Mitochondrial lineages in Ladin-speaking communities of the eastern Alps

Michele Stenico¹, Loredana Nigro¹ and Guido Barbujani^{2*}

¹*Dipartimento di Biologia, Università di Padova, via G. Colombo 3, 35121 Padova, Italy*

²*Dipartimento di Biologia, Università di Ferrara, via L. Borsari 46, 44100 Ferrara, Italy (bjg@ifeuniv.unife.it)*

European mitochondrial alleles cluster into five haplogroups. Haplogroup 2 is rare in general, but represents more than half of the few known sequences among Ladin speakers of the Alps. Here we describe DNA diversity in control region I of the hypervariable D-loop in 43 Ladins, and in 25 Italian speakers. Analysis of these data, and of previously published sequences, confirms a high degree of differentiation among Ladins and their geographical neighbours. This cannot be regarded as a simple effect of isolating factors, geographic or linguistic, as diversity is high within Ladin communities too. Rather, allele genealogies, population trees, and principal component analysis suggest a relationship between Ladin and Near Eastern samples. Two evolutionary hypotheses seem compatible with these findings. The view whereby Ladins could be descended from Palaeolithic inhabitants of the Alps is supported by the identification, in this study, of the probable ancestral haplotype of group 2, never previously observed in central Europe. Alternatively, a comparatively recent, Neolithic immigration of the ancestors of current Ladin speakers seems consistent with recent linguistic theories. In both cases, the number of lineages present, and their extensive diversity, are not compatible with a serious bottleneck in the Ladin population's history.

Keywords: population genetics; mtDNA; linguistic diversity; allele genealogies; Ladins

1. INTRODUCTION

The increasing amount of molecular data available is providing us with deeper insight into past evolutionary processes. However, by itself the simple accumulation of genetic data has seldom proved sufficient to answer the main open questions. Analyses of different polymorphisms, for example, may lead to different evolutionary inferences (e.g. Jorde *et al.* 1995), both because of sampling effects and because molecular evolution and demographic changes seldom occur at the same time-scale.

In humans, however, independent information on population history is provided by archaeology and linguistics. In particular, because biological and cultural differences tend to increase in isolation—the lesser the contacts the faster the divergence—linguistic barriers are often associated with increased genetic variation (Barbujani & Sokal 1991). In this study we investigate the genetic effects of geographic and linguistic subdivision in an area of the Alps where high mitochondrial differentiation was observed both between and within populations (Stenico *et al.* 1996). In the median networks (Bandelt *et al.* 1995) inferred from European mitochondrial allele genealogies, five groups of evolutionarily related haplotypes have been recognized (Richards *et al.* 1996). The oldest of these, haplogroup 2, the lineages of which coalesce at an ancestor living approximately 36 500 years ago, has frequencies lower than 20% all over Europe (Torroni *et al.* 1996), but greater than 60% in the only known Ladin-speaking samples (Stenico *et al.* 1996). The high internal diversity

observed among Ladins suggests genetic drift is not a sufficient explanation for that. Rather, the ancestors of these Ladin populations (but not of their linguistic relatives, the Swiss speakers of Romansch) were probably affected by gene flow processes that had a lesser impact upon their neighbours (Stenico *et al.* 1996). It was not known to what extent other Ladin-speaking communities share the same characteristics.

For this study we collected new samples of Italian and Ladin speakers. Sequences of German-speaking individuals (the main language group in the eastern Alps) from Stenico *et al.* (1996) were also considered. We looked for areas of agreement between our results and the hypotheses emerging from archaeological and linguistic studies, with special emphasis on the evolutionary relationships between the Ladins and the people who expanded from the Middle East in Palaeolithic and Neolithic times (see Cavalli-Sforza *et al.* 1994).

