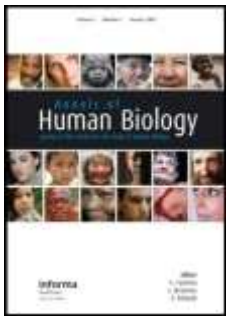

Origin and Genetic Differentiation of Three Mexican Native Groups (Purépechas, Triquis, and Mayas): Contribution of CODIS-STRs to the History of Human Populations of Mesoamerica

Journal:	*Annals of Human Biology*
Manuscript ID:	draft
Manuscript Type:	Research Paper
Keywords:	CODIS-STRs, Native groups, Mexico, Mesoamerica, Amerindians

Abstract

Background: CODIS-STRs have been scarcely analyzed in Native Mexican groups, both for human identification or anthropological purposes.

Aim: To analyze the genetic relationships and population structure among three Native Mexican groups from Mesoamerica.

Subjects and methods: 531 unrelated Native individuals from Mexico were PCR-typed for 15 and 9 autosomal STRs (Identifiler™ and Profiler™ kits, respectively), including five population samples: Purépechas from Mountain, Valley, and Lake; Triquis, and Mayas. Previously published STR data were included to the inter-population analyses.

Results: forensic statistical parameters were estimated by population. The majority of Native groups were not differentiated, excepting Triquis and Purépechas (Valley and Lake), attributable to their relative geographic and cultural isolation. Conversely, Purépechas-Mountain presented an elevated number of rare alleles, suggesting recurrent gene flow into this group. Interestingly, Huastecos and Yucatec Mayas were not differentiated, which is in agreement with the archeological hypothesis that Huastecos represent an ancestral Maya group. Interpopulation variability was five times larger in Natives than in Mestizos.

Conclusion: Results suggest European admixture has increased the similarity among Native Mexican groups. In addition, inconsistent clustering of Native groups by language or geography stresses the importance of serial founder effect and/or genetic drift to depict their present genetic relationships.

Introduction

Patterns of the current population structure provide an important source of data for inferences regarding recent demographic history. Genetic variation among human populations has shown that groups living on the same continent are relatively homogeneous (Bamshad et al., 2004). However, Native American populations exhibit considerable interpopulation variability indicating differences between populations from North and South America (Bortoloni et al., 2003; Mao et al., 2007; Wang et al., 2007). The pre-Columbian civilizations of the largest part of Mexico and Central America, conforming Mesoamerica, participated in the same universe of beliefs and rites; they shared a certain lifestyle –sedentary–, as well as social and political organization. A relative cultural homogeneity based on archaeological and anthropological data has been described (Duverger, 2007). However, also observed is a linguistic and genetic heterogeneity in Mesoamerica, shaped by both demographic and biological factors (Wang et al., 2008). In Mexico, the present number of indigenous population is 10.2 million, representing 9.6% of total Mexican population. There is a spread of 156,557 native settlements in 803 localities, in which >30% of the population speak an indigenous language. Using language as a criterion selection, it is possible to estimate that in Mexico there are >68 native groups >85 languages and variant dialects described until now (Cisneros, 2004; National Institute of Statistics, Geography, and Informatics-Mexico [INEGI], 2005); nearly 80% of this population is concentrated in eight Mexican States as follows: Chiapas; Oaxaca; Guerrero; Hidalgo; Yucatán; Campeche; Veracruz, and San Luis Potosí.

Among the Native Mexican groups analyzed in this work, Purépechas –also known as Tarascos– constituted one of the most important Mesoamerican cultures at the moment of Spanish contact, which came to control a vast area of western Mexico (70,000 km^2) including the State of

Michoacán and part of the states of Guanajuato, Guerrero, Jalisco, Colima, Querétaro, and Mexico. In point of fact, the Purépechas were one of the few groups that resisted the Aztec expansion prior to the Spanish Conquest (Michelet, 2001). They derived from admixture of different Chichimecas groups, a term referring to nomad hunters from Aridoamerica. According to the Relation and Chronicles of Michoacán, these groups went on pilgrimages the Aztecs and other Native groups from the mythic site, *Chicomoztoc*; they separated to the East and arrived at Michoacán, where they admixed with local Nahuas already settled in the Michoacán territory, giving rise the Native group known as pre-Tarascos (Kirchhof, 1956). Other sources claim they formed a social organization structured in shorts groups that arrived first at Zacápu and Naranxán in the state of Michoacán ca. 4,000 ybp; they eventually migrated and congregated at Pátzcuaro and contiguous Lakes (Jiménez-Moreno, 1948; Schöndube, 1996; Michelet, 1996, 2001). The second Native group analyzed in this study comprised the Triquis, who presumably originated in the Central Valley of Oaxaca State –probably Monte Alban–, and eventually were banished by the Zapotecans. Subsequently, they arrived at their actual location in the western Oaxaca mountain region nearly 2,000 ybp. At the beginning of the XV century, the Triquis were subjugated by the Aztecs and were forced to paid tribute (Lewin-Fisher and Sandoval-Cruz, 2007). At the time of the Spanish contact, the Triquis already constituted a cultural and linguistic island in the High Mixteca region of Oaxaca. Presently, the Triquis comprise two principal regions with cultural and linguistic differences: San Juan Copala (Low), and Chicahuaxtla (High); access to their territory is difficult due to its localization at confluence of the Sierra Madre Oriental with the Sierra Madre Occidental, comprising an extension of 500 km^2 (Huerta-Ríos, 1995; Lewin-Fisher and Sandoval-Cruz, 2007). Finally, we analyzed the Yucatec Mayas, who constituted one of the most important Mesoamerican cultures because of their ancient

cultural and scientific legacy. The Maya civilization inhabited a large area of southeastern Mexico and Central America, with a history of ca. 3,000 ybp. During this time, hundreds of dialects were spoken in these regions, generating nearly 44 different contemporary Mayan languages. Records and archeological data indicate that Pre-Columbian Mayas of the Yucatán Peninsula achieved two large migrations during the Late Classic and Early Post-Classic ages, including one from the Central Uplands of Mexico across the coastal plain of the Gulf of Mexico, and another, yet more ancient, from the Petén area in Maya Uplands at the South of Yucatan Peninsula (Nalda, 2005; Schmidt, 2007). The identity of these culture remains in force at present with the concurrence of at least three factors: the everyday use of the Mayan language; the permanence of religious rituals and customs, and a social organization of autonomous communities. Their social and political conditions were markedly inferior during the three centuries following the Spanish Conquest (Ruz, 2006).

To unravel the differentiation processes that generated the population's genetic heterogeneity, microsatellites –or Short tandem repeats (STRs)– constitute ideal polymorphic markers, whose relatively high mutation rate allows assessment of the biological diversity and elucidation of the history of human populations (Bosch et al., 2000; Zhivotovsky et al., 2003; Sahoo and Kashyap, 2005; Liu et al., 2006). In this context, we highlight the autosomal STRs included in the Combined DNA Index System (CODIS), which are widely used for human identification purposes. The correct interpretation of CODIS-STR-generated DNA profiles in forensic casework requires knowledge of the allele distribution and some statistical parameters in the population in which the system will be applied; thus, worldwide-population STR datasets have been generated for this purpose. In Mexico, despite the large number of Native groups, only a few molecular studies have been conducted with autosomal STR loci in these populations

(Rangel-Villalobos et al., 2000; Sánchez et al., 2005; Barrot et al., 2005; Ibarra-Rivera et al., 2008; González-Martín et al., 2008).

In this work, we obtained CODIS-STR population data in order to estimate statistical parameters of forensic importance of five population samples from the following three Native Mexican groups: Purépechas; Triquis, and Mayas. In addition, we analyzed the genetic relationships and population structure (AMOVA) in these native groups (clustered by geographic and linguistic criteria), including previously reported ancestral populations (African and European), Mestizos, and Natives from Mexico. Anthropological discussion addressed both Pre-Columbian records and the possible present-day effects of gene flow among these Native populations.

Methods

Population Sample

A total of 531 unrelated individuals from five indigenous communities were studied. Prior to the inclusion in our study, all volunteers signed an informed consent letter, according to the ethical guidelines of the Helsinki Declaration; they were classified into three Native Mexican groups: (i) 333 Purépechas from three areas of the western state of Michoacán, including the localities of Zipiajo (n = 168), Angahuan (n = 103), and Puácuaro (n = 62) from the Mountain, Valley, and Lacustrine Regions, respectively; these three population samples were analyzed individually; (ii) 108 Triquis from the District of San Juan Copala in the Mixteca region of the eastern state of Oaxaca, and (iii) 90 Mayas from different localities around Mérida, the largest city of the Yucatán peninsula, in Mexico's southeastern region. DNA was extracted from fresh blood samples by the salting-out method (Miller et al., 1988) and from buccal swabs by Chelex® 100 method (Walsh et al., 1991). For interpopulational analyses, we included previously published Native Mexican and Mestizo populations (Table I); their geographic location throughout the

Mexican Republic is presented in Figure 1. In addition, two worldwide population samples from Europe and Africa were included for this purpose. For Native groups, their linguistic classification is indicated in Figure 2 (Gordon, 2005; National Institute of Indigenous Languages-Mexico [INALI], 2008).

