UDC 575.
https://doi.org/10.2298/GENSR2103379A
Review article

X CHROMOSOMAL ANALYSIS IN POPULATION GENETICS AND FORENSIC SCIENCE: A MINI REVIEW

Azam ALI[1,2*], Nazim HUSSAIN[2], Muhammad Saqib SHAHZAD[3], INAMULLAH[2], Qurban ALI[2*]

[1]Institute of Molecular Biology & Biotechnology (IMBB), the University of Lahore, Lahore, Pakistan
[2]Centre of Applied Molecular Biology (CAMB), University of the Punjab, Lahore, Pakistan
[3]Forensic Sciences Department, University of Health Sciences, Khayaban-e-Jamia Punjab, Lahore, Pakistan

Ali A., N. Hussain, M. S. Shahzad, Inamullah, Q. Ali (2021). *X chromosomal analysis in population genetics and forensic science: a mini review*- Genetika, Vol 53, No.3, 1379-1386.

The human X chromosome analysis has been applied to decipher the genetic structure of populations for applications in medical genetics and for human identification, parentage analysis and kinship analysis. Although it has not been studied on vast level with regard to human populations with comparison to other of its counterparts like autosomal markers, Y chromosome and mtDNA yet it is important for great potential in studying oncology, various diseases and forensic science applications. In this mini review, a snapshot of X chromosomal properties as genetic marker has been entailed. The structure and potential multiplex oriented kits utilizing X chromosomal markers have been discussed. Moreover, concerns of different researchers over X chromosomal published data have been referred to point out need of analyzing X chromosomal markers to unravel their role in population genetics, medical genetics and human identification.

Keywords: X chromosome, genetic markers, kinship, forensic

X chromosome

The X chromosome has broad range of applications from causing diseases to human identity testing because of its distinctive inheritance mode than other genetic markers like mtDNA and Y chromosome. A healthy male carries single X chromosome while female has two copies of X chromosomes (SZIBOR *et al.*, 2003a). The study of X chromosome is very imperative in population studies and human identification. Various genetic based methods and DNA markers are being used now days to serve the purpose. Short Tandem Repeats (STRs) have been

Corresponding author: Azam Ali, Institute of Molecular Biology & Biotechnology (IMBB), the University of Lahore, Lahore, Pakistan, E-mail: azam.skyblue@gmail.com

widely explored in analyzing X chromosome for population studies. Other markers like SNPs and InDels have also been employed. Population studies play role in defining on the basis of its unique genetics and have appliances in medical genetics. Making use of population specific genetic segments different multiplex systems could be developed to aid criminal investigations. In this mini review, we shall look into various uses of X chromosomal analysis to serve population studies and human identification. Besides this, research gaps and future scope in terms of potential developments shall be discussed.

The scheme of X chromosomal inheritance makes it stand out in other marker system because of 100 % inheritance of genetic information from mother to son and daughter, from father to daughter and parental grandmother to granddaughter (BUTLER, 2011). In criminal investigations, it helps to resolve complicated kinship cases involving least one parent (mother). X chromosome based investigation, assists to resolve disputed paternity cases to a baby girl in motherless cases. It facilitates in testing half sister with father as common relative. The utility also extends in sorting out incest oriented paternity cases and grandparent grandchildren comparisons (BUTLER, 2015).

Genetic Map of X chromosome

The X chromosome was designated as `X` to illustrate unknown by dint of its presence as single unit made geneticists riddled of scenario as all others discovered had pairs (GUNTER, 2005). It extends to 153 million base pairs in length comprising near about 5% of total DNA present in Cell. A complete sequence of X chromosome was released including 99% reporting of euchromatin sequences. It harbored 1100 genes (ROSS *et al.*, 2005). But this number of protein coding genes varied 874 according to (NCBI) (GEER *et al.*, 2010), 893 according to (UniProt) (CONSORTIUM, 2018) and 852 according (Ensembl) (ROOS-ARAUJO, 2019) stretching to length of 153 Mb. The dissimilarity in counting gene number depends upon adoption of variable gene prediction methods. The X chromosome carries only 5 % of female genome while it drops down to 2.5% in males since being in single copy. It shares visible homology with Y chromosome (BUTLER, 2011).

