

# Global Patterns in Human Mitochondrial DNA and Y-Chromosome Variation Caused by Spatial Instability of the Local Cultural Processes

Vikrant Kumar<sup>1</sup>, Banrida T. Langstieh<sup>1,2</sup>, Komal V. Madhavi<sup>1</sup>, Vegi M. Naidu<sup>1,3</sup>, Hardeep Pal Singh<sup>1,4</sup>, Silpak Biswas<sup>1</sup>, Kumarasamy Thangaraj<sup>5</sup>, Lalji Singh<sup>5</sup>, B. Mohan Reddy<sup>1\*</sup>

**1** Biological Anthropology Unit, Indian Statistical Institute, Hubsiguda, Hyderabad, India, **2** Department of Anthropology, North Eastern Hill University, Shillong, India, **3** GSF Hematologikum, Medizinische Klinik III, Klinikum Grosshadern, München, Germany, **4** Kallam Anji Reddy Molecular Genetics Laboratory, L.V. Prasad Eye Institute, L.V. Prasad Marg, Banjara Hills, India, **5** Centre for Cellular and Molecular Biology, Hyderabad, India

**Because of the widespread phenomenon of patrilocality, it is hypothesized that Y-chromosome variants tend to be more localized geographically than those of mitochondrial DNA (mtDNA). Empirical evidence confirmatory to this hypothesis was subsequently provided among certain patrilocal and matrilineal groups of Thailand, which conforms to the isolation by distance mode of gene diffusion. However, we expect intuitively that the patterns of genetic variability may not be consistent with the above hypothesis among populations with different social norms governing the institution of marriage, particularly among those that adhere to strict endogamy rules. We test the universality of this hypothesis by analyzing Y-chromosome and mtDNA data in three different sets of Indian populations that follow endogamy rules to varying degrees. Our analysis of the Indian patrilocal and the matrilineal groups is not confirmatory to the sex-specific variation observed among the tribes of Thailand. Our results indicate spatial instability of the impact of different cultural processes on the genetic variability, resulting in the lack of universality of the hypothesized pattern of greater Y-chromosome variation when compared to that of mtDNA among the patrilocal populations.**

Citation: Kumar V, Langstieh BT, Madhavi KV, Naidu VM, Singh HP, et al. (2006) Global patterns in human mitochondrial DNA and Y-chromosome variation caused by spatial instability of the local cultural processes. *PLoS Genet* 2(4): e53. DOI: 10.1371/journal.pgen.0020053

## Introduction

The genetic patterns in human societies are often fashioned by their cultural practices. For example, it has been hypothesized that due to widespread phenomenon of patrilocality (a pattern of residence where the female spouse after marriage resides in the in-law's house) Y-chromosome variants tend to be more localized geographically than those of mitochondrial DNA (mtDNA) and the autosomes, and therefore high degree of inter-population genetic differences have been observed for the Y chromosome compared to the mtDNA [1–4]. Due to movement of females in patrilocal groups, the mtDNA diversity is assumed to be high within the populations and low between the populations, whereas the Y-chromosome diversity will be relatively low within the groups and high between the groups. This pattern is expected to be reversed in case of the matrilineal groups (a pattern of residence where the males after marriage reside in the in-law's house). Empirical evidence confirmatory to this hypothesis was subsequently provided by Oota et al. [5] among the three patrilocal and three matrilineal groups of Thailand. They found genetic diversity to be strikingly correlated with residence patterns suggesting the role of sex-specific patterns of migration in influencing the genetic patterns. In contrast, few other studies at the regional scale [6–8] show similar levels of differentiation for maternal and paternal lineages. Therefore, the patterns of genetic diversity at the local level may not reflect at the global scale, which is essentially an artifact of the sum total of differing local patterns. Concurrently, in a global survey, Wilder et al. [9] could not detect the signature of a

higher inter-population migration rate for females than for males. This is interpreted as due to lack of geographic stability of the behavioral customs of individual populations necessary to influence global genetic patterning. The norms governing the institution of marriage vary enormously among human populations of different regions or cultures [10–12], and different forms of social organization can impact patterns and levels of genetic diversity [13,14]. Therefore, the universality of the above hypothesis, i.e., the pattern of genetic variation vis-à-vis the residence pattern of spouses, is in question.

