AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 160:298–316 (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

Héctor Rangel-Villalobos,[1]* Víctor Manuel Martínez-Sevilla,[1] Gabriela Martínez-Cortés,[1] José Alonso Aguilar-Velázquez,[1] Martha Sosa-Macías,[3] Rodrigo Rubi-Castellanos,[4] and Antonio González-Martín[2]*

[1]*Instituto De Investigación En Genética Molecular, Centro Universitario De La Ciénega (CUCI-UdeG), Ocotlán, Jalisco, México*
[2]*Departamento De Zoología Y Antropología Física, Facultad De Biología, Universidad Complutense De Madrid, Madrid, Spain*
[3]*Unidad Durango (CIIDIR-IPN), Centro Interdisciplinario De Investigación Para El Desarrollo Integral Regional Del Instituto Politécnico Nacional, Durango, México*
[4]*Laboratorio De Genética, Centro De Investigaciones Regionales Hideyo Noguchi, Universidad Autónoma De Yucatán (UADY), Mérida, Yucatán*

KEY WORDS STRs; Mexico; Amerindians; gene drift; admixture; structure

ABSTRACT

Objective: To analyze the origin, structure, relationships, and recent admixture in Mexican Native groups based on 15 STRs commonly used in human identification.

Methods: We analyzed 39 Mexican Native population samples using STR databases based on the AmpF*l*STR® Identifiler kit ($n = 3{,}135$), including Mexican-Mestizos (admixed), European and African populations, as reference.

Results: Based upon effective population size (Ne) differences, Native groups were clustered into three regions: i) *Center-Southeast groups*, characterized by larger Ne, migration rate (Nm), genetic diversity (He), and relative homogeneity principally in the Yucatan Peninsula; ii) *Isolated southern groups* from Chiapas and Oaxaca, characterized by lower Ne, Nm, and He (i.e. higher isolation and genetic differentiation); iii) *North-Northwest groups*, which are similar to the previous group but are characterized by generating the widest gene flow barrier in the Pre-Hispanic Mexican territory, and currently by elevated admixture in some northern Native groups. Despite the relative congruence between genetic relationships with cultural, linguistic, geographic criteria, these factors do not explain the present-day population structure of Native groups, excepting in those linguistically related to the Mayan that show higher homogeneity. The Isolation by distance model was demonstrated at long distances (>1,500 km), whereas geographic isolation stands as a determining factor to avoid both non-indigenous admixture and bottleneck processes.

Conclusions: Different dynamics of gene flow and drift were observed among Mexican Native groups, highlighting the geographic barriers (mountains, canyons and jungle regions) as the main factor differentiating Pre-Hispanic populations, and eventually helping to avoid Post-European contact admixture and population bottleneck. Am J Phys Anthropol 160:298–316, 2016.

The Pre-Hispanic territory constituting the present-day Mexico was inhabited by Native American groups including three major cultural areas: Aridoamerica (Northeast and Baja California Peninsula), Oasisamerica (Northwest), and Mesoamerica (Center to Southeast). The first two areas include part of the United States of America (USA), whereas Mesoamerica covers approximately from Central Mexico until Central America and is characterized by a large number of complex Pre-Columbian societies that flourished in this area, such as the Olmec, Maya, Teotihuacan, Totonac, and Aztec or Mexica (Fiedel, 1992; López-Austin and López-Lujan, 2001). Since 7000 BC, Mesoamerican Native populations began domestication of corn, beans, squash and chile, as well as turkey and dog, which caused a transition from nomadic hunter-gatherers to sedentary agricultural communities that become highly stratified and interconnected societies. The Mesoamerican traditions were

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: CONACyT (to H.R.-V.); Grant number: 129693.

*Correspondence to: Héctor Rangel Villalobos, Instituto de Investigación en Genética Molecular, Centro Universitario de la Ciénega (CUCI-UdeG), Av. Universidad #1115, Col. Paso Blanco, 47810 Ocotlán, Jalisco, México. E-mail: hrangel13@hotmail.com or Antonio González Martín, Departamento de Zoología y Antropología Física, Facultad de Biología, Universidad Complutense de Madrid, 28040 Madrid, España. E-mail: antonio@bio.ucm.es

Received 29 August 2015; revised 21 January 2016; accepted 6 February 2016

DOI: 10.1002/ajpa.22969
Published online 4 March 2016 in Wiley Online Library (wileyonlinelibrary.com).

subjugated by the Spanish conquest in the XVI century, but their cultural heritages still survive among the indigenous peoples of this region (Cavalli-Sforza et al., 1994). Conversely, Aridoamerica is the name given to the wide cultural area extending north of the Mesoamerican limits, which has a dry and arid climate. Due to the severe conditions, Pre-Columbian populations of Aridoamerica were mostly nomadic, and were known as Chichimecas, meaning barbarian or uncivilized. Growing corn reached Aridoamerica approximately 2100 BC, which led to the creation of sedentary cultures, as the Hohokam and Mogollon that formed a region known as Oasisamerica (Fiedel, 1992; López-Austin and López-Lujan, 2001).

Currently, the Mexican population can be classified in two principal groups: Mestizos and Native Americans. Mestizos are the Spanish-speaking population result of nearly 500 years of admixture between Native Americans, Spaniards, and Africans, principally, they constitute most of the present day Mexican population (~90%) living in both urban and rural regions throughout the Mexican territory (INEGI, 2005; Martínez-Cortés et al., 2012). Conversely, the Native American groups represent about 10% of the total Mexican population, and there are 156,557 indigenous settlements in 803 localities in which >30% of the individuals speaks an Amerindian language; almost 80% of this Native population is concentrated throughout the following eight Mexican states: Chiapas; Oaxaca; Guerrero; Hidalgo; Yucatan; Campeche; Veracruz and San Luis Potosí (INEGI, 2005). Considering the use of language as a selection criterion, there are more than 68 indigenous groups and more than 85 languages and dialect variants in Mexico (Cisneros, 2004). Genetic studies in Mexican Native groups are relevant because there are more than 12 million of indigenous individuals living in ~20% of the Mexican territory (CDI, 2012). Analysis of this genetic pool offers the opportunity to rebuild part of the human history based on these Native American populations.

Ideally, noncoding or neutral genetic markers are chosen trying to exclude the selection process to explain patterns of human genetic diversity (Jobling et al., 2004). Among these markers highlight the Short tandem repeat (STR) loci due to their elevated polymorphism and simplicity of analysis, which also explain that they have become the markers of choice in human identification. The implementation of STRs in forensic casework requires population validation that includes estimation of different statistical parameters, such as allele frequencies, power of discrimination and exclusion, testing of the Hardy-Weinberg and linkage equilibrium agreement, among others (Butler, 2014). Therefore, a lot a STR databases from worldwide populations have been generated during the last years. However, this genetic information also is worthy for anthropological studies because allows analyzing the origin, structure, and relationships between human populations (Pérez-Lezaun et al., 1997a,b; Bosch et al., 2000; Sahoo and Kashyap, 2005; Kraaijenbrink et al., 2014). During the last decade, a relative large number of forensic STR databases of Mestizos from different Mexican states have been generated, and global analyses have been carried out (Rubi-Castellanos et al., 2009; Salazar-Flores et al., 2015). Nevertheless, in Mexican Native groups the progress has been slower and only in recent years the number of studied populations has increased importantly. For instance, the 15 STRs highly used for human identification purposes have been characterized in Choles (Sánchez et al., 2005), Otomies and Huastecos (Barrot et al., 2005), Tepehuas (González-Martín et al., 2008), Mayas, Purépechas and Triquis (Ibarra-Rivera et al., 2008; Martínez-Cortés et al., 2010), and 10 indigenous groups from Oaxaca (Quinto-Cortés et al., 2010). More recently 9 and 10 Native American groups from the North and Northwest, as well as from the West, Center, and Southeast of Mexico were reported, respectively (Rangel-Villalobos et al., 2013, 2014). However, until now this information has not been exploited to obtain anthropological information regarding the Native American populations who represent important cultural regions of Pre-Hispanic Mexico. Therefore, in this study we collected 39 STR databases (Identifier AmpF*l*STR® kit) from Mexican Native American groups to develop a global anthropological analysis, including a new dataset of Mayos from Sinaloa (North, Mexico). Particularly, we aimed to explore the Pre-Hispanic demography of Mexico based on the following features: i) genetic diversity, including effective population sizes (Ne); ii) bottleneck events; iii) genetic relationships; iv) gene flow based on the number of migrants per generation (Nm); v) barriers to gene flow; vi) population structure under different criteria (i.e. linguistic, geographic, and cultural); and vii) Post-Columbian admixture in Mexican Native groups.

MATERIAL AND METHODS

Population sample

A total of 3,135 unrelated volunteers from 39 Mexican Native groups were included in this work, in addition to three Mexican-Mestizo (n = 562), one African (n = 135), and two European (n = 315) populations, as ancestral references. Details of the total population sample comprising 4,147 individuals are detailed in Table 1. To our knowledge, there is not a consensus classification regarding the geographic regions of the Mexican territory; thus, for discussion purposes, we clustered the studied Mexican indigenous groups into five regions according to their geographic location (Fig. 1). In the Mayo population sample from Sinaloa (North, Mexico), and samples collected by our research group (Martínez-Cortés et al., 2010; Rangel-Villalobos et al. 2013, 2014), it was verified that individuals were born in an indigenous community and that they or their parents speak the corresponding Amerindian language. Prior to the inclusion in the study, volunteers signed an informed consent letter according to the ethical guidelines of the Helsinki Declaration. The anonymity of the volunteers included in this study was preserved at all time. The project was approved by the Ethical Research Committee of the Centro Universitario de la Ciénega, at the Universidad de Guadalajara. The linguistic classification proposed by Swadesh (1959) was taken into account for analysis and discussion purposes, but with small modifications according to some anthropological descriptions, as detailed in Supporting Information Figure S1.

Laboratory analyses

DNA was extracted from fresh blood samples by standard phenol-chloroform or salting-out method, as well as directly from dried blood spotted on FTA paper for PCR purposes. We used the human identification kit Identifiler™ (Applied Biosystems, Foster City, CA), which coamplify the following autosomal STR loci: D8S1179; D21S11; D7S820; vWA; D18S51; D3S1358; D13S317;

TABLE 1. General information of the Mexican Native groups, Mexican-Mestizo (admixed), and Ancestral reference population samples analyzed in this study

N	Mexican Amerindian population	Simple size	Code	Location	Mexican state	Region, country	Lat (N)	Long (W)	Linguistic group	Reference
1	Seri	28	Ser	Desemboque, Pitiquito	Sonora	North, Mexico	29.5	−112.39	Macro Yuma	Rangel-Villalobos et al., 2013
2	Mayo, Sonora	45	MyoS	Masiaca, Navojoa	Sonora	North, Mexico	26.76	−109.23	Macro-Nahua	Rangel-Villalobos et al., 2013
3	Mayo, Sinaloa	88	MyoSi	Mochicahui, EL Fuerte	Sinaloa	North, Mexico	25.95	−108.92	Macro-Nahua	This study
4	Guarijio	17	Gua	San Bernardo, Alamos	Sonora	North, Mexico	27.39	−108.84	Macro-Nahua	Rangel-Villalobos et al., 2013
5	Tarahumara	116	Tar	Choguita, Guachoci	Chihuahua	North, Mexico	27.71	−107.63	Macro-Nahua	Rangel-Villalobos et al., 2013
		88	Tar	Sierra Madre of Chihuahua[a]	Chihuahua	North, Mexico	26.82	−107.07	Macro-Nahua	Rangel-Villalobos et al., 2013
6	Tepehuano del Sur	123	Tep	Duraznitos, Mezquital	Durango	North, Mexico	24.19	−104.86	Macro-Nahua	Rangel-Villalobos et al., 2013
7	Mexicanero	84	Mex	Curachitos, Mezquital	Durango	North, Mexico	23.47	−104.39	Macro-Nahua	Rangel-Villalobos et al., 2013
8	Huichol, Durango	90	HuiD	Duraznitos, Mezquital	Durango	Northwest, Mexico	22.59	−104.92	Macro-Nahua	Rangel-Villalobos et al., 2013
9	Huichol, Jalisco	117	HuiJ	San Sebastian, Mezquitic	Jalisco	Northwest, Mexico	22.07	−104.05	Macro-Nahua	Rangel-Villalobos et al., 2014
10	Huichol, Nayarit	32	HuiN	Naranjo, De Ruiz	Nayarit	Northwest, Mexico	21.04	−104.81	Macro-Nahua	Rangel-Villalobos et al., 2013
11	Cora	94	Cora	Presidio de Reyes-Corapan, Ruiz	Nayarit	Northwest, Mexico	22.04	−104.94	Macro-Nahua	Rangel-Villalobos et al., 2013
12	Nahua, Mezcala	226	NahM	Mezcala	Jalisco	Center, Mexico	20.33	−103.01	Macro-Nahua	Rangel-Villalobos et al., 2014
13	Nahua, Tuxpan	74	NahT	Tuxpan	Jalisco	Center, Mexico	19.55	−103.36	Macro-Nahua	Rangel-Villalobos et al., 2014
14	Purepecha Valley	157	PurV	Angahuan	Michoacán	Center, Mexico	19.54	−102.22	Tarasco	Martínez-Cortés et al., 2010
15	Purepecha Mountain	167	PurM	Zipiajo	Michoacán	Center, Mexico	19.74	−101.55	Tarasco	Martínez-Cortés et al., 2010
17	Tepehua	36	Tph	Sierra Otomí-Tepehua	Hidalgo	Center, Mexico	20.33	−98.23	Macro-Maya	González-Martín et al., 2008
18	Huasteco	108	Huas	La Huasteca (Sierra Madre)	Hidalgo	Center, Mexico	21.01	−98.2	Macro-Maya	Barrot et al., 2005
16	Otomi, Valley	80	OtoV	Ixmiquilpan Valley	Hidalgo	Center, Mexico	20.48	−99.2	Macro-mixteca	Barrot et al., 2005
19	Popoloca	51	Popo	San José Miahuatlan	Puebla	South, Mexico	18.3	−97.31	Macro-Mixteca	Rangel-Villalobos et al., 2014
20	Amuzgo	30	Amu	San Pedro Musgos	Oaxaca	South, Mexico	16.65	−98.08	Macro-Mixteca	Quinto-Cortés et al., 2010
21	Chinanteco	40	Chin	San Lucas Ojitan	Oaxaca	South, Mexico	18.06	−96.4	Macro-Mixteca	Quinto-Cortés et al., 2010

