Int J Legal Med (2009) 123:85–87
DOI 10.1007/s00414-008-0294-7

SHORT COMMUNICATION

Haplotype analysis of two X-chromosome STR clusters in the Pakistani population

Muhammad Akram Tariq · Muhammad Farooq Sabir · S. Amer Riazuddin · Sheikh Riazuddin

Received: 2 June 2008 / Accepted: 27 October 2008 / Published online: 15 November 2008

Abstract Haplotype analysis of closely associated markers has proven to be a powerful tool in kinship analysis especially when X-chromosome short tandem repeats fail to resolve uncertainty in relationship analysis. Microsatellites located on the X chromosome show stronger linkage disequilibrium compared with autosomal microsatellites; hence, it is necessary to estimate the haplotype frequencies directly from population studies as linkage disequilibrium is population-specific. Here, we describe five markers residing in two clusters; cluster I harboring three STR markers DXS6801–DXS6809–DXS6789 and cluster II harboring two STR markers DXS7424–DXS101. A total of 302 male DNA samples of Pakistani descent were analyzed. Theoretically, 847 and 160 different combinations of haplotypes are possible in clusters I and II, but genotyping identified only 129 and 75 haplotypes, respectively. No evidence of linkage disequilibrium was detected, except for the pair (DXS6801–DXS6789), consistent with results obtained with the cluster I in a German population. Our results demonstrate that 83% haplotypes of cluster I and 65% haplotypes of cluster II show <1% frequency in the Pakistani population. This strongly suggests that haplotypes of these two clusters provide a powerful tool for kinship testing and relationship investigations.

Keywords Haplotyping · X chromosome · X-STR · Relationship testing · Pakistani population

Introduction

X chromosome short tandem repeats (X-STRs) are routinely used in parentage analysis and relationship investigations such as avuncular and first cousin relationships [1]. In addition to X-STRs, stable haplotypes of closely associated X-chromosome markers have proven to be a powerful tool in kinship analysis [2]. X-STRs show stronger linkage disequilibrium (LD) compared with autosomal STR since the X-chromosome recombines only in females [1]. LD between X-STR loci has proven to be population-specific and is further dependent on factors such as random drift, founder effect, mutation rates, selection, and population admixture [3, 4]. Hence, appropriate use of X-STR haplotypes requires that haplotype frequencies are estimated directly from the appropriate population samples.

Previously, two X-STR clusters have been developed for kinship testing; cluster I with DXS6801–DXS6809–DXS6789 located on chromosome Xq21 [1] and cluster II with DXS7424–DXS101 on Xq22 [5]. These markers possess high powers of discrimination (PD) compared with most X chromosome STRs, e.g., PD in females is 0.816, 0.953, 0.930, 0.871, and 0.940 for DXS6801, DXS6809, DXS6789, DXS7424, and DXS101, respectively [3, 6–14]. The five STR markers have been studied in different populations and allelic frequencies have been estimated. Here, we report haplotype analysis of these two X-STR

Electronic supplementary material The online version of this article (doi:10.1007/s00414-008-0294-7) contains supplementary material, which is available to authorized users.

This study was supported in part by Higher Education Commission and Ministry of Science and Technology Islamabad, Pakistan.

M. A. Tariq · M. F. Sabir · S. A. Riazuddin (✉) · S. Riazuddin (✉)
National Centre of Excellence in Molecular Biology,
University of the Punjab,
87 West Canal Bank Road,
Lahore 53700, Pakistan
e-mail: amerriazuddin@ncemb.org
e-mail: riaz@lhr.comsats.net.pk

clusters in the Pakistani population. Our results indicate that 83% of haplotypes of cluster I and 65% of haplotypes of cluster II show <1% frequency, which makes these clusters useful for haplotype analysis in kinship testing and relationship investigations.

Materials and methods

Informed consent according to the declarations of Helinsaki was obtained from all the participants prior to the study. Blood samples were collected from 302 unrelated Pakistani males. DNA was extracted using the phenol-chloroform method [15]. Primer sequences for DXS6801, DXS7424 and DXS6809 were obtained from the Genome Database (http://www.gdb.org). Primers for DXS6789 and DXS101 were designed by using primer3 software (http://primer3.sourceforge.net). All the primer pairs were analyzed by the AutoDimer software (http://www.cstl.nist.gov/div831/strbase/AutoDimerHomepage.htm) to avoid primer dimer formation in multiplex polymerase chain reaction (PCR) [16]. Forward primers of all five markers were labeled with fluorescein dyes at the 5′ ends. The forward primer of DXS6789 was labeled with NED, DXS101 and DXS6801 were labeled with VIC, and DXS7424 and DXS6809 were labeled with PET dye.

