

Study of Genetic Polymorphism at D21S11 and D21S215 Loci in the Jat Sikh Population of Punjab

I.S. Sidhu¹, K. Kaur¹, V.K. Sarhadi², D.S. Joshi³, R. Mukhopadhyaya³,
S.K. Mahajan³ and A.J.S. Bhanwer

1. Department of Human Genetics, Guru Nanak Dev University, Amritsar 143 005, Punjab India

2. Centre for Genetic Disorders, Guru Nanak Dev University, Amritsar 143 005, Punjab India

3. Molecular Biology and Agriculture Division, BARC, Trombay, Mumbai 400 085, Maharashtra, India

KEY WORDS Microsatellite; PCR; genetic polymorphism; DNA; primers

ABSTRACT Microsatellites, 1-6 bp long tandem repeats, are new class of polymorphic DNA markers and has many advantages over other DNA markers like minisatellites, VNTRs etc. due to which, these become markers of choice for studying genetic polymorphism at population level, forensic investigations, linkage studies etc. Genotyping of populations can provide genetic basis for affinities among different ethnic groups. Polymorphism revealed by microsatellites could be used to study the micro-evolutionary trends among populations. In the present study polymorphism at two microsatellites markers D21S11 (four bp repeat) & D21S215 (two bp repeat) studied among the Jat Sikh population of Punjab, India. Eight alleles of D21S11 and four alleles of D21S215 with heterozygosities 82.2% and 66.6% were observed respectively. Allelic polymorphism was compared with the data available for other Indian populations and Caucasoid population at D21S11 locus.

INTRODUCTION

The genetic information in humans is present in nucleus and mitochondria. The size of nuclear and mitochondrial genome is approximately 3×10^9 bp and 16569 bp, respectively. Almost all mitochondrial genome and <10% of the nuclear genome is coding. In earlier population biology studies, most of the research was based on phenotypes, so more emphasis was given to the coding region of the genome. The non coding DNA was considered as *junk DNA* with apparently no known function. However, with the advent of molecular genetic techniques, the term *junk DNA* has become obsolete as it has been discovered

that this non coding DNA exhibits high degree of genetic polymorphism which makes it an ideal genetic material for population genetic studies, forensic studies and for molecular diagnosis of various genetic disorders (Strachan and Read 1996).

Approximately 20% of the non coding DNA is composed of moderately to highly repetitive sequences, called *repetitive DNA*, in which repetitive elements are arranged in tandem, and are classified as satellite, minisatellite and microsatellite sequences, depending upon the size of the repeating sequence and degree of repetition in the genome. The microsatellites DNA form a significant proportion of the growing family of repetitive DNA sequences which are widely dispersed throughout the nuclear genome. If sequenced and amplified by the PCR, microsatellites may be used as highly informative, locus specific markers (Krawczak and Schmidtke 1994). These microsatellite markers have distinct advantages over other DNA and RFLP markers due to following reasons:

- i) they are easily typed by polymerase chain reaction (PCR). Thus less amount of DNA and less time is required for their typing which also reduces the cost of their analysis.
- ii) they are inherited in Mendelian co dominant fashion and thus can be easily traced back in a family or group of persons.
- iii) they exhibit extreme degree of genetic polymorphism and most of the microsatellite markers have more than 70% heterozygosity.
- iv) they are distributed throughout the human nuclear genome.

A large amount of data on traditional genetic markers for various Indian populations is available but scanty data is available on the polymorphism of microsatellite markers for Indian subcontinent. However, this type of baseline data is necessary to carry out medical genetic, population genetic and

*Corresponding Author and Address: Dr. A. J. S. Bhanwer, Professor and Head, Department of Human Genetics Guru Nanak Dev University Amritsar 143 005, Punjab, India
Phone: 0183-2258802-09 Ext. 3278
Fax: 0183-2258820 E-mail: ajsbhanwer@yahoo.com

forensic investigations using modern molecular genetic techniques. Keeping this in mind, an attempt to study DNA polymorphism in Jat Sikhs, a distinct endogamous group of India, was undertaken at two microsatellite markers of HC 21 i.e. D21S11 and D21S215.

