ज्ञान-विज्ञान विमुक्तये

UNIVERSITY GRANTS COMMISSION
BAHADUR SHAH ZAFAR MARG
NEW DELHI – 110 002.

Final Report (Feb.1, 2011-Jan.31, 2014)

of the work done on the Major Research Project

entitled

Study of genetic polymorphism of short tandem repeat (STR) loci in Punjabi population of North-west Punjab

(UGC Reference No. F.No.39-110/2010 (SR) dated Dec.27, 2010)

By

Dr. Badaruddoza
(Principal Investigator)
Department of Human Genetics

Department of Human Genetics

GURU NANAK DEV UNIVERSITY, AMRITSAR-143 005, (India)

(Established by the State Legislature Act No. 21 of 1969)

ਹਿਊਮਨ ਜਨੈਟਿਕਸ ਵਿਭਾਗ

ਗੁਰੂ ਨਾਨਕ ਦੇਵ ਯੂਨੀਵਰਸਿਟੀ, ਅੰਮ੍ਰਿਤਸਰ-143 005, ਭਾਰਤ

(ਰਾਜ ਵਿਧਾਨ ਸਭਾ ਦੇ 1969 ਦੇ ਐਕਟ ਨੰਬਰ 21 ਦੁਆਰਾ ਸਥਾਪਤ)

ਨੰਬਰ/No. 918 /HG

ਮਿਤੀ/Dated 08/04/14

To,
The Joint Secretary
University Grants Commission
Bahadur Shah Zafar Marg
New Delhi – 110 002

Dear Sir,

I am hereby enclosing two copies of the final report of the major research project entitled "Study of genetic polymorphism of short tandem repeat (STR) loci in Punjabi population of north-west Punjab" which was sanctioned by UGC. The reference number of the project was 39-110/2010 (SR) dated December 27, 2010. The duration of the project was from February 01, 2011 – January 31, 2014. The project has been successfully completed in all respects and many interesting results have been emerged. The details have been given in the report. Therefore, I deeply acknowledge UGC for giving financial support to carry out our research work and expect doing so in future also. We would be highly obliged if you acknowledge the receipt of the final project report.

With kind regards,

(Dr. Badaruddoza)
Principal Investigator, UGC Project
Department of Human Genetics
Guru Nanak Dev University
Amritsar – 143 005
Mobile - 98156-31536
Email: doza13@yahoo.co.in

Annexure - III

UNIVERSITY GRANTS COMMISSION
BAHADUR SHAH ZAFAR MARG
NEW DELHI – 110 002

STATEMENT OF EXPENDITURE IN RESPECT OF MAJOR RESEARCH PROJECT

1. Name of Principal Investigator: **Dr. Badaruddoza**
2. Dept. of University/College: **Department of Human Genetics, Guru Nanak Dev University**
3. UGC approval No. and Date: **F. No. 39-110/2010 (SR); Dec 27, 2010**
4. Title of the Research Project: **Study of genetic polymorphism of short tandem repeat (STR) loci in Punjabi population of North-west Punjab**
5. Effective date of starting the project: **1.2.2011**
6. a. Period of Expenditure: From **1.2.1011** to **31.1.2014**

   b. Details of Expenditure

S.No.	**Item**	**Amount Approved (Rs.)**	**Expenditure Incurred (Rs.)**
i.	Books & Journals	10,000	10,000
ii.	Equipment (Please enclose the quotation)	2,60,000	2,59,900
iii.	Contigency	75000	61,356
iv.	Field Work/Travel (Give details in proforma Annexure-VI)	-	-
v	Hiring services	-	
vi	Chemicals & Glassware	2,15,000	2,00,050
vii	Overhead	51,300	51,300
viii	Any other items (Please specify)	-	-
Total		**6,11,300**	**5,82,606**

c . Staff **Ms Manpreet Kaur**

Date of Appointment: **11.3.2011 (Afternoon)**

S.No.	Expenditure incurred	From	To	Amount Approved (Rs.)	Expenditure incurred (Rs.)
1.	Honarium to PI (Retired teachers) Rs. 12,000/- p.m.				
2.	Post-Doctoral Fellow Fellowship@12,000/-p.m.				
3.	Project Associate salary@Rs. 10,000/-p.m.				
4.	Project Fellow salary@Rs. 8000/- p.m.	**12.3.2011 to 31.1.2014**		**5,06,323**	**5,06,323**
Grand Total				**11,17,623**	**10,88,929**

1. It is certified that the appointment(s) have been made in accordance with the terms and conditions laid down by the Commission.

2. If as a result of check or audit objective, some irregularly is noticed, later date, action will be taken to refund, adjust or regularize the objected amounts.

3. Payment @ revised rates shall be made with arrears on the availability of additional funds.

4. It is certified that the grant of **Rs. 10,39,491/- (Rupees Ten lac thirty nine thousand four hundred and ninety one only)** received from the University Grants Commission under the scheme of support for Major Research Project entitled **Study of genetic polymorphism of short tandem repeat (STR) loci in Punjabi population of North-west Punjab** vide UGC letter **No. F. 39-110/2010** dated **27.12.2010** has been fully utilized for the purpose for which it was sanctioned and in accordance with the terms and conditions laid down by the University Grants Commission.
A final expenditure of **Rs. 10,88,929/- (Rupees Ten lac eighty eight thousand nine hundred and twenty nine only)** from total grant (Rs. 11,17,623) sanctioned has been made. An amount of **Rs. 49, 438/- (Rupees Forty nine thousand four hundred and thirty eight only)** has to be reimbursed by the UGC.

SIGNATURE OF PRINCIPAL INVESTIGATOR

REGISTRAR/PRINCIPAL
Registrar
Guru Nanak Dev University,
Amritsar.

Annexure - V

UNIVERSITY GRANTS COMMISSION
BAHADUR SHAH ZAFAR MARG
NEW DELHI – 110 002

Utilization certificate

It is certified that the grant of **Rs. 10,39,491/- (Rupees Ten lac thirty nine thousand four hundred and ninety one only**) received from the University Grants Commission under the scheme of support for Major Research Project entitled **Study of genetic polymorphism of short tandem repeat (STR) loci in Punjabi population** of **North-west Punjab** vide UGC letter **No. F. 39-110/2010** dated **27.12.2010** has been fully utilized for the purpose for which it was sanctioned and in accordance with the terms and conditions laid down by the University Grants Commission.

A final expenditure of **Rs. 10,88,929/- (Rupees Ten lac eighty eight thousand nine hundred and twenty nine only)** from total grant (Rs. 11,17,623) sanctioned has been made. An amount of **Rs. 49,438/- (Rupees Forty nine thousand four hundred and thirty eight only)** has to be reimbursed by the UGC.

SIGNATURE OF THE INVESTIGATOR

REGISTRAR/PRINCIPAL
Registrar. (Seal)
Guru Nanak Dev University,
Amritsar.

DEPUTY-PRINCIPAL REGISTRAR (ACCOUNTS)
(Seal)
Deputy Registrar (Accounts)
Guru Nanak Dev University,
Amritsar.

Annexure-VI

PROFORMA FOR SUPPLYING THE INFORMATION IN RESPECT OF THE STAFF APPOINTED UNDER THE SCHEME OF MAJOR RESEARCH PROJECT

UGC F. No. 39-110/2010 (SR) **YEAR OF COMMENCEMENT** 01-02-2011

TITLE OF THE PROJECT:

1.	Name of the Principal Investigator	**Dr. Badaruddoza**				
2.	Name of the University/College	**Guru Nanak Dev University, Amritsar**				
3.	Name of the Research Personnel Appointed	**Manpreet Kaur**				
4.	Academic qualification	S.No.	Qualification	Year	Marks	%age
		1.	**M.Sc**	**2010**	**1286/1800**	**71.4%**
		2.	**Pre-Ph.D course work**	**2010**	**CGPA-9.94/10.00**	**88.47%**
5.	Date of joining	**11.03.2011 (Afternoon)**				
6.	Date of Birth of Research Personnel	**Nov. 24, 1986**				
7.	Amount of HRA, if drawn	nil				
8.	Number of Candidate applied for the post	**Nine**				

CERTIFICATE

This is to certify that all the rules and regulations of UGC Major Research Project outlined in the guidelines have been followed. Any lapse on the part of the University will liable to terminate of the said UGC project.

Principal Investigator

Head of the Deptt.

Head
Department of Human Genetics
Guru Nanak Dev University
Amritsar-143005 (India)

Registrar/Principal

Registrar
Guru Nanak Dev University
Amritsar

Month-wise and year-wise detailed statement of expenditure towards salary of staff appointed under the project

Month &Year	Expenditure Statement
	Salary (Rs.) @ revised rates
12/3/2011 to 31/1/2012	
March 12, 2011-March 30, 2011	5161+3871= 9032
April, 2011	14,000
May, 2011	14,000
June, 2011	14,000
July, 2011	14,000
August, 2011	14,000
September, 2011	14,000
October, 2011	14,000
November, 2011	14,000
December, 2011	14,000
January, 2012	14,000
Total	**1,49,032**
1/2/2012 to 31/1/2013	
February, 2012	14,000
March, 2012	14,000
April, 2012	14,000
May, 2012	14,000
June, 2012	14,000
July, 2012	14,000
August, 2012	14,000
September, 2012	14,000
October, 2012	14,000
November, 2012	14,000
December, 2012	14,000
January, 2013	14,000
Total	**1,68,000**
1/2/2013 to 31/1/2014	
February, 2013	14,000
March 1, 2013- March 11, 2013	4,968
March 12, 2013- March 31, 2013	10,323
April, 2013	16,000
May, 2013	16,000
June, 2013	16,000
July, 2013	16,000
August, 2013	16,000
September, 2013	16,000
October, 2013	16,000
November, 2013	16,000
December, 2013	16,000
January, 2014	16,000
Total	**1,89,291**
Grand total	**5,06,323**

SIGNATURE OF PRINCIPAL INVESTIGATOR

REGISTRAR/PRINCIPAL

Registrar
Guru Nanak Dev University
Amritsar.

Annexure-IX

UNIVERSITY GRANTS COMMISSION
BAHADUR SHAH ZAFAR MARG
NEW DELHI- 110 002

PROFORMA FOR SUBMISSION OF INFORMATION AT THE TIME OF SENDING THE FINAL REPORT OF THE WORK DONE ON THE PROJECT

1. **TITLE OF THE PROJECT:** Study of genetic polymorphism of short tandem repeat (STR) loci in Punjabi population of North-west Punjab

2. **NAME AND ADDRESS OF THE PRINCIPAL INVESTIGATOR:** Dr. Badaruddoza

3. **NAME AND ADDRESS OF THE INSTITUTION:** Guru Nanak Dev University, Amritsar

4. **UGC APPROVAL LETTER NO. AND DATE**: 39-116/2010 (SR)

5. **DATE OF IMPLIMENTATION:** 01.02.2011

6. **TENURE OF THE PROJECT:** 3 years

7. **TOTAL GRANT ALLOCATED:** Rs. 11,17,623

8. **TOTAL GRANT RECEIVED:** Rs. 10,39,491

9. **FINAL EXPENDITURE:** Rs. 10,88,929

10. **TITLE OF THE PROJECT:** Study of genetic polymorphism of short tandem repeat (STR) loci in Punjabi population of North-west Punjab

11. **OBJECTIVES OF THE PROJECT:** The aims and objectives of the present study are:

(i) To analyze the genetic structure, relationship and gene flow in five endogamous population groups from the north-west districts of Punjab, based on variations of six STR loci (THO1, TPOX, CSF1PO, vWA, D7S820 and FGA).
(ii) To assess usefulness of STR markers in addressing genetic affinities and relationships among closely related populations.
(iii) To find out the genetic similarity and phylogenetic position of Punjabi population with respect to past history of admixture of foreign populations, especially, Caucasoid populations.
(iv) Besides above objectives, this study may generate powerful adjunct for further regional and worldwide population studies through meta-analysis.

12. **WHETHER OBJECTIVES WERE ACHIEVED:** Yes, all the objectives of the present study have been successfully achieved. The detailed study has been discussed below:

Introduction

India has served as a major corridor for the dispersal of human beings that started from Africa about 100,000 years ago (Mountain *et al*., 1995). Ethnic Indian population shows enormous social, cultural, linguistic, and genetic diversity due to its positioning on the crossroads of many historic and pre-historic human migrations nurtured to a large extent by the varied topography of the country (Cordaux *et al.*, 2004; Thanseem *et al.*, 2006; Nair *et al.*, 2011). Punjab has been considered the gateway of India from ages. Invaders like Alexander, Aryans, Persians, Greeks, Scythians, Parthians, Huns, Turks, Mongols, Afghans and Mughals came to India through North-West Punjab. Most of the successive invaders inroad into Punjab and settled here permanently, adjusted themselves to the new socio-cultural system and adopted the customs and traditions of the local land. Therefore, they merged into the indigenous population in process and became part of the culture of the Punjab. The present Punjabi population is the descendants of the various racial stocks which entered into it during the different stages of its history (Sekhon, 2000). Punjabi population possesses an exclusive genetic profile, primarily due to the many migratory events in this region which caused an extensive range of genetic diversity. Hence, the present study is an attempt to understand the effect of these influences on genetic differentiation, diversity and population structure of Punjab. In this study the genetic polymorphism of microsatellite loci in five Punjabi population groups of north-west Punjab was investigated and it was used to locate phylogenetic position of Punjabi population with other Indian and world populations, as well as, to explore the genetic relationships between various caste groups and their impact on gene pool. In this study six microsatellite markers: THO1, TPOX, CSF1PO, vWA, D7S820 and FGA have been analyzed.

