

***APOE* polymorphism distribution among Native Americans and related populations**

DARÍO A. DEMARCHI¹, FRANCISCO M. SALZANO²,
M. EUGENIA ALTUNA¹, MARILU FIEGENBAUM², K. HILL³,
A. M. HURTADO³, LUIZA T. TSUNETTO⁴, M. L. PETZL-ERLER⁴,
& MARA H. HUTZ²

¹Museo de Antropología, Facultad de Filosofía y Humanidades, Universidad Nacional de Córdoba, Córdoba, Argentina, ²Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, ³Department of Anthropology, University of New Mexico, Albuquerque, USA, and ⁴Departamento de Genética, Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, Brazil

(Received 16 August 2004; revised 17 February 2005; accepted 23 February 2005)

Abstract

Background: Apolipoprotein E (apoE, protein; *APOE*, gene) plays a central role in lipid metabolism. Three common alleles, *E*2*, *E*3* and *E*4* have quantitative effects on lipid and lipoproteins levels, which are major risk determinants of cardiovascular diseases in several populations. Given their clinical significance, it is of interest to know the distribution of *APOE* variants in populations from diverse ethnic groups, as well as to determine if this polymorphism presents variations that might be associated with given evolutionary factors.

Aim: We report the distribution of *APOE* polymorphisms in Native American populations from South America, comparing it with other native populations of the Americas and Siberia.

Subjects and methods: The sample consisted of 315 individuals from nine Native American populations living at subtropical latitudes of Argentina, Brazil and Paraguay. The extended analysis included 50 populations across South and North America, Greenland and Siberia. The geographic patterns of the variation were investigated through correlation analysis, spatial autocorrelation and analysis molecular of variance (AMOVA).

Results: The incidence of the most common allele (*APOE*3*) in the sample analysed ranged from 0.78 to 0.98. The second allele in prevalence, *APOE*4*, varied from 0.00 to 0.17. The rare allele *APOE*2* was found in five of the nine populations investigated. This variant was found in a male with both maternal and paternal Native American lineages, suggesting that this allele is present in Native Americans and hence should not be used as an indicator of admixture. *APOE*3* and *APOE*4* present, respectively, positive and negative associations with latitude, although the pattern is much more pronounced in the Northern Hemisphere than in South America. *APOE*2* increases its frequency with latitude but this pattern is statistically significant only in South America.

Conclusion: The overall *APOE* spatial pattern seems, in general, compatible with a directional demographic expansion which occurred in north-eastern Asia and much of the New World.

The *APOE**2 allele shows this pattern in South America but a random distribution in the Northern Hemisphere, suggesting that the possibility of selection should not be discarded.

Keywords: *Native Americans, APOE, spatial variation*

Introduction

Apolipoprotein E (apoE, protein; *APOE*, gene) plays a central role in lipid metabolism as a ligand for two cell-surface lipoprotein vectors, which mediate the cellular uptake of specific lipoproteins such as intermediate-density lipoprotein and chylomicron remnants (Mahley 1988, Mahley & Rall 2000). Three common alleles, *E**2, *E**3 and *E**4, coding for three protein isoforms, E2, E3 and E4, have quantitative effects on lipid and lipoproteins levels, which are the major risk determinants of cardiovascular disease (Kamboh et al. 1991). The *APOE**2 allele is usually associated with lower total cholesterol and low density lipoprotein (LDL), but higher high density lipoprotein (HDL) levels, whereas *APOE**4 is usually linked to opposite effects, besides being associated with Alzheimer disease in several populations (Daignon et al. 1988, Hallman et al. 1991, Saunders et al. 1993, Strittmatter et al. 1993, Andrade et al. 2000a, Crutcher 2004).

In an earlier paper, Marin et al. (1997) failed to detect the *APOE**2 allele in five Brazilian Amazonian tribes. Later, Andrade et al. (2000b) found a high heterogeneity of *APOE* polymorphism distribution in six other South Native American tribes, as well as the presence, although at very low frequencies, of the *APOE**2 allele in two of these populations.

In this study we investigated the distribution of *APOE* polymorphisms in nine Native American populations of Argentina, Brazil and Paraguay, all living at subtropical latitudes of South America. We also included data reported in the literature for other native populations from South America, North America, Greenland and Siberia, to verify the presence of a spatial structure that could suggest past migrations or a pattern suggestive of selection or any other evolutionary factor.

Subjects and methods

Populations

The sample consisted of 315 individuals from nine Native American populations living at subtropical latitudes of South America. Five populations are from the Argentinean part of the Gran Chaco region—Wichí from Formosa and from Chaco provinces, Toba from Formosa and Chaco provinces, and Pilagá. The other Argentinean population is the Mbyá-Guaraní, living in the subtropical forest of the province of Misiones. The Wichí and Toba of Formosa and Chaco were not pooled for analysis because, although sharing language and culture, each population constitutes an independent mating system, isolated by hundreds of kilometres of thorny forest. As for the three other populations, one is from southern (Kaingang), and another from central Brazil (Kaiowá/Ñandeva-Guaraní); while the Aché or Guayaki are from Paraguay. Sample sizes and language affiliations are presented in Table I.

Appropriate informed consent was obtained from all individuals whose data are presented here for the first time. In relation to the Brazilian Native samples these studies had been approved by the Brazilian Ethics National Committee (Resolution 123/98).

We also included in the study data reported in the literature for other native populations from South America, North America, Greenland and Siberia (Table II), to analyse the pattern and extent of genetic variation of the *APOE* polymorphisms at a wider scale.

Table I. *APOE* genotype distribution, heterozygosity, and linguistic classification in nine Native American populations from South America.

