Forensic Science International: Genetics 6 (2012) 136–142

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Short communication

# Distribution of Y chromosomal STRs loci in Mayan and Mestizo populations from Guatemala

Luis J. Martínez-González [a,*], María Saiz [a], María J. Álvarez-Cubero [a], Antonio Gómez-Martín [b], Juan C. Álvarez [a,d], Cristina Martínez-Labarga [c], José A. Lorente [a,d]

[a] Laboratory of Genetic Identification, Department of Legal Medicine, Universidad de Granada, Granada, Spain
[b] Savia Biotech S. A., Avda. Sudamérica, 94 04740 Roquetas de Mar, Almería, Spain
[c] Centre of Molecular Antropology for Ancient DNA Studies, Dipartimento di Biologia, Università di Roma "Tor Vergata", Rome, Italy
[d] Pfizer – Universidad de Granada – Junta de Andalucia Centre for Genomics and Oncological GENYO, Centro de Investigación Biomédica, Avda. del Conocimiento s/n, 18100, Armilla, Granada, Spain

ARTICLE INFO

*Article history:*
Received 15 September 2010
Received in revised form 18 February 2011
Accepted 2 April 2011

*Keywords:*
Genetic relationship
Admixture
Population genetics
Forensic science
DNA typing

ABSTRACT

In this study, a sample of 225 Guatemalan males, comprising 115 Mestizo-Guatemalan and 110 Mayan-Guatemalan, was typed for 17 Y-short tandem repeats (STRs) loci (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635, YGATA_H4.1 and DYS385a/b). The haplotype diversity ($H = 1$) and discrimination capacity (96.86%) were calculated. Analysis of molecular variance (AMOVA) demonstrated a low but significant interpopulation differentiation when compared with the results obtained when we confront the Mestizo and Mayan populations with the European populations.

Furthermore, the genetic variability and differences among the American, African, Asian, and European populations were analyzed with the software Statistica 9.1. In addition, the genetic distances were also calculated using other published data. Reynolds and Slatkiń genetic distance was visualized using the multidimensional scaling (MDS) analysis. All the analysis performed locates the Mayan population next to the Native American population, while Guatemalan-Mestizo population was found to be between these populations and the European population, similar to other Mestizo one.

The implementation of the estimation of individual ancestry proportions of the whole population sample showed the presence of two well-differentiated population groups.



## 1. Population

A total of 225 DNA samples was obtained from unrelated male donors from both the populations studied, comprising 115 Guatemalan-Mestizo inhabitants and 110 Guatemalan-Mayan individuals [Kaqchiquel (22), Ḱiché (37), Mam (27), and Ǫeqchi (24)] (see Fig. 1). Samples from Mayan individuals were obtained directly at their communities. About 25 μl of blood samples were spotted on FTA® Whatman paper under informed consent.

The population of Guatemala includes 12,727,566 individuals [1] and according to INE Guatemala, 48.92% of them are men [2]. Nowadays, the population of Guatemala can be divided into two principal groups: Native-American groups (40.1%) and the Spanish-speaking population (Mestizo and European descendents) (59.4%) [1]. The main Mayan ethnic groups living in the country are K'iche' (9.1%), Kaqchikel (8.4%), Q'eqchi' (6.3%), and Mam (7.9%) [1]. The great majority of Guatemalan-Mestizos are the result of admixture between these Mayan ethnic groups and Spaniards.

In order to perform the analysis of molecular variance (AMOVA), it was included a sample of 175 Spanish individuals previously analyzed.

## 2. Extraction

DNA extract was prepared for PCR analysis using standard FTA® protocols (Whatman, Clifton, NJ, USA) and 1.2 mm punch of FTA cards was used in each amplification.

## 3. PCR amplification

The samples were amplified using AmpFlSTR® Yfiler® kit including 17 loci (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635, YGATA_H4.1 and DYS385a/b) (Applied Biosystems, Foster City, CA) [3].

* Corresponding author. Tel.: +34 958249950; fax: +34 958246107.
*E-mail address:* luisjaviermg@ugr.es (L.J. Martínez-González).

1872-4973/$ – see front matter © 2011 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigen.2011.04.003


**Fig. 1.** Mayan and Mestizo distribution in Guatemala.

## 4. Typing

Alleles were separated and detected using an Applied Biosystems ABI 310 genetic analyzer. Fragment sizes were analyzed using the GeneMapper ID-X v.1.1 software (Applied Biosystems, Foster City, CA). The alleles were named according to the number of repeated units based on the sequenced allelic ladder (ISFG recommendations) [4].

## 5. Quality control

The regular participation of the lab in the proficiency testing of the GEP-ISFG Working Group (http://www.gep-isfg.org) and Proficiency testing of the GITAD (http://gitad.ugr.es/principal.htm) is certified.