MATERIALS AND METHODS

Markers Studied: *D21S11* is localised at 21q21, which is AT rich but gene poor region of the chromosome 21. It is 64cM apart from *D21S19* in the linkage map based on the assumption of equal recombination frequency in males and females but 43cM in male linkage map and 90cM in female linkage (Tanzi et al. 1988). This is a tetranucleotide (TCTA)_n repeat (Sharma and Litt 1992) and the exact sequence is: (TCTA)₄(TCTG)₆(TCTA)₃TA(TCTA)₃TC(ATCT)₁₀.

Thus it is evident from the sequence that *D21S11* is a complex microsatellite marker i.e. mixture of perfect and compound type of repeats (Webber 1990).

The other microsatellite marker *D21S215* is present on pericentric region of chromosome 21. It is a dinucleotide marker having repeat unit (GT)_n (Warren et al. 1992). Recombination frequency between *D21S215* and alphoid sequences of centromere is 0% (Peterson et al. 1992).

Population Studied: The present investigation was carried out on DNA samples of Jat Sikhs who were selected randomly from the state of Punjab. Approximately 60% of the Sikh population in Punjab consists of Jat Sikhs, thus making them the single largest endogamous group in Punjab. They are endogamous at caste level and exogamous at the (gotra) sub-caste level. Historical references note their presence here from 6th and 7th centuries AD. Major Todd classified Jats as one of the great Rajput tribes but General Cunningham held the view that the Rajput belonged to the original Aryan stock whereas the Jats belonged to the later wave of immigrants from the north-west, probably of Scythian race (Ibbetson 1984).

Sample Collection: 2ml of venous blood sample from 40 Jat Sikh individuals was collected in sterilised tubes containing 0.5 M EDTA as an anti coagulant. Proper care was taken to obtain blood samples of only unrelated healthy individuals. A brief family history of the individuals was asked to rule out any adoptions and the presence of any genetical disorder in the family. The blood samples were transported on ice to the laboratory and stored at -20°C till analysis.

Laboratory Analysis: The protocol of Miller et al. (1988) with slight modifications was used to extract DNA from these samples. The quality and quantity of DNA was checked by agarose electrophoresis and spectrophotometrically. The samples were then amplified by PCR using locus specific primers for these microsatellite markers. 50 ng of DNA was used for every PCR reaction.

The primers used for *D21S11* marker were same as were reported by Sharma and Litt (1992) sequences of which were.

Primer 1: 5' GTG AGT CAA TTC CCC AAG 3'

Primer 2: 5' GTTGTATTAGTCAATGTTCTCC3'

Following set of primers was used for PCR amplification of *D21S215* marker:

Primer 1: 5'-TCTAAAACA GTG TGT CTA GC-3'

Primer 2: 5'-GCT GAC GTG ACA GTT GTG AG-3'

One of the two primers, for each marker, was end labelled with ³²P γ - ATP. The labelled primer was added to the reaction mix so that 1 x 10⁵ cpm of radioactivity was present for every 20 μl of PCR cocktail. PCR was carried out on *Perkin-Elmer Thermal Cycler*.

PCR conditions consisted of initial denaturation step of 5 min at 94°C followed by 5 cycles at 94°C for 1 min; 61°C for 1 min; 72°C for 1 min; and next 25 cycles at 94°C for 40 sec; 61°C for 40 sec; 72°C for 40 sec. The last extension was done at 72°C for 5 minutes.

The PCR programme used for *D21S215* marker was same as that for *D21S11* marker except that the annealing temperature was 51°C.