Population	No. indiv.	Genotypes					H-W equil.* $p \pm SE$	h^\dagger	Linguistic classification‡
		E^*2/E^*3	E^*2/E^*4	E^*3/E^*3	E^*3/E^*4	E^*4/E^*4			
Wichí Formosa	34	3	0	29	2	0	1 ± 0.000	0.15	Mataco
Wichí Chaco	32	0	0	29	3	0	1 ± 0.000	0.09	Mataco
Toba Formosa	26	0	0	17	9	0	0.561 ± 0.017	0.35	Guaycurú
Toba Chaco	37	0	0	28	9	0	1 ± 0.000	0.24	Guaycurú
Pilagá	28	0	0	19	9	0	1 ± 0.000	0.32	Guaycurú
Mbyá-Guaraní	42	3	1	27	9	2	0.342 ± 0.016	0.31	Tupí-Guaraní
Kaingang	46	2	1	32	11	0	0.474 ± 0.014	0.30	Gê
Kaiowá/Nandeva-Guaraní	38	1	0	25	12	0	0.638 ± 0.011	0.34	Tupí-Guaraní
Aché	32	1	0	31	0	0	1 ± 0.000	0.03	Guayaki

*H-W equil.: test for Hardy-Weinberg equilibrium using the conventional Monte Carlo method (10 batches per analysis, 1000 permutations per batch, 10 000 total permutations). $p \pm SE$: probability $\pm SE$.

$^\dagger h$: heterozygosity (direct count).

‡ According to Loukotka (1968).

Table II. *APOE* allele frequency data.

Population	No. chromosomes	Allele frequencies			Coordinates*		Linguistic classification†	Country	Ref.‡
		<i>E</i> *3	<i>E</i> *4	<i>E</i> *2	Lat.	Long.			
Mbyá–Guarani	84	78	17	5	27.00	55.00	Tupi	Argentina	17
Kaingang	92	84	13	3	27.00	52.00	Kaingang	Brazil	17
Toba Formosa	52	83	17	0	26.20	58.20	Guaycuru	Argentina	17
Toba Chaco	74	88	12	0	26.00	60.50	Guaycuru	Argentina	17
Wichí Chaco	64	95	5	0	25.70	61.00	Mataco	Argentina	17
Pilagá	56	84	16	0	24.20	59.80	Guaycuru	Argentina	17
Wichí Formosa	68	93	4	3	24.00	62.00	Mataco	Argentina	17
Aché, Paraguay	62	98	0	2	23.30	55.30	Guayaqui	Paraguay	17
Kaiowá/Ñandeva–Guarani	76	83	16	1	23.00	54.30	Tupi	Brazil	17
Wichi Salta	86	94	2	4	22.17	62.42	Mataco	Argentina	3
Xavante	62	98	2	0	13.20	51.40	Ge	Brazil	3
Suruí	48	83	17	0	10.50	61.10	Tupi	Brazil	3
Zoró	60	70	30	0	10.20	60.20	Tupi	Brazil	3
Gavião	58	67	33	0	10.10	61.08	Tupi	Brazil	3
Kayapo	52	90	10	0	8.27	53.85	Ge	Brazil	2
Arara	42	93	7	0	4.00	53.30	Karib	Brazil	2
Wayana–Apalai	50	82	18	0	1.32	54.67	Karib	Brazil	2
Wayampi	52	58	42	0	1.00	53.00	Tupi	Brazil	2
Wai Wai	58	51	47	2	0.40	57.55	Karib	Brazil	3
Cayapa	182	72	28	0	−1.17	78.50	Chibcha	Ecuador	5
Baniwa	40	75	25	0	−1.55	68.73	Arawak	Brazil	4
Coreguaje	56	59	41	0	−2.00	76.00	Tucano	Colombia	6
Embera	50	86	14	0	−2.20	78.00	Paezan	Colombia	6
Nukak	40	63	37	0	−2.30	71.00	Maku	Colombia	6

Tule	60	91	6	3	-3.00	76.00	Cuna	Colombia	6
Yanomami	192	84	16	0	-3.00	64.00	Yanomami	Brazil	1
Yanomami	46	96	4	0	-3.00	64.00	Yanomami	Brazil	2
Yanomami	72	90	10	0	-3.00	64.00	Yanomami	Brazil	4
Wapishana	64	64	36	0	-3.07	60.03	Arawak	Brazil	4
Waunana	60	91	9	0	-4.00	77.00	Paezan	Colombia	6
Macushi	30	83	17	0	-4.00	60.50	Karib	Brazil	4
Guahibo	52	81	19	0	-5.30	68.00	Arawak	Colombia	6
Makiritare	54	89	11	0	-5.33	65.33	Karib	Venezuela	4
Yuco	60	100	0	0	-7.20	72.40	Karib	Venezuela	6
Butaregua	42	90	10	0	-9.00	72.30	Arawak	Venezuela	6
Ijka	60	81	19	0	-10.30	73.00	Chibcha	Colombia	6
Kogui	60	90	10	0	-11.00	74.00	Chibcha	Colombia	6
Mazatecan	150	90	10	0	-16.00	96.00	Mazatecan	Mexico	14
Maya	270	91	9	0	-19.00	88.00	Mayan	Mexico	15
Pima/Maricopa/Papago	3000	86	13	1	-35.00	112.30	Uto-Aztecan	Mexico	10
Seven tribes, Oklahoma	3054	83	14	3	-35.00	97.30	Uto-Aztecan	USA	10
Sioux	3028	88	11	1	-35.00	80.00	Siouan	USA	10
Buryat	250	80	17	3	-48.30	-115.00	Buriat	Mongolia	16
Evenki	248	84	15	1	-60.00	-120.00	Evenki	Russia	9
Alaska Natives	254	79	19	2	-61.00	150.00	Yuit	USA	7
Canadian Inuit	350	76	23	1	-65.00	130.00	Inuit	Canada	13
Greenland Inuit	156	77	23	0	-69.00	30.00	Inuit	Denmark	12
Greenland Inuit	196	78	21	1	-70.00	52.00	Inuit	Denmark	11
Greenland Inuit	200	79	18	3	-72.00	53.00	Inuit	Denmark	12
Greenland Inuit	266	76	23	1	-75.00	45.00	Inuit	Denmark	8

*Northern latitudes and eastern longitudes are presented as negative values.

