

Alu insertions versus blood group plus protein genetic variability in four Amerindian populations¹

J. BATTILANA*, S. L. BONATTO†, L. B. FREITAS*, M. H. HUTZ*, T. A. WEIMER*, S. M. CALLEGARI-JACQUES*, M. A. BATZER‡, K. HILL§, A. M. HURTADO§, L. T. TSUNETO¶, M. L. PETZL-ERLER¶ and F. M. SALZANO*

*Biosciences Institute, Federal University of Rio Grande do Sul, Porto Alegre, Brazil

†Biosciences Faculty, Pontifical Catholic University of Rio Grande do Sul, Porto Alegre, Brazil

‡Department of Biological Sciences, Biological Computation and Visualization Center, Louisiana State University, Baton Rouge, USA

§Department of Anthropology, University of New Mexico, Albuquerque, USA

¶Department of Genetics, Federal University of Paraná, Curitiba, Brazil

Received 22 February 2001; in revised form 5 July 2001; accepted 27 July 2001

Summary. *Background:* Do the population relationships obtained using DNA or blood group plus protein markers remain the same or do they reveal different patterns, indicating that the factors which influence genetic variation at these two levels of analysis are diverse? Can these markers shed light on the biological classification of the Aché, a Paraguayan tribe which only recently established more permanent contacts with non-Indians?

Subjects and methods: To consider these questions we typed 193 individuals from four Amerindian tribes in relation to 12 *Alu* polymorphisms (five of them never studied in these populations), while 22 blood group plus protein systems were studied among the Aché. These data were then integrated with those previously available (blood groups plus proteins) for the three other populations. DNA extraction and amplification, as well as the other laboratory procedures, were performed using standard methods currently in use in our laboratory. The genetic relationships were obtained using the D_A distance, and the trees were constructed by the neighbour-joining method, both developed by M. Nei and collaborators. Reliability of the trees was tested by bootstrap replications. Other population variability values were also determined using Nei's methods.

Results: *Alu* polymorphism was observed in all populations and for most of the loci; in the seven systems from which we could compare our results with those of other Amerindian groups agreement was satisfactory. Unusual findings on the blood group plus protein systems of the Aché were a very low (5%) *HP*1* frequency and the presence of the C^W phenotype in the Rh blood group. The intertribal patterns of relationship and other aspects of their variation were remarkably congruent in the two sets (*Alu*; blood group plus protein) of systems.

Conclusions: The answer to the first question posed above is affirmative. However, the problem of whether the Aché derived from a Gê group that preceded the Guarani colonization of Paraguay, or are just a differentiated Guarani group, could not be answered with the genetic information available; the second hypothesis seems more likely at present, but the point to be emphasized is the striking genetic distinctiveness of the Aché as compared to other Amerindians.

1. Introduction

Alu insertions, the most abundant and extensively studied mammalian class of SINEs (short interspersed elements) of repetitive sequences, represent approximately 10% of nuclear DNA in humans (Smit 1996). They are dimeric sequences of approximately 300 base pairs (bp) in length (Houck, Rinehart and Schmid 1979) originated from the 7SL RNA gene (Ullu and Tschudi 1984). *Alu* sequences are

¹ This paper is dedicated to our colleague Maria Helena L. P. Franco, who collaborated in this and other investigations performed by the Porto Alegre group, and whose untimely death is mourned by all of us.

postulated to be retrotransposons that have been inserted into the human genome via a single-stranded RNA intermediate generated by RNA polymerase III transcription (Weiner, Deininger and Efstratiadis 1986).

There are approximately 2000 Ya5/8 and 2000 Yb8 insertional elements randomly dispersed throughout the human genome. Many of these reflect recent retroposition events that have not yet been fixed within the human species (Batzer and Deininger 1991, Batzer, Stoneking, Alegria-Hartmann *et al.* 1994, Batzer, Rubin, Hellmann-Blumberg *et al.* 1995, Novick, Novick, Yunis *et al.* 1995, Arcot, De Angelis, Sherry *et al.* 1997, Stoneking, Fontius, Clifford *et al.* 1997, Novick, Novick, Yunis *et al.* 1998). Due to changes during the evolution of the source genes, there are at least 12 major *Alu* subfamilies, which may be classified as old, intermediate or young. The two oldest subfamilies, Jo and Jb, arose from independent retroposition events involving a single, ancestral source gene that occurred early in primate evolution (Kapitonov and Jurka 1996). The polymorphic *Alu* elements may serve as markers to elucidate various aspects of human evolution. There are many advantages for their use as genetic markers in human populations. It is highly improbable that a specific *Alu* sequence has retroposed into a particular site more than once in human evolution; and recent insertions represent stable polymorphisms that are rarely lost without leaving a trace. Because of these two properties, identity by descent and evidential stability due to incomplete deletion, *Alu* insertions are literally molecular fossils. Another attractive feature of *Alu*-based population studies is that the ancestral condition is defined by the lack of an *Alu* element. Knowledge of the original character state is useful in rooting phylogenies. A further advantage of *Alu* genetic markers is that *Alu* genotypes involving individual or multiple polymorphic loci can be determined by a rapid PCR-based assay.

For many decades there have been extensive investigations on the biological characteristics of South Amerindians (Salzano and Callegari-Jacques 1988). Although in recent years the studies have included investigations at the DNA level, most of the data have been gathered from morphological characteristics or traditional genetic markers, such as blood groups and proteins. Despite the considerable efforts employed, the patterns of genetic variability among South American native populations are not yet sufficiently clear (Clariá, Demarchi, Moreno Azorero *et al.* 1998).

The present study furnishes new information on 12 *Alu* insertion polymorphisms for four Amerindian groups, as well as blood group and protein genetic data for one of them (the Aché of Paraguay). Previous results concerning these latter systems were compiled for the three other populations, and the following questions were addressed: (a) Are the patterns of relationships involving these four groups the same, considering the two sets of characteristics? and (b) Can they provide information about how the Aché should be classified? There are competing hypotheses in this regard, some scholars considering them a Gê group that preceded the Guarani colonization of Paraguay, while others contend that they are just a differentiated Guarani group.

