Am. J. Hum. Genet. 47:867–875, 1990

Diversity of Some Gene Frequencies in European and Asian Populations. V. Steep Multilocus Clines

Guido Barbujani,* Geoffrey M. Jacquez,† and Laura Ligi*

*Dipartimento di Biologia, Universita' di Padova, Padua Italy; and †Department of Ecology and Evolution, State University of New York, Stony Brook

Summary

Regions of abrupt genetic change, which result from either rapid spatial change of selective pressures or limited admixture, were investigated in Europe and Asia on the basis of eight red cell markers typed in 960 samples. Two methods were employed, one based on genetic distances and one on evaluation of the first derivative of the surfaces representing allele-frequency variation. Genetic divergence tends to be maximal between populations that are separated by physical factors (mountain ranges and seas) but also separated by cultural barriers (different language affiliation). This suggests that mating isolation, rather than adaptive response to environmental change, accounts for spatially abrupt genetic change at the loci studied and that cultural differences associated with language contribute to isolating populations. Although selection may have determined two wide allele-frequency gradients, the genetic structure of European and Asian populations seems primarily to reflect isolation by distance when investigated on a small scale and migration patterns (or absence of migration) when investigated on a larger scale.

Introduction

Natural populations are a patchwork of isolates that either have diverged in loco because of genetic drift (or differential selection) or have evolved independently before coming into contact, separated often by narrow admixture zones. Gene flow is expected to blur random allele-frequency differences when the effective population size times the average migration rate exceeds one (Wright 1931). Effective population size and migration rate are difficult to measure directly (see Wood 1987), but estimates of their product (Wright 1951; Slatkin and Barton 1989) that are based on empirical measures of the standardized gene frequency variance (e.g., see Cavalli-Sforza 1966; Barbujani and Milani 1986) fall between 0.40 and 22, depending on the marker chosen; very few of them are smaller than one. Similarly, the estimates based on consanguineous marriage (reviewed in Morton 1982) range from 5.4 to 2,860. However, contrary to drift predictions, the frequencies of several markers differ substantially even between geographically close populations. Either sharp selective gradients or presence of local isolating factors may account for maintenance of such differences; the purpose of the present paper is to draw inferences on the responsible evolutionary mechanisms from the location of the zones where genetic change is sharpest.

Previous papers in this series have described genetic variation in humans from Europe and Asia on the basis of 10 markers of the red cell. We have shown that (1) geographical patterns of allele frequencies are different at different loci (Barbujani 1987*b*); (2) short-range genetic variation is gradual at almost all loci, in agreement with the predicted consequences of isolation by distance, but clines are wider than would be expected on the basis of individual migration distances (Barbujani 1988*b*); and (3) in only two cases is long-distance differentiation of populations simply clinal, as would be expected when a major migratory movement or a simple mechanism of differential selection affects gene frequencies; for most markers, more complex interactions of biological and demographic pressures must be

Received December 11, 1989; final revision received July 10, 1990.
Address for correspondence and reprints: Dr. Guido Barbujani, Dipartimento di Biologia, Universita' di Padova, Via Trieste 75, I-35121 Padua, Italy.

0002-9297/90/4705-015$02.00

postulated (Barbujani 1988*a*, 1988*b*; Farabegoli and Barbujani, in press). In the present study, we first located the areas of maximum genetic change per unit distance (steep multilocus clines, or "genetic boundaries"), on the basis of two approaches, outlined below. The geographical distribution of the genetic boundaries was analyzed to assess whether environmental diversity or limited admixture is likely to maintain them. When the latter seemed more plausible, we inquired into the nature of the factors limiting admixture.

The existence of zones of abrupt genetic change violates the assumption of stationarity (Mathéron 1970) underlying neutral models of population structure ("isolation by distance"; Wright 1943; Malécot 1948; Kimura and Weiss 1964; Morton et al. 1971; Barbujani 1987*a*). Stationarity implies that allele-frequency differences are statistically constant at constant distances between isolates, but large-scale human gene-frequency maps (Piazza et al. 1981*a*) show different rates of genetic change in different regions (Sokal 1988; Sokal et al. 1989*a*). As a consequence, the present study required unconventional approaches. The ratio between genetic and geographic distances at the boundaries between groups of populations was evaluated, and the zones of maximum slope were detected in the surfaces representing allele-frequency distributions (Barbujani et al. 1989). Finally, the results were examined in light of recent findings on human population structure.