†According to Loukotka (1968), Greenberg (1987), and Grimes (1992).

‡References: 1, Crews et al. (1993); 2, Marin et al. (1997); 3, Andrade et al. (2000b); 4, J.V. Neel (unpublished; partial data in Asakawa et al. 1985); 5, Scacchi et al. (1997); 6, Jaramillo-Correa et al. (2001); 7, Scheer et al. (1995); 8, De Knijff et al. (1992); 9, Kamboh et al. (1996); 10, Kataoka et al. (1996); 11, Boudreau et al. (1999); 12, Gerdes et al. (1996); 13, Hegele et al. (1997); 14, Gamboa et al. (2000); 15, Kamboh et al. (1991); 16, Tsunoda et al. (2002); 17, Present paper.

Laboratory methods

DNA was amplified by the PCR reaction, using the same conditions and oligonucleotide primers as those described by Maekawa et al. (1995). The amplification products were subsequently digested with *Hha*I under conditions recommended by the manufacturer. Genotypes were determined by electrophoresis of the digestion products in 8% native polyacrylamide gels (T8%, C5%), stained with ethidium bromide and exposed to UV.

Data analysis

Genotype and allele frequencies were determined by direct count. Possible departures from Hardy–Weinberg equilibrium were investigated by means of chi-square tests and the conventional Monte Carlo method (Guo & Thompson 1992). Heterogeneity among the studied populations was tested by the exact tests for population differentiation (Raymond & Rousset 1995).

The analysis of molecular variance (AMOVA) test measures the per cent of the total variance observed in the distribution of the populations that can be attributed to (1) the within-population variance, (2) the among-groups population variance and (3) the among-population within-group variance (Excoffier et al. 1992). The test was performed in this study at several hierarchical levels, using geographic and linguistic criteria for classifying groups. To detect any possible association between the *APOE* variation, latitude and longitude, total and partial Pearson's correlations were also calculated.

In addition, we subjected the allele frequencies to a spatial autocorrelation analysis, testing for positive or negative spatial association of the allele distribution with distance. We used Moran's *I* product-moment coefficient (Cliff & Ord 1973, Sokal & Oden 1978). One-dimensional correlograms were computed separately for South American populations (10 distance classes), and for the Northern Hemisphere populations (eight distance classes), calculating geographic distances between the localities as great circle distances. The plot of the autocorrelation coefficient *I* against distance is referred to as a correlogram, the overall significance of which is assessed through a Bonferroni test. A spatially random distribution results in a series of insignificant *I* values, at all distances. A decreasing set of *I* coefficients, from positive significant to negative significant, describes a cline, whereas a decreasing correlogram from positive significant to insignificant at large distances is expected for allele frequencies under isolation by distance, i.e. when genetic diversity reflects only genetic drift and short-range gene flow (Barbujani 1987).

All computations (including Pearson's correlation) were performed using the SPSS 8.0 for Windows, TFGA (Miller 1997), Arlequin 1.1 (Schneider et al. 1997) and SAAP (Wartemberg 1989) programs.

Results and discussion*New data*

The distribution of *APOE* genotypes in the nine population samples, as well as the Hardy–Weinberg test results, are presented in Table I. In all groups the genotypes observed are in agreement with those expected under Hardy–Weinberg equilibrium. Five different genotypes were found. The most common in the nine populations was *APOE**3/*APOE**3. The second genotype in prevalence is *APOE**3/*APOE**4. *APOE**2 was found associated with two different genotypes (*APOE**2/*APOE**3 and *APOE**2/*APOE**4). Two Mbyá-Guaraní showed the rare *APOE**4/*APOE**4, genotype. The Toba of Formosa and

Aché of Paraguay present, respectively, the highest and the lowest proportion of heterozygotes, whereas the Mbyá-Guaraní of northern Argentina were the only who showed all the genotypes observed in the whole sample.

The exact test for population differentiation indicates highly significant variation in the *APOE* allele distribution among the nine populations ($p = 0.0002 \pm 0.0002$).

In order to investigate if this variation was geographically or culturally patterned, the Palaeo-American speakers of the Chaco were compared with the Tropical Forest speakers of the East (following Loukotka 1968). The result did not show significant differences ($p = 0.1484 \pm 0.0158$), suggesting that the observed variation is not geographically or linguistically patterned at this level of analysis. Comparisons between cultural or linguistic groups within each of these two regions were not possible because of the small sample sizes.

Global analysis: Allele distribution

The allele frequencies of the studied populations are presented in Table II together with those available from other Native American populations as well as with other native populations from Greenland and Siberia. In agreement with Andrade et al. (2000b), we observed a highly heterogeneous prevalence of *APOE**4 in South America. For example, the allele is absent among the Aché of Paraguay, but reaches a frequency of 47% in the Wai Wai. *APOE**2 was found in five of the nine populations reported in this study. In South America, this allele was also found in the Wai Wai and in the Wichí of Salta province (Andrade et al. 2000b). Gerdes et al. (1996) suggested that *APOE**2 may have been absent in the groups who peopled the Americas in prehistoric times. However, this allele has been observed in 2% of North American Indians (Kataoka et al. 1996) and in these South Native American populations. Andrade et al. (2000b) proposed two possible explanations for these observations: (a) *APOE**2 was present in the founding Native American populations at very low frequencies, and either was lost from some groups during the process of tribalization or was not identified due to restricted sample sizes; or (b) the presence of *APOE**2 is due to admixture with non-Indians. Our results provide additional support for the first explanation, given that this allele was found in a Wichí male from Formosa which presents both maternal (mitochondrial DNA haplogroup B) and paternal (Y chromosome DYS199*T) Native American lineages (Demarchi et al. 2001, Demarchi & Mitchell 2004). It should be noted that, although this finding strongly supports this conclusion, it does not provide definitive evidence, since the *APOE**2 allele in this male could have come from a non-Native American ancestor who was simply not in the direct paternal or maternal line.