2. LANGUAGES IN THE ALPS

Three Indo-European languages are currently spoken in the eastern Italian Alps: German, Italian and Ladin. The first belongs to the Germanic family, whereas Ladin and Italian are Romance languages (Ruhlen 1991). Even closer linguistic relatives of Ladin are Romansch and Friulian, both spoken by small groups dwelling, respectively, in the mountains of eastern Switzerland and of north-eastern Italy. Virtually all Ladin speakers are bi- or trilingual, with German generally being their second language. This is in contrast with the view that Ladin could simply be an Italian dialect. Although German and Italian are now more

*Author for correspondence.

Table 1. *Other European samples of mitochondrial sequence data considered in the study*

(Two-letter codes indicate populations of the Alps.)

population	number of sequences	locality codes	source
Santa Lucia	10	LS	Stenico <i>et al.</i> (1996)
Val Gardena	10	LG	Stenico <i>et al.</i> (1996)
northern Italy, German speakers (2 localities)	20	GH, GJ	Stenico <i>et al.</i> (1996)
northern Italy, Italian speakers (2 localities)	20	IN, IS	Stenico <i>et al.</i> (1996)
Mocheni	10	GM	Stenico <i>et al.</i> (1996)
Austria	16	GA	Pult <i>et al.</i> (1994)
Switzerland, Romansch speakers	15	LR	Pult <i>et al.</i> (1994)
Switzerland, German speakers	48	GS	Pult <i>et al.</i> (1994)
Basques	45	BAS	Bertranpetit <i>et al.</i> (1995)
Bulgaria	30	BUL	Calafell <i>et al.</i> (1995)
Britain	100	BRI	Piercy <i>et al.</i> (1993)
Estonia	28	EST	Sajantila <i>et al.</i> (1995)
Finland	50	FIN	Sajantila <i>et al.</i> (1995)
Iceland	37	ICE	Sajantila <i>et al.</i> (1995)
Middle East	42	MEA	Di Rienzo & Wilson (1991)
Sardinia	69	SAR	Di Rienzo & Wilson (1991)
Tuscany, Italy	49	TUS	Francalacci <i>et al.</i> (1995)
Turkey	74	TUR	Calafell <i>et al.</i> (1996)

widespread in the Alps, the distribution of toponyms attests to a much broader diffusion of Ladin, or of a related language, in the early middle ages (Pellegrini 1991). Many members of the Ladin communities consider themselves as the 'original' inhabitants of the area, which appears difficult to assess based upon historical evidence (Pauli 1981). However, the Italian speakers are known to be the most recent immigrants. Levels of admixture among communities have never been quantified, but the geography of the area, characterized by mountain barriers and narrow valleys, suggests subdivision and isolation have long been substantial.

3. MATERIALS AND METHODS

(a) *The samples*

Sixty-eight specimens, from apparently unrelated individuals, were sampled for this study. Forty-three subjects were from the three valleys of the Alps where most Ladin speakers are concentrated, namely Badia, Fassa and Gardena (13, 14, and 16 individuals, respectively). In addition, 25 Italian speakers from a neighbouring area in Veneto were also sampled (see figure 4).

(b) *Control-region sequences*

DNA was extracted from fresh whole blood by the classical protocol (Maniatis *et al.* 1982) and mtDNA was amplified using the primers L15926 and H408. A sequence of 360 bp within the first hypervariable segment, between positions 16 024 and 16 383 of the Cambridge reference sequence (hereafter referred to as CRS; Anderson *et al.* 1981) was obtained by thermo sequenaseTM cycle sequencing (Amersham kit), using the L15997 and/or H16401 primers (Vigilant *et al.* 1991). Variable positions in the sequence are here indicated by the numbering of the CRS less 16 000.

(c) *Genetic diversity within samples*

In populations at equilibrium, levels of molecular variation are proportional to population size (N) and mutation rate (u).

The parameter describing their joint effects is $\theta = 2Nu$; Tajima's statistic, D , is the ratio of two independent estimates of θ , obtained, respectively, from the pairwise differences between sequences, and from the number of polymorphic sites in the sample (Tajima 1989a). Departures from the mutation-drift equilibrium result in discrepancies between estimates, with negative D values often reflecting population expansions (Tajima 1989b). Tajima's statistic was calculated separately for each sample, and jointly for groups of linguistically homogeneous samples.