Material and Methods

The Data

Eight unlinked polymorphic markers of the red cell were considered, for a total of 960 gene frequencies in European and Asian populations, each estimated from a sample of at least 50 individuals but generally many more (table 1). *PGM* and *PGD* map 45 cM apart and hence are virtually independent (Cook et al. 1974). Two markers included in previous studies of this series (*SOD* and *PGP*) were neglected, owing to the insufficient amount of information available. Details on preliminary processing of data can be found in the work of Barbujani and Milani (1986) and Farabegoli and Barbujani (in press). To achieve statistical independence of data, only one allele was considered at each locus. The values of the eight allele frequencies analyzed can be envisaged as eight discontinuous tridimensional surfaces, in which two dimensions are longitude and latitude and in which the third represents the allele frequency; discontinuity is due to the fact that samples are discrete in geographic space.

The Systemic Function

A first method for detecting areas of rapid change is based on the evaluation of a synthetic variable, called the "systemic function" (SF) by Womble (1951). The first derivatives of the allele-frequency surfaces attain their maxima where the slope of the surfaces (the rate of allele-frequency change) is highest. Allele-frequency boundaries are defined as those zones where such derivatives exceed an opportunely chosen threshold. This "wombling" approach required the following four steps:

1. Allele frequencies were transformed into quasi-continuous surfaces by interpolation. Hypothetical allele frequencies were computed at the nodes of a 58-row × 145-column grid, by using inverse squared distance weighting.
2. The first derivative of each interpolated surface was evaluated at the midpoint of each square formed by four adjacent nodes of the grid. We refer to these squares as "pixels." For each allele we constructed two matrices, containing the magnitude of this derivative and its direction, respectively. Together, these matrices define a vector field, describing the slope and orientation of the allele-frequency gradient (Barbujani et al. 1989).
3. The SF was computed by averaging magnitude and direction of the derivatives across all eight loci, by following the procedure described by Barbujani et al. (1989; also see Batschelet 1981, pp. 7–18).
4. These average vectors were judged significant if (1) they belonged to the upper 5% of the distribution of magnitudes; (2) they were connected by a king's move (by analogy to chess) to at least one other pixel belonging to the highest 5%; and (3) the orientation of the slope at any two potentially connected pixels did not differ by more than 30 degrees. In this way, random pixels showing a high slope in the SF were excluded, and only areas of consistently high allele-frequency variation became apparent.

Along with the approach described above, which will be referred to as SFO (systemic function, overall), a modified version (SFI; systemic function, surfaces individually analyzed) was followed. The areas of high slope in the individual surfaces were subjected to significance testing (step 4) *without* being averaged (step 3), yielding eight maps, each one referring to a single allele. The occurrences of each genetic boundary at the various loci were then counted. The binomial probability that a boundary occurs by chance at the same location in three or more surfaces is $(8!/3!5!)\ (0.05)^3\ (0.95)^5 = 0.0058$, when boundaries are defined as those zones

Table 1

Synopsis of Data Analyzed

Locus (symbol)	Chromosome Location[a]	No. of Samples	Allele Considered	Allele-Frequency Range
Glyoxalase (*GLO*)	6p11	116	GLO[2]	.410–.932
Esterase d (*ESD*)	13q14	121	ESD[2]	.059–.528
Phosphoglucomutase 1 (*PGM*)	1p22	168	PGM[1]	.398–.961
Acid phosphatase (*ACP*)	2p25	124	ACP[a]	.006–.610
Adenylate kinase (*AK*)	9q34	112	AK[2]	.000–.138
Adenosine deaminase (*ADA*)	20q13	108	ADA[2]	.019–.174
6-Phosphogluconate dehydrogenase (*PGD*)	1p36	126	PGD[a]	.763–.997
Glutamic-pyruvic transaminase (*GPT*)	16p11	85	GPT[1]	.265–.686
		960		

[a] Source: McKusick (1988).

across which gene-frequency differences fall into the highest 5% of the distribution of magnitudes. Therefore, the boundaries recognized in three or more individual surfaces are significant with less than 1% probability of type I error.

