





Article

Genetic Distribution of Five Spinocerebellar Ataxia Microsatellite Loci in Mexican Native American Populations and Its Impact on Contemporary Mestizo Populations

Rocío Gómez ¹ , Yessica S. Tapia-Guerrero ² , Bulmaro Cisneros ³, Lorena Orozco ⁴, César Cerecedo-Zapata ^{2,5}, Elvia Mendoza-Caamal ⁴, Gerardo Leyva-Gómez ⁶ , Norberto Leyva-García ², Luis Velázquez-Pérez ⁷ and Jonathan J. Magaña ^{2,8,*} 

¹ Department of Toxicology, CINVESTAV-IPN, Mexico City 07360, Mexico; mrgomez@cinvestav.mx

² Laboratory of Genomic Medicine, Department of Genetics, National Rehabilitation Institute-Luis Guillermo Ibarra Ibarra (INR-LGII), Mexico City 14389, Mexico; yessicasarai@gmail.com (Y.S.T.-G.); misael_207@hotmail.com (C.C.-Z.); nleyva@inr.gob.mx (N.L.-G.)

³ Department of Genetics and Molecular Biology, CINVESTAV-IPN, Mexico City 07360, Mexico; bcisnero@cinvestav.mx

⁴ Laboratory of Immunogenomics and Metabolic Diseases, National Genomic Medicine Institute (INMEGEN), Mexico City 14610, Mexico; lorozco@inmegen.gob.mx (L.O.); emendoza@inmegen.gob.mx (E.M.-C.)

⁵ Rehabilitation and Social Inclusion Center of Veracruz (CRIS-DIF), Xalapa, Veracruz 91097, Mexico

⁶ Department of Pharmacy, School of Chemistry, Universidad Nacional Autónoma de México (UNAM); Mexico City 04510, Mexico; leyva@quimica.unam.mx

⁷ Cuban Academy of Sciences, La Havana 10100, Cuba; velazq63@gmail.com

⁸ Department of Bioengineering, School of Engineering and Sciences, Tecnológico de Monterrey, Campus Ciudad de México (ITESM-CCM), Mexico City 14380, Mexico

* Correspondence: maganasm@hotmail.com; Tel.: +52-(55)-5999-1000 (ext. 14708)



Citation: Gómez, R.; Tapia-Guerrero, Y.S.; Cisneros, B.; Orozco, L.; Cerecedo-Zapata, C.; Mendoza-Caamal, E.; Leyva-Gómez, G.; Leyva-García, N.; Velázquez-Pérez, L.; Magaña, J.J. Genetic Distribution of Five Spinocerebellar Ataxia Microsatellite Loci in Mexican Native American Populations and Its Impact on Contemporary Mestizo Populations. *Genes* **2022**, *13*, 157. <https://doi.org/10.3390/genes13010157>

Academic Editor: Toshifumi Yokota

Received: 30 November 2021

Accepted: 13 January 2022

Published: 16 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Spinocerebellar ataxias (SCAs) conform a heterogeneous group of neurodegenerative disorders with autosomal dominant inheritance. Five of the most frequent SCAs are caused by a CAG repeat expansion in the exons of specific genes. The SCAs incidence and the distribution of polymorphic CAG alleles vary among populations and ethnicities. Thus, characterization of the genetic architecture of ethnically diverse populations, which have undergone recent admixture and demographic events, could facilitate the identification of genetic risk factors. Owing to the great ethnic diversity of the Mexican population, this study aimed to analyze the allele frequencies of five SCA microsatellite loci (SCA1, SCA2, SCA3, SCA6, and SCA7) in eleven Mexican Native American (MNA) populations. Data from the literature were used to compare the allelic distribution of SCA loci with worldwide populations. The SCA loci allelic frequencies evidenced a certain genetic homogeneity in the MNA populations, except for Mayans, who exhibited distinctive genetic profiles. Neither pathological nor large normal alleles were found in MNA populations, except for the SCA2 pre-mutated allele in the Zapotec population. Collectively, our findings demonstrated the contribution of the MNA ancestry in shaping the genetic structure of contemporary Mexican Mestizo populations. Our results also suggest that Native American ancestry has no impact on the origin of SCAs in the Mexican population. Instead, the acquisition of pathological SCA alleles could be associated with European migration.

Keywords: spinocerebellar ataxias; CAG repeats; Native American population; Mexican population; allelic distribution; large normal alleles

1. Introduction

The autosomal dominantly inherited polyglutamine spinocerebellar ataxias (polyQ-SCAs) constitute a group of neuromuscular disorders caused by a CAG trinucleotide repeat expansion in the coding region of different genes [1–3]. The abnormal expansion of CAG

repeats results, in turn, in the synthesis of mutant toxic proteins bearing a -polyglutamine tract (polyQ) [4]. Epidemiologically, SCA subtypes present a relative low frequency worldwide (3 in 100,000 inhabitants), with SCA3 being the most recurrent subtype, followed by SCA2, SCA1, SCA6, and SCA7; all of which are classified as polyQ-SCA [5,6]. Nonetheless, the distribution and prevalence of polyQ-SCAs vary between ethnicities and geographic regions. For instance, SCA2 is highly prominent in Cuba, while SCA3 and SCA7 have remarkable frequencies in Portuguese speaking countries and Mexico, respectively [5,7–9]. Studies to determine the prevalence of polyQ-SCAs and the length of CAG alleles in founder populations might help to elucidate the origin and spread of the expanded alleles. Founder effects have been found to be involved in the appearance and spread of SCAs in the American continents and the Caribbean [7]. Understanding of the genetic structure through SCA loci of Native American populations, which have undergone a recent admixture and several demographic events, is critical in order to clarify their diversity and demographic history, and to possibly identify relevant genetic factors for biomedical traits [10]. Nonetheless, the impact of the ethnic component on the contemporary genetic composition of MNA populations remains to be analyzed in depth. Specifically, under-representation of ethnic minorities, such as Native American populations, Latinos, and Africans, might cause a bias predicting disease risk in a given population [11].

The Mexican population presents a complex genetic architecture based on the admixture between Native American inhabitants (>50%), European migrants that arrived from Spain and Southern Europe (>40%), and West Africans brought to Mexico through the trans-Atlantic slave trade (<2%) [12]; such ethnic fusion originated the group known as “Mestizos” 500 years ago. Nevertheless, the contribution of Mexican Native American (MNA) populations to the study of disease-associated alleles remains unexplored. The present study aimed to analyze the influence of MNA populations in shaping the allelic distribution of five different SCA microsatellite loci (SCA1, SCA2, SCA3, SCA6, and SCA7) in the contemporary Mexican Mestizo populations. Further analysis of SCA loci data enabled us to infer genetic relatedness between MNA populations; nonetheless, we found no connection between the MNA population and the origin of SCA pathologies in Mexico. Our results support the hypothesis that European migration could have introduced pathological SCA alleles to the Mexican population.

