

Differential pattern of genetic variability at the DXYS156 locus on homologous regions of X and Y chromosomes in Indian population and its forensic implications

Sanjukta Mukerjee · Meeta Mukherjee · Tania Ghosh ·
D. Kalpana · Anil Kumar Sharma

Received: 13 February 2011 / Accepted: 31 October 2011 / Published online: 25 November 2011
© Springer-Verlag 2011

Abstract Sex determination is routinely performed in forensic casework using the amelogenin-based sex test. The human amelogenin gene resides on homologous regions of the sex chromosomes. However, a deletion in the *AmelY* locus may sometimes lead to gender misidentification. The pentanucleotide microsatellite DXYS156 maps to the pseudoautosomal region of both the sex chromosomes and helps in sex determination. This STR offers an advantage of being multi-allelic, with delimited and demographically restricted alleles for the X and Y chromosomes. Also, the Y-specific alleles can be discerned from their X chromosomal counterpart due to an adenine insertion in the (TAAAA)_n repeat units of the STR. The present study examines the differential variation pattern at the X and Y locus of this STR in unrelated males from linguistically and geographically diverse populations of India. The study also attempts to undertake a comparison between the two sex-determining markers through validation studies. Two population samples and few validation samples which showed erroneous results for the amelogenin locus produced alleles specific to each of the sex chromosomes at the DXYS156 locus. The error rate for the amelogenin locus was observed to be 0.27% in case of the population samples and 0.5% in case of validation samples.

Statistical parameters of forensic interest indicate that the DXYS156 locus is polymorphic and discriminating for the Indian population.

Keywords Sex determination · Amelogenin · DXYS156 · Pentanucleotide repeat · Indian population

Introduction

Determining the gender of human DNA samples is of prime importance in forensics. The amelogenin locus is routinely used for sex determination of a sample because it resides on homologous regions of the sex chromosomes [1]. PCR primers are designed spanning the 6-bp deletion on X chromosome to produce differentially sized amplicons (106 and 112 bp) for the X and Y chromosomes, respectively [2]. However, this method can sometimes misidentify a male as a female due to a deletion in the *AmelY* region. This deletion has been reported to occur at a frequency of 0.08% in Caucasian males, 1.8% in Indians, and as high as 8% in Sri Lankans [3–6]. Varying lengths of interstitial deletions ranging from 1.0 Mb to around 3.8 Mb have been observed on the pericentromeric region of the short arm of Y chromosome, which may include the amelogenin locus [7, 8]. The deletion in *AmelY* region does not have considerable consequence on enamel formation because of the presence of the X homologue but has important implications for forensic casework [9].

Several authors have suggested the inclusion of additional Y locus for conclusive sex test and to prevent gender mistyping. Chen et al. [10] first described and characterized the presence of a pentanucleotide repeat [(TAAAA)_n] DXYS156, which maps to both the sex chromosomes and can be used for genetic identity testing. The microsatellite

Electronic supplementary material The online version of this article (doi:10.1007/s00414-011-0646-6) contains supplementary material, which is available to authorized users.

S. Mukerjee · M. Mukherjee · T. Ghosh · D. Kalpana ·
A. K. Sharma (✉)

Ministry of Home Affairs, Government of India, Central Forensic
Science Laboratory, Directorate of Forensic Science Services,
30 Gorachand Road, Park Circus,
Kolkata, West Bengal 700014, India
e-mail: aks4n6@gmail.com

DXS156 has been localized by linkage analysis to a region on the long arm of X chromosome (Xq) which shows sequence homology with the short arm of Y chromosome (Yp) [11, 12]. The strictly sex-linked region offers a high degree of X–Y sequence homology which has been attributed to a recent transposition from Xq to Yp followed by Yp inversion during hominid evolution about 3–4 million years ago [13]. Because of the high sequence similarity, the polymorphisms in this STR can be detected by a single primer pair flanking the pentanucleotide repeat on both the sex chromosomes [10]. The pentanucleotide repeat is known to occur within a human long interspersed repetitive element (LINE) with a single TAAAA motif in the consensus sequence. It is still unclear whether both X and Y chromosomes contain this repeat as a part of homology or as independent insertions [10]. Unlike in the pseudoautosomal regions, no cross-over occurs between these regions of X and Y suggesting that the alleles of this STR stay specific to the respective chromosome [14]. The smaller and larger alleles have usually been assigned to the X and Y chromosomes, respectively.

