Patterns of Genetic Variation in Native America

DENNIS H. O'ROURKE,[1] ANNE MOBARRY,[1] AND BRIAN K. SUAREZ[2]

Abstract Allele frequencies from seven polymorphic red cell antigen loci (ABO, Rh, MN, S, P, Duffy, and Diego) were examined in 144 Native American populations. Mean genetic distances (Nei's D) and the fixation index F_{ST} are approximately equal for the North and South American samples but are reduced in the Central American geographic area. The relationship between genetic distance and geographic distance differs markedly across geographic areas. The correlation between geographic distance and genetic distance for the North and Central American data is twice as large as that observed for the South American samples. This geographic difference is confirmed in spatial autocorrelation analyses; no geographic structure is apparent in the South American data but geographic structure is prominent in North and Central American samples. These results confirm earlier observations regarding differences between North and South American gene frequency patterns.

Technological advances in two areas in the 1980s have provided new tools for assessing genetic variability in human populations at a level of resolution previously unattainable. Molecular biology restriction fragment length polymorphisms (RFLPs) and DNA sequence data have produced new and important insights into human evolution [e.g., Cann et al. (1987) and Di Rienzo and Wilson (1991)] and most recently have begun to be used to examine the evolutionary dynamics and biological history of Native American populations [e.g., Kidd et al. (1991), Wallace (this issue), Ward et al. (1991), Pääbo (1986), Pääbo et al. (1988), and Shearin et al. (1989)]. Simultaneous developments in computer technology, especially computer graphics, now permit visual display of complex patterns of variation not obvious in traditional forms of statistical analysis [e.g., Barnhill (1977, 1983), O'Rourke et al. (1986), and Piazza et al. (1981a)]. In conjunction with the new computer technologies, mul-

[1]Department of Anthropology, University of Utah, Salt Lake City, UT 84112.
[2]Departments of Psychiatry and Genetics, Washington University School of Medicine, St. Louis, MO 63110.

Human Biology, June 1992, Vol. 64, No. 3, pp. 417–434.

KEY WORDS: GENE FREQUENCIES, NATIVE AMERICA, AMERINDS, GEOGRAPHY, GENETIC DISTANCE

tivariable and multivariate methods for analyzing spatial pattern in population data have also proved beneficial [e.g., Sokal and Oden (1978a,b), Sokal and Wartenberg (1983), Sokal et al. (1986), Wartenberg (1985), and Barbujani (1987)].

Despite the recent advances in molecular genetic methods and markers, the most extensive genetic data on Native American populations for comparative purposes and for analysis by the newer analytical techniques are the traditional markers, especially red cell antigens. Recent summaries of variation in the highly polymorphic HLA loci (Black 1980, 1984, 1991; Black et al. 1983) and in the GM and KM systems [e.g., Schanfield, (this issue)] are available and will not be dealt with here.

Geography of Gene Frequencies

The earliest reports of gene frequency variation in Amerindian populations are those of G.A. Matson and colleagues on the Blood, Blackfoot, Chilcotin, and other groups of Montana and Alberta, Canada (Matson and Schrader 1933; Matson et al. 1936; Matson 1938). Since then, hundreds of reports of gene frequency variation in Amerindians of both North and South America have appeared, including several interpretive summaries (Spuhler 1979; Szathmary 1984; Salzano and Callegari-Jacques 1988; Black 1991; Barrantes et al. 1990; Crawford and Enciso 1982) focusing on specific geographic regions or marker systems. This wealth of genetic data, including the emerging data from RFLPs and DNA sequences, demonstrates the marked heterogeneity of Native American populations.

In a series of earlier articles, we attempted (1) to define the geographic pattern of gene frequency variation in native North America (Suarez, Crouse, and O'Rourke 1985; O'Rourke and Lichty 1989) and South America (O'Rourke and Suarez 1985) and (2) to address some of the hypotheses regarding evolutionary mechanisms that might have given rise to the observed patterns (Suarez, O'Rourke, and Crouse 1985; O'Rourke et al. 1985). The method used to capture the geographic distribution of gene frequencies in these studies was that of synthetic gene frequency maps (Menozzi et al. 1978; Piazza et al. 1981a; Suarez, Crouse, and O'Rourke 1985; O'Rourke and Suarez 1985; O'Rourke and Lichty 1989).

North America. In general, there are pronounced north-south clines in red cell antigen frequencies in native North America [e.g., Suarez, Crouse, and O'Rourke (1985)]. For some loci this may be the result of natural selection operating through mechanisms correlated with ecological zones, as indicated by strong association with patterns of climatic

variation (Piazza et al. 1981b; Ananthakrishnan and Walter 1972; O'Rourke et al. 1985), which also are characterized by latitudinal clines. For loci that show no association with ecological or climatic factors such clinal patterns must originate from other mechanisms, such as migration and admixture.

Spuhler (1979) documented the correlation of North Amerindian gene frequencies with language families, which might also impose a geographic pattern to the genetic variables. Recent work by Cavalli-Sforza et al. (1988) on the role of language affinities relative to the evolution of gene frequency patterns has occasioned considerable debate [e.g., Bateman et al. (1990), Zegura et al. (1990), Nichols (1990), and Wiley et al. (1990)]. Although not new, the study of the relationship between linguistic variation and genetic diversity has seen renewed interest in recent years.

Among all North American populations the Eskimo groups are clearly genetically distinct from Amerind populations, principally through their lower *ABO*O* and higher *ABO*A* allele frequencies. This distinctiveness is apparent in the visual patterning of synthetic gene frequency maps (Suarez, Crouse, and O'Rourke 1985; Menozzi et al. 1978) and in more conventional forms of analysis and would seem to confirm the separate and recent origin of Eskimo groups in the North American Arctic. Strong clinal distributions for individual gene frequencies and principal components of allele frequencies south of the Arctic were demonstrated through spatial autocorrelation analysis (O'Rourke and Lichty 1989).

