



Contents lists available at ScienceDirect

HOMO - Journal of Comparative Human Biology

journal homepage: [www.elsevier.com/locate/jchb](http://www.elsevier.com/locate/jchb)

## Evaluation of the contribution of D9S1120 to anthropological studies in Native American populations

J.A. Aguilar-Velázquez<sup>a</sup>, V. Manuel Martínez-Sevilla<sup>a</sup>, M. Sosa-Macías<sup>b</sup>,  
A. González-Martin<sup>c</sup>, J.F. Muñoz-Valle<sup>d</sup>, H. Rangel-Villalobos<sup>a,\*</sup>

<sup>a</sup> Instituto de Investigación en Genética Molecular, Centro Universitario de la Ciénega, Universidad de Guadalajara (CUCI-UdeG), Av. Universidad #1115, Ocotlán, Jalisco, México, CP 47810

<sup>b</sup> Centro Interdisciplinario de Investigación para el Desarrollo Integral Regional del Instituto Politécnico Nacional, Unidad Durango (CIIDIR-IPN), Durango, México

<sup>c</sup> Departamento de Zoología y Antropología Física, Universidad Complutense de Madrid (UCM), 28040 Madrid, Spain

<sup>d</sup> Instituto de Investigación en Ciencias Biomédicas, Centro Universitario en Ciencias de la Salud (CUCS-UdeG), Guadalajara, Jalisco, México

### A B S T R A C T

The D9S1120 locus exhibits a population-specific allele of 9 repeats (9RA) in all Native American and two Siberian populations currently studied, but it is absent in other worldwide populations. Although this feature has been used in anthropological genetic studies, its impact on the evaluation of the structure and genetic relations among Native American populations has been scarcely assessed. Consequently, the aim of this study was to evaluate the anthropological impact of D9S1120 when it was added to STR population datasets in Mexican Native American groups. We analyzed D9S1120 by PCR and capillary electrophoresis (CE) in 1117 unrelated individuals from 13 native groups from the north and west of Mexico. Additional worldwide populations previously studied with D9S1120 and/or 15 autosomal STRs (Identifier kit) were included for interpopulation analyses. We report statistical results of forensic importance for D9S1120. On average, the modal alleles were the Native American-specific allele 9RA (0.3254) and 16 (0.3362). Genetic distances between Native American and worldwide populations were estimated. When D9S1120 was included in the 15 STR population dataset, we observed improvements for admixture estimation in Mestizo populations and for representing congruent genetic relationships in dendrograms. Analysis of molecular variance (AMOVA) based on D9S1120 confirms that most of the genetic variability in the Mexican population is attributable to their Native American backgrounds, and allows the detection of significant intercontinental differentiation attributed to the exclusive presence of 9RA in America. Our findings demonstrate the contribution of D9S1120 to a better understanding of the genetic relationships and structure among Mexican Native groups.

### Introduction

Sample analysis of the *Human Genome Diversity Project-Centre d'Etude du Polymorphisme Humain* (HGDP-CEPH) panel with 377 microsatellites or short tandem repeats (STRs) conducted in five Native Amerindian populations (Pima, Maya, Colombian, Karitiana, and Surui) found an elevated frequency of a small allele (275 base pairs) at the tetranucleotide locus D9S1120, which was absent in

\* Corresponding author. Tel.: +52 392 9257112; fax: +52 392 9257112.

E-mail address: [hrangel13@hotmail.com](mailto:hrangel13@hotmail.com) (H. Rangel-Villalobos).

<https://doi.org/10.1016/j.jchb.2017.10.003>

Received 12 November 2016; Accepted 20 October 2017

0018-442X/ © 2017 Elsevier GmbH. All rights reserved.

47 other worldwide populations (Zhitovskiy et al., 2003). Based on the corresponding number of repeats, this private allele was identified as “9RA” (9 repeats allele). This allele was observed in Native American populations at an overall average frequency of 0.354 (0.301 and 0.471 in North and South America, respectively) (Schroeder et al., 2009; Wang et al., 2007; Zhitovskiy et al., 2003). The ubiquitous presence of 9RA in North and South American populations, including the Na-Dene and Aleut-Eskimo, and in related western Beringian groups suggested that all modern Native American populations originated from the same founding population (Schroeder et al., 2007).

In Mexico, the STR D9S1120 has only been analyzed in eight Mexican ethnic groups (Tarahumaras, Huichol, Purepecha, Mazatecas, Tzotzil, Tjolobal, Lacandon and Maya) and five Mestizo groups (Chihuahua, Jalisco, Veracruz, Chiapas and Yucatán), where the forensic and anthropological potential of this locus was assessed (Rangel-Villalobos et al., 2012). However, Mexico has a larger number of Amerindian populations, with over 68 ethnic groups representing 9.6% of the total population (Martínez-Cortes et al., 2010). Although 39 Native American population samples from Mexico were recently analyzed with 15 autosomal STRs widely used for human identification purposes (Rangel-Villalobos et al., 2016), D9S1120 was not included; thus, its –presumably– large anthropological potential has not been exploited.

In this work, we performed a deeper assessment of the anthropological relevance of D9S1120 in Native American groups. For this purpose, D9S1120 was analyzed in north and west Mexican Native populations. In addition to reporting statistical results of the forensic importance for D9S1120, we carried out an interpopulation analysis adding a 15 autosomal STR database from Mexican populations and two populations used as ancestral references (European and African). Comparison of these results (with and without D9S1120) allowed for an empirical evaluation of their significance for studying Native American groups.

## Materials and methods

### Samples studied

Genomic DNA was extracted from 1,117 unrelated individuals belonging to the following 13 Amerindian groups inhabiting the north and west regions of Mexico: Mayos from Sonora and Sinaloa states; Guarijío; Seri; Mexicanero; Huichol samples from Durango, Nayarit and Jalisco states; Cora; Tepehuano; Tarahumara; and Nahuas from the Mezcala and Tuxpan communities in the Jalisco state. Geographical locations, sample sizes, and abbreviations of these Native American groups are presented in Fig. 1.

All participants signed an informed consent form approved by the Research and Ethical Committee of the CUCiénega-UdeG (Universidad de Guadalajara), México. As properly described in the results section, for some interpopulation analyses the following two Mexican population groups were included (abbreviations indicated): 1) Mexican Mestizos (admixed): Chihuahua (Chi); Jalisco

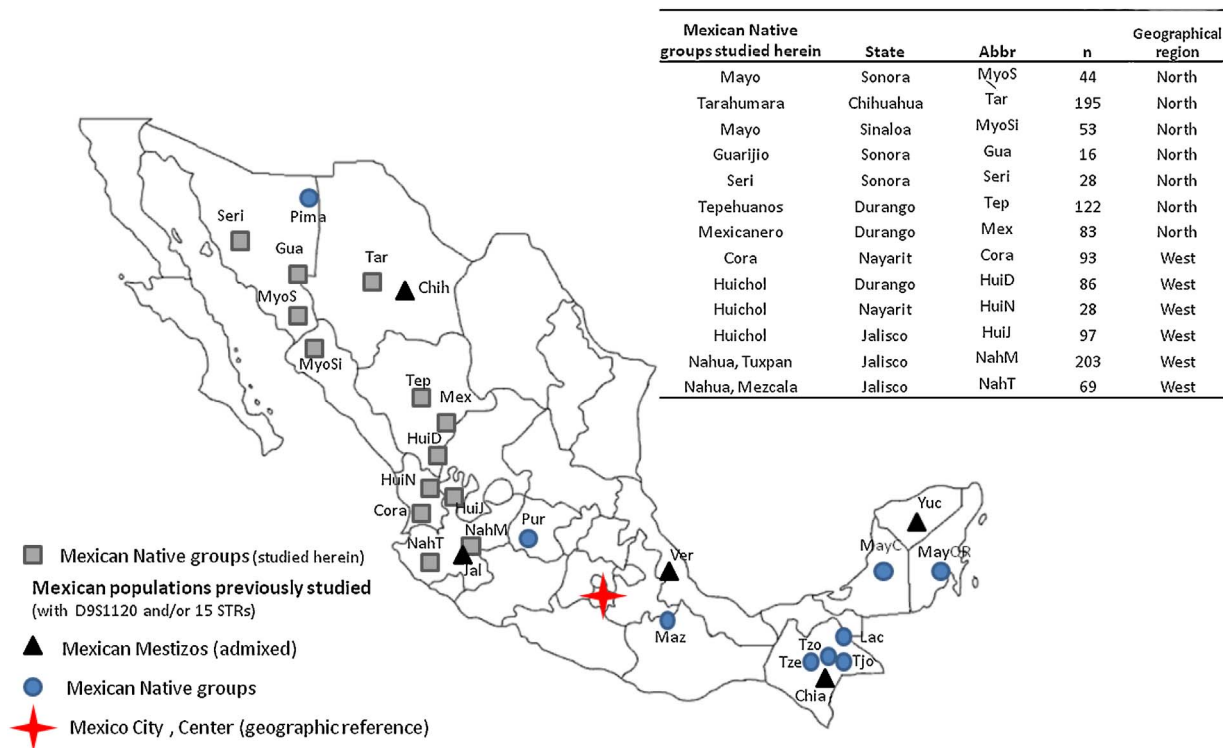


Fig. 1. Geographic locations, abbreviations, and sample sizes of ethnic groups from Mexico analyzed in this study. The location of additional Mexican groups (Native and Mestizos) is indicated in the map. Please check in Material and methods for additional abbreviations.

