International Congress Series 1239 (2003) 425–429

Sequence structure of 12 novel Y chromosome microsatellites and PCR amplification strategies

Anna González-Neira [a], Paula Sánchez-Diz [a], Leonor Gusmão [b], María Victoria Lareu [a], Angel Carracedo [a,*]

[a] *Institute of Legal Medicine, University of Santiago de Compostela, E-15705 Santiago de Compostela, Spain*
[b] *IPATIMUP, University of Porto, Porto, Portugal*

Abstract

In this work, we present the sequencing data for 12 recently reported Y STR loci (DYS434, DYS435, DYS436, DYS437, DYS438, DYS439, GATA A10, GATA 7.1 were implemented for their detection). Sequenced allelic ladders were constructed and a nomenclature for these new systems is proposed based on the sequence structure and following the ISFG recommendations. GATA A4 and DYS439 are likely to be the same STR (for this reason, the results were pooled). They have the same STR structure and the alleles are always identical in the same individuals. Sequence polymorphisms were observed in the GATA C4 and DYS437 STRs.

Keywords: Y chromosome; STRs; Sequence; Nomenclature

1. Introduction

The use of Y-STR polymorphisms has become common place in forensic laboratories [1,2]. Recently, several new Y chromosome microsatellites have been reported [3,4]. In this work, we report the sequence analysis of the different alleles of each STR locus, and the construction of sequenced allelic ladders. Based on the sequencing results obtained, a nomenclature is proposed for each system following the recommendations of the DNA Commission of the ISFG.

* Corresponding author. Tel.: +34-981-582327; fax: +34-981-580336.
E-mail address: apimlang@usc.es (A. Carracedo).

0531-5131/03
PII: S0531-5131(02)00397-7

Table 1
Sequencing results

G42671 (GATA A7.2); 174–190 bp; 13–17 alleles
Consensus structure:
*PrimerA.*AGACAGATATATCTAATAGGTAGAT**(GATA)$_2$**ATAGG**(TAGA)$_2$**AGATAGG**(TAGA)$_{9-13}$**CAGAT AAGAGAGAAACAGAAATATAGTGACACAGCA.*PrimerB*

G42674 (GATA A10); 150–170 bp; 17–22 alleles
Consensus structure:
*PrimerA.*CTTATCCA**(TTTA)$_2$**TT**(CATC)$_2$(TCTT)$_2$**TCTC**(TCCA)$_2$(TATC)$_{9-14}$**TAATCTATCATCTATC.
PrimerB

G42676 (GATA H4); 356–372 bp; 26–30 alleles
Consensus structure:
*PrimerA.*TGATACACATTGATACTTTCAGCACATCACTTGTATCCTAGGAATCATCATTAAAATGT-TATGCTGAGGAGAATTTCCAAATTTA**(AGAT)$_4$**CT**(ATAG)$_3$(GTAG)$_3$(ATAG)$_{8-12}$**AATGGATAGATTA **(GATG)$_2$ (ATAG)$_4$** ATAC**(ATAG)$_2$**GTGATTTATCCTGTTAAGTTGTTTAACAAGTGGGCTATG-TAAAATTTTACTAATAT TAAACATAAGTAGTTTGTAGATTTTC(TTAT)$_2$T.*PrimerB*

G42675 (GATA A7.1); 161–181 bp; 7–12 alleles
Consensus structure:
*PrimerA.*CAAGAAGAATTATCTAGGAAAGTCAAGACAGT(AGCA)$_2$CAAGAATACCAGAGGAATCTGA-CACCTCTGAC**(ATAG)$_{7-12}$**ATAGACAAATACA.*PrimerB*

G42673 (GATA C4); 250–270 bp; 22–27 alleles
Consensus structure:
ALLELE GATAC4.A
*PrimerA.*TGCTGCTGAATGGGAGCAGAAATGCCCAATGGAATGCTCTCTTGGCTTCTCACTTTGCATA-GAATC**(TCTA)$_4$(TGTA)$_2$(TCTA)$_2$(TGTA)$_2$(TCTA)$_{10-14}$**TCACATTTTCTTTATC**(CATT)$_2$** ATTGATGGA-TATTTG GGCTGGTTCC.*PrimerB*
ALLELE GATAC4.B
*PrimerA.*TGCTGCTGAATGGGAGCAGAAATGCCCAATGGAATGCTCTCTTGGCTTCTCACTTTGCATA-GAATC**(TCTA)$_4$(TGTA)$_2$(TCTA)$_2$(TGTA)$_2$(TCTA)$_2$(TGTA)$_2$(TCTA)$_{8-11}$**TCACATTTTCTTTATC **(CATT)$_2$** GATTGATGGATATTTGGGCTGGTTCC.*PrimerB*

