



PERINATOLOGÍA Y REPRODUCCIÓN HUMANA

www.elsevier.es/rprh



ORIGINAL

Análisis de las variaciones en el número de repeticiones de 5 marcadores ancestrales en donadores recurrentes en México



M.I. Zauza-Carrasco^a, F. Rosenfeld-Mann^b y H. Estrada-Juárez^{b,*}

^a Licenciatura en Bioquímica Diagnóstica, Facultad de Estudios Superiores Cuautitlán, Universidad Nacional Autónoma de México (UNAM), Ciudad de México, México

^b Coordinación de Hematología Perinatal, Instituto Nacional de Perinatología Isidro Espinosa de los Reyes, Ciudad de México, México

Recibido el 9 de octubre de 2015; aceptado el 21 de diciembre de 2015

Disponible en Internet el 22 de marzo de 2016

PALABRAS CLAVE

Repeticiones cortas en tándem;
Sistemas de Índice Combinado de ADN;
Población mexicana;
Ciencias forenses;
Grupo de ascendencia

Resumen Los microsatélites son secuencias altamente polimórficas de ADN constituidas por repeticiones cortas regularmente de 1-7 pares de bases y con más del 70% de heterocigocidad. El objetivo del artículo fue obtener las frecuencias alélicas de 5 loci por electroforesis capilar y compararlas con la base de datos *Combined DNA Index System* (CODIS) para determinar la eficacia de la técnica en la determinación ancestral.

Material y métodos: Se extrajo el ADN de 50 individuos no relacionados (G1). Se amplificaron 6 loci mediante PCR. Se identificó la longitud de los marcadores por electroforesis capilar y las frecuencias se compararon con las reportadas en CODIS (G2) para población latina.

Resultados: Las frecuencias de los polimorfismos TH01 y D2S1338 fueron las únicas similares a lo reportado en CODIS. El número de alelos del FGA fue distinto en ambos grupos, G1 no presentó 4 alelos reportados. El polimorfismo D18S51 en G1, presentó un alelo de 10.2 repeticiones no reportado en CODIS, y careció de 4 alelos. Para D16S539, los valores de nuestra población fueron distintos a los reportados, los alelos 7 y 17 presentes en nuestra población no están reportados en CODIS. Los alelos 14 y 15 reportados en CODIS no se observaron en nuestra población.

Conclusiones: En la comparación de las frecuencias de los polimorfismos en una muestra de origen mestizo mexicana con otra población latina, las diferencias encontradas nos muestran la necesidad de crear bases de datos de referencia de nuestra población, teniendo en cuenta el mestizaje que nos caracteriza.

© 2016 Instituto Nacional de Perinatología Isidro Espinosa de los Reyes. Publicado por Masson Doyma México S.A. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Autor para correspondencia.

Correo electrónico: higinio.estrada@inper.mx (H. Estrada-Juárez).

KEYWORDS

Short tandem repeat;
Combined DNA Index
System;
Mexican Americans;
Forensic sciences;
Ancestry

Analysis of variations in the number of repeats of 5 ancestry markers in recurrent donors in Mexico

Abstract Microsatellites are highly polymorphic DNA sequences, consisting of short repeats, regularly from 1 to 7 base pairs, and over 70% of heterozygosity. The objective of this study was to obtain allele frequencies of 5 loci by capillary electrophoresis and comparing them with the Combined DNA Index System database (CODIS) to determine the effectiveness of the technique in the ancestral determination.

Material and methods: DNA from 50 unrelated individuals (G1) was extracted. Six loci were amplified by PCR. The lengths of the markers were identified by capillary electrophoresis and the frequencies were compared with those reported in CODIS (G2) for Latin-Americans.

Results: The frequencies of the polymorphisms, TH01 and D2S1338, were the only ones similar to those reported in CODIS. The number of alleles of FGA was different in both groups, and G1 did not have 4 of the reported alleles. The D18S51 polymorphism in G1 presented with a 10.2 allele unreported in CODIS, and lacked 4 alleles. For D16S539, our population values were different from those reported, 7 and 17 alleles were present in our population, but were not reported in CODIS. Moreover, alleles 14 and 15 reported in CODIS were not observed in our population.

