



Revista Colombiana de Antropología

ISSN: 0486-6525

rca.icanh@gmail.com

Instituto Colombiano de Antropología e
Historia
Colombia

Duque, Constanza; Parra, MaríaVictoria; Valencia, Ana Victoria; Bedoya, Gabriel; Ruiz-Linares, Andrés

Comparación de cuatro programas utilizados en la determinación de la composición genética ancestral de la población antioqueña

Revista Colombiana de Antropología, vol. 48, núm. 1, enero-junio, 2012, pp. 233-257

Instituto Colombiano de Antropología e Historia
Bogotá, Colombia

Disponible en: <http://www.redalyc.org/articulo.oa?id=105025257006>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica
Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

COMPARACIÓN DE CUATRO PROGRAMAS

utilizados en la determinación de la composición genética ancestral de la población antioqueña

CONSTANZA DUQUE^A (mitoconnie@yahoo.com),

MARÍA VICTORIA PARRA^A, ANA VICTORIA VALENCIA^A, GABRIEL BEDOYA^A
Y ANDRÉS RUIZ-LINARES^B

Resumen

La población antioqueña posee una estructura genética compleja que refleja su historia de mezclas entre poblaciones europeas, africanas y amerindias. En este estudio se determinó la composición genética de 849 individuos de la población antioqueña, utilizando un panel de 75 marcadores informativos de ancestralidad y 4 programas que implementan métodos bayesianos para inferir la composición de la mezcla ancestral. La contribución europea (60%) fue dos veces mayor que la contribución amerindia (28%), ambas con una distribución amplia, y fue poca la contribución africana (12%). Se estimó el tiempo transcurrido desde el primer evento de mezcla, que fue de 11 generaciones por cada 100 cM, lo cual hace de esta población un recurso útil para implementar estudios de mapeo por mezcla. Se evaluó la utilidad del panel de marcadores, así como la robustez y utilidad de los diferentes programas utilizados.

PALABRAS CLAVE: mezcla genética, población antioqueña, marcadores informativos de ancestralidad.

COMPARISON OF FOUR SOFTWARES USED TO DETERMINING THE ANCESTRAL GENETIC COMPOSITION OF THE ANTIOQUIAN POPULATION

Abstract

The genetic structure of the population of Antioquia is complex and arises from the historical admixture of European, African and Amerindian populations. The genetic makeup of 849 Antioquian individuals was determined using a set of 75 ancestry informative markers and four computer programs based on Bayesian methods in order to infer the composition of ancestral admixture. A great European component (60%) was twice as high as the Amerindian contribution (28%), both with broad distributions, followed by a smaller African contribution (12%). The average number of generations since the admixture event was 11 generations per 100 cM, which makes this population a useful resource to implement admixture mapping studies.

The usefulness of the marker set as well as the power and helpfulness of the computer programs were assessed.

KEYWORDS: genetic admixture, Antioquian population, ancestry informative markers.

A Grupo de Genética Molecular, Instituto de Biología, Universidad de Antioquia, Medellín, Colombia.

B Department of Genetics, Evolution and Environment, University College London. Londres, Reino Unido.

INTRODUCCIÓN¹

La mezcla genética representa un experimento natural en antropología y epidemiología genética, equivalente a un cruce entre cepas endogámicas, para identificar genes de susceptibilidad a enfermedades complejas que difieren en el riesgo genético entre los grupos parentales. Además, los marcadores polimórficos y el desequilibrio de ligamiento (DL) pueden ser usados para inferir las bases genéticas de los rasgos de interés (Halder y Shriver 2003; Smith y O'Brien 2005).

1. El presente artículo es derivado de un proyecto de investigación. Constanza Duque, María Victoria Parra y Ana Victoria Valencia participaron igualmente en este trabajo.

Las poblaciones con una historia de mezcla reciente (veinte generaciones o menos), como la antioqueña, constituyen una muestra ideal para implementar estudios de mapeo por desequilibrio de ligamiento generado por la mezcla MALD (por su sigla en inglés, *mapping by admixture linkage disequilibrium*) (Smith y O'Brien 2005). La utilidad y precisión de este método se basan en un conocimiento exhaustivo del proceso de migración, del tamaño efectivo de la población fundadora, de la tasa de migración, del tiempo estimado desde el primer evento de mezcla, de la disponibilidad de marcadores informativos y de métodos estadísticos que combinan la información de estos marcadores para inferir la composición genética de las poblaciones (Bryc et ál. 2010; McKeigue 2005; Reich y Patterson 2005; Seldin 2007).

Las evidencias históricas y genéticas sugieren que la población antioqueña se estableció entre los siglos XVI y XVII, como consecuencia de un fuerte mestizaje entre hombres provenientes de España y mujeres indígenas, posiblemente de la etnia embera. Se llegó a esta conclusión cuando se encontró que el 90% del ADN mitocondrial o de linajes maternos es de origen amerindio, el 8% africano y el 2% europeo, y en cambio el 94% de la región no recombinante del cromosoma Y o de linajes paternos es de origen europeo, el 5% africano y 1% amerindio (Carvajal-Carmona et ál. 2000). Además, la administración y la sociedad coloniales impusieron presiones que favorecieron el establecimiento de familias entre mestizos, más que con mujeres amerindias, y así se redujo la posterior participación de las mujeres amerindias en la fundación y expansión de la población antioqueña (Sánchez 1977). Posteriormente se encontró, por medio del uso de marcadores

nucleares con alelos específicos en las poblaciones ancestrales, que la composición de esta población es en un $79\% \pm 6$ europea, $8\% \pm 3$ africana y $16\% \pm 6$ amerindia (Bedoya, Montoya et ál. 2006). Recientemente se evaluaron 13 poblaciones latinoamericanas, entre ellas 4 colombianas (Peque, Medellín, Cundinamarca y Pasto), y se encontró que el patrón tripoblacional se repite en la mayoría de las poblaciones estudiadas, con igual desviación en cuanto a la contribución de los sexos, pues en todos los casos la mezcla se formó con hombres europeos y mujeres amerindias o africanas, principalmente (Wang et ál. 2008).

En la población antioqueña han transcurrido aproximadamente veinte generaciones desde que se originó la mezcla; este hecho, sumado al aislamiento impuesto por la geografía del departamento de Antioquia, indujo durante dos siglos a la práctica del matrimonio endogámico (Bedoya, García et ál. 2006). Como consecuencia, en la población antioqueña persiste un DL de fondo significativo, que la hace comparable con otras poblaciones del mundo consideradas genéticamente homogéneas y aisladas (Service et ál. 2006). Este DL facilita la detección de locus involucrados en la susceptibilidad o la evolución de diferentes entidades.

Actualmente se cuenta con bases de datos de marcadores con gran capacidad para diferenciar los estados ancestrales de las diferentes poblaciones continentales y aun de subpoblaciones regionales, con respecto tanto a locus como a regiones cromosómicas. Estos marcadores se han denominado *marcadores informativos de ancestralidad* (AIM) (por su sigla en inglés, *ancestry informative markers*). El poder informativo de estos marcadores se basa en la diferencia en las frecuencias alélicas (δ) entre las poblaciones parentales putativas de una población mezclada. El valor absoluto de la diferencia en frecuencias alélicas es la medida más comúnmente usada de informatividad entre dos poblaciones parentales (International HapMap 2003; Price et ál. 2007; Shriver et ál. 2003; Smith et ál. 2004).

Para los estudios de mezcla, y especialmente para MALD, se han desarrollado diferentes algoritmos y programas que permiten calcular no solo la composición de mezcla promedio de una población, sino determinar el estado ancestral de cada uno de los cromosomas de un individuo en dicha población. En general, estos programas implementan combinaciones de métodos clásicos y bayesianos, usando datos genotípicos e información a

priori sobre las frecuencias alélicas en las poblaciones parentales y asumiendo el equilibrio de Hardy-Weinberg (EHW). Algunos de los más utilizados son Admixmap, Structure y Ancestrymap (Hoggart et ál. 2003; Hoggart et ál. 2004; Long 1991; Patterson et ál. 2004; Pritchard, Stephens y Donnelly 2000; Seldin, Pasaniuc y Price 2011).