2. Subjects and methods

General information about the four populations studied is presented in table 1 and figure 1. They live in the central-southern part of the continent, in Paraguay (Aché) and Brazil (the other three), speak languages classified in the Tupi and Gê-Kaingan linguistic groups, and the degree of non-Indian socio-cultural influence

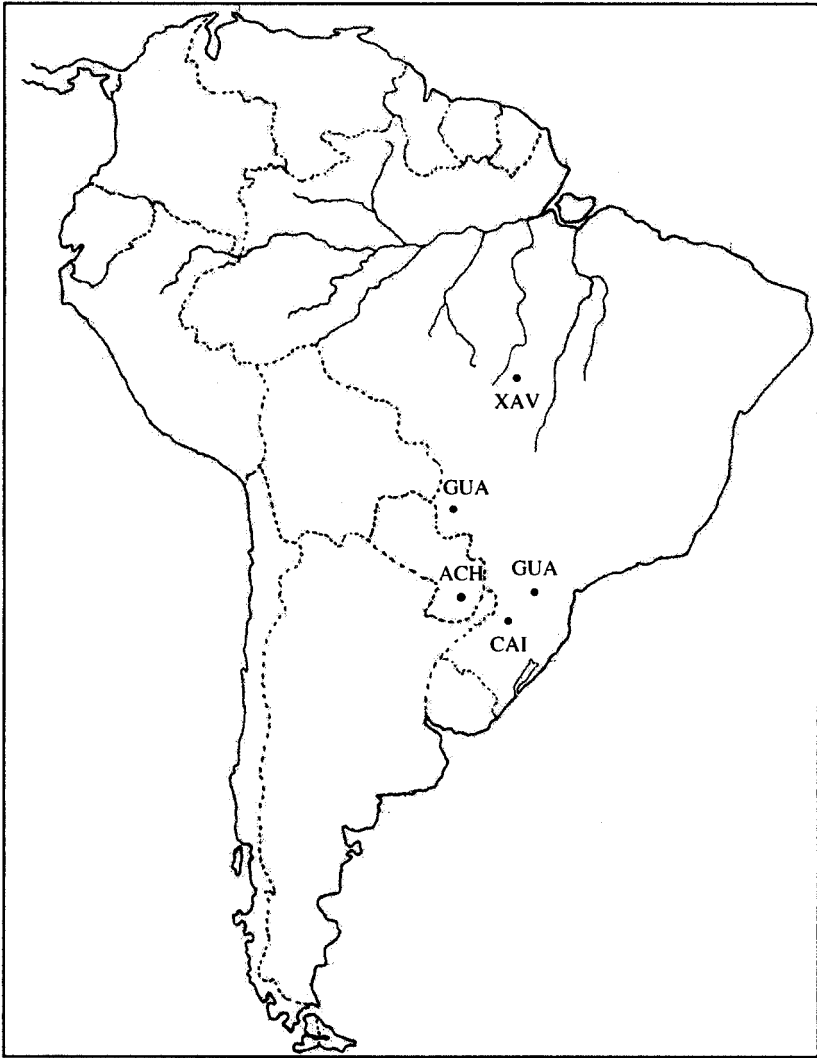


Figure 1. Geographic location of the four groups for which genetic data are reported here. ACH: Aché; CAI: Caingang; GUA: Guarani; XAV: Xavante.

received range from minimal (Aché), to moderate (Xavante), and extensive (Caingang, Guarani). Non-Indian ancestry, however, was estimated to be very small (0–7%) for all of them. Selected, relatively recent, additional information can be found in Hill and Hurtado (1996) for the Aché, Petzl-Erler, Luz and Sotomaior (1993) and Salzano, Callegari-Jacques, Weimer *et al.* (1997a) for the Caingang and Guarani, and Salzano, Franco, Weimer *et al.* (1997b) for the Xavante.

The blood samples were collected with anticoagulant, refrigerated shortly afterwards, and transported in this condition to Porto Alegre or Curitiba. The Aché samples were submitted to blood group typing immediately after arriving at the laboratory, using the DiaMed micro typing system of gel centrifugation and following the manufacturer's instructions. Haemoglobins were typed by horizontal starch

Table 1. Characterization of the groups studied.

Characteristics	Population			
	Aché (Guayaki)	Caingang	Guarani	Xavante
Localities	Arroyo Bandera Chupa-pou	Nonoai	Amambai Limão Verde Porto Lindo Rio das Cobras	Etéñitépa
Geographical coordinates	55°50'W, 23°30'S 56°30'W, 24°10'S	52°45'W, 27°20'S	55°12'W, 23°6'S 55°6'W, 23°12'S 54°30'W, 23°48'S 52°30'W, 25°20'S	51°40'W, 13°20'S
Language*	Guayaki	Kaingang	Guarani	Chavante
Linguistic group*	Tupi	Gê-Kaingan	Tupi	Gê-Kaingan
Non-Indian socio-cultural influence	minimal	extensive	extensive	moderate
Year of sample collection	1997	1975, 2000	1990, 1998	1990
Non-Indian ancestry (%)†	0	7	3	2

* According to Greenberg (1987).

† According to Salzano and Callegari-Jacques (1988) and Callegari-Jacques and Salzano (1999), using Szathmary and Reed's (1978) method.

gel electrophoresis employing the method described in Salzano and Tondo (1968); erythrocyte enzymes according to the techniques given in Harris and Hopkinson (1976); and serum proteins also using horizontal starch gel electrophoresis employing the buffer systems of Poulik (1957) for haptoglobin and transferrin and of Bowman and Bearn (1965) for the other two. Amido black 10B was used to stain albumin and transferrin, benzidine to stain haptoglobin, and orthodiansidine for ceruloplasmin.

DNA was extracted following standard procedures (Miller, Dykes and Polesky 1988, Lahiri and Nurnberger 1991). The oligonucleotide primers used for the *Alu* studies, as well as their annealing temperatures, have been described by Arcot, Fontius, Deininger *et al.* (1995a), Arcot, Wang, Weber *et al.* (1995b), Arcot, Adamson, Lamerdin *et al.* (1996), and Batzer, Arcot, Phinney *et al.* (1996). Polymerase chain reaction (PCR) amplification was carried out following Batzer and Deininger (1991) using a final volume of 25 μ L. Each sample was subjected to the following amplification procedures: 1 min at 94°C (denaturation), 2 min at the appropriate annealing temperature, 2 min at 72°C (extension) plus 5 min at 72°C (final extension) for 40 cycles. Fifteen microlitres of the PCR products were electrophoresed in 2% 1 \times TEB (Tris 90 mM, Acid Boric 90 mM, EDTA 0.5 M, pH 8.0, 0.002 mM, H₂O) agarose gels containing ethidium bromide, and the reaction products were directly visualized using ultraviolet fluorescence.

Allele frequencies were computed either by gene counting or using the MAXLIK program described by Reed and Schull (1968). A chi-square test for goodness of fit was used to verify if the observed genotype frequencies agreed with those expected under Hardy-Weinberg equilibrium using the Arlequin program (Schneider, Kueffer, Roessli *et al.* 1997). Inter-population heterogeneity was tested by the chi-square test of Roff and Bentzen (1989) for small numbers.

