Ethnicity, Gene Flow, and Population Subdivision in Limón, Costa Rica

L. Madrigal,[1]* B. Ware,[1] R. Miller,[1] G. Saenz,[2] M. Chavez,[2] and D. Dykes[3]

[1]*Department of Anthropology. University of South Florida. Tampa, Florida 33620*
[2]*Investigation Center of Abnormal Hemoglobins and Related Diseases, Universidad de Costa Rica, San José, Costa Rica*
[3]*Minneapolis War Memorial Blood Bank, Minneapolis, Minnesota*

KEY WORDS population structure; F_{st} statistics; heterozygosity; ethnicity

ABSTRACT In this paper we examine the effects of ethnicity on the gene flow between two groups living in Limón, Costa Rica. Our main interest is to determine if ethnicity has acted as a barrier to the exchange of genes, and if the groups have remained distinct genetically. We report the admixture estimates, F_{st} values, and inbreeding coefficients of the two samples. The data consist of blood samples and surnames obtained from 375 individuals. The subjects' two surnames were analyzed to determine the ethnicity of their parents (individuals carry their father's and mother's first surnames). We used the formula of Crow and Mange ([1965] Eugen Q 12:199–203) to compute F_t, F_n, and F_r with the surnames. Admixture estimates were computed for both groups using the computer program ADMIX.PAS kindly provided by Jeffrey Long. The estimates for the Hispanic-Limonense group are M1 = 0.5866 European, M2 = 0.3383 Amerindian, and M3 = 0.0751 African ancestry. For the Afro-Limonense group, the admixture estimates indicate M1 = 0.1047 European, M2 = 0.1357 Amerindian, and M3 = 0.7595 African ancestry. The F_{st} values are F_{st} = 0.00558 for the Hispanic group and F_{st} = 0.05137 for the Afro-Limonense group. These F_{st} values indicate that the Afro-Limonense group has experienced more genetic drift than has the other group, possibly as a result of its long history of isolation in Costa Rica. Indeed, when plotted along a scaled eigenvector R matrix of Caribbean gene frequencies, the two Limonense groups did not cluster with each other. Thus we conclude that the two ethnic groups have remained distinct breeding populations. Am J Phys Anthropol 114:99–108, 2001.

The study of ethnic or cultural barriers to gene flow has a long history in biological anthropology. Many studies give credence to the importance of language, cultural differences, and geographic distance as barriers to gene flow, and as determinants of genetic differentiation. Generally, research has focused on the genetic differences among several ethnic groups located over a large geographical area, e.g., in Italy (Barbujani and Sokal, 1991a,b; Barbujani et al., 1992), the Caucasus (Barbujani et al, 1994a,b), India (Sirajuddin et al, 1994), South America (Aguiar, 1992; Rickards et al., 1994), Central America (Barrantes et al., 1990), Hungary (Guglielmo and Beres, 1996), and Bougainville Island (Friedlaender, 1975; Relethford, 1985).

A review of the literature however, reveals a smaller record of research on cultural barriers to gene flow between ethnic groups living in the same settlement. Stevenson et al. (1983) document genetic differences among four linguistically defined ethnic groups residing in Milwaukee, indicating that these groups are relatively endogamous. An earlier paper (Tavares-Neto and Azevedo, 1978) reports significant differences in ABO blood-group frequencies among whites, mulattos, and blacks in Bahia, Brazil, indicating a degree of correspondence between genetic makeup and ethnicity. More recently, several studies investigated how religious differences and geographic distance affect the genetic structure of Ireland (Bittles and Smith, 1994; Relethford and Crawford, 1998; Smith et al., 1990).

It is possible to envision the effect of ethnicity on gene frequencies along a continuum between two extremes. Ethnicity could act as a total barrier, resulting in two distinct breeding populations, or it could be a nonsignificant cultural issue, resulting in one actual breeding population. In his review of the genetic structure of subdivided populations, Jorde (1980) noted that linguistic and cultural differences do provide significant barriers to gene flow.

Grant sponsor: Sigma Xi Grant-in-Aid; Grant sponsor: Pierre Stouse Memorial Fellowship; Grant sponsor: University of Kansas Summer Fellowship; Grant sponsor: Investigation Center of Abnormal Hemoglobins and Related Diseases, University of Costa Rica; Grant sponsor: Costa Rican Ministry of Health; Grant sponsor: Tony Facio Hospital, Limón. D.D. passed away after time of data anyalysis. See Crawford (1991) for an obituary.

*Correspondence to: L. Madrigal, Department of Anthropology, University of South Florida, Tampa, FL 33620.
E-mail: madrigal@luna.cas.usf.edu

Received 16 May 2000; accepted 2 October 2000.

The purpose of this paper is to determine with genetic and surname data if two groups which coexist in one settlement form one or two breeding populations. The two groups are also placed in the context of gene frequencies in the Caribbean. This will clarify if the two ethnicities cluster together (presumably as a result of frequent gene flow) or if they cluster with other groups with which they share a longer evolutionary history. The two ethnic groups have easily identified cultures: Afro-Limonenses speak English, have English names, are mostly Protestants, and have a distinct, Caribbean-oriented cuisine. Hispanic Limonenses speak Spanish, have Spanish names, are mostly Catholics, and have a Spanish-oriented cuisine (Chomsky, 1995). The population of Puerto Limón, therefore, is an excellent site to study ethnicity, as discussed by Bogin (1993), i.e., as a line defining the interaction between two groups of people who are all part of a larger group. Since interethnic matings are reported to be frequent (Madrigal, 1988; Purcell, 1993), the Limón population also provides a good opportunity to measure gene flow between two groups that proclaim to be ethnically distinct, but which engage in frequent mating.

MATERIALS AND METHODS

The population

Limón is the easternmost province of Costa Rica, stretching from the northern national border with Nicaragua, to the southern border with Panama. Limón City, or Puerto Limón, is the capital of the province, with approximately 60,000 inhabitants in the city and its suburbs (Purcell, 1993). Research for this paper was conducted in Limón City and in five adjacent suburbs: Westfalia, Cuba Creek, Matina, Zent, and Moin.