Global analysis: Geography and language

Table III presents the multivariate analysis of variance (AMOVA) performed using the data listed in Table II. As is a rule in human populations, most of the variance observed was due to the intrapopulation variability. At the first geographical level, considering all populations together (13 from Siberia, Greenland and North America, referred to as Northern Hemisphere; and 37 from South America), it can be seen that the among-groups variation is slightly smaller than the value for the among-populations/within-groups variance. Both values are low (although statistically significant), if one considers that the analysis included 50 populations living in different environments, latitudes and longitudes. This large-scale analysis is hiding different patterns of variation for each of the two major regions investigated. South America exhibits relatively high inter-population genetic

Table III. Analysis of the molecular variance for *APOE* polymorphism distribution among Native Americans and related populations using geography and language as criteria.

Region	Number of chromosomes	Number of samples	Number of groups	% Variance components*		
				Within populations	Among populations within groups	Among groups
<i>Geography</i>						
South America	2284	37	2	91.42	6.99	1.59†
Tropical Forests	1702	27	91.71	8.29	–	–
Southern Cone	582	10	97.71	2.29	–	–
Northern Hemisphere	11 422	13	2	97.75	0.35†	1.89
Mexico and USA	9502	5	99.59	0.41	–	–
Arctic populations	1920	8	99.88	0.12†	–	–
Total	13 706	50	4	96.81	1.85	1.34
<i>Language</i>						
South America	2284	37	14	91.76	5.99	2.25
Northern Hemisphere	11 422	13	8	98.75	0.14†	1.11
Total	13 706	50	22	97.28	1.85	0.87

*All values other than † are significant at the 1% level.

variation, mostly composed of the among-populations/within-groups component, over the among-groups variance. Separate analyses for each sub-region reveal that the Tropical Forest populations are substantially more variable than the Southern Cone groups. The pattern and extent of genetic variation in the Northern Hemisphere is quite different. Variation is much lower and mostly composed of the among-groups variation component.

As far as language is concerned, considering all populations together (grouped into 22 linguistic families, eight from the Northern Hemisphere and 14 from South America), the among-groups variance is less than half the value of the among-populations/within-groups variance. Both values are unexpectedly low for such a large number of samples. The differences in variation between South American and Northern Hemisphere populations are similar to those observed in the geographical analysis. That is, South America exhibits high variation, mostly due to the among-populations/within-groups variance, whereas in the Northern Hemisphere low variation is found, mostly due to the among-groups component.

Although the high inter-population genetic diversity among South American populations is well documented, the marked difference found here contrast with earlier reports. For example, O'Rourke et al. (1992), in a study based on seven blood groups, reported almost the same estimate of inter-group genetic differentiation in North America ($F_{ST}=0.0902$) and South America ($F_{ST}=0.0906$). It should be pointed out, however, that different genetic markers were used in the two studies; and the number of populations considered by us in the Northern Hemisphere was lower (predominantly composed of Uto-Aztec and Inuit speakers) than that of South America.

Global analysis: Correlations

The association between *APOE* allele frequencies, longitude and latitude was evaluated using Pearson's total and partial correlations (Table IV). As was found for the AMOVA analysis, South American and Northern Hemisphere populations present quite different patterns of variation. For the Northern Hemisphere, correlation analysis reveals an almost

Table IV. Total and partial Pearson's correlation values between the *APOE* allele distribution, latitude and longitude.

	South America			Northern Hemisphere		
	<i>APOE</i> *3	<i>APOE</i> *4	<i>APOE</i> *2	<i>APOE</i> *3	<i>APOE</i> *4	<i>APOE</i> *2
Latitude	0.220	-0.268	0.479*	0.921*	-0.916*	-0.235
Longitude	0.163	-0.173	0.113	0.095	-0.059	-0.188
Latitude (LON)†	0.271	-0.313	0.448*	0.926*	-0.925*	-0.203
Longitude (LAT)†	0.163	-0.173	0.113	-0.264	0.348	-0.145

†Partial correlations, keeping the other variable constant.

*Values significant at the 1% level.

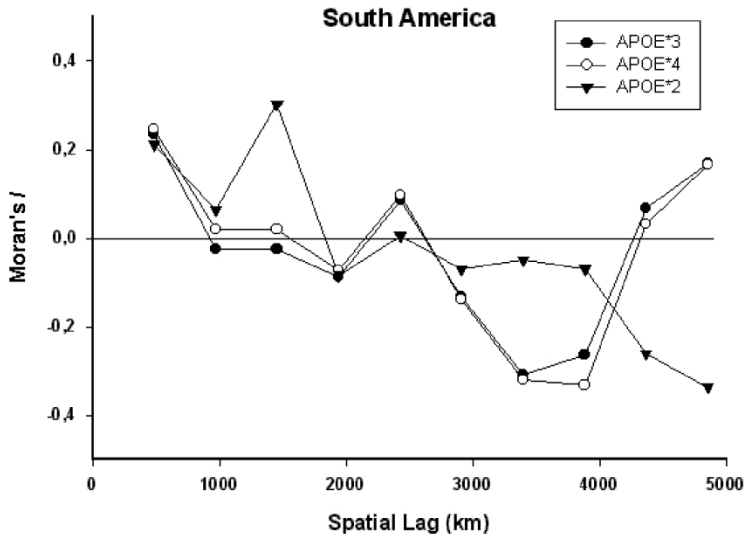


Figure 1. Pattern of spatial autocorrelation analysis observed for the *APOE* alleles in South America.

perfect association of the *APOE* distribution with latitude. *APOE**3 increases its frequency at high latitudes, whereas *APOE**4 decreases. There is no association between the *APOE**2 allele distribution and longitude.

In South American populations, on the other hand, although *APOE**3 and *APOE**4 also present, respectively, positive and negative associations with latitude, the pattern is weak and does not reach statistical significance. *APOE**2, however, markedly increases its frequency with latitude, this correlation being statistically significant. The overall picture does not change when we performed partial analyses (keeping longitude or latitude constant in each comparison), which may be interpreted as if both factors are acting independently.