Implicit in the above hypothesis is the assumption that the population boundaries are permeable, permitting male/female spouses to move across their respective populations and become part of the gene pool of the new population to which the other spouse belongs. Only in such a scenario can the expectations of the above hypothesis hold, either in patrilocal or matrilineal societies. This situation, broadly

**Editor:** Chris Tyler-Smith, Wellcome Trust Sanger Institute, United Kingdom

**Received** October 3, 2005; **Accepted** November 23, 2005; **Published** April 14, 2006

A previous version of this article appeared as an Early Online Release on February 23, 2006 (DOI: 10.1371/journal.pgen.0020053.eor).

**DOI:** 10.1371/journal.pgen.0020053

**Copyright:** © 2006 Kumar et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Abbreviations:** HVS1, hyper variable segment 1; mtDNA, mitochondrial DNA; Y-STR, Y-chromosome short tandem repeat

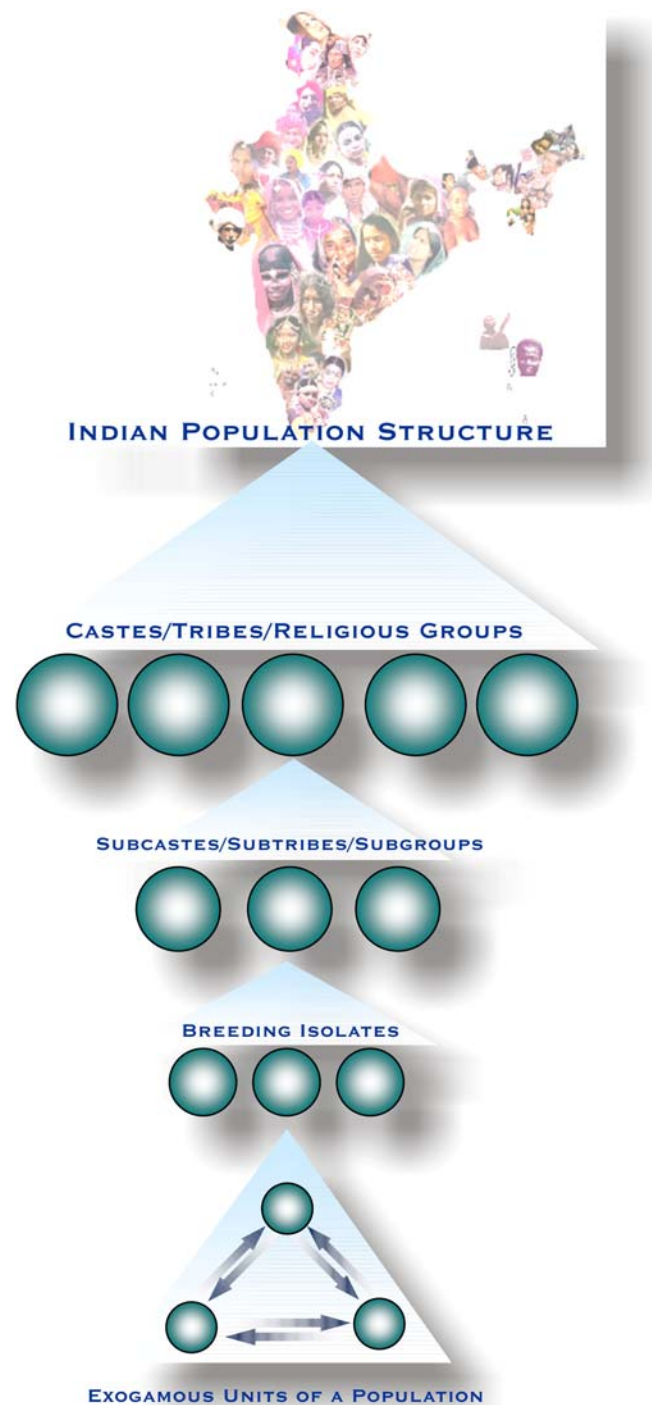
\* To whom correspondence should be addressed. E-mail: bmr@isical.ac.in