TABLE 1. Continued

N	Mexican Amerindian population	Simple size	Code	Location	Mexican state	Region, country	Lat (N)	Long (W)	Linguistic group	Reference
22	Chontal	29	Chon	Morro de Mazatan	Oaxaca	South, Mexico	16.1	−95.38	Macro-Yuma	Quinto-Cortés et al., 2010
23	Huave	29	Hua	San Dionisio del Mar	Oaxaca	South, Mexico	16.31	−94.75	Macro-Mixteca	Quinto-Cortés et al., 2010
24	Mazateco1	31	Maz1	San Felipe Jalapa de Diaz	Oaxaca	South, Mexico	18.07	−96.75	Macro-Mixteca	Quinto-Cortés et al., 2010
25	Mazateco2	80	Maz2	San Miguel Soyaltepec	Oaxaca	South, Mexico	18.24	−96.41	Macro-Mixteca	Rangel-Villalobos et al., 2014
26	Mixe	30	Mixe	San Pedro y San Pablo Ayutla	Oaxaca	South, Mexico	17.03	−96.07	Macro-Maya	Quinto-Cortés et al., 2010
27	Mixteco	30	Mixt	Santiago Juxtlahuaca	Oaxaca	South, Mexico	17.33	−98.01	Macro-Mixteca	Quinto-Cortés et al., 2010
28	Triqui	37	Triq	San Juan Copala	Oaxaca	South, Mexico	17.18	−97.96	Macro-Mixteca	Quinto-Cortés et al., 2010
29	Zapoteco, Istmo	30	ZapI	Ciudad Ixtepec	Oaxaca	South, Mexico	16.56	−95.1	Macro-Mixteca	Quinto-Cortés et al., 2010
30	Zapotecos, Valley	40	ZapV	Magdalena Teitipac	Oaxaca	South, Mexico	16.91	−96.56	Macro-Mixteca	Quinto-Cortés et al., 2010
31	Zoque	35	Zoq	Santa Maria Chipalapa	Oaxaca	South, Mexico	16.91	−94.68	Macro-Maya	Quinto-Cortés et al., 2010
32	Tzotzil	174	Tzo	San Juan Chamula	Chiapas	Southeast, Mexico	16.78	−92.69	Macro-Maya	Rangel-Villalobos et al., 2014
33	Tjolobal	75	Tjo	Las Margaritas	Chiapas	Southeast, Mexico	16.31	−91.98	Macro-Maya	Rangel-Villalobos et al., 2014
34	Tzeltal	20	Tze	Tenejapa	Chiapas	Southeast, Mexico	16.81	−92.51	Macro-Maya	Rangel-Villalobos et al., 2014
35	Lacandon	79	Lac	Ocosingo	Chiapas	Southeast, Mexico	16.9	−92.09	Macro-Maya	Rangel-Villalobos et al., 2014
36	Chol	104	Chol	Northeast Region	Chiapas	Southeast, Mexico	17.5	−91.98	Macro-Maya	Sánchez et al., 2005
37	Maya, Campeche[a]	48	MayC	Campeche	Campeche	Southeast, Mexico	20.15	−90.07	Macro-Maya	Ibarra-Rivera et al., 2008
		38	MayC	Campeche	Campeche	Southeast, Mexico	19.51	−91.08	Macro-Maya	Rangel-Villalobos et al., 2014
38	Maya, Yucatan[a]	113	MayY	Yucatan	Yucatan	Southeast, Mexico	20.93	−89.88	Macro-Maya	Ibarra-Rivera et al., 2008
39	Maya, Quintana Roo	78	MayQR	Felipe Carrillo Puerto	Quintana Roo	Southeast, Mexico	19.57	−88.03	Macro-Maya	Rangel-Villalobos et al., 2014
		144	MayQR	Chunhuhub	Quintana Roo	Southeast, Mexico	19.58	−88.59	Macro-Maya	Rangel-Villalobos et al., 2014
	Total	3,135								

	Mexican Mestizo population	Sample size	Code	Origin	Mexican State	Region, Country	Latitude (N)	Longitude (W)	Idiom	Reference
42	Chihuahua	162	Chih	Admixed	Chihuahua	North, Mexico	28° 37'	106° 03'	Spanish	Martínez-González et al., 2005

TABLE 1. Continued

	Mexican Mestizo population	Sample size	Code	Origin	Mexican State	Latitude (N)	Longitude (W)	Idiom	Reference
43	Jalisco	200	Jal	Admixed	Jalisco	20° 40'	103° 21'	Spanish	Rubi-Castellanos et al., 2009
44	Yucatán	200	Yuc	Admixed	Yucatán	20° 57'	89° 37'	Spanish	Rubi-Castellanos et al., 2009
	Total	562							

	Ancestral reference	Sample size	Code	Location	Country	Region, Country	Latitude (N)	Longitude (W)		Reference
40	European, pool	200	Eur	Central region	Portugal	Center, Portugal	38° 48'	−9° 14'	Portuguese	López et al., 2009
		115	Eur	Andalucia	Spain	South, Spain	37° 12'	−3° 33'	Spanish	Coudray et al., 2007
41	African reference	135	Afri	Equatorial Guinea	Equatorial Guinea	West, Africa	1° 27' (N)	10° 27' (E)	Spanish, French	Alves et al., 2005
	Total	450								

[a] Population pool obtained from different locations.

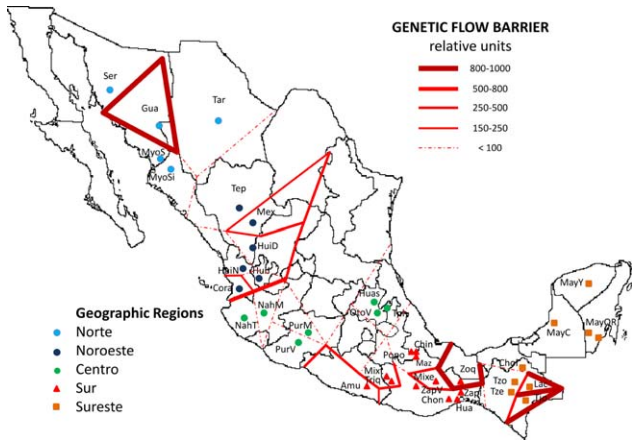

Fig. 1. Approximated geographic location of the Mexican Native groups analyzed in this study. Gene flow barriers detected between these populations are indicated. For abbreviations please check Table 1.

D5S818, FGA, CSF1PO, D19S433, TPOX, TH01, D16S539, and D2S1338. The amplified products were separated by capillary electrophoresis using the ABI Prism™ 310 and 3130 Genetic Analyzers following manufacturer recommendations. The allelic ladder provided with the kit and the software GeneMapper ID version 3.2 were utilized for STR genotyping.

Data analysis

Forensic parameters (Mayos from Sinaloa, North, Mexico). For the Mayo population sample (n = 88), we carried out descriptive analyses including estimation of allele frequencies and statistical parameters of forensic importance, as well as Hardy-Weinberg equilibrium and linkage disequilibrium (LD) exact tests (Table 1). For these tasks, we used the programs PowerStats (Tereba, 1999) and Genetic Data Analysis (GDA) version 1.1 (Weir, 1996), respectively. The following forensic parameters were reported by STR and the 15 loci system: Power of discrimination (PD), Power of exclusion (PE), Heterozygosity observed (H), Polymorphic informativity content (PIC), Typical paternity index (TPI), and Minimum allele frequencies (MAF).

Genetic diversity analysis. For the 39 Mexican Native American groups, the following genetic diversity statistics were calculated using the Excel complement GeneAlEx 6.5 (Peakall and Smouse, 2006, 2012): number of alleles (Na), number effective of alleles (Ne), observed heterozygosity (Ho), expected heterozygosity (He), inbreeding F-statistic (F). The theta value (θ) was calculated for all populations (as described by Friedlaender et al., 2008), which is linearly correlated with effective population size; thus, theta is appropriated to represent differences in effective population sizes among populations (e.g. a ratio of 2 between two populations indicate twice effective population size between populations). The Excel complement XL-STAT 3.07 for Windows was used to perform the non-parametric Kruskal-Wallis test checking significant differences among populations for θ and Ho, followed by Dunńs multiple comparison test. In addition, we estimated the number of alleles exclusively

observed in one population or "private alleles" with the program ADZE (Allelic Diversity AnalyZEr), which uses a rarefaction approach to trim unequal samples to the same standardized sample size (g), and to assess the sample size-corrected private allelic richness to any set of populations (Szpiech et al., 2008). Finally, genotype data were analyzed in Bottleneck 1.2.02 (Cornuet and Luikart, 1996) to reveal the possible occurrence of recent bottleneck in each Native American group.

Genetic relationship analyses. We evaluated the genetic relationships among Mexican Native groups by means of the following methods: 1) pairwise Fst distances using the program GeneAlEx 6.5 (Peakall and Smouse, 2006, 2012) and GDA 1.1 (Weir, 1996). This distance based on STR data was selected because it has been demonstrated that represents genetic differentiation patterns by drift, corresponding with both genetic and archeological records of human populations (Pérez-Lezaun et al., 1997, Calafel et al., 1998); 2) the pairwise Fst matrix was used to construct an unrooted neighbor joining (NJ) tree evaluating the branch́s consistence by means of 1,000 permutations with the package Phylip 3.695 (Felsenstein, 1989), program freely available at: http://evolution.gs.washington.edu/phylip.html. Similarly, we obtained a rooted NJ tree of Native American groups but including the three Mexican-Mestizo and ancestral populations with GDA 1.1 (Weir, 1996) and TreeView (Page, 1996) softwares. Subsequently, to simplify the landscape of genetic relationships we omitted "outlier" populations with large differentiation level, such as North and Northwest populations, plus Lacandones and Tojolabales; in this way, we obtained a Multidimensional Scaling (MDS) plot in the Excel-complement XL-STAT 3.07 for Windows; 3) the intensity of gene flow between Mexican Amerindian groups was assessed as the number of migrants per generation (Nm) according to the equation of Wright based on pairwise Fst values between populations. The parameter Nm was graphically represented in an Excel matrix based on gray-black intensity differences; 4) the genetic boundaries presence among indigenous populations throughout the Mexican territory, presumably representing Pre-Hispanic demographic events, was evaluated with the program Barrier version 2.2 (Manni et al., 2004). This program worked with the previously generated pairwise Fst matrix; 5) in order to investigate whether Isolation-by-distance (IBD) could explain genetic differentiation among indigenous populations, we revised the correlation between genetic and geographic distances among these groups (Ramachandran et al., 2005). Based upon genetic distances a $F_{ST}/(1-F_{ST})$ matrix was constructed (Rousset, 1997), and distances in kilometers between populations were estimated employing geographic coordinates with the program GeneAlEx 6.5. In addition to the correlation test using the complete dataset, subsets of the population sample were clustered to test the IBD model separately based on distance differences: <1,000 km, >1,000, and >1,500 km. The statistical significance of these correlations was evaluated by Mantel tests using the software tool zt (Bonnet and Van de Peer, 2002), and the Isolation by Distance Web Service (IBDWS) version 3.23 (Jensen et al., 2005), with the default parameters as provided by the web service.

