

# DENDOGRAMA BASADO EN LA FRECUENCIA DE APELLIDOS COMO INDICADOR DE AISLAMIENTO Y MIGRACIÓN EN EL ESTADO GUÁRICO, VENEZUELA

Álvaro Rodríguez Larralde

*Laboratorio de Genética Humana, Instituto Venezolano de Investigaciones Científicas,  
Caracas, Venezuela.*

## INTRODUCCIÓN

En las sociedades donde los apellidos son transmitidos de padres a hijos, aquéllos pueden ser comparados, en líneas muy generales, a genes del cromosoma Y. Yasuda *et al.* (1974), Zei *et al.* (1983) y Barrai *et al.* (1987), con datos del valle de Parma en Italia los primeros, de matrimonios consanguíneos en Cerdeña los segundos y de pueblos de diferentes tamaños de Japón e Italia los últimos, han demostrado que la distribución de los apellidos se ajusta a la de alelos neutros bajo efecto de las fuerzas de migración y fluctuación génica. En 1965, Crow y Mange publican su trabajo pionero sobre el uso de los apellidos para la estimación de la consanguinidad a través de la isonimia (apareamiento de individuos con apellidos iguales); desde de ese momento comienzan a realizarse numerosos estudios sobre estimación de estructura genética de poblaciones a partir de los apellidos presentes en ellas (Lasker 1985, para una revisión detallada del tema).

El uso de apellidos en este tipo de estudio tiene la ventaja sobre el de marcadores génicos de que se pueden obtener muestras de gran tamaño en poco tiempo y a costos relativamente bajos, tanto de poblaciones actuales como de las que ya no existen, con el fin de estudiar los cambios de la estructura genética en el tiempo. Por otra parte, hay que tener en cuenta sus limitaciones, ya que los apellidos

no son variables biológicas y pueden estar influidos por factores culturales.

Lasker (1977) propone el coeficiente  $R_i$  para estimar la relación entre poblaciones, basado en la frecuencia de los apellidos existentes en las mismas, siendo muy utilizado en el estudio de relaciones entre poblaciones europeas (Fuster 1986, Kaplan *et al.* 1978, Kuchemann *et al.* 1979, Lasker 1978, Lasker y Mascie-Taylor 1983, entre muchos otros) y poblaciones latinoamericanas (Lasker 1977, Rodríguez-Larralde 1989, Pinto-Cisternas *et al.* 1990a y 1990b).

En 1967, Cavalli-Sforza y Edwards proponen la distancia euclídeana como medida de distancia genética, siendo una función del ángulo entre los puntos cuyas coordenadas están dadas por las frecuencias génicas de las respectivas poblaciones. Si consideramos a los apellidos como alelos múltiples de un *locus* determinado, podemos utilizar sus frecuencias para estimar esta distancia, y medir así la relación entre poblaciones e inferir movimientos migratorios entre ellas. En el presente trabajo estimamos la distancia euclídeana entre los 32 municipios del estado Guárico, Venezuela, comparándolas con las distancias geográficas correspondientes infiriendo los movimientos migratorios dentro del estado.

## MATERIAL Y MÉTODOS

La distribución de los apellidos en 32 municipios del estado Guárico se obtuvo a partir del registro electoral venezolano, actualizado para 1991. Sólo se utilizó información sobre el primer apellido para que los resultados fuesen comparables con otros publicados por diversos autores, correspondientes a poblaciones no hispánicas. Con el fin de reducir el efecto que las migraciones recientes a las ciudades (Ducoff 1965) pueda tener sobre la distribución de los apellidos, sólo se incluyeron en el análisis aquellos individuos, sin distinción de sexo, nacidos antes de 1951; lo que garantiza una muestra formada por los mayores de 40 años para el momento de la actualización del registro, población que en líneas generales muestra más estabilidad que una constituida por sujetos más jóvenes.

