



Announcement of population data

Population genetics and mutational events at 6 Y-STRs in Korean population

Soon Hee Kim^{a,d}, Nam Ye Kim^{a,b}, Kyung Sook Kim^a, Jong Jin Kim^a, Jong Tae Park^b,
Ki Wha Chung^c, Myun Soo Han^a, Won Kim^{d,*}

^a DNA Analysis Division, National Institute of Scientific Investigation, Seoul 158-097, Republic of Korea

^b Department of Medicine, Jeonnam National University, Gwangju 501-746, Republic of Korea

^c Department of Biological Sciences, Kongju National University, Kongju 314-701, Republic of Korea

^d School of Biological Sciences, Seoul National University, San56-1, Sillim-dong, Gwanak-gu, Seoul 151-742, Republic of Korea

ARTICLE INFO

Article history:

Received 21 November 2007

Received in revised form 10 May 2008

Accepted 13 June 2008

Keywords:

Y-STR

Korean

Population data

Duplication

Microvariant

Mutation rate

ABSTRACT

Haplotype frequencies for 6 Y-STRs (DYS522, DYS533, DYS549, DYS570, DYS576 and DYS643) were determined in 539 unrelated Korean males. A total of 375 haplotypes were observed with the overall haplotype diversity of 0.9967. In DYS570, we found new point-3 microvariant allele series, i.e., 18.3, 19.3 and 20.3. DNA sequence analysis showed that the full repeat sequences were [TTTC]_nttc[TTTC]₅. Duplicated DYS643 allele was found and discussed on its forensic impact on the Y-STR interpretation of male–male mixtures. In 140 confirmed father/son pairs, one mutation was found in DYS576 with the overall mutation rate of 1.10×10^{-3} /locus/generation (95% CI 0.31– 1.89×10^{-3}).

© 2008 Elsevier Ireland Ltd. All rights reserved.

Population: Samples were collected from 539 unrelated Korean males and 140 father/son pairs (140 fathers, 152 corresponding sons) after obtaining informed consent to be involved in this study. Father's samples of the pairs were included in the 539 unrelated males. Samples were collected from all the major province of South Korea, i.e., Seoul–Gyeonggi, Chungcheong, Jeolla, Gyeongsang, Gangwon and Jeju. Father/son pairs were collected in Chungcheong and Jeolla province.

Extraction: DNA extraction was performed using a DNA IQ™ system (Promega).

PCR: The 6 Y-STRs of DYS522, DYS533, DYS549, DYS570, DYS576, and DYS643 were co-amplified using primer sequences described in a previously published paper [1]. The fluorescent dyes, Tamra (yellow), Fluorescein (blue), and JOE (green) were included on the primers. The PCR amplifications for Y-STR multiplex were performed in reaction volumes of 10 µl containing 1 × GeneAmp PCR Gold® buffer (Applied Biosystems), 1.5 mM MgCl₂, 200 µM dNTPs (Promega), 0.16 mg/ml BSA (Sigma), 1 unit of AmpliTaq Gold® DNA polymerase, the Y-STR primer mix, and 1 ng DNA template. The Y-STR primer mix contained 1 µM primer set of

DYS522 (fluorescent dye labeled with JOE), 1 µM DYS533 (Fluorescein), 0.75 µM DYS549 (Tamra), 0.3 µM DYS570 (JOE), 1 µM DYS576 (Fluorescein), and 0.2 µM DYS643 (Fluorescein). The cycle conditions were as follows: 95 °C for 10 min, followed by 28 cycles of 95 °C for 1 min, 55 °C for 1 min, 72 °C for 1 min, and a final extension at 60 °C for 1 h.

Genotyping and sequencing: The PCR products were run on the ABI3730 DNA Analyzer (Applied Biosystems). Automatic size calling was performed with GeneMapper® ID v3.1 and ILS600 (Promega) as the internal size standard. Two samples were directly sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) to verify the allele sequence. The Y-STR allele nomenclatures were designated by the number of repeats, according to the DNA commission of the International Society of Forensic Genetics (ISFG) [2].

Quality control: Proficiency testing of Collaborative Testing Service (CTS) was exercised. The same standard DNA 007 was included as internal control for each batch of genotyping.