The multipurpose STR DXYS156 offers an advantage over amelogenin in being multi-allelic, with demographically restricted alleles which may indicate the probable geographic origin of an individual [15, 16]. Its X and Y alleles can be unambiguously distinguished from each other due to an adenine insertion which has been reported to be Y specific [15]. Thus, DXYS156 locus would help to uncover the differential genetic variability occurring on homologous regions of X and Y chromosomes [17]. Since there is a dearth of data on the genetic variation of this pentanucleotide marker in Indian populations, the study would be a step towards setting up a database for the use of this marker in forensic casework.

India is the home to a number of extant endogamous populations which belong to the four major linguistic groups, namely, Indo-European (IE), Dravidian (DR), Tibeto-Burman (TB), and Austro-Asiatic (AA). IE speakers are concentrated in northern and central India while southern India is predominantly occupied by DR speakers. Northeastern India is the home to mainly TB speakers while AA speakers occupy the pockets of central and eastern India. Indian population is also socially stratified into castes and tribes. The strategic location at the crossroads of the African, the northern Eurasian, and the Oriental realm and the presence of linguistically, demographically, and socially diverse populations account for the high genetic diversity in India, next only to that of Africa [18, 19]. Since the genetic diversity within a geographical region is an indicator of the age of the populations' dwelling in that area, it is quite likely that India became peopled with early colonizers migrating out of Africa along a southern route [19]. Genetic markers are pivotal in tracing human migration patterns and

studying population stratification. Studying the variability of homologous markers on the sex-specific regions of X and Y chromosomes would provide an assessment of the role of sex-specific processes in shaping human population structures and the differential selective pressures acting on these markers [20]. Such studies can also complement those of mtDNA and Y chromosome for exemplifying the sex-specific patterns of population structure [20]. In the present study, we explore the genetic diversity existing among few Indian populations at the X and Y locus of DXYS156 marker and their genetic relationship with the populations of Africa and East and Southeast Asia.

Materials and methods

Sampling and subjects

Blood samples from 749 unrelated healthy male individuals were collected from 11 endogamous populations of India [Balmiki ($n=62$), Sakaldwipi Brahmin ($n=65$), Kanyakubja Brahmin ($n=78$), Konkanastha Brahmin ($n=71$), and Mahadev Koli ($n=65$) populations belonging to IE linguistic group; Iyengar ($n=66$), Kurumans ($n=67$), and Gond ($n=75$) populations from DR group; Tripuri ($n=65$) and Riang ($n=67$) populations of TB group, and Munda ($n=68$) population belonging to AA group] distributed over six geographical regions (north, east, south, west, central, and northeast). At the social level, five caste and six tribal populations have been studied. A detailed description of the populations selected for the study has been made elsewhere [21]. The blood samples were collected with an informed consent following the protocols approved by the institutional ethics committee.

DNA isolation and quantitation

Genomic DNA was isolated and quantified as described previously [21].

Amplification and genotyping

Approximately 1 ng of DNA from population samples was amplified in 10 μ l reaction for DXYS156 locus in GeneAmp[®] PCR System 9700 (Applied Biosystems, Foster City, CA) following the protocol of Chen et al. [10], with minor modifications. The experiments were conducted in accordance with quality control measures and using cell lines 9947A, 9948, and K562 from Promega Corp. (Promega, Madison) as standard reference materials [22]. The amplified products were electrophoresed on an ABI PRISM[®] 3100 Genetic Analyzer with POP4 polymer using GeneScan[™] LIZ-500 size standard (Applied

Biosystems, Foster City, CA). The data were analyzed with GeneMapper® ID Software v3.2 (Applied Biosystems, Foster City, CA). The samples were also genotyped for the amelogenin locus using AmpF/STR® Identifiler® (Applied Biosystems, Foster City, CA) [21].

Sequencing of alleles

The repeat number and structure of the alleles at this locus were determined as described previously [15, 17]. The X and Y chromosome alleles were separated on a 2% Ultrapure Agarose gel and eluted using the QIAquick® Gel Extraction kit (Qiagen Inc., Valencia, CA) following the manufacturer's protocol. The alleles were confirmed by bidirectional sequencing using Big Dye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) and published primers.