El estado Guárico ocupa el centro geográfico de Venezuela (Figura 1), con una superficie de 66 400 km<sup>2</sup> y una población de

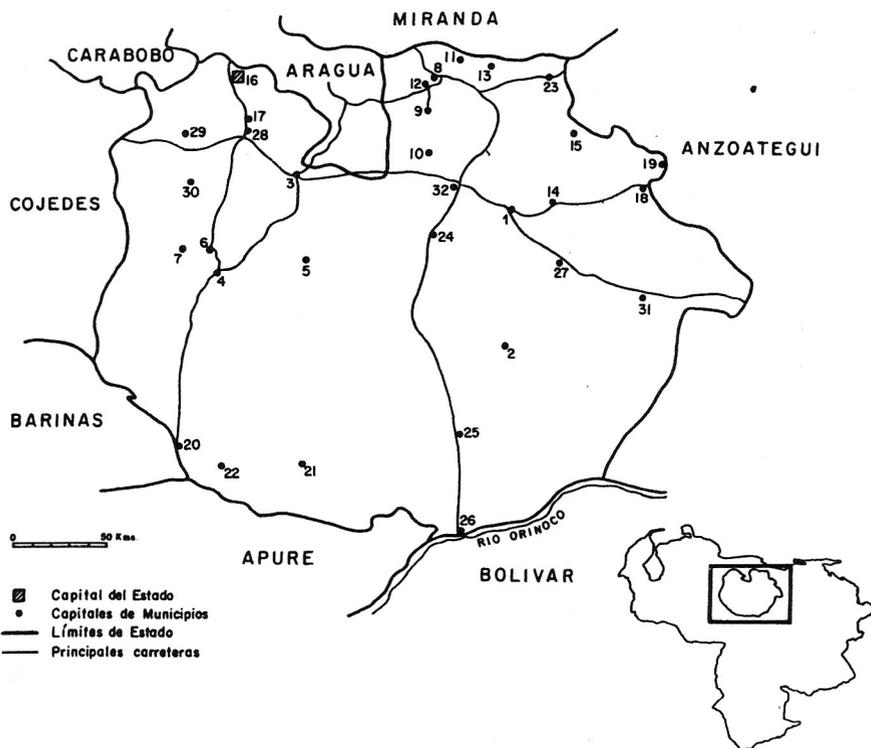


Figura 1. Ubicación geográfica del estado Guárico dentro de Venezuela y de la capital de los municipios dentro del estado. La identificación de los municipios está dada en el cuadro 1.

488 623 habitantes para el censo de 1990 (Oficina Central de Estadística e Informática 1991). Está cubierto por grandes llanuras y sólo en su porción nororiental se encuentran estribaciones montañosas de escasa altura de la Cadena del Interior de la Cordillera de La Costa. Su mayor riqueza es la agricultura y la ganadería. Está atravesado por dos importantes vías de comunicación: una carretera con dirección norte a sur, situada hacia el occidente del estado, que comunica una porción importante del sur del país con Caracas y los puertos en las costas del mar Caribe; otra con dirección este a oeste, por donde circula gran parte del transporte entre el oriente y el occidente de Venezuela. El estado tiene 32 municipios, cuyas capitales se encuentran identificadas con números arábigos en la figura 1. En

*Cuadro 1*  
 Capitales de los municipios del estado Guárico y los distritos a los que pertenecen

Capital	Distrito	i.d.*	n	n.a.
Valle de la Pascua	Infante	1	16 657	1 218
Espino	Infante	2	1 194	200
El Sombrero	Mellado	3	4 043	664
Calabozo	Miranda	4	17 044	1 552
El Calvario	Miranda	5	386	142
El Rastro	Miranda	6	505	202
Guardatinajas	Miranda	7	900	249
Altagracia de Orituco	Monagas	8	8 921	910
Lezama	Monagas	9	891	289
Libertad de Orituco	Monagas	10	338	151
San Francisco de Macaira	Monagas	11	717	208
San Rafael de Orituco	Monagas	12	727	211
Sabana Grande	Monagas	13	810	244
Tucupido	Rivas	14	6 557	607
San Rafael de Laya	Rivas	15	731	282
San Juan de los Morros	Roscio	16	15 757	1 735
Parapara	Roscio	17	828	257
Zaraza	Zaraza	18	9 417	764
San José de Unare	Zaraza	19	669	218
Camaguán	Camaguán	20	2 952	430
Cazorla	Camaguán	21	1 262	218
Guáyabal	Camaguán	22	1 984	294
San José de Guaribe	San José de Guaribe	23	1 823	388
Las Mercedes	Las Mercedes	24	2 890	503
Santa Rita	Las Mercedes	25	305	92
Cabruta	Las Mercedes	26	1 219	295
El Socorro	El Socorro	27	2 664	301
Ortiz	Ortiz	28	1 159	345
San Francisco de Tiznados	Ortiz	29	1 079	275
San José de Tiznados	Ortiz	30	760	235
Santa María de Ipire	Santa María de Ipire	31	2.547	323
Chaguaramas	Chaguaramas	32	1.465	336