Methodology

Sample size and design: A total of 516 samples were collected through simple random sampling method from five endogamous population groups of north-west border districts of Punjab such as Amritsar, Fazilka, Ferozepur, Gurdaspur, Pathankot, Tarn Taran and Malerkotla. The distribution of collected samples was 112 Jat Sikhs (higher caste of Sikh religion), 105 Mazhbi Sikhs (lower caste of Sikh religion), 89 Brahmins (higher caste of Hindu religion) and 92 Ramdasias/Valmikis (lower caste of Hindu religion) and 118 Muslims. A detailed questionnaire was prepared for the present study which included the family history of the individual along with other relevant details. An informed consent was duly signed by the subject taken.

Collection of blood samples: After the informed consent of the patient, 3 ml of venous blood was drawn with the help of a sterile disposable syringe and was immediately transferred to pre-labelled blood collecting vial containing 0.5M EDTA (as anticoagulant). All samples were transported on ice from the place of sample collection to laboratory and were stored at -20°C till further analysis.

DNA Extraction, PCR Amplification and STR genotyping: High molecular weight genomic DNA of collected samples was extracted using Phenol-chloroform method (Sambrook *et al.*, 1989) and was purified by ethanol precipitation. The genomic DNA was quantified using Nanodrop (Thermo Scientific™ 2000c, Pittsburgh, US). In the present study six STR markers were selected: THO1, TPOX, CSF1PO, vWA, D7S820 and FGA on the basis of their heterozygosity and their feasibility for rapid analysis (through PCR) in lab. Selection of the present STR markers has also been based on the global survey carried out by Perez-Lezaun *et al.* (1997). The PCR amplification was done on Mastercycler ® Personal (Eppendorf, Germany) using locus specific primer pairs (table 1). PCR program used was initial denaturation at 95°C for 5 min followed by 35 cycles of 94°C for 45 sec, 58°C for 45 sec, 72°C for 45 sec, followed by final extension at 72°C for 10 min for THO1, TPOX and CSF1PO loci done in a multiplex. Same protocol was used for other three loci also except annealing step which was 56°C for vWA, 53°C for D7S820 and 50°C for FGA. The PCR reaction mixture of 15 µl consisted of 50 ng of DNA, Taq polymerase buffer, 200 µmol/l dNTPs, and 1.0 U Taq polymerase and 5 pmol of each primer. PCR products were checked on 2.5% agarose gel and then analyzed on 10% native polyacrylamide gel electrophoresis (PAGE) (Maniatis *et al.* 1989) and various alleles were detected after proper silver staining.

Table 1**: STR markers along with the primer sequences used to amplify them and their respective PCR product size range.**

STR marker	Primers	Product size (bp)
THO1	For: ATTCAAAGGGTATCTGGGCTCTGG Rev: GTGGGCTGAAAAGCTCCCGATTAT	171-215
TPOX	For: ACTGGCACAGAACAGGCACTTAGG Rev: GGAGGAACTGGGAACCACACAGGTTA	216-264
CSF1PO	For: AACCTGAGTCTGCCAAGGACTAGC Rev: TTCCACACACCACTGGCCATCTTC	287-331
vWA	For: CCCTAGTGGATGATAAGAATAATC Rev: GGACAGATGATAAATACATAGGATGGATGG	122-182
D7S820	For: TGTCATAGTTTAGAACGAACTAACG Rev: CTGAGGTATCAAAAACTCAGAGG	198-234
FGA	For: GCCCCATAGGTTTTGAACTCA Rev: TGATTTGTCTGTAATTGCCAGC	172-212

Statistical Analysis: Various statistical analyses were performed on the STR data to accomplish the objectives of the study. The frequency of alleles for each STR was calculated for each genotype in the sample set through gene counting method. Average heterozygosity for each population were calculated through Nei's formula (Nei, 1973). The possible deviation from Hardy-Weinberg expectation (HWE) was determined by Fisher exact test and Likelihood ratio test using ARLEQUIN ver. 3.1. Genetic distance was estimated with the help of PHYLIP ver. 3.9 (Nei *et al.* 1983). The phylogenetic tree was drawn by Neighbor Joining (NJ) method using POPTREE ver. 2. Locus informativeness parameters, such as genotype match probability in two unrelated individuals (MP), power of discrimination between genotypes of two unrelated indivuduals (PD), and polymorphism information content (PIC) was calculated using PowerStats ver. 12 software (Promega). In order to provide a comprehensive picture of genetic similarities and differences of the studied five north-west Punjabi population groups, different comparisons with national and world-

wide data was also done through meta-analysis. An assessment of individual clustering was performed using the STRUCTURE ver 2.3.4.

Results

Allele frequencies in the population

The allele frequencies of THO1, TPOX, CSF1PO, vWA, D7S820 and FGA microsatellite markers among five north-west Punjabi population groups (Jat Sikh, Mazhbi Sikh, Brahmin, Ramdasia and Muslim) have been determined and presented in table 2. The allele repeat number varied from 5 (THO1) to 26 (FGA). The alleles having the maximum frequencies are presented in bold red letters for each population. The number of alleles ranged from 8 to 12 at six STR loci. Average number of observed alleles was 9.5. Overall, the maximum number of alleles has been observed in Ramdasia (50) and minimum in Muslims (42). The spectrum of allele frequency distribution for six STR markers among five populations is presented in figure 1. The shape of the distribution of allele frequencies is fairly uniform across most of the population groups, except Muslim for most of the loci, barring the few most polymorphic ones.

The exact test probabilities for HWE have been calculated among all the studied populations and loci (table 3). The analysis suggested some significant departures in certain loci and population groups. The results are as follows: the significant departure of THO1 among Muslims, TPOX among Jat Sikh and Ramdasia, CSF1PO among Brahmins, vWA among Jat Sikh and Mazhbi Sikh, D7S820 among all the five groups, FGA among Jat Sikh, Mazhbi Sikh, Brahmin and Muslim population groups has been observed. Therefore, maximum population departure from HWE may be due to certain amount of inbreeding among these population groups and a relatively small sample size in some cases may account for this departure.

Genomic diversity between population groups

The results of gene diversity analysis are presented in table 4 separately for each locus. The results included observed (H_{obs}) and expected heterozygosity (H_{exp}) values, genomic differentiation coefficient (G_{st}), inter-population fixation index (F_{st}) and inter-population estimates of gene flow (N_m). The highest observed heterozygosity for THO1 among Mazhbi Sikhs (0.8286), TPOX among Brahmins (0.8539), CSF1PO (0.7542) and vWA (0.7797) among Muslims, D7S820 among Ramdasias (0.7935) and FGA among Brahmins (0.9326) have been observed. Similarly, the highest expected heterozygosities have been found for locus THO1 among Jat Sikhs (0.8271), TPOX among Brahmins (0.8319), CSF1PO among Brahmins (0.8453) and vWA among Jat Sikhs (0.8988), D7S820 among Ramdasias (0.83) and FGA among Brahmins (0.8497). However, all the observed and expected heterozygosities among all the populations have not been found significantly different. In general, the average observed heterozygosity is lower than expected heterozygosity in six STR markers among the five population groups. In the present study the average sub-ethnic genomic differentiation (F_{st}) among five population groups of northwest Punjab was 0.0335. The CSF1PO has showed highest sub-ethnic differentiation (0.0649), whereas, the lowest (F_{st}) has been observed for FGA locus (0.013). The percentage of genomic diversity attributable to different populations relative to the total genomic diversity (G_{st}) varies between 6.1% for CSF1PO locus and 0.9% for D7S820. When all the loci are jointly considered, 3.0% of the total genomic diversity is attributable to the five population groups. The joint effects of gene flow among all these population groups with respect to different STR markers have also been calculated. The maximum gene flow (N_m=18.77%) has been observed in FGA, which is followed by THO1 (15.92%). The lowest amount of gene flow has

been observed in CSF1PO (3.6%). However, in general with all loci the gene flow has been observed to be 7.2% among these studied population groups.

Informativeness of the selected STR system

The informativeness of the six microsatellite markers have been presented in table 5. It includes the matching probability (MP), power of discrimination (PD), polymorphism information content (PIC), exclusion potential (PE) and paternity index (PI) of the six markers. The high level of intra-populational diversity resulted in a considerable discriminating power of this STR marker system. The PD values for almost all the markers lie in the range of 0.880-0.949. However, the highest power of discrimination has been observed for vWA (0.949) among Jat Sikh population group. The highest PE was observed for locus FGA (0.862) among Brahmin population group and lowest PE was observed for vWA (0.188) for Ramdasia population group. All the six markers (THO1, TPOX, CSF1PO, vWA, D7S820 and FGA) were highly polymorphic (PIC>0.5). In the total population sample there were no two individuals with the same genotype of the six marker system. Therefore, this marker system actually meets all the existing requirements to identification system.

Nei's Standard Genetic Distance among Five Population Groups

The pairwise comparison of genetic distances among five population groups is presented in table 6. The results showed the Brahmins (higher caste of Hindu religion) are more close to the Mazhbi Sikhs (lower caste of Sikh religion). However, it shows significant differences between Muslims and other groups, Ramdasia and Jat Sikh, Ramdasia and Brahmins.

Phylogenetic Reconstruction

To study the genetic relationship and differences existing among Indian populations, a population database of 5 Punjabi populations was compiled. Phylogenetic analysis was carried out for Jat Sikh, Mazhbi Sikh, Brahmin, Ramdasia and Muslim population groups. The phylogenetic assessments were made based on the allele frequency profile of 6 STR markers (THO1, TPOX, CSF1PO, vWA, D7S820 and FGA) studied in all 5 groups of northwest Punjabi population. A phylogram (figure 2) was generated using Neighbor Joining (NJ) method. Punjabi Mazhbi Sikh, Ramdasia and Brahmin group form a distinct cluster clearly separated from Jat Sikh and Muslim.

The present study has been compared with 26 Indian populations to get a more comprehensive picture of genetic similarity and differences existing between Indian populations. The criterion of selection was to cover the major geographical and ethnic groups from northern, southern, eastern and western India. Therefore, phylogenetic analysis was carried out in 31 populations (table 3). The phylogenetic assessments were done based on allele frequency profile of six STR markers studied in all 31 populations. The six STR loci included for the analysis are THO1, TPOX, CSF1PO, VWA, D7S820 and FGA. All present study populations except Muslims have clustered in one group, whereas, presently studied Muslim group clustered with South Indian Tamil populations. All the populations analyzed in the present study from different states formed separate and distinct clusters. Orissa, the eastern part of country along

with upper caste formed a distinct cluster. Similarly, the populations of Maharashtra have formed a separate cluster. One cluster was formed with the mixture of all populations from different regions. The important aspect of this phylogram suggested a high level of genetic diversity in different caste populations of India. The present study has also been compared with 14 world populations, which have been categorized into four major continents. The phylogram was constructed based on variance of allele frequency distribution of 6 STR markers (figure 4). The results reveals the present five population (Jat Sikh, Mazhbi Sikh, Brahmin, Ramdasia, Muslim) formed one distinct cluster, whereas, other continental populations have formed distinct and separate clusters.

To understand the extent of sub-structuring among five northwest Punjabi population, the present study has performed STRUCTURE analysis with different values of K. Simulation summary for K=2 and K=3 including the logarithm of estimated probability of data, the values of proportion of membership for each pre-defined five population groups in each of two or three clusters at the corresponding α values are given in table 7. Similarly, the pattern of sub-structuring among five groups of studied population is depicted in figure 5. The log probability values and the membership proportion of each group showed clear sub-structure among the population groups. In case of K=2 and K=3, Muslim group showed higher proportion of membership (0.677 and 0.397) to cluster 1. Ramdasia showed higher proportion of membership in cluster 2 with respect to K=2 (0.650) and K=3 (0.377), respectively. In case of cluster 3 of K=3, Mazhbi Sikh showed high values of membership proportion (0.362).

Discussion

All these five studied populations were settled in northwest Punjab since several generations. They showed socio-cultural as well as linguistic homogeneity (Sekhon, 2000). There are no such studies among these populations with respect to STR markers to investigate the phylogenetic position of these populations. This is perhaps the first comprehensive molecular genetic study attempted to investigate the genetic affinity and diversity among these closely related population groups in north-west Punjab.

