

Genetic polymorphism of ten MiniSTR loci in the population of Punjab Pakistan for forensic application

Muhammad Shafique · Muhammad Saqib Shahzad ·
Muhammad Adnan Shan · Azam Ali · Ziaur Rahman ·
Tayyab Husnain

Received: 11 July 2014 / Accepted: 14 November 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract Ten MiniSTR loci were analyzed on 250 unrelated individuals of Punjabi population from Punjab province of Pakistan. The product amplification range is from 65 to 280 bp. Allele frequency for a total of 98 observed alleles is from 0.002 to 0.41. Forensic and paternity statistical parameters were investigated including combined power of discrimination (PD), combined power of exclusion (PE), and cumulative probability of matching (PM) equaled to 0.99999999999824, 0.999824, and 1.75931×10^{-12} , respectively. These MiniSTRs on the basis of high degree polymorphism and fragment size reduction will be highly informative in most of the forensic cases where DNA of the samples is degraded, mass disasters, and dead body identification. Significant differences were observed in all loci when comparing with same STR loci in previously published different world populations.

Keywords MiniSTRs · Genetic polymorphism · Forensic · Punjabi population

In comparison to minisatellites, the small amplified fragment length of STRs facilitates its utility in the analysis of degraded DNA samples [1, 2]. Some existing high molecular weight markers have been converted into low molecular weight (<130 bp) “mini-STR” multiplexes simply by moving the primer binding sites closer to the STR repeat region (<http://www.cstl.nist.gov/div831/strbase/miniSTR.htm>) [3, 4]. The Punjab province of Pakistan is the most developed with approximately 56 % of the country's total population comprising different tribes, clans, and communities. Important within the Punjab include Jats, Arains, Gujjars, Awans, Rajputs, Gakhars, Khokhars, Sheikhs, Aheers, Kamboh, Mughals, Maliks, Niazi, Khosas, Dogars, Ranghars, mirani, Qureshis, and Syeds. Historically, these tribes were endogamous, but latterly, especially in the large cities, there is considerable intermarriage between members of different tribes, and differences are getting blurred [5].

Blood samples were collected from 250 unrelated volunteers with consent on the form approved by the ethical committee of the National Centre of Excellence in Molecular Biology, University of the Punjab Lahore Pakistan. DNA extraction was carried out using phenol chloroform procedure and quantified by Quantifiler™ Human DNA Kit on a 7500 SDS Real-Time PCR System (Applied Biosystems). Reference sequences were obtained from GenBank® (<http://www.ncbi.nlm.nih.gov>) [6] for PCR primers using web-based Primer3 [7] shown in Table 1. PCR amplification was performed using 1X PCR buffer, 1.5 mmol/L MgCl₂, 200 μmol/L dNTPs (Invitrogen), 1 μL of template DNA at 1 ng/μL, and 1 unit of Platinum DNA Taq polymerase (Life Technologies). Thermal cycling was performed with the GeneAmp 9700 (Applied

M. Shafique · Z. Rahman · T. Husnain
Forensic DNA Typing Laboratory, National Centre of Excellence in
Molecular Biology, University of the Punjab Lahore, Lahore 53700,
Pakistan

M. S. Shahzad (✉) · A. Ali
Forensic Science Section, Center for Research in Molecular
Medicine, The University of Lahore, Lahore, Pakistan
e-mail: Muhammad.shahzad@imbb.uol.edu.pk

M. A. Shan
Forensic Services Laboratory, Centre for Applied Molecular Biology,
Ministry of Science and Technology Lahore, Lahore 53700, Pakistan

Table 1 MiniSTR loci used in this study with primer sequences, gene bank accession, chromosomal location, and the size range

