INNSPUB

International Journal of Biosciences | IJB |
ISSN: 2220-6655 (Print) 2222-5234 (Online)
http://www.innspub.net
Vol. 9, No. 2, p. 159-163, 2016

RESEARCH PAPER OPEN ACCESS

Y-chromosome polymorphisms in two Pakistani ethnic groups and its relationship with neighbouring Indian populations

Muhammad Tariq*[1], Habib Ahmad[2], Inamullah[1], Umar Farooq[3], Tauseef Ahmad[4]

[1]*Department of Genetics, Hazara University, Mansehra, Pakistan*

[2]*Islamia College University, Peshawar, Pakistan*

[3]*Center for Human Genetics, Hazara University, Mansehra, Pakistan*

[4]*Department of Microbiology, Hazara University, Mansehra, Pakistan*

Key words: Pakistan, Ethnic groups, Y-chromosome, STRs, Haplotypes

http://dx.doi.org/10.12692/ijb/9.2.159-163 Article published on August 31, 2016

Abstract

The paternally inherited Y-chromosome short tandem repeats (Y-STRs), are excellent tool in inferring modern human evolutionary studies, genetic human identification and genetic genealogy. In the present study a total of 60unrelated male of two ethnic groups (Gujarsand Karlars) from Khyber Pakhtunkhwa province of Pakistan were analyzed using PowerPlex®12 Y-STR loci amplification system (DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 andDYS439), to investigate the genetic polymorphisms and to determine the genetic relationship of two Pakistani populations with three neighboring Indian populations of the same linguistic family. In Gujars a total of 21 haplotypes were identified, 15 of which were unique and 6 haplotype were shared among two or more individuals. The haplotype diversity among Gujars was 0.9701 ± 0.0171, while in Karlars population 19 haplotypes were identified among which 13 were unique haplotypes and 6 were shared among two or more individuals and the observed haplotype diversity value was 0.9563 ± 0.0213.R_{ST}pairwise analysis suggest close genetic relationship between three Indian populations (Andh, Naikpod, Pardhan), while genetically distinct from the two Pakistani Populations (Gujars and Karlars). Results also demonstrate that the 12 Y-STR loci analyzed were highly polymorphic in Gujars and Karlars populations and hence useful for forensic cases and population's genetic studies.

* **Corresponding Author:** Muhammad Tariq ✉ mtariqhu@gmail.com

Introduction

Pakistan and India are the two largest countries of south Asia that comprises populations of many diverse ethnic and linguistic groups, which results from the continuous foreign invasions in different time. The modern state of Pakistan also encompasses a diverse population made up of many ethnic groups. The country comprises greater than18 ethnic and 60 linguistic groups (Grimes, 1992), majority of them are the Punjabi speaking populations, which comprise complex admixture of many ethnic caste and groups (Ibbetson, 1883) such as the Gujar, Jats, Rajput and Arians etc. The northern parts of the country are mostly covered by the Himalayas, Karakoram and Hindu Kush mountains. Although the Hindu Kush Mountains in the northwest represent a significant geophysical feature separating Pakistan and Afghanistan, these mountains are riddled with an array of mountain passes that were used by a broad array of invaders, traders and travelers for entering the Indo-Pak Subcontinent during the Holocene and before. The southwestern region of the country is covered by flat land characterized by low-lying hills and deserts. Stone implements present near in the Soan River Valley of northern part of Pakistan provide the earliest Palaeolithic evidence of a human presence South Asia (Allchin and Allchin, 1982; Hussain, 1997). Although there is a dearth of human fossil remains, the stone implements of the Soan Culture provide incontrovertible evidence of a homin in presence in South Asia between 2.0 and 0.4 my (Wolpert, 2000). Given such antiquity, the Soan Culture was likely associated with some form of archaic Homo sapiens, either *Homo erectus* or *Homo heidelbergensis*.

Pakistan lies on the postulated southern coastal route followed by anatomically modern Homo sapiens out of Africa (Mellars, 2006), and so may have been inhabited by the modern humans at any point subsequent of 160 ky. as early as 6-7 thousand years ago. Evidence of large underground houses in Khyber Pakhtunkhwa province of Pakistan are found, but fossil evidence have been incomplete (Hussain, 1997).