\*i.d. = identificación de los municipios utilizada en el texto y en las figuras.

n= tamaño muestral.

n.a.= número de apellidos diferentes encontrados en el análisis.

el cuadro 1 aparecen los nombres de estas capitales y los números con que se identifican, además de los tamaños de las muestras y el número de apellidos diferentes utilizados en cada municipio para este análisis.

La distancia euclideana entre todos los pares de municipios se estimó de acuerdo con la siguiente fórmula:

$$D = \sqrt{1 - \cos \Theta} \quad , \quad \text{donde} \quad \cos \Theta = \sum \sqrt{p_{i1} p_{i2}} \quad ,$$

siendo  $p_{i1}$  y  $p_{i2}$  las frecuencias del apellido  $i$  en los municipios 1 y 2, respectivamente; la sumatoria es sobre todos los apellidos considerados en ambos municipios.

Para estimar la distancia geográfica entre pares de municipios, se tomó la capital de cada uno como punto de referencia y la distancia entre ellos se calculó utilizando la fórmula de Pitágoras:

$$d(km) = 110 \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

donde  $x_1$  y  $y_1$ , y  $x_2$  y  $y_2$  son las longitudes y latitudes de las capitales 1 y 2, respectivamente. El factor 110 representa el número de km que contiene un grado geográfico a  $10^\circ$  de latitud. Se utilizó la menor distancia entre dos puntos en lugar de la distancia real que los comunica a través de carreteras, por varias razones: *a)* la facilidad de su cálculo; *b)* se ha demostrado alta correlación entre ambas distancias en otros estudios (Crumpacker *et al.* 1976), y *c)* las llanuras del estado Guárico permiten que la mayoría de las carreteras sean rectas.

Se estimó la correlación de Pearson entre la matriz de distancias euclideanas y la de distancias geográficas, y su significación se estudió por el método recomendado por Smouse *et al.* (1986). A partir de la matriz de distancias euclideanas se construyó el dendrograma correspondiente, mediante el método de distancias promedio (Pielou 1984), representando gráficamente la relación entre los diferentes municipios (Figura 2).

Se identificaron los apellidos con distribución localizada dentro del estado, considerando apellido localizado aquellos con frecuencia mayor de 3% en no más de un municipio, y frecuencias menores de 0.5% en al menos 28 de los municipios restantes.

