advanced generation crosses are present (Nason and Ellstrand 1993).

The maximum-likelihood method of Nason and Ellstrand (1993) has limitations (Epifanio and Philipp 1997). First, genealogical class frequencies in populations, rather than individual class membership, are estimated. For example, any one individual in genotype category A could be a P_1, BP_1, or F_2. Therefore the common question, "Is this individual unhybridized, past or present?" is not answered by this method. Only those individuals of the F_2 class that fall into category S can be classified with certainty. Second, the model cannot accommodate crosses beyond the F_2 generation. Individuals from advanced generations will have multilocus genotypes that place them in one of the six categories, thereby biasing estimates of genealogical class frequencies. An extreme example of this is continuous backcrosses to the same parent (e.g., taxon A). By the fifth backcross (BC-5), more than 66% of individuals will likely be placed into parent category A using 13 codominant loci or dominant loci L_B (Boecklen and Howard 1997). The remaining individuals will almost certainly be placed into AI. These authors showed that more than 70 markers are needed before the probability of assigning a BC-5 to a hybrid class would exceed .95, although under this framework it would be incorrectly considered a first-generation backcross (BP_1).

Despite these limitations, maximum-likelihood methods for estimating genealogical origins in mixed hybrid populations should still be useful. If hybridization is a recent event, for example, due to population transfers (e.g., stocking), or if successful hybridization is restricted to the first or second generation of interbreeding (Avise 1994 and references therein), then the model of Nason and Ellstrand (1993) is applicable. Furthermore, Nason and Ellstrand (1993) showed that their model will often suggest the presence of advanced-generation hybrids by producing impossible estimates (i.e., frequencies <0 or >1.0) for some classes. In this case, class frequency estimates will be biased but advanced-generation hybrids will be detected, which may often be the goal of the research. The potential to significantly increase the number of diagnostic loci using dominant marker techniques will increase the power of maximum-likelihood methods and should advance the use of molecular data in hybrid studies.

From the Department of Fisheries and Wildlife, University of Minnesota, 1980 Folwell Ave., St. Paul, MN 55108.

References

Avise JC, 1994. Molecular markers, natural history, and evolution. New York: Chapman & Hall.

Avise JC and van den Avyle MJ, 1984. Genetic analysis of reproduction of hybrid white bass × striped bass in Savannah River. Trans Am Fish Soc 113:563–568.

Beismann H, Barker JHA, Karp A, and Speck T, 1997. AFLP analysis sheds light on distribution of two *Salix* species and their hybrid along a natural gradient. Mol Ecol 6:989–993.

Boecklen WJ and Howard DJ, 1997. Genetic analysis of hybrid zones: numbers of markers and power of resolution. Ecology 78:2611–2616.

Campton DE, 1990. Application of biochemical and molecular markers to analysis of hybridization. In: Electrophoretic and isoelectric focusing techniques in fisheries management (Whitmore DH, ed). Boca Raton, FL: CRC Press; 241–264.

Crawford DJ, Brauner S, Cosner MB, and Stuessy TF, 1993. Use of RAPD markers to document the origin of the intergeneric hybrid *Margyracaena skottsbergi* (Rosaceae) on the Juan Fernandez Islands. Am J Bot 80:89–92.

Epifanio JM and Philipp DP, 1997. Sources for misclassifying genealogical origins in mixed hybrid populations. J Hered 88:62–65.

Fritsch P and Rieseberg LH, 1996. The use of random amplified polymorphic DNA (RAPD) in conservation genetics. In: Molecular genetic approaches in conservation (Smith TB and Wayne RK, eds). New York: Oxford University Press; 54–73.

Greene BA and Seeb JE, 1997. SINE and transposon sequences generate high-resolution DNA fingerprints, "SINE prints," that exhibit faithful Mendelian inheritance in pink salmon (*Oncorhynchus gorbuscha*). Mol Mar Biol Biotechnol 6:328–338.

Gupta M, Chyi Y-S, Romero-Severson J, and Owen JL, 1994. Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. Theor Appl Genet 89:998–1006.

Nason JD and Ellstrand NC, 1993. Estimating the frequencies of genetically distinct classes of individuals in hybridized populations. J Hered 84:1–12.

Philipp DP, Childers WF, and Whitt GS, 1983. A biochemical genetic evaluation of northern and Florida subspecies of largemouth bass. Trans Am Fish Soc 112: 1–20.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, and Zabeau M, 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414.

Welsh J and McClelland M, 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res 18:7213–7218.

Williams DJ, Kazianis S, and Walter RB, 1998. Use of random amplified polymorphic DNA (RAPD) for identification of largemouth bass subspecies and their intergrades. Trans Am Fish Soc 127:825–832.

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, and Tingey SV, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535.

Wolfe AD, Xiang Q-Y, and Kephart SR, 1998. Assessing hybridization in natural populations of *Penstemon* (Scrophulariaceae) using hypervariable intersimple sequence repeat (ISSR) bands. Mol Ecol 7:1107–1125.

Zietkiewicz E, Rafalski A, and Labuda D, 1994. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20:176–183.

Is There Really Natural Selection Affecting the *I* Frequencies (Long Hair) in the Brazilian Cat Populations?

M. Ruiz-Garcia

The scientific literature on cat genetics contains a presumed typical example of natural selection affecting *I* frequencies (long hair) in 16 Brazilian cat populations. It has been observed that the hotter and more tropical the climate in Brazil, the lower the values of *I* frequencies in the cat populations. Nevertheless, this study of some new cat populations in Latin America showed that all of them, independent of the climate, had high or very high *I* frequencies. I postulate that an alternative migrational-historical hypothesis exists that explains the correlation between the *I* frequencies and climate characteristics (which are correlated with the latitude) without using natural selection explanations concerning the appearance of the *I* allele in Brazil.

Seven of 12 genes coding for coat characters such as color, tabby, and length, as well as certain skeletal anomalies have been studied in many domestic cat populations (*Felis catus*) worldwide (e.g., Ahmad et al. 1980; Lloyd 1985; Lloyd and Todd 1989; Ruiz-Garcia 1991, 1994, 1997b; Ruiz-Garcia et al. 1995, 1998, 1999). These loci have significant spatial patterns in different areas of the world (Ruiz-Garcia 1994, 1997b), showing a close relationship to historical and commercial human migrations (Ruiz-Garcia and Alvarez 1996; Todd 1977). Artificial human selection has been shown to have had little influence on the genetic profiles of the stray cat populations studied in different cities. Clark (1975) showed, for instance, that the hu-

