Forensic Parameters and Genetic Variation of 15 Autosomal STR Loci in Mexican Mestizo Populations from the States of Yucatan and Nayarit

Lizbeth Gonzalez-Herrera[a],*, LourdesVega-Navarrete[b], Cinthia Roche-Canto[c], Jorge Canto-Herrera[a], Dany Virgen-Ponce[d], Gustavo Moscoso-Caloca[d], Erick Delgado-Nájar[d], Betzabet Quintanilla-Vega[e] and Ricardo M. Cerda-Flores [f,g]

[a]*Laboratorio de Genética, Centro de Investigaciones Regionales, Universidad Autónoma de Yucatán, Mérida, Yucatán, México*

[b]*Centro de Estudios e Investigaciones ANIGEN, Distrito Federal, México*

[c]*Procuraduría General de Justicia-Estado de Yucatán, Mérida, Yucatán, México*

[d]*Procuraduría General de Justicia-Estado de Nayarit, Tepic, Nayarit, México*

[e]*Departmento de Toxicología, CINVESTAV-IPN, México*

[f]*Unidad de Biología Molecular, Genómica y Secuenciación, Centro de Investigación y Desarrollo en Ciencias de la Salud, Universidad Autónoma de Nuevo León (UANL), México*

[g]*Facultad de Enfermería, UANL, México*

Abstract: Allele frequencies and six forensic parameters for autosomal short tandem repeats (STR), as well as the genetic structure in two Mexican mestizo populations were determined. A total of 356 unrelated individuals from the states of Yucatan, at Southern Mexico (N=120) and Nayarit at Pacific Coast (N=236) were PCR-typed for 15 STR loci using the Power Plex 16 Promega kit. Genotype distribution per each locus was in agreement with Hardy-Weinberg expectations for the 15 STR in both populations. Penta E and FGA were the STR with the highest heterozygousity index, 0.8893 in Yucatan, and 0.9057 in Nayarit. Penta E also showed the highest power of discrimination in both populations, with 0.972 in Yucatan and 0.980 in Nayarit. For both populations, the combined power of exclusion and power of discrimination were higher than 0.9999999. Based on a principal component analysis, Yucatan and Nayarit were located between Spain and Mexican indigenous populations (Mayas from Campeche and Yucatan). Both populations were located in two separate clusters, suggesting different genetic structure. Our report contributes with STR data in Mexican populations, specially the allele frequencies for Penta D and Penta E in Yucatan and provides population data in Nayarit where no STR database was available.

Keywords: STR, population data, allele frequencies, mestizo, mexicans.

1. POPULATION

Mexican-mestizos are the result of admixture of Spaniards, Amerindians, and Africans. Mestizos are the most widely distributed population in both urban and rural regions throughout the country and constitute nearly 93% of the total population. Yucatan is located at Southeastern Mexico and has a high density of population with Mayan ascendancy, whereas Nayarit is located at Pacific Coast of Mexico and shows an important influence from the ethnical cultures of the Coras and Huichols. The population was composed of 356 unrelated healthy Mestizo volunteers residing in Yucatan (N= 120) and Nayarit (N= 236), who were selected as Mestizo with the criterion that they did not belong to some specific Mexican ethnic groups and only those individuals who for two generations, including their own had been born in Mexico. A Mexican Mestizo is defined as someone born in Mexico who is a descendant of the original autochthonous inhabitants of the region and of individuals, mainly Spaniards of Caucasian and/or Black origin, who came to America during the 16th century. They were recruited by opened invitation and from casework of routine paternity testing. All participants signed informed consent according to ethical issues of the Declaration of Helsinki.

2. DNA EXTRACTION

DNA Extraction was performed by collecting one punch of bloodstains on FTA Classic Cards (Whatman) and bucal cells. DNA was purified using FTA purification reagent (Whatman WB 120204) in a Maxwell 16 Promega, following the manufacturer's directions.