Multiplex PCR was performed in 20-μl reaction volumes containing 2 ng of genomic DNA, 75 mM Tris HCl, 20 mM $(NH)_2$ SO_4, 2 mM $MgCl_2$, 200 μM each dNTP, and three units of Taq DNA polymerase. Primer concentrations of 0.085, 0.0825, 0.225, 0.375, and 0.75 μM for DXS6789, DXS6801, DXS101, DXS7424, and DXS6809, respectively, were used in the multiplex PCR reactions. Amplification was performed in a Thermal Cycler 2700 System (Applied Biosystems, Foster City, CA, USA) and the PCR program consisted of a denaturation step at 95°C for 5 min followed by 28 cycles consisting of denaturing at 94°C for 1 min, annealing at 57°C for 1 min, and extension at 65°C for 2 min, with a final extension at 60°C for 60 min.

Electrophoresis was performed in a 16-capillary ABI PRISM 3100 (Applied Biosystems) Genetic Analyzer. Briefly, samples were prepared by adding 1 μl of PCR product to 11.8 μl deionized formamide (Applied Biosystems) and 0.2 μl GeneScan 500 LIZ size standard (Applied Biosystems). Results were analyzed using GeneScan Analysis software 3.7 (Applied Biosystems). Alleles were designated according to the suggestion of Edelmann et al. for DXS6789, DXS7424, DXS101, DXS6801, and DXS6809 and according to the recommendations of Szibor et al for standard DNA 9947A (Promega, Madison, WI, USA) [7, 8, 17–19]. The common alleles from each marker were sequenced, and these sequenced alleles were used as an allelic ladder.

The presence of new alleles was confirmed by bidirectional sequencing. Primers were designed using the primer3 program (http://primer3.sourceforge.net). Amplifications were performed in 25 μl reactions containing 50 ng of genomic DNA, 2.5 μl 10X Gene Amp PCR buffer II, 8 pmol of each primer, 2.5 mM dNTP, 2.5 mM $MgCl_2$, and 0.2 U *Taq* DNA polymerase. The PCR consisted of a denaturation step at 96°C for 5 min, followed by 40 cycles each consisting of 96°C for 45 s followed by 57°C for 45 s and 72°C for 1 min. The PCR products were used for bidirectional sequencing using Big Dye Terminator Ready reaction mix according to the manufacturer's instructions (Applied Biosystems). Sequencing products were resuspended in 10 μl of formamide and denatured at 95°C for 5 min. Sequencing was performed on an ABI PRISM 3100 Automated sequencer (Applied Biosystems). Sequencing results were assembled using ABI PRISM sequencing analysis software version 3.7 (Applied Biosystems) and analyzed with Chromas software (http://www.technelysium.com.au/chromas.html).

Haplotype frequencies were estimated from genotypes of male DNA samples (http://www.causascientia.org/math_stat/ProportionCI.html). LD tests were performed for all the pairs of loci of the two clusters using ARLEQUIN version 3.1 [20]. Allelic frequencies were calculated at each locus by PowerStatsV12 program (http://www.promega.com/geneticidtools/). Mean exclusion chance in Duos (MEC_D), mean exclusion chance in trios (MEC_T), PD for females (PD_F), and PD in males (PD_M) were calculated by using chromosome X web software (http://www.chrx-str.org) [21].

Results and discussion

Here, we report haplotyping results of two X-chromosome (X-STR) clusters, cluster I (DXS6801–DXS6809–DXS6789) and cluster II (DXS7424–DXS101), in the Pakistani population. A total of 302 DNA samples from unrelated Pakistani males were analyzed in a pentaplex PCR. Our results demonstrate that 83% of haplotypes of cluster I and 65% of haplotypes of cluster II have <1% frequency in the Pakistani population. This strongly suggests that haplotypes of these clusters provide a powerful tool for kinship testing and relationship investigations.

In cluster I, we identified seven alleles for marker DXS6801 and 11 alleles for markers DXS6809 and DXS6789, respectively (Supplemental Table 1). Hence, theoretically, 847 different combinations of haplotype are possible, but genotyping identified only 129 different haplotypes. Most of the haplotypes (83%) had <1% frequency in the Pakistani population. The two most abundant haplotypes (11–33–20 and 11–34–20) illustrated frequencies of 5% each, respectively. Due to space

limitations, all haplotypes observed, along with their frequencies, are given in Supplemental Table 2.

In cluster II, we identified 16 alleles for marker DXS101 and 10 alleles for DXS7424 (Supplemental Table 1). Consequently, 160 different combinations of haplotypes are possible, while genotyping identified only 75 different haplotypes. The majority of the haplotypes (65%) had <1% frequency in the Pakistani population. The two most abundant haplotypes (15–24 and 16–24) showed frequencies of 8% and 6%, respectively. Due to space limitations, all haplotypes observed along with their frequencies are given in Supplemental Table 3.

The allelic frequencies and other parameters of all the five X-linked STRs are shown in Supplemental Table 4. A total of 7–16 alleles for each locus and, altogether, 55 alleles were observed. The allele distribution of all five STRs in the Pakistani population is similar to other reported populations; however, a rare allele 9 of DXS7424 was observed in the Pakistani population that was only reported in Italian and German populations [17, 22].