STR locus	Primer sequence (5' to 3')	Gene bank accession	Chromosome location	MiniSTR size (bp)	Product size (bp) in STR kits	Size reduced (bp)
CSF1PO [8]	F 6FAM-acagtaactgccctcatagatag R ggtcagacccctgtctaaagta	X14720	5q33.1	84–121	304–341 (Identifier)	220
D18S51 [8]	F VIC-tgagtgacaattgagacct R gtttacaataacagttgctactatt	AP001534	18q21.33	107–190	262–345 (ProPlus)	155
D7S820 [8]	F 6FAM-gaacacttgcatagtttagaacgaac R tcatgacagaatgaccca	AC004848	7q21.11	135–173	255–293 (ProPlus)	120
D2S1338	F VIC-gagccagtggtttggaac R tcctaccagaatgccagtc	AC010136	2q35	219–272	307–360 (Identifier)	88
TPOX [8]	F VIC-cttaggggaacctcacigaatg R gtcctgtcagcgtttattgc	M68651	2p25.3	65–93	222–250 (Cofiler)	157
D5S818 [8]	F PET-ggggtatttctcttcttgg R aacattgtatcttctctgtatcctatttat	AC008512	5q23.2	106–144	134–172 (ProPlus)	28
FGA [8]	F NED-aaataaaataggcatatttacaagc R gctgagtgattgtctgtaattg	M64982	4q28	141–282	214–355 (ProPlus)	73
D13S317 [8]	F NED-tctgacccatctaagccta R cagacagaaagatagatagatgattga	AL353628	13q31.1	95–123	216–244 (Identifier)	121
D21S11 [8]	F PET-attccccagtggaatgc R ggtatagatagctgagatagatagacga	AP000433	21q21.1	155–210	185–240 (ProPlus)	30
D16S539	F PET-aagtccagatcctcgtgt R gcggttctgtgcatcctgt	AC024591	16q24.1	235–275	264–304 (PP16)	29

Table 2 Allelic frequency distribution, forensic, and paternity statistical parameters of 10 MiniSTR loci investigated in the population of central Punjab Pakistan ($n=250$)

Allele	CSF1PO	D18S51	D7S820	D2S1338	TPOX	D5S818	FGA	D13S317	D21S11	D16S539
7	0.002	—	0.046	—	0.002	0.002	—	0.008	—	—
8	—	—	0.248	—	0.41	0.008	—	0.16	—	0.06
9	0.016	—	0.088	—	0.122	0.03	—	0.076	—	0.142
10	0.254	0.008	0.22	—	0.076	0.096	—	0.084	—	0.108
11	0.308	0.018	0.242	—	0.358	0.388	—	0.244	—	0.282
12	0.348	0.094	0.14	—	0.03	0.33	—	0.304	—	0.242
13	0.064	0.092	0.016	—	0.002	0.136	—	0.098	—	0.134
14	0.008	0.27	—	—	—	0.01	—	0.026	—	0.028
15	—	0.204	—	—	—	—	—	—	—	0.004
16	—	0.126	—	0.008	—	—	—	—	—	—
17	—	0.084	—	0.116	—	—	—	—	—	—
18	—	0.038	—	0.142	—	—	0.008	—	—	—
19	—	0.04	—	0.16	—	—	0.032	—	—	—
20	—	0.01	—	0.126	—	—	0.13	—	—	—
21	—	0.012	—	0.036	—	—	0.148	—	—	—
21.2	—	0.004	—	—	—	—	0.008	—	—	—
22	—	—	—	0.046	—	—	0.16	—	—	—
22.2	—	—	—	—	—	—	0.008	—	—	—
23	—	—	—	0.158	—	—	0.158	—	—	—
24	—	—	—	0.11	—	—	0.188	—	—	—
24.2	—	—	—	—	—	—	0.006	—	—	—
25	—	—	—	0.08	—	—	0.112	—	—	—
25.2	—	—	—	—	—	—	0.002	—	—	—
26	—	—	—	0.018	—	—	0.028	—	0.002	—
27	—	—	—	—	—	—	0.012	—	0.02	—
28	—	—	—	—	—	—	—	—	0.124	—
29	—	—	—	—	—	—	—	—	0.146	—
29.2	—	—	—	—	—	—	—	—	0.002	—
30	—	—	—	—	—	—	—	—	0.194	—
30.2	—	—	—	—	—	—	—	—	0.022	—
31	—	—	—	—	—	—	—	—	0.04	—
31.2	—	—	—	—	—	—	—	—	0.144	—
32	—	—	—	—	—	—	—	—	0.008	—
32.2	—	—	—	—	—	—	—	—	0.216	—
33	—	—	—	—	—	—	—	—	0.004	—
33.2	—	—	—	—	—	—	—	—	0.068	—
34.2	—	—	—	—	—	—	—	—	0.008	—
36.2	—	—	—	—	—	—	—	—	0.002	—
MP	0.135	0.042	0.073	0.029	0.149	0.133	0.038	0.069	0.042	0.062
PD	0.864	0.957	0.927	0.970	0.851	0.867	0.961	0.930	0.957	0.938
PIC	0.661	0.824	0.772	0.865	0.626	0.667	0.844	0.772	0.833	0.781
PE	0.447	0.636	0.629	0.637	0.315	0.447	0.715	0.577	0.683	0.486
PI	1.736	2.778	2.717	2.778	1.316	1.736	3.571	2.358	3.205	1.894
HO	0.712	0.820	0.816	0.820	0.620	0.712	0.860	0.788	0.844	0.736
HE	0.715	0.841	0.801	0.877	0.682	0.711	0.861	0.799	0.851	0.808
P	0.555	0.989	0.238	0.024	0.022	0.671	0.371	0.522	0.048	0.231