The clear geographic (clinal) structure seen in gene frequency data from northern North America is less clear in similar data on Central American groups [see, e.g., Suarez, Crouse, and O'Rourke (1985, Figures 2 and 3)]. In this regard the geographic patterning of gene frequency data in Central America is intermediate between the patterns observed for the North and South American continents taken individually.

South America. As is the case with North America, population genetic studies of South America have a long and productive history. Three recent works have attempted to summarize the patterns of genetic variation for the continent (Black 1991; Salzano and Callegari-Jacques 1988; O'Rourke and Suarez 1985). The results of these syntheses are not altogether concordant. 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. Rather, they concluded that "the geography of gene frequencies in South Amerindians seems to be irregular, nonpredictable and derived principally from isolated populations drifting independently" (p. 24). 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. Similarly, O'Rourke and Suarez (1985) found no relationship between gene frequency distributions and local climatic conditions, whereas Salzano and Callegari-Jacques (1988) reported significant correlations between some gene frequencies and monthly temperature range.

More recently, Black (1991) has argued for some geographic structure to allele frequencies of the HLA system in South Amerindian groups. In this case he suggests that recent population history and linguistic affiliation are more important than generally accredited in previous analyses of this type. Barrantes et al. (1990) also report geographic structure in Central American gene frequency data and the important role of linguistic diversity in this pattern.

These alternative inferences regarding the patterning of gene frequency variation in South America indicate the complexity of South Amerindian population genetics, emphasize the difference between North and South America with respect to the geography of gene frequencies, and suggest that additional analyses are warranted. The remainder of this article focuses on a closer examination of the geographic patterning of gene frequencies in Native American populations.

Methods

The data come from the seven red cell antigen loci used to construct genetic maps in earlier publications (Suarez, Crouse, and O'Rourke 1985; O'Rourke and Suarez 1985). Inclusion of population samples in the analysis required sample sizes greater than 20 and reported frequencies for at least 5 of the 7 systems examined (ABO, Rh, MN, S, P, Duffy, and Diego). If there were missing data, geographic distances were calculated between a population with a missing frequency and all other populations and a distance weighted average was computed. This method assigns greatest weight to geographically proximal populations and assumes little contribution from distant ones. In the North American samples 29 populations required the estimation of frequencies for a single system (Suarez, Crouse, and O'Rourke 1985), whereas in the South American samples only 6 frequencies needed to be estimated (O'Rourke and Suarez 1985).

For the present analyses 3 geographic areas are defined: North America, with 44 populations (including 7 Eskimo samples); Central America, with 30 groups; and South America, with 70 samples. The geographic distinction between North America and Central America is taken to be 22° north latitude. Figure 1 illustrates the geographic location of the North and Central American samples, and Figure 2 gives the location of the 70 South American populations.

Allele frequencies for the seven loci examined are used to compute Nei's (1985) genetic distance between all pairs of populations in each

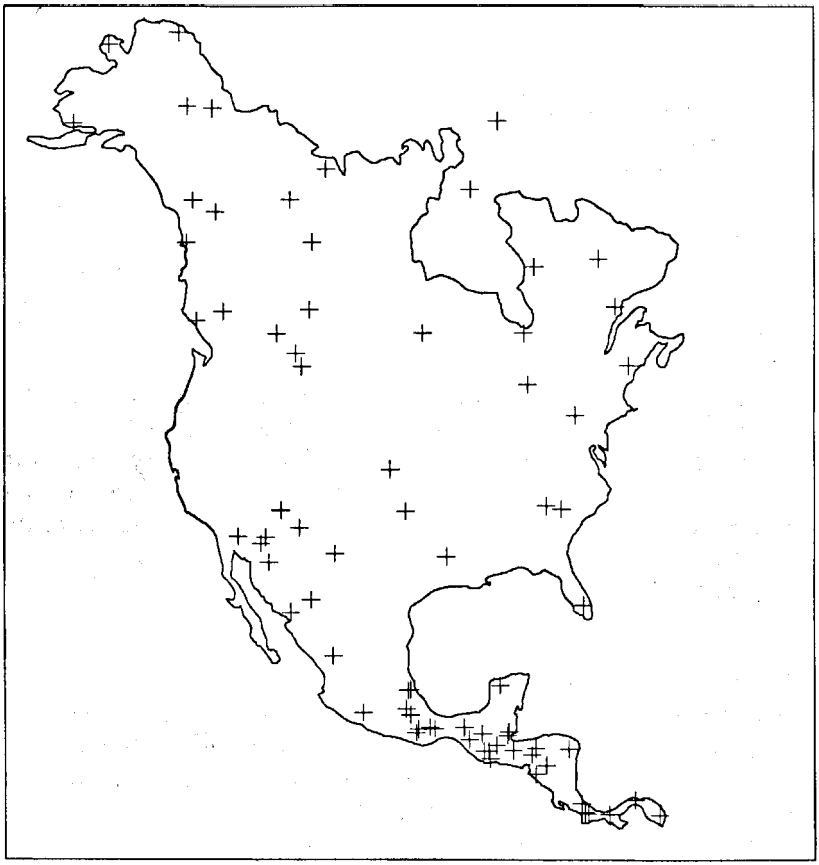

Figure 1. Geographic locations of 44 North American and 30 Central American populations used in analysis.

of the three geographic regions. Nei's genetic distance statistic is defined as

$$D = J_{XY}/(J_{XX}J_{YY})^{1/2}, \tag{1}$$

where J_{XY} is the probability of randomly selecting an allele identical in state from two separate populations (X and Y) and J_{XX} and J_{YY} are the probabilities that two randomly sampled alleles from the same population are identical. The genetic distances are related to the corresponding great circle arc distances between samples by simple least-squares regression.

