phys Anthropol 105:109–119

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

Melton T, Clifford S, Martinson J, Batzer M, Stoneking M (1998) Genetic evidence for the pro-Austronesian homeland in Asia: mtDNA and nuclear DNA variation in Taiwanese aboriginal tribes. Am J Hum Genet 63:1807–1823

Merriwether DA, Friedlaender JS, Mediavillla J, Mgone C, Gentz F, Ferrell RE (1999) Mitochondrial DNA variation is an indicator of Austronesian influence in Island Melansia. Am J Phys Anthropol 110:243–270

Mountain JL, Hebert JM, Bhattacharyya S, Underhill PA, Ottolenghi C, Gadgil M, Cavalli-Sforza LL (1995) Demographic history of India and mtDNA-sequence diversity. Am J Hum Genet 56:979–992

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Passarino G, Semino O, Modiano G, Santachiara-Benerecetti A (1993) COII/ $tRNA^{lys}$ Intergenic 9-bp deletion and other mtDNA markers clearly reveal that the Tharus (Southern Nepal) have oriental affinities. Am J Hum Genet 53:609–618

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

Richards M, Oppenheimer S, Sykes B (1998) mtDNA suggests Polynesian origins in Eastern Indonesia. Am J Hum Genet 63: 1234–1236

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9:552–569

Santos M, Ward RH, Barrantes R (1994) mtDNA variation in the Chibcha Amerindian Huetar from Costa Rica. Hum Biol 66: 963–977

Schneider S, Roessli D, Excoffier L (1999) Arlequin ver 2.0: a software for population genetic data analysis. University of Geneva, Switzerland

Schurr TG, Ballinger SW, Gan Y-Y, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, et al (1990) Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. Am J Hum Genet 46:613–623

Shields GF, Hecker K, Voevoda MI, Reed JK (1992) Absence of the Asian-specific region V mitochondrial marker in native Beringians. Am J Hum Genet 50:758–765

Soodyall H, Vigilant L, Hill AV, Stoneking M, Jenkins T (1996) MtDNA control-region sequence variation suggests multiple independent origins of an " Asian-specific" 9-bp deletion in sub-Saharan Africans. Am J Hum Genet 58:595–608

Su B, Xiao J-H, Underhill P, Deka R, Zhang W-L, Akey J, Huang W et al. (1999) Y-chromosome evidence for a northward migration of modern humans into eastern Asia during the last Ice Age. Am J Hum Genet 65:1718–1724

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

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitution in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526

Thomas MG, Cook EC, Miller KW, Waring MJ, Hagelberg E (1998) Molecular instability in the COII/ $tRNA^{lys}$ intergenic region of the human mitochondrial genome: multiple origins of the 9-bp deletion and heteroplasmy for expanded repeats. Philos Trans R Soc Lond Biol 353:955–965

Torroni A, Petrozzi M, Santolamazza P, Sellitto D, Cruciani F, Scozzari R (1995) About the "Asian"-specific 9-bp deletion of mtDNA. Am J Hum Genet 57:507–508

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. Science 253:1503–1507

Ward RH, Frazier BL, Dew-Jager K, Pääbo S (1991) Extensive mitochondrial diversity within a single Amerindian tribe. Proc Natl Acad Sci USA 88:8720–8724

Ward RH, Redd AJ, Valencia D, Frazier B, Pääbo S (1993) Genetic and linguistic differentiation in the Americas. Proc Natl Acad Sci USA 90:10663–10667

Watkins WS, Bamshad M, Dixon ME, Bhaskara Rao B, Naidu JM, Reddy DG, Prasad BVR, et al (1999) Multiple origins of the mtDNA 9-bp deletion in populations of south India. Am J Phys Anthropol 109:147–158

Wrischnik LA, Higuchi RG, Stoneking M, Erlich H, Arnheim N, Wilson AC (1987) Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. Nucleic Acids Res 15:529–542

Yao Y-G, Lü X-M, Luo H-R, Li W-H, Zhang Y-P (2000) Gene admixture in the silk road of China – evidence from mtDNA and melanocortin 1 receptor polymorphism. Genes Genet Syst (in press)