TÜBİTAK

Turkish Journal of Medical Sciences

http://journals.tubitak.gov.tr/medical/

Research Article

Turk J Med Sci
(2018) 48: 611-614
© TÜBİTAK
doi:10.3906/sag-1801-21

SE33 locus as a reliable genetic marker for forensic DNA analysis systems

Munir Ahmad BHINDER[1,2,*], Muhammad Yasir ZAHOOR[3], Haleema SADIA[2], Muhammad QASIM[4], Rukhsana PERVEEN[2], Ghulam Murtaza ANJUM[5], Muhammad IQBAL[6], Najeeb ULLAH[1], Wasim SHEHZAD[3], Muhammad TARIQ[7], Ali Muhammad WARYAH[8]

[1]Department of Human Genetics and Molecular Biology, University of Health Sciences, Lahore, Pakistan
[2]Center for Applied Molecular Biology, University of the Punjab, Lahore, Pakistan
[3]Institute of Biochemistry and Biotechnology, University of Veterinary and Animal Sciences, Lahore, Pakistan
[4]Department of Bioinformatics and Biotechnology, Government College University, Faisalabad, Pakistan
[5]School Education Department, Government of Punjab, Sargodha, Pakistan
[6]Department of Biochemistry and Biotechnology, The Islamia University of Bahawalpur, Bahawalpur, Pakistan
[7]Division of Science and Technology, University of Education, Township, Lahore, Pakistan
[8]Department of Molecular Biology and Genetics, Medical Research Center,
Liaquat University of Medical and Health Sciences, Jamshoro, Pakistan

Received: 03.01.2018 • **Accepted/Published Online:** 15.04.2018 • **Final Version:** 014.06.2018

Background/aim: Genetic variation, an authentic tool of individual discrimination, is being used for forensic investigations worldwide. A missing result for even one out of 13–17 markers leads to an inconclusive report. Additional reliable markers are required to compensate such deficiencies. The SE33 locus has high genetic variability in different populations and is being used in forensic investigation systems in some countries. The purpose of the study was to assess the viability of use of the SE33 locus as a supportive marker for forensic DNA profiling.

Materials and methods: Amplification of the SE33 locus was performed using the PowerPlex ES Monoplex System SE33 (Promega). After genotyping 204 Pakistani individuals, different genetic and forensic parameters for the SE33 locus were studied.

Results: Genotyping of the SE33 locus revealed a total of 43 alleles including 3 novel alleles. Significant values of different forensic and genetic parameters including power of discrimination, power of exclusion, and polymorphism information content were observed.

Conclusions: Addition of the SE33 locus in forensic DNA profiling may help to produce conclusive reports where results are inconclusive due to degraded evidence samples. The SE33 locus can confidently be used for Pakistani and neighboring populations having common ancestors from Iran to Central Asia, the Middle East, India and Turkey.

Key words: STR loci, SE33, allele frequency, Pakistani population, forensic DNA analysis, Hardy–Weinberg equilibrium

1. Introduction

Forensic DNA evaluation is a powerful tool for human identification in criminal investigations and parentage analyses (1). Sometimes a degraded or compromised DNA sample is available for evaluation, which often results in dropout of the larger molecular weight loci, therefore producing an incomplete profile. Many forensic testing labs hesitate to issue conclusive reports in the case of absence of data of even a single one of 13–17 loci used for analysis. In this scenario, in addition to customary STR markers, other highly polymorphic STR markers may be useful to produce conclusive reports for such criminal or parental investigations without compromising the power of discrimination and power of exclusion. New polymorphic STR markers are therefore needed to be validated in different populations and regions. This study is an effort to augment forensic evaluation systems, especially in Pakistan and in Asia at large. This population study is important as Pakistan is a hub of different ethnic groups and shares diverged ancestors. The Pakistani population is an admixture of Central/South Asian and European, East Asian, Greek, and Caucasian ancestry.

SE33 or the human beta-actin related pseudogene (HUMACTBP2) locus has been considered a highly discriminating locus due to its high mutation rate, ensuing in multiple length variants. It is a complex STR marker comprising a basic AAAG motif with high power of discrimination in forensic DNA analysis. The degree of

* Correspondence: munirbhinder@yahoo.com

concordance of STR alleles among commercial kits is very important when comparing and analyzing STR profiles. Due to high variability definite concordance is difficult to obtain when different primer pairs are employed (2).