The genetic relationships between the four populations were evaluated using the D_A genetic distance of Nei, Tajima and Tateno (1983). These authors demonstrated that this method shows a good performance in all tree-making procedures, generally giving a larger probability value and a smaller deviation from the true values in simulation experiments. It also shows a better performance than others for closely related populations, such as those of humans. Trees were constructed by the neighbour-joining method (Saitou and Nei 1987). Additionally, three-dimensional plots based on principal coordinate and principal components analysis were also obtained (Sneath and Sokal 1973). The reliability of the trees was tested by bootstrap replications (Hedges 1992). For the genetic distances calculations the DISPAN (Ota 1993) and NTSYS (Rohlf 1987) programs were used. Average heterozygosities, associated standard errors, and other population variability values were determined according to Nei (1987).

3. Results and discussion

Information about the presence of the 12 *Alu* insertions in the four population groups is presented in table 2. Due to a series of circumstances (DNA availability or quality) sample sizes varied in relation to the sites studied (Ach e: 31–75; Caingang: 40–50; Guarani: 24–35; Xavante: 29–33), but on the whole they can be considered as fairly representative of the groups under study. Polymorphism was observed in all populations, and for most of the loci. Exceptions are 4.75, fixed in all tribes; *APO*, uniformly present in the Ach e and Xavante; *ACE*, in the Ache, and *FXIIB* in the Xavante. 4.25, 4.32 and 4.65 showed the lowest insertion frequencies, 4.65 being completely absent in the Ach e. Most distributions were in Hardy–Weinberg equilibrium. Some of them, however, showed departures from it. This is not unexpected, since in these small, endogamous communities, mating is far from random. Also, due to the number of comparisons made, it is expected that some may show such pattern just by chance; and finally, due to the small sample sizes, the position of just a few individuals in a given category may greatly influence the probability values. Extensive repeated tests were performed in the samples with the unexpected distributions, confirming the typings previously obtained.

For five insertions (4.75, 3.23, 4.59, 4.65, 4.32) these are the first results obtained in Amerindians. For the other seven the number of populations and individuals sampled is variable, and a summary is given in table 3. *TPA25* (16 populations, 593 individuals) is the most studied in North and Central America, while for South America *APO* (respectively 19 and 713) is the most studied. Generally our results in these seven systems agree with those obtained elsewhere, the largest difference occurring in 4.25 (our average, 9%, North plus Central America average, 21%). This also agrees with the fact that, as is shown in table 3, the differences in prevalences in the north plus central and south areas are also small.

Twenty-two blood group and protein systems were investigated among the Ach e, and the results are presented in table 4. They were monomorphic for 13 of them (ABO, Kell, Lutheran, haemoglobin (two loci), glucose-6-phosphate dehydrogenase, phosphogluconate dehydrogenase, phosphoglucomutase 2, adenylate kinase, acid phosphatase, transferrin, ceruloplasmin, albumin), in accordance with previous studies performed in South American Indians. Unusual findings were the low (5%) *HP*1* frequency, and the presence of the C^W phenotype in the Rh blood group.

Table 2. Presence of *Alu* insertions among members of four South American Indian populations.

<i>Alu</i> insertion	Genotype*	Aché			Caingang			Guarani			Xavante		
		No. of individuals	Allele frequencies	No. of individuals	Allele frequencies	No. of individuals	Allele frequencies	No. of individuals	Allele frequencies	No. of individuals	Allele frequencies		
FXIIB	+	45	0.782	32	0.872	27	0.935	30	1.000				
	+	21		11		4		0					
	-	5		0		0		0					
HW eq.		$p: 0.296 \pm 0.001$		$p: 1.000 \pm 0.000$		$p: 1.000 \pm 0.000$		NA					
4.75	+	74	1.000	40	1.000	27	1.000	29	1.000				
	+	0		0		0		0					
	-	0		0		0		0					
HW eq.		NA		NA		NA		NA					
MABDI	+	10	0.581	32	0.706	9	0.394	10	0.532				
	+	16		1		8		13					
	-	5		13		16		8					
HW eq.		$p: 1.000 \pm 0.000$		$p: 0.000 \pm 0.000$		$p: 0.009 \pm 0.000$		$p: 0.471 \pm 0.001$					
3.23	+	61	0.908	9	0.420	20	0.800	17	0.712				
	+	7		24		8		13					
	-	3		17		2		3					
HW eq.		$p: 0.008 \pm 0.000$		$p: 1.000 \pm 0.000$		$p: 0.308 \pm 0.001$		$p: 1.000 \pm 0.000$					
A25	+	0	0.013	0	0.037	0	0.097	1	0.234				
	+	2		3		6		13					
	-	73		38		25		18					
HW eq.		$p: 1.000 \pm 0.000$		$p: 1.000 \pm 0.000$		$p: 1.000 \pm 0.000$		$p: 0.652 \pm 0.001$					
TPA25	+	34	0.866	17	0.675	17	0.710	6	0.417				
	+	3		20		10		13					
	-	4		3		4		11					
HW eq.		$p: 0.001 \pm 0.000$		$p: 0.487 \pm 0.002$		$p: 0.219 \pm 0.001$		$p: 0.705 \pm 0.001$					

4.59	+	+	29	0.968	31	0.854	16	0.652	15	0.717
	+	-	2		8		11		13	
	-	-	0		2		6		2	
HW eq.			$p: 1.000 \pm 0.000$		$p: 0.181 \pm 0.001$		$p: 0.137 \pm 0.001$		$p: 1.000 \pm 0.000$	
4.65	+	+	0	0.000	2	0.159	0	0.021	0	0.078
	+	-	0		9		1		5	
	-	-	31		30		23		27	
HW eq.			NA		$p: 0.240 \pm 0.001$		$p: 1.000 \pm 0.000$		$p: 1.000 \pm 0.000$	
APO	+	+	36	1.000	37	0.963	32	0.941	33	1.000
	+	-	0		3		0		0	
	-	-	0		0		2		0	
HW eq.			NA		$p: 1.000 \pm 0.000$		$p: 0.001 \pm 0.000$		NA	
4.32	+	+	7	0.198	2	0.250	1	0.130	2	0.242
	+	-	11		20		5		11	
	-	-	45		26		21		18	
HW eq.			$p: 0.001 \pm 0.000$		$p: 0.700 \pm 0.001$		$p: 0.356 \pm 0.001$		$p: 1.000 \pm 0.000$	
PV92	+	+	51	0.855	28	0.793	20	0.783	23	0.813
	+	-	16		9		7		6	
	-	-	2		4		3		3	
HW eq.			$p: 0.619 \pm 0.001$		$p: 0.449 \pm 0.001$		$p: 0.100 \pm 0.001$		$p: 0.047 \pm 0.001$	
ACE	+	+	76	1.000	17	0.543	25	0.829	15	0.683
	+	-	0		16		8		11	
	-	-	0		13		2		4	
HW eq.			NA		$p: 0.043 \pm 0.001$		$p: 0.235 \pm 0.001$		$p: 0.417 \pm 0.002$	

*The presence and absence of the *Alu* repeat are denoted by + and -, respectively.