INSERT TABLE I, FIGURE 1 AND FIGURE 2

PCR amplification and genotyping

We used the Profiler Plus™ and Identifiler™ kits from Applied Biosystems (Foster City, CA, USA), which are designed for co-amplification of the following autosomal STR loci: D8S1179; D21S11; D7S820; vWA; D18S51; D3S1358; D13S317; D5S818, and FGA (Profiler Plus™ PCR kit). Additionally, CSF1PO, D19S433, TPOX, TH01, D16S539, and D2S1338 were analyzed in Purépechas (Identifiler™ PCR kit). The amplified products were separated by capillary electrophoresis using the ABI Prism™ 310 Genetic Analyzer following manufacturer recommendations. The allelic ladder provided with the kit and GeneMapper ID software version 3.2 were utilized for genotyping.

Data analyses

Allele distribution and statistical parameters of forensic importance were computed with the PowerStats program (Tereba, 1999). For each population sample, Hardy-Weinberg expectations and two-loci equilibrium were verified by exact tests with a 95% Confidence interval (95% CI) with the Genetic Data Analysis (GDA) program version 1.1 (Lewis and Zaykin, 2001). Bonferroni correction was applied to evaluate these *p*-values according to the loci-number of Profiler and Identifiler kits ($p < 0.0055$ and $p < 0.0033$, respectively). Gene flow among Native groups was assessed as the number of migrants per generation (Nm) according to the equation of Wright (Wright, 1951). In addition, we estimated the following parameters of genetic diversity in

each Native group: (i) mean allele number, (ii) average expected heterozigosity, and (iii) number of alleles exclusively observed in one population or "rare alleles".

For interpopulational analysis, we included STR data from previously published populations described in Table I. For consistent comparison, data of only 9 STR loci included in the Profiler™ kit analyzed in all these populations were employed for this purpose. Genetic differentiation was evaluated by normalized F_{ST} distances and pairwise F_{ST} *p*-values, computed with the Arlequin 3.1 software (Excoffier et al., 2005). Bonferroni correction was implemented to evaluate multiple F_{ST} *p*-values by population. F_{ST} distance was selected because represent genetic differentiation patterns by drift, corresponding with both genetic and archeological records of human populations (Pérez-Lezaún et al. 1997). Genetic distances were displayed on a Multidimensional scaling (MDS) plot to explore the genetic relationships among populations with the SPSS for Windows program version 10.0. Analysis of molecular variance (AMOVA) was carried out placing Mestizos and Natives populations in different clusters based on geography and linguistic classification, as properly described in the text. Additionally, we utilized Spatial analysis of molecular variance (SAMOVA), which is similar to the traditional AMOVA, to define accurate population groups that as such geographically are genetically homogeneous, and groups sufficiently differentiated from each other (Dupanloup et al., 2002).

To establish whether decrease of homozygosity (or increase of heterozygosity) reflects European admixture in Native groups, we reviewed correlation of the decrease of homozygosity with the genetic distance between each group and the southwestern Spanish population. For each Native group, this European admixture marker (a decrease in homozygosity) was correlated with its geographic distance and altitude to the nearest Mexican-Mestizo population. Thus, the final purpose was to verify whether geographic distance and altitude influence European admixture in these Native groups. In order to investigate whether Isolation-by-distance (IBD) could explain

genetic differentiation among Native populations, we revised the correlation between genetic and geographic distances among these groups (Ramachandran et al., 2005). The statistical significance of these correlations was evaluated by the Mantel test. Distances in km between populations were computed employing geographic coordinates with the Great Circle Calculator program (http://www.gb3pi.org.uk/great.html). Concurrently, we examined possible landscapes of genetic and geographic differentiation processes by means of AIDA program software (Bertorelle and Barbujani, 1995).

Results

Statistical Parameters of Forensic Importance and Genetic Diversity

Allele distribution and statistical parameters of forensic importance of the Native Mexican groups Purépechas (West), Triquis (South), and Mayas (South-East) are shown in Supplementary Tables S1-S5. In general, for all five Native population samples, genotype distribution by locus and two loci combination were in agreement with Hardy-Weinberg and linkage equilibrium, respectively. Only two loci displayed significant *p*-values for HWE test after applying the Bonferroni correction: D3S1358 in Purépechas-Lake, and D7S580 in Triquis; these *p*-values were close to the Bonferroni limit and represented unique events by population (Tables S1-S5). Therefore, they do not support immigration or endogamy processes in these Native groups; thus, we did not consider they deserve further discussion. The combined Power of discrimination (PD) and Power of exclusion (PE) for both STR systems were ≥0.9999 and ≥0.99752, respectively. The genetic diversity parameters of these groups are graphically presented in Table II. Purépechas-Mountain had the largest number of rare (private) alleles with six, followed by the Purépecha-Valley and Mayas, with three rare alleles each Native group. Thus, the three Purépecha population samples jointly presented 11 rare alleles. For the mean allele number, again Purépechas-Mountain had the maximum value, followed by Mayas and Choles. Finally, the

average of expected heterozigosity pointed out Otomi-Sierra, Choles and Purépechas-Mountain, respectively, as the Native groups with larger genetic diversity, whereas the smallest value was observed in Triquis.

INSERT TABLE II

Genetic Differentiation among Populations

The MDS plots based on pairwise Fst values (Figure 3) shows the genetic relationships among populations. The stress values for both MDS plots (Figure 3A and 3B) were 0.10100 and 0.11430 respectively. Therefore, indicates that the data represent an appropriate configuration in their spatial distribution. As could be expected, Mexican Mestizos displayed a closer genetic relationship with the European population than the Native groups (Figure 3A). Additional discussion concerning genetic differentiation among Mexican Mestizos will be omitted, considering that this has been conducted in a recent report (Rubi-Castellanos et al., 2009). Regarding Native groups, Triquis and Purépechas from Valley and Lake presented significant differences with all Mestizo populations included herein (data not shown), which can be inferred analyzing the MDS plot between populations (Figure 3A). This result suggests low European admixture in these three Native populations, contrasting with a previous observation of elevated European admixture in Purépechas in view of their high heterozygosity and similar STR allele frequencies to western Mestizos (p >0.05) (Rangel-Villalobos et al., 2000); the low number of markers and the small size (n = 25) and geographical origin of the Purépecha population sample previously studied appear to be relevant in explaining this difference.

INSERT FIGURE 3

Conversely, the Tepehuas, Otomíes-Sierra, Otomíes-Valley, Mayas, and Choles were genetically closer to Mestizos from Central and southeastern regions, including the Valley of Mexico, Hidalgo, Puebla, Veracruz, and Yucatán (Figure 3A). This result suggests the presence of certain

European admixture level in these Native populations, as previously reported for the Chol population sample (González-Martín et al., 2008), which here was the closest Native group to some Mestizo population, in this case Puebla (Figure 3A). Concurrently, pairwise comparisons showed non-significant differentiation among Tepehuas, Otomíes-Sierra, Otomíes-Valley, Mayas, and Choles (Table III).

INSERT TABLE III

The correlation was not significant between homozygosity in Native groups and the increase of genetic distance to the Spanish population of reference (r^2= 0.587; p = 0.0550), indicating that homozygosity was not a suitable European admixture marker (plot not shown). This conclusion was confirmed when correlation test was repeated without Triquis, the most differentiated Native group, diminishing the estimated correlation (r^2 = 0.072; p = 0.3320). Therefore, posterior correlations with altitude and geographic distance respect to the nearby Mestizos were not carried out.

Genetic structure (AMOVA)

Analysis molecular of variance (AMOVA) tests consistently demonstrated that the majority of genetic variability for the 9 STR system in Mexican populations is at the intrapopulational level (F_{IT} = 98.8–99.3%), which was moderately significant. Conversely, inter-population variability in Native groups was nearly five times larger than in Mestizos (F_{ST} = 1.25 vs. 0.26%), and extremely significant (Table IV). The following AMOVA test clustering Mestizos vs. Native groups indicated low internal consistency –or high heterogeneity– into these clusters, because the genetic differentiation among populations into groups was larger than the differentiation among groups (0.61 vs. 0.38%), both of these significant (Table IV).

INSERT TABLE IV

Finally, a set of AMOVA tests was carried out exclusively in Native groups, which were clustered according to linguistic and geographic criteria (Table IV). Results revealed that on increasing linguistic criteria for clustering Native groups (stock and family, particularly), differentiation among groups also increased (F_{CT} = 0.2–0.62%), decreasing differentiation among populations into groups (F_{SC} = 1.10–0.74%).

