A Mitochondrial Revelation of Early Human Migrations to the Tibetan Plateau Before and After the Last Glacial Maximum

Zhendong Qin,[1] Yajun Yang,[1,2] Longli Kang,[1,3] Shi Yan,[4] Kelly Cho,[5] Xiaoyun Cai,[1] Yan Lu,[1] Hongxiang Zheng,[1] Dongchen Zhu,[1] Dongmei Fei,[1] Shilin Li,[1] Li Jin,[1,2,4*] Hui Li[1*], and The Genographic Consortium

[1]*State Key Laboratory of Genetic Engineering and Ministry of Education Key Laboratory of Contemporary Anthropology, School of Life Sciences and Institutes of Biomedical Sciences, Fudan University, Shanghai 200433, China*
[2]*China Medical City Institute of Health Sciences, Taizhou, Jiangsu 225300, China*
[3]*Department of Medicine, Tibet Nationality College, Xianyang, Shaanxi 712082, China*
[4]*Chinese Academy of Sciences and Max Planck Society Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China*
[5]*Department of Genetics, Yale School of Medicine, New Haven, CT 06520*

KEY WORDS mitochondrial DNA; complete sequence; the Tibetan Plateau; last glacial maximum

ABSTRACT As the highest plateau surrounded by towering mountain ranges, the Tibetan Plateau was once considered to be one of the last populated areas of modern humans. However, this view has been tremendously changed by archeological, linguistic, and genetic findings in the past 60 years. Nevertheless, the timing and routes of entry of modern humans into the Tibetan Plateau is still unclear. To make these problems clear, we carried out high-resolution mitochondrial-DNA (mtDNA) analyses on 562 Tibeto-Burman inhabitants from nine different regions across the plateau. By examining the mtDNA haplogroup distributions and their principal components, we demonstrated that maternal diversity on the plateau reflects mostly a northern East Asian ancestry. Furthermore, phylogeographic analysis of plateau-specific sublineages based on 31 complete mtDNA sequences revealed two primary components: pre-last glacial maximum (LGM) inhabitants and post-LGM immigrants. Also, the analysis of one major pre-LGM sublineage A10 showed a strong signal of post-LGM population expansion (about 15,000 years ago) and greater diversity in the southern part of the Tibetan Plateau, indicating the southern plateau as a refuge place when climate dramatically changed during LGM. Am J Phys Anthropol 000:000–000, 2010.

The Tibetan Plateau, the "Roof of the World," is the highest plateau on the earth with an average elevation of more than 4,000 m, covers more than 2,500,000 km^2 of plateaus and mountains in central Asia and is surrounded by towering mountain ranges (Himalayas on the south, Karakoram on the west, Kunlun on the north-

Additional Supporting Information may be found in the online version of this article.

The Genographic Consortium includes: Janet S. Ziegle (Applied Biosystems, Foster City, California, United States); Pandikumar Swamikrishnan (IBM, Somers, New York, United States); Asif Javed, Laxmi Parida, Daniel E. Platt, and Ajay K. Royyuru (IBM, Yorktown Heights, New York, United States); Lluis Quintana-Murci (Institut Pasteur, Paris, France); R. John Mitchell (La Trobe University, Melbourne, Victoria, Australia); Danielle A. Badro, Marc Haber, and Pierre A. Zalloua (Lebanese American University, Chouran, Beirut, Lebanon); Syama Adhikarla, Arun Kumar Ganesh Prasad, Ramasamy Pitchappan, Kavitha Valampuri John, and Arun Varatharajan Santhakumari (Madurai Kamaraj University, Madurai, Tamil Nadu, India); Christoff J. Erasmus, Angela Hobbs, and Himla Soodyall (National Health Laboratory Service, Johannesburg, South Africa); Elena Balanovska and Oleg Balanovsky (Research Centre for Medical Genetics, Russian Academy of Medical Sciences, Moscow, Russia); Chris Tyler-Smith (The Wellcome Trust Sanger Institute, Hinxton, United Kingdom); Daniela R. Lacerda and Fabrício R. Santos (Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil); Pedro Paulo Vieira (Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil); Jaume Bertranpetit, David Comas, Begoña Martínez-Cruz, and Marta Melé (Universitat Pompeu Fabra, Barcelona, Spain); Christina J. Adler, Alan Cooper, Clio S. I. Der Sarkissian, and Wolfgang Haak (University of Adelaide, South Australia, Australia); Matthew E. Kaplan and Nirav C. Merchant (University of Arizona, Tucson, Arizona, United States); Colin Renfrew (University of Cambridge, Cambridge, United Kingdom); Andrew C. Clarke and Elizabeth A. Matisoo-Smith (University of Otago, Dunedin, New Zealand); Matthew C. Dulik, Jill B. Gaieski, Amanda C. Owings, and Theodore G. Schurr (University of Pennsylvania, Philadelphia, Pennsylvania, United States); David F. Soria Hernanz and R. Spencer Wells (National Geographic Society, Washington, District of Columbia, United States).

Grant sponsor: National Geographic Society; Grant sponsor: IBM; Grant sponsor: The Waitt Family Foundation; Grant sponsor: National Outstanding Youth Science Foundation of China; Grant number: 30625016; Grant sponsor: National Science Foundation of China; Grant numbers: 30760097, 30771180, 30890034; Grant sponsor: 863 Program; Grant number: 2007AA02Z312; Grant sponsor: Natural Science Foundation of Shanghai; Grant number: 10ZR1402200; Grant sponsor: Shanghai Leading Academic Discipline Project; Grant number: B111; Grant sponsor: China Postdoctoral Science Foundation; Grant number: 200902208.

*Correspondence to: Li Jin or Hui Li, School of Life Sciences, Fudan University, Shanghai 200433, China.
E-mail: JinLi.Fudan@gmail.com or LiHui.Fudan@gmail.com
Received 21 February 2010; accepted 3 May 2010

DOI 10.1002/ajpa.21350
Published online in Wiley InterScience (www.interscience.wiley.com).

west, and Qilian on the northeast). Despite its inhospitable environment, the Tibetan Plateau is now occupied by over seven million people (2000 census), mostly indigenous Tibetans. When and how modern human conquered the Tibetan Plateau is one of the most interesting questions of human evolution, however, little was known about the peopling of Tibet.

Some efforts have been made to estimate the age of the peopling of Tibet in the fields of archeology and linguistics. Archeological records of Late Paleolithic age dated the human presence on the plateau back to about 20–30 thousand years ago (KYA) (Aldenderfer and Yinong, 2004). However, the sparseness and discontinuity of archeological findings on this plateau have led to ambiguous explanations of the demographic history of indigenous Tibetans. In contrast to the old history of the Tibetans suggested by archeology, linguistic studies (Van Driem, 1998, 2001, 2002, 2005) postulated complex models and favored a much more recent Neolithic peopling of the plateau. To clarify the inconsistency of the timing by different disciplines, more solid evidences are required, such as genetic diversity in Tibet.

During the past two decades, genetic markers (especially mitochondrial and Y chromosomal markers) were proved extremely useful in tracing population history in East Asia (Ding et al., 2000; Jin and Su, 2000; Zhang et al., 2007). However, previous genetic studies on the Tibetans also issued inconsistent results. Some studies using classical genetic traits (Du et al., 1997), autosomal microsatellite markers (Gayden et al., 2009; Kang et al., 2010) and mitochondrial DNA (mtDNA) (Torroni et al., 1994) suggested a north Asian origin of Tibetans. While evidences from the Y chromosomal *Alu* insertion (YAP) marker revealed much more intricate stories for the origin of Tibetan peoples (Hammer et al., 1997; Qian et al., 2000; Su et al., 2000; Shi et al., 2008). The YAP polymorphism was enriched on the Tibetan Plateau, Japan, and Andaman islands, but almost absent in other regions. This contradiction has been attributed to different demographic histories between patrilineal and matrilineal genetic pools, but it is still highly debated concerning the resolution of the limited markers used and the incomplete sampling in each study. More detailed works about Tibetan populations both on mitochondrial DNA and Y chromosome DNA are definitely required.

Maternally inherited mtDNA plays an important role in studying modern human migrations (Wallace et al., 1999; Pakendorf and Stoneking, 2005) and probable selection effects (Ruiz-Pesini et al., 2004; Ingman and Gyllensten, 2007). Over the past few years, genetic studies about mtDNA of populations residing on and around the Tibetan Plateau have been conducted (Qian et al., 2001; Torroni et al., 1994; Wen et al., 2004b; Yao and Zhang, 2002), which led to a number of important insights into the genetic history of Tibetans and their adaptive process. However, most of these studies focused on the short sequence of mitochondrial control region, lacking power to distinguish detailed population histories. The recent development of mtDNA analysis based on complete sequencing made it possible to reconstruct maternal phylogeny of Tibetan populations thus may shed light on their genetic structure and population history. Gu et al. (2008, 2009) first used this analysis in comparing mtDNA genomes of Tibetans and Han Chinese and proposed possible selective effects in Tibetan populations. However, they did not have enough Tibetan population samples and therefore did not draw any conclusions on the origin of the Tibetans. Actually, those previous studies all suffered from insufficient population sample coverage of the plateau (a maximum of three population samples), lacking power of detecting substructure among Tibetan populations.

A recent article about mtDNA genome variation within six regional Tibetan populations has revealed successful Late Paleolithic settlement on the Tibetan Plateau (Zhao et al., 2009). In this study, we further sampled eleven Tibeto-Burman populations from nine different geographic regions of the Tibetan Plateau. On the basis of our increased population coverage and comprehensive mtDNA information with complete mtDNA sequences, we traced the origin and the expansion of Tibetan populations on the plateau. Our findings indicate that Tibetan maternal gene pool consists of both pre- and post-LGM components. In addition, we found a signal of post-LGM population expansion on the plateau.

MATERIALS AND METHODS

Population samples and DNA extraction

Samples were collected by buccal swab from 562 anonymous and unrelated volunteers living on in the Tibetan Plateau with appropriate informed consent under protocols approved by the relevant institutional review boards. These population samples included nine Tibetan populations (46 individuals from Ngari, 58 from Nagqu, 59 from Shigatse, 59 from Lhasa, 61 from Chamdo, 56 from Shannan, 53 from Nyingchi, 44 from Yushu of Qinghai, 55 from Garze of western Sichuan), one Monba population (51 from Nyingchi), and one Lhoba population (20 from Shannan). The sampling locations were illustrated in Figure 1A. Genomic DNA was extracted using Silica/GuSCN method (Supprting Information S1).

Sequencing the control region and typing SNPs in the coding regions

The hypervariable segment I (HVS-I) of the control region was amplified by primers L15974 and H16488 (Yao et al., 2002). Purified PCR products were sequenced using the BigDye terminator cycle sequencing kit and an ABI 3130XL genetic analyzer (Applied Biosystems). Samples with a poly-cytosine (poly-C) tract caused by the transition at nucleic position (np) 16189 were sequenced in both directions. Overall, 562 HVS-I sequences have been submitted to GenBank. A SNaPshot assay was used for typing SNPs in the coding regions to resolve haplogroup determination. This assay was designed as a multiplex panel including 21 coding region SNPs and one length variation marker (for information of markers and primers see Supprting Information Table S2). All haplogroup F (10310G) samples were screened at np 12406 using enzyme *Hpa*I digesting. Some newly defined haplogroup diagnostic SNPs were checked by directly sequencing of appropriate fragments. The mtDNA haplogroup nomenclature used here was following with PhyloTree.org mtDNA tree Build 6 (28 September, 2009) (van Oven and Kayser, 2009) and previous studies (Tanaka et al., 2004; Kong et al., 2003, 2006; Derenko et al., 2007; Soares et al., 2008).