Access to the data: Complete data can be obtained from supplementary data in electronic version of the journal.

Analysis of data: Allele frequencies were calculated from the observed haplotypes. The diversity of Y-STR haplotypes was calculated according to Nei [3]. Mutation rate was calculated by the number of mutation per locus per meiosis. Analysis of

* Corresponding author. Tel.: +82 2 880 6695; fax: +82 2 872 1993.

E-mail address: wonkim@plaza.snu.ac.kr (W. Kim).

Molecular Variance (AMOVA), R_{ST} values and the corresponding P values were estimated using the Arlequin v3.0 software [4].

Results: Allele frequency and haplotypes are shown in Supplementary Tables 1 and 2, respectively.

Other remarks: Y-STR profiles for standard DNA 007 resulted in 11–13–13–17–19–11 (DYS522–DYS533–DYS549–DYS570–DYS576–DYS643). Of the 539 Korean male individuals, a total of 375 haplotypes for the 6 Y-STRs were detected. Of these, 293 haplotypes were found in single individuals. The most frequent haplotype was found in 19 individuals. The haplotype diversity for the 6 Y-STRs was 0.9967 and a discrimination capacity was 0.6957.

Three novel microvariant alleles at DYS570 were observed in 10 individuals. Microvariant alleles 18.3, 19.3 and 20.3 at DYS570 were not found in other studies including this locus [1,5–7], whereas alleles 13.2, 14.2, 17.2 and 22.2 in DYS570 had been described in a population from Germany, The Netherlands and Turkey [5]. The variant alleles observed at DYS570 have been confirmed by sequencing. A 18.3 allele revealed the irregular repeat unit [TTTC]₁₃ttc[TTTC]₅. Seven 19.3 allele gave the same sequence [TTTC]₁₄ttc[TTTC]₅. And two 20.3 allele showed the same sequence [TTTC]₁₅ttc[TTTC]₅. First deletion process of a T (ΔT), and then divergence through the gain or loss of a repeat unit [TTTC] may give rise to [TTTC]_nttc[TTTC]₅ allele series over the time scale of Korean evolution.

The individual with duplicated allele (11–12) in DYS643 also have duplicated allele (22–23) at DYS390 using popular Y-filer™ kit (Applied Biosystems). According to BLAT (<http://genome.ucsc.edu/cgi-bin/hgBlat>), DYS390 is located within a 15.784 Mb region at Yq11.221 and DYS643 located within a 15.935 Mb at Yq11.221 on Y chromosome. The occurrence of the duplicated DYS643 and DYS390 alleles can be explained by simultaneous duplication, because duplicated loci are close to one another (<1 Mb). Duplicated alleles at two Y-STRs (DYS389I–DYS389II or DYS389II–DYS439) within the same individuals occur 7 times among 174 duplication cases in the YHRD database (<http://www.yhrd.org> release of November 2007). One haplotype from the YHRD database showed duplications at three Y-STRs (DYS389I–DYS389II–DYS439). Butler et al. observed duplication in up to four Y-STRs (DYS389I–DYS389II–DYS437–DYS439) from a single individual [8]. So far, those duplicated Y-STRs reported are located within the *azoospermia factor a (AZFa)* region at Yq11.21, therefore, it was suggested that the *AZFa* region that includes these Y-STRs might have been duplicated together. The present case is in agree with the current criteria of the forensic interpretation of male–male mixtures, because the question of whether locus duplication or two male mixtures can be determined by the context of Y-STRs location along the Y-chromosome.

We performed the direct analysis of confirmed 140 father/son pairs with 292 samples (probability > 99.99%) at 6 Y-STRs and one case of mutation occurred at DYS576 with allele 19 mutating to 20. In our study, the locus-specific mutation rate estimates varied between 0 and 6.58×10^{-3} /locus/generation based on 152 meiosis events per locus. The overall mutation rate in our sample (1.10×10^{-3} /locus/generation; 95% confidence interval 0.31– 1.89×10^{-3}) was not significantly different from that previously

reported: 1.57×10^{-3} /locus/generation [9], and 2.80×10^{-3} /locus/generation [10].