Spatial autocorrelation analysis was used to assess geographic structure in the allele frequency data (Sokal and Oden 1978b; Sokal and

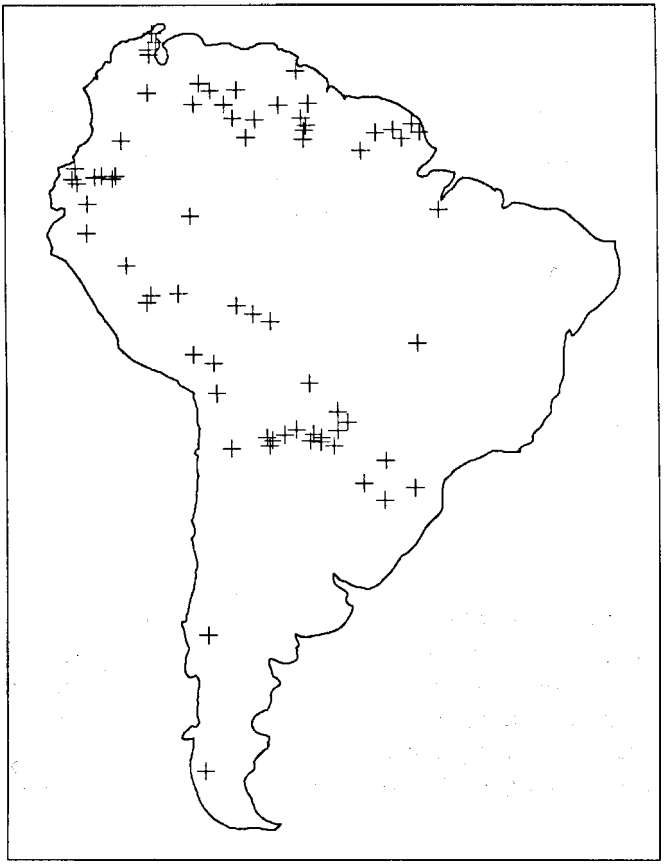

Figure 2. Geographic locations of 70 South American populations used in analysis.

Wartenberg 1983; Sokal et al. 1986). Moran's I, the measure of spatial autocorrelation, is defined as

$$I = n \Sigma w_{ij} z_i z_j / W \Sigma z_i^2 \quad (2)$$

where n is the number of populations under analysis, W is the connectivity matrix with elements w_{ij}, and the z terms represent the deviation of the allele frequency of each population from the total population mean for that frequency (Cliff and Ord 1981). Results of the autocorrelation analyses are presented as spatial correlograms with 12 distance classes. The distance classes are equidistant and represent approximately 625 km for each class.

Table 1. Summary Statistics for Amerindian Genetic Data

Statistic	*North America*	*South America*	*Central America*
Number of populations	44	70	30
Mean genetic distance (Nei's *D*)	0.055	0.062	0.029
Range	0.003–0.203	0.003–0.275	0.002–0.134
F_{ST}	0.0902	0.0906	0.0517
Correlation: Geographic vs. genetic distance			
Transformed data (square-root)	0.362	0.174	0.308
Untransformed data	0.331	0.127	0.319
Mantel permutation test (untransformed data)	0.337[a]	0.128[b]	0.304[c]

a. $p = 0.004$.
b. $p = 0.020$.
c. $p = 0.008$.

Results

Genetic and Geographic Distances. Summary statistics for the analyses performed on the geographic and genetic distances are given in Table 1. The qualitative comparison of genetic distances and fixation indexes across the geographic areas is informative. Mean genetic distance is slightly greater among the South American samples than among the North American groups, indicating somewhat greater heterogeneity among South Amerindians. Conversely, average genetic distances are noticeably smaller among the Central American samples. This pattern is also reflected in the range of genetic distances by geographic area, where South America has the greatest range of *D* values and Central America has by far the most restricted range of genetic distances.

Similarly, the F_{ST} values show nearly the same pattern. Here, degree of population differentiation appears to be nearly identical in North and South America but reduced in Central America. This pattern is concordant with that seen for genetic distances and is not unexpected, given the smaller geographic area of Central America and the density of observations from this region.

The relationship between genetic and geographic distances between pairs of North American populations is shown in Figure 3. Both axes of Figure 3 have been transformed to the square-root of the original distances. This avoids clumping of points near the origin because of the large number of population pairs that are geographically close and therefore have a relatively small genetic distance between them. The least-squares regression line of the transformed variables is also shown in Figure 3. The visual impression of association between geographic and genetic

