



Genetic portrait of North-West Indian population based on X chromosome *Alu* insertion markers

Gagandeep Singh^{1,2} · Harkirat Singh Sandhu³ · Yellapu Srinivas⁴ · Kawaljit Matharoo² · A. J. S. Bhanwer²

Received: 14 October 2019 / Accepted: 11 December 2019 / Published online: 18 December 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

In the present study, allele frequencies and forensic parameters of four ethnic groups (Brahmin, Khatri, Jat Sikh, and Scheduled Caste) of Punjab, India, at 10 *Alu* insertions of X chromosome were calculated. Six *Alu* markers were observed to be highly polymorphic with no significant deviations from Hardy–Weinberg equilibrium and no linkage disequilibrium present in any marker. Multidimensional plot showed higher genetic affinity of studied populations with Asian populations. Overall, the tested markers were reliable and were found suitable in human forensics and population genetic studies.

Keywords X chromosome · *Alu* insertions · Population genetics · Polymorphism · Human forensics

The X chromosome provides a valuable source of information for population genetics and serves as a unique tool in forensic studies as it integrates the desirable features of other commonly used genetic markers [1]. Till date, no literature was available to assess the variability of *Alu* polymorphisms of the X chromosome in population groups from Punjab, North-West India. Due to its geographical location, the present day North-West Indian population might offer unique insights into historical migration events. Moreover, the exposure of this region's population with the different ethnic background people from the pan-world might have created a mosaic-like pattern in their genomic ancestry [2]. Therefore, in this pioneer study, a set of 10 X chromosome *Alu* markers in four Indo-European speaking ethnic groups (Brahmin, Jat Sikh, Khatri, and Scheduled Caste) of Punjab, North-West India were analyzed.

A total of 379 (175 males and 204 females) unrelated individual blood samples were obtained from four different population groups (Brahmin, Khatri, Jat Sikh, and Scheduled Caste) of Punjab, India, with prior consent. The study was approved by the institutional ethical committee of Panjab University. Complete information regarding genomic DNA extraction methods and studied ethnic groups including their geographic origin, ethnicity, pedigree, inclusion, and exclusion criteria were previously reported [3, 4]. All PCR amplifications were carried out in 10- μ l reactions for a set of 10 *Alu* insertions of X chromosome (Ya5DP4, Ya5DP3, Ya5DP77, Yb8DP49, Ya5491, Ya5NBC37, Yb8NBC102, Yb8NBC634, Ya5DP62, and Ya5DP13). Primer sequences and amplification conditions were obtained from previous study [5]. Negative and positive controls were carried out in all PCR runs to monitor the efficiency and reliability of the PCR reactions. Genotypes were identified by electrophoresis on 2–3% agarose stained with ethidium bromide and visualized under UV transilluminator.

Allele frequencies data of the studied population groups were calculated by direct gene counting method. Online software X-chromosome STR homepage (<http://www.chrx-str.org>) was used to estimate statistical parameters of forensic relevance [6]. The Bayesian clustering analysis implemented in the program STRUCTURE v2.3 [7] was used to detect population structure within the studied populations assuming population admixture model and correlated allele frequencies within the population. Heat maps were constructed based on F_{ST} and D_A values using R statistical software v3.3. For easier

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00414-019-02238-3>) contains supplementary material, which is available to authorized users.

✉ Gagandeep Singh
gsgaba@gmail.com

¹ Department of Anthropology, Panjab University, Chandigarh, India

² Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India

³ Icahn School of Medicine at Mount Sinai, Tisch Cancer Institute, 1470 Madison Avenue, New York, NY, USA

⁴ Wildlife Institute of India, Dehradun, India

visualization of the distance genetic relationship, a multidimensional scaling (MDS) plot of the pairwise F_{ST} values was depicted using SPSS v16.0.

