Journal of Forensic and Legal Medicine 37 (2016) 45–49

Contents lists available at ScienceDirect

Journal of Forensic and Legal Medicine

journal homepage: www.elsevier.com/locate/jflm

Short report

Baluchi and Pakhtun population data of 9 X-chromosomal short tandem repeat loci

Muhammad Israr [a, b, c, *], Ahmad Ali Shahid [a], Allah Rakha [b], Mian Sahib Zar [a], Muhammad Saqib Shahzad [d], Ziaur Rahman [a], Tayyab Husnain [a]

[a] National Centre of Excellence in Molecular Biology, University of the Punjab, 87-West Canal Bank Road, Lahore 53700, Pakistan
[b] Department of Forensic Sciences, University of Health Sciences, Lahore, Pakistan
[c] Department of Forensic Studies, University of Swat, Swat, Pakistan
[d] Department of Forensic Sciences, The University of Lahore, Lahore, Pakistan

ARTICLE INFO

Article history:
Received 15 April 2015
Received in revised form
8 September 2015
Accepted 24 October 2015
Available online 2 November 2015

Keywords:
Forensic science
DNA typing
X-chromosome short tandem repeats
Baluchi population
Pakhtun population
Population genetics

ABSTRACT

Baluchistan is the largest province of Pakistan in terms of area, constituting approximately 44% of the country's total land mass, and the smallest in terms of population, being home to less than 5% of the country's population. Khyber Pakhtunkhwa (KPK) formerly called North-West Frontier Province is located in the north-west of Pakistan having an estimated 13.4% of total population of Pakistan in which Pakhtuns are the major ethnic group. A total of 250 samples from Baluchi population and 250 samples from Pakhtun population were typed for 9 X-chromosomal STR markers: DXS101, DXS6789, DXS7132, DXS7423, DXS7424, DXS8378, GATA31E08, GATA172D05 and HPRTB along with sex typing locus, Amelogenin. A total of 59 alleles were found in Baluchi population while 61 alleles were found in Pakhtun population. This is the first study of the two populations based on these markers and the population data can be used as reference database for Baluchi and Pakhtun populations.

1. Introduction

Forensic scientists have developed, standardized and discarded many techniques for human identification, from fingerprinting to DNA profiling, until the data basing of core short tandem repeats (STR) loci. Since then, millions of profiles are generated and it is very likely that STRs will be the workhorses for the near future.[1,2]

To alleviate the problems associated with analyzing DNA from degraded samples a new set of STR primers known as Miniplexes were designed by moving the primers closer to the repeat region leaving the extra sequences out.[3,4] Using shorter amplicons in polymerase chain reaction (PCR), improvement has been reported in obtaining results from forensic evidence or a mass disaster site having degraded specimens.[5]

The DNA typing has played a pivotal role to establish the paternity of child which is utmost priority for support, inheritance right and other social benefits of a child. STRs located on X chromosome are powerful marker for complex kinship testing such as deficiency paternity testing when the disputed child is a female.[6,7] X-STRs are routinely used in parentage analysis and relationship investigations such as avuncular and first cousin relationships. X-STRs have also advantage over autosomal STRs for paternity cases involving close blood relatives as alternative putative fathers and in deficiency paternity cases, i.e. when the DNA sample from putative father is not available and DNA from paternal relative has to be analyzed instead.[8] Further, X-linked STRs can be used to solve sibling ship status, without using father's DNA, of two females having the same biological father.[9,10] X-STRs can determine the relationship of grandmother/granddaughter as granddaughter theoretically has to carry at least one allele in common with the grandmother.[11] In forensic analysis of mixed stains, X-STRs are helpful to identify the female DNA.[12,13] X-STR markers are low size markers and can efficiently be used for degraded DNA analysis.

This study reports 9 mini X-STRs (DXS101, DXS6789, DXS7132, DXS7423, DXS7424, DXS8378, GATA31E08, GATA172D05 and HPRTB) along with sex typing locus Amelogenin, data of Baluchi and Pakhtun populations of Pakistan using an ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems).