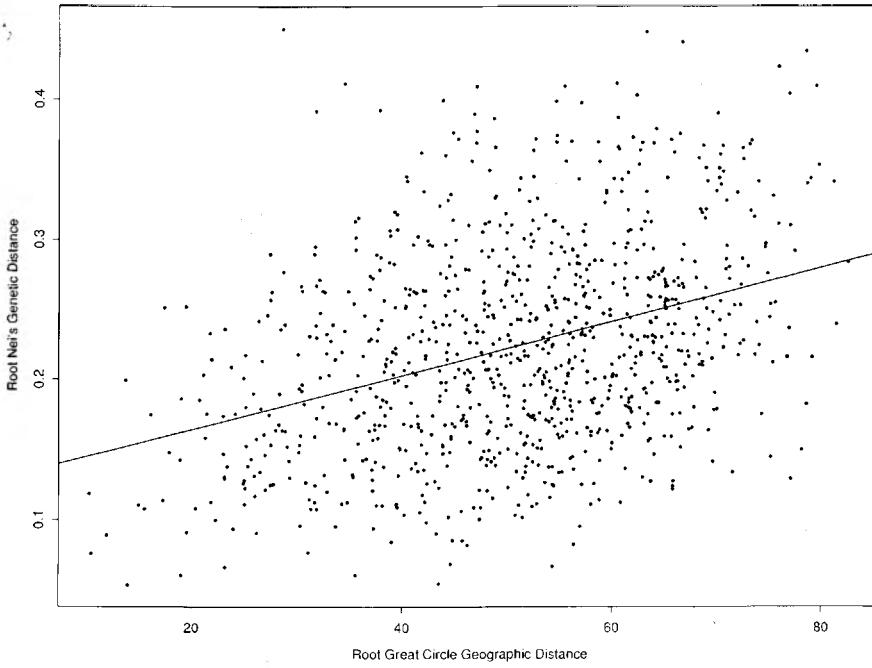

Figure 3. Scatter plot of square-root of geographic and genetic distances for North Amerindian populations. Solid line is fitted least-squares regression.

distance given by the distribution of points in Figure 3 is characterized by a correlation of $r = 0.362$ on the transformed variables or $r = 0.331$ on the untransformed data (see Table 1).

In contrast, the relationship between geographic and genetic distance in South America is given in Figure 4. The points in the scatter plot are more randomly distributed than those in Figure 3, and indeed the least-squares regression line is nearly flat. The correlation defining the relationship between the two distance measures among South American groups is $r = 0.174$ for the transformed data illustrated in Figure 4. Note that significance levels have not been assigned to the correlations based on either the untransformed or the root-transformed data in Table 1. The number of pairwise comparisons far exceeds the available degrees of freedom such that the requirement of independence of observations is not met. Thus standard significance levels are not appropriate or meaningful.

The Mantel permutation procedure (Mantel 1967), however, does permit evaluation of such significance levels (Legendre and Fortin 1989), and these are also provided in Table 1 for the untransformed distance values. The correlations obtained from the Mantel procedure are between

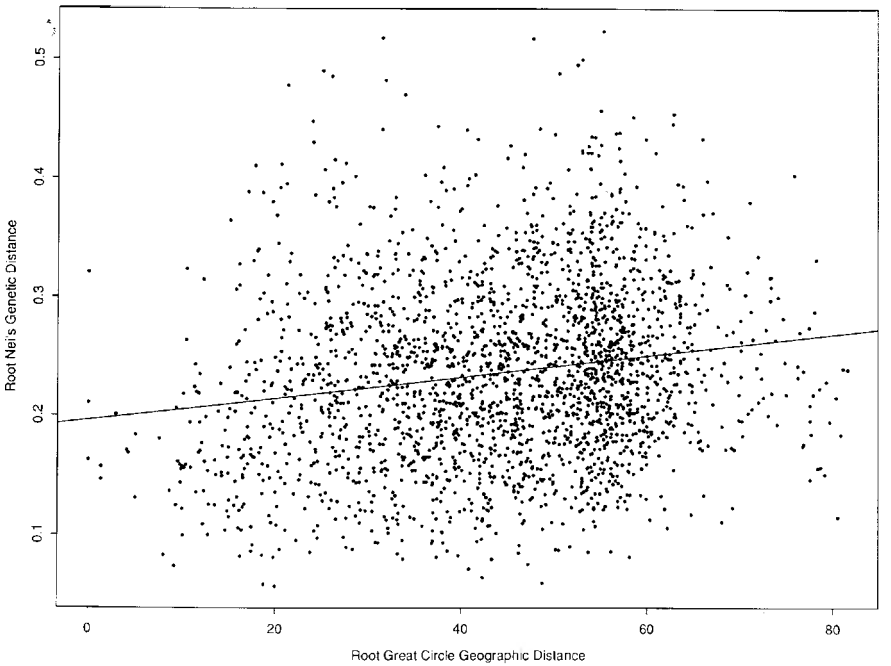

Figure 4. Scatter plot of square-root of geographic and genetic distances for South Amerindian populations. Solid line is fitted least-squares regression.

the triangular genetic and geographic distance matrices. Note that the relatively low correlation for the South American distances is statistically significant at the $\alpha = 0.02$ level; the correlation for the North American matrices is larger and, despite the substantially smaller number of samples, statistically significant at the $\alpha = 0.004$ level. Irrespective of the significance levels, the magnitude of difference between the measures of distance association in North America versus South America suggests that substantial differences in genetic structure are present.

As noted earlier, synthetic gene frequency maps suggest that the geographic distribution of gene frequencies in Central America differs from geographic patterns to the north and shares some similarities with the South American continent. Accordingly, 30 Central American populations located south of 22° north latitude through Panama were examined. Results of the distance analysis on these groups are displayed in Figure 5.

