Int J Legal Med (2018) 132:149–151
DOI 10.1007/s00414-017-1642-2

POPULATION DATA

Evaluation of InnoTyper® 21 in a sample of Rio de Janeiro population as an alternative forensic panel

R. S. Moura-Neto[1] · I. C. T. Mello[2] · R. Silva[2] · A. P. C. Maette[2] · C. G. Bottino[3] · A. Woerner[4] · J. King[4] · F. Wendt[4] · B. Budowle[4,5]

Received: 24 April 2017 / Accepted: 4 July 2017 / Published online: 26 July 2017
© Springer-Verlag GmbH Germany 2017

Abstract The use of bi-allelic markers such as retrotransposable element insertion polymorphisms or Innuls (for insertion/null) can overcome some limitations of short tandem repeat (STR) loci in typing forensic biological evidence. This study investigated the efficiency of the InnoTyper® 21 Innul markers in an urban admixed population sample in Rio de Janeiro ($n = 40$) and one highly compromised sample collected as evidence by the Rio de Janeiro police. No significant departures from Hardy-Weinberg equilibrium were detected after the Bonferroni correction ($\alpha' \approx 0.05/20$, $p < 0.0025$), and no significant linkage disequilibrium was observed between markers. Assuming loci independence, the cumulative random match probability (RMP) was 2.3×10^{-8}. A lower mean Fis value was obtained for this sample population compared with those of three North American populations (African-American, Southwest Hispanic, US Caucasian). Principal component analysis with the three North American populations and one from 21 East Asian population showed that African Americans segregated as an independent group while US Caucasian, Southwest Hispanic, East Asian, and Rio de Janeiro populations are in a single large heterogeneous group. Also, a full Innuls profile was produced from an evidence sample, despite the DNA being highly degraded. In conclusion, this system is a useful complement to standard STR kits.

Keywords Forensic genetics · Transposable element markers · Innuls · Rio de Janeiro · Brazil

Short tandem repeat (STR) loci are the primary markers used to characterize forensic biological evidence and are core markers for all national DNA databases. Despite their appealing features, the amplicon size of many STR loci is greater than 200 bp, and these large loci are not ideal for analysis of highly degraded and low template DNA samples [1]. Also, STR loci have relatively high mutation rates that may complicate kinship analyses. An alternative to overcome the limitations of STRs is the use of bi-allelic markers, such as single nucleotide polymorphisms (SNPs) and insertion/deletion (InDels) polymorphisms. These markers can be detected in shorter amplicons and have low mutation rates. Indeed, Kayser and de Knijff [2] point out that SNPs and InDels are alternative markers that perform substantially better than STRs for the analysis of degraded DNA samples. A type of marker similar in nature to InDels is retrotransposable element (RE) insertion polymorphisms or Innuls (for insertion/null) [3]. The Innuls polymorphism consists of long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs or Alu insertion) [3]. Innuls are highly abundant in the human genome and extremely stable once inserted [4–9].

Electronic supplementary material The online version of this article (doi:10.1007/s00414-017-1642-2) contains supplementary material, which is available to authorized users.

✉ R. S. Moura-Neto
rodrigomouraneto@ufrj.br

1 Laboratório de Biologia Molecular Forense, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

2 Instituto de Biofisica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

3 Policia Civil do Estado do Rio de Janeiro, Instituto de Pesquisas e Perícias em Genética Forense, Rio de Janeiro, RJ, Brazil

4 Center for Human Identification, University of North Texas Health Science Center, Fort Worth, TX, USA

5 Center of Excellence in Genomic Medicine Research (CEGMR), King Abdulaziz University, Jeddah, Saudi Arabia

Previous attempts to use Alu insertions for identity testing exploited the size difference between insertion and null alleles by amplifying the entire region with the same forward and reverse primers positioned in the flanking region [10]. The insertion allele would be 200–400 bp larger than the null allele and could be detected by electrophoresis based on size differences. This approach can be useful for paternity testing and some population studies, where DNA is not limited or compromised. However, the large size difference between amplicons of the no-insertion (null) and insertion alleles will impact amplification efficiency and increase allele dropout during the PCR, which is a limitation for forensic samples (i.e., preferential amplification favoring the smaller allele amplicon and degradation resulting possibly in dropout of the insertion allele). Thus, the use of REs as a multiplexed marker system had not been embraced for the analysis of forensic samples [11]. The InnoTyper® 21 system was designed to detect the presence or absence of specific Alu insertions by positioning primers at the junction of the insertion site to overcome the allele size differences. This strategy generates small amplicons of almost the same size and can type highly degraded DNA samples [10–12].

Validation and population studies have been conducted in only four distinct population groups [13]. However, the performance of this panel of Innuls needs to be evaluated in additional populations for better use of these markers for human identification.

In this study, the efficiency of the InnoTyper® 21 markers was tested in an urban admixed population sample in Rio de Janeiro and was compared with that of North America populations and one from East Asia to enhance information to support the worldwide application of this system. In addition, the ability to obtain a full Innuls profile was compared with that of STR typing of a highly compromised biological evidence.

