caulay et al. 1997 [in press]). A more extensive pedigree study by Bendall et al. (1996) confirms the orthodox mtDNA mutation rate.

We do agree with the comment by Bianchi and Bailliet that more research is needed in order to understand the mutational mechanisms acting on mtDNA and, specifically, on np 16519; for example, it is intriguing that np 16519 is virtually "frozen" in some haplogroups, such as group B in Amerinds and group T in Caucasoids, and in both the Caucasoid and the Amerind branches of group X.

PETER FORSTER,[1] ROSALIND HARDING,[3] ANTONIO TORRONI,[4] AND HANS JÜRGEN BANDELT[2]
[1]*Heinrich-Pette-Institut and* [2]*Mathematisches Seminar, Universität Hamburg, Hamburg;* [3]*Institute of Molecular Medicine, John Radcliffe Hospital, Oxford; and* [4]*Dipartimento di Genetica e Biologia Molecolare, Università di Roma "La Sapienza," Rome*

References

Bailliet G, Rothhammer F, Carnese FR, Bravi CM, Bianchi NO (1994) Founder mitochondrial haplotypes in Amerindian populations. Am J Hum Genet 55:27–33

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

Batista O, Kolman CJ, Bermingham E (1995) Mitochondrial DNA diversity in the Kuna Amerinds of Panama. Hum Mol Genet 4:921–929

Bendall KE, Macaulay VA, Baker JR, Sykes BC (1996) Heteroplasmic point mutations in the human mtDNA control region. Am J Hum Genet 59:1276–1287

Bianchi NO, Rothhammer F (1995) Reply to Torroni and Wallace. Am J Hum Genet 56:1236–1238

Chen Y-S, Torroni A, Excoffier L, Santachiara-Benerecetti AS, Wallace DC (1995) Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. Am J Hum Genet 57:133–149

Easton RD, Merriwether DA, Crews DE, Ferrell RE (1996) mtDNA variation in the Yanomami: evidence for additional New World founding lineages. Am J Hum Genet 59: 213–225

Forster P, Harding R, Torroni A, Bandelt H-J (1996) Origin and evolution of Native American mtDNA variation: a reappraisal. Am J Hum Genet 59:935–945

Graven L, Passarino G, Semino O, Boursot P, Santachiara-Benerecetti S, Langaney A, Excoffier L (1995) Evolutionary correlation between control region sequence and restriction polymorphisms in the mitochondrial genome of a large Senegalese Mandenka sample. Mol Biol Evol 12:334–345

Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, Tajima K (1993) Peopling of the Americas, founded by four major lineages of mitochondrial DNA. Mol Biol Evol 10:23–47

Howell N, Bindoff LA, McCullough DA, Kubacka I, Poulton J, Mackey D, Taylor L, et al (1991*a*) Leber hereditary optic neuropathy: identification of the same mitochondrial ND1 mutation in six pedigrees. Am J Hum Genet 49:939–950

Howell N, Kubacka I, Mackey DA (1996) How rapidly does the human mitochondrial genome evolve? Am J Hum Genet 59:501–509

Howell N, Kubacka I, Xu M, McCullough DA (1991*b*) Leber hereditary optic neuropathy: involvement of the mitochondrial ND1 gene and evidence for an intragenic suppressor mutation. Am J Hum Genet 48:935–942

Macaulay VA, Richards MB, Forster P, Bendall KE, Watson E, Sykes BC, Bandelt H-J (1997) mtDNA mutation rates—no need to panic. Am J Hum Genet 61 (in press)

Merriwether DA, Hall WW, Vahlne A, Ferrell RE (1996) mtDNA variation indicates Mongolia may have been the source for the founding population for the New World. Am J Hum Genet 59:204–212

Scozzari R, Cruciani F, Santolamazza P, Sellitto D, Cole DEC, Rubin LA, Labuda D, et al (1997) mtDNA and Y chromosome–specific polymorphisms in modern Ojibwa: implications about the origin of their gene pool. Am J Hum Genet 60:241–244

Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, et al (1996) Classification of European mtDNAs from an analysis of three European populations. Genetics 144:1835–1850

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, et al (1993*a*) Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet 53:563–590

Torroni A, Sukernik RI, Schurr TG, Starikovskaya YB, Cabell MF, Crawford MH, Comuzzie AG, et al (1993*b*) mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. Am J Hum Genet 53: 591–608

Torroni A, Wallace DC (1995) mtDNA haplogroups in Native Americans. Am J Hum Genet 56:1234–1236

Vigilant L (1990) Control region sequences from African populations and the evolution of human mitochondrial DNA. PhD thesis, University of California, Berkeley

Address for correspondence and reprints: Dr. Peter Forster, Heinrich-Pette-Institut, Universität Hamburg, Martinistrasse 52, 20251 Hamburg, Germany. E-mail: forster@hpi.uni-hamburg.de

0002-9297/97/6101-0037$02.00

Am. J. Hum. Genet. 61:247–251, 1997

Paleolithic and Neolithic Lineages in the European Mitochondrial Gene Pool

To the Editor:

In a recent analysis of lineage groups derived from European and Middle Eastern samples of mtDNA D-loop sequences, Richards et al. (1996) have stated that most extant European mtDNA lineages predate the Neolithic

agricultural expansion, with only minor genetic contributions from the Middle East. They conclude that the spread of agriculture was an essentially indigenous development with very little demic diffusion. A reanalysis of their data confirms both the lack of genetic diversity within Europe and the sharp difference between Europe and the Middle East; this conflicts, however, with other results, which are based on autosomal and Y-chromosome frequencies and which indicate a more gradual cline from the Middle East to the extreme west of Europe, similar to that found in the archaeological record. This discrepancy casts doubt on the conclusions of Richards et al., which are based on mtDNA sequence data, but it can be resolved by consideration of the high mutation rates in the D-loop and the differential patterns of male and female gene flow due to cultural practices such as virilocality and hypergamy.

The introduction of D-loop analysis by Allan Wilson and his colleagues (Vigilant et al. 1989, 1991; also see Horai and Hayasaka 1990) was an attempt to exploit high sequence variation to obtain evolutionary information by the direct sequencing of one or two small DNA segments. After some years of experience, however, doubts have arisen that D-loop analysis is a useful tool for the analysis of population similarities. In this laboratory we have had discordance between trees generated by use of the mtDNA D-loop and those generated by use of other markers. Branches leading to populations outside Africa seem definitely shorter and less distinguishable than those of African populations (Mountain et al. 1995). Similar results have been published by Jorde et al. (1995). On the positive side, one must acknowledge that D-loop sequence comparisons did generate the demographic-growth model on the basis of the distribution of the number of nucleotide differences between pairs of individuals (Slatkin and Hudson 1991; Harpending and Rogers 1992), but this can probably be done with any set of closely linked markers. The recent study by Richards et al. (1996) is another example in which D-loop analysis is misleading for the study of very simple problems.

Richards et al. (1996, table 4), in an article with the same title as that of this letter, analyze sequences of segment I of the D-loop of 757 individuals from 15 European countries, plus others from Turkey and the Middle East, and they provide a 15×6 (and a 15×5) table of frequencies of lineage groups. The full haplotypes do not add much further information beyond that derived from the lineage groups, because of the noise generated by high mutation rates in the D-loop.

The authors analyze the statistical significance of the differences between pairs of populations by a permutation test, and the only clear result is that the Middle East data are significantly different from the rest, which derive from all of Europe plus Turkey. The only other consistent difference found is between Basques and the other populations. This is a well-known result, although with nuclear markers there is ample evidence that other populations tested in this sample, such as Sardinians, show greater difference from the rest of Europe than do Basques, whereas other mtDNA data (Bertranpetit et al. 1995) show no significant difference between Basques and the rest of Europe. This latter result may reflect the small number of Basques and Sardinians tested with mtDNA.

