

Estudio genético multiloci en la población colombiana I. Desequilibrio de ligamiento para el sistema CcEe.

Díaz J.A.¹, Bustos I.²

Resumen

Una muestra de 11.113 individuos de la población colombiana, caracterizados para el sistema sanguíneo Rh con los sueros anti-C, anti-c, anti-E y anti-e, se distribuyeron en diez regiones de acuerdo con su lugar de nacimiento y dos generaciones, padres e hijos; se estimó el desequilibrio de ligamiento por el método del haplotipo para el sistema CcEe. Todos los valores de desequilibrio calculados fueron negativos y altos, debido a la baja asociación encontrada de la gameta CE; además, en todas las regiones y en la población total para las dos generaciones, se encontró que el valor de desequilibrio es cercano al máximo valor de desequilibrio posible.

Summary

A sample of 11,113 individuals Colombians, taken from ten regions according to place of birth spanning two generations, parents and children, were characterised for the Rh factor using anti-C, anti-c, anti-E and anti-e serum.

The linkage disequilibrium was estimated using the haplotype method for the CcEe system. The estimated disequilibrium values were high and negative; this was caused by the low association of the CE gamete. Furthermore, in all the regions, for the total sample and for both generations, the disequilibrium value was found to be near the maximum possible disequilibrium.

El desequilibrio de ligamiento o desequilibrio gamético, se define como la condición en la cual las frecuencias gaméticas de los sistemas considerados en conjunto no coinciden con los valores esperados; esta situación se genera por la asociación no aleatoria de los genes situados en diferentes loci (1).

Existen diferentes mecanismos que pueden generar desequilibrio permanente en una población: la interacción epistática entre los loci involucrados, la deriva genética y la selección natural (2); así mismo, la migración entre pobla-

ciones con frecuencias génicas diferentes y la subdivisión de una población (3).

En este trabajo se estudia el desequilibrio de ligamiento para el complejo CcEe del grupo sanguíneo Rh y los factores que inciden en el desequilibrio de ligamiento de la población total, especialmente, la presencia en el país de regiones con diferente grado de mezcla de los grupos étnicos que confluyeron en Colombia.

Materiales y métodos

Una muestra de 11.113 individuos de la población colombiana, caracterizada por el sistema sanguíneo Rh con los antisueros anti-C, anti-c, anti-E y anti-e, fueron distribuidos en diez regiones geográficas, de acuerdo con el lugar de nacimiento y en dos generaciones, padres e hijos (tabla 1).

¹ Profesor catedrático, Universidad Pedagógica Nacional. Departamento de Biología, Santafé de Bogotá.

² Profesora asociada. Universidad Nacional de Colombia, Departamento de Biología, Santafé de Bogotá.

Tabla 1. Número de individuos estudiados por regiones y generaciones

Región	Departamento	Generación	n
1	Cundinamarca	1	1.974
		2	1.121
2	Boyacá	1	1.461
		2	694
3	Antioquia	1	718
		2	356
4	Caldas	1	632
	Quindío	-	-
	Risaralda	2	292
5	Tolima	1	839
		2	390
6	Norte de Santander	1	346
		2	156
7	Meta	1	269
	Caquetá	-	-
	Casanare	2	160
8	Chocó	1	373
	Valle	-	-
	Cauca	-	-
	Nariño	2	170
9	Guajira	1	96
	Cesar	-	-
	Magdalena	2	41
10	Atlántico	-	-
	Bolívar	1	674
	Córdoba	-	-
	Sucre	2	351
	San Andrés	-	-
Padres		1	7.382
Hijos		2	3.731
Total			11.113

Para determinar el ajuste al equilibrio de Hardy y Weinberg de las frecuencias fenotípicas del sistema CcEe en cada región y en la población total, se emplearon las estadísticas X_T y X_C (4); además, se utilizó el chi-cuadrado de homogeneidad (X_H) para determinar si los diferentes grupos (regiones y generaciones) presentaban una distribución fenotípica similar (5).

El desequilibrio de ligamiento D_{AB} se estimó en las diferentes regiones para las dos generaciones por el método del haplotipo:

$$D_{AB} = G4^{1/2} - ((G3 + G4)(G2 + G4))^{1/2} \quad (6)$$

donde:

G2: frecuencia fenotípica de los individuos A-bb

G3: frecuencia fenotípica de los individuos aaB-

G1: frecuencia fenotípica de los individuos A-B-

G4: frecuencia fenotípica de los individuos aabb

Para la estandarización de D_{AB} , se empleó el método de máxima intensidad relativa de D_{AB} (D'):

$$D' = \frac{D_i}{D_{i(max)}} \quad (7)$$

donde:

D_i : coeficiente de desequilibrio de ligamiento en la i-ésima subpoblación.

