Hereditary Genetics

Dogra, et al., Hereditary Genet 2015, 4:1
DOI: 10.4172/2161-1041.1000142

Research Article Open Access

Population Genetics for Autosomal STR Loci in Sikh Population of Central India

Dogra D[1], **Shrivastava P**[2*], **Chaudhary R**[3], **Gupta U**[2] **and Jain T**[2]

[1]*Department of Biotechnology, Barkatullah University, Bhopal 462023, Madhya Pradesh, India*

[2]*DNA Fingerprinting Unit, State Forensic Science Laboratory, Sagar 470001, Madhya Pradesh, India*

[3]*Biotechnology Division, Department of Zoology, Government Motilal Vigyan Mahavidyalaya, Bhopal 462023, Madhya Pradesh, India*

[*]**Corresponding author:** Pankaj Shrivastava, DNA Fingerprinting Unit, State Forensic Science Laboratory, Sagar-470001, Madhya Pradesh, India, Tel: 94243 71946, E-mail: pankaj.shrivastava@rediffmail.com

Rec date: February 4, 2015**, Acc date:** February 23, 2015, **Pub date:** February 26, 2015

Abstract

This study is an attempt to generate genetic database for three endogamous populations of Sikh population (Arora, Jat and Ramgariha) of Central India. The analysis of eight autosomal STR loci (D16S539, D7S820, D13S317, FGA, CSF1PO, D21S11, D18S51, and D2S1338) was done in 140 unrelated Sikh individuals. In all the three studied populations, all loci were in Hardy -Weinberg equilibrium except at locus FGA in Ramgariha Sikh and locus D16S539 in Arora Sikh. An analysis of molecular variance (AMOVA) showed 1% variation among the three studied populations. The close genetic relationship between Jat and Ramgariha Sikh population were confirmed in the MDS Plot generated from the pairwise genetic distances.

Keywords: Ramgariha dikh; Jat sikh; Arora sikh; Central India; Short tandem repeats; Population data

Introduction

Microsatellite markers are most suited for the genetic structural assessment of a population due to ease of use, co-dominant inheritance, high polymorphism and mutation rate [1,2]. Use of multiplex polymerase chain reaction based technology has made the task of identification easier and this has emerged as the dominant conclusive identification method in forensic investigation and for anthropological studies. India is a country rich in ethnic, cultural and linguistic variant groups. Although a considerable amount of information on polymorphism at microsatellite loci in humans is now available, but such studies are confined to limited groups [3-13]. Like most other Indians, Sikh is endogamous by caste and exogamous by sub caste [14]. Sikhism is India's fourth-largest religion and has existed for over 500 years, beginning with the birth of its founder Guru Nanak dev in the late 15th century C.E. in the Punjab region of what is today in India and Pakistan. The Sikhs community has a stronghold in the state of Punjab; roughly 60% of the population belongs to the Sikh faith. The state of Madhya Pradesh (MP) comprises about 1.9% of the Sikh population [15]. There is necessary to fill a big lacuna with information about the genetic diversity of Sikh Population. A very few number of genetic studies on Sikh population have been carried out around the world [16-23]. However, no STR marker based study on Sikh populations of central India has been reported in the literature till date. Therefore the present data would be used in the forensics and individual identification for these selected population groups and these genetic data would enrich the genetic informational resource. In the present study 140 unrelated individuals of three studied Sikh population Arora (n=40), Ramgariha (n=50) and Jat (n=50) were taken for analysis from MP, India on the nine microsatellite loci which are D13S317 (13q22-31), D7S820 (7q11.21-22), Amelogenin (X:p22.1-22.3;Y:p11.2), D2S1338 (2q35-37.1), D21S11 (21q11.2-q21), D16S539 (16q24-qter), D18S51 (18q21.3), CSF1PO (5q33.3-34) and FGA (4q28). These loci were chosen for two reasons. Firstly, they consist of repetitions of tetranucleotide repeat units and are therefore less prone to slippage of polymerase during enzymatic amplification [24]. Secondly they are located on different chromosomes so there is no possibility of mitotic recombination as they are present far apart from each other. All studied loci are substantially unlinked, which make them ideal tools to study genomic variation.