Sequencing whole mtDNA genome

Whole mtDNA genome sequencing was performed for a selection of 31 samples following methods described by

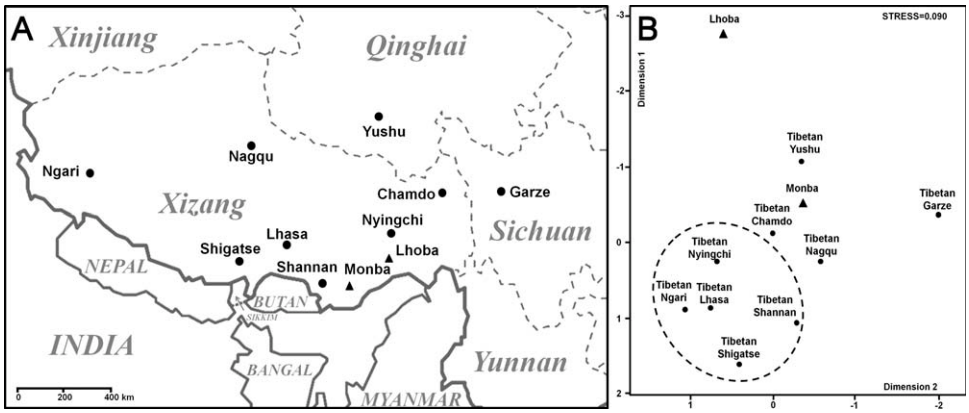

Fig. 1. A good correlation between the geographic locations and mtDNA HVS-I region diversities of the Tibetan samples. (**A**) Geographic locations of the 11 indigenous populations in the Tibetan plateau. (**B**) Two-dimensional MDS plot based on an *F*st distance matrix calculated from 440-bp-length sequences of mtDNA HVS-I region.

Torroni et al. (2001). Sequencing trace files were edited and aligned with the DNASTAR software (DNASTAR, Inc.). Mutations were scored relative to the revised Cambridge Reference Sequence (rCRS) (Andrews et al., 1999).

Data analysis

For HVS-I sequence data; median joining networks (Bandelt et al., 1995) were constructed using the free software package Network 4.1 (Fluxus-engineering.com). Descriptive statistical indexes, the Tajima's *D* (Tajima, 1989) and Fu's *Fs* (Fu, 1997) neutrality tests, pairwise distances between populations (F_{ST}), mismatch distribution, and analysis of molecular variance (AMOVA) (Excoffier et al., 1992) were calculated using Arlequin software (version 3.11) (Excoffier et al., 2005). The mutation model used here was Kimura two-parameter model with a Gamma shape parameter (alpha) of 0.26.

For complete sequences, we first constructed ML phylogenetic trees with PhyML3.0 (Ronquist and Huelsenbeck, 2003) under HKY+G mutation model with an alpha parameter of 0.12 (Macaulay et al., 2005) for 100 boostrap runs. A consensus tree was obtained through these 100 bootstap trees by CONSENSUS implemented in PHYLIP package (version 3.69) (Felsenstein, 1989). Then the assumption of a molecular clock was tested with the PAML package (Yang, 1997) under the same evolutionary model. The null hypothesis of a molecular clock cannot be rejected ($P = 0.797$). The complete mtDNA phylogeny was constructed and verified with several runs using Network (version 4.5) (Fluxus-engineering.com). The highly variable site 16519 and the length variation in the poly-C stretches between np 16180–16193 and 309–315 were discarded for phylogeny construction. The coalescence times were estimated with the ρ statistics (Forster et al., 1996), and standard errors were calculated following Saillard et al. (2000). For estimating the time to the most recent common ancestor (TMRCA) of each cluster, we used Mishmar rate (Saillard et al., 2000) and modified Mishmar rate (Perego et al., 2009) for coding region sequences (from np 577 to 16023); modified Kivisild rate (Perego et al., 2009) and Soares synonymous rate for synonymous mutations; and Soares rate (Soares et al., 2009) for complete mitochondrial genomes (all the substitutions excluding the 16519 mutations and the 16182C, 16183C, and 16194C). One hundred and forty-nine additional complete sequences from the literature were employed for tree reconstruction and age estimation (Ingman et al., 2000; Mishmar et al., 2003; Tanaka et al., 2004; Macaulay et al., 2005; Starikovskaya et al., 2005; Kong et al., 2003, 2006; Derenko et al., 2007; Ingman and Gyllensten, 2007; Tamm et al., 2007; Bilal et al., 2008; Hartmann et al., 2008; Soares et al., 2008; Chandrasekar et al., 2009; Zhao et al., 2009).

Principal-component analysis and multidimensional scaling plot

Principal-component (PC) analysis was performed using mtDNA haplogroup frequencies as input vectors by SPSS15.0 software (SPSS). Nonparametric multidimensional scaling (MDS) analysis based on F_{ST} statistics calculated from HVS-I sequences was also performed using SPSS15.0 software (SPSS) to visualize relationships among Tibetan populations and other Asian populations around. Population data of mtDNA diversity in East (Qian et al., 2001; Yao et al., 2002; Yao and Zhang, 2002; Wen et al., 2004a,b, 2005; Li et al., 2007), North (Horai et al., 1996; Kong et al., 2003; Derenko et al., 2007; Volodko et al., 2008), Central (Yao et al., 2000, 2004; Comas et al., 2004; Quintana-Murci et al., 2004; Heyer et al., 2009), South (Cordaux et al., 2003; Fornarino et al., 2009), and Southeast (Fucharoen et al., 2001; Hill et al., 2006, 2007) Asia were retrieved from the literature and included in our comparative analysis.

RESULTS

MtDNA haplogroup profiles

Detailed sequence variations and haplogroup assignments of 562 mtDNAs from 11 populations in this study are presented in Supporting Information Table S3. A total of 48 haplogroups or paragroups (unclassified lineages within a clade marked with an asterisk [*]) were observed in our samples, all within the three principal non-African macrohaplogroups: M, N, and R (under N). Table 1 presents the haplogroup frequencies of the studied populations. The majority of the mtDNA lineages belong to eastern Eurasian groups. Only 2.4% mtDNAs can be traced for their origins to western or southern Eurasia, including one haplogroup J1b1 sample, one U2,

TABLE 1. Haplogroup frequencies of mtDNA of the 11 indigenous populations on the Tibetan plateau

	Tibetan										
Haplogroup	Yushu (44)	Nagqu (58)	Chamdo (61)	Ngari (46)	Nyingchi (53)	Shannan (56)	Shigatse (59)	Lhasa (59)	Garze (55)	Monba (51)	Lhoba (20)
A4	13.6	6.9	1.6	8.7	9.4	5.4	5.1	3.4	10.9	23.5	5.0
A7	...	...	...	...	...	...	...	...	...	...	10.0
A10	4.5	1.7	1.6	10.9	3.8	16.1	13.6	5.1	1.8	...	...
B*	...	...	1.6	...	...	...	...	...	...	...	...
B4*	4.5	...	3.3	...	...	3.6	...	...	3.6	2.0	...
B4a	4.5	1.7	...	2.2	...	...	...	...	...	2.0	...
B5b	...	...	1.6	...	...	1.8	...	...	3.6	...	...
C*	...	...	...	...	...	...	...	...	...	3.9	...
C4	4.5	...	1.6	2.2	3.8	7.1	6.8	...	...	...	...
C4d	2.3	3.4	1.6	2.2	1.9	1.8	...	...	5.5	...	5.0
C5	...	...	1.6	...	...	...	...	...	...	...	...
D2b	...	...	1.6	...	1.9	...	...	...	...	...	...
D4	15.9	6.90	6.56	4.35	15.09	14.29	6.78	10.17	12.73	1.96	15
D4j1	...	3.45	1.64	4.35	3.77	3.57	1.69	...	3.64	1.96	...
D5	...	3.4	3.3	...	...	3.6	...	1.7	...	2.0	...
D5a2	4.5	3.4	1.6	4.3	3.8	...	...	1.7	...	3.9	15.0
F*	2.3	...	...	...	...	...	...	...	...	...	...
F1	...	1.7	1.6	2.2	1.9	...	...	1.7	...	3.9	...
F1a	2.3	...	...	...	...	...	...	...	...	...	...
F1b	...	8.6	8.2	2.2	7.5	3.6	13.6	15.3	7.3	13.7	5.0
F2a	...	...	4.9	...	...	...	1.7	...	1.8	...	...
G*	2.3	6.9	3.3	...	3.8	...	1.7	1.7	1.8	...	5.0
G1a1	...	...	...	...	...	...	1.7	1.7	...	...	...
G2	2.3	1.7	3.3	4.3	...	1.8	1.7	6.8	5.5	5.9	...
G3	2.3	6.9	4.9	2.2	1.9	...	...	3.4	9.1	5.9	10.0
J1b1	...	...	...	...	...	...	1.7	...	...	...	...
M*	6.8	...	...	...	1.9	...	1.7	1.7	1.8	5.9	10.0
M7b2	2.3	...	...	...	...	...	...	...	...	...	...
M7c	...	...	...	...	...	...	...	...	1.8	...	...
M10	2.3	...	1.6	2.2	1.9	...	...	...	...	...	...
M11a	...	...	1.6	...	1.9	...	...	...	...	...	...
M13*	...	1.7	...	...	...	...	...	...	1.8	2.0	...
M13a	12.3	1.7	1.6	...	...	...	...	...	7.3	...	...
M13b	...	3.4	1.6	4.3	3.8	1.8	1.7	...	3.6	3.9	5.0
M62	...	5.2	1.6	...	1.9	5.4	1.7	3.4	...	...	...
M8a	...	...	1.6	...	...	...	...	1.7	3.6	...	...
M9a	2.3	6.9	13.1	23.9	13.2	8.9	25.4	25.4	1.8	11.8	...
M9d	6.8	6.9	13.1	10.9	13.2	16.1	6.8	8.5	1.8	3.9	10.0
M9e	...	8.6	...	...	...	...	...	...	...	...	...
N*	...	1.7	...	...	...	1.8	...	...	...	...	...
N9a	...	...	...	...	1.9	...	...	...	...	...	...
R*	6.8	...	...	2.2	1.9	...	...	...	...	...	...
R9b	...	1.7	...	...	...	...	...	...	...	...	...
T1	...	...	...	4.3	...	...	.....	3.6	...	...	
U2	...	...	1.6	...	...	...	...	...	...	...	...
U7	...	3.4	...	...	...	...	3.4	1.7	...	2.0	...
U5	2.3	...	...	...	...	...	...	...	...	...	...
Z	...	1.7	3.3	...	...	3.6	1.7	5.1	1.8	...	...

one U5, six U7, and four T1 samples. This pattern is consistent with the notion that Himalayas served as a barrier to gene flow (Gayden et al., 2007, 2009). Among the eastern Eurasian component, 67.6% belongs to macrohaplogroup M and its derived haplogroups (such as C, D4, D5, G, M*, M7, M8, M9, M10, M11, M13, and Z), and 30.1% belongs to macrohaplogroup N and its nested haplogroups (A, B, F, N9a, R9b, N*, and R*). Within macrohaplogroup M, haplogroup D and M9 are the most common lineages, accounting for 20.3% and 13.6% of all samples, respectively. The most prevalent haplogroups within macrohaplogroup N, haplogroup A and F represented 8.6% and 11.5% of the studied lineages. Although the Tibetan Plateau is located in southwestern part of China, the majority of maternal gene pool on the plateau showed a strong similarity with north Asian populations, which was evident by the high proportion of north Asian-prevalent haplogroups (Tanaka et al., 2004; Derenko et al., 2003, 2007), such as haplogroups A, C, D4, D5, F1b, M9, and G. It is also worth noting that haplogroup B, which is one of the most common lineages in southern and eastern Asian populations, exhibited low frequency at the southern part of the plateau, accounting for only 5.4% in Shannan Tibetans, 4.0% in Monba, and it was not found in Nyingchi, Shigatse, and Lhasa Tibetans, nor in Lhoba.