(d) *Other samples considered*

Other mitochondrial sequence data were used in this study (listed in table 1). Because they came from the same valley, in all comparisons among populations the Ladin sequences from Selva Val Gardena of Stenico *et al.* (1996) were pooled with the 16 sequences from Gardena valley obtained in this study.

(e) *Genetic distances*

Sequence differences between individuals were quantified by Kimura's two-parameter distance (Kimura 1980). Because mutation rates vary across sites, a Γ distribution of mutation probabilities was assumed (Wakeley 1993), whose a parameter was set at 0.47, in agreement with approximations derived by Wakeley (1993) for the first hypervariable region of the mitochondrial genome. The distances calculated accordingly by the MEGA package (Kumar *et al.* 1993) will be referred to as Γ -distances.

Genetic distances between populations were also Γ -distances, estimated between individuals of different samples, and corrected for within-sample variation. If the average Γ -distances within samples A and B are m_a and m_b , and the average distance between members of the two samples is I_{ab} , the distance between populations is

$$D_{ab} = I_{ab} - \frac{m_a + m_b}{2}.$$

(f) Evolutionary trees

The inferred genetic relationships among samples and individuals were represented in tree form using the neighbour-joining method (Saitou & Nei 1987). To root the trees, an African population, western Pygmies (Vigilant *et al.* 1991), and an individual from the same population (HMSTDL001) were chosen as outgroups.

(g) Principal component analysis

Mitochondrial variation was also summarized by principal component analysis. The plot of the first two principal components allows one to represent population relationships in a two-dimensional plane. As a preliminary step, we calculated for each European sample the frequencies of the haplogroups recognized by Richards *et al.* (1996), treating the CRS as a separated entity. In this way, the overall number of haplogroups considered was six. Their frequencies were normalized, and synthetic variables were calculated from the correlation matrices between normalized frequencies.

(h) Genetic boundaries

The sampling localities of the Alps were connected by a Delaunay network (Brassel & Reif 1979), a graphical method to define adjacent points on a map. Lines of maximum genetic differentiation were traced perpendicular to the edges of the network associated with the highest D_{ab} genetic distances, and continuous lines, or genetic boundaries (Barbujani *et al.* 1989), were then drawn according to Monmonier's (1973) algorithm.

4. RESULTS

Fifty-seven distinct region-I haplotypes were identified among the 68 individuals sequenced. Overall, 68 sites were polymorphic. The CRS, the most common mitochondrial sequence in Europe, was found only once among Ladins, in Fassa. One individual from Gardena differed from the CRS only by a transition in position 126. This sequence is considered the common molecular ancestor to all haplogroup 2 haplotypes, and it has been previously described in the Near East (Di Rienzo & Wilson 1991), but neither in eastern nor in central Europe. As for the frequency of haplogroup 2, Veneto had 12.0%, Fassa 35.7%, Badia 24.9%, whereas in Gardena we observed a combined frequency of 63.0% for the 26 individuals of this and of the previous study.

The average pairwise difference between sequences of the same sample varies from 4.7 (Badia) to 5.6 (Veneto), with Fassa and Gardena showing the same value, 5.1. Although lower than comparable estimates in Africa (e.g. Graven *et al.* 1995), this level of mitochondrial differentiation is higher than commonly observed in Europe, even within urban populations (see Stenico *et al.* 1996). Similar amounts of internal diversity are present in all other samples from the Alps except Mocheni, where the average is 1.4. Tajima's test statistics are negative for all populations of the Alps, although, owing to the limited sample sizes, statistical significance is reached only in linguistic aggregates (German speakers: $D = -1.81$; Italian speakers: $D = -2.06$; Ladin speakers: $D = -2.04$, all $p < 0.05$).