Materials and Methods

Sample collection

Venous blood from a total of 140 unrelated healthy individuals from three endogamous groups (50 Ramgariha, 50 Jat sikh and 40 Arora sikh) of Sikh population from Bhopal and Raisen district of Madhya Pradesh, India were taken on FTA card.

DNA extraction

A 1.2 mm punch from a dried sample spot on FTA paper was taken in a PCR tube. FTA purification reagent (200 µl) was added to PCR tube, incubated for 5 minutes at room temperature and then continuously agitated by using a pipette. This process was repeated thrice with FTA purification reagent and twice with 100 µl TE-buffer. Finally the entire unspent TE buffer was removed and discarded by pipetting and the disc was allowed to dry at room temp for overnight and was directly used for PCR amplification.

PCR amplification

Multiplexed PCR amplifications of the 9 STR loci: D16S539, D13S317, D7S820, CSF1PO, FGA, D21S11, D2S1338, D18S51 and amelogenin was performed using AmpFlSTR® MiniFilerTM PCR amplification kit (Applied Biosystem, Foster city, CA, USA). The PCR reagents have been standardized in the laboratory for consistency of results. PCR was performed by taking the ½ reaction volume of the manufacturer's recommended protocol [25] by using 9700 thermal

cycler (Applied Biosystems, USA). For one 1.2 mm washed punch of FTA paper the PCR mix was comprised of Reaction Buffer - 5.0 µL, Primers - 2.5 µL, MQ water - 5.0 µL to make final volume 12.5 µL.

Genotyping of amplified fragments

The PCR products were genotyped using multicapillary electrophoresis with POP-4 polymer in ABI Prism Avant 3100 Genetic Analyzer (Applied Biosystem, Foster city, CA, USA) according to the manufacturer's protocol provided with the kit and the data was analyzed using Gene Mapper Software v3.5 (Applied Biosystem, Foster city, CA, USA) to designate alleles by comparison with the allelic ladder supplied with the kit. Peak detection threshold was set to 50 RFUs for allele designation. All steps were according to the laboratory internal standards and respective kit controls.

Analysis of the data

Allele frequencies of the 8 STR loci were calculated by GenAlEx 6.5 software [26]. Several Statistical parameters of forensic importance like the power of discrimination (PD), polymorphism information content (PIC), matching probability (PM) and power of exclusion (PE) were calculated using the Excel PowerStats spreadsheet program [27].

PD is the probability that two randomly chosen persons would not have matching DNA profiles [28] and CPD is used to prove that selected loci can be safely used to establish DNA based database for the studied population. CPM describes the possibility of finding two individuals with the same genotype in the population is almost null. Heterozygosity is a measure of genetic variation within a population. High heterozygosity values for a breed may be due to long term natural selection for adaptation, to the mixed nature of the breeds or to historic mixing of strains of different populations. A low level of heterozygosity may be due to isolation with the subsequent loss of unexploited genetic potential. Observed heterozygosity is defined as the % of loci heterozygous per individual. Low observed heterozygosity values indicate inbreeding and may be a departure from HWE too. Observed and expected heterozygosities and Hardy-Weinberg equilibrium (HWE) using exact test were calculated using Arlequin v3.5 [29]. The same package of software was used for calculating AMOVA among the three endogamous studied populations. The genetic affinities among the studied three populations were observed by plotting dendogram using POPTREE software [30]. Nei's genetic distances for the studied populations with other published Indian population data [31-33] was calculated by using POPTREE software [30] which were graphically summarized using Principal Component Analysis (PCA) plot generated using Past v3.02a software [34] to visualize population affinities.