Global analysis: Spatial autocorrelation

The correlogram for South America shows some geographic structure (Figure 1). As assessed through the Bonferroni test, this pattern is not significant for *APOE**3, and only marginally significant for *APOE**4 ($p = 0.041$). Both alleles show almost identical patterns,

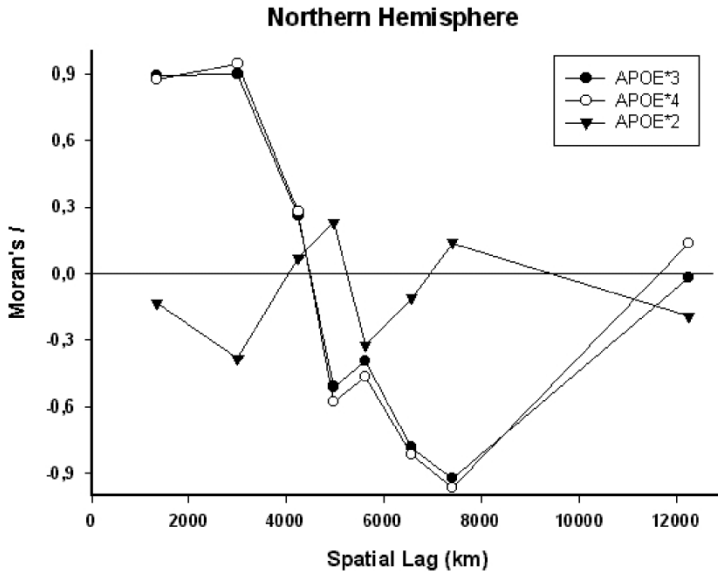


Figure 2. Pattern of spatial autocorrelation analysis observed for the *APOE* alleles in the Northern Hemisphere.

displaying positive and significant Moran *I* values at low distances (0–500 km) and negative autocorrelations at distances between 2500 and 3500 km, without displaying a monotonic decline at intermediate distances. Beyond 3500 km there is a positive but non-significant autocorrelation. *APOE*2* presents the highest overall Bonferroni test significance ($p = 0.014$) and a slightly different pattern, with significant positive autocorrelations extending up to the third spatial lag (but with a depression at the second lag). At intermediate distances the autocorrelation values tend to negative values and above 3500 km the values increase significantly. The overall pattern is clinal, with populations at the extremes of the geographical range showing the highest divergence and intermediate samples showing intermediate characteristics.

The correlogram for the Northern Hemisphere reveals a much stronger clinal pattern for *APOE*3* and *APOE*4*, with highly significant positive autocorrelations for short distance classes decreasing to highly negative values (close to -1) between 4500 and 8000 km (Figure 2). Beyond this distance the values tend to zero. Conversely, the *APOE*2* allele shows a random spatial distribution.

The overall spatial pattern just described, which entirely agrees with the correlation analysis, is compatible with a directional demographic expansion affecting the north-east of Asia and much of the New World. The *APOE*2* allele shows this pattern in South America but a random distribution in the Northern Hemisphere. On the other hand, Scacchi et al. (1997) suggested that selection due to a common genotype–environment interaction could have played an important role in favouring a higher frequency of the *APOE*4* allele in populations with a low-cholesterol diet, like Native Americans, Sub-Saharan Africans, and aborigines from Oceania (Hallman et al. 1991, Gerdes et al. 1996). The *APOE*4* carriers have a higher absorption of cholesterol at the intestinal level; thus, individuals in non-Western populations carrying this allele could have been favoured because it could be useful in re-balancing cholesterol levels, which would otherwise

be too low (Kesäniemi et al. 1987, Scacchi et al. 1997). This finding, however, would explain just the *APOE**4 distribution in the Northern Hemisphere.

The two main mechanisms commonly invoked to explain geographic gradients in gene frequencies are gene flow and natural selection (Fix 1997). Theory predicts that effects resulting from natural selection can be distinguished from migration by considering the number of loci involved. Natural selection is expected to act on a specific locus as a particular phenotype is selected by a particular environmental factor and should not affect different loci identically (Cavalli-Sforza & Bodmer 1971). In contrast, clines due to gene flow may involve multiple loci. Correct inference in human population genetics, however, is often not so simple. Several authors have attempted to summarize the patterns of genetic variation for South America, with sometimes divergent results. By constructing synthetic gene frequency maps from red cell antigen frequencies, O'Rourke and Suarez (1985) concluded that there were no geographic trends in the South American gene frequency data. In contrast, Salzano and Callegari-Jacques (1988), using multiple regression methods and an expanded set of genetic markers, found significant north-south clines for some alleles. In another paper, O'Rourke et al. (1992), investigating the genetic variation of seven blood groups by means of spatial autocorrelation analyses, failed to find any evidence of geographic structure in South America, but observed a prominent geographic structure in North and Central America. Fagundes et al. (2002), on the other hand, based on different sets of data and genetic markers (serum proteins, red cell enzymes, immunoglobulin systems, and mitochondrial DNA), obtained uneven results. Whereas they found low association between geography and blood groups plus protein systems, three of the four mitochondrial haplogroups gave evidence of spatial structure. They concluded that these differences could be explained by the fact that females seem to be more mobile than males.

Another complication in the interpretation of the spatial pattern observed for the *APOE* variants is that waves of colonizing kin-structured founder groups (as most of the South American populations are) that would replace previous occupants of the land may also produce gene gradients, mimicking demic diffusion or natural selection (Fix 2004). Furthermore, because kin-structured founder effects are highly stochastic, not all loci will be fixed and different alleles may vary in their spatial structure. In consequence, to discriminate among these alternatives, other types of information (physical, archaeological, epidemiological, demographic, social) are required (Fix 2004).

The main message of our study is that the factors influencing the genetic variability, at least in this system, may be clearly different in South America as compared to the Northern Hemisphere.