Population summary statistics

Internal population diversity indices and results of Tajima's *D* and Fu's *Fs* neutrality tests are presented in Table 2. All the studied populations exhibited high and

TABLE 2. Diversity indices and results of neutrality tests of the 11 indigenous populations on the Tibetan Plateau, based on 440-bp-length sequences (from np 16024–16463) of HVS-I region

Population	n[a]	H[b]	K[c]	S[d]	Pi[e] (SE)	θ_k (95% CI)	Tajima' $D(P)$[f]	Fu's $Fs(P)$[f]
Lhoba	20	0.9842 (0.0205)	17	34	6.884 (3.776)	50.68 (18.75–149.75)	−1.11 (0.125)	−7.86
Monba	51	0.9702 (0.0125)	33	45	6.138 (2.968)	39.38 (22.64–69.37)	−1.32 (0.072)	−20.15
Tibetan_Ngari	46	0.9778 (0.0125)	35	51	6.594 (3.172)	65.06 (34.90–125.38)	−1.51	−24.69
Tibetan_Chamdo	61	0.9896 (0.0065)	51	68	6.411 (3.783)	143.83 (78.19-–276.80)	−1.91	−25.17
Tibetan_Garze	55	0.9886 (0.0059)	42	50	6.194 (2.988)	79.44 (44.62–145.71)	−1.48	−25.22
Tibetan_Lhasa	59	0.9813 (0.0114)	49	63	6.145 (2.964)	133.26 (72.24–257.01)	−1.87	−25.23
Tibetan_Nyingchi	53	0.9855 (0.0080)	41	56	5.901 (2.863)	81.25 (44.92–151.86)	−1.80	−25.30
Tibetan_Nagqu	58	0.9897 (0.0056)	45	61	6.635 (3.178)	90.36 (51.05–164.96)	−1.70	−25.12
Tibetan_Yushu	44	0.9937 (0.0067)	39	60	6.937 (3.324)	160.77 (73.29–382.20)	−1.76	−25.08
Tibetan_Shigatse	59	0.9468 (0.0219)	38	59	6.338 (3.048)	45.06 (26.86–76.43)	−1.71	−24.60
Tibetan_Shannan	56	0.9812 (0.0089)	41	53	6.135 (2.962)	67.46 (38.68–120.46)	−1.60	−25.24

[a] Sample size.
[b] Haplotype diversity.
[c] Number of different haplotypes.
[d] Number of segregating sites.
[e] Average number of pairwise differences.
[f] All P values are <0.05 except where noted.

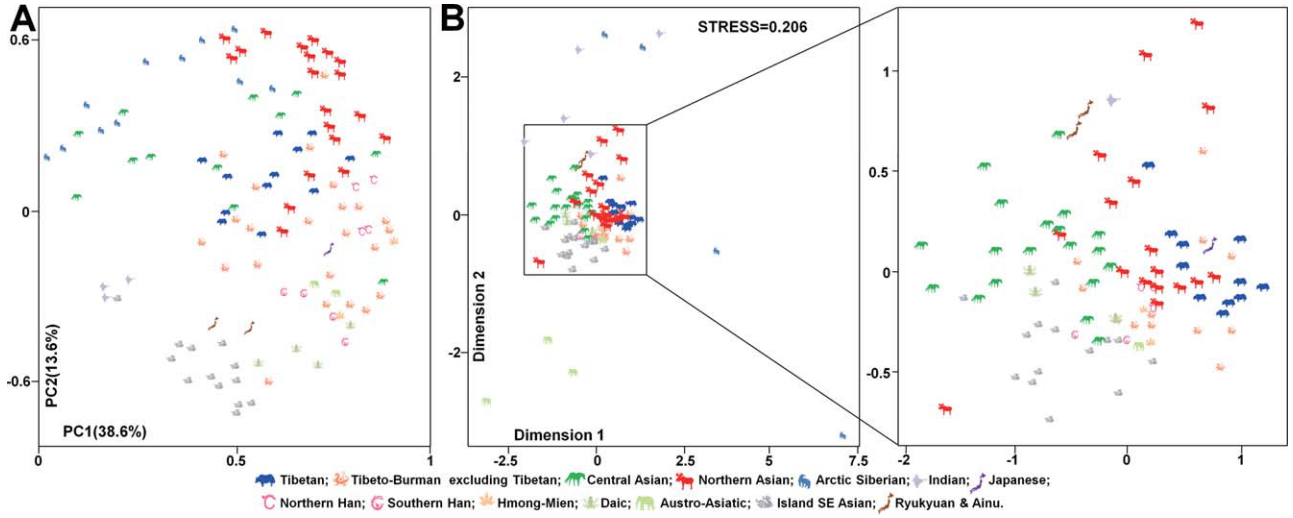

Fig. 2. Principal component plot (**A**) and multidimensional scaling (**B**) of 112 populations residing in eastern Eurasia. Populations are grouped with either geographical distribution or linguistic classification as shown in the plot. [Color figure can be found in the online issue, which is available at www.interscience.wiley.com]

similar diversity levels. The haplotype diversity (H) ranged from 0.947 to 0.994, and the mean number of pairwise differences observed (π) ranged from 5.90 to 6.94. Most population samples yielded significant negative values for both Tajima's D and Fu's Fs neutrality tests, suggesting historical population expansion. Lhoba and Monba population showed significant negative Fu's Fs value, but the results of Tajima's D test were not significantly different from zero. This contradiction perhaps resulted from mutation-rate heterogeneity of the HVS-I region, which can influence the signature of population expansion in Tajima's D test as previously reported (Tanaka et al., 2004; Derenko et al., 2003, 2007).

PC analysis and MDS plot

Samples from the 11 populations in current study and 101 populations residing in eastern Eurasia were included in principal component analysis. All of these 112 population samples were plotted using the first two PCs (Fig. 2A), which accounted for 38.6% and 13.6% of the total variation, respectively. These first two PCs mainly reflected geographic distributions of the populations. Tibetan populations clustered closely with northern Asian populations, northern Han Chinese, and other Tibeto-Burman populations residing in Nepal and India. This close relationship may indicate a common ancestor of their maternal gene pools. In contrast, Tibeto-Burman populations from southern China scattered widely and part of them clustered closely to populations from southern East Asia, reflecting strong influences of southern East Asian inhabitants on maternal diversity of southern Tibeto-Burman populations as we previous reported (Wen et al., 2004b).

The strong affinity of Tibetans and northern Asian populations was also reflected in MDS plot (Fig. 2B). Central Asian populations were segregated from the Tibetans clearly by the northern Asian populations in the first dimension of the MDS plot. Southern Tibeto-Burman populations distributed closely in the plot with southern Han Chinese, Daic populations, and Southeast Asian populations.

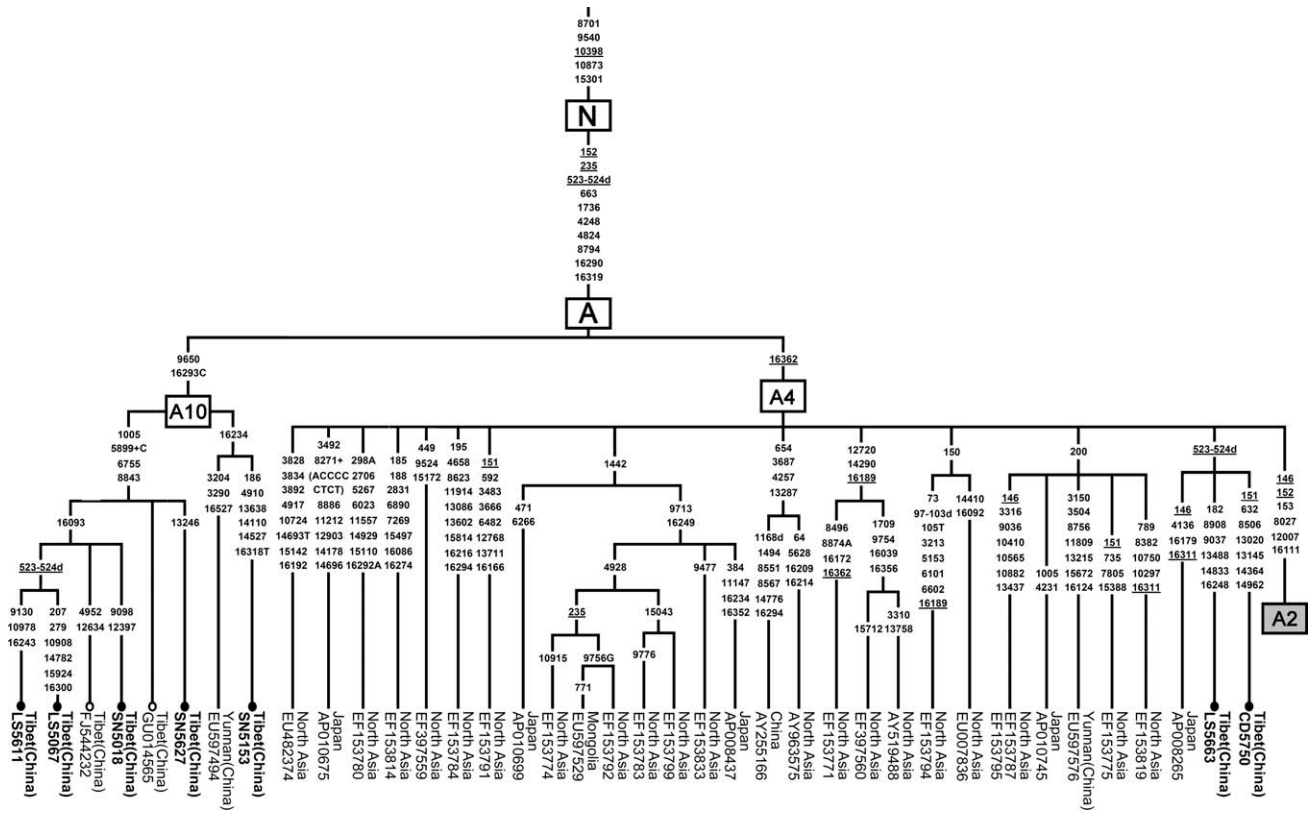

Fig. 3. The phylogenetic tree of haplogroup A complete mtDNA sequences. The tree is rooted in macrohaplogroup N. Mutations are scored relative to the rCRS (Andrews et al., 1999). Transitions are presented as mutation positions, and only transversions are specified. Deletions are indicated by a "d," and insertions are indicated by a "+" followed with the inserted nucleotides. All the recurrent mutations are underlined. TMRCAs are also shown beside nodes. For sequences taken from the literature, accession numbers are indicated followed with their geographic locations. New sequences generated through this study are in bold.

The MDS plot of the populations on the plateau (Fig. 1B) showed a south-north pattern in terms of geographic location. Most of the Tibetan populations residing at the southern part of the plateau were grouped to the right upper quadrant. AMOVA showed that the difference between the southern and northern groups is small (0.80%) but significant ($P = 0.004$). Lhoba was separated from other populations in the first dimension. Considering that Adi, which shares common ancestry with Lhoba indicated by both ethnologic and linguistic information (Lewis, 2009), was composed of almost 100% of the Y chromosome haplogroup O3a3c-M134, while Tibetan and Monba populations retained a high proportion of haplogroup D-YAP+ (Su et al., 2000; Cordaux et al., 2004), it may not be surprising to observe such a great distance between Lhoba and the other Tibetan populations.

Tibeto-Burman associated mtDNA lineages

One of the most significant results of this study is revealing of the Tibeto-Burman associated mtDNA lineages. These lineages include our redefined subhaplogroups A10, C4d, and M13b, as well as the recently identified haplogroup M62 (Chandrasekar et al., 2009).