Pakistani population is polygenetic and as such represents an amazing amalgamation of various ethnic groups and cultures. Located at the extreme periphery of northwestern South Asia occupies a unique position on the cultural and historical map of the world immediately adjacent to West Asia, Central Asia, South Asia and the age-old causeway to the Far East, the Great Silk Road.

The inheritance pattern of the Y chromosome has been used widely to reconstruct of human population histories and the dispersal of modern humans out of Africa and into the rest of the world (Barger *et al.,* 2012). Y-chromosome DNA variations were first reported in 1985 (Casanova *et al*., 1985; Lucotte and Ngo, 1985). The non-recombining Y-chromosome (NRY) of the human Y chromosome, transmitted exclusively by the father along the part line, without any change except by mutation generation after generation. Thus, the Y-chromosome preserves a relatively straightforward record of the genetic history of a specific population. The Y-chromosome short tandem repeats (STRs) preserve a sufficient degree of variability among individuals within a population coupled with a high degree of geographical differentiation, which makes them especially suited for forensic and evolutionary studies (Kim *et al.,* 2012). Large number of STR markers have been identified on the human Y-chromosome (Karafet *et al*., 2008). The availability these STRs markers during the last decade has greatly assisted genome-based investigations of diverse populations worldwide. However, there are no studies specifically conducted on some major ethnic groups of Pakistan. This study explain the genetic polymorphism of two Pakistani ethnic groups from Khyber Pakhtunkhwa province and to determine their relationship with three neighbouring Indian Populations of the same linguistic family using Y-STRs markers.

Materials and methods

Sampling collection and DNA analysis

Mouthwash samples were collected from a total of 60 unrelated male from Gujars (n=30) and Karlars (n=30) ethnic populations residing in Khyber Pakhtunkhwa province of Pakistan were collected (Table 1).

Informed consent from all participants of this study were taken before sampling. The sampling donors were explained the aims and objectives of the study. DNA was extracted from mouthwash samples using a standard Phenol: chloroform procedure. The isolated DNA samples were used forco-amplification of 12 Y-STR loci following the conditions recommended by the manufacturer of the commercial Power Plex Y-STRs system (Promega).

The amplified fluorescently labeled PCR products were separated by capillary gel electrophoresis on an ABIP rism 3130 genetic analyzer (Applied Biosystems). The genotypes were read using Gene Mapper software v 3.1 (Applied Biosystems). The quality of gene typing was controlled using the standard set of alleles of all STRs("ladder") supplied within the Power Plex 12 Y-STR system; the "ladder" were loaded in each gene typing cycle (in each run).

Table 1. Comparative populations for Y-STRs NRY analysis.

Ethnic Group	Sample Size (N)	Samples location	Language Group	References
Gujars	30	KP, Pakistan	Indo-European	This Study
Karlar	30	KP, Pakistan	Indo-European	This Study
Andh	53	southern India	Indo-European	Thanseem *et al.,* 2006
Naikpod	68	southern India	Indo-European	Thanseem *et al.,* 2006
Pardhan	128	southern India	Indo-European	Thanseem *et al.,* 2006
Total	309			

Statistical Analysis

Haplotype diversity, which is analogues to gene diversity andmean number of pairwise differences were calculated for both Gujars and Karlars populations using Arlequin v 3.5.1.2 (Excoffier and Lischer, 2010). Arlequin v 3.5.1.2 (Excoffier and Lischer, 2010), was also utilized to determine the Pairwise genetic distance (R_{ST}) between a given pair of populations based on six Y-STR loci (DYS19, DYS389I, DYS389II, DYS390, DYS391 and DYS393). The resulted R_{ST} values were visualized in the Multidimensional Scaling plot (MDS), using SPSS V 21 software to show similarities and differences among two Pakistani ethnic groups and the three reference Indian populations (Andh, Naikpod, Pardhan) (see Table 1) (Thanseem *et al.,* 2006). An analysis of molecular variance (AMOVA) was also carried out using P-value generated from1,000 permutation copies and the R_{ST} genetic distance, which takes into account the probability of recurrent mutation for calculating distance using the same Arlequin v 3.5.1.2 software (Excoffier and Lischer, 2010).

