

Genetic data and *de novo* mutation rates in father-son pairs of 23 Y-STR loci in Southern Brazil population

Nicole Nascimento Da Fré · Rodrigo Rodenbusch ·
André Zoratto Gastaldo · Erin Hanson ·
Jack Ballantyne · Clarice Sampaio Alho

Received: 2 June 2014 / Accepted: 29 October 2014 / Published online: 13 November 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract We evaluated haplotype and allele frequencies, as well as statistical forensic parameters, for 23 Y-chromosome short tandem repeats (STRs) loci of the PowerPlex®Y23 system (DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635, Y-GATA-H4, DYS481, DYS533, DYS549, DYS570, DYS576, DYS643) in a sample of 150 apparently healthy males, resident in South Brazil. A total of 150 different haplotypes were identified. The highest gene diversity (GD) was observed for the single locus marker DYS570 (GD=0.7888) and for a two-locus system DYS385 (GD=0.9009). We also examined 150 father-son pairs by the same system, and a total of 13 mutations were

identified in the 3450 father-son allelic transfers, with an overall mutation rate across the 23 loci of 3.768×10^{-3} (95 % CI: 3.542×10^{-3} to 3.944×10^{-3}). In all cases there was only one locus mutated with gain/loss of repeats in the son (5 one-repeat gains, and 7 one-repeat and 1 two-repeat losses); we observed no instances of mutations involving a non-integral number of repeats.

Keywords Y-STR · Genetic marker · DNA · Mutation rate · Paternity

Short tandem repeat (STR) population data

Samples stored on FTA paper were obtained from 150 father-son pair donors from paternity cases enrolled in Rio Grande do Sul, Brazil, a region that was essentially colonized by Portuguese, German, and Italian immigrants in the 18 century [1]. The biological relationship of all father-son pairs was previously confirmed by autosomal STR analysis (AmpF/STR® Identifiler® PCR Amplification Kit; Applied Biosystems; Life Technologies, USA), with paternity index values >10,000. The age of the father at the time of son's birth was noted. All of the participants signed informed consent forms, and this study was carried out in accordance to approved ethical standards (CAAE 11096212.6.0000.5336; Protocol # 180.121). We used the nomenclature recommended by the DNA Commission of the International Society of Forensic Genetics [2], as well as the recommendations for publication of population data for forensic purposes [3].

Genomic DNA samples were purified from dried blood samples preserved on FTA cards (Whatman Bioscience, Cambridge, UK) and amplified using the PowerPlex® Y23 System kit (Promega Corporation, Madison, WI) which include 23 Y-STR loci: DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439,

Electronic supplementary material The online version of this article (doi:10.1007/s00414-014-1101-2) contains supplementary material, which is available to authorized users.

N. N. Da Fré · R. Rodenbusch · A. Z. Gastaldo · C. S. Alho (✉)
Laboratório de Genética Humana e Molecular, Pontifícia
Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6681,
Prédio 12C, Porto Alegre, RS 90619-900, Brazil
e-mail: csalho@puers.br

N. N. Da Fré · R. Rodenbusch · A. Z. Gastaldo
Laboratório de Investigação de Paternidade, Fundação Estadual de
Produção em Pesquisa do Rio Grande do Sul, Av. Ipiranga, 5400,
Porto Alegre, RS 90610-000, Brazil

E. Hanson · J. Ballantyne
Graduate Program in Biomolecular Science, University of Central
Florida, P.O. Box 162366, Orlando, FL 32816-2366, USA

E. Hanson · J. Ballantyne
Department of Chemistry, University of Central Florida, Building 5,
4000 Central Boulevard, P.O. Box 162366, Orlando,
FL 32816-2366, USA

E. Hanson · J. Ballantyne
National Center for Forensic Science, P.O. Box 162367, Orlando,
FL 32816-2367, USA

DYS448, DYS456, DYS458, DYS635, Y-GATA-H4, DYS481, DYS533, DYS549, DYS570, DYS576, and DYS643 [4, 5]. PCR products were separated and genotyped by capillary electrophoresis on an ABI PRISM® 3130xl Genetic Analyzer (Life Technologies) and analyzed by GeneMapper® ID software v3.2 (Life Technologies, USA). All mutational event cases were confirmed by re-analysis. The mutations were confirmed by a second typing by AmpF/STR® Yfiler® PCR Amplification Kit (Applied Biosystems; Life Technologies, USA) whenever the locus involved was present in this different system. A quality control check was performed using the proficiency testing of the Y-STR Haplotyping Quality Assurance Exercise. The 23 Y-STRs haplotype data for the sample of father-son pairs from Rio Grande do Sul (RS), Brazil, will be available in the YHRD database (<http://www.yhrd.org>). Allele and haplotype frequencies were determined using the software Arlequin, version v.3.5.1.3 [6]. Haplotype diversity (HD), power of matching (PM), power of discrimination (PD), and discrimination capacity (DC) were calculated as described elsewhere [7]. Y-STR haplotype frequencies and diversity were calculated treating loci DYS385a and DYS385b as a single haplotype, and locus DYS389I after subtracting the value of DYS389I. The age of the fathers (at the time of son's birth) with and without mutations was compared by a Mann–Whitney *U* test.