SE33 has been reported for hypervariability and high power of individualization in some populations, becoming a useful forensic tool in paternity testing and human identification (3–6). The SE33 locus has been added to the new European standard set of STR loci used in forensic evaluation systems (7). In the present study, the SE33 locus has been genotyped in four major subpopulations of Pakistan. This study demonstrates the viability of the use of the SE33 locus as a potential forensic STR marker to supplement the forensic evaluation system in Pakistan and neighboring Asian populations.

2. Materials and methods

2.1. Sample collection and DNA extraction

This project was based on population genetics and was approved by the institutional ethics committee. Pakistan is a federation of four provinces accommodating four main subpopulations: Punjabi, Sindhi, Balochi, and Pakhtun. After obtaining informed consent, a total of 204 blood/saliva samples were collected: 56 from the Punjabi population, 50 from the Balochi population, 45 from the Sindhi population, and 53 from the Pakhtun population. DNA from blood/saliva samples was extracted using the DNA IQ System (Promega, Madison, WI, USA). DNA was quantified by spectrophotometric analysis.

2.2. PCR amplification and genotyping of SE33 locus

Amplification of the SE33 locus was performed using the PowerPlex ES Monoplex System SE33 (Promega) kit with a GeneAmp 9600 Thermal Cycler (PE Applied Biosystems, Foster City, CA, USA) following the manufacturer's instructions. Amplicons were subjected to capillary electrophoresis using an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems) and genotypes were assigned using GeneMapper ID Software (Applied Biosystems). Alleles were assigned according to the allelic ladder of the SE33 locus.

2.3. Statistical analysis

Statistical analyses were performed using PowerStats (Promega) and GenePop (8) software. Allele frequencies, heterozygosity, power of discrimination, power of exclusion, and Hardy–Weinberg equilibrium (HWE) were estimated.

3. Results

Genotyping of the SE33 locus revealed a total of 43 alleles including 3 novel alleles in the study population. Allele frequencies, shown in the Figure, revealed that 19 was the most frequent allele, while 13 different alleles, including three novel alleles, 19.1, 20.1, and 22.1, were found infrequently. Statistical analysis revealed the SE33 locus as a reliable marker in the study population. Values of different forensic and genetic parameters of the SE33 locus in the Pakistani population are shown in Table 1. Values of power of discrimination, power of exclusion, and polymorphism information content were 0.991, 0.769, and 0.95, respectively. Power of discrimination, power of exclusion, and heterozygosity values of the SE33 locus in our population are comparable with different world populations as shown in Table 2.

4. Discussion

Pakistan, located in South Asia, is the sixth most populous country in the world with a population exceeding 200 million and an area of 796,095 km^2 (excluding Gilgit-

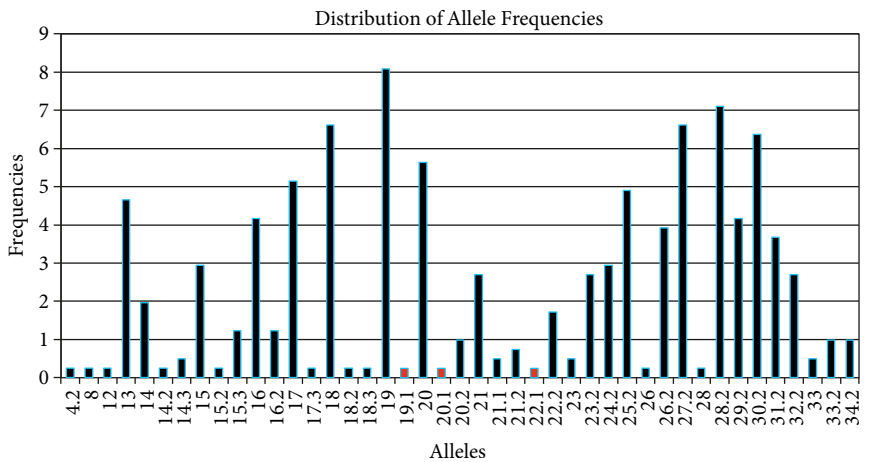

Figure. Distribution of allele frequencies for the SE33 locus in the Pakistani population (red bars indicate novel alleles).