Landscapes of Genetic and Geographic Differentiation Patterns

Although the geographic distance (km) and genetic differentiation (F_{ST}) among Native Mexican groups was not correlated (r^2 = –0.0167; p = 0.4300), the correlation plot allowed shaping three different population clusters, representing 1) Purépechas, Otomíes, Huastecos, and Tepehuas, 2) Mayas and Choles, and 3) Triquis (Figure 4). In the correlation test by cluster, only the geographically more remote native groups (Mayas and Choles) presented a significant correlation (r^2 = –0.5095; p = 0.0040). Concurrently, analysis with AIDA software displayed a slight pattern observed in IBD processes. Despite this, few significant values (4/9) could support the aforementioned differentiation model in Native groups from Mexico. Interestingly, the most significant value in the AIDA autocorrelogram plot appears to represent the geographical distance of the Triquis; subsequent analysis without this dataset clearly generates a random differentiation pattern (plot not shown). Moreover, although autocorrelation values representing Mayas and Choles (800–1,400 km) decreased from positive to negative, only one of these four points was significant (Figure 4); these classes include pairwise comparisons between Mayas and Choles with all Mexican Native populations from Hidalgo and Michoacán states.

INSERT FIGURE 4

Discussion

Statistical Parameters of Forensic Importance

The correct application of CODIS-STRs for human identification purposes requires that allele frequencies and forensic statistical parameters be estimated in the population where the genetic system will be employed (Evett and Weir, 1998). Particularly, genetic data of these widely employed STR systems are scarce in Native Mexican groups; as observed, these populations have a distinctive distribution regarding the admixed Mexican Mestizos, supporting the establishing of local STR databases. In this context, our results are important because they support the confident employment of the respective STR system for DNA profile interpretation in forensic casework.

AMOVA and Genetic Differentiation among Populations

The non-differentiation observed between Tepehuas, Otomíes-Sierra, Otomíes-Valley, Mayas, and Choles (Table III), inferred as those with larger European component (Figure 3A), suggests that this could be acting as a homogenizing factor that has increased similarity among Native American populations. A similar observation has been reported in three of the seven indigenous groups studied with the Polymarker system (PM) including Mixteca Alta, Mixteca Baja, and Nahuas of Xochimilco (Buentello-Malo et al., 2003). This is in agreement with the AMOVA results indicating lower differentiation among Mexican Mestizos regarding Native groups; consequently, admixture occurring after European contact with New World populations came to diminish Native population genetic differentiation, previously generated by processes such as serial founder effect and random genetic drift as described for human populations (Ramachandran et al., 2005; Zhang and Dolan, 2008). Unfortunately, we could not use homozygosity as European admixture marker in these Native American populations. Probably the homozygosity usefulness diminished by a similar –although probably low– admixture level in the mentioned Mesoamerican Native groups. Finally, to estimate correctly the presence of European and/or African admixture in these groups, a deeper analysis with further loci would be needed (i.e., with Ancestry informative markers [AIMS]).

The larger genetic differentiation among populations into groups than among groups (Table IV), is consistent with the proposal of heterogeneity as a major characteristic of Mexican populations (Bonilla et al., 2005; Wang et al., 2008), although in contrast with a previous report claiming genetic homogeneity for seven Native Mexican groups based on five PM-system loci (Buentello-Malo et al., 2003); unfortunately, the authors did not apply a significance test to evaluate F_{ST}. The greater resolution power of the 9 STRs to disclose population genetic structure –with respect to the PM system– could explain the contrasting conclusions of these studies in Mexican populations.

The poor quality of both linguistic and geographic (SAMOVA test) criteria for clustering Native groups was particularly noteworthy because in all cases, differentiation among populations into groups was significant ($p = 0.0000$). Taken together, these results emphasize the importance of the differentiation processes that acted upon Native American populations (Wang et al., 2007). Results of AIDA software and correlation tests indicated that, at the geographical level of these Native groups is not possible to invoke a simple population pattern of genetic differentiation. Therefore, more complex evolutionary landscapes could fit better to explain the genetic differentiation presently observed among Native groups from Mesoamerica, such as Isolation by Migration (IM) models (Hey, 2005; Kitchen et al., 2008).

With respect to the genetic relationships among Native groups, we omitted discussing Otomíes from the Valley and Sierra, Tepehuas, and Huastecos (central region) because this has been previously addressed (González-Martín et al., 2004, 2008). Particularly, caution must be taken respect to the lack of differentiation of the Tepehuas respect to the majority of Native groups (Table III), because this population sample had many STR data lost and was relatively small (n= 47); consequently, discussion about the Tepehuas genetic relationships will be avoided.

Therefore, we present a particular discussion of the results concerning the population samples studied herein:

Purépechas

The MDS plot (Figure 3B) in conjunction with the significant F_{ST} *p*-values (Table III) depicted the Triquis and Purépechas from the Lake and Valley as the most differentiated Native groups, respectively; these were probably influenced by cultural and geographic isolation, and the small effective population size of these groups, promoting differentiation processes as random genetic drift. In agreement with this differentiation, the Purépecha language has been described as an isolated dialect that is not related with any other linguistic family from Mexico (INALI, 2008) (Figure 2). In addition, some authors have suggested that Purépechas received one or several migrations from Peru that landed on the Pacific Coast in the Mexican state of Michoacán; because they possess a distinctive archeology, anthropology, culture, and language (Ruiz, 1891; Peñaloza et al., 2001). However, this asseveration is difficult to confirm, bearing in mind that the Purépechas rarely touched or lived on the coast; in addition, historical, archeological and anthropological records are not sufficient for supporting this theory (Michelet, 2001; Márquez-Joaquín, 2007). Conversely, Purépechas from the Mountain presented the largest quantity of rare alleles, without a significant increment in genetic diversity (Table II). Although for STRs we could not apply a neutrality test to evaluate the excess of rare alleles respect to the mutation-drift equilibrium expectation, it has been demonstrated the excess of rare alleles is consequence of population amalgamation (Chakraborty et al., 1988), and particularly this effect has been observed in Native American populations by means of mitochondrial DNA (mtDNA), implying recurrent and high levels of gene flow (Fuselli, 2003). Concurrently, preliminary studies of Native American paternal lineages defined by the mutation M3 (Páez-Riberos et al., 2006), allowed to propose that the even distribution of Y-STR haplotypes throughout the network

joining tree including different Native Mexican groups is consequence of the Pre-Columbian multiethnic origin of this Native group rather than of European admixture (via Mestizos). Similarly, mtDNA-haplogroups have revealed that Purépechas present an intermediate position between two clusters in a principal components plot (Peñaloza-Espinoza et al., 2007). In brief, our results are in agreement with the hypothesis that Purépechas is an ancient and cystic group in their own territory that, once it was shaped by different Native groups spread out in western Mesoamerica, most of this Native group remained in the same place and had low admixture with other Mesoamerican groups (Jiménez-Moreno, 1948; Schöndube, 1996; Michelet, 1996, 2001). However, important differences in gene flow could exist, as observed in the Purépechas-Mountain population sample respect to those of the Valley and Lake (Table II). Finally, to explain the present genetic background of this group, the recent Purépecha gene flow should not be disregarded, considering that census data (2000–2005) recorded that a total of 1,498 Purépecha speakers living in Michoacán state migrated mainly to the states of Jalisco, Baja California, and Mexico, and to the U.S.A. (INEGI, 2007).

Mayas

In agreement with their same linguistic affiliation within the Maya-Totonaco group, Mayanse stock, and Maya family, Yucatec Mayas were not differentiated from Huastecos and Choles (Figure 2; Table III). However, Huastecos showed significant differences with Choles, probably attributable to the higher genetic differentiation of Huastecos, and to recent gene flow that Choles have received from other ethnic groups (probably Highlands central groups), and/or from Mexican Mestizos (Alejos-García and Martínez-Sánchez, 2007). This non-differentiation between Huastecos and Yucatec Mayas is important because is in agreement with the hypothesis that Huastecos could represent an ancestral Maya group that separated and remained in the Huasteca zone during migrations occurring 3,000 ybp (Ekholm, 1944). Concurrently, the non-

differentiation between the nearby Maya groups (from Yucatán and Choles from Campeche) with Otomíes could be indicative of gene flow among these central and southeastern native groups, as a consequence of multiple human movements and arrangements throughout Mesoamerica since the fall of Teotihuácan up to the Early Post-Classic age (1,200–700 ybp), especially in the central highlands and Maya region (Nalda, 2005). This controversial theory of Toltec migration to Yucatán is supported by historical, archeological, pictography, social, and political organization, as well as the religion and militarism present in peninsular Mayas (Morley, 1946). In this context, based on 9 STR data, we estimated an elevated migration rate for these Native groups from central and southeastern regions of Mexico (Nm = 38.8). Similarly, Y-linked markers have displayed an elevated migration rate throughout these regions (Nm = 24.76), increasing homogeneity among these Native groups (Rangel-Villalobos et al., 2008). Additionally, the influence of gene flow on Native groups from southeastern Mexico is supported by archeological references concerning Pre-Columbian Mayas, who carried out several migration stages especially during the Late Classic and Early Post-Classic age (Nalda, 2005). In fact, multiple dates have catalogued this age as a "dynamic era" of Maya history (Soustelle, 1993; Nalda, 2005; Schmidt, 2007; Ibarra-Rivera et al., 2008).