The results of allele frequency variations at six STR loci revealed that there is underlying microsatellite diversity among all these population groups. The content of deviation of studied loci from Hardy-Weinberg equilibrium among five studied population groups at the sub-structural level is indicative of their unique population genetic structure. For instance, Jat Sikhs and Muslims deviate from HWE at maximum number of loci (4 and 3, respectively). This scenario in case of Jat Sikhs and Muslims is probably due to their endogamous nature of marriage structure or smaller sample size.

The least average heterozygosity values among Jat Sikh (58.93%) for CSF1PO locus might be explained by their smaller sample size, strategic location and preferential marriage practice among the clans prohibiting external gene flow (Koley et al., 2007; Krithika et al., 2008; Noor et al., 2009). Interestingly the maximum average heterozygosity value (93.26%) was observed among Brahmin population for FGA locus. However in general, the exact test of population differentiation for STR markers showed moderate diversity among all five studied groups. The least significant difference between Mazhbi Sikh (lower caste of Sikh religion) and Brahmin (higher caste of Hindu religion) might be due to the fact that these two populations have formed from single population in earliest settlement of this region. However, this requires further validation (Eaaswarkhanth et al., 2010; Chaubey et al., 2011).

The ethno-historical backgrounds of these study population groups are similar and developed through fission-fusion process of inter-caste groups and separation of this group has been done in recent past. This hypothesis is also supported by low average Gst values with respect to all STR markers (0.09% to 0.11%). This is an indication of low degree of genetic differentiation among these groups. This low degree of differentiation among these groups is further substantiated by the clustering pattern obtained from STRUCTURE analysis. A close cluster of Mazhbi Sikh and Brahmins, observed in phylogenetic tree corroborates with their geographic proximity and origin of these two groups. The separation of Ramdasia, Mazhbi Sikh and Brahmins in the phylogeny despite belonging to the same ancestral group could be the consequence of migration of a few close kin-groups from their ancestral population (Majumder et al., 1999; Gaikwad et al., 2006; Chaubey et al., 2011). Overall, the low average Gst values, close clustering in STRUCTURE analysis, phylogeny and low Fst values support the recent formation of all five studied groups from a common ancestral group (Misra, 2001; Nair et al., 2011). However, this observation needs to be speculated as the increase in the number of samples and microsatellite loci may affect the observation (Krithika et al., 2008).

India is supposed to have arrival in several waves of migration during different time periods from eastern-, sourthern- and central-Asian regions and have localized in various regions of this sub-continent (Watkins et al., 2005; Gaikwad et al., 2006; Mukherjee et al., 2009). However, migration from central Asian region entered India through northwest Punjab and settled down, which created regional genetic differences with populations of rest of India. This is also validated from the results obtained from the phylogenetic tree of 31 different Indian population groups which showed the formation of clusters based on geographic proximity of the populations. For example, all the present five population groups of northwest Punjab clustered together and showed clear genetic differentiation from other populations in India. This is also true when present population was compared with other world populations. This probably suggested common genetic affinity and origin of five studied population groups of northwest Punjab.

Conclusions

Overall, five Punjabi speaking population groups such as Jat Sikh, Mazhbi Sikh, Brahmin, Ramdasia and Muslim from northwest Punjab are regionally well differentiated and exhibit strong genetic affinity based on their origin, settlement and their shared ethno-historic background. However, a clear picture will probably emerge from the analysis of increased number of informative genetic markers and population size.

References

Chaubey, G., Metspalu, M. Choi, Y., Mägi, R., Romero, I.G., Soares, P., van Oven, M., Behar, D.M., Rootsi, S., Hudjashov, G., Mallick, C.B., Karmin, M., Nelis, M., Parik, J., Reddy, A.G., Metspalu, E., van Driem, G., Xue, Y., Tyler-Smith, C., Thangaraj, K., Singh, L., Remm, M., Richards, M.B., Lahr, M.M., Kayser, M., Villems, R. and Kivisild, T. (2011). Population Genetic Structure in Indian Austroasiatic Speakers: The Role of Landscape Barriers and Sex-Specific Admixture. *Molecular Biology Evolution* **28**(2): 1013–1024.

Cordaux, R., Aunger, R., Bentley, G., Nasidze, I., Sirajuddin, S.M. and Stoneking, M. (2004). Independent origins of Indian caste and tribal paternal lineages. *Current Biology* **14**: 231-235.

Eaaswarkhanth, M., Haque, I., Ravesh, Z., Romero, I.G., Meganathan, P.R., Dubey, B., Khan, F.A., Chaubey, G., Kivisild, T., Tyler-Smith, C., Singh, L. and Thangaraj, K. (2010). Traces of sub-Saharan and Middle Eastern lineages in Indian Muslim populations. *European Journal of Human Genetics* **18**: 354–363.

Gaikwad, S., Vasulu, T.S. and Kashyap, V.K. (2006). Microsatellite diversity reveals the interplay of language and geography in shaping genetic differentiation of diverse proto-australoid populations of west-central India. *American Journal of Physical Anthropology* **129**: 260-267.

Koley, R.N. (2006). Sociocultural Aspects of Adi Panggi. Itanagar, India: Directorate of Research,Government of Arunachal Pradesh.

Krithika, S., Maji, S. and Vasulu, T.S. (2008). A Microsatellite Guided Insight into the Genetic Status of Adi, an Isolated Hunting-Gathering Tribe of Northeast India. *PLoS ONE* **3**(7): e2549.

Majumder, P.P., Roy, B., Banerjee, S., Chakraborty, M., Dey, B., Mukherjee, N., Roy, M., Thakurta, P.G., Sil, S.K. (1999). Human-specific insertion/deletion polymorphism in Indian populations and their possible evolutionary implications *European Journal of Human Genetics* **7**: 435-446.

Maniatis, T., Sambrook, J. and Fristsch, E.F. (1989). *Molecular Cloning- A Laboratory Manual*. Cold Spring Harbor, NY, Cold Spring Harbor University Press.

Mountain JL, Hebert JM, Bhattacharyya S, Underhill PA, Ottolenghi C, Gadgil M, Cavalli-Sforza LL (1995) Demographic history of India and mtDNA-sequence diversity. *American Journal of Human Genetics* **56**: 979-92.

Mukherjee, M.B., Tripaty, V., Colah, R.B., Solanki, P.K., Ghosh, K., Reddy, B.M. and Mohanty, D. (2009). Microsatellite diversity among the primitive tribes of India. *Indian Journal of Human Genetics* **15**: 114-120.

Nair, P.S., Geetha, A. Jagannath, C. (2011). Y-short tandem repeat haplotype and paternal lineage of the Ezhava population of Kerala, south India *Croatian Medical Journal* **52**: 344-350

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of National Academy of Sciences (USA)* **70**: 3321-3323.

Nei, M., Tajima, F. and Tateno, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19**: 153-170.

Noor, S., Ali, S., Eaaswarkhanth, M. and Haque, I. (2009). Allele frequency distribution for 15 autosomal STR loci in Afridi Pathan population of Uttar Pradesh, India. *Legal Medicine* **11**: 308–311.

Perez-Lezaun, A., Calafell, F., Mateu, E., Comas, D., Bosch, E. and Bertranpetit, J. (1997). Allele frequencies for 20 microsatellites in a worldwide population survey. *Human Heredity* **47**: 189-196.

Sambrook J, Fritsch EF, Maniatis T 1989. Molecular cloning: a laboratory –manual" 2nd ed. New York, NY: Cold Spring Harbor Laboratory Press.

Sekhon, I.S. (2000). *Punjabis – The People, their History, Culture and Enterprise.* Cosmo Publication: pp 7-18.

Thanseem, I., Thangaraj, K., Chaubey, G., Singh, V.K., Bhaskar, L.V., Reddy, B.M., Reddy, A.G. and Singh, L. (2006). Genetic affinities among the lower castes and tribal groups of India: inference from Y chromosome and mitochondrial DNA. *BMC Genetics* **7**: 42.

Watkins, W.S., Prasad, B.V.R., Naidu, J.M., Rao, B.B., Bhanu, B.A., Ramachandran, B., Das, P.K., Gai, P.B., Reddy, P.C., Reddy, P.G., Sethuraman, M., Bamshad, M.J. and Jorde, L.B. (2005). Diversity and Divergence among the tribal populations of India. *Annals of Human Genetics* **69**: 680-692.

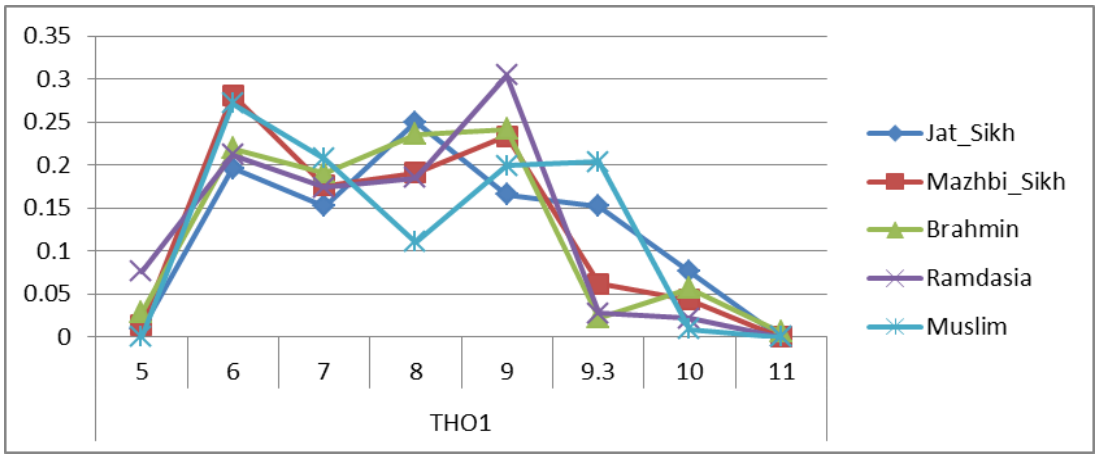
Table 2. Allele frequency data of THO1, TPOX, CSF1PO, vWA and D7S820 microsatellite markers among four north west Punjabi population groups

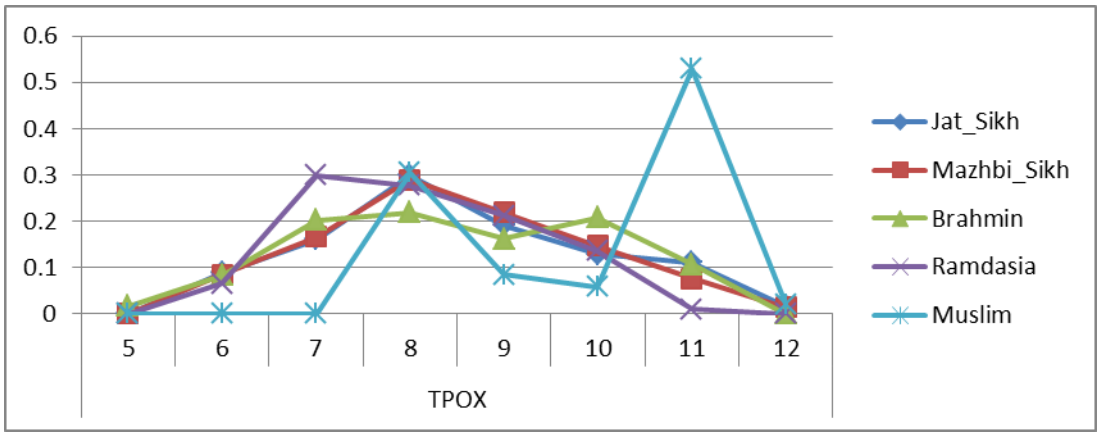
Locus		Populations				
		Jat Sikh	**Mazhbi Sikh**	**Brahmin**	**Ramdasia**	**Muslim**
THO1	5	0.0089	0.0143	0.0281	0.0761	0.0000
	6	0.1964	**0.2810**	0.2191	0.2120	**0.2712**
	7	0.1518	0.1762	0.1910	0.1739	0.2076
	8	**0.2500**	0.1905	0.2360	0.1848	0.1102
	9	0.1652	0.2333	**0.2416**	**0.3043**	0.1992
	9.3	0.1518	0.0619	0.0225	0.0272	0.2034
	10	0.0759	0.0429	0.0562	0.0217	0.0085
	11	0.0000	0.0000	0.0056	0.0000	0.0000
TPOX	5	0.0000	0.0000	0.0169	0.0000	0.0000
	6	0.0893	0.0857	0.0843	0.0652	0.0000
	7	0.1607	0.1667	0.2022	**0.2989**	0.0000
	8	**0.2991**	**0.2905**	**0.2191**	0.2772	0.3051
	9	0.1920	0.2190	0.1629	0.2120	0.0847
	10	0.1295	0.1476	0.2079	0.1359	0.0593
	11	0.1116	0.0762	0.1067	0.0109	**0.5297**
	12	0.0179	0.0143	0.0000	0.0000	0.0212
CSF1PO	7	0.0000	0.0000	0.0000	0.0163	0.0085
	8	0.0045	0.0000	0.0056	0.0109	0.0085
	9	0.0045	0.0143	0.0169	**0.3043**	0.0254
	10	0.2054	0.1000	0.1292	0.0489	0.2669
	11	0.2723	0.1952	0.1461	0.2989	0.2542
	12	**0.4464**	0.2714	**0.2135**	0.1413	**0.3983**
	13	0.0670	**0.2762**	0.1067	0.1359	0.0254
	14	0.0000	0.1048	0.1854	0.0380	0.0127
	15	0.0000	0.0381	0.1854	0.0054	0.0000
	16	0.0000	0.0000	0.0112	0.0000	0.0000
vWA	11	0.0357	0.0333	0.0000	0.0217	0.0000
	12	0.0536	0.0667	0.0112	0.0163	0.0000
	13	0.0223	0.0429	0.0112	0.0217	0.0169
	14	0.1473	0.1381	0.0899	**0.3424**	0.1271
	15	0.1250	0.0857	0.2303	0.0326	0.0720
	16	**0.1607**	0.2333	0.1798	0.1033	0.1356
	17	0.0848	**0.2905**	**0.2921**	0.1957	**0.3729**
	18	0.1027	0.0476	0.0787	0.1467	0.1398
	19	0.1071	0.0619	0.0787	0.0761	0.0636
	20	0.0804	0.0000	0.0281	0.0380	0.0720
	21	0.0759	0.0000	0.0000	0.0054	0.0000
	23	0.0045	0.0000	0.0000	0.0000	0.0000