Subhaplogroup A10

Haplogroup A is one of the most common haplogroups in northern and eastern Asia with a moderate frequency from 5 to 10% (Derenko et al., 2007), nesting eight subhaplogroups (A2–A9) (van Oven and Kayser, 2009). This haplogroup was found at 14.4% on average in our Tibetan population samples, with the highest frequency in Monba (20.8%) and the lowest in Tibetans from Chamdo (3.3%). Only two previously proposed subclades (A4 and A7) were detected, where A7 was only found in two individuals of Lhoba with identical HVS-I motifs. On the other hand, a distinguishing HVS-I motif (16290-16293C-16319) was presented at a high frequency (39.5%) of all haplogroup A samples (ranging from 0% in Monba and Lhoba to 72.7% in Tibetans from Shigatse). Our database search for similar HVS-I led to 27 mtDNA sequences (Horai et al., 1996; Wen et al., 2004a,b; Yao et al., 2002, 2004). Half of these samples belonged to Tibeto-Burman populations, indicating their strong ethnic association. We assigned this sub-haplogroup as A10. To further investigate haplogroup A on the plateau, complete genome sequencing was performed in seven Tibetan mtDNAs of haplogroup A. Five of these selected samples (two from Lhasa Tibetans, three from Shannan Tibetans) represented our newly identified A10 subclade, and the other two belonged to A4 sublineage. All these new sequences and selected sequences taken from literature (Ingman et al., 2000; Starikovskaya et al., 2005; Derenko et al., 2007) were used to reconstruct a haplogroup A phylogeny, illustrated in Figure 3. The phylogenetic analysis demonstrated that A10 subclade split from the root of haplogroup A very early and formed a distinct lineage by one coding-region transition at np 9650. Four of our A10 samples harbored three additional

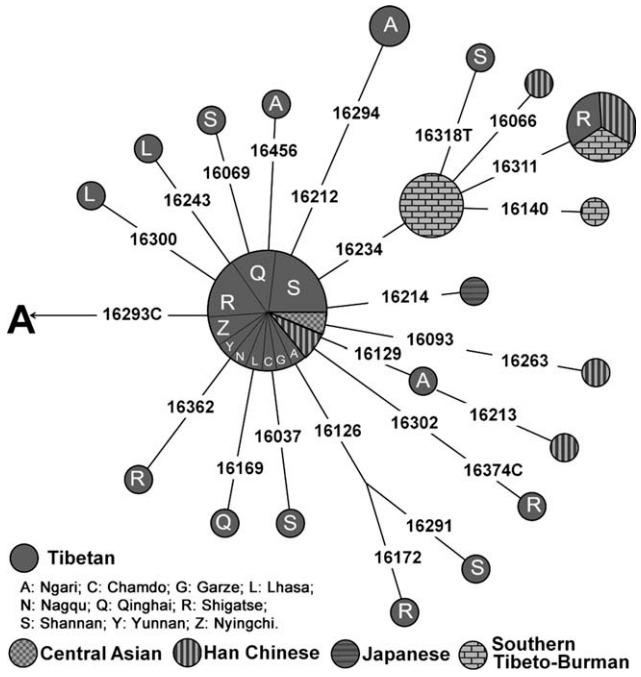

Fig. 4. Median-joining network of the newly identified subhaplogroup A10, based on HVS-I sequences between the region np 16024–16463.

mutations (at np 1005, 6755, 8843), named as A10a; one sequence from Shannan Tibetans and the only one published A10 full sequence from the Naxi population (neighboring the Tibetans) living in Yunnan China (Hartmann et al., 2008) had a control-region substitution at site 16234, assigned as A10b. It is noteworthy that sublineage A10b has been found outside the Tibetan Plateau (mostly in Tibeto-Burman speaking populations residing in Yunnan China), suggesting that these people migrated southward from the plateau. This observation is concordant with our previous conclusion about the genetic structure of Tibeto-Burman populations (Wen et al., 2004b). The Tibetan A4 sequences represented as novel haplotypes within the subhaplogroup A4 shared no coding-region mutation with any previously published A4 sequences (16362 is the only marker defining A4).

Network analysis of HVS-I sequences of lineage A10 revealed a star-like pattern and thus showed a signal of population expansion on the plateau (Fig. 4), which was also confirmed by unimodal mismatch distribution in this subhaplogroup (data not shown). The A10 lineage expansion time was estimated to 15,100 ± 4,700 YBP. Importantly, database searching also identified a similar HVS-I motif of 16242-16293C-16319 in Mongolia (Kolman et al., 1996) and southern Siberia (Derenko et al., 2003) populations. However, previously published complete sequences (Starikovskaya et al., 2005; Derenko et al., 2007) grouped these samples to northern Asian specific lineage A8, and thus they were excluded in Network analysis.

Haplogroup M62

Haplogroup M62 was first reported as a novel haplogroup in northeast India (Chandrasekar et al., 2009). However, recent published 13 complete sequences of M62 from Tibet indicated that it is Tibetan specific (Zhao et al., 2009). In this study, we confirmed this specificity by intensively coding region mutation screening. Three M62 samples were also selected to be complete sequenced to update the M62 phylogeny, which was illustrated in Figure 5. We revised the classification of haplogroup M62, as having three subclade: M62a, characterized by one control-region (203) transition; M62b, characterized by three transitions (3693, 6305, and 7364) and one transversion (187T); and M62c, characterized by one back mutation at np 204. Two Indian M62 sequences shared five transitions (146, 310, 4763, 9935, and 16147) with our sample from Shigatse (RKZ5474), nesting in M62c branch. Actually, these Indian sequences were also belonged to Tibeto-Burman speaking populations, indicating their origin in the Tibetan Plateau.

Subhaplogroup C4d

Haplogroup C is another common lineage that is widespread in East Asia and Siberia and is one of the founder lineages among Native Americans (Torroni et al., 1993). The entire C haplogroup is characterized by the HVS-I motif 16223-16327 and five mutations in the coding region (3552, 9545, 11914, 13263, and 14318). This haplogroup was detected in populations of the Tibetan Plateau with considerable frequencies (average 4.7%). Almost 40% (11/28) of our haplogroup C samples harbored a specific HVS-I motif 16093-16298-16327 and missed one of the characteristic mutations at np 16223. We sequenced two complete mtDNA genomes (LZ5332, SN5515) belonging to this type and performed phylogenetic analysis. The results demonstrated that these two sequences clustered closest with sublineage C4, but lacked an Adenine (A) insertion at np 2232, which is one of the four coding-region mutations (2232A ins., 6026, 11969, 15204) defining subhaplogroup C4 Starikovskaya et al., 2005). Additional screening at np 2232 in all these eleven samples bearing HVS-I motif 16093-16298-16327 confirmed that none of them had 2232A insertion. Searching for homologous HVS-I motifs in databases resulted only one identical subject from Qinghai Tibetans (Wen et al., 2004b), thus suggesting that this haplotype may be Tibetan specific. We then reconstructed the phylogeny of subhaplogroup C4 with our newly sequenced samples (LZ5332, SN5515, and SN5668) and all previously published C4 mtDNA genomes (Fig. 6). Subhaplogroup C4 was redefined here and was characterized by only three coding-region mutations (6026, 11969, 15204) and two control-region mutations (249 and 16327). We referred to this Tibetan-specific lineage as subhaplogroup C4d. TMRCA of subhaplogroup C4d was calculated to be 20,600 ± 6,800 years ago. It should be noted that two complete sequences share three additional mutations (7100, 12780, and 15236) in coding region, also indicating a significant divergence within the Tibetan-specific C4d branch.

Subhaplogroup M13b

Tibetan distinctiveness has also been found in haplogroup M13. This haplogroup is presented in East Asian populations at very low frequency and has been found sporadically in northern Asia (Fedorova et al., 2003; Pakendorf et al., 2003; Derenko et al., 2007). On the plateau, the frequency of haplogroup M13 is remarkable, accounting for 4.3% (24/562) samples. Moreover, 58.3% (14/24) of these M13 samples harbored a distinguishing

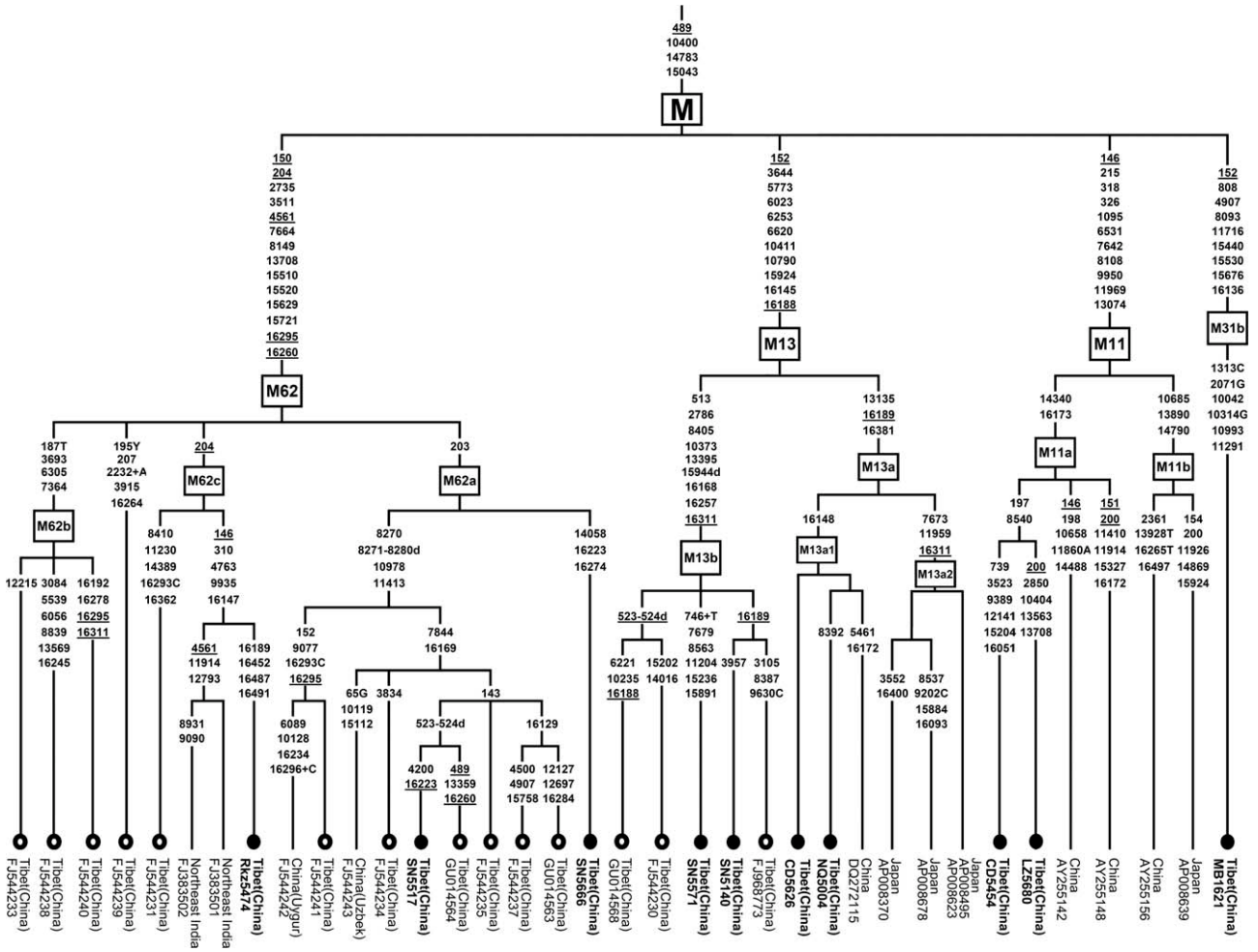

Fig. 5. The phylogenetic tree of complete mtDNA sequences of haplogroup M11, M13, and M62. The tree is rooted in macrohaplogroup M. For additional information, see the Figure 3 legend.

HVS-I motif: 16145-16168-16223-16257-16311. Extensive literature search revealed that this specific lineage was only observed in populations of Tibeto-Burman linguistic branch residing closely to Tibetan Plateau, e.g., one Tibetan (Yao and Zhang, 2002), one Bai and two Pumi (Wen et al., 2004b) from Yunnan of China, one Lisu and one White Karen individual from northern Thailand (Oota et al., 2001) (Supprting Information Table S4). Thus we defined this special mtDNA as M13b. We sequenced four complete Tibetan mtDNA genomes and compared them with all published M13 complete sequences (Fig. 5). A Mongolian sample (MG50) (Kong et al., 2006) nested with two Tibetan samples (CD5626, NQ5004) and formed a subclade, M13a1, characterized by control-region mutation at np 16148. The Japanese mtDNAs clustered into another branch, M13a2, characterized by two coding-region mutations (7673 and 11959). It was noteworthy that the two Tibetan M13b mtDNAs (SN5140, SN5571) differed from the rest of haplogroup M13 by six coding-region and four control-region transitions. The divergence time of M13a and M13b was calculated as 21,100 ± 6,100 years. With the exclusive distribution of subhaplogroup M13b in Tibeto-Burman populations, this deep divergence time may indicate the late Pleistocene modern human settlement and long-time isolation.