Admixmap es un programa multipropósito que modela la mezcla poblacional con datos de genotipos multilocus y datos fenotípicos. Está basado en un acercamiento que combina métodos

2 Nota de los editores: El término es una adaptación del inglés *frequentist*. Aunque algunos autores prefieren usar la expresión *métodos frecuentistas*, aquí se ha preferido el término más extendido en español, *métodos frecuentistas*.

bayesianos y frecuentistas² que estiman la mezcla de cada individuo y la distribución de la mezcla individual en la población. Para ello toma como distribución a priori las variables no observadas,

como la mezcla de cada uno de los gametos en la población y el estado ancestral de los alelos de cada locus en cada cromosoma, para cada uno de los individuos tipificados, y genera una distribución posterior de dichas variables dados los genotipos multilocus observados, por medio de una simulación de MCMC (por su sigla en inglés, Markov-Chain-Monte-Carlo) (Hoggart et ál. 2003; Hoggart et ál. 2004).

El programa Structure, introducido por Pritchard et ál. en el 2000 (Falush, Stephens y Pritchard 2003; Montana y Pritchard 2004; Pritchard, Stephens y Donnelly 2000), es un método basado en la agrupación con un acercamiento bayesiano, que genera la distribución posterior de los estados ancestrales de cada locus en un cromosoma (probabilidades de transición) para cada genotipo observado, por medio de un modelo de MCMC basado en el muestreo de Gibbs. El modelo está condicionado a la probabilidad a priori de que los alelos de un marcador (AIM) se hayan derivado de una de las poblaciones ancestrales; para calcular esta probabilidad se asume que los genotipos en la población de estudio están en EHW. De esta manera, con las frecuencias alélicas de los marcadores (AIM) en las poblaciones ancestrales se calculan las probabilidades de asignar un genotipo observado para cada marcador a cada una de las poblaciones ancestrales, cuando se tienen datos genotípicos de estas (Pritchard, Stephens y Donnelly 2000). El modelo también puede asumir que hay K poblaciones parentales desconocidas, caracterizadas por las frecuencias alélicas de cada locus, cada uno de ellos en EHW. Este

algoritmo también puede ser aplicado para inferir estructuración poblacional utilizando datos de genotipos de marcadores no ligados, pero no permite ajustar regresiones con la proporción de mezcla. El método puede ser aplicado a diversos tipos de marcadores, como RFLP, SNP, STR, etc., puesto que no asume una tasa de mutación determinada.

El programa Ancestrymap versión 2.0 estima la ancestría a lo largo del genoma, en una muestra de población con mezcla reciente entre dos grupos étnicos. El programa usa datos genotípicos de marcadores escogidos estratégicamente, por su diferencia en las frecuencias alélicas entre dos poblaciones parentales. Idealmente, estas frecuencias deben ser conocidas. La inclusión de controles no es necesaria para identificar genes asociados a enfermedades, pero es útil para obtener estimaciones robustas de las frecuencias alélicas en las poblaciones ancestrales, lo cual incrementa el poder del análisis (Patterson et ál. 2004).

MATERIALES Y MÉTODOS

Se recolectó información de 849 individuos residentes en el municipio de Medellín y su área metropolitana (municipio de Bello, al norte, y municipio de Envigado, al sur). Los individuos se captaron a través de la consulta externa de endocrinología del Hospital Universitario San Vicente de Paúl, Susalud EPS, Comfenalco, Comfama y en lugares donde el municipio de Medellín ha establecido programas de educación y prevención para grupos de la tercera edad.

Previo consentimiento informado, se realizó una encuesta de origen, para clasificarlos como *antioqueños*. Con al menos cuatro bisabuelos procedentes de esta región, se tomaron muestras de 10 ml de sangre y se extrajo el DNA por el método de fenol-cloroformo. Para la genotipificación de SNP se utilizó una técnica de alto rendimiento, iPLEX Sequenom Mass Array, usando las condiciones experimentales descritas por Tang et ál. (1999).

Para calcular el componente ancestral autosómico, amerindio, africano y europeo se incluyeron en el análisis genotipos de 75 marcadores AIM en un total de 849 individuos. Once de estos marcadores se eligieron a partir de los reportados por Parra, Shriver y Molokhia (Molokhia et ál. 2003; Parra et ál. 1998; Shriver

et ál. 2003). Los demás marcadores se eligieron con base en los mapas de SNP informativos para poblaciones latinas (Mao et ál. 2007; Martínez-Marignac et ál. 2007; Price et ál. 2007; Tian et ál. 2006). Se obtuvieron los genotipos en las poblaciones no mezcladas (parentales): europeos-americanos de Baltimore y Chicago ($n = 77$) y de la colección CEU de Utah ($n = 60$), españoles de Valencia ($n = 31$), africanos del este de Ghana ($n = 52$) y la colección YRI de Nigeria ($n = 60$), y amerindios de poblaciones mazahua ($n = 22$), zapoteca ($n = 60$), mixteca ($n = 23$) y mixe ($n = 29$) (tabla 1).

Para incluir estos marcadores se tuvo en cuenta su poder informativo con respecto a las poblaciones parentales (las diferencias en frecuencias alélicas entre las poblaciones putativas parentales). Con un promedio de diferencia entre europeos y africanos de 0,579; entre africanos y amerindios de 0,375 y entre

europeos y amerindios de 0,440, y espaciados en el genoma, al menos cada 10 cM (es decir, con una longitud mayor a 10 Mb)³

3 cM (centi Morgan): unidad de distancia genética entre loci. Mb (megabase): unidad de longitud de ADN.

en los cromosomas del 1 hasta el 22. Entre los 75 marcadores tipificados, 64 fueron informativos para determinar la contribución europea y africana, 36 fueron útiles para discriminar entre las poblaciones africanas y amerindias y 55 para discriminar entre la europea y la amerindia (tabla 1). Se obtuvieron datos completos para el 92% de los genotipos, y quedó una proporción de datos faltantes del 8%, de la cual se espera que no afecte las estimaciones de las proporciones de mezcla.

TABLA 1. FRECUENCIAS ALÉLICAS DE 75 AIM EN POBLACIONES PARENTALES Y POBLACIÓN ANTIOQUEÑA, Y DELTA (δ) DE FRECUENCIAS ENTRE POBLACIONES PARENTALES DE LOS MARCADORES TIPIFICADOS EN LA MUESTRA

Locus	Crn	Posición física pb	*A1	Frecuencias alélicas parentales				* δ de frecuencias alélicas			
				Europeos	Africanos	Amerindios	Eur-afr.	Afr-amer.	Eur-amer.	Antioquia	
rs6603855	1	16205847	C	0,07	0,88	0,08	0,81	0,8	0,01	0,17	
rs1498111	1	35310701	A	0,96	0,32	0,28	0,64	0,04	0,68	0,7	
rs11208538	1	65101310	C	0,17	0,86	0,8	0,69	0,06	0,63	0,48	
rs2366125	1	112251103	A	0,89	0,75	0,09	0,14	0,67	0,8	0,58	
rs11240089	1	144285330	C	0,12	0,9	0,52	0,78	0,38	0,4	0,33	
rs2814778	1	155987756	G	0	1	0,01	1	0,99	0,01	0,11	
rs3138521 (AT3)	1	170618399	Ins	0,28	0,86	0,06	0,58	0,8	0,22	0,3	
rs10753176	1	174683642	A	0,4	0,97	0,78	0,57	0,18	0,39	0,56	
rs2288697	2	23771820	A	0,03	0,01	0,85	0,03	0,84	0,82	0,27	
rs12469563	2	40806881	C	0,03	0	0,76	0,03	0,76	0,74	0,26	
rs2625051	2	131346589	A	0,16	0,93	0,19	0,77	0,74	0,03	0,24	
rs1000141	2	234024347	A	0,51	0,06	0,12	0,45	0,06	0,39	0,36	
rs1869868	3	46869992	A	0,8	0,13	0,9	0,68	0,77	0,1	0,79	
rs13067741	3	115646846	C	0,2	0,86	0,82	0,65	0,04	0,61	0,46	
rs4677637	3	195956225	A	0,7	0,14	0,25	0,57	0,12	0,45	0,41	
rs12642527	4	22733807	G	0,59	0,02	0,04	0,57	0,02	0,55	0,38	
rs563671	4	43299648	C	0,85	0,17	0,47	0,68	0,3	0,38	0,6	