(Jal); Veracruz (Ver); Chiapas (Chia); Yucatan (Yuc); 2) Mexican Native groups: Purepechas (Pur); Mazatecas (Maz); Tzeltales (Tze); Tzotziles (Tzo); Tojolabales (Tjo); Lacandonones (Lac); Mayas from Campeche (MayC) and Quintana Roo (MayQR) states. Geographic locations of these populations throughout the Mexican territory are also indicated in Fig. 1. These groups were included because genotypes for D9S1120 (Rangel-Villalobos et al., 2012) and/or 15 STRs of the Identifier kit (Rangel-Villalobos et al., 2016) were available for interpopulation analyses. Similarly, the following worldwide populations were included for interpopulation comparison based on D9S1120: Yakut (Yak) (Rosenberg, 2006), Chokchi (Chok), Koryaks (Kor) (Schroeder et al., 2007) from Siberia; Colombians (admixed) (Col), Mulalos (Mul), Awa-Kuaiker (Awa), Coyaima (Coya), and Pijao from Colombia; Pimas from Mexico; as well as European (Eur) and African (Afr) populations (Phillips et al., 2008; Alves et al., 2005; Lópes et al., 2009; Coudray et al., 2007).

### Genotyping

DNA was extracted from fresh blood samples by the standard salting-out method (Miller et al., 1988). For amplification of D9S1120, we used primers and conditions described by Phillips et al. (2008). The PCR products were separated by capillary electrophoresis using the ABI Prism 310. D9S1120 genotypes were obtained using the GeneMapper ID software, version 3.2 (Applied Biosystems, Foster City, CA, USA). D9S1120 allele calling was done according to the repeat structure and size (bp) as described by Phillips et al. (2008), which follows the International Society of Forensic Genetics guidelines for STR analysis.

### Statistical analysis

Allele and genotype frequencies were estimated by the gene-counting method for each population. Forensic statistical parameters (PE: Power of exclusion; PD: Power of discrimination; PIC: Polymorphism information content; Typical IP: Paternity Index) were calculated with the Powerstats software (Tereba, 1999). Heterozygosity (He), Hardy-Weinberg equilibrium (HWE) expectations by population, and coancestry coefficients were determined as genetic distances between populations with the GDA v1.1 software (Lewis and Zaykin, 2001). Coancestry distances were represented in a Neighbor-Joining (NJ) tree using GDA and the Treeview v1.0 software (Page, 1996). Analysis of molecular variance (AMOVA), pairwise  $F_{ST}$  and  $F_{ST}$   $p$ -values between populations were determined by Arlequin v3.5 software (Excoffier and Lischer, 2010).  $F_{ST}$  distances were graphically displayed on a Multidimensional scaling (MDS) plot with the SPSS v19 (IBM corp., 2010).

### Ancestry analysis

The STRUCTURE software, a model-based clustering program designed to estimate the number ( $K$ ) of clusters has been used. This software determines posterior probability  $Pr(X/K)$  that fits with the data employed (Falush et al., 2003). In order to analyze the D9S1120 contribution, two supervised analyses were performed based on: 1) D9S1120 dataset and; 2) D9S1120 plus the 15 STR loci of the Identifier kit. We used one European (Eur) and one African (Afr) population samples as ancestral references. The admixture model with correlation between allele frequencies across clusters was chosen (Falush et al., 2003; Falush et al., 2007; Pritchard et al., 2000). The assessed number of clusters ( $K$ ) ranged from two to eight, and twenty-five independent runs were carried out for each  $K$ . We used a 50,000-iteration burn-in period followed by 10,000 iterations. The Structure-Harvester online program determining the best  $K$  according to the Evanno's method with the runs obtained was used (Earl and von Holdt, 2012). Finally, we employed the CLUMPAK program to visualize the best run for each  $K$  (Kopelman et al., 2015).

## Results

We estimated allele and genotype frequencies, in addition to statistical parameters of forensic importance estimated by population (Table 1). With the exception of the Huichol groups from Nayarit and Jalisco states, genotype distributions of the remaining population samples were in agreement with HWE. Nine D9S1120 alleles have been found and they are: 9, 12, 13, 14, 15, 16, 17, 18 and 19. This entire set of alleles was observed only in Nahuas from Tuxpan. Conversely, Cora and Seri showed only five alleles. For the entire sample, the modal alleles were 9 (33.5%) and 16 (33.6%). However, allele 9 was the most frequent among five groups: Mayo from Sinaloa, Huichol from Jalisco, Cora, Tepehuano, and Nahua from Mezcala, whereas allele 16 was the most frequent in the rest of the samples. Notably, for the first time, we found the allele 12 in three Mexican Native groups: Tarahumaras from Chihuahua state, Mayo groups from Sonora and Sinaloa states, and Nahuas from Tuxpan and Mezcala communities located in the Jalisco state. The heterozygosity ranged from 71% to 35.7% in Nahuas from Tuxpan and Huichols from Nayarit and Durango, respectively (Table 1). Thirty-three different genotypes were found in the dataset. As could be expected, genotypes which included the modal alleles were the most frequent: 9/16 (29.9%), 16/16 (13.3%) and 9/9 (12.9%) (Table 1).

Pairwise  $F_{ST}$  distances and  $F_{ST}$   $p$ -values based on D9S1120 among 26 Mexican groups and 11 other worldwide populations were estimated (Appendix 1).  $F_{ST}$  distances were graphically displayed in an MDS plot (Fig. 2). Among Mexican groups (Fig. 2a), although Mestizo (admixed) groups were not clearly differentiated from Native groups, Mestizos were relatively close to each other. Similarly, genetic distances based on D9S1120 among worldwide populations allow clustering Native American and Mestizo groups, whereas ancestral (European and African) and those highly differentiated populations remained in peripheral positions of the MDS plot (Mulalos, Yakut and Lacandonones) (Fig. 2b).

AMOVA showed that interpopulation differentiation based on D9S1120 was significant in most cases (Table 2). The second highest  $F_{ST}$  value was observed among worldwide populations ( $F_{ST} = 4.3\%$   $p = 0.016$ ). Although differentiation among all Mexican

Table 1

Allele and genotype frequencies and forensic parameters of D9S1120 in 13 ethnic groups from Mexico. For abbreviations see Material and methods (Statistical analysis).