Results and Discussion

The genetic variation in allele frequency distribution at 8 STR loci and statistical analysis of forensic parameters for Arora, Jat and Ramgariha Sikh population are shown in Tables 1, 2 and 3 respectively. The common pattern of allele distribution was observed at all the studied loci which may be due to their practice of endogamy as a social rule. The distribution of the most common allele (MCA) and least common allele (LCA) in the three endogamous caste is presented in Table 4, when the significance level was corrected by the Bonferroni method [35] and P values <0.05/8=0.00625 were considered statistically significant only two deviation persisted one at locus FGA for Ramgariha sikh and other for locus D16S539 in Arora Sikh. Molecular variance analysis (AMOVA) was conducted to understand the intra and interpopulation variations in the three Sikh endogamous populations (Figure 1).

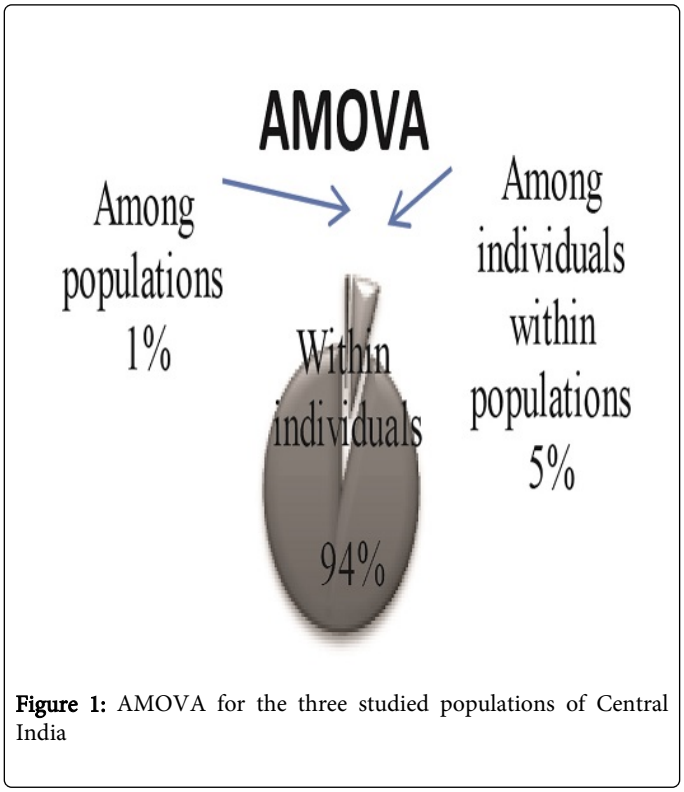

Figure 1: AMOVA for the three studied populations of Central India

The genetic relation among them is shown in (Figure 2). Jat and Ramgariha Sikh showed close affinity than the Arora Sikh. PCA plot (Figure 3) shows comparison of the three studied populations to other published population of India. The study shows that all the three populations showed significant differences from the tribal population. This finding is similar to the early reports [33-34] on caste and tribal population, which indicate that social stratification has played a major role in shaping the genetic diversity of India.

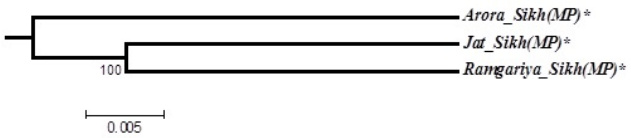

Figure 2: Genetic relationship among the three studied endogamous groups of Sikh population of Central India based on genetic distance