Genetic Distances

Since the interpolation of data necessary for wombling is likely to smooth the allele-frequency surfaces, an independent confirmation of the results was sought. In principle, we could compare the gene frequencies between adjacent localities and detect the pairs of localities between which the rate of genetic change is highest. The zones between these localities should correspond to the genetic boundaries detected by wombling. The problem with this approach is in comparing the eight sets of allele frequencies, since the available samples differ in number and location among loci. In wombling, interpolation increased the number of data points; for the alternative approach, the solution chosen was to reduce them. The data referring to spatially close populations were clumped, so as to create 28 "pseudolocalities," each one having a complete set of allele frequencies, because it included at least one sample (4.3 samples on the average) of individuals typed for each of the eight genes. In defining such geographic groups of populations, only the criterion of spatial closeness was followed; the degree of ethnic similarity among populations within the same group was disregarded. The allele-frequency values chosen to represent each group were the median allele frequencies among the samples pooled within the group. The geographic barycenter of each group was calculated as the unweighted average of the latitudes and longitudes of the samples of interest, whose number ranged from 18 to 74 (table 2).

In the next phase of analysis, the 28 pseudolocalities were joined by Delaunay triangulation (a connection graph; Green and Sibson 1977; fig. 1), and their distances were computed along this network. Sixty-six lines of contact, perpendicular to the links between adjacent geographical groups, were thus defined. A matrix of genetic distances among groups was then calculated, based on Cavalli-Sforza and Edwards' (1967) statistic. The rate of genetic change per distance unit was finally calculated as the genetic distance divided by the respective geographic distance, for all pairs of pseudolocalities directly connected in the Delaunay network. We shall refer to the approach based on genetic distances as "GD."

The 66 lines separating pseudolocalities were then ranked according to the rate of genetic change they displayed, and Monmonier's (1973) maximum-difference method was applied. This algorithm traced gene-frequency boundaries by linking the lines that separate the most sharply differentiated population groups. From the edges of Delaunay's network for which the rate of genetic change is highest, the boundaries were drawn across the highest-rate edges of the network until they had reached either the limits of the map or another preexisting boundary.

Table 2

Median Allele Frequencies in 28 Groups of Populations

Population Group[a]	Allele Frequency								No. of Samples
	GLO[2]	ESD[2]	PGM[1]	ACP[a]	AK[2]	ADA[2]	PGD[a]	GPT[1]	
HIB	.567	.121	.755	.311	.041	.050	.976	.498	43
BRI	.553	.112	.775	.332	.035	.052	.983	.528	42
NWI	.558	.133	.780	.374	.051	.101	.978	.531	24
FRB	.562	.123	.725	.280	.031	.041	.986	.514	32
ALP	.571	.108	.791	.344	.038	.061	.976	.524	22
GER	.555	.112	.779	.346	.038	.056	.973	.511	34
ITN	.591	.135	.712	.302	.029	.079	.978	.549	40
SAR	.551	.121	.763	.264	.013	.061	.989	.300	20
ITS	.611	.152	.708	.279	.032	.084	.966	.552	25
SIC	.612	.153	.708	.272	.036	.086	.961	.508	38
SFD	.623	.092	.689	.348	.033	.074	.969	.638	34
EEU	.566	.096	.743	.361	.038	.058	.974	.504	41
HUN	.680	.104	.759	.317	.018	.085	.984	.448	26
BGR	.653	.089	.696	.220	.039	.072	.967	.561	23
MEA	.630	.196	.715	.302	.029	.131	.967	.544	28
IRA	.723	.195	.684	.203	.053	.087	.960	.476	53
INW	.754	.228	.713	.281	.100	.111	.966	.390	38
INS	.770	.253	.715	.250	.063	.143	.979	.594	24
INE	.787	.319	.724	.228	.047	.108	.975	.405	63
BTH	.812	.240	.736	.290	.002	.111	.957	.445	29
IMV	.815	.299	.745	.344	.003	.092	.957	.438	31
FEA	.835	.290	.742	.226	.031	.074	.941	.415	32
CHS	.826	.354	.668	.284	.001	.054	.929	.417	18
CHN	.883	.315	.686	.201	.008	.050	.900	.597	36
JAS	.932	.370	.757	.195	.000	.023	.912	.590	24
JAN	.924	.364	.786	.280	.000	.029	.935	.604	30
ESU	.633	.142	.683	.332	.044	.096	.972	.572	36
ASU	.790	.113	.801	.400	.012	.097	.955	.611	74
Total									960