Limón is a distinct cultural center in Costa Rica as a result of its Afro-Caribbean population. This group is a product of the immigration of mostly Jamaican workers beginning in the late 1800s to Limón. The laborers were brought to work for the construction of a railroad from the Central Valley to Limón (Bryce-Laporte, 1962; Casey, 1979; Chomsky, 1995; Duncan, 1972; Stewart, 1967; Melendez, 1972). Having a language, religion, and other cultural traits different from those of the rest of Costa Rica, and having been subjected to racist and discriminatory practices and laws, the Jamaican migrants remained isolated from the wider population. In fact, before the 1920s, few Costa Ricans would venture to the Limón area, mostly in fear of the high incidence of malaria. Certainly, the genetic makeup of the Afro-Limonenses allowed them to survive and succeed in a most difficult tropical ecology, specifically, under endemic malarial conditions (Madrigal, 1989).

Only with increasing in-migration of the larger, Hispanic population starting in the 1920s, did the migrants begin to interact more widely with the rest of the country (Purcell, 1993). The migration of the Hispanic population out of the Central Valley to the Atlantic coast was a reflection of what Guillén (1989) refers to as an increasing proletarization process of the Costa Rican work force. Thus, Costa Rican society started undergoing a process of social differentiation during the first decades of the 20th century, a process that resulted in an outmigration of landless workers from the main population centers towards the coasts (Guillén, 1989). The 1948 National Constitution signaled the beginning of the incorporation of the Afro-Caribbean population into the nation. This constitution legally recognized the once-Jamaicans as Costa Rican citizens, and incorporated laws against racial discrimination (Fernandez and Mendez, 1973). Starting that year, there was an influx of Afro-Limonenses to the cities of the Central Valley, in search of better jobs and education.

The migration of Hispanic Costa Ricans to Limón has only increased in the recent past. According to Purcell (1993), although Costa Rica does not use ethnic categories in its national statistics, a 1973 estimate of the population in the Limón province indicated that there were 49% white (Hispanic), 46% black, 3% Amerindian, and 2% Chinese. However, by 1986, blacks were reported to constitute less than 25% of the Limón population (Purcell, 1993).

With increased migration of Hispanic Costa Ricans to Limón , the proportion of interethnic marriages has also increased. No specific statistics are available on interethnic marriage, since Costa Rica does not use ethnic categories. However, when Purcell (1993) performed his survey of 218 households, 6.5% of all unions were mixed, with 45.2% of respondents approving of such unions. The percentage of interethnic matings, however, is likely to be much higher than that of interethnic marriages. The family structure of Afro-Limonenses has been characterized as having unstable male-female unions, and a relatively high proportion of out-of-wedlock births. Thus, while the national average of couples living in common-law unions in 1982 was 4.6%, it was 12.2% in Limón (Chavez, 1982). Madrigal (1988) also observed very frequent unions between members of the two ethnic groups.

Data

Blood samples and interviews were obtained from a total of 400 individuals at the Limón hospital, and in door-to door visits. The samples were analyzed for hemoglobin and G-6PD phenotypes at the Investigation Center of Abnormal Hemoglobins and Related Diseases (CIHATA) of the University of Costa Rica (see Madrigal, 1989 for laboratory methods). The samples were also centrifuged and preserved at the Limón Hospital, and shipped weekly for analyzes at the Minneapolis War Memorial Blood Bank in Minneapolis. These samples were analyzed for the following systems: GCI, Bf, ACP, PGM1, and PLG (see Dykes et al., 1983 for laboratory methods). Unfortunately, some samples were lost during shipping to the USA, so only 375 were available for genetic analysis.

During the interview, subjects were asked to ethnically classify themselves. The only two terms chosen by the subjects were black (245 individuals) and white (145 individuals). The former will be referred to as Afro-Limonenses (AL) and the latter as Hispanic-Limonenses (HL). Other information collected in the interview included basic data such as place of birth, number of siblings, number of offspring produced, and, for females, reproductive history (Madrigal, 1994, 1995).

Surname methodology

Our study on genetic differences between the two ethnic groups of Limón can be enriched by investigating the language of subjects' surnames. In Limón, last names pinpoint the ethnic membership of the individual, since they are either English or Spanish. Thus, in order to investigate the flow of genes between the two groups, we looked at the frequency of Spanish surnames in the Afro-Limonense, and the frequency of English surnames in the Hispanic group.

In Costa Rica, as in other Spanish-speaking countries, individuals carry two surnames, the first inherited from the father, the second inherited from the mother (Sans, 2000). Thus, using an individual's two surnames, it is possible to compute the inbreeding coefficient by isonymy. For this purpose, we use the formulae of Crow and Mange (1965) to compute total inbreeding (F_t), nonrandom inbreeding (F_n), and random inbreeding (F_r) (see Madrigal and Ware, 1997, 1999 for a discussion and application of this methodology). One important caveat is that in Costa Rica, if an individual does not have a legally acknowledged father at birth, he/she carries his/her mother's first surname, but twice. Thus, an isonymy study of the subjects' two surnames may pinpoint illegitimacy as well as inbreeding. With the purpose of determining if isonymy is more frequent in either ethnic group, we test with a simple X^2 test the hypothesis that the frequency of isonymous individuals is independent of ethnicity.

Population structure methodology

Madrigal et al. (2000) describe how the gene frequencies were computed. For each group, heterozygosity of loci found to be in equilibrium (Nei and Roychoudhury, 1974), and the overall heterozygosity were computed following Harpending and Chasko (1976).

Admixture estimates were computed for both groups using the computer program ADMIX.PAS, kindly provided by Jeffrey Long (see Long, 1991; Long and Smouse, 1983). A tri-hybrid model was used, using data from West Africa, Spain, and Amerindian gene frequencies as the parental populations. For several systems, however, there were no available data for the American Indian, or the Spanish components. In those cases, the parental data were obtained from those populations deemed to be as close to the parental ones as possible. The parental gene frequencies, and their sources, are listed in the Appendix. The PLG system was not used for the admixture estimates because adequate frequencies for the three parental populations were not found.