man taste for specific mutant colors was not reflected in the genetic profile of the cat population of Glasgow. Nevertheless, it has been particularly difficult to establish exact geographical patterns of and influences on one gene, l (long hair). Todd et al. (1974) put forward that " . . . populations which are unarguably related in other aspects, show great disparity in their l frequencies." At first glance one would think that this allele would be favorably selected for in cold climates, and conversely would be negatively selected for in hot climates. It would therefore correspond to our expectations to find high frequencies of l in populations in very cold climates, such as Leningrad [$q(l) = 0.64$] and Alma Ata [$q(l) = 0.56$] in the former Soviet Union, or Inverness [$q(l) = 0.52$] in Scotland, for example. It would be paradoxical, however, that populations in places such as Cyprus [$q(l) = 0.50$], Jericho, Israel [$q(l) = 0.42$], and Phoenix, Arizona [$q(l) = 0.51$], with very high recorded temperatures, would have high frequencies of l. They are, in fact, much higher than other populations with notably colder climates, such as Poznan [$q(l) = 0$], Bialowieza [$q(l) = 0.21$], and Wroclaw [$q(l) = 0$] in Poland or in Iceland [$q(l) = 0.14$]. This has motivated the study of this characteristic in different areas of the world by several investigators. Lloyd (1983, 1985), for example, showed that a significantly negative correlation existed between the average minimum temperatures in 35 populations from the Atlantic coast of North America during January (winter) and the frequencies of l ($r = -0.44$). However, there was no significant correlation ($r = -0.08$) with respect to the average maximum temperature in July (summer). Still, the most famous example where the existence of natural selection affecting l frequencies was postulated was that described by Watanabe (1984) in Brazil. The author observed that the hotter and more tropical the climate, the lower the values of $q(l)$ for 16 cat populations studied in Brazil ($r = -0.95$), concluding that l was strongly unfavorable in tropical climates. This finding has been recognized as a clear example of natural selection affecting a genetic character for a morphological trait in cat populations [see, e.g., Klein (1993), Lloyd (1987), Lloyd and Todd (1989), and Todd and Lloyd (1984), among others]. However, the analysis of other Latin American cat populations has allowed me to postulate an alternative hypothesis that would explain the correlation between $q(l)$ and climate characteristics according to lati-

Table 1. Some mutant allele frequencies of European (Spain, Portugal, and Italy), Hispanic settlements in the United States, and Latin American cat populations (of Spanish and Portuguese origin)

Populations	Locus							
	n	O	a	t^b	d	l	S	W
Spain								
Barcelona (1989)	709	0.16	0.70	0.27	0.27	0.14	0.27	0.004
Granollers	158	0.17	0.71	0.16	0.08	0.21	0.23	0.003
Girona	159	0.22	0.67	0.24	0.12	0.15	0.29	0.003
L'Estartit	100	0.19	0.81	0.23	0.00	0.15	0.31	0.000
Llansa	75	0.26	0.74	0.15	0.24	0.00	0.20	0.000
R Castell (1989)	319	0.24	0.80	0.17	0.22	0.15	0.28	0.007
R Castell (1994)	228	0.20	0.78	0.26	0.26	0.13	0.27	0.009
Sitges	204	0.13	0.74	0.32	0.25	0.20	0.25	0.006
Vilanova	118	0.20	0.72	0.22	0.23	0.09	0.35	0.005
Tarragona	216	0.21	0.68	0.38	0.15	0.00	0.29	0.000
Benidorm (1990)	207	0.23	0.70	0.55	0.16	0.00	0.28	0.005
Alicante	335	0.23	0.78	0.50	0.24	0.00	0.31	0.006
Murcia	171	0.18	0.74	0.62	0.08	0.08	0.38	0.012
Cadiz	64	0.16	0.71	0.45	0.25	0.00	0.23	0.010
Mahon (Balearics)	475	0.30	0.80	0.18	0.38	0.12	0.14	0.003
Villacarlos (Balearics)	226	0.24	0.75	0.13	0.44	0.17	0.18	0.002
Ciudadela (Balearics)	510	0.21	0.73	0.22	0.35	0.08	0.27	0.004
P.Majorca (Balearics)	475	0.20	0.72	0.36	0.35	0.27	0.23	0.004
Ibiza (Balearics)	273	0.24	0.76	0.30	0.23	0.19	0.29	0.010
Vigo	228	0.23	0.80	0.36	0.14	0.19	0.35	0.004
Santiago Compostela	120	0.24	0.75	0.40	0.16	0.16	0.37	0.000
Tenerife (Canary)	146	0.29	0.80	0.32	0.34	0.25	0.41	0.003
Puerto Cruz (Canary)	126	0.16	0.83	0.33	0.38	0.13	0.32	0.000
Portugal and Islands								
Lisbon	371	0.07	0.65	0.45	0.27	0.09	0.21	0.020
Porto	256	0.14	0.77	0.43	0.30	0.23	0.29	0.010
Terceira (Azores)	160	0.31	0.66	0.37	0.34	0.16	0.45	0.030
Faial/Pico (Azores)	109	0.25	0.57	0.38	0.34	0.17	0.34	0.010
S. Miguel (Azores)	159	0.23	0.63	0.50	0.34	0.11	0.29	0.000
Madeira	133	0.18	0.78	0.35	0.41	0.00	0.39	0.010
Mindelo (C Verdes)	206	0.12	0.77	0.60	0.19	0.00	0.34	0.000
Praia (C Verdes)	168	0.23	0.83	0.47	0.11	0.00	0.44	0.010
Italy								
Rome	480	0.09	0.66	0.49	0.34	0.10	0.31	0.010
Venice (1991)	145	0.10	0.56	0.27	0.34	0.19	0.20	0.014
San Remo	148	0.04	0.58	0.48	0.32	0.29	0.27	0.000
Rimini	518	0.13	0.68	0.38	0.41	0.22	0.26	0.012
Riccione	130	0.11	0.59	0.44	0.42	0.27	0.28	0.008
Hispanic settlements in the United States								
Denver (Colorado)	286	0.20	0.84	0.26	0.38	0.35	0.29	0.010
Lubbock (Texas)	265	0.31	0.79	0.36	0.33	0.44	0.24	0.000
Dallas (Texas)	311	0.25	0.67	0.27	0.32	0.44	0.18	0.000
Denton (Texas)	311	0.25	0.81	0.27	0.33	0.46	0.22	0.010
Mineral Wells (Texas)	311	0.31	0.73	0.34	0.29	0.52	0.20	0.000
Houston (Texas)	294	0.25	0.69	0.29	0.29	0.35	0.19	0.000
Richmond (California)	107	0.19	0.77	0.27	0.33	0.35	0.28	0.030
San Francisco (California)	195	0.27	0.79	0.33	0.32	0.36	0.31	0.030
Humboldt County (California)	238	0.27	0.75	0.51	0.35	0.30	0.22	0.030
Hispanic settlements in Latin America								
Los Mochis (Mexico)	141	0.31	0.71	0.39	0.24	0.31	0.32	0.010
Mexico City (Mexico)	170	0.16	0.62	0.23	0.29	0.57	0.29	0.020
Caracas (Venezuela)	164	0.13	0.79	0.32	0.10	0.33	0.33	0.009
Willemstadt (Curacao)	151	0.14	0.84	0.19	0.19	0.28	0.34	0.020
Havana (Cuba)	334	0.30	0.72	0.24	0.14	0.62	0.39	0.022
Bogota (Colombia)	1105	0.19	0.86	0.18	0.35	0.34	0.21	0.009
Ibague (Colombia)	147	0.24	0.82	0.11	0.37	0.32	0.23	0.003
Bucaramanga (Colombia)	240	0.16	0.87	0.11	0.32	0.29	0.24	0.002
Cali (Colombia)	258	0.23	0.84	0.18	0.47	0.33	0.32	0.005
Pasto (Colombia)	210	0.20	0.82	0.10	0.37	0.41	0.29	0.000
Santiago (Chile)	126	0.13	0.76	0.42	0.51	0.59	0.33	0.036
Buenos Aires (1992)	295	0.27	0.82	0.31	0.45	0.40	0.28	0.021
Buenos Aires (1996)	675	0.21	0.79	0.29	0.43	0.41	0.29	0.016

tude, leaving aside the possible existence of selection in Brazil.

