Velasco E, Valero C, Valero A, Moreno F, Hernandez-Chico C (1996) Molecular analysis of the SMN and NAIP genes in Spanish spinal muscular atrophy (SMA) families and correlation between number of copies of cBCD541 and SMA phenotype. Hum Mol Genet 5:257–263

Viollet L, Bertrandys S, Bueno-Brunialiti AL, Lefebvre S, Burlet P, Clermont O, Cruard C, et al (1997) cDNA isolation expression and chromosomal localization of the mouse survival motor neuron gene (SMN). Genomics 40:185–188

Wang CH, Xu J, Carter TA, Ross BM, Dominski MK, Bellcross CA, Penchaszadeh GK, et al (1996) Characterization of survival motor neuron (SMN^T) gene deletions in asymptomatic carriers of spinal muscular atrophy. Hum Mol Genet 5: 359–365

Address for correspondence and reprints: Dr. Arthur Burghes, Department of Neurology, 654 Upham Drive, Columbus, OH 43210. E-mail: burghes.1@osu.edu

0002-9297/98/6202-0036$02.00

Am. J. Hum. Genet. 62:488–491, 1998

Evidence for Paleolithic and Neolithic Gene Flow in Europe

To the Editor:

In recent Letters to the Editor, Cavalli-Sforza and Minch (1997) and Richards et al. (1997) discuss the relative contributions of the first Paleolithic colonizers of Europe, and of later Neolithic immigrants, to the gene pool of current Europeans. Using the method of median networks (Bandelt et al. 1995), Richards et al. (1996) demonstrated that most mitochondrial lineages coalesce at ancestors who presumably lived in the Paleolithic period, which, in Europe, means >10,000 years ago. Through an analysis of the geographic distribution of these lineages, they reached the conclusion that most mitochondrial alleles spread in Europe prior to the Neolithic period. Two implications of this finding were that (1) farming was essentially a local development, the spread of which was not accompanied by extensive gene flow, and (2) the gradients of allele frequencies described in many studies (starting with Menozzi et al. [1978] and reviewed in Cavalli-Sforza et al. [1994]) were not due to a Neolithic demic diffusion from the Near East (Ammerman and Cavalli-Sforza 1984), as is generally believed. Richards et al. (1996) interpreted the results of a simulation study of various population-expansion mechanisms (Barbujani et al. 1995) as supporting a Paleolithic origin of these clines.

Cavalli-Sforza and Minch (1997) argued that sequences of the mtDNA hypervariable region are not suitable for reconstructing evolutionary processes at this scale, because the high mutation rates at some sites cause an excess of random noise. In addition, a high female mobility might have blurred some previously existing geographic patterns. They suggested that a figure of ~25% might realistically represent the contribution of Neolithic immigrants to the gene pool of Europeans, because, in principal-component analyses of allele frequencies, a clinal component accounts for one quarter of the genetic variance (Menozzi et al. 1978; Piazza et al. 1995). If that were the case, there would be little overall disagreement; given the approximate nature of any such estimates, the figure (15%) proposed by Richards et al. (1997) may not differ significantly. We would like to suggest a third possibility—namely, that the available mitochondrial data do not contradict a much larger Neolithic contribution and that envisaging the current European gene pool as essentially a product of an Upper Paleolithic colonization may create more problems than it solves.

There are four traditional reasons to believe that there was a major Neolithic contribution to the European gene pool: (1) the continentwide gradients of allele frequencies; (2) their correlation with the archaeological record; (3) their overlapping with areas defined by linguistic criteria; and (4) their similarity to the gradients theoretically predicted under, or generated in simulation studies of, a model of demic diffusion. None of these pieces of evidence is proof, but in this field there is little that one can really prove. The point, at this stage, is to find the simplest explanation that accounts for most (or, possibly, for all) observed population characteristics. Of course, speaking of Paleolithic versus Neolithic processes is an oversimplification of phenomena that were certainly more complicated. However, such a highly schematic opposition is useful for the sake of clarity.