2. Materials and Methods

2.1. Subjects

The field work was conducted in nine different states of Mexico, namely Chihuahua (CHH), Guerrero (GRO), Michoacán (MIC), Morelos (MOR), Oaxaca (OAX), San Luis Potosí (SLP), Sonora (SON), Yucatán (YUC), and Veracruz (VER) (Figure 1). Blood samples were collected from 566 unrelated individuals who identified themselves as Mexican Native Americans (MNA) (Figure 1). These individuals belong to 11 different ethnic groups, representing five linguistic families. They have lived in the same community at least during the last three generations (parents and grandparents), further speaking the Native languages own their community. Prior studies have demonstrated that the individuals analyzed herein are carriers of $95 \pm 5.7\%$ of NA ancestry [13]. The Research Committee of the National Rehabilitation Institute approved the data and sample collection. All MNA individuals signed an informed consent letter before sample obtention. To respect their identity, the ethnic names used in the present study correspond to their self-identification, which were obtained from the National Institute of the Indigenous Peoples.

2.2. Molecular Genetic Analysis

Genomic DNA was isolated from peripheral blood leukocytes with the Genra Puregene Blood kit (QIAGEN, Hilden, NRW, Düsseldorf, Germany), following the recommendations of the manufacturer. DNA samples (15 ng) were amplified using a Verity 96-Well Fast Thermal Cycler (Applied Biosystems, Carlsbad, CA, USA), with the reaction conditions and chimeric primers described previously [9,14]. The resulting amplicons were characterized

by capillary electrophoresis on an ABI Prism 3730XL system (Applied Biosystems, Foster City, CA, USA), following the manufacturer's instructions, as previously reported [9]. Those samples genotyped for SCA2 and SCA7 who exhibited a homozygous genotype, were further analyzed using a triplet repeat primed assay [15] in order to accurately determine the presence of large CAG (>130) alleles associated with SCA pathologies.

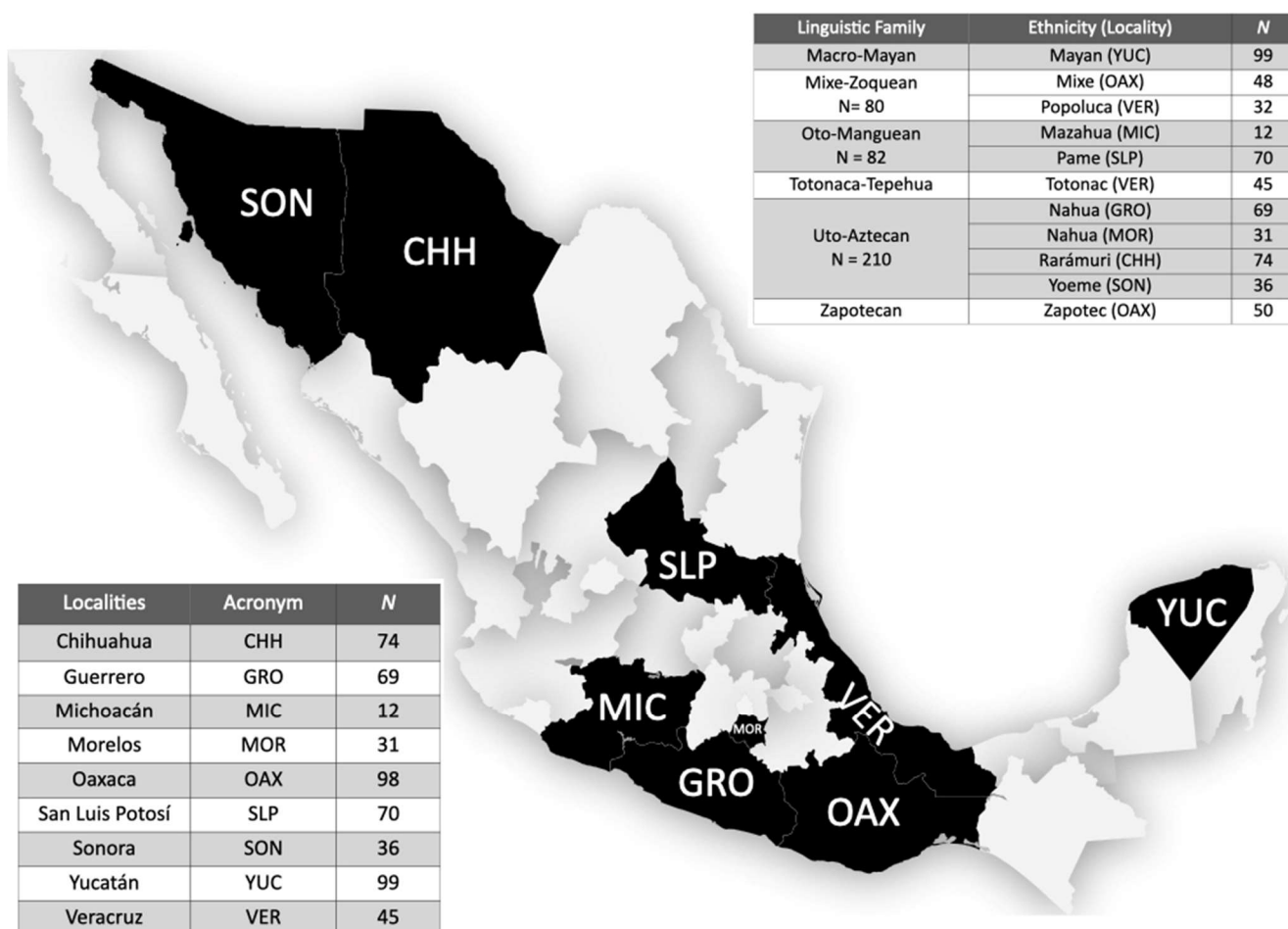


Figure 1. Map of Mexico showing the geographical origin of Mexican Native American populations included in the present study. The states where MNA populations live are shaded in black. *N*-number of subjects.

2.3. Statistical Analyses

The allele frequency was evaluated by analysis of molecular variance (AMOVA) and population differentiation analyses (R_{ST} values) using Arlequin v.3.5 [16]. The R_{ST} values were assessed using 10,000 permutations, and then visualized in a multidimensional scaling plot (MDS) using SPSS v.20 [17]. *p*-values were adjusted for false discovery rates in the R software [18,19]. The Hardy–Weinberg expectation (HW) was calculated with Weir and Cockerham's *F* statistics using Genèx v.4.05.2 with 10,000 permutations [20].

2.4. Comparative Data

Data from the literature were used to compare the allelic distribution of SCA loci between MNA and Mestizo populations with worldwide populations (Supplementary Information, Table S1). The studies where the allelic frequency was not adequately reported (i.e., bar plots) were excluded from this analysis. Thirty-seven worldwide populations ($n = 7532$ individuals) were included in the database for further analyses. Such a database was built through random sampling so as to avoid bias.