Population Allele	MayS n = 44	MaySi n = 53	Gua n = 16	Ser n = 28	Mex n = 83	HuiD n = 86	HuiN n = 28	HuiJ n = 97	Cor n = 93	Tep n = 122	Tar n = 195	NahM n = 203	NahT n = 69	Total n = 1117
9	0.2386	0.3585	0.3197	0.2500	0.3214	0.3373	0.3333	0.4186	0.3750	0.4536	0.2921	0.3284	0.2578	0.3254
12	0.0114	0.0094	-	-	-	-	-	-	-	-	0.0025	0.0075	0.0208	0.0054
13	-	-	-	-	-	0.0060	-	0.0058	-	0.0052	-	0.0149	0.0104	0.0041
14	0.0227	0.0283	0.0205	0.0313	-	0.0422	0.0108	0.0058	0.0357	0.0103	0.0173	0.0224	0.1016	0.0333
15	0.1591	0.1415	0.2254	0.0938	0.0179	0.1446	0.1237	0.2151	0.1607	0.2268	0.2673	0.1567	0.1458	0.1845
16	0.4091	0.2925	0.3648	0.5313	0.4107	0.3735	0.4462	0.3023	0.2500	0.2629	0.3168	0.2612	0.3281	0.3362
17	0.0682	0.0943	0.0492	0.0938	0.2143	0.0422	0.0806	0.0523	0.1786	0.0412	0.0990	0.1791	0.1302	0.0927
18	0.0909	0.0660	0.0205	-	0.0357	0.0542	0.0054	-	-	-	0.0050	0.0149	0.0052	0.0171
19	-	0.0094	-	-	-	-	-	-	-	-	-	0.0149	-	0.0014
<b>Genotype</b>														
9/9	0.0681	0.1509	0.125	0.1071	0.1084	0.1976	0.2500	0.1855	0.0860	0.1229	0.1282	0.1133	0.1014	0.1287
9/10	-	-	-	-	-	-	-	-	-	-	-	-	0.0289	0.0018
9/12	-	-	-	-	-	-	-	-	-	-	0.0153	-	0.0144	0.0037
9/14	0.0227	-	-	-	0.024	-	-	-	-	-	0.0102	0.0098	-	0.0075
9/15	0.0454	0.1320	0.0625	-	0.1445	0.1744	0.1200	0.2268	0.0860	0.0983	0.0461	0.1034	0.0724	0.1033
9/16	0.1818	0.1132	0.1250	0.3214	0.2168	0.2441	0.1071	0.2680	0.3333	0.2377	0.1282	0.1674	0.2173	0.2086
9/17	0.0681	0.0943	-	-	0.0120	0.0232	0.0357	0.0412	0.0645	0.0409	0.0512	0.0689	0.0869	0.0516
9/18	0.0227	0.0566	-	-	0.0602	-	-	-	-	0.0163	-	0.0049	0.0144	0.0084
9/19	-	0.0188	-	-	-	-	-	-	-	-	-	-	0.0289	0.0018
10/11	-	-	-	-	-	-	-	-	-	-	0.0051	-	-	0.0009
10/14	-	-	-	-	-	-	-	-	-	-	0.0102	-	-	0.0018
12/12	-	-	-	-	-	-	-	-	-	-	0.0051	-	-	0.0009
12/14	-	-	-	-	-	-	-	-	-	-	0.0102	-	-	0.0018
12/15	-	-	-	-	-	-	-	-	-	-	-	0.0049	-	0.0009
12/16	0.0227	0.0188	-	-	-	-	-	-	-	-	0.0051	-	-	0.0018
13/13	-	-	-	-	-	-	-	-	-	-	-	-	0.0144	0.0009
13/14	-	-	-	-	0.0120	-	-	0.0010	-	-	0.0051	-	-	0.0028
13/15	-	-	-	-	-	-	-	0.0010	-	-	0.0051	-	-	0.0009
13/16	-	-	-	-	-	0.0116	-	-	-	-	0.0102	-	-	0.0028
14/14	-	-	-	-	-	-	0.0357	-	-	0.0081	0.0512	0.0049	0.0144	0.0131
14/15	-	-	-	-	0.0240	0.0116	-	0.0010	-	-	0.0051	0.0098	-	0.0065
14/16	0.0227	0.0377	-	-	-	-	-	-	0.0215	0.0245	0.0461	-	0.0144	0.0159
14/17	-	-	-	-	0.0120	-	-	-	-	-	0.0205	0.0049	-	0.0056
14/18	-	0.0188	-	-	-	-	-	-	-	-	-	-	-	0.0009
15/15	0.0454	0.0377	-	-	0.0120	0.0116	0.0714	0.0721	-	0.0573	0.0666	0.0788	0.0579	0.0516
15/16	0.0909	0.0377	0.1250	0.0357	0.0722	0.1046	0.0714	0.0618	0.1182	0.1967	0.0871	0.197	0.0724	0.1193
15/17	0.0227	0.0188	0.1250	-	0.012	0.0116	-	0.0010	0.0322	0.0163	0.0102	0.0591	0.0289	0.0234
15/18	0.0681	0.0188	-	-	0.012	-	-	-	0.0107	0.0245	-	0.0049	0.0144	0.0103
16/16	0.1818	0.1509	0.4375	0.1428	0.1807	0.1046	0.1428	0.0927	0.1935	0.1065	0.1589	0.1034	0.0434	0.1334
16/17	0.0454	0.0566	0.1250	0.1071	0.0481	0.0465	0.0357	0.0010	0.0430	0.0491	0.0512	0.0541	0.1159	0.0516
16/18	0.0909	0.0188	-	0.0357	0.0481	-	-	-	-	-	-	0.0049	-	0.0075
17/17	-	-	-	0.0714	-	0.0116	0.1428	0.0010	0.0107	-	0.0564	0.0049	0.0579	0.0234
17/18	-	0.0188	-	0.0357	-	-	-	-	-	-	0.0102	-	-	0.0028
<b>Forensic parameters</b>														
He	0.7045	0.6604	0.5000	0.6786	0.6988	0.3571	0.3571	0.6392	0.6989	0.7049	0.5260	0.6931	0.7101	
PD	0.8946	0.9021	0.7969	0.8240	0.8730	0.8421	0.8597	0.8224	0.8200	0.8564	0.9104	0.8802	0.8986	
PIC	0.6996	0.7144	0.5885	0.6196	0.6733	0.6258	0.6945	0.6139	0.6098	0.6577	0.7438	0.6838	0.7347	
PE	0.4353	0.3697	0.1875	0.3959	0.4264	0.3410	0.0899	0.3405	0.4266	0.4359	0.2113	0.4176	0.4441	
IPT	1.6923	1.4700	1.0000	1.5556	1.6600	1.3900	0.7778	1.3857	1.6607	1.6944	1.0549	1.6290	1.7250	
EHW	0.8464	0.3256	0.2002	0.5985	0.5662	0.2984	0.0001	0.0121	0.7720	0.1056	0.0501	0.0804	0.0544	

groups (Mestizos plus Amerindian groups) and among Mexican Amerindians was similar ( $F_{ST} = 2.3 - 2.2\%$ , respectively), differentiation among Mexican Amerindians ( $F_{ST} = 2.2\%$   $p = 0.0082$ ) was around five times larger than among Mestizos ( $F_{ST} = 0.44\%$ ;  $p = 0.187$ ). This result indicates that most of the genetic variability of the Mexican population is attributable to their Native American groups. On the other hand, AMOVA of Amerindians vs. Mestizos showed a significant differentiation both between groups ( $F_{CT} = 1.16\%$ ;  $p = 0.0043$ ) and within groups ( $F_{SC} = 2.0\%$ ;  $p = 0.0075$ ).

When Mexican Amerindians were clustered into regions (north, west and south), differentiation among populations within groups ( $F_{SC} = 2.0\%$ ;  $p = 0.0074$ ) and among groups ( $F_{CT} = 0.3\%$   $p = 0.0012$ ) were significant. On the other hand, when genetic structure was assessed between continent clustering populations from America (Mexican, Colombian and Brazilian populations), Europe, and Africa, differentiation among groups increased ( $F_{CT} = 5.82\%$ ;  $p = 0.0231$ ). This result correlates with the significant intercontinental differentiation for D9S1120 by the exclusive presence of 9RA in America.

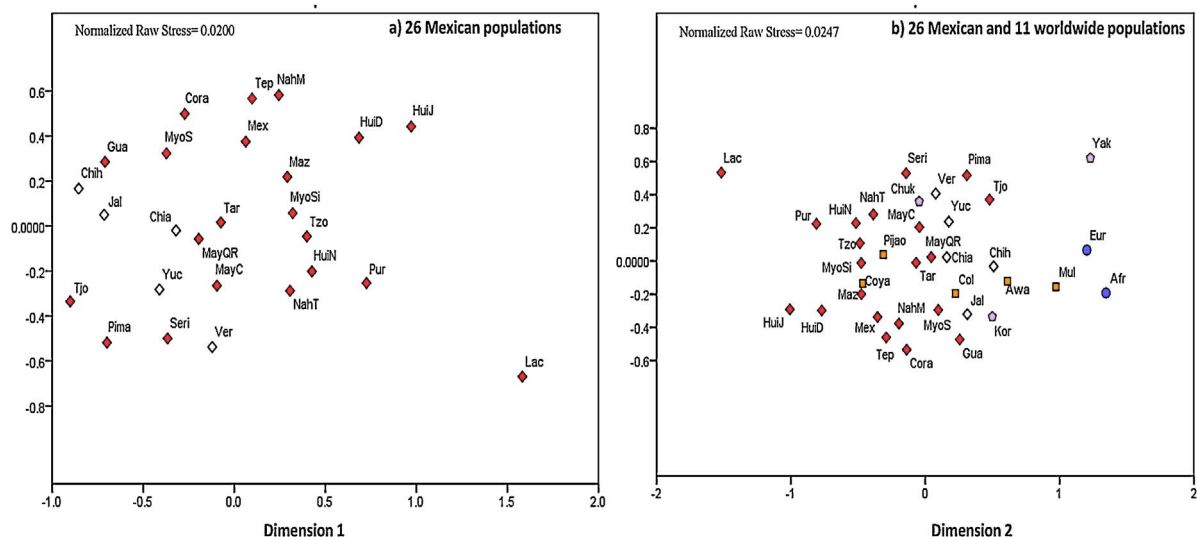


Fig. 2. MDS plot representing  $F_{st}$  genetic distances among Mexican groups (a) and adding worldwide populations (b). Abbreviations can be checked in Fig. 1, and in Material and methods. The Surui group was omitted from the MDS plot due to their extreme peripheral position.

Table 2

Analysis of molecular variance (AMOVA) tests based on D9S1120 in Mexican and worldwide populations.

No. Pop	No. Groups	Population/Cluster	Within populations (%) Fit	Among populations (%) $F_{st}$	Among groups (%) Fct	Among pop within groups (%) Fsc
37	1	World populations <sup>*</sup>	95.7 p = 0.3612	4.3 p = 0.0160	-	-
26	1	Mexican populations (Amerindian + Mestizos)	97.7 p = 0.0365	2.3 p = 0.0085	-	-
21	1	Mexican Amerindian <sup>‡</sup>	97.8 p = 0.3637	2.2 p = 0.0082	-	-
5	1	Mexican Mestizos <sup>†</sup>	99.6 p = 0.3742	0.4 p = 0.0016	-	-
26	2	Mexican Amerindian <sup>‡</sup> vs. Mexican-Mestizos <sup>†</sup>	96.8 p = 0.3650	-	1.16 p = 0.0043	2.0 p = 0.0075
21	3	Mexican Amerindian: North vs. West vs. South <sup>¶</sup>	97.6 p = 0.3637	-	0.30 p = 0.0012	2.0 p = 0.0074
34	3	American populations <sup>£</sup> vs. Europe vs. Africa	91.1 p = 0.3612	-	5.82 p = 0.0231	3.1 p = 0.0123

\* World populations: full population data set.

‡ Mexican Amerindians: MyoS, MyoS<sub>i</sub>, Tep, Gua, Ser, Mex, Cora, HuiD, HuiN, HuiJ, NahM, NahT, Tar, Pur, Maz, Tzo, Tjo, Lac, MayC, MayQR and Pima.

† Mexican Mestizos: Yuc, Chia, Ver, Chih and Jal.

¶ 1) North: Tar, Pima, MyoS, MyoS<sub>i</sub>, Tep, Seri, Gua; 2) West: Mex, Cora, HuiD, HuiN, HuiJ, NahM, NahT, Pur; 3) South: Maz, Tzo, Tjo, Lac, MayC and MayQR.