Genetic structure analysis. The STR genotype dataset was further examined with the software STRUCTURE 2.3.2 (Pritchard et al., 2000; Falush et al., 2003, 2007) using the admixture model with correlation between allele frequencies across clusters. The number of clusters (K) presented ranged from 2 to 7; for each K 50 independent runs were carried out. Results from $K=5$ and $K>7$ were omitted because they gave practically the same or unclear panorama than the already obtained, probably due to the limitations of the 15 STR loci used for this analysis. We utilized a 20,000-iteration burn-in period followed by 20,000 iterations. In addition, we assessed and detected the number of genetic groups (K) that best fit to the data by means of the program Structure-harvester (Earl and vonHoldt, 2012). For the Mexican Native population dataset, we removed all information concerning geographic label (unsupervised analysis). In order to detect the probable presence of European or African admixture in the Native American population samples, we included ancestral references and Mexican Mestizo populations (Table 1) and we performed a supervised analysis identifying the genetic pool of parental populations in the genetic dataset. A total of 50 repetitions were assessed following identical settings as in the unsupervised method. Based on this analysis, we estimated the proportion of Post-Columbian admixture in the Mexican Native American groups, which was plotted with the software Excel. For discussion purposes of the admixture results by region, nearby Maya populations from Chiapas (Tzo, Tjo, Tze, and Lac) and those closer to the Yucatan Peninsula (Chol, MayC, MayY, MayQR) were clustered in separated groups (MayChiap and MayPenin, respectively). Analysis molecular of variance (AMOVA) was carried out in Mexican indigenous groups with the software ARLEQUIN 3.1 (Excoffier et al., 2005). For this purpose, populations were clustered under different criteria such as geography, linguistic classification, and historical records as properly described in the text. Finally, we utilized the Spatial analysis of molecular variance (SAMOVA) to define geographically and genetically homogeneous population groups, as well as those groups sufficiently differentiated from each other (Dupanloup et al., 2002).

RESULTS

Forensic parameters in Mayos from Sinaloa group

Allelic distribution and statistical parameters of forensic importance of the 15 STRs were estimated in Mayos from Sinaloa, and all 15 loci were in Hardy-Weinberg equilibrium (Supporting Information Table S1). Allelic association between pairs of loci was discarded by the linkage disequilibrium (LD) exact tests (result not shown). The combined PD and PE for the system of 15 STRs were >99.99%, sufficient to solve most of the human identification cases. This information validates the employment of the 15-STR system in forensic casework and paternity cases in this indigenous community (Supporting Information Table S1) (Butler, 2014). In general, availability of STRs databases in Mexican Native American groups potentially improves the interpretation of genetic evidence in forensic casework and, consequently, the administration of justice in these historically vulnerable and discriminated populations. However, because this work has demographic and

anthropological perspective, the forensic findings above described will not be further discussed.

Genetic diversity parameters

Different parameters of genetic diversity and *P* values of the bottlenecks test are described in Table 2. Although heterozygosity was not different between populations ($P = 0.65$; Kruskal–Wallis test), Lacandones were peculiar by presenting the smallest diversity values (Ho, He, Nea, and theta) and largest inbreeding (*F*). Conversely, comparison of theta (θ) indicated significant differences in effective population size (Ne) between populations (Supporting Information Table S2). For discussion purposes, we present graphically the comparison of theta indicating three main population clusters that are partially consistent with some geographical regions, as shown in Figure 2. The North and Northwest groups in addition to some isolated populations from Chiapas and Oaxaca showed the lowest theta values among the Mexican Native groups. Conversely, these values increase slightly in Maya groups from the Yucatan Peninsula. Although bottleneck effect was detected in most of the Mexican Native groups (64%) (Table 2), between regions the proportion displayed differences: North (3/5 = 60%), Northwest (1/6 = 16.7%), Center (5/7 = 71.4%), Southern (9/13 = 69.2%), and Southeast (6/8 = 75%). The comparison of private allele number (sample-size corrected) is presented between regions (Fig. 3) and between populations into each region (Supporting Information Fig. S2). The Northwest and North displayed the lowest number of private alleles, respectively. In opposition, the Center presented the highest number, and the South and Southeast had intermediate values very similar between each other.

Genetic relatedness

Based on pairwise Fst distances between the studied populations (Supporting Information Table S3), we obtained and unrooted NJ tree where the most differentiated populations (outliers) are indicated (Fig. 4a): native groups from the North and Northwest regions plus Lacandones and Tojolabales. In general, the population clustering shown in the NJ tree is in agreement with linguistic and geographic criteria, such as the nearby Huicholes and Tepehuanos from a common linguistic trunk; they are also adjacent to Tarahumaras, Tepehuanes, Mexicaneros, and Coras whose cultural and linguistic similarity is well known (Supporting Information Fig. S1). In order to clarify the genetic relationships among the majority of the Mexican Native groups and Mestizos, in Figure 4b we obtained a MDS plot omitting the outlier populations that allows distinguishing better closer populations. In this MDS plot some populations from Oaxaca were located in the periphery, suggesting a relative higher degree of isolation. Conversely, the position of the Mestizo populations seems to be related to the ancestral proportion those presumably with higher Amerindian and European ancestry (Yucatan and Chihuahua, respectively) were located toward the center and periphery, respectively (Fig. 4b). This is in agreement with previous studies in Mexican Mestizos (Rubi-Castellanos et al., 2009; Salazar-Flores et al., 2015).

Gene flow estimation (nm)

The number of migrants per generation (Nm) between populations was plotted in a matrix with different intensities of gray shades (Fig. 5). In counterpart, geographic barriers to gene flow between the Mexican Native groups were represented graphically (Fig. 1). Both results showed—again—the largest isolation of Lacandones and Tojolabales (Southeast, Mexico) followed by populations from the North and Northwest regions and some groups from Oaxaca, respectively. As could be expected, this is in agreement with differences in effective population size (Fig. 2). Although the analysis of the genetic differentiation between the complete dataset of Native populations slightly seems to fit to the IBD model (Supporting Information Fig. S3), a deeper analysis of subsets of samples clarifies that IBD model fits only when populations are separated by large distances, particularly >1,500 km ($r^2 = 0.10765$; $P = 0.0002$) (Supporting Information Fig. S3).

Population structure and admixture

The STRUCTURE results of the unsupervised and supervised analyses with the total population dataset are shown in Figure 6a,b. Three genetic groups was the number that best fit for the data in the unsupervised analysis. This genetic structure involved two Native American components representing the North/Northwest (purple) and Center/Southeast (orange) regions besides to the non-Native American (sky blue) component predominant in the ancestral populations and Mexican Mestizos (Fig. 6a; $K = 3$). However, the European and African components were not differentiated between each other, probably due to the larger differentiation level among Mexican Native groups than the observed between the parental populations used as reference, as below confirmed by the AMOVA results (Supporting Information Table S4). Therefore, admixture detected can be generically described as "no-indigenous" component. The following structure results ($K \geq 4$) make evident the differentiation of Lacandones and Tojolabales from the rest of populations. The next genetic structure levels presented suggest some peculiar genetic components that allow clustering some North and Northwest populations, in addition to indicate a peculiar genetic structure of Tzotziles and Purépechas Valley, and Nahuas from Mezcala (Fig. 6a; $K = 6$ and 7).

Interestingly, Structure-harvester results for the supervised analysis indicate that $K = 7$ provides the best fit to explain the population structure of the studied population sample. Although unsupervised and supervised analyses practically offer the same panorama, the latter allow separating the European and African population references (Fig. 6b; $K = 7$). Thus, the following genetic components related to some specific populations could be described: 1) European (sky blue); 2) African (orange); 3) Lacandon-Tojolabal (bright pink); 4) Tarahumara-Guarijios (fandango); 5) Huichol-Tepehuano (pale pink); 6) Nahuas Mezcala (green) 7) Purepechas Valley and Tzotziles (purple) (Fig. 6b; $K = 7$). Subsequently, we estimated the admixture proportion in each region and Mexican Native population (Fig. 7); for simplicity, this was based on the no-indigenous component obtained by the supervised analysis (Fig. 6b; $K = 3$). By region, the largest admixture was observed in the North (23.9%) and Mayas from the Yucatan Peninsula (20.1%).

TABLE 2. Descriptive statistics of the genetic diversity in 39 Mexican Native groups[a]: sample size (n), number of alleles (Na), number effective of alleles (Nea), heterozygosity observed (Ho), heterozygosity expected (He), inbreeding coefficient (F), teta (θ) as comparative indicator of effective population size (Ne), and Bottleneck test P value (P-B)

Region	North					Northwest						Center							South	
Population, mean	Ser	MyoS	MyoSi	Gua	Tar	Tep	Mex	HuiD	HuiJ	HuiN	Cora	NahM	NahT	PurV	PurM	OtoV	Tph	Huas	Popo	Amu
N	28.0	45.0	88.0	17.0	204.0	123.0	84.0	90.0	117.0	32.0	93.9	225.9	73.6	156.8	166.8	79.3	35.5	106.9	51.0	30.0
Na	6.07	7.07	8.40	5.53	8.33	6.80	6.93	6.87	6.47	6.20	7.33	9.60	9.00	8.00	9.07	7.53	6.60	7.87	7.20	6.13
Nea	3.55	4.18	4.54	3.56	4.21	3.59	3.85	3.85	3.62	3.69	4.03	4.21	4.42	4.07	4.48	4.36	3.95	4.06	4.30	3.82
Ho	0.69	0.76	0.74	0.71	0.72	0.68	0.73	0.72	0.67	0.70	0.71	0.72	0.72	0.75	0.73	0.74	0.73	0.73	0.76	0.73
He	0.69	0.72	0.74	0.69	0.73	0.69	0.71	0.70	0.69	0.69	0.72	0.73	0.75	0.72	0.74	0.74	0.72	0.72	0.74	0.72
F	0.01	−0.05	0.00	−0.02	0.02	0.02	−0.04	−0.02	0.03	−0.02	0.00	0.02	0.04	−0.03	0.02	0.00	−0.01	0.00	−0.02	−0.01
θ	4.72	6.07	6.99	4.67	6.26	4.70	5.33	5.24	4.73	4.86	5.78	6.38	7.21	6.10	7.08	6.82	5.83	6.08	7.01	5.99
P-B[b]	*0.0008*	*0.0124*	*0.0027*	*0.0004*	*0.0109*	*0.0386*	*0.0181*	*0.0730*	*0.2078*	*0.0003*	*0.0102*	*0.00003*	*0.00004*	*0.0102*	*0.0006*	*0.01508*	*0.00168*	*0.00168*	*0.0882*	*0.0003*

Region	South											Southeast								
Population, mean	Chin	Chon	Hua	Maz1	Maz2	Mixe	Mixt	Triq	ZapI	ZapV	Zoq	Tzo	Tjo	Tze	Lac	Chol	MayC	MayY	MayQR	Global
n	40.0	29.0	29.0	31.0	80.0	30.0	30.0	37.0	30.0	40.0	35.0	174.0	75.0	20.0	79.0	103.8	85.9	112.6	222.0	80.3
Na	6.80	7.13	6.87	6.73	7.40	6.53	7.00	6.40	6.93	6.60	6.33	7.13	6.40	5.80	6.33	8.73	8.33	8.67	9.00	7.23
Nea	4.19	4.27	4.27	3.97	4.23	3.60	4.23	3.68	4.29	4.24	3.47	3.95	3.58	3.66	3.26	4.53	4.35	4.41	4.56	4.03
Ho	0.74	0.73	0.74	0.74	0.72	0.70	0.71		0.74	0.74	0.68	0.71	0.70	0.69	0.60	0.74	0.71	0.73	0.75	0.72
He	0.74	0.74	0.73	0.72	0.73	0.69	0.73	0.69	0.73	0.73	0.68	0.72	0.68	0.69	0.66	0.75	0.74	0.75	0.76	0.72
F	−0.01	0.01	−0.01	−0.03	0.02	−0.01	0.03	0.00	−0.01	−0.02	−0.01	0.01	−0.02	0.00	0.08	0.01	0.04	0.03	0.01	0.00
θ	6.65	7.01	6.14	5.67	6.59	4.67	6.19	4.68	6.56	6.32	4.30	5.83	4.49	4.75	3.75	7.81	6.70	7.46	8.12	5.80
P-B[b]	*0.0832*	*0.00003*	*0.00006*	*0.0013*	*0.0479*	*0.00009*	*0.0004*	*0.0003*	*0.0013*	*0.3591*	*0.0000*	*0.0034*	*0.0151*	*0.0084*	*0.010*	*0.0013*	*0.0006*	*0.0006*	*0.0027*	

[a] Please refer to Table 1 for abbreviations.
[b] For illustrative purposes, cells shadow intensity is related with the significance of the Bottleneck (*P* value).