X Chromosomal Short Tandem Repeats

The portion of DNA having repeat units ranging from 2bp to 7bp are known as STR also recognized as SSR and microsatellite. They have become more ideal choice to study and apply for genetic analysis because of easy amplification and devoid of differential amplification. As they provide high discriminatory power for human identification enabling STRs as valuable for DNA based identification. According to human genome consortium 2001 they make up of 3% human genome. Short tandem repeats could be observed throughout genome and are found as one STR after every 10,000 nucleotides. Application of STRs to study X chromosome complements autosomal STRs and mtDNA. It helps to resolve complex kinship analysis half sister deficient paternity cases. While analyzing trios involving daughters, we could get high mean exclusion chance (MEC) using X chromosomal STR markers (YANG *et al.*, 2017). In 2007, Professor Szibor from institute of legal medicine in Magdeburg (Germany) along with colleagues Sandra Hering (Dresden, Germany) and Jeanette Edelmann (Leipzig, Germany) introduced X chromosomal typing to forensic DNA typing (SZIBOR, 2007). They built a website

(http://www.chrx-str.org/) presenting an X chromosomal research in forensic science with account on STR markers. One of the purposes of website was to provide STRs discovered on X chromosome along with haplotype based published data from several countries. The website enlists 55 STRs and consists of data issued in peer reviewed scientific journals. The researchers can also submit STR based population data after meeting with set standards. Professor Butler (BUTLER, 2011) has mentioned 33 STRs from X chromosome has been used by forensic community. The commercial kits available so far (See Table 1) having X chromosomal STRs are Menotype Argus 8 kit including sex determining STR amelogenin with eight X chromosomal STRs. While Investigator ArgusX-12 kit (Biotype, Germany) consists of 12 X chromosomal STRs, D21S11 and amelogenin suitable for kinship, paternity and individual identification (LIU *et al.*, 2011). Another kit X Decaplex (an in house kit) developed by adding 10 X chromosomal markers through collaborative effort accomplished by Spanish and Portugeese ISFG working group (GEP-ISFG) (GUSMÃO *et al.*, 2009). The AGCUX19 is relatively new commercial kit with 19 X chromosomal STRs mainly used so far to study population samples from Chinese Kazakh, Uyghur and Guanzhong Han (HE *et al.*, 2017). The detailed composition of previously developed X chromosomal STR kits with number of markers in each kit based on previous studies have been given in Table 1 (HE *et al*., 2017).

Table 1. The complete information of X-STR markers and sets corresponding to linkage group compositions of previous studies based developed commercial kits (HE et al., 2017)

Commercial Kit Name	Number of X chromosomal STR markers	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	**Set 7**
Mentype®Argus X-UL PCR Amplification Kit	4	DXS8378	DXS7132	HPRTB	DXS7423			
Mentype® Argus X-8 PCR Amplification Kit	8	DXS8378, DXS10135	DXS7132, DXS10134	HPRTB,DXS10101	DXS7423, DXS10074			
Investigator Argus X-12 PCR Amplification Kit	12	DXS8378, DXS10135, DXS10148	DXS7132, DXS10134, DXS10079	HPRTB,DXS10101,SXS10103	DXS7423, DXS10074 ,DXS10146			
X-Decaplex (Inhouse-Kit)	10	DXS8378, DXS9902	DXS7132, DXS9898, DXS6809, DXS6789	DXS7133, GATA172D05	GATA31E08, DXS7423			
AGCU X19 kit	19	DXS8378, DXS10135, DXS10148	DXS10159, DXS10162, XS10164	DXS7132, DXS10079, DXS10074,DXS10075	DXS6809, DXS6789	DXS7424, DXS101	DXS10103,HPRTB,DXS10101	DXS7423,DXS10134