## Synopsis

In most human societies, women traditionally move to their husband's home after marriage, and these societies are thus "patrilocal," but in a few "matrilocal" societies, men move to their wife's home. These social customs are expected to influence the patterns of genetic variation. They should lead to a localization of male-specific Y-chromosomal variants and wide dispersal of female-specific mitochondrial DNA variants in patrilocal societies and vice versa in matrilocal societies. These predicted patterns have indeed been observed in previous studies of populations from Thailand. Indian societies, however, are endogamous, so marriage should always take place within a population, and these different patterns of genetic variation should not build up. The authors have now analyzed ten patrilocal and five matrilocal Indian populations, and find that there is indeed little difference between the patrilocal and matrilocal societies. The authors therefore conclude that patterns of genetic variation in humans are not universal, but depend on local cultural practices.

speaking, approximates to isolation by distance mode of gene diffusion. On the other hand, for populations bound by rigid endogamy rules with their boundaries absolutely impermeable, neither patrilocality nor matrilocality can make any difference to their genetic variability, be it Y-chromosome or mtDNA, since the movement is restricted to within a population. The Indian subcontinent with its unique population structure and strictly defined endogamous castes, tribes, and religious groups is a case in point (Figure 1). The marriage interactions are restricted within an endogamous population consisting of the number of exogamous units/clans between which marriages take place. We directly test the universality of the hypothesis delineated above and attempt to assess the spatial stability of the local cultural processes necessary to influence global patterning in two stages. In the first stage, we analyzed Y-chromosome short tandem repeat (Y-STR) and mtDNA hyper variable segment 1 (HVS1) sequence data from two groups of Indian tribes, comprised of five populations each, belonging to a broad linguistic family and with similar socio-economic status. The genetic data were obtained from the same set of populations and individuals making it appropriate for comparison. The populations included in this study are Maram, Khyntiam, Pnar, Bhoi, and WarKhasi, the five matrilocal Khasi tribes of Meghalaya in the Northeastern part of India; and Asur, Bhumij, Kharia, Munda, and Santhal, the five patrilocal Mundari tribes of Eastern India, who along with the matrilocal Khasis, belong to the broad Austro-Asiatic linguistic family. At the second stage, to gauge the consistency in the genetic patterns within broad regional or cultural context, the same set of genetic data were generated on the five Dravidian language-speaking patrilocal caste populations from Andhra Pradesh (Akutota, Kapu, Panta, Pokanati, and Vanne) of Southern India and compared with the Austro-Asiatic matrilocal tribes.

The structure of populations considered in this study is characterized by numerous endogamous groups cohabiting as islands with no or negligible gene flow between them. Therefore, as the marital boundary of each population is impermeable, we intuitively expect that the pattern of genetic variability may not strictly follow the expectations of the aforesaid hypothesis, either in patrilocal or matrilocal

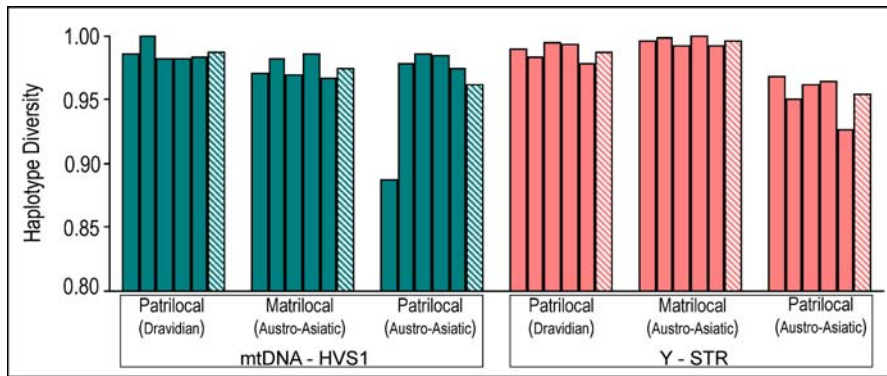


**Figure 1.** Schematic Representation of Indian Population Structure Characterized by Movement of Spouses Only within but Not among the Endogamous Groups

Each circle represents a population and its size represents the hierarchy. While the populations until the breeding isolates are all endogamous, the exogamous units refer to clans/lineages within a breeding isolate/population.