Materials and Methods

In order to demonstrate the absence of natural selection affecting the locus L in cats in Brazil, the allele frequencies of seven loci controlling fur color, tabby, and length were studied in the cities of La Havana ($n = 334$) in Cuba; Bogotá ($n = 1105$), Ibagué ($n = 147$), Bucaramanga ($n = 240$), Cali ($n = 257$), and Pasto ($n = 210$) in Colombia; Santiago ($n = 126$) in Chile;

Table 1. Continued

Populations	Locus							
	n	*O*	*a*	*t*^b	*d*	*l*	*S*	*W*
Brazil								
Porto Alegre	489	0.16	0.68	0.26	0.28	0.27	0.31	0.020
Curitiba	327	0.17	0.71	0.25	0.21	0.27	0.35	0.020
Sao Paulo	1164	0.22	0.71	0.26	0.18	0.35	0.42	0.030
Rio de Janeiro (1984)	1545	0.24	0.74	0.32	0.27	0.20	0.38	0.030
Rio de Janeiro (1996)	232	0.14	0.74	0.23	0.30	0.31	0.37	0.012
Belho Horizonte	859	0.18	0.71	0.26	0.25	0.22	0.36	0.020
Campo Grande	512	0.19	0.64	0.27	0.20	0.24	0.29	0.020
Brasilia	492	0.26	0.63	0.28	0.19	0.18	0.31	0.010
Cuiaba	302	0.16	0.66	0.20	0.12	0.10	0.29	0.040
Salvador	959	0.20	0.57	0.42	0.14	0.16	0.47	0.020
Rio Branco	235	0.25	0.49	0.29	0.06	0.06	0.34	0.000
J. Norte	503	0.33	0.60	0.24	0.21	0.05	0.39	0.000
Teresina	994	0.18	0.69	0.33	0.26	0.01	0.35	0.030
Fortaleza	1254	0.20	0.66	0.27	0.29	0.08	0.50	0.020
S. Luis	1323	0.24	0.66	0.19	0.24	0.00	0.53	0.010
Manaus	993	0.18	0.68	0.32	0.06	0.06	0.47	0.030
Belem	909	0.22	0.65	0.25	0.14	0.05	0.42	0.010

a new sample of cats in Buenos Aires (n = 675) in Argentina; a new sample in Rio de Janeiro (n = 232) in Brazil; and two nonoverlapping samples in the Canary Islands (one from Lloyd 1989, unpublished, in Tenerife City, n = 146; one from Puerto de la Cruz, Tenerife Island, n = 126). Each population was extensively sampled to minimize local effects that could cause deviations in the allele frequencies. The cats sampled were alley cats, or "pseudowild." Previously reported samples that were included in this analysis were from Rio de Janeiro (Watanabe 1984) and Buenos Aires (Kajon et al. 1992). The phenotypes of the individuals were recorded from direct observation. The genetic nomenclature used is in accordance with the Committee on Standardized Genetic Nomenclature for Cats (1968). The genetic characteristics studied included the sex-linked gene [O, o; Orange (epistatic to the observation of the A locus) versus non-orange] and the nonlinked autosomal loci: A [A, a; agouti versus non-agouti (epistatic to the observation of the T locus)], T (t^+, t^b, T^a, striped or mackerel tabby versus blotched tabby versus Abyssinian tabby), D (D, d; nondilution versus dilution), L (L, l; short hair versus long hair), S (S, s; piebald white spotting versus non-white spotting), and W [W, w; dominant white (epistatic to all the other colors) versus normal color]. For the characteristics of these genes see Robinson (1977). The frequency of the allele orange was calculated using a differential equation (Ahmad et al. 1980). The autosomic recessive frequencies (q) were calculated as the square roots of the observed phenotypic frequencies, while the dominant frequencies (p) were taken as 1 − q.

The genetic relationships between Brazilian populations reported by Watanabe (1981, 1984) were analyzed in this study, matching them against other populations. The populations were grouped in two ways: (1) The first group consisted of 40 cat populations including the Latin American (both of Spanish and Portuguese origins), southwestern United States, and two Canary Island populations. (2) A second group included 80 populations, among them the 40 populations of the group just described as well as a group of North American cat populations from the United States and Canada (reported by Lloyd and Todd 1989) of probable British origin, and 20 European populations from the countries of origin of all these American populations (Kajon et al. 1992; Lloyd and Todd 1989; Ruiz-Garcia, 1990a–d, 1993, 1994, 1997b). In order to analyze the relationship between these cat populations, two kinds of analysis were carried out. The first was to obtain matrices of genetic distances between pairs of populations. The genetic distances used were four: the Nei standard genetic distance (Nei 1978), the Cavalli-Sforza and Edwards (1967) chord distance, the distance of Prevosti (1974), and the DA distance (Nei et al. 1983). With these matrices, different dendrograms were constructed in order to explain the overall genetic relationships between all of these American and European populations. The algorithms used were the UPGMA (Sneath and Sokal 1973), WPGMA (using the recommendation of Pamilo 1990), COMPLETE, and neighbor-joining (Saitou and Nei 1987). To determine the reliability of the trees generated, three statistical methods were applied: the interior branch and Rzhestsky and Nei (1992) tests (Li 1989), the Felsenstein (1985) bootstrap test, and the cophenetic correlation coefficient (Sneath and Sokal 1973). The trees that showed the best statistics for reliability are shown here. All of the analyses were performed both including and excluding the locus L to determine its influence on the relationships found between the populations studied. The second analysis was a canonical analysis of populations. This separates groups of populations along axes of high discrimination power using the Mahalanobis square distance, and is based on the fulfillment of two hypotheses: (1) that there is homogeneity between all covariance matrices corresponding to the population groups (maximum likelihood test), and (2) that the means of the k groups are significantly different [Wilks' Λ test, and the associate value of the Fisher–Snedecor F test by means of the approximation of Rao (1951)]. Subsequently, a canonical transformation, the eigenvalues, the significance of the first canonical axes with the Bartlett's test, and the radius of the confidence regions (for a 90% level) were calculated. In this canonical population analysis the following groups were employed: (1) Buenos Aires (two samples), (2) Mexico (three samples), (3) Venezuela and Curacao (two samples), (4) southern Brazil (six samples), (5) northern Brazil (10 samples), (6) Colombia (five samples), (7) Canary Islands (two samples), and separately, La Havana, Santiago, and the second sample from Rio de Janeiro.