The Y-STR Haplotype Reference Database [5] accession number for the population sample "Guatemala [Mestizo]" $n = 115$ is YA003664; and the YHRD accession number for Guatemala, Mayan, $n = 110$ are Chimaltenango, Guatemala [Maya, Kaqchiquel] YA003663, Santa Cruz del Quiché, Guatemala [Maya, Ḱiché] YA003666,Alta Verapaz, Baja Verapaz, Petén e Izabal, Guatemala [Maya, Qeqchí] YA003662 and Northeast, Guatemala [Maya, Mam] YA003665 (http://www.yhrd.org/).

## 6. Analysis of data

Haplotype frequencies were calculated by the gene counting method. Haplotype diversities were calculated according to Nei [6,7] using Arlequin v3.0 software [6]. Previously published haplotypes and frequencies data from several populations were compiled, which comprised South American (Toba, Argentina [8]; El Beni, Bolivia [9]; Rio de Janeiro, Brazil [10]; Rio Grande do Sul, Brazil [11]; Mestizo, Ecuador [12]; and Kichwa, Ecuador [12]), North American (African descendent; European descendent; Hispanic descendent) [13], African (Algeria [14]; Equatorial Guinea [15]; Ovambo, Namibia [16]), Asian (Japan [17]; Korea [18]; Chinese Tibetan [19]; Chaoshan, China [20]), and European (Spain [21]; Portugal [22]; Italy [23]; Croatia [24]; Serbia [25]; and Austria [26]). To study the ancestry relationships, six markers (DYS385a/b, DYS439, DYS456, DYS458 and DYS635) with the highest mutation rates were eliminated [27]. A correspondence analysis was calculated with Statistica 9.1 (Statsoft Inc. Tulsa USA) to see the association between allele frequencies and populations. Arlequin v3.0 software was used to calculate the distributions of the observed allele frequencies in each group, F-statistics and Reynolds and Slatkiń genetic distances. Arlequin v3.0 software was also used to test the hypothesis of a random distribution of individuals between the pairs of populations with an exact test of population differentiation and to perform AMOVA [28]. To represent the distances in a more appropriate way, genetic distances (Reynolds and Slatkiń genetic-distance matrix) between the American and European populations were graphically summarized by nonmetric multidimensional scaling (NM-MDS) [29], using the software SSPS v.15.0. NM-MDS uses an iterative process to transform a similarity/dissimilarity matrix into distances represented in a Euclidean $n$-dimensional space. Four markers were eliminated owing to their high mutation rates (DYS439, DYS458, DYS456 and DYS635) [27].

Furthermore, the admixtured proportions of the Mayan and South West European populations (90% Spanish and 10% Portugal data) in Guatemalan-Mestizo population were estimated by means of the weighted least squares (WLS), ADMIX.PAS program [30] and ADMIX95 program [31]. We implemented the estimation of individual ancestry proportions with the program STRUCTURE 2.3.1 [32–34] http://pritch.bsd.uchicago.edu/structure.html. Replicate runs of STRUCTURE using different burn-in period and interactions were obtained, and here, we have reported the estimations with a burn-in period of 50,000 interactions followed by an additional 100,000 interactions ($K = 1$, 2 and 3) with all the populations. We used two models of the available options, namely admixture and no admixture models.

## 7. Results

A total of 225 samples from two Guatemalan populations, Mayan and Mestizo, were investigated in this study. All the 225 haplotypes obtained were different, and therefore, in our case, the haplotype diversity was 1.0. See electronic supplementary material, Tables S1 and S2 for haplotype distribution of the Y-chromosome obtained among the 225 individuals analyzed. The distributions of the observed allele frequencies in Mestizo and Mayan populations from Guatemala for 17 short tandem repeats (STRs) loci studied are shown in Tables 1 and 2. These tables show us the power of discrimination (PD) of the markers studied. The most informative marker was found to be DYS458 in both the populations. The less descriptive locus was observed to be DYS393 in the Guatemalan-Mestizo population and DYS437 in the Guatemalan-Mayan population. The combined PD in both the populations was higher than 0.999999.

## 8. Other remarks

Population genetic variation studies have demonstrated that there is an overall low level of differentiation in human populations [35]; however, local factors, such as geography and differential settlement can greatly enhance genetic discontinuity.

In this study, we have carried out an extensive analysis of Y-chromosomal diversity in Guatemalan populations, examining 17 STRs. These loci have allowed us to study the level of variability and substructure of the actual populations of Guatemala. AMOVA revealed that the highest variation is mainly within the populations. Both the Guatemalan populations displayed a low but significant interpopulation differentiation when compared with the results obtained when we confronted the Mestizo and Mayan populations with the European populations. The results are presented in Table 3.