An admixed sample from Rio de Janeiro ($n = 40$) was typed with a panel of 20 Innul markers (Supplementary Material and Methods), plus the amelogenin locus (InnoTyper® 21). A sample electropherogram of this multiplex is shown in Supplementary Figure S1. The peak high ratio (PHR), one of the parameters needed for mixture analyses, was taken from average heterozygous ratios of each marker across all samples. Values of PHR varied from 0.723 to 0.869, with a mean value of 0.798 ± 0.147, which is comparable to STR systems. To estimate DNA frequencies with 95% of confidence, we calculated the minimum allele frequency of 0.045 for a bi-allelic system according to Budowle et al. [14]. Also, according to Chakraborty [15], a population sample size of 36 would be more than adequate to detect alleles and estimate allele frequencies for a bi-allelic marker (at $p = 0.05$ and $\alpha = 0.05$). All alleles, on all loci, showed frequencies well above that minimum value, even considering the standard error (Supplementary Table S1). No departures from HWE were detected for all Innul markers except for SB19.12 (Supplementary Table S1). This finding is consistent with departures expected by chance, after the Bonferroni correction ($\alpha' \approx 0.05/20$, $p < 0.0025$).

The significant linkage disequilibrium (LD) between marker pairs, measured as R^2, was assessed using Fisher's exact test with 10,000 permutations [16]. For the InnoTyper® 21 markers, there were 190 possible pairwise comparisons. A total of six pairs had a detectable LD at the 0.025 ($\alpha'' \approx 0.05/2$, $p < 0.025$) level in the Rio de Janeiro sample population (Supplementary Table S2). This proportion of detectable LD was less than expected by chance (~9 to 10 pairs). The data support the use of the product rule to calculate the combined genotype frequencies.

The power of an Innuls panel is related to the number of markers with a random match probability (RMP) near or below 0.4 (considering the ideal value of $p = q = 0.5$, RMP is 0.375). For all markers, the RMP varied between 0.36 and 0.52 in the population tested (Supplementary Table S1). Nine Innuls are above the threshold (i.e., RMP of 0.4.) (Supplementary Figure S2). The cumulative RMP was 2.3×10^{-8} for the Rio de Janeiro population, assuming loci independence and no substructure effect. This value is similar to that reported for the North American population groups [13].

The Rio de Janeiro population was compared with three North American populations: African-American ($n = 207$); Southwest Hispanic ($n = 40$); and US Caucasian ($n = 205$) and one from East Asia ($n = 44$) [13]. Wright's F_{st} was estimated to assess population substructure [17]. The overall F_{st} was 0.0516. For the five populations, the F_{st} pairwise values are provided for the populations in Supplementary Table S3 and this parameter was used to generate principal component analysis (PCA) which showed a comparable genetic distance among all five population samples (Supplementary Figure S3). The African-Americans segregated as an independent group. The US Caucasian, Southwest Hispanic, East Asian, and the Rio de Janeiro populations are in a single large heterogeneous group, indicating that the markers constitute a suitable system for human identification (HID) to be used with distinct ethnic groups and, also, with admixed populations.

To evaluate population differentiation due to substructure, Wright's F_{is} was estimated. The F_{is} value was −0.010622 (Supplementary Table S1) which is lower than that reported for the North American sample populations [13]. Also, individual ancestry proportions were estimated with Structure v2.3.4. The composition of those populations was assumed to be at least two ($K = 2$), since the estimated natural logarithm probability LnP was found to be higher (Supplementary Figure S4). The African-American (AFA) group is different from all others, from $K = 2$ to 7, and the USC, HIS, ASI, and RIO showed the same pattern of substructure (Supplementary Figure S5). This result suggests that those markers can reveal different population compositions between African and non-African.

Using the Quantifiler Trio, the DNA degradation index was equal to 204, and the sample was analyzed with InnoTyper® 21 and PowerPlex® Fusion. In addition, the alleged mother was analyzed as a reference sample for body identification purposes. Supplementary Figure S6 shows the STR profile with only eight markers, all below 180 bp, being fully or partially successfully typed, using a threshold of 150 rfu. This highly degraded DNA sample produced a full Innul profile (Supplementary Figure S7), although there was some allelic imbalance and noise, validation studies performed previously had optimized the PCR to reduce allelic dropout [13]. The likelihood ratio (LR) of the STR system results was 73.7, with a posterior probability (PP) of 98.7%. The Innuls system provided a LR = 49.4 (PP = 98%). Neither system provided sufficient discrimination to obtain an acceptable conclusion, but the combined LR (assuming loci independence) was 3644, with PP = 99.97%, which strongly supports the genetic relationship. This outcome indicates that Innuls are a useful set of markers and can complement conventional STRs, especially for challenged biological evidence such as those involving the identification of human remains, hair shafts, paraffin-embedded tissues, and other sample types [18–20].

The InnoTyper® 21 system is a highly discriminating nuclear DNA detection system for human identification. Since the InnoTyper® 21 system was designed to generate small product fragments (between 63 and 123 bp), degraded samples may yield more typing results with this system than with STR loci. Brown et al. [13] observed that InnoTyper® 21 recovered more alleles than MiniFiler on compromised samples. Thus, the InnoTyper® 21 system is well-suited for kinship analysis of degraded human remains and genetic studies in admixed populations.