[a] HIB = Hiberic peninsula; BRI = British isles; NWI = Norway and Iceland; FRB = France and Benelux; ALP = Switzerland, Austria, and Southern Germany; GER = western Germany; ITN = northern Italy; SAR = Sardinia; ITS = southern Italy; SIC = Sicily; SFD = Sweden, Finland, and Denmark; EEU = East Germany, Poland, and Czechoslovakia; HUN = Hungary; BGR = Yugoslavia, Balkans, and Greece; MEA = Middle East; IRA = Iran and Afghanistan; INW = northwestern India; INS = southern India and Sri Lanka; INE = northeastern India, Nepal, and Bangladesh; BTH = Burma and Thailand; IMV = Indonesia, Malaysia, and Vietnam; FEA = Far East; CHS = southern China and Taiwan; CHN = northern China and Korea; JAS = southern Japan; JAN = northern Japan; ESU = European USSR; and ASU = Asian USSR.

Results

Joint analysis of eight allele-frequency surfaces (SFO; examples of these surfaces for different datasets are given in Sokal et al. 1989*a*) yielded the map shown in figure 2. The highest values of the SF formed numerous clusters. Some of them fell outside the area in which allele frequencies were measured and were regarded as artifacts of interpolation. Once these were eliminated, 10 clusters were recognized as genetic boundaries and were assigned arbitrary codes, from S1 to S10. In addition, 21 zones of rapid genetic change were recognized by SFI; five of them were significant according to the previously described criterion and were designated as boundaries S11, S12, S15, S16, and S24 (fig. 2). All the boundaries recognized by SFO were significant as well. The results of the GD approach are described in the legend to figure 3.

Gene-frequency boundaries appear to isolate relatively small areas, rather than subdividing Europe and Asia into large regions. Substantial genetic divergence is described mostly in Asia by SF and mainly in Europe but also in Asia by GD. Five regions differ sharply from adjacent areas, namely, Sardinia (boundaries G1 and S16), Hungary (G2 and S1), Japan (G4, G6, and G9),

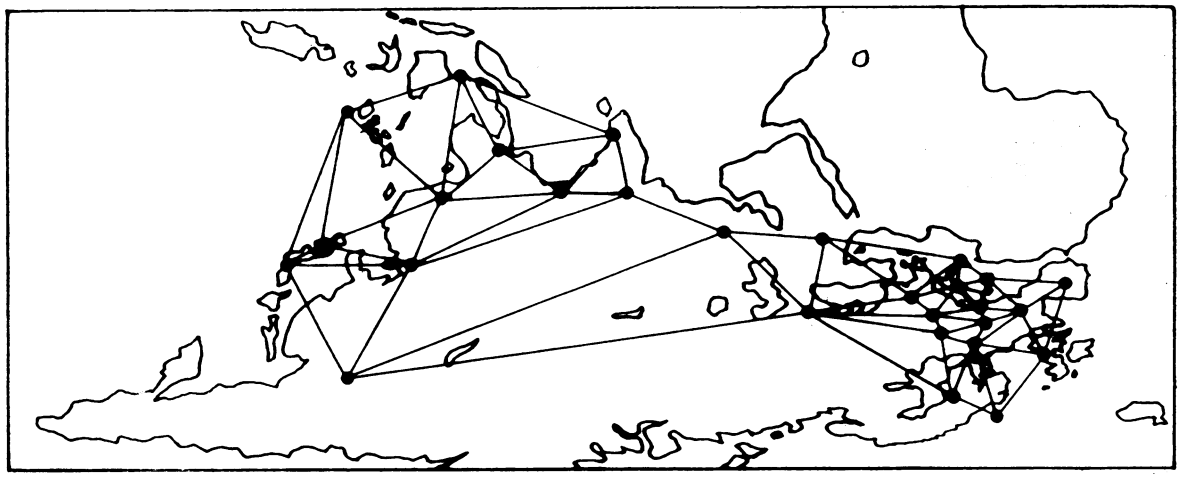

Figure 1 Delaunay triangulation between population groups

eastern Siberia (G4 and S15), and, probably, Sri Lanka (S10). The last-mentioned area does not appear significantly differentiated under the GD approach, because its populations were pooled with south Indian populations; however, there is a significant boundary between the INS group and the rest of India (G5), and wombling locates the area of most rapid change in the Palk Strait (S10).