$D_{i(max)}$: máximo valor que D_i puede alcanzar en la i-ésima subpoblación.

El desequilibrio de la población total, se estimó empleando la ecuación propuesta para poblaciones subdivididas.

$$D_T = \bar{D} + Cov(P_A, P_B) \quad (3)$$

donde:

\bar{D} : coeficiente de desequilibrio promedio de todas las subpoblaciones.

$Cov(P_A, P_B)$: covarianza de las frecuencias génicas.

P_A : frecuencia génica del alelo A

P_B : frecuencia génica del alelo B

La prueba de la hipótesis $D_{AB}=0$ se llevó a cabo mediante los estadísticos Tr^2 y el coeficiente de correlación (r).

La descomposición del desequilibrio de la población colombiana, se hizo empleando el método de partición del desequilibrio total (DT).

$$(N - 1)S_T = (I - 1)S_A + (N - I)S_W \quad (5).$$

En este método se asigna el valor de 1 a los fenotipos CC y EE, 1/2 a los fenotipos Cc y Ee y 0 a los fenotipos cc y ee. Las covarianzas para las frecuencias génicas por regiones (S_i), dentro de regiones (S_w), entre regiones (S_A) y para la población total (S_T), se calcularon a partir de las siguientes ecuaciones:

$$S_i = \frac{\sum (X_{ij} - \bar{X}_i)(Y_{ij} - \bar{Y}_i)}{J_i - 1}$$

$$S_w = \frac{\sum (J_i - 1)S_i}{\bar{N} - 1}$$

$$S_A = \frac{\sum J_i(\bar{X}_i - \bar{X}_T)(\bar{Y}_i - \bar{Y}_T)}{I - 1}$$

$$S_T = \frac{\sum (\bar{X}_{ij} - \bar{X}_T)(\bar{Y}_{ij} - \bar{Y}_T)}{N - 1}$$

donde:

X_{ij} o Y_{ij} : valor genotípico del J-ésimo individuo en la i-ésima región para los sistemas C,c o E,e respectivamente.

\bar{X}_i o \bar{Y}_i : valor genotípico promedio de la i-ésima región para los sistemas C,c o E,e respectivamente.

N : número total de individuos de la población.

J_i : número de individuos en la i-ésima región.

I : número de regiones.

\bar{X}_T o \bar{Y}_T : valor genotípico promedio de la población total para los sistemas C,c o E,e respectivamente.

La decisión sobre la igualdad de S_w , S_A y S_T a cero, se realizó mediante la prueba de hipótesis $H_0: r = 0$.

Resultados

Las frecuencias fenotípicas encontradas para cada una de las generaciones del sistema bivariado CcEe en las regiones y en la población total estudiada, se alejan en forma altamente significativa de las frecuencias fenotípicas esperadas (tablas 2 y 3; X_C y X_T); sin embargo, existe homogeneidad entre las generaciones dentro de cada región y en la población total (X_{HG}), a excepción de las regiones 7 y 9 las que mostraron heterogeneidad entre padres e hijos. De otro lado, las regiones dentro de cada generación son altamente heterogéneas (X_{HR}).

Todos los valores de D_{AB} mostraron ser negativos y diferentes de cero (tabla 4) y los valores de máxima intensidad relativa de D_{AB} (D') para las dos generaciones son próximos a uno, lo cual indica que los valores de desequilibrio para cada región se encontraron próximos al máximo valor posible de desequilibrio (D).

Los valores de covarianza dentro de las regiones (S_w), entre regiones (S_A) y en la población total (S_T), son diferentes de cero en forma altamente significativa (tabla 5).

Discusión

La ausencia del fenotipo CCEE, así como las bajas frecuencias de los CCEe y CcEE, pueden ser explicadas por la baja frecuencia de la gameta CE en la población colombiana. Estas bajas frecuencias serán mantenidas en la población debido al ligamiento íntimo existente entre los loci del sistema sanguíneo Rh.

Es posible que estén actuando o hayan actuado factores selectivos sobre la gameta CE, acción que presentaría, además, un efecto acumulativo de dosis sobre los cigotos y/o individuos que presenten más números de alelos C y E es decir que la presencia en los cigotos o individuos de la gameta CE acompañada de otra gameta que presente los alelos C y/o E, hace a éstos más susceptibles al efecto de la selección, que a los cigotos o individuos en donde la gameta CE esté acompañada por la gameta ce.