Triquis

Triquis had the lowest average of genetic diversity (h = 0.6953) and the most distant MDS-plot position, suggesting that additional and/or more profound genetic differentiation processes have occurred in this population (i.e., inbreeding, founder effect, etc.). Demographic data indicate that total Triqui population throughout the Mexican territory is relatively small (~25,000 inhabitants), and recently a certain fraction has migrated to the States of Morelos, Veracruz, and Sonora, and to Mexico City, in addition to the U.S.A (Lewin-Fisher and Sandoval-Cruz, 2007). Particularly,

the Triqui territory of the Lower Region (that belongs to the San Juan Copala, the origin of the population sample) is a small town with scarce communication with Mexican Mestizos or nearby native groups (i.e. Mixtecos), aided by their rugged geographic location in an abrupt, difficult-access mountainous region. In addition, they have a cultural commitment to maintain their language and traditions, and limited confidence in persons from the outside (Huerta-Ríos, 1995; Lewin-Fisher and Sandoval-Cruz, 2007). Therefore, both geographic and cultural aspects have operated simultaneously, probably since Pre-Columbian times, to shape the current differentiation of this Native group.

Conclusion

The CODIS-STR data here obtained validate the use of these markers for human identification purposes in these Native Mexican groups. A significant differentiation of Triquis and Purépechas from Valley and Lake was demonstrated, attributable to their relative geographic and cultural isolation. Although a relative homogeneity was detected among Mesoamerican groups, particularly those inferred with higher European admixture, the large interpopulational variability rendered it impossible to shape consistent population clusters, stressing the importance of serial founder effect and genetic drift to depict their genetic relationships. Concurrently, geographic and/or linguistic elements constituted a limited tool for explaining their current genetic relationships, presumably due to the complex historic and demographic events of the human populations from Mesoamerica, both prior to and after the Spanish Contact.

Acknowledgments

We thank to all the volunteers who apportioned blood samples for this study, and to the Consejo Nacional de Ciencia y Tecnología (CONACyT) for fellowship to G. M-C, and for the research grant N° 48710 to H. R-V.

Literature cited

Alejos-García J, Martínez-Sánchez NE. 2007. Choles. México: CDI, Inc. 47p.

Bamshad M, Wooding S, Salisbury BA, Stephens JC. 2004. Deconstructing the relationship between genetics and race. Nat Rev Genet 5:598-609.

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 119:111-5.

Bertorelle G, Barbujani G. 1995. Analysis of DNA diversity by spatial autocorrelation. Genetics 140:811-19.

Bonilla C, Gutierrez G, Parra EJ, Kline C, Shriver MD. 2005. Admixture analysis of a rural population of the state of Guerrero, Mexico. Am J Phys Anthropol 128:861-69.

Bortolini MC, Salzano FM, Thomas MG, Stuart S, Nasanen SP, Bau CH, Hutz MH, Layrisse Z, Petzl-Erler ML, Tsuneto LT, Hill K, Hurtado AM, Castro-de-Guerra D, Torres MM, Groot H, Michalski R, Nymadawa P, Bedoya G, Bradman N, Labuda D, Ruiz-Linares A. 2003. Y-chromosome evidence for differing ancient demographic histories in the Americas. Am J Hum Genet 73:524-39.

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 8:360-66.

Buentello-Malo L, Peñaloza-Espinosa RI, Loeza F, Salamanca-Gomez F, Cerda-Flores RM. 2003. Genetic structure of seven Mexican indigenous populations based on five polymarker loci. Am J Hum Biol 5:23-8.

Cerda-Flores RM, Budowle B, Jin L, Barton SA, Deka R, Chakraborty R. 2002. Maximum likelihood estimates of admixture in Northeastern Mexico using 13 short tandem repeat loci. Am J Hum Biol 14:429-39.

Chakraborty R, Smouse PE, Neel JV. 1988. Population amalgamation and genetic variation: observations on artificially agglomerated tribal populations of Central and South America. Am J Hum Genet 43:709–25.

Cisneros IH. 2004. Situación de los pueblos indígenas en México. En: Derechos humanos de los pueblos indígenas en México. México: CDHDF. 9p.

Dupanloup I, Schneider S, Excoffier L. 2002. A simulated annealing approach to define the genetic structure of populations. Mol Ecol 11:2571-81.

Duverger C. 2007. El primer mestizaje, la clave para entender el pasado mesoamericano. México: Santillana, Inc. 740p.

Evett IW, Weir BS. 1998. Interpreting DNA evidence; statistical genetics for forensic scientists. Sinauer Associates, Inc., Sunderland, MA. 79p.

Ekholm GF. 1944. Excavations at Tampico and Panuco in the Huasteca, México. Nueva York: Anthropological Papers of The American Museum of Natural History. v. 38. 5p.

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.

Fuselli S, Tarazona-Santos E, Dupanloup I, Soto A, Luiselli D, Pettener D. 2003. Mitochondrial DNA Diversity in South America and the Genetic History of Andean Highlanders. Mol Biol Evol 20:1682-91.

Gamero JJ, Romero JL, González JL, Arufe MI, Cuesta MI, Corte-Real F, Carvalho M, Anjos MJ, Vieira DN, Vide MC. 2000. A study on ten short tandem repeat systems: African immigrant and Spanish population data. Forensic Sci Int 110:167-77.

González-Martín A, Barrot C, Ortega M, Brant-Casadevall C, Moreno P, Rangel-Villalobos H, Gorostiza A, Huguet E, Corbella J, Gene M. 2004. Variabilidad genética y mestizaje en una

comunidad indígena Otomí, México. Actas XIII Congreso de la Sociedad Española de Antropología Biológica: 553-62.

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 20:605-13.

Gordon RG. 2005. Ethnologue: Languages of the World, Fifteenth edition. Dallas, Tex.: SIL International. Online version: http://www.ethnologue.com/.

Gorostiza A, González-Martín A, Ramírez CL, Sánchez C, Barrot C, Ortega M, Huguet E, Corbella J, Gené M. 2007. Allele frequencies of the 15 AmpF/Str Identifiler loci in the population of Meztitlán (Estado de Hidalgo), México. Forensic Sci Int 166:230-32.

Hey J. 2005. On the number of New World founders: a population genetic portrait of the peopling of the Americas. PLoS Biol 3:e193.

Huerta Ríos C. 1995. Los triquis. En: Etnografía contemporánea de los pueblos indígenas de México, región Pacífico Sur. México: INI. 24p.

Ibarra-Rivera L, Mirabal S, Regueiro MM, Herrera RJ. 2008. Delineating genetic relationships among the Maya. Am J Phys Anthropol 135:329-47.

INEGI. 2005. Instituto Nacional de Estadística Geografía e Informática. México http://www.inegi.gob.mx

INEGI. 2007. Instituto Nacional de Estadística, Geografía e Informática. "Estadística a propósito del día internacional de las poblaciones indígenas". Datos de hablantes de la lengua Purépecha. http://www.inegi.gob.mx/inegi/contenidos/espanol/prensa/Contenidos/estadisticas/2007/purepecha07.pdf

INALI. 2008. Instituto Nacional de Lenguas Indígenas. México http://www.inali.gob.mx

Jiménez-Moreno W. 1948. Historia antigua de la zona tarasca. México: Sociedad Mexicana de Antropología, México. p. 146-55.

Kitchen A, Miyamoto MM, Mulligan CJ. 2008. A three-stage colonization model for the peopling of the Americas. PLoS ONE 3:e1596.

Kirchhoff P. 1956. "Relación de Michoacán como fuente para la historia de la sociedad y cultura tarascas". En: reproducción facsímil del Ms C IV 5 de El Escorial con transcripción, introducción y notas por José Tudela, Madrid, Aguilar. 3p.

Lewin-Fisher P, Sandoval Cruz F. 2007. Triquis de Oaxaca. México: CDI, Inc. 47p.

Lewis PO, Zaykin D. 2001. Genetic Data Analysis (GDA): Computer program for the analysis of allelic data. Version 1.0 (d16c). Programa disponible por los autores en internet, http://lewis.eeb.uconn.edu/lewishome/software.html.

Licea-Cadena RA, Rizzo-Juárez RA, Muñiz-Lozano E, Páez-Riberos LA, Rangel-Villalobos H. 2006. Population data of nine STRs of Mexican-Mestizos from Veracruz (Central South-Eastern, Mexico). Leg Med 8:251-2.

