

Original Paper

Human
HereditHum Hered 2008;66:1–9
DOI: [10.1159/000114160](https://doi.org/10.1159/000114160)Received: March 21, 2007
Accepted after revision: August 27, 2007
Published online: January 28, 2008

Maternal Footprints of Southeast Asians in North India

Kumarasamy Thangaraj^a Gyaneshwer Chaubey^{a, b, n} Toomas Kivisild^{b, c} Deepa Selvi Rani^a
 Vijay Kumar Singh^a Thanseem Ismail^a Denise Carvalho-Silva^d Mait Metspalu^b L.V.K.S. Bhaskar^a
 Alla G. Reddy^a Sarat Chandra^e Veena Pande^f B. Prathap Naidu^a Niharika Adarsh^g
 Abhilasha Verma^h Inaganti Amara Jyothiⁱ Chandana Basu Mallick^j Nidhi Shrivastava^k
 Ragala Devasena^l Babita Kumari^m Amit Kumar Singhⁿ Shailendra Kumar Dhar Dwivediⁿ
 Shefali Singhⁿ Geeta Raoⁿ Pranav Guptaⁿ Vartika Sonvane^o Kavita Kumari^m Afsar Basha^p
 K.R. Bhargaviⁱ Albert Lalremruata^q Arvind Kumar Gupta^j Gurukamal Kaur^r K.K. Reddy^s
 A. Papa Rao^s Richard Villems^b Chris Tyler-Smith^d Lalji Singh^{a*}

^aCentre for Cellular and Molecular Biology, Hyderabad, India; ^bEstonian Biocentre and Tartu University, Tartu, Estonia;
^cLeverhulme Centre of Human Evolutionary Studies, University of Cambridge, Cambridge, ^dThe Wellcome Trust
 Sanger Institute, Hinxton Cambs, UK; ^eNagaland University, Nagaland, ^fKumaun University, Nainital, ^gBIMR College of
 Life Science, Gwalior, ^hBundelkhand University, Jhansi, ⁱJJ College of Arts and Science, Pudukottai, ^jAPS University,
 Rewa, ^kGGD University, Bilaspur, ^lDS College of Arts and Science for women, Perambalur, ^mTM Bhagalpur University,
 Bhagalpur, ⁿVBS Purvanchal University, Jaunpur, ^oBarkatullah University, Bhopal, ^pSt. Josph's PG College, Kurnool,
^qLoyola College, Chennai, ^rGuru Nanak Dev University, ^sS. V. University, Tirupati, India

Key Words

South Asia · 9bp indel · mtDNA · Haplogroup

Abstract

We have analyzed 7,137 samples from 125 different caste, tribal and religious groups of India and 99 samples from three populations of Nepal for the length variation in the COII/tRNA^{Lys} region of mtDNA. Samples showing length variation were subjected to detailed phylogenetic analysis based on HVS-I and informative coding region sequence variation. The overall frequencies of the 9-bp deletion and insertion variants in South Asia were 1.9 and 0.6%, respectively. We have also defined a novel deep-rooting haplogroup M43 and identified the rare haplogroup H14 in Indian populations carrying the 9-bp deletion by complete mtDNA sequencing. Moreover, we redefined haplogroup M6 and dissected it into two well-defined subclades. The presence of haplogroups F1 and B5a in Uttar Pradesh sug-

gests minor maternal contribution from Southeast Asia to Northern India. The occurrence of haplogroup F1 in the Nepalese sample implies that Nepal might have served as a bridge for the flow of eastern lineages to India. The presence of R6 in the Nepalese, on the other hand, suggests that the gene flow between India and Nepal has been reciprocal.

Copyright © 2008 S. Karger AG, Basel

Introduction

Sequence variation of mitochondrial DNA (mtDNA) has been widely studied to assess genetic relatedness at both the species and population levels. In the consensus sequence of the human mtDNA [1] there are two copies

K. Thangaraj, G. Chaubey, T. Kivisild and D. Selvi Rani contributed equally to the paper.

KARGER

Fax +41 61 306 12 34
 E-Mail karger@karger.ch
www.karger.com

© 2008 S. Karger AG, Basel
 0001-5652/08/0661-0001\$24.50/0

Accessible online at:
www.karger.com/hhe

Dr. Lalji Singh/Dr. K. Thangaraj
 Centre for Cellular and Molecular Biology, Uppal Road
 Hyderabad-500 007 (India)
 Tel. +91 40 2716 0789, Fax +91 40 2716 0591
 E-Mail lalji@ccmb.res.in, thangs@ccmb.res.in

Table 1. Frequencies of 9-bp deletion/insertion in tribal and caste populations of India (present study)