*Address correspondence to this author at the Laboratorio de Genética, Centro de Investigaciones Regionales, Universidad Autónoma de Yucatán, Mérida, Yucatán, México; Tel: (52) 999 929 5755; Fax: (52) 999 923 6120; E-mail: lizbeth@uady.mx

3. PCR

Amplification was performed by multiplex PCR in a Perkin-Elmer model 9700 thermal cycler using the commercial kit PowerPlex 16 System (Promega).

4. TYPING

Separation of amplified DNA fragments was performed employing multicapillary electrophoresis in an ABI Prism 3100 Avant Genetic Analyzer (Applied Biosystems) and ILS (Internal Lane Standard) to each sample. Genotyping was performed after comparison of STR fragments to the allelic ladder and internal size standard using GeneMapper ID v3.2 software. All steps were according to the Laboratory internal control standards and kit controls and proficiency testing provided by Promega Corporation.

5. STATISTICAL ANALYSIS

The direct gene counting method was used to ascertain STR allele frequencies. Parameters of population genetics including Matching Probability (MP), Power of Discrimination (PD), Polymorphic Information Content (PIC), Power of Exclusion (PE) and Typical Paternity Index (TPI) were calculated with the PowerStats 1.2 Software [1]. The Hardy–Weinberg equilibrium deviation for each locus was verified with the exact test by means of 2000 permutations, using the Arlequin software package [2]. Bonferroni's correction was used for Hardy–Weinberg equilibrium test, which assumes that a 0.05 significance level obtained for 15 tests (one per locus) yields an actual significance threshold of 0.0033.

The R x C contingency table exact test [3] was used to compare allele distribution for each STR of the CODIS found in Yucatan and Nayarit, assessing population differentiation with other ethnically related populations; such as Mexican Mestizos from Central Mexico [4], Mexico city [5], Valley of Mexico [6], Nuevo Leon [7], Jalisco, Puebla, and a Yucatan population previously tested with Identifiller kit [8]. A principal component analysis was performed among the populations mentioned above, except Central Mexico, using SPSS-PC software, considering the Allele frequencies of 10 STR: D18S51, D21S11, TH01, D3S1358, FGA, D8S1179, vWA, D7S820, D13S317, D5S818, because only these STR are reported for the representative ancestral populations used in this study: the Southwestern region of Spain [9], three Southwestern Native-American communities (Apache, Navajo and Pueblo), which represented Native-American ancestral allele frequencies [10] and two Mayan native populations from Yucatan and Campeche [11].

6. RESULTS AND DISCUSSION

Allele frequencies of 15 autosomal STR of the Power Plex system for Mexican Mestizos from Yucatan and Nayarit are shown in Tables **1** and **2**, respectively. Genotype distribution per each locus was in agreement with Hardy–Weinberg expectations for all 15 STR after Bonferroni´s correction. Six forensic parameters of the 15 autosomal STR for each population are shown in Table **3**. Combined power of exclusion and combined power of discrimination were higher than 0.9999999, both in Yucatan and Nayarit. The combined matching probability based on the two most common alleles through the 15 STR for each population was

1.87×10^{-13} or 1 in 5.3 $X10^{12}$ (1 in 5 trillion) in Yucatan and 2.13×10^{-14} or 1 in 4.69 x 10^{13} (1 in 46 trillion) in Nayarit. Since previously reported Mexican populations used Identifiler system, this is the first report of the allele frequencies for Penta D and Penta E in Mexicans using Power plex 16 kit. The 15 loci are highly polymorphic in these Mestizo populations and the power of exclusion and the power of discrimination values are comparable with other cosmopolitan databases [4, 8, 10].

