

Duarte Nuno Vieira • Anthony Busuttil
Denis Cusack • Philip Beth
Editors



Acta Medicinae
Legalis et Socialis

L. Cainé^{1,2,3}, F. Pinheiro^{3,4,5}

¹ *North Branch of the National Institute of Legal Medicine, Oporto, Portugal*

² *Faculty of Medicine, University of Oporto, Portugal*

³ *Centre for Forensic Sciences, Coimbra, Portugal*

⁴ *Biomedical Sciences Institute Abel Salazar, Oporto, Portugal*

⁵ *Fernando Pessoa University, Oporto, Portugal*

THE IMPORTANCE OF STR SPECIFIC DATABASES: A COMPARATIVE STUDY

Abstract: The STR allele frequencies have been used in studies of human populations to assess their genetic composition, variability, relationships and the evolutionary factors to which they are subjected. The Brazilian population is composed of individuals from different ethnical origins, varying according to the geographical region. Santa Catarina population is mainly originated from Portuguese (Azores archipelago). In this study, a Neighbor-Joining tree was constructed based on genetic distances of Nei, using the allele frequencies of the CODIS STR loci, from Santa Catarina population and other 23 different populations, including some of other ethnic groups. The results indicate the requirement of an extensive genetic study in order to implement population specific databases for forensic purposes.

Introduction

Short Tandem Repeat (STR) loci are widely used in individual identification, parentage testing and forensic casework, because of their higher discriminating information, sensitivity and reproducibility. Various multiplex PCR systems have been developed which allow the simultaneous amplification of several STR loci, they are rapid and powerful tools in forensic genetics. In this study we have analyzed the specific core of 13 STRs used to generate a nationwide USA DNA database, called the FBI (Federal Bureau of Investigation) Combined DNA Index System (CODIS).

The Brazilian population is composed of individuals from different ethnical origins, varying according to the geographical region. The Santa Catarina colonization began in the 16th century with the arrival of the Portuguese. The Azorean immigrants came later, in the 18th century, sent by the Portuguese king to complete the process. Despite the fact that a considerable amount of literature on the allele frequencies of these STRs is available, no data have been published about comparisons between different Brazilian states. Appropriate databases are imperative, because the variation in STR allele proportions between populations has been previously reported, reflecting different ancestral gene pools.

A comparative study, for the same markers, between Santa Catarina population and others was performed, being the main goal to demonstrate the need of the establishment of specific population databases for application in forensic and paternity investigations. A larger comparative study was not performed, because of the lack of additional Brazilian published population data.

Materials and Methods

Blood stains from 185 unrelated, autochthonous healthy donors from Santa Catarina were collected, with a major ethnic background considered to be Caucasian, 160 of them were used to perform the study previously published (Caine et al., 2003). All the methodology was made according to the afore mentioned study. Additionally, we used the PHYLIP version 3.6c software package (Felsenstein, 1986), to obtain a Neighbor-Joining tree (Nei and Saitou, 1987) from the genetic distances of Nei (Nei, 1972). The statistical robustness of their nodes was tested through a bootstrap approach (Efron, 1982; Felsenstein, 1985). Loci were resampled with replacement of 1000 iterations. The tree was visualized on TREE VIEW version 1.6.6 program (Page, 1996). For comparisons the following populations from the literature were used: Australia (Bagdonavicius et al., 2002), Azores (Velosa et al., 2002), Rio de Janeiro, Brazil (Góes et al., 2004), Mato Grosso, Brazil (Silva et al., 2004), Greece (Skitsa et al., 2003), Hong Kong (Law et al., 2002), Italy (Garofano et al., 1998), India (Sahoo and Kashyap, 2002), Japan (Hashiyada et al., 2003), Jamaica and Trinidad & Tobago (Budowle et al., 2001), Madeira (Fernandes et al., 2002), Morocco (Jauffrit et al., 2003), North Africa (Farfán et al., 2001), Portugal (Abrantes et al., 2004), Paraguay (Espin et al., 2002), Peru (Pérez et al., 2003), Poland (Pepinski et al., 2001), Scotland (Goodwin et al., 2001), Spain (Paredes et al., 2003), Swiss (Gehrig et al., 1999), Turkey (Akbasak et al., 2001) and Vietnam (Shimada et al., 2002).

Results

The observed allele frequencies (n=185), for the 13 core STR loci are similar to the frequencies obtained before (n=160), data not shown. The Neighbor-Joining tree based on Nei's genetic distances is represented in figure 1.

Discussion

A comparison of Santa Catarina population with the others showed minor divergences with Caucasian populations. On the other hand, significant differences were found between Santa Catarina and other Brazilian (Rio Janeiro and Mato Grosso Sul), South American, Asiatic and Negroid populations. It can be seen that Caucasian populations, mostly European appear on the same cluster that contains Santa Catarina. A possible reason for this is the higher contribution of Europeans to the state formation. The closeness of Santa Catarina with Azores, Madeira and Portugal, may reflect their

Portuguese origin. The Brazilian Mato Grosso Sul population together with other South American populations (Peru and Paraguay) included in this analysis are grouped on a second cluster. A probable cause for this could be their geographic proximity.