HW eq.: Exact test for Hardy-Weinberg equilibrium (probability value plus or minus its standard error). NA: Not applicable.

Table 3. Characteristics of previous studies involving seven *Alu* insertions performed in Amerindians.

Geographical region and statistical characteristics	Loci						
	<i>FXIIB</i>	<i>MABD1</i>	<i>A25</i>	<i>TPA25</i>	<i>APO</i>	<i>PV92</i>	<i>ACE</i>
North and Central America							
No. of samples	10	2	2	16	10	10	10
No. of individuals	323	101	101	593	323	323	323
Lowest frequency	0.50	0.45	0.21	0.29	0.90	0.57	0.44
Highest frequency	1.00	0.46	0.21	0.66	1.00	0.99	0.89
Mean	0.84	0.45	0.21	0.55	0.97	0.75	0.70
South America							
No. of samples	17	4	—	21	19	19	17
No. of individuals	665	313	—	668	713	704	508
Lowest frequency	0.53	0.42	—	0.12	0.58	0.42	0.57
Highest frequency	1.00	0.65	—	0.93	1.00	1.00	0.98
Mean	0.90	0.54	—	0.53	0.97	0.87	0.79

Sources: Batzer *et al.* (1994); Barley, Blackwood, Carter *et al.* (1994); Tishkoff, Ruano, Kidd *et al.* (1996); Stoneking *et al.* (1997); Novick *et al.* (1998); Oliveira (1999); Rupert, Devine, Monsalve *et al.* (1999); Tishkoff, Pakstis, Stoneking *et al.* (2000).

Three other blood group and protein studies had been conducted among the Ache (or Guayaki, a term that the Ache consider as derogatory). They have been reported by Matson, Sutton, Swanson *et al.* (1968), Brown, Gajdusek, Leyshon *et al.* (1974), and Clariá *et al.* (1998), and involved all southern, while our data were obtained in northern groups. Respectively 16, 15 and 4 allele distributions could be compared considering the present data and those obtained by these authors, and significant differences were observed in the MNSs, Rh, Lewis and haptoglobin systems only. Generally, the most marked diversity was found between the frequencies reported here and those given by Brown *et al.* (1974). We have found high (53% vs 28%) *L*Ns* (and complementary low *L*Ms*), as well as high (82% vs 60%) *RH*CDe* (compensated by low *RH*cDE*) frequencies as compared to those obtained by the indicated authors. In the Lewis system, our finding of 67% only of *LEWIS*Le*, against 100% found by Matson *et al.* (1968), is much more in accordance with other studies using this marker (Salzano and Callegari-Jacques 1988). Finally, as far as we can ascertain, the Aché *HP*1* frequency observed here is the lowest reported so far in South American Indians. In the other studies with this group previously mentioned the values were also low (22% reported both by Matson *et al.* 1968 and Brown *et al.* 1974, against a general average, for South American Indians, of 60% ± 14%, Santos, Ribeiro-dos-Santos, Guerreiro *et al.* 1998), but not as low as the present figure. Repeated typings were also performed to confirm our findings, and they provided uniformly the same result.

Are the *Alu* differences found among the four tribes under consideration congruent with those observed for the blood group plus protein systems? To answer this question we assembled previous results available for the latter (displayed in the Appendix) to establish the comparison. These data were obtained from the same communities studied here, with the exception of the Guarani, who had been previously surveyed in a more southerly place (see figure 1 and table 1).

The two sets of genetic distances are shown in table 5, and the derived dendrograms in figure 2. As can be seen, they depict essentially the same picture, the Caingang and Xavante (who speak languages of the same linguistic group) clustering

Table 4. Blood group and protein genetic systems studied among the Aché.

System	Phenotype	No. of individuals	Alleles or haplotypes	Frequency	HW* χ^2
ABO	O	99	<i>ABO*O</i>	1.000	NA
MNSs	MS	1	<i>L*MS</i>	0.222	25.09
	MSs	3	<i>L*Ms</i>	0.247	3 d.f.
	Ms	7	<i>L*Ns</i>	0.531	$p < 0.001$
	MNS	0			
	MNSs	39			
	MNs	32			
	NS	0			
	NSs	0			
P	P1	97	<i>P*1</i>	0.858	NA
	P2	2			
Rh	CC ^W De	1	<i>RH*CDE</i>	0.011	2.95
	CDE	1	<i>RH*CDe</i>	0.813	2 d.f.
	CDEe	0	<i>RH*C^WDe</i>	0.005	$p > 0.20$
	CDe	66	<i>RH*cDE</i>	0.105	
	CcDE	0	<i>RH*cDe</i>	0.066	
	CcDEe	16			
	CcDe	12			
	cDE	2			
	cDEe	1			
	cDe	0			
Kell	K-	99	<i>KELL*K-</i>	1.000	NA
Duffy	a+b+	34	<i>Fy*A</i>	0.828	4.25
	a+b-	65			1 d.f. $p < 0.05$
Lewis	a-b+	88	<i>LEWIS*Le</i>	0.667	NA
	a-b-	11			
Lutheran	a-b+	99	<i>Lu*b</i>	1.000	NA
Haemoglobin	A	99	<i>HB*A</i>	1.000	NA
	A2	99	<i>HB*A2</i>	1.000	NA
Glucose-6-phosphate-dehydrogenase	B	45 F 54 M	<i>G6PD*B</i>	1.000	NA
Phosphogluconate dehydrogenase	A	99	<i>PGD*A</i>	1.000	NA
Phosphoglucomutase 1	1-1	48	<i>PGM1*1</i>	0.707	0.53
	2-1	44			1 d.f.
	2-2	7			$p > 0.30$
Phosphoglucomutase 2	1-1	99	<i>PGM2*1</i>	1.000	NA
Phosphoglycolate phosphatase	1-1	62	<i>PGP*1</i>	0.788	0.11
	2-1	32			1 d.f.
	2-2	5			$p > 0.70$
Adenylate kinase	1-1	99	<i>AK*1</i>	1.000	NA
Acid phosphatase	B	99	<i>ACP*B</i>	1.000	NA
Esterase D	1-1	83	<i>ESD*1</i>	0.919	0.76
	2-1	16			1 d.f.
	2-2	0			$p > 0.30$
Glyoxalase 1	1-1	5	<i>GLO*1</i>	0.364	12.35
	2-1	62			1 d.f.
	2-2	32			$p < 0.001$
Haptoglobin	1-1	1	<i>HP*1</i>	0.050	2.45
	2-1	8			1 d.f.
	2-2	90			$p > 0.10$
Transferrin	C	99	<i>TF*C</i>	1.000	NA
Ceruloplasmin	B	99	<i>CP*B</i>	1.000	NA
Albumin	A	99	<i>ALB*A</i>	1.000	NA

* HW: Hardy-Weinberg (chi-square test for equilibrium). NA: Not applicable.