Results and Discussion

Table 1 and Figures 1–3 show the basis for the apparent association between the frequencies of l and climatic factors in Brazil without having to resort to the explanation of selection such as has been offered up until now. Figure 1a shows the analysis based on the application of the WPGMA algorithm with Cavalli-Sforza and Edwards chord distance, including the L locus for 40 populations. Figure 1b shows the tree derived from the neighbor-joining method with DA distance, including the L locus for 40 populations. Figure 1c shows the result from the COMPLETE algorithm with Cavalli-Sforza and Edwards chord distance, without the L locus for 40 populations. They show clearly that upon analyzing the genetic relationships between the Latin
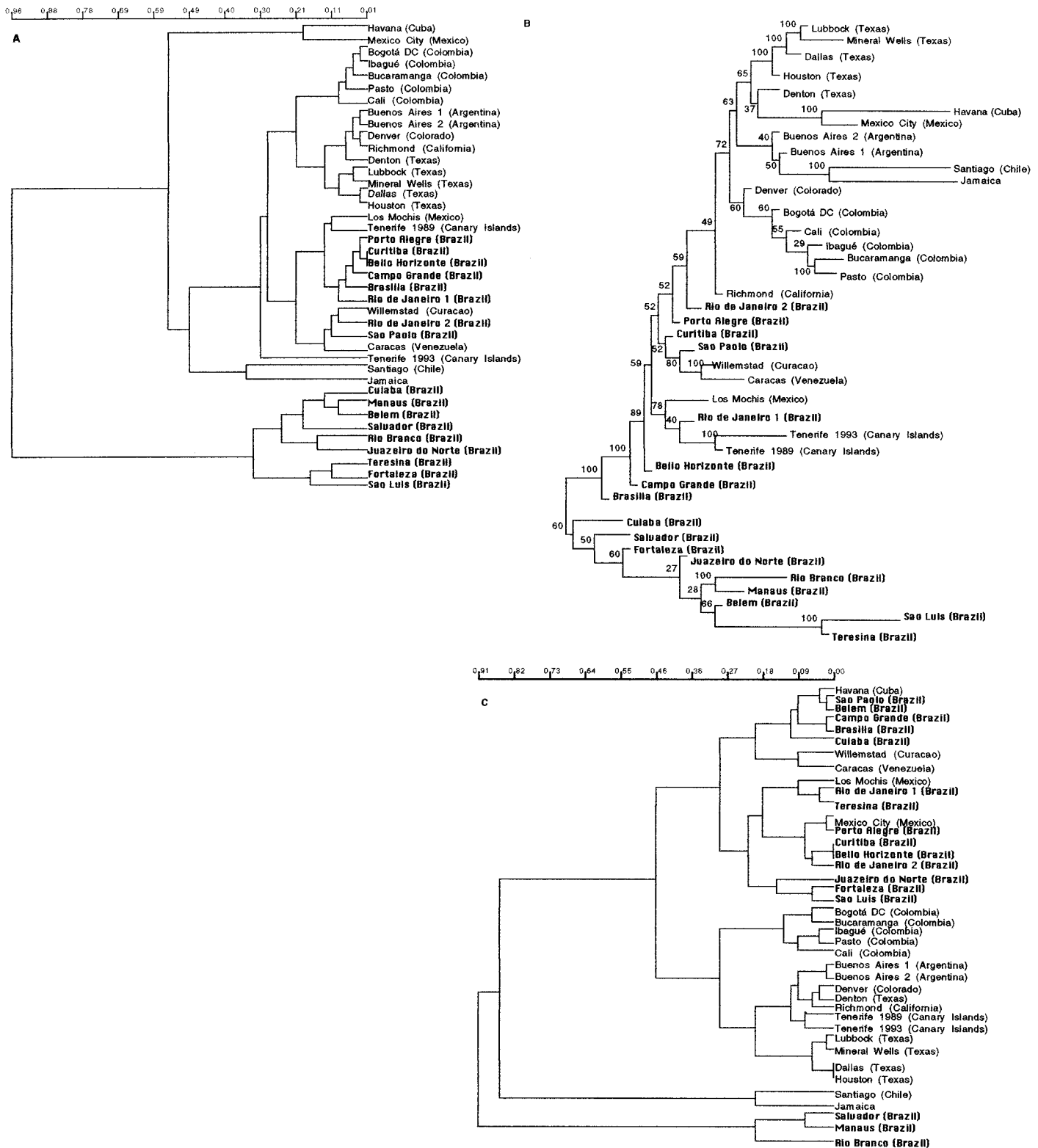
Figure 1. **(a)** WPGMA phenetic analysis from Cavalli-Sforza and Edwards (1967) chord distance of Hispanic settlements in the United States, Hispanic America, two Canary Island, and Brazilian cat populations (40 populations) with the inclusion of the *L* locus. Cophenetic correlation coefficient, $r = 0.71$; approximate Mantel *t* test: $t = 8.79$, $P <$ .0000; out of 1,000 random permutations: one-tail probability is p[random $Z >$ observed Z] = 0.001. **(b)** Neighbor-joining tree with DA distance (Nei et al. 1983) of 40 populations with the inclusion of the *L* locus. The numbers in the figure are the bootstrap (1,000) percentages. **(c)** COMPLETE phenetic analysis from Cavalli-Sforza and Edwards (1967) chord distance of 40 populations without the inclusion of the *L* locus. Cophenetic correlation coefficient, $r = 0.69$; approximate Mantel *t* test: $t = 7.17$, $P <$ .0000; out of 1,000 random permutations: one-tail probability is p[random $Z >$ observed Z] = 0.001. The dendrograms shown are those with the better cophenetic correlation coefficients, better percentages of Felsenstein bootstraps, and better Rzhestsky and Nei (1992) statistics.
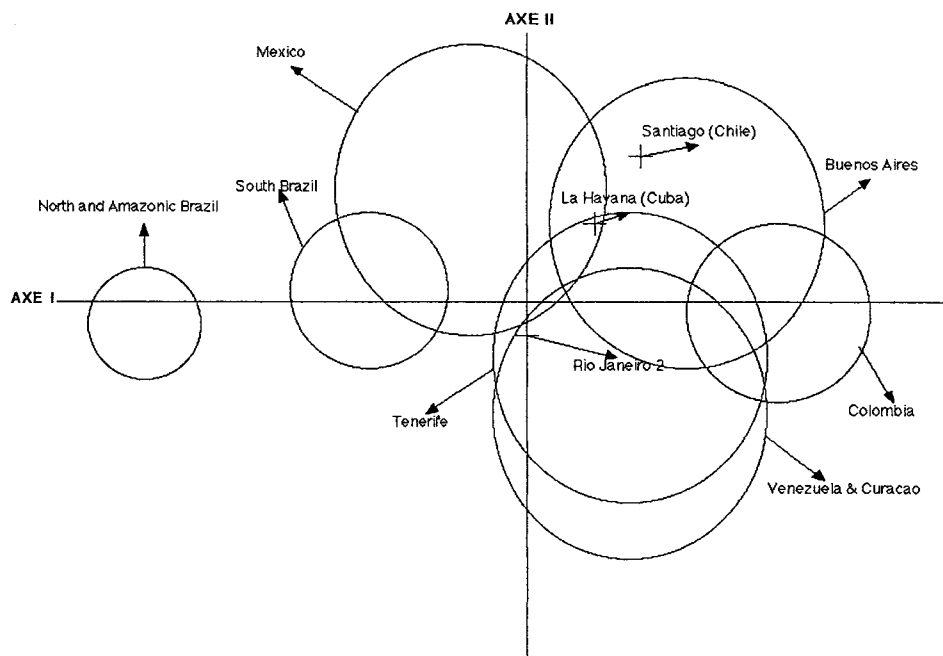
Figure 2. Canonical analysis of populations. The first and second axes explain 90.09% of the variation.

American cat populations, the Brazilian populations do not form a homogeneous group. A group of Brazilian populations was observed which showed more of a genetic similarity to the Hispanic populations, such as Caracas (Venezuela), Willemstadt (Curacao), and Los Mochis (Mexico), than to the other group of Brazilian populations. A second group had marked differences from the first, and had no particular resemblance to the Hispanic populations. The composition of these two groups of Brazilian populations is not geographically random, however. The group resembling the Hispanic populations is located in southern Brazil (Porto Alegre, Curitiba, Sao Paulo, and Rio de Janeiro). These populations have the highest *l* frequencies in Brazil (0.22–0.35). On the contrary, the other group of Brazilian populations was made up of coastal populations from northern Brazil and extended inland into the Amazon regions, namely Salvador, Sao Luis, Rio Branco, J. Norte, Teresina, Fortaleza, Manaus, and Belem, and is characterized by very low or null *l* frequencies (0–0.16). Populations from Campo Grande, Brasilia, Bello Horizonte, and Cuiaba are found to cluster differently with both groups with regard to the algorithmic techniques and genetic distances employed.