---

* Corresponding author. Department of Forensic Studies, University of Swat, Odigram Swat, Pakistan.
E-mail address: israr@uhs.edu.pk (M. Israr).

http://dx.doi.org/10.1016/j.jflm.2015.10.006
1752-928X/

1.1. General information about the populations

Baluchistan is a province of south-western Pakistan. It is the largest province of Pakistan in terms of area, constituting approximately 44% of the country's total land mass, and the smallest in terms of population, being home to less than 5% of the country's population. It has an eventful history dating back to the Stone Age. Recent research and archaeological excavations have revealed 9000 years old civilization (http://www.balochistan.gov.pk/index.php). Khyber Pakhtunkhwa (KPK) formerly called North-West Frontier Province is located in the north-west of Pakistan. The province has an estimated population of about 21 million, or 13.4% of total population of Pakistan (www.census.gov.pk). Pakhtuns or Pathans make up the largest ethnic group in province, historically have been living in the areas for centuries (http://www.khyberpakhtunkhwa.gov.pk/aboutus/History.php).

2. Materials and methods

2.1. Populations samples collection

After the study was approved by the Institutional Review Board of CEMB (Centre of Excellence in Molecular Biology), Blood samples were collection from 250 individuals (100 males, 150 females) from Baluchistan and 250 individuals (150 males and 100 females) from KPK after their informed consent were obtained. All the individuals in this study were unrelated with each other and were residents of the area since as far as 3 generations back. This was by self-declaration as no genealogy records are maintained in these parts of the world.

2.2. DNA extraction and quantification

Phenol chloroform DNA extraction was performed on each of the blood sample as reported in the literature.[14] DNA concentration was estimated through NanoDrop ND-1000 (NanoDrop Technologies, Wilmington, DE).

2.3. PCR amplification and genotyping

The STRs were amplified in multiplex set which included 9 STRs (DXS7132, DXS8378, DXS101, DXS6789, DXS7424, GATA31E08, DXS7423, GATA172D05 and HPRTB) plus Amelogenin.[15–19] The Polymerase chain reaction (PCR) was carried out on ABI GeneAmp® PCR System 2700 using temperature conditions as given in Supplementary Fig. 1. The reaction was carried out in a 10 µL reaction consisting of 1 unit of AmpliTaq Gold polymerase (ABI), 200 µM of each dNTP, 2 mM $MgCl_2$, 1X (NH_4SO_2) PCR Buffer (ABI), 1 ng of genomic DNA, distilled water and appropriate amount of labeled primers (Supplementary Table 1).

The capillary electrophoresis was performed using 1 µL of PCR product, 13.6 µL Hi-Di™ Formamide (ABI), and 0.4 µL GeneScan® 500-LIZ™ size standard (ABI). Using a 3130xl Genetic Analyzer (ABI), at 3 kV the samples were injected for 10 s. In a Performance Optimized Polymer (POP-4), electrophoresis was carried out for 25 min at 15 kV and run temperature of 60 °C. ABI 3130xl Data Collection Software application v3.0 was used for collection of data. Results were analyzed using GeneMapper software v3.7 (ABI).

2.4. Data analysis

Allele frequencies for each locus were calculated for both males and females collectively using a spreadsheet program. For Hardy–Weinberg Exact Test and Linkage Disequilibrium, Arlequin v3.5 was used.[20] PowerMarker v3.1 was used for Gene Diversity,[21] Polymorphic Information Content (PIC), and Heterozygosity. The ChrX-STR website was used for calculation of Paternity Index, Power of Exclusion, Power of Discrimination in males (PDm) and females (PDf), and Mean Exclusion Chance (MEC) in trios involving a daughter and in father/daughter duos.[7]

Alleles were assigned as recommended by ISFG (International Society of Forensic Genetics) through comparison with standard DNA 9947A. Allele nomenclature and allele ranges were according to already reported literature.[7] This manuscript strictly follows the guidelines for publication of population data requested by the journal.

Table 1
Allele frequency distribution of 9 X-STRs in Baluchi population.