State	Populations	Linguistic affiliation	Sample size	9-bp deletion		9-bp insertion	
				number	frequency	number	frequency
Tamil Nadu	Paliyar (T)	Dravidian	71	-	-	-	-
	Pulaiyar (T)	Dravidian	51	-	-	-	-
	Kota (T)	Dravidian	50	-	-	-	-
	Toda (T)	Dravidian	83	-	-	-	-
	Kurumba (T)	Dravidian	94	-	-	-	-
	Irula (T)	Dravidian	25	-	-	-	-
	Malaikuravar (C)	Dravidian	24	-	-	-	-
Andhra Pradesh	Koyas (T)	Dravidian	93	1	0.006	-	-
	Naidu (C)	Dravidian	78	-	-	-	-
	Vysya (T)	Dravidian	70	-	-	-	-
	Thoti (T)	Dravidian	137	-	-	-	-
	Chenchu (T)	Dravidian	243	18	0.074	-	-
	Nayak(C)	Dravidian	50	-	-	-	-
	Jatapu(C)	Dravidian	39	-	-	-	-
	Pattpu (T)	Dravidian	196	-	-	17	0.087
	Mala (C)	Dravidian	44	-	-	-	-
	Lambadi(T)	Dravidian	83	-	-	-	-
	Agnikula Kshatriya (C)	Dravidian	38	-	-	-	-
	Valmiki (C)	Dravidian	83	-	-	-	-
	Telagas (T)	Dravidian	48	-	-	-	-
	Erukula (T)	Dravidian	121	15	0.124	-	-
	Yanadi (T)	Dravidian	96	18	0.188	-	-
	Pradhan (T)	Dravidian	26	-	-	-	-
	Vadera (C)	Dravidian	23	-	-	-	-
	Naiyakpod (C)	Dravidian	130	3	0.023	-	-
Konda (C)	Dravidian	40	-	-	-	-	
Rajasthan	Meena (T)	Indo-European	54	-	-	-	-
	Megwal (T)	Indo-European	36	-	-	-	-
	Bhil (T)	Indo-European	29	-	-	-	-
Madhya Pradesh	Kondar (T)	Indo-European	49	-	-	-	-
	Gond (T)	Dravidian	55	-	-	-	-
	Manjhi (T)	Indo-European	51	-	-	-	-
	Bhilala (T)	Indo-European	50	-	-	-	-
	Kol (T)	Indo-European	88	-	-	-	-
Uttar Pradesh	Tharu (T)	Indo-European	154	3	0.019	2	0.013
	Dushadh (T)	Indo-European	82	4	0.049	-	-
	Habura (T)	Indo-European	89	1	0.011	-	-
	Harijan (C)	Indo-European	50	1	0.020	-	-
	Brahmin (C)	Indo-European	50	-	-	-	-
Punjab	Ramgarhia (C)	Indo-European	112	-	-	-	-
Nagaland	Ao-naga(T)	Tibeto-Burman	67	-	-	5	0.075
	Naga Sema (T)	Tibeto-Burman	53	-	-	-	-
	Chakhes naga (T)	Tibeto-Burman	55	-	-	-	-

Table 1 (continued)

State	Populations	Linguistic affiliation	Sample size	9-bp deletion		9-bp insertion	
				number	frequency	number	frequency
Jharkhand	Oraon (T)	Dravidian	97	–	–	2	0.021
	Munda (T)	Austro-Asiatic	24	–	–	2	0.083
	Ho (T)	Austro-Asiatic	48	–	–	–	–
	Pahariya (T)	Austro-Asiatic	20	2	0.100	–	–
	Asur (T)	Austro-Asiatic	48	1	0.021	–	–
	Cheek-Barik (T)	Indo-European	45	–	–	–	–
	Santhal (T)	Austro-Asiatic	25	–	–	2	0.080
	Lohra (T)	Indo-European	22	–	–	–	–
	Manjhi (T)	Indo-European	51	–	–	–	–
	Murmu (T)	Indo-European	27	–	–	–	–
Chhattisgarh	Kharia (T)	Austro-Asiatic	78	–	–	–	–
	Bahelia (C)	Indo-European	51	–	–	–	–
	Halba (T)	Indo-European	33	–	–	–	–
	Bhunjiya (T)	Indo-European	21	–	–	–	–
Uttaranchal	Buxas (T)	Indo-European	48	–	–	–	–
	Tharu (T)	Indo-European	46	1	0.022	–	–
Total (Present Study)	–		3,844	68	–	30	–
Total (Thangaraj et al. 2005)			3,293	71	0.022	12	0.004
Sum Total			7,137	139	0.019	42	0.005

T = Tribe; C = caste.

of the 9-bp motif cccctcta in the non-coding region V [2–4]. Deletion of one copy has been reported in many populations worldwide at different frequencies [2–18]. Previous studies have demonstrated multiple and independent origins of the 9-bp deletion on different mtDNA haplogroup backgrounds in populations of different continental affiliation such as Australian Aborigines [16] South Indians [18–22] and Sub-Saharan Africans [15]. However, in Central and East Asian populations, the 9-bp deletion is predominantly associated with a single haplogroup, B [23–29].

India is culturally and geographically a highly heterogeneous country. The caste and tribal groups of India are considered socially and culturally the most stratified of all known societies in human history. Indian populations are structured further by their linguistic and religious affiliations. More than 60% of the present day Indian maternal lineages are affiliated with mtDNA haplogroup M. Most of the sub-clades of these Eurasian founder haplogroups are autochthonous to South Asia [for a recent review see 30]. South East Asia is conven-

tionally divided into two cultural, linguistic and geographic regions, i.e. Insular Southeast Asia – the island or peninsular countries of Malaysia, Singapore, Brunei, The Philippines and East Timor; and Mainland Southeast Asia – the countries of Thailand, Laos, Burma, Cambodia and Vietnam. Tibeto-Burman, Austronesian and Austro-Asiatic are the major language families of this area.