The procedure for computing admixture estimates (Long, 1991) allowed us to obtain an estimator of F_{st}, based on the variance of allele frequencies with respect to their expectations generated by the admixture model. These F_{st} values pinpoint the importance of genetic drift in the populations' evolutionary history. Besides the admixture estimates and the F_{st} values, the procedure allows for the computation of the mean squared error (MSE). This represents the proportion of allele-frequencies variation not explained by the admixture model. Alternatively, the explained portion of the variation can be expressed in terms of an R^2, where $R^2 = 1 - \text{MSE}$ (Long et al., 1991). In other words, R^2 quantifies the proportion of variance in the hybrid population's allele frequencies, which is accounted for by the admixture model (Long and Smouse, 1983).

It was of interest to place both Limonense populations in the wider Afro-Caribbean gene-frequency maps obtained by previous researchers. Particularly important was to compare our results with those of researchers in other Afro-Central American populations (Crawford, 1983, 1984; Crawford et al., 1981, 1984; Devor et al., 1984). To this end, the allele frequencies of our populations and other Afro-Caribbean groups graciously provided by Martinez-Labarga et al. at http://www.uniroma2.it/biologia/lab/anthromol/collab.htm were analyzed using the R matrix methodology proposed by Harpending and Jenkins (1973). The South American and parental gene frequencies provided by Martinez-Labarga et al. were not used in this paper. Rather, we used the parental populations described above, and decided to concentrate on the Caribbean only, to the exclusion of South America. Therefore, only 16 populations are considered in the R matrix analysis: African parental, Amerindian parental, Spanish parental, Afro-Limonenses, Hispanic Limonenses, Panamanian (an average of the frequencies provided by Martinez-Labarga et al.), Bluefields, Black Caribs of St. Vincent island, Creoles St. Vincent, Black Caribs Livingston (Guatemala), Black Caribs Belize, Creoles Belize, Haiti, Trinidad, Mulatos Cuba, and Negros Cuba.

RESULTS

Gene frequencies

The gene frequencies of Afro-Limonenses and Hispanic-Limonenses are displayed in Tables 1 and 2. The differences in sample sizes were caused by loss of specimens during shipping to the USA. Madrigal et al. (2000) discuss these gene frequencies at length.

Heterozygosity estimates

Table 3 shows that the average heterozygosities were virtually identical in the two groups (for the

TABLE 1. Gene frequencies in Afro-Limonenses[1]

Systems	Phenotypes	Observed numbers	Expected frequencies	Alleles	Frequencies
Gel	1 1	1	0.01	1	0.01
	1 2	0	0.19	2	0.095
	2 2	2	0.9025	1F	0.68
	1F	0	1.36	1S	0.17
	1F 1F	49	46.24	1A1	0.03
	1 1S	0	0.34	1C10	0.015
	1S 1S	1	2.89		
	1 1A1	0	0.06		
	1 1C10	0	0.03		
	2 1F	8	12.92		
	2 1S	7	3.23		
	2 1A1	0	0.57		
	2 1C10	0	0.285		
	1F 1S	23	23.12		
	1F 1A1	4	4.08		
	1F 1C10	3	2.04		
	1S 1A1	2	1.02		
	1S 1C10	0	0.51		
	1A1 1C10	0	0.09		
	1C10 1C10	0	0.02		
	1A1 1A1	0	0.09		
	Total	100			
$X^2 = 13.95$, df = 15, $P > 0.05$ (rare variants = dropped the 1 1 phenotype)					
BF	F F	33	34.7	F	0.5423
	F S	56	53.71	S	0.4197
	S S	20	20.78	F1	0.0338
	F F1	6	4.3258	S07	0.0042
	F1 F1	0	0.13		
	F07	0	0.537		
	S07 S07	1	0.002		
	S F1	2	3.31		
	S07	0	0.413		
	F1 S07	0	0.0335		
	Total	118			
$X^2 = 2.51$, df = 6, $P > 0.05$ (rare variants = dropped the S07 S07 phenotype)					
PGMI	1+ 1+	92	89.3	1+	0.6305
	1+ 1−	49	54.0	1−	0.1903
	1− 1−	10	8.14	2+	0.1305
	1+ 2+	38	37.0	2−	0.0487
	2+ 2+	4	3.8		
	1+ 2−	13	13.72		
	1− 2+	10	11.18		
	2− 2−	0	0.54		
	1− 2−	7	4.16		
	2+ 2−	3	2.86		
	Total	226			
$X^2 = 3.66$, df = 6, $P > 0.05$					
ACP	A A	13	10.75	A	0.2234
	A B	67	74.0	B	0.7697
	B B	132	129.0	C	0.0023
	A C	0	0.215	R	0.0023
	C C	0	0.001	A′	0.0023
	B C	1	0.7525		
	R R	1	0.0016		
	A R	0	0.258		
	A′ A′	0	0.001		
	B R	0	0.8815		
	C R	0	0.00258		
	A A′	0	0.215		
	B A′	1	0.7525		
	C A′	0	0.00215		
	R A′	0	0.00215		
	Total	215			
$X^2 = 2.93$, df = 10, $P > 0.05$ (rare variants = dropped R R phenotype)					
Hb	A A	185	183.4	A	0.8797
	A S	43	44.0	S	0.1054
	S S	2	2.6	C	0.0126
	A C	4	5.3	F	0.0023
	C C	0	0.04		
	A F	0	0.948		
	S C	2	0.616		
	F F	0	0.001		
	S F	1	0.114		
	C F	0	0.014		
	Total	237			
$X^2 = 11.38$, df = 6, $P > 0.05$					

TABLE 1. (continued)

Systems	Phenotypes	Observed numbers	Expected frequencies	Alleles	Frequencies
PLG	1 1	67	67.8	1	0.7743
	1 2	36	35.595	2	0.2035
	2 2	5	4.633	3	0.0089
	1 B	1	0.791	B	0.0044
	B B	0	0.0022	D	0.0089
	2 B	0	0.202		
	D D	0	0.0089		
	1 D	2	1.5594		
	3 3	0	0.0089		
	2 D	0	0.4068		
	B D	0	0.0088		
	1 3	2	1.5594		
	2 3	0	0.4068		
	B 3	0	0.0088		
	D 3	0	0.0178		
	Total	113			
$X^2 = 1.42$, df = 10, $P > 0.05$					
G6PD				(average both sexes)	
Males	A+	15	12.64	A+	0.16
	A−	6	4.345	A−	0.055
	B+	54	58.46	B+	0.74
	B−	1	1.58	B−	0.02
	Variant	3	1.975	Variant	0.025
	Total	79			
Females	A+ A+	12	3.8656		
	A+ A−	0	2.6576		
	A− A−	4	0.453		
	A+ B−	1	0.9664		
	B− B−	4	0.0604		
	A+ B+	12	35.7568		
	B+ B+	115	82.6876		
	A+ Variant	2	1.208		
	A− B+	0	12.2914		
	A− Variant	0	0.41525		
	B− B+	0	4.4696		
	B− Variant	0	0.151		
	B+ Variant	1	5.587		
	Var. Var.	0	0.0944		
	A− B−	0	0.3322		
	Total	151			
$X^2 = 88.62$, df = 17, $P < 0.001$ (rare variants = collapsed A− A−, A+ B−, B− B−, A− B−, A− Var., B− Var., and Var. Var.)					