3. Results

3.1. Allele Frequencies and Genetic Diversity of Five SCA Loci in Mexican Native American Populations

The CAG repeat allele frequencies of the SCA1, SCA2, SCA3, SCA6, and SCA7 loci were estimated in Mexican Native American (MNA) populations belonging to six linguistic families. The allelic distribution of these microsatellite loci was found to vary to a certain extent between MNA populations (Supplementary Information, Tables S2–S6). Briefly, SCA1 and SCA3 loci showed the highest diversity of CAG repeat alleles, with 13 and 18 different alleles in MNA. In contrast, the SCA7 locus exhibited the lowest diversity, with only eight different alleles compared with the remaining SCA loci. Interestingly, all SCA alleles showed a lower diversity in MNA than the other worldwide populations. For SCA1, Mixes and Totonacs showed the highest number of CAG alleles ($k = 10$ alleles), while Mazahuas and *Rarámuris* (Tarahumaras) exhibited the lowest diversity ($k = 6$ and $k = 7$, respectively). Interestingly, the 31 CAG repeats allele was predominant in 7 out of 11 MNA populations analyzed. With respect to the SCA3 locus, Mayans ($k = 13$) and Nahuas-GRO ($k = 11$) exhibited the highest diversity of CAG repeat alleles, with the 23 CAG repeat alleles being the most frequent in all of the ethnicities genotyped. For SCA2, the most frequent allele was 22 CAG repeats in all of the MNA populations, while Mayans ($k = 9$) showed the highest number of CAG alleles. For SCA6, only 10 CAG alleles were identified in MNA, and *Rarámuris* exhibited the highest diversity ($k = 8$). Finally, for SCA7, the 10 CAG repeats allele was the predominant one in all of the 14 MNA analyzed. Of note, Nahuas-HGO were set apart from Nahuas-MOR, considering the SCA2 locus genetic data ($R_{ST} = 0.132$; p -value = 0.006). Overall, the SCA loci allelic frequencies of the MNA populations were similar to that reported for the contemporary Mexican mestizo population [9]. These findings imply that the genetic structure of the contemporary Mexican population has been shaped, at least in part, by the Native American ancestry.

Regarding the genetic structure of the SCA loci in MNA populations, we found a remarkable deviation from HW in eight out of eleven populations genotyped (Supplementary Information, Table S7) due to homozygous excess (F_{IS} range: 0.123–0.253).

3.2. Absence of SCA Pathological Alleles in Mexican Native Americans Populations

To provide insight into the origin of SCAs in the Mexican population, we search for the presence of SCA pathological alleles in MNA populations. None of the MNA populations genotyped harbored genetic vulnerability to SCAs, apart from a subject belonging to the Zapotec population, who carried a pre-mutated allele of 35 CAG repeats at the SCA2 locus (Supplementary Information, Table S3). As SCA2 large normal alleles were rarely found in MNA populations (<3.5%), the case mentioned above appears to be an isolated event. Large-normal alleles of the SCA1 and SCA3 loci were found in almost all the MNA populations genotyped (frequencies of <5% and <1%, respectively), while SCA6 and SCA7 large normal alleles were not detected.

3.3. Genetic Interplay between Mexican Native American Populations

To unveil the genetic interconnection between MNA populations, the SCA loci genotyping data were analyzed using analysis of molecular variance (AMOVA). Significant differences among MNA populations, as well as between MNA populations, were found (global variation source) (Table 1). Then, diverse model simulations were run to evidence the relative impact of a given population on the global genetic diversity of MAN populations. Interestingly, the removal of the Mayan population caused significant differences between populations to disappear (percentage of variation 0.023, p -value = 0.946). Although the removal of Nahuas-HGO decreased the variation between populations (1.896%), such differences remained significant (p -value ≤ 0.0001). Next, MNA populations were divided according to their spoken language (Uto-Aztecan (Nahuas, Tarahumaras and Yaquis), Oto-Manguean (Zapotecos and Pames), Totonaca-Tepehua (Totonacas), Mixe- Zoquean (Mixes), and Macro-Mayan (Mayas)) and their geographic origin. Neither the linguistic

nor the geographic criteria impacted the genetic diversity of MNA populations. data not shown). Next, to explore the diversity patterns of MNA populations in more detail, the SCA loci genetic data were analyzed using a population differentiation test based on R_{ST} values (Supplementary Information, Tables S8–S14). MNA populations were categorized by ethnicity, geography, and language family criteria. The ethnicity criterion plotted Mays apart from the remaining populations (Figure 2A), even when the Mestizo populations from the Central Valley of Mexico (CVM) and Tlaltetela (VER) were included (Figure 2B); these findings highlight the distinctive genetic background of Mayans. Mayans, Mazahuas, CVM Mestizos, Nahuas-Guerrero, and Tlaltetela Mestizos were scattered around a well-defined group positioned in the plot center, composed by Yaquis, Totonacs, Nahuas-Morelos, Pames, Popolucas, Zapotecs, Raramuris, and Mixes (Figure 2B). Utilization of the geographic criterion reinforced the genetic differences between Mayans and the remaining populations (Supplementary Information, Figure S1), with differences between Mayans and North populations being the most prominent ($p \approx 0.039$). On the other hand, applying of the language affiliation criterion resulted in a cluster conformed by the Mixe-Zoquean, Oto-Manguean, Totonaca-Tepehuan, Uto-Aztecan, and Zapotec linguistic families, all of which were set up apart from the Macro-Mayan family (Supplementary Information, Figure S2). MDS plots show the analysis of each SCA microsatellite separately, which are depicted in Supporting Information (Supplementary Information, Figures S3–S6).

Table 1. Genotyping data from five different SCA loci of contemporary Mexican Native American populations subjected to the analysis of molecular variance.

Source of Variation	Variation (%)	<i>p</i> -Value
Global		
Among populations	3.29	0.004
Among individuals within population	52.43	≤ 0.0001
Within individuals	44.27	≤ 0.0001
Geography		
Among populations	4.02	0.0001
Among individuals within population	52.04	≤ 0.0001
Within individuals	43.94	≤ 0.0001
Language		
Among populations	3.99	≤ 0.0001
Among individuals within population	51.98	≤ 0.0001
Within individuals	44.03	≤ 0.0001

Geography: Centre (Nahuas-Morelos); North-East (Pames and *Rarámuris*); North-West (*Yoemem*); South (Mixes, Nahuas-Guerrero, Popolucas, Totonacs, and Zapotecs); South-East (Mayans); West (Mazahuas). Language: Macro-Mayan (Mayans); Mixe-Zoquean (Mixes and Popolucas); Oto-Manguean (Mazahuas and Pames); Totonaca-Tepehuan (Totonacs); Uto-Aztecan (Nahuas, *Rarámuris*, and *Yoemem*); Zapotecan (Zapotecs).