The amplified gene products were loaded on 0.4 mm thick 5% denaturing polyacrylamide gel for typing various alleles of these markers. The gel was pre-run for 30 min to remove impurities. The PCR products were also denatured at 94°C for 5 minutes and then immediately ice-chilled and then promptly loaded. The ATGC sequencing ladder was also loaded at regular intervals of 6 wells in the gel so as to size the alleles which could differ by even a single base pair. *USB Sequenase-2 PCR product sequencing kit* was used to prepare sequencing ladder which was labelled with ³⁵S radioisotope. PAGE was performed at constant power of 55 watts. The gels were dried and then autoradiography was done to determine the size of various alleles of both the microsatellite markers.

Statistical Analysis: The allele frequency of all the alleles observed in the sample of population analyzed was calculated using the following formula (Bodmer and Cavalli-Sforza 1976)

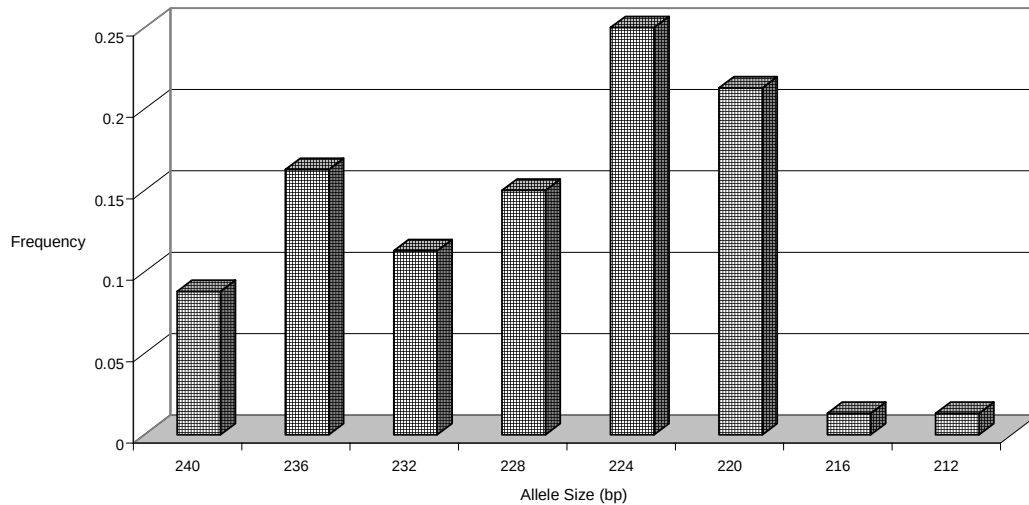


Fig. 1. Frequency distribution at D21S11 locus in Jat Sikh population

$$\text{Frequency of } i^{\text{th}} \text{ allele} = \frac{\text{Number of "i" alleles in the sample}}{\text{Total number of alleles in the sample}}$$

The observed heterozygosity was calculated by counting the number of total heterozygotes in the sample while the expected heterozygosity was calculated according to the formula (Roychoudhury 1988)

$$\text{Het.} = 1 - \sum p_i^2$$

where p_i is the frequency of the i^{th} allele.

RESULTS

In the present study, a total of 80 chromosomes

for marker D21S11 and 52 chromosomes for marker D21S215 were studied. 8 alleles of D21S11 were observed in the samples analysed with their allele size of 212 to 240 bp while for microsatellite marker D21S215, 4 alleles were observed in the size range of 166 bp-172 bp. The expected heterozygosities for D21S11 and D21S215 microsatellites markers was found out to be 82.2% and 66.6%.

Figure 1 shows the frequency of each of the 8 alleles. 4 out of 8 alleles occur at a frequency of 15% or more while allele having 232 bp with 11.3%. The most frequent allele was 224 bp allele (25%) followed closely by 220 bp allele (21.3%). Out of