Other X Chromosomal Markers

The markers like Insertion-Deletion (INDELs) and Single Nucleotide Polymorphism (SNP) are found within X chromosome and could be used to devise multiplex schemes. A study was conducted to analyze 14 SNPs on X chromosome and was found highly useful for low throughput applications (LI *et al.*, 2010). A multiplex (25plex) consisting of 25 SNPs from X chromosome was developed and found informative (TOMAS *et al.*, 2010). A panel of 62 X linked SNPs was tested on four population samples and observed as highly discriminative and useful for forensic identification and kinship analysis (STEPANOV *et al.*, 2016b). A multiplex system of 66 X chromosomal SNPs (XSNPid) was proposed for human identification and reported as highly polymorphic with enormous multiplex capacity along with time and cost efficacy (STEPANOV *et al.*, 2016a). The use of biallelic marker like SNPs poses problem in result interpretation like in case of multiple donor interpretation because of overlap making it hard for personal identification. This is one of the reason behind lack of interest in exploring X chromosomal SNPs (GOMES *et al.*, 2020).

Another multiplex system making use of 33 X chromosomal insertion/deletion was reported (FREITAS *et al.*, 2010). One more study examining 30 X chromosomal indels from DIPlex to test utility in forensic science was reported by typing 200 unrelated individuals from Northeast Italy (TURRINA *et al.*, 2011). Pursuing the trail to test new markers next came the 32 X chromosomal indels in single PCR test were shown to be informative and polymorphic in major human population groups (PEREIRA *et al.*, 2012). An INDEL system including 18 X chromosomal biallelic markers was tested on Han population from China and was found reliable in deficient relationship cases where other markers failed to give conclusive results (ZHANG *et al.*, 2015). The 1000 genome database was explored to select 10 X Multi INDEL and investigated in Chinese populations for applications in relationship testing and human identification (FAN *et al.*, 2015). To investigate the efficiency of X chromosomal INDELs, 21 biallelic markers were tested on one German and three Baltic population groups. This system proved to be useful in population studies to see association on genetic level and in forensic science applications (EDELMANN *et al.*, 2016). The applications of X chromosome were further explored to optimize 33 X chromosome associated INDELs using 320 individuals from Argentina and was found valuable for criminal investigations (CAPUTO *et al.*, 2017). While utilizing the X chromosome, another kind of marker (Haploblock) was assessed for its potential use in forensic and population genetic studies. The haploblocks are genomic segments with little recombination rate and have capacity to be used in forensic investigations. No commercial kits have been developed having INDELs. One of the possible reason of non-availability of INDELs based kits is lack of X chromosomal applications in forensic genetics as compared to other autosomal markers. In a study performed by Fan and his colleagues presented substitute approach to facilitate forensic genetic based studies having various INDELs that were tightly linked , amplified through single pair of PCR primers and have new of marker (FAN *et al.*, 2015). In that study 24, STRs were included during haploblock selection to ensure the increment in resolution for research. Keeping in view the size of X chromosome, only four haploblocks were selected to carry out the population genetic studies. Along with Alu repeats, the potential of X chromosomal haploblocks still to be explored by studying different population samples from different geographical regions to establish their credibility. While analyzing the difference at

genetic level North-East Argentina Populations, (DI SANTO MEZTLER *et al.*, 2019) the Found the usefulness of Alu repeats for the comparison of distantly related population samples.

Concerns over Published data

This section showcases mistreatment of parameters not costumed for X choromosomal data, unwanted use of bioinformatics giving rise to incompatible results and inaccuracies go unchecked by reviewers, readers and editors. Some experts in field have reviewed the published population genetic and forensic related data and have suggested some precautions to be considered for the publication of data specifically X chromosomal data.

According to experts , data generated from X chromosomal markers requires strict quality checking as it is usually done in case of its matching parts like mtDNA, Y chromosome and autosomal data. Its distinct mode of transmission, being present in form of single copy and sex linking poses a bit complication in analyzing the data (CARRACEDO *et al.*, 2010). They pointed out some errors in literature of X chromosome of conceptual and as of calculation type (BUCKLETON *et al.* 2016). Some researchers have highlighted issue of misusing ideas and terms (presented as of interest for forensic), ultimately given rise to a lot of miscalculations.