Luna-Vazquez A, Vilchis-Dorantes G, Aguilar-Ruiz MO, Bautista-Rivas A, Rojo-Nava AL, Rios-Barrios E, Rangel-Villalobos H. 2005. Population data for 15 loci (Identifiler Kit) in a sample from the Valley of Mexico. Leg Med 7:331-3.

Liu H, Prugnolle F, Manica A, Balloux F. 2006. A geographically explicit genetic model of worldwide human-settlement history. Am J Hum Genet 79:230-7.

Mao X, Bigham AW, Mei R, Gutierrez G, Weiss KM, Brutsaert TD, Leon-Velarde F, Moore LG, Vargas E, McKeigue PM, Shriver MD, Parra EJ. 2007. Genome-wide admixture mapping panel for Hispanic/Latino Populations. Am J Hum Genet 80:1171-8.

Márquez-Joaquín P. 2007. ¿Tarascos o Purhépechas? Voces sobre antiguas y nuevas discusiones en torno al gentilicio michoacano. Universidad Michoacana de San Nicolás de Hidalgo/Universidad Intercultural indígena de Michoacán/El Colegio de Michoacán. Fondo Editorial Morevallado.

Martínez-González LJ, Martínez-Espin E, Fernández-Rosado F, Moguel MA, Entrala C, Álvarez 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 50:236-8.

Michelet D. 1996. El origen del reino tarasco protohistórico: la cuenca de Zacapu. Arqueología Mexicana 4:24-7.

Michelet D. 2001. La zona occidental en el posclásico. En: editorial Porrua. Historia Antigua de México. México. México: INAH e IIA (UNAM). 38p.

Miller SA, Dykes DD, Polesky HF. 1988. A simple salting out procedure for extracting DNA from human nucleated cell. Nucleic Acids Res 16:1215-6.

Morley SG. 1946. The Ancient Maya. Stanford University Press, Stanford.

Nalda E. 2005. Clásico Terminal (750-1050 D.C.) y Posclásico en el área México maya: Colapso y reacomodos. Arqueología mexicana 13: 30-9.

Páez-Riberos LA, Muñoz-Valle JF, Figuera LE, Nuño-Arana I, Sandoval-Ramírez L, González-Martín A, Ibarra B, Rangel-Villalobos H. 2006. Y-linked haplotypes in Amerindian chromosomes from Mexican populations: genetic evidence to the dual origin of the Huichol tribe. Leg Med 8:220-5.

Peñaloza-Espinoza RI, Delgado P, Arenas-Aranda D, Barrientos C, Buentello-Malo L, Loeza F, Salamanca F. 2001. (AC)n dinucleotide repeat polymorphism in 5' β-globin gene in Native and Mestizo Mexican populations. Hum Biol. 73:885-90.

Peñaloza-Espinosa RI, Arenas-Aranda D, Cerda-Flores RM, Buentello-Malo L, González-Valencia G, Torres J, Álvarez B, Mendoza I, Flores M, Sandoval L, Loeza F, Ramos I, Muñoz L, Salamanca F. 2007. Characterization of mtDNA haplogroups in 14 Mexican indigenous populations. Hum Biol 79:313-20.

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

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. PNAS 102:15942-7.

Rangel-Villalobos H, Rivas F, Sandoval L, Ibarra B, Garcìa-Carvajal ZY, Cantú JM, Figuera LE. 2000. Genetic variation among four Mexican Populations (Huichol, Purépecha, Tarahumara, and Mestizo) revealed by two VNTRs and four STR's. Hum Biol 72:983-95.

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 135:448-61.

Rubi-Castellanos R, Anaya-Palafox M, Mena-Rojas E, Bautista-España D, Muñoz-Valle JF, Rangel-Villalobos H. 2008. 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). FSI:Genetics (online).

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 Jan 12 [Published ahead of print].

Ruz MH. 2006. Mayas, primera parte. México: CDI, Inc. 91p.

Ruiz E. 1891. Michoacán. Paisajes, tradiciones y leyendas. Oficina Tipográfica de la Secretaria de Fomento, México.

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.

Schmidt Peter J. 2007. Los toltecas de Chichén Itzá, Yucatán. Arqueología mexicana 15:64-8.

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.

Schöndube OB. 1996. Los tarascos. Arqueología Mexicana 4:14-21.

Soustelle J. 1993. La familia otomí-pame del México central. México: CEMCA/FCE. 40p.

Tereba A. 1999. Tools for analysis of population statistics. Profiles in DNA. Promega Corp. The evaluation of forensic evidence.

Walsh PS, Metzger DA, Higuchi R, 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing forensic material. Biotechniques10:506-13.

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.

Wang S, Ray N, Rojas W, Parra MV, Bedoya G, Gallo C, Poletti G, Mazzotti G, Hill K, Hurtado AM, Camrena B, Nicolini H, Klitz W, Barrantes R, Molina JA, Freimer NB, Bortolini MC, Salzano FM, Petzl-Erler ML, Tsuneto LT, Dipierri JE, Alfaro EL, Bailliet G, Bianchi NO,

Llop E, Rothhammer F, Excoffier L, Ruiz-Linares A. 2008. Geographic patterns of genome admixture in Latin American Mestizos. PLoS Genet 4:e1000037.

Wright S. 1951. The genetical structure of populations. Ann Eugen 15:323-54.

Zhang W, Dolan ME. 2008. Exploring the Evolutionary History of the Differentially Expressed Genes between Human Populations: Action of Recent Positive Selection. Evol Bioinform Online 15:171-9.

Zhivotovsky LA, Rosenberg NA, Feldman MW. 2003. Features of Evolution and Expansion of Modern Human, Inferred from Genome-wide Microsatellites Markers. Am J Hum Genet 72:1171-86.

FIGURE LEGENDS

Figure 1. Geographical locations of Mexican populations studied herein, and those used for comparison purposes. The previously published populations are indicated by black stars (Mestizos) and black points (Native American groups). Black triangles indicate populations reported in this study.

Figure 2. Linguistic classification of Native Mexican populations used for interpopulational analyses (Gordon, 2005; INALI, 2008). Underlined groups are reported on in this study.

Figure 3. Multidimensional scaling (MDS) plot based on normalized F_{ST} distances between (A) Mestizos, Native American, and Ancestral populations (European and African); (B) only Native American groups. See Table I for description of abbreviations.

Figure 4. Overlapped plots representing correlation between geographical and genetic distances (black lines) and the AIDA autocorrelogram (grey lines). Correlation plot displays the following three groups: Purépechas, Otomíes, Huastecos, and Tepehuas (black circles); Triquis (black squares), and Mayas and Choles (black triangles). In the autocorrelogram, filled diamonds indicate significant *p*-values ($p < 0.05$).

Table I. Description of the Mexican and Worldwide populations used for interpopulation analysis.

Population	Abbr.	Sample size	Geographical Origin	Reference
Native American				
Chol	Chol	106	Campeche State	Sánchez et al., 2005
Tepehua	Tep	47	Hidalgo State	González-Martín et al., 2008
Otomi Sierra	OtoS	83	Hidalgo State	Barrot et al., 2005
Otomi Valley	OtoV	82	Hidalgo State	Barrot et al., 2005
Huastecos	Hua	133	Hidalgo State	Barrot et al., 2005
Maya	May	90	Yucatan State	This study
Triqui	Tri	108	Oaxaca State	This study
Purépechas	Pur	333	Michoacán State (Mich)	
Zipiajo	Pur M	168	Zipiajo, Mich (Mountain)	This study
Angahuan	Pur V	103	Angahuan, Mich (Valley)	This study
Puacuaro	Pur L	62	Puacuaro, Mich (Lake)	This study
Mestizos				
Chihuahua	Chi	162	North Central	Martínez-González et al., 2005
Nuevo León	NL	143	North East	Cerda-Flores et al., 2002
Jalisco	Jal	309	West	Rubi-Castellanos et al., 2008
Veracruz	Ver	170	Central	Licea-Cadena et al., 2006
Valley of Mexico	Mex	242	Central	Luna-Vázquez et al., 2005
Hidalgo	Hid	106	Central	Gorostiza et al., 2007
Puebla	Pue	313	Central	Rubi-Castellanos et al., 2008
Yucatán	Yuc	262	South-East	Rubi-Castellanos et al., 2008
Worldwide				
European	Eur	138	Southern Spain	Gamero-Lucas et al., 2000
African	Afr	132	North Africa	Gamero-Lucas et al., 2000

Table II. Parameters of Genetic Diversity based on nine CODIS-STRs estimated in ten Native Mexican groups

Native Americans	Number of rare alleles	Mean allele number	Average of expected heterozigosity
Chol	2	8.889	0.7641
Purépecha-Mountain	6	9.556	0.7635
Purépecha-Valley	3	7.778	0.7352
Purépecha-Lake	2	7.444	0.7412
Tepehua	0	7.111	0.7483
Otomi-Sierra	1	8.333	0.7663
Otomi-Valley	0	7.889	0.7546
Huasteco	1	8.333	0.7405
Triquis	2	8.000	0.6953
Mayas	3	9.111	0.7566

Table III. Pairwise normalized F_{ST} distances (below diagonal), and F_{ST} *p*-values* (above diagonal) among 10 Native Mexican groups (See Table I for description of abbreviations).