The two approaches agree also in locating genetic boundaries west of Iran (G4 and S11), west of China (G4, G6, S5, and S24; the upper part of S2 may indicate the same phenomenon; the position of boundaries is approximate in this scarcely sampled zone), between India and Indochina (G5, G10, and S4), between Borneo and the Philippines (G10 and S8), and west of New Guinea (G10 and S9). Note that the SF and GD approaches are not a form of cluster analysis. Populations not separated by significant genetic boundaries are not necessarily similar; genetic variation in the area between them is gradual in mode but not necessarily small in extent.

Discussion

By and large, the results obtained through the two approaches overlap. Some inconsistencies may be due

Figure 2 Genetic boundaries in Europe and Asia, as recognized by the SFO (1–10) and SFI (11, 12, 15, 16, and 24) approaches. These lines are approximately orthogonal to the direction of the vectors representing the slope of the gene-frequency surfaces. The genetic systems contributing to the various boundaries under SFI are as follows: S11—*GLO, PGM,* and *ADA;* S12—*GLO, ACP, AK,* and *ADA;* S15—*GLO, ESD,* and *ACP;* S16—*GLO, ADA,* and *GPT;* and S24—*ESD, ACP,* and *PGD.*

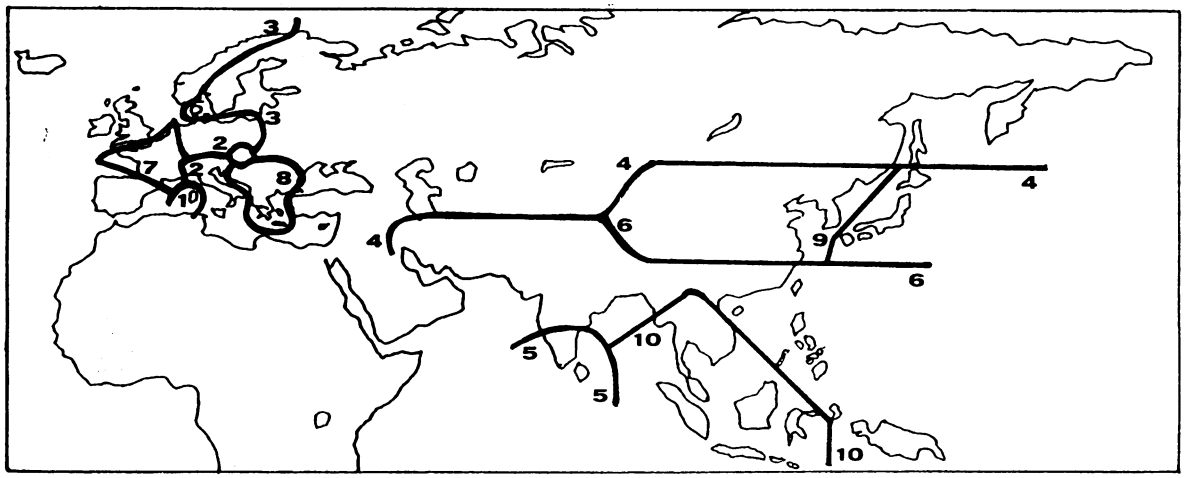

Figure 3 Genetic boundaries in Europe and Asia, as recognized on the basis of the GD approach. The 10 most significant boundaries, ranked from the sharpest, are as follows: G1—all around Sardinia (1); G2—along the western and northern Alps and then all around Hungary (2); G3—between Norway and Sweden, in the North Sea, and east of Poland and Czechoslovakia (3); G4—a long boundary running east-west south of Siberia and west of Iran (4); G5—all around southern India (5); G6—south of Japan and northern China (6); G7—all around France (7); G8—all around Greece and the Balkans (8); G9—between Japan and China (9); G10—all around the populations of Southeast Asia (10).

to the different preliminary treatment of data. Wombling is applied to interpolated surfaces, which are smoother than the real distributions of allele frequencies. The effect of interpolation may be substantial when substantial variation occurs within areas as small as a single pixel (1.5 degrees in the present study). As a consequence, SFO and SFI may have missed some boundaries detected by GD in extensively sampled areas of Europe and the Far East. Examples are G3 and parts of G1, G2, and G9.