Alleles	DXS101	DXS6789	DXS7132	DXS7423	DXS7424	DXS8378	GATA172D05	GATA31E08	HPRTB
6							0.1850		
7								0.2475	
8							0.1400	0.0150	
9						0.0450	0.1025	0.1325	
10						0.3325	0.2425	0.2775	
11						0.5025	0.2350	0.2625	0.0075
12			0.1025	0.0300	0.0925	0.0750	0.0950	0.0650	0.1225
13			0.2850	0.4150	0.0725	0.0450			0.3475
14		0.1025	0.3850	0.3875	0.1750				0.3175
15		0.0575	0.1825	0.0950	0.2575				0.1375
16		0.0225	0.0375	0.0725	0.2650				0.0675
17			0.0075		0.0825				
18		0.0300			0.0550				
19		0.2625							
20		0.1900							
21	0.1275	0.2400							
22	0.0425	0.0725							
23	0.1925	0.0150							
24	0.2850	0.0075							
25	0.1000								
26	0.1275								
27	0.0925								
28	0.0325								

Table 2
Different forensic efficiency parameters for Baluchi population.

Marker	DXS101	DXS6789	DXS7132	DXS7423	DXS7424	DXS8378	GATA172D05	GATA31E08	HPRTB	Mean
Gene Diversity:	0.8272	0.798	0.7079	0.6524	0.8223	0.651	0.8169	0.7647	0.734	0.7399
Heterozygosity:	0.5	0.6304	0.6304	0.587	0.2121	0.5652	0.6957	0.8261	0.75	0.6397
PIC:	0.8084	0.771	0.6592	0.5884	0.8023	0.5892	0.7908	0.7254	0.6896	0.6979
Power of Exclusion (PE):	0.651259	0.629348	0.466675	0.372047	0.616166	0.322944	0.62292	0.545737	0.491744	
Paternity Index (PI):	0.086192	0.091926	0.137902	0.168968	0.095416	0.187118	0.093624	0.114679	0.130319	
[a]PD female:	0.949313	0.941976	0.878038	0.824066	0.937615	0.796644	0.938124	0.909584	0.889706	
[b]PD male:	0.827617	0.816149	0.724195	0.662065	0.809168	0.625764	0.812753	0.770642	0.739363	
MEC Krüger:	0.661871	0.639941	0.485447	0.40141	0.626118	0.366295	0.625898	0.550794	0.509617	
MEC Kishida:	0.806646	0.791926	0.678301	0.600331	0.7832	0.56234	0.785824	0.732831	0.697	
MEC Desmarais:	0.806646	0.791926	0.678301	0.600331	0.7832	0.562461	0.785938	0.732831	0.697	
MEC Desmarais Duo:	0.691712	0.672939	0.536532	0.455305	0.661326	0.416931	0.66371	0.598602	0.557869	
Hardy–Weinberg:	**0**	**0.0008**	0.163	**0.032**	**0**	0.138	0.192	0.295	**0.027**	

MEC Krueger, Mean Exclusion Chance for Autosomal markers typed in trios involving mother, child and putative father.
MEC Kishida, Mean Exclusion Chance for ChrX markers which covers trios including a daughter.
MEC Desmarais, Mean Exclusion Chance for ChrX markers in trios involving daughters.
MEC Desmarais, Mean Exclusion Chance for father/daughter duos lacking maternal genotype information.
Deviations from Hardy-Weinberg equilibrium are shown in bold.
[a] Only female's data is used for the analysis.
[b] Only male data is used for the analysis.

Allele frequencies were calculated by simple counting. Allelic ladder was constructed by mixing all the hemizygous male samples having all the different alleles for a particular locus. This mixture was then amplified in thermocylcer with Polymerase, buffer and the specific primers for that locus. The ladder was sequenced for confirmation of repeat numbers. The allele fragments were numbered by comparing with standard DNA 9947A.

3. Results and discussion

A total of 59 alleles were found at all the 9 loci tested in Baluchi population. Maximum of 10 alleles were observed at locus DXS6789 while minimum of 5 alleles were found at DXS7423 and DXS8378 (Table 1). Minimum gene diversity of 0.651 was observed at locus DXS8378 while maximum of 0.8272 was observed at locus DXS101. Locus GATA31E08 was the most heterozygous with 0.8261 while the least heterozygous was DXS7424. Polymorphic Information Content (PIC) ranged from 0.5884 at locus DXS7423 to 0.8084 at locus DXS101. The combined power of discrimination in females is 1.16×10^{-10} while in male it is 9.71×10^{-7}. The population deviates from Hardy–Weinberg at locus DXS101, DXS6789 and DXS7424 even after Bonferroni's correction for multiple testing (Table 2). The reason for this may be the intermarriages in the same cast. These families are living in these areas for very long time and usually do not marry outside family, which might be the reason for a limited gene pool and deviation from Hardy–Weinberg at some loci.