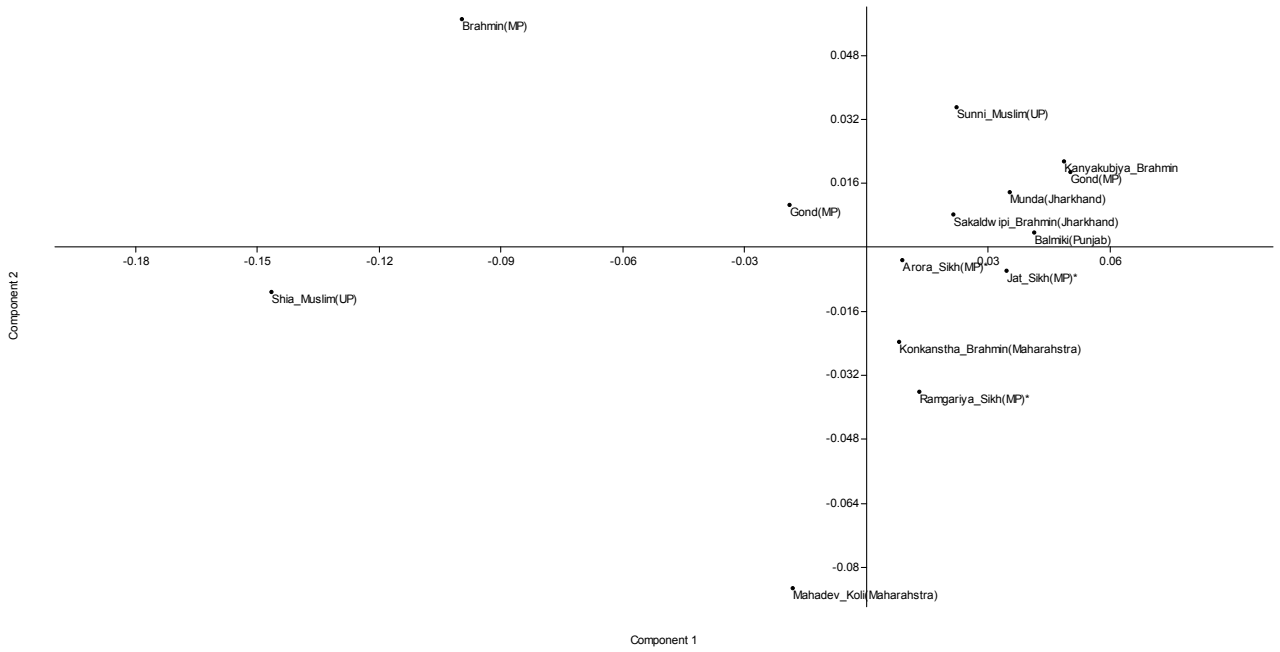

Figure 3: PCA plot for comparison of studied populations with other published populations based on Nei's genetic distance

Allele/n	D13S317	D7S820	D2S1338	D21S11	D16S539	D18S51	CSF1PO	FGA
7	0.02							
8	0.19	0.34						
9	0.17	0.06			0.18		0.02	
10	0.11	0.22			0.11	0.01	0.28	
11	0.13	0.18			0.32	0.02	0.28	
12	0.31	0.18			0.24	0.12	0.34	
13	0.03	0.01			0.13	0.08	0.07	
14	0.04	0.01			0.02	0.22	0.01	
15						0.27		
16			0.02			0.12		
17			0.07			0.07		
18			0.17			0.03		0.02
19			0.19			0.01		0.06
20			0.14					0.13
21			0.08			0.02		0.09
22			0.07			0.02		0.11

23			0.13			0.01		0.17
24			0.08					0.19
25			0.04					0.16
26			0.01					0.02
27				0.04				0.03
28				0.11				
29				0.17				
30				0.21				0.02
30.2				0.04				
31				0.01				
31.2				0.11				
32.2				0.24				
33.2				0.07				
PM	0.074	0.108	0.049	0.082	0.094	0.059	0.165	0.048
PD	0.926	0.892	0.951	0.918	0.906	0.941	0.835	0.952
PIC	0.782	0.731	0.861	0.817	0.744	0.818	0.671	0.852
PE	0.527	0.675	0.599	0.755	0.527	0.562	0.715	0.675
Ho	0.76	0.84	0.8	0.88	0.76	0.78	0.86	0.84
He	0.815	0.775	0.883	0.845	0.786	0.845	0.729	0.875
P-value	0.134	0.312	0.011	0.012	0.501	0.129	0.358	0.282
PM - matching probability, Power of Discrimination, PIC- Polymorphism Information Content, PE- Power of Exclusion, Ho- Observed Heterozygosity, He- Expected Heterozygosity, P -values- Hardy–Weinberg equilibrium (P< 0.00625)								

Table 1: Allele frequency distribution and forensic parameters for 8 autosomal STR loci investigated in Arora Sikh population of Central India