As illustrated in Figure 5, the relationship between geographic distance and genetic distance in Central America is intermediate between the two major continental samples, although the Central American pattern is more similar to the North American pattern. The correlation be-

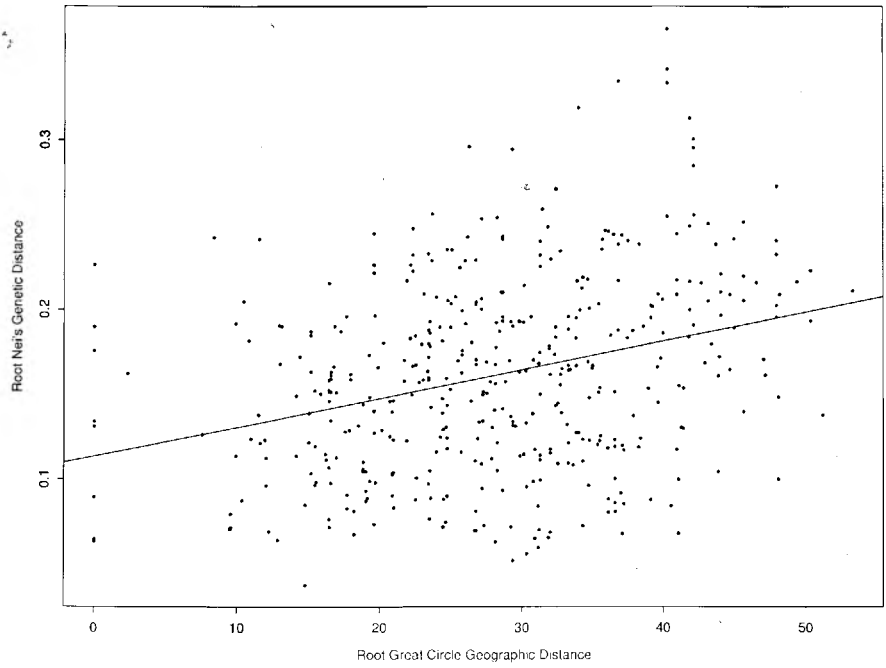

Figure 5. Scatter plot of square-root of geographic and genetic distances for Central Amerindian populations. Solid line is fitted least-squares regression.

tween the root-transformed geographic and genetic distances ($r = 0.308$) for these Central American populations more nearly approximates the value obtained in the North American samples than in the South American populations (Table 1). The Mantel permutation test on the Central American data also indicates a significant association between genetics and geography in these data.

Discussion

The results raise two issues regarding patterns of genetic variation in Native America: (1) the minimal association between genetic and geographic distances among South Amerindian populations compared with North and Central American groups and (2) the origin of disparate results regarding gene frequency clines in South America by different investigators. It is useful to consider these issues in reverse order.

In our earlier work (O'Rourke and Suarez 1985) we found little evidence for gene frequency clines in South American red cell antigen frequencies. Conversely, Salzano and Callegari-Jacques (1988) report

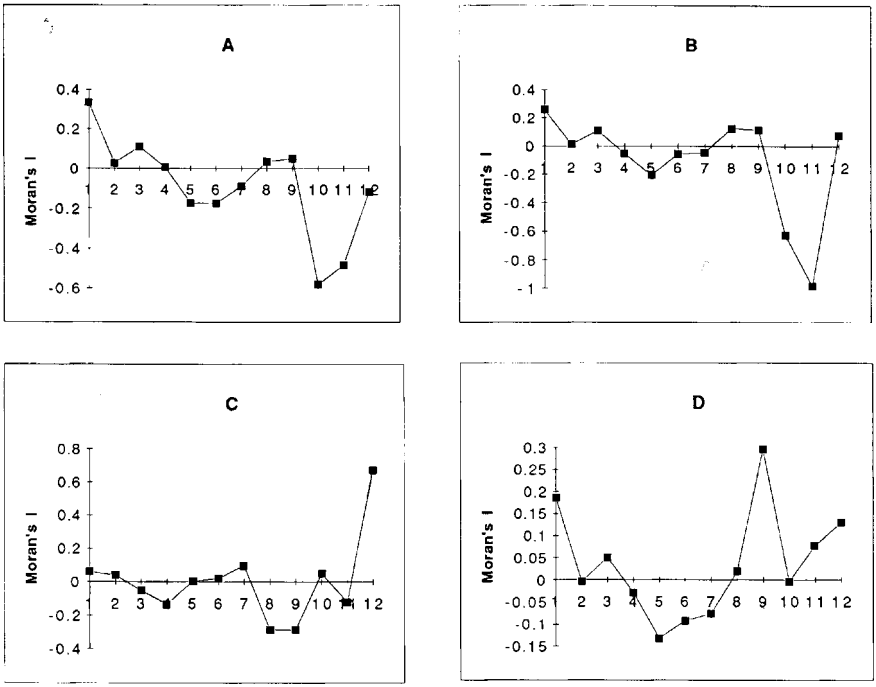

Figure 6. Spatial correlograms for (A) *RH*R1*, (B) *RH*R2*, (C) *Di*^{a}, and (D) *FY*A* allele frequencies observed in South Amerindian populations.

prominent clines for specific alleles on the continent. There are several reasons for this discrepancy. First, Salzano and Callegari-Jacques used a slightly expanded series of markers, and the strongest evidence for clines in allele frequencies occur for those loci not included in our data (e.g., Lewis and GC). Nevertheless, Salzano and Callegari-Jacques report clines for the *RH*R1* and *RH*R2* alleles of the RH system, for the Duffy locus, and, after pooling multiple samples from single populations, for the Diego locus.

Although we found evidence for clinal distributions for these variables in North America (Suarez, Crouse, and O'Rourke 1985; O'Rourke and Lichty 1989), neither synthetic gene frequency maps nor the observed relationship between genetic and geographic distance (Figure 4 and Table 1) based on our data suggests similar patterns in South America. Figure 6 presents spatial correlograms for the two RH alleles and alleles for Duffy and Diego over 12 equidistant distance classes.

As expected, the correlograms for the two RH alleles (Figures 6A,B) are similar. For *RH*R1* Moran's *I* is statistically significant at 6 distance classes (1, 3, 5, 6, 10, and 11), whereas for *RH*R2* only 5 values reach

significance (1, 3, 5, 10, and 11). More important, the clinal pattern of statistically significant positive values at short distances and monotonically decreasing values to large negative ones at greater distances is absent. Hence there is no evidence for clinal distribution of these alleles in the sample populations.