With all respect for nonparametric tests, there is some merit in the use of the parametric χ^2 test. The theoretical sampling distribution behind it—the positive binomial—has a strong and distinguished background in genetic applications of this kind. In the G version (Sokal and Rohlf 1981), χ^2 is also more resistant to the effect of small absolute frequencies, which, if anything, would tend too easily to give significant results.

For the full data table given by Richards et al. in their table 4, with 15 populations (including Turkey and the Middle East) and five lineages, the χ^2 is 97.60 with 56 df, $P = .0049$. When the most distant and least relevant population, that of the Middle East, is removed, χ^2 becomes 56.23 with 52 df, and P is now .32, not significant. When the split of lineage 2 into 2A and 2B is introduced, the χ^2 including the Middle Eastern population is 113.02 with 70 df, again highly significant ($P = .00086$). When the Middle Eastern population is eliminated, the χ^2 (82.37 with 65 df) is again not significant ($P = .073$).

It seems inevitable to conclude that, with the numbers of individuals examined, there is, in Richards et al.'s data, no proof of genetic variation among European populations, apart from the difference between the Middle Eastern population and all others considered by the authors. The rather general conclusions drawn by Richards et al. (1995, p. 197)—namely, that "the majority of modern Europeans are descended from the settlement of Europe by anatomically modern humans during the Upper Paleolithic" and that "the overall demographic influence on modern Europeans [of farmers from the Middle East] is relatively small"—are not warranted by their data.

Richards et al. (1995, p. 197) also state that the papers that have suggested the Neolithic immigration "have emphasized, though not precisely quantified, the genetic contribution of the Neolithic immigrants." This is not correct. In the first such paper (Menozzi et al. 1978), based on 38 genes, the first principal component was shown to account for 27% of the total genetic variation; in the second such paper, based on 95 genes (Cavalli-Sforza et al. 1994), it accounted for 28% of the variation. In the third paper (Piazza et al. 1995), which includes more data from the Caucasus, the percentage of total variation explained is 26%. This percentage is not—or not necessarily—the same as the genome pro-

portion contributed by farmers; but it is probably not very far from it, in which case it is clear that the proportion of genes contributed by Neolithic farmers, although not the absolute majority of the European genome, is certainly the largest influence.

There are good reasons, however, why one would find a difference between gene flow measured by mtDNA and that measured by autosomes. The former is tied to the migration of women. Y chromosomes, markers of which are becoming available, reflect the migration of men. A geographic map of two Y-chromosome markers that show considerable difference between the east and west of Europe has been published recently (Semino et al. 1996). The geographic gradient, across Europe, of the two markers is rather similar to that observed from the pattern of the first principal component of autosomes. In figure 1 we compare the latter pattern (fig. 1*B*, from the most recent data set [i.e., Piazza et al. 1995]), with that of the first principal component of the Y-chromosome markers, which we calculated on the basis of the Semino et al. data (fig. 1*C*). There is a high correlation, and both patterns are highly correlated with the dates of first arrival of Neolithic farmers, as inferred from archaeological observations on the spread of farming in Europe (fig. 1*A*).