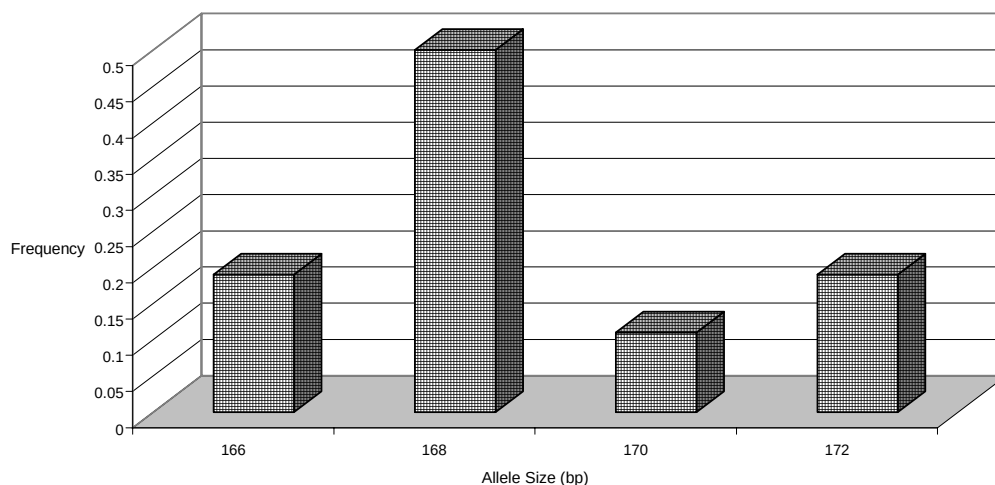


Fig. 2. Allele frequency distribution at D21S215 locus in Jat Sikh population

the 4 alleles observed for marker D21S215, the allele with 168 bp size was found out to be maximum frequent i.e. 50% and alleles with 166bp and 170 bp were present in 19% of the samples analyzed for this markers (Fig.2).

DISCUSSION

Genotyping of populations can provide genetic basis for affinities among different ethnic groups. Studies have indicated that allelic frequency distribution for hypervariable markers may vary considerably for different social and ethnic groups (Deka et al. 1995). The Indian subcontinent is a rich mixture of endogamous groups, which are maintaining their genetic identity due to high degree of endogamy. In order to fully understand the nature and extent of genetic heterogeneity of the Indian sub continent, an extensive database on different polymorphic DNA markers is needed. Although extensive data on various blood group markers, protein and enzyme polymorphic markers and HLA etc. is already available, however, only scanty data on DNA markers is available for North-West India. Earlier studies on D21S11 polymorphism (Bhanwer et al. 1997) and (Mahajan et al. 1998) in India were done on general population consisting of individuals from various endogamous groups. Thus they are not a true representation of the genetic structure for any specific endogamous group. The present study is an effort to generate a basic gene frequency distribution data at two microsatellite loci (D12S11 & D21S215) in Jat Sikhs a distinct endogamous group of Punjab.

The observed heterozygosity for D21S215 marker was 65%, which was quiet low as compared to other reported microsatellite markers. This low degree of heterozygosity at this locus may be due

Table 1: Allele size, number of alleles and heterozygosity at D21S11 locus in different populations

	Population groups studied		
	Caucasian ¹	Indian ²	Jat Sikh ³
Chromosomes studied	72	317	80
Alleles observed	12	11	8
Allele size	172-264 bp	208-248 bp	212-240 bp
Heterozygosity	0.9	0.86	0.822

¹ Sharma and Litt (1992), ² Mahajan et al. (1998),

³Present study

to the following reasons:

- i) Microsatellite D21S215 is closely linked (0cM) to the alphoid sequences of centromere of

Table 2: Allele size and gene frequency of different alleles of D21S11 locus in different population groups

Allele length(nt)	Allele frequency observed		
	Caucasian population ¹	Indian Population ²	Jat Sikh population ³
264	0.006	x	x
248	0.013	0.02	x
244	x	0.069	x
240	0.117	0.158	0.088
236	0.026	0.133	0.163
232	0.065	0.079	0.113
228	0.097	0.152	0.150
224	0.032	0.177	0.250
220	0.247	0.177	0.213
218	x	0.006	x
216	x	0.02	0.013
212	0.221	x	0.013
208	x	0.006	x
204	0.156	x	x
196	0.013	x	x
172	0.006	x	x