There is a limited clustering of geographically or linguistically close individuals in the neighbour-joining tree of the eastern Alps sequences (figure 1). A cluster comprising 40 individuals, 30 of them Ladin speakers,

corresponds to sequences that Richards *et al.* (1996) classified as haplogroup 2. Within it, most sequences of two Ladin localities, Gardena and Santa Lucia, fall in a subcluster corresponding to haplogroup 2B, and most sequences from Fassa fall in another subcluster, corresponding to haplogroup 2A. Sequences from Badia are scattered in the tree.

Genetic distances between populations range from -0.008 (between two Italian-speaking samples) to 0.228 (between German-speaking Mocheni and one Ladin-speaking sample). Occasional negative distances are expected when within-sample diversity is high, as is the case in human mitochondrial studies. After adding a constant (which does not affect the successive results) a population tree was constructed (figure 2). Its short branches show that populations do not differ much. The only exception seems to be a group including Gardena and Santa Lucia, showing higher divergence, and appearing genetically related to the Middle East population; a third Ladin sample, Badia, falls in the same cluster. The closest evolutionary relatives of these populations seem to be the German-speaking populations of the Italian Alps. In addition, essentially all populations of the Alps occupy the upper part of the dendrogram, whose lower part contains an apparently random aggregation of the other European samples.

In the plot of the first two principal components, jointly accounting for 67% of the overall variance (figure 3), three Ladin localities and the Near Eastern sample show negative values of the first component. All other samples form a loose cluster in the right part of the graph. The three Ladin samples are then separated from the Middle Eastern sample along the second axis, where the Alps populations, regardless of their language, appear genetically distinct from all other European populations. In the neighbour-joining tree, the outlier among Ladin samples was Fassa, whereas here it is Badia. This, and other limited discrepancies are due to the fact that molecular differences between alleles of the same haplogroup are irrelevant for this kind of principal component analysis, but affect the topology of the evolutionary tree.

Zones of sharp genetic differentiation separate two Ladin localities, Santa Lucia (figure 4, boundary 1) and Gardena (figure 4, boundaries 2 and 3), from the other samples. Part of genetic boundary 2 divides the north-western from the south-eastern region, and surrounds the geographic and linguistic isolate of Mocheni. Because this method is not associated with a test of significance, the number of boundaries one can trace is arbitrary; we chose to identify only three, given the limited number of samples available. The fourth boundary, however, would separate another Ladin locality, Badia, from all other samples. The genetic differentiation of Ladin samples, each of them from one another and all of them but Fassa from the non-Ladin samples, is confirmed by this geographically constrained clustering method.

5. DISCUSSION

In a previous paper (Stenico *et al.* 1996) some mitochondrial haplotypes that are rare or absent elsewhere were shown to reach high frequencies in two Ladin communities. The observed, high internal diversity ruled

Figure 2. Neighbour-joining tree of European and Near Eastern populations. F -distances between populations were calculated based on Kimura's two-parameter model (ratio transitions to transversions=49:1). Populations coded as in figure 1, bar Veneto (IV).

out the possibility that those communities could have simply lost different alleles under severe isolation. Indeed, population genetics theory demonstrates that small isolates are unlikely to maintain high genetic diversity. Therefore, with the exception of Mocheni, the populations of the eastern Italian Alps seemed to have evolved under a more complicated set of factors, probably including different contributions from different immigrating groups of females.

The present study shows that haplogroup 2 is common among Ladin speakers in general, and not only in the two localities studied by Stenico *et al.* (1996). Frequency differences exist, with the widest valleys (Fassa and Badia) showing less uncommon mitochondrial characteristics than Gardena and the isolated village of Santa Lucia. As shown by Tajima's tests, in no language group are populations at mutation-drift equilibrium.

It seems that the Ladin linguistic unit can also be regarded, by and large, as a reasonably compact evolutionary unit. Ladin speakers share with their neighbours the non-equilibrium distribution of mitochondrial sequence variation, and differ from them in the frequency of one set of evolutionarily related haplotypes. Two open questions are where these haplotypes come from, and when? The first question is simpler to answer. The overall similarity observed here between Ladin and Middle Eastern populations, and the high frequency of haplogroup 2 alleles in the Middle East (Sykes *et al.* 1996; Richards *et al.* 1996), suggest an oriental origin of these genes. Alternatively, an origin somewhere near the Alps followed by an eastward spread is conceivable in principle, but is not supported by any known demographic process affecting Europe (Cavalli-Sforza *et al.* 1994).