Comparison of allele frequencies between Yucatan and Nayarit showed significant differences for the STR THO1, D5S81, D7S820, D16S539, D8S1179, TPOX and FGA (p<0.05). Comparison of allele distribution between Yucatan and other Mexican populations reveled, interestingly, that although THO1 was the STR with the least number of alleles in Yucatan, it was able to determine differentiation with other Mexican populations (p<0.03), given that THO1 showed different alleles within populations. Moreover, D7S820 displayed differentiation (p<0.003), except for Nuevo Leon and Yucatan previously reported with Identifiler kit, whereas vWA showed to be homogenously distributed between Yucatan and the other populations, except for Mayans from Campeche (p=0.004). As expected, the studied sample of Yucatan did not show significant differences with the sample of Yucatan previously reported with Identifiler genotyping [8] (p>0.05), suggesting a homogenous distribution between these two samples of Yucatan population and underlying the correspondence in allele frequencies of the two human identity testing systems. At this respect, we contributed to extend the allele frequencies for Penta E and Penta D in Mestizos from Yucatan wit this report.

Noticeable, major differences in allele distribution were found in both populations of Yucatan and Nayarit when compared with two samples of Mayan autochthonous people from Yucatan and Campeche. Mestizos and Mayans from Yucatan showed significantly different allele distribution for the STR THO1, D21S11, D18S51, D5S818, D7S820 and TPOX, suggesting that these markers might be the carriers of genetic interchange during admixture process, and the potential contributors of Mayan ancestry. In the case of Nayarit, allele frequencies of the CODIS-STR allowed a clear differentiation with both Mayan native populations (p<0.03), except for D3S1358, D21S11 and CSF1PO for Yucatan and D18S51 for Campeche. Allele frequencies of CODIS-STR were homogenously distributed between Nayarit and Nuevo Leon and between Nayarit and Jalisco (p>0.05), except for CSF1PO. Differences observed between Yucatan and Nayarit with other ethnically related Mexican populations, might reflect that Mestizo populations include the most cosmopolitan states of Mexico, such as Mexico city, Jalisco, Puebla, Northeast (Monterrey, Nuevo Leon), among them.

The principal component analysis showed that the first two components accounted for 95.09% of total sample variation (Fig. **1**), and suggesting three clusters, one formed by Mestizo populations from Pacific Coast-North Mexico including Nuevo León, Nayarit and Jalisco. A second one having closely related populations from Central-South Mexico, including populations from Yucatan, Mexico city, Valley of Mexico and Puebla, and the third one having closely related native populations, including Mayans from Yucatan and Campeche and the Amerindians. Mestizo populations were

Table 1.   Allele Frequencies for 15 Autosomal STR in 120 Mexican Mestizos from Yucatan

Alleles	D3S1358	THO1	D21S11	D18S51	PENTA E	D5S818	D13S317	D7S820	D16S539	CSF1PO	PENTA D	VWA	D8S1179	TPOX	FGA
2.2											.0042				
5					.0083						.0083				
6		.4508													
7		.2333			.0708	.1042		.0042		.0083	.0083				
8		.0625			.0166	.0125	.0625	.0500	.0042		.0146			.4625	
8.3		.0042													
9		.0875			.0042	.0500	.2583	.0541	.1333	.0458	.2042		.0042	.0500	
9.2				.0042											
9.3		.1917													
10				.0042	.0292	.0625	.1250	.3125	.2833	.2208	.2667		.0625	.0292	
10.3								.0083							
11				.0083	.0542	.4750	.1625	.2417	.2625	.3166	.1542		.0708	.2583	
12				.1167	.2417	.2125	.2042	.2542	.2458	.3125	.1333		.1250	.1917	
13	.0042			.1041	.0750	.0833	.1250	.0625	.0625	.1000	.175		.3458	.0083	
14	.0500			.1541	.1042		.0583	.0125	.0083	.0042	.0292	.0708	.2083		
15	.4833			.1667	.0958		.0042					.0708	.1500		
16	.2875			.0708	.0917							.3625	.0292		
17	.1166			.2083	.0708							.3000	.0042		
18	.0542			.0750	.0333							.1375			.0083
19	.0042			.0583	.0292							.0375			.0500
20				.0167	.0416							.0125			.0375
21				.0042	.0292							.0083			.0750
22				.0042	.0042										.1125
23				.0042											.1125
24								0.0021							.1375
25															.2542
26															.1708
27			.0208												.0292
28			.0583												.0083
29			.2833												.0042
30			.2500												
30.2			.0167												
31			.0667												
31.2			.0750												
32			.0167												
32.2			.1458												
33			.0042												
33.2			.0500												
34.2			.0125												