Comparación de cuatro programas utilizados en la determinación de la composición genética ancestral de la población antioqueña

Locus	Crm	Posición física pb	*A1	Frecuencias alélicas parentales				*δ de frecuencias alélicas			
				Europeos	Africanos	Amerindios		Eur.-afr.	Afr.-amer.	Eur.-amer.	Antioquia
rs7041 (GCHaeIII) ^a	4	72983369		0,16	0,85	0,34		0,7	0,51	0,18	0,49
rs7041 (GCStyl) ^a	4	72983369		0,61	0,07	0,54		0,54	0,47	0,07	0,79
rs1921877	4	85319832	A	0,24	0,88	0,75		0,64	0,13	0,51	0,54
rs1525760	4	117492984	C	0,14	0,92	0,06		0,78	0,86	0,08	0,2
rs11723316	4	184790487	A	0,44	0,89	0,91		0,45	0,02	0,47	0,62
rs2278354	5	10501978	T	0,9	0,28	0,24		0,62	0,04	0,66	0,64
rs16897585	5	28355305	A	0	0	0,78		0	0,78	0,78	0,19
rs874973	5	72773651	T	0,91	0,11	0,91		0,8	0,8	0	0,75
rs227282	5	132041078	G	0,26	0,99	0,87		0,73	0,13	0,61	0,48
rs10499012	6	97874312	A	0	0,05	0,83		0,05	0,77	0,83	0,19
rs4305737	6	145093287	A	0,23	0,89	0,96		0,65	0,07	0,73	0,55
rs768994	6	159369757	T	0,91	0,93	0,21		0,02	0,72	0,7	0,71
rs3055	6	169884144	C	0,67	0,12	0,11		0,56	0	0,56	0,43
rs2965404	7	21522742	C	0,22	0,94	0,76		0,72	0,18	0,54	0,46
rs2010269	7	50897172	C	0,78	0,14	0,33		0,64	0,19	0,45	0,63
rs3808013	7	103392153	G	0,73	0,1	0,2		0,63	0,1	0,53	0,55
rs17161479	7	139434810	C	0,04	0	0,6		0,04	0,6	0,57	0,19
rs285 (LPL)	8	19859469	T	0,49	0,97	0,44		0,48	0,53	0,05	0,46
rs1038545	8	92001312	C	0,65	0,01	0,2		0,65	0,19	0,46	0,47
rs2124036	8	126717316	C	0,26	0,89	0,58		0,63	0,31	0,32	0,41

Locus	Crn	Posición física pb	*A1	Frecuencias alélicas parentales				* δ de frecuencias alélicas			
				Europeos	Africanos	Amerindios		Eur.-afr.	Afr.-amer.	Eur.-amer.	Antioquia
rs913700	9	460086	A	0,82	0,16	0,23		0,66	0,07	0,59	0,56
rs4484797	9	18250419	G	0,89	0,86	0,09		0,03	0,77	0,8	0,66
rs644490	9	122882224	C	0,15	0,98	0,58		0,82	0,39	0,43	0,38
rs660592	10	13188273	T	0,11	0,82	0,46		0,71	0,36	0,36	0,32
rs2394931	10	74336071	T	0,07	0,81	0,03		0,74	0,78	0,04	0,15
rs992528	10	118152572	A	0,09	0,91	0,72		0,82	0,19	0,64	0,41
rs1800498 (DRD2)	11	112796798	T	0,67	0,14	0,05		0,54	0,09	0,63	0,44
rs3138522 (APOA)	11	116217675	Indel	0,93	0,42	0,98		0,51	0,56	0,05	0,9
rs10840686	12	8152917	C	0,74	0,12	0,39		0,62	0,27	0,35	0,52
rs10876851	12	54471828	A	0,03	0,84	0,64		0,81	0,2	0,6	0,29
rs2304439	12	92573458	A	0,94	0,32	0,49		0,62	0,16	0,45	0,77
rs9316044	13	43079611	C	0,73	0,2	0,43		0,54	0,24	0,3	0,52
rs252544 (RB1)	13	47776293	C	0,32	0,93	0,18		0,61	0,75	0,14	0,34
rs11156978	14	20994513	C	0,46	0,03	0,06		0,43	0,03	0,4	0,28
rs8009244	14	54713505	A	0,89	0,14	0,91		0,74	0,77	0,03	0,79
rs2295903	14	66704071	C	0,03	0,68	0,62		0,65	0,06	0,59	0,28
rs730570	14	100212643	A	0,86	0,19	0,06		0,67	0,13	0,8	0,52
rs2714758	15	23030430	A	0,98	0,1	0,98		0,88	0,88	0,01	0,87
rs1800404 (OCA2)	15	25909368	A	0,75	0,12	0,49		0,63	0,37	0,26	0,59
rs4924116	15	35086443	C	0,91	0,99	0,17		0,08	0,82	0,74	0,72

Comparación de cuatro programas utilizados en la determinación de la composición genética ancestral de la población antioqueña

Locus	CrM	Posición física pb	*A1	Frecuencias alélicas parentales				*δ de frecuencias alélicas			
				Europeos	Africanos	Amerindios		Eur.-afr.	Afr.-amer.	Eur.-amer.	Antioquia
rs1426654	15	46213776	A	1	0,03	0,06		0,98	0,03	0,94	0,68
rs734780	15	87365962	T	0,08	0,74	0,73		0,65	0	0,65	0,41
rs9937557	16	8244391	C	0,96	0,92	0,23		0,03	0,69	0,72	0,73
rs3138523 (PV92)	16	81469612	Indel	0,15	0,23	0,79		0,07	0,57	0,64	0,37
rs959071	17	19082819	G	0,89	0,02	0,28		0,87	0,26	0,61	0,63
rs1917913	18	5732915	A	0,06	0,66	0,48		0,6	0,18	0,42	0,23
rs3911730	18	66022323	C	0,08	0,93	0,13		0,85	0,8	0,05	0,17
rs1833791	19	9661576	C	0,81	0,15	0,4		0,66	0,25	0,41	0,65
rs3138524 (Sb19.3)	19	22559998	Indel	0,9	0,42	0,65		0,49	0,23	0,26	0,72
rs2216595	19	38225132	C	0,86	0,24	0,3		0,63	0,07	0,56	0,5
rs1418032	20	2025744	C	0,24	0,94	0,59		0,7	0,35	0,35	0,56
rs6034866	20	17551728	A	0,07	0,93	0,15		0,87	0,78	0,08	0,18
rs878522	20	45050583	A	0,62	0,01	0,05		0,61	0,04	0,57	0,39
rs387098	20	61625038	C	0,17	0,99	0,11		0,82	0,89	0,07	0,29
rs722098	21	15607469	A	0,85	0,12	0,25		0,73	0,14	0,59	0,53
rs915750	21	36803410	A	0,95	0,22	0,28		0,74	0,06	0,68	0,59
rs9606607	22	15943021	C	0,59	0,07	0,2		0,52	0,13	0,39	0,38
rs727563	22	40191877	G	0,22	0,87	0,96		0,65	0,08	0,74	0,52
								Promedio			
								0,58			

Indel = inserción o delección del marcador; CrM = número de cromosoma; la posición física del marcador está reportada en pares bases (pb); (* A1) = las frecuencias reportadas corresponden al alelo A1; (*δ) delta de las frecuencias está reportado en valores absolutos. Eur. = europeos; afr. = africanos; amer. = amerindios [a] = un locus 3 alelos.