Results and discussion

Haplotype Diversity

Haplotypes distributions among the two ethnic groups in this study was presented in table 2. In Gujars a total of 21 haplotypes were identified for all the 12 Y-STRs loci (DYS389I, DYS390, DYS389II, DYS19, DYS385a, DYS385b-DYS393-DYS391-DYS439-DYS392-DYS437-DYS438), of which 15 were unique haplotypes and 6 haplotypes were shared among two or more individuals. The most frequent haplotypes among Gujars was (13-24-30-16-11-14-13-11-10-11-14-11) occurring in 4 individuals, while haplotype (12-22-28-14-13-16-11-10-11-14-15-11) was found among 3 individuals. Karlars population exhibit a total of 19 haplotypes, 13 of which were unique and 6 haplotypes were shared among two or more individuals. The most common haplotype among Karlars was (13-24-30-16-11-14-13-10-10-11-14-11) occurring in 5 individuals, while haplotypes (13-24-30-16-11-14-13-11-10-11-14-11) and (13-25-30-16-11-14-13-11-11-11-14-11) both occurring in identical 3 individuals respectively. The two Pakistani populations were characterized by high haplotypic diversities; the least value were found among the Karlars (0.9563 ± 0.0213) and highest value among the Gujars (0.9701 ± 0.0171) (Table 2).

The highest value of haplotype diversity among the Gujars (0.97) shows that there were few shared haplotypes. The mean number of pairwise differences value was high among Gujars (7.733333 ± 3.706028), while the least value (6.728736 ± 3.263364) was found among Karlars population. These results indicate a high degree of Y-chromosome STRs diversity both within Gujars and Karlars Pakistani ethnic populations. Previous study on Pakistani ethnic groups showed that Y-chromosome diversity was structured by geography and not by ethnicity (Qamar *et al.,* 2002). With the exception of Hazara population from Baluchistan all ethnic groups shown to have similar Y-chromosome variation and have close relationship with south Asian and Middle Eastern males (Qamar *et al.,* 2002). This study describes the Y-chromosome diversity of Gujars and Karlars populations of Khyber Pakhtunkhwa province of Pakistan for the first time. We also explored the genetic composition of these two populations and their genetic diversity were correlated with three other populations of the same linguistic family from neighboring India (Table 1).

Comparative data analysis

To explore the relationship of two Pakistani ethnic groups and three Indian ethnic groups, R_{ST} values based on only six Y-STRs loci (DYS19, DYS389I, DYS389II, DYS390, DYS391 and DYS393), were estimated and displayed in the multidimensional scaling (MDS) plot (Table 3 and Fig. 1). The stress value (goodness of fit statistic) was 0.00492 of the MDS plot result. The MDS plot provide clear picture about all the five populations. The three Indian populations (Andh, Naikpod and Pradhan), were loosely cluster within the right side of the plot, while they were significantly different from the two Pakistani (Gujars and Karlars) populations. The Gujars was an outlier, situated at the bottom right-hand quadrant of the MDS plot, while the Karlars populations was situated at the top upper quadrant of the plot were also significantly different from each other. The genetic diversity found among these five populations revealed that they belong to diverse cultural heritage as well as their complex genetic histories.

Table 2. Descriptive statistic of two ethnic populations based on 12 Y-chromosome STRs loci.

Ethnic Groups	Gujars	Karlars
No. of samples	30	30
No. of Haplotypes	21	19
Unique Haplotypes	15	13
Shared Haplotypes	6	6
Haplotype Diversity	0.9701 ± 0.0171	0.9563 ± 0.0213
Mean number of pairwise differences	7.733333 ± 3.706028	6.728736 ± 3.263364

Table 3. R_{ST} values from Y-STR haplotypes (Six loci); RST values (Lower Matrix) and P-values (Upper Matrix).

	Gujars	Karlar	Andh	Naikpod	Pardhan
Gujars	*	0.25225 ± 0.0326	0.00000 ± 0.0000	0.00000 ± 0.0000	0.00000 ± 0.0000
Karlar	0.16838	*	0.02703 ± 0.0139	0.00000 ± 0.0000	0.00000 ± 0.0000
Andh	0.51744	0.43149	*	0.00000 ± 0.0000	0.00000 ± 0.0000
Naikpod	0.48668	0.41941	0.02514	*	0.00000 ± 0.0000
Pardhan	0.45685	0.41744	0.04922	0.00244	*

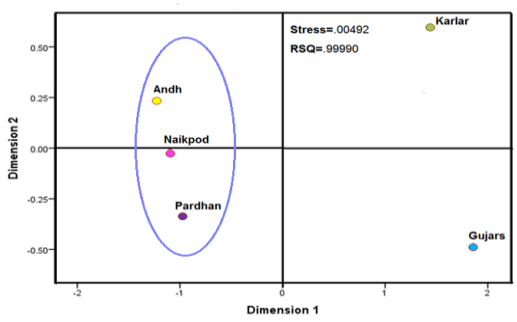

Fig. 1. Multidimensional scaling plot of the genetic distances estimates for all the five ethnic populations; the three Indian Populations are enclosed by circle.