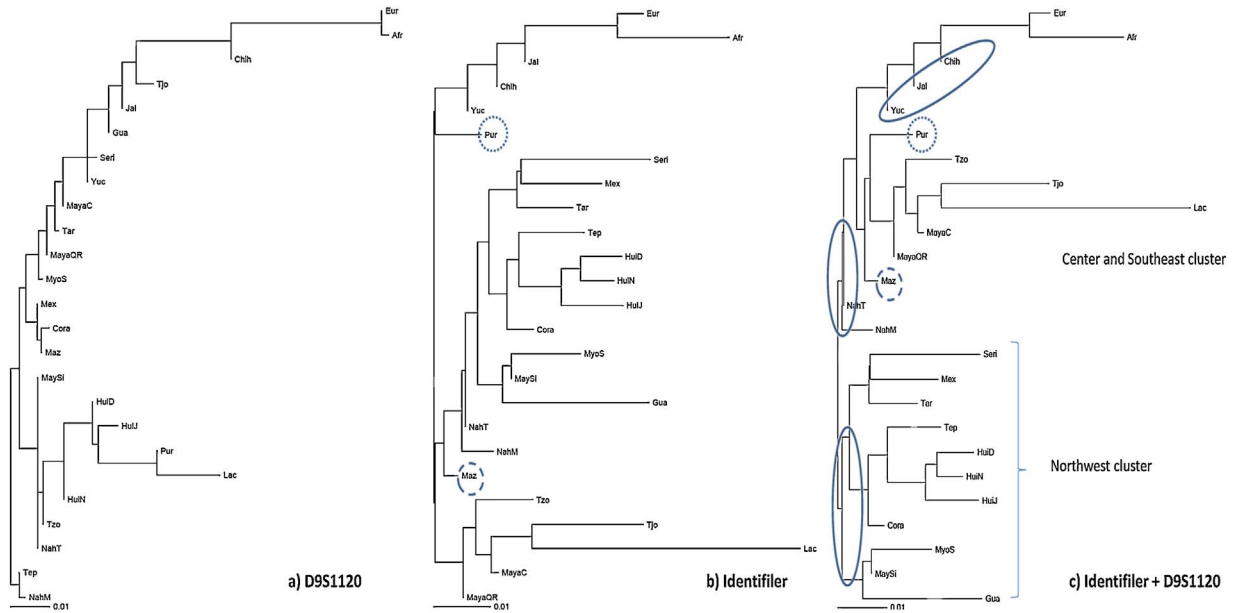
£ American populations: Mexican Amerindians and Mestizos, Colombians, Mulalos, Awa-Kuaiker, Pijao, Coyaima and Surui.

In order to compare the impact of D9S1120 in the analysis of genetic relationships among 26 Mexican groups, we generated different NJ trees with a) D9S1120, b) 15 autosomal STRs (Identifiler), and c) D9S1120 plus Identifiler (Fig. 3abc; Fig. 4ab). As expected, many inconsistencies were observed when an NJ tree was based only on D9S1120 (Fig. 3a). This outlook improved when Identifiler was employed (Fig. 3b and 3a). However, when D9S1120 and Identifiler were employed together, a significant improvement was achieved regarding this interpopulation landscape taking into account some previous genetic and genomic studies that will be discussed below.

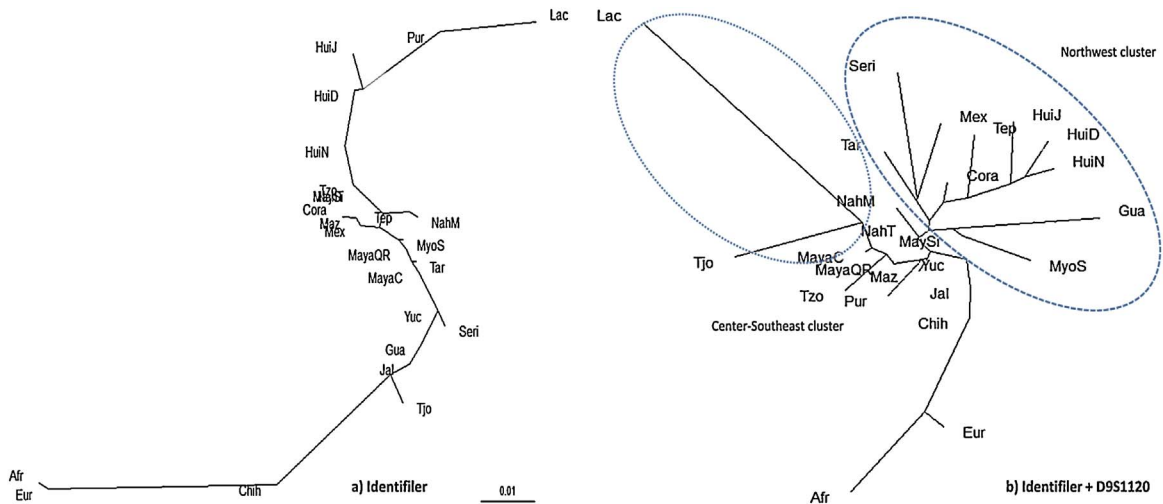
Finally, we carried out a Structure analysis for these 26 Mexican groups including European and African populations as ancestral references based on Identifiler with and without D9S1120 (Fig. 5). Interestingly, Structure-harvester results for both supervised analyses indicate that  $K = 5$  provides the best fit for the explanation of the population structure of this studied population sample. However, in both cases, the same clusters were observed in most isolated groups that are discussed below. Although both analyses offer practically the same position, the main advantage offered by the inclusion of D9S1120 to the 15 STR dataset is that it allowed for differentiating the European and African populations (light pink in  $K = 6$ ; Fig. 5b), which were considered into the same population structure level when only 15 STRs were used for the test (from  $K = 2$  to  $K = 8$ ).

## Discussion

In this study, the STR locus D9S1120 was characterized in 13 Amerindian groups from the north and west of Mexico. Although we report for the first time the allele 12 in Mexican Native groups, the allele 10 previously found in Purépechas was not detected herein



**Fig. 3.** Rooted NJ tree of Mexican population plus African and European populations based on: a) D9S1120, b) Identifier, and c) D9S1120 plus Identifier. For abbreviations see Fig. 1, and Material and methods. Some populations/roots are indicated in circles for discussion purposes (please, see the text).

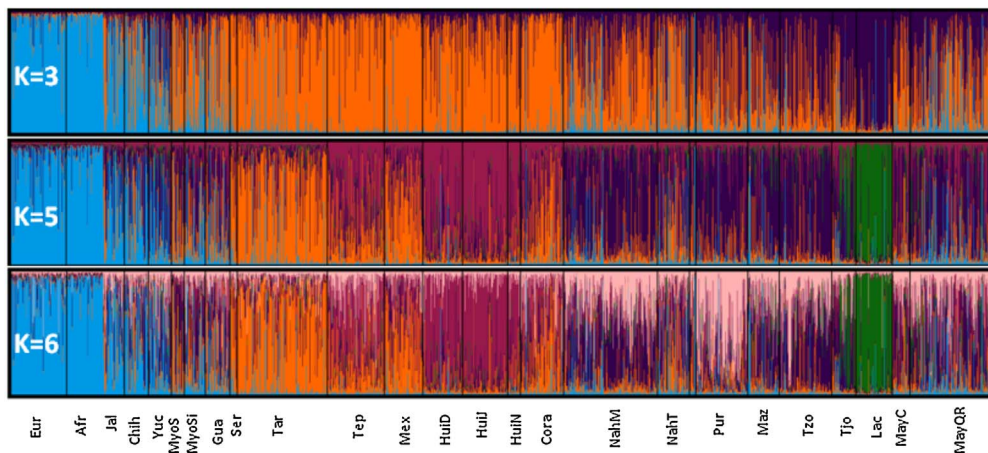


**Fig. 4.** Unrooted NJ tree of Mexican population plus African and European populations based on: a) Identifier, and b) Identifier plus D9S1120. For abbreviations see Fig. 1, and Material and methods. Some populations/roots are indicated in circles for discussion purposes (please, see the text).

(Rangel-Villalobos et al., 2012). The average frequency of 9RA (0.3254) was in line with previous reports (Rangel-Villalobos et al., 2012; Schroeder et al., 2007; Wang et al., 2007); however, differences were observed in the average frequency of 9RA between the north-west (studied herein) and center-southeast Amerindian groups (Rangel-Villalobos et al., 2012) (0.3254 vs. 0.3820;  $p < 0.0001$ ). Firstly, the Pre-Columbian genetic structure probably explains the interpopulation variability regarding the 9RA frequency, which could have been determined by differences in geographic isolation and genetic drift among Mexican Native groups, as was recently described (Rangel-Villalobos et al., 2016). In addition, 9RA frequency has surely been influenced by recent Post-European contact with Native American groups, involving constant acculturation and admixture processes with Mestizo groups geographically adjacent (Martinez-Cortes et al., 2013; Moreno-Estrada et al., 2014; Sandoval et al., 2012). Genotype frequency distribution was in agreement with HWE expectations for all populations, excepting for Huichol from Nayarit and Jalisco states (Table 1). This finding is probably explained by genetic structure among Huichol communities and endogamy, as suggested by the high homozygosity levels previously observed in the Huichol group (Rangel-Villalobos et al., 2013).

