Diachronic and Synchronic Genetic Analysis of Ancient Piedmont Population

S. Vai[1], M. Lari[1], S. Ghirotto[2], E. Rizzi[3], A. Achilli[4], A. Olivieri[5], E. Bedini[6], B. Lippi[7], E. Rabino Massa[8], A. Piazza[9], A. Torroni[5], G. Barbujani[2], D. Caramelli[1]

[1] Dip. di Biologia Evoluzionistica, Università di Firenze.Via del Proconsolo 12, 50122 Firenze. E-mail: stefania.vai@unifi.it
[2] Dip. di Biologia ed Evoluzione, Università di Ferrara.Via Borsari 46, 44100 Ferrara
[3] Institute for Biomedical Technologies, (ITB – CNR). Via F.lli Cervi 93, Segrate (Milano)
[4] Dip. di Biologia Cellulare e Ambientale. Università degli Studi di Perugia. Via Elce di Sotto, 06123 Perugia
[5] Dip. di Genetica e Microbiologia, Università di Pavia. Via Ferrata 1, 27100 Pavia
[6] Anthropozoologica, Via Grande 82, 57123 Livorno
[7] Dip. di Biologia. Università di Pisa. Via Volta 4, 56126 Pisa
[8] Dip. Scienze della Vita e Biologia dei Sistemi. Università di Torino. Via Accademia Albertina 13, 10123 Torino
[9] Dip. di Genetica, Biologia e Biochimica. Università di Torino. Via Santena 19, 10126 Torino

KEY WORDS: ancient DNA, mtDNA, human, Piedmont.

Introduction

Italy is a heterogenic country both by a genetically and linguistic point of view.
This heterogeneity has its origin in the ancient structure of the peninsula (Piazza *et al.*, 1988), in isolation by geographical and linguistic boundaries (Barbujani and Sokal, 1991a, b) and in migrations with different effects in different areas (Achilli *et al.*, 2007). Previous studies on ancient human DNA showed that, in different Italian regions cases of continuity or discontinuity can be recognized (Belle *et al.*, 2006; Ghirotto *et al.*, 2010). Consequently, a singular approach on each region or sub-region is necessary to understand the Italian population structure in the past and its impact on the present. In this work, we present the first ancient genetic data for the Piedmont region. Our analyses were carried out on 173 samples from different provinces covering a chronological range of fifteen centuries. Particularly, an intensive sampling was dedicated to Longobards, within the Early Middle Ages (VI-VIII century), which encompasses the so-called "period of migrations" of Germanic peoples.
Our aim is to verify the genetic relationship between populations that occupied the same territory in different times, to understand if cultural differences between human groups correspond also to biological differences, and if the archeologically hypothesized migrations in the Early Middle Age had an effective North-European ancestry, or if there could only be a case of acculturation of local people. Our sampling is the largest performed so far in such kind of studies, both in terms of number of specimens and temporal range.

Materials and Methods

Analyses were carried out on 173 samples from the provinces of Turin, Cuneo and Alessandria. Samples were divided into six historical and cultural groups: i) 6 Protohistorical samples (including Bronze and Iron Ages samples), ii) 20 Roman samples (III-V century), iii) 6 Ostrogoths samples, iv) 104 Longobards samples, v) 20 Merovingians samples (the three Early Middle Ages cultural groups which appeared in Piedmont on VI-VIII centuries), and finally, vi) 17 Late Middle Ages and beginning of Renaissance samples (X-XIV century).
Some of the specimen were not washed or manipulated in any way, however complete information on the handling history of each specimen was available, and the DNA sequences of all the archaeologists and the geneticists who have been in contact with the samples were typed, hence any possible sources of modern contamination can be traced down.
DNA extraction, PCR amplification and analysis of the PCR products were performed in separate laboratory rooms of the Florence Molecular Anthropology laboratory. After brushing and irradiating each bone surface (1 hr under ultraviolet light), all fragments were manually powdered in a mortar. DNA was extracted from the powdered bone by means of a silica-based protocol (modified from Caramelli *et al.*, 2008). Two µl of DNA extracted from the bone were amplified by PCR (Guimaraes *et al.*, 2009). The 360 bp long HVR-I was subdivided into three overlapping fragments using the following primer pairs: L15995/H16132; L16107/H16261; L16247/H16402. Specimens were handled using disposable mask, gloves and sterile laboratory coats in areas where no contemporary DNA had ever been studied. The DNA extraction and the setting up of PCR reactions of ancient DNA templates were carried out in two different rooms under two different hoods, daily irradiated with UV rays (254 nm). All sterile tubes, filtered tips, sterile reagents and solutions were disposable or exclusively dedicated to ancient DNA. Different sets of pipettes were used for DNA extraction, PCR amplification and analysis of the PCR products. Negative controls were included in each set of extractions and amplifications.