DOI: 10.1371/journal.pgen.0020053.g001

groups. All three groups of populations have contiguous geographic distribution in their respective areas, which provide opportunity for exchange of mates, if the social norms permit, thus providing ideal study frame.



**Figure 2.** Haplotype Diversity in mtDNA (Green) and Y-STR (Pink) and Their Mean (Shaded Bar) in Five Dravidian and Five Austro-Asiatic Patrilocal and Five Austro-Asiatic Matrilocal Populations

From left to right, the Dravidian patrilocal groups (mtDNA sample size and Y-STR sample size) are Akhutota (32, 21), Kapu (22,16), Panta (37, 21), Pokanati (59, 25), and Vanne (32, 23); the Austro-Asiatic matrilocal groups are Maram (72, 58), Khyntiam (95, 82), Pnar (69, 40), Bhoi (34, 30), and Warkhasi (31, 23); the Austro-Asiatic patrilocal groups are Asur (30, 28), Bhumij (40, 39), Kharia (21, 13), Munda (23, 23), and Santhal (39, 38).

DOI: 10.1371/journal.pgen.0020053.g002

## Results

Within-group mtDNA diversity (Figure 2) is similar (Mann-Whitney U test,  $p = 0.690$ ) for matrilocal Khasi tribes (0.975) versus patrilocal Mundari tribes (0.962), although the mean within-group Y-chromosome diversity of patrilocal Mundari tribal groups (0.954) is significantly lower (Mann-Whitney U test,  $p = 0.008$ ) when compared with matrilocal Khasi (0.995). However, when we compare the patrilocal Dravidian caste groups with the matrilocal Khasi tribes we found similar and non-significant difference in the level of within-group diversity for both mtDNA ( $p = 0.056$ ) and Y-chromosome ( $p = .095$ ). The average values of genetic distance (Table 1) reflecting inter-group diversity (although smaller for mtDNA and larger for Y-chromosome among patrilocal Mundari groups than for matrilocal Khasi groups) are not statistically significantly different. Likewise, the average genetic distances in the Dravidian patrilocal groups are smaller for mtDNA and larger for Y-chromosome but not significantly so when compared with the matrilocal Khasi groups.

The index of probability of identity, which gives a quantitative measure of haplotype sharing between a pair of populations, further suggests, as against the hypothesis, that

the degree of Y-chromosome haplotype sharing (Table 2), although not significant, is substantially higher among the patrilocal Mundari groups when compared with the Matrilocal Khasi tribes, whereas the degree of mtDNA haplotype sharing is almost identical for both groups. On the other hand, we observe a very low level of mtDNA haplotype sharing among the patrilocal Dravidian groups compared with the matrilocal Khasi groups, while the level of Y-chromosome haplotype sharing is similar for both the groups. As per the hypothesis, a relatively lower degree of mtDNA haplotype sharing and greater degree of Y-chromosome haplotype sharing is expected among the matrilocal groups compared with the patrilocal groups. Overall, the results are not consistent with the universality of the hypothesis in question.

## Discussion

The foregoing analysis of the results does not reflect higher migration rate of females and males, respectively, in the patrilocal and matrilocal populations, suggesting that the pattern of residence of the spouses has no bearing on the mtDNA and Y-chromosome variability in the populations, in which sex-specific migrations implicit in the hypothesis are

**Table 1.** Average Genetic Distance and Their Standard Error Based on mtDNA HVS1 and Y-STR among the Matrilocal and Patrilocal Groups

Genetic Distances	Patrilocal (Austro-Asiatic; Mundari) Average $\pm$ SE	Matrilocal (Austro-Asiatic; Khasi) Average $\pm$ SE	Patrilocal (Dravidian) Average $\pm$ SE	Mann-Whitney U Test (p: Two-Tailed <sup>a</sup> )	
				Mundari versus Khasi	Dravidian versus Khasi
$R_{st}$ (Y-STR)	0.100 $\pm$ 0.002	0.055 $\pm$ 0.001	0.114 $\pm$ 0.003	0.123	0.474
$D_A$ (mtDNA)	0.128 $\pm$ 0.103	0.200 $\pm$ 0.095	0.142 $\pm$ 0.087	0.165	0.971

Genetic distances ( $d_A$  and  $R_{st}$ ) and SE, based on 1,000 bootstrap replicates, were calculated using MEGA (<http://www.megasoftware.net/mega3/mega.html>) and RSTCALC (<http://helios.bto.ed.ac.uk/evolgen/rst/rst.html>).

