Legal Medicine 19 (2016) 43–46

Contents lists available at ScienceDirect

Legal Medicine

journal homepage: www.elsevier.com/locate/legalmed

Short Communication

# Detection of a novel X-chromosomal short tandem repeat marker in Xq28 in four ethnic groups

Takeki Nishi [a,*], Takako Nakamura [b], Katuya Honda [b]

[a] Department of Forensic Medicine, The Jikei University School of Medicine, 3-25-8 Nishi-shinbashi, Minato-ku, 105-8461 Tokyo, Japan
[b] Department of Legal Medicine, Faculty of Medicine, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8575, Japan

ARTICLE INFO

*Article history:*
Received 28 July 2015
Received in revised form 28 December 2015
Accepted 27 January 2016
Available online 27 January 2016

*Keywords:*
X chromosome
Short tandem repeat
X-STR
STR sequence
Population

ABSTRACT

DNA testing of X-chromosomal short tandem repeat (X-STR) polymorphisms has been the focus of attention in several studies, mainly due to its applicability in the investigation of complex kinship cases. Studies of X-STR in analyses of DNA sequences, population studies and DNA testing applications have been reported. We performed detection and population genetic study of a novel tetranucleotide X-STR locus in the present study. We identified a unique X-STR locus consisting of two tetranucleotides in Xq28. Although the STR is a simple tetranucleotide, its polymorphism was comparatively high [polymorphism information content (PIC) = 0.7140] in Japanese subjects. In addition, the STR varied in structure among ethnic groups. We conclude that this locus will be useful for forensic DNA testing and anthropological studies.



## 1. Introduction

Several autosomal and gonosomal short tandem repeats (STRs) are useful tools for determining individual identities. Currently, forensic DNA testing with autosomal STR and Y-chromosomal STR (Y-STR) is commonly used to identify individuals. Commercially available kits for both chromosomal STRs are commonly used in forensic DNA testing. There are many reports of population and comparative studies using Identifiler (Applied Biosystems, Foster City, CA, USA) [1–3], PowerPlex (Promega Corporation, Madison, WI, USA) [2,3], Globalfiler (Applied Biosystems) [4,5], Yfiler (Applied Biosystems) [6,7] and Powerplex Y (Promega) [7].

DNA testing of X-chromosomal STR (X-STR) polymorphisms has been the focus of attention in several studies, mainly due to its applicability in the investigation of complex kinship cases, such as deficient-paternity cases [8,9]. Analyses of DNA sequences [10–13], population genetic studies and multiplex polymerase chain reactions (PCR) for application in DNA testing [14–16] continue to be reported. Recently, some studies of X-STR linkage groups have been reported. A commercially available kit for X-STR multiplex PCR, investigator Argus X-12 kit (Qiagen GmbH, Hilden, Germany), contains X-STR linkage groups. As for linkage groups carrying X-STR, five regions, p22, centromere, q12, q26 and q28, have been reported [9,17–19]. In this study, we performed the detection of a novel tetranucleotide X-STR locus in Xq28 and a population genetic study of four ethnic groups.

## 2. Materials and methods

### 2.1. Sample and DNA extraction

With the approval of an ethics committee of the University of Tsukuba, samples were collected from 610 unrelated individuals of four nationalities. They included 249 Japanese (150 males and 99 females), 93 Mongolians (81 males and 12 females), 88 Caucasian-Americans (33 males and 55 females) and 180 native American Colombians (87 males and 93 females). DNA was extracted with a Maxwell 16 Instrument (Promega). The extracted DNA was quantitated with a UV spectrophotometer (DU Series 700; Beckman Coulter, CA, USA). This study involving human material or human data was approved by the Institutional Review Board of the University of Tsukuba.

### 2.2. Detection, sequence analysis and PCR products

For the sequence analysis, the study adopted 100 Japanese (100 males), 93 Mongolians (81 males and 12 females), 88 Caucasian-Americans (33 males and 55 females), and 180 native American Colombians (87 males and 93 females). A heterozygote DNA samples from females were analyzed after separating it with polyacrylamide electrophoresis.