The timing of the arrival of haplogroup 2 alleles in the Alps is more difficult to define with a reasonable degree of

Figure 3. Principal component analysis of mitochondrial diversity in European and Near Eastern populations. Samples coded as in figure 1, bar Veneto (IV). The first two principal components represent, respectively, 37.8% and 29.2% of the overall variance.

confidence. There is little doubt that two main demographic expansions from the Near East, approximately along the same routes (Richards *et al.* 1997), contributed to the peopling of Europe; their relative importance, however, has yet to be established (Richards *et al.* 1996; Cavalli-Sforza & Minch 1997; Barbujani *et al.* 1998). The first process corresponded to the initial, Palaeolithic spread of *Homo sapiens sapiens*, and the second one was the Neolithic dispersal of early farmers. These were large-scale phenomena, and it may well be that some localities, especially those in the mountains, were affected only to a limited extent by the major, continent-wide demographic transitions. Therefore, one could reformulate the question above by asking which of the following processes is most likely: (i) haplogroup 2 was common among the first Europeans, and then it was diluted in most groups, but less so among Ladins; (ii) haplogroup 2 was typical of Neolithic farmers, whose genes gave a greater-than-average contribution to the genetic make-up of Ladins; or (iii) haplogroup 2 was carried to the Alps by recent immigrants whose mitochondrial genes had a strong Middle Eastern component.

Sykes *et al.* (1996) observed that two subgroups of haplogroup 2 seem to have appeared less than 12 000 years ago in Europe but earlier on in the Middle East, and suggested that they spread with the Neolithic farmers. Conversely, another subset characterized by two transitions at 126 and 294, called IIC in Sykes *et al.* (1996) and 2B in Richards *et al.* (1996), appears to be more than 30 000 years old both in Europe and in the Middle East, which was regarded as evidence for a Palaeolithic spread of the individuals carrying it. Because most alleles of the Ladins fall in subgroup 2B, one could conclude that ancestral characteristics of the first Europeans were maintained in the Ladin valleys; this is what we called process (i). The presence among Ladins of the probable ancestral haplotype of subgroup 2B (carrying a transition at 126) is another element in favour of this view, although not proof, because the age of haplotype and the age of the population where it is found are not necessarily the same. However, the Alps were not suitable for continuous human presence until 15 000 years ago, and the oldest human remains near the area we studied can be dated at 12 000–13 000 years ago (Broglio

Figure 4. Map of the eastern Alps, with the approximate indication of the origin of the samples. Samples coded as in figure 1, bar Veneto (IV). A Delaunay network was superimposed, across which the main lines of genetic differentiation are shown, and numbered according to their relative importance. Figures along the edges of the network are F -distances, $\times 1000$.

1993). Therefore, a common origin of current Ladin speakers in the Palaeolithic implies that their ancestral group maintained cohesion through long evolutionary times while moving across different localities.

The alternative view, whereby haplogroup 2 appeared comparatively recently in the Alps, stems from the proposal of a relationship between Ladin and some Slavic languages (Alinei 1996). This view is not prevalent among linguists (e.g. Ruhlen 1991). However, if we follow it for the sake of the argument, the historical scenario emerging is one in which mining technologies, introduced in the Alps around 4000 years ago, were carried by immigrants from eastern Europe, or perhaps simply attracted them. The model does not specify whether these immigrants were also the first farmers expanding in Europe (process (ii)), or if they reached the Alps through a more localized migration movement (process (iii)). The latter appears

unlikely, because the mtDNA sequences present at high frequencies among Ladins have also been found, albeit rarely, all over Europe.