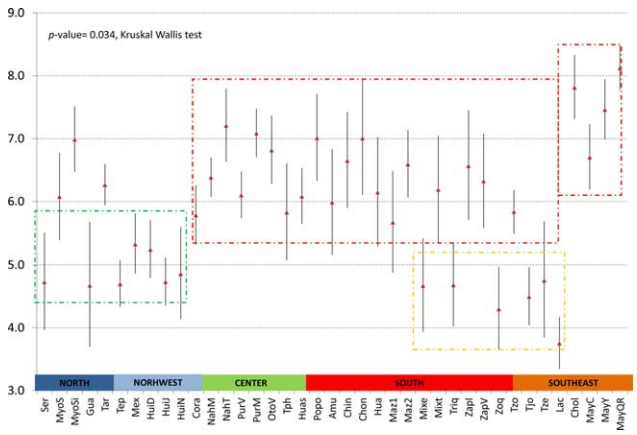

Fig. 2. Comparison of theta value (representing Ne) between Mexican Native groups which were clustered for discussion purposes. For abbreviations please check Table 1.

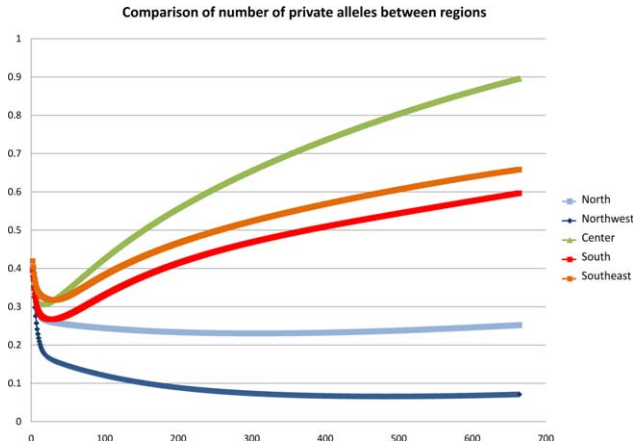

Fig. 3. Private allele diversity (*Y*-axis) considering a standardized sample size (*X*-axis) compared between Mexican Native groups clustered by region (based on ADZE program). For abbreviations please check Table 1.

Conversely, Mayas from Chiapas (9.4%) and the Northwest region (8.7%) displayed the lowest admixture. By population, Mayos from Sonora (35.5%) and Huicholes from Jalisco (6%) showed the maximum and minimum admixture levels, respectively. In general, these results suggest the presence of nonindigenous admixture in all the present-day Mexican Native groups with important differences between some geographic regions. However, limitations of the STR loci to perform this task compel careful interpretation of these preliminary admixture estimates.

The population structure also was analyzed by AMOVA including "main" population groups (Native Americans, Mestizos, and ancestral populations) (Supporting Information Table S4). Mexican Native groups present the largest interpopulation differentiation, sligtly larger than the ancestral reference (Fst = 2.69 vs. 2.47%), whereas Mestizos displayed the lowest. By region, the largest was shown in the North and the lowest in the Center (Fst = 2.74 vs. 1.44%). Interestingly, omitting the most differentiated populations (Lacandones and Tojolabales) was evident that the interpopulation differentiation between Mexican Native groups actually is lower in Mesoamerica (Fst = 1.43%), and particularly in the Southeast region (Fst = 0.82%) (Supporting Information Table S4). Comparison between main the population groups demonstrates that the genetic differentiation between Mexican Native groups and Mestizos is low and non-significant (Fct = 0.26%; $P = 0.2909$). The following AMOVA results analyze the genetic structure of Mexican Native groups under diverse geographic, cultural and linguistic criteria (Supporting Information Table S5). Genetic differentiation into-populations and into-groups were significant in all cases. Interestingly, differentiation between groups (Fct) was not significant ($P > 0.05$) between linguistically related Macro-Mayan groups clustered either by state or by geographic region. Finally, SAMOVA tests at different levels ($K = 2$ to 20) separated to each Mexican Native group independently; thus, any geographic population cluster was evident with this test (data not shown).

DISCUSSION

To our knowledge, although this is the widest study based on STRs regarding Mexican Native groups from a genetic-demographic perspective, some considerations should be taken into account for interpretation of these findings: 1) uncontrollable sampling bias, such as the inbreeding that is easily inferred by the elevated frequency of particular surnames, but is difficult to detect in indigenous communities without familial records. 2) The small population sample size for some Mexican Native groups, such as Guarijios ($n = 17$) and Tzeltals ($n = 20$). For the latter issue, is worthy considering the conclusions of Shriver et al. (1995) who stated that the genetic distance variances based on STRs shows no significant changes after 25 individuals (50 chromosomes). In addition, their value is demonstrated by the congruency of the established genetic relationships (Fig. 4), which support their inclusion in this work. 3) Origin of population samples coming from one or few communities must be taken into account when particular historical processes are inferred, such as bottleneck, genetic drift, among others. For this reason, samples' geographic origins are indicated in Table 1. 4) The possibility of admixture estimation bias, as a result of ancestral reference limitations and the poor ability of STRs to perform this task. 5) Bottleneck inference, because this test does not quantify—and rectify—effects such as admixture, inbreeding, and population sample size; in addition, this test does not indicate the period of time the bottleneck event presumably occurred neither (Cornuet and Luikart, 1996). 6) Unfortunately, the Pre-Hispanic human movement is partially known; thus, some findings could not have an immediate or obvious explanation. In brief, the genetic landscape offered herein represents an approximation that requires to be contrasted with additional genetic and non-genetic evidence to become significant, under considerations like the abovementioned. The subsequent discussion arises from the comparison regarding the effective population size (Ne) between populations (Fig. 2), as the basic parameter that promotes differentiation by genetic drift, reducing the genetic variability, and eventually modifying the genetic relationships. This parameter and geography were used for clustering Mexican Native groups and make easy to discuss the observed findings.

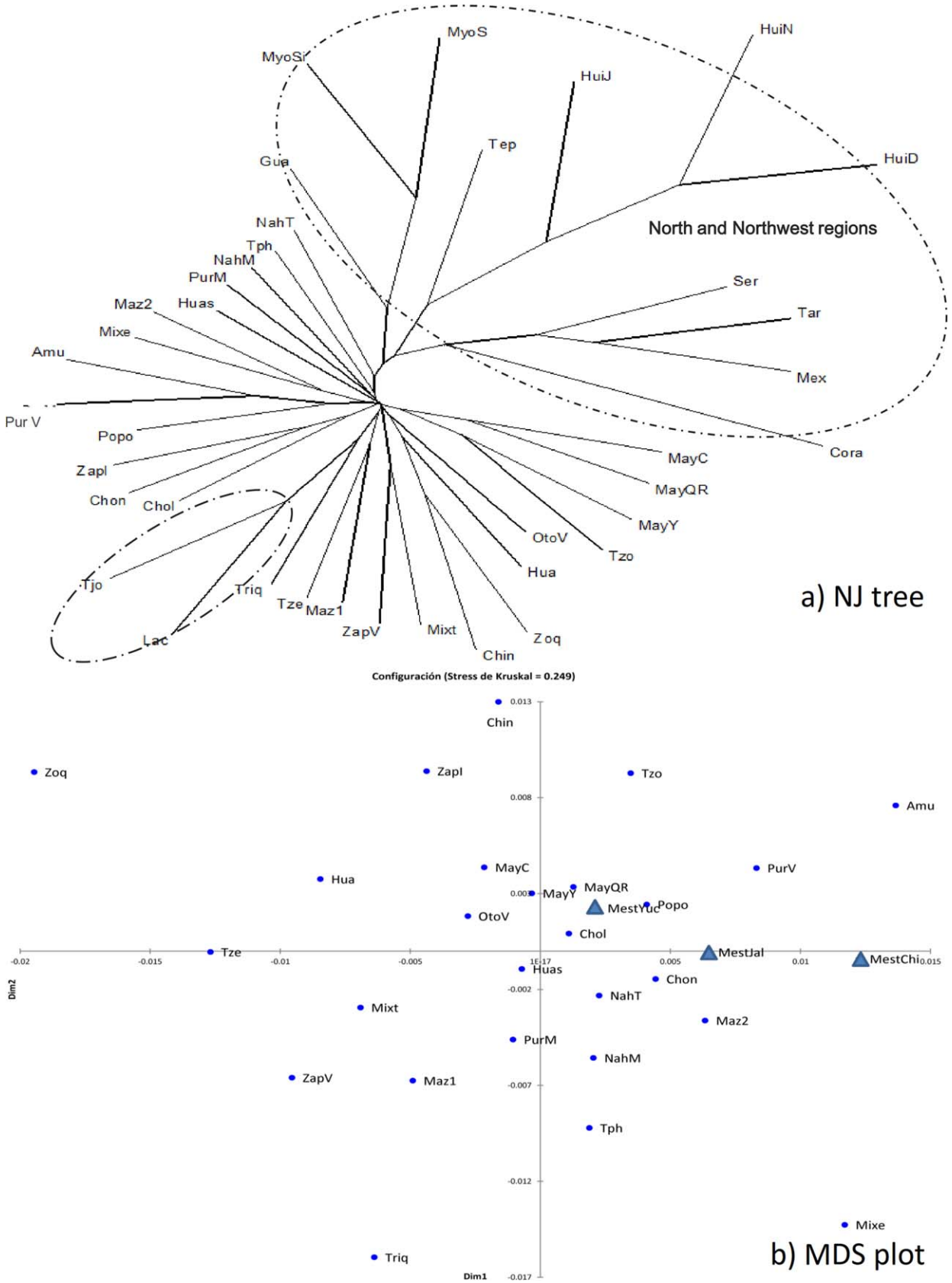

Fig. 4. Genetic distances between Mexican Native groups represented in a NJ tree (**a**) and in MDS plot (**b**) omitting outlier populations. For abbreviations please check Table 1.

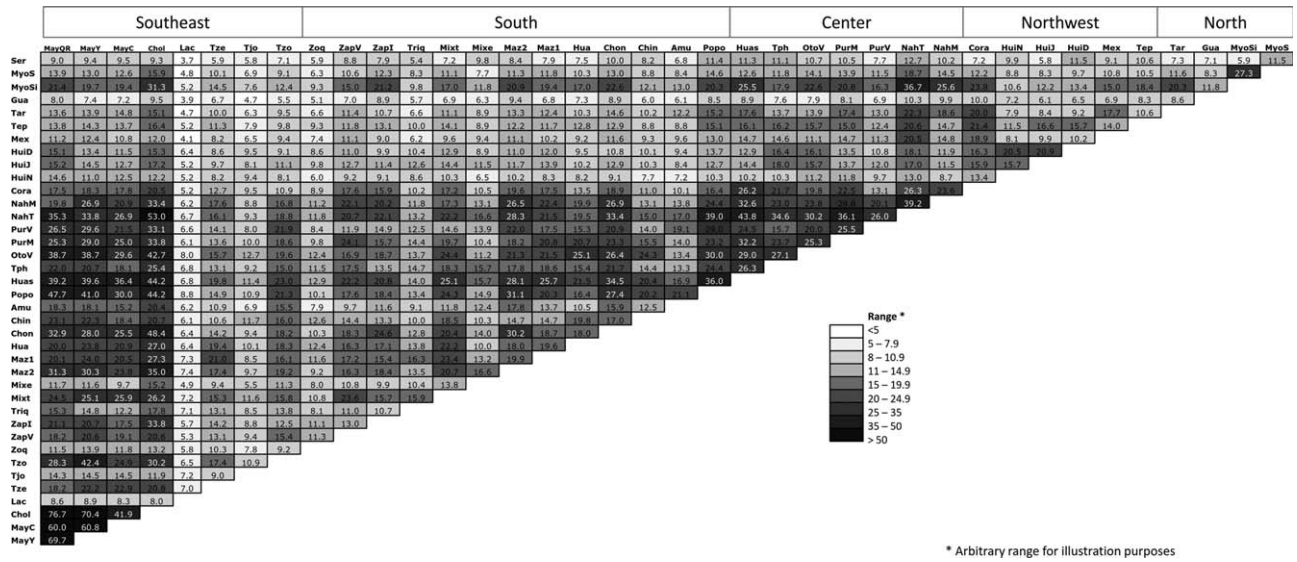

Fig. 5. Pairwise estimates of the number of migrants (Nm) between the studied Mexican Native groups. Arbitrary black/gray tones are indicated in the upper matrix to illustrate the gene flow intensity. For abbreviations please check Table 1.