	PurM	PurV	PurL	Tep	OtoS	OtoV	Hua	Chol	May	Tri
PurM	*****	*0.0000*	*0.0000*	0.27051	0.00098	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*
PurV	0.1455	*****	0.97168	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*
PurL	0.13914	0.02352	*****	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*
Tep	0.05192	0.10701	0.11114	*****	0.97656	0.82812	0.03711	0.98047	0.99023	*0.0000*
OtoS	0.0671	0.12189	0.10988	0.01866	*****	0.0459	0.00391	0.01172	0.01172	*0.0000*
OtoV	0.12184	0.1492	0.15905	0.0285	0.05858	*****	*0.0000*	0.02637	*0.0000*	*0.0000*
Hua	0.10651	0.11673	0.1234	0.06391	0.06808	0.1084	*****	0.00098	0.0127	*0.0000*
Chol	0.10254	0.13507	0.13314	0.02244	0.06342	0.06287	0.07892	*****	0.05566	*0.0000*
Tri	0.07013	0.10072	0.09182	0.01649	0.06462	0.09041	0.06123	0.05601	*****	*0.0000*
May	0.21182	0.275	0.29854	0.11466	0.1966	0.24037	0.19435	0.20016	0.20298	*****

* Bonferroni correction indicated significance at $p < 0.0056$

Table IV. AMOVA and SAMOVA tests in Mexican populations based on 9 CODIS-STRs

MEXICAN POPULATIONS	Nº Pop	Nº Groups	Into populations F_{IT} (%)	Inter populations F_{ST} (%)	
Mestizos	8	1	99.27; *p*= 0.04203	F_{ST} = 0.26%; *p*= 0.0000	
Native Americans	10	1	98.83; *p*= 0.02542	F_{ST} = 1.25%; *p*= 0.0000	
MESTIZO/NATIVE AMERICANS			**Into populations F_{IT} (%)**	**Among groups F_{CT} (%)**	**Populations into Groups F_{SC} (%)**
Mestizos *vs*. Native Americans	18	2	98.72; *p*= 0.0000	0.38; *p*= 0.0000	0.61; *p*= 0.0000
NATIVE AMERICANS GROUPED					
Linguistic Group classification [a]	10	3	98.77; *p*= 0.0332	0.20; *p*= 0.0449	1.10; *p*= 0.0000
Linguistic Stock classification [a]	10	5	98.71; *p*= 0.0273	0.62; *p*= 0.0000	0.74; p= 0.0000
Linguistic Family classification [a]	10	6	98.76; *p*= 0.0263	0.56; *p*= 0.0048	0.75; p= 0.0000
Geographic location [b]	10	5	98.74; *p*= 0.0293	0.63; *p*= 0.0000	0.71; *p*= 0.0000
Geographic location [c]	10	4	98.69; *p*= 0.0273	0.68; *p*= 0.0000	0.70; *p*= 0.0000

a. See linguistic classification criteria in Figure 2
b. May, Chol *vs.* Hua *vs.* PurM, PurV, PurL *vs.* Tri *vs.* Tep, OtoS, OtoV.
c. May, Chol, Hua *vs.* PurM, PurV, PurL *vs.* Tri *vs.* Tep, OtoS, OtoV.

Figure 1. Geographical locations of Mexican populations studied herein, and those used for comparison purposes. The previously published populations are indicated by black stars (Mestizos) and black points (Native American groups). Black triangles indicate populations reported in this study.
150x103mm (300 x 300 DPI)

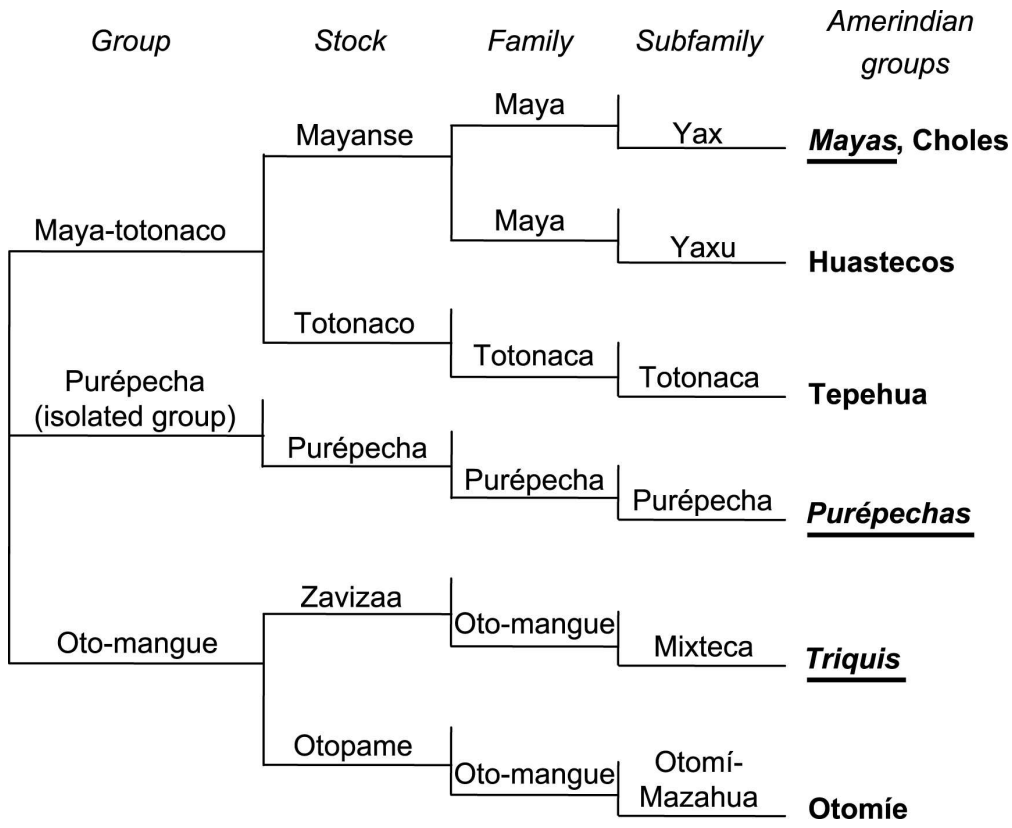

Figure 2. Linguistic classification of Native Mexican populations used for interpopulational analyses (Gordon, 2005; INALI, 2008). Underlined groups are reported on in this study.
145x117mm (300 x 300 DPI)

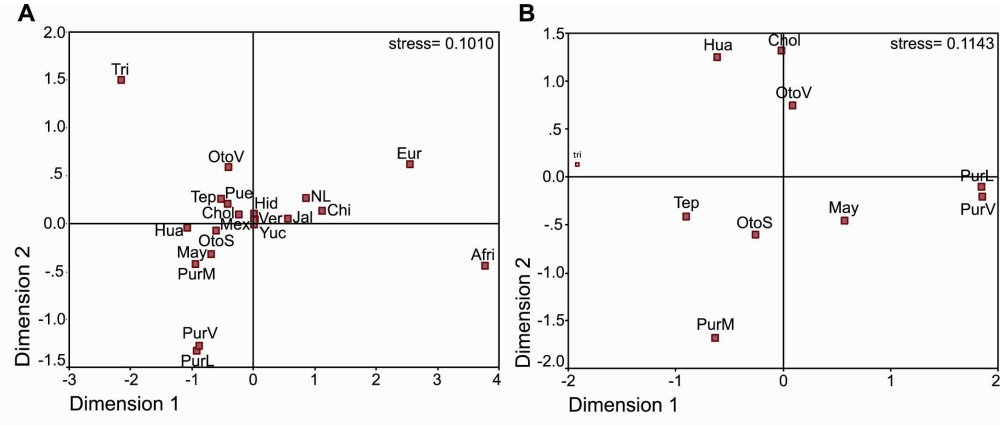

Figure 3. Multidimensional scaling (MDS) plot based on normalized FST distances between (A) Mestizos, Native American, and Ancestral populations (European and African); (B) only Native American groups. See Table I for description of abbreviations.
176x74mm (300 x 300 DPI)