The same analysis excluding the *L* locus shows a similar perspective, although it is slightly less clear. The northern Brazilian populations (Salvador, Manaus, Rio Bran-co, Sao Luis, Fortaleza, and J. Norte) are less related to the Hispanic American populations than are the southern Brazilian populations (particularly Porto Alegre, Curitiba, and the two samples from Rio de Janeiro). A canonical analysis of populations is shown in Figure 2. This analysis showed a Wilks $\Lambda = 0.0002$ and $F = 4.63$ with 70 and 100 df, being $F = 1.43$ with $\alpha = 0.05$. Consequently the hypothesis that the representative group means are equal was rejected as expected. The two first canonical axes explained 90.09% of the variability. All the Hispanic American groups were clearly related. The southern Brazilian group and the new sample from Rio de Janeiro were also highly related to the Hispanic American groups. To the contrary, the northern Brazilian and Amazon groups were isolated from the other groups analyzed.

The main point which allows a new nonselectionist hypothesis of an historic-migrational character is the following. Ruiz-Garcia (1997a,c) and Ruiz-Garcia et al. (1998, 1999) observed that a constant feature of the Hispanic American cat populations was the high, or very high, *l* frequencies [e.g., Bogotá (0.34), Buenos Aires (0.41), Mexico City (0.57), Santiago, Chile (0.59), and La Havana (0.62)]. These values are much higher than those found in Spanish populations, mostly with values of $q(l)$ between 0 and 0.20 (Ruiz-García 1990c,d, 1991, 1994, 1997b). Since in the entire geographic range of the His-

panic American cat population, from California, Colorado, and Texas, to Argentina and Chile (more than 7000 km), the climatic characteristics are extremely diverse, such high and constant $q(l)$ values cannot be attributed to the action of natural selection, either in favor or against, after the formation of the original Hispanic American populations. Neither could genetic drift explain the systematic occurrence of such high *l* frequencies in all of the Hispanic American populations. The Spaniards have a great admiration for long-haired cats, a trait with low frequencies in Spain (Ruiz-García M, unpublished observations). It is quite likely that during the period when the Americas were being colonized, the *l* frequencies were lower still. A migrational selection could therefore have occurred due to the novelty of this character (Todd 1977, 1978). This hypothesis is much more parsimonious than the existence of human selection a posteriori at a time when the populations in question in Latin America had become important from a demographic point of view. As shown by Anderson and Jenkins (1979), Morrill and Todd (1978), and Ruiz-García (1991), once the human population of a locality approaches 30,000, the cat population may be large enough to be refractory against changes in the allele frequencies.

The populations from southern Brazil resembled the Hispanic populations more than those of the rest of Brazil. This could have originated from the establishment of a gene flow of a certain magnitude between those Brazilian populations and populations of Spanish origin near the southern frontier of Brazil. The commerce between southern Brazil and the Hispanic colonies in Uruguay, Paraguay, and the area of Rio de La Plata in Argentina was very intense from 1713 onward. In fact, the Portuguese succeeded in setting up some commercial colonies on the Rio de La Plata. For instance, the first explorer in current Paraguay was the Portuguese Alejo García in 1525 searching for the "Silver Mountain." Later the area was conquered by the Spaniards Juan de Salazar y Espinosa, Alvaro Nuñez de Vaca, and Domingo Martinez de Irala. Juan de Salazar founded Asunción in 1537, and coming from Buenos Aires, Alvaro Nuñez and Domingo Martínez founded small colonies in the south of today's Brazil. In 1588 the Spaniard Jesuits founded numerous "Missions" or "Reducciones," where they congregated hundreds of Guaraní Indian families and lodged Spaniard colonists. They trav-
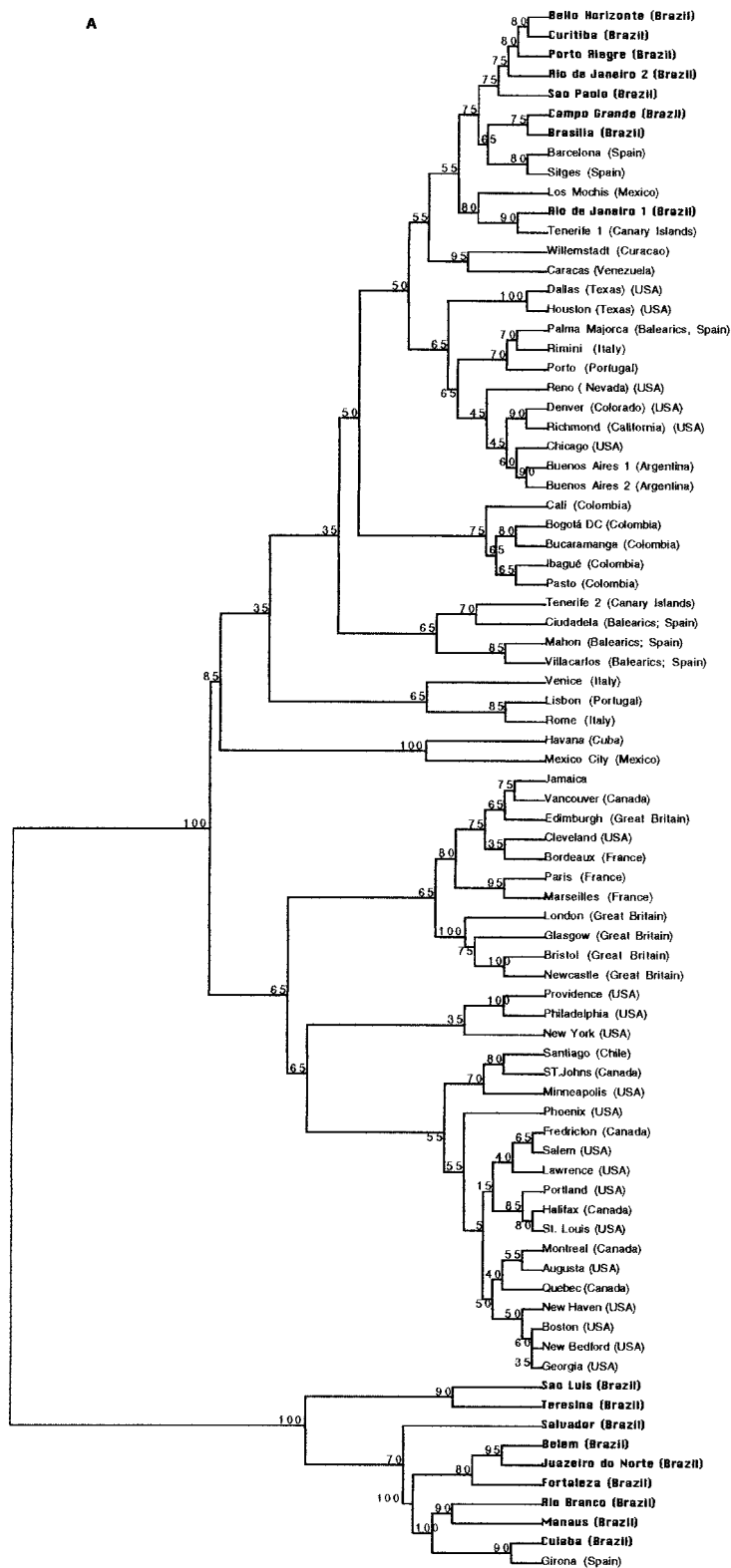
A