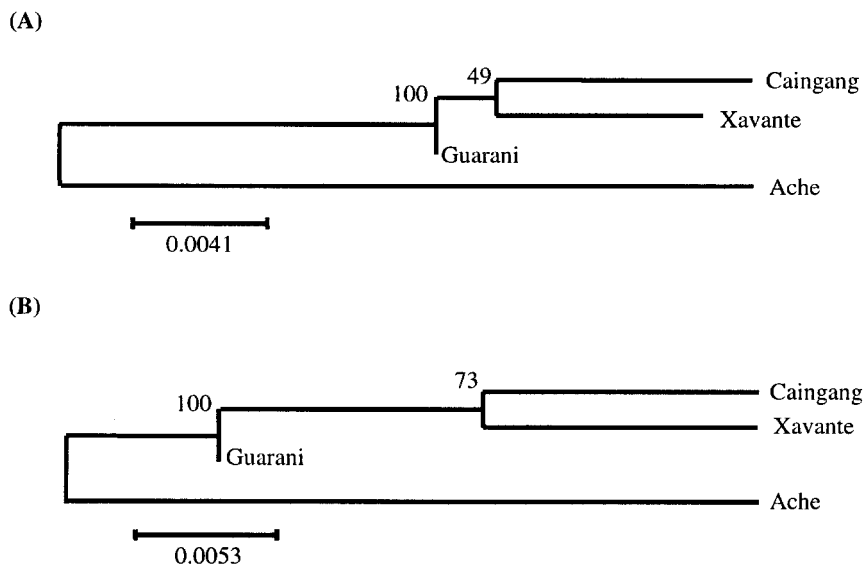


Figure 2. Dendrograms obtained with the neighbour-joining method using the genetic distances displayed in table 5 and mid-point rooting. (a) Protein plus blood group systems, (b) *Alu* elements.

Table 5. Matrix of genetic distances. Above diagonal: blood group plus protein systems; below diagonal: *Alu* insertion frequencies.

	Aché	Caingang	Guarani	Xavante
Aché	—	0.0410	0.0314	0.0400
Caingang	0.0499	—	0.0091	0.0138
Guarani	0.0300	0.0214	—	0.0072
Xavante	0.0564	0.0211	0.0148	—

together. In both cases the Aché remain far from the others, with the Guarani occupying an intermediate position. Essentially the same results were obtained with the principal coordinate and principal components analyses (data not shown).

Average heterozygosities considering the *Alu* insertions and the blood group plus protein systems results showed essentially the same patterns, with the Aché having somewhat lower values (0.15 for both sets), while the three other groups are more uniform (0.21–0.23; 0.25–0.29, respectively). But since the standard errors are high, the differences are statistically non-significant. Total variability, and the amount of it that is due to inter-population differences were similar considering the *Alu* polymorphisms (0.26; 8%) and the blood group plus protein results (0.23; 10%).

We are now in a position to answer the two questions posed in the Introduction: (a) yes, the intertribal patterns of relationship and other aspects of their variation show excellent congruence in the two sets of systems; and (b) due to the marked genetic peculiarities of the Aché we cannot decide between the two hypotheses concerning their classification, but presently the view that they are a differentiated Guarani group seems more likely. The point to be emphasized, however, is their distinctiveness in relation to the other Amerindians in general. Unpublished results we have in other genetic systems are pointing in the same direction, and a previous analysis made by Salzano and Callegari-Jacques (1988) showed that in a dendrogram obtained comparing 58 South American Indian groups uniformly studied for seven

genetic systems (MNSs, P, Rh, Duffy, Kidd, Diego and haptoglobin) the Aché clearly differentiated from all of them with a single exception. They clustered with the Parakanã, who live 2500 km away, in the Amazon region, but who also speak a Tupi-Guarani language and have, like the Aché, light skin.

Acknowledgements

Thanks are due to the Fundação Nacional do Índio (FUNAI) for permission to study the Indians and help in the field, and to the Fundación Bertoni, for logistic assistance. The Indian leaders and the subjects of the investigation were appropriately informed about the aims of the study and gave their approval, which is gratefully acknowledged. The investigation was also approved by the Brazilian National Ethics Commission (CONEP Resolution no. 123/98). Financial support was provided by Programa de Apoio a Núcleos de Excelência (PRONEX, Brazil), Financiadora de Estudos e Projetos (FINEP, Brazil), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS, Brazil), Louisiana Board of Regents Millennium Trust Health Excellence Fund HEF (2000-05)-05 and HEF (2000-05)-01, National Science Foundation (NSF, USA), and National Institutes of Health (NIH, USA).

Appendix

Blood group and protein genetic systems allele or haplotype frequencies observed among the Caingang, Guarani and Xavante*.