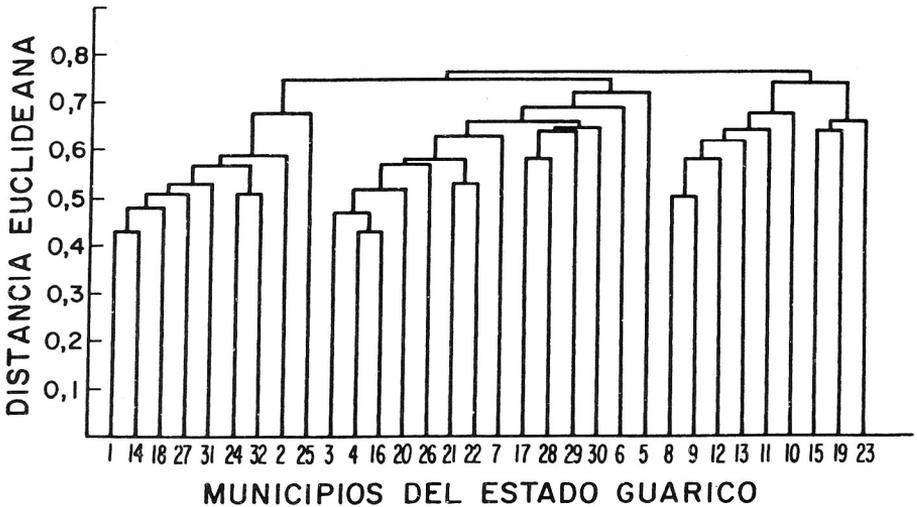


Figura 2. Dendrograma del estado Guárico elaborado con base en la distancia euclídeana entre municipios, estimada con la frecuencia de los apellidos presentes en ellos. La identificación de los municipios está dada en el cuadro 1.

## RESULTADOS

El número total de individuos estudiados dentro del estado fue de 109 180, que representa 22% de la población del mismo. El número total de apellidos diferentes identificado en estos individuos fue de 4 470. Los municipios más poblados fueron Valle de La Pascua (identificado con el número 1), Calabozo (número 4) y San Juan de los Morros (número 16), la capital del estado. Son ellos también los que presentaron mayor número de apellidos diferentes. Los municipios más pequeños fueron El Calvario (número 5), con 386 individuos y 142 apellidos diferentes; Libertad de Orituco (número 10), con 338 individuos y 151 apellidos, y Santa Rita (número 25), con 305 individuos y 92 apellidos diferentes.

La distancia geográfica promedio entre todos los posibles pares de municipios fue de 138 km, con un valor máximo de 300 km, entre los municipios 19 y 20, y un valor mínimo de 11 km, entre los municipios 17 y 28. La distancia euclídeana promedio entre los 496

posibles pares de municipios fue de 0.6646 unidades, con un valor máximo de 0.8344, entre los municipios 15 y 25, y un valor mínimo de 0.4261 observado entre los municipios 4 y 16. El municipio más aislado resultó ser el 25, con una distancia euclidea promedio, con respecto a todos los demás, de 0.7611 unidades (con desviación típica de 0.0448), seguido del número 5, con una distancia promedio de 0.7270 unidades (d.t.= 0.0457). Los que resultaron en promedio más relacionados con el resto fueron los municipios 1 y 16, con distancias euclideas promedios de 0.5825 y 0.5864 unidades, respectivamente (desviaciones típicas de 0.0793 y 0.0715).

La correlación entre la matriz de distancias geográficas y la de distancias euclideas fue de 0.40, altamente significativa, indicando que existe aislamiento por distancia y que aquellos municipios más cercanos geográficamente tienden a compartir más apellidos que los más distantes.

El dendrograma elaborado con la distancia euclidea entre municipios muestra tres grupos bien diferenciados (Figura 2): el primero está formado por nueve municipios localizados hacia el sureste del estado, siete de los cuales están sobre carreteras importantes que se dirigen hacia el oriente del país. Los dos últimos municipios que se incorporan a este grupo son el 2 y el 25, que, como se observa en la figura 1, son los más aislados de este conglomerado y no están sobre las carreteras mencionadas.

El segundo grupo de municipios está integrado por catorce, situados de norte a sur, al occidente del estado, a excepción del municipio 26 que se encuentra en el extremo sur, hacia el centro (Figura 1). Dentro de este grupo se observa un subgrupo formado por los municipios más poblados de esta zona (3, 4 y 16), al que se van incorporando los municipios del sur: 20, 26, 21 y 22. Otro subgrupo está conformado por los municipios del norte: 17, 28, 29 y 30. El último municipio que se incorpora a este grupo es el 5, bastante aislado del resto, como se observa en la figura 1.

El tercer grupo está formado por nueve municipios situados hacia la zona montañosa del noreste del estado. Consta de dos subgrupos, uno constituido por los seis municipios más occidentales del conglomerado, donde el municipio 10, el más aislado de esta zona, es el último en incorporarse al subgrupo, y el otro formado por los tres municipios más orientales.