Recently, we have shown that populations from the Nicobar Islands that speak the Mon-Khmer branch of the Austro-Asiatic language family carry the highest frequency of the 9-bp deletion thus far reported among South Asian populations [18]. In contrast, we did not detect the 9-bp deletion among 302 Mundari samples, another Austro-Asiatic speaking group from mainland India [18]. Detailed molecular characterization of the Nicobarese samples using coding region SNPs classified them under haplogroup B5a1a [18, 31], which is also common in China and other South East Asian populations [23–29]. In contrast, studies on Mundari populations have not so far detected any lineages with the 9-bp deletion [18, 32,

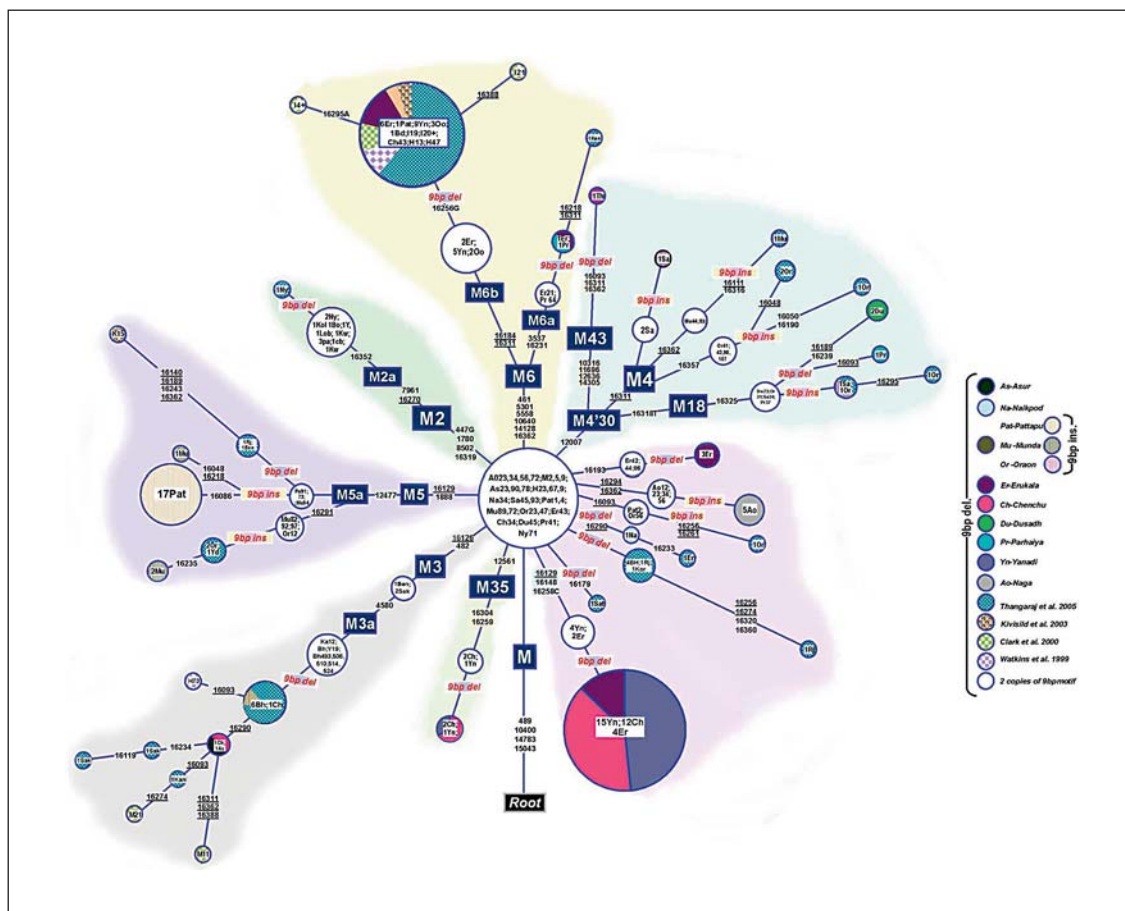


Fig. 1. Phylogeography of macrohaplogroup M, constructed on the basis of coding region SNPs and HVSI sequence information. The white circles represent the control samples (with 2 copies of the 9-bp repeat) (our unpublished data) while the colored circles represent samples with one or three copies of the 9-bp repeat. Solid-colored circles represent new data while circles with different color patches represent samples published elsewhere [18–22, 32]. All the recurrent mutations are underlined.

33]. These findings suggest that the Austro-Asiatic populations of the Indian subcontinent and Nicobar Islands may have distinct genetic sources [18]. Given the limited number of Mundari populations covered in our previous survey, we have now analyzed mtDNA diversity in 3,844 additional individuals from India, including 243 Mundari speakers from six populations not investigated previously, and, for the first time, populations from Nepal, to determine the genetic affinities of the population of mainland India with South East Asian populations, and to infer the origins of the 9-bp deletion/insertion variants in the Indian subcontinent.