Gene flow and genetic relationships in the Mesoamerican core

Based on theta value (Fig. 2), it is clear that most of the Mexican Native groups of the Mesomerican core (from the Center to Southeast) present similar effective population size (Ne) and increased number of private alleles respect to the rest of regions (Fig. 3). The existence of important gene flow (Nm) was estimated among Mayas of the Yucatan Peninsula (Southeast) and different Native groups from the Center and South (Fig. 5), displaying a low and nonsignificant structure between Mayan groups linguistically related (Supporting Information Table S5). However, some groups from Chiapas and Oaxaca displayed larger isolation and differentiation represented as strong barriers (Fig. 1) and even as a particular genetic structure level (i.e. Lacandones and Tojolobales, followed by Tzotziles) (Fig. 6). The larger genetic differentiation of these groups also indicated by their lowest genetic diversity values (Table 2) and peripheral position in the MDS plot (Fig. 4b). This finding in the MDS plot has been interpreted as supporting a demographic history characterized by large genetic drift effects, as described among Mexican Native groups based on mtDNA analysis (González-Martín et al., 2015). In fact, the impact of the most isolated Native groups (i.e. Lacandones and Tojolabales) in the genetic structure was evident when they were omitted to quantify the interpopulation variability, which diminished 30.6% in the Mesomerican core (Fst = 2.06 to 1.43%) and 62% in the Southeast region (Fst = 2.13 to 0.82%) (Supporting Information Table S4). In brief, omitting isolated groups from Chiapas and Oaxaca, our results support the existence gene flow among populations from the Center to the Southeast of Mesoamerica.

The Mesomerican gene flow hypothesis is reinforced by the Olmec origin of the proto-Mayan language that would suggest a close genetic relationship between Mayas (Southeast) and Huastecos (Center) (Coe and Kootz, 2002; Sharer and Traxler, 2006). Interestingly, in this study Huastecos displayed the lowest genetic distances with Choles and Mayas from Yucatan and Quintana Roo (Fst = 0.006; Supporting Information Table S3), supporting an ancestral relationship between these groups. This conclusion is in agreement with previous studies with nine STRs (Martínez-Cortés et al., 2010) and with the genomic analysis that suggest facilitating routes (Moreno-Estrada et al., 2014), such as the Atlantic corridor represented by the Gulf of Mexico coasts and with the genetic barriers estimated herein (Fig. 1). In general, this panorama is in agreement with the lack of differentiation described for Mesoamerica respect to Native American groups throughout the continent (Wang et al., 2007), and the high population density concomitant to the agricultural development (Fiedel, 1992; Cavalli-Sforza et al., 1994). Similarly, this finding partially harmonizes with the elevated human mobility described between groups from Oaxaca (Quinto-Cortés et al., 2010). Although our conclusions could be biased by the relatively high admixture level detected in Native groups from the Central region and Yucatan Peninsula (Fig. 7), probably this was not significant given its consistence with conclusions obtained in the genomic study where admixture effects were controlled (Moreno-Estrada et al., 2014).

Interestingly, the aforementioned hypothesis was discarded in a previous study with 15 STRs due to the clear differentiation between Mayas and Huastecos (Ibarra-Rivera et al., 2008). In addition to the larger number of Mexican Native groups studied in this study, the opposing conclusions could be explained by the large number of missing STR data of the Huasteco's database used in both studies (Barrot et al., 2005), which could have eclipsed the actual genetic relationships between populations in the previous study (Ibarra-Rivera et al., 2008). Therefore, the STR database purification carried out in this work (omitting incomplete DNA profiles) seems to be critical to avoid obtaining different conclusions from the same populations.

Gene flow between Maya groups

Southeast region. The Mayas of the Yucatan Peninsula and Choles had the largest Ne among the Mexican

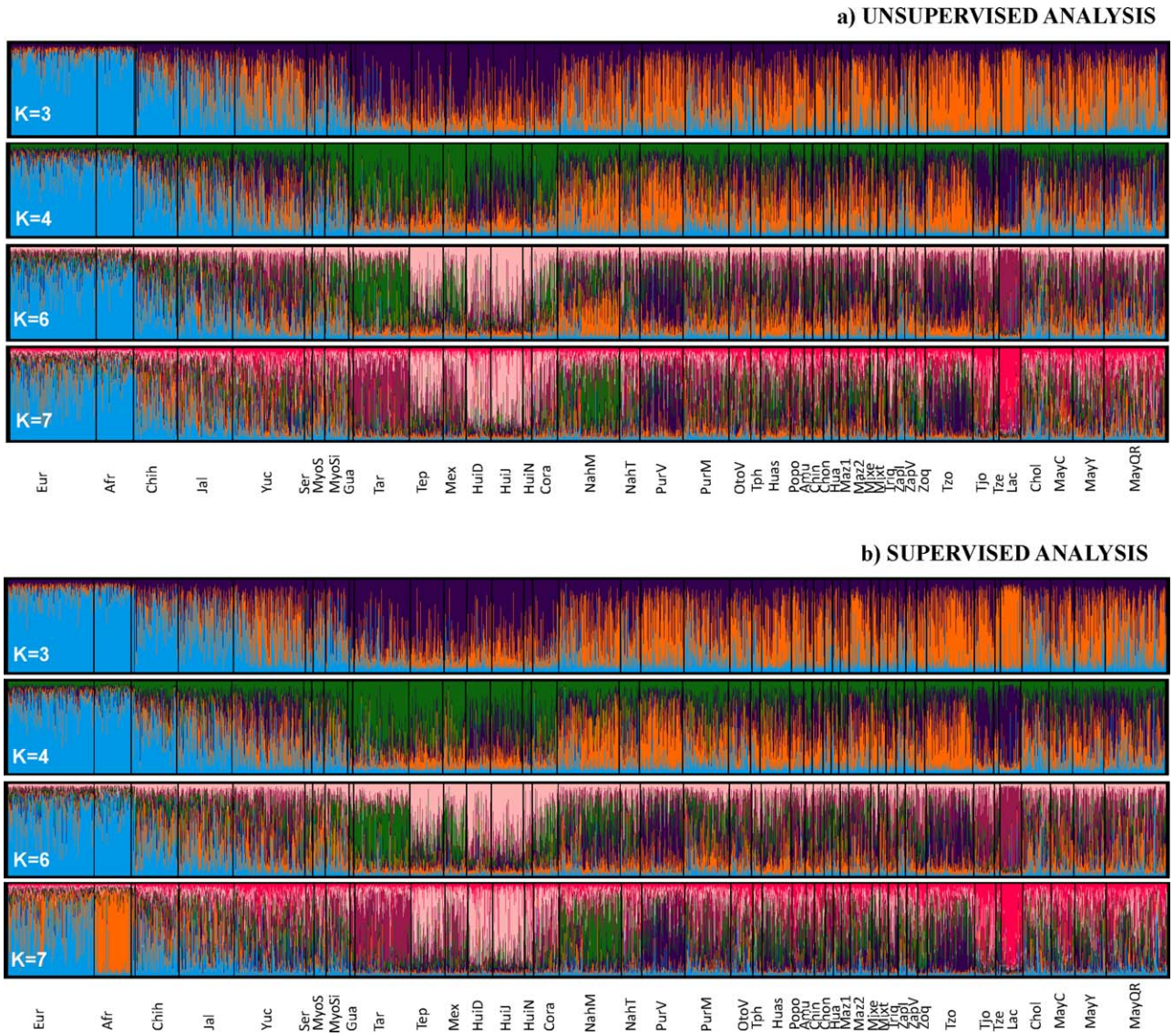

Fig. 6. (**a**) Unsupervised and (**b**) supervised structure analyses with the complete population sample. Structure plots from K = 3 to K = 7 are presented. $K = 5$ plot was omitted because gives the same information than K = 4. For abbreviations please check Table 1.

Native groups (Fig. 2). This value relates to their highest Nm values (Fig. 5), lack of barriers (Fig. 1), elevated He (Table 2), close genetic similarity (Fig. 4), and the lowest Fst value in the Southeast region (omitting Lacandones and Tojolobales) (Supporting Information Table S4) that explain the similarity between groups linked to the Mayan language (Supporting Information Table S5). This panorama is consistent with historical records because Mayas are described as the largest group that inhabited Mesoamerica with at least 3,500 to 4,000 years old, covering roughly a half of Mesoamerica by the time of the Spanish conquest, from the eastern border of the Isthmus of Tehuantepec in Mexico to Honduras and El Salvador, and whose descendants now represent 7.5 million of people who speak more than 28 Mayan languages (Ibarra-Rivera et al., 2008). Altogether, our results support the theories of extensive gene flow and trade throughout the Maya empire (Sharer and Traxler, 2006), as previously claimed by Ibarra-Rivera et al. (2008). However, this work improves the Pre-Hispanic perspective of this conclusion into a broader context throughout the Mexican territory.

On the other hand, bottleneck signals were detected in Peninsular Mayas and Chol groups (Table 2). Among the relevant historical background about this finding, it should be recalled that during the Pre-Classic era in the Maya empire (400 BC–100 AD) emerged the like-chiefdom political organization, leading to the birth of a ruling elite and development of city-state government system, which incited rivalries and warfare for control trade routes between the Maya highlands and lowlands (Sharer and Traxler, 2006). In the Classic period, it becomes a dynamic entity that performs a series of expansions and contractions, whereas for the Terminal Classic (AD 800–900) there are massive declines in population size that caused the abandonment of central and southern Maya territories, consistent with the bottleneck detected (Table 2). Explanations for this

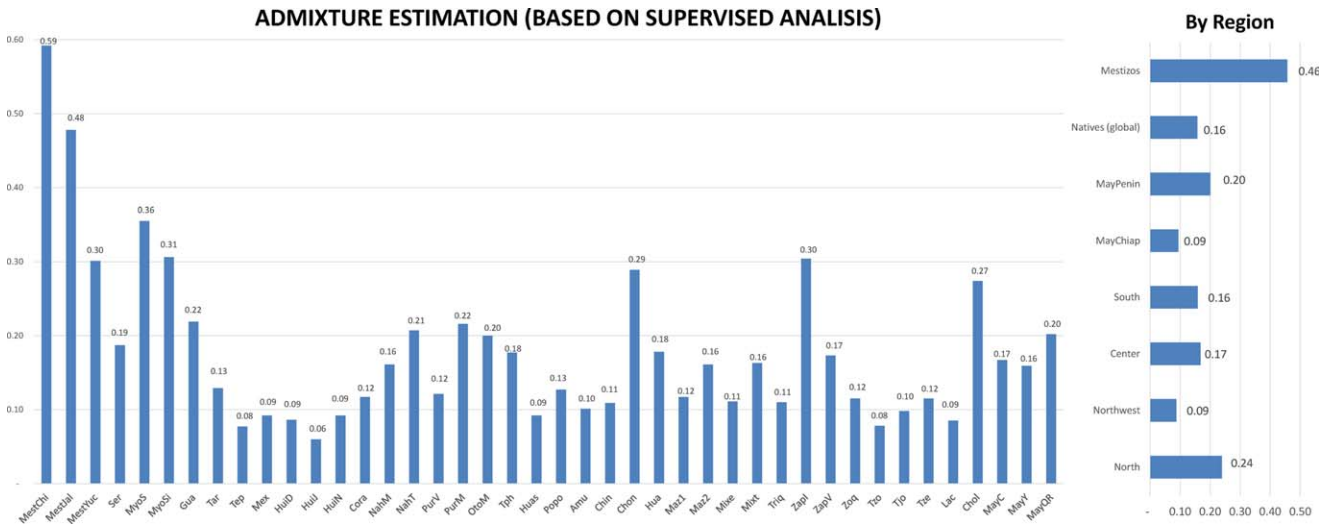

Fig. 7. Admixture proportion (%) in Mexican Mestizos and Native groups (by population and by geographic regions). Mayas from Chiapas (MayChiap) and those from the Yucatan Peninsula (MayPenin) were separated for discussion purposes.

phenomenon involve natural disasters, climate change, famine, conflict escalation, environmental exploitation and ecological collapse, although there is no agreed explanation for this historical fact (Sharer and Traxler, 2006). In brief, different historical events could be involved to explain the bottleneck effects detected in Maya populations.