Conclusions: The results of our study enable us to compare the frequencies of the polymorphisms in a sample of Mexican with other Latin Americans. The differences suggest the need to create databases of reference for our population, taking into account the race mixture that characterises us.

© 2016 Instituto Nacional de Perinatología Isidro Espinosa de los Reyes. Published by Masson Doyma México S.A. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

En su libro *El origen de las especies*, Charles Darwin escribió: «Probablemente todos los seres orgánicos que hayan vivido nunca en esta tierra han descendido de alguna forma primordial». ¹ Es así que nuestro ADN puede ser prácticamente idéntico al de otras personas, sin embargo, hay regiones heredadas que pueden variar de una persona a otra; estas secuencias se denominan «polimorfismo» ².

Las regiones polimórficas del ADN tienen un alto poder de discriminación, por ejemplo permiten con gran fiabilidad excluir a un individuo que puede ser falsamente acusado en un crimen ³. El tamaño del genoma humano es de aproximadamente 3.2 Gb, de este, el 0.1% es diferente entre las personas, lo que permite la variabilidad genética entre grupos étnicos ⁴.

Los microsatélites (o STR por sus siglas en inglés: *short tandem repeat*) son repeticiones cortas en tándem y son un tipo de polimorfismos con secuencias repetidas de 2-7 pares de bases ⁵, que varían en longitud entre los individuos y generalmente existen muchos alelos diferentes dentro de una población ⁶.

Estos loci son propensos a mutaciones debido a que son susceptibles de deslizarse durante la replicación del ADN ⁷. Esta tasa alta de mutación ha sido un factor clave en la determinación de la capacidad informativa de los STR en los estudios genéticos. Han sido utilizados ampliamente durante los últimos 10 años para investigar la variación genética a través de poblaciones en muchas especies ⁸.

La diversidad genética (entre individuos de un mismo grupo o entre grupos diferentes) puede ser evaluada a partir

de la distribución de las frecuencias alélicas de los marcadores genéticos. Las variaciones en estas secuencias permiten estimar el grado de divergencia entre las poblaciones, cuya variabilidad puede deberse a diversos factores ^{9,10}.

En medicina forense los STR desempeñan un papel importante. Las bases de datos de ADN tanto de EE. UU. (*Combined DNA Index System*) como de países europeos (*European Standard Set*, *UK Core Loci*, *Interpol Standard Set of Loci*) se basan únicamente en estos loci para crear huellas genéticas ¹¹. La naciente comunidad de la genealogía genética utiliza estos loci para desarrollar bases de datos que contienen linajes de cientos de miles de individuos ¹².

El marcador TH01 es una repetición simple tetranucleótida (TCAT) con alelos que van desde 3 a 14 repeticiones, localizada en el intrón 1 (es decir, 01) del gen de la tirosina hidroxilasa en 11p15.5 ¹³.

FGA es una repetición compuesta tetranucleótida (CTTT)_n flanqueada a ambos lados por repeticiones degeneradas (TTCC)_n que se encuentra en el tercer intrón del locus de la cadena alfa del fibrinógeno humano en 4q31.3. Los alelos reportados varían en tamaño desde 12.2 a 51.2, que abarca más de 35 repeticiones. Una delección 2 bp, desde la pérdida de un CT, en la región justo antes del motivo de repetición central es responsable de las microvariantes alélicas X.2 que son muy frecuentes en este sistema STR ¹³.

D18S51 es una repetición sencilla AGAA en el brazo largo del cromosoma 18. Los alelos varían de tamaño de 5.3 a 40 repeticiones. Más de 70 alelos han sido reportados, siendo uno de los más polimórficos de uso común ¹³.

D16S539 es una repetición simple de 4 nucleótidos (GATA) que van de 4 a 17 unidades localizado en 16q24.1 ¹³. D2S1338

Tabla 1 Descripción de las secuencias de los iniciadores para los STR autosómicos utilizados. Se indica la unidad de repetición y el número de alelos existentes