Statistical analyses

The estimation–maximization algorithm was employed for computing maximum likelihood estimates of allele frequencies for DXYS156X and DXYS156Y loci using Arlequin program v3.1 [23]. Gene diversity, population pairwise genetic distances (F_{ST}), and analysis of molecular variance (AMOVA) among the studied populations were also computed using Arlequin program. For the evaluation of forensic efficiency of the marker, various statistical parameters were calculated with chromosome X website software (<http://www.chrx-str.org>) [24].

Validation studies

Ten paternity case samples of male individuals showing the absence of Y allele peak and two samples of Sakaldwipi Brahmin population having X allele peak absent [21] at the amelogenin locus of AmpF/STR® Identifiler® (Applied Biosystems, Foster City, CA) were employed for validation of DXYS156 STR for forensic casework. The Y allele deficient samples were also typed for 17 Y STR loci using AmpF/STR® Yfiler® (Applied Biosystems, Foster City, CA) for confirmation of the presence of Y chromosome. DNA isolated from various body fluids (blood stain, semen stain, and saliva), postmortem tissue samples, and swabs from surfaces (buccal, vaginal, and fingerprint) were also utilized for the validation of the marker. DNA isolation was followed by purification and concentration with Microcon® 100 to evade the problem of PCR inhibition. The same procedure was undertaken for determining the template concentration of validation samples as done for the population samples. Also, DNA from different male and female samples were mixed in various ratios and used for mixture analysis interpretation studies.

Results and discussion

The analyses of 749 X and Y chromosomes in 11 populations revealed 11 different alleles (5–15.1) spanning a range of 130 to 181 bp. The allele frequencies are presented in Table 1. Alleles were assigned to their respective chromosomes according to Cali et al. and Karafet et al. [15, 17]. For the DXYS156X locus, the highest frequency was observed for the allele 7, concordant with the finding that this allele is highly frequent in non-African populations [17]. The allele 12.1 showed highest frequency in all the studied populations at the DXYS156Y locus. The long allele Y14.1 was observed in high frequency in the populations of TB linguistic family inhabiting northeastern India while allele Y15.1 was found exclusively in these populations. These alleles have been reported to occur in East Asian populations [16, 17], suggestive of a genetic signature of the eastern part of Asia. Previous Y chromosomal studies have suggested that speakers of the TB language family arrived on waves of migration from southern China through the northeastern corridor of India [25]. This passage has also served as a geographic barrier for human migrations, thereby creating a genetic discontinuity and isolation between the populations of northeast and those of Indian mainland [25]. The overall allelic distribution pattern of DXYS156 in India depicts a blend of the African and Asian data. This is consistent with the finding that an early wave of exodus from Africa, along the coastline of India to south Asia was responsible for the colonization of south Asia and parts of southeastern Asia [19, 26].

The gene diversity values for the X chromosome locus were found to be smaller as compared to those for the Y chromosome locus in most of the studied populations. The highest value of gene diversity for the DXYS156X locus was observed in the Sakaldwipi Brahmin population (0.470) while for the DXYS156Y locus was observed in the Kurumans population (0.594). Endogamous populations inhabiting the same geographical areas were found to harbor comparable total gene diversity. The variation in the genetic diversity among the populations at the X and Y locus is mentioned in Table 1 (a) and (b) (supplementary data). The results of population pairwise F_{ST} at DXYS156X showed that only three populations had significant genetic distances from the remaining studied populations, while at DXYS156Y, most of the populations were significantly distinct from each other. Results of the exploration of genetic multiplicity by AMOVA at X and Y chromosome loci have been shown in Table 2 (a) and (b) (supplementary data).

The substantial genetic diversity at DXYS156Y locus compared to that at DXYS156X locus in Indian populations is in congruence with the astounding Y diversity in Asians

Table 1 Allele frequencies and forensic parameters for the studied population for the locus DXYS156