<sup>a</sup>Calculated on the basis of genetic distance matrices.

SE, standard error.

DOI: 10.1371/journal.pgen.0020053.t001

**Table 2.** Index of Probability of Identity Based on mtDNA HVS1 and Y-STR among the Patrilocal and Matrilocal Groups

Genetic Markers	Patrilocal (Austro-Asiatic; Mundari) Average	Matrilocal (Austro-Asiatic; Khasi) Average	Patrilocal (Dravidian) Average	Mann-Whitney U Test ( $p$ : Two-Tailed)	
				Mundari versus Khasi	Dravidian versus Khasi
Y-STR	0.0116	0.0023	0.0011	0.260	0.029
mtDNA	0.0132	0.0117	0.0031	0.626	0.045

DOI: 10.1371/journal.pgen.0020053.t002

confined within the endogamous groups and do not usually transect the caste/tribal boundaries. However, a weak and non-significant trend of greater inter-group variation in Y-chromosome and lower variation in mtDNA in case of patrilocal groups, and greater mtDNA and lower Y-chromosome inter-group variation in matrilocal groups, which is consistent with the hypothesis, is observed. Nevertheless, the magnitude of differences, either intra- or inter-population observed in our study, are substantially smaller than what has been observed by Oota et al. [5] in Thailand, despite a relatively small number of samples and populations. The non-significant differences in the mean values of the genetic distances could have been due to two reasons: (1) either to small sample size; hence lacking sufficient power to correctly reject the null hypothesis, or (2) to small number of Y-STRs, which may not have adequate resolution. Therefore, we calculated power of the Mann-Whitney U test for the given sample sizes in the study and the results suggest that the test has  $> 99\%$  power, even at  $\alpha = 0.001$ , both for mtDNA and Y-chromosome. Additional analysis based on 15 Y-STRs suggests, contrary to the hypothesis, that the average genetic distance among the patrilocal groups was quite low ( $0.0469 \pm 0.0009$ ), albeit non-significantly ( $p = 0.1$ ), as compared with the matrilocal groups ( $0.1024 \pm 0.0024$ ). Therefore, the hypothesized correlation of genetic diversity with the sex-specific migration patterns may not be applicable to the Indian situation, although it is observed elsewhere in certain populations whose marital boundaries are probably permeable.

One of the questions raised by Wilder et al. [9] is the extent to which local cultural practices influence genetic patterns at the regional and global scale. The groups we have considered in the present study have different cultural norms governing the rules of marriages compared to those studied by Oota et al. [5]; hence we find variation in the genetic patterns. Even within India, we find variation in the pattern depending on whether we compare the matrilocal Khasi tribes with the patrilocal Mundari tribal groups or with the patrilocal Dravidian caste groups. For example, the index of probability of identity shows very low values for both mtDNA and Y-chromosome haplotype sharing among the Dravidian castes when compared with the Austro-Asiatic tribes, either Mundari or Khasi (Table 2). This pattern is observed because the caste populations of India are considered to follow endogamy very strictly; hence their marital boundaries are highly rigid compared with the marital boundaries of the Indian tribes, particularly from Northeast India, suggesting the impact of varying cultural practices pertaining partic-

ularly to marriage, resulting in variable genetic patterns. Results of our study taken together with the previous studies, that have [6–8] or have not [1,3–5] detected sex-specific migration, suggest that the local cultural processes do not have spatial stability required to influence global patterning. Perhaps due to this, Wilder et al. [9] did not observe higher migration rates of females vis-à-vis males at the continental level, although most of the populations of the world follow patrilocality [15]. Therefore, the hypothesis of greater Y-chromosome vis-à-vis mtDNA variability due to patrilocality is not universal, as it can only be selectively applicable to populations with cultural norms that permit inter-group marriages; not to, for example, highly endogamous Indian populations. Pertinent to this are the two recent large-scale Indian studies [16,17] wherein the lack of spatial structure in the quantitative biological variables—anthropometry and dermatoglyphics— and traditional genetic markers was inferred to be consistent with population structure characterized by numerous endogamous groups cohabiting as islands with no or negligible gene flow between them; the monotonic decline in the spatial autocorrelation expected under the model of contiguous diffusion of genes is not evident in those data.