Acknowledgments This work was supported in part by research grants from FAPERJ, CNPq, CAPES, and INMETRO. Also, we would like to acknowledge the support of Innovagen and InnoGenomics.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Butler JM, Shen Y, McCord BR (2003) The development of reduced size STR amplicons as tools for analysis of degraded DNA. J Forensic Sci 48:1054–1064 Paper ID JFS2003043_485
2. Kayser M, de Knijff P (2011) Improving human forensics through advances in genetics, genomics and molecular biology. Nat Rev Genet 12:179–192. doi:10.1038/nrg2952
3. Batzer MA, Deininger PL (2002) Alu repeats and human genomic diversity. Nat Rev Genet 3:370–379. doi:10.1038/nrg798
4. Gill P, Werrett DJ, Budowle B, Guerrieri R (2004) An assessment of whether SNPs will replace STRs in national DNA databases—joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDAM). Sci Justice 44:51–53. doi:10.1016/S1355-0306(04)71685-8
5. LaRue BL, Sinha SK, Montgomery AH, Thompson R, Klaskala L, Ge J, King J, Turnbough M, Budowle B (2012) INNULs: a novel design amplification strategy for retrotransposable elements for studying population variation. Hum Hered 74:27–35. doi:10.1159/000343050
6. Sudhir K, Carter A (2010) Method for genetic detection using interspersed genetic elements. US patent # 7,794,983
7. Feng Q, Moran JV, Kazazian HH Jr, Boeke JD (1996) Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87:905–916
8. Houck CM, Rinehart FP, Schmid CW (1979) A ubiquitous family of repeated DNA sequences in the human genome. J Mol Biol 132:289–306
9. Kazazian HH, Moran JV (1998) The impact of L1 retrotransposons on the human genome. Nat Genet 19:19–24. doi:10.1038/ng0598-19
10. Mamedov IZ, Shagina IA, Kurnikova MA, Novozhilov SN, Shagin DA, Lebedev YB (2010) A new set of markers for human identification based on 32 polymorphic Alu insertions. Eur J Hun Genet 18:808–814. doi:10.1038/ejhg.2010.22
11. Hughes-Stamm R, Ashton KJ, van Daal A (2011) Assessment of DNA degradation and the genotyping success of highly degraded samples. Int J Legal Med 125:341–348. doi:10.1007/s00414-010-0455-3
12. Batzer MA, Stoneking M, Alegria-Hartman M, Bazan H, Kass DH, Shaikh TH, Novick GE, Ioannou PA, Scheer WD, Herrera RJ (1994) African origin of human-specific polymorphic Alu insertions. Proc Natl Acad Sci U S A 91:12288–12292
13. Brown H, Thompson R, Murphy G, Peters D, LaRue B, Kiing J, Montgomery AH, Carroll M, Baus J, Sinha S, Wendt FR, Song B, Chakraborty R, Budowle B, Sinha SK (2017) Development and validation of a novel multiplexed DNA analysis system, InnoTyper®21. Forens Sci Int Genet 29:80–99. doi:10.1016/j.fsigen.2017.03.017
14. Budowle B, Monson KL, Chakrabort R (1996) Estimating minimum allele frequencies for DNA profile frequency estimates for PCR-based loci. Int J Legal Med 108:173–176
15. Chakraborty R (1992) Sample size requirements for addressing the population genetic issues of forensic use of DNA typing. Hum Biol 64:141–159
16. Oostdik K, Lenz K, Nye J, Schelling K, Yet D, Bruski S, Strong J, Buchanan C, Sutton J, Linner J, Frazier N, Young H, Matthies L, Sage A, Hahn J, Wells R, Williams N, Price M, Koehler J, Staples M, Swango KL, Hill C, Oyerly K, Duke W, Katzilierakis L, Ensenberger MG, Bourdeau JM, Sprecher CJ, Krenke B, Storts DR (2014) Developmental validation of the PowerPlex® fusion system for analysis of casework and reference samples: a 24-locus multiplex for new database standards. Forensic Sci Int Genet 12:69–76. doi:10.1016/j.fsigen.2014.04.013
17. Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370. doi:10.1086/667904
18. Campbell CD, Eichler EE (2013) Properties and rates of germline mutations in humans. Trends Genet 29:575–584. doi:10.1016/j.tig.2013.04.005
19. van den Berge M, Wiskerke D, Gerretsen RRR, Tabak J, Sijen T (2016) DNA and RNA profiling of excavated human remains with varying postmortem intervals. Int J Legal Med 130:1471–1480. doi:10.1007/s00414-016-1438-9
20. Sinha S, Murphy G, Brown H, Montgomery A, Carrol M, Tabak J (2015) Retrotransposable elements: novel and sensitive DNA markers and their application in human identity. Forensic Sci Int Genet Supp Ser 5:e627–e629. doi:10.1016/j.fsigss.2015.10.005