Fuente: Mao et ál. [2007], Martínez-Margnac et ál. [2007], Molokhia et ál. [2003], Parra et ál. [1998], Price et ál. [2007], Shriver et ál. [2003], Tian et ál. [2006].

ANÁLISIS ESTADÍSTICO

Con los genotipos de los AIM se calcularon los parámetros poblacionales: las frecuencias alélicas y el equilibrio de Hardy Weinberg (mediante el programa Genepop v. 3.1). En el programa Admixmap v. 3.8 se implementó la prueba de dispersión de frecuencias alélicas para evaluar el ajuste de las frecuencias especificadas en el modelo, con base en la probabilidad predictiva posterior. La dispersión se evalúa para cada subpoblación en cada locus, y de manera global sobre todos los loci. Los resultados son valores *p* bayesianos, valores muy pequeños de la probabilidad posterior predictiva pueden indicar la falta de ajuste. Con este programa también se estimó el parámetro de la suma de intensidades, que corresponde a modelar el proceso estocástico de transición de los estados ancestrales en los cromosomas híbridos. La suma de intensidades es modelada de manera global por una distribución de *Poisson* y mide la frecuencia de dichas transiciones. Este parámetro se puede interpretar como el promedio del número de generaciones ocurridas desde el evento de mezcla.

La composición de la mezcla en cada individuo (europea, africana y amerindia), así como el promedio de cada componente ancestral en la muestra, se obtuvieron usando diferentes programas; el programa Admixmap v. 3.8 para Windows, el programa Structure v. 2.2 y el programa Ancestrymap v. 2.0. Con el programa Admix. Pas (amablemente suministrado por su autor, J. C. Long) se estimó la proporción de mezcla en la población, mediante una aproximación de mínimos cuadrados ponderados; con este programa solo se puede calcular la proporción de mezcla total (Long 1991).

En el programa Admixmap v. 3.8 se implementó un modelo de no dispersión de las frecuencias alélicas de las poblaciones no mezcladas, en el que se asume que las frecuencias alélicas de las poblaciones mezcladas corresponden a las de las poblaciones ancestrales usadas para obtener los parámetros a priori, en una población mezclada, producto de tres poblaciones ancestrales (europea, africana y amerindia). El programa estima las frecuencias alélicas de las poblaciones parentales y de la población mezclada con remuestreo simultáneo (Hoggart et ál. 2003; Hoggart et ál. 2004).

En el programa Structure se implementó un modelo de frecuencias alélicas correlacionadas, en el que se asume que las

frecuencias en las poblaciones son similares, lo cual se acerca más a la realidad en poblaciones estrechamente relacionadas. Por lo tanto, si una población es muy divergente, el modelo de frecuencias correlacionadas puede lograr mejores inferencias cuando se omite esta población. El parámetro α especifica si los individuos pueden ser asignados con mayor probabilidad $\Pr(X/K)$ a una población o si pueden ser asignados a varias poblaciones con probabilidad similar: $\Pr(X/K) = 0,5$ para $K = 2$; $\Pr(X/K) = 0,33$ para $K = 3$. Para obtener estimativos confiables de $\Pr(X/K)$, los parámetros del modelo deben alcanzar el equilibrio antes de finalizar la longitud de la simulación, la cual se determina realizando simulaciones previas, para determinar la longitud de las iteraciones y de la simulación en la cadena de Markov. Los parámetros se obtuvieron con 20.000 iteraciones (Pritchard y Rosenberg 1999; Pritchard, Stephens y Donnelly 2000).

En el programa Ancestrymap se explora la distribución posterior de varios parámetros, dados los datos, y se obtienen estadísticas que dan cuenta de la incertidumbre de esos parámetros desconocidos mediante un número suficiente de iteraciones. El programa minimiza el número de iteraciones requerido usando un algoritmo que maximiza los valores esperados para escoger los valores más cercanos a los reales. La estimación de la mezcla individual a lo largo del genoma es promediada entre todos los individuos para identificar las regiones genómicas con mayor contenido de una de las poblaciones parentales. El algoritmo usa un modelo de Markov escondido, en el que el estado ancestral escondido es inferido a partir de los genotipos (Patterson et ál. 2004).

RESULTADOS Y DISCUSIÓN

Las frecuencias alélicas de 75 AIM en las poblaciones parentales y en la población antioqueña se muestran en la tabla 1. Menos del 5% de los marcadores presentaron desviaciones significativas del EHW con una prueba de χ^2 y después de corregir por Bonferroni las comparaciones múltiples.

La prueba de dispersión implementada en Admixmap para determinar el ajuste de las frecuencias alélicas actuales de las poblaciones parentales, con respecto a las frecuencias alélicas de las poblaciones ancestrales que participaron en la mezcla,

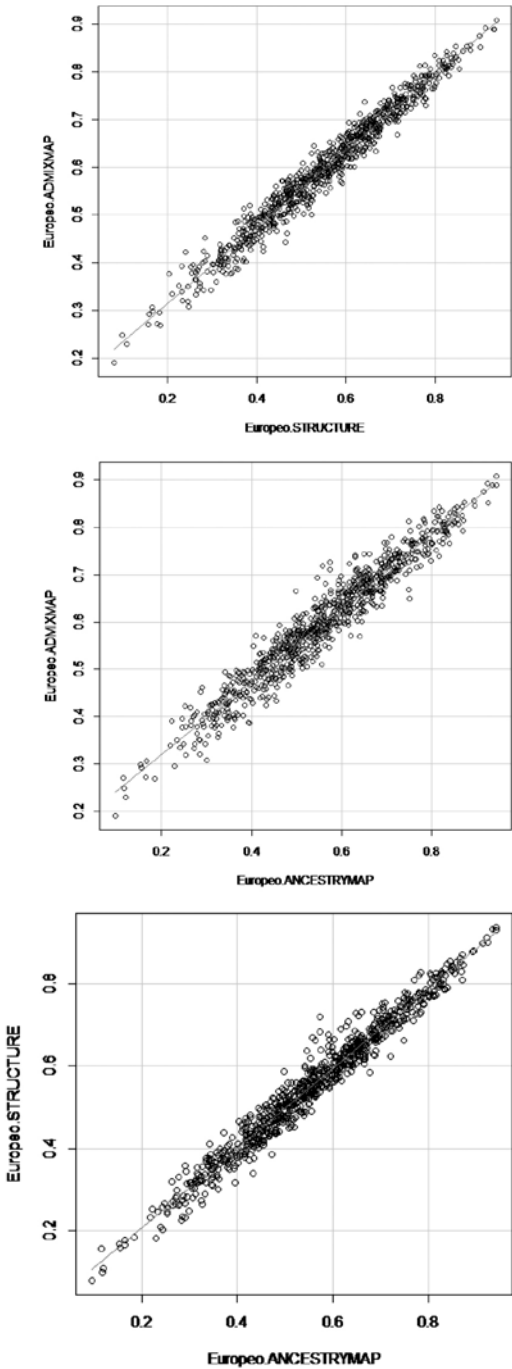
mostró valores de probabilidad posterior predictiva (valor p bayesiano) de 0,0315, 0,368 y 0,071 para europeos, africanos y amerindios respectivamente. La dispersión significativa de las frecuencias se asume cuando se encuentran valores de p muy bajos que pueden reflejar la falta de ajuste (Hoggart et ál. 2003; Pritchard, Stephens y Donnelly 2000).