Other MtDNA lineages presented on the Tibetan plateau

Haplogroup M9

One major component of mtDNA pool on the plateau is represented by haplogroup M9, which distinguished the Tibetans from other East Asian populations. As previously reported (Kivisild et al., 2002; Yao et al., 2002; Soares et al., 2008), haplogroup M9 encompassed two subclades: E and M9a. While subhaplogroup E is restrictedly distributed in Island Southeast Asia (ISEA) and Taiwan (Trejaut et al., 2005; Hill et al., 2007; Soares et al., 2008), lineage M9a was widely presented in mainland East Asia and reached its greatest frequency and diversity in Tibet (Torroni et al., 1994; Tanaka et al., 2004). In this study, the average frequency of haplogroup M9a on the Tibetan Plateau was 21.6% with the highest frequencies in the ethnic Tibetans from Ngari (34.8%), Shigatse (32.2%), and Lhasa (32.8%). Sublineage M9a has been detected in central (Yao et al., 2004), northern (Derenko et al., 2007), and eastern (Tanaka et al., 2004; Soares et al., 2008) Asian populations, but all in low frequencies (<5%). In addition, it has been reported that the deepest branches of haplogroup M9 (pre-M9a) were found in Indochina, Mainland China, Taiwan, and the

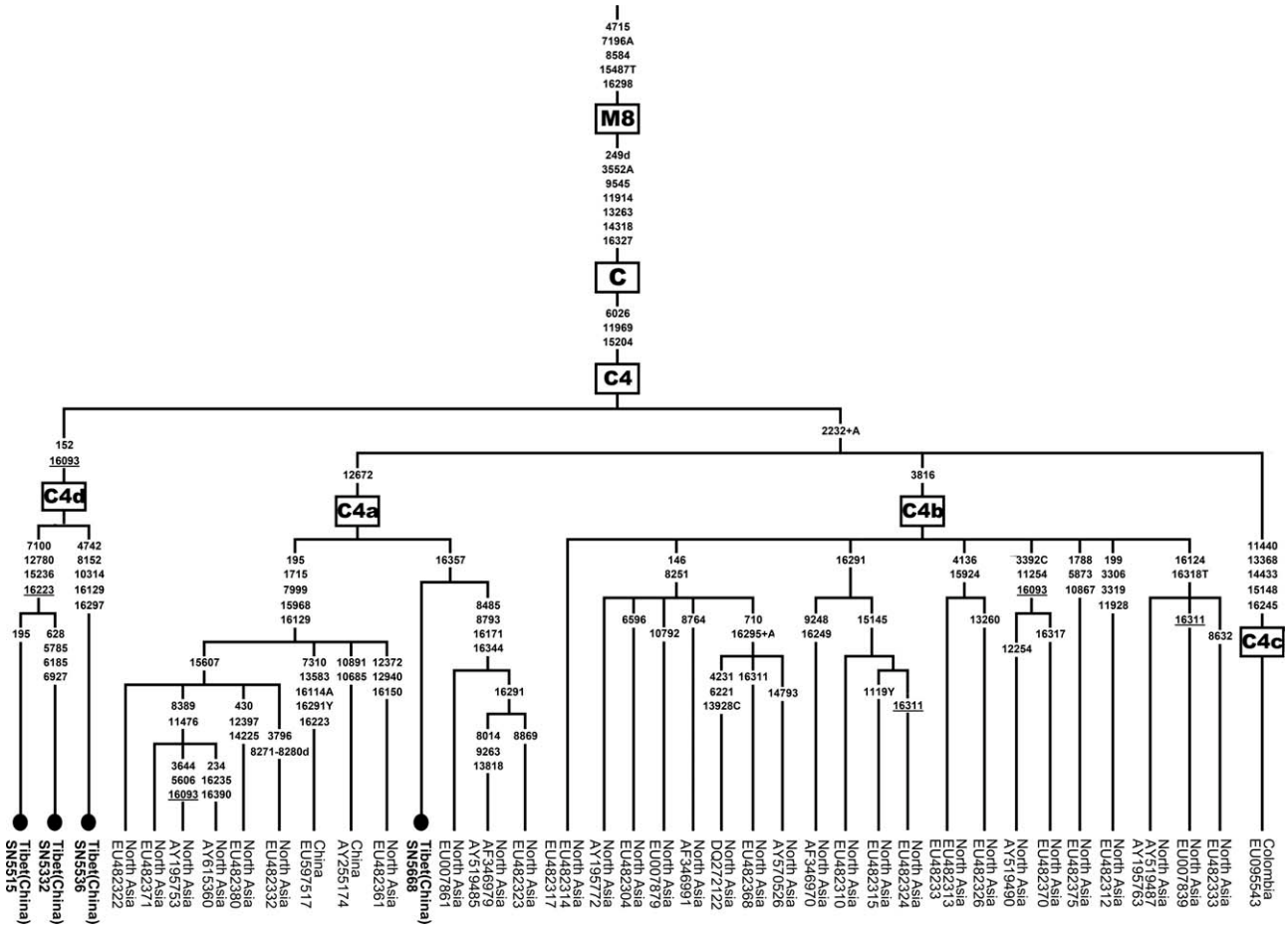

Fig. 6. The phylogenetic tree of complete mtDNA sequences of subhaplogroup C4. The tree is rooted in macrohaplogroup M. For additional information, see the Figure 3 legend.

next deepest in ISEA (Soares et al., 2008). On the plateau, M9 haplogroup is represented as four distinct HVS-I motif types: 16223-16234-16362 (pre-M9a), 16223-16234-16316-16362 (M9a), 16158-16223-16234-16362 (M9d), and 16145-16223-16234. To further assess the lineage M9 variations found in mitochondrial gene pool of the Tibetan Plateau, we selected 10 samples of all the above four different motifs and sequenced their complete mtDNA genomes. Combining all published haplogroup M9 (excluding sublineage E) mtDNA genomes (Tanaka et al., 2004; Kong et al., 2006; Bilal et al., 2008; Soares et al., 2008; Chandrasekar et al., 2009), and our nine newly collected samples, we reconstructed a tree of 56 complete sequences (Fig. 7). According to this updated phylogenetic tree, we defined one new M9 subclade and its diagnostic coding SNPs as: M9e (7256C). Interestingly, two sequences (CD5630, LS5629) harboring 16223-16234-16362 HVS-I motif, which have been designated as pre-M9a previously, fall into subclades M9d and M9a, respectively. Additional screening of mutations at np 9242, 12362, and 7256 in these pre-M9a samples revealed that they all belong to either M9a or M9d subclades. Notably we found that the only published M9d sequence (EF093545) from Southeast Asia nested in Tibeto-Burman samples, and the M9d frequency was much higher on the Tibetan Plateau (8.9%) than in Southeast Asia. Thus this M9d sequence may indicate a recent gene flow from Tibet to Southeast Asia.

Haplogroup D2b and D4j

The distribution of subhaplogroup D2b ("D2a" in (Derenko et al., 2007)) in Asian populations indicated a southern Siberian rather than Beringian origin of haplogroup D2 lineages. In our samples from the Tibetan Plateau, we found three mtDNA sequences (LZ1572, LZ1600, and CD5479) bearing similar haplogroup D2 HVS-I motif (16129-16223-16271-16362). Moreover, our results from complete mtDNA sequencing and selected variation checking revealed that two of these sequences (LZ1600 and CD5479) were similar to the southern Siberian-specific D2b sublineage but lacked the substitution at np 5004, which is one of the four mutations (at np 195, 5004, 9181, and 16092) defining subhaplogroup D2b as previously reported (Derenko et al., 2007). The updated phylogeny of D2b was illustrated in Figure 8. These two Tibetan sequences showed strong affinity with the southern Siberian D2b sequences. The other sequence (LZ1572), bearing transitions at np 5262 and 11969, grouped into D4j1 cluster (Fig. 8), which was newly identified in Tibeto-Burman populations from Northeast India (Chandrasekar et al., 2009). Screening

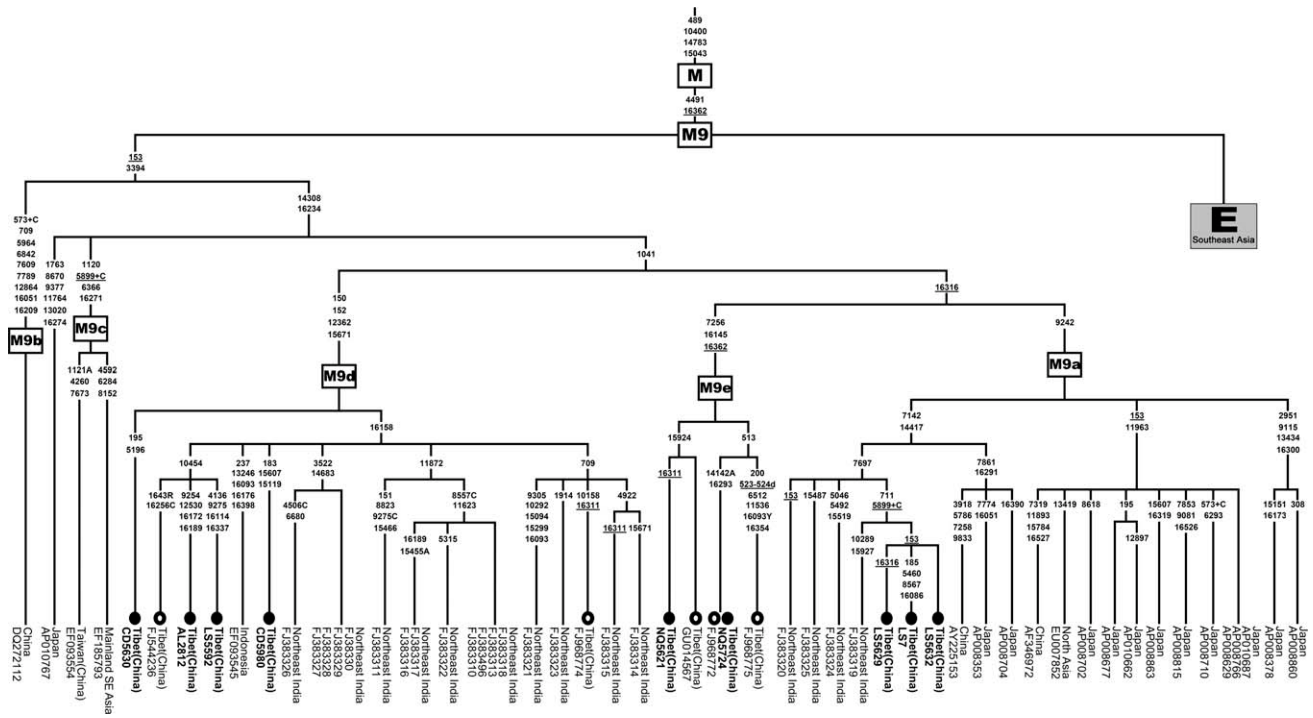

Fig. 7. The phylogenetic tree of complete mtDNA sequences of haplogroup M9. The sequences of subhaplogroup E are not presented in the phylogeny. The tree is rooted in macrohaplogroup M. For additional information, see the Figure 3 legend.

at np 5262 revealed 16% (12/75) of haplogroup D4 samples on the plateau belonged to D4j1.

Haplogroup M11 and M31b

We have also completely sequenced three M* genomes that cannot be further classified. Two sequences (LZ5454 and CD5680) fell into haplogroup M11 distinguished by seven coding-region mutations (1095, 6531, 7642, 8108, 9950, 11969, and 13074) and four mutations in HVS-II region (146, 215, 318, and 326). It should be noted that these two sequences showed deep divergence with each other, pointing to a remarkable divergence on the plateau. Another M* sample (MB1621) from the Monba population belonged sublineage Tibeto-Burman populations from Northeast India (Chandrasekar et al., 2009). Screening at np.

Time estimate

The coalescence time and variation computed from representative lineages on the plateau (A10, C4d, M62, M13, M13b, M11, M9a, M9d, M9e, and D2b) were given in Table 3. The old age and the restricted geographic distribution of haplogroup M62 on the Tibetan Plateau indicated that this haplogroup may have evolved in suit on the plateau before LGM, which have also been suggested by other study recently (Zhao et al., 2009). Other haplogroups such as A10, C4d, and M13b also showed strong association with the Tibetans. These lineages outside the plateau in very low frequency are either in the Tibeto-Burman ethnic groups which shared common ancestry with the Tibetans or in the populations living with the Tibetans (Supprting Information Table S3). The coalescence time estimates of the A10 and C4d subclades were also around 20KYA. However, the synonymous mutation methods issued more recent estimates, indicating the limitations of this method (Soares et al., 2009). Although most methods indicated the oldness of A10, C4d, and M13b, we cannot yet rule out the possibility of the post-LGM appearances of these lineages on the plateau, therefore, these lineages can only be regarded as putative pre-LGM lineages.

The ages of the three M9 subclades (M9a, M9d, and M9e) fell into early Holocene (about 12KYA), showing a remarkable gap with the ages of the putative pre-LGM lineages (around 20KYA). The Tibetan M9a samples all harbored three additional coding region mutations (7142, 7697, and 14417), indicating the recent appearance of this lineage in Tibetan populations. We also estimated the age of the D2b cluster as about 10KYA. The Tibetan samples of D2b nested in the Siberian prevalent lineages, indicating recent gene flows between the two regions.