It seems then that both process (i) and process (ii), namely a Palaeolithic and a not-so-recent Neolithic population genealogy, can be reconciled with the observed patterns of mitochondrial diversity, even though no simple mechanism is sufficient to account for all the features of mitochondrial DNA diversity in the Alps. Future analyses of ancient DNA may clarify the picture. For the moment, it is clear that, although numerically small (a few tens of thousand individuals), the Ladin-speaking isolates of the Italian Alps are both internally differentiated and genealogically related. This suggests they evolved from a common pool of ancestors without passing through severe bottlenecks. These ancestors (whatever their original geographical location) formed a single group long before the current linguistic identity was

established. As is customary for human populations (Barbujani *et al.* 1997), admixture has reduced, but not completely erased, the differences between them and their geographical neighbours. Gardena and Santa Lucia were presumably more isolated, whereas the frequencies of haplogroup 2 in Badia and Fassa (still high, but comparable to the frequencies observed among German speakers) suggest that these localities were exposed to higher rates of female-related gene flow.

Blood samples were provided by Mario Chiocchetti, director of the Institut Cultural Ladin 'Majon di Fashegn', and by Mario Picci. Laura Vanelli explained to us various aspects of the linguistic structure of the Alps. Giulietta Di Benedetto, Giorgio Bertorelle and Lounès Chikhi helped us in various phases of this work and critically read the manuscript. We thank them all. This study was supported by the Italian CNR grant no. 96.01182.PF36.