At all of the 9 loci, 61 alleles were found in the Pakhtun population of KPK. Maximum of 10 alleles were observed at locus DXS6789 while minimum of 5 alleles were observed at DXS7423 (Table 3). Minimum gene diversity of 0.56 was observed at locus DXS8378 while maximum of 0.8075 was observed at locus DXS6789. Locus GATA31E08 was the most heterozygous with 0.80 while the least heterozygous was DXS7424 with 0.3684. Polymorphic Information Content (PIC) ranged from 0.5202 at locus DXS8378 to 0.7781 at locus DXS6789 (Table 4). The combined power of

Table 3
Allele frequency distribution of 9 X-STRs in Pakhtun population.

Alleles	DXS101	DXS6789	DXS7132	DXS7423	DXS7424	DXS8378	GATA172D05	GATA31E08	HPRTB
6							0.0413		
7								0.2586	
8							0.1736	0.0172	
9						0.0661	0.0826	0.1724	
10						0.1983	0.4132	0.2586	
11						0.5372	0.1405	0.2414	
12			0.0796	0.0833	0.0932	0.1570	0.1488	0.0517	0.0769
13			0.3274	0.4083	0.0678	0.0165			0.2479
14		0.1525	0.4071	0.3917	0.1102	0.0248			0.3846
15		0.0678	0.1239	0.0833	0.2034				0.2308
16		0.0169	0.0531	0.0333	0.3390				0.0342
17					0.1356				0.0256
18		0.0169	0.0088		0.0508				
19		0.1864							
20		0.0932							
21	0.0661	0.2797							
22	0.0496	0.1356							
23	0.3471	0.0339							
24	0.1818	0.0169							
25	0.1818								
26	0.0826								
27	0.0496								
28	0.0331								
29	0.0083								

Table 4
Different forensic efficiency parameters for Pakhtun population.

Marker	DXS101	DXS6789	DXS7132	DXS7423	DXS7424	DXS8378	GATA172D05	GATA31E08	HPRTB	Mean
Gene Diversity:	0.7738	0.8075	0.7701	0.635	0.6524	0.56	0.7413	0.7788	0.6731	0.6767
Heterozygosity:	0.6	0.7895	0.6842	0.7	0.3684	0.6	0.8	0.75	0.5789	0.6371
PIC:	0.738	0.7781	0.7349	0.5679	0.5921	0.5202	0.7042	0.7425	0.6101	0.6293
Power of Exclusion (PE):	0.581663	0.665802	0.432129	0.379722	0.579909	0.33363	0.510126	0.554009	0.472357	
Paternity Index (PI):	0.104713	0.082429	0.148753	0.166276	0.105193	0.183019	0.124895	0.112356	0.136164	
[a]PD female:	0.931056	0.952826	0.862221	0.82967	0.929146	0.823624	0.905475	0.912239	0.881189	
[b]PD male:	0.790573	0.835143	0.702494	0.667449	0.789615	0.633961	0.75021	0.775288	0.727673	
MEC Krüger:	0.607703	0.67432	0.46035	0.409694	0.601916	0.402438	0.542572	0.556349	0.491739	
MEC Kishida:	0.765374	0.815147	0.653106	0.607709	0.763023	0.59145	0.718081	0.737906	0.683024	
MEC Desmarais:	0.765489	0.815147	0.653225	0.607709	0.763023	0.59157	0.718081	0.738023	0.683024	
MEC Desmarais Duo:	0.64003	0.703094	0.510109	0.46279	0.636412	0.443062	0.581725	0.604497	0.541901	
Hardy–Weinberg:	0.183	0.701	0.158	0.802	**0.005**	1	0.745	**0.007**	0.368	

MEC Krueger, Mean Exclusion Chance for Autosomal markers typed in trios involving mother, child and putative father.
MEC Kishida, Mean Exclusion Chance for ChrX markers which covers trios including a daughter.
MEC Desmarais, Mean Exclusion Chance for ChrX markers in trios involving daughters.
MEC Desmarais, Mean Exclusion Chance for father/daughter duos lacking maternal genotype information.
Deviations from Hardy-Weinberg equilibrium are shown in bold.