Tabla 2. Frecuencias fenotípicas bivariadas para el complejo CcEe, por regiones, para padres.

REGION	CCEE	CCEe	CCee	CcEE	CcEe	Ccee	ccEE	ccEe	ccee	XT	XC
1	0,00000	0,00456	0,23354	0,00101	0,24215	0,25937	0,06130	0,13070	0,06738	774,56 **	780,57 **
2	0,00000	0,00274	0,22279	0,00137	0,26352	0,22861	0,09309	0,12663	0,05613	787,87 **	740,69 **
3	0,00000	0,00279	0,19359	0,00000	0,18245	0,32173	0,04596	0,12813	0,12535	178,01 **	171,97 **
4	0,00000	0,00000	0,19937	0,00158	0,18671	0,33386	0,03797	0,12184	0,11867	143,91 **	146,01 **
5	0,00000	0,00000	0,26103	0,00000	0,21573	0,27175	0,04768	0,12038	0,08343	310,36 **	301,55 **
6	0,00000	0,00289	0,18497	0,00000	0,23699	0,31214	0,06647	0,12717	0,06936	123,75 **	126,66 **
7	0,00372	0,00000	0,24907	0,00743	0,28253	0,26394	0,04833	0,08665	0,04833	100,27 **	106,52 **
8	0,00000	0,00268	0,21716	0,00266	0,18499	0,28150	0,03753	0,10992	0,16354	81,07 **	73,72 **
9	0,00000	0,02083	0,18750	0,00000	0,18750	0,27083	0,01042	0,17708	0,14583	16,50 **	12,44 *
10	0,00000	0,00742	0,15579	0,00297	0,15579	0,30861	0,04748	0,16172	0,16024	126,51 **	118,38 **
$X_{HR} = 334,719$											
Total	0,00013	0,00325	0,21850	0,00135	0,22256	0,27553	0,05919	0,12869	0,09076	2526,79 **	2460 **

** Diferencias altamente significativas

* Diferencias significativas

NS Diferencias no significativas

Tabla 3. Frecuencias fenotípicas bivariadas para el complejo CcEe, por regiones, para hijos.

REGION	CCEE	CCEe	CCee	CcEE	CcEe	Ccee	ccEE	ccEe	ccee	X _T	X _C	X _{HG}
1	0,00000	0,00089	0,24246	0,00178	0,24086	0,26405	0,05531	0,13024	0,06423	437,31 **	446,57 **	4,16 NS
2	0,00000	0,00432	0,23199	0,00144	0,28098	0,19020	0,08646	0,14553	0,05908	360,22 **	351,30 **	5,71 NS
3	0,00000	0,00000	0,21348	0,00000	0,20225	0,30056	0,04494	0,11517	0,12360	95,88 **	93,63 **	2,63 NS
4	0,00000	0,00000	0,16781	0,00342	0,21575	0,33562	0,02740	0,12329	0,12671	49,61 **	51,95 **	2,98 NS
5	0,00000	0,00256	0,24615	0,00000	0,23590	0,28718	0,03077	0,12564	0,07179	114,08 **	123,84 **	5,41 NS
6	0,00000	0,00000	0,19231	0,00000	0,23718	0,29485	0,07051	0,11538	0,08974	58,09 **	56,15 **	1,32 NS
7	0,00000	0,01250	0,28125	0,00000	0,23125	0,20000	0,05625	0,13750	0,08125	70,91 **	63,65 **	11,67 **
8	0,00000	0,00588	0,18235	0,00000	0,15294	0,33527	0,04118	0,12363	0,15882	34,55 **	30,84 **	3,53 NS
9	0,00000	0,02439	0,09756	0,00000	0,21951	0,46341	0,04873	0,12195	0,02439	15,95 **	13,95 *	11,20 **
10	0,00000	0,00285	0,17379	0,00285	0,14815	0,26496	0,03419	0,18803	0,18519	71,22 **	59,39 **	5,61 NS
$X_{HR} = 206,332$												
Total	0,00013	0,00268	0,22112	0,00134	0,22862	0,26588	0,05333	0,13535	0,09166	1211,76 **	1220,00 **	4,45 NS

** Diferencias altamente significativas

* Diferencias significativas

NS Diferencias no significativas

Tabla 4. Desequilibrio de ligamiento (D) e intensidad relativa de D (D') para el complejo CcEe, por regiones y generaciones.