REFERENCES

- Alinei, M. 1996 *Le origini*. Bologna: Il Mulino.
- Anderson, S. (and 13 others) 1981 Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457–465.
- Bandelt, H.-J., Forster, P., Sykes, B. C. & Richards, M. B. 1995 Mitochondrial portraits of human populations using median networks. *Genetics* **141**, 743–753.
- Barbujani, G., Bertorelle, G. & Chikhi, L. 1998 Evidence of Palaeolithic and Neolithic gene flow in Europe. *Am. J. Hum. Genet.* (In the press.)
- Barbujani, G., Magagni, A., Minch, E. & Cavalli-Sforza L. L. 1997 An apportionment of human microsatellite diversity. *Proc. Natn. Acad. Sci. USA* **94**, 4516–4519.
- Barbujani, G., Oden, N. L. & Sokal, R. R. 1989 Detecting regions of abrupt change in maps of biological variables. *Syst. Zool.* **38**, 376–389.
- Barbujani, G. & Sokal, R. R. 1991 Zones of sharp genetic change in Europe are also linguistic boundaries. *Proc. Natn. Acad. Sci. USA* **87**, 1816–1819.
- Bertranpetit, J., Sala, J., Calafell, F., Underhill, P. A., Moral, P. & Comas, D. 1995 Human mitochondrial DNA and the origin of Basques. *Ann. Hum. Genet.* **59**, 63–81.
- Brassel, K. E. & Reif, D. 1979 A procedure to generate Thiessen polygons. *Geogr. Anal.* **11**, 289–303.
- Broglio, A. 1993 Mountain sites in the context of the North-east Italian Upper Palaeolithic and Mesolithic. *Preistoria Alpina* **28**, 293–310.
- Calafell, F., Underhill, P., Tolun, A., Anglicheva, D. & Kalaydjieva, L. 1996 From Asia to Europe: mitochondrial DNA sequence variability in Bulgarians and Turks. *Ann. Hum. Genet.* **60**, 35–49.
- Cavalli-Sforza, L. L., Menozzi, P. & Piazza, A. 1994. *The history and geography of human genes*. Princeton University Press.
- Cavalli-Sforza, L. L. & Minch, E. 1997 Paleolithic and neolithic lineages in the European mitochondrial gene pool. *Am. J. Hum. Genet.* **61**, 247–251.
- Di Rienzo, A. & Wilson, A. C. 1991 Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Natn. Acad. Sci. USA* **88**, 1597–1601.
- Francalacci, P., Bertranpetit, J., Calafell, F. & Underhill, P. 1995 Sequence diversity of the control region of mitochondrial DNA in Tuscany, and its implications for the peopling of Europe. *Am. J. Phys. Anthropol.* **100**, 443–460.
- 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. *Molec. Biol. Evol.* **12**, 334–345.
- Jorde, L., Bamshad, M. J., Watkins, W. S., Zenger, R., Fraley, A. E., Krakowiak, P. A., Carpenter, K. D., Soodyall, H., Jenkins, T. & Rogers, A. R. 1995 Origins and affinities of modern humans: A comparison of mitochondrial and nuclear genetic data. *Am. J. Hum. Genet.* **57**, 523–538.
- Kimura, M. 1980 A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Molec. Evol.* **16**, 111–120.
- Kumar, S., Tamura, K. & Nei, M. 1993 *MEGA: molecular evolutionary genetic analysis*, version 1.01. Pennsylvania State University Press.
- Maniatis, T., Fitch, E. F., Sambrook, J. 1982 *Molecular cloning: a laboratory manual*. New York: Cold Spring Harbor Laboratory Press.
- Monmonier, M. 1973 Maximum-difference barriers: an alternative numerical regionalization method. *Geogr. Anal.* **3**, 245–261.
- Pauli, L. 1981 *Die Alpen in Frühzeit und Mittelalter: Die archäologische Entdeckung einer Kulturlandschaft*. Munich: Beck.
- Pellegrini, G. B. 1991 *La genesi del Retoromanzo (o Ladino)*. Tübingen: Niemeyer.
- Piercy, R., Sullivan, K. M., Benson, N. & Gill, P. 1993 The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. *Int. J. Leg. Med.* **106**, 85–90.
- Pult, I., Sajantila, A., Simainen, J., Georgiev, O., Schaffner, W. & Pääbo, S. 1994 Mitochondrial DNA sequences from Switzerland reveal striking homogeneity of European populations. *Biol. Chem. Hoppe Seyler* **375**, 837–840.
- Richards, M., Corte-Real, H., Forster, P., Macaulay, V., Wilkinson-Herbots, H., Demaine, H., Papiha, S., Hedges, R., Bandelt, H.-J. & Sykes, B. 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., Hedges, R., Forster, P. & Bandelt, H.-J. 1997 Paleolithic and neolithic lineages in the European mitochondrial gene pool. Reply to Cavalli-Sforza and Minch. *Am. J. Hum. Genet.* **61**, 251–254.
- Ruhlen, M. 1991 *A guide to the world's languages*, vol. 1 (*Classification*), 2nd edn. London: Edward Arnold.
- Saitou, N. & Nei, M. 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molec. Biol. Evol.* **4**, 406–426.
- Sajantila, A. (and 12 others) 1995 Genes and languages in Europe: an analysis of mitochondrial lineages. *Genome Res.* **5**, 42–52.
- Stenico, M., Nigro, L., Bertorelle, G., Calafell, F., Capitanio, M., Corrain, C. & Barbujani, G. 1996 High mitochondrial sequence diversity in linguistic isolates of the Alps. *Am. J. Hum. Genet.* **59**, 1363–1375.
- Sykes, B., Corte-Real, H. & Richards, M. 1996 Palaeolithic and neolithic contributions to the European gene pool. In *Molecular biology and human diversity* (ed. A.J. Boyce & C. G. N. Mascie-Taylor), pp. 130–140. Cambridge University Press.
- Tajima, F. 1989a Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.
- Tajima, F. 1989b The effect of change in population size on DNA polymorphism. *Genetics* **123**, 597–601.
- Torroni, A., Huoponen, K., Francalacci, P., Petrozzi, M., Morelli, L., Scozzari, R., Obinu, D., Savontaus, M.-L. & Wallace, D. C. 1996 Classification of European mtDNAs from an analysis of three European populations. *Genetics* **144**, 1835–1850.
- Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. & Wilson, A. C. 1991 African populations and the evolution of human mitochondrial DNA. *Science* **253**, 1503–1507.
- Wakeley, J. 1993 Substitution rate variation among sites in hyper-variable region I of human mitochondrial DNA. *J. Molec. Evol.* **37**, 613–623.