One might expect that data based on autosomes,

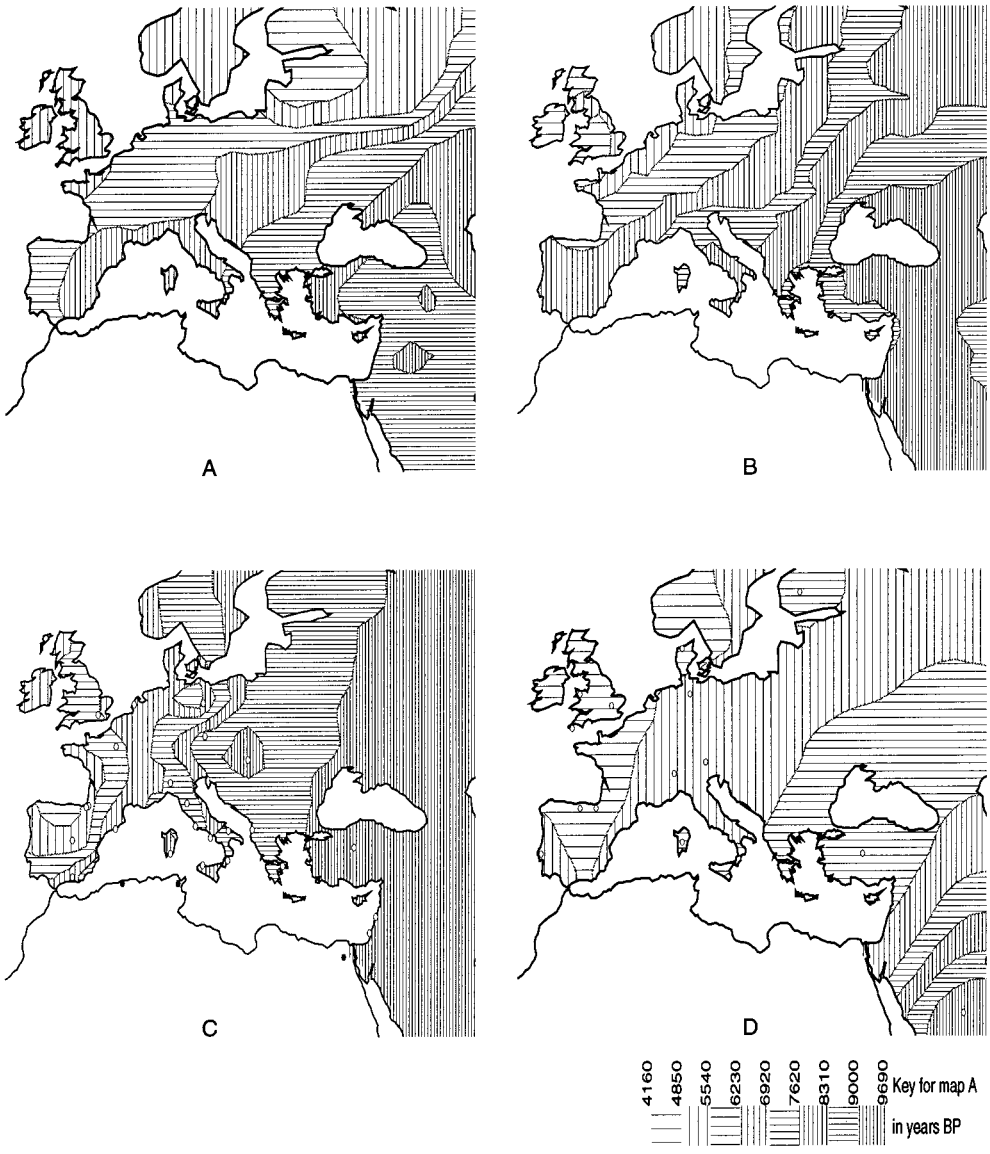

Figure 1 *A,* Spread of agriculture in the Neolithic, based on ^{14}C dating at 106 archeological sites. *B,* First principal component for 91 autosomal markers (30% of variance explained); geographic coverage was modified for comparison with other figures. *C,* First principal component for 2 Y-chromosome microsatellite polymorphisms (93%). *D,* First principal component for 63 mtDNA single-nucleotide polymorphisms (23%). The circles in *C* and *D* indicate locations of data samples.