Table 2.   Allele Frequencies for 15 Autosomal STR in 236 Mexican Mestizos from Nayarit

Alleles	D3S1358	THO1	D21S11	D18S51	PENTA E	D5S818	D13S317	D7S820	D16S539	CSF1PO	PENTA D	VWA	D8S1179	TPOX	FGA
2.2											0.0084				
5		0.0021			0.0106						0.0021				
6		0.2606												0.0042	
7		0.3263		0.0021	0.0911	0.0403		0.0127		0.0021	0.0021	0.0021		0.0042	
8		0.0996			0.0233	0.0170	0.0678	0.1358	0.0297	0.0042	0.0170		0.0127	0.5233	
9		0.0890		0.0021	0.0064	0.0551	0.2161	0.0678	0.0975	0.0339	0.1589		0.0084	0.0614	
9.2		0.0021													
9.3		0.1759													
10		0.0191		0.0064	0.0340	0.0424	0.0932	0.2817	0.1610	0.2288	0.1992		0.0762	0.0614	
11				0.0170	0.0720	0.4428	0.2373	0.3050	0.2945	0.2754	0.1398	0.0021	0.0636	0.2140	
12				0.1631	0.1843	0.3072	0.2034	0.1716	0.2797	0.3665	0.1610		0.0741	0.1208	
13	0.0042			0.1059	0.0810	0.0890	0.1165	0.0233	0.1144	0.0720	0.2076		0.3538	0.0021	
14	0.0593			0.1590	0.0911	0.0042	0.0551	0.0041	0.0148	0.0042	0.0847	0.0636	0.2691		
15	0.4259			0.1547	0.1123		0.0021			0.0042	0.0084	0.1165	0.0911		
15.2												0.0084			
16	0.2691			0.0869	0.0975						0.0021	0.3072	0.0423		
17	0.1483			0.1377	0.0572							0.2924			
18	0.0784			0.0657	0.0211							0.1250			0.0148
19	0.0064			0.0403	0.0275							0.0614			0.0657
19.2															0.0021
20				0.0211	0.0381							0.0127			0.0890
20.2															0.0021
21				0.0064	0.0233										0.1165
21.2															0.0064
22				0.0085	0.0064			0.0021							0.1250
23				0.0085											0.1483
24					0.0064										0.1399
25				0.0021											0.1313
25.2			0.0021												
26			0.0021												0.1186
27			0.0106												0.0170
28			0.0678												0.0105
29			0.2161												0.0042
29.2			0.0021												
30			0.3008												
30.2			0.0148												

Table 2. contd…..

Alleles	D3S1358	THO1	D21S11	D18S51	PENTA E	D5S818	D13S317	D7S820	D16S539	CSF1PO	PENTA D	VWA	D8S1179	TPOX	FGA
31			0.0636												
31.2			0.1207												
32			0.0170												
32.2			0.1229												
33.2			0.0381												
34.1			0.0021												
34.2			0.0106												

Table 3.   Forensic Parameters for Two Mexican Mestizo Populations from Yucatan and Nayarit