Materials and Methods

Approximately 5–10 ml of blood was collected from 7,137 individuals belonging to 125 endogamous populations of India, which includes the 3,293 samples reported in our previous work [18]. In addition, 99 samples from three Nepalese populations were also analyzed. All samples were collected with the written informed consent of the donors. The non-coding intergenic region V between the genes COII and $tRNA^{Lys}$ in the mtDNA was amplified [34] and size-fractionated in 8% polyacrylamide gel electrophoresis. HVSI and Single Nucleotide Polymorphisms (SNPs, from the coding region) were typed by sequencing and RFLP analysis (supplementary table 1, www.karger.com/doi/10.1159/000114160) [35, 36] in the samples carrying either the deletion or insertion of the 9-bp motif. Individuals with ambiguous haplogroup affiliation were subjected to further sequencing of informative coding region stretches. To minimize errors, both

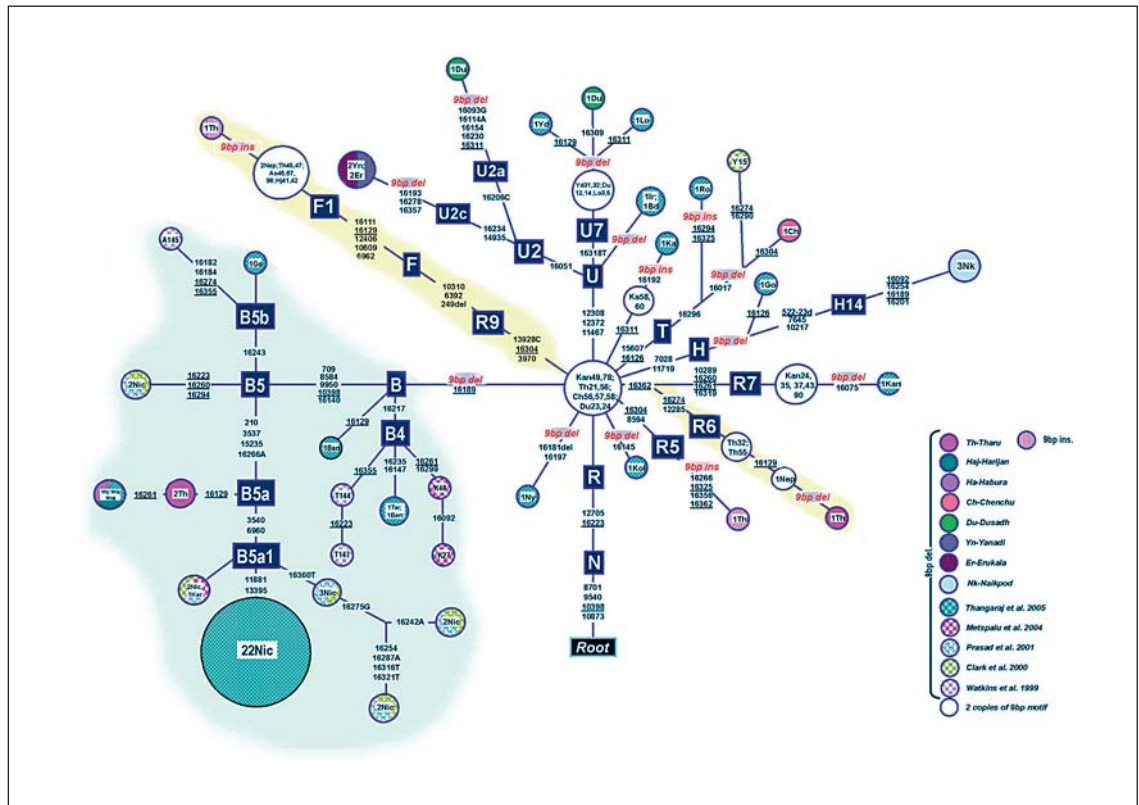


Fig. 2. Phylogeography of macrohaplogroup N constructed on the basis of coding region SNPs and HVS-I sequence information. The white circles represent control samples (with 2 copies of the 9-bp repeat) (our unpublished data). Solid-colored circles represent new data while circles with different color patches represents samples published elsewhere [18–22, 32]. All the recurrent mutations are underlined.

strands were double-sequenced. Phylogenetic relationships between the observed haplotypes were reconstructed with the NETWORK program (version 4.1) [37] [www.fluxus-engineering.com].

Results and Discussion

Of the 7,137 individuals studied (present data and previous work) 139 individuals had the 9-bp deletion (1.94%) whereas 42 individuals had the insertion (0.59%) (supplementary table 1). The highest frequency (~20%) of the 9-bp deletion was observed in the Yanadi, a Dravidian-speaking population from southern districts of Andhra Pradesh. Nine populations showed the presence of the 9-bp insertion and the highest frequency (8.7%) was also observed in a Dravidian-speaking population, from Andhra Pradesh (table 1, supplementary table 1).

Both Mishmar et al. [38] and Ruiz-Pesini et al. [39] have postulated a role for natural selection in shaping the mtDNA distribution at a global level. They classified the South and South East Asian haplogroup B, the African haplogroup L0a [15] and the Australian variants [16] as ‘tropical’ haplogroups, which are associated with the 9-bp deletion. In contrast, none of the haplogroups found in Europe or Northeast Asia had high frequencies of the 9-bp deletion and were not classified as ‘tropical’ [38, 39]. Surprisingly, the distribution appears different in the Indian sub-continent, where the frequency of the 9-bp deletion is generally low, except in Nicobar Islanders [18–22, 33; this study].