Se identificaron cinco apellidos: Aquino y Cadenas, ubicados en el municipio 5, con frecuencias de 4.7 y 3.4%, respectivamente, y Delgado, Guarán y Salcedo, localizados en el municipio 25, con frecuencias de 3.3, 4.9 y 5.2%, respectivamente.

## DISCUSIÓN

Relethford (1982) manifiesta que una correlación positiva y significativa entre distancia genética y distancia geográfica indica la existencia de un efecto de la distancia geográfica sobre la variación genética entre las poblaciones. Por otra parte, Jorde (1989) declara que la tendencia a observar correlaciones bajas entre distancia geográfica y aquella estimada a través de apellidos, en poblaciones recientemente fundadas, mientras se observan valores mayores en grupos bien establecidos, refleja que existe una acumulación en el tiempo del efecto del aislamiento por distancia. En Guárico la correlación entre estas distancias fue de 0.40 ( $p < 0.001$ ), indicando que, en general, los municipios más cercanos comparten más apellidos en común que aquellos más distantes geográficamente, lo que se refleja en menores distancias euclidianas estimadas a través de los apellidos, en el primer caso. Existe una variación en los apellidos asociada con la distancia, de donde podríamos inferir una variación genética. Si comparamos el valor de la correlación obtenida en Guárico con la obtenida en otros estados de Venezuela, utilizando la misma metodología y fuente de información, vemos que aquélla es sólo ligeramente mayor a la observada en Falcón (0.39), y menor a la observada en Yaracuy (0.44), Lara (0.49), Nueva Esparta (0.49) y Mérida (0.65) (Rodríguez-Larralde 1993, Rodríguez-Larralde *et al.* 1993). La alta correlación observada en Mérida posiblemente se deba a las barreras geográficas provocadas por la cordillera de Los Andes, que lo atraviesa. En Nueva Esparta, un estado insular, además de las barreras geográficas existentes, podría haber barreras culturales, como se discute en otra publicación (Rodríguez-Larralde *et al.* 1993). En Guárico, el hecho de estar cubierto por grandes llanuras y la ausencia hasta hace pocas décadas de buenas vías de comunicación hacen que la principal barrera geográfica sea la distancia; sin embargo, parece que ha habido menos aislamiento dentro de este

estado que lo observado en Lara, Nueva Esparta y Mérida, posiblemente, entre otras causas, por ser vía de paso importante entre el oriente y el occidente, y entre el norte y el sur del país.

En el dendrograma de la figura 2 se destacan tres zonas migratorias dentro del estado: la occidental, la suroriental y la nororiental. Los municipios de esta última, ubicada en la zona montañosa de Guárico, son los que se incorporan de último al dendrograma, a una distancia euclídeana de 0.765 unidades, indicando que son los más aislados del estado.

En el grupo de municipios de la zona occidental, los que aparecen más relacionados son 3, 4 y 16, que necesariamente no son los más cercanos geográficamente, pero sí los más poblados de la zona; centros de atracción de población de todo el estado, lo que explica la estrecha relación entre ellos. En este mismo grupo llama la atención la presencia del municipio 26, que geográficamente está más cercano a los municipios del primer grupo, al suroriente de Guárico. Nuestra interpretación es que 26 se relaciona con los municipios suroccidentales a través de movimientos migratorios por los ríos Orinoco y Apure, ya que la carretera pavimentada a Cabruta (número 26) es reciente, posterior a 1958, y la principal vía de comunicación de este municipio era a través del Orinoco. Camaguán, Cazorla y Guayabal (números 20, 21 y 22, respectivamente), todos al suroccidente del estado, se encuentran relativamente cerca del río Apure, afluente del Orinoco.