Alleles	D3S1358	THO1	D21S11	D18S51	PENTA E	D5S818	D13S317	D7S820	D16S539	CSF1PO	PENTA D	VWA	D8S1179	TPOX	FGA
						YUCATAN									
H	0.6650	0.7204	0.8193	0.8666	0.8893	0.7051	0.8269	0.7700	0.7688	0.7414	0.8140	0.7462	0.7894	0.6793	0.8517
PM	0.158	0.119	0.059	0.038	0.028	0.131	0.058	0.102	0.097	0.140	0.068	0.102	0.075	0.162	0.044
PD	0.842	0.881	0.941	0.962	0.972	0.869	0.942	0.898	0.903	0.860	0.932	0.898	0.925	0.838	0.956
PIC	0.61	0.68	0.79	0.85	0.88	0.67	0.80	0.74	0.73	0.71	0.78	0.71	0.77	0.63	0.84
PE	0.262	0.428	0.612	0.778	0.679	0.553	0.728	0.553	0.568	0.679	0.615	0.441	0.711	0.424	0.728
TPI	1.18	1.67	2.59	4.62	3.16	2.22	3.75	2.22	2.31	3.16	2.61	1.71	3.53	1.65	3.75
HWE (p)	0.021*	0.886	0.149	0.321	0.215	0.465	0.931	0.130	0.699	0.011*	0.412	0.323	0.602	0.465	0.764
						NAYARIT									
H	0.8082	0.82477	0.8727	0.8635	0.8625	0.7242	0.8118	0.7934	0.7724	0.7659	0.8732	0.8162	0.8233	0.7404	0.9057
PM	0.137	0.091	0.054	0.032	0.020	0.147	0.058	0.090	0.082	0.124	0.049	0.086	0.094	0.171	0.031
PD	0.863	0.909	0.946	0.968	0.980	0.853	0.942	0.910	0.918	0.876	0.951	0.914	0.906	0.829	0.969
PIC	0.67	0.73	0.80	0.86	0.90	0.64	0.80	0.74	0.75	0.68	0.81	0.75	0.74	0.61	0.87
PE	0.513	0.477	0.597	0.964	0.755	0.470	0.613	0.536	0.597	0.443	0.613	0.582	0.566	0.349	0.860
TPI	2.02	1.85	2.49	3.33	4.18	1.83	2.60	2.13	2.49	1.72	2.60	2.39	2.29	1.41	7.31
HWE (p)	0.2480	0.5840	0.7500	0.0175*	0.9290	0.3710	0.0825	0.6070	0.1450	0.8675	0.4680	0.8510	0.5730	0.0600	0.0015*

Heterozygousity index (H), Matching Probability (MP), Power of Discrimination (PD), Polymorphic Information Content (PIC), Power of Exclusion (PE) and Typical Paternity Index (TPI), Bonferroni correction (p=0.003)*

intermediate to the ancestral Spanish and Amerindian populations. Our findings are consistent with differences in admixture components based on the trihybrid model (Amerindian, Caucasian and/or African). For Mexico, differences in genetic admixture have been described along the country, where northern states would have a higher Caucasian component than southeastern populations, characterized by a higher Amerindian component [12]. This gradient has also been demonstrated with ancestral-component estimation in Mexican Mestizos based on 13 CODIS STRs, suggesting

