

Brief Communication

Mitochondrial Diversity in Linguistic Isolates of the Alps: A Reappraisal

CRISTIANO VERNESI, SILVIA FUSELLI, LOREDANA CASTRÌ, GIORGIO BERTORELLE,
AND GUIDO BARBUJANI^{1,2}

Abstract In Stenico et al. (1996) we reported unusually high levels of mitochondrial diversity in the Alps. In particular, two communities of Ladin speakers appeared the most extreme European mitochondrial outliers at that time. Recently, it has been observed that some rare nucleotide substitutions occur repeatedly among those sequences, raising the possibility of systematic sequencing errors. No biological material was left from the previous study, and hence we had to sample new individuals from the same communities. Here, we present the HVSI sequence variation, along with haplogroup assignment based on restriction fragment length polymorphism (RFLP), in 20 Ladin speakers of Colle Santa Lucia. None of the new sequences displays substitutions at the sites viewed as problematic. However, Ladins still show high levels of mtDNA diversity, both within their community and with respect to other Europeans, and they can still be considered one of the main European mitochondrial outliers.

In a previous study (Stenico et al. 1996), we described unusually high levels of mitochondrial DNA (mtDNA) diversity in seven communities of the Alps. With the exception of one group, the Mocheni, whose sequences suggested a recent founder effect, variation of the hypervariable segment I, HVSI, was high both within and among populations, so much so that a language isolate, the Ladins, appeared at that time the most extreme European mitochondrial outlier. We interpreted those results as suggesting that, despite their current small number (30,000 to 35,000 individuals: http://www.ethnologue.com/show_language.asp?code=LLD), Ladin speakers are descended from an originally large and genetically variable population, which had been isolated from its neighbors of the plains. In particular, a high frequency of an elsewhere-rare haplogroup, *T* (Richards et al. 1996), was apparent among Ladins. The presence of a T motif [16126, 16292, 16294, according to the numbering in Anderson et al. (1981)] in an individual who lived in the same area 6000 years ago suggests some degree of genealogic continuity, at least from the earliest Neolithic period (Di Benedetto et al. 2000).

¹All authors are affiliated with the Dipartimento di Biologia, Università di Ferrara, Ferrara, Italy.

²Corresponding author, G. Barbujani. E-mail: bjg@unife.it.

Recently, Hans-Jurgen Bandelt (personal communication) informed us that nucleotide substitutions at six sites, 16085, 16106, 16221, 16313, 16315, and 16246, occur repeatedly in the sequences of Stenico et al. (1996), although they are rarely seen in other European samples, and never in the same branches of the mitochondrial network. He attributed those mutations ('phantom' mutations; Bandelt et al. 2001) to systematic sequencing errors.

The DNA extraction and typing for Stenico et al. (1996) was carried out in Padua, in the laboratory headed by Loredana Nigro, who passed away in 1998. Recovering the original electropherograms, or the blood samples for resequencing, proved impossible. We then decided to collect new samples from the same communities. Here, we report on the HVSI sequences from 20 Ladin speakers (from all of whom informed consent was obtained) of Colle Santa Lucia, whose maternal grandmothers had been local residents, and who were not aware of any consanguinity along their female lines. Given the increasing diffusion of the Italian language and the small census size (480 individuals in 1991; www.dossier.net/pop.htm), collecting larger numbers of unrelated Ladins in that community appears hardly possible.

Materials and Methods

DNA Analysis. DNA was extracted from 200 μ L of fresh blood with the Genomix kit (Amersham Pharmacia) following the manufacturer's instructions. To avoid the occurrence of mutations that appear only in some batches of polymerase chain reaction (PCR) and sequencing reactions, we amplified HVSI at least twice in each individual (for PCR conditions and primers see Stenico et al. 1998), each time from an independent DNA extract from whole fresh blood. We then performed several cycle-sequencing reactions (Big Dye Terminator kit, Applied Biosystems) of both the L and the H strand, from independent amplicons. In addition, all individuals were typed for specific restriction polymorphic sites, which define haplogroups *M*, *L3*, *HV*, *H*, *J*, *T*, *pre-JT*, *I*, and *UK*, according to Torroni et al. (1996) and Macaulay et al. (1999).