[a] Only female's data is used for the analysis.

[b] Only male data is used for the analysis.

discrimination in females is 9.4×10^{-10} while in male it is 4.02×10^{-6}. The population does not deviate from Hardy–Weinberg equilibrium at all loci except DXS7424 and GATA31E08 but after Bonferroni's correction, there was no deviation at any loci.

Nei's distances were estimated based on allele frequency distribution using Poptree v.2 software.[22] Based on these markers, the two populations in this study tend to be closer as compared to other previously reported Pakistani populations like Punjabi and Sindhi.[15] However, these populations were closer to each other as compared to other Indian populations. They were closest to Riang and Tripuri populations of India and the farthest from Sakaldwipi Brahmin and Balmiki of India[23] (Fig. 1).

To the best of our knowledge, this is the first study involving these 9 STRs in Baluchi and Pakhtun populations. Although, there is another study which reported 13 X-STR loci in Pakistani population of which 7 markers are the same as used here.[18] However, as Baluchis are only 5% while KPK comprises only 13% of the total population of Pakistan, hence a meaningful comparison is not possible with the earlier study. X-STRs included in this study span the whole X-chromosome and represent all the 4-linkage groups categorized previously by Szibor and colleagues.[24] For all pair of loci, the exact test for LD was performed in Baluchi and Pakhtun populations in this study. No evidence of LD was detected in all pair of markers. This is consistent with a previous study on generic Pakistani population did not report any such linkage disequilibrium.[25] All markers of this study are recommended for forensic practice in KPK population. However, since Baluchi population deviates from Hardy–Weinberg at locus DXS101, DXS6789 and DXS7424 even after Bonferroni's correction for multiple testing, more research is needed before these 3 markers are to be used for forensic purposes in this population.

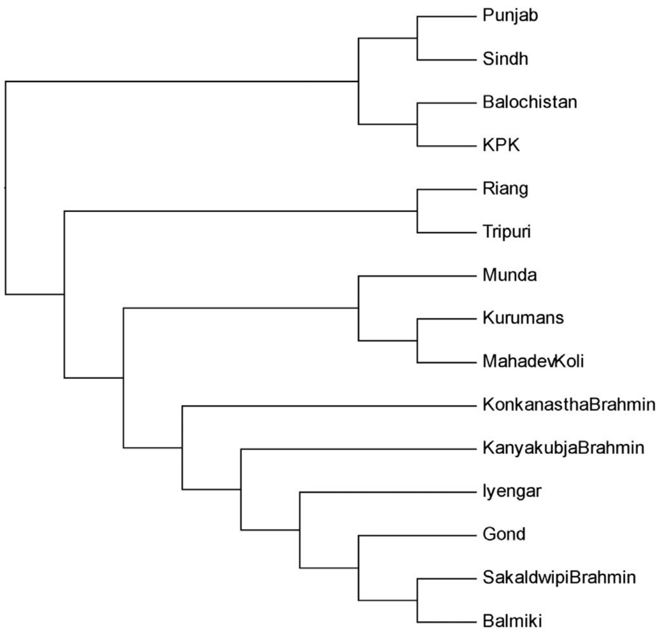

Fig. 1. Pakistani and Indian populations based on X-STR frequencies.

Author's contributions

MI, ZR and AAS conceived the study, perform and analyzed the data. AR, MSZ helped in collection of samples and data analysis. MSZ and TH helped in drafting and critically reviewing the manuscript. All the authors read and approved the manuscript.

Conflict of interests

The authors declare that they have no competing interests.

Funding

None.

Ethical approval

None.

Acknowledgments

The authors would like to acknowledge all individuals who participated in this study. The authors would like to thank Dr. Shaheen N. Khan, Obaid Ullah, SM Ibrahim and M. Akram Tariq for their support in this study. This work was partially supported by Higher Education Commission (HEC), Islamabad, Pakistan.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jflm.2015.10.006.