Given the generally low frequency we find in our large South Asian sample, it seems unlikely that the 9-bp deletion is advantageous in tropical regions. Its high frequency in certain populations is better explained by random genetic drift (e.g. a founder effect involving haplogroup

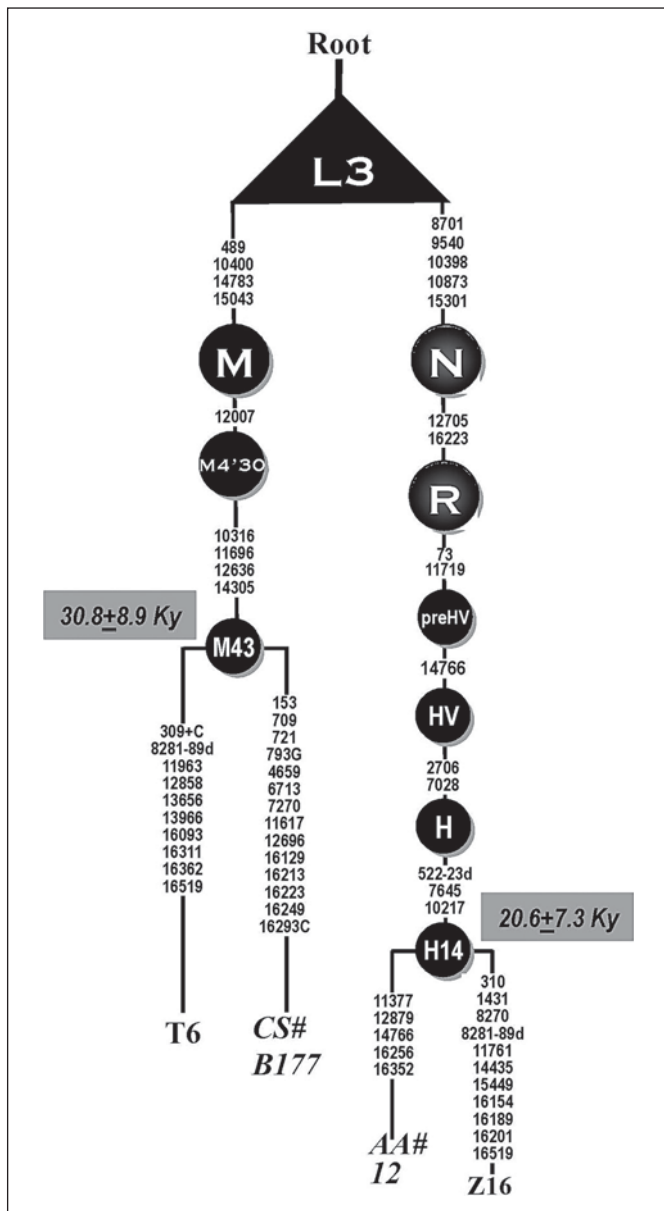


Fig. 3. The most parsimonious tree for newly identified deep-rooting haplogroup M43 and a rare haplogroup H14 [44] based on complete mtDNA sequencing. Mutations are scored after comparing with r-CRS [1]. Suffixes are transversions. Samples taken from published sources [43, 44] are indicated by the name of the first author. T and Z refer Tharu and Naikpod populations respectively.

B5a1 in the Nicobarese) resulting in different frequencies of different lineages in different regions, where some of the lineages harbor phylogenetically deep deletions that render them more frequent. In contrast to the 9-bp deletion, the insertion polymorphism has not reached high

frequency in any of the populations of our present study. Nevertheless, Thangaraj et al. [18] reported that the insertion, like the deletion, has occurred repeatedly on multiple mtDNA lineages. While most of the insertion/deletion events examined in this study are autochthonous to India, i.e. they are found on mtDNA lineages specific to India, a few can still be attested as having evolved elsewhere and only recently been carried into India. Besides the Nicobarese B5a, several other sub-clades of haplogroup B are found in India at detectable but overall low frequency (fig. 1). In India, the insertion is rarer than the deletion (supplementary table 1, fig. 2). We did not find any individual with four or more copies of the 9-bp motif, which may suggest that more copies have a lower fitness or that they are structurally unstable. The frequency distribution of mtDNA haplogroups associated with the 9-bp indel in mainland India is shown in figure 2.

By additional genotyping of informative SNPs in the coding region in the samples reported in our present and previous studies [18], we classified the samples with the 9-bp ins/del polymorphism according to known mtDNA haplogroups (supplementary table 1). Our results suggest that there have been at least 21 independent occurrences of the 9-bp ins/del polymorphism in macrohaplogroup M (fig. 3) and 15 independent incidents in macrohaplogroup R background (fig. 1). Our survey of published complete mtDNA sequence data of European populations shows a 3% (16/527) 9-bp deletion frequency [40–42] on the background of multiple European-specific haplogroups which is by and large consistent with our findings for India. This comparison further strengthens the view that the elevated frequency of the 9-bp deletion in some previously reported populations from the tropics can be explained by limited sampling and the effect of genetic drift rather than requiring explanations involving adaptive selection to the tropical environment.

We identified here a novel deep-rooting mtDNA clade M43, based on the complete sequencing of a 9-bp-deleted Tharu individual from Uttaranchal state of India (fig. 4). This sample shared four coding region mutations with individual B177 of Sun et al. [43]. We redefined haplogroup M6 by five coding and a single control region substitution. The fine dissection of this haplogroup suggests two well-defined subclades M6a and M6b. Subclade, M6b occupies one branch among the two major founder clades of South Indian 9-bp-deleted samples; the another major clad is unclassified M* (fig. 3). Moreover, for the first time we have identified the rare West Eurasia-specific haplogroup H14 [44, 45] in Indian population (fig. 4). The complete sequencing of the Indian branch