Allele/n	D13S317	D7S820	D2S1338	D21S11	D16S539	D18S51	CSF1PO	FGA
7	0.020							
8	0.190	0.340						
9	0.170	0.060			0.180		0.020	
10	0.110	0.220			0.110	0.010	0.280	
11	0.130	0.180			0.320	0.020	0.280	
12	0.310	0.180			0.240	0.120	0.340	
13	0.030	0.010			0.130	0.080	0.070	
14	0.040	0.010			0.020	0.220	0.010	
15						0.270		
16			0.020			0.120		
17			0.070			0.070		
18			0.170			0.030		0.020

19			0.190			0.010		0.060
20			0.140					0.130
21			0.080			0.020		0.090
22			0.070			0.020		0.110
23			0.130			0.010		0.170
24			0.080					0.190
25			0.040					0.160
26			0.010					0.020
27				0.040				0.030
28				0.110				
29				0.170				
30				0.210				0.020
30.2				0.040				
31				0.010				
31.2				0.110				
32.2				0.240				
33.2				0.070				
PM	0.074	0.108	0.049	0.082	0.094	0.059	0.165	0.048
PD	0.926	0.892	0.951	0.918	0.906	0.941	0.835	0.952
PIC	0.782	0.731	0.861	0.817	0.744	0.818	0.671	0.852
PE	0.527	0.675	0.599	0.755	0.527	0.562	0.715	0.675
Ho	0.760	0.840	0.800	0.880	0.760	0.780	0.860	0.840
He	0.815	0.775	0.883	0.845	0.786	0.845	0.729	0.875
P-value	0.134	0.312	0.011	0.012	0.501	0.129	0.358	0.282

PM - matching probability, Power of Discrimination, PIC- Polymorphism Information Content, PE- Power of Exclusion, Ho- Observed Heterozygosity, He- Expected Heterozygosity, P -values- Hardy–Weinberg equilibrium (P< 0.00625)

Table 2: Allele frequency distribution and forensic parameters for 8 autosomal STR loci investigated in Jat Sikh population of Central India

Allele/n	D13S317	D7S820	D2S1338	D21S11	D16S539	D18S51	CSF1PO	FGA
7	0.031	0.052						
8	0.333	0.271			0.094			
9	0.135	0.021			0.229			
10	0.167	0.198			0.042		0.219	
11	0.073	0.24			0.302	0.01	0.26	
12	0.146	0.219			0.24	0.083	0.417	
13	0.094				0.083	0.083	0.083	
14	0.021				0.01	0.26	0.021	

15						0.281		
16						0.135		
17			0.073			0.063		
18			0.219			0.042		
19			0.25			0.021		0.063
20			0.125					0.083
21			0.052			0.01		0.063
22			0.021			0.01		0.115
23			0.135					0.167
24			0.073					0.26
25			0.042					0.146
26								0.063
27			0.01	0.01				0.021
28				0.115				
29				0.208				
29.2				0.01				
30				0.219				0.021
31				0.01				
31.2				0.052				
32				0.021				
32.2				0.313				
33.2				0.042				
PM	0.077	0.098	0.055	0.083	0.09	0.077	0.161	0.057
PD	0.923	0.902	0.945	0.917	0.91	0.923	0.839	0.943
PIC	0.782	0.743	0.821	0.764	0.748	0.791	0.652	0.834
PE	0.622	0.475	0.546	0.475	0.271	0.703	0.546	0.35
Ho	0.813	0.729	0.771	0.729	0.583	0.854	0.771	0.646
He	0.814	0.787	0.849	0.801	0.789	0.823	0.711	0.859
P-value	0.268	0.222	0.293	0.359	0.015	0.493	0.579	0.001

PM - matching probability, Power of Discrimination, PIC- Polymorphism Information Content, PE- Power of Exclusion, Ho- Observed Heterozygosity, He- Expected Heterozygosity, P -values- Hardy–Weinberg equilibrium (P< 0.00625)

Table 3: Allele frequency distribution and forensic parameters for 8 autosomal STR loci investigated inRamgariya Sikh population of Central India

Conclusion

The present study involves the development of forensic databases for indigenous populations in Central India. The autosomal STRs indicate general allelic homogeneity in the three endogamous populations and it can be suggested that a single database based on combined random samples may be sufficient in addressing the forensic needs.