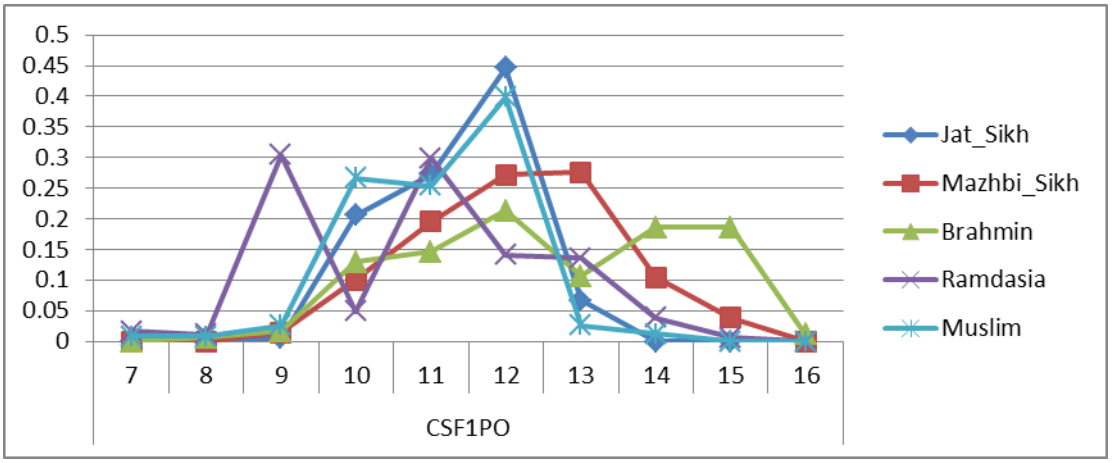
Contd. on next page

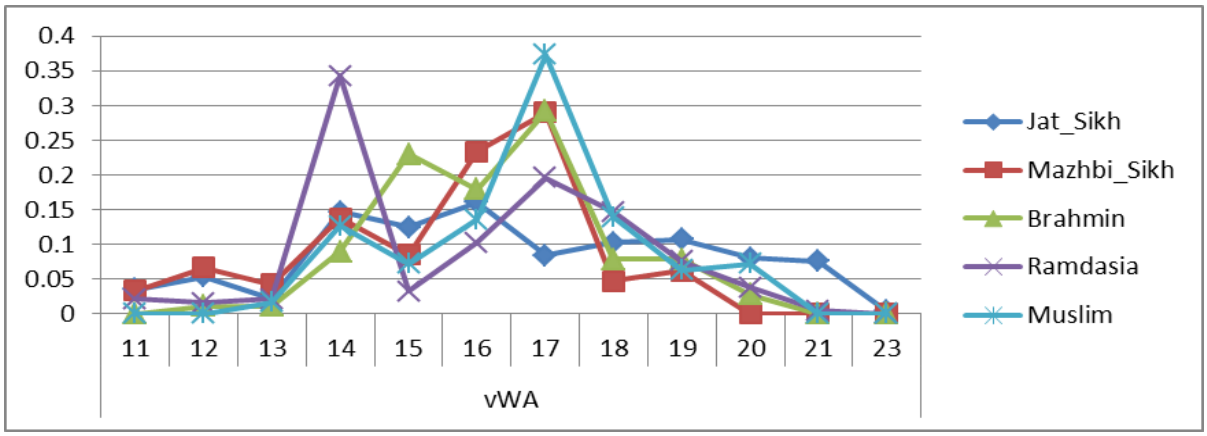
D7S820	7	0.0357	0.1190	0.0506	0.1576	0.0254
	8	0.1250	**0.2286**	0.1461	0.1359	0.1229
	9	0.0848	0.0810	0.1067	0.1467	0.1398
	10	**0.3571**	0.1952	**0.2921**	**0.2337**	**0.2924**
	11	0.1920	0.2095	0.1910	**0.2337**	0.1992
	12	0.1339	0.1524	0.1573	0.0652	0.2119
	13	0.0491	0.0143	0.0562	0.0272	0.0085
	14	0.0223	0.0000	0.0000	0.000	0.0000
FGA	16	0.0000	0.0095	0.0000	0.0109	0.0000
	17	**0.3125**	0.0095	0.0225	0.0109	0.0000
	18	0.0268	0.0381	0.0337	0.0652	0.0000
	19	0.0536	0.0238	0.0337	0.0163	0.0508
	20	0.0893	0.1143	0.1236	0.0815	0.1271
	21	0.2098	**0.2571**	0.1910	**0.2609**	0.1653
	22	0.1339	0.0952	0.1461	0.1033	0.1271
	23	0.2589	0.2095	**0.2528**	0.2065	0.1314
	24	0.1071	0.1333	0.1180	0.1304	0.0975
	25	0.0893	0.1095	0.0787	0.1141	**0.2881**
	26	0.0000	0.0000	0.0000	0.0000	0.0127

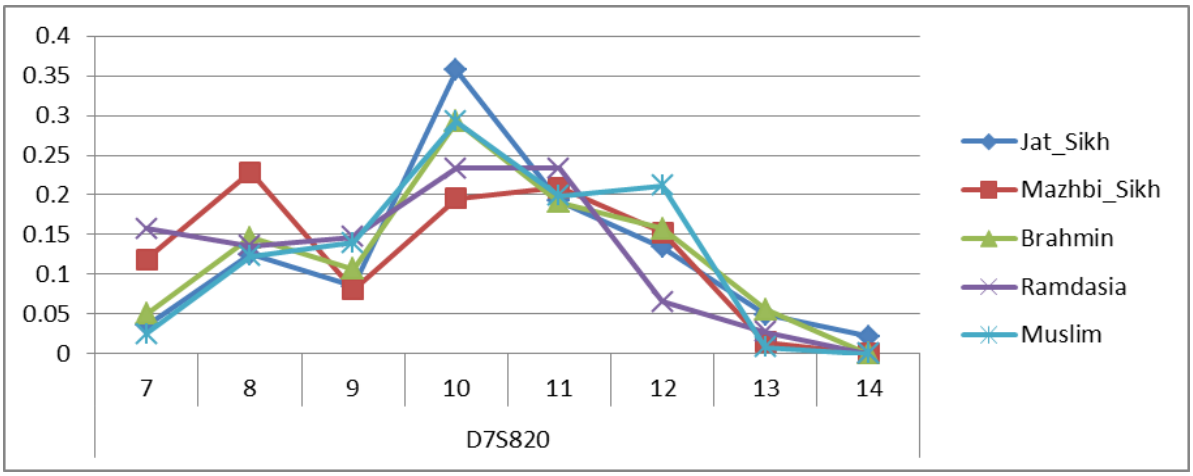
Figure 1: Frequency polygons showing the distribution of alleles of selected STR markers (THO1, TPOX, CSF1PO, vWA, D7S820 and FGA) across the five endogamous Punjabi population groups








0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
11
12
13
14
15
16
17
18
19
20
21
23
vWA
Jat_Sikh
Mazhbi_Sikh
Brahmin
Ramdasia
Muslim


0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
7
8
9
10
11
12
13
14
D7S820
Jat_Sikh
Mazhbi_Sikh
Brahmin
Ramdasia
Muslim

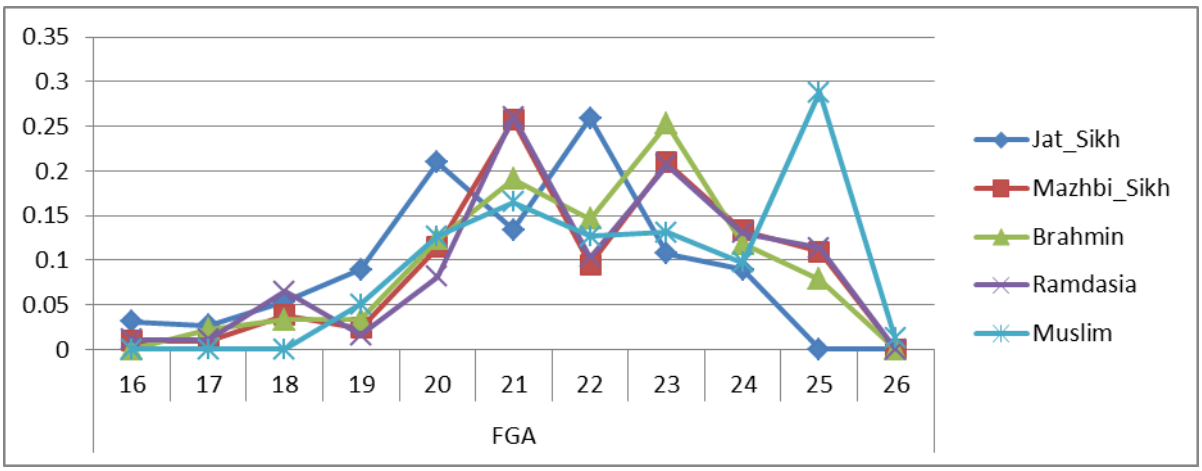

0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
16
17
18
19
20
21
22
23
24
25
26
FGA
Jat_Sikh
Mazhbi_Sikh
Brahmin
Ramdasia
Muslim

Table 3: Probability (p) of departures of Hardy-Weinberg Equilibrium using Exact test among all studied population groups and loci.

Locus	Jat Sikh	Mazhbi Sikh	Brahmin	Ramdasia	Muslim
THO1	0.5073	0.7804	0.1676	0.2486	0.0013
TPOX	0.0001	0.4334	0.2070	0.0063	0.6666
CSF1PO	0.8900	0.0683	0.0000	0.8914	0.0859
vWA	0.0014	0.0371	0.2771	0.8153	0.8900
D7S820	0.0000	0.0000	0.0036	0.0000	0.0007
FGA	0.0084	0.0002	0.0183	0.6321	0.0000

Table 4. Observed (H_{obs}) and expected heterozygosity (H_{exp}) values, genomic differentiation coefficient (G_{st}), inter-population fixation index (F_{st}) and inter-population estimates of gene flow (N_m) among five caste groups of north-west Punjab

Locus	Parameter	Jat Sikh (n=112)	Mazhbi Sikh	Brahmin (n=89)	Ramdasia (n=92)	Muslim (n=118)	Mean
THO1	H_{obs}	0.7143	0.8286	0.7528	0.6630	0.8136	0.7545
	H_{exp}	0.8271	0.8011	0.8061	0.7997	0.7969	0.8062
	F_{st}						0.0155
	G_{st}						0.011
	N_m						15.9153
TPOX	H_{obs}	0.6429	0.7429	0.8539	0.6957	0.6525	0.7176
	H_{exp}	0.8176	0.8125	0.8319	0.7745	0.6205	0.7714
	F_{st}						0.0618
	G_{st}						0.058
	N_m						3.7952
CSF1PO	H_{obs}	0.5893	0.6952	0.7079	0.7283	0.7542	0.6950
	H_{exp}	0.6860	0.7969	0.8453	0.7839	0.7099	0.7644
	F_{st}						0.0649
	G_{st}						0.061
	N_m						3.6007
vWA	H_{obs}	0.6875	0.7714	0.5730	0.5000	0.7797	0.6623
	H_{exp}	0.8988	0.8292	0.8170	0.8116	0.7989	0.8311
	F_{st}						0.0335
	G_{st}						0.029
	N_m						7.2100
D7S820	H_{obs}	0.7679	0.7143	0.7640	0.7935	0.7373	0.7554
	H_{exp}	0.7978	0.8295	0.8242	0.8300	0.8014	0.8166
	F_{st}						0.0134
	G_{st}						0.009
	N_m						18.4612
FGA	H_{obs}	0.9196	0.9048	0.9326	0.9239	0.8983	0.9158
	H_{exp}	0.8466	0.8439	0.8497	0.8464	0.8349	0.8443
	F_{st}						0.0131
	G_{st}						0.010
	N_m						18.7660
General F_{st}							0.0335
General G_{st}							0.030
Mean N_m							7.2192