PCR products were cloned using TOPO TA Cloning Kit (Invitrogen) according to the manufacturer's instructions. Screening of white recombinant colonies was accomplished by PCR, transferring the colonies into a 40 µl reaction mix (67mM Tris HCl (pH 8.8), 2 mM MgCl2, 1 µM of each primer, 0.125 mM of each dNTP, 0.75 units of Taq Polymerase) containing M13 forward and reverse universal primers. After 5 min at 92°C, 30 cycles of PCR (30 s at 90°C, 1 min at 50°C, 1 min at 72°C) were carried out and clones with inserts of the expected size were identified by agarose gel electrophoresis. After purification of these PCR products with Microcon PCR devices (Amicon), a volume of 3 ml was cycle-sequenced following the BigDye Terminator kit (Applied Biosystems) supplier's instructions. The sequence was determined using an Applied BioSystems 3100 DNA sequencer. Different clones were sequenced for each individual.

The sequences obtained were aligned and compared across clones. The consensus sequences were finally compared with a database of European mitochondrial sequences and analysed by a Median Joining Network (Bandelt *et al.*, 1999) to test whether the obtained ancient sequences had phylogenetic sense. A Median Joining Network was also constructed to show how the ancient piedmonts sequences were phylogenetically linked to each other.

The character weights, inversely proportional to the frequency of mutation, were set according the observations of Macaulay *et al.* (1997), Richards *et al.* (1998), Bandelt *et al.* (2000). For each group, variability at the intra-population level was investigated using Arlequin 3.5.1.2 (Excoffier, 2005) by calculation of number of different haplotypes, haplotype diversity (or heterozygosity) and mean number of pairwise differences (MPWD). Inter-population analysis was carried out by comparing the six ancient piedmonts groups with a dataset of 69 populations from Europe, Nord Africa, Near East and Central Asia. Matrices of pairwise Fst distances, haplotype sharing (HS), Nei's D and D_A indexes were estimated, using Kimura-2p model (Jin and Nei, 1990; Kimura, 1980). A Multidimensional Scaling were obtained from the Fst matrix, using R STATS package (R-DevelopmentCoreTeam, 2006).

Results

360-bp for mtDNA control region were sequenced for 49 samples (Vai, 2011). Table 1 shows the results for each ancient group and standard and genetic diversity indexes. Figure 1 shows how the sequences are linked in the Median Joining Network. The most represented haplotype is the CRS, shared by almost all the historical-cultural groups. Inter-population HS tables are too large to be showed, however significant results could be highlight only for the Longobards group, because its high number of individual gives consistency to data. Longobards show the highest HS values with Austria, Cornwales, Estonia, France, Russia and, in Italy, with Bologna and Vercelli. There are no contemporary population, neither groups which divided the same territory with Longobards, showing so high HS values with them. MDS graphics, obtained from Fst indexes, show the Merovingian group always located in a peripheral position respect to all the other populations. Figure 2

	Protohistorical	Roman	Ostrogoths	Longobards	Merovingians	Late Medieval
No. of individuals	3	5	2	28	6	5
No. of haplotypes	3	4	2	19	3	5
No. of usable loci	360	360	360	360	360	360
No. of polymorphic sites	6	6	2	23	5	16
Number of observed transitions	5	5	2	22	5	16
Number of observed transversions	1	1	0	2	0	0
Mean number of pairwise differences	4,23+/-2,86	2,49+/-1,61	2,06+/-1,77	3,85+/-1,99	2,31+/-1,46	7,07+/-4,00
Gene diversity	1,0+/- 0,27	0,90+/- 0,16	1,00+/- 0,50	0,96+/- 0,02	0,73+/- 0,16	1,00+/- 0,13

Tab. 1. *Standard and genetic diversity indexes for the ancient Piedmontese groups.*