Southern native groups with high isolation and genetic drift

A number of Mexican Native groups from Chiapas and Oaxaca displayed lower effective population size (Ne) and genetic flow (Nm) (Figs. 2 and 5), presenting higher differentiation than the nearby Native populations from the Mesoamerican core (Figs. 4 and 6). Interestingly, the geographic isolation of these populations in the Mountains of the Sierra Madre from Oaxaca and Chiapas emerges as determinant to explain these characteristics. Although we present a general discussion of these findings, we avoid repeating discussions before accomplished in the previous work with Native groups from Oaxaca (Quinto-Cortés et al., 2010), source of the STR dataset included herein.

Lacandones and Tojolabales. As below described, these groups displayed the highest isolation and differentiation level in Mesoamerica that result in a particular population structure level (Fig. 6). This result seems involving a common genetic origin supported by their shared classification into the Maya-Chiapaneco linguistic group (Supporting Information Fig. S1). However, a scarce interaction between these groups has been described, and even different geographic origins are implied: Tojolabales from the Cuchumatanes of Guatemala, and Lacandones from the Yucatan Peninsula and the Guatemalan Petén (Erosa-Solana, 1995; Ruz, 1995). Therefore, a common origin would not be the unique (or principal) explanation to their relative genetic similarity. In addition to their Mayan ancestry, their similar social organization system based on kinship to form family groups (clans) probably has been determinant to result in a shared extreme differentiation degree (Fig. 6); although the highest degree of geographic dispersion (isolation), including polygyny and inbreeding, could have accentuated the genetic differentiation in Lacandones (Erosa-Solana, 1995; Ruz, 1995).

Tzotziles and Tzeltales. Linguistically they are related into the Maya-Chiapaneco group (Supporting Information Fig. S1); they live closely in the highlands of Chiapas and presented positive bottleneck signals (Table 2). Their shared history is characterized by rebellions and insurrections, such as the *castas war*, a social movement that began in 1847 against the Mexican Creoles and Mestizos in the Yucatan Peninsula, which cost about a quarter of a million of lives, ending officially in 1901 with the occupation of the Maya capital of Chan Santa Cruz by troops of the Mexican federal army (Duch, 1998). Spaniard conquers submitted Tzotziles and Tzeltales by means of reductions and indigenous concentrations in "encomiendas" or communities where they were forced to pay tribute. In fact, at the beginning of the XVIII century these indigenous communities were on the edge of economic and demographic collapse, which eventually originated the first rebellion in the region (Robledo-Hernández, 1995). In brief, bottleneck signals detected in Tzotziles and Tzeltales are historically justified.

Zoques and Mixes from Oaxaca. Some historical descriptions link to these Native groups as Olmec descendants, and linguistically as members of the Popoluca-Mixe-Zoque group within the Maya-macro group (Supporting Information Fig. S1). In addition to the reduced effective population size (Ne) and elevated genetic differentiation (Figs. (1 and 2), and 4), bottleneck also was detected in both groups (Table 2). In Zoques, bottleneck could be related to the conquest and reduction history of the Zoque's territory, by acculturation, disease, and forced labor; including a massive relocation of 12,000 individuals due to the eruption of the Chichonal volcano that has increased significantly its geographical dispersion (Uriel del Carpio, 1995). In contrast, in

Mixes their history is characterized by conflicts to defend their land and autonomy; in fact, they never were conquered by their enemies, including Zapotecs, Aztecs, Mixtecs or Spaniards. For this reason, the strategy taken by the Spaniards to conquer Mixes during the middle of the century XVI was different, they penetrated through religion; the evangelization process was slow, creating a religious syncretism between "Christian" and "pagan" that is still observable (Reyes-Gómez, 1995). Probably, their geographic isolation in highlands of the "Sierra Mixe" and the Isthmus of Tehuantepec has been crucial for the genetic differentiation and isolation processes, besides the aforesaid political and cultural protection. To explain the bottleneck signals detected, recent demographic declines were described in Mixes during 1921 to 1930, and again in 1950 (Valdés, 1988); although the precise causes are unknown, serious interethnic conflicts that happened during those periods have been implied (Reyes-Gómez, 1995).

Triquis. This Mexican Native group from Oaxaca presented a reduced effective population size and high genetic differentiation (Figs. (1 and 2), and 4), as well as positive bottleneck signal (Table 2). These characteristics probably are related to their linguistic and cultural isolation, probably helped by their geographic isolation. In fact, they have been described as an "island" surrounded by the high and low Mixteca region, but not inhabited by Mixtecos, Mestizos (admixed), or another Native groups. This isolation process seems to be reinforced by communal land tenure and social organization with unilineal descent groups, forming both outbred and inbred clans (Huerta Ríos, 1995).

Genetic findings in the central region of Mesoamerica

The Mexican Native groups from the central region presented the largest number of private alleles (Fig. 3), elevated effective population size (Fig. 2), gene flow (Fig. 5), and relatively close genetic relationships (Fig. 4), similar to different non-isolated groups from the South and Southeast of Mexico characterized by weak genetic barriers (Fig. 1). Based on these characteristics and geographic location, we formed two central population clusters to perform the discussion: 1) Center-West, closer to the Pacific Ocean, which includes to Purépechas and Nahuas; and 2) Center, closer to the Gulf of Mexico that comprises to Huastecos, Tepehuas, and Otomies (Fig. 1). Again, we avoided, as much as possible, to discuss the same anthropological findings observed in the previous papers where the STR data were originally reported (Table 1).

Purépechas and Nahuas. The above described characteristics in Purépechas could imply diverse historical facts: i) multiethnic origin, because records indicate that they result from admixture at the ending of the century XII between different Native groups, including Chichimecas, Nahuas, and Pre-Tarascos groups (Argueta Villamar, 1995). ii) Social development level, which explains why Purépechas became the second most largest Prehispanic state in Mesoamerica at the moment of the Spanish Conquest, being one of the most important societies of the Post-Classic period like opponents of the Mexicas, who never were able that Purépechas pay tribute (Argueta Villamar, 1995). Another interesting finding was the bottleneck detected in Purépechas-Mountain (Table 2) that is in agreement with the population decrease detected in this group, because in the year of 1,500 a total of 200,000 inhabitants were calculated in the Diocesis of Michoacán area, which reduced to ~92,000 due to the war, genocide, diseases, and migration, principally. Eventually in 1920 and 1930 around 33,600 Purépecha inhabitants were censed in Michoacán (Argueta Villamar, 1995). Similarly, Nahuas (together with Mayas) are descendants from large Mesomerican civilizations, which correlates with relative large gene flow (Fig. 5), effective population size (Fig. 2), and low genetic differentiation (weak barriers) (Fig. 1), as detected in these central populations. In fact, these characteristics are concordant with the smaller proportions of runs of homozigosity (ROH) detected in Nahuas in the previous genomic analysis (Moreno-Estrada et al. 2014), supporting our interpretation. In addition, a main genetic finding is the bottleneck effect detected in Nahuas (Table 2), which can be explained by the bloody conquest process carried out in this area by the Spanish Conquer Nuño de Guzmán. This process originated many complains that the Spanish Crown decided to prosecute and convict him, taking the provincial government and sending back Nuño de Guzman to Spain as prisoner, where he died imprisoned in the castle of Torrejón de Velasco in 1544 (Riva-Palacio, 1991).

Otomi, Tepehuas, and Huastecos. In addition to the relatively close genetic relationship among these central Native groups from the state of Hidalgo, Tepehuas and Huastecos showed bottleneck effects (Table 2). In Tepehua this could be related to their demographic size (roughly 9,500 people) and the high migration rate to the USA and rapid acculturation process (Hernández-Montes and Heiras-Rodríguez, 2004). Similarly, this fact explains why they present the lowest number of private alleles in the central region (Supporting Information Fig. S3) and reduced gene flow (Fig. 5). In Huastecos, historical descriptions indicate that, because of its geographic location as output way to the sea, they were most affected in the early years after the conquest because they were the subject of numerous abuses of the Spaniards. From 1523 to 1532 they were subject to slavery to supply labor to the West Indies. In those 10 years two indigenous rebellions happened, resulting 400 "principal" individuals and 60 "caciques" of indigenous origin murdered. From the 16th century the most notable changes were the demographic decline and dispossession of land to Huasteca communities (CDI, 2009a). In Otomies, despite the historical descriptions of the impact of diseases such as smallpox in this group, not bottleneck was observed. The precedent that could reconcile this finding is that Otomies were valuable contributors to the Spaniards, helping to evangelizing Chichimeca peoples, which could have helped to avoid major demographic declines in Otomies, as observed throughout Mesoamerica during and after the conquest (Vázquez-Valdivia, 1995).

Distinctive characteristics of Mexican native groups from the north and northwest

In general, these two regions were characterized by reduced effective population size (Fig. 2), low genetic

diversity (Table 1) and gene flow (Fig. 5), which correlate with higher differentiation and population structure among the Mesomerican populations (Figs. 4a and 6). This scenario is concordant with stronger genetic drift effects originating the widest barrier to human gene flow in this Pre-Hispanic territory that separates north and northwest populations from the rest of Mesoamerica (Fig. 1). This effect probably is helped by geographic barriers consisting of mountains and canyons of the Sierra Madre Occidental where these individuals live. In addition, the dry climatic conditions in great part of these regions cause the habitat become less suitable for agriculture, which explains the lower population density facilitating genetic drift differentiation of these groups (Cavalli-Sforza et al., 1994). Genetic drift effects can be seen as longer branches for these groups in the NJ tree (Fig. 4a), particular genetic structure levels (Fig. 6a), and their peripheral location in the MDS plot (unshown plot). Conversely, the central plateau of Mexico enjoyed a temperate climate that favored the early development of agriculture and training of Pre-Hispanic empires (Fiedel, 1992; López-Austin and López-Lujan, 2001).

The north of Mesomerica

The STRUCTURE results allow clustering the northern Mexican Native groups: Seri and Mayo regarding Guarijio and Tarahumara (Fig. 6; $K = 7$). At the same time, Seri and Guarijio were characterized by a reduced genetic diversity (Table 2), effective population size (Fig. 2), and gene flow (Fig. 5) represented by strong genetic barriers only comparable with those separating to Lacandones and Tojolabales (Fig. 1). Conversely, Mayos displayed the largest proportion of admixture among all the Mexican Native groups (mean 33.1%; Fig. 7), which also was supported by their increased genetic diversity levels (Table 2), and descriptions indicating their daily contact with Mestizos and easy access to their communities (Aguilar-Zeleny, 1995b). Interestingly, this value is even higher than the observed in Mestizos from Yucatan (30.1%). As a result, the nonindigenous admixture level in the North was the largest among all the regions of Mesoamerican regions considered in this work (23.9%) (Fig. 7), which is expected considering some historical facts: i) the low population density of the northern Mexican territory since Pre-Hispanic times (Nárez, 2000); ii) the demographic decrease of Native population due to overwork, warfare, and epidemic diseases, principally (Serrano-Sánchez, 1995); and iii) the economic development brought by the Spaniards who formed large cities and implemented productive activities (e.g. mining), which increased simultaneously the population density and admixture, as has been described in previous studies in Mexican Mestizos (Rubi-Castellanos et al., 2009; Martínez-Cortés et al., 2012).

Seris. They presented one of the highest genetic differentiation levels among the studied Mexican Native groups that suggests genetic drift effects and relative isolation by restricted gene flow (Figs. 1 and 4). This differentiation also is consistent with their linguistic classification into the Macro-Yuma group, distinct to the Native neighbors that belong to the Macro-Nahua group (Uto-Aztecan) (Supporting Information Fig. S1). However, this result also involves to their nomadic lifestyle without religious-farming culture, where bosses or authority are not recognize, hampering their integration after the Spanish conquest. This led to the near annihilation of the Seri group that gradually was confined to the most inhospitable part of its territory (Pérez-Ruíz, 1995), which is in agreement with the bottleneck signal detected herein (Table 2).

Guarijios and Mayos. The close genetic relationship between Guarijios and Mayos observed in the NJ tree (Fig. 4a) is in agreement with the geographic origin of the Guarijio's population sample (Sonora state), and their close linguistic relationship into the Taracahita family of the Uto-Aztecan group (Supporting Information Fig. S1). Interestingly, despite the above described high non-Native admixture level detected in Mayos (Fig. 7), the genetic, linguistic, and historical congruence of these results suggests that admixture does not alter significantly our inferences concerning the studied Mexican Native groups.

Guarijios are described as a link between Tarahumaras from Chihuahua and Mayos from Sonora, because in both states there are Guarijíos (Aguilar-Zeleny, 1995a). Particularly, the close—and probable ancestral—relationship between Tarahumaras and Guarijios was confirmed by a common genetic structure component observed in Figure 6 ($K = 7$). The reduced effective population size detected in Guarijios (Fig. 2) suggests a greater differentiation by genetic drift aided by its geographic-isolation (Fig. 5). This fact could be correlated with the historical record scarcity implying a low representation of Guarijios over other Mexican Native groups from the North region. Interestingly, the bottleneck detected herein in Guarijios (Table 2) possibly is related to the fission that they suffered in colonial times after a severe Spanish repression causing that part of the population were displaced from the Chihuahua state towards the Sonora mountains (Aguilar-Zeleny, 1995a).