Table 5. Informativeness of the five microsatellite markers studied

Locus	Parameter	Jat Sikh	Mazhbi Sikh	Brahmin	Ramdasia	Muslim
THO1	**MP**	0.085	0.117	0.082	0.106	0.111
	PD	0.915	0.883	0.918	0.894	0.889
	PIC	0.79	0.73	0.76	0.75	0.69
	PE	0.784	0.607	0.495	0.347	0.407
	PI	4.73	2.56	1.93	1.40	1.59
TPOX	**MP**	0.070	0.071	0.072	0.109	0.225
	PD	0.930	0.929	0.928	0.891	0.775
	PIC	0.78	0.78	0.80	0.73	0.55
	PE	0.345	0.498	0.703	0.422	0.359
	PI	1.40	1.94	3.42	1.64	1.44
CSF1PO	**MP**	0.158	0.080	0.081	0.110	0.180
	PD	0.842	0.920	0.919	0.890	0.820
	PIC	0.62	0.76	0.81	0.74	0.65
	PE	0.278	0.421	0.441	0.473	0.517
	PI	1.22	1.64	1.71	1.84	2.03
vWA	**MP**	0.051	0.066	0.095	0.104	0.092
	PD	0.949	0.934	0.905	0.896	0.908
	PIC	0.88	0.80	0.78	0.78	0.77
	PE	0.409	0.547	0.260	0.188	0.562
	PI	1.60	2.19	1.17	1.00	2.27
D7S820	**MP**	0.082	0.065	0.070	0.094	0.094
	PD	0.918	0.935	0.930	0.906	0.906
	PIC	0.77	0.80	0.79	0.80	0.76
	PE	0.541	0.451	0.534	0.587	0.488
	PI	2.15	1.75	2.12	2.42	1.90
FGA	**MP**	0.090	0.098	0.076	0.057	0.106
	PD	0.910	0.902	0.924	0.943	0.894
	PIC	0.82	0.82	0.82	0.82	0.81
	PE	0.836	0.805	0.862	0.844	0.792
	PI	6.22	5.25	7.42	6.57	4.92

MP-matching probability; PD-power of discrimination; PIC-polymorphism information content; PE-exclusion potential; PI-paternity index.

Table 6. Distance matrix of five studied population groups

Population	Jat Sikh	Mazhbi Sikh	Brahmin	Ramdasia	Muslim
Jat Sikh		0.1169	0.1258	0.1938	0.1786
Mazhbi Sikh			0.0717	0.1315	0.2241
Brahmin				0.1792	0.2382
Ramdasia					0.3661
Muslim					

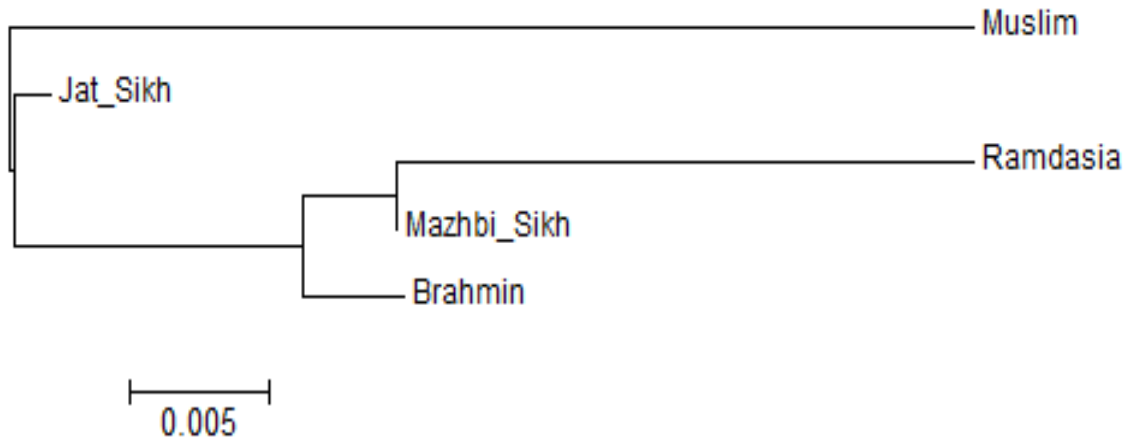


Figure 2: Figure showing genetic affinity among the five studied Punjabi population groups (Jat Sikh, Mazhbi Sikh, Brahmin, Ramdasia and Muslim) using Neighbor Joining (NJ) method

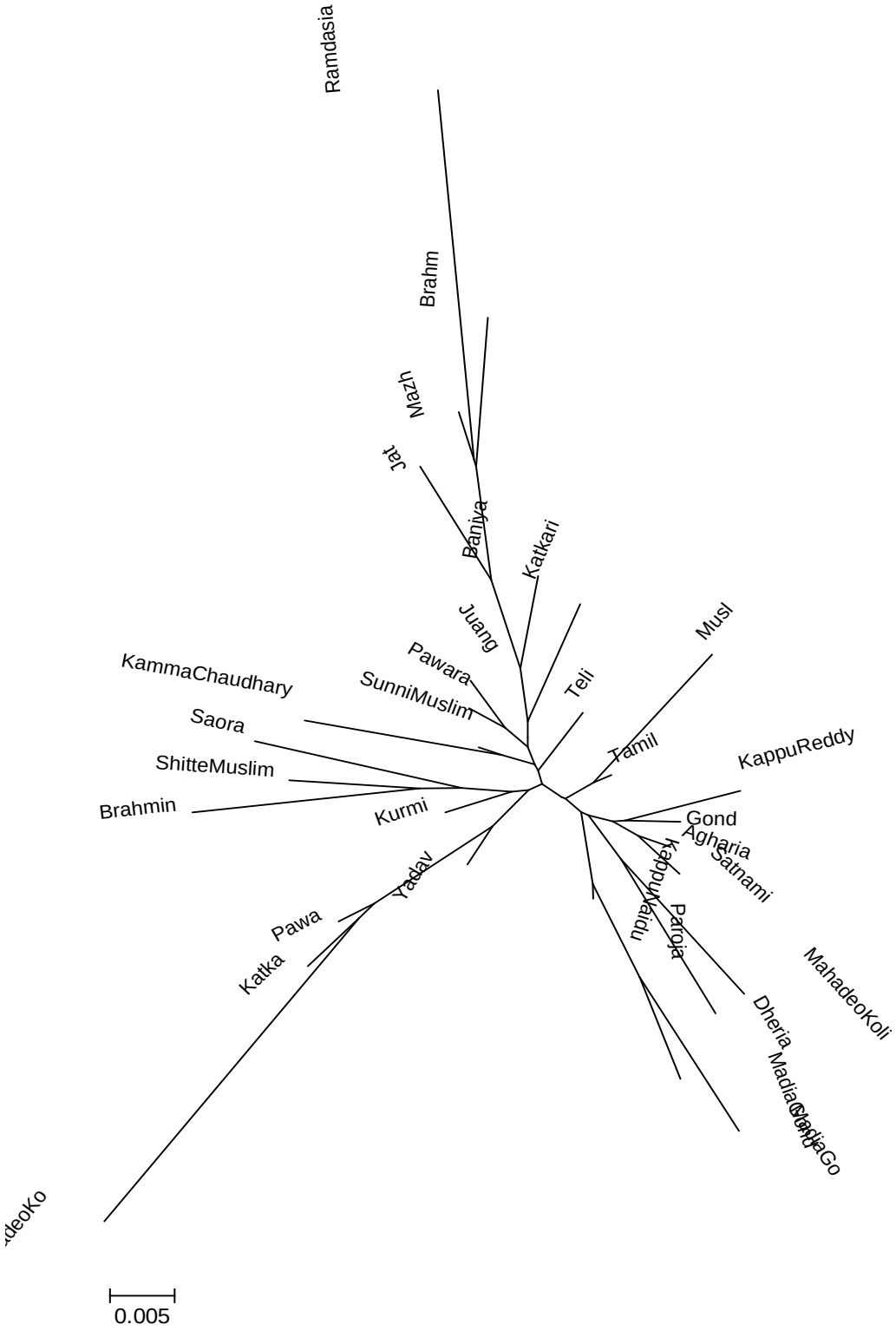


Figure 3: Figure showing NJ tree for genetic affinity among studied populations with other populations in India using Neighbor Joining (NJ) method

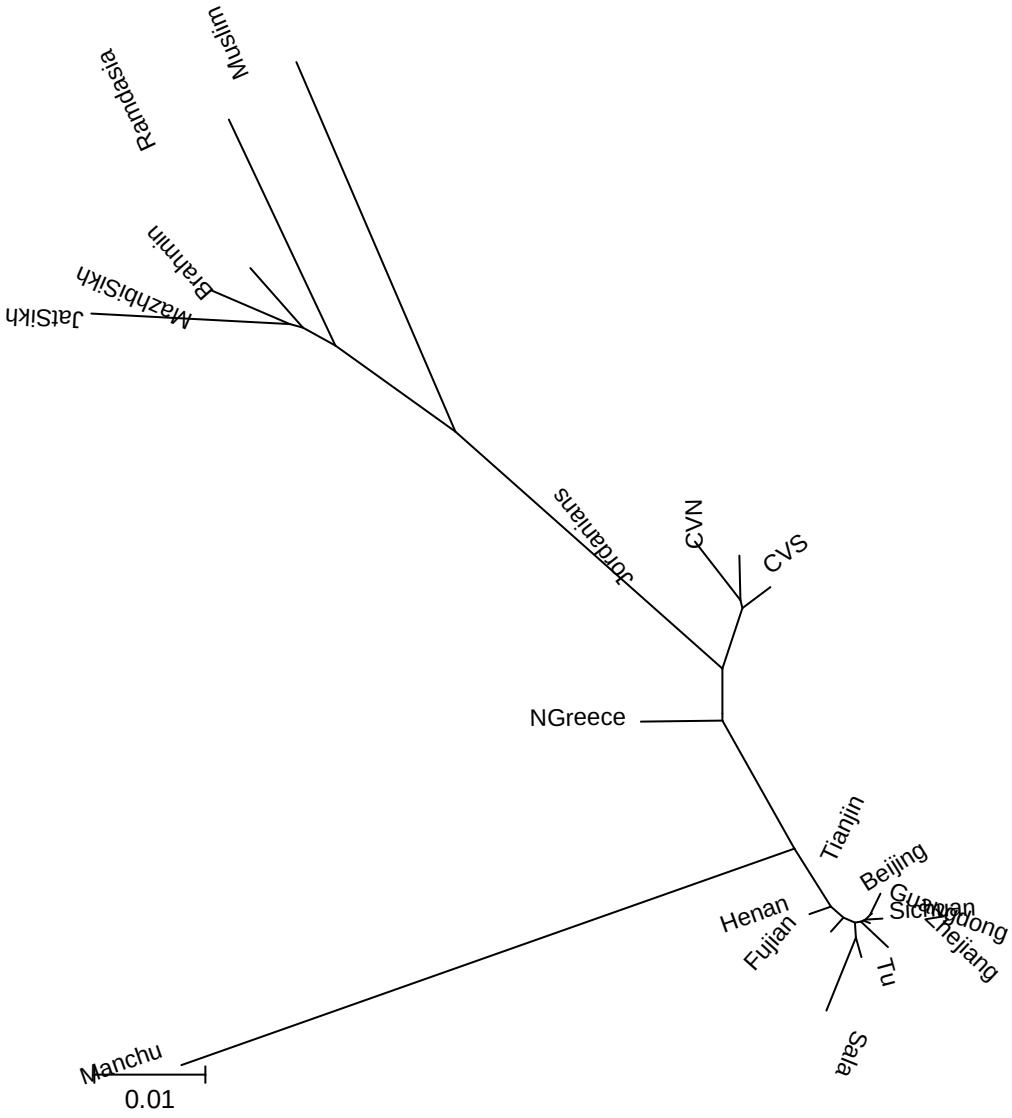


Figure 4: Figure showing Neighbor Joining tree for genetic affinity among studied populations with other international populations using Neighbor Joining (NJ) method

Table 7: Membership proportion of each pre-defined five northwest Punjabi population groups in each of K clusters and logarithm of estimated probability of data

(Admixture Model with Correlated Allele Frequencies; Simulations of data for five northwest Punjabi population groups)

	K=2		K=3		
Inferred Clusters	**1**	**2**	**1**	**2**	**3**
Jat Sikh	0.575	0.425	0.365	0.329	0.307
Mazhbi Sikh	0.408	0.592	0.315	0.323	0.362
Brahmin	0.433	0.567	0.305	0.341	0.354
Ramdasia	0.350	0.650	0.297	0.377	0.327
Muslim	0.677	0.323	0.397	0.292	0.311
Ln Prob	-19039.9			-11902.9	
α value	0.9388			1.0061	