The parameters usually studied and calculated for X chromosome, are regarded as standard statistics (generally demanded by reviewers) consist of Expected heterozygosity (HET), typical paternity index (TPI), Polymorphic information content (PIC), Power of discrimination PD, Power of Exclusion PE and Mean exclusion chance (MEC) are usually specified for autosomal markers. Some of those are used for X chromosomal data analysis, especially a factor calculated for males and females discretely known as PD (DESMARAIS *et al.*, 1998). In spite of the fact, that PE is not for maternity, it is regularly calculated for daughters (BRENNER, 1990; SZIBOR *et al.*, 2003b). Another term TPI, which is population based and devised for analysis of autosomal marker, relies on expected heterozygosity (BRENNER, 1990; FERRAGUT *et al.*, 2019). Higher TPI value manifests the high value of heterozygosity for a given marker in population. It is suggested to revisit the grounds on basis of which this value is used .Moreover, researchers found discrepancy as in many studies including X chromosomal markers, lowest TPI were reported corresponding to highest HET (TILLMAR *et al.*, 2017).

While going through the published articles relevant to X chromosomal data, 31 out of 52 were observed to have parameters computations, which are specific for autosomal markers. Out of 52, twenty one were having outmoded parameters (HET and PD) for males. The absence of simulation based computations of markers to assess the informative capability was observed in literature as it is recommended by ISFG for X chromosomal based kinship analysis 30. The extensive increase in population studies utilizing X chromosomal STRs mark the need of evaluation of X STR nomenclature for commonly used polymorphism to ensure the accuracy and authenticity of data (GOMES *et al.*, 2020). Some of the genetic studies were not having genetic variation data for some major population groups from Africa and Asia. The first version of X chromosomal kit for commercial purpose, the Investigator Argus X-12 (Qiagen) was mostly utilized on European people for the characterization of markers in it and was devoid of genetic data from other population groups. In the light of upper mentioned mishandlings in data presentation as evident from many publications, there is a need of imposition of quality checks, pre publication quality controls and curative systems for post publications.

CONCLUDING REMARKS

There is need to conduct more population studies using X chromosomal markers to investigate this unique population genetic pattern and apply to forensic science, parentage and kinship based investigations. More in depth studies to explore it will be beneficial to scientists studying sex linked diseases. Especially, scientists studying oncology try to optimize marker panels to evaluate microsatellite instability (MSI) and loss of heterozygosity. To accomplish this, they frequently tend to consider Chr X STRs. Researches could be conducted to see full potential of other markers like mini STRs, Copy Number Variations (CNVs) and X chromosomal haploblocks for applications in population genetic studies and forensic investigations.

Received, July 27th, 2020
Accepted March 18th, 2021

REFERENCES

BRENNER, C. (1990): Paternity index calculations in single locus hypervariable DNA probes: validation and other studies. Proceedings for the international symposium on human identification 1989. Madison, WI 1990, Promega: 21-53.

BUCKLETON, J., J., CURRAN, J., GOUDET, D., TAYLOR, A., THIERY *et al.* (2016): Population-specific FST values for forensic STR markers: A worldwide survey. Forensic Science International: Genetics, *23*: 91-100.

BUTLER, J. M. (2011): Advanced topics in forensic DNA typing: methodology. Academic press.

BUTLER, J.M. (2015): US initiatives to strengthen forensic science & international standards in forensic DNA. Forensic Science International: Genetics, *18*: 4-20.

CAPUTO, M., M., AMADOR, S., SANTOS, D., CORACH (2017): Potential forensic use of a 33 X-InDel panel in the Argentinean population. Int. J. Legal Med., *131*: 107-112.

CARRACEDO, Á., J.M., BUTLER, L., GUSMÃO, W., PARSON, L., ROEWER *et al.* (2010): Publication of population data for forensic purposes. Forensic Sci. Int.: Genetics, *4*: 145-147.

CONSORTIUM, U. (2018): UniProt: the universal protein knowledgebase. Nuc. Ac. Res., *46*: 2699.

DESMARAIS, D., Y., ZHONG, R., CHAKRABORTY, C., PERREAULT, L., BUSQUE (1998): Development of a highly polymorphic STR marker for identity testing purposes at the human androgen receptor gene (HUMARA). J. Forensic Sci., *43*: 1046-1049.