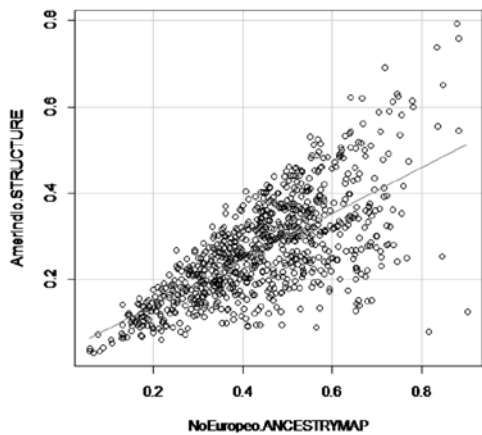
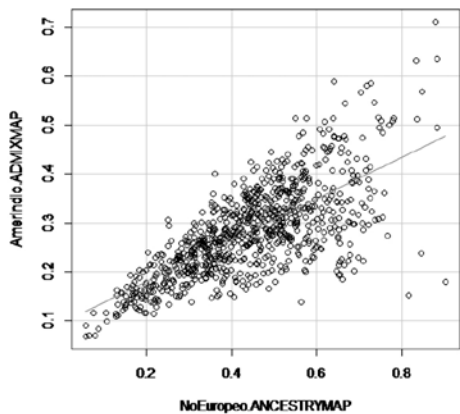
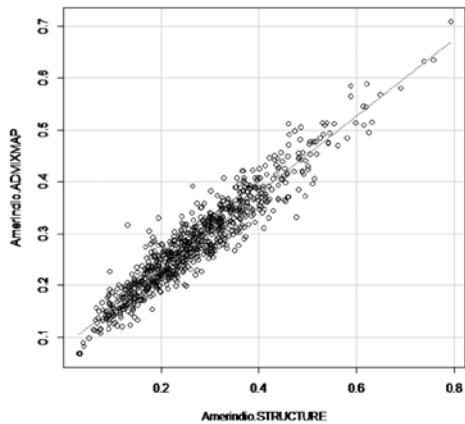
El cálculo de las proporciones de mezcla con poca incertidumbre, dado por los bajos niveles de dispersión de las frecuencias alélicas, depende en gran parte de cuánto se relacionen las que se usan como parentales con respecto a las de las poblaciones ancestrales reales (Price et ál. 2007). Por lo tanto, los valores obtenidos indican un buen ajuste (no dispersión) de las frecuencias alélicas en las poblaciones no mezcladas que se usaron para obtener los parámetros a priori, y sugieren que el panel de 75 AIM utilizado fue adecuado para calcular las proporciones de mezcla en esta población.

El parámetro de la suma de intensidades fue de 11 por cada 100 cM con un IC de 95% (9,0 - 13,3). Lo anterior implica que el promedio de tiempo en que ocurrió la mezcla es de al menos 11 generaciones, desde el primer evento de mezcla. Es probable que la variación observada en la estimación de este parámetro esté influenciada por diferentes niveles de flujo génico, en comparación con otras poblaciones latinoamericanas; por ejemplo, en la población mexicana el valor estimado para este parámetro fue de 6,7 por 100 cM con IC de 95% (5,7 - 8,0) (Martínez-Marignac et ál. 2007), en una población afroamericana fue de 6,0 por 100 cM (Patterson et ál. 2004) y en una población hispana, de 8 por 100 cM (Hoggart et ál. 2004). Las poblaciones relativamente aisladas presentan tiempos mayores desde el evento de mezcla, mientras que poblaciones vecinas de grandes poblaciones amerindias o de regiones con reciente migración europea muestran tiempos más cortos.

El parámetro de la suma de intensidades es directamente proporcional a la densidad de marcadores necesarios para encontrar variantes en genes con efecto sobre un determinado rasgo. Se estimó que en la población mexicana es necesario tipificar 1.400 marcadores informativos de ancestría para localizar genes que confieren riesgo para enfermedades en la población amerindia y europea, bajo un modelo de flujo génico continuo, y que son necesarios al menos 2.000 individuos para detectar un OR de al menos 1,5. La historia demográfica de la población antioqueña

Comparación de cuatro programas utilizados en la determinación de la composición genética ancestral de la población antioqueña





Comparación de cuatro programas utilizados en la determinación de la composición genética ancestral de la población antioqueña

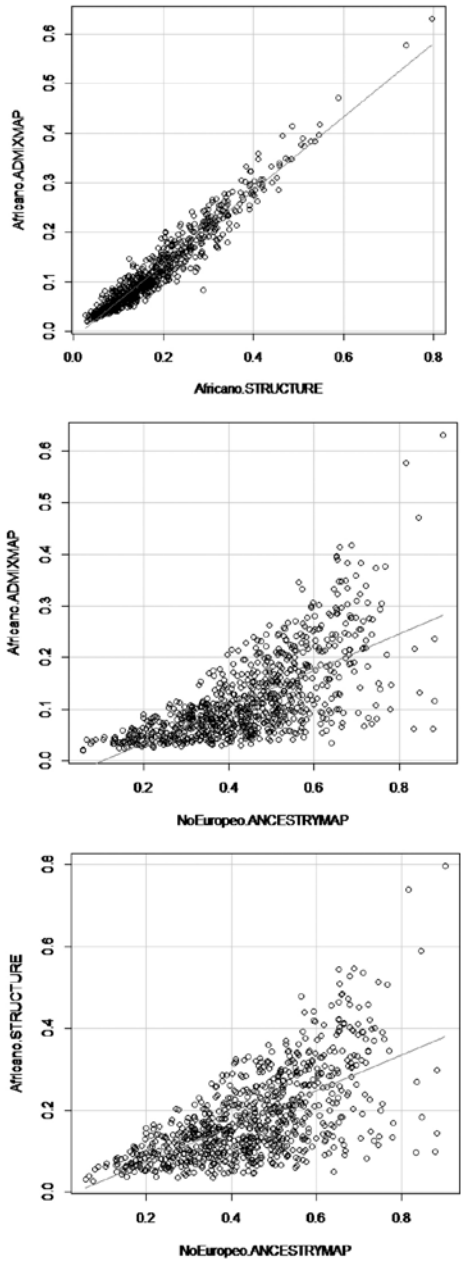


FIGURA 1. COMPARACIÓN DE LOS PROGRAMAS UTILIZADOS PARA ESTIMAR MEZCLA INDIVIDUAL

Fuente: Esta investigación.

sugiere que el valor encontrado es consistente con un modelo de un único evento de mezcla y que el número de marcadores necesarios para implementar MALD puede ser igual o menor al que se requiere en la población mexicana (Martínez-Marignac et ál. 2007).

Usando el panel de 75 AIM altamente informativo se caracterizó la proporción de la mezcla europea, africana y amerindia en la población total y para cada individuo. Las proporciones de mezcla individual calculadas con los diferentes programas utilizados fueron muy similares (figura 1). Se calculó una correlación de Pearson para cada componente de mezcla entre pares de programas; los coeficientes de correlación se muestran en la tabla 2. Los análisis indican una muy alta correlación para los componentes europeo, africano y amerindio calculados con los programas Admixmap y Structure. Sin embargo, al analizar este componente entre estos programas y Ancestrymap la correlación disminuye. Los promedios de los componentes ancestrales que se obtuvieron con los diferentes programas utilizados se muestran en la tabla 3. Lo anterior sugiere que los algoritmos implementados en los programas Admixmap y Structure son similares y robustos para determinar las proporciones de mezcla poblacional e individual. No obstante, este tipo de análisis en las poblaciones mezcladas producto de tres o más poblaciones parentales no es apropiado con algoritmos como el de Ancestrymap, debido a que solo se pueden asumir dos poblaciones parentales y con el programa Admix.Pas no es posible obtener proporciones de mezcla individual.

TABLA 2. COEFICIENTES DE CORRELACIÓN DE PEARSON PARA CADA COMPONENTE DE MEZCLA INDIVIDUAL ENTRE LOS DIFERENTES PROGRAMAS UTILIZADOS

Componente	Admixmap vs. Structure	Admixmap vs. Ancestrymap	Structure vs. Ancestrymap
Europeo	*0,983	*0,968	*0,983
Africano	*0,967	**0,713	**0,699
Amerindio	*0,950	**0,713	**0,699

* Coeficientes de correlación.