3.4. Comparison of the SCA Loci Allelic Distribution between Mexican Native American and Worldwide Populations

In order to have a broader landscape about the SCA loci allelic distribution, our findings were contrasted with previous reports from worldwide populations, using a population differentiation analysis based on R_{ST} values (Supplementary Information, Table S1). Concerning SCA1, two well-defined clusters were observed; the first included the Mestizos and all the MNA populations, except for *Yoemem*, which was placed in the second group encompassing European populations (Supplementary Information, Table S8 and Figure S7). The worldwide comparison of the SCA2 allelic distribution rendered a dense cluster in the center of the plot encompassing all populations, except for Yoruba (Supplementary Information, Table S9 and Figure S8). Of note, Zapotec, CVM-Mestizo, and Cuban populations exhibited a closed genetic distance ($R_{ST} \approx -0.008$, p -value = 0.732) (Figure 3).

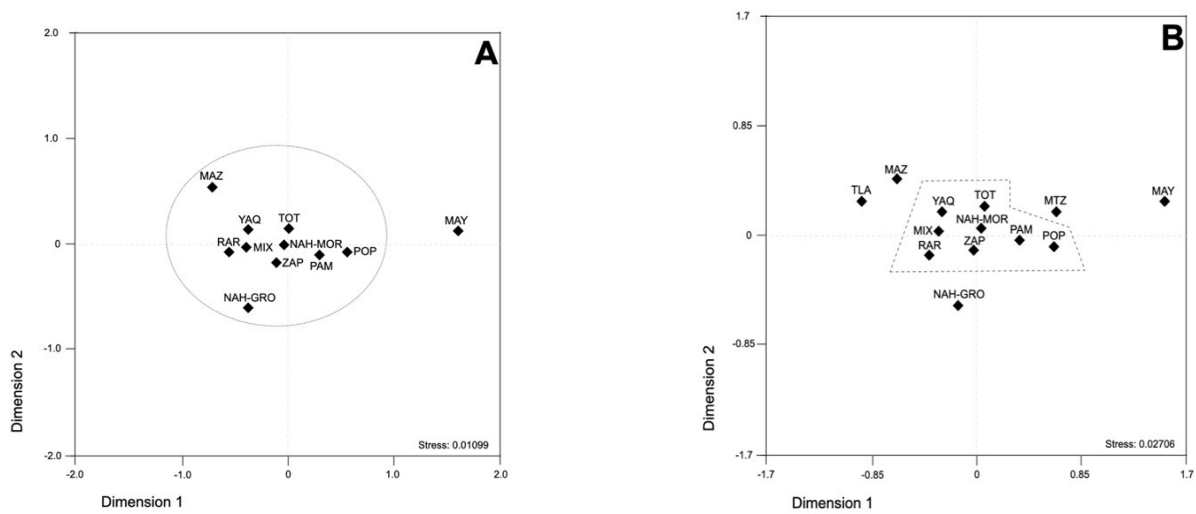


Figure 2. Multidimensional scale plot of *RST* values using five SCA microsatellite loci data, and the ethnicity criterion for contemporary Mexican Native Americans (A) and contemporary Mexican Native Americans and the Mexican Mestizos from the Central Valley of Mexico and Tlaltetela Veracruz (B). Mayans-MAY; Mazahuas-MAZ; Mixes-MIX; Nahuas-Guerrero-NAH-GRO; Nahuas-Morelos-NAH-MOR; Pames-PAM; Popolucas-POP; Rarámuris (Tarahumaras)-RAR; Yoemem (Yaquis)-YOE; Mexican Mestizos from the Central Valley of Mexico—CMV; Mexican Mestizos from Tlaltetela, Veracruz-TLA. Dotted lines enclose those populations that show genetic connections between them (no significant genetic distances); *p*-values were adjusted with the method of false discovery rates in R-software.

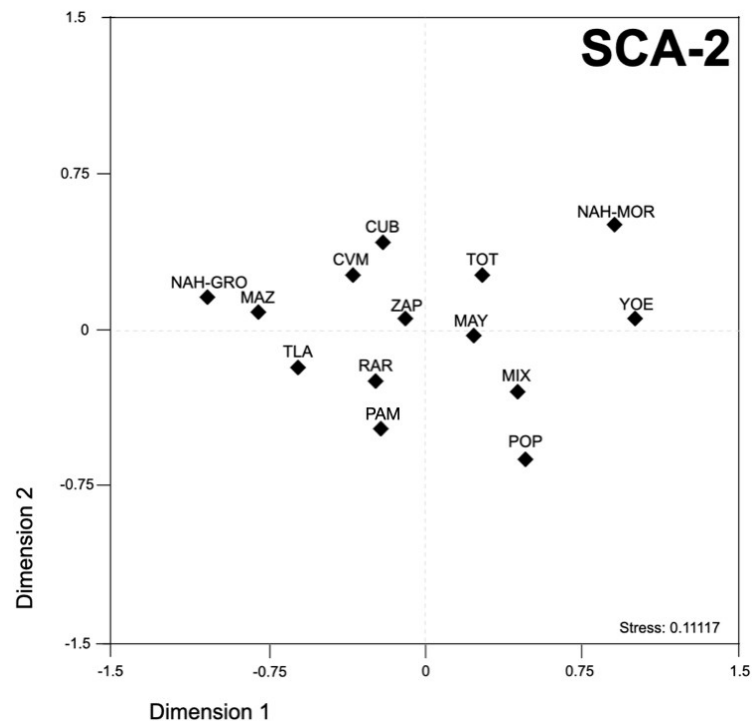


Figure 3. Multidimensional scale plot of *RST* values using the SCA2 locus data from Mexican Native American, Mestizos, and Cuban populations. Mayans-MAY; Mazahuas-MAZ; Mixes-MIX; Nahuas-Guerrero-NAH-GRO; Nahuas-Morelos-NAH-MOR; Pames-PAM; Popolucas-POP; Rarámuris (Tarahumaras)-RAR; Yoemem (Yaquis)-YOE; Mexican Mestizos from the Central Valley of Mexico-CMV; Mexican Mestizos from Tlaltetela, Veracruz-TLA; Cuba-CUB. *p*-values were adjusted with the method of false discovery rates in R-software.

Analysis of the SCA3 showed a loose cluster in the center of the plot composed by both MNA and the worldwide population, with *Rarámuris* being plotted separately (Supplementary Information, Table S10 and Figure S9). It is worth to note the genetic proximity of Mestizos (CVM and Tlaltetela) with Finland and Portugal populations. The comparative analysis for the SCA6 locus failed to evidence any genetic proximity between MNA populations and the European ones (Supplementary Information, Table S11 and Figure S10). Nonetheless, this analysis confirmed the remarkable genetic distance between Nahuas from Guerrero and the remaining MNA populations (Supplementary Information, Figure S10). Finally, the comparison of the allelic distribution patterns of SCA7 sustained the distinctive genetic diversity of Mayans and evidenced the genetic proximity between Mestizos from Tlaltetela and Finland populations (Supplementary Information, Table S12 and Figure S11).