A total of 150 father-son pairs with full profiles were analyzed in this study. Allele/genotype frequencies and gene diversity values calculated for each locus in the sons are listed in Supplemental Table 1 (the six new loci of the PowerPlex Y-23 kit are underlined and in gray box). The highest gene diversity was observed for the single locus marker DYS570 (GD=0.7888) and for the two-locus system DYS385 (GD=0.9009). We identified a total of 150 different haplotypes in a sample. The genotyping with the 23 Y-STR loci using PowerPlex® Y23 System kit resulted in a greater discrimination capacity when compared to the previously published studies in the same RS population with 11 Y-STR loci (DC=84.0 %) [8], and with 17 Y-STR loci (DC=96.9 %) [9]. Thirteen mutations were identified in the 3450 father-son allelic transfers, with an overall mutation rate across the 23 loci of 3.768×10^{-3} (95 % CI, 3.542×10^{-3} to 3.944×10^{-3}). This rate varied from other Y-STRs studies in other populations as is showed in Supplemental Table 2, but this difference was not statically significant considering an overall proportion of 2.432×10^{-3} ($P=0.116$, Binomial-Bilateral Test; see Supplemental Table 2). Based in the data from Supplemental Table 2, we compare the proportion of mutations per meiosis (father to son transmission of the whole Y chromosome) calculated by the number of mutations within the whole Y chromosome divided by the number of the father-son pairs analyzed; as expected, a larger number of mutations is encountered when the Y-STR system have more loci (Supplemental Fig. 1).

In the 23 Y-STR panel, we observed a not so different number of gains ($N=5$) and losses ($N=8$) of repeats in the son; this proportions were similar to global proportions of 55.2 % of gains ($P=0.225$) and 44.8 % of losses ($P=0.350$, Binomial-Bilateral Test; see Supplemental Table 2). Some variations among studies might be the result of differences in sample size and/or in the number of analyzed loci. In all cases, there was only one single locus mutated per pair with gain/loss of repeats in the son; five were gain of one-repeat (single-step), seven were loss of one-repeat, and in one case there was a two-repeat loss (locus DYS481) (Supplemental Table 3). Ninety-two percent (12/13) were mutations that resulted in one-repeat difference, which is consistent with the concept that due to strand slippage during replication, the majority of mutations comprise single-step repeat gain or loss [10].

In order to confirm the paternity of these 13 alleged father-son pairs, the AmpF/STR® Identifiler® PCR Amplification Kit (Applied Biosystems; Life Technologies, USA) results were reviewed. None of these father-son pairs was found to have observable *de novo* mutations with the 15 autosomal STRs. The combined paternity index (CPI) values for each sample pair using the 15 autosomal AmpF/STR® Identifiler® PCR Amplification Kit STR loci genotyping results considering father-son pair, as well as the CPI with mothers' profile included were shown in Supplemental Table 3. In these cases, the CPI for father-son pair were above 4.2×10^2 ; and when the mother's profile is included for each pair, the CPI was above 8.3×10^6 . Supplemental Table 3 shows also the age of the fathers involved in the observed mutation events. No significant differences between mutation event and the father's age at the birth of child was observed ($P=0.7970$; Mann–Whitney *U* Test) since the average age of the fathers involved in mutation events was 27.8 years (SD=8.21); and for no-mutation transmissions, the average was 29.0 years (SD=9.07). A similar result was noticed in a study with 1766 father-son pairs (15,894 meioses) in nine Y-STR loci [11]. However, it is in contrast with a collaborative work with 3026 father-son pairs in 17 Y-STR loci [12] where the authors found that mutation rate increased with the age ($P<0.001$; Mann–Whitney *U* Test) and with subsequent work that analyzed 1730 father-son pairs by AmpF/STR® 17Yfiler® PCR amplification kit [13]. Variations among these and other studies are in Supplemental Table 4 and might be explained by the differences in the number of Y-STR analyzed loci or by the still small group of fathers who transmit mutations (about 0.2 % of all evaluated men).