* Corresponding author.
*E-mail address:* mrpork24@jikei.ac.jp (T. Nishi).

http://dx.doi.org/10.1016/j.legalmed.2016.01.010
1344-6223/

We used the University of California, Santa Cruz (UCSC) Genome Browser (http://www.genome.ucsc.edu/) for the detection of several tetranucleotide repeats on the X chromosome. New primers were designed using PRIMER3 (http://www.genome.wi.mit.edu/cgi-bin/primer/primer3) software. PCR reactions were performed in a volume of 20 μl with a reaction mix containing 5–15 ng of genomic DNA, 2× AmpliTaq Gold® 360 Master Mix (Applied Biosystems) and a primer set. Primer sequences are listed in Table 1. The amplification conditions were optimised in a Veriti 96-well Thermal Cycler (Applied Biosystems) and consisted of initial denaturation at 95 °C for 10 min, followed by 28 cycles of denaturing at 95 °C for 30 s, annealing at 60 °C for 30 s, extension at 72 °C for 30 s and a final extension at 72 °C for 7 min. PCR for sequencing was performed using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Primers used for sequence analysis were the same as those used for PCR. Excessive dye was removed using the BigDye XTerminator Purification Kit (Applied Biosystems). The product was sequenced using the ABI 3130 Genetic Analyser (Applied Biosystems).

### 2.3. Base length analysis

For the base length analysis, the study adopted 249 Japanese (150 males 99 females), 93 Mongolians (81 males and 12 females), 88 Caucasian-Americans (33 males and 55 females), and 180 native American Colombians (87 males and 93 females).

PCR products were separated using an ABI 3130 Genetic Analyser (Applied Biosystems) in which 1 μl of PCR product was mixed with 14.5 μl of Hi-Di formamide and 0.5 μl of GeneScan-600LIZ size standard (Applied Biosystems). Gene Mapper ID software v3.2 (Applied Biosystems) was used to analyse the data. K562(8–13) and 9947(7–11/7–13) of human cell line DNA with already-known repetition number confirmed by the sequence analysis were used as control.

### 2.4. Statistical analysis

We calculated several forensic statistical parameters using previously described methods: polymorphism information content (PIC), homozygosity (h), heterozygosity (Het), power of exclusion (PE), power of discrimination (PD) in females and males and mean exclusion chance (MEC) for trios involving daughters and for father/daughter duos lacking maternal genotype information [20–24].

## 3. Results

We identified a new tetranucleotide locus approximately 100 kbp downstream from DXS10011, which is located in the USCG Genome Browser GRCh38 X-chromosome 151,866,487 bp (accession number LC050343-LC50360). It was a combination of two tetranucleotide repeats, TGCC and TTCC. In four ethnic groups, this unique repeat showed 5–9 TGCC and 10–22 TTCC repeat units. The control sample, K562, carried $(TGCC)_8(TTCC)_{13}$ (Table 2) and 9947 carried $(TGCC)_7(TTCC)_{11}$ and $(TGCC)_7(TTCC)_{13}$.

The analysis of DNA sequences and allele frequencies of four ethnic groups is shown in Table 3. The highest allele frequency for TGCC was of allele 8 in all ethnic groups. Allele 5 of TGCC was observed in Mongolians, Americans and Colombians but not in Japanese, and allele 9 was observed in Americans and Colombians. The frequencies of alleles 6, 7 and 8 of TGCC in Japanese, Americans and Colombians were approximately the same. In contrast, in Mongolians, the frequency of allele 6 was 0.2095 and that of allele 7 was 0.2857. TTCC in Japanese, Mongolians and Americans showed the highest frequency of approximately 0.5 for allele 15. The second highest frequency was 0.2 for allele 14. However, the allele frequencies of TTCC in Colombians differed sharply: 0.4249 for allele 14 and 0.3479 for allele 15. The TTCC allele numbers in Japanese and Mongolians widely ranged from 10 to 21. In contrast, it ranged in Americans from 10 to 18 and in Colombians from 12 to 16.