References

1. Gill P, Werrett DJ, Budowle B, Guerrieri R. An assessment of whether SNPs will replace STRs in national DNA databases—joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDAM). *Sci Justice* 2004;**44**:51–3.
2. Muller K, Klein R, Miltner E, Wiegand P. Improved STR typing of telogen hair root and hair shaft DNA. *Electrophoresis* 2007;**28**:2835–42.
3. Butler JM, Shen Y, McCord BR. The development of reduced size STR amplicons as tools for analysis of degraded DNA. *J Forensic Sci* 2003;**48**: 1054–64.
4. Coble MD, Butler JM. Characterization of new miniSTR loci to aid analysis of degraded DNA. *J Forensic Sci* 2005;**50**:43–53.
5. Drabek J, Chung DT, Butler JM, McCord BR. Concordance study between Miniplex assays and a commercial STR typing kit. *J Forensic Sci* 2004;**49**: 859–60.
6. Butler JM. *Advanced topics in forensic DNA typing: methodology*. Academic Press; 2011.
7. Szibor R, Hering S, Edelmann J. A new Web site compiling forensic chromosome X research is now online. *Int J Legal Med* 2006;**120**: 252–4.
8. Barbaro A, Cormaci P. X-STR typing for an identification casework. *Int Congr Ser* 2006;**1288**:513–5.
9. Toni C, Presciuttini S, Spinetti I, Domenici R. Population data of four X-chromosome markers in Tuscany, and their use in a deficiency paternity case. *Forensic Sci Int* 2003;**137**:215–6.
10. Toni C, Presciuttini S, Spinetti I, Rocchi A, Domenici R. Usefulness of X-chromosome markers in resolving relationships: report of a court case involving presumed half sisters. *Int Congr Ser* 2006;**1288**:301–3.
11. Edelmann J, Lessig R, Klintschar M, Szibor R. Advantages of X-chromosomal microsatellites in deficiency paternity testing: presentation of cases. *Int Congr Ser* 2004;**1261**:257–9.
12. Shin SH, Yu JS, Park SW, Min GS, Chung KW. Genetic analysis of 18 X-linked short tandem repeat markers in Korean population. *Forensic Sci Int* 2005;**147**: 35–41.
13. Szibor R, Lautsch S, Plate I, Beck N. Population data on the X chromosome short tandem repeat locus HumHPRTB in two regions of Germany. *J Forensic Sci* 2000;**45**:231–3.
14. Sambrook J, Russel DW. *Molecular cloning: a laboratory manual*. New York: Cold Spring Harbor Laboratory Press; 2001.
15. Israr M, Shahid AA, Rahman Z, Zar MS, Shahzad MS, Husnain T, et al. Development and characterization of a new 12-plex ChrX miniSTR system. *Int J Legal Med* 2014;**128**:595–8.
16. Asamura H, Sakai H, Kobayashi K, Ota M, Fukushima H. MiniX-STR multiplex system population study in Japan and application to degraded DNA analysis. *Int J Legal Med* 2006;**120**:174–81.
17. Edelmann J, Deichsel D, Hering S, Plate I, Szibor R. Sequence variation and allele nomenclature for the X-linked STRs DXS9895, DXS8378, DXS7132, DXS6800, DXS7133, GATA172D05, DXS7423 and DXS8377. *Forensic Sci Int* 2002;**129**: 99–103.
18. Tariq M, Ullah O, Riazuddin S, Riazuddin S. Allele frequency distribution of 13 X-chromosomal STR loci in Pakistani population. *Int J Legal Med* 2008;**122**: 525–8.
19. Haas-Rochholz H, Weiler G. Additional primer sets for an amelogenin gene PCR-based DNA-sex test. *Int J Legal Med* 1997;**110**:312–5.
20. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 2010;**10**:564–7.
21. Liu K, Muse SV. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 2005;**21**:2128–9.
22. Takezaki N, Nei M, Tamura K. POPTREE2: software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. *Mol Biol Evol* 2010;**27**:747–52.
23. Mukerjee S, Ghosh T, Kalpana D, Mukherjee M, Sharma AK. Genetic variation of 10 X chromosomal STR loci in Indian population. *Int J Legal Med* 2010;**124**: 327–30.
24. Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, Krause D. Use of X-linked markers for forensic purposes. *Int J Legal Med* 2003;**117**:67–74.
25. Tariq MA, Sabir MF, Riazuddin SA, Riazuddin S. Haplotype analysis of two X-chromosome STR clusters in the Pakistani population. *Int J Legal Med* 2009;**123**: 85–7.