[1] ACP1, acid phosphatase 1; BF, properdin factor; PGM1, phosphoglucomutase 1 (subtypes); GcI, group-specific component (subtypes); Hb, hemoglobin; G6PD, glucose-6-phosphate dehydrogenase.

Afro-Limonenses average h = 0.42, and for the Hispanic-Limonenses average h = 0.41). These results are expected in light of the fact that African-derived populations are not expected to have higher heterozygosities for classic genetic markers (Relethford, 1997; Relethford and Jorde, 1999).

Analysis of surnames

For various reasons, it was possible to obtain the two surnames from only 204 Afro-Limonense, and of 132 Hispanic-Limonense subjects. The hypothesis that ethnicity (as reported by the subjects themselves) is independent from surname language is strongly rejected ($X^2 = 237.9$, df = 2, $P = 0.01$). Only 4 out of 132 Hispanics (1.56%) had one English surname, and none had two English surnames. In contrast, 15 out of 204 (4.46%) of Afro-Limonenses had one Spanish surname, and 24 out of 204 (7.14%) had two Spanish surnames (see Table 4).

The hypothesis that ethnicity is independent of isonymy is also strongly rejected ($X^2 = 9.688$, df = 1, $P = 0.002$; see Table 5). Thus, isonymy is significantly more frequent among Afro-Limonenses. A more in-depth study of isonymy using the formulae of Crow and Mange (1965) to compute total inbreeding (F_t), nonrandom inbreeding (F_n), and random inbreeding (F_r) confirms this result. Thus, the total inbreeding is higher in the Afro-Limonense ($F_t = 0.072577$) group than in the Hispanic one ($F_t = 0.036162$). In both groups, the nonrandom component of inbreeding by isonymy accounts for most of the total inbreeding ($F_n = 0.071394$ and $F_r = 0.001274$ in the Afro-Limonense group, and $F_n = 0.034513$ and $F_r = 0.001707$ in the Hispanic group).

Admixture and F_{st} estimates

The admixture estimates (Table 6) for the Hispanic population indicate a very similar contribution of the Amerindian and European parental populations to the formation of this gene pool. This contradicts the popular notion that "white" Costa Ricans are "mostly Europeans." Indeed, the Euro-

TABLE 2. Gene frequencies in Hispanic-Limonenses[1]

Systems	Phenotypes	Observed numbers	Expected frequencies	Alleles	Frequencies
GcI	2 2	2	2.5	2	0.188
	2 1F	15	8.7	1 F	0.33
	2 1S	8	13.0	1S	0.482
	1F 1F	10	7.6		
	1S 1S	24	16.5		
	1S 1F	13	23.74		
	Total	72			
$X^2 = 15.61$, df = 3, $0.001 < P < 0.05$					
BF	F F	4	3.5	F	0.219
	F F1	1	0.67	S	0.753
	F1 F1	0	0.023	F1	0.021
	F S	23	24.08	S07	0.007
	SF1	2	2.31		
	S S	42	41.4		
	S07	0	0.004		
	F S07	0	0.219		
	F1 S07	0	0.0219		
	S S07	1	0.7799		
	Total	73			
$X^2 = 0.663$, df = 6, not significant					
PGMI	1+ 1+	32	31.1	1+	0.527
	1− 1−	6	6.0	1−	0.232
	2+ 2+	2	2.3	2+	0.143
	2− 2−	1	1.07	2−	0.098
	1+ 1−	25	27.4		
	2+ 2−	2	3.14		
	1+ 2+	19	16.9		
	1+ 2−	10	11.57		
	1− 2+	7	7.43		
	1− 2−	8	5.09		
	Total	112			
$X^2 = 2.80$, df = 6, $P > 0.05$					
ACP	A B	40	45.2115	A	0.281
	A C	2	0.8115	B	0.706
	A A	11	9.0	C	0.013
	B B	60	56.8		
	C B	1	2.1		
	C C	0	0.019		
	Total	114			
$X^2 = 3.56$, df = 3, $P > 0.05$					
Hb	A A	124	123.97	A	0.988
	A S	2	2.0	S	0.008
	A C	1	1.0	C	0.004
	Total	127			
$X^2 = 0.018$, df = 3, not significant					
PLG	1 1	48	46.594	1	0.783
	1 2	20	23.446	2	0.197
	2 2	5	2.9596	3	0.007
	1 E	2	1.5	E	0.013
	1 3	1	0.88436		
	EE	0	0.013		
	2 E	0	0.388		
	3 3	0	0.004		
	2 3	0	0.205		
	E 3	0	0.014		
	Total	76			
$X^2 = 2.84$, df = 6, $P > 0.05$					
G6PD					(average both sexes)
Males	A+	1	0.936	A+	0.026
	A−	1	0.648	A−	0.018
	B+	34	34.416	B+	0.956
	Total	36			
Females	A+ A+	1	0.061		
	A+ A−	0	0.084		
	A− A−	0	0.029		
	A+ B+	2	4.47		
	A− B+	1	3.096		
	B+ B+	86	82.25		
	Total	90			
$X^2 = 7.073$ (for both sexes), df = 6, $P > 0.05$ (rare variants = collapsed A+ A+, A+ A− and A− A− phenotypes)					

[1] ACP1, acid phosphatase 1; BF, properdin factor; PGM1, phosphoglucomutase 1 (subtypes); GcI, group-specific component (subtypes); Hb, hemoglobin; G6PD, glucose-6-phosphate dehydrogenase.