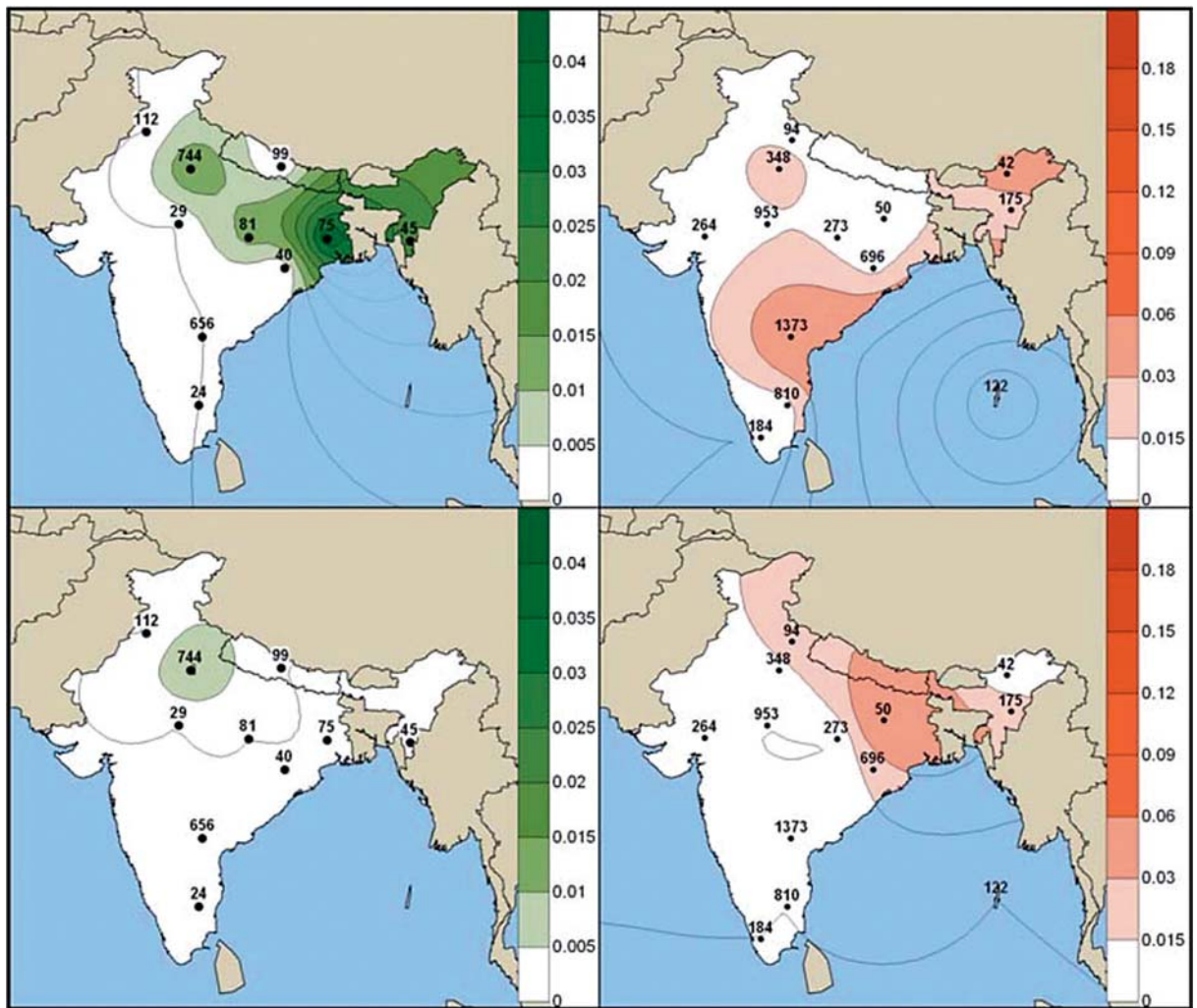


Fig. 4. The 9-bp ins/del spatial distribution in Indian populations. The upper and lower panels show the distribution of 9-bp deletion and insertion variants, respectively. The green and orange colors depict the distribution in caste and tribal populations, respectively. Numbers next to data point indicate sample sizes.

suggests a deep split between the Indian and European lineages ~20 KYA.

After screening a total of 545 Mundari-speaking individuals from mainland India, including 302 samples from our previous study, we found that the 9-bp deletion polymorphism was observed in 0.5% of the analyzed samples, in contrast to our previous study [18] where no samples with the deletion were detected and only a single individual with the 9bp-insertion was found. However, while all Nicobarese with the 9-bp deletion belonged to haplogroup B5a, the new Mundari samples with the 9-bp deletion polymorphism were found in different backgrounds within the macrohaplogroup M. These findings

suggest independent occurrences of the 9-bp deletion event even among the Mundari population. Accordingly, these results support our previous hypothesis that the Mundari (an Austro-Asiatic speaking population) from mainland India did have an independent migration/origin compared with the Mon-Khmeric Austro-Asiatic speaking groups from the Nicobar Islands [18, 31]. In addition to that, the 262 samples (220 from this work (table 1) and 42 Nishi from our previous study) from Tibeto-Burman speaking groups of North Eastern India did not show the presence of haplogroup B either.

The presence of haplogroups B5a and F1 in the Tharu of Eastern Uttar Pradesh (near the Indo-Nepal border)

points to past human movements from East Asia to India, perhaps through Nepal. In order to investigate this possibility further, we sequenced the HVS-1 in 99 samples from Nepal and searched for the 9-bp polymorphism. We did not find any Nepali sample either belonging to haplogroup B5a or carrying the 9-bp deletion/insertion polymorphism. However, we detected haplogroups R6 and F1, which suggest reciprocal gene flow. The presence of haplogroup F1 in the Nepalese implies that Nepal could have served as a bridge for migrations from East Asia to India. On the other hand, haplogroup R6 in Nepal suggests gene flow from India to Nepal.