Table 1. Forensic and genetic parameters for the SE33 locus in the Pakistani population.

	Punjabi	Balochi	Sindhi	Pakhtun	Pakistani population (combined)
Matching probability	0.023	0.025	0.032	0.022	0.009
Expressed as 1 in	43.6	40.3	31.2	44.6	113.1
Power of discrimination	0.977	0.975	0.968	0.978	0.991
Polymorphism information content	0.94	0.94	0.94	0.95	0.95
Power of exclusion	0.745	0.715	0.818	0.807	0.769
Typical paternity index	4.00	3.57	5.63	5.30	4.43
Allele frequencies					
Homozygotes	12.5%	14.0%	8.9%	9.4%	11.3%
Heterozygotes	87.5%	86.0%	91.1%	90.6%	88.7%
Total Alleles	112	100	90	106	408

Table 2. Comparison of forensic and genetic parameters for the SE33 locus between Pakistani and other populations.

Forensic parameters	Present study	German	Portuguese	Polish	Colombian	Turk	Chinese Han	Japanese	Hungarian	Moroccan
Power of discrimination	0.991	0.987	0.995	0.991	0.994	0.9773	0.993	0.993	0.9876	0.9851
Power of exclusion	0.769	0.894	0.903	0.888	0.8223	0.7991	-	-	0.8892	0.8832
Heterozygosity	0.887	0.951	0.947	0.945	0.933	0.8535	0.944	0.916	0.9458	0.8652
Hardy–Weinberg equilibrium	0.0042	0.331	Deviation from HWE	-	0.0631	0.058	-	-	0.064	0.052

Baltistan and Kashmir). The country has a warmer climate during most parts of the year (9). This region has ancient multicultural history from the Neolithic to the Bronze Age Indus Valley Civilization and remained under occupation of rulers of different cultures and faiths, including Hindus, Indo-Greeks, Muslims, Turco-Mongols, Afghans, and Sikhs.

Pakistan is bordered by India in the east, China in the northeast, Afghanistan in the west, and Iran in the southwest while in the north it is separated by a narrow strip of Afghanistan from Tajikistan. It also has a marine border with Oman (http://www.pbs.gov.pk). The Pakistani population has several ethnic groups, generally based on provincial boundaries, i.e. Punjabi, Sindhi, Balochi, and Pakhtun. These subpopulations have strong relations and historical ties with their respective bordering countries. The Punjabi population shares a lineage with the Indian Punjab and has a mixed ancestry of different ethnic groups, as most of the invaders of the Indian subcontinent traveled through Punjab, mixing with inhabitants (10). The Balochi population shows ancestry of an Aleppo population that migrated via Iran into Pakistan. The lineage history of the ancient Sindhi population shows links with Greeks and Hindus. Later on, some other ethnic groups mixed due to the settlement of Arabs, Turks, and Persians after the arrival of Islam in this area. The Pakhtun population residing in Khyber Pakhtunkhwa shares ancestry, ethnicity, and geography with the Afghan population.

The values of power of discrimination, power of exclusion, and heterozygosity for the SE33 locus have revealed this locus as a significant and valuable marker for forensic DNA analysis. The SE33 locus has also been studied in many European (German, Portuguese, Polish, and Hungarian), some Asian (Turk, Chinese Han, and Japanese), African (Moroccan), and American (Colombian) populations and has been found to be pertinent for forensic use. Power of discrimination for the SE33 locus in the Pakistani population is higher than other loci already studied in this population (11,12). Similarly, the power of exclusion and heterozygosity for the SE33 locus are also observed to be higher than those of the other loci being used in forensic investigation in Pakistan. High values of discrimination power, exclusion power, and heterozygosity for the SE33 locus may be helpful in forensic DNA analysis systems of Pakistan and neighboring countries. The use of the SE33 locus is also

significant in the region due to the strategic importance of this region in the war against terrorism.

The HWE exact test value for the SE33 locus in the Pakistani population is 0.0042. A p-value less than 0.05 indicates that the population is not in HWE. Mutation, migration, genetic drift, and nonrandom mating between individuals of a population are the factors usually affecting HWE. The deviation of the Pakistani population from HWE may be due to the presence of varied subpopulations in Pakistan. These subpopulations prefer consanguineous marriages or marriages within their own caste and clan, avoiding mating with other subpopulations. A second possible reason for deviation from HWE may be the fact that Pakistan had hosted one of the largest refugee populations in the world for about three decades. The Great Migration, resulting from the 1947 partition of British India, may also be linked to this deviation.