Data Analysis. Indices of standard and molecular diversity and genetic distances, d_{ab} (Nei 1987), were calculated with ARLEQUIN (ver. 2.00; Schneider et al. 2000). Negative d_{ab} values, which occur when within-population diversity exceeds the diversity between two populations, were corrected by adding a constant quantity to each value in the matrix. A neighbor-joining tree (Saitou and Nei 1987) was then constructed by means of PHYLIP (ver. 3.57c; Felsenstein 1995).

The average genetic distance, d_{L-E} , between Ladins and the European populations in Stenico et al. (1996) was compared to the average genetic distance between random pairs of the same European populations, d_{E-E} . Individuals were randomly extracted with replacement from each population for 1000 times, and d_{L-E} and d_{E-E} were computed and compared each time.

Results

Diversity Indices. The observed polymorphic sites, along with the haplogroup assignments based on restriction fragment length polymorphism, are reported in Table 1. None of the new specimens displayed nucleotide substitutions at the six sites listed by Bandelt. As was already the case for the Ladins of Stenico et al. (1996), the present sample shows a high mean pairwise difference (MPD) between sequences, 6.48, among the highest values of MPD in a data set of nearly 3000 sequences in 34 populations from Europe, the Near East, and the Caucasus (updated from Simoni et al. 2000). The gene diversity is also high for a European population, 0.92. Both findings are not expected in small communities, where many individuals tend to be related along the female lines of descent.

Population Comparisons. The topology of the tree (Figure 1B) is very similar to that in Figure 5 of Stenico et al. (1996), which we also report here (Figure 1A). There are fewer populations in Figure 1B because all samples of Stenico et al. have been removed, and the Swiss groups speaking the same language have been put together. Since intrapopulation diversity is high, the external position of the Ladins can only reflect an extensive differentiation between the Ladins and the other populations of the Alps. In populations with low long-term effective size, one would expect a reduction in both MPD and gene diversity. Therefore, there is no reason to suspect that a founder effect or a bottleneck may account for the evident Ladin divergence. In 979 out of 1000 distance matrices generated by

Table 1. Mitochondrial HVSI Sequences in 20 Ladin Speakers from Colle Santa Lucia

Haplotype Identification	<i>n</i>	HVSI Haplotype	RFLPs Haplotype	Haplogroup
<i>SLU1</i>	1	189 263 270	+12308 <i>Hinf</i> I	<i>U5b*</i>
<i>SLU2</i>	3	182C 183C 189 249 274	+12308 <i>Hinf</i> I	<i>U1a</i>
<i>SLU3</i>	4	129 148 192 223 294	+10032 <i>Alu</i> I	<i>I</i>
<i>SLU4</i>	2	224 311	+12308 <i>Hinf</i> I	<i>K</i>
<i>SLU8</i>	2	261	-14766 <i>Mse</i> I -7025 <i>Alu</i> I	<i>H</i>
<i>SLU9</i>	2	093 224 311	+12308 <i>Hinf</i> I	<i>K</i>
<i>SLU10</i>	3	126 163 186 189 294	-11251 <i>Tsp</i> 509I +13366 <i>Bam</i> HI +15606 <i>Alu</i> I	<i>T1</i>
<i>SLU13</i>	1	051 129C 182C 183C 189 270 362	+ 12308 <i>Hinf</i> I	<i>U2</i>
<i>SLU14</i>	1	093 129 258 316	-14766 <i>Mse</i> I -7025 <i>Alu</i> I	<i>H</i>
<i>SLU16</i>	1	069 126 222	-11251 <i>Tsp</i> 509I -13704 <i>Bst</i> OI	<i>J</i>

Note: *n* = number of individuals with that sequence. The HVS haplotype is defined by the substitutions observed with respect to the Cambridge reference sequence; the position refers to the original numbering, minus 16,000. A letter, when present, indicates the observed nucleotide.