In conclusion, we report a novel autochthonous deep-rooting M43 lineage, redefine haplogroup M6 and identify a rare West Eurasia-specific haplogroup H14 in Indian population (fig. 4). Based on the co-presence of haplogroup F1 in North Indian and Nepali populations, but the lack of B5a in the latter, it seems likely that the migration of Southeast Asian maternal lineages to North India has occurred not only through the corridor of Nepal but also has involved gene flow along the southern slopes of

the Himalayas within the Indian sub-continent. The occurrence of R6 in the Nepalese suggests gene flow from India to Nepal. Further analyses of the B5a and F1 lineages in larger samples from Nepal and Eastern Uttar Pradesh should shed further light on the origin and spread of the Tharu and other populations of North India. Studies of mtDNA and Y-chromosomal markers in Nepalese populations would be useful for the investigation of this putative gene flow into India.

Acknowledgements

We are gratefully acknowledged Council of Scientific and Industrial Research, Govt. of India for their financial support to L.S. and K.T. for analyzing the Indian samples. We are thankful to all anonymous blood donors. Nepalese samples were collected as part of the European Science Foundation EUROCORES Programme OMLL, which was supported by funds from the Arts and Humanities Research Board and the EC Sixth Framework Programme under Contract no. ERAS-CT-2003-980409. C.T.-S. was supported by The Wellcome Trust.

References

- Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N: Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 1999;23:147.
- Horai S, Hayasaka K: Evolutionary implications of mitochondrial DNA polymorphisms in human populations; in Vogel F, Sperling K (eds): *Human Genetics. Proceedings of the 7th International Congress.* Berlin, Springer, 1987, pp 177–181.
- Hertzberg M, Mickleson KN, Serjeantson SW, Prior JF, Trent RJ: An Asian-specific 9-bp deletion of mitochondrial DNA is frequently found in Polynesians. *Am J Hum Genet* 1989;44:504–510.
- Stoneking M, Wilson AC: Mitochondrial DNA; in Hill A, Serjeantson S (eds): *The Colonization of Pacific: A Genetic Trial.* Oxford University Press Oxford England, 1989, pp 215–245.
- Schurr TG, Ballinger SW, Gan YY, Hodge JA, Merriwether DA, Lawrence DN, et al: Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies suggesting they derived from four primary maternal lineages. *Am J Hum Genet* 1990;46:613–623.
- Ward RH, Frazier BL, Dew-Jager K, Paabo S: Extensive mitochondrial diversity within a single Amerindian tribe. *Proc Natl Acad Sci USA* 1991;88:8720–8724.
- Ward RH, Redd A, Valencia D, Frazier B, Paabo S: Genetic and linguistic differentiation in the Americas. *Proc Natl Acad Sci USA* 1993;90:10663–10667.
- Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC: African populations and the evolution of human mitochondrial DNA. *Science* 1991;253:1503–1507.
- Ballinger SW, Schurr TG, Torroni A, Gan YY, Hodge JA, Hassan K, Chen K-H, Wallace DC: Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. *Genetics* 1992;130:139–152.
- Harihara S, Hirai M, Suutou Y, Shimizu K, Omoto K: Frequency of a 9-bp deletion in the mitochondrial DNA among Asian populations. *Hum Biol* 1992;64:161–166.
- Torroni A, Schurr TG, Yang C, Szathmary EJ, Williams RC, Schanfield MS, et al: Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. *Genetics* 1992;130:153–162.
- Hagelberg E, Clegg JB: Genetic polymorphisms in prehistoric Pacific Islanders determined by analysis of ancient bone DNA. *Proc R Soc Lond B Biol Sci* 1993;252:163–170.
- Hagelberg E, Quevedo S, Turbon D, Clegg JB: DNA from ancient Easter Islanders. *Nature* 1994;369:25–26.
- Redd AJ, Takezaki N, Sherry ST, McGarvey ST, Sofro AS, Stoneking M: Evolutionary history of the COII/tRNA^{Lys} intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. *Mol Biol Evol* 1995;12:604–615.
- Soodyall H, Vigilant L, Hill AV, Stoneking M, Jenkins T: MtDNA control-region sequence variation suggests multiple independent origins of an 'Asian-specific' 9-bp deletion in sub-Saharan Africans. *Am J Hum Genet* 1996;58:595–608.
- Betty DJ, Chin-Atkins AN, Croft L, Sraml M, Eastale S: Multiple independent origins of the COII/tRNA [Lys] intergenic 9-bp mtDNA deletion in aboriginal Australians. *Am J Hum Genet* 1996;58:428–433.
- Yao YG, Watkins WS, Zhang YP: Evolutionary history of the mtDNA 9-bp deletion in Chinese populations and its relevance to the peopling of east and southeast Asia. *Hum Genet* 2000;107:504–512.
- Thangaraj K, Sridhar V, Kivisild T, Reddy AG, Chaubey G, et al: Different population histories of the Mundari- and Mon-Khmer-speaking Austro-Asiatic tribes inferred from the mtDNA 9-bp deletion/insertion polymorphism in Indian populations. *Hum Genet* 2005a;116:507–517.