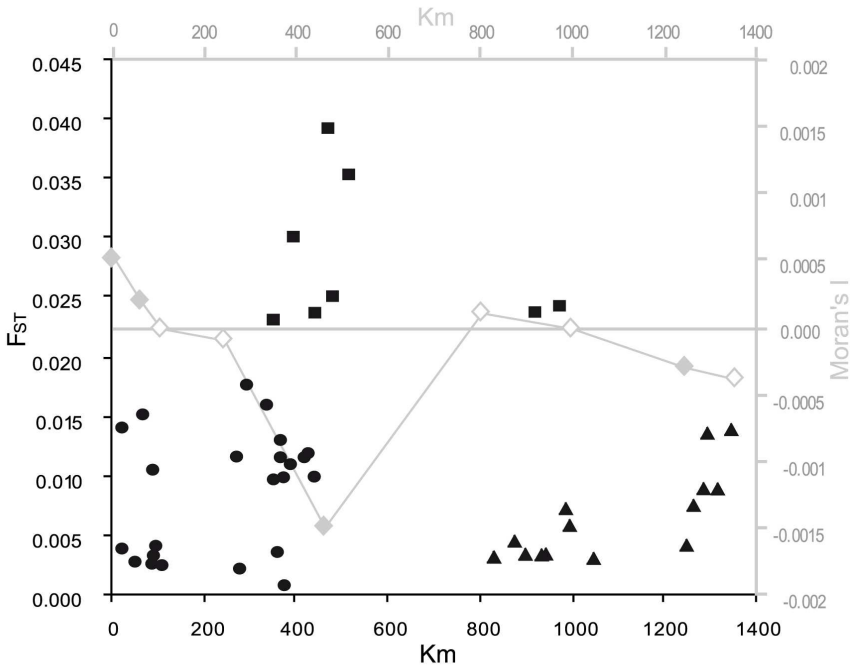

Figure 4. Overlapped plots representing correlation between geographical and genetic distances (black lines) and the AIDA autocorrelogram (grey lines). Correlation plot displays the following three groups: Purépechas, Otomíes, Huastecos, and Tepehuas (black circles); Triquis (black squares), and Mayas and Choles (black triangles). In the autocorrelogram, filled diamonds indicate significant p-values ($p < 0.05$).
170x121mm (300 x 300 DPI)

Supplementary Tables

S1. Allele frequency distribution for 15 STR loci (Amp/STR® Identifiler™), and statistical parameters of forensic importance in Purépechas of Zipiajo (Mountain).

Allele	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
6						0.2589						0.0029			
6.3						0.0059									
7						0.5595								0.1160	
8			0.1084			0.0357	0.0208	0.0029				0.4613			
9	0.0029		0.0542			0.0238	0.3452	0.0299				0.0386	0.0029	0.0952	
9.3						0.1160									
10	0.1131		0.1777	0.2365			0.1517	0.3233		0.0059		0.0089	0.0239	0.0238	
11	0.0357		0.3162	0.2814			0.1428	0.2634				0.2291	0.0119	0.5476	
11.2										0.0059					
12	0.0952		0.3192	0.4161	0.0089		0.1875	0.3083		0.0297		0.2410	0.0688	0.1577	
12.2										0.0029					
13	0.4613		0.0090	0.0568	0.0029		0.1101	0.0628		0.2440		0.0178	0.1137	0.0535	
13.2										0.1250					
14	0.2113		0.0090	0.0029	0.0238		0.0416			0.2381	0.0238		0.1586	0.0059	
14.2										0.1636			0.0029		
15	0.0773		0.0030		0.5714			0.0089		0.0476	0.1398		0.2006		
15.2			0.0030							0.0714					
16	0.0029			0.0029	0.2410				0.0238	0.0416	0.4107		0.0898		
16.2										0.0238					
17				0.0029	0.0922				0.0565		0.1964		0.1047		
18					0.0297				0.0506		0.1815		0.1197		0.0180
19					0.0297				0.4017		0.0476		0.0149		0.0572
20									0.1458				0.0509		0.0722
21									0.0238				0.0119		0.1024
22									0.0416				0.0179		0.1024
23									0.1428						0.1475
24		0.0029							0.0267				0.0029		0.1144
24.2		0.0089													
25									0.0773						0.2289
26									0.0059				0.0029		0.0662
27															0.0873
28		0.0654													0.0030
29		0.2142							0.0029						
30		0.1726													
30.2		0.0089													
31		0.1041													
31.2		0.1220													
32		0.0178													
32.2		0.1547													

33.2		0.1160													
34		0.0059													
34.2		0.0059													
MAF	0.0162	0.0175	0.0169	0.0156	0.0151	0.0146	0.0169	0.0171	0.0165	0.0179	0.0178	0.0153	0.0181	0.0158	0.0173
PD	0.8826	0.9581	0.8911	0.8403	0.7966	0.7966	0.9227	0.8606	0.9277	0.9464	0.8724	0.8495	0.9687	0.8421	0.9679
PE	0.4413	0.6281	0.5149	0.3507	0.2923	0.2327	0.5408	0.5597	0.4797	0.6739	0.6623	0.3144	0.6837	0.3872	0.5681
TPI	1.7143	2.7097	2.0244	1.4153	1.2537	1.1053	2.1538	2.2568	1.8667	3.1111	3.0000	1.3125	3.2115	1.5273	2.3056
PIC	0.6788	0.8397	0.7114	0.6307	0.5580	0.5508	0.7595	0.6757	0.7608	0.8098	0.6998	0.6202	0.8678	0.6170	0.8579
H	0.7083	0.8155	0.7530	0.6467	0.6012	0.5476	0.7679	0.7784	0.7321	0.8393	0.8333	0.6190	0.8443	0.6726	0.7831
HWE*	0.6069	0.1310	0.0602	0.0488	0.7339	0.5225	0.1202	0.5839	0.1169	0.3724	0.0758	0.4887	0.0480	0.2751	0.1622

MAF: minimum allele frequency; PD: power of discrimination; PE: power of exclusion; TPI = typical paternity index; PIC: polymorphism information content; H: heterozigosity expected; HWE: Hardy-Weinberg equilibrium test (p-value). * Bonferroni correction to evaluate HWE test (p< 0.0033)

S2. Allele frequency distribution for 15 STR loci (Amp*f*STR® Identifiler™), and statistical parameters of forensic importance in Purépechas from Angahuan (Valley).

Allele	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
5						0.0049									
5.3						0.0049									
6						0.3398									
7						0.3495								0.0728	
8			0.0588				0.0147					0.5097			
9	0.0049		0.0049	0.0050		0.0243	0.3431	0.0340				0.0049		0.0146	
9.3						0.2718									
10	0.0534		0.2108	0.1634		0.0049	0.2255	0.2573		0.0098				0.0194	
11	0.0291		0.3775	0.3911	0.0049		0.1863	0.2670				0.1990	0.0147	0.5971	
12	0.0340		0.3284	0.3317			0.0931	0.3883		0.0098		0.2767	0.0833	0.2961	
13	0.4612		0.0196	0.1040	0.0049		0.0882	0.0534		0.2647			0.0784		
13.2										0.1569					
14	0.2864			0.0050			0.0490			0.2647	0.0686	0.0097	0.2255		
14.2										0.0490					
15	0.1311				0.4757					0.1029	0.0539		0.1078		
15.2										0.0294					
16					0.3932				0.0049	0.0833	0.4167		0.1912		
16.2										0.0294					
17					0.0825				0.0728		0.3137		0.1912		
18					0.0388				0.0097		0.1275		0.0441		0.0248
19									0.4029		0.0196		0.0098		0.1535
20									0.2039				0.0049		
21									0.0049				0.0196		0.2178
22									0.0728				0.0294		0.1436
22.2		0.0049													
23									0.1796						0.0297
24									0.0388						0.1436
24.2		0.0922													
25									0.0097						0.1337
26															0.1436
27															0.0099
28		0.0388													
29		0.1505													
30		0.1553													
31		0.1019													
31.2		0.1845													
32		0.0049													
32.2		0.1845													
33.2		0.0583													
34.2		0.0243													
MAF	0.0254	0.0304	0.0259	0.0258	0.0246	0.0257	0.0286	0.0260	0.0275	0.0286	0.0269	0.0256	0.0284	0.0238	0.0294
PD	0.8455	0.9557	0.8499	0.8566	0.7705	0.8378	0.9045	0.8453	0.8977	0.9381	0.8416	0.7620	0.9479	0.7179	0.9527
PE	0.3974	0.8014	0.4222	0.3881	0.3298	0.4268	0.6623	0.4574	0.5919	0.6623	0.5182	0.4119	0.6434	0.2595	0.6978

TPI	1.5606	5.1500	1.6452	1.5303	1.3553	1.6613	3.0000	1.7758	2.4524	3.0000	2.0400	1.6094	2.8333	1.1704	3.3667
PIC	0.6352	0.8449	0.6461	0.6455	0.5343	0.6241	0.7460	0.6545	0.7189	0.7892	0.6561	0.5574	0.8301	0.4821	0.8289
H	0.6796	0.9029	0.6961	0.6733	0.6311	0.6990	0.8333	0.7184	0.7961	0.8333	0.7549	0.6893	0.8235	0.5728	0.8515
HWE *	0.5272	0.3438	0.0877	0.4143	0.8747	0.8478	0.8218	0.1703	0.5975	0.9131	0.2729	0.0183	0.0230	0.5562	0.8851

MAF: minimum allele frequency; PD: power of discrimination; PE: power of exclusion; TPI = typical paternity index; PIC: polymorphism information content; H: heterozigosity expected; HWE: Hardy-Weinberg equilibrium test (p-value). * Bonferroni correction to evaluate HWE test ($p < 0.0033$)

S3. Allele frequency distribution for 15 STR loci (Amp*f*STR® Identifiler™), and statistical parameters of forensic importance in Purépechas from Puacuaro (Lake).

