Forensic Science International: Genetics 6 (2012) e131–e133

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Letter to the Editor

## Sequence variation of mitochondrial DNA control region in North Central Venezuela

Dear editor,

The analysis of the complete mitochondrial DNA control region (mtDNA-CR) has become increasingly common in the last years, in part due to general interest in the greater discriminatory power and phylogenetic signal provided by entire control region data [1]. In countries like those of the Americas, with known historical admixture between genetically different groups, characterization of the mtDNA diversity is critical for understanding regional heterogeneity and substructure. These types of population genetic features are, in turn, important for the proper application of statistics in mtDNA-based forensic casework.

Venezuelan population is the product of a complex admixture, through a process that was not homogeneous throughout the country [2]. Caracas, Venezuela's capital located in the north-central region, has been an important economic center for people coming from across the country and from abroad. As a result, its population is very heterogeneous in its socio-economic composition, with evident genetic and cultural admixture as well. Here, a population of the North Central region of Venezuela (NCR) was studied to contribute to the development of a high quality database of entire mtDNA-CR sequences.

Blood samples from 101 unrelated voluntary donors born in the north-central Region (NCR) of Venezuela and living in the city of Caracas were analyzed. After obtaining informed consent, blood samples were collected and DNA extracted according to conditions previously described [3]. The complete control region between positions 16024–573 and the adjacent 5′ portion between positions 15878 and 16023 in rCRS [4,5] were PCR-amplified, sequenced and analyzed following the strategies outlined in Brandstätter et al. [6] and Irwin et al. [7]. ISFG guidelines for mtDNA were followed [8]. An additional quality control check was performed by EMPOP [9] (http://www.empop.org).

Haplogroup membership was initially allocated employing Haplogrep [10] and Phylotree version 12 [11]. Haplotypes assigned to haplogroups A–D were reviewed and their clade membership revised and/or refined according to available phylogeographical evidence (see footnotes in ESM Table S1). Arlequin v3.1 software [12] was employed to calculate nucleotide diversity, sequence diversity and mean number of pairwise differences; the random match probability was calculated as was previously described [13]. Indels in poly-C stretches at positions 16184–16193, 303–309, 568–573, and the 514–523 AC dinucleotide variation were ignored in statistical evaluations of diversity [14,15]. Haplogroups frequencies were used to calculate pairwise differences between populations. An AMOVA analysis was performed to detect population differentiation or substructure using Arlequin v3.1 software.

Sequence data for the 5′ extended complete control region of our sample set from NCR are given in ESM Table S1. Eighty-three haplotypes were defined when C-stretch polymorphisms were included. Seventy-two of the lineages (86.7%) were unique. The analysis of the complete control region identified a greater number of haplotypes in our sample (83) than the 79 haplotypes previously reported by Lander et al. in their dataset of 100 [16].

High values of sequence diversity (0.9939), nucleotide diversity ($0.0137 \pm 0.007$) and mean number of pairwise differences ($16.6049 \pm 7.454$) were calculated from the data. Consequently the random match probability is low (2.08%) and the power of discrimination high (0.9791).

The diversity values obtained for our sample considering the entire region control are slightly higher than those reported for Caracas (HVSI + II + III, 0.986 [16]) and other admixed Latin American populations such as Argentina (northern, 0.906; central, 0.937 and southern, 0.878) [17], and they are also higher than values ranging from 0.897 to 0.955 (HVSI + II) reported for indigenous peoples in central and south America. The values reported here are comparable to values of 0.9985 reported for African Americans [18] and of 0.993 for Santa Catarina, Brazil [14]. The heterogeneous geographic origin of the inmigrants who participated in the admixture process in America, may explain these high values. The diversity estimates for the mtDNA-CR, and the sequence diversity values indicate that the use of mitochondrial DNA in the NCR is very informative for forensic casework, since the probability of discriminating two maternal lineages is high.