The correlograms for the Duffy and Diego alleles (Figures 6C–D) are equally interesting. Neither presents any evidence of clinal distribution but, for Duffy in particular, the correlograms suggest extreme heterogeneity of allele frequencies over geographic space. It is worth noting that in Salzano and Callegari-Jacques's (1988) multiple regression analyses of allele frequencies on latitude and longitude, variation at the Duffy locus was significantly associated with latitude when all samples were included in the analysis but was significantly associated with longitude when multiple reports from the same group were pooled. Diego frequencies were found to be significantly associated with longitude only in the pooled sample analysis. This reversal of implied clinal direction for the Duffy frequencies suggests caution in interpreting geographic pattern from these analytical methods. Indeed, a constant problem of assessing patterns of geographic variation in biological characters is the spatial distribution of the samples (O'Rourke and Lichty 1989). This is particularly true of South American gene frequencies.

In South America reported frequencies for most markers form a geographic C on the continent (see Figure 2). There is a relatively dense series of samples across the equatorial north of the continent in the Amazon and Orinoco river basins, a similar grouping of samples running north to south along the Andes, and, finally, a group of samples spanning northern Argentina, Paraguay, and southern Brazil. For most marker systems, data outside this genetic arc are simply not available. Such nonuniformity of sample distribution is not generally observed in North American samples but can be an analytical problem when examining South American data. Regression using latitude and/or longitude and allele frequencies may be particularly susceptible to this distributional anomaly. The nondirectional spatial autocorrelation method is one approach to circumvent the problem, and the results prove consistent with inferences made from the surfaces generated in synthetic gene frequency maps.

It should be noted that the geographic and spatial analyses presented here are also sensitive to sample distribution changes. For example, if the North American and Central American groups are combined into a single sample, as was done in earlier analyses [cf. Suarez, Crouse, and O'Rourke (1985) and O'Rourke et al. (1985)], the observed correlation between geographic and genetic distances is $r = 0.50$, a highly statistically significant value. This is an increase of 50% over the observed correlation between distance matrices in these samples treated

separately (see Table 1). Moreover, similar analyses performed on a subset of the South American groups for which local climatic data are available resulted in a correlation between geographic and genetic distances of only $r = 0.09$, a nonsignificant value by the Mantel test. Although the correlations varied widely in these analyses, as did the observed significance levels, the basic inference remains the same: There is little evidence for geographic structure in South Amerindian gene frequency data. This is not true for corresponding data in North America.

In this context, the analyses presented here and those of Salzano and Callegari-Jacques (1988) can be viewed as complementary rather than conflicting. Salzano and Callegari-Jacques found some evidence for geographic structure in specific allele frequencies, whereas we do not. However, the sensitivity of these analyses to sample composition may be the principal factor in this difference. Any investigation of patterns of gene frequency variation over broad geographic areas must weigh the importance of maximizing the number of samples versus the number of markers. Salzano and Callegari-Jacques (1988) examined variation in a larger series of genetic systems at the expense of number of populations sampled, whereas we examined fewer marker loci in exchange for increasing the number of population samples and hence the geographic representation of the continent. It is important to note that Salzano and Callegari-Jacques (1988) do not find widespread evidence of geographic structure in their data. Rather, they observe geographic structure for only a few select allele frequencies. Thus the basic inference of minimal geographic structure in allele frequency among South Amerindian groups appears to be essentially concordant across different analytical approaches.

The demonstrated lack of correspondence between geographic and genetic distances in the South American data is of considerable theoretical interest and import. Population genetic theory predicts a decline in genetic identity with increasing geographic distance (Wright 1943; Malécot 1969). Such isolation by distance is not apparent in the South American data, although it is apparent for the same allele frequencies in North America and Central America. Why should such a situation obtain in only one geographic region?

Several possible scenarios can be imagined. First, colonization of the South American continent may be sufficiently recent that geographic structure in gene frequencies simply has not had sufficient time to develop, as it has in North America. This hypothesis has little to recommend it because archeological data indicate an ancient presence on the continent and there is no evidence of a more recent replacement migration. Second, migration among South American groups may be low such that extreme differentiation occurs between adjacent groups and drift results in similarity between unrelated populations at great distances. Although plausible and consistent with our earlier conclusion that on a con-

tinental scale South America appears to be a collection of populations drifting independently (O'Rourke and Suarez 1985), this scenario is problematic as well. There is little evidence that migration has been attenuated among South Amerindians and considerable evidence that travel along river systems has, in fact, enhanced communication between groups. Indeed, it is surprising that comparatively recent protohistoric migrations (e.g., the Quechua out of the Andean region and Tupi speakers in southeastern South America) have not resulted in detectable clines reflecting those movements.

Nevertheless, isolation and resultant drift may be relevant to the observed pattern. If the peopling of the New World is viewed as proceeding from north to south, the geographic constriction that is Central America would have become relatively more densely populated early in the southward migration of colonizing peoples. Such an increase in population in Central America may have resulted in a blockage to further migrants entering South America from the north (R.H. Ward, personal communication, 1991). Thus drift operating on those few small migrant groups to South America before the blockage may have contributed to the spatial heterogeneity observed in modern populations of South Amerindians.