Acknowledgements

This research was supported by the Brazilian Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) through a fellowship awarded to D. A. Demarchi who was also partially funded by a Third World Academy of Sciences (TWAS) grant. Other financial sponsors were the Programa de Apoio a Núcleos de Excelência (PRONEX), and Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS). D. A. Demarchi is a Scientific Investigator Career member of CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina).

References

- Andrade FM, Larrandaburu M, Callegari-Jacques SM, Gastaldo G, Hutz MH. 2000a. Association of apolipoprotein E with plasma lipids and Alzheimer's disease in a southern Brazilian population. *Braz J Med Biol Res* 33:529–537.
- Andrade FM, Coimbra Jr CEA, Santos RV, Goicoechea A, Carnese FR, Salzano FM, Hutz MH. 2000b. High heterogeneity of Apolipoprotein E gene frequencies in South American Indians. *Ann Hum Biol* 27:29–34.
- Asakawa J-I, Takahashi N, Rosenblum BB, Neel JV. 1985. Two-dimensional gel studies of genetic variation in the plasma proteins of Amerindians and Japanese. *Hum Genet* 70:222–230.
- Barbujani G. 1987. Autocorrelation of gene frequencies under isolation by distance. *Genetics* 117:777–782.
- Boudreau DA, Scheer WD, Malcolm GT, Mulvad G, Pedersen HS, Jul E. 1999. Apolipoprotein E and atherosclerosis in Greenland Inuit. *Atherosclerosis* 145:207–219.
- Cavalli-Sforza LL, Bodmer WF. 1971. *The genetics of human populations*. San Francisco, California: Freeman.
- Cliff AD, Ord JK. 1973. *Spatial autocorrelation*. London: Pion.
- Crews DE, Kamboh MI, Mancilha-Carvalho JJ, Kottke B. 1993. Population genetics of apolipoprotein A-4, E, and H polymorphisms in Yanomami Indians of north-western Brazil: Associations with lipids, lipoproteins, and carbohydrate metabolisms. *Hum Biol* 65:211–224.
- Crutcher KA. 2004. Apolipoprotein E is a prime suspect, not just an accomplice, in Alzheimer's disease. *J Mol Neurosci* 23:181–188.
- Davignon J, Gregg RE, Sing CF. 1988. Apolipoprotein E polymorphism and atherosclerosis. *Arteriosclerosis* 8:1–21.
- Demarchi DA, Panzetta-Dutari MG, López De Basualdo M, Motran CC, Marcellino AJ. 2001. Mitochondrial DNA haplogroups in Amerindian populations from the Gran Chaco. *Am J Phys Anthropol* 115:199–203.
- Demarchi DA, Mitchell RJ. 2004. Genetic structure and gene flow in Gran Chaco populations of Argentina: Evidence from Y chromosome markers. *Hum Biol* 76:413–429.
- De Knijff P, Johansen LG, Rosseneu M, Frants RR, Jespersen J, Havekes LM. 1992. Lipoprotein profile of a Greenland Inuit population. Influence of anthropometric variables. APOE and A4 polymorphism, and lifestyle. *Arterioscler Thromb* 12:1371–1379.
- Excoffier L, Smouse P, Quattro J. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* 131:479–491.
- Fagundes NJ, Bonatto SL, Callegari-Jacques SM, Salzano FM. 2002. Genetic, geographic, and linguistic variation among South American Indians: Possible sex influence. *Am J Phys Anthropol* 117:68–78.
- Fix AG. 1997. Gene frequency clines produced by kin-structured founder effects. *Hum Biol* 69:663–673.
- Fix AG. 2004. Kin-structured migration: Causes and consequences. *Am J Hum Biol* 16:387–394.
- Gamboa R, Hernandez-Pacheco G, Hesiquio R, Zuñiga I, Massó, F, Montano LF, Ramos-Kuri M, Estrada J, Granados J, Vargas-Alarcon G. 2000. Apolipoprotein E polymorphism in Indian and Mestizo populations of Mexico. *Hum Biol* 72:975–981.
- Gerdes LU, Gerdes C, Hansen PS, Klausen IC, Faergeman O, Dyerberg J. 1996. The apolipoprotein E polymorphisms in Greenland Inuit in its global perspective. *Hum Genet* 98:546–550.
- Greenberg JH. 1987. *Language in the Americas*. Stanford, California: Stanford University Press.
- Grimes BF. 1992. *Ethnologue. Languages of the World*. Dallas: Summer Institute of Linguistics.
- Guo SW, Thompson EA. 1992. Performing the exact test for Hardy–Weinberg proportion for multiple alleles. *Biometrics* 48:361–372.
- Hallman DM, Boerwinkle E, Saha N, Sandholzer C, Menzel HJ, Csazar A, Utermann G. 1991. The apolipoprotein E polymorphisms: A comparison of allele frequencies and effects in nine populations. *Am J Hum Genet* 49:338–349.
- Hegele RA, Young TK, Connelly PW. 1997. Are Canadian Inuit at increased risk for coronary heart disease? *J Mol Med* 75:364–370.
- Jaramillo-Correa JP, Keyeux G, Ruiz-Garcia M, Rodas C, Bernal J. 2001. Population genetic analysis of the genes APOE, APOB (3/VNTR) and ACE in some Black and Amerindian communities from Colombia. *Hum Hered* 57:14–33.
- Kamboh MI, Weiss KM, Ferrell RE. 1991. Genetic studies of human apolipoproteins. XVI. APOE polymorphisms and cholesterol levels in the Mayans of the Yucatan Peninsula, Mexico. *Clin Genet* 39:26–32.
- Kamboh MI, Crawford MH, Aston CE, Leonard WR. 1996. Population distributions of APOE, APOH, and APOA4 polymorphisms and their relationships with quantitative plasma lipid levels among the Evenki herders of Siberia. *Hum Biol* 68:231–243.