DI SANTO MEZTLER, G.P., L.A., GLESMANN, M.E., ESTEBAN, S., DEL PALACIO, M.G., MÉNDEZ *et al.* (2019): Comparative Study of 10 X-STR Markers in Populations of Northeast Argentina. Human Biol., *91*: 9-20.

EDELMANN, J., M., KOHL, J., DRESSLER, A., HOFFMANN (2016): X-chromosomal 21-indel marker panel in German and Baltic populations. Int. J. Legal Med., *130*: 357-360.

FAN, G., Y., YE, H., LUO, Y., HOU (2015): Screening of Multi-InDel markers on X-chromosome for forensic purpose. Forensic Science International: Genetics Supplement Series, *5*: e42-e44.

FERRAGUT, J., N., PINTO, A., AMORIM, A., PICORNELL (2019): Improving publication quality and the importance of Post Publication Peer Review: The illustrating example of X chromosome analysis and calculation of forensic parameters. Forensic Sci. Int.: Genetics, *38*: e5-e7.

FREITAS, N.S., R.L., RESQUE, E.M., RIBEIRO-RODRIGUES, J.F., GUERREIRO, N.P., SANTOS *et al.* (2010): X-linked insertion/deletion polymorphisms: forensic applications of a 33-markers panel. Int. J. Legal Med., *124*: 589-593.

GEER, L. Y., A. MARCHLER-BAUER, R. C. GEER, L. HAN, J. HE *et al.* (2010): The NCBI biosystems database. Nuc. Ac. Res., *38*: D492-D496.

GOMES, I., N. PINTO, S. ANTÃO-SOUSA, V. GOMES, L. GUSMÃO *et al.* (2020): Twenty years later: a comprehensive review of the X chromosome use in forensic genetics. Frontiers in Genetics: 926.

GUNTER, C. (2005): She moves in mysterious ways. Nature, *434*: 279-280.

GUSMÃO, L., P. SÁNCHEZ-DIZ, C. ALVES, I. GOMES, M. T. ZARRABEITIA *et al.* (2009): A GEP-ISFG collaborative study on the optimization of an X-STR decaplex: data on 15 Iberian and Latin American populations. Int. J. Legal Med., *123*: 227-234.

HE, G., Y. LI, X. ZOU, P. LI, P. CHEN *et al.* (2017): Forensic characteristics and phylogenetic analyses of the Chinese Yi population via 19 X-chromosomal STR loci. Int. J. Legal Med., *131*: 1243-1246.

LI, L., C. LI, S. ZHANG, S. ZHAO, Y. LIU *et al.* (2010): Analysis of 14 highly informative SNP markers on X chromosome by TaqMan® SNP genotyping assay. Forensic Sci. Int.: Genetics, *4*: e145-e148.

LIU, Q.-L., D.-J. LU, X.-G. LI, H. ZHAO, J.-M. ZHANG *et al.* (2011): Development of the nine X-STR loci typing system and genetic analysis in three nationality populations from China. International J. Legal Med., *125*: 51-58.

PEREIRA, R., V. PEREIRA, I. GOMES, C. TOMAS, N. MORLING *et al.* (2012): A method for the analysis of 32 X chromosome insertion deletion polymorphisms in a single PCR. Int. J. Legal Med., *126*: 97-105.

ROOS-ARAUJO, D. (2019): Investigation of Xq chromosomal variation in relation to migraine, pp. Queensland University of Technology.

ROSS, M. T., D. V. GRAFHAM, A. J. COFFEY, S. SCHERER, K. MCLAY *et al.* (2005): The DNA sequence of the human X chromosome. Nature, *434*: 325-337.

STEPANOV, V., K. VAGAITSEVA, V. KHARKOV, A. CHEREDNICHENKO, A. BOCHAROVA (2016a): Panel of X-linked single-nucleotide polymorphic markers for DNA identification (XSNPid) based on multiplex genotyping by multilocus PCR and MALDI-TOF mass spectrometry. Mol. Biol., *50*: 387-397.