DXYS156 Allele	Balmiki		Sakaldwipi Brahmin		Kanyakubja Brahmin		Konkanastha Brahmin		Mahadev Koli		Iyengar		Kurumans		Gond		Tripuri		Riang		Munda	
	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
5																						
7	0.807	0.707	0.707	0.910	0.910	0.718	0.892	0.892	0.849	0.849	0.015	0.015	0.030	0.030	0.907	0.907	0.939	0.939	0.925	0.925	0.912	0.912
8	0.145	0.154	0.154	0.051	0.051	0.141	0.077	0.077	0.091	0.091	0.045	0.045	0.164	0.164	0.040	0.040	0.046	0.046	0.045	0.045	0.044	0.044
9	0.032	0.108	0.108	0.026	0.026	0.141	0.031	0.031	0.045	0.045			0.030	0.030	0.053	0.053	0.046	0.046	0.015	0.015	0.015	0.015
9.1															0.030							
11	0.016	0.031	0.031	0.013	0.013																	0.029
11.1	0.081	0.081	0.046	0.077	0.077	0.211	0.211	0.211	0.108	0.108	0.182	0.182	0.313	0.313	0.067	0.067	0.067	0.067	0.030	0.030	0.030	0.030
12.1	0.871	0.871	0.800	0.897	0.897	0.662	0.662	0.662	0.538	0.538	0.773	0.773	0.553	0.553	0.880	0.880	0.880	0.880	0.354	0.354	0.313	0.313
13.1	0.048	0.048	0.154	0.026	0.026	0.127	0.127	0.127	0.354	0.354	0.030	0.030	0.104	0.104	0.040	0.040	0.040	0.040	0.015	0.015	0.015	0.015
14.1											0.015	0.015	0.015	0.015	0.013	0.013	0.013	0.013	0.600	0.600	0.597	0.597
15.1															0.031	0.031	0.031	0.031	0.031	0.031	0.060	0.060
Null allele																						
GD	0.332	0.236	0.470	0.339	0.170	0.451	0.508	0.200	0.582	0.274	0.374	0.274	0.374	0.374	0.594	0.176	0.222	0.119	0.522	0.143	0.549	0.168
Total GD	0.645	0.705	0.705	0.593	0.593	0.742	0.698	0.698	0.665	0.665	0.744	0.744	0.744	0.744	0.602	0.602	0.663	0.663	0.676	0.676	0.574	0.574
PIC	0.298	0.219	0.426	0.301	0.162	0.403	0.446	0.186	0.491	0.254	0.328	0.254	0.328	0.334	0.516	0.166	0.210	0.113	0.422	0.137	0.467	0.161
Total PIC	0.573	0.650	0.650	0.504	0.504	0.696	0.644	0.644	0.598	0.598	0.699	0.699	0.699	0.699	0.517	0.517	0.598	0.598	0.616	0.616	0.477	0.477
PD _M	0.326	0.232	0.464	0.334	0.168	0.445	0.501	0.197	0.574	0.269	0.368	0.269	0.368	0.369	0.585	0.173	0.219	0.116	0.514	0.142	0.541	0.165
PE	0.075	0.039	0.158	0.079	0.022	0.144	0.189	0.029	0.260	0.051	0.096	0.051	0.096	0.096	0.273	0.023	0.035	0.011	0.243	0.016	0.226	0.021
PI	0.337	0.384	0.268	0.333	0.416	0.406	0.249	0.401	0.213	0.366	0.316	0.366	0.316	0.315	0.208	0.414	0.390	0.442	0.200	0.429	0.229	0.417

GD Gene Diversity, PIC Polymorphism Information Content, PD_M Power of Discrimination in males, PE Power of Exclusion, PI Paternity Index

rather than Africans [16, 17]. The mutation rate at Y chromosome is several folds higher than that at the X chromosome, favoring the accumulation of mutations in the repeat units. There is evidence that mutation rate enhances with the increase in the number of repeat units [27, 28]. Also, the smaller effective population size of the Y chromosome for each generation makes the mutations occurring on this chromosome more easily heritable [16]. Sex-specific processes such as lower migration rates and lower effective population size of males may also account for the discordant pattern of genetic variability on the sex chromosomes [20]. Previous Y chromosomal studies on Indian populations also reveal higher genetic differentiation at these loci due to the prevalent practices of restricted male social mobility [29]. Earlier studies analyzing the genetic structure among Indian ethnic populations have shown that they have resulted from admixture of four or five ancestral populations [29, 30]. There exists considerable genetic heterogeneity between castes and tribes, which might be attributed to their diverse origins, population histories, and restricted geographical territories, with a high degree of confounding occurring between language and geography [26]. The variations among the studied populations at DXYS156 computed by AMOVA also provide a reflection of previous observations. The genetic differentiation observed at the social level is indicative of little or no exchange of genetic material between tribal and caste populations [29, 30].