System	Allele or haplotype frequency	Caingang (n = 35)†	Guarani (n = 99)†	Xavante (n = 85)†
ABO	<i>ABO*O</i>	1.00	1.00	1.00
MNSs	<i>L*MS</i>	0.36	0.29	0.29
	<i>L*Ms</i>	0.36	0.35	0.43
	<i>L*NS</i>	0.12	0.09	0.16
	<i>L*Ns</i>	0.16	0.27	0.12
P	<i>P*1</i>	0.39	0.59	0.52
Rh	<i>RH*CDE</i>	0.01	0.01	0.00
	<i>RH*CDe</i>	0.44	0.60	0.56
	<i>RH*cDE</i>	0.39	0.22	0.16
	<i>RH*cDe</i>	0.08	0.14	0.28
	<i>RH*cde</i>	0.08	0.03	0.00
Kell	<i>KELL*k</i>	1.00	1.00	1.00
Duffy	<i>FY*A</i>	0.46	0.58	0.45
Haemoglobin	<i>HB*A</i>	1.00	1.00	1.00
Glucose-6-phosphate-dehydrogenase	<i>G6PD*B</i>	1.00	1.00	1.00
Phosphogluconate dehydrogenase	<i>PGD*A</i>	1.00	1.00	1.00
Phosphoglucomutase 1	<i>PGM1*1</i>	0.86	0.82	0.85
Phosphoglucomutase 2	<i>PGM2*1</i>	1.00	1.00	1.00
Adenylate kinase	<i>AK*1</i>	1.00	1.00	1.00
Acid phosphatase	<i>ACP*B</i>	0.89	0.88	0.78
Esterase D	<i>ESD*1</i>	0.89	0.49	0.63
Glyoxalase 1	<i>GLO*1</i>	0.25	0.18	0.22
Haptoglobin	<i>HP*1</i>	0.69	0.59	0.55
Transferrin	<i>TF*C</i>	1.00	1.00	0.85
Ceruloplasmin	<i>CP*B</i>	1.00	1.00	1.00
GC	<i>GC*1</i>	0.34	0.42	0.74
Albumin	<i>ALB*A</i>	1.00	1.00	1.00

* Based on data reported as follows: Caingang: Salzano and Shreffler (1966); Salzano, Callegari-Jacques, Franco *et al.* (1980); Guarani: Salzano *et al.* (1997a); Xavante: Salzano *et al.* (1997b).

† Modal number of individuals studied.

References

- ARCOT, S. S., FONTIUS, J. F., DEININGER, P. L., and BATZER, M. A., 1995a, Identification and analysis of a 'young' polymorphic *Alu* element. *Biochimica et Biophysica Acta*, **1263**, 99–102.
- ARCOT, S. S., WANG, Z., WEBER, J. L., DEININGER, P. L., and BATZER, M. A., 1995b, *Alu* repeats: a source for the genesis of primate microsatellites. *Genomics*, **29**, 136–144.
- ARCOT, S. S., ADAMSON, A. W., LAMERDIN, J. E., KANAGY, B., DEININGER, P. L., CARRANO, A. V., and BATZER, M. A., 1996, *Alu* fossil relics—distribution and insertion polymorphism. *Genome Research*, **6**, 1084–1092.
- ARCOT, S. S., DE ANGELIS, M. D., SHERRY, S. T., ADAMSON, A. W., LAMERDIN, J. E., DEININGER, P. L., CARRANO, A. V., and BATZER, M. A., 1997, Identification and characterization of two polymorphic Ya5 *Alu* repeats. *Mutation Research Genomics*, **382**, 1–5.
- BARLEY, J., BLACKWOOD, A., CARTER, N. D., CREWS, D. E., CRUICKSHANK, J. K., JEFFERY, S., OGUNLESI, A. O., and SAGNELLA, G. A., 1994, Angiotensin converting enzyme insertion/deletion polymorphism: association with ethnic origin. *Journal of Hypertension*, **12**, 955–956.
- BATZER, M. A., and DEININGER, P. L., 1991, A human-specific subfamily of *Alu* sequences. *Genomics*, **9**, 481–487.
- BATZER, M. A., STONEKING, M., ALEGRIA-HARTMANN, M., BAZAN, H., KASS, D. H., SHAIKH, T. H., NOVIK, G. E., IOANNOU, P. A., SCHEER, W. D., HERRERA, R. J., and DEININGER, P. L., 1994, African origin of human-specific polymorphic *Alu* insertions. *Proceedings of the National Academy of Sciences, USA*, **91**, 12288–12292.
- BATZER, M. A., RUBIN, C. M., HELLMANN-BLUMBERG, U., ALEGRIA-HARTMANN, M., LEEFLANG, E. P., STERN, J. D., BAZAN, H., SHAIKH, T. H., DEININGER, P. L., and SCHMID, C. W., 1995, Dispersion and insertion polymorphism in two small subfamilies of recently amplified human *Alu* repeats. *Journal of Molecular Biology*, **247**, 418–427.
- BATZER, M. A., ARCOT, S. S., PHINNEY, J. W., ALEGRIA-HARTMANN, M., KASS, D. H., MILLIGAN, S. M., KIMPTON, C., GILL, P., HOCHMEISTER, M., IOANNOU, P. A., HERRERA, R. J., BOUDREAU, D. A., SCHEER, W. D., KEATS, B. J. B., DEININGER, P. L., and STONEKING, M., 1996, Genetic variation of recent *Alu* insertions in human populations. *Journal of Molecular Evolution*, **42**, 22–29.
- BOWMAN, B. H., and BEARN, A. G., 1965, The presence of sub-units in the inherited group specific protein of human serum. *Proceedings of the National Academy of Sciences, USA*, **53**, 722–729.
- BROWN, S. T., GADJUSEK, D. C., LEYSHON, W. C., STEINBERG, A. G., BROWN, K. S., and CURTAIN, C. C., 1974, Genetic studies in Paraguay: blood group, red cell and serum genetic patterns of the Guayaki and Ayore Indians, Mennonite settlers and seven other Indian tribes of the Paraguayan Chaco. *American Journal of Physical Anthropology*, **41**, 317–343.
- CALLEGARI-JACQUES, S. M., and SALZANO, F. M., 1999, Brazilian Indian/non-Indian interaction and their effects. *Ciência e Cultura*, **51**, 166–174.
- CLARIÀ, D. M., DEMARCHI, D. A., MORENO AZORERO, R., and GARDENAL, C. N., 1998, Protein polymorphism in three South Amerindian populations. *Annals of Human Biology*, **25**, 581–588.
- GREENBERG, J. H., 1987, *Language in the Americas* (Stanford: Stanford University Press).
- HARRIS, H., and HOPKINSON, D. A., 1976, *Handbook of Enzyme Electrophoresis in Human Genetics* (Amsterdam: North Holland).
- HEDGES, S. B., 1992, The number of replications needed for accurate estimation of the bootstrap P value in phylogenetic studies. *Molecular Biology and Evolution*, **9**, 366–369.
- HILL, K., and HURTADO, A. M., 1996, *Ache Life History* (New York: Aldine de Gruyter).
- HOUCK, C. M., RINEHART, F. P., and SCHMID, C. W., 1979, An ubiquitous family of repeated DNA sequences in the human genome. *Journal of Molecular Biology*, **132**, 289–306.
- KAPITONOV, V., and JURKA, J., 1996, The age of *Alu* subfamilies. *Journal of Molecular Evolution*, **42**, 59–65.
- LAHIRI, D. K., and NURNBERGER, J. I., 1991, A rapid non-enzymatic method for preparation of HMW DNA from blood for RFLP studies. *Nucleic Acids Research*, **19**, 5444.
- MATSON, G. A., SUTTON, H. E., SWANSON, J., and ROBINSON, A., 1968, Distribution of blood groups among Indians in South America. VI. In Paraguay. *American Journal of Physical Anthropology*, **29**, 81–98.
- MILLER, A. S., DYKES, D. D., and POLESKY, H. F., 1988, A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, **16**, 1215.
- NEI, M., 1987, *Molecular Evolutionary Genetics* (New York: Columbia University Press).
- NEI, M., TAJIMA, F., and TATENO, Y., 1983, Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *Journal of Molecular Evolution*, **19**, 153–170.
- NOVICK, G. E., NOVICK, C. C., YUNIS, J., YUNIS, E., MARTINEZ, K., DUNCAN, G., TROUP, G. M., DEININGER, P. L., STONEKING, M., BATZER, M. A., and HERRERA, R. J., 1995, Polymorphic human specific *Alu* insertions as markers for human identification. *Electrophoresis*, **16**, 1596–1601.

