International Congress Series 1239 (2003) 287–290

Y-chromosome short tandem repeat polymorphisms: a comparison between humans and chimpanzees

Leonor Gusmão [a,*], Annabel Gonzaléz-Neira [b], Cíntia Alves [a], António Amorim [a,c], Angel Carracedo [b]

[a] *Instituto de Patologia e Immunologia Molecular da Universidade do Porto, IPATIMUP, R. Dr. Roberto Frias, s/n, 4200 Oporto, Portugal*
[b] *Institute of Legal Medicine, Universidade de Santiago de Compostela, E-15705 Santiago de Compostela, Spain*
[c] *Faculdade de Ciências, Universidade do Porto, Praça Gomes Teixeira, 4050 Oporto, Portugal*

Abstract

Y-chromosome specific microsatellites, DYS434, DYS435, DYS436, DYS437, DYS438, DYS439, GATA A10, A7.1, A7.2, C4, and H4, were typed for fragment length and sequenced in chimpanzees (*Pan troglodytes*). Primers described for GATA A4 were found to amplify the same region as reported for DYS439. Moreover, the forward primer only matches the repeat flanking region in 14 of the 28 base pairs, this being responsible for a very weak amplification. Therefore, this system was not included in this study. The analysis of the repeat and sequence structure observed in chimpanzee and human Y-chromosomes allows evolutionary comparisons to be made as well as providing a basis for improving Y STR nomenclature.

Keywords: Y-chromosome specific STRs; Sequence data; Nomenclature

1. Introduction

The importance of the establishment of a nomenclature has already been emphasised by the groups working in the forensic field, since standardisation of nomenclature is crucial to allow for second opinions, proficiency testing and exchange of data.

The aim of this work was to make a comparative analysis of the sequence structure of the Y STR loci (DYS434, DYS435, DYS436, DYS437, DYS438, DYS439, GATA A10, A7.1, A7.2, C4, and H4) in humans and in chimpanzees. We believe that this approach

* Corresponding author. Fax: +351-2255-70799.

0531-5131/03
PII: S0531-5131(02)00285-6

Table 1
Y STR sequences

```
DYS434
Hum     :P1-ataggcgtgactgcatatcagtcactc(taat)2(ctat)8-11 ---------- aa-P2
Chim    :P1-.............g.........(taat)1(ctat)3  ctgt(ctat)2cca-P2
DYS435
Hum     :P1-tggcaactgaaggacagtaagtacacactgtccacagccagggttgtccagagaaacagccaataagatgtgtgg
Chim    :P1-............................c.........................................

        atgtg(tgga)11-13cagacagctatagaaatacagatagataaaccaatgagtagattatagatagac-P2
        .....(tgga)7  ..................................................-P2
DYS436
Hum :  P1-aaggaaagcagg---(gtt)10-12,15ttatcacctctaacgcagctcgtccctttttactgtgtcctgcttcc-P2
Chim : P1-............atg(gtt)10-11  .....................a.....................-P2
DYS437
Hum     :P1-gcccatccgg----(tcta)8-10(tctg)1-2(tcta)4tcatctatcatctgtgaatgatgtctat
Chim    :P1-.........(tctg)0-1(tcta)8-10tttg(tcta)4.........................

       ctacttatctatgaatgatatttatctgtggttatctatctatctata-P2
       ..............................................-P2
DYS438
Hum   :P1-acagtata(ttttc)1(tttta)0-1(ttttc)8-13tatttgaaatggagtttcactcttgttgcccaggctgaaatgca
Chim  :P1-.......(ttttc)3,5(ttttg)0-1(ttttc)1  ........................................

       atggtgtgatctcgactcaccacaacctccacttcccaggttcaagcgattctcctgcatcagcctcccaggtagctg-P2
       ...............................................................................-P2
DYS439
Hum  :P1-tcttctcgagttgttatggttttaggtctaacatttaagtctttaatctatcttgaattaatagattcaaggt(gata)2
Chim :P1-.......a....t.....................................................................

 tacagatagatagatacataggtggagacagatagatgataaatagaa(gata)9-14 gaaagtataagtaaagagatgatgggt-P2
 ........c.........................c...........(gata)10-12.........................-P2
GATA-A7
Hum  :P1-caagaagaattatctaggaaagtcaagacagtagcaagcacaagaataccagaggaatctgacacctctgacata
Chim :P1-.....................................................................t.......

     (gata)9-12  atagacaaatacataataaatgataggcagaggatagatgatatggatagacagatatatctaataggt
     (gata)8,10-11.........................................g.........................

     agat(gata)2ataggtagatagaagatagg(taga)11-14(caga)1  taagagagaaacagaaatatagtgacacagca-P2
     ..............................(taga)10-12(caga)1-2..............................-P2
GATA-A10
Hum     :P1-cttatccatttatttattcatccatctctttctttctc(tcca)2(tatc)9-14taatc(tatc)1atctatc-P2
Chim    :P1-.........................g...................(tatc)4  .....(tatc)2.......-P2
GATA-C4
Hum  :P1-tgctgctgaatgggagcagaaatgcccaatggaatgctctcttggcttctcactttgcatagaatc(tcta)2
Chim :P1-.....................................................................a....------

      [(tcta)2(tgta)2]2,3(tcta)8-12tcacattttctttatccattcattgattgatggatatttgggctggttcc-P2
      -----------------(tcta)9-12.................................g.............c..-P2

GATA-H4
Hum   :P1-tgatacacattgatactttcagcacatcacttgtatcctaggaatcatcattaaaatgttatgctgaggagaatttc
Chim  :P1-.....................................................................................

     caaattta(agat)4ctat(agat)2(aggt)3(agat)8-12agaatggatagattagatggatga(atag)4
     ........................(aggt)6(agat)4  ........................(atag)7-11

     ------------------------------(atac)1(atag)2gtgatttatcctgttaagttgtttaacaag
      (atac)0-1(atag)0-1(atac)1(atag)2(atac)1(atag)2.............................

     tgggctatgtaaaattttactaatatttaaacataagtagtttgtagattttcttatttatt-P2
    ..............................................................-P2
```