Figure 3. **(a)** UPGMA phenetic analysis with DA distance with 80 populations with the inclusion of the *L* locus. Cophenetic correlation coefficient, $r = 0.62$; approximate Mantel *t* test: $t = 32.66$, $P < .0000$; out of 1,000 random permutations: one-tail is p[random $Z >$ observed Z] = 0.001. The numbers in the figure are the bootstrap (1,000) percentages. **(b)** WPGMA phenetic analysis from Prevosti (1974) genetic distance of 80 populations without the inclusion of the *L* locus. Cophenetic correlation coefficient, $r = 0.68$; approximate Mantel *t* test: $t = 25.81$, $P < .0000$; out of 1,000 random permutations: one-tail is p[random $Z >$ observed Z] = 0.001. The dendrograms shown are those which offered better cophenetic correlation coefficients, better Felsenstein bootstraps, and better Rzhestsky and Nei (1992) statistics.

eled throughout the Paraguay and Paraná rivers as far as the frontier of southern Brazil. The tremendous ability of the Guaraní Indians in various jobs encouraged the Portuguese to enslave entire populations, which were made to travel into southern Brazil (1690–1767). Also, numbers of poor Portuguese colonists in Paraguay and Uruguay followed the Spanish Missions along their way.

When the Jesuits were expelled from the Spanish Empire in 1767 the majority of Guaraní Indian families, poor Portuguese colonists, and one-sixth of the Spanish colonists in Paraguay emigrated to the state of Sao Paulo (southern Brazil) to work on the rice and maize plantations and herding cattle. In Uruguay there was also a relationship between the Spanish and the Portuguese populations. The first settlements in Uruguay were established by the Portuguese in Colonia in 1680. Later Spain established another colony in this area, Montevideo in 1726, and the area was finally controlled by the Spanish in 1777. In 1811 and 1816, the Portuguese invaded a large portion of Uruguay and relations between this country and southern Brazil have been very important ever since. There was another Hispanic–Portuguese connection in southern Brazil: In 1680–1710, a small and as yet undeveloped harbor, Rio de Janeiro, started to gain importance. The finding of gold was of key importance to this little harbor, and many Spanish ships arrived from the Canary Islands. By 1763 Rio de Janeiro had grown into a very important city. The Canary Islands were and remain a key part of the commercial routes between Spain and Latin America. During the last decades of the century, more than 25,000 ships sailed to these islands.

One more point is very important to better explain the Spanish relationship with Rio de Janeiro. In 1650 Potosí (Bolivia) was the second largest city of the Western World (160,000 inhabitants), after London. The reason was the mining of silver. The Spaniards arrived in Potosí following this route: southern Spain–Canary Islands–Dominican Republic–Panamá (Puerto de Dios, Porto Belo, and Panamá City)–Guayaquil (Ecuador)–Perú (Callao and Lima). Once the silver had been loaded in Potosí, one of the most important routes back to Spain was through Bolivia (previously called Alto Perú)–northern Argentinean deserts–the Argentinean cities of Salta, Jujuy, Córdoba, Tucumán, La Plata (in Spanish, silver) and finally, to the Port of Buenos Aires. All of these cities were founded
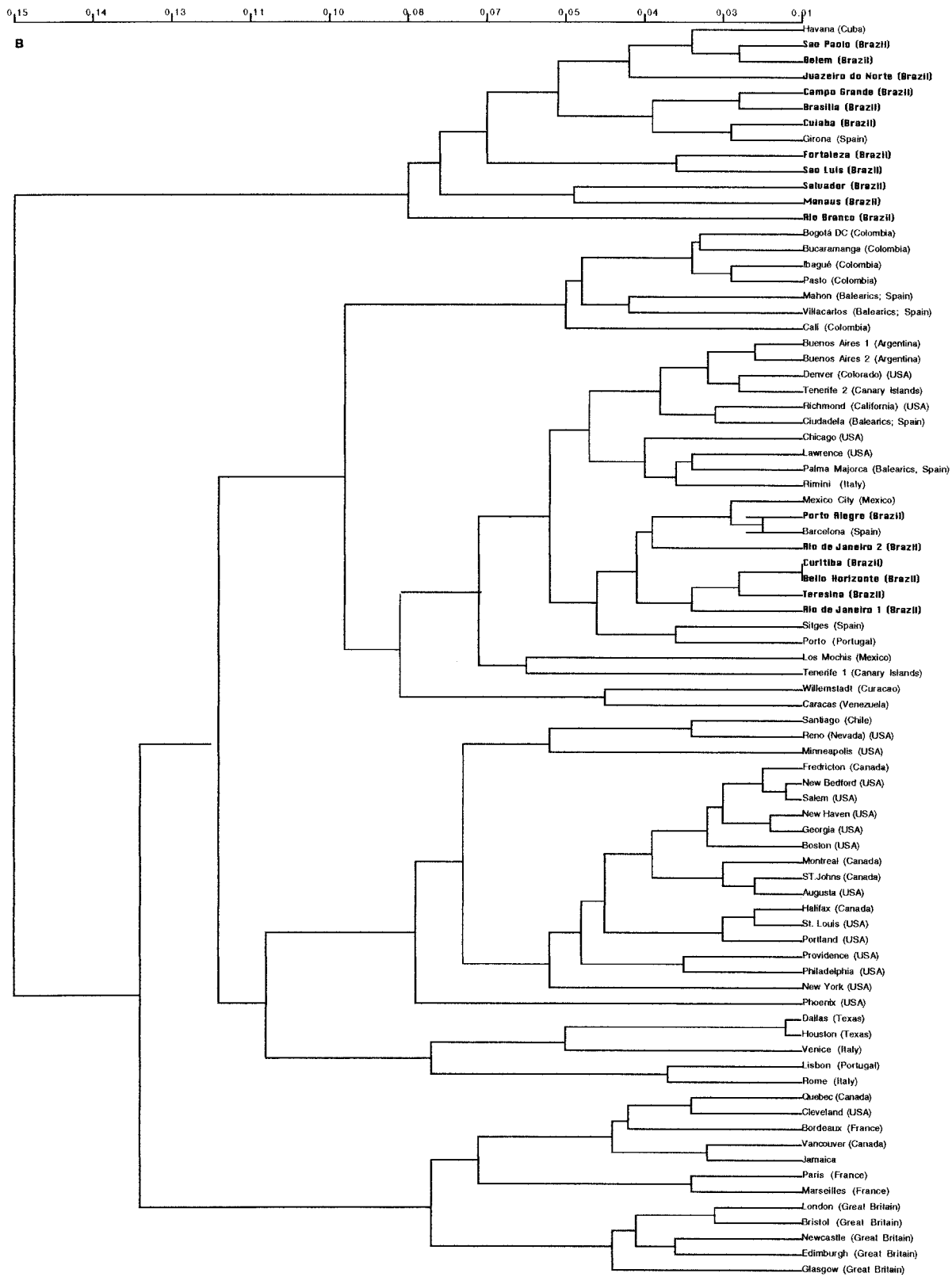
Figure 3. Continued.

during the years of the silver trade. Many of the ships carrying silver sailed from Buenos Aires to Spain via Rio de Janeiro (Fernández de Oviedo 1944; Perrottet 1990). It could have been that this intense contact with Hispanic populations, where the frequencies of *l* were high was generated by migrational selection, based on novelty, from their original populations in Spain, promoting an increase in the frequencies of *l* in the populations of southern Brazil.