TABLE 3. Heterozygosity estimates

System	Corrected heterozygosity (hi)
Afro-Limónese	
GcI	0.500
BF	0.530
PGMI	0.548
ACP	0.358
PLG	0.360
Hb	0.215
hi average	0.42
Hispanic-Limónese	
GcI	0.627
BF	0.387
PGMI	0.641
ACP	0.424
PLG	0.350
Hb	0.023
hi average	0.41

TABLE 4. Language of surnames in both ethnic groups[1]

	Ethnicity					
	Afro		Hispanic		Observed	
Language of surnames	n	%	n	%	N	%
Two English surnames	165	49.40	0	0.00	165	49.40
Two Spanish surnames	24	7.14	128	37.80	152	44.94
One English/one Spanish surname	15	4.46	4	1.19	19	5.65
Totals	204	61.01	132	38.99	336	100.00

[1] Contingency chi-square test yields a chi-square value of X^2 = 237.867, df = 2, P = 0.001.

TABLE 5. Frequency of isonymous surnames by ethnicity[1]

	Ethnicity					
	Afro		Hispanic		Observed	
Isonymy	n	%	n	%	n	%
Isonymous	59	17.86	19	5.65	79	23.51
Nonisonymous	145	43.15	113	33.33	257	76.49
Totals	204	61.01	132	38.99	336	100.00

[1] Contingency chi-square test yields a chi-square value of X^2 = 9.688, df = 1, P = 0.002.

TABLE 6. Admixture estimates and standard errors[1]

Ancestral population	Admixture estimate	SE of admixture
I. Hispanic Limonense Sample		
European	0.5866	0.069
Amerindian	0.3383	0.0653
African	0.0751	0.0434
II. Afro Limonense Sample		
European	0.1047	0.1482
Amerindian	0.1357	0.1437
African	0.7595	0.1266

[1] HL MSE = 0.009, R^2 = 0.991, F_{st} = 0.00558. AL MSE = 0.053, R^2 = 0.9467, F_{st} = 0.05137.

pean component for the Hispanic group is less than 60% (M1 = 0.5866), with 34% deriving from the Amerindian (M2 = 0.3383), and 7% from the African parental groups (M3 = 0.0751). The admixture estimates for the Afro-Limonense group indicate that it derives mostly from the African parental population (M3 = 0.7595), with very similar contributions from the European (M1 = 0.1047) and the Amerindian (M2 = 0.1357) parental groups. The F_{st} of the Afro group (F_{st} = 0.05137) is higher than that of the Hispanic group (F_{st} = 0.00558), and indicates a greater importance of gene drift in the evolution of the Afro-Limonense group. As proposed by Long et al. (1991), the admixture estimates can be evaluated by the proportion of variance of allele frequencies explained by the admixture model. For both groups, the R^2 is very high, indicating that the model is satisfactory (R^2 = 0.9467 for the Afro-Limonense group, and R^2 = 0.991 for the Hispanic group). The high standard errors of the admixture estimates probably reflect the relatively small number of loci used for the calculations.

R matrix results

The percentage of variation explained by the first two scaled eigenvectors is 77% of the total variation (the first one explains 47% and the second one 30% of the variance). A plot of the populations along the first two eigenvectors basically displays a range of the importance of African ancestry in all the populations. Along this range, the Amerindian parental population is farthest removed from the African parental population. The position of the Hispanic-Limonense group is indicative of its high Amerindian component and its low African contribution. Indeed, it is the population closest to the Amerindian parental population. The population closest to the HL group is the Black Caribs from St. Vincent Island, who according to Crawford et al. (1984) "gravitated toward the Amerindians" in its GcI frequencies. The Afro-Limonenses, in contrast, are clearly aligned with other Afro-Caribbean groups. Admittedly, they are the Afro-derived group closest to the Hispanic Limonenses (after the Black Caribs from St. Vincent, as noted above).

DISCUSSION

New World populations of African ancestry have been of great anthropological interest (Da Silva et al., 1999; Martinez-Labarga et al., 1999). Afro-Caribbean human groups have received attention in part because of the interaction of natural selection and gene flow in shaping their population structure. The slave trade brought African populations to the New World. However, it was their resistance to malaria which allowed them to be so successful in the Caribbean and Atlantic Central American coast (Crawford, 1983). The importance of gene flow, however, is not restricted to the initial arrival of the African slave groups, but it continued (and continues today) in the form of large population movements. Particularly important for Central American countries was the construction of the Panama Canal, and in Costa Rica, of the Atlantic-coast railroad. In this manner, a group of Afro-Jamaican workers were brought to Limón, Costa Rica, where