The GD approach, on the other hand, is applied to groups of populations, and genetic boundaries falling within a group cannot be detected. In addition, owing to the clumping of data, the pseudolocalities thus defined are more distant from each other than are the original localities sampled, and the rates of genetic change per distance unit are correspondingly decreased, but not to the same extent for all pairs of groups. This reduces the sensitivity of the GD approach when data are averaged across wide areas (e.g., in some sections of Asia) and may account for nondetection of (*a*) parts of boundaries S2 and S3 and (*b*) boundaries S10 and S12. Other boundaries have been displaced; for instance, the genetic differences between eastern Siberia populations and the Chinese population (S15) are reflected in the easternmost section of boundary G4.

Regardless of the statistical method employed, however, it is clear that several allele frequencies vary rapidly and simultaneously in several zones of Europe and Asia. Even in the presence of linkage disequilibrium (which we could neither rule out nor assess), the occurrence of steep clines at different loci in the same localities is not consistent with independent responses of different genetic systems to environmental pressures (Slatkin 1975). Alternatively, multilocus clines may result from demographic phenomena, which in principle should affect all loci in the genome equally (Slatkin 1985, 1987). It may be concluded that the genetic boundaries detected in the present study are likely to reflect isolation and that they correspond to zones across which admixture is limited.

Further support for this view comes from the nonrandom spatial distribution of the observed genetic boundaries. In various cases they can be immediately associated with physical obstacles to gene flow. They occur around most of the sampled islands, such as Sardinia (but not Sicily; see Beretta et al. 1986), Sri Lanka, the Philippines, Japan, and Borneo, as well as in southern Asia and Europe, where migration is constrained by mountain ranges, arid zones, and arms of sea. Conversely, genetic boundaries are not detected in the plains of northern and central Asia.

In general, the zones of abrupt genetic change in Europe and Asia seem maintained by geographical barriers; some of them, however, occur in regions where there are no obvious physical constraints to population admixture (parts of G2, G3, G4, G5, G7, and G8). On the other hand, five of these boundaries do coin-

cide with the lines of contact between linguistically differentiated populations (G2 around Magyar speakers, G4 in the zone between Iranian and Arabic speakers, G5 between Indo-Aryan and Dravidian speakers, G7 between Romance and Germanic speakers, and G8 around southern Slavs and Greek speakers). This finding should be interpreted with some caution, as the GD approach fixes potential boundaries at arbitrary locations, and since some populations of Asia could not be assigned to a unique language family. Nevertheless, these results (1) confirm the similarity between patterns of genetic and linguistic differentiation (Cavalli-Sforza et al. 1988) and (2) suggest that language differences may be sufficient to limit gene flow even in the absence of other isolating factors, yielding a sudden decrease of genetic similarity (as predicted by theory on dispersal barriers; Slatkin 1973; Endler 1977, pp. 80–88; Nagylaki 1988; Nagylaki and Barcilon 1988). Perhaps the main factor impairing gene flow at linguistic boundaries is not the language difference by itself but rather the associated cultural differences, which presumably increase the degree of reproductive isolation (Jorde 1980; Barrantes et al. 1990). However, in a study aimed at quantifying some social and cultural traits, no correlation was found between linguistic and cultural distances (Chakraborty et al. 1976). Therefore, the relative importance of language and other related cultural traits in constraining admixture still remains to be clarified.

As previously mentioned, recent studies (Cavalli-Sforza et al. 1988; Sokal 1988; Sokal et al. 1988, 1989*b*, 1990; Barbujani and Sokal 1990) have pointed out that indices of genetic variation in humans correlate widely with measures of linguistic differentiation and that zones of sharp linguistic and genetic change overlap in Europe. In the absence of detailed ethnohistorical information, it is impossible to test whether this association is simply due to the parallel evolution of linguistic and genetic differences or whether factors associated with language contribute to genetic divergence among populations. The results of the present study support the view that the latter may be the case when multilocus genetic differences are observed, despite the absence of other factors known to maintain them.