** El coeficiente de correlación con Ancestrymap entre europeo y no europeo (africano y amerindio).

Fuente: Esta investigación.

Comparación de cuatro programas utilizados en la determinación de la composición genética ancestral de la población antioqueña

TABLA 3. PROPORCIONES DE MEZCLA GENÉTICA OBTENIDAS CON LOS PROGRAMAS ADMIX-MAP, STRUCTURE, ANCESTRYMAP Y ADMIX.PAS

Software	Europeo	Africano	Amerindio
Admixmap	0,60 ± 0,016	0,12 ± 0,007	0,28 ± 0,009
Structure	0,56 ± 0,022	0,15 ± 0,009	0,29 ± 0,013
Ancestrymap	0,56 ± 0,024	0,44 ± 0,024	
Admix.Pas	0,58 ± 0,014	0,13 ± 0,013	0,20 ± 0,016

Nota: Los valores son promedios más o menos el error estándar.

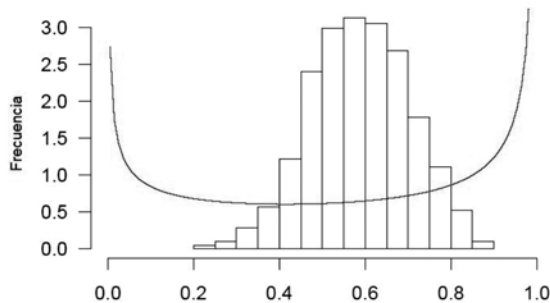
Fuente: Esta investigación.

Las proporciones de mezcla promedio calculadas en esta muestra de población antioqueña con el programa Admixmap muestran, de manera consistente con otros reportes (Bedoya, Montoya et ál. 2006; Carvajal-Carmona et ál. 2000; Carvajal-Carmona et ál. 2003; Gómez-Pérez et ál. 2010), la contribución relativa de tres poblaciones putativas parentales, con una contribución europea (60%) dos veces mayor que la contribución amerindia (28%), ambas con una distribución amplia, y una escasa contribución africana (12%). Comparando los promedios de mezcla obtenidos en esta muestra con varios reportes sobre la población antioqueña (Bedoya, Montoya et ál. 2006; Carvajal-Carmona et ál. 2000; Carvajal-Carmona et ál. 2003; Gómez-Pérez et ál. 2010), se observa una diferencia substancial en la estimación de las proporciones de mezcla, en particular para europeos y amerindios. Algunas razones que explican las diferencias con este estudio son: primero, en estos estudios la estimación del mestizaje se basó en un número diferente de marcadores AIM; y segundo, las características de la muestra son diferentes. Sin embargo, el método bayesiano implementado en los programas Admixmap, Structure y Ancestrymap, y el método de mínimos cuadrados ponderados del programa Admix.Pas mostraron ser robustos y consistentes para estimar los componentes de mezcla con poca incertidumbre.

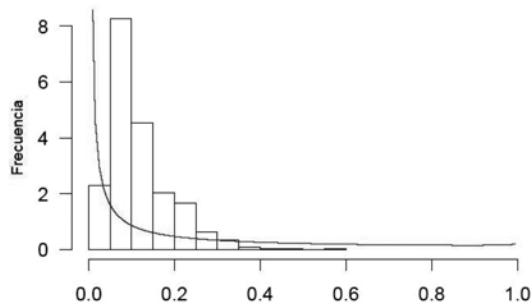
Las proporciones de mezcla observadas en la población antioqueña son similares a las reportadas para una población hispana del Valle de San Luis (Colorado), en la cual se encontró un 63% de contribución europea, un 34% de amerindio y un 3% de africano (Bonilla et ál. 2004); pero difieren de los reportes acerca de otras poblaciones de Latinoamérica y de Colombia. Por ejemplo, en un estudio realizado en una población mexicana se reportó una

contribución europea de 30%, amerindia del 64% y africana del 5% (Martínez-Marignac et ál. 2007), y entre los individuos de una población de Cartagena (departamento de Bolívar) se encontró una mayor contribución europea (43%), una considerable contribución africana (38,3%) y muy poca contribución amerindia (18,6%) (Vergara et ál. 2009). Estas diferencias pueden ser atribuidas a los diferentes patrones demográficos propios de cada población, como la densidad de la población amerindia en la época en que llegaron los colonizadores y la cantidad de europeos y africanos que migraron a cada región (Price et ál. 2007).

Las proporciones de mezcla individual obtenidas con el programa Admixmap se muestran en las figuras 2A, 2B y 2C. La distribución de los componentes ancestrales individuales fue la siguiente: entre 20 y 90% para europeo, entre 0 y 35% para africano y entre 10 y 60% para amerindio. Estos patrones de distribución de la mezcla individual en la población antioqueña son bastante amplios, principalmente en europeos y amerindios.



A. Proporción de mezcla individual europea



B. Proporción de mezcla individual africana

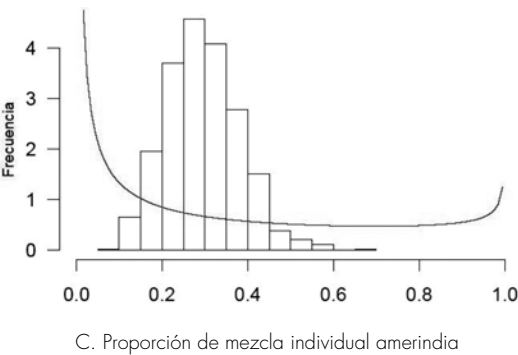


FIGURA 2. DISTRIBUCIÓN DE LA MEZCLA INDIVIDUAL EN LA MUESTRA ESTIMADA CON EL PROGRAMA ADMIXMAP, CONTRIBUCIONES (A) EUROPEA, (B) AFRICANA Y (C) AMERINDIA.
Fuente: Esta investigación.

Para implementar estudios de MALD tendientes a identificar loci de riesgo para enfermedades con prevalencias diferentes entre las poblaciones parentales es importante que las poblaciones muestren una distribución amplia del componente ancestral parental que confiere mayor riesgo para una entidad específica; lo ideal es que esta distribución esté entre el 10 y el 90% (Patterson et ál. 2004; Price et ál. 2007). Estos resultados sustentan la importancia de las poblaciones mezcladas como recurso útil para abordar el mapeo de genes para enfermedades con prevalencias diferentes entre poblaciones.

CONCLUSIONES

Se comprobó la robustez y utilidad de diferentes programas para modelar la mezcla poblacional; las diferencias observadas son inherentes a los algoritmos implementados en cada uno de ellos. En este estudio se comprobó también el poder informativo del panel de marcadores utilizados. Este panel de marcadores fue adecuado para estimar las proporciones de mezcla individual, pero no es suficiente para aportar información de utilidad epidemiológica. Las diferencias observadas en las proporciones de mezcla estimadas en este estudio con respecto a otros reportes pueden explicarse por la cantidad de marcadores utilizados, su

informatividad y, además, el tamaño y las características demográficas de las poblaciones. El promedio de tiempo estimado desde el primer evento de mezcla fue de 11 generaciones por cada 100 cM, lo cual implica que el DL generado por la mezcla no ha decaído considerablemente. Además, la distribución de mezcla individual amerindia y europea en la población antioqueña es suficientemente amplia. Lo anterior hace de la población antioqueña una muestra interesante para implementar estudios de mapeo por mezcla o MALD, principalmente en rasgos con diferente prevalencia entre europeos y amerindios. Esto también podría ser una constante en otras poblaciones latinas que valdría la pena explorar.