Allele frequencies for the 10 *Alu* insertions in four studied populations are shown in Supplementary Fig. 1. Six X *Alu* insertions (Ya5DP3, Yb8DP49, Ya5NBC37, Yb8NBC102, Yb8NBC634, Ya5DP62) were observed to be polymorphic in all the populations, whereas the remaining are monomorphic in at least one studied population (Ya5DP4 in Khatri and Scheduled Caste, Ya5DP13 in all the populations for the absence of insertion; and Ya5DP77 in Brahmin and Ya5491 in all the populations for the insertion). Significant deviations from HWE were found in few *Alu* insertion polymorphisms in female samples and no significant linkage disequilibrium was present in any pair of the *Alu* markers in male samples. Gene diversity and parameters of forensic relevance calculated for each marker and population were presented in Supplementary Table 1. Ya5NBC37 displayed the highest heterozygosity (0.3895) and Ya5DP4 the lowest (0.0263). Likewise, the most diverse population seems to be the Khatri ($H = 0.1087$) followed by the Scheduled Caste (0.0891), Jat Sikh (0.0571), and Brahmin (0.0509). Polymorphic information content (PIC) values were in the range of 0.0104 to 0.3608 with a mean value of 0.1521. Additionally, the maximum value of power of exclusion (PE) was 0.1645 at Ya5NBC37 marker, whereas the minimum was 0.0001 at Ya5DP4 locus. The Bayesian clustering analysis suggested the presence of two genetic clusters ($K = 2$) in the studied populations. As shown in Supplementary Fig. 2, at $K = 2$, the Khatri population was separated from the remaining three populations with some admixture from other populations. Similarly, with an increase in K ($K = 2-4$), the studied populations shared mixed membership in different color components with no sub structuring in the studied groups. The results are in accordance with the previous studies on the ethnic groups of North-West Indian populations using different sets of molecular markers [8, 9].

To explore the genetic differences, the studied populations were compared with previously published data on populations from different continents [11–15] (Supplementary Table 2) using pairwise F_{ST} and Nei's D_A genetic distance measures. The observed F_{ST} values were in the range of 0.09 to 0.832 with maximum variation found between 0.4 and 0.8. Similarly, North-West Indian populations in this study have moderate and equal proportions (F_{ST} values 0.5–0.7) of genetic differentiation with other populations elsewhere (heat maps shown in Supplementary Fig. 3a). Nei's D_A distance values were in the range of 0.004 to 0.61, with maximum findings between 0.2 and 0.5 (heat maps shown in Supplementary Fig. 3b).

In the MDS plot (Supplementary Fig. 4), all the Asian populations appeared closer to European and Amerindian populations were placed in between African and European

population clusters. Also the four ethnic groups of this study form a closer cluster, suggesting a homogenous genetic entity. In conclusion, the results obtained from these X chromosome *Alu* elements comprise a reliable set of genetic markers that could assist in human forensic genetic investigations. This work also highlights the importance of new studies on additional populations of the Indian subcontinent, as no comparable data exist in literature on X chromosome *Alu* insertions to complete the genetic portrait of the Indian population.

This study follows the guidelines of the International Journal of Legal Medicine for the publication of population data [10].

Acknowledgments The authors are grateful to all of the donors for providing blood samples and to the people who contributed to the collection. They are especially thankful to Dr. Jim Lemon for helping with R codes needed for the genetic analysis.

Funding information This study is financially supported by the Department of Biotechnology (DBT), Government of India, through project grant BT/PR2722/Med/13/121/2001 (DBT, India) to Bhanwer AJS at Guru Nanak Dev University Amritsar.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical statement The study was conducted in compliance with ethical standards.

Informed consent All samples were collected after obtaining written informed consent of participants.

References

- Schaffner SF (2004) The X chromosome in population genetics. *Nat Rev Genet* 5(1):43–51
- Ahloowalia BS (2009) Invasion of the genes genetic heritage of India. Strategic Book 273 Publishing
- Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16(3):1215
- Singh G, Talwar I, Sharma R, Sandhu HS, Matharoo K, Bhanwer AJ (2016) Analysis of ANKK1 (rs1800497) and DRD2 (rs1079597, rs1800498) variants in five ethnic groups from Punjab, North-West India. *Gene* 584(1):69–74
- Callinan PA, Hedges DJ, Salem AH, Xing J, Walker JA, Garber RK, Watkins WS, Bamshad MJ, Jorde LB, Batzer MA (2003) Comprehensive analysis of Alu-associated diversity on the human sex chromosomes. *Gene* 317:103–110
- Szibor R, Hering S, Edelman J (2006) A new web site compiling forensic chromosome X research is now online. *Int J Legal Med* 120(4):252–254
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155(2):945–959
- Singh G, Sandhu HS, Sharma R, Srinivas Y, Matharoo K, Singh M, Bhanwer AJ (2019) Genetic variation and population structure of

- five ethnic groups from Punjab, North-West India: analysis of MHC class I polymorphic Alu insertions (POALINs). *Gene*. 701:173–178
9. Singh G, Talwar I, Sandhu HS, Matharoo K, Bhanwer AJ (2017) Genetic dissection of five ethnic groups from Punjab, North-West India—a study based on autosomal markers. *Legal Med* 26:25–32
 10. Poetsch M, Bajanowski T, Pfeiffer H (2012) The publication of population genetic data in the international journal of legal medicine: guidelines. *Int J Legal Med* 126:489–490

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.