The northwest of Mesomerica

One peculiar feature of the Northwest region is the lowest proportion of nonindigenous admixture and populations where bottleneck was detected ($1/6 = 16.67\%$) respect to the rest of Mesoamerica (>60%) (Table 2). Their geographic isolation and the lack of economic interest for their territories located into inaccessible mountains and canyons seems to have been a significant protective factor to be conquered and engaged in political conflicts, wars, and epidemics suffered throughout Mesoamerica after the Spanish contact. In agreement with this explanation, the most isolated Native groups from Chiapas (Lacandones and Tojolabales) neither presented bottleneck signals (Table 2) and showed low nonindigenous admixture level (8.5 and 9.8%, respectively), inferior to the global admixture level estimated in Mexican Native groups (15.7%) (Fig. 7).

Tarahumara, Cora, and Mexicanero. A relative close genetic relationship was observed among populations from the North (Tarahumaras) and Northwest (Coras and Mexicaneros) (Fig. 4a), in addition to a common genetic component that is clearly appreciated in green in the Structure supervised analysis (Fig. 6; $K = 6$), which is more frequent in Tarahumaras. Therefore, for discussion purposes this genetic component will be called Tarahumara ancestry. Although linguistically the

Tarahumara, Cora and Mexicanero groups are classified into the same Uto-Aztecan group, they belong to different families (Supporting Information Fig. S1): Taracahita, Corachol and Nahuatl, respectively (Aguilar-Zeleny, 1995a,b; CDI, 2009c). Consequently, based on these linguistic criteria, a closer genetic relationship would be expected among Tarahumara, Guarijío, and Mayos, and between Huichol and Cora groups, as partially is observed in Figures 4a and 6. On the other hand, Mexicaneros geographically are closer to Tepehuanos (Fig. 1), but under linguistic criteria they are closer to Nahuas (Supporting Information Fig. S1). Although previous studies and recent historical records did not offer an obvious explanation to these findings, the consistence of our results by genetic distances and Structure analyses (Figs. 4a and 6) allow suggesting an ancestral relationship between North and Northwest populations based on the presence of the "Tarahumara" component in the Northwest (Fig. 6). This "Tarahumara founder effect" probably imply a North-West migration process, in agreement with the coastal routes suggested for the America's peopling (Wang et al., 2007; Reich et al., 2012). Despite the large number of wars and rebellions registered in Tarahumaras since the Spanish contact (Heras, 1995), they did not present positive bottleneck signal (Table 2), probably helped by their geographic isolation in mountains and canyons of the Sierra Madre from Chihuahua, like most the Mexican Native groups from the Northwest region.

Huichol and Tepehuanos. The observed genetic relationship between Huichol and Tepehuano groups (Figs. 4a and 6) is consistent with their linguistic affiliation (Supporting Information Fig. S1), and with their geographic proximity into the canyons and mountains of the Sierra Madre Occidental (Fig. 1). Interestingly, only the Huichol of Nayarit group showed positive bottleneck signal throughout the Northwest region. Although we were not able to obtain historic records concerning this specific Huichol group, we found that the southwestward passage of the Spanish conquer Nuño de Guzman was through Nayarit, leaving a destruction trail that could have affected specifically this Huichol group (CDI, 2009b), explaining the bottleneck signal detected herein.

Correlation between genetic relationships and other aspects

Despite the general congruence between the genetic relationships with cultural, linguistic, and geographic criteria, the formal evaluation of these hypotheses was not significant in most of cases (Supporting Information Table S5). Probably this is due to the high genetic differentiation of the Mexican Native groups described in previous studies based on different genetic systems (Wang et al., 2007; Sandoval et al., 2009, 2012; Reich et al., 2012; Moreno-Estrada et al., 2014). As previously noted, the main exception involves to the Mayan groups (Supporting Information Table S5), which is in agreement with their previously stated larger similarity between themselves (Ibarra-Rivera et al., 2008; Moreno-Estrada et al. 2014), and relative lack of differentiation between Mesoamerican populations (Wang et al., 2007). Finally, the correlation between genetic and geographic distances in Mesoamerica only was significant at large distances >1,500 km (Supporting Information Fig. S3), as previously found in Mexican Native groups analyzed with Y-linked loci (Rangel-Villalobos et al., 2008).

CONCLUSIONS

In brief, the analysis of 39 Mexican Native population samples based on 15 STRs showed the following characteristics: i) a low genetic differentiation level an elevated gene flow in Mayan groups from the Yucatan Peninsula and central groups; pattern interrupted by the presence of more isolated—and differentiated—groups from Chiapas and Oaxaca; ii) The Native groups from North and Northwest regions had the largest genetic differentiation levels, setting the widest genetic gene flow barrier in Mesoamerica; iii) the IBD model in Mexican Native groups only adjusted at long distances (>1,500 km); iv) geographic isolation stands as a determining factor to avoid either nonindigenous admixture and population bottleneck processes; v) despite the congruence between the estimated genetic relationships with cultural, linguistic, geographic criteria, most of these do not explain the observed population structure, excepting by the Mayan groups linguistically related that show a relative homogeneity.

ACKNOWLEGMENTS

The authors thank the indigenous volunteers included in this study, and to the authors who kindly shared STR datasets of Mexican Native groups, making possible to carry out the performed analyses. The authors thank the technical advice of Joel Salazar-Flores for using the Structure software at the beginning of this project.

LITERATURE CITED

Aguilar-Zeleny A. 1995a. Guarijíos, en: Etnografía contemporánea de los pueblos indígenas de México (Región Norte). México, DF: Instituto Nacional Indigenista (INI). p 13–51.

Aguilar-Zeleny A. 1995b. Mayos, en: Etnografía contemporánea de los pueblos indígenas de México (Región Norte). México, DF: Instituto Nacional Indigenista (INI). p 83–129.

Alves C, Gusmão L, López-Parra AM, Soledad Mesa M Amorim A, Arroyo-Pardo 2005. STR allelic frequencies for an African population sample (Equatorial Guinea) using AmpF*l*STR Identifiler 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). México, DF: Instituto Nacional Indigenista (INI).

Barrot C, Sánchez C, Ortega M, González-Martín A, Brand-Casadevall C, Gorostiza A, Huguet E, Corbella J, Gené M. 2005. Characterization of three Amerindian populations from Hidalgo State (Mexico) by 15 STR-PCR polymorphisms. Int J Legal Med 119:111–115.

Bosch E, Calafell F, Pérez-Lezaun A, Clarimón J, Comas D, Mateu E, Martínez-Arias R, Morera B, Brakez Z, Akhayat O, Sefiani A, Hariti G, Cambon-Thomsen A, Bertranpetit J. 2000. Genetic structure of north-west Africa revealed by STR analysis. Eur J Hum Genet 8:360–366.

Bonnet E, Van de Peer Y. 2002. zt: a software tool for simple and partial Mantel tests. J Stat Softw 7:1–12.

Butler JM. 2014. Advanced topics in forensic DNA typing: interpretation. London UK: Elsevier Academic Press.

Calafell F, Shuster A, Speed WC, Kidd JR, Kid KK. 1998. Short tandem repeat polymorphism evolution in humans. Eur J Hum Genet 6:38–49.

Cavalli-Sforza, Menozzi P, Piazza A. 1994. The history and geography of human genes. New Jersey, UK: Princenton University Press.

Comisión Nacional para el Desarrollo de los pueblos indígenas (CDI). 2009a. Huastecos de san luis potosi– Teenek (monografías) CDI (Ed.) Mexico DF. Available at: http://www.cdi.gob.mx/index.php?option=com_content&view=article&id=593:huastecos-de-san-luis-potosi-teenek&catid=54:monografias-de-los-pueblos-indigenas& Itemid=62. Last accessed August 16, 2015.

Comisión Nacional para el Desarrollo de los pueblos indígenas (CDI). 2009b. Huicholes - Wirraritari o Wirrárika. CDI (Ed.) Mexico DF. Available at: http://www.cdi.gob.mx/index.php?option=com_content&view=article&id=596:huicholes-wirraritari-o-wirrarika-&catid=54:monografias-de-los-pueblos-indigenas& Itemid=62. Last accessed August 16, 2015.

Comisión Nacional para el Desarrollo de los pueblos indígenas (CDI). 2009c. Mexicaneros. CDI (Ed.) Mexico DF. Available at: http://www.cdi.gob.mx/index.php?option=com_content&view=article&id=621:mexicaneros&catid=54:monografias-de-los-pueblos-indigenas& Itemid=62. Last accessed August 16, 2015.

Comisión Nacional para el Desarrollo de los pueblos indígenas. 2012. Los pueblos indígenas de México. CDI (Ed.) Mexico DF. Available at: http://www.cdi.gob.mx/index.php?option=com_content&view=article& id=1387&Itemid=24. Last accessed August 16, 2015.

Cisneros IH. 2004. Situación de los pueblos indígenas en México. In: Derechos humanos de los pueblos indígenas en México. México: Comisión de los Derechos Humanos del Distrito Federal (CDHDF). p 9.

Coe MD, Kootz R. 2002. Mexico: from the Olmec to the Aztecs. New York: Thames & Hudson.

Cornuet JM, Luikart G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001–2014.

Coudray C, Calderon R, Guitard E, Ambrosio B, González-Martín A, Dugoujon JM. 2007. Allele frequencies of 15 tetrameric short tandem repeats (STRs) in Andalusians from Huelva (Spain). Forensic Sci Int 168:1–4.

Duch J. 1998. Yucatán en el tiempo: enciclopedia alfabética. Merida, México: Yucatán Inversores Cares (Ed.).

Dupanloup I, Schneider S, Excoffier L. 2002. A simulated annealing approach to define the genetic structure of populations. Mol Ecol 11:2571–2581.

Earl DA, vonHoldt BM. 2012. Structure harvester: a website and program for visualizing structure output and implementing the Evanno method. Conservation Genet Resour 4:359–361.

Erosa-Solana E. 1995. Lacandones, en: Etnografía contemporánea de los pueblos indígenas de México (Región Sureste). México, DF: Instituto Nacional Indigenista (INI).

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50.

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587.

Falush D, Stephens M, Pritchard JK. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7:574–578.

Felsenstein J. 1989. PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics 5: 164–166.

Fiedel SJ. 1992. Prehistory of the Americas, 2nd ed. UK: Cambridge University Press.

Friedlaender JS, Friedlaender FR, Reed FA, Kidd KK, Kidd JR, Chambers GK, Lea RA, Loo JH, Koki G, Hodgson JA, Merriweter DA, Weber JL. 2008. The genetic structure of pacific islanders. PLoS Genet 4:e19.

González-Martín A, Gorostiza A, Rangel-Villalobos H, Acunha V, Barrot C, Sánchez C, Ortega M, Gené M, Calderón R. 2008. Analyzing the genetic structure of the Tepehua in relation to other neighbouring Mesoamerican populations. A study based on allele frequencies of STR markers. Am J Hum Biol 20:605–613.

González-Martín A, Gorostiza A, Regalado-Liu L, Arroyo-Peña S, Tirado S, Nuño-Arana I, Rubi-Castellanos R, Sandoval K, Coble MD, Rangel-Villalobos H. 2015. Demographic history of indigenous populations in Mesoamerica based on mtDNA sequence data. PLos One 10:e0131791.

Heras M. 1995. Los Tarahumaras, en: Etnografía contemporánea de los pueblos indígenas de México (Región Norte). México, DF: Instituto Nacional Indigenista (INI). p 405–479.

Hernández Montes M, Heiras Rodríguez CG. 2004. Tepehuas. Pueblos indígenas del México contemporáneo. México: CONADEPI.

Huerta Ríos C. 1995. Los Triquis, en: Etnografía contemporánea de los pueblos indígenas de México (Región Pacífico Sur). México, DF: Instituto Nacional Indigenista (INI).

Ibarra-Rivera L, Mirabal S, Regueiro MM, Herrera RJ. 2008. Delineating genetic relationships among the Maya. Am J Phys Anthropol 135:329–347.

INEGI. 2005. National Institute of Statistics, Geography, and Informatics. INEGI (Ed), Mexico, DF. Available at: http://www.inegi.gob.mx. Last accessed August 16, 2015.

Jensen JL, Bohonak AJ, Kelley ST. 2005. Isolation by distance, web service. BMC Genet 6:13. Available at: http://ibdws.sdsu.edu/. Last accessed August 16, 2015.