AGRADECIMIENTOS

Los autores agradecen al Comité para el Desarrollo de la investigación (CODI) de la Universidad de Antioquia (Proyecto de Sostenibilidad de Grupos 2009-2011 y CPT 0218) y al Instituto para el Desarrollo de la Ciencia y la Tecnología (Colciencias, código 111504012986).

REFERENCIAS

- BEDOYA, GABRIEL, JENNY GARCÍA, PATRICIA MONTOYA, WINSTON ROJAS, MARÍA EUGENIA AMÉZQUITA, IVÁN SOTO, MARÍA CECILIA LÓPEZ, JORGE OSPINA-DUQUE Y ANDRÉS RUIZ-LINARES. 2006. "[Isonymy Analysis between 2 Populations in Northwestern Colombia]". *Biomédica* 26 (4): 538-545.
- BEDOYA, GABRIEL, PATRICIA MONTOYA, JENNY GARCÍA, IVÁN SOTO, STEPHANE BOURGEOIS, LUIS CARVAJAL, DAMIAN LABUDA, VÍCTOR ÁLVAREZ, JORGE OSPINA, PHILIP. W. HEDRICK Y ANDRÉS RUIZ-LINARES. 2006. "Admixture Dynamics in Hispanics: A Shift in the Nuclear Genetic Ancestry of a South American Population Isolate". *Proceedings of the National Academy of Sciences* 103 (19): 7234-7239.
- BONILLA, CAROLINA, ESTEBAN J. PARRA, CARRIE L. PFAFF, SONIA DIOS, JULIE A. MARSHALL, RICHARD F. HAMMAN, ROBERT E. FERRELL, CLIVE L. HOGGART, PAUL M. McKEIGUE Y MARK D. SHRIVER. 2004. "Admixture in the Hispanics of the San Luis Valley, Colorado, and Its Implications for Complex Trait Gene Mapping". *Annals of Human Genetics* 68, parte 2: 139-153.

- BRYC, KATARZYNA, ADAM AUTON, MATTHEW R. NELSON, JORGE R. OKSENBURG, STEPHEN L. HAUSER, SCOTT WILLIAMS, ALAIN FROMENT, JEAN M. BODO, CHARLES WAMBEBE, SARAH A. TISHKOFF Y CARLOS. D. BUSTAMANTE. 2010. "Genome-wide Patterns of Population Structure and Admixture in West Africans and African Americans". *Proceedings of the National Academy of Sciences* 107 (2): 786-791.
- CARVAJAL-CARMONA, LUIS G., ROEL OPHOFF, SUSAN SERVICE, JAANA HARTIALA, JULIO MOLINA, PEDRO LEÓN, JORGE OSPINA, GABRIEL BEDOYA, NELSON FREIMER Y ANDRÉS RUIZ-LINARES. 2003. "Genetic Demography of Antioquia (Colombia) and the Central Valley of Costa Rica". *Human Genetics* 112 (5-6): 534-541.
- CARVAJAL-CARMONA, L. G., IVÁN D. SOTO, NICOLÁS PINEDA, DANIEL ORTIZ-BARRIENTOS, CONSTANZA DUQUE, JORGE OSPINA-DUQUE, MARK MCCARTHY, PATRICIA MONTOYA, VÍCTOR M. ÁLVAREZ, GABRIEL BEDOYA Y ANDRÉS RUIZ-LINARES. 2000. "Strong Amerind/white Sex Bias and a Possible Sephardic Contribution among the Founders of a Population in Northwest Colombia". *American Journal of Human Genetics* 67 (5): 1287-1295.
- FALUSH, DANIEL, MATTHEW STEPHENS Y JONATHAN K. PRITCHARD. 2003. "Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies". *Genetics* 164 (4): 1567-1587.
- GÓMEZ-PÉREZ, LUIS, MIGUEL A. ALFONSO-SÁNCHEZ, ANA M. PÉREZ-MIRANDA, SUSANA GARCÍA-OBREGÓN, JUAN J. BUILES, MARÍA L. BRAVO, MARIAN M. DE PANCORBO Y JOSÉ A. PENA. 2010. "Genetic Admixture Estimates by Alu Elements in Afro-Colombian and Mestizo Populations from Antioquia, Colombia". *Annals of Human Biology* 37 (4): 488-500. doi: 10.3109/03014460903433810.
- HALDER, INDRANI Y MARK D. SHRIVER. 2003. "Measuring and Using Admixture to Study the Genetics of Complex Diseases". *Human Genomics* 1 (1): 52-62.
- HOGGART, CLIVE J., ESTEBAN J. PARRA, MARK D. SHRIVER, CAROLINA BONILLA, RICK A. KITTLES, DAVID G. CLAYTON Y PAUL M. McKEIGUE. 2003. "Control of Confounding of Genetic Associations in Stratified Populations". *American Journal of Human Genetics* 72 (6): 1492-1504.
- HOGGART, CLIVE J., MARK D. SHRIVER, RICK A. KITTLES, DAVID G. CLAYTON Y PAUL M. McKEIGUE. 2004. "Design and Analysis of Admixture Mapping Studies". *American Journal of Human Genetics* 74 (5): 965-978.
- INTERNATIONAL HAPMAP CONSORTIUM. 2003. "The International HapMap Project". *Nature* 426 (6968): 789-796.
- Long, Jeffrey C. 1991. "The Genetic Structure of Admixed Populations". *Genetics* 127: 417-428.

- MAO, XIANYUN, ABIGAIL W. BIGHAM, RUI MEI, GERARDO GUTIÉRREZ, KEN M. WEISS, TOM D. BRUTSAERT, FABIOLA LEÓN-VELARDE, LORNA G. MOORE, ENRIQUE VARGAS, PAUL M. McKEIGUE, MARK D. SHRIVER Y ESTEBAN J. PARRA. 2007. "A Genomewide Admixture Mapping Panel for Hispanic/Latino Populations". *American Journal of Human Genetics* 80 (6): 1171-1178.
- MARTÍNEZ-MARIGNAC, VERÓNICA L., ADÁN VALLADARES, EMILY CAMERON, ANDREA CHAN, ARJUNA PERERA, RACHEL GLOBUS-GOLDBERG, NIELS WACHER, JESÚS KUMATE, PAUL McKEIGUE, DAVID O'DONNELL, MARK D. SHRIVER, MIGUEL CRUZ Y ESTEBAN J. PARRA. 2007. "Admixture in Mexico City: Implications for Admixture Mapping of Type 2 Diabetes Genetic Risk Factors". *Human Genetics* 120 (6): 807-819.
- McKEIGUE, PAUL M. 2005. "Prospects for Admixture Mapping of Complex Traits". *American Journal of Human Genetics* 76 (1): 1-7.
- MOLOKHIA, MARIAM, CLIVE HOGGART, ALAN L. PATRICK, MARK SHRIVER, ESTEBAN PARRA, J. YE, ALAN J. SILMAN Y PAUL M. McKEIGUE. 2003. "Relation of Risk of Systemic Lupus Erythematosus to West African Admixture in a Caribbean Population". *Human Genetics* 112 (3): 310-318.
- MONTANA, GIOVANNI Y JONATHAN K. PRITCHARD. 2004. "Statistical Tests for Admixture Mapping with Case-control and Cases-only Data". *American Journal of Human Genetics* 75 (5): 771-789.
- PARRA, ESTEBAN J., AMY MARCINI, JOSHUA AKEY, JEREMY MARTINSON, MARK A. BATZER, RICHARD COOPER, TERRENCE FORRESTER, DAVID B. ALLISON, RANJAN DEKA, ROBERT E. FERRELL Y MARK D. SHRIVER. 1998. "Estimating African American Admixture Proportions By Use of Population-specific Alleles". *American Journal of Human Genetics* 63 (6): 1839-1851.
- PATTERSON, NICK, NEIL HATTANGADI, BARTON LANE, KIRK E. LOHMEUILLER, DAVID A. HAFNER, JORGE R. OKSENBERG, STEPHEN L. HAUSER, MICHAEL W. SMITH, STEPHEN J. O'BRIEN, DAVID ALTSHULER, MARK J. DALY Y DAVID REICH. 2004. "Methods for High-Density Admixture Mapping of Disease Genes". *American Journal of Human Genetics* 74 (5): 979-1000.
- PRICE, ALKES L., NICK PATTERSON, FULI YU, DAVID R. COX, ALICJA WALISZEWSKA, GAVIN J. McDONALD, ARTI TANDON, CHISTINE SCHIRMER, JULIE NEUBAUER, GABRIEL BEDOYA, CONSTANZA DUQUE, ALBERTO VILLEGAS, MARÍA CATIRA BORTOLINI, FRANCISCO M. SALZANO, CARLA GALLO, GUIDO MAZZOTTI, MARCELA TELLO-RUIZ, LAURA RIBA, CARLOS A. AGUILAR-SALINAS, SAMUEL CANIZALES-QUINTEROS, MARTA MENJIVAR, WILLIAM KLITZ, BRIAN HENDERSON, CHRISTOPHER A. HAIMAN, CHERYL WINKLER, TERESA TUSIE-LUNA, ANDRÉS RUIZ-LINARES Y DAVID REICH. 2007. "A Genomewide Admixture Map for Latino Populations". *American Journal of Human Genetics* 80 (6): 1024-1036.
- PRITCHARD, JONATHAN K. Y NOAH A. ROSENBERG. 1999. "Use of Unlinked Genetic Markers to Detect Population Stratification in Association Studies". *American Journal of Human Genetics* 65 (1): 220-228.