Recent evidence suggests that extreme genetic diversity between lineages in founding populations may be expected (Ward et al. 1991). The postulated small number of founders for Native America [e.g., Wallace and Torroni (this issue) and Schurr et al. (1990)] may be more appropriate for the colonizers of South America than for the New World in general [cf. Ward et al. (1991)]. This scenario of early Central American population density and subsequent differentiation coupled with the linearity of the Central American landmass would be consistent with the association among allele frequencies, geography, and linguistic affinity among the Chibchan speakers reported by Barrantes et al. (1990). As noted earlier, the degree of differentiation between populations is approximately equivalent throughout North America and South America but slightly depressed for the Central American groups, as indicated by the F_{ST} values in Table 1. This is not particularly unexpected, given the smaller geographic area of Central America, particularly in the context of an early increase in population density.

One further character that may influence gene frequency patterns must also be mentioned. A number of researchers have noted the relationship between genetic variation and linguistic affinity [Spuhler 1979; Crawford and Enciso 1982; Salzano and Callegari-Jacques 1988; Cavalli-Sforza et al. 1988; Black 1991; but see Black et al. (1983)]. Given the tremendous linguistic diversity in native South America, the role of language on gene frequency variation may be substantial. In a preliminary analysis Mobarry (1991) found that discriminant function analysis of these

South Amerindian gene frequencies results in nearly 85% of populations being correctly classified by language family. Moreover, this high percentage of correct classification is not improved and is usually eroded by inclusion of additional variables, such as geographic location or measures of local ecology [cf. O'Rourke and Suarez (1985)]. Such preliminary results suggest a fair correspondence between gene frequency variation and language in these data.

Finally, it is worth noting that only red cell antigen frequencies have been treated here. Geographic patterning in the GM, KM, or highly polymorphic HLA systems might lead to alternative inferences. This emphasizes the genetic diversity in Native American populations, the complexity of the evolutionary mechanisms that structure this variation, and the need for increasingly detailed work to clarify the action of alternative evolutionary scenarios. It is likely that the power of the emerging molecular markers for resolving questions of anthropological and evolutionary interest will lead the way in resolving a number of long-standing problems in this area. Identifying a common set of molecular markers informative in Native American groups that can be assessed in all samples will hasten progress in understanding the evolutionary genetics of Amerindian populations.

Acknowledgments We are pleased to acknowledge the computing assistance and expertise of Alan Lichty. Lynn Jorde kindly shared software for obtaining great circle distances, and Alan Rogers provided stimulating discussion, pertinent advice, and programming assistance. Spatial analyses and the Mantel tests were performed using the R Package for Multivariate Data Analysis, developed and provided by Pierre Legendre. We are grateful for their contributions. This work was supported in part by a Faculty Development Grant from the University of Utah.

Received 3 September 1991; revision received 6 November 1991.

Literature Cited

Ananthakrishnan, R., and H. Walter. 1972. Some notes on the geographical distribution of the human red cell acid phosphatase phenotypes. *Humangenetik* 15:177–181.

Barbujani, G. 1987. Autocorrelation of gene frequencies under isolation by distance. *Genetics* 117:777–782.

Barnhill, R.E. 1977. Representation and approximation of surfaces. In *Mathematical Software III,* J.R. Rice, ed. New York: Academic Press, 69–120.

Barnhill, R.E. 1983. A survey of the representation and design of surfaces. *IEEE Comp. Graph. Appl.* 3:9–16.

Barrantes, R., P.E. Smouse, H.W. Mohrenweiser, H. Gershowitz, J. Azofeifa, T.D. Arias, and J.V. Neel. 1990. Microevolution in lower Central America: Genetic characterization of the Chibcha-speaking groups of Costa Rica and Panama, and a consensus taxonomy based on genetic and linguistic affinity. *Am. J. Hum. Genet.* 46:63–84.

Bateman, R., I. Goddard, R. O'Grady, V.A. Funk, R. Mooi, W.J. Kress, and P. Cannell. 1990. Speaking of forked tongues: The feasibility of reconciling human phylogeny and the history of language. *Curr. Anthropol.* 31:1–24.

Black, F.L. 1980. HLA antigens in South American Indians. *Tissue Antigens* 16:368–376.

Black, F.L. 1984. Interrelationships between Amerindian tribes of Lower Amazonia as manifest by HLA haplotype disequilibria. *Am. J. Hum. Genet.* 36:1318–1331.

Black, F.L. 1991. Reasons for failure of genetic classifications of South Amerind populations. *Hum. Biol.* 63:763–774.

Black, F.L., F.M. Salzano, L.L. Berman, Y. Gabbay, Y.A. Weimer, M.H.L.P. Franco, and J.P. Pandy. 1983. Failure of linguistic relationships to predict genetic distances between the Waiapi and other tribes of Lower Amazonia. *Am. J. Phys. Anthropol.* 60:327–335.

Cann, R.L., M. Stoneking, and A.C. Wilson. 1987. Mitochondrial DNA and human evolution. *Nature* 325:31–36.

Cavalli-Sforza, L.L., A. Piazza, P. Menozzi, and J. Mountain. 1988. Reconstruction of human evolution: Bringing together genetic, archaeological, and linguistic data. *Proc. Natl. Acad. Sci. USA* 85:6002–6006.

Cliff, A.D., and J.K. Ord. 1981. *Spatial Processes*. London: Pion.

Crawford, M.H., and V. Bach Enciso. 1982. Population structure of circumpolar groups of Siberia, Alaska, Canada, and Greenland. In *Current Developments in Anthropological Genetics*, vol. 2, *Ecology and Population Structure*, M.H. Crawford and J.H. Mielke, eds. New York: Plenum, 51–91.

Di Rienzo, A., and A.C. Wilson. 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 88:1597–1601.

Kidd, J.R., F.L. Black, K.M. Weiss, I. Balazs, and K.K. Kidd. 1991. Studies of three Amerindian populations using nuclear DNA polymorphisms. *Hum. Biol.* 63:775–794.