STEPANOV, V., K. VAGAITSEVA, V. KHARKOV, A. CHEREDNICHENKO, A. BOCHAROVA *et al.* (2016b): Forensic and population genetic characteristics of 62 X chromosome SNPs revealed by multiplex PCR and MALDI-TOF mass spectrometry genotyping in 4 North Eurasian populations. Legal Medicine, *18*: 66-71.

SZIBOR, R. (2007): The X chromosome in forensic science: past, present and future. Mol. Forensics: 103-126.

SZIBOR, R., J. EDELMANN, S. HERING, I. PLATE, H. WITTIG *et al.* (2003a): Cell line DNA typing in forensic genetics—the necessity of reliable standards. Forensic Sci. Int., *138*: 37-43.

SZIBOR, R., M., KRAWCZAK, S., HERING, J., EDELMANN, E., KUHLISCH *et al.* (2003b): Use of X-linked markers for forensic purposes. Int. J. Legal Med., *117*: 67-74.

TILLMAR, A. O., D., KLING, J.M., BUTLER, W., PARSON, M., PRINZ *et al.* (2017): DNA Commission of the International Society for Forensic Genetics (ISFG): Guidelines on the use of X-STRs in kinship analysis. Forensic Sci. Int.: Genetics, *29*: 269-275.

TOMAS, C., J.J., SANCHEZ, J.A., CASTRO, C., BØRSTING, N., MORLING (2010): Forensic usefulness of a 25 X-chromosome single-nucleotide polymorphism marker set. Transfusion, *50*: 2258-2265.

TURRINA, S., G., FILIPPINI, D., DE LEO (2011): Forensic evaluation of the Investigator DIPplex typing system. Forensic Science International: Genetics Supplement Series, *3*: e331-e332.

YANG, X., X., ZHANG, J., ZHU, L., CHEN, C., LIU *et al.* (2017): Genetic analysis of 19 X chromosome STR loci for forensic purposes in four Chinese ethnic groups. Sci. Rep., *7*: 1-11.

ZHANG, S., K., SUN, Y., BIAN, Q., ZHAO, Z., WANG *et al.* (2015): Developmental validation of an X-Insertion/Deletion polymorphism panel and application in HAN population of China. Sci. Rep., *5*: 1-7.

X HROMOZOMSKA ANALIZA U POPULACIONOJ GENETICI I FOREZNICI: MALI PREGLEDNI RAD

Azam ALI[1,2*], Nazim HUSSAIN[2], Muhammad Saqib SHAHZAD[3], Inamullah[2], Qurban ALI[2*]

[1]Institut za molekularnu biologiju i biotehnologiju (IMBB), Univerzitet u Lahoru, Lahor, Pakistan
[2]Centar za primenjenu molekularnu biologiju (CAMB), Univerzitet u Pendžabu, Lahor, Pakistan
[3]Departman za forezniku, Univerzitet za zdravstvene nauke, Khayaban-e-Jamia Pendžab, Lahor, Pakistan

Izvod

Analiza humanog X hromozoma primenjena je za dešifrovanje genetske strukture populacija za primenu u medicinskoj genetici i za identifikaciju ljudi, analizu roditeljstva i analizu srodstva. Iako nije proučavan na velikom nivou u odnosu na ljudsku populaciju u poređenju sa drugim njegovim pandanima poput autozomnih markera, Y hromozoma i mtDNK, ipak je važan za veliki potencijal u proučavanju onkologije, različitih bolesti i primena forenzičke nauke. U ovom mini pregledu, napravljen je snimak X hromozomskih svojstava kao genetskog markera. Raspravljalo se o strukturi i potencijalnim kompletima orijentisanim na multipleks koji koriste X hromozomske markere. Štaviše, zabrinutost različitih istraživača u vezi sa objavljenim podacima o X hromozomima ukazuje na potrebu analize X hromozomskih markera kako bi se otkrila njihova uloga u populacionoj genetici, medicinskoj genetici i identifikaciji ljudi.

Primljeno 27.VII.2020.
Odobreno 18. III. 2021.