Acknowledgement

Authors are thankful to Director, State Forensic Science Laboratory, Sagar, MP, India for providing genotyping facility. The study was

supported by the I.M.B.I.B.E Award Fellowship grant from M.P. Biotechnology Council, Bhopal, India sanctioned to Devika Dogra.

Locus	Arora Sikh		Jat Sikh		Ramgariha Sikh	
	MCA	LCA	MCA	LCA	MCA	LCA
CSF1PO	11	9	12	14	12	14
D2S1338	19	16,26	19	26	19	27
D7S820	8	7,13	8	13,14	8	9
D13S317	12	15,16	12	7	8	14
D18S51	13,15	9	15	10,19,23	15	11,21,22
D21S11	28	27,29.2,31	32.2	31	32.2	27,29.2,31
D16S539	11	10	11	14	11	14
FGA	23,24	18	24	18,26,30	24	27,30

Table 4: Distribution of the most common allele (MCA) and least common allele (LCA) in the three endogamous Sikh populations of Central India

References

1. Tautz D, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res 12: 4127-4138.
2. Weber JL, May PE (1989) Abundant class of Human DNA polymorphisms which can be typed using the polymerase chain reaction. Amer J Human Genet 44: 388- 396.
3. Deka R Shriver MD, Yu LM, Ferrell RE, Chakraborty R (1995) Intra- and inter-population diversity at short tandem repeat loci in diverse populations of the world. Electrophoresis 16: 1659-1664.
4. Deka R, Shriver MD, Yu LM, Heidrich EM, Jin L, et al. (1999) Genetic variation at twenty-three microsatelliteloci in sixteen human populations. J Genet 78: 99–121.
5. Mountain JL, Hebert JM, Bhattacharyya S, Underhill PA, Ottolenghi C, et al. (1995) Demographic history of India and mtDNA-sequence diversity. Am J Hum Genet 56: 979-992.
6. Papiha SS, Mastana SS, Purandare CA, Jayasekara R, Chakraborty R (1996) Population genetic study of three VNTR loci (D2S44, D7S2 and D12S11) in five ethnically defined populations of the Indian subcontinent. Hum Biol 68: 819-835.
7. Mukherjee N, Majumder PP, Roy B, Roy M, Dey B, et al. (1999) Variation at 4 short tandem repeat loci in 8 population groups of India. Hum Biol 71: 439-446.
8. Panneerchelvam S, Vanaja N, Baskar D, Sivapriya V, Damodaran C (2001) Distribution of alleles of 12 STR loci in Tamilpopulation (South India). Forensic Sci Int 119: 126–128.
9. Reddy BH, Sun G, Luis JR, Crawford MH, Hemam NS, et al. (2001) Genomic diversity at thirteen short tandem repeat loci in a substructured caste population, Golla, of southern Andhra Pradesh, India. Hum Biol 73: 175-190.
10. Agrawal S, Müller B, Bharadwaj U, Bhatnagar S, Khan F (2002) Allele frequencies of microsatellite repeat loci in Bhargavas, Chaturvedis, and Brahmins of North India. J Forensic Sci 47: 678-681.
11. Agrawal S, Müller B, Bharadwaj U, Bhatnagar S, Khan F (2002) Distribution of allele frequencies of six STR markers in north Indians. J Forensic Sci 47: 682-685.
12. Agrawal S, Müller B, Bharadwaj U, Bhatnagar S, Khan F (2002) Evaluation of six short tandem repeat loci in forensics: north Indian populations. J Forensic Sci 47: 686-689.
13. Das B, Ghosh A, Chauhan PS, Seshadri M (2002) Genetic polymorphism study at four minisatellite loci (D1S80, D17S5, D19S20, and APOB) among five Indian population groups. Hum Biol 74: 345-361.
14. Singh P, Singh M, Bhatnagar DP, Kaur T, Mastana S (2008) Apolipoprotien C3 (SstI) Gene variability in Northwest India: A Global Perspective. Int J Hum Genet 8: 51-60.
15. http://www.censusindia.net/
16. Chattopadhyay PK, Cleef S, Duelmer M, Henke L, Henke J (2000) Human autosomal short tandem repeat types in Jat Sikhs from North India. Forensic Sci Int 113: 29-32.
17. Sidhu IS, Kaur K, Sarhadi VK, Joshi DS, Mukhopadhaya M, et al. (2003) Study of genetic polymorphism at D21S11 and D21S215 loci in the Jat Sikh population of Punjab. Int J Hum Genet 3: 45-50.
18. Babita K, Usha D (2004) HLA Antigen distribution in Sikhs from Punjab, India. Int J Hum Gent 4: 111-113.
19. Singh J, Sengupta S (2004) Some Morphogenetic and behavioral traits among the Assamese Sikh. Anthropologist 6: 253-255.
20. Teck TC, Kook SC, Badruddin N, Panneerchelvam S, Norazmi MN (2005) Autosomal STR polymorphism in Sikh population of Peninsular Malaysia. J Forensic Sci 50: 1223-1224.
21. Nagy M, Henke L, Henke J, Chatthopadhyay PK, Völgyi A, et al. (2007) Searching for the origin of Romanies: Slovakian Romani, Jats of Haryana and Jat Sikhs Y-STR data in comparison with different Romani populations. Forensic Sci Int 169: 19-26.
22. Chahal SM, Virk RK, Kaur S, Bansal R (2011) Social stratification in the Sikh population of Punjab (India) has a genetic basis: evidence from serological and biochemical markers. Genet Test Mol Biomarkers 15: 543-556.
23. Sanghera DK, Been LF, Ralhan S, Wander GS, Mehra NK, et al. (2011) Genome-wide linkage scan to identify loci associated with type 2 diabetes and blood lipid phenotypes in the Sikh Diabetes Study. PLoS One 6: e21188.
24. He L, Morris S, Lennon A, St Clair DM, Porteous DJ, et al. (1996). A genome-wide search for linkage in a large bipolar family: comparison of genotyping accuracy using di-and tetranucleotide repeat microsatellite markers. Psychiatr Genet 6: 123-129.
25. Shrivastava P, Mishra N, Sharma NC, Trivedi VB, Negi DS, et al. (2013) Autosomal STR genotyping Analysis of juvenile delinquents of Madhya Pradesh: A pilot study. Advanced Biotech 13: 20-24.