As for the gradients detected for roughly one third of the alleles studied in Europe (Sokal et al. 1989) (point 1), few doubt that they result from some form of population movement. Indeed, random genetic drift alone cannot generate nonrandom patterns on such a broad scale, and major selective effects on many independent loci appear unlikely (Ammerman and Cavalli-Sforza 1984; but see also Fix 1996). The problem is *when* those movements took place. As Richards et al. (1996, 1997) pointed out, the correlation with archaeological gradients (point 2) and, specifically, with the first evidence of farming activities (Sokal et al. 1991) now seems less cogent. Indeed, evidence is emerging that, not only in the Neolithic but also in the Paleolithic period, the main population movements occurred along a southeast-northwest axis (Richards et al. [1997] and references therein). If so, whatever the relative importance of the two temporal phases, both should have determined similar clines of gene frequencies. On the contrary, however, if it were shown that Paleolithic populations moved

largely in other directions, the longitudinal clines should still be taken as evidence for a major Neolithic contribution to the European gene pool.

What does not seem equally easy to reconcile with both views is the linguistic evidence (point 3). Renfrew (1987) proposed that the genes of Anatolian and Near Eastern populations, the technologies for farming and animal breeding, and the Indo-European languages were brought to Europe in the course of the same expansion, starting some 10,000 years ago. Renfrew (1991) also suggested that, in three other linguistically related areas (where Afro-Asiatic, Altaic, and Elamo-Dravidian languages are or were spoken), farming technologies could have spread at once with languages and with the people who spoke them. When a large-scale analysis of genetic data was performed, all four areas identified by Renfrew showed, for most genes, highly significant clinal patterns (with a lower statistical significance for the Afro-Asiatic–speaking area), some of which were not apparent if linguistic affiliations were disregarded. The geographic limits of those clines corresponded to major language barriers, and similar clines were not observed for language families that were not supposed to have spread during the Neolithic period (Barbujani and Pilastro 1993; Barbujani et al. 1994).

In synthesis, gene-frequency clines correspond to linguistic areas that, presumably, were established in the Neolithic period, not only in Europe, but in three other regions of Eurasia and North Africa, where clines were sought on the basis of linguistic evidence and were detected only when the different linguistic groups were analyzed separately. The main language families of Eurasia are unlikely to have spread in earlier time periods; several linguists are very reluctant to accept dates much earlier than 5,000 years ago (Coleman 1988). Therefore, the correlation between gene frequencies and languages seems difficult to explain if one assumes a limited genetic contribution of Neolithic farmers. One should imagine that, in four regions of Eurasia and North Africa, the routes of spread of four Neolithic language groups overlapped by chance with the routes followed by the first Paleolithic colonizers; this does not seem a parsimonious hypothesis.

Simulation results (point 4) also seem to be in better agreement with a comparatively recent dispersal of the Europeans' ancestors. Rendine et al. (1986) and Barbujani et al. (1995) showed that Neolithic population expansions that started in the Near East and were accompanied by little admixture do produce clines that are not observed under alternative models. Although the relative contributions of Paleolithic and Neolithic groups are difficult to calculate from the published results, in Rendine et al. (1986), population sizes were increased 3–25 times in the process, largely because of demographic growth among Neolithic farmers. As we understand it, this means that the Paleolithic contribution can be estimated to be between one third and a much lower value. The second study showed that simulated allele-frequency patterns significantly resemble the observed ones if the expanding population largely or entirely replaced any pre-existing settlers (Barbujani et al. 1995). Therefore, all simulation results obtained so far are certainly consistent with diffusion of Neolithic populations who were ancestral to most current Europeans. Conversely, the alternative view, which involves a series of founder effects during the initial colonization of Europe and which is conceivable in principle (Barbujani et al. 1995; Richards et al. 1996), has not yet been formally tested.