which average the migration of the two sexes, would be intermediate between data based on mtDNA and data based on the Y chromosome. This is very approximately true, since the first principal component of the mtDNA data of Richards et al. (fig. 1*D,* which uses all 63 polymorphic sites) shows a pattern slightly similar to those given by archaeology, autosomes, and Y chromosomes, although the gradient is almost flat over most of Europe, except for a sharp pole in the Middle East and its diffuse opposite in the extreme west of Europe. The Y-chromosome data are limited to three alleles, two of which have probably extreme behavior, and information on a wider range of variants would be necessary. The mtDNA D-loop is probably plagued both by noise, which is due to excessively high mutation rates, and by an unknown factor, mentioned above and probably affecting all mtDNA data, which decreases genetic distances among populations outside Africa. One may speculate that this is due, at least in part, to heteroplasmy, which will determine a segregation lag of recently appearing mutants. This is likely to affect particularly those populations that have separated more recently from the rest, and thus non-African populations are more likely to show lower divergence among themselves than are African populations. There have been repeated observations pointing to the presence of heteroplasmy of mtDNA (e.g., the latest one, in Bendall et al. 1996). But it is not known whether the average number of mtDNA chromosomes per cell during the germinal cell cycle can result in heteroplasmy sufficiently high to explain the depression of genetic distances among non-African populations by segregation lag of new mutants.

The very low heterogeneity observed for mtDNA among European populations is likely to have other causes as well, tied to the special pattern of female migration compared with that of males. Two factors seem potentially important in the human species. One of them is a tendency, at marriage, for women to migrate more than men, in that it is more often women who relocate to join their spouse; in anthropological terminology, marriage is more often than not patri- or virilocal. This is believed to have been true even for hunter-gatherers (Ember 1978; Hewlett 1996), as well as for farmers, in whom patrilocality is a consequence of preferential inheritance of the land by sons. This makes women, on average, genetically more mobile than men, even though their average daily displacement may be less than that of men. Another factor that may have been especially active during the spread of farmers is female hypergamy. This is the condition in which the chance of marrying into a higher social class is greater for women than for men; in traditional societies—for example, in many Indian castes—this opportunity is often available only to women. Hypergamy is still noted today in societies in which the spread of farmers among hunter-gatherers is still happening—for example, in the tropical forest of central Africa (Cavalli-Sforza 1986, pp. 406–411). Both patrilocality and hypergamy, as well as abduction of women, which was frequent in antiquity and is still observed—for example, among the Yanomama—can increase the gene flow tied to womens' migration and hence of mtDNA, over that of autosomes or Y chromosomes. Most probably for the same reasons, Y chromosomes seem to show a greater geographic clustering than is seen in mtDNA trees, although comparisons are still limited and indirect (Ruiz-Linares et al. 1996; Underhill et al. 1996).

One of the problems with the historical genetics of Europe is that it has the lowest genetic variation, a third of that of the most variable continents, when measured on the basis of the ratio of genetic variation to long-distance geographic variation (Cavalli-Sforza et al. 1994, p. 122). To make matters worse, the average genetic difference between non-African countries, according to analysis of the mtDNA D-loop, is only approximately a third of that between African populations (Jorde et al. 1995; Mountain et al. 1995). The amount of noise generated in mtDNA by mutation makes this variation even less attractive as a basis on which to calculate divergence of populations; its main remaining attraction is that it is the only current source of measurement of female migration. Thus it is not surprising that the evolutionary analysis of 10 species on the basis of the sequences of the D-loop has not given satisfactory trees and that only the sequence of the complete mtDNA has proved reasonably adequate for establishment of the evolutionary tree (Cummings et al. 1995).

L. L. Cavalli-Sforza and E. Minch
Department of Genetics
Stanford University Medical Center
Stanford

References

Bendall KE, Macaulay VA, Baker JR, Sykes BC (1996) Heteroplasmic point mutations in the human mtDNA control region. Am J Hum Genet 59:1276–1287

Bertranpetit J, Sala J, Calafell F, Underhill PA, Moral P, Comas D (1995) Human mitochondrial DNA variation and the origin of Basques. Ann Hum Genet 59:63–81

Cavalli-Sforza LL (ed) (1986) African Pygmies. Academic Press, Orlando

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

Cummings MP, Otto SP, Wakeley J (1995) Sampling properties of DNA sequence data in phylogenetic analysis. Mol Biol Evol 12:814–822

Ember CR (1978) Myths about hunter-gatherers. Ethnology 17:439–448

Hewlett B (1996) Cultural diversity among African Pygmies. In: Kent S (ed) Cultural diversity among twentieth-century foragers. Cambridge University Press, Cambridge, pp 215–244