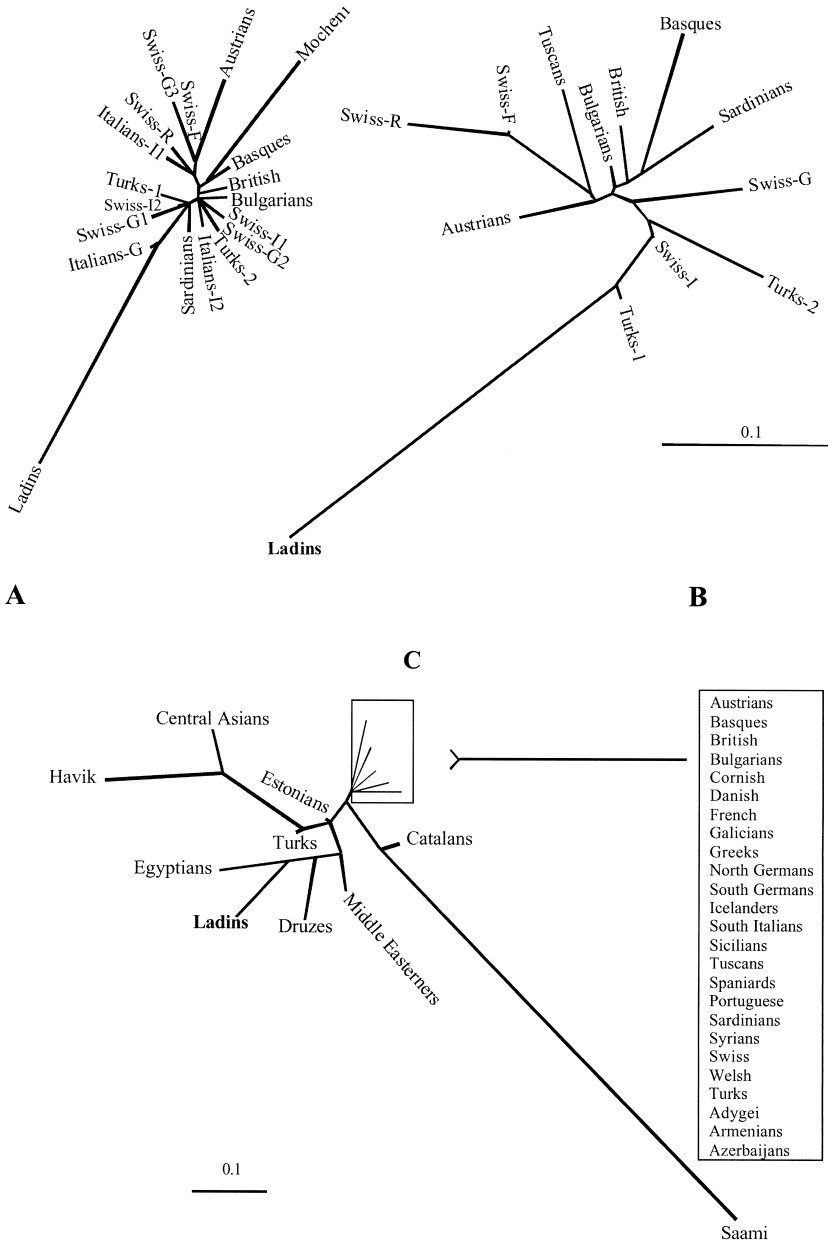


Figure 1. Mitochondrial neighbor-joining trees of European (**A**, **B**) and Eurasian (**C**) populations. **A:** Ladins, Mocheni, and other Eastern Alps samples from Stenico et al. (1996); the letters accompanying the Italian and Swiss samples refer to the language spoken: G = German, I = Italian, F = French, R = Romansch. **B:** Same as above; all samples from Stenico et al. (1996) have been removed, and Ladins come from the present study. The Basque and Austrian samples now incorporate recent data from Parson et al. (1998) and Corte-Real et al. (1996).

bootstrapping individuals within each population, the average distance, d_{L-E} , between the Ladins and the remaining populations is greater than the average distance, d_{E-E} , among the 12 populations of the NJ tree in Figure 1B. Therefore, there is reason to believe that the position of the Ladins in the evolutionary tree (Figure 1B) does not just reflect sampling effects.

In the neighbor-joining tree (Figure 1C), most European populations form a tight cluster, here in the upper right-hand corner of the tree. Saami are the most evident outlier, whereas the Ladins fall in another well-differentiated cluster, including only populations of the Eastern and Southern Mediterranean shores. Ladins of Colle Santa Lucia share with Egyptians, Druzes, and Near Easterners several haplotypes of groups K, I, and JT, which have only been observed at low frequencies over the rest of Europe (Richards et al. 1996; 2000).