26. Peakall R, Smouse PE (2006) GENALEX6: genetic analysis in excel. Population genetic software for teaching and research. Mol Ecol Notes 6: 288-95.
27. Tereba A (1999) Tools for analysis of population statistics. Profiles DNA 2: 14-16.
28. Fisher RA (1951) Standard calculations for evaluating a blood-group system. Heredity (Edinb) 5: 95-102.
29. Excoffier L, Laval G, Schneider S (2007) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 1: 47-50.
30. Takezaki N, Nei M, Tamura K (2010) POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. Mol Biol Evol 27: 747-752.
31. Ghosh T, Kalpana D, Mukerjee S, Mukherjee M, Sharma AK, et al. (2011) Genetic diversity of autosomal STRs in eleven populations of India. Forensic Sci Int Genet 5: 259-261.
32. Dubey B, Meganathan PR, Eaaswarkhanth M, Vasulu TS, Haque I (2009) Forensic STR profile of two endogamous populations of Madhya Pradesh, India. Leg Med (Tokyo) 11: 41-44.
33. Eaaswarkhanth M, Dubey B, Ramakodi Meganathan P, Noor S, Haque I (2009) Microsatellite diversity delineates genetic relationships of Shia and Sunni Muslim populations of Uttar Pradesh, India. Hum Biol 81: 427-445.
34. Hammer O, Harper DAT, Ryan PD (2001) PAST: Palaeontological Statistics software package for education and data analysis. Palaeontologia Electronica 4: 9.
35. Bland JM, Altman DG (1995) Multiple significance tests: the Bonferroni method. BMJ 310: 170.