4. Discussion

In this study, we report our knowledge, for the first time, on the allelic distribution of five different SCA loci in Mexican Native American (MNA) populations. Further analysis of this data enabled us to ascertain the impact of Native American ancestry on the contemporary Mexican population structure, which could provide some clues for the geographic origin of SCA pathologies in Mexico.

The SCA loci allelic frequencies found in MNA populations resembled that of the contemporary Mexican population [9], which indicates that Native American ancestry shaped, at least in part, the genetic structure of the latter population. However, a low diversity of SCA loci alleles was found in MNA populations compared with the contemporary Mestizo population [9], implying that the interethnic admixture among Spanish, Native Americans, and African lineages might be the mechanism underlying the high diversity of SCA alleles observed in the contemporary Mexican population.

Departure from the HW equilibrium due to heterozygote deficiency was not an unexpected result in MNA populations, considering the fact that this phenomenon is a signature of inbreeding populations. The MNA genetic architecture seems to be influenced by geographic barriers that keep them isolated [21]. MNA populations are also deeply rooted in their traditions, maintaining thus their native language and cultural practices, despite 500 years of miscegenation [22]. Both features have favored the contemporary Natives to maintain their ancestral connections and to settle down in geographic areas fairly close to those of their ancestors [21]. We speculate that these ancient practices might have promoted the inheritance of almost identical alleles, thereby increasing homozygosity and causing HW departure.

4.1. Searching for Large SCA Alleles in Mexican Native Americans

It is thought that SCAs prevalence in a given population is driven by the presence of intermediate and/or large normal alleles [23,24]. Consistent with this, we found a low frequency of large normal alleles and virtually no pathological alleles in MNA populations, apart from a pre-mutated SCA2 allele (CAG₃₅) in the Zapotec group. Thus, the presence of SCA pathologies in the Mexican population, with a high prevalence of SCA2 and SCA7, could be the consequence of founding effects, as previously proposed [7]. Hypothetically, European settlers, who were predominantly Spaniards, could introduce the ancestral SCAs chromosomes to Mexico. The significant gene flow from Europe to Mexico has been demonstrated by the analysis of uniparental markers of the Y-chromosome non-recombining region (NRY), because miscegenation was sex-biased due to the arrival of only 6% of female migrants to the Americas during colonization [25]. NRY lineages carried by European colonizers (i.e., R1b, E1b, G, J, and I) have been found in Mexican Mestizos [26–29]. In fact, European ancestry has previously been linked to different SCA subtypes in the Americas (i.e., SCA2, SCA3, and SCA7) [5]. With respect to the presence of a SCA2 permutated allele in the Zapotec population, we hypothesized that this Native group might play a role in the origin of SCA2 in Mexico, by spreading pre-mutated alleles during

the colonial period, because Oaxaca was a trade bridge connecting itself with other Mexican states (Campeche, Chiapas, and Veracruz), as well as with Guatemala and Cuba during the Viceroyalty of Mexico (1521–1821) [30]. Taking into consideration that colonization occurred first in Cuba (1511) and later in Mexico (1519) [31], and that SCA2 is highly prevalent in the former country, we speculate that Europeans who carried premutation alleles arrived firstly to Cuba and then to Oaxaca. Further haplotype analysis using uniparental markers as well as runs of homozygosity and identity-by-descent analyses are required to test this assumption. It is thought that European migration is implicated in the introduction of SCAs pathological alleles into the American continents and the Caribbean [7]. In fact, founder effects are the driving force behind the origin of SCAs worldwide. For instance, specific founder effects are responsible for the high prevalence of SCA3 observed in the Portuguese/Azorean island of Flores and Northeastern mainland [32].

4.2. Genetic Landscape of Mexican Native American Populations Using SCA Microsatellite Loci

Currently, diverse MNA populations coexist in the geographically wealthy territory of Mexico. The interplay among MNA populations through time could have tailored their present genetic structure. Thus, characterization of the genetic markers in these populations could broaden the picture about the origin and spread of the SCA expanded alleles. Furthermore, the study of genetic markers with biomedical relevance, such as the SCAs microsatellite loci, could help to trace the origin of these pathologies, as well as to identify the genetic factors relevant for biological processes of the MNA and Mestizo populations of today. Nonetheless, these studies should be reinforced by the analysis of ancestral markers.

The genetic distance of MNA populations depicted minor differences between them, exhibiting a subtle percentage of variation (3.29%), which is consistent with other studies [33]. The analysis of the SCA1 microsatellite loci separated the *Yoemem* population from six out of eleven MNA populations studied. This finding could be explained by the European ancestry, not necessarily Spanish, acquired by its connection with USA populations [34]. The heterozygous excess of SCA3 loci ($F_{IS} = -0.186$) supports this hypothesis. Likewise, evaluation of SCA2 and SCA3 loci failed to provide a clear population differentiation of MNA populations. Noteworthy, analysis of the SCA2 locus showed a certain divergence between Nahuas populations ($R_{ST} = 0.132$; p -value = 0.006), which could be the result of a differential genetic structure between them; Nahuas-GRO exhibited high genetic diversity ($k = 4$), while Nahuas-MOR showed an endogamy behavior ($k = 2$). It could be argued that the gene flow of Nahuas-GRO with European and African ancestries [26,35–37] and with neighbor ethnicities occurred to a greater extent than that of Nahuas MOR. On the other hand, the diagnosis of SCA6 microsatellite clustered all MNA populations, with the exception of Nahuas-GRO and Mayans. Such grouping suggests a continuous gene-flow between MNA groups and might also indicate the absence of cultural boundaries [22]; the latter is supported by the linguistic connection that took place between all indigenous groups (except Mayans). In fact, the genetic connection among Meso- and Arido-American populations has been extensively reported [12,13,22,38,39]. Furthermore, the genetic fusion between most of the MNA populations could reflect the hegemony of the Aztec Empire [22]. Although Nahuas-GRO and Mayans exhibited a distinctive diversity between one another ($p = 0.028$), the genetic distance between them was not considerable ($R_{ST} = 0.037$). Similar results have been previously reported using NRY markers [27]. It is worth mentioning that both populations have undergone a distinctive gene flow from Europe through the Dutch, English, and French pirate invasion that occurred at the Yucatan Peninsula between the XVI and XVIII centuries [27,40,41].