Locus	Secuencia del primer (5'3')	Unidad de repetición	Número de alelos existentes
TH01	F 6FAM-CCTGTTCTCCCTTATTTCCTC R GTTCTTTGGGAACACAGACTCCATGGTG	TCAT	20
FGA	F 6FAM-AAATAAAATTAGGCATATTTACAAGC R GCTGAGTGATTTGTCTGTAATTG	CTTT – TTCC	31
D18S51	FVIC-TGAGTGACAAATTGAGACCTT R GTCTTACAATAACAGTTGCTACTATT	AGAA	27
D16S539	F NED-ATACAGACAGACAGACAGGTG R GCATGTATCTATCATCCATCTCT	GATA	15
D2S1338	F PET-TGGAAACAGAAATGGCTTGG R GATTGCAGGAGGGAAGGAAG	TGCC-TTCC	23

Fuente: Buttler y Hill¹³.

es una repetición de tetranucleótidos compuesta que se encuentra en el brazo largo del cromosoma 2. Los alelos van desde 10 a 31 repeticiones¹³.

Metodología

Se colectaron muestras de sangre de 50 individuos sanos, donadores recurrentes, no relacionados entre ellos, nacidos en México y de padres mexicanos (G1)

El ADN se extrajo mediante el ensayo comercial *High Pure PCR Template Preparation Kit V. 20* (Roche Mannheim Alemania, n.º de catálogo: 11796828001).

Las secuencias se amplificaron por PCR punto final. El iniciador sentido se marcó en el extremo 5' con 6-FAM para tirosin hidroxilasa (TH01), cadena alfa de fibrinógeno (FGA) y amelogénina, el D18S51 con PET, el D16S539 con NED y el D2S1338 con VIC, para su posterior identificación por electroforesis capilar (tabla 1).

Las amplificaciones se realizaron individualmente para cada par de iniciadores, en un volumen de reacción de 10 µL usando una mezcla maestra que contenía 7.5 U de *Hot Start Master Mix* (Qiagen, Cat. 203445), 0.5 µM de cada par de iniciadores, 50 ng de ADN y se ajustó el volumen final con agua grado biología molecular. Para amplificación se usaron las siguientes condiciones: 95 °C por 10 min, 32 ciclos: 94 °C por 1 min, 55 °C por 1 min, 72 °C por 1 min, 60 °C por 45 min.

La electroforesis capilar se llevó a cabo en un secuenciador Abi 3500, con un capilar de 50 cm, usando el polímero POP 7 y determinando el peso molecular mediante GS 600 LIZ.

La longitud de los marcadores se comparó con las frecuencias reportadas en *Combined DNA Index System* (CODIS) para población latina (G2) (tabla 2).

Resultados

El polimorfismo de la TH01 presentó el mismo número de repeticiones en ambos grupos, sin embargo, las frecuencias variaron en cada una de ellas. El alelo con 6 repeticiones fue el más frecuente en G1 (0.306), mientras que el de 7 lo fue en G2 (0.296). En ambos casos el alelo con 10 repeticiones fue el menos frecuente (0.031 y 0.0085 respectivamente).

El número de alelos del FGA fue distinto en ambas poblaciones, el G1 no presentó 4 alelos reportados para población latina (17, 22.2, 23.2 y 29), en ninguno de los 2 grupos se observó el alelo de 11 repeticiones que sí ha sido reportado en otras poblaciones, por ejemplo africanos, asiáticos o caucásicos. El alelo de 23 repeticiones fue el más frecuente en G1 (0.194), mientras que los alelos 18 y 28 fueron los menos prevalentes (0.01 en ambos casos). Comparando los datos de G2 el alelo de 22 es el más frecuente (0.1653) y los alelos de 17, 28 y 29 repeticiones son los menos frecuentes (0.0021 en los 3 casos).

El polimorfismo D18S51 en G1 presentó un alelo de 10.2 repeticiones no reportado en CODIS con una frecuencia de 0.1224, asimismo, careció de 4 alelos reportados para población latina (14.2, 21, 22, 24). Las frecuencias observadas en G1 y G2 para cada alelo fueron distintas. El alelo más frecuente en ambos grupos fue el de 13 repeticiones (0.1939 y 0.1229 respectivamente), mientras que el menos frecuente fue el de 20 y 17 para G1 (0.02) y para G2 fueron 10, 14.2 y 24 (0.0021 en todos los casos).