Haplogroups found in our dataset showed a strong female Amerindian contribution to the NCR population (66.7%), with all of the major pan-American haplogroups present. Haplogroup A2 was most common (36.6%), followed by C1 (13.9%), B4b and B2 (7.9%), D1 (6.9%) and D4h3a (1%). African haplogroups reached a frequency of 26%, while the West Eurasians lineages were present at only 8% (Table 1). In order to determine whether or not the NCR sample exhibited the same pattern of mtDNA heterogeneity by socioeconomic level that has been previously reported for Caracas, the NCR sample was stratified by socioeconomic levels using the criteria of Martínez et al. [19] and then re-evaluated for haplogroup composition. About 48% of the sample was classified in the middle socioeconomic strata, while the low and high strata were each represented by about 26% of the sample. The distribution of haplogroups among the three socioeconomic levels showed no statistically significant differences (data not shown). There was, however, a greater prevalence of Eurasian haplogroups in the upper stratum, as previously reported by Martínez et al. [19]. The overall haplogroup composition of the NCR sample reflects generally low European female migration to Venezuela. The higher frequency of African and Amerindian maternal lineages, on the other hand, can be explained by their historical contribution during colonial times. The exception to this general pattern, found in the upper socioeconomic level of the Caracas population where the occurrence of European maternal lineages was greater, is the result of recent preferential migration towards this stratum [19].

1872-4973/$ – see front matter 
doi:10.1016/j.fsigen.2011.11.004

**Table 1**
Observed values and percentages (%) of mitochondrial DNA haplogroups in NCR, Caracas, Lara and Oriente states, Venezuela.

| Haplogroups | NCR (This Study) | Caracas Lander et al. [16] | Caracas Castro de Guerra et al. [20] | | Lara Castro de Guerra et al. [20] | Oriente Castro de Guerra et al. [20] |
|---|---|---|---|---|---|---|
| | | | High[a] | Low[a] | | |
| A | 37 (37) | 34 (34) | 13(25) | 19 (38) | 18 (22) | 54 (43) |
| B | 8 (8) | 5 (5) | 2 (4) | 8 (16) | 29 (36) | 22 (18) |
| C | 14 (14) | 20 (20) | 6 (12) | 9 (18) | 5 (6) | 13 (10) |
| D | 8 (8) | 8 (8) | 4 (8) | 3 (6) | 8 (10) | 5 (4) |
| L (×L3) | 17 (17) | 11 (11) | 6 (11) | 4 (8) | 15 (19) | 13 (10) |
| L3 | 9 (9) | 8 (8) | 5 (10) | 3 (6) | 5 (6) | 10 (8) |
| H | 5 (5) | 3 (3) | 7 (14) | 3 (6) | 0 | 1 (1) |
| Others | 3 (3) | 11 (11) | 8 (16) | 1 (2) | 1(1) | 8 (6) |
| Total | 101 | 100 | 51 | 50 | 81 | 125 |

[a] Socieconomic strata.

**Table 2**
Pairwise differences between NCR, Caracas, Lara and Oriente states, Venezuela.

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 0.00000 | | | | | |
| 2 | −0.00133 | 0.00000 | | | | |
| 3 | 0.08573** | 0.05549** | 0.00000 | | | |
| 4 | 0.00013 | 0.01000 | 0.07919** | 0.00000 | | |
| 5 | 0.04339** | 0.06846** | 0.15811** | 0.03193* | 0.00000 | |
| 6 | 0.00750 | 0.02730* | 0.11858** | −0.00926 | 0.03467** | 0.00000 |

1: NCR (This Study), 2: Caracas (Lander et al. [16]), 3: Caracas-High Strata. 4: Caracas-Low Strata, 5: Lara, 6: Oriente (Castro de Guerra et al. [20]). Fst = 0.04071.
* $p < 0.05$.
** $p < 0.001$

The distribution of haplogroups found in NCR (Table 1) was compared with previously reported RFLP data from Caracas, the Venezuelan western state of Lara and the eastern states represented by Sucre, Monagas and Nueva Esparta [19,20]. The results indicate that the haplogroup distributions in the capital region, NCR (this study), Caracas (Lander et al. [16]), and the low socioeconomic strata of Caracas [19] are similar (Table 2). However, the sample of the high socioeconomic level of Caracas shows significant differences with the rest of the capital region due to a higher proportion of European haplogroups. Lara State also shows significant differences in the distribution of haplogroups when compared to all samples of the capital and eastern region of the country. The coefficient of differentiation obtained for all samples suggests genetic heterogeneity or substructuring (Fst = 0.040, $p < 0.001$). These results are consistent with the complexity of admixture reported for Venezuela [2,20].These results suggest that the sample set analyzed for NCR can be regarded as representative of the distribution of mitochondrial haplotypes in the capital region of Venezuela and can be used as an appropriate reference database for forensic casework in this region. However, due to the great diversity observed within Caracas, interpretations are complicated. In this regard, is recommended: (1) to increase the sample size for a better representation of diversity, (2) further characterization of mtDNA distributions based on socioeconomic level may justify the application of statistical models with the appropriate corrections [21–23]. Similarly, due to the heterogeneity reported for the Venezuelan population and the differential distribution of haplogroups between the capital, eastern and the samples from the western portion of the country, we suggest the need for a national reference mtDNA database, in which each region could be considered separately. Such a database would contribute to a thorough understanding of the diversity patterns within Venezuela at subpopulation levels, and help establish appropriate interpretational guidelines for each forensic case analyzed.