¹Sharma and Litt (1992) ²Mahajan et al. (1998), ³ Present study

chromosome 21 (Warren et al. 1992) and (Peterson et al. 1992). From this linkage disequilibrium, it is probable that this marker will exhibit the same behaviour as exhibited by alphoid sequences of centromere during cell division. (Marcais et al. 1990) analysed the alphoid sequences of chromosome 13 and of chromosome 21 of more than 30 unrelated individuals and found that alphoid sequences are inherited in Mendelian fashion and no meiotic recombination was detected in over 150 meioses scored in their study. Crossing over during meiosis ensures genetic uniqueness leading to increased heterozygosity and absence of crossing over to a greater extent favours homozygosity. This could be one of the reasons for low heterozygosity and an increased homozygosity of this genetic marker in our population.

- ii) Statistically due to small allele number (i.e. only 4 alleles have been observed at this locus in this population) and uneven distribution of these alleles (i.e. frequency of allele 168 bp is very high (50%) as compared to other alleles) could also be one of the factors attributing to

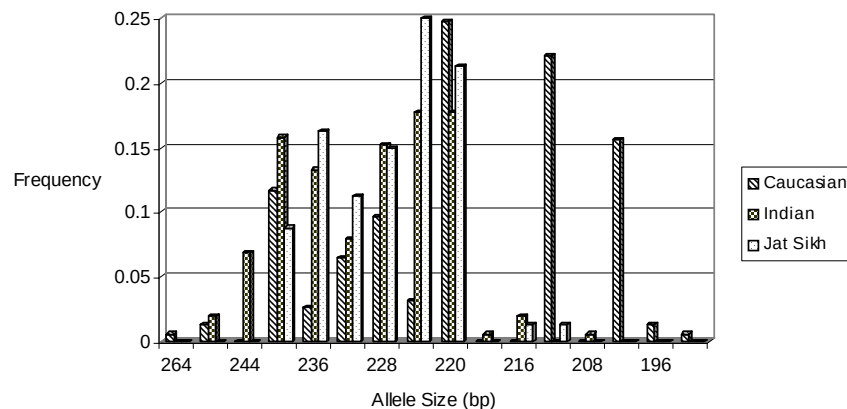


Fig. 3. Comparative allele frequency distribution at D21S11 locus

the observed homozygosity. Because with less number of alleles small number of genotypic combinations can be possible this may lead to high homozygosity.

The data generated by the present investigation at D21S215 locus could not be compared with other populations as no such report was found on population data for D21S215 even after extensive literature search.

The allelic polymorphism at the D21S11 locus was first identified in Caucasian population by Sharma and Litt in 1992. The data on number of alleles observed, allele size and heterozygosity observed at D21S11 locus is summarised in table 1. The size of various alleles at D21S11 locus alongwith their gene frequency observed in the present study and other studies is reported in table 2. The results are compared and also represented graphically in figure 3.

The comparative results indicate that these populations (Caucasian, Indian and Jat Sikh) may have their origin from a common ancestral group as all have partially overlapping frequency distribution of the alleles at this locus. It is also evident that allele at 244 bp reported by Mahajan et al. (1998) with a frequency of 6.9%, is absent both from Caucasoid population and Jat Sikh population. Similarly, the allele with 212 bp reported by Sharma and Litt (1992), is also observed in Jat Sikh population but not reported for Indian population by Mahajan et al. (1998). The allele with 218 and 244 bp, reported for Indian population, is not observed among the Jat Sikh samples studied as well in Caucasoid population studied by Sharma and Litt (1992). Absence of the allele with size 264 bp, 248 bp, 204 bp, 196 bp and 172 bp from Jat Sikh

population may indicate the absence of these alleles in the founder subgroup of Jat Sikh population. However, presence of allele having 216 bp in Indian, and Jat Sikh population may be due to some kind of genetic intermixing of these populations at some stage of their evolution.