Forensic parameters for the STR are shown in Table 1. High values of Polymorphism Information Content (PIC), Power of Discrimination (PD), and Power of Exclusion (PE) were observed in the Kurumans, Konkanastha Brahmin, and Mahadev Koli populations. This STR was found to be polymorphic in all populations, with the southern and western regions of India displaying considerable diversity and polymorphisms. Alleles specific to X and Y chromosomes were observed for all validation samples which had failed to give positive result for the amelogenin sex test (Figs. 1 and 2, supplementary data). Samples having high detectable DNA (Table 3, supplementary data) produced peaks with high relative fluorescence units. For templates with low DNA concentration (<10 pg/ μ l), complete amplification was achieved by increasing the number of PCR cycles to 34. The procedure has also been described elsewhere [31]. Genotyping of mixture samples (containing both male and female DNA) showed the presence of alleles of both X and Y loci. The higher peak height and relative fluorescence units of the X chromosomal allele accounted for its occurrence in both the samples (Fig. 3, supplementary data). Further interpretation of such samples was done by typing additional X and Y STRs. Significantly high values were obtained for all the forensic parameters showing that this STR can be efficient-

ly used for sex determination, especially in situation where a failure is observed in amelogenin-based sex test.

Conclusions

The present study shows that the pentanucleotide STR DXYS156 is polymorphic for the Indian population. It can contribute towards DNA profiling and also serve as an additional sex-determining marker. Higher genetic variability was observed at the Y chromosomal locus than at its sex chromosomal counterpart in the studied populations. Few glimpses of genetic affinity of Indian populations with the neighboring ones are indicative of the central role played by India in colonization of south and southeastern Asia. This STR may be used to complement other Y chromosomal and mtDNA markers for a better understanding of the peopling of India. A database of this STR in the Indian population needs to be established for its further use in forensic identity testing and evolutionary studies. Additional studies on X–Y homologous markers may be undertaken to have an insight into the role of sex-specific processes in shaping the genetic structure of human populations.

Acknowledgments The study was funded by the Directorate of Forensic Science Services, Ministry of Home Affairs, Government of India, New Delhi. The authors are thankful to all the blood sample donors for their cooperation. Mukerjee S. was provided fellowship by the Directorate of Forensic Science Services for carrying out the research work. The authors would like to thank Patricia Balaesque and the anonymous reviewers for providing their useful comments for the improvement of the manuscript.

Conflict of interest The authors declare no conflict of interest.

References

1. Nakahori Y, Takenaka O, Nakagome Y (1991) A human X-Y homologous region encodes amelogenin. *Genomics* 9:264–269
2. Mannucci A, Sullivan KM, Ivanov PL, Gill P (1994) Forensic application of a rapid and quantitative DNA sex test by amplification of the X-Y homologous gene amelogenin. *Int J Legal Med* 106:190–193
3. Thangaraj K, Reddy AG, Singh L (2002) Is the amelogenin gene reliable for gender identification in forensic casework and prenatal diagnosis? *Int J Legal Med* 116:121–123
4. Kashyap VK, Sahoo S, Sitalaximi T, Trivedi R (2006) Deletions in the Y-derived amelogenin gene fragment in the Indian population. *BMC Med Genet* 7:2350–2357
5. Cadenas AM, Regueiro M, Gayden T, Singh N, Zhivotovskiy LA, Underhill PA, Herrera RJ (2007) Male amelogenin dropouts: phylogenetic context, origins and implications. *Forensic Sci Int* 166:155–163
6. Chang YM, Burgoyne LA, Both K (2003) Higher failures of amelogenin sex test in an Indian population group. *J Forensic Sci* 48:1309–1313
7. Lattanzi W, Giacomo MC, Lenato GM, Chimienti G, Voglino G, Resta N, Pepe G, Guanti G (2005) A large interstitial deletion