This paper follows the guidelines for publications of the journal [1]. The haplotypes reported in the present work will be made available via EMPOP (www.empop.org) and GenBank upon publication.

## Acknowledgements

The authors would like to thank Erin Gorden, Lisa Bedford, Christopher May and Laura Dolezal for assistance with data analysis and review, along with the US National Institute of Justice (Interagency Agreement 2005-DN-R-086) for funding this work. The opinions and assertions contained herein are solely those of the authors and are not to be construed as official or as views of the US Department of Defense or the US Department of the Army. In Venezuela funds were partially granted by IVIC and FONACIT, Grant 2008000919.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fsigen.2011.11.004.

## References

[1] A. Carracedo, J.M. Butler, L. Gusmão, W. Parson, L. Roewer, P.M. Schneider, Publication of population data for forensic purposes, Forensic Sci. Int. Genet. 4 (2010) 145–147.
[2] A. Rodríguez-Larralde, D. Castro de Guerra, M. González Coira, J. Morales, Frecuencia génica y porcentaje de mezcla en diferentes áreas geográficas de Venezuela, de acuerdo a los grupos Rh y ABO, Interciencia 26 (2001) 8–12.
[3] A. Simmons, G. Rodríguez-Arroyo, A. Rodríguez-Larralde, Admixture estimates based on ABO, Rh and nine STRs in two Venezuelan regions, Ann. Hum. Biol. 34 (2007) 56–67.
[4] S. Anderson, A.T. Bankier, B.G. Barrell, M.H. de Bruijn, A.R. Coulson, J. Drouin, I.C. Eperon, D.P. Nierlich, B.A. Roe, F. Sanger, P.H. Schreier, A.J. Smith, R. Staden, I.G. Young, Sequence and organization of the human mitochondrial genome, Nature 290 (1981) 457–465.
[5] R.M. Andrews, I. Kubacka, P.F. Chinnery, R.N. Lightowlers, D.M. Turnbull, N. Howell, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA, Nat. Genet. 23 (1999) 147.
[6] A. Brandstätter, C.T. Peterson, J.A. Irwin, S. Mpoke, D.K. Koech, W. Parson, T.J. Parsons, Mitochondrial DNA control region sequences from Nairobi (Kenya): inferring phylogenetic parameters for the establishment of a forensic database, Int. J. Legal Med. 118 (2004) 294–306.