REGION	P A D R E S				H I J O S			
	D	D'	T _{r2}	r	D _{AB}	D'	T _{r2}	r
1	-0,12164	0,99028	621,73 **	-0,55 **	-0,12419	1,02823	375,90 **	-0,56 **
2	-0,13914	1,00188	549,60 **	-0,61 **	-0,13125	0,91532	226,86 **	-0,60 **
3	-0,08395	0,92379	126,62 **	-0,44 **	-0,07377	0,77915	48,02 **	-0,44 **
4	-0,08159	0,91421	108,38 **	-0,43 **	-0,06212	0,69649	28,48 **	-0,37 **
5	-0,10482	0,96254	217,94 **	-0,53 **	-0,10366	0,95460	100,11 **	-0,51 **
6	-0,12261	1,06063	111,61 **	-0,53 **	-0,09921	0,87703	33,27 **	-0,51 **
7	-0,10957	0,83049	69,32 **	-0,53 **	-0,10825	0,86090	40,36	-0,54 **
8	-0,04941	0,57521	23,96 **	-0,36 **	-0,06930	0,87894	22,31 **	-0,38 **
9	-0,06688	0,75261	10,76 **	-0,34 **	-0,18179	1,69299	30,61 **	-0,49 **
10	-0,08007	0,94768	107,75 **	-0,39 **	-0,07685	0,92954	49,34 **	-0,38 **
Total	-0,10868	0,28495	1713,43 **	-0,50 **	-0,10663	0,27820	825,83 **	-0,50 **

** Diferencias altamente significativas

* Diferencias significativas

NS Diferencias no significativas

Tabla 5. Coeficientes de covarianza para la población colombiana.

COVARIANZA	PADRES	HIJOS
S _w	-0,0546	-0,0535
S _A	0,2863	0,1241
S _T	-0,0541	-0,0531
r _w	-0,4692 **	-0,4755 **
r _A	-0,4717 **	-0,4786 **
r _T	-0,5043 **	-0,5011 **

** Diferencias altamente significativas

* Diferencias significativas

NS Diferencias no significativas

Los valores negativos de D son debidos a la baja asociación de algunas de las gametas en estado de acoplamiento (7, 8); en el caso de la población colombiana, se encontró baja asociación de la gameta CE, como se señaló anteriormente con la baja frecuencia de la gameta CE.

Los valores calculados de DAB en la población colombiana, mostraron la presencia de un alto

grado de desequilibrio en cada una de las generaciones. El valor de D calculado en una población de individuos adultos es considerablemente mayor al valor estimado de D en una población en estado cigótico (2); contrario a esto, en las regiones 8 y 9, el valor estimado de D_{AB} es considerablemente menor en la generación paterna. Esta diferencia en el valor de desequilibrio entre las dos generaciones de una misma región puede ser debida al momento en que actúa la selección y al efecto que ésta ocasiona sobre el desequilibrio.

La presencia de valores negativos del coeficiente de desequilibrio para el sistema CcEe en todas las regiones, es ocasionada por la baja frecuencia de la gameta CE en la población colombiana.

La población colombiana está conformada por regiones genéticas heterogéneas para el sistema CcEe; sin embargo, existe la tendencia general a disminuir el desequilibrio para el sistema CcEe.

Los valores de r_T y r_w son similares a los encontrados en los indígenas Yanomama del Brasil (5),

e indican el efecto del ligamiento de los loci C y E sobre el desequilibrio. De otro lado, el valor de r_A encontrado en este trabajo, muestra el efecto que tiene la subdivisión de la población colombiana en regiones altamente heterogéneas sobre el desequilibrio.

El desequilibrio encontrado en la población colombiana para el sistema CcEe es ocasionado por el ligamiento íntimo entre los loci C y E, por el efecto de la selección y por la existencia en la población colombiana de subpoblaciones genéticamente heterogéneas.

Referencias

1. **Ohta T.** Linkage disequilibrium with the island model. *Genet* 1982; 101:139.
2. **Cadien JD, Hassett MJ, Frielaender JF.** Linkage disequilibrium in Bougainville Island. *Amer J Phys Antrop* 1978; 48:297.
3. **Nei M, Li W.** Linkage disequilibrium in subdivided populations. *Genetics* 1973; 75:213.
4. **Sokal.** *Biometría. Principios y métodos básicos en la investigación biológica.* H. Blume Ediciones 1979.
5. **Smouse P, Neel J.** Multivariate analysis of gametic disequilibrium in the Yanomama. *Genetics* 1977; 85:733.
6. **Degos L.** La repartition antropologique des genes HLA et la dynamique des populations. In: Daisset J, Pla M. 1985.
7. **Thomson G.** The effect of selected locus on linked neutral loci. *Genetics* 1977; 85:753.
8. **Dausset J, Pla M.** HLA le complexe a jeur d'histocompatibilite de l'homme. 1985.