Discussion

We could not replicate all results of the previous study in another sample from the same population, and it is indeed possible that sequencing errors occurred in Stenico et al. (1996). Although those errors may have affected 24 sequences at worst, at this stage we feel it is safer to exclude all 70 sequences published in Stenico et al. (1996) from further analyses. At the same time, we can confirm the main message of that study, namely that the Ladin-speaking inhabitants of the Alps do display a high internal mitochondrial diversity, and that they are one of the main European mitochondrial outliers. Further typing of other Ladin communities is in progress.

The unusually high mitochondrial diversity observed both in Stenico et al. (1996) and in this study suggests that current Ladin speakers are descended in reproductive isolation from a large and genetically variable ancestral population. The current language barriers cannot have been established longer than 15 centuries ago, and so it is necessary to conclude that other isolating mechanisms, presumably related to geographic barriers, contributed to maintaining unusual mitochondrial features in the Eastern Italian Alps.

Acknowledgments We thank Hans-Jürgen Bandelt for pointing out to us the problems with the Ladin mitochondrial sequences. Isabelle Dupanloup helped us with the development of the software for data analysis.

Received 13 December 2001; revision received 14 July 2002.

Literature Cited

Anderson, S., T. Bankier, B.G. Barrel et al. 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465.

- Bandelt, H.J., P. Lahermo, M.B. Richards et al. 2001. Detecting errors in mtDNA data by phylogenetic analysis. *Int. J. Legal. Med.* 115:64–69.
- Corte-Real, H.B., V.A. Macaulay, M.B. Richards et al. 1996. Genetic diversity in the Iberian Peninsula determined from mitochondrial sequences analysis. *Ann. Hum. Genet.* 60:331–350.
- Di Benedetto, G., I. Nasidze, M. Stenico et al. 2000. Mitochondrial DNA sequences in prehistoric human remains from the Alps. *Eur. J. Hum. Genet.* 8:669–677.
- Felsenstein, J. 1995. PHYLIP ver. 3.57c. Seattle, WA: Department of Genetics, University of Washington.
- Macaulay, V., M. Richards, E. Hickey et al. 1999. The emerging tree of west Eurasian mtDNAs: A synthesis of control-region sequences and RFLPs. *Am. J. Hum. Genet.* 64:232–249.
- Parson, W., T.J. Parsons, R. Scheithauer et al. 1998. Population data for 101 Austrian Caucasian mitochondrial DNA d-loop sequences: Application of mtDNA sequence analysis to a forensic case. *Int. J. Legal. Med.* 111:124–132.
- Richards, M., H. Corte-Real, P. Forster et al. 1996. Paleolithic and Neolithic lineages in the European mitochondrial gene pool. *Am. J. Hum. Genet.* 59:185–203.
- Richards, M.B., V.A. Macaulay, H.J. Bandelt et al. 1998. Phylogeography of mitochondrial DNA in Western Europe. *Ann. Hum. Genet.* 62:241–260.
- Richards, M., V. Macaulay, E. Hockey et al. 2000. Tracing founder lineages in the Near Eastern mtDNA pool. *Am. J. Hum. Genet.* 67:1251–1276.
- Saitou, N., and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406–425.
- Schneider, S., D. Roessli, and L. Excoffier. 2000. ARLEQUIN: A software for population genetics data analysis. Version 2.0. Department of Anthropology, University of Geneva, Switzerland.
- Simoni, L., F. Calafell, D. Pettener et al. 2000. Geographic Patterns of mtDNA Diversity in Europe. *Am. J. Hum. Genet.* 66:262–278.
- Stenico, M., Nigro, L., G. Bertorelle et al. 1996. High mitochondrial sequence diversity in linguistic isolates of the Alps. *Am. J. Hum. Genet.* 59:1363–1375.
- Stenico, M., L. Nigro, and G. Barbujani. 1998. Mitochondrial lineages in Ladin-speaking communities of the eastern Alps. *Proc. R. Soc. Lond. B Biol. Sci.* 265(1396):555–561.
- Torrioni, A., K. Huoponen, P. Francalacci et al. 1996. Classification of European mtDNAs from an analysis of three European populations. *Genetics* 144:1835–1850.