Higher values of power of discrimination, power of exclusion, and heterozygosity of the SE33 locus and a smaller value of matching probability show the substantial potential of SE33 to be added to forensic investigation systems in Pakistan and neighboring populations. Addition of the SE33 locus in forensic DNA analysis systems can provide additional information of different forensic parameters to compensate incomplete profiles generated due to an inadequate quantity of DNA in degraded evidence samples. This additional information may help to produce conclusive reports in those cases where a report is inconclusive due to deficiency of data of one or two loci. Therefore, SE33 can be designated as an imperative marker to aid forensic investigation systems in Pakistan and neighboring ancestor-sharing populations in Iran, Central Asia, the Middle East, India, and Turkey.

Acknowledgment
We are grateful to the Higher Education Commission of Pakistan for funding Project # SRGP 234 as a Startup Research Grant.

References

1. Daud S, Shahzad S, Shafique M, Bhinder MA, Niaz M, Naeem A, Azam M, Rehman Z, Husnain T. Optimization and validation of PCR protocol for three hypervariable regions (HVI, HVII and HVIII) in human mitochondrial DNA. Adv Life Sci 2014; 1: 165-170.
2. Davis C, Ge J, King J, Malik N, Weirich V, Eisenberg AJ, Budowle B. Variants observed for STR locus SE33: a concordance study. Forensic Sci Int-Genet 2012; 6: 494-497.
3. Lászik A, Sótonyi P, Rand S, Hohoff C. Frequency data for the STR locus ACTBP2 (SE33) in eight populations. Int J Legal Med 2001; 115: 94-96.
4. Schmid D, Anslinger K, Rolf B. Allele frequencies of the ACTBP2 (SE33), D18S51, D8S1132, D12S391, D2S1360, D3S1744, D5S2500, D7S1517, D10S2325 and D21S2055 loci in a German population sample. Forensic Sci Int-Genet 2005; 151: 303-305.
5. Paredes M, Laverde L. Analysis of the hypervariable microsatellite locus HUMACTBP2 (SE33) in Colombian population. In: 16th International Symposium on Human Identification; 29 September 2005; Grapevine, TX, USA. Poster #23.
6. Butler JM, Hill CR, Kline MC, Duewer DL, Sprecher CJ, McLaren RS, Rabbach DR, Krenke BE, Storts DR. The single most polymorphic STR locus: SE33 performance in U.S. populations. Forensic Sci Int-Genet 2009; 2: 23-24.
7. Green RL, Lagacé RE, Oldroyd NJ, Hennessy LK, Mulero JJ. Developmental validation of the AmpFℓSTR® NGM SElect™ PCR Amplification Kit: a next-generation STR multiplex with the SE33 locus. Forensic Sci Int-Genet 2013; 7: 41-51.
8. Rousset F. Genepop'007: A complete reimplementation of the Genepop software for Windows and Linux. Mol Ecol Resources 2008; 8: 103-106.
9. Bhinder MA, Arshad MW, Zahoor MY, Shehzad W, Tariq M, Shabbir MI. Junctional epidermolysis bullosa (non-Herlitz type). J Coll Physicians Surg Pak 2017; 27: 308-310.
10. Kaur M, Badaruddoza. Genetic polymorphism of five STR markers among four groups of Punjabi population in North-West Punjab, India. Int J Hum Genet 2014; 14: 1-7.
11. Tariq MA, Ullah O, Riazuddin SA, Riazuddin S. Allele frequency distribution of 13 X-chromosomal STR loci in Pakistani population. Int J Legal Med 2008; 122: 525-528.
12. Tahir MA, Herrera RJ, Khan AA, Kashyap VK, Duncan G, Barna C, Budowle B, Rowold DJ, Sinha S, Amjad M. Distribution of HLA-DQA1, polymarker, CSF1PO, vWA, TH01, TPOX, D16S539, D7S820, D13S317, and D5S818 alleles in East Bengali and West Punjabi populations from Indo-Pak Subcontinent. J Forensic Sci 2000; 45: 1320-1323.