The observed mutation rates at 23 Y-STR loci were 6.7×10^{-2} (in DYS458, DYS358a/b, and DYS456), 1.3×10^{-3} (in DYS481, DYS570, and DYS390), and 2×10^{-3} in DYS576. In 16 Y-STR loci there were no mutations. Since mutation rate is believed to increase with the length of allele repeat size [10], we recorded the allele category size (the alleles were

categorized into 25, 50, and 25 % quantiles for short, moderate, and long sizes, respectively); all mutations occurred in the progenitor alleles (father's alleles) with moderate size. We observed no instances of mutations involving a non-integral number of repeats, like du- or triplication, deletion, microvariant, etc. in our sample. Knowledge of the mutation rates of the loci described above will improve the exclusion criteria in paternity testing and the interpretation of the DNA profiles in the population of Southern Brazil. These data show that the present set of 23 Y-STR markers studied with PowerPlex® Y23 System kit is highly polymorphic and discriminative in the Brazilian southernmost state population.

References

- Instituto Brasileiro de Geografia e Estatística-IBGE (2000) Brasil 500 anos de povoamento. IBGE, Rio de Janeiro
- Gusmão L, Butler JM, Carracedo A, Gill P, Kayser M, Mayr WR, Morling N, Prinz M, Roewer L, Tyler-Smith C, Schneider PM, International Society of Forensic Genetics (2006) DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. *Int J Legal Med* 120(4):191–200
- Carracedo A, Butler JM, Gusmão L, Parson W, Roewer L, Schneider PM (2010) Publication of population data for forensic purposes. *Forensic Sci Int Genet* 4(3):145–147
- Davis C, Ge J, Sprecher C, Chidambaram A, Thompson J, Ewing M, Fulmer P, Rabbach D, Storts D, Budowle B (2013) Prototype PowerPlex® Y23 system: a concordance study. *Forensic Sci Int Genet* 7(1):204–208
- Thompson JM, Ewing MM, Frank WE, Pogemiller JJ, Nolde CA, Koehler DJ, Shaffer AM, Rabbach DR, Fulmer PM, Sprecher CJ, Storts DR (2013) Developmental validation of the PowerPlex® Y23 system: a single multiplex Y-STR analysis system for casework and database samples. *Forensic Sci Int Genet* 7:240–250
- Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10(3):564–567
- Willuweit S, Roewer L (2007) International forensic Y chromosome user group, Y chromosome haplotype reference database (YHRD): update. *Forensic Sci Int Genet* 1(2):83–87
- Leite FP, Callegari-Jacques SM, Carvalho BA, Kommers T, Matte CH, Raimann PE, Schwengber SP, Sortica VA, Tsuneto LT, Petzl-Erler ML, Salzano FM, Hutz MH (2008) Y-STR analysis in Brazilian and South Amerindian populations. *Am J Hum Biol* 20: 359–363
- Schwengber SP, Kommers T, Matte CH, Raimann PE, Carvalho BA, Leite FP, Medeiros MA, Souza LF, Castro CS, Chassot FG, Bonatto SL (2009) Population data of 17 Y-STR loci from Rio Grande do Sul state (South Brazil). *Forensic Sci Int Genet* 4:e31–e33
- Ge J, Budowle B, Aranda XG, Planz JV, Eisenberg AJ, Chakraborty R (2009) Mutation rates at Y chromosome short tandem repeats in Texas populations. *Forensic Sci Int Genet* 3(3):179–184
- Dupuy BM, Stenersen M, Egeland T, Olaisen B (2004) Y-chromosomal microsatellite mutation rates: differences in mutation rate between and within loci. *Hum Mutat* 23(2):117–124
- Gusmão L, Sánchez-Diz P, Calafell F, Martín P, Alonso CA, Alvarez-Fernández F, Alves C, Borjas-Fajardo L, Bozzo WR, Bravo ML, Builes JJ, Capilla J, Carvalho M, Castillo C, Catanesi CI, Corach D, Di Leonardo AM, Espinheira R, Fagundes de Carvalho E, Farfán MJ, Figueiredo HP, Gomes I, Lojo MM, Marino M, Pinheiro MF, Pontes ML, Prieto V, Ramos-Luis E, Riancho JA, Souza Góes AC, Santapa OA, Sumita DR, Vallejo G, Vidal Rioja L, Vide MC, Vieira da Silva CI, Whittle MR, Zabala W, Zarrabeitia MT, Alonso A, Carracedo A, Amorim A (2005) Mutation rates at Y chromosome specific microsatellites. *Hum Mutat* 26(6):520–528
- Goedbloed M, Vermeulen M, Fang RN, Lembring M, Wollstein A, Ballantyne K, Lao O, Brauer S, Krüger C, Roewer L, Lessig R, Ploski R, Dobosz T, Henke L, Henke J, Furtado MR, Kayser M (2009) Comprehensive mutation analysis of 17 Y-chromosomal short tandem repeat polymorphisms included in the AmpF/STR Yfiler PCR amplification kit. *Int J Legal Med* 123(6):471–482