The present haplotype data were compared with those available for the same set of 6 Y-STRs from Germany, The Netherlands and Turkish population [5]. Population pairwise comparison showed that the Korean population has significant differences from Berlin ($R_{ST} = 0.00286$, $P = 0.00901$), Dresden ($R_{ST} = 0.00348$, $P = 0.00000$), Hamberg ($R_{ST} = 0.00745$, $P = 0.00000$), Munich ($R_{ST} = 0.00574$, $P = 0.00000$), Rostock ($R_{ST} = 0.00471$, $P = 0.00000$), Sorbs ($R_{ST} = 0.01340$, $P = 0.00000$) in Germany, Dutch population from Leiden ($R_{ST} = 0.00295$, $P = 0.00000$), and Turkish population living in Berlin ($R_{ST} = 0.00341$, $P = 0.00901$).

This paper follows the guidelines for publication of population data requested by the journal [2,11].

Acknowledgments

This work was supported by a grant from the Korean Ministry of Science and Technology (M10640020002-06N4002-00210).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fsigen.2008.06.004.

References

- [1] S.K. Lim, Y. Xue, E.J. Parkin, C. Tyler-Smith, Variation of 52 new Y-STR loci in the Y Chromosome Consortium worldwide panel of 76 diverse individuals, *Int. J. Legal Med.* 121 (2007) 124–127.
- [2] L. Gusmão, J.M. Butler, A. Carracedo, P. Gill, M. Kayser, W.R. Mayr, N. Morling, M. Prinz, L. Roewer, C. Tyler-Smith, P.M. Schneider, DNA commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis, *For. Sci. Int.* 157 (2006) 187–197.
- [3] M. Nei, *Molecular Evolutionary Genetics*, Columbia University Press, New York, 1987, pp. 176–179.
- [4] L. Excoffier, G. Laval, S. Schneider, Arlequin ver. 3.0: an integrated software package for population genetics data analysis, *Evol. Bioinform. Online* 1 (2005) 47–50.
- [5] H. Rodig, L. Roewer, A. Gross, T. Richter, P. de Knijff, M. Kayser, W. Brabetz, Evaluation of haplotype discrimination capacity of 35 Y-chromosomal short tandem repeat loci, *For. Sci. Int.* 174 (2008) 182–188.
- [6] H. Asamura, H. Sakai, M. Ota, H. Fukushima, MiniY-STR quadruplex systems with short amplicon lengths for analysis of degraded DNA samples, *For. Sci. Int.: Genet.* 1 (2007) 56–61.
- [7] J.M. Butler, A.E. Decker, P.M. Vallone, M.C. Kline, Allele frequencies for 27 Y-STR loci with U.S. Caucasian, African American, and Hispanic samples, *For. Sci. Int.* 156 (2006) 250–260.
- [8] J.M. Butler, A.E. Decker, M.C. Kline, P.M. Vallone, Chromosomal duplications along the Y-chromosome and their potential impact on Y-STR interpretation, *J. For. Sci.* 50 (2005) 853–859.
- [9] B. Budowle, M. Adamowicz, X.G. Aranda, C. Barna, R. Chakraborty, D. Cheswick, B. Dafeo, A. Eisenberg, R. Frappier, A.M. Gross, C. Ladd, H.-S. Lee, S.C. Milne, C. Meyers, M. Prinz, M.L. Richard, G. Saldanha, A.A. Tierney, L. Viculis, B.E. Krenke, Twelve short tandem repeat loci Y chromosome haplotypes: genetic analysis on populations residing in North America, *For. Sci. Int.* 150 (2005) 1–15.
- [10] M. Kayser, L. Roewer, M. Hedman, L. Henke, J. Henke, S. Brauer, C. Krüger, M. Krawczak, M. Nagy, T. Dobosz, R. Szibor, P. de Knijff, M. Stoneking, A. Sajantila, Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs, *Am. J. Hum. Genet.* 66 (2000) 1580–1588.
- [11] P. Lincoln, A. Carracedo, Publication of population data of human polymorphisms, *For. Sci. Int.* 110 (2000) 3–5.