Analysis of molecular variance (AMOVA)

Analysis of molecular variance (AMOVA) test based on was carried out to examine the amount of variations present between two Pakistani and three Indian ethnic groups (Table 4) AMOVA analysis shows that about 41.31% variations was found among groups, 2.75% among populations within groups, whereas the reminder 55.94 % genetic variations accounted for within populations.

Table 4. AMOVA results for the Gujars (n=30), Karlars (n=30) from Pakistan and Indian populations including Andh (n=53), Naikpod (n=68 and Pardhan (n=128). Investigated Y-STRs loci were DYS19, DYS389I, DYS389II, DYS390, DYS391 and DYS393.

Groups	Percentage of Variation	P-Value
Among groups	41.31	0.10459 ± 0.01041
Among populations within groups	2.75	0.00000 ± 0.00000
Within populations	55.94	0.00000 ± 0.00000

Conclusions

In summary the results of this study revealed a high degree of Y-chromosome STRs diversity both within Gujars and Karlars Pakistani ethnic populations, and among the three Indian ethnic populations. The Rst results shows that the populations from India are more closely related to each other, while distant from the two Pakistani populations. The genetic diversity found among these five populations revealed that they belong to diverse cultural background as well as their complex genetic histories.

Acknowledgements

This study is part of the PhD dissertation project of the principal author. We thankall the participants of the two ethnic groups for providing us their samples and make this work possible. We also thank Higher education commission of Pakistan for providing financial support for this study.

References

Allchin B, Allchin R. 1982. The Rise of Civilization in India and Pakistan. Cambridge University Press.

Barger B, Niederstatter H, Erhart D, Gassner C, Gassner H, Schennach H, Parson W. 2012. Tracking the iceman's scent by high resolution mapping of y haplogroup G in Tyrol (Austria). DNA in Forensics. 8th international Forensic Y-User Workshop.

Casanova M, Leroy P, Boucekkine C, Weissenbach J, Bishop C, Fellous M, Purrello M, Fiori G, Siniscalco M. 1985. A human Y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance. Science **230**, 1403-1406.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources **10,** 564-567.

Grimes BF. 1992. "Ethnologue: Languages of the World," 12th ed., Summer Institute of Linguistics, Inc., Dallas, Texas, USA.

Hussain JA. 1997. History of the peoples of Pakistan towards independence. Oxford University press, Karachi, Pakistan.

Ibbetson D. 1883. Punjab Caste. Sang-e-Meel publications, Lahore 9-16.

Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome Research **185,** 830-838.

Kim S, Kim K, Han M, Kim W. 2012. Forensic genetic analysis for the AmpF1STR Yfiler system in the Korean Population. DNA in Forensics 8th International Forensic Y-User Workshop.

Lucotte G, Ngo NY. 1985. p49f, A highly polymorphic probe, that detects Taq1 RFLPs on the human Y chromosome. Nucleic Acids Research **13(22)**, 8285.

Mellars P. 2006. Going east: new genetic and archaeological perspectives on the modern human colonization of Eurasia. Science **313(5788),** 796-800.

Qamar R, Ayub Q, Mohyuddin A, Helgason A, Mazhar K, Mansoor A, Zerjal T, Tyler-Smith C, Mehdi SQ. 2002. Y-chromosomal DNA variation in Pakistan. The American Journal of Human Genetics **70(5),** 1107-1124.

Thanseem I, Thangaraj K, Chaubey G, Singh VK, Bhaskar LV, Reddy BM, Reddy AG, Singh L. 2006. Genetic affinities among the lower castes and tribal groups of India: inference from Y chromosome and mitochondrial DNA. BMC genetics **7(1),** 1.

Wolpert S. 2000. A new history of India. Oxford University Press, New York.