Genetic distances (Fig. 2) and pairwise comparisons (Appendix 1) showed a relative genetic homogeneity for D9S1120 among Native American groups, except for some isolated populations such as Lacandones, Tojolabales, Huichol from Jalisco and Nayarit, Mulaloes from Colombia, and particularly Surui from Brazil that was omitted from the MDS plots due to its far peripheral position

## a) Supervised Analysis from IDENTIFILER kit



## b) Supervised Analysis from IDENTIFILER + D9S1120

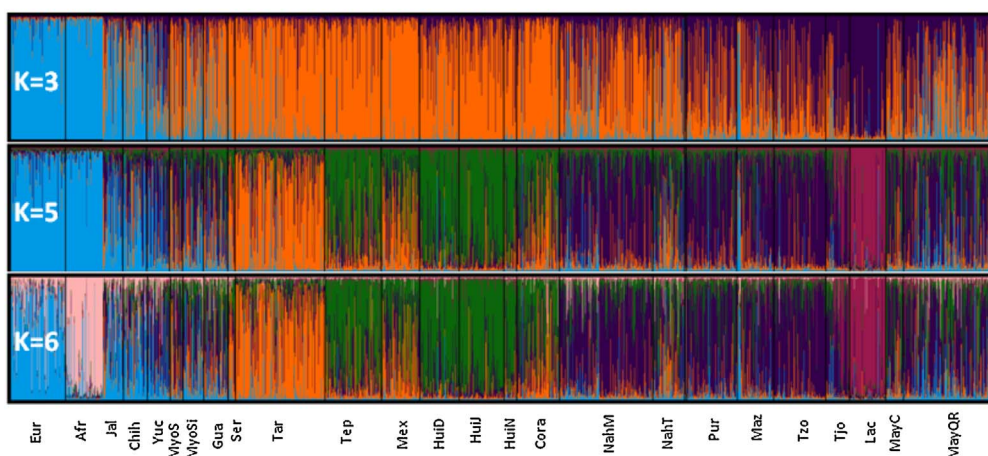


Fig. 5. Structure plots of supervised analysis based on: a) Identifiler, and b) Identifiler plus D9S1120. For abbreviations see Fig. 1, and Material and methods. Coloured figure on line with the description of colours in the Discussion.

(Fig. 2). These findings are in line with previous historical and genetic reports (Schroeder et al., 2007, Wang et al., 2007). Based on D9S1120, Huichol from Nayarit were differentiated from those of Jalisco and Durango (Appendix 1). Interestingly, this could be explained by the bottleneck detected with 15 autosomal STRs in Huichol from Nayarit, which would have promoted greater genetic drift effects (Rangel-Villalobos et al., 2016).

The AMOVA has shown that most of the genetic variability in the Mexican population is attributable to the Native American groups (Table 2). This is compatible with the conclusion obtained for different genetic systems, such as autosomal STRs (Rubí-Castellanos et al., 2009), mitochondrial SNPs (Martínez-Cortés et al., 2013), and genome-wide SNPs claiming that the Mexican population –largely admixed– recapitulates the Native American genetic structure (Moreno-Estrada et al., 2014). On the other hand, when Mexican Amerindians were clustered by geographic criteria, non-consistent groups were formed because significant differentiation of populations within groups was observed. This result is attributable to strong genetic drift effects causing significant population differentiation among Native American groups, as described in previous reports (Moreno-Estrada et al., 2014; Rangel-Villalobos et al., 2016). Interestingly, we have detected a significant intercontinental differentiation for D9S1120, which can mostly be attributed to the exclusive presence of 9RA in America.

After the inclusion of D9S1120 in the Identifiler STR population dataset for the NJ tree construction, the following new features were observed: 1) the position of Chihuahua (the northeast group) was closer to the European reference (Fig. 3bc), which is more consistent with the ancestral gradient of European admixture in Mestizos throughout the Mexican territory. This fact has been described in previous reports with different genetic systems including autosomal STRs (Rubí-Castellanos et al., 2009), Y-STRs (Salazar-Flores et al., 2010), Y-SNPs (Martínez-Cortés et al., 2012) and genome-wide SNPs (Silva-Zolezzi et al., 2009). 2) The Purepecha group changed from the branch closer to Mestizos, toward the branch clustering Native groups from the center and southeast (Fig. 3bc), which is consistent with their Amerindian origin and genetic relationships previously inferred among Mexican Native groups based on autosomal STRs (Rangel-Villalobos et al., 2016) and genome-wide SNPs (Moreno-Estrada et al., 2014). 3)

Similarly, the Mazateca group was closer to the Mayas than to Nahua groups (Fig. 3c), as expected by their geographical proximity (Fig. 1) and by the previous genetic and genomic studies that included this Native group (Moreno-Estrada et al., 2014; Rangel-Villalobos et al., 2016). 4) Main branches of the tree were more consistent with the cited genetic and genomic studies in Mexican Native groups, clustering better the north and west populations regarding those from the center and southeast (Fig. 3c). 5) Similarly, the strongest genetic drift attributed to the more differentiated Mexican Native groups was better represented by the branch length in unrooted NJ trees after the inclusion of D9S1120 (Fig. 4b). This is expected in some Mexican Native groups such as Lacandones, Tojolabales, and those from the north and west (mostly studied herein), in accordance with previous results and conclusions (Rangel-Villalobos et al., 2016). These findings demonstrate that when D9S1120 is included in the Identifiler STR population dataset, the anthropological resolution is improved for ancestry estimation in Mestizo (admixed) populations and for representing genetic relationships in NJ trees among Native American groups.

Regarding the Structure analysis, in both cases well-defined clusters were observed in  $K \geq 5$ : 1) a Tarahumara cluster (Tar or orange, on line) with descendent frequency from north to west population but absent in south and southeast groups; 2) a Huichol component (see Huij or dark pink in Fig. 5a and green in Fig. 5b, on line) that is partially viewed in geographically close populations (Tepehuano, Mexicanero and Cora groups); 3) the observed cluster in the Maya and Nahua groups (Maz and NahM, NahT or dark purple, on line). Moreover, this component was also observed in the Purepecha group. This could be due to the relation with Nahua groups according to historical records (Argueta-Villamar, 1995). Finally, 4) the last cluster from the most isolated Lacandon population (Lac or green in Fig. 5a, and dark pink in Fig. 5b, on line), and partially observed in Tojolobal group. These assumptions based on 15 forensic STR markers plus D9S1120 are in full agreement with the genome-wide record (Moreno-Estrada et al., 2014). Nevertheless, the main difference observed when D9S1120 was included in the Identifiler STR population dataset was the differentiation of the African ancestry from the European one ( $K \geq 6$ ; Fig. 5). Differentiation between the African and European ancestries is a critical step for admixture analysis in both Latin American and Native American populations. Particularly, the presence of the African component in Mexican population is historically well known (Aguirre-Beltrán, 1988; Barquera and Acuña-Alonso, 2012), and it has been detected in different previous genetic studies in both Mexican Mestizos (Rubí-Castellanos et al., 2009; Silva-Zolezzi et al., 2009; Salazar-Flores et al., 2015) and Mexican Native groups (Rangel-Villalobos et al., 2016; Moreno-Estrada et al., 2014; Solé-Llusa et al., 2015).

In conclusion, we showed that D9S1120 contributes to obtaining better anthropological inferences in Native American populations when it is included in STR genotype datasets. Even individually, D9S1120 allows for some general –but congruent– inferences about the population structure in Mexican populations. Therefore, the inclusion of D9S1120 is recommended for population genetic studies with Native American groups.

## Acknowledgments

The authors thank the generous participation of indigenous volunteers to this study. We thank CONACYT for the financial support to HRV for this project (CONACYT grant 129693), and for the scholar fellowship to JAAV. Authors declare no conflicts of interest.

## Appendix A. Genetic distances $F_{ST}$ (below the diagonal) and $F_{ST}$ p-value (above the diagonal) based on D9S1120 between 26 Mexican and 11 worldwide populations

		Eur	Afr	MayS	MaySi	Tep	Gua	Seri	Mex
1	Eur	—	0.2182	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	Afr	0.0016	—	0.0000	0.0000	0.0000	0.0273	0.0000	0.0000
3	MyoS	0.0506	0.0543	—	0.1727	0.1455	0.4727	0.1182	0.3727
4	MyoSi	0.0997	0.1127	0.0093	—	0.1909	0.0818	0.1636	0.4364
5	Tep	0.0831	0.0825	0.0050	0.0057	—	0.1091	0.0000	0.2909
6	Gua	0.0571	0.0615	-0.0016	0.0301	0.0190	—	0.3546	0.1636
7	Ser	0.0888	0.1112	0.0207	0.0179	0.0389	0.0064	—	0.0182
8	Mex	0.0932	0.0993	0.0007	-0.0008	0.0009	0.0099	0.0221	—
9	Cora	0.0907	0.0950	0.0055	0.0130	0.0085	-0.0066	0.0115	0.0019
10	HuiD	0.1322	0.1372	0.0300	0.0035	0.0053	0.0505	0.0492	0.0089
11	HuiN	0.1025	0.1221	0.0278	-0.0043	0.0150	0.0473	0.0185	0.0142
12	HuiJ	0.1570	0.1634	0.0478	0.0103	0.0159	0.0755	0.0677	0.0201
13	NahM	0.0697	0.0701	0.0142	0.0111	0.0020	0.0354	0.0472	0.0140
14	NahT	0.0884	0.1054	0.0234	-0.0007	0.0173	0.0419	0.0178	0.0161
15	Tar	0.0561	0.0692	0.0096	0.0087	0.0143	0.0171	0.0204	0.0111
16	Pur	0.1384	0.1560	0.0452	0.0041	0.0295	0.0623	0.0250	0.0240
17	Maz	0.1041	0.1128	0.0068	-0.0076	-0.0016	0.0198	0.0185	-0.0056
18	Tzo	0.0969	0.1087	0.0152	-0.0051	0.0104	0.0391	0.0211	0.0077
19	Tjo	0.0460	0.0611	0.0242	0.0469	0.0493	0.0003	0.0018	0.0427
20	Lac	0.2261	0.2524	0.1089	0.0379	0.0845	0.1423	0.0749	0.0707
21	MayC	0.0542	0.0742	0.0065	0.0024	0.0194	0.0192	0.0015	0.0130