The exact test for LD was performed for all pairs of the five STR markers (DXS6801, DXS6809, DXS6789, DXS7424, and DXS101). No evidence of LD was detected, except for the pair (DXS6801–DXS6789), which showed significant association (p=0.0204), consistent with results obtained with the cluster DXS6801–DXS6809–DXS6789 in Germans [1]. Although an estimated p value of 0.0204 was observed for markers DXS6801–DXS6789, this p value is above the significance level of 0.0166 (obtained after Bonferroni correction), suggesting no evidence of association between the above mentioned pair of STR markers.

Conclusions

In conclusion, the data presented here strongly suggests that the haplotypes of these clusters provide a powerful tool for haplotype analysis in kinship testing and relationship investigations. Human population characteristics at the genetic level are integral to both forensic biology and population genetics and, to the best of our knowledge, this is the first study estimating the haplotypes frequency for X-STR markers in the Pakistani population.

Acknowledgements The authors are grateful to all individuals for their participation in this study. This work was supported in part by Higher Education Commission, Islamabad, Pakistan and Ministry of Science and Technology, Islamabad, Pakistan.

References

1. Szibor R, Hering S, Kuhlisch E, Plate I, Demberger S, Krawczak M, Edelmann J (2005) Haplotyping of STR cluster DXS6801–DXS6809–DXS6789 on Xq21 provides a powerful tool for kinship testing. Int J Legal Med 119:363–369
2. Hering S, Augustin C, Edelmann J et al (2006) DXS10079, DXS10074 and DXS10075 are STRs located within a 280-kb region of Xq12 and provide stable haplotypes useful for complex kinship cases. Int J Legal Med 120:337–345
3. Lee HY, Park MJ, Jeong CK et al (2004) Genetic characteristics and population study of 4 X-chromosomal STRs in Koreans: evidence for a null allele at DXS9898. Int J Legal Med 118:355–360
4. Pereira R, Gomes I, Amorim A, Gusmao L (2007) Genetic diversity of 10 X chromosome STRs in northern Portugal. Int J Legal Med 121:192–197
5. Edelmann J, Hering S, Kuhlisch E, Szibor R (2002) Validation of the STR DXS7424 and the linkage situation on the X-chromosome. Forensic Sci Int 125:217–222
6. Edelmann J, Szibor R (2001) DXS101: a highly polymorphic X-linked STR. Int J Legal Med 114:301–304
7. Edelmann J, Deichsel D, Plate I, Kaser M, Szibor R (2003) Validation of the X-chromosomal STR DXS6809. Int J Legal Med 117:241–244
8. Edelmann J, Szibor R (2005) Validation of the X-linked STR DXS6801. Forensic Sci Int 148:219–220
9. Asamura H, Sakai H, Kobayashi K, Ota M, Fukushima H (2006) MiniX-STR multiplex system population study in Japan and application to degraded DNA analysis. Int J Legal Med 120:174–181
10. Shin KJ, Kwon BK, Lee SS et al (2004) Five highly informative X-chromosomal STRs in Koreans. Int J Legal Med 118:37–40
11. Shin SH, Yu JS, Park SW, Min GS, Chung KW (2005) Genetic analysis of 18 X-linked short tandem repeat markers in Korean population. Forensic Sci Int 147:35–41
12. Gu S, Li S (2006) X-chromosome STRs analysis of Ewenke ethnic population. Forensic Sci Int 158:72–75
13. Liu Q, Li S (2006) Patterns of genetic polymorphism at the 10 X-chromosome STR loci in Mongol population. Forensic Sci Int 158:76–79
14. Kang L, Li S (2006) X-chromosome STR polymorphism of Luoba ethnic group living in Tibet (SW China). Forensic Sci Int 156:88–90
15. Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor
16. Butler JM, Ruitberg CM, Vallone PM (2001) Capillary electrophoresis as a tool for optimization of multiplex PCR reactions. Fresenius J Anal Chem 369:200–205
17. Edelman J, Hering S, Michael M et al (2001) 16 X-chromosome STR loci frequency data from a German population. Forensic Sci Int 124:215–218
18. Bar W, Brinkmann B, Budowle B et al (1997) DNA recommendations. Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. International Society for Forensic Haemogenetics. Int J Legal Med 110:175–176
19. Szibor R, Edelmann J, Hering S et al (2003) Cell line DNA typing in forensic genetics—the necessity of reliable standards. Forensic Sci Int 138:37–43
20. Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinformatics Online 1:47–50
21. Szibor R, Hering S, Edelmann J (2006) A new web site compiling forensic chromosome X research is now online. Int J Legal Med 120:252–254
22. Robino C, Giolitti A, Gino S, Torre C (2006) Development of two multiplex PCR systems for the analysis of 12 X-chromosomal STR loci in a northwestern Italian population sample. Int J Legal Med 120:315–318