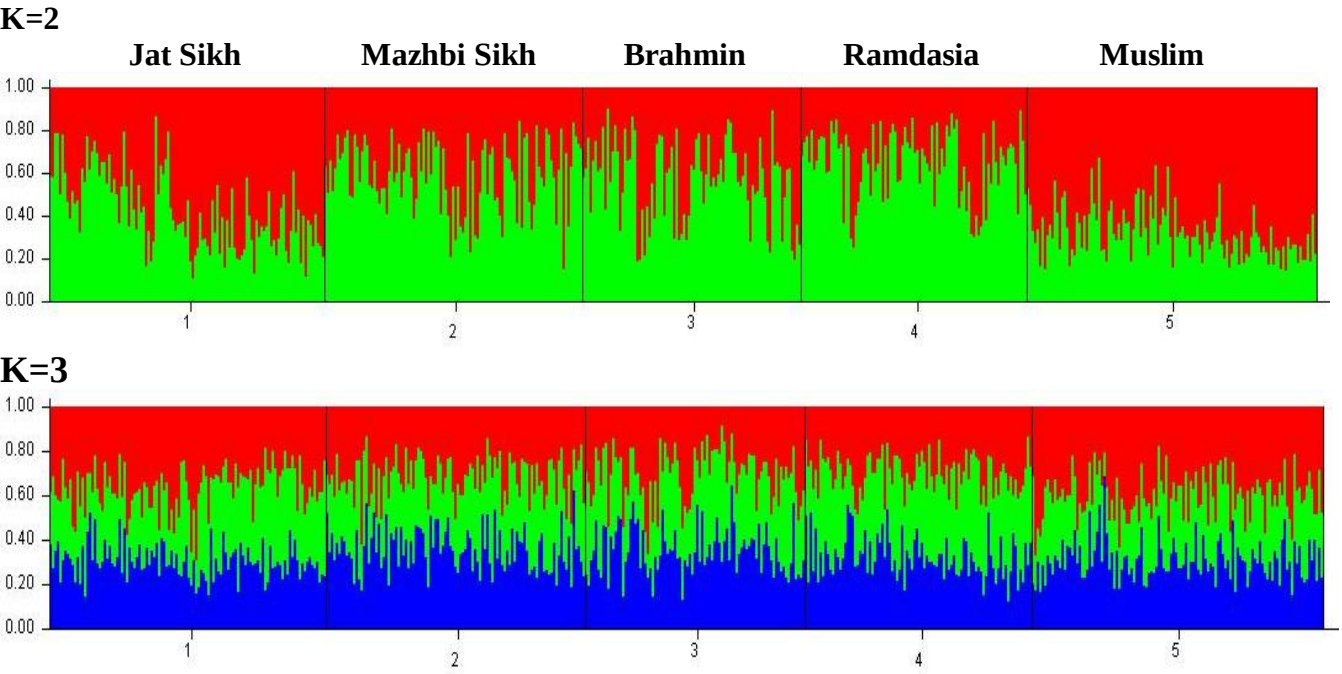


Figure 5: Bar plot estimation figures of five northwest Punjabi population groups, inferred from STRUCTURE analysis

13. **ACHIEVEMENTS FROM THE PROJECT:** The five study population groups of northwest Punjab showed socio-cultural as well as linguistic homogeneity. There are no such studies among these population groups with respect to STR markers to investigate the phylogenetic position of these populations. This is perhaps the first comprehensive molecular genetic study attempted to investigate the genetic affinity and diversity among these closely related population groups in north-west Punjab.

14. **SUMMARY OF THE FINDINGS:** The major foreign population movements in India took place via Punjab and they merged into the indigenous population in process and became part of the culture of the Punjab. The present Punjabi population is the descendants of the various racial stocks which entered into it during the different stages of its history and it has influenced the total genetic structure of Punjabi population. In the present study the genetic polymorphism of six microsatellite loci (THO1, TPOX, CSF1PO, vWA, D7S820 and FGA) in five Punjabi population groups of north-west Punjab namely, Jat Sikh, Majbi Sikh, Brahmin, Ramdasia and Muslim was investigated and it was used to locate phylogenetic position of Punjabi population with other Indian and world populations, as well as, to explore the genetic relationships between various caste groups and their impact on gene pool. All selected groups practice endogamy. A total of 518 individuals including 112 Jat Sikhs, 105 Majbi Sikhs, 89 Brahmins, 92 Ramdasias and 118 Muslims were recruited from northwest districts of Punjab. The allele repeat number varied from 5 (THO1) to 26 (FGA) among the studied population groups. The exact test suggested some significant departures in certain loci and population groups. THO1 among Muslims, TPOX among Jat Sikh and Ramdasia, CSF1PO among Brahmins, vWA among Jat Sikh and Mazhbi Sikh, D7S820 among all the five groups, FGA among Jat Sikh, Mazhbi Sikh, Brahmin and Muslim population groups showed significant departure from HWE. All the observed and expected heterozygosities among all the populations have not been found significantly different. The average sub-ethnic genomic differentiation (F_{st}) among five population groups of northwest Punjab was 0.0335. The CSF1PO has showed highest sub-ethnic differentiation (0.0649), whereas, the lowest (F_{st}) has been observed for FGA locus (0.013). It was observed that 3.0% of the total genomic diversity (G_{st}) is attributable to the five studied population groups. The maximum gene flow (N_m=18.77%) has been observed in FGA. The high level of intra-populational diversity resulted in a considerable discriminating power of this STR marker system. The PD values for almost all the markers lied in the range of 0.880-0.949. All the six markers (THO1, TPOX, CSF1PO, vWA, D7S820 and FGA) were found to be highly polymorphic (PIC>0.5). In the total population sample there were no two individuals with the same genotype of the six marker system. Therefore, this marker system actually meets all the existing requirements to identification system. The phylogenetic assessments were made based on the allele frequency profile of 6 STR markers. Punjabi Mazhbi Sikh, Ramdasia and Brahmin group formed a distinct cluster clearly separated from Jat Sikh and Muslim. The present study has been compared with 26 Indian populations to get a more comprehensive picture of genetic similarity and differences existing between Indian populations. All present study populations except Muslims have clustered in one group, whereas, presently studied Muslim group clustered with South Indian Tamil populations. The present study has also been compared with 14 world populations, which have been categorized into four major continents. The present five populations (Jat Sikh, Mazhbi Sikh, Brahmin, Ramdasia, Muslim) formed one distinct cluster, whereas, other continental populations have formed distinct and separate clusters. To understand the extent of sub-structuring among five northwest Punjabi populations, the present study has performed STRUCTURE analysis with different values of K. The log probability values and the membership proportion of each group showed clear sub-structure among the population groups. In case of K=2 and K=3, Muslim group showed higher proportion of membership (0.677 and 0.397) to cluster 1. Ramdasia showed higher proportion of membership in cluster 2 with respect to K=2 (0.650) and K=3 (0.377), respectively. In case of cluster 3 of

K=3, Mazhbi Sikh showed high values of membership proportion (0.362). Overall, five Punjabi speaking population groups such as Jat Sikh, Mazhbi Sikh, Brahmin, Ramdasia and Muslim from northwest Punjab are regionally well differentiated and exhibit strong genetic affinity based on their origin, settlement and their shared ethno-historic background. However, a clear picture will probably emerge from the analysis of increased number of informative genetic markers and population size.

15. **CONTRIBUTION TO THE SOCIETY:** Analysis of the present data has shown that all humans are generally homogeneous and genetic variation and homogeneity tends to be shared widely among populations. Different genetic structure is expected from geographically isolated human populations. Substantial overlap of the study group has been observed through their dendrogram and structure analysis. Therefore, present study has provided useful information in biomedical context. As the foreign populations have entered India through this region, the gene flow might have influenced human behavior and allele frequencies which are often compared to other populations. The present study may produce useful evolutionary insight of the present Punjabi population.

    This study has accumulated the population genetic data of several STRs in Punjabi population. This data could also be used in forensic and individual identification for this population which is assumed to be highly heterogenic population. It has helped to locate phylogenetic position of Punjabi population with other Indian and world population. This population genetic data could enrich Indian genetic informational resources.

16. **WHETHER ANY PH.D ENROLLED/PRODUCED OUT OF THE PROJECT:** Yes (1)
    Name of the Ph.D. fellow: Manpreet Kaur
    Ph.D. Topic: Population Genetic Analysis of Five North West Punjabi Endogamous Groups using Microsatellite Markers
    Status: The Ph.D. is at the last stage of completion. The data compilation and analysis are ongoing.

17. NO. OF PUBLICATIONS OUT OF THE PROJECT:

Papers publications

1. Manpreet Kaur and Badaruddoza (2014). Genetic Polymorphism of Five STR markers among four groups of Punjabi population in North-West Punjab, India. *International Journal of Human Genetics* ~~(in press).~~ 14(1): 1-7.
2. Badaruddoza and Manpreet Kaur (2014). Genetic affinities among five endogamous groups of Punjabi population. *Meta Gene* (communicated).
3. Badaruddoza and Manpreet Kaur (2014). Genetic structure of North-Indian Punjabi population based on autosomal microsatellite loci. *Genetics and Molecular Biology* (communicated).
4. Manpreet Kaur and Badaruddoza (2014). Population data of 6 autosomal microsatellite markers and 3 Y-chromosomal markers among five groups of Punjabi population in Northwest Punjab, India. *Molecular Biology International* (communicated).

Posters presentations

1. Manpreet Kaur and Badaruddoza. "Genetic variation of five polymorphism STR markers (THO1, TPOX, D7S820, CSF1PO and vWA) in five Punjabi population groups". International Conference on Human Genetics & 39th Annual Meeting of Indian Society of Human Genetics- Healthy Genes - Healthy Life held at Ahmedabad Management Institute, Ahmedabad from Jan. 22-15, 2014.

(PRINCIPAL INVESTIGATOR)

(REGISTRAR/PRINCIPAL)
(Seal)
Registrar,
Guru Nanak Dev University,
Amritsar.

DEPARTMENT OF HUMAN GENETICS
GURU NANAK DEV UNIVERSITY, AMRITSAR
PUNJAB - 143005

UGC Project: Study of Genetic Polymorphism of Short Tandem Repeat (STR) Loci in Punjabi Population of North-West Punjab

ETHNOGRAPHIC INFORMATION

1. Sample No. _____ 2. Code _____
3. Name of individual _____
4. Age/D.O.B. _____ 5. Sex _____
6. Religion _____ 7. Caste _____
8. Ancestral Place (Village/Town) _____
9. Address with phone number _____

SOCIO-ECONOMIC STATUS AND EDUCATION

10. Occupation: Labour/Agriculture/Service (private/govt./retired)/self-employed/others
11. Monthly Income (Rs.): Below 2,000/ 2,000 – 5,000/ 5,000 -8,000/ 8,000 - 10,000/ Above 10,000
12. Education: Illiterate/Below primary/Primary/Secondary/ Higher secondary/Graduate/Post-graduate/Professional/Others

ANTHROPOMETRIC AND PHYSIOMETRIC MEASUREMENTS

13. Height(cm) _____ 14. Weight(kg) _____
15. BMI _____
16. Waist Circumference(cm) _____ 17. Hip Circumference(cm) _____
18. WHR _____
19. SBP (mm/Hg) (a) _____ (b) _____ (Avg) _____
20. DBP (mm/Hg) (a) _____ (b) _____ (Avg) _____
21. MBP (mm/Hg) (a) _____ (b) _____ (Avg) _____
22. Pulse Rate (a) _____ (b) _____ (Avg) _____
23. Pulse Pressure (a) _____ (b) _____ (Avg) _____

24. Any existing disorder/disease: _____

25. Brief pedigree:

Date: _____ Signature of Investigator

CONSENT

I have been explained the possible risks and benefits and also have understood the purpose of the study for which my blood sample, physical examination, measurements and other information is being sought by the Department of Human Genetics, Guru Nanak Dev University, Amritsar. I am free from any pressure whatsoever and hereby giving my own consent for the same. I have no objection for all types of analysis of my sample for non-profit research purpose foe acquisition of knowledge for the benefit of the mankind.