The Mayan population was also segregated from five out of eleven MNA populations genotyped upon analysis of the SCA7 microsatellite locus. The genetic distance between Mayans and Mixes, *Yoemem* and Zapotecans, exhibited only a marginal significance, which could be due to the geographic closeness between Yucatan and Oaxaca [42]. Both the whole-genome sequencing and haplotyping of NRY markers (i.e., R1b, 10%; G-M201, 6%;

I-M258, 3%; and J-M305, 1%) support the European admixture in Mayans [13,40], while the European ancestry of Nahuas-GRO is low (4.2%) [35,36]. On the other hand, the *Yoemem* diaspora to the United States of America could explain, at least in part, the partial genetic segregation from Mayans [34]. On the other hand, the genetic distance between Mayans and Nahuas-GRO and *Rarámuris* were marked. A portion of these differences could be attributed to the linguistic families to which they belong. While *Rarámuris* belong to Uto-Aztecan, Mayans belong to the Macro-Mayan family [43]. With respect to Nahuas-GRO, their prominent African ancestry could lead them to acquire a distinctive genetic profile (~4%) [26,35,37].

The limitations of this study are that the SCA loci alleles are not the suitable genetic tool to reconstruct demographic and cultural histories of populations, because they are not neutral markers. Furthermore, differences in the sample size between populations could bias our conclusions.

5. Conclusions

In summary, our findings highlight the importance of analyzing MNA populations in order to better understand the influence of Native American ancestry on modelling the genetic structure of contemporary Mexican Mestizo populations. Furthermore, this study suggests that Native Americans had no influence on the origin of the most prevalent SCAs in Mexico, reinforcing the idea that the introduction of the pathological alleles into Mexico could be associated with European migrations.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/genes13010157/s1>. Figure S1: Multidimensional scale plot of R_{ST} values estimated from five SCA loci in MNA and Mestizos using geographic criterion. Figure S2: Multidimensional scale plot of R_{ST} values estimated from five SCA loci in MNA and Mestizos using the linguistic affiliation as criterion. Figure S3: Multidimensional scale plot of R_{ST} values estimated from the SCA1 locus in MNA and Mestizo populations. Figure S4: Multidimensional scale plot of R_{ST} values estimated from the SCA3 locus in MNA and Mestizo populations. Figure S5: Multidimensional scale plot of R_{ST} values estimated from the SCA6 locus in MNA and Mestizo populations. Figure S6: Multidimensional scale plot of R_{ST} values estimated from the SCA7 locus in MNA and Mestizo populations. Figure S7: Multidimensional scale plot of R_{ST} values estimated from the SCA1 locus in MNA and Mestizos and its comparison with worldwide populations. Figure S8: Multidimensional scale plot of R_{ST} values estimated from the SCA2 locus in MNA and Mestizos and its comparison with worldwide populations. Figure S9: Multidimensional scale plot of R_{ST} values estimated from the SCA3 locus in MNA and Mestizos and its comparison with worldwide populations. Figure S10: Multidimensional scale plot of R_{ST} values estimated from the SCA6 locus in MNA and Mestizos and its comparison with worldwide populations. Figure S11: Multidimensional scale plot of R_{ST} values estimated from the SCA7 locus in MNA and Mestizos and its comparison with worldwide populations. Table S1: Comparison data obtained from different worldwide populations. Table S2: Allele frequencies of SCA1 found in the contemporary MNA populations. Table S3: Allele frequencies of SCA2 found in the contemporary MNA populations. Table S4: Allele frequencies of SCA3 found in the contemporary MNA populations. Table S5: Allele frequencies of SCA6 found in the contemporary MNA populations. Table S6: Allele frequencies of SCA1 found in the contemporary MNA populations. Table S7: Hardy–Weinberg expectation of the contemporary MNA populations. Table S8: R_{ST} values estimated from SCA1 locus in the contemporary MNA populations and other worldwide populations. Table S9: R_{ST} values estimated from SCA2 locus in the contemporary MNA populations and other worldwide populations. Table S10: R_{ST} values estimated from SCA3 locus in the contemporary MNA populations and other worldwide population. Table S11: R_{ST} values estimated from SCA6 locus in the contemporary MNA populations and other worldwide populations. Table S12: R_{ST} values estimated from SCA7 locus in the contemporary MNA populations and other worldwide populations. Table S13: R_{ST} values estimated from all SCA loci in the contemporary MNA populations and other worldwide populations. Table S14: R_{ST} values estimated from all SCA loci in the contemporary MNA populations and Mexican populations.