DYS434; 110–122 bp; 8–11 alleles
Consensus structure:
*PrimerA.*ATAGGCGTGACTGCATATCAGTCACTCTAATTA**(ATCT)$_{8-11}$**ATAA.*PrimerB*

DYS436 ; 127–142 bp; 10–15 alleles
Consensus structure:
*PrimerA.*AGGAAAGCAGG**(GTT)$_{10-15}$**TTATCACCTCTAACGCAGCTCGTCCCTTTTACTGTGTCCTGCT-TCC.*PrimerB*

DYS438; 216–236 bp; 9–13 alleles
Consensus structure:
*PrimerA.*ACAGTATA**(TTTTC)$_{9-13}$**TATTTGAAATGGAGTTTCACTCTTGTTGCCAGGCTGAAATG-CAATGG TGTGATCTCGACTCACCACAACCTCCACTTCCCAGGTTCAAGCGATTCTCCTGCAT-CAGCCTCCCAGGTAGCTG.*PrimerB*

Table 1 (*continued*)

DYS435; 218–226 bp; 15–17 alleles
Consensus structure:
*PrimerA.*TGGCAACTGAAGGACAGTAAGTACACACTGTCCACAGCCAGGGTTGTCCAGAGAAACAGC-CAATAAGATGTGTGGATGTG**(TGGA)$_{11-13}$**CAGACAGCTATAGAAATAC **(AGAT)$_2$** AAACCAATGAG-TAGATT **(ATAG)$_2$**AC.*PrimerB*

DYS439; 240–252 bp; 19–22 alleles
Consensus structure:
*PrimerA.*TCTTCTCGAGTTGTTATGGTTTTAGGTCTAACATTTAAGTCTTTA**(ATCT)$_2$**TGAATTAATA-GATTCAAGGT**(GATA)$_2$**TAC**(AGAT)$_3$**ACATAGGTGGAGAC**(AGAT)$_2$**GATAAATAGA**(AGAT)$_{10-13}$** AGAAAGTATA AGTAAAGAGATGATGGGT.*PrimerB*

DYS437; 184–192 bp; 16–18 alleles
Consensus structure:
ALLELE DYS437.A
*PrimerA.*GCCCATCCGG**(TCTA)$_{8-10}$(TCTG)$_2$(TCTA)$_4$**TCATCTATCATCTGTGAATGATG**(TCTA)$_2$** CTTATCTATGAATGATATTTATCTGTGGT(TATC)$_3$ TATA.*PrimerB*
ALLELE DYS437.B
*PrimerA.*GCCCATCCGG**(TCTA)$_{9-11}$(TCTG)$_1$(TCTA)$_4$**TCATCTATCATCTGTGAATGATG**(TCTA)$_2$** CTTATCTATGAATGATATTTATCTGTGGT(TATC)$_3$TATA.*PrimerB*

2. Material and methods

2.1. DNA samples

Samples were taken from unrelated individuals from major population groups (Europeans, Asians, Afro-Americans and Afrocaribbeans). Genomic DNA extraction: phenol-chloroform procedure and quantification using either the Quantiblot kit (PE) or fluorescence detection with DyNAQuant 200 (APB).

2.2. Amplification conditions

Primer sequences are described by White et al. [3] and Ayub et al. [4].

Multiplex I (7 STRs): GATA A10 (0.12 μM FAM), GATA A7.2 (0.08 μM TET), GATA A7.1 (1.8 μM HEX), GATA C4 (0.08 μM FAM), DYS438 (0.6 μM FAM), DYS439 (0.32 μM TET) and GATA H4 (0.12 μM FAM).