Horai S, Hayasaka K (1990) Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. Am J Hum Genet 46:828–842

Jorde LB, Bamshad MJ, Watkins WS, Zenger R, Fraley AE, Krakowiak PA, Carpenter KD, et al (1995) Origins and affinities of modern humans: a comparison of mitochondrial and nuclear genetic data. Am J Hum Genet 57:523–538

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

Mountain JL, Hebert JM, Bhattacharyya S, Underhill PA, Ottolenghi C, Gadgil M, Cavalli-Sforza LL (1995) Demographic history of India and mtDNA-sequence diversity. Am J Hum Genet 56:979–992

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

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

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9:552–569

Ruiz-Linares A, Nayar K, Goldstein D, Hebert JM, Seielstad MT, Underhill PA, Feldman MW, et al (1996) Geographical clustering of human Y-chromosome haplotypes. Ann Hum Genet 60:401–408

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

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129:555–562

Sokal RR, Rohlf FJ (1981) Biometry. WH Freeman, San Francisco

Underhill PA, Jin L, Zemans R, Oefner PJ, Cavalli-Sforza LL (1996) A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. Proc Natl Acad Sci USA 93:196–200

Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC (1989) Mitochondrial DNA sequences in single hairs from a southern African population. Proc Natl Acad Sci USA 86:9350–9354

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. Science 253:1503–1507

Address for correspondence and reprints: Dr. L. L. Cavalli-Sforza, Department of Genetics, MS-5120, Stanford University School of Medicine, Stanford, CA 94305-5120. E-mail: cavalli@lotka.stanford.edu

0002-9297/97/6101-0038$02.00

Am. J. Hum. Genet. 61:251–254, 1997

Reply to Cavalli-Sforza and Minch

To the Editor:

In a recent paper (Richards et al. 1996), we used a phylogeographic approach to infer that most (>85%) of the mtDNA control region (D-loop) variation in present-day Europeans has an ancient ancestry within Europe, coalescing during the Upper Paleolithic. This seems to be in contrast with earlier principal-component analyses of nuclear-gene frequencies in Europe, widely interpreted as evidence for a substantial Neolithic settlement from southwest Asia, which overwhelmed the Mesolithic hunter-gatherers. This apparent conflict has engendered the response by Cavalli-Sforza and Minch. They make criticisms of our treatment of the data in particular and of the reliability of mitochondrial control-region sequences in general, both of which criticisms we will address below. It is worth noting at the outset, however, their new suggestion that the proportion of the variation accounted for by the first principal component (26%) is "probably not very far" from the proportion of genes contributed by Neolithic newcomers to the European gene pool. Were this correct, it might seem that there could be little room for debate, since we could agree that the genetic contribution of the newcomers, while not insignificant, was relatively minor. However, there is more to the issue than this.

With regard first to their specific criticisms of our paper, it is precisely because there is little of interest to be learned from population-based comparisons using a single locus that we adopted a genealogical approach. There was—and apparently still is—a basic misunderstanding concerning the way in which mtDNA and Y-chromosome sequences should be analyzed for population studies. Traditionally, nuclear-allele frequency data have been the target of investigation, but, because recombination operates on such data in every generation, such analyses are inevitably restricted to coarse-grained summary statistics at the population level (diversity measures, population trees, principal-component maps, etc.). The resulting loss of information is then compensated in part by taking a large number of such loci into consideration. With mtDNA (or, for that matter, any other single locus), this approach is bound to be rather uninformative, and it is no surprise that earlier reports of European mtDNA diversity (Pult et al. 1994; Bertranpetit et al. 1995) were unable to detect significant structure. Table 2 in our earlier paper testifies to the futility of applying diversity measures between populations to mtDNA. We evidently did not emphasize this clearly enough in our paper, leading Cavalli-Sforza and Minch to miss our point and to reiterate this unhelpful test scenario by use of table 4 in our previous paper.