The comparative analysis of this marker indicates that the founder subgroups of Indian populations studied by Mahajan et al. (1998) and that of Jat Sikhs may be different, but this is only a preliminary indication about the origin of Jat Sikhs, which needs to be confirmed by employing a data of large battery of microsatellite markers in Jat Sikhs and other Indian populations. If at any stage the indication provided by the present study is confirmed, this may confirm the concept of General Cunningham who held that Jats belong to a wave of immigrants that is different from the wave for populations belonging to original Aryan stock.

REFERENCES

- Bhanwer AJS, Mukhopadhyaya R, Shettigar SKG, Gangabagirathi R, Mahajan SK, Joshi DS, Singh Inderdeep, Kaur Kirandeep 1997. The allelic distribution of microsatellite loci. D21S11 and D21S215 in the Punjabi Population. XXII. Annual Conference of Indian Society of Human Genetics, Bangalore. Jan. 23-25
- Bodmer WF, Cavalli-Sforza LL 1976. Genes in population In: WF Bodmer and LL Cavalli-Sforza (Eds.): *Genetics, Evolution and Man*. San Francisco: W.H. Freeman and Company.
- Deka R, Shriver MD, Yu LM, Ferrel RE, Chakraborty R 1995. Intra- and inter population diversity at short tandem repeat loci in diverse populations of the world. *Electrophoresis*, **16**: 1659-1664.
- Ibbetson D 1984. The Jat, Rajput and allied castes. In: *Punjab Castes*. Delhi: Neeraj Publishing House. pp 97-163.

- Krawczak M, Schmidtke J 1994. *DNA Fingerprinting*. Oxford: BIOS Scientific Publishers Ltd.
- Mahajan SK, Mukhopadhyaya R, Gangabhairathi R, Mathur M, Shettigar SKG, Desai U, Joshi DS 1998. Microsatellite polymorphism on HC21 in Indian population. In: JR Singh (Ed.): *Current Concepts in Human Genetics*. Amritsar: Guru Nanak Dev University Press.
- Marcais B, Bellis M, Gerrard A, Pages M, Boublik Y, Roizes G 1990. Structural organization and polymorphism of alpha satellite DNA sequences of chromosome 13 and 21 as revealed by Pulse field gel electrophoresis. *Hum Genet*, **86**: 311-316.
- Miller SA, Dykes DD, Polesky HF 1988. A simple salting out procedure for extraction of DNA from human nucleated cells. *Nucleic Acid Res*, **16**: 1215.
- Peterson MB, Frantzen M, Antonarakis SE, Warren AC, Van Broeckhoven C, Chakravarti A, Cox TK, Lund C, Olsen B, Poulsen H et al. 1992. Comparative study of microsatellites and cytogenetic markers for detecting the origin of non disjoined chromosome 21 in Down syndrome. *Am J Hum Genet*, **51**: 516-525.
- Roychoudury AK 1988. Human population genetics In: KC Malhotra (Ed.): *Statistical Methods in Human Population Genetics*. Pune: ISHG, IIBRAD Calcutta: ISI.
- Sharma V, Litt M 1992. Tetranucleotide repeat polymorphism at D21S11 locus. *Hum Mol Genet*, **1**: 67.
- Strachan T, Read AP 1999. *Human Molecular Genetics*. Oxford: BIOS Scientific Publishers Ltd.
- Tanzi RE, Haines JL, Watkins PC, Stewart GD, Wallace MR, Hallowell R, Wong C, Wexler NS, Connolly PM, Gusella JF 1988. Genetic linkage map of human chromosome 21. *Genomics*, **3**: 129-136 .
- Warren AC, Petersen MB, Van Hul W, McInnis MG, Van Broeckhoven C, Cox TK, Chakravarti A, Antonarakis SE 1990. D21S215 is a (GT)_n polymorphic marker close to centromeric alphoid sequences on chromosome 21. *Genomics*, **13**: 1365-1367.
- Webber JL 1990. Human DNA and methods of analysis. *Curr Opin Biotechnol*, **1**: 166-171.