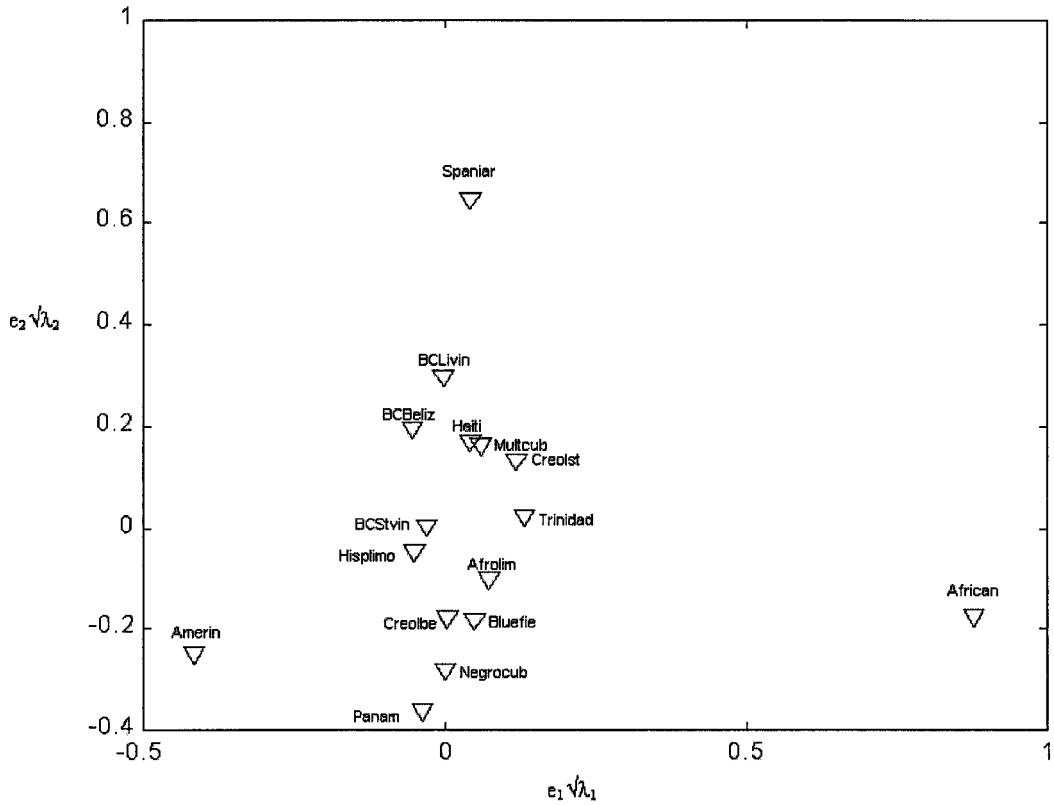

Fig. 1. Plot of 16 populations along two-scaled eigenvectors (77% of total variation explained). Populations plotted: African, African parental; Amerin, Amerindian parental; Spaniar, Spanish parental; Afrolim, Afro-Limonenses; Hisplimo, Hispanic Limonenses; Panam, Panama; Bluefie, Bluefields; BCStvin, Black Caribs of St. Vincent island; Creolst, Creoles St. Vincent; BCLivin, Black Caribs Livingston (Guatemala); BCBeliz, Black Caribs Belize; Creobe, Creoles Belize, Haiti, Trinidad; Multcub, Mulatos Cuba; Negrocub, Negros Cuba (see text for sources of data).

they remained despite promises that they would eventually return to Jamaica. It is thus that the evolution of the Limón population should be understood: in the context of the wider Afro-Caribbean, as well as in the context of Costa Rica population movements (Chomsky, 1995). Indeed, although the Afro-Caribbean group in Limón remained rather isolated during the first part of the 20th century, a large movement of the wider Costa Rican population resulted in the establishment of the Hispanic ethnic group of Limón, and allowed gene flow between both ethnic groups.

We acknowledge that the number of loci sampled for this paper is small. Further sampling of the Limón population, which would include more loci, should complement this project, as discussed below. Moreover, we acknowledge that some of the parental groups are less than ideal. This reflects the lack of information of PGM1 and GCI *subtypes* in Central American native groups.

The admixture estimates reveal a sizable contribution of the Amerindian, and a small contribution of the African parental populations to the Hispanic group. These results contradict the popular notion that Costa Ricans are "mostly European," and suggest a low level of gene flow from the Afro-Limonense group into the Hispanic one. The genetic data also indicate that the Afro-Limonense sample is overwhelmingly of African origin and has received few genes from the Hispanic group. When compared to a table showing admixture estimates of 14 Afro-Latin American samples, the Afro-Limonense group is second in African ancestry to only one sample (Sans, 2000).

The F_{st} values demonstrate clearly the differential importance of genetic drift in the evolution of these two groups. Given that the F_{st} is computed from the variance of allele frequencies with respect to their expectations from the admixture model, it is to be expected that the F_{st} of the Afro group (F_{st} = 0.05137) is higher than that of the Hispanic group (F_{st} = 0.00558). Since F_{st} measures the "reduction in heterozygosity of a subpopulation due to random genetic drift" (Hartl, 1988), a higher F_{st} value indicates that genetic drift has been more important in the evolution of the Afro-Limonense breeding population. The F_{st} value of the AL probably reflects the group's historically small size and its relative isolation from the Costa Rican population at large. This F_{st} may also reflect the fact that the AL group is a sample taken from a Jamaican group, which was itself a sample taken from Africa. That the R^2 is lower for the Afro-Limonenses also supports the notion that they have experienced more genetic drift.

The relative isolation of the Afro-Limonense sample suggested by the F_{st} is also indicated by the

analysis of surnames. The higher coefficient of inbreeding by isonymy can be interpreted as reflecting the fact that the Afro-Limonense subjects were the offspring of a more endogamous breeding population. However, we acknowledge that it may also reflect the high level of out-of-wedlock births reported to occur in this group (since offspring of unmarried mothers have isonymous surnames). Although we are not in a position to partition the coefficient of inbreeding by isonymy into the contribution due to higher illegitimacy and that due to high endogamy, the inclusion of surnames complements and enriches the genetic analysis. The distribution of surnames virtually mirrors the R matrix plot position of the groups, in that there is a clear separation of ethnicity by language of surnames: only 1.56 of Hispanics carried one English surname, and just under 12% of Afro-Limonenses carried one or two Spanish names. These percentages indicate that the offspring of such biethnic unions tend to be classified as belonging to the Afro-Limonense group.

Since Madrigal (1988) and Purcell (1993) both observed frequent interethnic unions during their fieldwork in the 1980s and 1970s, respectively, we were surprised that the two groups, when plotted along the scaled R matrix eigenvectors, did not cluster with each other. The plot clearly clustered the Afro-Limonense group with other Afro-Caribbean groups and the Hispanic-Limonense group with the St. Vincent Black Caribs.

This apparent discrepancy (both groups do not cluster, yet they engaged in gene flow) can be resolved by noting that the people whose blood samples and surnames were obtained by Madrigal (1988) were at least in their 40s, but most of them were in their 50s, 60s, and 70s. In contrast, the people whom Madrigal (1988) and Purcell (1993 and personal communication) observed in frequent interethnic unions were in their teens and 20s. Thus, it is likely that the subjects whose blood samples and surnames are here analyzed were not the offspring of the liberal interethnic mating observed by Madrigal (1988) and Purcell (1993). Rather, it appears that they were the offspring of mostly within-ethnic mating. The observations by Madrigal (1988) and Purcell (1993) lead us to predict that the offspring of the frequent interethnic unions they observed in the field will have less-clearly-genetically demarcated boundaries. It would be of great interest to sample the adult Limón population in about 15 years, and determine if the genetic differences between the two groups have become more blurred.