Tanto en el grupo de municipios occidentales como en el de los surorientales, los municipios más aislados geográficamente, ya sea por la distancia a la que se encuentran en relación con el resto del conglomerado o por estar alejados de las vías principales de comunicación, son los últimos en incorporarse al grupo. Así tenemos al municipio 5 del grupo occidental y a los municipios 2 y 25 del grupo suroriental, indicándonos la menor migración que ha habido entre ellos y el resto de los municipios del grupo correspondiente. Los municipios 5 y 25 son también los que muestran mayor distancia euclídeana promedio con el resto de municipios, de 0.7270 y 0.7611, respectivamente, confirmando nuevamente su relativo aislamiento. Es de hacer notar que éstos son municipios con poca población y sólo en ellos aparecen localizados apellidos con frecuencias altas (superiores a 3%), que en el resto del estado tienen frecuencias muy bajas, sugi-

riendo pocos movimientos migratorios que los involucren. Lo contrario sucede con los municipios 1 y 16, que muestran las menores distancias euclidianas promedio al compararlos con los otros municipios: incluyen las principales ciudades del estado, centros de atracción de población de diferentes partes de Guárico y del país en general, lo que hace que compartan muchos apellidos con los otros municipios.

En resumen, la distribución de los apellidos en el estado Guárico ha permitido estudiar el aislamiento y la migración dentro del mismo. Se establecieron dentro del estado tres zonas con mayor intercambio migratorio: la nororiental, la más aislada, la occidental y la suroriental. Dentro de esta última está situado Cabruta, municipio que resultó estar más relacionado con los suroccidentales que con los orientales, posiblemente por los movimientos migratorios a través de los ríos Orinoco y Apure. La identificación de apellidos con frecuencia mayor de 3% en los municipios El Calvario (número 5) y Santa Rita (número 25), y prácticamente ausentes en los otros municipios del estado, confirma el aislamiento relativo de los mismos.

### *Agradecimientos*

Se agradece al Consejo Supremo Electoral habernos facilitado la copia del registro electoral para fines de investigación. A Jesús Quiroz, del Servicio de Computación del Instituto Venezolano de Investigaciones Científicas, por ayudarnos con parte del procesamiento de los datos; a Sergio Arias C., por la revisión y comentarios al manuscrito.

### RESUMEN

Los apellidos han sido utilizados en genética de poblaciones humanas para obtener un estimado rápido y económico de la estructura genética de las poblaciones y de la relación entre ellas. En este trabajo se presenta un dendrograma elaborado con base en la distancia euclidiana estimada con la frecuencia de los apellidos de los 32 municipios del estado Guárico, Venezuela, tomados de dos en dos. La fuente de información usada fue el registro electoral actualizado en 1991. La muestra consistió en 109 180 sujetos que compartieron 4 470 apellidos diferentes. La correlación entre las transformaciones logarítmicas de las distancias euclidiana y geográfica fue de 0.40, altamente significativa, indicando que existe aislamiento por distancia

y que aquellos municipios más cercanos geográficamente comparten más apellidos que aquellos más distantes. El dendrograma muestra tres grupos de municipios bien diferenciados: el primero comprende nueve municipios situados hacia el sureste de Guárico, el segundo está formado por 14 municipios situados al oeste del estado y el último grupo comprende nueve municipios situados al noreste. Los dos primeros grupos se unen a una distancia euclidiana de 0.75 unidades, y el último se une con estos dos a una distancia de 0.76 unidades. Los resultados confirman que los apellidos pueden utilizarse en estudios de migración y aislamiento en poblaciones humanas.

PALABRAS CLAVE: isonimia, migración, apellidos, Venezuela.

### ABSTRACT

Family names have been applied as a tool in human population genetics, for obtaining a fast and economic estimate of the genetic structure of human groups, and the relations among their members. This paper presents a dendrogram with the euclidian distances, estimated from the frequency of the family names, taken two by two, in 32 *municipios* in the Venezuelan state of Guárico. Our data base was the electoral registry, updated in 1991, which included 109 180 persons with 4 470 different family names. The highly significant correlation between the logarithmic transformations of the euclidian and geographical distances was 0.40. This indicates that there is isolation by distance, and that those *municipios* that are geographically close, share more family names than the most distant ones. The dendrogram shows three well differentiated groups of *municipios*. The first is made up of 9 *municipios*, Southwest of Guárico; the second of 14, in the West; and the last of 14 to the Northeast of the capital city. The first two join at an euclidian distance of 0.75 units, and are joined by the third one at a distance of 0.76 units. Our results confirm the potencial use of family names in migration and isolation estimates of human groups.