Legendre, P., and M.-J. Fortin. 1989. Spatial pattern and ecological analysis. *Vegetatio* 80:107–138.

Malecot, G. 1969. *The Mathematics of Heredity*. New York: W.H. Freeman.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27:209–220.

Matson, G.A. 1938. Blood groups and ageusia in Indians of Montana and Alberta. *Am. J. Phys. Anthropol.* 24:81–89.

Matson, G.A., and H.F. Schrader. 1933. Blood grouping among the "Blackfeet" and "Blood" tribes of American Indians. *J. Immunol.* 25:15–163.

Matson, G.A., P. Levine, and H.F. Schrader. 1936. Distribution of the sub-groups of A and the M and N agglutinogens among the Blackfeet Indians. *Proc. Soc. Exp. Biol.* 35:46–47.

Menozzi, P., A. Piazza, and L.L. Cavalli-Sforza. 1978. Synthetic maps of human gene frequencies in Europeans. *Science* 201:786–792.

Mobarry, A. 1991. Genetic classification and linguistic affinity in Native America. Unpublished.

Nei, M. 1985. *Molecular Evolutionary Genetics*. New York: Columbia University Press.

Nichols, J. 1990. Comment in "Discussion and Criticism." *Curr. Anthropol.* 31:313–314.

O'Rourke, D.H., B.K. Suarez, and J.D. Crouse. 1985. Genetic variation in North Amerindian populations: Covariance with climate. *Am. J. Phys. Anthropol.* 67:241–250.

O'Rourke, D.H., and B.K. Suarez. 1985. Patterns and correlates of genetic variation in South Amerindians. *Ann. Hum. Biol.* 13:13–31.

O'Rourke, D.H., A.S. Lichty, and B.K. Suarez. 1986. The geography of gene frequencies in American Indians. *Am. J. Phys. Anthropol.* 69:249.

O'Rourke, D.H., and A.S. Lichty. 1989. Spatial analysis and gene frequency maps of Native North American populations. *Coll. Antropol.* 13:73–84.

Pääbo, S. 1986. Molecular genetic investigations of ancient human remains. *Cold Spring Harbor Symp. Quant. Biol.* 51:441–446.

Pääbo, S., J.A. Gifford, and A.C. Wilson. 1988. Mitochondrial DNA sequences from a 7000-year old brain. *Nucleic Acids Res.* 16:9775–9787.

Piazza, A., P. Menozzi, and L.L. Cavalli-Sforza. 1981a. The making and testing of geographic gene frequency maps. *Biometrics* 37:635–659.

Piazza, A., P. Menozzi, and L.L. Cavalli-Sforza. 1981b. Synthetic gene frequency maps of man and selective effects of climate. *Proc. Natl. Acad. Sci. USA* 78:2638–2642.

Salzano, F.M., and S.M. Callegari-Jacques. 1988. *South American Indians: A Case Study in Evolution*. Oxford: Clarendon Press.

Schurr, T.G., S.W. Ballinger, Y.-Y. Gan, J.A. Hodge, D.A. Merriwether, D.N. Lawrence, and D.C. Wallace. 1990. Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. *Am. J. Hum. Genet.* 46:613–623.

Shearin, N.L., E.J. King, and D.H. O'Rourke. 1989. DNA preservation in Precolumbian remains from the American Southwest. *Hum. Evol.* 4:263–270.

Sokal, R.R., and N.L. Oden. 1978a. Spatial autocorrelation in biology 1: Methodology. *Biol. J. Linn. Soc.* 10:199–228.

Sokal, R.R., and N.L. Oden. 1978b. Spatial autocorrelation in biology 2: Some biological implications and four applications of evolutionary and ecological interest. *Biol. J. Linn. Soc.* 10:229–249.

Sokal, R.R., and D.E. Wartenberg. 1983. A test of spatial autocorrelation analysis using an isolation-by-distance model. *Genetics* 105:219–237.

Sokal, R.R., P.E. Smouse, and J.V. Neel. 1986. The genetic structure of a tribal population, the Yanomama Indians. XV. Patterns inferred by autocorrelation analysis. *Genetics* 114:259–287.

Spuhler, J.N. 1979. Genetic distances, trees, and maps of North American Indians. In *The First Americans: Origins, Affinities and Adaptations,* W.S. Laughlin and A.B. Harper, eds. New York: Gustav Fisher, 135–183.

Suarez, B.K., J.D. Crouse, and D.H. O'Rourke. 1985. Genetic variation in North Amerindian populations: The geography of gene frequencies. *Am. J. Phys. Anthropol.* 67:217–232.

Suarez, B.K., D.H. O'Rourke, and J.D. Crouse. 1985. Genetic variation in North Amerindian populations: Association with cultural complexity. *Am. J. Phys. Anthropol.* 67:233–239.

Szathmary, E.J.E. 1984. Peopling of northern North America: Clues from genetic studies. *Acta Antropogen.* 8:79–109.

Ward, R.H., B.S. Frazier, K. Dew, and S. Pääbo. 1991. Extensive mitochondrial diversity within a single Amerindian tribe. *Proc. Natl. Acad. Sci. USA* 88:8720–8729.

Wartenberg, D. 1985. Multivariate spatial correlation: A method for exploratory geographical analysis. *Geogr. Anal.* 17:263–283.

Wiley, E.O., A. Comuzzie, and M. Bamshad. 1990. Comment in "Discussion and Criticism." *Curr. Anthropol.* 31:314–315.

Wright, S. 1943. Isolation by distance. *Genetics* 28:114–138.

Zegura, S., W.H. Walker, K.K. Stout, and J.D. Diamond. 1990. Comment in "Discussion and Criticism." *Curr. Anthropol.* 31:420–426.