Haplotype analysis of TGCC and TTCC is described in Table 4. $(TGCC)_8(TTCC)_{15}$ (0.27) showed the highest frequency in Japanese, followed by $(TGCC)_8(TTCC)_{14}$ (0.14), $(TGCC)_6(TTCC)_{15}$ (0.12) and $(TGCC)_7(TTCC)_{15}$ (0.11). The frequencies of TGCC-TTCC in Mongolians were $(TGCC)_8(TTCC)_{15}$ (0.2285), $(TGCC)_8(TTCC)_{14}$ (0.1333) and $(TGCC)_7(TTCC)_{15}$ (0.1142). These results of haplotype analysis were rarely different from each allele frequency. In Americans, $(TGCC)_8(TTCC)_{15}$ (0.2447) showed the highest frequency, followed by $(TGCC)_6(TTCC)_{15}$ (0.1258), $(TGCC)_6(TTCC)_{14}$ (0.0839). The combination of allele 8 with the highest frequency of TGCC and allele 14 with the second highest frequency of TTCC occurred in low frequency, 0.0559. The frequencies for TGCC–TTCC in Colombians were $(TGCC)_6(TTCC)_{14}$ (0.2454), $(TGCC)_8(TTCC)_{15}$ (0.2344) and $(TGCC)_8(TTCC)_{14}$ (0.1392). These exceeded the frequency of

**Table 1**
Primer sequences used in this study.

| Primers | Primer sequence | Position of the primer (bp)[a] |
|---|---|---|
| Forward | 5′-TCTGACTACAGCACTTATCAAAGACA-3′ | 15,18,66,358 |
| Reverse | 5′-AAGGAGTTCACTTTCATCACGTC-3′ | 15,18,66,658 |

[a] Location on X-chromosome (USCG genome browser, GRCh38).

**Table 2**
Sequence of K562 control sample.

| Accession number LC050358 |
|---|
| TCTGACTACAGCACTTATCAAAGACACTAGTTTCAGACTT<br>TTAAACAAGATTCAGTCAGATTTTCTATGTAGATTATAAT<br>AATGTCTGCAAATAAAGGGTTTTTATTTTTCTTTTTCAATC<br>TACATGCTTGCCTGCCTGCCTGCCTGCCTGCCTGCC<br>TTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTT<br>CCTTCCTTCCTTCTTCTTCCCTTATTTTTCTTCTGGCTTTA<br>TTGCAGTGGTTAGAATCTCTAGTACAATGTTAAATAGACG<br>TGATGAAAGTGAACTCCTTATCTT |

**Table 3**
Allele frequency of tetranucleotides in four ethnic populations.

| | Japanese $n = 100$ | Mongolians $n = 93$ | Americans $n = 88$ | Columbians $n = 180$ |
|---|---|---|---|---|
| *TGCC allele* | | | | |
| 5 | | 0.038 | 0.014 | 0.004 |
| 6 | 0.280 | 0.210 | 0.287 | 0.315 |
| 7 | 0.180 | 0.286 | 0.182 | 0.168 |
| 8 | 0.540 | 0.467 | 0.490 | 0.484 |
| 9 | | | 0.028 | 0.029 |
| *TTCC allele* | | | | |
| 10 | | 0.010 | 0.007 | |
| 11 | 0.010 | | 0.014 | |
| 12 | 0.020 | 0.010 | 0.028 | 0.011 |
| 13 | 0.030 | 0.019 | 0.077 | 0.088 |
| 14 | 0.200 | 0.257 | 0.210 | 0.425 |
| 15 | 0.500 | 0.438 | 0.462 | 0.348 |
| 16 | 0.160 | 0.171 | 0.126 | 0.121 |
| 17 | 0.020 | 0.010 | 0.056 | |
| 18 | | 0.010 | 0.021 | 0.004 |
| 19 | 0.010 | 0.019 | | |
| 20 | 0.040 | 0.038 | | 0.004 |
| 21 | 0.010 | 0.019 | | |