Multiplex II (4 STRs): DYS434 (0.2 μM FAM), DYS435 (0.05 μM TET), DYS436 (0.05 μM HEX), DYS437 (0.2 μM HEX).

Both multiplexes were amplified under identical conditions. The reaction mix contained 200 μM of each dNTPs, 2 mM $MgCl_2$, 2 U Taq Gold polymerase (PE) and 5 ng of DNA template in 25 μl final volume. After a denaturation step at 94 °C for 10 min, a touchdown PCR protocol was performed over 12 cycles beginning at 94 °C for 1 min, 64 °C for 1 min and 68 °C for 2 min. The annealing temperature was decreased by 0.5 °C in each cycle. These 12 cycles were followed by a further 18 cycles at 94 °C for 1 min, 58 °C for 1 min and 68 °C for 2 min and a final extension step at 68 °C for 60 min. Amplification detection was performed using an ABI 377 automatic sequencer and Genescan 2.1 analysis software (PE).

2.3. Sequencing conditions

The total of number of sequenced samples was 170. Before sequencing, PCR products were purified with Centricon 100 columns (APB). A dideoxy cycle sequencing reaction was carried out using ABI dRhodamine Dye Terminator kit (PE). The products were purified with Centri-Sep Spin Columns (Princeton Separations), dried in a vacuum centrifuge, re-suspended in 4 μl of Genescan blue formamide loading buffer, and run in a 6% denaturing gel on an ABI 377 sequencer. The results were analyzed using Sequence Analysis 3.3 and Sequence Navigator 1.01 (PE).

3. Results and discussion

3.1. Sequencing data and nomenclature

Sequencing results and the consensus structure of the Y chromosome STRs described by White et al. [3] and Ayub et al. [4] are shown in Table 1. GATA A4 and DYS439 are likely to be the same STR. They have the same STR structure and the alleles are always identical in the same individuals. The only difference is the primer design. However, there are some inconsistencies in the sequences reported in GenBank for the 5′ flanking region of both systems. This requires further investigation and repeat sequence analysis of the clones where these STRs were observed.

3.2. Nomenclature

In all cases, except for three loci, the nomenclature is based on the sum of the total number of repeated units including constant and variable stretches in agreement with the recommendations of the DNA Commission of the International Society for Forensic Genetics. [5]. This would prevent nomenclature problems if further variation is found in the constant stretches when further population studies are carried out. The exceptions to this general rule are the systems GATA H4, GATA A7.1 and DYS437 where some constant repeated units are not included because of the possibility of designing new intermediate primers between these stretches and the variable stretches.

In conclusion, we strongly believe that these new STRs represent a very valuable contribution to the analysis of Y chromosome variation not only in a forensic setting. Some of the loci seem to be more robust than the classical Y STRs that are already being used in most forensic laboratories. Further studies should establish the discrimination rates for different ethnic groups and continue to optimize PCR performance. We will then have a larger panel of well-characterized Y STRs and the option to select the optimal loci for use in forensic casework.

Acknowledgements

This work was supported by grants from the Xunta de Galicia XUGA 20806B97 and PIPGIDT99PXI20807B. The technical assistance of Meli Rodriguez is highly appreciated.

References

[1] M.A. Jobling, A. Pandya, C. Tyler-Smith, The Y chromosome in forensic analysis and paternity testing, Int. J. Legal Med. 110 (1997) 118–124.
[2] L. Gusmao, M.J. Brión, A. González-Neira, M.V. Lareu, A. Carracedo, Y chromosome specific polymorphisms in forensic analysis, Legal Med., (1999) 155–160.
[3] P. White, O. Tatum, L. Deaven, J. Longmire, New, male-specific microsatellite markers from the human Y chromosome, Genomics 57 (1999) 433–437.
[4] Q. Ayub, A. Mohyuddin, R. Qamar, K. Mazhar, T. Zerjal, Q. Mehdi, C. Tyler-Smith, Identification and characterisation of novel human Y-chromosomal microsatellites from sequence database information, Nucleic Acids Res. 28 (2) (2000) e8.
[5] P. Lincoln, B. Olaisen, W. Bär, W. Mayr, A. Carracedo, B. Budwole, P. Gill, DNA commission recommendations on STR typing and nomenclature, Forensic Sci. Int. 81 (1997) 181–184.