Can one say that allele frequencies support a Neolithic model, whereas DNA data support a Paleolithic model? We do not think so. The main finding of the study by Richards et al. (1996) is that most mutations that characterize the mitochondrial haplotypes of present-day Europeans occurred in the Paleolithic period. We do not question that conclusion, but we do not think that the age of a group of haplotypes can be mechanically equated to the age of the population from which they came, especially if these haplotypes are also found elsewhere. Certain alleles of the HLA-DRB1 locus arose millions of years ago and are shared by our species and others (see Ayala 1995). Their presence in European populations does not imply that Europe was colonized prior to the separation between humans and chimps. Similarly, suppose that some Europeans colonize Mars next year: If they successfully establish a population, the common mitochondrial ancestor of their descendants will be Paleolithic. But it would not be wise for a population geneticist of the future to infer from that a Paleolithic colonization of Mars.

The average coalescence time of two sequences sampled from two diverging populations is, in general, older, or much older, than the split of the groups. Unless a group colonizing a new territory passes through a strong and long-lasting bottleneck, part of its initial diversity will be maintained (Nei et al. 1975; Tajima 1983). Therefore, the coalescence times inferred from samples of its descendants will be close to the coalescence times of the population of origin, and these times will consistently overestimate the age of the derived populations. Clearly, inferences from population history must be based on measures of genetic diversity *between populations*, not *between molecules*. Among the appropriate statistics is Nei's d_{AB} distance—the average sequence difference between the haplotypes of two different samples, diminished by the average pairwise difference within samples (Nei 1987). In populations at equilibrium, this distance is linearly related, through the mutation rate, to the divergence time between pairs of populations. If one calculates d_{AB} from the mitochondrial data of Rich-

ards et al. (1996) and uses the divergence rate proposed in that article (1/10,500 years), Middle Easterners appear to be separated from all other European samples by <4,000 years. Although recent gene flow has probably reduced the levels of population differentiation, these figures do not point to population splits that predate the Neolithic period in Europe.

A link between molecules and populations can be established. However, to do that, one should focus on haplogroups that are geographically restricted to the area of interest—Europe, in our case—a point that Richards et al. (1996) did not thoroughly consider. Among the haplogroups recognized by Richards et al., three have not yet been observed outside Europe; they coalesce, respectively, at $6,000 \pm 2,000$ (haplogroup 2A-C), $12,500 \pm 6,000$ (haplogroup 2A-W), and $17,400 \pm 2,000$ (haplogroup 4) years ago. Once again, these dates are not in conflict with a mostly Neolithic origin of the European populations.

When considered at the population level, other DNA polymorphisms suggest a recent separation of European groups. In a study of seven DNA polymorphisms, we used a measure of microsatellite diversity (Goldstein et al. 1995) to estimate times since population divergence in Europe. Geographic clines were broad and significant, and none of the comparisons between European and Near Eastern populations led to estimates $\geqslant$10,000 years, with the single, predictable exception of Saami (Lapps) (Chikhi et al., in press). We are more than ready to accept that future analyses of larger numbers of loci may somewhat modify this conclusion. However, at present, all the molecular data analyzed at a continental scale appear fully consistent with a major Neolithic phenomenon that left its marks on present levels of population diversity. The European distribution of two Y-chromosome polymorphisms is in nice agreement with this conclusion (Semino et al. 1996). It is quite possible that the gene pools of certain isolated groups, Saami (Lahermo et al. 1996), Ladins (Stenico et al. 1998), and Basques (Calafell and Bertranpetit 1994) among them, contain a greater proportion of alleles derived from the first, Paleolithic colonizers. But this seems the exception, not the rule.