Date: _____ Signature/Thumb Impression of Respondent

 Int J Hum Genet, 14(1): 1-7 (2014)

Genetic Polymorphism of Five STR Markers among Four Groups of Punjabi Population in North-West Punjab, India

Manpreet Kaur and Badaruddoza

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

KEYWORDS STR Loci. Genetic Diversity. Jat Sikh. Majbi Sikh. Brahmin. Ramdasia

ABSTRACT Historically, Punjab has served a major passage way for human migration to Indian subcontinent and has influenced the genetic structure of extant populations in Punjab. Therefore, Punjabi population possesses an exclusive genetic profile primarily due to the many migratory events in this region which caused an extensive range of genetic diversity. Hence, the present study is an attempt to understand the effect of these influences on genetic differentiation, diversity and population structure of Punjab. Genetic polymorphism at five highly polymorphic short tandem repeat loci (STR) is studied in four endogamous population groups of Punjab, India. The studied groups included Jat Sikh, Majbi Sikh, Brahmin and Ramdasia to evaluate their significance in human identification and genetic study. All selected groups practice endogamy and a total of 358 individuals belonging to these four endogamous groups were studied for five highly polymorphic with greater power of exclusion STR loci: THO1, TPOX, D7S820, CSF1PO and vWA. The highest observed heterozygosity was found in Ramdasia population for almost all the markers except vWA which had highest observed heterozygosity among Majbi Sikh population group. In this study, the average sub-ethnic differentiation (Fst) among the four populations of north-west Punjab was 0.0821. The marker with the highest contribution to interpopulation genetic difference was observed to be vWA.

INTRODUCTION

The multiallelic and hyper variable nature of STRs make them highly informative markers to study the population genetic structure and evolutionary relationship between human populations, forensic sciences and human gene mapping (Nakamura et al. 1987; Deka et al. 1995; Shazia et al. 2009; Ferdous et al. 2010;He and Guo 2013; Liu et al. 2013; Vieira et al. 2013; Soltyszewski et al. 2014). Therefore, in the present study, the researchers report allele frequency data of five autosomal polymorphic microsatellite loci (THO1, TPOX, D7S820, CSF1PO and vWA) from four distinct endogamous groups of Punjabi population in India to obtain a reference population genetic database.

Population Information

Ethnic Punjabi population shows enormous cultural, linguistic, and genetic diversity due to its positioning on the crossroads of many historic and pre-historic human migrations (Sekhon 2000). The hierarchical caste system dominates the social structure of the Punjabi populations. The origin of the caste system in India is a matter of debate with many linguists and anthropologists suggesting that it began with the arrival of Indo-European speakers from Central Asia about 3500 years ago. Also, Indian castes have been found to be more closely related to the Central Asians than to the Indian tribal groups (Cordaux et al. 2004; Nair et al. 2011). Punjabi population possesses an exclusive genetic profile primarily due to the many migratory events in this region which caused an extensive range of genetic diversity. Hence, the present study is an attempt to understand the effect of these influences on genetic differentiation, diversity and population structure of Punjab. Till date no study has reported the microsatellite diversity among these population groups in Punjab. Of all the states in India, Punjab has the highest number of Scheduled Caste group. They constitute an impressive 28.9% of the total population of Punjab. The caste groups selected for the present study were Jat Sikh (higher caste of Sikh religion), Majbi Sikh (lower caste of Sikh religion), Brahmins (higher caste of Hindu religion) and Ramdasia/Valmiki (lower caste of Hindu religion) with their informed consents. All selected groups practice a high degree of endogamy (Sekhon 2000).

Address for correspondence
Dr. Badaruddoza
Department of Human Genetics;
Guru Nanak Dev University;
Amritsar143 005, Punjab, India
E-mail: doza13@yahoo.co.in

Jat Sikh: Jats are the biggest group in terms of numbers (66%) among Sikh religion and recognized as higher caste group. They own more than 80%, of available agricultural land in Punjab. Besides agriculture, which is their signature trade, Jat Sikhs are now very well educated and have taken up various professions. They often reside in the rural areas, and are economically influential in the state. Jat Sikhs are known for their lively spirit and easy-going nature. The Jats have probably originated as one of the late immigrants to the subcontinent.

Majbi Sikh: They are mainly found in the Punjab, Kashmir and Rajastan regions (Sekhon 2000) and are considered lower caste group in Sikh religion. In the glorious Sikh Regiment in Indian army, Majbi Sikhs were recruited in good numbers due to their bravery, physical strength and self-sacrifice. Their houses are generally located on the outskirts of the town and are economically very poor. The urban Mazbhis have made social and economic progress over the years, however, poverty and illiteracy is still rampant among them.

Brahmin: They are recognized as the highest class group among Hindu religion. In Punjab most of the Brahmins are Saraswat Brahmins and they were recognized as the intellectual and priestly class from ancient civilization to till date. They are highly respected and honored for creating the world's oldest literary and religious traditions. They were the original propagators of the revered texts such as the Vedas and the Upanishads. They have also excelled as educators, law makers, scholars, doctors, warriors, writers, poets, land owners and politicians.

Ramdasia: They generally belong to socially low class society among the Hindu religion. The Ramadasia community has traditionally been relegated to the most menial labour with negligible possibility of upward mobility and also subject to social disadvantages and exclusion in comparison to the wider community. More than 80% of Ramadasias in Punjab are living in villages.

A very few number of genetic studies (Badaruddoza et al. 2007, 2008; Badaruddoza and Sudhir 2012; Saini et al. 2012) have been carried out on Punjabi population. However, no STR marker based study on present study populations have been reported in the literature till date. Therefore, the present data would be used in forensic and individual identification for these selected population groups and these genetic data would enrich Indian genetic informational resources. The study was approved by Ethical Research Committee of Guru Nanak Dev University, Amritsar, Punjab.

MATERIAL AND METHODS

Sample Collection

3ml venous blood samples were obtained from a total of 358 (112 Jat Sikh, 105 Majbi Sikh, 71 Brahmin and 70 Ramdasia) unrelated healthy subjects through simple random sampling from four endogamous groups of north-west border districts of Punjab such as Amritsar, Fazilka, Ferozepur, Gurdaspur, Pathankot and Tarn Taran. Blood was immediately transferred to pre-labelled blood collecting vial containing 0.5M EDTA (as anticoagulant). All samples were transported on ice from the place of sample collection to laboratory and were stored at -20°C till further analysis.

DNA Extraction, PCR Amplification and STR Genotyping

Genomic DNA was extracted using phenol-chloroform method (Sambrook et al. 1989) and quantified using NanoDrop™ 2000/2000c Spectrophotometer (Thermo Scientific™, Pittsburgh, US). In the present study five STR markers were selected: THO1, TPOX, CSF1PO, vWA and D7S820 on the basis of their heterozygosity and their feasibility for rapid analysis (through PCR) in lab. Selection of the present STR marker has also been based on the global survey carried out by Perez-Lezaun et al. (1997).

The PCR amplification was done on Mastercycler ® Personal (Eppendorf, Germany). using locus specific primer pairs (Table 1). PCR program used was initial denaturation at 95°C for 5 min followed by 35 cycles of 94°C for 45 sec, 58°C for 45 sec, 72°C for 45 sec, followed by final extension at 72°C for 10 min for THO1, TPOX and CSF1PO loci done in a multiplex. Same protocol was used for other two loci also except the annealing step which was 56°C for vWA and 53°C for D7S820. The PCR reaction mixture of 15 µl consists of 50 ng of DNA, Taq polymerase buffer, 200 µmol/l dNTPs, and 1.0 U Taq polymerase and 5 pmol of each primer. PCR products were checked on 2.5% agarose gel and then ana-

lyzed on 10% native polyacrylamide gel electrophoresis (PAGE) and various alleles were detected after proper silver staining.

Statistical Analysis

Various statistical analyses were performed on the STR data to accomplish the objectives of the study. The allele frequencies, observed and expected heterozygosity, genotype distribution to the Hardy-Weinberg equilibrium was performed using GENEPOP software pakage (version 3.3). Locus informativeness parameters, such as matching probability (MP), power of discrimination (PD), polymorphism information content (PIC), power of exclusion (PE) and paternity index (PI) were analyzed using PowerStat v12 software (Promega Corporation, USA). The phylogenetic tree was drawn by unweighted group-method with arithmetic mean (UPGMA). Pair wise genetic distances were calculated based on the allele frequencies of STR loci using Nei's method.

RESULTS

The distributions of the observed allele frequencies for five STR loci in four Punjabi population groups are compared in Table 2. Seven for THO1, eight for TPOX and D7S820 each, 10 for CSFIPO and 12 for vWA different alleles were observed. The maximum allele repeats for THO1, TPOX, D7S820, CSF1PO and vWA have been observed 9, 8, 10, 12 and 16 respectively in Jat Sikh; 9, 8, 8, 13 and 17 respectively in Majbi Sikh; 6, 8, 11, 14-15 and 17 respectively in Brahmin; 9, 7, 10, 13 and 17 respectively in Ramdasia groups. It is also noticed that 8, 9 and 17 allele repeats are most common in all four groups. Different populations contain alleles with repeated number varying from 5 to 23. In all four populations, the distributions of allele frequencies are bimodal with major peak in 8 and secondary peak may be found at alleles 9 and 17. The highest frequency (44.64%) observed for allele 12 for CSFIPO locus among Jat Sikh, whereas, the lowest frequency (0.45%) observed for allele 8 for the same locus among Jat Sikh.

For all the markers, the researchers determined the observed and expected heterozygosity values (Table 3). The highest observed heterozygosities were found in Ramdasia group for four loci (0.8462, 0.7692, 0.9231 and 0.7692 for THO1, TPOX, D7S820 and CSF1PO loci respectively). However, the observed heterozygosity varied from 0.5893 in Jat Sikh for CSFIPO locus to 0.9231 in Ramdasia for D7S820 locus. In general, the average observed heterozygosity is lower than expected heterozygosity in five STR markers in four population groups. In this study the average sub-ethnic differentiation (Fst) among four population groups of northwest Punjab was 0.0821. The D7S820 marker has showed small inter-population differences (F_{st}=0.0145), whereas, vWA showed the highest contribution to inter-population genetic differences (F_{st}=0.1577).

With respect to PIC value, THO1, TPOX, D7S820, CSF1PO and vWA loci were classified as highly informative markers (PIC >0.70) (Table 4). All five studied loci have considerable discriminating power, PD values lying in the range from 0.834 for THO1 in Ramdasia to 0.949 for vWA in Jat Sikh. Therefore, in total sample there were no two individuals with the same genotype of the five marker system.

Based on the genetic distance, dendrogram was constructed to depict the genetic affinities (Fig. 1) among four population groups. The dendrogram showed the low genetic distance bem-

Table 1: STR markers with primer sequences and product sizes

STR marker	*Primers*	*Product size (bp)*
HUMTHO1	For: ATTCAAAGGGTATCTGGGCTCTGG	
	Rev: GTGGGCTGAAAAGCTCCCGATTAT	171-215
HUMTPOX	For: ACTGGCACAGAACAGGCACTTAGG	
	Rev: GGAGGAACTGGGAACCACACAGGTTA	216-264
HUMCSF1PO	For: AACCTGAGTCTGCCAAGGACTAGC	
	Rev: TTCCACACACCACTGGCCATCTTC	287-331
HUMVWA	For: CCCTAGTGGATGATAAGAATAATC	
	Rev: GGACAGATGATAAATACATAGGATGGATGG	122-182
D7S820	For: TGTCATAGTTTAGAACGAACTAACG	
	Rev: CTGAGGTATCAAAAACTCAGAGG	198-234

Table 2: Allele frequency data of THO1, TPOX, CSF1PO, vWA and D7S820 microsatellite markers among four Punjabi population groups (n=358)

Locus	*Allele*	*Populations*			
		Jat Sikh (n=112)	*Majbi Sikh (n=105)*	*Brahmin (n=71)*	*Ramdasia (n=70)*
THO1	5	0.0089	0.0147	0.0373	0.0962
	6	0.1964	0.2843	0.2910	0.1154
	7	0.1518	0.1765	0.2537	0.2115
	8	0.2500	0.1912	0.1866	0.1923
	9	0.3170	0.2892	0.1567	0.3269
	9.3	0.0759	0.0441	0.0672	0.0577
	10	0.0000	0.0000	0.0075	0.0000
TPOX	5	0.0000	0.0000	0.0217	0.0000
	6	0.0893	0.0882	0.1087	0.0577
	7	0.1607	0.1667	0.2609	0.3654
	8	0.2991	0.2794	0.2826	0.3269
	9	0.1920	0.2206	0.2101	0.1731
	10	0.1295	0.1520	0.0870	0.0769
	11	0.1116	0.0784	0.0217	0.0000
	12	0.0179	0.0147	0.0072	0.0000
D7S820	7	0.0357	0.1188	0.0652	0.0000
	8	0.1250	0.2327	0.1884	0.1154
	9	0.0848	0.0693	0.1087	0.2308
	10	0.3571	0.1931	0.1377	0.3269
	11	0.1920	0.2129	0.2464	0.2115
	12	0.1339	0.1584	0.1957	0.0577
	13	0.0491	0.0149	0.0580	0.0577
	14	0.0223	0.0000	0.0000	0.0000
CSF1PO	7	0.0000	0.0000	0.0000	0.0577
	8	0.0045	0.0000	0.0072	0.0385
	9	0.0045	0.0147	0.0217	0.1346
	10	0.2054	0.0980	0.0725	0.1731
	11	0.2723	0.1863	0.1377	0.1346
	12	0.4464	0.2745	0.1304	0.1538
	13	0.0670	0.2794	0.1377	0.1923
	14	0.0000	0.1078	0.2391	0.0962
	15	0.0000	0.0392	0.2391	0.0192
	16	0.0000	0.0000	0.0145	0.0000
vWA	11	0.0367	0.0300	0.0000	0.0000
	12	0.0550	0.0650	0.0145	0.0000
	13	0.0229	0.0400	0.0000	0.0192
	14	0.1514	0.1350	0.1232	0.0192
	15	0.1239	0.0800	0.0290	0.0000
	16	0.1560	0.2350	0.2391	0.2308
	17	0.0780	0.3000	0.3768	0.3077
	18	0.1055	0.0500	0.1014	0.2885
	19	0.1055	0.0650	0.0870	0.0962
	20	0.0826	0.0000	0.0290	0.0385
	21	0.0780	0.0000	0.0000	0.0000
	23	0.0046	0.0000	0.0000	0.0000

tween Majbi (lower caste group of Hindu religion) and Brahmin (higher caste group of Hindu religion). The tree also reflected the ethnic background of the four populations. Considering the linguistics and genetic affinities, the four groups have same origin with recent separation from common stock. However, the Majbi and Brahmin groups showed extensive gene flow between each others.