- encompassing the *amelogenin* gene on the short arm of the Y chromosome. *Hum Genet* 116:395–401
8. Jobling MA, Lo CC, Turner DJ, Bowden GR, Lee AC, Xue Y, Carvalho-Silva D, Hurler ME, Adams SM, Chang YM, Kraaijenbrink T, Henke J, Guanti G, McKeown B, van Oorschot RA, Mitchell RJ, de Knijff P, Tyler-Smith C, Parkin EJ (2007) Structural variation on the short arm of the human Y chromosome: recurrent multigene deletions encompassing *amelogenin* Y. *Hum Mol Genet* 16:307–316
 9. Kastelic V, Budowle B, Drobni K (2009) Validation of SRY marker for forensic casework analysis. *J Forensic Sci* 54:551–555
 10. Chen H, Lowther W, Avramopoulos D, Antonarakis D (1994) Homologous loci DXYS156X and DXYS156Y contain a polymorphic pentanucleotide repeat (TAAAA)_n and map to human X and Y chromosomes. *Human Mutat* 4:208–211
 11. Page D, Martinville BD, Barker D, Wyman A, White R, Francke U, Botstein D (1982) Single-copy sequence hybridizes to polymorphic and homologous loci on human X and Y chromosomes. *Proc Natl Acad Sci USA* 79:5352–5356
 12. Kim HS, Crow TJ (1998) Human DXYS156 of pentanucleotide repeat (TAAAA)_n: chromosomal localization by somatic hybrid mapping and sequencing analysis. *Cytogenet Cell Genet* 83:54–55
 13. Schwartz A, Chan DC, Brown LG, Alagappan R, Pettay D, Disteche C, McGillivray B, de la Chapelle A, Page DC (1998) Reconstructing hominid Y evolution: X-homologous block, created by X-Y transposition, was disrupted by Yp inversion through LINE-LINE recombination. *Hum Mol Genet* 7:1–11
 14. Freije D, Helms C, Watson MS, Donis-Keller H (1992) Identification of a second pseudoautosomal region near the Xq and Yq telomeres. *Science* 258:1784–1787
 15. Cali F, Forster P, Kersting C, Mirisola MG, D'Anna R, De Leo G, Romano V (2002) DXYS156: a multi-purpose short tandem repeat locus for determination of sex, paternal and maternal geographic origins and DNA fingerprinting. *Int J Legal Med* 116:133–138
 16. Kersting C, Hohoff C, Rolf B, Brinkmann B (2001) Pentanucleotide short tandem repeat locus DXYS156 displays different patterns of variations in human populations. *Croat Med J* 42:310–314
 17. Karafet T, de Knijff P, Wood E, Ragland J, Clark A, Hammer MF (1998) Different patterns of variation at the X- and Y-chromosome-linked microsatellite loci DXYS156X and DXYS156Y in human populations. *Hum Biol* 70:979–992
 18. Gadgil M, Joshi NV, Shambhu Prasad UV, Manoharan S, Patil S (1998) Peopling of India. In: Balasubramanian D, Rao N (eds) *The Indian human heritage*. Universities Press, Hyderabad, pp 100–129
 19. Majumder PP (2010) The human genetic history of South Asia. *Curr Biol* 20:184–187
 20. Balaesque P, Manni F, Dugoujon JM, Crousau-Roy B, Heyer E (2006) Estimating sex-specific processes in human populations: are X-Y homologous markers an effective tool? *Heredity* 96:214–221
 21. Ghosh T, Kalpana D, Mukerjee S, Mukherjee M, Sharma AK, Nath S, Rathod VR, Thakar MK, Jha GN (2010) Genetic diversity of autosomal STRs in eleven populations of India. *Forensic Sci Int Genet* 5:259–261
 22. Szibor R, Edelmann J, Hering S (2003) Cell line DNA typing in forensic genetics—the necessity for reliable standards. *Forensic Sci Int* 138:37–43
 23. Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform (Online)* 1:47–50
 24. Szibor R, Hering S, Edelmann J (2005) A new website compiling forensic chromosome X research is now online. *Int J Legal Med* 120:252–254
 25. 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
 26. Majumdar PP (2008) Genomic inferences on peopling of south Asia. *Curr Opin Genet Development* 18:280–284
 27. Primmer CR, Saino N, Moller AP, Ellegren H (1996) Directional evolution in germline microsatellite mutations. *Nat Genet* 13:391–393
 28. Brinkmann B, Klitschar M, Neuhuber F, Hühne J, Rolf B (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *Am J Hum Genet* 62:1408–1415
 29. Basu A, Mukherjee N, Roy S, Sengupta S, Banerjee S, Chakraborty M, Dey B, Roy M, Roy B, Bhattacharyya NP, Roychoudhury S, Majumdar PP (2003) Ethnic India: a genomic view, with special reference to peopling and structure. *Genome Res* 13:2277–2290
 30. Indian Genome Variation Consortium (2008) Genetic landscape of the people of India: a canvas for disease gene exploration. *J Genet* 87:3–20
 31. Kalpana D, Ghosh T, Mukerjee S, Mukherjee M, Sharma AK (2011) Pentaplex typing of new European Standard Set (ESS) STR loci in Indian Population. *Forensic Sci Int Genet*. doi:10.1016/j.fsigen.2011.07.003