In synthesis, we believe that any model that suggests a largely Paleolithic origin of the European gene pool should incorporate an explanation of (1) the correlation between genetic data and linguistic patterns that are highly unlikely to have been established prior to the Neolithic period; (2) the simulation results showing that clines are generated by population processes supposed to have occurred in the European Neolithic period; and (3) the admittedly limited, but internally consistent, evidence from studies of microsatellite diversity. On the contrary, all these findings, as well as the distribution of mitochondrial diversity (to the clarification of which Richards et al. have contributed), can easily fit within a model of extensive demographic replacement associated with the dispersal of Near Eastern farming populations.

Acknowledgments

We thank Hans Bandelt, Peter Forster, Luca Cavalli-Sforza, and Eric Minch for giving us access to their unpublished letters and for discussing them with us. We also thank Italo Barrai for fruitful discussion. Of course, these individuals do not share all the views expressed here.

Guido Barbujani,[1] Giorgio Bertorelle,[1,2] and Lounès Chikhi[1,2]

[1]*Department of Biology, University of Ferrara, Ferrara, and*
[2]*Department of Biology, University of Padova, Padova*

References

Ammerman AJ, Cavalli-Sforza LL (1984) The Neolithic transition and the genetics of populations in Europe. Princeton University Press, Princeton

Ayala FJ (1995) The myth of Eve: molecular biology and human origins. Science 270:1930–1936

Bandelt HJ, Forster P, Sykes BC, Richards MB (1995) Mitochondrial portraits of human populations using median networks. Genetics 141:743–753

Barbujani G, Pilastro A (1993) Genetic evidence on origin and dispersal of human populations speaking languages of the Nostratic macrofamily. Proc Natl Acad Sci USA 90: 4670–4673

Barbujani G, Pilastro A, DeDomenico S, Renfrew C (1994) Genetic variation in North Africa and Eurasia: Neolithic demic diffusion vs Paleolithic colonization. Am J Phys Anthropol 95:137–154

Barbujani G, Sokal RR, Oden NL (1995) Indo-European origins: a computer-simulation test of five hypotheses. Am J Phys Anthropol 96:109–132

Calafell F, Bertranpetit J (1994) Principal component analysis of gene frequencies and the origin of Basques. Am J Phys Anthropol 93:201–215

Cavalli-Sforza LL, Menozzi P, Piazza A (1994) The history and geography of human genes. Princeton University Press, Princeton

Cavalli-Sforza LL, Minch E (1997) Paleolithic and Neolithic lineages in the European mitochondrial gene pool. Am J Hum Genet 61:247–251

Coleman C (1988) Comment on *Archaeology and Language: The Puzzle of Indo-European Origins* by C Renfrew. Curr Anthropol 29:449–453

Chikhi L, Destro-Bisol G, Pascali V, Baravelli V, Dobosz M, Barbujani G. Clinical variation in the nuclear DNA of Europeans. Hum Biol (in press)

Fix A (1996) Gene frequency clines in Europe: demic diffusion or natural selection? J R Anthropol Inst 2:625–643

Goldstein DB, Ruiz-Linares A, Cavalli-Sforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. Genetics 139:463–471