ACKNOWLEDGMENTS

The support of E. Mohs, M. Coto, and M. Barrenechea is gratefully acknowledged.

APPENDIX. Parental gene frequencies[1]

Allele	Parental frequencies		
	African	Amerind	Spaniard
acp1*a	0.185	0.062	0.324
acp1*b	0.79	0.813	0.624
acpa*c	0.0	0.0	0.052
acpa*r	0.025	0.125	0.0
PGM1*1+	0.7585	0.485	0.621
PGM1*1−	0.0905	0.342	0.114
PGM1*2+	0.123	0.02	0.211
PGM1*2−	0.028	0.153	0.054
G6PD*B+	0.638	0.1	0.997
G6PD*A+	0.226	0.0	0.0
G6PD*A−	0.136	0.0	0.0
HB*A	0.88	0.1	0.9995
HB*S	0.104	0.0	0.0005
HB*C	0.013	0.0	0.0

[1] Sources of parental populations: BF, West African from Parra et al. (1995), Spanish from Roychoudhury and Nei (1988), and Central American Indian from Cavalli-Sforza et al. (1994). ACP, a West African average computed from data from Roychoudhury and Nei (1988), Spanish and Central American Indian from Roychoudhury and Nei (1988). PGM1, a West African average computed from data from Roychoudhury and Nei (1988), Spanish and North American Indian (no data available for subtypes from Central American groups) from Roychoudhury and Nei (1988). GCI, a West African average computed from data from Roychoudhury and Nei (1988) and Parra et al. (1995), Roman (instead of Spanish) and North American Indian (no data available for subtypes from Central American groups) from Roychoudhury and Nei (1988). Hb, a West African average computed from data from Roychoudhury and Nei (1988), Portugal (instead of Spain) and Central-American Indian from Roychoudhury and Nei (1988). G-6PD, a West African average computed from data from Roychoudhury and Nei (1988), Spain and South American Indian from Roychoudhury and Nei (1988).

LITERATURE CITED

Aguiar GFS. 1992. Effects of demographic and ethnohistorical factors on average heterozygosities of South Amerindians. Am J Phys Anthropol 88:299–308.

Barbujani G, Sokal RR. 1991a. Genetic population-structure of Italy. 1. Geographic patterns of gene frequencies. Hum Biol 63:253–272.

Barbujani G, Sokal RR. 1991b. Genetic population-structure of Italy. 2. Physical and cultural barriers to gene flow. Am J Hum Genet 48:398–411.

Barbujani G, Vian P, Fabbris L. 1992. Cultural barriers associated with large gene-frequency differences among Italian populations. Hum Biol 64:479–495.

Barbujani G, Whitehead GN, Bertorelle G, et al. 1994a. Testing hypotheses on processes of genetic and linguistic change in the Caucasus. Hum Biol 66:843–864.

Barbujani G, Nasidze IS, Whitehead GN. 1994b. Genetic diversity in the Caucasus. Hum Biol 66:639–668.

Barrantes R, Smouse PE, Mohrenweiser HW, et al. 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.

Bittles AH, Smith MT. 1994. Religious differentials in postfamine marriage patterns, Northern Ireland, 1840–1915. 1. Demographic and isonymy analysis. Hum Biol 66:59–76.

Bogin B. 1993. Biocultural studies of ethnic groups. In: Lasker G, Mascie C, editors. Research strategies in human biology: field and survey studies. Cambridge: Cambridge University Press. p 33–61.

Bryce-Laporte R. 1962. Social relations and cultural persistence (change) among Jamaicans in a rural area of Costa Rica. Doctoral dissertation. University of Puerto Rico.

Casey J. 1979. Limón: 1880–1940. San Jose: Editorial Costa Rica.

Cavalli-Sforza LL, et al. 1994. The history and geography of human genes. Princeton, NJ: Princeton University Press.