22	<b>MayQR</b>	0.0480	0.0577	0.0014	0.0069	0.0111	0.0115	0.0109	0.0100
23	<b>Pima</b>	0.0539	0.0793	0.0230	0.0362	0.0555	0.0103	-0.0095	0.0403
24	<b>Jal</b>	0.0481	0.0571	0.0006	0.0278	0.0254	-0.0154	0.0059	0.0156
25	<b>Chih</b>	0.0097	0.0151	0.0046	0.0361	0.0260	0.0090	0.0334	0.0304
26	<b>Ver</b>	0.0479	0.0698	0.0230	0.0176	0.0343	0.0356	0.0108	0.0335
27	<b>Chia</b>	0.0348	0.0458	-0.0013	0.0093	0.0115	0.0078	0.0110	0.0115
28	<b>Yuc</b>	0.0456	0.0637	0.0111	0.0175	0.0279	0.0066	-0.0032	0.0220
29	<b>Col</b>	0.0222	0.0184	-0.0146	0.0114	-0.0054	0.0097	0.0409	0.0054
30	<b>Mul</b>	-0.0011	0.0011	0.0324	0.0740	0.0571	0.0468	0.0822	0.0691
31	<b>Coya</b>	0.1025	0.1113	0.0088	-0.0057	0.0015	0.0197	0.0160	-0.0027
32	<b>Pijao</b>	0.0857	0.1014	0.0065	-0.0093	0.0043	0.0177	0.0055	-0.0020
33	<b>Awa</b>	0.0120	0.0125	0.0155	0.0403	0.0243	0.0390	0.0637	0.0405
34	<b>Surui</b>	0.4933	0.5338	0.4059	0.2957	0.3211	0.5591	0.4217	0.3144
35	<b>Yak</b>	0.0102	0.0319	0.0787	0.1149	0.1177	0.0837	0.0797	0.1228
36	<b>Chuk</b>	0.0490	0.0709	0.0222	0.0102	0.0272	0.0499	0.0246	0.0309
37	<b>Kor</b>	0.0297	0.0314	-0.0015	0.0422	0.0295	-0.0202	0.0177	0.0247

	<b>Cora</b>	<b>HuiD</b>	<b>HuiN</b>	<b>HuiJ</b>	<b>NahM</b>	<b>NahT</b>	<b>Tar</b>	<b>Pur</b>	<b>Maz</b>	<b>Tzo</b>
1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.1636	0.0091	0.0364	0.0000	0.0364	0.0000	0.0546	0.0000	0.2455	0.0636
4	0.0727	0.2364	0.7182	0.1091	0.0636	0.4455	0.1455	0.2636	0.8636	0.9182
5	0.0455	0.1000	0.0546	0.0182	0.1818	0.0000	0.0000	0.0000	0.4727	0.0182
6	0.5636	0.0182	0.1091	0.0000	0.0273	0.0273	0.1000	0.0000	0.1818	0.0182
7	0.1091	0.0000	0.1182	0.0000	0.0091	0.0364	0.0455	0.0364	0.0909	0.0182
8	0.2546	0.0636	0.1818	0.0000	0.0182	0.0182	0.0091	0.0000	0.8000	0.0273
9	—	0.0000	0.0727	0.0000	0.0000	0.0000	0.0000	0.0000	0.2636	0.0000
10	0.0213	—	0.3182	0.7455	0.0000	0.0273	0.0000	0.0273	0.4727	0.0818
11	0.0253	0.0055	—	0.1546	0.2000	0.9727	0.2273	0.8273	0.6091	0.6364
12	0.0380	-0.0033	0.0102	—	0.0000	0.0000	0.0000	0.0091	0.1818	0.0182
13	0.0233	0.0109	0.0102	0.0200	—	0.0091	0.0000	0.0000	0.1818	0.0091
14	0.0262	0.0135	-0.0108	0.0203	0.0124	—	0.0364	0.3364	0.2455	0.5182
15	0.0182	0.0274	0.0079	0.0396	0.0145	0.0084	—	0.0000	0.1182	0.0000
16	0.0336	0.0118	-0.0072	0.0142	0.0298	0.0019	0.0283	—	0.1909	0.0909
17	0.0023	-0.0031	-0.0004	0.0051	0.0076	0.0043	0.0114	0.0066	—	0.6000
18	0.0204	0.0074	-0.0044	0.0138	0.0113	-0.0009	0.0147	0.0045	-0.0023	—
19	0.0241	0.0778	0.0465	0.1025	0.0514	0.0401	0.0275	0.0637	0.0451	0.0472
20	0.0922	0.0470	0.0262	0.0397	0.0856	0.0386	0.0799	0.0147	0.0474	0.0376
21	0.0189	0.0315	-0.0004	0.0457	0.0170	-0.0011	0.0012	0.0168	0.0089	0.0031
22	0.0126	0.0270	0.0091	0.0417	0.0110	0.0074	0.0057	0.0264	0.0079	0.0078
23	0.0301	0.0797	0.0363	0.1026	0.0578	0.0303	0.0230	0.0522	0.0419	0.0376
24	0.0059	0.0549	0.0417	0.0779	0.0364	0.0355	0.0157	0.0581	0.0236	0.0343
25	0.0293	0.0620	0.0416	0.0839	0.0224	0.0340	0.0136	0.0703	0.0366	0.0374
26	0.0365	0.0471	0.0045	0.0618	0.0237	0.0033	0.0099	0.0256	0.0258	0.0140
27	0.0136	0.0325	0.0108	0.0495	0.0096	0.0081	0.0019	0.0321	0.0110	0.0102
28	0.0163	0.0465	0.0134	0.0653	0.0270	0.0110	0.0068	0.0318	0.0200	0.0186
29	0.0146	0.0256	0.0248	0.0435	-0.0069	0.0190	0.0068	0.0512	0.0092	0.0124
30	0.0737	0.1029	0.0792	0.1258	0.0450	0.0670	0.0389	0.1170	0.0785	0.0740
31	0.0036	0.0003	-0.0001	0.0090	0.0099	0.0043	0.0122	0.0069	-0.0099	-0.0010
32	0.0057	0.0051	-0.0097	0.0146	0.0083	-0.0059	0.0000	0.0012	-0.0085	-0.0051
33	0.0489	0.0577	0.0405	0.0758	0.0110	0.0342	0.0199	0.0744	0.0425	0.0380
34	0.3485	0.2716	0.3343	0.2451	0.3163	0.2923	0.3098	0.2364	0.3245	0.2726
35	0.1165	0.1626	0.1066	0.1891	0.0977	0.0898	0.0726	0.1413	0.1277	0.1070
36	0.0438	0.0363	-0.0039	0.0475	0.0118	-0.0050	0.0064	0.0195	0.0212	0.0053
37	0.0110	0.0703	0.0604	0.0975	0.0395	0.0504	0.0250	0.0804	0.0362	0.0469