that admixture proportions throughout Mexico, displayed a clear North-to-South increase of the Amerindian component (38.3–75.7%), with a proportional decrease of the European component (50.3–8.5%), and vice versa [8, 13]. Conversely, African ancestry remained relatively low and constant across populations (9.4–18.5%) [8, 13]. The Mexican Genome Diversity Project based on a HapMap-like database with SNP frequencies in Mexicans confirmed the presence of considerable variability between Mexican Mestizo subpopulations and suggested that this diversity is mainly related to a differ-
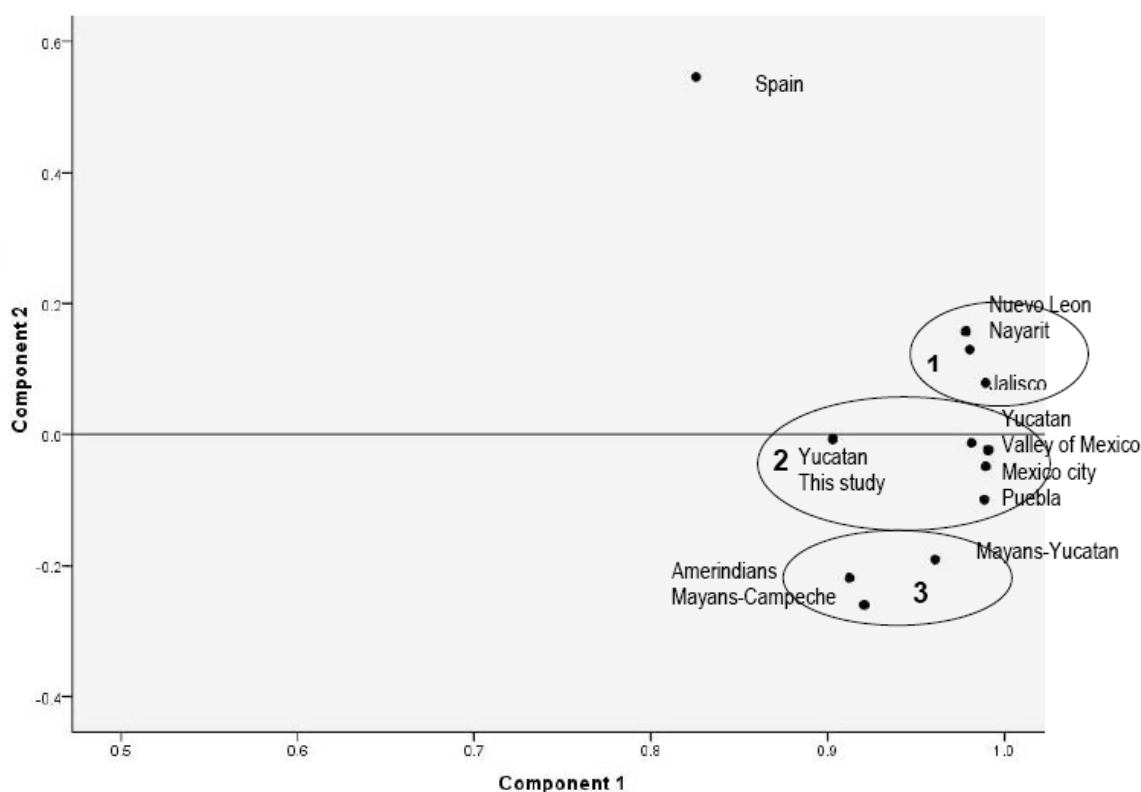
Fig. (1). Principal component analysis: Cluster 1: Mexican mestizo populations from Pacific Coast-North Mexico. Nayarit (this study), Jalisco [8] and Nuevo Leon [7]. Cluster 2, Mexican mestizo populations from Central-South Mexico: Yucatan (this study), Yucatan tested with identifiler [8], Puebla [8], Mexico city [5], Valley of Mexico [6]; Cluster 3, Amerindian populations: Mayan native populations from Yucatan and Campeche [11] and native american amerindians [10].

ential distribution of Amerindian and European ancestral components, particularly for Mestizos from Yucatan, who have a distinctive Mayan Ameridian contribution [14]. The regional differences in contributions from different gene pools in Mexico remarks the need to establish corresponding local databases to improve the impact of STR loci on forensic computations in such admixed populations.

Our results suggested that STR provide a powerful discrimination and support the importance of the generation of local databases in Mexico for reliable DNA profile interpretation when STR are being currently used in forensic casework, mainly in Nayarit and Southern Mexico, where there are not databases. Availability of genetic data on these loci should increase our understanding of population structure in the various Mexican populations and improve the statistical interpretation of DNA forensic and parentage testing results.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Tereba, A. *Tools for analysis of population statistics*. Profiles in DNA. Promega Corporation; **1999**.