- PRITCHARD, JONATHAN K., MATTHEW STEPHENS Y PETER DONNELLY. 2000. "Inference of Population Structure Using Multilocus Genotype Data". *Genetics* 155 (2): 945-959.
- REICH, DAVID Y NICK PATTERSON. 2005. "Will Admixture Mapping Work to Find Disease Genes?". *Philosophical Transactions of the Royal Society: Biological Sciences* 360 (1460): 1605-1607.
- SÁNCHEZ ALBORNOZ, NICOLÁS. 1977. *La poblacion de America Latina: desde los tiempos precolombinos*. Madrid: Alianza.
- SELDIN, MICHAEL F. 2007. "Admixture Mapping as a Tool in Gene Discovery". *Current Opinion in Genetics and Development* 17 (3): 177-181.
- SELDIN, MICHAEL F., BOGDAN PASANIUC Y ALKES L. PRICE. 2011. "New Approaches to Disease Mapping in Admixed Populations". *Nature Reviews Genetics* 12 (8): 523-528.
- SERVICE, SUSAN, JOSEPH DEYOUNG, MARÍA KARAYIORGOU, J. LOUW ROOS, HERMAN PRETORIOUS, GABRIEL BEDOYA, JORGE OSPINA, ANDRÉS RUIZ-LINARES, ANTONIO MACEDO, JOANA A. PALHA, PETER HEUTINK, YURII AULCHENKO, BEN OOSTRA, CORNELIA VAN DUJIN, MARJO RIJTTA JARVELIN, TEPPO VARILLO, LYNETTE PEDDLE, PROTON RAHMAN, GIOVANNA PIRAS, MARÍA MONNE, SARAH MURRAY, LUANA GALVER, LEENA PELTONEN, CHIARA SABATTI, ANDREW COLLINS Y NELSON FREIMER. 2006. "Magnitude and Distribution of Linkage Disequilibrium in Population Isolates and Implications for Genome-wide Association Studies". *Nature Genetics* 38 (5): 556-560.
- SHRIVER, MARK D., ESTEBAN J. PARRA, SONIA DIOS, CAROLINA BONILLA, HEATHER NORTON, CELINA JOVEL, CARRIE PFAFF, CECILY JONES, AISHA MASSAC, NEIL CAMERON, ARCHIE BARON, TABITHA JACKSON, GEORGE ARGYROPOULOS, LI JIN, CLIVE J. HOGGART, PAUL M. McKEIGUE Y RICK A. KITTLES. 2003. "Skin Pigmentation, Biogeographical Ancestry and Admixture Mapping". *Human Genetics* 112 (4): 387-399.
- SMITH, MICHAEL W. Y STEPHEN J. O'BRIEN. 2005. "Mapping by Admixture Linkage Disequilibrium: Advances, Limitations and Guidelines". *Nature Reviews Genetics* 6 (8): 623-632.
- SMITH, MICHAEL W., NICK PATTERSON, JAMES A. LAUTENBERGER, ANN L. TRUELOVE, GAVIN J. McDONALD, ALICJA WALISZEWSKA, BAILEY D. KESSING, MICHAEL J. MALASKY, CHARLES SCAFE, ERNEST LE, PHILIP L. DE JAGER, ANDRÉ A. MIGNAULT, ZENG YI, GUY DE THE, MYRON ESSEX, JEAN L. SANKALE, JASON H. MOORE, KWABENA POKU, JOHN P. PHAIR, JAMES J. GOEDERT, DAVID VLAHOV, SCOTT M. WILLIAMS, SARAH A. TISHKOFF, CHERYL A. WINKLER, FRANCISCO M. DE LA VEGA, TREVOR WOODAGE, JOHN J. SNINSKY, DAVID A. HAFLE, DAVID ALTSHULER, DENNIS. A. GILBERT, STEPHEN J. O'BRIEN Y DAVID REICH. 2004. "A

- High-density Admixture Map for Disease Gene Discovery in African Americans". *American Journal of Human Genetics* 74 (5): 1001-1013.
- TANG, KAI, DONG J. FU, DOMINIQUE JULIEN, ANDREAS BRAUN, CHARLES R. CANTOR Y HUBERT KOSTER. 1999. "Chip-based Genotyping by Mass Spectrometry". *Proceedings of the National Academy of Sciences* 96 (18): 10016-10020.
- TIAN, CHAO, DAVID A. HINDS, RUSSELL SHIGETA, RICK KITTLES, DENNIS G. BALLINGER Y MICHAEL F. SELDIN. 2006. "A Genomewide Single-nucleotide-polymorphism Panel with High Ancestry Information for African American Admixture Mapping". *American Journal of Human Genetics* 79 (4): 640-649. doi: S0002-9297(07)63074-2 [pii].
- VERGARA, CANDELARIA, LUIS CARABALLO, DILIA MERCADO, SILVIA JIMÉNEZ, WINSTON ROJAS, NICHOLAS RAFAELS, TRACEY HAND, MÓNICA CAMPBELL, YUJUNG J. TSAI, LI GAO, CONSTANZA DUQUE, SERGIO LÓPEZ, GABRIEL BEDOYA, ANDRÉS RUIZ-LINARES Y KATHLEEN C. BARNES. 2009. "African Ancestry is Associated with Risk of Asthma and High Total Serum IgE in a Population from the Caribbean Coast of Colombia". *Human Genetics* 125 (5-6): 565-579.
- WANG, SIJIA, NICOLAS RAY, WINSTON ROJAS, MARÍA V. PARRA, GABRIEL BEDOYA, CARLA GALLO, GIOVANNI POLETTI, GUIDO MAZZOTTI, KIM HILL, ANA M. HURTADO, BEATRIZ CAMRENA, HUMBERTO NICOLINI, WILLIAM KLITZ, RAMIRO BARRANTES, JULIO A. MOLINA, NELSON B. FREIMER, MARÍA CATIRA BORTOLINI, FRANCISCO M. SALZANO, MARÍA L. PETZL-ERLER, LUIZA T. TSUNETO, JOSÉ E. DIPIERRI, EMMA L. ALFARO, GRACIELA BAILLIET, NÉSTOR O. BIANCHI, ELENA LLOP, FRANCISCO ROTHHAMMER, LAURENT EXCOFFIER Y ANDRÉS RUIZ-LINARES. 2008. "Geographic Patterns of Genome Admixture in Latin American Mestizos". *PLoS Genetics* 4 (3): e1000037.