Allele	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
6						0.3548									
7						0.3306								0.121	
8			0.0738									0.4655			
9	0.0081		0.0082			0.0161	0.379	0.0323							
9.2						0.0081									
9.3						0.2903									
10	0.0403		0.1393	0.1475			0.25	0.2016		0.0088				0.0323	
11	0.0484		0.4016	0.3852			0.1532	0.2581				0.2328		0.5726	
12	0.0565		0.377	0.3607			0.0565	0.4274		0.0175		0.3017	0.0702	0.2661	
12.2										0.0088					
13	0.3952			0.0984			0.129	0.0726		0.2368			0.0614	0.0081	
13.2										0.1667					
14	0.2581			0.0082	0.0565		0.0323	0.0081		0.2018	0.0776		0.2544		
14.2										0.0263					
15	0.1855				0.4758					0.1579	0.0862		0.0614		
15.2										0.0175			0.0088		
16	0.0081				0.379				0.0242	0.114	0.3966		0.2018		0.0082
16.2										0.0439					
17					0.0806				0.0484		0.2845		0.2193		
18					0.0081				0.0242		0.1121		0.0614		0.0082
19									0.4194		0.0431		0.0351		0.1311
20									0.2258				0.0088		
21									0.0081				0.0088		0.1885
22									0.0484						0.1475
23									0.1774				0.0088		0.041
24									0.0242						0.1393
24.2		0.0726													
25															0.1557
26															0.1557
27															0.0246
28		0.0081													
29		0.1532													
30		0.1694													
30.2		0.0081													
31		0.0806													
31.2		0.1613													
32.2		0.2661													
33.2		0.0484													
34.2		0.0323													
MAF	0.0413	0.0452	0.041	0.0442	0.0452	0.0422	0.0446	0.0405	0.0436	0.0483	0.049	0.043	0.0483	0.0387	0.0478
PD	0.8954	0.9454	0.8272	0.8315	0.6738	0.8096	0.8871	0.8663	0.8897	0.9406	0.8401	0.7907	0.9394	0.7508	0.9465
PE	0.4184	0.6416	0.3632	0.5455	0.6416	0.4693	0.6111	0.3710	0.5521	0.6121	0.6847	0.3624	0.6121	0.2683	0.7323
TPI	1.6316	2.8182	1.4524	2.1786	2.8182	1.8235	2.5833	1.4762	2.2143	2.5909	3.2222	1.4500	2.5909	1.1923	3.8125

PIC	0.6957	0.8159	0.6104	0.6334	0.5479	0.6128	0.7121	0.6554	0.6996	0.8130	0.6947	0.5653	0.8076	0.5266	0.8378
H	0.6935	0.8226	0.6557	0.7705	0.8226	0.7258	0.8065	0.6613	0.7742	0.8070	0.8448	0.6552	0.8070	0.5806	0.8689
HWE*	0.7216	0.7078	0.6751	0.8954	0.0029	0.2366	0.7843	0.5635	0.6443	0.3687	0.2684	0.8256	0.2848	0.2654	0.5605

MAF: minimum allele frequency; PD: power of discrimination; PE: power of exclusion; TPI = typical paternity index; PIC: polymorphism information content; H: heterozigosity expected; HWE: Hardy-Weinberg equilibrium test (p-value). * Bonferroni correction to evaluate HWE test ($p < 0.0033$)

S4. Allele frequency distribution for 9 STR loci (AmpF*l*STR® Profiler Plus™), and statistical parameters of forensic importance in the Triquis.

Allele	D851179	D21511	D75820	D351358	D135317	VWA	D18S51	D55818	FGA
7								0.01389	
8	0.00463		0.08333						
9					0.35648			0.06019	
10	0.25463		0.28704		0.08333			0.04167	
11	0.05093		0.36111		0.11574			0.67593	
12	0.08333		0.19907	0.00926	0.18056		0.09722	0.20833	
13	0.43519		0.06944		0.15741		0.13426		
13.2							0.00463		
14	0.09722			0.06019	0.10185	0.00926	0.10185		
15	0.06481			0.65741	0.00463	0.03241	0.15741		
16	0.00926			0.20833		0.63889	0.06019		
17				0.06019		0.22685	0.24537		
18				0.00460		0.06481	0.10648		
19						0.01852	0.00463		0.09722
19.2									0.00463
20						0.00926			0.00463
21							0.02315		0.02315
22							0.00463		0.07407
23							0.01389		0.09722
24							0.02315		0.41204
25							0.00926		0.18056
26							0.01389		0.09259
28									0.00463
29		0.18056							0.00926
30		0.29630							
31		0.18056							
31.2		0.12963							
32		0.04167							
32.2		0.10648							
33.2		0.04167							
34.2		0.02315							
FAM	0.02544	0.02861	0.02302	0.02289	0.02631	0.02252	0.02733	0.02118	0.02528
PD	0.87439	0.92995	0.88580	0.70799	0.91735	0.74811	0.96313	0.69239	0.91598
PE	0.50976	0.77278	0.28208	0.27135	0.59218	0.24090	0.68032	0.15010	0.49406
TPI	2.00000	4.50000	1.22727	1.20000	2.45454	1.12500	3.17647	0.91525	1.92857
PIC	0.68569	0.79083	0.69145	0.47252	0.75666	0.48727	0.84607	0.44894	0.73843
H	0.75000	0.88889	0.59259	0.58333	0.79629	0.55556	0.84259	0.45370	0.74074
HWE*	0.65042	0.15033	0.00423	0.08302	0.47083	0.84357	0.86308	0.29691	0.38377

MAF: minimum allele frequency; PD: power of discrimination; PE: power of exclusion; TPI = typical paternity index; PIC: polymorphism information content; H: heterozigosity expected; HWE: Hardy-Weinberg equilibrium test (p-value).
* Bonferroni correction to evaluate HWE test (p< 0.0055).

S5. Allele frequency distribution for 9 STR loci (AmpF*l*STR® Profiler Plus™), and statistical parameters of forensic importance in the Mayas.

Allele	D8S1179	D21S11	D7S820	D3S1358	D13S317	VWA	D18S51	D5S818	FGA
7			0.00556					0.07778	
8			0.02222		0.02222			0.00556	
9	0.00556		0.03889		0.29444			0.03333	
10	0.03889		0.20000		0.18889		0.01111	0.07778	
11	0.04444		0.34444		0.15000		0.01111	0.57778	
12	0.11667		0.35556		0.21111		0.07222	0.17778	
13	0.39444		0.03333		0.06667	0.00556	0.09444	0.05000	
13.2							0.00556		
14	0.27778			0.05556	0.06667	0.10556	0.18889		
15	0.10000			0.56111		0.05000	0.15000		
16	0.00556			0.26667		0.39444	0.13889		
17	0.01667			0.07778		0.28889	0.14444		
18				0.02778		0.12778	0.11667		0.01111
19				0.01111		0.02222	0.04444		0.05556
20						0.00556	0.01667		0.05000
21							0.00556		0.11111
21.2									0.00556
22									0.05556
23									0.10000
24									0.17778
25									0.24444
26		0.00556							0.15556
26.2									0.00556
27		0.00556							0.02222
28		0.03333							0.00556
29		0.23333							
29.2		0.01667							
30		0.23333							
30.2		0.05000							
31		0.10000							
31.2		0.08889							
32.2		0.14444							
33		0.00556							
33.2		0.08333							
FAM	0.03108	0.03220	0.02966	0.02661	0.03284	0.02966	0.03319	0.02844	0.03319
PD	0.88960	0.94840	0.86099	0.79410	0.91510	0.87460	0.95830	0.81430	0.95310
PE	0.57860	0.66229	0.46347	0.22940	0.70580	0.46350	0.72800	0.36290	0.72800
TPI	2.36840	3.00000	1.80000	1.09760	3.46150	1.80000	3.75000	1.45160	3.75000
PIC	0.70260	0.82285	0.66039	0.55140	0.77300	0.68930	0.85770	0.58610	0.83650
H	0.78890	0.83330	0.72220	0.54440	0.85560	0.72220	0.86670	0.65560	0.86670
HWE	0.44304	0.17097	0.12665	0.47690	0.06516	0.19504	0.05293	0.25293	0.26113

MAF: minimum allele frequency; PD: power of discrimination; PE: power of exclusion; TPI = typical paternity index; PIC: polymorphism information content; H: heterozigosity expected; HWE: Hardy-Weinberg equilibrium test (p-value).
* Bonferroni correction to evaluate HWE test (p< 0.0055).