DISCUSSION

Common ancestry of the Northern Asians and the Tibetan populations

Analyses of 562 mtDNA sequences from 11 Tibeto-Burman populations residing in different regions of the Tibetan Plateau showed that the maternal variation on plateau was largely contributed by northern Asian-prevalent haplogroups, thus testifying a common maternal ancestry of the Tibetan populations and northern Asian populations (Torroni et al., 1994; Gayden et al., 2009). This notion was confirmed by both PC analysis and MDA plots, where the Tibetan populations clustered closely to the northern Asian groups and separated from central and southern Asian populations. However, unlike the northern Asian populations, the Tibetan populations contents very low frequencies of western Eurasian

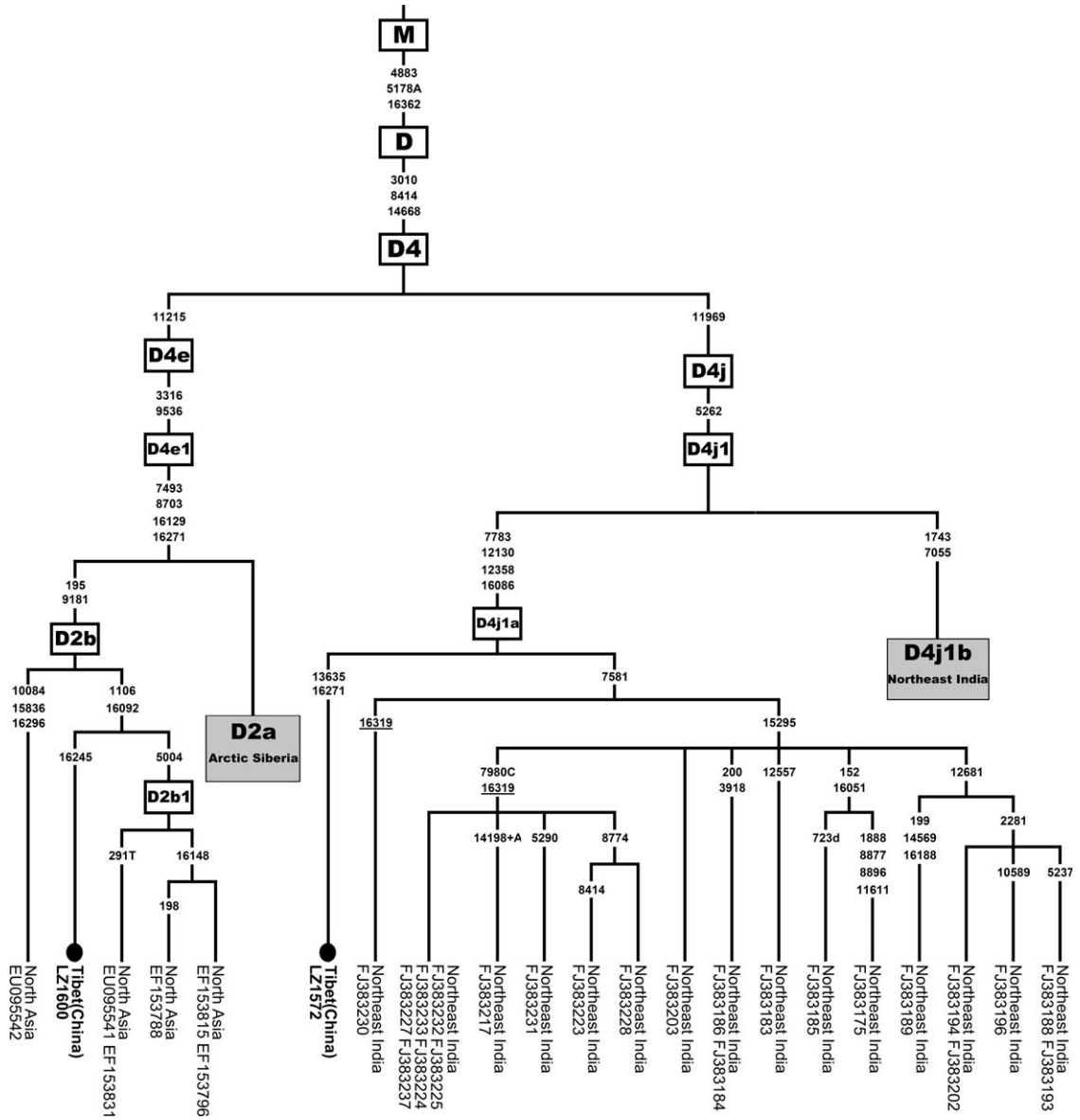

Fig. 8. The phylogenetic tree of complete mtDNA sequences of subhaplogroup D2 and D4j. The tree is rooted in macrohaplogroup M. For additional information, see the Figure 3 legend.

genetic components, indicating that the Tibetan Plateau is more like a "genetic capsulate" area. Moreover, extremely low frequencies of the South Asian lineages on the plateau reflected the strong "barrier effect" of the Himalayas between the Indian subcontinent and the Tibetan Plateau (Gayden et al., 2007; Kang et al., 2010).

Evidences of the pre-LGM human activity on the plateau

Archeological sites of more than 20,000 years old (Huang, 1994; Zhang and Li, 2002; Yuan et al., 2007) supported a pre-LGM human activity on the Tibetan Plateau. Notably, some of these pre-LGM Paleolithic sites were found in the northern part of the plateau, where there are very few current inhabitants living with harsh climate, indicating that most regions of the plateau was inhabited in the pre-LGM period. During the LGM, snow accumulated in the Tibetan Plateau and developed into greater glacier and permafrost (Lehmkuhl and Haselein, 2000; Zheng et al., 2003), which might have "closed" most of the plateau, resulting in a large decrease of the local human population.

Recently, Zhao et al. (2009) claimed that an infrequent haplogroup (M16) in the Tibetan populations may represent the genetic relics of the Late Paleolithic inhabitants on the plateau. In our study, we also found some old lineages (M62, A10, and C4d) which may be the remains of pre-LGM inhabitants, providing evidences for LGM survivals. However, the age of a single haplogroup may not be the age of a population demographic event. If the people who colonized the area carried part of the original diversity not just one founder, the TMRCA of a haplogroup can be much older than the colonization event, and this haplogroup can in some cases have reached much lower frequency in the source population because of drift. Conversely if the founding group had a small effective population size, the TMRCA of a haplogroup can be much

TABLE 3. Age estimations of some haplogroups found in Tibetan Plateau based on different calibration rates

Haplogroup	n	Mismar rate[a]		Modified Mismar rate[b]		Modified Kivisild rate[c]		Soares synonymous rate[d]		Soares complete genome rate[e]	
		$\rho \pm \sigma$	T(ky)	$\rho \pm \sigma$	T(ky)	$\rho \pm \sigma$	T(ky)	$\rho \pm \sigma$	T(ky)	$\rho \pm \sigma$	T(ky)
M62	18	4.75 ± 0.88	24.42 ± 4.54	4.75 ± 0.88	21.90 ± 4.07	2.94 ± 0.89	22.53 ± 6.81	2.94 ± 0.89	23.21 ± 7.02	8.61 ± 1.06	23.67 ± 6.06
M11	6	5.33 ± 1.20	27.41 ± 6.18	5.33 ± 1.20	24.59 ± 5.54	2.17 ± 0.73	16.58 ± 5.56	2.17 ± 0.73	17.08 ± 5.73	8.17 ± 1.48	22.38 ± 8.43
C4d	3	4.33 ± 1.45	22.27 ± 7.47	4.33 ± 1.45	19.98 ± 6.70	2.33 ± 0.88	17.85 ± 6.75	2.33 ± 0.88	18.40 ± 6.95	8.00 ± 2.05	21.89 ± 11.63
A10	8	5.00 ± 1.15	25.70 ± 5.89	5.00 ± 1.15	23.05 ± 5.28	1.63 ± 0.67	12.43 ± 5.15	1.63 ± 0.67	12.81 ± 5.31	6.88 ± 1.28	18.67 ± 7.16
M13	12	5.00 ± 0.82	25.7 ± 4.20	5.00 ± 0.82	23.05 ± 3.76	1.75 ± 0.61	13.39 ± 4.64	1.75 ± 0.61	13.80 ± 4.78	9.00 ± 1.15	24.81 ± 6.61
M13b	5	2.80 ± 0.85	14.39 ± 4.36	4.33 ± 1.45	12.91 ± 3.91	1.00 ± 0.60	7.65 ± 4.59	1.00 ± 0.60	7.88 ± 4.73	4.00 ± 1.06	10.61 ± 5.69
M9a	19	3.11 ± 0.70	15.96 ± 3.60	3.11 ± 0.70	14.32 ± 3.23	1.32 ± 0.53	10.07 ± 4.09	1.32 ± 0.53	10.37 ± 4.21	4.74 ± 0.83	12.64 ± 8.50
M9d	7	2.00 ± 0.83	10.28 ± 4.28	2.00 ± 0.83	9.22 ± 3.84	0.57 ± 0.57	4.37 ± 4.37	0.57 ± 0.57	4.51 ± 4.51	4.57 ± 1.37	12.18 ± 7.42
M9e	5	1.20 ± 0.85	6.17 ± 4.36	1.20 ± 0.85	5.53 ± 3.91	0.40 ± 0.28	3.06 ± 2.16	0.40 ± 0.28	3.15 ± 2.23	3.20 ± 0.94	8.43 ± 4.97
D2b	7	1.86 ± 1.22	9.55 ± 6.27	1.86 ± 1.22	8.56 ± 5.63	0.71 ± 0.71	5.46 ± 5.46	0.71 ± 0.71	5.46 ± 5.46	3.86 ± 1.21	10.22 ± 6.49

Coalescence times were estimatd based on rho(ρ) statistics with different mutation rates: Coding region substitutions.
[a]Mismar rate,
[b]coding region substitutions Modified Mismar rate,
[c]coding region synonymous transitions Modified Kivisild rate,
[d]coding region synonymous substitutions Soares synonymous rate, and
[e]complete genome substitution Soares rate. n indicates number of sequences, σ indicates stander error, and CI indicates confidence interval.

younger than the colonization time. In this study, time estimates of several haplogroups resulted in similar old ages of around 22,000 years, making it less possible to be an over estimate of the colonization time. Giving the inhospitable environment of the Tibetan Plateau and the population decreasing during LGM, the effective population size on the plateau might be very small at least in part of the history; therefore, the colonization time of the plateau could be older than our estimates. This consequence mostly supports the pre-LGM colonization of the plateau suggested by archeology findings.

Multiple origins of Tibetan populations

The exclusive distribution of mtDNA sublineages on the plateau enabled us to estimate the colonization time of the plateau. Some of these plateau-specific lineages (A10, C4d, and M62) seem to have evolved on the plateau for a long time, indicating that the first entry of modern human into the Tibetan Plateau might happened before LGM deteriorated the plateau climate substantially (Aldenderfer and Yinong, 2004). In this respect, subclade A10 provided further information about the pre-LGM settlers. Subhaplogroup A10 showed the highest frequency and diversity in Shannan and Shigatse of south Tibet. Considering the present suitable environment of south Tibet, especially that of Shannan, (Aldenderfer and Yinong, 2004), the warm valleys in south Tibet might have been the refugium for the pre-LGM settlers during LGM. In addition, this haplogroup showed a signal of population expansion, dating back to ~15 KYA, as a consequence of environment change after the LGM period.

Phylogeographic analysis of another major haplogroup on the plateau, M9, revealed a considerable genetic component of post-LGM migrants. However, the ancestral haplotypes of M9 was not found in the Tibetan populations, but in the populations of Southeast Asia, Japan, and coastal China (Jiangsu and Shandong province in China, author's unpublished data). Moreover, another lineage nested in haplogroup M9, subclade E, has a Pleistocene origination in east Sunda of Southeast Asia (Derenko et al., 2007; Soares et al., 2008). Combining the geographic distribution information of the M9 diversity, we supported the hypothesis that haplogroup M9 originated in Southeast Asia more than ~50,000 years ago and evolved into the subhaplogroup E and M9a-M9d, and M9a-M9d subsequently migrated northward probably during Late Pleistocene (Soares et al., 2008). Totally, the presence of only derived sublineages in the Tibetan Plateau indicates a counter-clockwise dispersal in mainland East Asia as previously proposed (Chaix et al., 2008). The high frequencies of sublineages M9a and M9d on the plateau may result from either demographic effect or selective effect (Gu et al., 2008), which requires further investigation.