## Materials and Methods

Blood samples from 636 individuals belonging to 15 populations were obtained for the above populations during 2000–2003 with informed written consent; DNA was extracted. The names of the populations along with their sample size are given in Figure 2. We analyzed 350 base pairs of the HVS1 of the mtDNA control region corresponding to positions 16050–16400 and six Y-STR loci (DYS19, DYS389I, DYS389b, DYS390, DYS391, and DYS393). Allele length for DYS389b was obtained by subtracting the allele length of DYS389I from DYS389II. The HVS1 sequences have been submitted to GenBank and are also available from the authors, as are the Y-STR data. To measure within-group variability we estimated haplotype diversity [18] for the HVS1 sequences and Y-STR haplotypes (Table S1), and calculated  $d_A$  distances [19] for the HVS1 sequences using the number of different sites model, and  $R_{ST}$  for the Y-STR haplotypes [20] as measures of between-group diversity. Further, we computed an index of probability of identity [21], which gives a quantitative measure of haplotype sharing between a pair of populations. To ascertain, for the given sample sizes, that the test has enough power at  $\alpha = 0.05$ – $0.001$ , we computed power required for the Mann-Whitney U test. For this purpose, we decreased the sample sizes by 15% and used this sample size to compute power required for a  $t$ -test. This rule is based on the lower bound for the asymptotic relative efficiency (ARE) of the Mann-Whitney U test versus the  $t$ -distribution, which is 0.864. This says that no matter what the distribution is, the ARE of the Mann-Whitney U test can never be worse than 0.864 for a reasonable broad class of probability distributions. Inverting that gives an increase in the sample size by a factor of 1.157, and therefore the sample sizes were reduced by 15% [22]. To increase the

resolution, in addition to the six Y-STRs, we typed nine more Y-STRs (DYS388, DYS426, DYS437, DYS438, DYS439, DYS447, DYS448, DYS460, and H4; Table S2) in three populations each of Mundari patrilineal groups (Bhumij, Munda, and Santhal) and Khasi matrilineal groups (Khyriam, Maram, and Pnar) and recomputed genetic distances based on 15 Y-STR loci.

## Supporting Information

**Table S1.** Y-Chromosome Haplotypes Based on Six Y-STRs for 15 Populations

Found at DOI: 10.1371/journal.pgen.0020053.st001 (946 KB DOC).

**Table S2.** Y-Chromosome Haplotypes Based on Nine Y-STRs for Six Populations

Found at DOI: 10.1371/journal.pgen.0020053.st002 (387 KB DOC).

## Accession Numbers

The GenBank (<http://www.ncbi.nlm.nih.gov>) accession numbers for the sequence discussed in this paper are HVSI (AY72095–AY721592).