What can we infer about human microevolution from the studies in this series? Isolation by distance seems to have caused gradual, nonclinal variation (see Jorde 1980; Cavalli-Sforza 1984; Wijsman and Cavalli-Sforza 1984; Sokal et al. 1989*a*) over areas much larger than individual dispersal distances. Evidence suggesting differential selection exists only for *GLO* and *ESD*. The former is linked with other clinally distributed markers (the HLA loci) which are the likely target of selective pressures (Hedrick et al. 1986), and the latter shows a north-south gradient of allele frequencies (Barbujani 1988*b*) resembling clines that have been attributed to some form of climatic selection (Piazza et al. 1981*b*). Directional migration may account for other areas of clinal variation (Menozzi et al. 1978; Sokal and Menozzi 1982) but has led to gradual genetic change only within limited regions. These regions are separated by steep multilocus clines, which seem due mainly to barriers in the physical as well as in the cultural environment. A similar scheme of interaction between demographic and cultural factors in determining geographical allele-frequency variation has been called the "modified gene-flow" model (as opposed to a model of unconstrained gene flow) by Harding and Sokal (1988). On the whole, the papers in this series concur with the results of comparable studies (e.g., see Smouse and Long 1988) in indicating that the patterns of human allele frequencies tend to reflect past gene flow (Slatkin 1989), its absence, or the anisotropies of its pattern.

Acknowledgments

This study was supported by the Italian Ministry of Education 60% research fund. We wish to thank Peter Smouse and Junhyong Kim for critical reading of the manuscript, Lynn Jorde for providing several references and reprints, and R. R. Sokal, G. A. Danieli, R. Veronesi-Martuzzi, A. Russo, and C. Iodice for comments and suggestions.

References

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

——— (1987*b*) Diversity of some gene frequencies in European and Asian populations. III. Spatial correlogram analysis. Ann Hum Genet 51:345–353

——— (1988*a*) Detecting and comparing the direction of gene-frequency gradients. J Genet 67:129–140

——— (1988*b*) Diversity of some gene frequencies in European and Asian populations. IV. Genetic population structure assessed by the variogram. Ann Hum Genet 52: 215–225

Barbujani G, Milani F (1986) Diversity of some gene frequencies in European and Asian populations: effects of longitude. J Hum Evol 15:61–69

Barbujani G, Oden NL, Sokal RR (1989) Detecting areas of abrupt change in maps of biological variables. Syst Zool 38:376–389

Barbujani G, Sokal RR (1990) Zones of sharp genetic change in Europe are also linguistic boundaries. Proc Natl Acad Sci USA 87:1816–1819

Barrantes R, Smouse PE, Mohrenweiser HW, Gershowitz H, Azofeifa J, Arias TD, Neel JV (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

Batschelet E (1981) Circular statistics in biology. Academic Press, London

Beretta M, Mazzetti P, Frosina G, Schilirò G, Russo A, Russo G, Barrai I (1986) Population structure of eastern Sicily. Hum Hered 36:379–387

Cavalli-Sforza LL (1966) Population structure and human evolution. Proc Soc R Lond [Biol] 164:362–379

——— (1984) Isolation by distance. In: Chakravarti A (ed) Human population genetics: the Pittsburgh symposium. Van Nostrand-Reinhold, New York, pp 229–248

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. Am J Hum Genet 19:233–257

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

Chakraborty R, Blanco R, Rothhammer F, Llop E (1976) Genetic variability in Chilean Indian populations and its association with geography, language, and culture. Soc Biol 23:73–81

Cook PJL, Robson EB, Buckton KE, Jacobs PA, Polani PE (1974) Segregation of genetic markers in families with chromosome polymorphisms and structural rearrangements involving chromosome 1. Ann Hum Genet 37:261–274

Endler JA (1977) Geographic variation, speciation, and clines. Princeton University Press, Princeton, NJ

Farabegoli A, Barbujani G. Diversity of some gene frequencies in European and Asian populations. VI. Geographical patterns of PGM and ACP. Hum Hered (in press)

Green PJ, Sibson R (1977) Computing Dirichlet tessellations in the plane. Comput J 21:168–173

Harding RM, Sokal RR (1988) Classification of European language families by genetic distance. Proc Natl Acad Sci USA 85:9370–9372