DISCUSSION

The present study was undertaken to produce population genetic database based on STR

Table 3: Observed and expected heterozygosity values and genetic differentiation coefficients in four Punjab population groups (n=358)

Locus	*Parameter*	*Jat Sikh (n=112)*	*Majbi Sikh (n=105)*	*Brahmin (n=71)*	*Ramdasia (n=70)*	*Mean*
THO1	H_{obs}	0.6518	0.7941	0.7463	0.8462	0.7596
	H_{exp}	0.7730	0.7694	0.7915	0.8009	0.7837
	F_{st}					0.0307
TPOX	H_{obs}	0.6429	0.7451	0.7246	0.7692	0.7205
	H_{exp}	0.8140	0.8122	0.7933	0.7345	0.7885
	F_{st}					0.0862
D7S820	H_{obs}	0.7679	0.7129	0.8841	0.9231	0.8220
	H_{exp}	0.7942	0.8231	0.8332	0.7903	0.8102
	F_{st}					0.0145
CSF1PO	H_{obs}	0.5893	0.6961	0.6522	0.7692	0.6767
	H_{exp}	0.6829	0.7928	0.8307	0.8756	0.7955
	F_{st}					0.1493
vWA	H_{obs}	0.6789	0.7700	0.6812	0.6154	0.6864
	H_{exp}	0.8954	0.8208	0.7715	0.7722	0.8150
	F_{st}					0.1577
General F_{st}						0.0821

Table 4: Informativeness of the five microsatellite markers studied

Locus	*Parameter*	*Jat Sikh*	*Majbi Sikh*	*Brahmin*	*Ramdasia*
THO1	MP	0.086	0.114	0.089	0.166
	PD	0.914	0.886	0.911	0.834
	PIC	0.74	0.73	0.75	0.75
	PE	0.367	0.599	0.523	0.687
	PI	1.46	2.50	2.06	3.25
TPOX	MP	0.070	0.069	0.101	0.154
	PD	0.930	0.931	0.899	0.846
	PIC	0.78	0.78	0.76	0.67
	PE	0.345	0.501	0.467	0.543
	PI	1.40	1.96	1.82	2.17
D7S820	MP	0.082	0.066	0.070	0.115
	PD	0.918	0.934	0.930	0.885
	PIC	0.77	0.79	0.80	0.74
	PE	0.541	0.448	0.763	0.843
	PI	2.15	1.74	4.31	6.50
CSF1PO	MP	0.158	0.081	0.070	0.077
	PD	0.842	0.919	0.930	0.923
	PIC	0.62	0.76	0.80	0.84
	PE	0.278	0.422	0.358	0.543
	PI	1.22	1.65	1.44	2.17
vWA	MP	0.051	0.066	0.105	0.118
	PD	0.949	0.934	0.895	0.882
	PIC	0.88	0.79	0.73	0.72
	PE	0.396	0.545	0.393	0.310
	PI	1.56	2.17	1.55	1.30

MP-matching probability; PD-power of discrimination; PIC-polymorphism information content; PE-exclusion potential; PI-paternity index

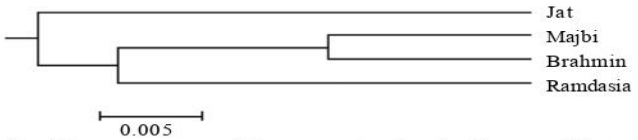


Fig. 1. Nei's genetic distances was used to generate the dendogram illustrating the phylogenetic relationship of four studied population groups

loci among four Punjabi caste groups to compare allele frequencies. Presently, there has been increasing use of microsatellite loci to understand the genetic relationship between closely related populations (Chu et al. 1998; Reddy et al. 2001, 2005; Krithika et al. 2008). In the present study the researchers used five autosomal STR microsatellite markers to understand the population structure and variation of four caste populations distributed in north-west Punjab. These STR loci are highly polymorphic within each of the population. The distributions of allele frequencies are moderately uniform across the population, suggesting relative homogeneity among the four castes of Punjabi population. No deviation was observed of the studied loci from Hardy Weinberg Equilibrium (HWE). The least average heterozygosity value among Jat Sikh for CSF1PO loci (~59%), THO1 (~65%) and TPOX (~64%) might be explained by their preferential marriage practices among clans prohibiting external gene flow. In the present study, D7S820 was the only marker showing very small differences among the population studied, whereas, the marker with highest contribution to inter-population genetic differences was vWA. Furthermore, not many studies have been done in these population groups to compare the present results (Reddy et al. 2005; Khan et al. 2007; Noor et al. 2009). The results suggested that the present marker system actually meets all the existing requirements and can be used for DNA typing and population studies.

The phylogentic analyses of four caste populations showed closer proximity between Majbi Sikh which is lower caste of Sikh religion and Brahmin which is higher caste of Hindu religion. This probably suggested their common genetic affinity and possible admixture between these two populations. Therefore, the clustering between Brahmin and Majbi Sikh also points towards the possibility of the central Asian immigrants appointed themselves to predominantly belonged to caste of higher rank and subsequently, the lower caste was admixed with them. The similar type of clustering between higher and lower caste has been observed in many north Indian population (Bamshad et al. 2001; Khan et al. 2007). However, the results should be confirmed in further study. Overall, Jat Sikh population (higher caste) are well differentiated with other neighboring populations. Bamshad et al. (2003) expressed the view that a large number of microsatellite loci are required to differentiate populations. However, the present five autosomal STR markers that are highly polymorphic and widely used for forensic investigations, for investigating local population structure and to construct the evolutionary relationship between different groups of population (Langsteieh et al. 2004; Reddy et al. 2005).

CONCLUSION

The phylogenetic analysis of four endogamous caste populations of the north-west Punjab based on five STR loci have revealed the information about the genetic affinities between Jat Sikh (higher caste of Sikh religion), Majbi Sikh (lower caste of Sikh religion), Brahmin (higher caste of Hindu religion) and Ramdasia (lower caste of Hindu religion). The present results have demonstrated that intra-population differences were marginal, however, there was a definite pattern of genetic variation found in four different caste populations.

ACKNOWLEDGEMENTS

The study is supported by the UGC major project "Study of Genetic Polymorphism of Short Tandem Repeat (STR) loci in Punjabi Population of North-west Punjab" sanctioned to Dr. Badaruddoza (F. No. 39-110/2010SR).

REFERENCES

Badaruddoza, Bhanwer AJS, Rambani M, Singh R, Matharoo K, Bamezai RNK 2008. YAP element among an endogamous human isolate in Punjab. *Int J Hum Genet*, 8: 269-271.

Badaruddoza, Bhanwer AJS, Kaur R, Singh R, Matharoo K, Bamezai RNK 2007. Study of DYS 390 Polymorphism among *Khatri* Population of Punjab in comparison to other Indian and world population. *Int J Hum Genet*, 7: 263-266.

Badaruddoza, Sudhir N 2012. No evidence for association between ACE Gene Insertion (I)/Deletion (D) polymorphism and hypertension in North Indian Punjabi population. *Int J Hum Genet*, 12: 179-185.

Bamshad MJS, Kivisild T, Watkins WS, Dixon ME, Ricker CE, Rao BB 2001. Genetic evidence on the origins of Indian caste populations. *Genome Res*, 11: 994-1004.

Bamshad MJS, Wooding WS, Watkins CT, Ostler, Batzer MA, Jorde LB 2003. Human population genetic structure and inference of group membership. *Am J Hum Genet*, 72: 578–589.

Chu JY, Huang W, Kuang SQ, Wang JM, Xu JJ, Chu ZT, Yang ZQ, Lin KQ, Li P, Wu M, Geng ZC, Tan

CC, Du RF, Jin L 1998. Genetic relationship of populations in China. *Evolution*, 95: 11763-11768.
Cordaux RR, Aunger G, Bentley I, Nasidze SM, Sirajuddin, Stoneking M 2004. Independent origins of Indian caste and tribal paternal lineages. *Curr Biol,* 14: 231-235.
Deka RM, Shriver LMY, Ferrel RE, Chakraborty R 1995. Intra and inter population diversity at short tandem repeat loci in diverse population of the world. *Electrophoresis*, 16: 1659-1664.
Ferdous AME, Ali S, Alam M, Hasan M, Hossain T, Akhteruzzaman S 2010. Allele frequencies of 10 autosomal STR loci from Chakma and Tripura tribal population in Bangladesh. *Mol Bio Int*, 8: 12-23.
He J, Guo F 2013. Genetic variation of fifteen autosomal STR loci in a Manchu population from Jilin province, northeast China. *Forensic Sci Int Genet*, 7: e45-46.
Nakamura Leppert M, O'Connell P, Wolff R, Holm T, CulverM, Martin C, Fujimoto E, Hoff M, Kumlin E et al. 1987. Variable number of tandem repeats markers for human gene mapping. *Science*, 235: 1616-1622.
Khan F, Pandey AK, Tripathi M, Talwar S, Bisen PS, Borkar M, Agrawal, S 2007. Genetic affinies between endogamous and inbreeding populations of Uttar Pradesh. *BMC Genetics*, 8: 12-22.
Krithika S, Maji S, Vasulu TS 2008. A microsatellite guided insight into the genetic status of Adi, an isolated hunting-gathering tribe of northeast India. *PLoS One*, 3: e2549.
Langstieh BT, Reddy BM, Thangaraj K, Kumar V, Singh L 2004. Genetic diversity and relationships among the tribes of Meghalaya compared to other Indian and Continental populations. *Hum Biol*, 76: 569-590.
Liu J, Guo L, Qi R, Li SY, Yin JY, Zhang W, Sun Z, Tian X, Gao B 2013. Allele frequencies of 19 autosomal STR loci in Manchu population of China with phylogenetic structure among worldwide populations. *Gene*, 529: 282-287.
Nair SP, Geetha A, Jagannath C 2011. Y-short tandem repeat haplotype and paternal lineage of the Ezhava population of Kerala, south India. *Croat Med J*, 52: 344-350.
Noor S, Ali S, Eaaswarkhanth M, Haque I 2009. Allele frequency distribution for 15 autosomal STR loci in Afridi Pathan population of Uttar Pradesh, India. *Legal Med*, 11: 308-311.
Reddy BM, Sun G, Dutta R, Deka R 2001. STR data for AmplFlSTR Profiler Plus loci among Golla population of Southern Andhra Pradesh, India. *J Forensic Sci*, 46: 734-735.
Reddy BM, Naidu VM, Madhavi VK, Thangaraj K, Langstieh BT, Venkataramana P, Kumar V, Singh L 2005. STR data for the Amp F/STR Profiler plus loci among 27 populations of different social hierarchy from southern part of Andhra Pradesh, India. *Forensic Sci Int,* 149: 81-97.
Saini JS, Kumar A, Matharoo K, Sokhi J, Badaruddoza, Bhanwer AJS 2012. Genomic diversity and affinities in population groups of north west India: An analysis of Alu insertion and a single nucleotide polymorphism. *Gene*, 511: 293-299
Sambrook J, Fritsch EF, Maniatis T 1989. *Molecular Cloning: A Laboratory–Manual.* New York, NY: Cold Spring Harbor Laboratory Press.
Sekhon IS 2000. *Punjabis: The People, their History, Culture and Enterprise.* New Delhi: Cosmo Publication.
Shazia AP, Nithya, Seshadri M 2009. Genetic variation of polymorphic *NOS* STR locus in ten Indian population groups. *Tehetnka,* 45: 1-4.
Soltyszewski I, Pepinski W, Wolanska-Nowak P, Maciejewska A, Paszkowska R et al. 2014. Polish population data on 15 autosomal STRs of AmpFlSTR NGM PCR kit. *Forensic Sci Int Genet*, 9: 142-149.
Vieira1 TC, Silva DM, Gigonzac MAD, Ferreira VL, Gonçalves MW, da Cruz AD 2013. Allelic frequencies and statistical data obtained from 15 STR loci in a population of the Goiás State. *Genet Mol Res*, 12: 23-27.