The Asiatic populations can also be found in this group; however in a different branch. On the other hand, the Brazilian Rio Janeiro population is grouped with populations of Negroid origin. Assuming a high degree of admixture of this population (Portuguese, African descendants and native Brazilians), it was not surprising to find it in this group.

Conclusion

The genetic differences found between all compared populations, namely the Brazilian ones, grouped in different clusters, support the importance of developing specific local databases of the reference populations. In Brazil this necessity is essential because of the enormous population heterogeneity and the vast area occupied by this country.

References

- ABRANTES D, LIMA G, CAINÉ L, PONTES L, PINHEIRO F, Allele Frequencies of 15 Loci Using AmpFISTR Identifiler Kit in a Northern Portuguese Population. *J Forensic Sci* 49(4): 842-844, 2004.
- AKBASAK B, BUDOWLE B, REEDER D, REDMAN J, KLINE N, Turkish population data with the CODIS multiplex short tandem repeat loci. *Forensic Sci Int* 123:227-229, 2001.
- BAGDONAVICIUS A, TURBETTE G, BUCKLETON J, WALSH S, Western Australian Sub-Population Data for the Thirteen AMPFISTR® Profiler Plus™ and COfiler™ STR Loci. *J Forensic Sci* 47(5):1149-1153, 2002.
- BUDOWLE B, SHEA B, NIEZGODA S, CHAKRABORTY R, CODIS STR data from 41 sample populations. *J Forensic Sci* 46(3):453-489, 2001.
- CAINE L, CORTE-REAL F, ANJOS M, CARVALHO M, SERRA A, ANTUNES H, VIDE M, Allele Frequencies of 13 Loci in the Santa Catarina Population of Southern Brazil. *J Forensic Sci* 48(4):901-902, 2003.
- EFRON B, *The Jackknife, the Bootstrap, and other resampling plans.* Society for Industrial and Applied Mathematics. Philadelphia, 1992.
- ESPIN E, ROSADO F, ALVAREZ J, ENTRALA C, LORENTE J, DUARTE M, VILLANUEVA E, BUDOWLE B, Paraguayan Population Data on the Fifteen STR Loci Included in PowerPlex 16™ Kit. *J Forensic Sci* 48(1):253-255, 2002.
- FARFÁN M, PRIETO V, TORRES Y, LÓPEZ-SOTO M, SANZ P, STR data for the AmpFISTR Profiler Plus and COfiler loci from Maghreb (North Africa). *Forensic Sci Int* 121:199-200, 2001.
- FELSENSTEIN J, Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783-791, 1985.

- FELSENSTEIN J, PHYLIP – Phylogeny inference package (version 3.5c). *Cladistics* 5:164-166, 1996.
- FERNANDES A, BREHM A, ALVES C, GUSMÃO L, AMORIM A, Genetic profile of the Madeira Archipelago population using the new PowerPlex®16 System Kit. *Forensic Sci Int* 125:281-283, 2002.
- GAROFANO L, PIZZAMIGLIO M, VECCHIO C, LAGO G, FLORIS T, D'ÉRRICO G, BREMBILLA G, ROMANO A, BUDOWLE B, Italian population data on thirteen short tandem repeat loci: HUMTH01, D21S11, D18S51, HUMVWFA31, HUMFIBRA, D8S1179, HUMTPOX, HUMCSF1PO, D16S539, D7S820, D13S317, D5S818, D3S1358. *Forensic Sci Int* 97:53-188, 1998.
- GEHRIG C, HOCHMEISTER M, BORER V, DIRNHOFER R, BUDOWLE B, Swiss Caucasian Population Data for 13 STR Loci Using AmpFISTR Profiler Plus and COfiler PCR Amplification Kits. *J Forensic Sci* 44(5):1035-1038, 1999.
- GÓES A, SILVA D, GIL E, SILVA M, PEREIRA R, CARVALHO E, Allele frequencies data and statistic parameters for 16 STR loci – D19S433, D2S1338, CSF1PO, D16S539, D7S820, D21S11, D18S51, D13S317, D5S818, FGA, Penta E, TH01, vWA, D8S1179, TPOX, D3S1358 – in the Rio de Janeiro population, Brazil. *Forensic Sci Int* 140:131-132, 2004.
- GOODWIN W, SCOLAR C, LINACRE A, 13 STR loci frequency data from a Scottish population. *Forensic Sci Int* 116:187-188, 2001.
- HASHIYADA M, ITAKURA Y, NAGASHIMA T, NATA M, FUNAYAMA M, Polymorphism of 17 STRs by multiplex analysis in Japanese population. *Forensic Sci Int* 133:250-253, 2003.
- JAUFFRIT A, AMRI H, AIRAUD F, ANDRÉ M, HERBERT O, TROTTIER G, GIRAUDET S, RICHARD C, CHVENTRÉ A, MOISAN J, DNA Short Tandem Repeat Profiling of Morocco. *J Forensic Sci* 48(2):458-459, 2003.
- LAW M, TO K, HO S, PANG B, WONG L, WUN H, YAU S, CHAN K, STR data for the PowerPlex™ 16 loci for the Chinese population in Hong Kong. *Forensic Sci Int* 129:64-67, 2002.
- NEI M, Genetic distance between populations. *The American Naturalist* 106:283-292, 1972.
- NEI M, SAITOU N, The Neighbor-Joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406-425, 1987.
- PAGE M, TreeView: An application to display phylogenetic trees on personal computers. *Computer Application in the Biosciences* 12:357-358, 1996.
- PAREDES M, CRESPILO M, LUQUE J, VALVERDE J, STR frequencies for the PowerPlex®16 System Kit in a population from Northeast Spain. *Forensic Sci Int* 135:75-78, 2003.
- PEPINSKI W, JANICA J, SKAWRONSKA M, NIEMCUNOWICZ-JANICA A, SOLTYSZEWSKI I, Population genetics of 15 STR loci in the population of Podlasie (NE Poland). *Forensic Sci Int* 124:226-227, 2001.
- PÉREZ L, HAU J, IZARRA F, OCHOA O, ZUBIATE U, GARCÍA O, Allele frequencies for the 13 CODIS STR loci in Peru. *Forensic Sci Int* 132:164-165, 2003.
- SAHOO S and KASHYAP V, Allele Frequency Data for PowerPlex 16 Loci in Four Major Populations of Orissa, India. *J Forensic Sci* 47(4):912-915, 2002.
- SHIMADA I, BRINKMANN B, TUYEN N, HOHOFF C (2002) Allele frequency data for 16 STR loci in the Vietnamese population. *Int J Legal Med* 116:246-248, 2002.