MP matching probability, PD power of discrimination, PIC polymorphism information content, PE power of exclusion, PI paternity index, HO observed heterozygosity, HE expected heterozygosity, P probability values of exact tests for Hardy–Weinberg equilibrium

Table 3 Comparison of power of discrimination and power of exclusion for the ten MiniSTR loci in the various world populations

	CSF1PO	D18S51	D7S820	D2S1338	TPOX	D5S818	FGA	DI3S317	D21S11	DI6S539	CPD	CPE	CMP
Kashmiri [11] (PD/PE)	0.881/0.342	0.958/0.739	0.931/0.629	–	0.826/0.301	0.879/0.342	0.955/0.755	0.922/0.513	0.949/0.771	0.917/0.675	0.99999999892121	0.99973977	1.07879E-10
Pakistani [12] Northern Areas	0.868/0.389	0.964/0.726	0.924/0.597	–	0.841/0.339	0.889/0.489	0.964/0.685	0.925/0.531	0.952/0.726	0.925/0.608	0.99999999938044	0.999641031	6.1955E-11
Morocco [13]	0.869/0.485	0.968/0.643	0.927/0.561	0.956/0.740	0.875/0.400	0.888/0.430	0.962/0.544	0.905/0.477	0.956/0.643	0.917/0.544	0.99999999997514	0.999721363	2.48521E-12
Bangladeshi [14]	0.862/0.454	0.946/0.669	0.928/0.603	0.967/0.736	0.873/0.469	0.897/0.543	0.967/0.723	0.925/0.623	0.944/0.649	0.937/0.638	0.99999999997977	0.999939009	2.02239E-12
Indonesian [14]	0.883/0.474	0.938/0.642	0.930/0.555	0.953/0.706	0.779/0.334	0.913/0.562	0.964/0.750	0.916/0.574	0.966/0.721	0.924/0.592	0.99999999996414	0.999912876	3.58558E-12
Nepal [15]	0.847/0.406	0.932/0.634	0.908/0.530	0.955/0.671	0.773/0.290	0.878/0.326	0.950/0.653	0.929/0.581	0.944/0.616	0.925/0.616	0.99999999983728	0.999659727	1.62719E-11
Pachtouns [16]	0.889/0.515	0.958/0.658	0.918/0.515	0.963/0.814	0.807/0.493	0.838/0.609	0.952/0.894	0.917/0.410	0.951/0.609	0.909/0.493	0.99999999992143	0.999963225	7.85628E-12
Afghanistan													
Iran [17]	0.868/0.418	0.963/0.662	0.921/0.459	0.968/0.701	0.822/0.438	0.894/0.481	0.959/0.662	0.900/0.562	0.955/0.538	0.916/0.449	0.99999999996680	0.999717639	3.32019E-12
Iraq [18]	0.871/0.322	0.963/0.781	0.908/0.410	0.962/0.578	0.811/0.579	0.875/0.314	0.956/0.528	0.911/0.379	0.947/0.512	0.903/0.361	0.99999999992064	0.999314605	7.93643E-12
Turkey [19]	0.817/0.568	0.954/0.662	0.904/0.762	0.962/0.482	0.818/0.428	0.907/0.630	0.946/0.662	0.913/0.510	0.941/0.662	0.917/0.482	0.999999999988042	0.999904198	1.19581E-11
Saudi Arabia [20]	0.86/0.444	0.959/0.721	0.904/0.543	–	0.816/0.401	0.864/0.498	0.953/0.708	0.91/0.592	0.939/0.643	0.902/0.547	0.999999999651313	0.999592735	3.48687E-10
Uygur China [21]	0.886/0.491	0.965/0.703	0.925/0.579	0.97/0.721	0.829/0.346	0.899/0.547	0.958/0.699	0.935/0.629	0.947/0.679	0.933/0.573	0.99999999998497	0.999915293	1.50309E-12
Azerbaijan [22]	0.868/0.470	0.968/0.627	0.932/0.604	0.972/0.735	0.823/0.287	0.901/0.435	0.961/0.637	0.914/0.486	0.958/0.664	0.925/0.499	0.99999999998551	0.999738558	1.4492E-12
Balmiki India [23]	0.848/0.348	0.947/0.628	0.919/0.49	0.954/0.749	0.835/0.416	0.891/0.393	0.949/0.516	0.903/0.49	0.948/0.874	0.886/0.416	0.999999999984168	0.999846233	1.58315E-11
Punjabi Pakistan	0.864/0.447	0.957/0.636	0.927/0.630	0.970/0.636	0.850/0.315	0.866/0.447	0.961/0.714	0.930/0.577	0.957/0.683	0.937/0.486	0.99999999998241	0.999823924	1.75931E-12