[2]   Excoffier, L.; Laval, G.; Schneider, S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinform. Online*, **2005**, *1*, 47-50.

[3]   Raymond, M.L.; and Rousset, F. An exact test for population differentiation. *Evolution*., **1995**, *49*, 1280-1283.

[4]   Hernandez-Gutierrez, S.; Hernandez-Franco, P.; Martinez-Tripp, S.; Ramos-Kuri, M.; Rangel-Villalobos, H. STR data for 15 loci in a population sample from the central region of Mexico. *Forensic Sci. Int. Genet*., **2005**, *151*, 97-100.

[5]   Juarez-Cedillo, T.; Zuñiga, J.; Acuña-Alonzo, V.; Perez-Hernandez, N.; Rodríguez-Perez, J.M.; Barquera, R.; Gallardo, G.J.; Sanchez-Arenas, R.; García-Peña, M.C.; Granados, J.; Vargas-Alarcón, G. Genetic admixture and diversity estimations in the Mexican Mestizo population from Mexico City using 15 STR polymorphic markers. *Forensic Sci. Int. Genet*., **2008**, *2*, e37-e39.

[6]   Luna-Vazquez, A.; Vilchis-Dorantes, G.; Aguilar-Ruiz, M.; Bautista-Rivas, A.; Rojo-Nava, A.; Rios-Barrios, E.; Rangel-Villalobos, H. Population data for 15 loci (Identifiler Kit) in a sample from the Valley of Mexico. *Legal Med*., **2005**, *7,* 331-333.

[7]   Cerda-Flores, R.M.; Budowle, B.; Jin, J.; Barton, S.A.; Deka, R.; Chakraborty, R. Maximum likelihood estimates of admixture in Northeastern Mexico using 13 short tandem repeat loci. *Am. J. Hum. Biol*., **2002**, *14*, 429-439.

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

[9]   Gamero, J.J.; Romero, J.L.; Gonzalez, J.L.; Arufe, M.I.; Cuesta, M.I.; Corte-Real, F.; Carvalho, M.; Anjor, M.J.; Vieira, D.N.; Vide, M.C. A study of ten short tandem repeat system: African immigrant and Spanish population data. *Forensic Sci. Int*., **2000**, *110*, 167-177.

[10] Budowle, B.; Shea, B.; Niezgoda, S.; Chakraborty, R. CODIS STR loci data from 41 sample populations. *J. Forensic Sci*., **2001**, *46*, 453-489.

[11] Ibarra-Rivera, L.; Mirabal, S.; Regueiro, M.; Herrera, R. Delineating Genetic Relationships among the Maya. *Am. J. Phys. Anthropol.*, **2008**, *135*, 329-347.

[12]   Gorodesky, C.; Alaez, C.; Vazquez-García, M.N.; De la Rosa, G.; Infante, E.; Balladares, S.; Toribio, R.; Pérez-Luque, E.; Muñoz, L. The genetic structure of Mexican Mestizos of different locations: Tracking back their origins through MHC genes, blood group systems, and microsatellites. *Hum. Immunol.*, **2001**, *62*, 979-991.

[13]   Rubi-Castellanos, R.; Martínez-Cortes, G.; Muñoz-Valle, J.; González-Martín, A.; Cerda-Flores, R. M.; Anaya-Palafox, M.; Rangel-Villalobos, H. Pre-Hispanic Mesoamerican demography approximates the present-day ancestry of Mestizos throughout

the territory of Mexico. *Am. J. Phys. Anthropol.*, **2009**, *139*, 284-294.

[14]   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.; Goya, R.; Hernandez-Lemus, E.; Davila, C.; Barrientos, E.; March, S.; Jimenez-Sanchez, G. Analysis of genomic diversity in Mexican Mestizo populations to develop genomic medicine in Mexico. *Proc. Natl. Acad. Sci. USA.*, **2009**, *106*, 8611-8616.