Chavez R. 1982. Los recursos fisicos de la zona atlantica de Costa Rica. University of Costa Rica Irregular Publication.
Chomsky A. 1995. Afro-Jamaican traditions and labor organizing on United Fruit Company plantations in Costa Rica, 1910. J Soc Hist 28:837–855.
Crawford M. 1983. The anthropological genetics of the Black Caribs (Garifuna) of Central America and the Caribbean. Yrbk Phys Anthropol 26:161–192.
Crawford M. 1984. Problems and hypotheses: an introduction. In: Crawford M, editor. Black Caribs, a case study in biocultural adaptation. New York: Plenum Press. p 1–9.
Crawford M. 1991. Obituary: Dale David Dykes (1945–1990). Hum Biol 63:3–8.
Crawford M, Gonzalez NL, Schanfield MS, Dykes DD, Skradski K, Polesky H. 1981 The Black Caribs (Garifuna) of Livingston, Guatemala: genetic markers and admixture estimates. Hum Biol 53:87–103.
Crawford M, Dykes DD, Skradsky K, Polesky H. 1984. Blood group, serum protein, and red cell enzyme polymorphisms, and admixture among the Black Caribs and Creoles of Central America and the Caribbean. In: Crawford M, editor. Black Caribs, a case study in biocultural adaptation. New York: Plenum Press. p 303–333.
Crow J, Mange AP. 1965. Measurement of inbreeding from the frequency of marriages between persons of the same surname. Eugen Q 12:199–203.
Da Silva J, Bortolini W, Meyer M, Salzano D, Elion F, Krishnamoorthy J, Schneider R, De Guerra M, Layrisse D, Castellano Z, Weimer H, Zago MA. 1999. Genetic diversity of two African and sixteen South American populations determined on the basis of six hypervariable loci. Am J Phys Anthropol 109:425–437.
Devor E, Crawford MH, Bach-Enciso V. 1984. Genetic population structure of the Black Caribs and Creoles. In: Crawford M, editor. Black Caribs, a case study in biocultural adaptation. New York: Plenum Press. p 365–379.
Duncan Q. 1972. El negro Antillano: inmigracion y presencia. In: Melendez C, Duncan Q, editors. El negro en Costa Rica. San Jose: Editorial Costa Rica. p 89–128.
Dykes D, Nelson M, Polesky H. 1983. Distribution of plasminogen allotypes in 8 populations of the Western hemisphere. Electrophoresis 4:417–420.
Fernandez F, Mendez H. 1973. El Negro en la historia de la politica Costarricense. Thesis, History and Geography Department and the School of Law, University of Costa Rica.
Friedlaender J. 1975. Patterns of human variation. The demography, genetics, and phenetics of Bougainville Islanders. Cambridge, MA: Harvard University Press.
Guillén D. 1989. Costa Rica. Instituto de Investigaciones Dr. José María Luis Mora. Guadalajara, Mexico: Universidad de Guadalajara Press.
Guglielmino CR, Beres J. 1996. Genetic structure in relation to the history of Hungarian ethnic groups. Hum Biol 68:335–355.
Harpending H, Chasko W Jr. 1976. Heterozygosity and population structure in southern Africa. In: Giles E, Friedlaender J, editors. The measures of man: methodologies in biological anthropology. Cambridge, MA: Peabody Museum Press. p 214–129.
Harpending HC, Jenkins T. 1973. Genetic distance among southern African populations. In: Crawford M, Workman P, editors. Method and theory in anthropological genetics. Albuquerque: University of New Mexico Press. p 177–199.
Hartl D. 1988. A primer of population genetics. 2nd ed. Sunderland, MA: Sinauer Associates.
Jorde L.1980. The genetic structure of subdivided human populations. In: Mielke J, Crawford M, editors. Theory and methods. New York: Plenum Press. p 135–209.
Long J. 1991. The genetic structure of admixed populations. Genetics 127:417–426.
Long J, Smouse PE. 1983. Intertribal gene flow between the Ye'cuana and Yanomama: genetic analysis of an admixed village. Am J Phys Anthropol 61:411–422.
Long J, Williams RC, McAuley JE, Medis R, Partel R, Tregellas WM, South SF, Rea AE, McCormick SB, Iwaniec U. 1991. Genetic variation in Arizona Mexican Americans: estimation and interpretation of admixture proportions. Am J Phys Anthropol 84:141–157.
Madrigal L. 1988. Hemoglobin genotype and fertility in a malarial environment: Limón, Costa Rica. Ph.D. dissertation, Department of Anthropology, University of Kansas.
Madrigal L. 1989. Hemoglobin genotype, fertility and the malaria hypothesis. Hum Biol 61:311–325.
Madrigal L. 1994. Twinning rates in admixed Costa Rican populations. Am J Hum Biol 6:215–218.
Madrigal L. 1995. Differential fertility of mothers of twins and mothers of singletons: a study in Limón, Costa Rica. Hum Biol 67:779–787.
Madrigal L, Ware B. 1997. Inbreeding in Escazu, Costa Rica (1800–1840, 1850–1899): isonymy and ecclesiastical dispensations. Hum Biol 69:703–714.
Madrigal L, Ware B. 1999. Mating pattern and population structure in Escazu, Costa Rica: a study using marriage records. Hum Biol 71:963–975.
Madrigal L, Saenz G, Dykes D. 2001. The frequency of twinning in two Costa Rican ethnic groups: an update. Am J Hum Biol 13 (in press).
Martinez-Labarga C, Rickards O, Scacchi R, Corbo RM, Biondi G, Pena JA, De Vieira CV, Guevara AEB, Santurino MSM, De Stefano GF. 1999. Genetic population structure of two African-Ecuadorian communities of Esmeraldas. Am J Phys Anthropol 109:159–174.
Martinez-Labarga C, Rickards O, Scacchi R, Corbo RM, Biondi G, Peña JA, de Vieira V, Guevara AEB, Mesa Santurino MS, De Stefano GF. At http://www.uniroma2.it/biologia/lab/anthromol/collab.htm
Melendez C. 1972. Aspectos sobre la inmigracion Jamaicana. In: Melendez C, Duncan Q, editors. El negro en Costa Rica. Editorial Costa Rica.
Nei M, Roychoudhury AK. 1974. Sampling variances of heterozygosity and genetic distance. Genetics 76:379–390.
Parra EJ, Ribeiro JCT, Caeiro JLB, et al. 1995. Genetic structure of the population of Cabo Verde (West Africa)—evidence of substantial European admixture. Am J Phys Anthropol 97:381–389.
Purcell TW. 1993. Banana fallout. Los Angeles: Center for African-American Studies.
Relethford JH. 1985. Isolation by distance, linguistic similarity, and the genetic structure on Bougainville Island. Am J Phys Anthropol 66:317–326.
Relethford JH. 1997. Mutation rate and excess African heterozygosity. Hum Biol 69:785–792.
Relethford JH, Crawford MH. 1998. Influence of religion and birthplace on the genetic structure of Northern Ireland. Ann Hum Biol 25:117–125.
Relethford J, Jorde LB. 1999. Genetic evidence for larger African population size during recent human evolution. Am J Phys Anthropol 108:251–260.
Rickards O, Tartaglia M, Martinez-Labarga C, et al. 1994. Genetic characterization of the Cayapa Indians of Ecuador and their genetic-relationships to other Native-American populations. Hum Biol 66:1127–1127.
Roychoudhury A, Nei M. 1988. Human polymorphic genes. New York: Oxford University Press.
Sans M. 2000. Admixture studies in Latin America: from the 20th to the 21st century. Hum Biol 72:155–177.
Sirajuddin SM, Duggirala R, Crawford MH. 1994. Population-structure of the Chenchu and other south Indian tribal groups—relationships between genetic, anthropometric, dermatoglyphic, geographic, and linguistic distances. Hum Biol 66:865–884.
Smith MT, Williams WR, Mchugh JJ, et al. 1990. Isonymic analysis of postfamine relationships in the ards peninsula, NE Ireland—effects of geographical and politicoreligious boundaries. Am J Hum Biol 2:245–254.
Stevenson J, Brown RJ, Schanfield MS. 1983. Surname analysis as a sampling method for recovery of genetic information. Hum Biol 55:219–225.
Stewart W. 1967. Keith y Costa Rica. San Jose: Editorial Costa Rica.
Tavares-Neto J, Azevedo ES.1978. Family names and ABO blood group frequencies in a mixed population of Bahia, Brazil. Hum Biol 50:361–367.