- NOVICK, G. E., NOVICK, C. C., YUNIS, J., YUNIS, E., MAYOLO, P. A., SCHEER, W. D., DEININGER, P. L., STONEKING, M., YORK, D. S., BATZER, M. A., and HERRERA, R. J., 1998, Polymorphic *Alu* insertions and the Asian origin of Native American populations. *Human Biology*, **70**, 23–39.
- OLIVEIRA, S. F., 1999, Inserções *Alu* em populações indígenas da Amazônia Brasileira. Ph.D. Thesis, Universidade de São Paulo, São Paulo.
- OTA, T., 1993, *DISPAN: Genetic and Phylogenetic Analysis* (University Park, PA: Institute of Molecular Evolutionary Genetics, Pennsylvania State University).
- PETZL-ERLER, M. L., LUZ, R., and SOTOMAIOR, V. S., 1993, The HLA polymorphism of two distinctive South-American Indian tribes: the Kaingang and the Guarani. *Tissue Antigens*, **41**, 227–237.
- POULIK, M. D., 1957, Starch gel electrophoresis in a discontinuous system of buffers. *Nature*, **180**, 1477–1479.
- REED, T. E., and SCHULL, W. J., 1968, A general maximum likelihood method estimation program. *American Journal of Human Genetics*, **20**, 579–580.
- ROHLF, F. J., 1987, *NTSYS-pc. Numerical Taxonomy and Multivariate Analysis System for the IBM PC Microcomputer (and Compatibles)* (Setauket: Applied Biostatistics Inc.).
- ROFF, D. A., and BENTZEN, P., 1989, The statistical analysis of mitochondrial DNA polymorphisms: χ^2 and the problem of small samples. *Molecular Biology and Evolution*, **6**, 539–545.
- RUPERT, J. L., DEVINE, D. V., MONSALVE, M. V., and HOCHACHKA, P. W., 1999, Angiotensin-converting enzyme (ACE) alleles in the Quechua, a high altitude South American native population. *Annals of Human Biology*, **26**, 375–380.
- SAITOU, N., and NEI, M., 1987, The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406–425.
- SALZANO, F. M., and SHREFFLER, D. C., 1966, The Gc polymorphism in the Caingang Indians of Brazil. *Acta Genetica et Statistica Medica*, **16**, 242–247.
- SALZANO, F. M., and TONDO, C. V., 1968, Hemoglobin types of Brazilian Indians. *American Journal of Physical Anthropology*, **28**, 355–360.
- SALZANO, F. M., and CALLEGARI-JACQUES, S. M., 1988, *South American Indians: A Case Study in Evolution* (Oxford: Clarendon Press).
- SALZANO, F. M., CALLEGARI-JACQUES, S. M., FRANCO, M. H. L. P., HUTZ, M. H., WEIMER, T. A., SILVA, R. S., and DA ROCHA, F. J., 1980, The Caingang revisited: blood genetics and anthropometry. *American Journal of Physical Anthropology*, **53**, 513–524.
- SALZANO, F. M., CALLEGARI-JACQUES, S. M., WEIMER, T. A., FRANCO, M. H. L. P., HUTZ, M. H., and PETZL-ERLER, M. L., 1997a, Electrophoretic protein polymorphisms in Kaingang and Guarani Indians of southern Brazil. *American Journal of Human Biology*, **9**, 505–512.
- SALZANO, F. M., FRANCO, M. H. L. P., WEIMER, T. A., CALLEGARI-JACQUES, S. M., MESTRINER, M. A., HUTZ, M. H., FLOWERS, N. M., SANTOS, R. V., and COIMBRA, C. E. A. JR., 1997b, The Brazilian Xavante Indians revisited: new protein genetic studies. *American Journal of Physical Anthropology*, **104**, 23–34.
- SANTOS, S. E. B., RIBEIRO-DOS-SANTOS, A. K. C., GUERREIRO, J. F., SANTOS, E. J. M., WEIMER, T. A., CALLEGARI-JACQUES, S. M., MESTRINER, M. A., FRANCO, M. H. L. P., HUTZ, M. H., and SALZANO, F. M., 1998, New protein genetic studies in six Amazonian Indian populations. *Annals of Human Biology*, **25**, 505–522.
- SMIT, A. F., 1996, The origin of interspersed repeats in the human genome. *Current Opinion in Genetics and Development*, **6**, 743–748.
- SCHNEIDER, S., KUEFFER, J.-M., ROESSLI, D., and EXCOFFIER, L., 1997, Arlequin ver 1.1: a software for population data analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva.
- SNEATH, P. H. A., and SOKAL, R. R., 1973, *Numerical Taxonomy* (San Francisco: W. H. Freeman).
- STONEKING, M., FONTIUS, J. J., CLIFFORD, S. L., SOODYALL, H., ARCOT, S. S., SAHA, N., JENKINS, T., TAHIR, M., DEININGER, P. L., and BATZER, M. A., 1997, *Alu* insertion polymorphisms and human evolution: evidence for a larger population size in Africa. *Genome Research*, **7**, 1061–1071.
- SZATHMAY, E. J. E., and REED, T. E., 1978, Calculation of the maximum amount of gene admixture in a hybrid population. *American Journal of Physical Anthropology*, **48**, 29–34.
- TISHKOFF, S. A., RUANO, G., KIDD, J. R., and KIDD, K. K., 1996, Distribution and frequency of a polymorphic *Alu* insertion at the plasminogen activator locus in humans. *Human Genetics*, **97**, 759–764.
- TISHKOFF, S. A., PAKSTIS, A. J., STONEKING, M., KIDD, J. R., DESTRO-BISOL, G., SANJANTILA, A., DEINARD, A. S., SIRUGO, G., JENKINS, T., KIDD, K. K., and CLARK, A. G., 2000, Short tandem-repeat polymorphism/*Alu* haplotype variation at the PLAT locus: implications for modern human origins. *American Journal of Human Genetics*, **67**, 901–925.
- ULLU, E., and TSCHUDI, C., 1984, *Alu* sequences are processed 7SL RNA genes. *Nature*, **312**, 171–172.
- WEINER, A. M., DEININGER, P. L., and EFSTRATIADIS, A., 1986, Nonviral retroposons: genes, pseudo-genes, and transposable elements generated by the reverse flow of the genetic information. *Annual Review of Biochemistry*, **55**, 631–661.