[7] J.A. Irwin, J.L. Saunier, K.M. Strouss, K.A. Sturk, T.M. Diegoli, R.S. Just, M.D. Coble, W. Parson, T.J. Parsons, Development and expansion of high quality control region databases to improve forensic mtDNA evidence interpretation, Forensic Sci. Int. Genet. 1 (2007) 154–157.
[8] W. Bär, B. Brinkmann, B. Budowle, A. Carracedo, P. Gill, M. Holland, P.J. Lincoln, W. Mayr, N. Morling, B. Olaisen, P.M. Schneider, G. Tully, M. Wilson, DNA Commission of the International Society for Forensic Genetics: Guidelines for mitochondrial DNA typing, Int. J. Legal Med. 113 (2000) 193–196.
[9] A. Brandstätter, H. Niederstätter, M. Pavlic, P. Grubwieser, W. Parson, Generating population data for the EMPOP database - an overview of the mtDNA sequencing and data evaluation processes considering 273 Austrian control region sequences as example, Forensic Sci. Int. 166 (2007) 164–175.
[10] A. Kloss-Brandstaetter, D. Pacher, S. Schöenherr, H. Weissensteiner, R. Binna, G. Specht, F. Kronenberg, HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups, Hum. Mutat. 32 (2011) 25–32, http://www.haplogrep.uibk.ac.at.
[11] M. van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation, Hum. Mutat. 30 (2009) E386–E394, http://phylotree.org.
[12] L.G. 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.
[13] M. Stoneking, D. Hedgecock, R.G. Higuchi, L. Vigilant, H.A. Erlich, Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes, Am. J. Hum. Genet. 48 (1991) 370–382.
[14] L. Palencia, L. Valverde, A. Alvarez, L.M. Cainé, S. Cardoso, M.A. Alfonso-Sánchez, M.F. Pinheiro, M.M. de Pancorbo, Mitochondrial DNA diversity in a population from Santa Catarina (Brazil): predominance of the European input, Int. J. Legal Med. 124 (2010) 331–336.
[15] A.O. Tillmar, M.D. Coble, T. Wallerström, G. Holmlund, Homogeneity in mitochondrial DNA control region sequences in Swedish subpopulations, Int. J. Legal Med. 124 (2009) 91–98.
[16] N. Lander, M.G. Rojas, M.A. Chiurillo, J.L. Ramírez, Haplotype diversity in human mitochondrial DNA hypervariable regions I–III in the city of Caracas (Venezuela), Forensic Sci. Int. Genet. 2 (2008) e61–e64.
[17] M.C. Bobillo, B. Zimmermann, A. Sala, G. Huber, A. Röck, H.-J. Bandelt, D. Corach, W. Parson, Amerindian mitochondrial DNA haplogroups predominate in the population of Argentina: towards a first nationwide forensic mitochondrial DNA sequence database, Int. J. Legal Med. 124 (2010) 263–268.
[18] T.M. Diegoli, J.A. Irwin, R.S. Just, J.L. Saunier, J.E. O'Callaghan, T.J. Parsons, Mitochondrial control region sequences from an African American population sample, Forensic Sci. Int. Genet. 4 (2009).
[19] H. Martínez, A. Rodríguez-Larralde, M.H. Izaguirre, D. Castro de Guerra, Admixture estimates for Caracas, Venezuela, based on autosomal, Y-chromosome, and mtDNA markers, Hum. Biol. 79 (2007) 201–213.
[20] D. Castro de Guerra, C. Figuera Pérez, M.H. Izaguirre, E. Arroyo Barahona, A. Rodríguez Larralde, M. Vívenes de Lugo, Gender differences in ancestral contribution and admixture in Venezuelan populations, Hum. Biol. 83 (2011) 345–361.
[21] D.W. Gjertson, C.H. Brenner, M.P. Baur, A. Carracedo, F. Guidet, J.A. Luque, R. Lessig, W.R. Mayr, V.L. Pascali, M. Prinz, P.M. Schneider, N. Morling, ISFG: Recommendations on biostatistics in paternity testing, Forensic Sci. Int. Genet. 1 (2007) 223–231.
[22] J.U. Palo, M. Hedman, I. Ulmanen, M. Lukka, A. Sajantila, High degree of Y-chromosomal divergence within Finland-forensic aspects, Forensic Sci. Int. Genet. 1 (2007) 120–124.
[23] A. Salas, H.J. Bandelt, V. Macaulay, M.B. Richards, Phylogeographic investigations: the role of trees in forensic genetics, Forensic Sci. Int. 168 (2007) 1–13.

D. Castro de Guerra[a,*]
C. Figuera Pérez[a]
C.M. Bravi[b]
J. Saunier[c]
M. Scheible[c]
J. Irwin[c]
M.D. Coble[c]
A. Rodríguez-Larralde[a]
[a]*Laboratorio de Genética Humana, Centro de Medicina Experimental, Instituto Venezolano de Investigaciones Científicas, Caracas 1020 A, Venezuela*
[b]*IMBICE, CCT-La Plata CONICET, (1900) La Plata, Argentina*
[c]*Armed Forces DNA Identification Laboratory, Rockville, MD 20850, United States*

*Corresponding author. Tel.: +58 2125041087;
fax: +58 2125041086
*E-mail address:* dinorah_castro@hotmail.com
(D. Castro de Guerra)

22 December 2010