- Kataoka S, Robbins DC, Cowan LD, Go O, Yeh JL, Devereux RB, Fabsitz RR, Lee ET, Welty TK, Howard BV. 1996. Apolipoprotein E polymorphism in American Indians and its relation to plasma lipoproteins and diabetes. *Arterioscler Thromb Vasc Biol* 16:918–925.
- Kesäniemi YA, Ehnholm C, Miettinen TA. 1987. Intestinal cholesterol absorption efficiency in man is related to apoprotein E phenotype. *J Clin Invest* 80:578–581.
- Loukotka C. 1968. Classification of South American Indian languages. Los Angeles: Latin American Center, University of California.
- Maekawa B, Cole TG, Seip L, Bylund D. 1995. Apolipoprotein E genotyping methods for the clinical laboratory. *J Clin Lab Anal* 9:63–69.
- Mahley RW. 1988. Apolipoprotein E: Cholesterol transport protein with expanding role in cell biology. *Science* 240:622–629.
- Mahley RW, Rall SC. 2000. Apolipoprotein E: Far more than a lipid transport protein. *Annu Rev Genomics Hum Genet* 1:507–537.
- Marin GB, Tavella MH, Guerreiro JF, Santos SEB, Zago MA. 1997. Absence of the E*2 allele of apolipoprotein in Amerindians. *Braz J Genet* 20:741–743.
- Miller MP. 1997. Tools for population genetic analyses (TFPGA) 1.3: A Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by the author.
- O'Rourke DH, Suarez BK. 1985. Patterns and correlates of genetic variation in South Amerindians. *Ann Hum Biol* 13:13–31.
- O'Rourke DH, Mobarry A, Suarez BK. 1992. Patterns of genetic variation in Native America. *Hum Biol* 64:417–434.
- Raymond ML, Rousset F. 1995. An exact test for population differentiation. *Evolution* 49:1280–1283.
- Salzano FM, Callegari-Jacques SD. 1988. South American Indians: A case study in Evolution. Oxford: Clarendon Press.
- Saunders AM, Strittmatter WJ, Schmechel D, George-Hyslop PH, Pericak-Vance MA, Joo SH, Rosi BL, Gusella JF, Crapper-MacLachlan, DR, Alberts MJ. 1993. Association of Apolipoprotein E (allele E4) with late-onset familial and sporadic Alzheimer's disease. *Neurology* 43:1467–1472.
- Scacchi R, Corbo RM, Rickards O, Mantuano E, Guevara A, De Stefano GF. 1997. Apolipoprotein B and E genetic polymorphisms in the Cayapa Indians of Ecuador. *Hum Biol* 69:375–382.
- Scheer WD, Boudreau DA, Malcolm GT, Middaugh JP. 1995. Apolipoprotein E and atherosclerosis in Alaska Natives. *Atherosclerosis* 114:197–202.
- Schneider S, Kueffer JM, Roessli D, Excoffier L. 1997. Arlequin, version 1.1: A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva.
- Sokal RR, Oden NL. 1978. Spatial autocorrelation in biology. 2. Some biological implications and four applications of evolutionary and ecological interest. *Biol J Linn Soc* 10:229–249.
- Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance MA, Enghild J, Sanvesen GS, Roses AD. 1993. Apolipoprotein E: High avidity binding to B-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer's disease. *Proc Natl Acad Sci USA* 90:1977–1981.
- Tsunoda K, Harihara S, Dashnyam B, Semjidmaa D, Yamaguchi Y, Tanabe Y, Sakai N, Sato A, Sato K. 2002. Apolipoprotein E and H polymorphisms in Mongolian Buryat: Allele frequencies and relationship with plasma lipid levels. *Hum Biol* 74:659–671.
- Wartemberg D. 1989. SAAP. Spatial Autocorrelation Analysis Program. Department of Environmental and Community Medicine. Robert Wood Johnson Medical School, Piscataway, NJ, USA.

Resume. *Arrière plan:* L'alipoprotéine E (protéine apoE; gene APOE) joue un rôle central dans le métabolisme des lipides. Trois allèles communs, E*2, E*3 et E*4 ont des effets quantitatifs sur les niveaux des lipides et des lipoprotéines, qui sont des déterminants de risque majeurs pour les maladies cardiovasculaires dans diverses populations. Au vu de leur importance clinique, il est intéressant de connaître la distribution des variantes APOE dans des populations ethniques diverses, ainsi que de déterminer si ce polymorphisme présente des variations qui pourraient être associées à certains facteurs évolutifs.

But: On décrit la distribution du polymorphisme APOE dans des populations amérindiennes d'Amérique du Sud et on la compare avec celles d'autres populations autochtones des Amériques et de Sibérie.

Sujets et Méthodes: L'échantillon est formé par 315 individus provenant de neuf populations amérindiennes vivant dans les latitudes subtropicales d'Argentine, du Brésil et du Paraguay. L'analyse étendue englobe 50 populations d'Amérique du Nord et du Sud, du Groenland et de Sibérie. La forme de la variation géographique a été analysée par des analyses de corrélation, par autocorrélation spatiale et par analyse de la variance moléculaire (AMOVA).

Résultats: L'incidence de l'allèle le plus commun (APOE*3) dans l'échantillon analysé varie de 0,78 à 0,98. Le second allèle en prévalence APOE*4 varie de 0,00 à 0,17. L'allèle rare APOE*2 est rencontré dans cinq des neufs populations étudiées. Cette variante est trouvée chez un homme ayant une double filiation maternelle et paternelle autochtone, ce qui suggère que cet allèle soit présent chez les amérindiens et ne puisse donc être utilisé comme indicateur de métissage. APOE*3 et APOE*4 présentent des associations respectivement positive et négative avec la latitude, bien que cela soit beaucoup plus prononcé dans l'hémisphère nord qu'en Amérique du Sud. APOE*2 accroît sa fréquence avec la latitude, mais cette modalité n'est statistiquement significative qu'en Amérique du Sud.