CPD combined power of discrimination, CPE combined power of exclusion, CMP combined matching probability

Table 4 Linkage disequilibrium P values between pairwise loci comparison in the Punjabi population of Pakistan

	D18S51	D7S820	D2S1338	TPOX	D5S818	FGA	D13S317	D21S11	D16S539
CSF1PO	0.0716	<i>0.0417</i>	<i>0.0443</i>	<i>0.0353</i>	0.0549	0.0578	0.0685	0.0632	<i>0.0449</i>
D18S51		0.0688	0.1193	0.0735	0.0855	0.1066	0.0923	0.1017	0.0700
D7S820			0.0791	<i>0.0483</i>	<i>0.0461</i>	0.0720	0.0644	0.0737	0.0630
D2S1338				<i>0.0406</i>	0.0550	0.1089	0.0794	0.0879	0.0595
TPOX					<i>0.0445</i>	0.0635	<i>0.0459</i>	0.0669	<i>0.0389</i>
D5S818						0.0787	<i>0.0477</i>	0.0773	<i>0.0455</i>
FGA							0.0773	0.1160	0.0678
D13S317								0.0792	0.0543
D21S11									0.0948

Pairwise linkage disequilibrium showed significant values at $P < 0.05$ indicated in italics

Biosystems) followed by initial denaturation for 5 min at 95 °C, 30 cycles at 94 °C for 45 s, 60 °C for 1 min, 72 °C for 45 s with final extension at 60 °C for 30 min in 9600 emulation mode. Genotyping of PCR amplicons was performed on the ABI 3130xl Genetic Analyzer, and data was analyzed using GeneMapper v3.2 program (Applied Biosystems) in comparison with MiniSTR allelic ladder which was constructed by amplification of the dilution from Identifiler™ allelic ladder (Applied Biosystems) [8] and control DNA 9947A.

The allelic frequencies, parentage, and forensic statistical parameters for ten MiniSTR loci were calculated using modified PowerStats v1.2 [9] shown in Table 2. The exact chi-square test for the Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium tests of the ten MiniSTR loci to pair up with each other were calculated using the PowerMarker V3.25 [10]. The power of discrimination (PD) ranged from 0.8509 for the TPOX locus to 0.9700 for the D2S1338 locus, and the power of exclusion (PE) ranged from 0.3155 for TPOX locus to 0.7147 for FGA locus. This panel of ten MiniSTR loci characterizes a very high value of combined power of discrimination (PD), combined power of exclusion (PE), and cumulative probability of matching (PM) equaled to 0.99999999999824, 0.999824, and 1.75931×10^{-12} , respectively shown in Table 3. Comparing with other world populations CPD, CPE, and CPM at each locus has been shown in Table 3. Linkage disequilibrium test revealed significant results at 12 pairs of locus to locus, out of 45 pairwise comparison ($P < 0.05$) indicated in italics as shown in Table 4.

Concluding the analysis, on the basis of high polymorphism, high power of exclusion, lower probability of matching, and reduced fragment size, these ten MiniSTRs have shown to be significantly revealing in the Punjabi population for regional differences and forensic relevance. This paper follows the guidelines for publication of population data requested by the International Journal of Legal Medicine [24, 25].