Hum: human [3–5].
Chimp: results from the 10 chimpanzee samples used in this work.

might help to improve Y STR nomenclature as it provides a deeper insight into the variability of the polymorphisms concerned.

2. Material and methods

DNA samples from 10 male chimpanzees (*Pan troglodytes*) were kindly provided by Prof. Dr. W.R. Mayr from the University Clinic for Blood Group Serology and Transfusion Medicine (University of Vienna, Austria).

The PCR amplification was performed using 5 ng genomic DNA in a 25-μl reaction volume comprising 1.5 mM $MgCl_2$, 1 U Taq Gold polymerase (PE), 200 μM of each dNTP and 0.15 μM of each primer. The primers used were those described by Ayub et al. [1] and by White et al. [2]. Since GATA A7.1 is adjacent to A7.2, these two STRs were amplified with

Table 2
Consensus structure of 11 human Y STRs studied in this work

Locus	Previously published data on the repeat structure in humans	Proposed repeat structure in humans
Y434	$(CTAT)_n$ [1,3] $(TCTA)_n$ [5] $(ATCT)_n$ [4]	$(TAAT)_2(CTAT)_n$
Y435	$(TGGA)_n$ [1,5] $(TGGA)_nN_{19}(AGAT)_2N_{17}(ATAG)_2$ [4]	$(TGGA)_n$
Y436	$(GTT)_n$ [1,4,5]	$(GTT)_n$
Y437	$(TCTA)_n(TCTG)_2(TCTA)_4$ [1,3] $(TCTA)_1(TCAA)_{0-1}(TCTA)_n(TCTG)_{1-2}(TCTA)_4$ [5] $(TCTA)_n(TCTG)_{1-2}(TCTA)_4N_{23}(TCTA)_2$ [4] $(GATA)_4(GACA)_2(GATA)_n$ [6]	$(TCTA)_n(TCTG)_m(TCTA)_4$
Y438	$(TTTTC)_n$ [1,4,6] $(TTTTC)_1(TTTTA)_{0,1}(TTTTC)_n$ [3,5]	$(TTTTC)_1(TTTTA)_{0,1}(TTTTC)_n$
Y439	$(GATA)_n$ [1,5] AGAT [3] $(ATCT)_2N_{20}(GATA)_2N_3(AGAT)_3N_{14}$ $(AGAT)_2N_{10}(AGAT)_n$ [4] $(GATA)_2N_4(GATA)_3N_{14}(GATA)N_3$ $(GATA)N_7(GATA)_n$ [6]	$(GATA)_n$
A7.1	$(ATAG)_n$ [4] $(GATA)_n$ [2]	$(GATA)_n$
A7.2	$(GATA)_2N_5(TAGA)_2N_7(TAGA)_n$ [4] $(GATA)_n$ [2]	$(TAGA)_n(CAGA)_1$
A10	$(TTTA)_2N_2(CATC)_2(TCTT)_2N_4(TCCA)_2(TATC)_n$ [4] $(GATA)_n$ [2]	$(TCCA)_2(TATC)_n$
C4	$(TCTA)_4(TGTA)_2(TCTA)_2(TGTA)_2(TCTA)_2(TGTA)_{0,2}$ $(TCTA)_nN_{16}(CATT)_2$ [4] $(GATA)_3(CATA)_3(GATA)_2(CATA)_2(GATA)_n$ [2]	$(TCTA)_4(TGTA)_2(TCTA)_2(TGTA)_2$ $(TCTA)_2TGTA)_{0,2}(TCTA)_n$
H4	$(AGAT)_4N_2(ATAG)_3(GTAG)_3(ATAG)_nN_{13}(GATG)_2N_1$ $(ATAG)_4N_4(ATAG)_2$ [4] $(GATA)_nN_{14}(GATA)_2N_4(GATA)_4$ [2]	$(AGAT)_4CTAT(AGAT)_2(AGGT)_3$ $(AGAT)_nN_{24}(ATAG)_4(ATAC)_1(ATAG)_2$
H4.1 H4.2		$(AGAT)_4CTAT(AGAT)_2(AGGT)_3(AGAT)_n$ $(ATAG)_4(ATAC)_1(ATAG)_2$