Jobling MA, Hurles ME, Tyler-Smith C. 2004. Human evolutionary genetics: origins, peoples and disease. New York: Garland Science.

Kraaijenbrink T, van der Gaag KJ, Zuniga SB, Xue Y, Carvalho-Silva DR, Tyler-Smith C, Jobling MA, Parkin EJ, Su B, Shi H, Xiao CJ, Tang WR, Kashyap VK, Trivedi R, Sitalaximi T, Banerjee J, Karma Tshering of G, Tuladhar NM, Opgenort JR, van Driem GL, Barbujani G, de Knijff P. 2014. A linguistically informed autosomal STR survey of human populations residing in the greater Himalayan region. PLoS One 9:e91534.

Lópes V, Serra A, Gamero J, Sampaio L, Balsa F, Oliveira C, Batista L, Corte-Real F, Vieira DN, Vide MC, Anjos MJ, Carvalho M. 2009. Allelic frequency distribution of 17 STRs from Identifiler and PowerPlex-16 in Central Portugal area and the Azores archipelago. Forensic Sci Int Genet 4:e1–e7.

López-Austin A, López-Lujan L. 2001. El pasado indígena. Fondo de cultura económica, El Colegio de México. México, DF: Fideicomiso Historia de las Américas.

Manni F, Guérard E, Heyer E. 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by "Monmonier's algorithm". Hum Biol 76:173–190.

Martínez-Cortés G, Nuño-Arana I, Rubi-Castellanos R, Vilchis-Dorantes G, Luna-Vázquez A, Coral-Vázquez RM, Canto-Cetina T, Salazar-Flores J, Muñoz-Valle JF, Sandoval-Mendoza K, López Z, Gamero-Lucas JJ, Rangel-Villalobos H. 2010. Origin and genetic differentiation of three Native Mexican groups (Purepechas, Triquis and Mayas): contribution of CODIS-STRs to the history of human populations of Mesoamerica. Anns Hum Biol 37:801–819.

Martínez-Cortés G, Salazar-Flores J, Fernández-Rodríguez LG, Rubi-Castellanos R, Rodríguez-Loya C, Velarde-Félix JS, Muñoz-Valle JF, 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-González LJ, Martínez-Espin E, Fernández-Rosado F, Moguel MA, Entrala C, Alvarez JC, Lorente JA, Budowle B. 2005. Mexican population data on fifteen STR loci (Identifiler kit) in a Chihuahua (North Central Mexico) sample. J Forensic Sci 50:236–238.

Moreno-Estrada A, Gignoux CR, Fernández-López JC, Zakharia F, Sikora M, Contreras AV, Acuña-Alonzo V, Sandoval K, Eng C, Romero-Hidalgo S, Ortiz-Tello P, Robles V, Kenny EE, Nuño-Arana I, Barquera-Lozano R, Macín-Pérez G, Granados-Arriola J, Huntsman S, Galanter JM, Via M, Ford JG, Chapela R, Rodriguez-Cintron W, Rodríguez-Santana JR, Romieu I, Sienra-Monge JJ, del Rio Navarro B, London SJ, Ruiz-Linares A, Garcia-Herrera R, Estrada K, Hidalgo-Miranda A, Jimenez-Sánchez G, Carnevale A, Soberón X, Canizales-Quinteros S, Rangel-Villalobos H, Silva-Zolezzi I, Burchard EG, Bustamante CD. 2014. The genetics of Mexico recapitulates Native American substructure and impacts biomedical traits. Science 344:1280–1285.

Nárez, J. 2000. Aridamérica y Oasisamérica. In: Manzanilla L, López-Luján L, editors. Historia Antigua de México, Vol. I: El Mexico Antiguo, sus áreas culturales, los orígenes y el horizonte Preclásico. México, DF: Instituto Nacional de Antropología e Historia (INAH), Coordinación de Humanidades del Instituto de Investigaciones Antropológicas de la UNAM y Miguel Angel Porrua. p 121–157

Page RD. 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12: 357–358.

Peakall R, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295.

Peakall R, Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. Bioinformatics 28:2537–2539.

Pérez-Lezaun A, Calafell F, Mateu E, Comas D, Bosch E, Bertranpetit J. 1997a. Allele frequencies for 20 microsatellites in a worldwide population study. Hum Hered 47:187–186.

Pérez-Lezaun A, Calafell F, Mateu E, Comas D, Ruiz-Pacheco R, Bertranpetit J. 1997b. Microsatellite variation and the differentiation of modern humans. Hum Genet 99:1–7.

Pérez-Ruíz ML. 1995. Seris. In: Etnografía contemporánea de los pueblos indígenas de México (Región Noroeste). México, DF: Instituto Nacional Indigenista (INI). p 367–401

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure from multilocus genotype data. Genetics 155: 945–959.

Quinto-Cortés CD, Arriola LA, García-Hughes G, García-López R, Molina DP, Flores M, Palacios R, Piñero D. 2010. Genetic characterization of indigenous peoples from Oaxaca, Mexico, and its relation to linguistic and geographic isolation. Hum Biol 82:409–432.

Ramachandran S, Deshpande O, Roseman CC, Rosenberg NA, Feldman MW, Cavalli-Sforza LL. 2005. Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. Proc Natl Acad Sci USA 102:15942–15947.

Rangel-Villalobos H, Muñoz-Valle JF, González-Martín A, Gorostiza A, Magaña MT, Páez-Riberos LA. 2008. Genetic admixture, relatedness, and structure patterns among Mexican populations revealed by the Y-chromosome. Am J Phys Anthropol 135:448–461.

Rangel-Villalobos H, Martínez-Sevilla VM, Salazar-Flores J, Martínez-Cortez G, Muñoz-Valle JF, Galaviz-Hernández C, Lazalde-Ramos BP, Sosa-Macías M. 2013. Forensic parameters for 15 STRs in eight Amerindian populations from the north and west of Mexico. Forensic Sci Int Genet 7:e62–e65.

Rangel-Villalobos H, Muñoz-Rivas CD, Martínez-Sevilla VM, Nuño-Arana I, Rubi-Castellanos R, Martínez-Cortés G. 2014. Forensic evaluation of the AmpF*l*STR Identifiler kit in nine Mexican Native populations from the Pre-Columbian Mesoamerican region. Int J Legal Med 128:467–468.

Reich D, Patterson N, Campbell D, Tandon A, Mazieres S, Ray N, Parra MV, Rojas W, Duque C, Mesa N, García LF, Triana O, Blair S, Maestre A, Dib JC, Bravi CM, Bailliet G, Corach D, Hünemeier T, Bortolini MC, Salzano FM, Petzl-Erler ML, Acuña-Alonzo V, Aguilar-Salinas C, Canizales-Quinteros S, Tusié-Luna T, Riba L, Rodríguez-Cruz M, López-Alarcón M, Coral-Vázquez R, Canto-Cetina T, Silva-Zolezzi I, Fernández-Lopez JC, Contreras AV, Jímenez-Sánchez G, Gómez-Vázquez MJ, Molina J, Carracedo A, Salas A, Gallo C, Poletti G, Witonsky DB, Alkorta-Aranburu G, Sukernik RI, Osipova L, Fedorova SA, Vasquez R, Villena M, Moreau C, Barrantes R, Pauls D, Excoffier L, Bedoya G, Rothhammer F, Dugoujon JM, Larrouy G, Klitz W, Labuda D, Kidd J, Kidd K, Di Rienzo A, Freimer NB, Price AL, Ruiz-Linares A. 2012. Reconstructing Native American population history. Nature 488:370–374.

Reyes-Gómez L. 1995. Mixes. In: Etnografía contemporánea de los pueblos indígenas de México (Región Transísmica). México, DF: Instituto Nacional Indigenista (INI). p 169–207.

Riva-Palacio V. 1991 (1880). México a través de los siglos, Vol. II. España: Ballescá y Cía, Editores, Barcelona.

Robledo-Hernández G. 1995. Tzotziles y Tzeltales. In: Etnografía contemporánea de los pueblos indígenas de México (Región Sureste). México, DF: Instituto Nacional Indigenista (INI). p 187–216.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation-by-distance. Genetics 145:1219–1228.

Rubi-Castellanos R, Anaya-Palafox M, Mena-Rojas E, Bautista-España D, Muñoz-Valle JF, Rangel-Villalobos H. 2009. Genetic data of 15 autosomal STRs (Identifiler kit) of three Mexican Mestizo population samples from the states of Jalisco (West), Puebla (Center), and Yucatan (Southeast). Forensic Sci Int Genet 3:e71–e76.

Rubi-Castellanos R, Martínez-Cortés G, Muñoz-Valle JF, González-Martín A, Cerda-Flores R, 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 135:448–461.

Ruz MH (1995) Tojolabales. In: Etnografía contemporánea de los pueblos indígenas de México (Región Sureste). México, DF: Instituto Nacional Indigenista (INI). p 129–181.

Sahoo S, Kashyap VK. 2005. Influence of language and ancestry on genetic structure of contiguous populations: a microsatellite based study on populations of Orissa. BMC Genet 6:4.

Salazar-Flores J, Zuñiga-Chiquette F, Rubi-Castellanos R, Álvarez-Miranda JL, Zetina-Hérnandez A, Martínez-Sevilla VM, González-Andrade F, Corach D, Vullo C, Álvarez JC, Lorente JA, Sánchez-Diz P, Herrera RJ, Cerda-Flores RM, Muñoz-Valle JF, Rangel-Villalobos H. 2015. Admixture and genetic relationships of Mexican Mestizos regarding Latin American and Caribbean populations based on 13 CODIS-STRs. Homo 66:44–59.

Sánchez C, Barrot C, Ortega M, González-Martin A, Gorostiza A, Corbella J, Huguet E, Gené M. 2005. Genetic diversity of 15 STRs in Choles from Northeast of Chiapas (Mexico). J Forensic Sci 50:1–3.

Sandoval K, Buentello-Malo L, Peñaloza-Espinosa R, Avelino H, Salas A, Calafell F, Comas D. 2009. Linguistic and maternal genetic diversity are not correlated in Native Mexicans. Hum Genet 126:521–531.

Sandoval K, Moreno-Estrada A, Mendizabal I, Underhill PA, López-Valenzuela M, Peñaloza-Espinosa R, López-López M, Buentello-Malo L, 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.

Serrano-Sánchez, C. 1995. 500 años de historia: La conquista y el mestizaje biológico en México. In: Ochoa L, editor. Conquista, transculturación y mestizaje, raíz y origen de México. México DF: Instituto de Investigaciones Antropológicas, UNAM. p 37–45.

Sharer RJ, Traxler LP. 2006. The ancient Maya. Stanford, CA: Stanford University Press.

Shriver MD, Jin L, Boerwinkle E, Deka R, Ferrell RE, Chakraborty R. 1995. A novel measure of genetic distance for highly polymorphic tandem repeat loci. Mol Biol Evol 12:914–920.

Swadesh M. 1959. Mapas de clasificación lingüística de México y las Américas. Instituto Nacional de Antropología e Historia. Serie antropológica no 8. México, DF: Cuadernos.

Szpiech ZA, Jakobsson M, Rosenberg NA. 2008. ADZE: a rarefaction approach for counting alleles private to combinations of populations. Bioinformatics 24:2498–2504.

Tereba A. 1999. Tools for analysis of population statistics. Profiles in DNA. Promega Corp. Madison, WI.

Uriel del Carpio C. 1995. Zoques de Chiapas. In: Etnografía contemporánea de los pueblos indígenas de México (Región Sureste). México, DF: Instituto Nacional Indigenista (INI).

Vázquez-Valdivia. 1995. Otomíes del Valle de Mezquital, Hidalgo, en: Etnografía contemporánea de los pueblos indígenas de México (Región Centro). México, DF: Instituto Nacional Indigenista (INI). p 217–289.

Valdés LM. 1998. El perfil demográfico de los indios mexicanos. México: Siglo XXI/UNAM/CIESAS. p 19.

Wang S, Lewis CM, Jakobsson M, Ramachandran S, Ray N, Bedoya G, Rojas W, Parra MV, Molina JA, Gallo C, Mazzotti G, Poletti G, Hill K, Hurtado AM, Labuda D, Klitz W, Barrantes R, Bortolini MC, Salzano FM, Petzl-Erler ML, Tsuneto LT, Llop E, Rothhammer F, Excoffier L, Feldman MW, Rosenberg NA, Ruiz-Linares A. 2007. Genetic variation and population structure in Native Americans. PLoS Genet 3: e185.

Weir BS. 1996. Genetic Data Analysis. 2nd ed. Sinauer Associates, Sunderland, Massachusetts.