In northern Brazil the process of colonization was, in contrast, very different. Discovered by the Portuguese Pedro Alvares Cabral in 1500, the most important populations founded were located in northern Brazil, with Salvador de Bahía being the first capital of Brazil (one of the cat populations studied). The important crop behind this development was sugar, for which the shoots for the plantations had been brought from the Madeira Islands (African coast). This activity was sustained for nearly two centuries, during which southern Brazil remained forgotten. The populations of the north coast and inland Brazil were not affected by the Hispanic influence, since their $q(l)$ frequencies are null or very low. These Brazilian populations are therefore probably constituted almost exclusively of Atlantic Portuguese influences, coming from the Islands of Madeira, Cape Verde, and the Azores, where the *l* frequencies are null or very low. Alternatively they could have their origin in some continental Portuguese populations which have not yet been studied.

The tree displayed in Figure 3a shows the analysis based on the application of the UPGMA algorithm with DA distance, including the *L* locus, for 80 populations, and Figure 3b shows the corresponding WPGMA algorithm with Prevosti distance, excluding the *L* locus, for the 80 populations, revealing that some Spanish populations were mainly clustered with the populations from southern Brazil, especially some Catalan populations such as Barcelona, Sitges, and Girona, which had previously been postulated to be Spanish populations with probable ancient genetic profiles (Kajon et al. 1992; Ruiz-Garcia 1988, 1990a–d, 1991, 1994). The southern Brazilian populations are therefore more similar to certain Spanish populations than to recently sampled Portuguese populations such as Lisbon and Porto, confirming the proposed hypothesis.

From the historical and genetic evidence gathered so far, the apparent correlation between the frequencies of *l* and geographic latitude in Brazil in fact reveals a geographical correlation with the migrational and historical factors that were involved in the formation of the colonial settlements in Brazil. This correlation there-fore does not result from the occurrence of natural selection, but rather from the particular location of the migrational and historical relationships, which by chance, in this case, coincide with the location, according to latitude, of the Brazilian populations and their genetic constitution.

One more group of results supports the conclusion that natural selection does not necessarily act against the *l* allele in tropical populations: data for the *l* allele in the cat population of La Havana (Cuba), a place of tropical climate, is one of the highest worldwide (0.62). Also, data from other recently sampled tropical Caribbean populations, such as Santo Domingo [Dominican Republic, $q(l) = 0.534$]; Veracruz [Mexican Carib, $q(l) = 0.472$], and Acapulco [Mexico, $q(l) = 0.441$], have $q(l)$ values greater than 0.40, which could not be expected under the selection hypothesis (Ruiz-Garcia M, unpublished data). Additional studies on the subject would, however, be necessary in order to explain why there is apparently no migrational selection favoring *l* in the case of the Portuguese, as seems to have occurred for the Spanish migration. It could have been that the *l* allele was not systematically present in some Portuguese cities which were the sources of cats during the colonization of the Americas. Another possibility would be that in Portugal there were at least two different gene pools, one very similar or identical to certain Spanish populations postulated to be ancient, and which were the origin of southern Brazilian populations studied, and another(s) of different characteristics, which might have been the origin of certain Portuguese islander and northern and Amazonian Brazilian cat populations. Only a thorough study of the Portuguese cat populations would allow a more precise historical hypothesis.

I conclude that the *l* allele can in fact be favorably selected for in cold or very cold climates (Lloyd 1983, 1985), but its behavior is neutral in temperate, hot, or very hot climates, such as the tropics.

From Unidad Genética (Genética de Poblaciones-Biología Evolutiva), Departamento de Biología, Pontificia Universidad Javeriana, Cra 7a No 43-82, Bogotá DC, Colombia, and CIGEEM, Barcelona, Spain. The author

References

Ahmad M, Blumenberg B, and Chaudhary MF, 1980. Mutant allele frequencies and genetic distance in cat populations of Pakistan and Asia. J Hered 71:323–330.

Anderson MM and Jenkins SH, 1979. Gene frequencies in the domestic cats of Reno, Nevada: confirmation of a recent hypothesis. J Hered 70:267–269.

Cavalli-Sforza LL and Edwards AWF, 1967. Phylogenetic analysis: models and estimation procedures. Evolution 21:550–570.

Clark JM, 1975. The effects of selection and human preference on coat colour gene frequencies in urban cats. Heredity 35:195–210.

Committee on Standardized Genetic Nomenclature for Cats, 1968. Standardized genetic nomenclature for the domestic cat. J Hered 59:39–40.

Felsenstein J, 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791.

Fernández de Oviedo J, 1944. Historia general y natural de las Indias. Ed. Guaranía. Asunción.

Kajon A, Centron D, and Ruiz-Garcia M, 1992. Gene frequencies in the cat population of Buenos Aires, Argentina, and the possible origin of this population. J Hered 83:148–152.

Klein K, 1993. Population genetics and gene geography. In: Genetikii koshkii (Ruvinki A and Borodin PM, eds). Moscow: Russian Academy of Science; 118–150.

Li WH, 1989. A statistical test of phylogenies estimated from sequence data. Mol Biol Evol 6:424–435.

Lloyd AT, 1983. Population genetics of domestic cats (*Felis catus* L.) in New England and the Canadian Maritime Provinces: an investigation of the historical immigration hypothesis (PhD dissertation). Boston: Boston University.

Lloyd AT, 1985. Geographic distribution of mutant alleles in domestic cat populations of New England and the Canadian Maritimes. J Biogeog 12:315–322.

Lloyd AT, 1987. Cats from history and history from cats. Endeavour 11:112–115.

Lloyd AT and Todd NB, 1989. Domestic cat gene frequencies. A catalogue and bibliography. Newcastle upon Tyne: Tetrahedron Publications.

Morrill RB and Todd NB, 1978. Mutant allele frequencies in the domestic cats of Denver, Colorado. J Hered 69:131–134.

Nei M, 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590.

Nei M, Tajima F, and Tateno Y, 1983. Accuracy of estimated phylogenetic trees from molecular data. Gene frequency data. J Mol Evol 19:153–170.

Pamilo P, 1990. Statistical tests of phenograms based on genetic distances. Evolution 44:689–697.

Perrottet T, 1990. South America. Singapore: Hofer Press.

Prevosti A, 1974. La distancia genética entre poblaciones. Misc. Alcobé. Publicacions de l'Universitat de Barcelona; 109–118.

Rao CR, 1951. Advanced statistical methods in biometric research. Darien, CT: Hafner Publishing.

Robinson R, 1977. Genetic for cat breeders. Oxford: Pergamon Press.

Ruiz-Garcia M, 1988. Frecuencias alélicas mutantes en una población de gatos domésticos urbanos (Barcelona) y en una población de gatos rurales (Castelldefels rural) en Cataluña, España. Genét Ibér 40:157–187.

Ruiz-Garcia M, 1990a. Frecuencias alélicas en la población de gatos domésticos de la isla de Menorca (Baleares): diferentes modelos de evolución colonizadora. Evol Biol 4: 307–342.