The high frequency of Y chromosome polymorphic Alu insertion (YAP) in Tibet makes the origins of the Tibetans intriguing and controversial. Although multiple origins of the Tibetan people have been proposed by different studies in the last decade (Qian et al., 2000; Su et al., 2000), the sources and routes presented in such claims are highly debated (Qian et al., 2000; Su et al., 2000; Thangaraj et al., 2005; Shi et al., 2008). Central Asia has been considered as the one of the main sources of YAP (Qian et al., 2000; Su et al., 2000; Gayden et al., 2007). In this study, however, considering the extremely

low frequency of western or central Asian mtDNA haplogroups (2.3%), it is less likely that Central Asian is a major contributor, at least in the maternal aspect. Recently, Shi et al. (2008) demonstrated two independent Paleolithic dispersal events of modern human into East Asia of ~50 KYA (marked by Y haplogroup D-M174) and ~30 KYA (marked by Y haplogroup O-M175 and its derivatives) (Shi et al., 2005, 2008). Haplogroup D and O are both dominant Y haplogroups in Tibet, indicating multiple migrations into Tibet. The routes Shi et al. proposed of the two ancient human dispersals are consistent with our findings of pre- and post-LGM migrations to the Tibetan Plateau. This concurrence, however, needs to be investigated in our further study.

Possible bias in time estimates

The genetic time estimates may be affected by several factors. The current statistic method for time estimate is still developing. A recent study showed that molecular dating with the rho statistic could produce biased results with large asymmetrical variances (Cox, 2008). Also demographic histories (such as bottlenecks, founder events and changes in effective population size) can particularly distort date estimates with the rho statistic (Cox, 2008; Endicott et al., 2009). Therefore, our conclusion may be revised in case that bias in time estimate exists, and Tibetan populations may have demographic histories different from our hypothesis. Furthermore, mutation rates to be used in mtDNA dating are yet to be determined. Researchers have demonstrated that Mishimar mutation rate always gave much older dates for East Asian populations (Endicott et al., 2009). However, we used multiple mutation rates to eliminate the bias, and the results are fairly consistent. Thus, we interpreted our data as pre- and post-LGM Tibetan plateau inhabitants with cautions. To make more solid conclusion, we need to generate sufficient Tibetan mtDNA complete sequences and analyze data with several different methods (Endicott and Ho, 2008; Ho and Endicott, 2008; Sores et al., 2009).

ACKNOWLEDGMENTS

We thank all of the sample donors and those who helped with sample collection for making this work possible.

LITERATURE CITED

Aldenderfer M, Yinong Z. 2004. The prehistory of the Tibetan Plateau to the seventh century A.D.: perspectives and research from China and the West since 1950. J World Prehist 18:1–55.

Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet 23:147.

Aris-Brosou S, Excoffier L. 1996. The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. Mol Biol Evol 13:494–504.

Bandelt HJ, Forster P, Sykes BC, Richards MB. 1995. Mitochondrial portraits of human populations using median networks. Genetics 141:743–753.

Bilal E, Rabadan R, Alexe G, Fuku N, Ueno H, Nishigaki Y, Fujita Y, Ito M, Arai Y, Hirose N, Ruckenstein A, Bhanot G, Tanaka M. 2008. Mitochondrial DNA haplogroup D4a is a marker for extreme longevity in Japan. PLoS One 3:e2421.

Chaix R, Austerlitz F, Hegay T, Quintana-Murci L, Heyer E. 2008. Genetic traces of east-to-west human expansion waves in Eurasia. Am J Phys Anthropol 136:309–317.

Chandrasekar A, Kumar S, Sreenath J, Sarkar BN, Urade BP, Mallick S, Bandopadhyay SS, Barua P, Barik SS, Basu D, Kiran U, Gangopadhyay P, Sahani R, Prasad BV, Gangopadhyay S, Lakshmi GR, Ravuri R R, Padmaja K, Venugopal PN, Sharma MB, Rao VR. 2009. Updating phylogeny of mitochondrial DNA macrohaplogroup M in India: dispersal of modern human in South Asian Corridor. PLoS One 4:e7447.

Comas D, Plaza S, Wells RS, Yuldaseva N, Lao O, Calafell F, Bertranpetit J. 2004. Admixture, migrations, and dispersals in Central Asia: evidence from maternal DNA lineages. Eur J Hum Genet 12:495–504.

Cordaux R, Saha N, Bentley GR, Aunger R, Sirajuddin SM, Stoneking M. 2003. Mitochondrial DNA analysis reveals diverse histories of tribal populations from India. Eur J Hum Genet 11:253–264.

Cordaux R, Weiss G, Saha N, Stoneking M. 2004. The Northeast Indian Passageway: a barrier or corridor for human migrations? Mol Biol Evol 21:1525–1533.

Cox M. 2009. Accuracy of molecular dating with the Rho Statistic: deviations from coalescent expectations under a range of demographic models. Hum Biol 80:335–357.

Derenko M, Malyarchuk B, Grzybowski T, Denisova G, Dambueva I, Perkova M, Dorzhu C, Luzina F, Lee HK, Vanecek T, Villems R, Zakharov I. 2007. Phylogeographic analysis of mitochondrial DNA in northern Asian populations. Am J Hum Genet 81:1025–1041.

Derenko MV, Grzybowski T, Malyarchuk BA, Dambueva IK, Denisova GA, Czarny J, Dorzhu CM, Kakpakov VT, Miscicka-Sliwka D, Wozniak M, Zakharov IA. 2003. Diversity of mitochondrial DNA lineages in South Siberia. Ann Hum Genet 67:391–411.

Ding YC, Wooding S, Harpending HC, Chi HC, Li HP, Fu YX, Pang JF, Yao YG, Yu JG, Moyzis R, Zhang Y. 2000. Population structure and history in East Asia. Proc Natl Acad Sci USA 97:14003–14006.

Du R, Xiao C, Cavalli-Sforza LL. 1997. Genetic distances between Chinese populations calculated on gene frequencies of 38 loci. Sci China C Life Sci 40:613–621.

Endicott P, Ho SYW. 2008. A Bayesian evaluation of human mitochondrial substitution rates. Am J Hum Genet 82:895–902.

Endicott P, Ho SYW, Metspalu M, Stringer C. 2009. Evaluating the mitochondrial timescale of human evolution. Trends Ecol Evol 24:515–521.

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491.

Fedorova SA, Bermisheva MA, Villems R, Maksimova NR, Khusnutdinova EK. 2003. [Analysis of mitochondrial DNA haplotypes in yakut population]. Mol Biol (Mosk) 37:643–653.

Fornarino S, Pala M, Battaglia V, Maranta R, Achilli A, Modiano G, Torroni A, Semino O, Santachiara-Benerecetti SA. 2009. Mitochondrial and Y-chromosome diversity of the Tharus (Nepal): a reservoir of genetic variation. BMC Evol Biol 9:154.

Forster P, Harding R, Torroni A, Bandelt HJ. 1996. Origin and evolution of Native American mtDNA variation: a reappraisal. Am J Hum Genet 59:935–945.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915–925.

Fucharoen G, Fucharoen S, Horai S. 2001. Mitochondrial DNA polymorphisms in Thailand. J Hum Genet 46:115–125.

Gayden T, Cadenas AM, Regueiro M, Singh NB, Zhivotovsky LA, Underhill PA, Cavalli-Sforza LL, Herrera RJ. 2007. The Himalayas as a directional barrier to gene flow. Am J Hum Genet 80:884–894.

Gayden T, Mirabal S, Cadenas AM, Lacau H, Simms TM, Morlote D, Chennakrishnaiah S, Herrera RJ. 2009. Genetic insights into the origins of Tibeto-Burman populations in the Himalayas. J Hum Genet 54:216–223.

Gu ML, Wang YJ, Shi L, Jiang F, Qiu MJ, Lin KQ, Tao YF, Huang XQ, Liu B, Chu JY. 2008. [Comparative analysis of the complete mitochondrial genome between Tibetan and Han population]. Zhonghua Yi Xue Yi Chuan Xue Za Zhi 25:382–386.

Gu ML, Wang YJ, Shi L, Zhang YB, Chu JY. 2009. [Comparison on mitochondrial ATP6. ATP8 and Cyt b genes between Chinese Tibetans in three different zones: detecting the signature of natural selection on mitochondrial genome]. Yi Chuan 31:147–152.

Hammer MF, Spurdle AB, Karafet T, Bonner MR, Wood ET, Novelletto A, Malaspina P, Mitchell RJ, Horai S, Jenkins T, Zegura SL. 1997. The geographic distribution of human Y chromosome variation. Genetics 145:787–805.

Hartmann A, Thieme M, Nanduri LK, Stempfl T, Moehle C, Kivisild T, Oefner PJ. 2009. Validation of microarray-based resequencing of 93 worldwide mitochondrial genomes. Hum Mutat 30:115–122.

Heyer E, Balaresque P, Jobling M, Quintana-Murci L, Chaix R, Segurel L, Aldashev A, Hegay T. 2009. Genetic diversity and the emergence of ethnic groups in Central Asia. BMC Genet 10:49.

Hill C, Soares P, Mormina M, Macaulay V, Clarke D, Blumbach PB, Vizuete-Forster M, Forster P, Bulbeck D, Oppenheimer S, Richards M. 2007. A mitochondrial stratigraphy for island southeast Asia. Am J Hum Genet 80:29–43.

Hill C, Soares P, Mormina M, Macaulay V, Meehan W, Blackburn J, Clarke D, Raja JM, Ismail P, Bulbeck D, Oppenheimer S, Richards M. 2006. Phylogeography and ethnogenesis of aboriginal Southeast Asians. Mol Biol Evol 23:2480–2491.

Ho SYW, Endicott P. 2008. The crucial role of calibration in molecular date estimates for the peopling of the Americas. Am J Hum Genet 83:142–146.

Horai S, Murayama K, Hayasaka K, Matsubayashi S, Hattori Y, Fucharoen G, Harihara S, Park KS, Omoto K, Pan IH. 1996. mtDNA polymorphism in East Asian populations, with special reference to the peopling of Japan. Am J Hum Genet 59:579–590.

Huang W. 1994. The prehistoric human occupation of the Qinghai-Xizang plateau. Gottinger Geographische Abhandlungen 95:201–219.

Ingman M, Gyllensten U. 2007. Rate variation between mitochondrial domains and adaptive evolution in humans. Hum Mol Genet 16:2281–2287.

Ingman M, Kaessmann H, Paabo S, Gyllensten U. 2000. Mitochondrial genome variation and the origin of modern humans. Nature 408:708–713.

Jin L, Su B. 2000. Natives or immigrants: modern human origin in east Asia. Nat Rev Genet 1:126–133.

Kang LL, Li SL, Gupta S, Zhang YG, Liu K, Zhao JM, Jin L, Li H. 2010. Genetic structures of the Tibetans and the Deng people in the Himalayas viewed from autosomal STRs. J Hum Genet 55:270–277.

Kivisild T, Tolk HV, Parik J, Wang Y, Papiha SS, Bandelt HJ, Villems R. 2002. The emerging limbs and twigs of the East Asian mtDNA tree. Mol Biol Evol 19:1737–1751.

Kolman CJ, Sambuughin N, Bermingham E. 1996. Mitochondrial DNA analysis of Mongolian populations and implications for the origin of New World founders. Genetics 142:1321–1334.

Kong QP, Bandelt HJ, Sun C, Yao YG, Salas A, Achilli A, Wang CY, Zhong L, Zhu CL, Wu SF, Torroni A, Zhang YP. 2006. Updating the East Asian mtDNA phylogeny: a prerequisite for the identification of pathogenic mutations. Hum Mol Genet 15:2076–2086.

Kong QP, Yao YG, Liu M, Shen SP, Chen C, Zhu CL, Palanichamy MG, Zhang YP. 2003a. Mitochondrial DNA sequence polymorphisms of five ethnic populations from northern China. Hum Genet 113:391–405.

Kong QP, Yao YG, Sun C, Bandelt HJ, Zhu CL, Zhang YP. 2003b. Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences. Am J Hum Genet 73:671–676.