Address for correspondence: Prof. Francisco M. Salzano, Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, 91 501-970 Porto Alegre, RS, Brazil. email: francisco.salzano@ufrgs.br

Zusammenfassung. *Hintergrund:* Sind die Verwandtschaftsbeziehungen zwischen Populationen die anhand der DNA ermittelt wurden die gleichen wie die, welche anhand der Blutgruppe und der Proteinmarker gefunden wurden? Zeigen sich Unterschiede, die darauf hinweisen, dass die Faktoren, welche die genetische Variabilität auf diesen zwei Niveaus der Analyse beeinflussen, verschieden sind? Können diese Marker bei der biologischen Klassifikation der Aché, einem paraguayischen Stamm, der erst vor kurzem ständigen Kontakte mit Nicht-Indianern aufnahm, genutzt werden?

Material und Methoden: Zur Klärung dieser Fragen wurden 193 Individuen von vier Amerindianischen Stämmen in Bezug auf 12 *Alu* Polymorphismen typisiert (fünf davon wurden in diesen Populationen bisher noch nie untersucht), während 22 Blutgruppen und Proteinsysteme der Aché untersucht wurden. Diese Daten wurden dann zu jenen Daten (Blutgruppen und Proteine) in Beziehung gesetzt, welche vorher für die drei anderen Populationen vorhandenen waren. DNA Extraktion und Amplifikation sowie andere Laboranalysen wurden in unserem Labor mit aktuell gebräuchlichen Standardmethoden durchgeführt. Die genetischen Verwandtschaften wurden mittels D_A -Distanz ermittelt und die Stammbäume durch die neighbour-joining Methode konstruiert, welche von M. Nei *et al.* entwickelt wurde. Die Zuverlässigkeit der Stammbäume wurde durch Bandenreplikationen geprüft. Andere Variabilitäten bei den Populationen wurden auch mit Nei's Methoden bestimmt.

Ergebnisse: *Alu* Polymorphismen wurden in allen Populationen und für die meisten Loci beobachtet; in den sieben Systemen, in denen wir unsere Resultate mit denen anderer Amerindianischer Gruppen vergleichen konnten, war die Übereinstimmung zufriedenstellend. Ungewöhnliche Befunde bei der Blutgruppe und den Proteinsystemen der Aché waren eine sehr niedrig (5%) *HP*1* Frequenz und das Vorhandensein eines C^W Phänotypes bei der Rh Blutgruppe. Die Verwandtschaftsbeziehungen zwischen den Stämmen und andere Aspekte ihrer Variabilität waren bei den zwei Systemen (*Alu*; Blutgruppe und Protein) bemerkenswert übereinstimmend.

Schlussfolgerungen: Die Antwort auf die erste der aufgeworfenen Fragen ist positiv. Jedoch lässt sich das Problem ob die Aché von einer Gê Gruppe abstammen, welche der Guarani Besiedlung von Paraguay voranging, oder eine unterschiedliche Guarani Gruppe darstellen, nicht mit den verfügbaren genetischen Informationen lösen. Die zweite Hypothese scheint zur Zeit wahrscheinlicher, aber ein Punkt, der hervorgehoben werden muss, ist die auffallende genetische Besonderheit der Aché verglichen mit anderen Amerindians.

Résumé. *Arrière-plan:* Est-ce que les relations entre populations observées au moyen d'ADN ou de groupes sanguins plus marqueurs protéiques demeurent identiques, ou bien révèlent-elles des dispositions différentes qui indiqueraient que les facteurs qui influencent la variation génétique à ces deux niveaux sont différents? Ces marqueurs peuvent-ils aider à répertorier les Aché, une ethnie paraguayenne ni n'a que récemment établi des contacts plus fréquents avec des non indiens?

Sujets et méthodes: Afin de répondre à ces questions, on a typé 193 individus de 4 ethnies amérindiennes pour 12 polymorphismes *Alu* (dont 5 n'ont jamais été étudiés dans ces populations), tandis que 22 groupes sanguins et sériques ont été étudiés chez les Aché. Ces données ont été ensuite intégrées à celles qui étaient déjà disponibles (groupe sanguins et sériques) pour les trois autres populations. L'extraction et l'amplification de l'ADN ainsi que d'autres procédures de laboratoire ont été réalisées par les méthodes standard habituellement utilisées par notre laboratoire. Les associations génétiques ont été obtenues au moyen de la distance D_A et les dendrogrammes construits par la méthode de regroupement de voisinage établie par M. Nei et collaborateurs. La validité des dendrogrammes a été testée par répliquations croisées. D'autres valeurs de variabilité populationnelle ont également été déterminées par les méthodes de Nei.

Résultats: Le polymorphisme *Alu* est observé dans toutes les populations et pour la plupart des loci. Dans sept systèmes à partir desquels nous avons pu comparer nos résultats avec ceux d'autres groupes amérindiens, l'agrément a été satisfaisant. Des résultats inhabituels sur les groupes sanguins et sériques des Aché portaient sur une fréquence très basse (5%) de *HP*1* et sur la présence du phénotype C^W dans le système Rh. Les schémas intertribaux d'association et d'autres aspects de leur variation, étaient remarquablement congruents dans les deux séries de systèmes (*Alu*/groupes sanguins et sériques).

Conclusions: La réponse à la première question est affirmative. Par contre, le problème de savoir si les Aché sont des dérivés d'un groupe Gê qui a précédé la colonisation du Paraguay par les Guarani ou sont seulement un groupe guarani différencié, n'a pas pu être résolu par les données génétiques disponibles. La seconde hypothèse paraît à présent plus probable, mais le point qui doit être éclairci est celui de la frappante singularité des Aché par rapport aux autres amérindiens.