Ruiz-Garcia M, 1990b. Frecuencias alélicas en la población de gatos domésticos de Palma de Mallorca e Ibiza y relaciones genéticas con otras poblaciones de gatos Europeos y Norteafricanos. Evol Biol 4:189–216.

Ruiz-Garcia M, 1990c. Mutant allele frequencies in domestic cat populations in Catalonia, Spain, and genetic relationships between Spanish and English colonial cat populations. Genetica 82:209–214.

Ruiz-Garcia M, 1990d. Mutant allele frequencies in domestic cat populations on the Spanish Mediterranean coast and genetic distances from other European and North African cat populations. Genetica 82:215–221.

Ruiz-Garcia M, 1991. Más sobre la genética de poblaciones de Felis catus en la costa Mediterránea Española: un análisis de la estructura genética de las poblaciones naturales de gatos. Evol Biol 5:227–283.

Ruiz-Garcia M, 1993. Analysis of the evolution and genetic diversity within and between Balearic and Iberian cat populations. J Hered 84:173–180.

Ruiz-Garcia M, 1994. Genetic profiles from coat genes of natural Balearic cat populations: an eastern Mediterranean and North African origin. Genet Sel Evol 26: 39–64.

Ruiz-Garcia M, 1997a. Características genéticas de las poblaciones de gatos domésticos (Felis catus) en las Américas. I. Estructura espacial inducida por colonización. Brazil J Biol (in press).

Ruiz-Garcia M, 1997b. Genetic relationships among some new cat populations sampled in Europe: a spatial autocorrelation analysis. J Genet 76:1–24.

Ruiz-Garcia M, 1997c. Perfiles genéticos de las poblaciones de gatos domésticos de La Habana (Cuba) y de Bogotá (Colombia) y posibles origenes Europeos de esas poblaciones. I: no existencia de paralelismo con el modelo colonizador británico. Misc Zool (in press).

Ruiz-Garcia M and Alvarez D, 1996. The use of the domestic cat as an extragenic marker of the historical and commercial human movements. Brazil J Genet 19:184.

Ruiz-Garcia M, Barrera MI, and Alvarez D, 1998. Genetic relationships between Caribbean and South-American cat populations: confirmation of the "Historical Migration Hypothesis." Genet Sel Evol (submitted).

Ruiz-Garcia M, Campos HA, Alvarez D, Kajon A, and Diaz S, 1999. Mutant allele frequencies of cat populations in Latin America, Havana (Cuba), Bogotá (Colombia), Ibagué (Colombia), Santiago (Chile) and Buenos Aires (Argentina): genetic relationships with other American and European cat populations. J Hered (submitted).

Ruiz-Garcia M, Ruiz S, and Alvarez D, 1995. Perfiles genéticos de poblaciones de gatos domésticos (Felis catus) de la provincia de Girona (Catalunya, NE, España) y posibles relaciones genéticas con otras poblaciones europeas occidentales. Misc Zool 18:169–196.

Rzhestsky A and Nei M, 1992. A simple method for estimating and testing minimum-evolution trees. Mol Biol Evol 9:945–967.

Saitou N and Nei M, 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425.

Sneath PH and Sokal RR, 1973. Numerical taxonomy. San Francisco: W. H. Freeman.

Todd NB, 1977. Cats and commerce. Sci Am 237:100–107.

Todd NB, 1978. An ecological, behavioural genetic model for the domestication of the cat. Carnivore 1:52–60.

Todd NB and Lloyd AT, 1984. Mutant allele frequencies in the domestic cats of Portugal and the Azores. J Hered 75:495–497.

Todd NB, Robinson R, and Clark JM, 1974. Gene frequencies in domestic cats of Greece. J Hered 65:227–231.

Watanabe MA, 1981. Mutant allele frequencies in the domestic cats of Sao Paulo, Brasil. Carnivore Genet Newslett 4:168–177.

Watanabe MA, 1984. Estudo populacional da cor de pelagem de gato domestico (Felis catus L.) em dezesseis localidades do Brasil (PhD dissertation). Sao Paulo: Universidade de Sao Paulo.

Corresponding Editor: Stephen J. O'Brien

Improved Estimation of the Proportion of Triploids in Populations with Diploid and Triploid Individuals

M. S. Ridout

We consider the estimation of the proportion of triploids in populations of plants or animals in which diploid and triploid individuals coexist, using data from electrophoretic analysis of isozyme or microsatellite markers. Individuals that have three distinct alleles at a locus are unambiguously triploid. However, other individuals cannot be classified with certainty as diploid or triploid, unless allelic dosage can be determined reliably. This is impossible for microsatellite markers, and for many isozyme markers. We therefore present a maximum likelihood method of estimating the proportion of triploids based only on the presence or absence of different alleles.

Populations that include both diploid and triploid individuals occur in various plant and animal species. Estimation of the proportion of triploid individuals within such mixed populations is important for population genetic studies and for ecological studies (Krieger and Keller 1998). One approach to this problem is to use phenotypic data from electrophoretic analysis of isozyme or microsatellite markers. These techniques identify the distinct alleles that are present. However, determination of allelic dosage is generally impossible with microsatellite markers. Dosage can sometimes be determined for isozyme markers, but is often unreliable. Clearly an individual that has three distinct alleles at a locus is necessarily triploid. However, if dosage cannot be determined reliably, an individual that has only two distinct alleles may be diploid, or it may be a triploid that has two identical alleles. Similarly, an individual that has only one distinct allele may be either diploid or triploid. Thus unambiguous classification of individuals as diploid or triploid is impossible in general. Krieger and Keller (1998) showed that it is possible, nonetheless, to estimate the proportion of triploids from phenotypic data, provided that the same alleles occur with the same frequencies in diploids and triploids and provided that genotype frequencies are in Hardy–Weinberg equilibrium.

In this article we demonstrate that there are advantages in using likelihood methods to address this problem. We show that the maximum likelihood estimate of the proportion of triploids is almost always preferable to the estimator of Krieger and Keller (1998). Another advantage of the likelihood method is that it leads naturally to a test of one of the key assumptions, namely that the alleles occur at the same frequency in diploids and triploids.

Triploidy Estimates

Single Locus with Three Alleles

We consider a locus with three alleles (A, B, and C) in a mixed population of diploids and triploids in which the proportion of triploids is t. The allele frequencies, a, b, and c, are assumed to be the same for diploids and triploids and genotype frequencies are assumed to be related to allele frequencies according to the Hardy–Weinberg law. There are seven phenotypic classes, as shown in Table 1 [identical to Table 1 in Krieger and Keller (1998)]. ABC individuals are necessarily triploid and may therefore be termed "overt triploids." Other individuals may be diploid or triploid.

The method of Krieger and Keller (1998) involves the following steps

1. Estimate allele frequencies a, b, and c on the assumption that all individuals, apart from the overt triploids, are diploid.

2. Estimate t by equating the observed number of overt triploids (n_7) to the expected number, based on the current estimates of a, b, and c.

3. Calculate revised estimates a, b, and c, based on the current estimate of t.

Steps 2 and 3 are alternated until the estimates converge, or until the estimate of t exceeds one, in which case the estimate is taken as one. Krieger and Keller (1998) show that convergence is usually achieved in a few iterations.

An alternative method of estimating t is by maximum likelihood. Given a sample of N independent individuals, the distribution of the numbers of individuals in the