GATA A7.1 forward and GATA A7.2 reverse primers. After a 95 °C pre-incubation step for 10 min, PCR amplification was performed in 30 cycles at 94 °C for 30 s, 56 °C for 30 s and 72 °C for 1 min, followed by a 20 min at 70 °C.

PCR amplified fragments were purified with Microspin S-300 HR columns (Pharmacia) and sequenced using ABI Big Dye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems). The products were ran on an ABI 377 sequencer (Applied Biosystems) and analysed using the Data Collection Software 377-18.

3. Results and discussion

Sequencing results for the 10 chimpanzee samples and comparisons with the previous results in humans are shown in Table 1.

In all the analysed loci, the size variation found within human or chimpanzee sequences was restricted to the main variable repeat unit or it was found in the repeats adjacent to the variable stretch.

In the flanking regions of all loci (further away than 7 bp from the variable stretch), 16 tetra-repeats were found and when compared in humans and chimpanzees, no size variation was found. The stability observed in these sequences implies that no variation is expected with the accumulation of new population data. Hence, these repetitive structures should not be considered in the STR allele nomenclature.

The nomenclature proposed is summarised in Table 2 and compared with previous ones.

This nomenclature is consistent with the existing data for humans and chimpanzees and with the ISFG guidelines and we strongly recommend its use to prevent further confusion arising in the field.

Acknowledgements

This work was supported by FCT "Programa de Financiamento Plurianual de Unidades de I&D. Programa Operacional Ciência, Tecnologia e Inovação (POCTI). Quadro Comunitário de Apoio III".

References

[1] Q. Ayub, A. Mohyuddin, R. Qamar, K. Mazhar, T. Zerjal, S.Q. Mehdi, C. Tyler-Smith, Nucleic Acids Res. 2 (2000) e8.
[2] P.S. White, O.L. Tatum, L.L. Deaven, J.L. Longmire, Genomics 57 (1999) 433–437.
[3] L. Gusmão, C. Alves, A. Amorim, Ann. Hum. Genet. 65 (2001) 285–291.
[4] A. González-Neira, M. Elmoznino, M.V. Lareu, P. Sánchez-Diz, M. Prinz, A. Carracedo, Forensic Sci. Int. (2001) in press.
[5] Y.P. Hou, J. Zhang, Y.B. Li, J. Wu, S.Z. Zhang, M. Prinz, Forensic Sci. Int. 118 (2001) 153–159.
[6] P. Grignani, G. Peloso, P. Fattorini, C. Previderè, Int. J. Leg. Med. 114 (2000) 125–129.