Para el marcador D16S539, los valores de nuestra población fueron muy distintos a los reportados en CODIS. Los alelos de 7 y 17 repeticiones presentes en nuestra población no están reportados en CODIS para latinos. Por otra parte, los alelos 14 y 15 reportados en CODIS no se observaron en nuestra población. Los alelos más frecuentes fueron el de 11 (0.316) en G1 y de 12 en G2 (0.278). Los menos frecuentes fueron los de 13 y 17 (0.010) en G1 y 15 (0.0021) en G2.

En el caso del marcador D2S1338 se observó el mismo número de alelos en las mismas frecuencias reportadas por CODIS para la población latina. El alelo 19 (0.214 y 0.193) fue el más frecuente en ambos casos. El alelo menos frecuente en G2 fue el de 26 (0.0169), para G1 además del 26 también fue el de 21 (0.0102).

Discusión

Cuando un individuo es el resultado de una mezcla de ascendencias representa un mosaico genético.

Dentro del proyecto de los 1,000 genomas, la tipificación mexicana (MXL) corresponde a una muestra de Los Ángeles, California. Sin embargo la población hispana incluye

Tabla 2 Frecuencias de los marcadores analizados en donadores recurrentes comparadas con los datos reportados en la base de datos CODIS

	TH01		FGA		D18S51		D16S539		D2S1338	
	G1	G2	G1	G2	G1	G2	G1	G2	G1	G2
6	0.306	0.239								
7	0.296	0.297					0.020			
8	0.092	0.091					0.092	0.019		
9	0.133	0.146					0.153	0.140		
9.3	0.143	0.218								
10					0.020	0.002	0.265	0.150		
10.2					0.122					
11					0.122	0.015	0.316	0.265		
12					0.163	0.114	0.133	0.278		
13					0.194	0.123	0.010	0.133		
14					0.071	0.161		0.013		
14.2						0.002				
15					0.153	0.159		0.002		
16					0.051	0.125			0.051	0.030
17				0.002	0.020	0.125	0.010		0.143	0.169
18			0.010	0.013	0.031	0.078			0.082	0.081
19			0.061	0.081	0.031	0.047			0.214	0.193
20			0.061	0.085	0.020	0.028			0.143	0.127
21			0.133	0.153		0.008			0.010	0.032
22			0.163	0.165		0.011			0.051	0.057
22.2				0.004						
23			0.194	0.121					0.194	0.140
23.2				0.004						
24			0.112	0.142		0.002			0.071	0.076
25			0.112	0.119					0.031	0.078
25.2										
26			0.102	0.061					0.010	0.017
27			0.041	0.044						
28			0.010	0.002						
29				0.002						
H _{obs}	0.837	0.907	0.685	0.881	0.837	0.907	0.779	0.788	0.878	0.852

los diversos grupos étnicos de América Latina; aunque las contribuciones genéticas significativas se pueden remontan hasta los mismos grupos indígenas, incluye también ascendencia europea y de poblaciones de África occidental. Dentro de esta amalgama genética cada población latina estará determinada por su propio origen étnico. Esta composición génica implica una comprensión más profunda de la evolución humana, ya que la diversidad genética entre las poblaciones latinas es particular en cada caso.

La población mexicana se conoce por ser una mezcla de múltiples características que desciende principalmente de linajes indígenas mesoamericanos y europeos, principalmente españoles, pero también tiene un origen africano que se deriva en gran parte de la trata de esclavos transatlántica, principalmente de África occidental, desde Senegal a Nigeria y Angola; por ejemplo, los mexicanos en promedio tienen menor ascendencia africana que los brasileños¹⁴.

El uso de marcadores de longitud variable es una estrategia viable, capaz de diferenciar el ADN de distintos individuos; y dado que la probabilidad de su coincidencia al azar es superior a $1/1 \times 10^{15}$, lo convierte en una metodología aplicable en otras investigaciones, por ejemplo para

la identificación de ADN fetal a partir de sangre periférica materna.

Nuestros resultados son diferentes a las frecuencias reportadas en la base CODIS para latinos; creemos que nuestra población de estudio es muy homogénea al tratarse únicamente de mexicanos con padres mexicanos, sin embargo, esto solo refleja las características de una muestra de habitantes de la Ciudad de México; aunque no intencional, sí presenta un sesgo, lo que a su vez nos obliga a tratar de obtener una muestra representativa de la población mexicana.