- 19 Watkins WS, Bamshad M, Dixon ME, Rao BB, Naidu JM, et al: Multiple origins of the mtDNA 9-bp deletion in populations of South India. *Am J Phys Anthropol* 1999;109:147–158.
- 20 Clark VJ, Sivendren S, Saha N, Bentley GR, Auger R, et al: The 9-bp deletion between the mitochondrial lysine tRNA and COII genes in tribal populations of India. *Hum Biol* 2000;72:273–285.
- 21 Prasad BV, Ricker CE, Watkins WS, Dixon ME, Rao BB, et al: Mitochondrial DNA variation in Nicobarese Islanders. *Hum Biol* 2001;73:715–725.
- 22 Kivisild T, Rootsi S, Metspalu M, Mastana S, Kaldma K, Parik J, et al: The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. *Am J Hum Genet* 2003;72:313–332.
- 23 Nishimaki Y, Sato K, Fang L, Ma M, Hasekura H, Boettcher B: Sequence polymorphism in the mtDNA HVS-I region in Japanese and Chinese. *Leg Med* 1999;1:238–249.
- 24 Yao YG, Kong QP, Bandelt H-J, Kivisild T, Zhang YP: Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *Am J Hum Genet* 2002;70:635–651.
- 25 Yao YG, Nie L, Harpending H, Fu YX, Yuan ZG, Zhang YP: Genetic relationship of Chinese ethnic populations revealed by mtDNA sequence diversity. *Am J Phys Anthropol* 2002;118:63–76.
- 26 Fucharoen G, Fucharoen S, Horai S: Mitochondrial DNA polymorphisms in Thailand. *J Hum Genet* 2001;46:115–125.
- 27 Tsai LC, Lin CY, Lee JC, Chang JG, Linacre A, Goodwin W: Sequence polymorphism of mitochondrial D-loop DNA in the Taiwanese Han population. *Forensic Sci Int* 2001;119:239–247.
- 28 Kong QP, Yao YG, Sun C, Bandelt H-J, Zhu CL, Zhang YP: Phylogeny of East Asian Mitochondrial DNA Lineages Inferred from Complete Sequences. *Am J Hum Genet* 2003;73:671–676.
- 29 Kong QP, Yao YG, Liu M, Shen SP, Chen C, et al: Mitochondrial DNA sequence polymorphisms of five ethnic populations from northern China. *Hum Genet* 2003;113:391–405.
- 30 Chaubey G, Metspalu M, Kivisild T, Villems R: Peopling of South Asia: investigating the caste-tribe continuum in India. *Bioessays* 2007;29:91–100.
- 31 Thangaraj K, Chaubey G, Kivisild T, Reddy AG, Singh VK, Rasalkar AA, Singh L: Reconstructing the origin of Andaman Islanders. *Science* 2005;13:996.
- 32 Roychoudhury S, Roy S, Basu A, Banerjee R, Vishwanathan H, Usha Rani MV, Sil SK, Mitra M, Majumder PP: Genomic structures and population histories of linguistically distinct tribal groups of India. *Hum Genet* 2001;109:339–350.
- 33 Metspalu M, Kivisild T, Metspalu E, Parik J, Hudjashov G, et al: Most of the extant mtDNA boundaries in South and Southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BMC Genet* 2004;5:26.
- 34 Thomas MG, Cook CE, Miller KW, Waring MJ, Hagelberg E: 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. *Phil Trans R Soc Lond B Biol Sci* 1998;353:955–965.
- 35 Rieder MJ, Taylor SL, Tobe VO, Nickerson DA: Automating the identification of DNA variations using quality-based fluorescence re-sequencing: analysis of the human mitochondrial genome. *Nucleic Acids Res* 1998;26:967–973.
- 36 Quintana-Murci L, Semino O, Bandelt H-J, Passarino G, McElreavey K, Santachiara-Benerecetti AS: Genetic evidence of an early exit of Homo sapiens sapiens from Africa through eastern Africa. *Nat Genet* 1999;23:437–441.
- 37 Bandelt HJ, Forster P, Rohl A: Median joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 1999;16:37–48.
- 38 Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, et al: Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci USA* 2003;100:171.
- 39 Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC: Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* 2004;303:223–226.
- 40 Finnila S, Lehtonen MS, Majamaa K: Phylogenetic network for European mtDNA. *Am J Hum Genet* 2001;68:1475–1484.
- 41 Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, Olefsky JM, Beal MF, Davis RE, Howell N: Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am J Hum Genet* 2002;70:1152–1171.
- 42 Kivisild T, Shen P, Wall DP, Do B, Sung R, Davis K, Passarino G, Underhill PA, Scharfe C, Torroni A, Scozzari R, Modiano D, Coppa A, de Knijff P, Feldman M, Cavalli-Sforza LL, Oefner PJ: The role of selection in the evolution of human mitochondrial genomes. *Genetics* 2006;172:373–387.
- 43 Sun C, Kong Q-P, Palanichamy MG, Agrawal S, Bandelt H-J, Yao Y-G, Khan F, Zhu C-L, Chaudhuri TK, Zhang Y-P: The dazzling array of basal branches in the mtDNA macrohaplogroup M from India as inferred from complete genomes. *Mol Biol Evol* 2006;3:683–690.
- 44 Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, Cruciani F, Zeviani M, Briem E, Carelli V, et al: The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *Am J Hum Genet* 2004;75:910–918.
- 45 Roostalu U, Kutuev I, Loogväli EL, Metspalu E, Tambets K, et al: Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. *Mol Biol Evol* 2007;24:436–448.