- SILVA D, CROUSE C, CHAKRABORTY R, GÓES A, CARVALHO E, Statistical analyses of 14 short tandem repeat loci in Brazilian populations from Rio de Janeiro and Mato Grosso do Sul states for forensic and identity testing purposes. *Forensic Sci Int* 139:173-176, 2004.
- SKITSA I, SALAS A, LAREU M, CARRACEDO A, STR – CODIS typing in Greece. *Forensic Sci Int* 137:104-106, 2003.
- VELOSA R, FERNANDES A, BREHM A, Genetic profile of the Açores Archipelago population using the new PowerPlex®16 System Kit. *Forensic Sci Int* 129:68-71, 2002.

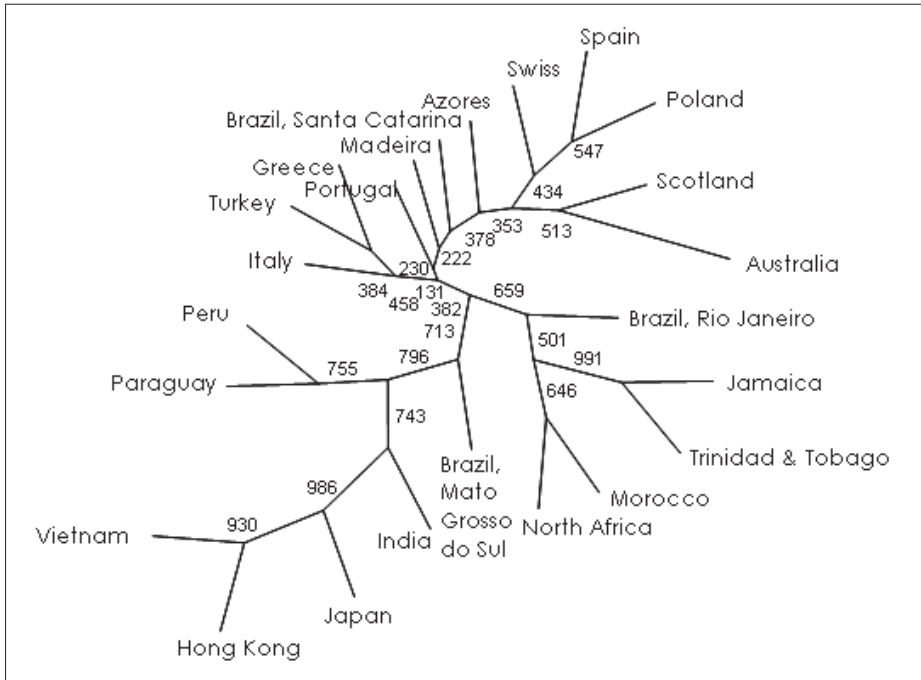


Figure 1 – A Neighbor-Joining tree based on Nei's genetic distances for the 13 STRs loci in 24 populations. The numbers indicate the bootstrap value.