Lahermo P, Sajantila A, Sistonen P, Lukka M, Aula P, Peltonen

L, Savontaus ML (1996) The genetic relationship between the Finns and the Finnish Saami (Lapps): analysis of nuclear DNA and mtDNA. Am J Hum Genet 58:1309–1322
Menozzi P, Piazza A, Cavalli-Sforza LL (1978) Synthetic maps of human gene frequencies in Europeans. Science 201: 786–792
Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York
Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. Evolution 29: 1–10
Piazza A, Rendine S, Minch E, Menozzi P, Mountain J, Cavalli-Sforza LL (1995) Genetics and the origin of European languages. Proc Natl Acad Sci USA 92:5836–5840
Rendine S, Piazza A, Cavalli-Sforza LL (1986) Simulation and separation by principal components of multiple demic expansions in Europe. Am Nat 128:681–706
Renfrew C (1987) Archaeology and language: the puzzle of Indo-European origins. Jonathan Cape, London
———(1991) Before Babel: speculations on the origins of linguistic diversity. Camb Archaeol J 1:3–23
Richards M, Côrte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, Demaine A, Papiha S, et al (1996) Paleolithic and Neolithic lineages in the European mitochondrial gene pool. Am J Hum Genet 59:185–203
Richards M, Macaulay V, Sykes B, Pettit P, Forster P, Hedges R, Bandelt HJ (1997) Reply to Cavalli-Sforza and Minch. Am J Hum Genet 61:251–254
Semino O, Passarino G, Brega A, Fellous M, Santachiara-Benerecetti AS (1996) A view of the Neolithic demic diffusion in Europe through two Y chromosome–specific markers. Am J Hum Genet 59:964–968
Sokal RR, Harding RM, Oden NL (1989) Spatial patterns of human gene frequencies in Europe. Am J Phys Anthropol 80:267–294.
Sokal RR, Oden NL, Wilson C (1991) New genetic evidence for the spread of agriculture in Europe by demic diffusion. Nature 351:143–145
Stenico M, Nigro L, Barbujani G (1998) Mitochondrial lineages in Ladin-speaking communities of the eastern Alps. Proc R Soc Lond B Biol Sci 265:1–7
Tajima F (1983) Evolutionary relationships of DNA sequences in finite populations. Genetics 105:437–460

Address for correspondence and reprints: Dr. Guido Barbujani, Dipartimento di Biologia Università di Ferrara, via L. Borsari 46, I-44100 Ferrara, Italy. E-mail: bjg@ifeuniv.unife.it

0002-9297/98/6202-0037$02.00

Am. J. Hum. Genet. 62:491–492, 1998

Reply to Barbujani et al.

To the Editor:

We agree entirely with Barbujani et al. (1998 [in this issue]) that the age of a group of haplotypes cannot be mechanically equated to the age of the population from which they come and that such an uncritical equation would artificially elevate the estimated age of the population under study. However, our analysis (Richards et al. 1996, p. 194) focuses not simply on haplogroups, but on haplotypes within haplogroups. Such an analysis depends critically on the correct identification—by cross-population comparison of lineages—of all of the major founder haplotypes, which can then be used as a baseline from which to date the founder events associated with each cluster of haplotypes. This is exemplified in our paper by the identification of a number of distinct founder haplotypes in lineage group 2A, picked out on the basis of their presence as shared ancestral nodes in the European and Near Eastern phylogenies, which root deeply (during the Upper Paleolithic period) in the Near Eastern data but which have accumulated only a small amount of variation—equivalent to ~10,000 years or so—within Europe. This suggests, to us, expansion into Europe from the Near East during the Neolithic period. Of the various lineage clusters that we identified in Europe and the Near East, only group 2A showed this pattern; other clusters did not show evidence of recent founder events within Europe.

We believe that a phylogeographic analysis such as this—which is indeed based on molecules rather than on populations—is capable of a much finer resolution than one based on distance statistics, such as that suggested by Barbujani et al. (1998). Moreover, the particular statistic used is misleading, as it is based on a model of populations of constant size at mutation-drift equilibrium, which is patently unsuitable for application to Europe and the Near East. However, an important weakness of our published analysis is the meager volume of comparative data from the Near East: essentially 42 individuals, mostly from the Arabian peninsula. In subsequent work, we are extending the analysis to a much larger sample from southwestern Asia, to improve our confidence that most founder haplotypes have been identified. We aim to report our conclusions in the near future.

MARTIN RICHARDS AND BRYAN SYKES
Department of Cellular Science
Institute of Molecular Medicine
University of Oxford
Oxford

References

Barbujani G, Bertorelle G, Chikhi L (1998) Evidence for Paleolithic and Neolithic gene flow in Europe. Am J Hum Genet 62:000–000 (in this issue)
Richards M, Côrte-Real H, Forster P, Macaulay V, Wilkinson-