Author Contributions: Conceptualization, J.J.M.; methodology, R.G. and Y.S.T.-G.; software, R.G.; validation, R.G., L.O., and E.M.-C.; formal analysis, R.G. and J.J.M.; investigation, R.G., Y.S.T.-G., C.C.-Z. and G.L.-G.; clinical analysis, C.C.-Z., N.L.-G. and L.V.-P.; resources, J.J.M.; data curation, R.G.; writing—original draft preparation, R.G. and J.J.M.; writing—review and editing, B.C., G.L.-G. and J.J.M.; visualization, B.C.; supervision, B.C., L.V.-P. and J.J.M.; project administration, J.J.M.; funding acquisition, J.J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a CONACYT grant (grant number 258043) to Jonathan J. Magaña.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and was approved by the Research Committee of the National Rehabilitation Institute (protocol INR-15/12).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Acknowledgments: We thank the enthusiastic participation of the Native American populations; their collaboration made this study possible. We thank Emilio Martínez-Cruz, Francisco Cid, and Hernán Cortés help with the logistics of sample collection and technical assistance.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Benarroch, L.; Bonne, G.; Rivier, F.; Hamroun, D. The 2021 version of the gene table of neuromuscular disorders (nuclear genome). *Neuromuscul. Disord.* **2020**, *30*, 1008–1048. [[CrossRef](#)]
2. Buijsen, R.A.M.; Toonen, L.J.A.; Gardiner, S.L.; van Roon-Mom, W.M.C. Genetics, Mechanisms, and Therapeutic Progress in Polyglutamine Spinocerebellar Ataxias. *Neurother. J. Am. Soc. Exp. Neurother.* **2019**, *16*, 263–286. [[CrossRef](#)]
3. Gardiner, S.L.; Boogaard, M.W.; Trompet, S.; de Mutsert, R.; Rosendaal, F.R.; Gussekloo, J.; Jukema, J.W.; Roos, R.A.C.; Aziz, N.A. Prevalence of Carriers of Intermediate and Pathological Polyglutamine Disease-Associated Alleles Among Large Population-Based Cohorts. *JAMA Neurol.* **2019**, *76*, 650–656. [[CrossRef](#)] [[PubMed](#)]
4. Borbolla-Jiménez, F.V.; Del Prado-Audelo, M.L.; Cisneros, B.; Caballero-Florán, I.H.; Leyva-Gómez, G.; Magaña, J.J. New Perspectives of Gene Therapy on Polyglutamine Spinocerebellar Ataxias: From Molecular Targets to Novel Nanovectors. *Pharmaceutics* **2021**, *13*, 1018. [[CrossRef](#)] [[PubMed](#)]
5. Klockgether, T.; Mariotti, C.; Paulson, H.L. Spinocerebellar ataxia. *Nat. Rev. Dis. Prim.* **2019**, *5*, 24. [[CrossRef](#)]
6. de Oliveira Scott, S.S.; Pedrosa, J.L.; Barsottini, O.G.P.; França-Junior, M.C.; Braga-Neto, P. Natural history and epidemiology of the spinocerebellar ataxias: Insights from the first description to nowadays. *J. Neurol. Sci.* **2020**, *417*, 117082. [[CrossRef](#)]
7. Rodríguez-Labrada, R.; Martins, A.C.; Magaña, J.J.; Vazquez-Mojena, Y.; Medrano-Montero, J.; Fernandez-Ruiz, J.; Cisneros, B.; Teive, H.; McFarland, K.N.; Saraiva-Pereira, M.L.; et al. Founder Effects of Spinocerebellar Ataxias in the American Continents and the Caribbean. *Cerebellum* **2020**, *19*, 446–458. [[CrossRef](#)]
8. Magaña, J.J.; Gómez, R.; Maldonado-Rodríguez, M.; Velázquez-Pérez, L.; Tapia-Guerrero, Y.S.; Cortés, H.; Leyva-García, N.; Hernández-Hernández, O.; Cisneros, B. Origin of the spinocerebellar ataxia type 7 gene mutation in Mexican population. *Cerebellum* **2013**, *12*, 902–905. [[CrossRef](#)]
9. Magaña, J.J.; Tapia-Guerrero, Y.S.; Velázquez-Pérez, L.; Cerecedo-Zapata, C.M.; Maldonado-Rodríguez, M.; Jano-Ito, J.S.; Leyva-García, N.; González-Piña, R.; Martínez-Cruz, E.; Hernández-Hernández, O.; et al. Analysis of CAG repeats in five SCA loci in Mexican population: Epidemiological evidence of a SCA7 founder effect. *Clin. Genet.* **2014**, *85*, 159–165. [[CrossRef](#)] [[PubMed](#)]
10. Sirugo, G.; Williams, S.M.; Tishkoff, S.A. The Missing Diversity in Human Genetic Studies. *Cell* **2019**, *177*, 26–31. [[CrossRef](#)] [[PubMed](#)]
11. Popejoy, A.B.; Fullerton, S.M. Genomics is failing on diversity. *Nature* **2016**, *538*, 161–164. [[CrossRef](#)] [[PubMed](#)]
12. Moreno-Estrada, A.; Gignoux, C.R.; Fernández-López, J.C.; Zakharia, F.; Sikora, M.; Contreras, A.V.; Acuña-Alonzo, V.; Sandoval, K.; Eng, C.; Romero-Hidalgo, S.; et al. Human genetics. The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science* **2014**, *344*, 1280–1285. [[CrossRef](#)]
13. Romero-Hidalgo, S.; Ochoa-Leyva, A.; Garcíarrubio, A.; Acuña-Alonzo, V.; Antúnez-Argüelles, E.; Balcazar-Quintero, M.; Barquera-Lozano, R.; Carnevale, A.; Cornejo-Granados, F.; Fernández-López, J.C.; et al. Demographic history and biologically relevant genetic variation of Native Mexicans inferred from whole-genome sequencing. *Nat. Commun.* **2017**, *8*, 1005. [[CrossRef](#)]
14. Dorschner, M.O.; Barden, D.; Stephens, K. Diagnosis of five spinocerebellar ataxia disorders by multiplex amplification and capillary electrophoresis. *J. Mol. Diagn.* **2002**, *4*, 108–113. [[CrossRef](#)]
15. Cagnoli, C.; Stevanin, G.; Michielotto, C.; Gerbino Promis, G.; Brussino, A.; Pappi, P.; Durr, A.; Dragone, E.; Viemont, M.; Gellera, C.; et al. Large pathogenic expansions in the SCA2 and SCA7 genes can be detected by fluorescent repeat-primed polymerase chain reaction assay. *J. Mol. Diagn.* **2006**, *8*, 128–132. [[CrossRef](#)] [[PubMed](#)]
16. Excoffier, L.; Lischer, H.E.L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **2010**, *10*, 564–567. [[CrossRef](#)]