## References

- Salem AH, Badr FM, Gaballah MF, Paabo S (1996) The genetics of traditional living. Y-chromosomal and mitochondrial lineages in the Sinai Peninsula. *Am J Hum Genet* 59: 741–743.
- Seielstad MT, Minch E, Cavalli-Sforza LL (1998) Genetic evidence for a higher female migration rate in humans. *Nat Genet* 20: 278–280.
- Perez-Lezaun A, Calafell F, Comas D, Mateu E, Bosch E, et al. (1999) Sex-specific migration pattern in Central Asian populations, revealed by analysis of Y-chromosome short tandem repeats and mtDNA. *Am J Hum Genet* 65: 208–219.
- Oota H, Kitano T, Jin F, Yuasa I, Wang L, et al. (2002) Extreme mtDNA homogeneity in continental Asian populations. *Am J Phys Anthropol* 118: 146–153.
- Oota H, Settheatham-Ishida W, Tiwawech D, Ishida T, Stoneking M (2001) Human mtDNA and Y-chromosome variation is correlated with matrilineal versus patrilineal residence. *Nat Genet* 29: 20–21.
- Mesa NR, Mondragon MC, Soto ID, Parra MV, Duque C, et al. (2000) Autosomal, mtDNA, and Y-chromosome diversity in Amerinds: Pre- and post-Columbian patterns of gene flow in South America. *Am J Hum Genet* 67: 1277–1286.
- Al-Zahery N, Semino O, Benuzzi G, Magri C, Passarino G, et al. (2003) Y-chromosome and mtDNA polymorphisms in Iraq, a crossroad of the early human dispersal and of post-Neolithic migrations. *Mol Phylogenet Evol* 28: 458–472.
- Fuselli S, Tarazona-Santos E, Dupanloup I, Soto A, Luiselli D, et al. (2003) Mitochondrial DNA diversity in South America and the genetic history of Andean highlanders. *Mol Biol Evol* 20: 1682–1691.
- Wilder JA, Kingan SB, Mobasher Z, Pilkington MM, Hammer MF (2004) Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of females versus males. *Nat Genet* 36: 1122–1125.
- Levinson D (1996) *Encyclopedia of world cultures*. New York: MacMillan. 3855 p.
- Korotayev AV (2003) Form of marriage, sexual division of labor, and postmarital residence in cross-cultural perspective: A reconsideration. *J Anthropol Res* 59: 69–89.
- Marlowe FW (2004) Marital residence among foragers. *Curr Anthropol* 45: 277–284.
- Austerlitz F, Heyer E (1998) Social transmission of reproductive behavior increases frequency of inherited disorders in a young-expanding population. *Proc Natl Acad Sci U S A* 95: 15140–15144.
- Excoffier L, Schneider S (1999) Why hunter-gatherer populations do not show signs of Pleistocene demographic expansions. *Proc Natl Acad Sci U S A* 96: 10597–10602.
- Burton ML, Moore CC, Whiting JWM, Romney AK (1996) Regions based on social structure. *Curr Anthropol* 37: 87–123.
- Reddy BM, Demarchi DA, Malhotra KC (2001) Patterns of variation in a caste-cluster of Dhangars of Maharashtra, India. *Coll Antropol* 25: 425–442.
- Reddy BM, Demarchi DA, Bharati S, Kumar V, Crawford MH (2004) Patterns of ethnic, linguistic, and geographic heterogeneity of palmar interdental ridge counts in the Indian subcontinent. *Hum Biol* 76: 211–228.
- Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *J Mol Evol* 19: 153–170.
- Nei M (1987) *Molecular evolutionary genetics*. New York: Columbia University Press. 333 p.
- Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139: 457–462.
- Melton T, Peterson R, Redd AJ, Saha N, Sofro AS, et al. (1995) Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. *Am J Hum Genet* 57: 403–414.
- Lehmann EL (1975) *Nonparametrics: Statistical methods based on ranks*. San Francisco: Holden-Day. 457 p.

## Acknowledgments

This work is essentially a part of three different plan projects of the Indian Statistical Institute (ISI), Kolkata being carried out by BMR in collaboration with Centre for Cellular and Molecular Biology (CCMB), Hyderabad, India. Thanks are due to Directors of both the ISI and CCMB for logistic support. We are grateful to a large number of anonymous subjects from different parts of India who volunteered to give blood samples. We are also grateful to the anonymous reviewers whose comments helped in improved presentation of the results, and to T. Krishnan, former Professor of the ISI, for statistical advice in computing power of the Mann-Whitney U test.

**Author contributions.** BMR conceived and designed the experiments. VK, BTL, KVM, VMN, HPS, and SB performed the experiments. VK and BMR analyzed the data. KT, LS, and BMR contributed reagents/materials/analysis tools. VK, BTL, and BMR collected samples. KT and LS commented on the draft of the manuscript. KT helped in preparing the diagrams. VK and BMR wrote the paper.

**Funding.** This work was funded by the Indian Statistical Institute.

**Competing interests.** The authors have declared that no competing interests exist. ■