References

- Hummel S, Schultes T, Bramanti B, Herrmann B (1999) Ancient DNA profiling by megaplex amplifications. *Electrophoresis* 20(8): 1717–1721
- Clayton TM, Whitaker JP, Fisher DL, Lee DA, Holland MM, Weedn VW, Maguire CN, DiZinno JA, Kimpton CP, Gill P (1995) Further validation of a quadruplex STR DNA typing system: a collaborative effort to identify victims of a mass disaster. *Forensic Sci Int* 76(1):17–25
- Edwards A, Civitello A, Hammond HA, Caskey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am J Hum Genet* 49(4):746–756
- Alonso A, Andelinovic S, Martin P, Sutlovic D, Erceg I, Huffine E et al (2001) DNA typing from skeletal remains: evaluation of multiplex and megaplex STR systems on DNA isolated from bone and teeth samples. *Croat Med J* 42(3):260–266
- Wikeley JM, Zafar RR (1968) Punjabi Musalmans. Book House
- Ruitberg CM, Reeder DJ, Butler JM (2001) STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res* 29:320–322
- Rozen S, Skaletsky HJ (1998) Primer3 (primer 3 [www.cgi v 0.2](http://www.cgi.v0.2)); Code available at <http://www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi>
- Butler JM, Shen Y, McCord BR (2003) The development of reduced size STR amplicons as tools for analysis of degraded DNA. *J Forensic Sci* 48(5):1054–1064
- Tereba A (1999) Tools for analysis of population statistics. Profiles in DNA, vol 2. Promega Corporation, Madison, pp 14–16
- Liu K, Muse SV (2005) Power marker: integrated analysis environment for genetic marker data. *Bioinformatics* 21(9):2128–2129
- Rakha A, Yu B, Hadi S, Li S (2008) Genetic analysis of Kashmiri Muslim population living in Pakistan. *Legal Med* 10(4):216–219
- Rakha A, Yu B, Hadi S, Sheng-bin L (2009) Population genetic data on 15 autosomal STRs in a Pakistani population sample. *Legal Med* 11(6):305–307
- Ossmani HE, Talbi J, Bouchrif B, Chafik A (2009) Allele frequencies of 15 autosomal STR loci in the southern Morocco population with phylogenetic structure among worldwide populations. *Legal Med* 11(3):155–158
- Dobashi Y, Kido A, Fujitani N, Hara M, Susukida R, Oya M (2005) STR data for the AmpFLSTR Identifiler loci in Bangladeshi and Indonesian populations. *Legal Med* 7(4):222–226
- Ota M, Droma Y, Basnyat B, Katsuyama Y, Asamura H, Sakai H, Fukuhsima H (2007) Allele frequencies for 15 STR loci in Tibetan populations from Nepal. *Forensic Sci Int* 169(2):234–238

16. Di Cristofaro J, Buhler S, Temori SA, Chiaroni J (2012) Genetic data of 15 STR loci in five populations from Afghanistan. *Forensic Sci Int Genet* 6(1):e44–e45
17. Shepard EM, Herrera RJ (2006) Iranian STR variation at the fringes of biogeographical demarcation. *Forensic Sci Int* 158(2):140–148
18. Barni F, Berti A, Pianese A, Boccellino A, Miller MP, Caperna A, Lago G (2007) Allele frequencies of 15 autosomal STR loci in the Iraq population with comparisons to other populations from the middle-eastern region. *Forensic Sci Int* 167(1):87–92
19. Erkol Z, Tuğ A, Alakoç YD, Elma C, Büken B, Cetinyürek A, Erkol H (2007) STR data for the AmpF/STR identifier loci from an old settlement in northwestern Turkey. *Forensic Sci Int* 173:238–240
20. Alshamali F, Alkhayat AQ, Budowle B, Watson ND (2005) STR population diversity in nine ethnic populations living in Dubai. *Forensic Sci Int* 152(2):267–279
21. Dong L, Liu XX, Zhang HH, Bo R, Wei G, Zhang LS (2010) Genetic data for 15 STR loci in Uyghur ethnic group of northwest China. *Rom J Legal Med* 1:59–62
22. Nasibov E, Bulbul O, Jabrayili G, Zorlu T, Shahzad MS, Cengiz S, Filoglu G, Altuncul H, Sadixov G (2013) Allele frequencies of 15 STR loci in Azerbaijan population. *Forensic Sci Int Genet* 7(3):e99–e100
23. Ghosh T, Kalpana D, Mukerjee S, Mukherjee M, Sharma AK, Nath S, Jha GN (2011) Genetic diversity of autosomal STRs in eleven populations of India. *Forensic Sci Int Genet* 5(3):259–261
24. Parson W, Roewer L (2010) Publication of population data of linearly inherited DNA markers in the International Journal of Legal Medicine. *Int J Legal Med* 124:505–509
25. Poetsch M, Bajanowski T, Pfeiffer H (2012) The publication of population genetic data in the International Journal of Legal Medicine: guidelines. *Int J Legal Med* 126:489–490