### REFERENCIAS

BARRAI, I. *ET AL.*

1987 Surnames in Ferrara: distribution, isonymy and levels of inbreeding, *Annals of Human Biology*, 14: 415-423.

CAVALLI-SFORZA, L.L. Y A.W.F. EDWARDS

1967 Phylogenetic analysis models and estimation procedures, *American Journal of Human Genetics*, 19: 233-257.

CROW, J.F. Y A.P. MANGE

- 1965 Measurement of inbreeding from the frequency of marriages between persons of the same surname, *Eugenics Quarterly*, 12: 199-203.

CRUMPACKER, D.W. ET AL.

- 1976 Air distance versus road distance as a geographical measure for studies on human population, *Geographical Analysis*, 8: 215-223.

DUCOFF, L.J.

- 1965 The role of migration in the demographic development of Latin America, *Millbank Memorial Fund Quarterly*, 49: 489-210.

FUSTER, V.

- 1986 Relationship by isonymy and migration pattern in Northwest Spain, *Human Biology*, 58: 391-406.

JORDE, L.B.

- 1989 Inbreeding in the Utah Mormons: an evaluation of estimates based on pedigrees, isonymy and migration matrices, *Annals of Human Genetics*, 53: 339-355.

KAPLAN, B.A. ET AL.

- 1978 Communality of surnames: a measure of biological interrelationship among thirtyone settlements in Upper Val Varaita in the Italian Alps, *American Journal of Physical Anthropology*, 49: 251-256.

KUCHEMANN, C.F. ET AL.

- 1979 Historic changes in the coefficient of relationship by isonymy among the populations of the Otmoor villages, *Human Biology*, 51: 63-77.

LASKER, G.W.

- 1977 A coefficient of relationship by isonymy: a method for estimating the genetic relationship between populations, *Human Biology*, 49: 489-493.

- 1978 Relationships among the Otmoor villages and surrounding communities as inferred from surnames contained in the current register of electors, *Annals of Human Biology*, 5: 105-111.

- 1985 *Surnames and Genetic Structure*, Cambridge University Press, Cambridge.

LASKER, G.W. Y C.G.N. MASCIE-TAYLOR

- 1983 Surnames in five English villages: relationship to each other, to surrounding areas and to England and Wales, *Journal of Biosocial Sciences*, 15: 25-34.

OFICINA CENTRAL DE ESTADÍSTICA E INFORMÁTICA

- 1991 *Censo 90, Estado Guárico*, Oficina Central de Estadística e Informática, Caracas.

## PIELOU, E.C.

- 1984 *The Interpretation of Ecological Data. A Primer on Classification and Orientation*, Wiley, New York.

## PINTO-CISTERNAS, J. ET AL.

- 1990a Comparison of two Venezuelan populations through the coefficient of relationship by isonymy ( $R_i$ ), *Human Biology*, 62: 413-419.
- 1990b Comparison of Lasker's coefficient of relationship in a Venezuelan town in two different periods, *Annals of Human Biology*, 17: 305-314.

## RELETHFORD, J.H.

- 1982 Isonymy and population structure of Irish isolates during the 1890's, *Journal of Biosocial Sciences*, 14: 241-247.

## RODRÍGUEZ-LARRALDE, A.

- 1989 Relationship between 17 Venezuelan counties through communality of surnames, *Human Biology*, 61: 31-44.
- 1993 Genetic distance estimated through surname frequencies of 37 counties from the State of Lara, Venezuela, *Journal of Biosocial Sciences*, 25: 101-110.

## RODRÍGUEZ-LARRALDE, A. ET AL.

- 1993 Isonymy structure in four Venezuelan States, *Annals of Human Biology*, 20: 131-145.

## SMOUSE, P.E. ET AL.

- 1986 Multiple regression and correlation extensions of the Mantel test of matrix correspondence, *Systematic Zoology*, 35: 627-632.