Lehmkuhl F, Haselein F. 2000. Quaternary paleoenvironmental change on the Tibetan Plateau and adjacent areas (Western China and Western Mongolia). Quat Int 65:121–145.

Lewis MP, editor. 2009. Ethnologue: languages of the world, 16th ed. Dallas: SIL International. Online version: http://www.ethnologue.com/.

Li H, Cai X, Winograd-Cort ER, Wen B, Cheng X, Qin Z, Liu W, Liu Y, Pan S, Qian J, Tan CC, Jin L. 2007. Mitochondrial DNA diversity and population differentiation in southern East Asia. Am J Phys Anthropol 134:481–488.

Macaulay V, Hill C, Achilli A, Rengo C, Clarke D, Meehan W, Blackburn J, Semino O, Scozzari R, Cruciani F, Taha A, Shaari NK, Raja JM, Ismail P, Zainuddin Z, Goodwin W, Bulbeck D, Bandelt HJ, Oppenheimer S, Torroni A, Richards M. 2005. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. Science 308: 1034–1036.

Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S, Brandon M, Easley K, Chen E, Brown MD, Sukernik RI, Olckers A, Wallace DC. 2003. Natural selection shaped regional mtDNA variation in humans. Proc Natl Acad Sci USA 100:171–176.

Oota H, Settheetham-Ishida W, Tiwawech D, and Stoneking TIM. 2001. Human mtDNA and Y-chromosome variation is correlated with matrilocal versus patrilocal residence. Nature Genet 29(Brief Communication):20–21.

Pakendorf B, Stoneking M. 2005. Mitochondrial DNA and human evolution. Annu Rev Genomics Hum Genet 6:165–183.

Pakendorf B, Wiebe V, Tarskaia LA, Spitsyn VA, Soodyall H, Rodewald A, Stoneking M. 2003. Mitochondrial DNA evidence for admixed origins of central Siberian populations. Am J Phys Anthropol 120:211–224.

Perego UA, Achilli A, Angerhofer N, Accetturo M, Pala M, Olivieri A, Kashani BH, Ritchie KH, Scozzari R, Kong QP, Myres NM, Salas A, Semino O, Bandelt HJ, Woodward SR, Torroni A. 2009. Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. Curr Biol 19:1–8.

Qian Y, Qian B, Su B, Yu J, Ke Y, Chu Z, Shi L, Lu D, Chu J, Jin L. 2000. Multiple origins of Tibetan Y chromosomes. Hum Genet 106:453–454.

Qian YP, Chu ZT, Dai Q, Wei CD, Chu JY, Tajima A, Horai S. 2001. Mitochondrial DNA polymorphisms in Yunnan nationalities in China. J Hum Genet 46:211–220.

Quintana-Murci L, Chaix R, Wells RS, Behar DM, Sayar H, Scozzari R, Rengo C, Al-Zahery N, Semino O, Santachiara-Benerecetti AS, Coppa A, Ayub Q, Mohyuddin A, Tyler-Smith C, Qasim Mehdi S, Torroni A, McElreavey K. 2004. Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. Am J Hum Genet 74:827–845.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.

Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC. 2004. Effects of purifying and adaptive selection on regional variation in human mtDNA. Science 303:223–226.

Saillard J, Forster P, Lynnerup N, Bandelt HJ, Norby S. 2000. mtDNA variation among Greenland Eskimos: the edge of the Beringian expansion. Am J Hum Genet 67:718–726.

Shi H, Dong YL, Wen B, Xiao CJ, Underhill PA, Shen PD, Chakraborty R, Jin L, Su B. 2005. Y-chromosome evidence of southern origin of the East Asian-specific haplogroup O3-M122. Am J Hum Genet 77:408–419.

Shi H, Zhong H, Peng Y, Dong YL, Qi XB, Zhang F, Liu LF, Tan SJ, Ma RL, Xiao CJ, Wells RS, Jin L, Su B. 2008. Y chromosome evidence of earliest modern human settlement in East Asia and multiple origins of Tibetan and Japanese populations. BMC Biol 6:45.

Soares P, Ermini L, Thomson N, Mormina M, Rito T, Röhl A, Salas A, Oppenheimer S, Macaulay V, Richards MB. 2009. Correcting for purifying selection: an improved human mitochondrial molecular clock. Am J Hum Genet 84:740–759.

Soares P, Trejaut JA, Loo JH, Hill C, Mormina M, Lee CL, Chen YM, Hudjashov G, Forster P, Macaulay V, Bulbeck D, Oppenheimer S, Lin M, Richards MB. 2008. Climate change and postglacial human dispersals in southeast Asia. Mol Biol Evol 25:1209–1218.

Starikovskaya EB, Sukernik RI, Derbeneva OA, Volodko NV, Ruiz-Pesini E, Torroni A, Brown MD, Lott MT, Hosseini SH, Huoponen K, Wallace DC. 2005. Mitochondrial DNA diversity in indigenous populations of the southern extent of Siberia, and the origins of Native American haplogroups. Ann Hum Genet 69:67–89.

Starikovskaya YB, Sukernik RI, Schurr TG, Kogelnik AM, Wallace DC. 1998. mtDNA diversity in Chukchi and Siberian Eskimos: implications for the genetic history of Ancient Beringia and the peopling of the New World. Am J Hum Genet 63:1473–1491.

Su B, Xiao C, Deka R, Seielstad MT, Kangwanpong D, Xiao J, Lu D, Underhill P, Cavalli-Sforza L, Chakraborty R, Jin L. 2000. Y chromosome haplotypes reveal prehistorical migrations to the Himalayas. Hum Genet 107:582–590.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Tamm E, Kivisild T, Reidla M, Metspalu M, Smith DG, Mulligan CJ, Bravi CM, Rickards O, Martinez-Labarga C, Khusnutdinova EK, Fedorova SA, Golubenko MV, Stepanov VA, Gubina MA, Zhadanov SI, Ossipova LP, Damba L, Voevoda MI, Dipierri JE, Villems R, Malhi RS. 2007. Beringian standstill and spread of Native American founders. PLoS One 2:e829.

Tanaka M, Cabrera VM, Gonzalez AM, Larruga JM, Takeyasu T, Fuku N, Guo LJ, Hirose R, Fujita Y, Kurata M, Shinoda K, Umetsu K, Yamada Y, Oshida Y, Sato Y, Hattori N, Mizuno Y, Arai Y, Hirose N, Ohta S, Ogawa O, Tanaka Y, Kawamori R, Shamoto-Nagai M, Maruyama W, Shimokata H, Suzuki R, Shimodaira H. 2004. Mitochondrial genome variation in eastern Asia and the peopling of Japan. Genome Res 14:1832–1850.

Thangaraj K, Chaubey G, Kivisild T, Reddy AG, Singh VK, Rasalkar AA, Singh L. 2005. Reconstructing the origin of Andaman Islanders. Science 308:996.

Torroni A, Miller JA, Moore LG, Zamudio S, Zhuang J, Droma T, Wallace DC. 1994. Mitochondrial DNA analysis in Tibet: implications for the origin of the Tibetan population and its adaptation to high altitude. Am J Phys Anthropol 93:189–199.

Torroni A, Rengo C, Guida V, Cruciani F, Sellitto D, Coppa A, Calderon FL, Simionati B, Valle G, Richards M, Macaulay V, Scozzari R. 2001. Do the four clades of the mtDNA haplogroup L2 evolve at different rates? Am J Hum Genet 69:1348–1356.

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo CM, Wallace DC. 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet 53:563–590.

Trejaut JA, Kivisild T, Loo JH, Lee CL, He CL, Hsu CJ, Lee ZY, Lin M. 2005. Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations. PLoS Biol 3:e247.

Van Driem G. 1998. Neolithic correlates of ancient Tibeto-Burman migrations. In: Blench R, Spriggs M, editors. Archaeology and language II. London: Routledge. p 67–102.

Van Driem G. 2001. Languages of the Himalayas: an ethnolinguistic handbook of the Greater Himalayan Region, 2 vols. Leiden, The Netherlands: Brill.

Van Driem G. 2002. Tibeto-Burman phylogeny and prehistory: languages, material culture, and genes. In: Bellwood P, Renfrew C, editors. Examining the farming/language dispersal hypothesis. Oxford: McDonald Institute for Archaeological Research, University of Cambridge. p 233–249.

Van Driem G. 2005. Tibeto-Burman vs Indo-Chinese: implications for population geneticists, archeologists and prehistorians. In: Sagart L, Blench R, Sanches-Mazas A, editors. The peopling of East Asia: putting together archaeology, linguistics and genetics. London: Routledge Curzon. p 81–106.

van Oven M, Kayser M. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat 30:e386–e394.

Volodko NV, Starikovskaya EB, Mazunin IO, Eltsov NP, Naidenko PV, Wallace DC, Sukernik RI. 2008. Mitochondrial genome diversity in arctic Siberians, with particular reference to the evolutionary history of Beringia and Pleistocenic peopling of the Americas. Am J Hum Genet 82:1084–1100.

Wallace DC, Brown MD, Lott MT. 1999. Mitochondrial DNA variation in human evolution and disease. Gene 238:211–230.

Wen B, Li H, Gao S, Mao X, Gao Y, Li F, Zhang F, He Y, Dong Y, Zhang Y, Huang W, Jin J, Xiao C, Lu D, Chakraborty R, Su B, Deka R, Jin L. 2005. Genetic structure of Hmong-Mien speaking populations in East Asia as revealed by mtDNA lineages. Mol Biol Evol 22:725–734.

Wen B, Li H, Lu D, Song X, Zhang F, He Y, Li F, Gao Y, Mao X, Zhang L, Qian J, Tan J, Jin J, Huang W, Deka R, Su B, Chakraborty R, Jin L. 2004a. Genetic evidence supports demic diffusion of Han culture. Nature 431:302–305.

Wen B, Xie X, Gao S, Li H, Shi H, Song X, Qian T, Xiao C, Jin J, Su B, Lu D, Chakraborty R, Jin L. 2004b. Analyses of genetic structure of Tibeto-Burman populations reveals sex-biased admixture in southern Tibeto-Burmans. Am J Hum Genet 74:856–865.

Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci 13:555–556.

Yao YG, Kong QP, Bandelt HJ, Kivisild T, Zhang YP. 2002a. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. Am J Hum Genet 70:635–651.

Yao YG, Kong QP, Wang CY, Zhu CL, Zhang YP. 2004. Different matrilineal contributions to genetic structure of ethnic groups in the silk road region in china. Mol Biol Evol 21:2265–2280.

Yao YG, Lu XM, Luo HR, Li WH, Zhang YP. 2000. Gene admixture in the silk road region of China: evidence from mtDNA and melanocortin 1 receptor polymorphism. Genes Genet Syst 75:173–178.

Yao YG, Nie L, Harpending H, Fu YX, Yuan ZG, Zhang YP. 2002b. Genetic relationship of Chinese ethnic populations revealed by mtDNA sequence diversity. Am J Phys Anthropol 118:63–76.

Yao YG, Zhang YP. 2002. Phylogeographic analysis of mtDNA variation in four ethnic populations from Yunnan Province: new data and a reappraisal. J Hum Genet 47:311–318.

Yuan B, Huang W, Zhang D. 2007. New evidence for human occupation of the northern Tibetan Plateau. China during the Late Pleistocene. Chin Sci Bull 52:2675–2679.

Zhang D, Li S. 2002. Optical dating of Tibetan human hand- and footprints: an implication for the palaeoenvironment of the last glaciation of the Tibetan plateau. Geophys Res Lett 29:1609.

Zhang F, Su B, Zhang YP, Jin L. 2007. Genetic studies of human diversity in East Asia. Philos Trans R Soc Lond B Biol Sci 362:987–995.

Zhao M, Kong QP, Wang HW, Peng MS, Xie XD, Wang WZ, Jiayang Duan JG, Cai MC, Zhao SN, Cidanpingcuo, Tu YQ, Wu SF, Yao YG, Bandelt HJ, Zhang YP. 2009. Mitochondrial genome evidence reveals successful Late Paleolithic settlement on the Tibetan Plateau. Proc Natl Acad Sci USA 106:21230–21235.

Zheng Y, Gao S, Wang S, Xue B, Liu H, Zeng X. 2003. Simulations of LGM climate of East Asia by regional climate model. Sci China Ser D Earth Sci 46:753–764.