		Tjo	Lac	MyaC	MyaQR	Pima	Jal	Chi	Ver	Chia
1	<b>Eur</b>	0.0000	0.0000	0.0000	0.0000	0.0091	0.0000	0.0273	0.0000	0.0000
2	<b>Afr</b>	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0455	0.0000	0.0000
3	<b>MyoS</b>	0.0091	0.0000	0.1909	0.3182	0.0909	0.4273	0.2364	0.0364	0.5727
4	<b>MyoSi</b>	0.0000	0.0000	0.2636	0.1000	0.0455	0.0182	0.0091	0.0727	0.1364
5	<b>Tep</b>	0.0000	0.0000	0.0000	0.0091	0.0000	0.0000	0.0000	0.0000	0.0546
6	<b>Gua</b>	0.3182	0.0000	0.1000	0.1091	0.3273	0.9000	0.1909	0.0727	0.2182
7	<b>Ser</b>	0.2818	0.0000	0.3364	0.1091	0.6909	0.3273	0.0000	0.2091	0.1455
8	<b>Mex</b>	0.0000	0.0000	0.0546	0.0182	0.0000	0.0909	0.0091	0.0091	0.0727
9	<b>Cora</b>	0.0000	0.0000	0.0182	0.0000	0.0273	0.1182	0.0091	0.0000	0.0273
10	<b>HuiD</b>	0.0000	0.0000	0.0091	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
11	<b>HuiN</b>	0.0000	0.0364	0.4636	0.1818	0.0636	0.0182	0.0182	0.3455	0.1546
12	<b>HuiJ</b>	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
13	<b>NahM</b>	0.0000	0.0000	0.0273	0.0000	0.0000	0.0000	0.0000	0.0273	0.0636
14	<b>NahT</b>	0.0000	0.0000	0.4182	0.0818	0.0000	0.0000	0.0091	0.3273	0.1091
15	<b>Tar</b>	0.0000	0.0000	0.4182	0.0182	0.0364	0.0091	0.0000	0.1091	0.3636
16	<b>Pur</b>	0.0000	0.0000	0.0364	0.0000	0.0000	0.0000	0.0000	0.0091	0.0000
17	<b>Maz</b>	0.0091	0.0000	0.1000	0.0818	0.0546	0.0455	0.0000	0.0364	0.0546
18	<b>Tzo</b>	0.0000	0.0000	0.2636	0.0273	0.0273	0.0000	0.0000	0.0546	0.0546
19	<b>Tjo</b>	—	0.0000	0.0091	0.0182	0.6818	0.3000	0.0727	0.0636	0.0818
20	<b>Lac</b>	0.1390	—	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
21	<b>MayC</b>	0.0136	0.0673	—	0.7273	0.3000	0.1000	0.1000	0.7909	0.8182
22	<b>MayQR</b>	0.0155	0.0813	-0.0043	—	0.0909	0.1182	0.0818	0.2273	0.9909
23	<b>Pima</b>	-0.0074	0.1159	0.0011	0.0121	—	0.2818	0.0727	0.2818	0.2546
24	<b>Jal</b>	0.0021	0.1291	0.0111	0.0084	0.0036	—	0.1909	0.0364	0.1818
25	<b>Chih</b>	0.0136	0.1499	0.0104	0.0062	0.0182	0.0065	—	0.0727	0.6182
26	<b>Ver</b>	0.0151	0.0810	-0.0082	0.0026	0.0046	0.0244	0.0136	—	0.4909
27	<b>Chia</b>	0.0101	0.0958	-0.0068	-0.0055	0.0075	0.0042	-0.0022	-0.0016	—
28	<b>Yuc</b>	-0.0017	0.0937	-0.0065	0.0001	-0.0070	0.0030	0.0070	-0.0039	-0.0041
29	<b>Col</b>	0.0309	0.1301	0.0055	-0.0049	0.0366	0.0084	-0.0119	0.0149	-0.0100
30	<b>Mul</b>	0.0480	0.2056	0.0411	0.0341	0.0552	0.0397	0.0011	0.0393	0.0220
31	<b>Coya</b>	0.0416	0.0475	0.0083	0.0081	0.0385	0.0231	0.0365	0.0239	0.0110
32	<b>Pijao</b>	0.0294	0.0427	-0.0055	0.0000	0.0209	0.0163	0.0253	0.0063	0.0010
33	<b>Awa</b>	0.0423	0.1565	0.0199	0.0130	0.0493	0.0332	-0.0034	0.0186	0.0044
34	<b>Surui</b>	0.4555	0.1740	0.3772	0.3260	0.4869	0.4275	0.4510	0.3934	0.3913
35	<b>Yak</b>	0.0381	0.2315	0.0529	0.0576	0.0359	0.0636	0.0280	0.0360	0.0450
36	<b>Chuk</b>	0.0322	0.0723	-0.0093	0.0009	0.0196	0.0351	0.0139	-0.0135	-0.0027
37	<b>Kor</b>	-0.0008	0.1664	0.0198	0.0112	0.0080	-0.0118	-0.0023	0.0314	0.0050

	Yuc	Col	Mul	Coy	Pijao	Awa	Surui	Yak	Chuk	Kor
1	0.0000	0.0455	0.4182	0.0000	0.0000	0.0455	0.0000	0.1000	0.0000	0.0182
2	0.0000	0.1727	0.3273	0.0000	0.0000	0.1182	0.0000	0.0000	0.0000	0.0364
3	0.0546	0.7091	0.0091	0.1727	0.1909	0.0727	0.0000	0.0000	0.0546	0.3000
4	0.0636	0.2546	0.0000	0.8455	0.9364	0.0273	0.0000	0.0000	0.1636	0.0091
5	0.0000	0.5636	0.0000	0.2818	0.1818	0.0091	0.0000	0.0000	0.0000	0.0091
6	0.1909	0.3273	0.0182	0.1364	0.1000	0.0273	0.0000	0.0091	0.0182	0.8182
7	0.5364	0.0455	0.0000	0.0818	0.2000	0.0000	0.0000	0.0091	0.0818	0.1091
8	0.0000	0.2818	0.0000	0.6636	0.4273	0.0000	0.0000	0.0000	0.0091	0.0455
9	0.0364	0.1273	0.0000	0.2000	0.1727	0.0000	0.0000	0.0000	0.0000	0.2000
10	0.0000	0.1364	0.0000	0.3636	0.2273	0.0091	0.0000	0.0000	0.0091	0.0091
11	0.1455	0.2818	0.0000	0.4091	0.9000	0.0273	0.0000	0.0000	0.6636	0.0182
12	0.0000	0.0273	0.0000	0.0818	0.0909	0.0000	0.0000	0.0000	0.0182	0.0000
13	0.0000	0.4909	0.0000	0.0364	0.1636	0.1818	0.0000	0.0000	0.1455	0.0091
14	0.0455	0.1455	0.0000	0.2818	0.7909	0.0091	0.0000	0.0000	0.6182	0.0000
15	0.1273	0.4000	0.0000	0.0273	0.4909	0.0546	0.0000	0.0000	0.2727	0.0636
16	0.0091	0.0546	0.0000	0.0455	0.2818	0.0000	0.0000	0.0000	0.0818	0.0000
17	0.0364	0.3727	0.0000	1.0000	0.8818	0.0091	0.0000	0.0000	0.0727	0.0273
18	0.0000	0.2364	0.0000	0.3727	0.7000	0.0182	0.0000	0.0000	0.2182	0.0000
19	0.4182	0.1000	0.0000	0.0000	0.0182	0.0091	0.0000	0.0000	0.0091	0.3182
20	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
21	0.7546	0.2182	0.0000	0.0909	0.5455	0.0182	0.0000	0.0000	0.6273	0.0000

22	0.4273	0.5455	0.0000	0.0546	0.4727	0.0727	0.0000	0.0000	0.4091	0.1000
23	0.7000	0.0727	0.0000	0.0182	0.1091	0.0000	0.0000	0.0000	0.1636	0.2000
24	0.3273	0.3091	0.0000	0.0364	0.0727	0.0091	0.0000	0.0000	0.0364	0.8727
25	0.1546	0.7091	0.3091	0.0000	0.0000	0.5546	0.0000	0.0455	0.1182	0.3546
26	0.6636	0.1818	0.0000	0.0364	0.2182	0.0636	0.0000	0.0364	0.9091	0.0273
27	0.5818	0.7364	0.0455	0.0818	0.2909	0.2000	0.0000	0.0000	0.4455	0.2182
28	—	0.2364	0.0000	0.0273	0.2364	0.0091	0.0000	0.0182	0.3455	0.1273
29	0.0128	—	0.2818	0.2546	0.1727	0.8455	0.0000	0.0364	0.2455	0.3273
30	0.0382	0.0010	—	0.0000	0.0000	0.4364	0.0000	0.0455	0.0091	0.0091
31	0.0181	0.0117	0.0782	—	0.7909	0.0000	0.0000	0.0000	0.0455	0.0273
32	0.0042	0.0094	0.0640	-0.0075	—	0.0091	0.0000	0.0000	0.2182	0.0000
33	0.0244	-0.0187	-0.0030	0.0435	0.0331	—	0.0000	0.0091	0.1546	0.0273
34	0.3962	0.5603	0.5150	0.3021	0.3374	0.4945	—	0.0000	0.0000	0.0000
35	0.0405	0.0592	0.0230	0.1222	0.0974	0.0395	0.6067	—	0.0182	0.0000
36	0.0028	0.0053	0.0344	0.0203	0.0024	0.0080	0.4375	0.0440	—	0.0182
37	0.0078	-0.0008	0.0247	0.0355	0.0309	0.0234	0.5620	0.0507	0.0438	—