A diferencia de otras regiones del mundo en donde el colonialismo europeo dio lugar a una disminución significativa de la población nativa, en México los indígenas no solo contribuyeron en el proceso de mestizaje, sino que actualmente constituyen un segmento importante de la población¹⁵. Estudios recientes han mostrado que el promedio de ancestría indígena es del 53%, mientras que el 42 y 5% son de origen europeo y africano, respectivamente¹⁶.

En 2014 Moreno-Estrada et al.¹⁷ presentaron los resultados de la genotipificación de 1,000 personas, representada por 20 grupos indígenas y 11 poblaciones mestizas, sin

embargo, la creación de una base de datos mexicana debería incluir, además de los grupos endémicos, a aquellos que comparten el territorio con zonas urbanas, ya que su aporte genético da lugar a las nuevas poblaciones mestizas metropolitanas, de esta forma podríamos evaluar con mucha mayor resolución el impacto de los componentes indígenas en la población.

Conclusión

En este trabajo presentamos los resultados de un estudio preliminar que nos permitió comparar las frecuencias de los polimorfismos de amplitud variable en una muestra de origen mestiza mexicana con otra población latina. Las diferencias encontradas aquí nos muestran la necesidad de crear bases de datos de referencia de nuestra población, teniendo en cuenta el mestizaje que nos caracteriza.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Bibliografía

1. Passarge E. Atlas de genética. Madrid: Editorial Panamericana; 2010.
2. Checa-Caratechea MA. Polimorfismos genéticos: Importancia y aplicaciones. *Rev Inst Nal Enf Resp Mex*. 2007;20:213–21.
3. Oria-Hernández J, Reyes-Vivas H, Romero-Álvarez I, Velázquez-López I. Genética forense: la ciencia al servicio de la justicia. *Mensaje Bioquímico*. 2007;31:45–67.
4. Cobián A, Eguiarte LE. Estructura y complejidad del genoma humano. *Ciencias*. 2002;68:56–64.
5. Gill P. Role of short tandem repeat DNA in forensic casework in the UK-past, present, and future perspectives. *BioTechniques*. 2002;32:366–8.
6. Chen CM, Sio CP, Lu YL, Chang HT, Hu CH, Pai TW. Identification of conserved and polymorphic STRs for personal genomes. *BMC genomics*. 2014;15 Suppl 10:S3–6.
7. Willems T, Gymrek M, Highnam G, Mittelman D, Erlich Y, Genomes Project C. The landscape of human STR variation. *Genome Res*. 2014;24:1894–904.
8. Pemberton TJ, Sandefur CI, Jakobsson M, Rosenberg NA. Sequence determinants of human microsatellite variability. *BMC genomics*. 2009;10:612–5.
9. Tracey M. Short tandem repeat-based identification of individuals and parents. *Croatian Med J*. 2001;42:233–8.
10. Gusmao L, Alves C, Amorim A. Molecular characteristics of four human Y-specific microsatellites (DYS434, DYD437, DYS438, DYS439) for population and forensic studies. *Ann Hum Genet*. 2001;65 Pt 3:285–91.
11. Kayser M, de Knijff P. Improving human forensics through advances in genetics, genomics and molecular biology. *Nature Rev Genet*. 2011;12:179–92.
12. Khan R, Mittelman D. Rumors of the death of consumer genomics are greatly exaggerated. *Genome Biol*. 2013;14:139–42.
13. Butler JM, Hill CR. Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. *Forensic Sci Rev*. 2012;24:15–26.
14. Lopez-Beltran C, Deister VG. [Scientific approaches to the Mexican mestizo]. *Historia, Ciencias, Saude-Manguinhos*. 2013;20:391–410.
15. Moreno A, Sandoval K. Diversidad genómica en México: Pasado indígena y mestizaje. *Cuicuilco*. 2013;20:249–75.
16. Silva-Zolezzi I, Hidalgo-Miranda A, Estrada-Gil J, Fernandez-Lopez JC, Uribe-Figueroa L, Contreras A, et al. Analysis of genomic diversity in Mexican Mestizo populations to develop genomic medicine in Mexico. *Proc Natl Acad Sci USA*. 2009;106:8611–6.
17. Moreno-Estrada A, Gignoux CR, Fernandez-Lopez JC, Zakharia F, Sikora M, Contreras AV, et al. Human genetics. The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science*. 2014;344:1280–5.