Hedrick PW, Thomson G, Klitz W (1986) Evolutionary genetics: HLA as an exemplary system. In: Karlin S, Nevo E (eds) Evolutionary processes and theory. Academic Press, Orlando, FL, pp 583–606

Jorde L (1980) The genetic structure of subdivided human populations: a review. In: Mielke JH, Crawford MH (eds) Current developments in anthropological genetics: theory and methods. Plenum, New York, pp 135–208

Kimura M, Weiss GH (1964) The stepping-stone model of population structure and the decrease of genetic correlation with distance. Genetics 49:561–576

McKusick VA (1988) Mendelian inheritance in man. Johns Hopkins University Press, Baltimore and London

Malécot G (1948) Les mathématiques de l'hérédite. Masson, Paris

Mathéron G (1970) La théorie des variables regionalisées, et ses applications. Centre de Morphologie Mathematique, Fontainebleu, France

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

Monmonier M (1973) Maximum-difference barriers: an alternative numerical regionalization method. Geogr Anal 3:245–261

Morton NE (1982) Kinship and inbreeding in populations of Middle Eastern origin and controls. In: Crawford MH, Mielke JH (eds) Current developments in anthropological genetics: ecology and population structure. Plenum, New York, pp 449–466

Morton NE, Yee S, Harris DE, Law R (1971) Bioassay of kinship. Theor Popul Biol 2:507–524

Nagylaki T (1988) The influence of spatial inhomogeneities on neutral models of geographic variation. I. Formulation. Theor Popul Biol 33:291–310

Nagylaki T, Barcilon V (1988) The influence of spatial inhomogeneities on neutral models of geographic variation. II. The semi-infinite linear habitat. Theor Popul Biol 33:311–343

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

——— (1981*b*) Synthetic gene frequency maps of man and selective effects of climate. Proc Natl Acad Sci USA 78:2638–2642

Slatkin M (1973) Gene flow and selection in a cline. Genetics 75:733–756

——— (1975) Gene flow and selection in a two-locus system. Genetics 81:789–802

——— (1985) Gene flow in natural populations. Annu Rev Ecol Syst 16:393–430

——— (1987) Gene flow and the geographic structure of natural populations. Science 236:787–792

——— (1989) Population structure and evolutionary progress. Genome 31:196–202

Slatkin M, Barton NH (1989) A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43:1349–1368

Smouse PE, Long JC (1988) A comparative F-statistics analysis of the genetic structure of human populations from lowland South America and highland New Guinea. In: Weir BS, Eisen EJ, Goodman MM, Namkoong G (eds) Proceedings of the Second International Conference on Quantitative Genetics. Sinauer, Sunderland, Mass, pp 32–46

Sokal RR (1988) Genetic, geographic, and linguistic distances in Europe. Proc Natl Acad Sci USA 85:1722–1726

Sokal RR, Harding RM, Oden NL (1989*a*) Spatial patterns of human gene frequencies in Europe. Am J Phys Anthropol 80:267–294

Sokal RR, Menozzi P (1982) Spatial autocorrelation of HLA frequencies in Europe supports demic diffusion of early farmers. Am Nat 119:1–17

Sokal RR, Oden NL, Legendre P, Fortin MJ, Kim J, Thomson BA, Vaudor A, et al (1990) Genetics and language in European populations. Am Nat 135:157–175

Sokal RR, Oden NL, Legendre P, Fortin MJ, Kim J, Vaudor A (1989*b*) Genetic differences among language families in Europe. Am J Phys Anthropol 79:489–502

Sokal RR, Oden NL, Thomson BA (1988) Genetic change across language boundaries in Europe. Am J Phys Anthropol 76:337–361

Wijsman EM, Cavalli-Sforza LL (1984) Migration and genetic population structure with special reference to humans. Annu Rev Ecol Syst 15:279–301

Womble WH (1951) Differential systematics. Science 114: 315–322

Wood JW (1987) The genetic demography of the Gainj of Papua New Guinea. II. Determinants of effective population size. Am Nat 129:165–187

Wright S (1931) Evolution in Mendelian populations. Genetics 16:97–159

——— (1943) Isolation by distance. Genetics 28:114–138

——— (1951) The genetical structure of populations. Ann Eugen 15:323–354