## References

- Alves, C., Gusmão, L., López-Parra, A.M., Soledad Mesa, M., Amorim, A., Arroyo-Pardo, 2005. STR allelic frequencies for an African population sample (Equatorial Guinea) using AmpFISTR Identifier and Powerplex 16 kits. *Forensic Sci. Int.* 148, 239–242.
- Argueta-Villamar, A., 1995. Purépechas, en: *Etnografía contemporánea de los pueblos indígenas de México (Región Centro)*. Instituto Nacional Indigenista (INI), México, México, DF.
- Barquera, R., Acuña-Alonso, V., 2012. The African colonial migration into Mexico: history and biological consequences. In: Crawford, M.H., Campbell, B.C. (Eds.), *Causes and Consequences of Human Migration, An Evolutionary Perspective*. Cambridge University Press, Cambridge, pp. 201–223.
- Coudray, C., Calderon, R., Guitard, E., Ambrosio, B., González-Martín, A., Dugoujon, J.M., 2007. Allele frequencies of 15 tetrameric short tandem repeats (STRs) in Andalusians from Huelva (Spain). *Forensic Sci. Int.* 168, 1–4.
- Earl, D., von Holdt, B., 2012. Structure Harvester: a website and program for visualizing Structure output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361.
- Excoffier, L., Lischer, H., 2010. Arlequin suite ver3.5: a new series of programs to perform population genetics analyses under Linux and Windows. 3.5 ed. *Mol. Ecol. Resour.* 10, 564–567. <http://dx.doi.org/10.1111/j.1755-0998.2010.02847.x>.
- Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164, 1567–1587.
- Falush, D., Stephens, M., Pritchard, J.K., 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes* 7, 574–578.
- IBM Corp, 2010. IBM SPSS Statistics for Windows. v19.0. Armonk, NY.
- Lewis, P.O., Zaykin, D., 2001. GDA (Genetic Data Analysis): Computer Program for the Analysis of Allelic Data. v1.1. University of Connecticut, Storrs. <http://phylogeny.uconn.edu/software/>.
- Lópes, V., Serra, A., Gamero, J., Sampaio, L., Balsa, F., Oliveira, C., Batista, L., Corte-Real, F., Vieira, D.N., Vide, M.C., Anjos, M.J., Carvalho, M., 2009. Allelic frequency distribution of 17 STRs from Identifier and PowerPlex-16 in Central Portugal area and the Azores archipelago. *Forensic Sci. Int. Genet.* 4, 1–7.
- Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., Mayrose, I., 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* 15, 1179–1191.
- Martínez-Cortés, G., Nuño-Arana, I., Rubí-Castellanos, R., Vilchis-Dorantes, G., Luna-Vázquez, A., Coral-Vázquez, R.M., Canto-Cetina, T., Salazar-Flores, J., Muñoz-Valle, J.F., Sandoval-Mendoza, K., Lopez, Z., Gamero-Lucas, J.J., Rangel-Villalobos, H., 2010. Origin and genetic differentiation of three Native Mexican groups (Purépechas, Triquis and Mayas): contribution of CODIS-STRs to the history of human populations of Mesoamerica. *Ann. Hum. Biol.* 37, 801–819.
- Martínez-Cortés, G., Salazar-Flores, J., Fernandez-Rodríguez, I.G., Rubí-Castellanos, R., Rodríguez-Loya, C., Velarde-Felix, J.S., Muñoz-Valle, J.F., Parra-Rojas, I., Rangel-Villalobos, H., 2012. Admixture and population structure in Mexican-Mestizos based on paternal lineages. *J. Hum. Genet.* 57, 568–574.
- Martínez-Cortés, G., Salazar-Flores, J., Haro-Guerrero, J., Rubí-Castellanos, R., Velarde-Felix, J.S., Muñoz-Valle, J.F., Lopez-Casamichana, M., Carrillo-Tapia, E., Canseco-Avila, I.M., Bravi, C.M., Lopez-Armenta, M., Rangel-Villalobos, H., 2013. Maternal admixture and population structure in Mexican-Mestizos based on mtDNA haplogroups. *Am. J. Phys. Anthropol.* 151, 526–537.
- Miller, S.A., Dykes, D.D., Polesky, H.F., 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16, 1215.
- Moreno-Estrada, A., Gignoux, C.R., Fernández-López, J.C., Zakharia, F., Sikora, M., Contreras, A.V., Acuña-Alonso, V., Sandoval, K., Eng, C., Romero-Hidalgo, S., Ortiz-Tello, P., Robles, V., Kenny, E.E., Nuño-Arana, I., Barquera-Lozano, R., Macín-Pérez, G., Granados-Arriola, J., Huntsman, S., Galanter, J.M., Via, M., Ford, J.G., Chapela, R., Rodríguez-Cintrón, W., Rodríguez-Santana, J.R., Romieu, I., Sienna-Monge, J.J., Navarro, B.D.R., London, S.J., Ruiz-Linares, A., Garcia-Herrera, R., Estrada, K., Hidalgo-Miranda, A., Jimenez-Sanchez, G., Carnevale, A., Soberón, X., Canizales-Quinteros, S., Rangel-Villalobos, H., Silva-Zolezzi, I., Burchard, E.G., Bustamante, C.D., 2014. The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science* 344, 1280–1285.
- Page, R.D.M., 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. *CABIOS* 12, 357–358.
- Phillips, C., Rodríguez, A., Mosquera-Miguel, A., Fondevila, M., Porras-Hurtado, L., Rondon, F., Salas, A., Carracedo, A., Lareu, M.V., 2008. D9S1120, a simple STR with a common Native American-specific allele: forensic optimization, locus characterization and allele frequency studies. *Forensic Sci. Int. Genet.* 3, 7–13.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Rangel-Villalobos, H., Martínez-Sevilla, V.M., Martínez-Cortés, G., Aguilar-Velázquez, J.A., Sosa-Macias, M., Rubí-Castellanos, R., González-Martín, A., 2016. Importance of the geographic barriers to promote gene drift and avoid pre- and post-Columbian gene flow in Mexican native groups: Evidence from forensic STR Loci. *Am. J. Phys. Anthropol.* 160, 298–316.
- Rangel-Villalobos, H., Martínez-Sevilla, V.M., Salazar-Flores, J., Martínez-Cortés, G., Muñoz-Valle, J.F., Galaviz-Hernández, C., Lázalde-Ramos, B.P., Sosa-Macias, M., 2013. Forensic parameters for 15 STRs in eight Amerindian populations from the north and west of Mexico. *Forensic Sci. Int. Genet.* 7, 62–65.
- Rangel-Villalobos, H., Sanchez-Gutierrez, V.M., Botello-Ruiz, M., Salazar-Flores, J., Martínez-Cortés, G., Muñoz-Valle, J.F., Phillips, C., 2012. Evaluation of forensic and anthropological potential of D9S1120 in Mestizos and Amerindian populations from Mexico. *Croat. Med. J.* 53, 423–431.
- Rosenberg, N.A., 2006. Standardized subsets of the HGDP-CEPH human genome diversity cell line panel, accounting for atypical and duplicated samples and pairs of close relatives. *Ann. Hum. Genet.* 70, 841–847.
- Rubí-Castellanos, R., Martínez-Cortés, G., Muñoz-Valle, J.F., Gonzalez-Martín, A., Cerda-Flores, R.M., Anaya-Palafox, M., Rangel-Villalobos, H., 2009. Pre-Hispanic Mesoamerican demography approximates the present-day ancestry of Mestizos throughout the territory of Mexico. *Am. J. Phys. Anthropol.* 139, 284–294.
- Salazar-Flores, J., Dondego-Aldape, R., Rubí-Castellanos, R., Anaya-Palafox, M., Nuño-Arana, I., Canseco-Avila, L.M., Flores-Flores, G., Morales-Vallejo, M.E., Barojas-

- Pérez, N., Muñoz-Valle, J.F., Campos-Gutiérrez, R., Rangel-Villalobos, H., 2010. Population structure and paternal admixture landscape on present-day Mexican-Mestizos revealed by Y-STR haplotypes. *Am. J. Hum. Biol.* 22, 401–409.
- Salazar-Flores, J., Zuñiga-Chiquette, F., Rubí-Castellanos, R., Alvarez-Miranda, J.L., Zetina-Hernandez, A., Martinez-Sevilla, V.M., Gonzalez-Andrade, F., Corach, D., Vullo, C., Alvarez, J.C., Lorente, J.A., Sanchez-Diz, P., Herrera, R.J., Cerda-Flores, R.M., Muñoz-Valle, J.F., Rangel-Villalobos, H., 2015. Admixture and genetic relationships of Mexican Mestizos regarding Latin American and Caribbean populations based on 13 CODIS-STRs. *HOMO - J. Comp. Hum. Biol.* 66, 44–59.
- Sandoval, K., Moreno-Estrada, A., Mendizabal, I., Underhill, P.A., Lopez-Valenzuela, M., Peñaloza-Espinosa, R., Lopez-Lopez, M., Buentello-Malo, I., Avelino, H., Calafell, F., Comas, D., 2012. Y-chromosome diversity in Native Mexicans reveals continental transition of genetic structure in the Americas. *Am. J. Phys. Anthropol.* 148, 395–405.
- Schroeder, K.B., Jakobsson, M., Crawford, M.H., Schurr, T.G., Boca, S.M., Conrad, D.F., Tito, R.Y., Osipova, L.P., Tarskaia, L.A., Zhadanov, S.I., Wall, J.D., Pritchard, J.K., Malhi, R.S., Smith, D.G., Rosenberg, N.A., 2009. Haplotypic background of a private allele at high frequency in the Americas. *Mol. Biol. Evol.* 26, 995–1016.
- Schroeder, K.B., Schurr, T.G., Long, J.C., Rosenberg, N.A., Crawford, M.H., Tarskaia, L.A., Osipova, L.P., Zhadanov, S.I., Smith, D.G., 2007. A private allele ubiquitous in the Americas. *Biol. Lett.* 3, 218–223.
- Silva-Zolezzi, I., Hidalgo-Miranda, A., Estrada-Gil, J., Fernández-López, J.C., Uribe-Figueroa, L., Contreras, A., Balam-ortiz, E., Del Bosque-Plata, L., Lara, C., Velázquez-Fernández, D., Goya, R., Hernández-Lemus, E., Dávila, C., Barrientos, E., March, S., Jiménez-Sánchez, G., 2009. Analysis of genomic diversity in Mexican Mestizo populations to develop genomic medicine in Mexico. *Proc. Natl Acad. Sci. USA* 106, 8611–8616.
- Solé-Llussà, A., Gorostiza, A., Rubí-Castellanos, R., Acunha-Alonzo, V., Gené, M., Rangel-Villalobos, H., González-Martín, A., 2015. Exploring the African genetic influence in the first settlement founded by African slaves in America. *Anthropol. Sci.* 123, 121–128.
- Tereba, A., 1999. Tools for analysis of population statistics. *Profiles in DNA* 2, 14–16.
- Wang, S., Lewis, C.M., Jakobsson, C.M., Ramachandran, S., Ray, N., Bedoya, G., Rojas, W., Parra, M.V., Molina, J.A., Gallo, C., Mazzotti, G., Poletti, G., Hill, K., Hurtado, A.M., Labuda, D., Klitz, W., Barrantes, R., Bortolini, M.C., Salzano, F.M., Petzl-Erler, M.L., Tsuneto, L.T., Llop, E., Rothhammer, F., Excoffier, L., Feldman, M.W., Rosenberg, N.A., Ruiz-Linares, A., 2007. Genetic variation and population structure in Native Americans. *PLoS Genet.* 3, 185.
- Zhivotovskiy, L.A., Rosenberg, N.A., Feldman, M.W., 2003. Features of evolution and expansion of modern humans, inferred from genomewide microsatellite markers. *Am. J. Hum. Genet.* 72, 1171–1186.