17. Nie, N.; Hull, C.; Bent, D.J.C.S.C. *SPSS*, Version 20. IBM statistical package for the social sciences. Computer Software: Chicago, IL, USA, 2011.
18. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289–300. [[CrossRef](#)]
19. R Foundation for Statistical Computing. R: A Language and Environment for Statistical Computing. 2019. Available online: <http://www.r-project.org> (accessed on 29 August 2021).
20. Belkhir, K.; Chikhi, L.; Raufaste, N.B.F. *Logiciel Sous Windows TM Pour la Génétique des Populations*, GENETIX 4.05; Université Montpellier II: Montpellier, France, 2004.
21. García-Ortiz, H.; Barajas-Olmos, F.; Contreras-Cubas, C.; Cid-Soto, M.Á.; Córdova, E.J.; Centeno-Cruz, F.; Mendoza-Caamal, E.; Cicerón-Arellano, I.; Flores-Huacuja, M.; Baca, P.; et al. The genomic landscape of Mexican Indigenous populations brings insights into the peopling of the Americas. *Nat. Commun.* **2021**, *12*, 5942. [[CrossRef](#)] [[PubMed](#)]
22. Gómez, R.; Vilar, M.G.; Meraz-Ríos, M.A.; Véliz, D.; Zúñiga, G.; Hernández-Tobías, E.A.; Figueroa-Corona, M.D.P.; Owings, A.C.; Gaieski, J.B.; Schurr, T.G. Y chromosome diversity in Aztlan descendants and its implications for the history of Central Mexico. *iScience* **2021**, *24*, 102487. [[CrossRef](#)] [[PubMed](#)]
23. Takano, H.; Cancel, G.; Ikeuchi, T.; Lorenzetti, D.; Mawad, R.; Stevanin, G.; Didierjean, O.; Dürr, A.; Oyake, M.; Shimohata, T.; et al. Close associations between prevalences of dominantly inherited spinocerebellar ataxias with CAG-repeat expansions and frequencies of large normal CAG alleles in Japanese and Caucasian populations. *Am. J. Hum. Genet.* **1998**, *63*, 1060–1066. [[CrossRef](#)]
24. Laffita-Mesa, J.M.; Almaguer-Mederos, L.E.; Kourí, V.; Bauer, P.O.; Vázquez-Mojena, Y.; Cruz Mariño, T.; Velázquez-Pérez, L. Large normal alleles and SCA2 prevalence: Lessons from a nationwide study and analysis of the literature. *Clin. Genet.* **2014**, *86*, 96–98. [[CrossRef](#)]
25. Ongaro, L.; Scliar, M.O.; Flores, R.; Raveane, A.; Marnetto, D.; Sarno, S.; Gnechi-Ruscione, G.A.; Alarcón-Riquelme, M.E.; Patin, E.; Wangkumhang, P.; et al. The Genomic Impact of European Colonization of the Americas. *Curr. Biol.* **2019**, *29*, 3974–3986.e4. [[CrossRef](#)] [[PubMed](#)]
26. González-Sobrino, B.Z.; Pintado-Cortina, A.P.; Sebastián-Medina, L.; Morales-Mandujano, F.; Contreras, A.V.; Aguilar, Y.E.; Chávez-Benavides, J.; Carrillo-Rodríguez, A.; Silva-Zolezzi, I.; Medrano-González, L. Genetic Diversity and Differentiation in Urban and Indigenous Populations of Mexico: Patterns of Mitochondrial DNA and Y-Chromosome Lineages. *Biodemography Soc. Biol.* **2016**, *62*, 53–72. [[CrossRef](#)]
27. Santana, C.; Noris, G.; Meraz-Ríos, M.A.; Magaña, J.J.; Calderon-Aranda, E.S.; Muñoz, M.d.L.; Gómez, R. Genetic analysis of 17 Y-STRs in a Mestizo population from the Central Valley of Mexico. *Hum. Biol.* **2014**, *86*, 289–312. [[CrossRef](#)]
28. Saiz, M.; Alvarez-Cubero, M.J.; Lorente, J.A.; Alvarez, J.C.; Martínez-Gonzalez, L.J. Genetic structure in the paternal lineages of South East Spain revealed by the analysis of 17 Y-STRs. *Sci. Rep.* **2019**, *9*, 5234. [[CrossRef](#)]
29. Martínez-González, L.J.; Martínez-Espín, E.; Álvarez, J.C.; Albardaner, F.; Rickards, O.; Martínez-Labarga, C.; Calafell, F.; Lorente, J.A. Surname and Y chromosome in Southern Europe: A case study with Colom/Colombo. *Eur. J. Hum. Genet.* **2012**, *20*, 211–216. [[CrossRef](#)]
30. Márquez, E.A. *Zapotecos del Istmo de Tehuantepec*; Comisión Nacional para el Desarrollo de los Pueblos Indígenas: Ciudad de México, Mexico, 2007; ISBN 978-970-753-093-5.
31. Encyclopaedia Britannica Inc. The Spanish Conquest. Available online: <http://www.britannica.com> (accessed on 7 September 2021).
32. Martins, S.; Sequeiros, J. Origins and Spread of Machado-Joseph Disease Ancestral Mutations Events. *Adv. Exp. Med. Biol.* **2018**, *1049*, 243–254. [[CrossRef](#)]
33. Andrés, A.M.; Lao, O.; Soldevila, M.; Calafell, F.; Bertranpetit, J. Dynamics of CAG repeat loci revealed by the analysis of their variability. *Hum. Mutat.* **2003**, *21*, 61–70. [[CrossRef](#)]
34. Moctezuma-Zamarron, J.L. Yaquis. In *Pueblos Indígenas del Mexico Contemporaneo*; Comisión Nacional para el Desarrollo de los Pueblos Indígenas: Ciudad de México, México, 2007.
35. Cahua-Pablo, J.Á.; Cruz, M.; Tello-Almaguer, P.V.; Del Alarcón-Romero, L.C.; Parra, E.J.; Villerías-Salinas, S.; Valladares-Salgado, A.; Tello-Flores, V.A.; Méndez-Palacios, A.; Pérez-Macedonio, C.P.; et al. Analysis of admixture proportions in seven geographical regions of the state of Guerrero, Mexico. *Am. J. Hum. Biol. Off. J. Hum. Biol. Counc.* **2017**, *29*, e23032. [[CrossRef](#)] [[PubMed](#)]
36. Bonilla, C.; Gutiérrez, G.; Parra, E.J.; Kline, C.; Shriver, M.D. Admixture analysis of a rural population of the state of Guerrero, Mexico. *Am. J. Phys. Anthropol.* **2005**, *128*, 861–869. [[CrossRef](#)] [[PubMed](#)]
37. Silva-Zolezzi, I.; Hidalgo-Miranda, A.; Estrada-Gil, J.; Fernandez-Lopez, J.C.; Uribe-Figueroa, L.; Contreras, A.; Balam-Ortiz, E.; del Bosque-Plata, L.; Velazquez-Fernandez, D.; Lara, C.; et al. Analysis of genomic diversity in Mexican Mestizo populations to develop genomic medicine in Mexico. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 8611–8616. [[CrossRef](#)] [[PubMed](#)]
38. Sandoval, K.; Moreno-Estrada, A.; Mendizabal, I.; Underhill, P.A.; Lopez-Valenzuela, M.; Peñaloza-Espinosa, R.; Lopez-Lopez, M.; Buentello-Malo, L.; Avelino, H.; Calafell, F.; et al. Y-chromosome diversity in Native Mexicans reveals continental transition of genetic structure in the Americas. *Am. J. Phys. Anthropol.* **2012**, *148*, 395–405. [[CrossRef](#)] [[PubMed](#)]
39. Reich, D.; Patterson, N.; Campbell, D.; Tandon, A.; Mazieres, S.; Ray, N.; Parra, M.V.; Rojas, W.; Duque, C.; Mesa, N.; et al. Reconstructing Native American population history. *Nature* **2012**, *488*, 370–374. [[CrossRef](#)] [[PubMed](#)]

40. Perez-Benedico, D.; La Salvia, J.; Zeng, Z.; Herrera, G.A.; Garcia-Bertrand, R.; Herrera, R.J. Mayans: A Y chromosome perspective. *Eur. J. Hum. Genet.* **2016**, *24*, 1352–1358. [[CrossRef](#)]
41. Jorge, V.O. Piratas en tierra adentro. Estrategia defensiva de una ciudad novohispana. Siglos XVI al XVIII. In *Millars Espai i Història*; Universitat Jaume I: Castelló de la Plana, España, 2003; Volume 26. [[CrossRef](#)]
42. Martínez-Cortés, G.; Nuño-Arana, I.; Rubi-Castellanos, R.; Vilchis-Dorantes, G.; Luna-Vázquez, A.; Coral-Vázquez, R.M.; Canto-Cetina, T.; Salazar-Flores, J.; Muñoz-Valle, J.F.; Sandoval-Mendoza, K.; et al. Origin and genetic differentiation of three Native Mexican groups (Purépechas, Triquis and Mayas): Contribution of CODIS-STRs to the history of human populations of Mesoamerica. *Ann. Hum. Biol.* **2010**, *37*, 801–819. [[CrossRef](#)]
43. Rangel-Villalobos, H.; Muñoz-Valle, J.F.; González-Martín, A.; Gorostiza, A.; Magaña, M.T.; Páez-Riberos, L.A. Genetic admixture, relatedness, and structure patterns among Mexican populations revealed by the Y-chromosome. *Am. J. Phys. Anthropol.* **2008**, *135*, 448–461. [[CrossRef](#)]