Conclusion: La distribution spatiale de APOE est en général compatible avec une expansion démographique directionnelle telle qu'elle s'est produite en Asie du Nord-Est et dans la plus grande partie du Nouveau Monde. L'allèle APOE*2 exprime cette distribution en Amérique du Sud mais présente une distribution aléatoire dans l'hémisphère nord, suggérant ainsi que l'intervention de la sélection ne doit pas être écartée.

Zusammenfassung. *Hintergrund:* Apolipoprotein E (apoE, Protein; *APOE*, Gen) spielt eine zentrale Rolle im Lipidstoffwechsel. Drei häufige Allele *E*2*, *E*3* und *E*4*, haben quantitative Effekte auf Lipid- und Lipoproteinspiegel, die wesentlich für das Risiko kardiovaskulärer Erkrankungen in zahlreichen Populationen verantwortlich sind. Angesichts ihrer klinischen Bedeutung ist die Kenntnis der Verteilung von APOE-Varianten in Populationen bestehend aus verschiedenen ethnischen Gruppen von Interesse, sowie die Bestimmung, ob dieser Polymorphismus Variationen beinhaltet, die mit bekannten evolutionären Faktoren assoziiert sein könnten.

Ziel: Wir berichten über die Verteilung von APOE-Polymorphismen bei Indianischen Populationen Südamerikas und vergleichen sie mit anderen eingeborenen Bevölkerungsgruppen aus Amerika und Sibirien.

Probanden und Methoden: Die untersuchte Gruppe bestand aus 315 Personen von neun Indianischen Bevölkerungsgruppen aus den subtropischen Breiten Argentiniens, Brasiliens und Paraguays. Der gesamte Umfang der Analyse umfasste 50 Bevölkerungsgruppen aus Süd- und Nordamerika, Grönland und Sibirien. Die geographischen Muster der Variation wurde mittels Korrelationsanalyse, räumlicher Autokorrelation und Analyse der molekularen Variation (AMOVA) untersucht.

Resultate: Die Inzidenz des häufigsten Allels (*APOE*3*) der untersuchten Gruppe schwankte zwischen 0,78 und 0,98. Das zweithäufigste Allel, *APOE*4*, schwankte zwischen 0,00 und 0,17. Das seltene Allel *APOE*2* wurde in fünf der neun untersuchten Bevölkerungsgruppen gefunden. Diese Variante wurde bei einem Mann gefunden, der sowohl von mütterlicher als auch von väterlicher Seite Indianischer Abstammung war. Das legt nahe, dass dies Allel bei Indianern angetroffen wird und deshalb nicht als Indikator für genetische Beimischungen verwendet werden sollte. *APOE*3* und *APOE*4* zeigen jeweils positive, bzw. negative Korrelationen mit der geographischen Breite. Allerdings ist dies Muster in der nördlichen Hemisphäre deutlicher ausgeprägt als in Südamerika. Die Häufigkeit von *APOE*2* steigt mit der geographischen Breite an, aber dies Muster ist nur in Südamerika statistisch signifikant.

Zusammenfassung: Das Gesamtmuster der räumlichen APOE-Verteilung scheint im Allgemeinen mit einer gerichteten demographischen Expansion vereinbar, die sich im Nordosten Asiens und in großen Teilen der Neuen Welt abgespielt hat. Das APOE*2-Allel zeigt dies Muster nur in Südamerika, während es auf der nördlichen Hemisphäre zufallsverteilt ist, was nahelegt, die Möglichkeit von Selektion nicht auszuschließen.

Resumen. *Antecedentes:* La apolipoproteína E (proteína Apo E; gen APOE) juega un papel central en el metabolismo lipídico. Tres alelos comunes, el E*2, E*3 y E*4, tienen efectos cuantitativos sobre los niveles lipídicos y de las lipoproteínas, los cuales son factores principales de riesgo para las enfermedades cardiovasculares en muchas poblaciones. Dada su importancia clínica, es interesante conocer la distribución de las variantes APOE en poblaciones de diversos grupos étnicos, así como determinar si su polimorfismo presenta variaciones que pudieran estar asociadas con determinados factores evolutivos.

Objetivo: Se muestra la distribución de los polimorfismos APOE en poblaciones de nativos americanos de Sudamérica, y se compara con otras poblaciones nativas de América y de Siberia.

Sujetos y métodos: La muestra consistió en 315 individuos de nueve poblaciones nativas americanas, que viven en latitudes subtropicales de Argentina, Brasil y Paraguay. El análisis completo incluía 50 poblaciones a lo largo de Sudamérica y Norteamérica, Groenlandia y Siberia. Los patrones geográficos de variación se investigaron mediante un análisis de correlación, de auto-correlación espacial y un análisis de varianza molecular (AMOVA).

Resultados: En la muestra analizada, la frecuencia del alelo más común (APOE*3) oscilaba entre 0,78 y 0,98. El segundo alelo en prevalencia, el APOE*4, variaba entre 0,00 y 0,17. El alelo raro APOE*2 se encontró en cinco de las nueve poblaciones investigadas. Esta variante fue encontrada en un varón con ascendencia nativa americana, tanto paterna como materna, lo que sugiere que este alelo está presente en los nativos americanos y, por lo tanto, no debería utilizarse como un indicador de mestizaje. Los alelos APOE*3 y APOE*4 presentan, respectivamente, asociaciones positivas y negativas con la latitud, aunque el patrón es mucho más pronunciado en el Hemisferio Norte que en Sudamérica. El alelo APOE*2 incrementa su frecuencia con la latitud, pero este patrón solo es estadísticamente significativo en Sudamérica.

Conclusión: El patrón espacial total del APOE parece, en general, compatible con una expansión demográfica direccional que ocurrió en el nordeste asiático y en gran parte del Nuevo Mundo. El alelo APOE*2 muestra este patrón en Sudamérica, pero una distribución aleatoria en el Hemisferio Norte, lo que sugiere que no debe descartarse la posibilidad de que haya existido selección.