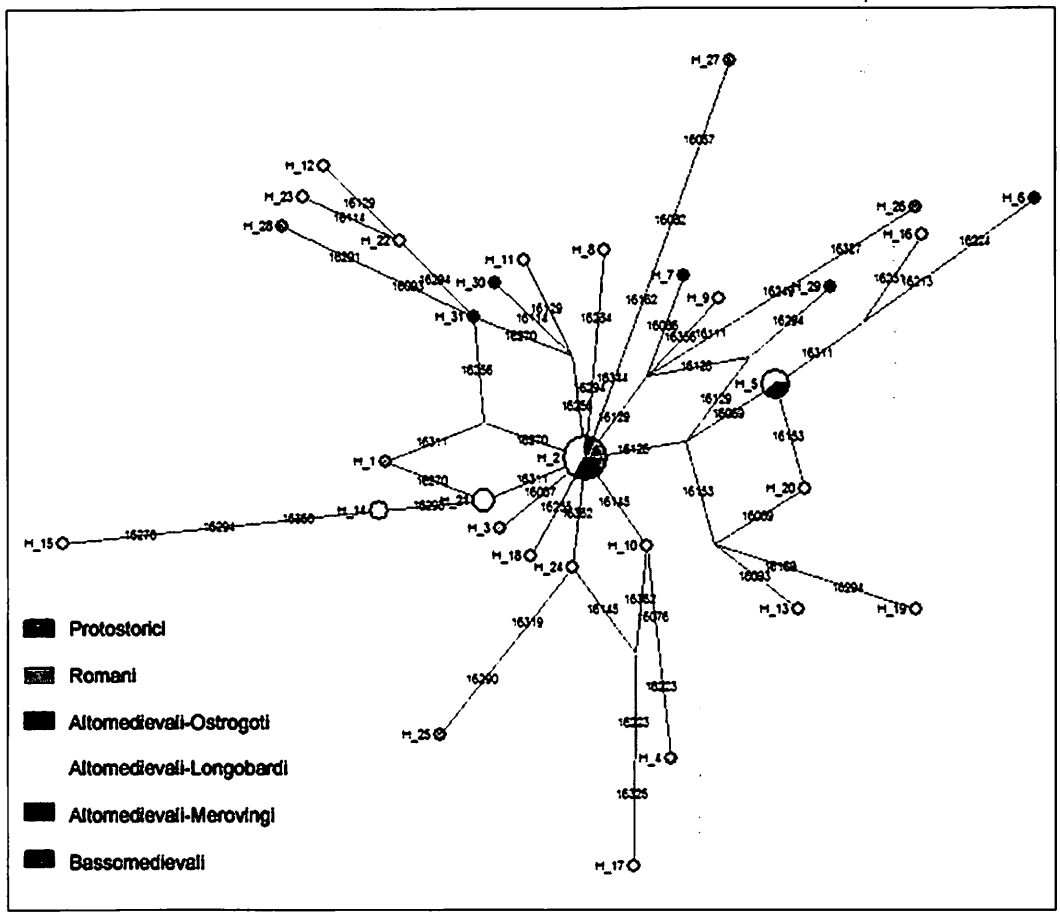

Fig. 1. *Median Joining Network. Colored nodes represent haplotypes found in the ancient samples. Nodes dimention is proportional to the haplotype frequency. Different colors highlight in which temporal and cultural facies the haplotype is present (see legend). Red numbers indicate the nucleotide polymorphic positions.*

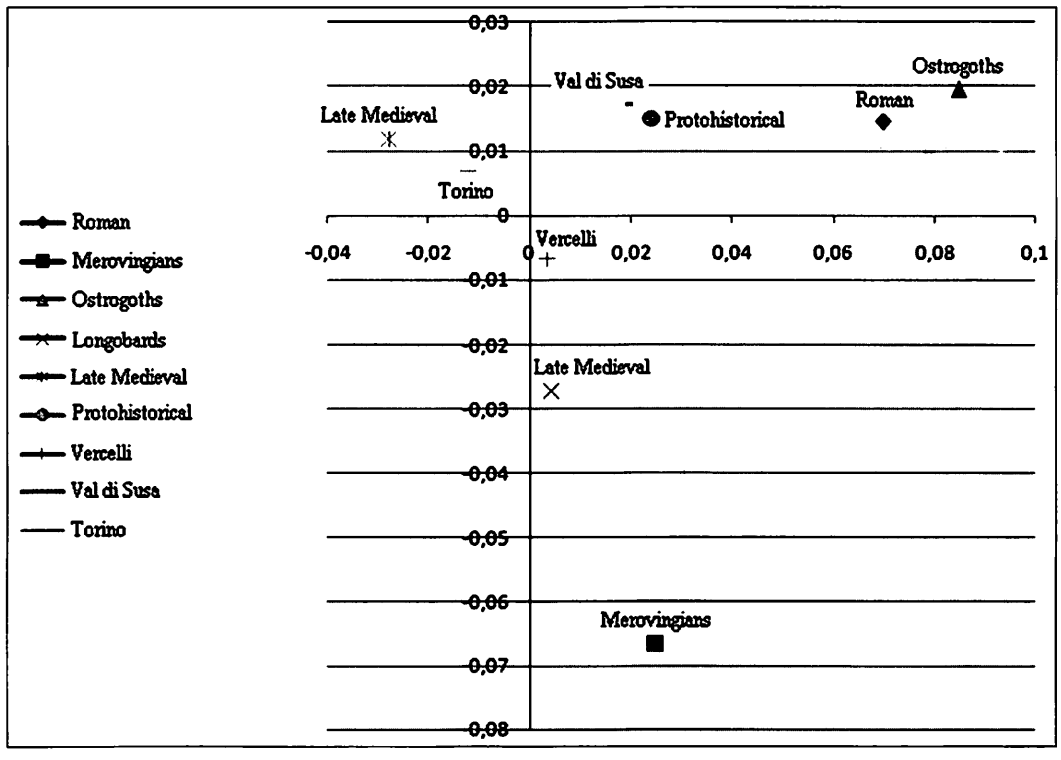

Fig. 2. *MDS representing Fst genetic distances between the ancient groups and the modern Piedmontese samples (stress value: 0,17).*

represents the MDS with the distances between the six groups of ancient and the modern Piedmontese population. The protohistorical group, in particular, is positioned near the modern Val di Susa inhabitants. The Mean Pairwise Differences (MPWD), expressed by Nei's D_A index, are calibrated to the internal diversity of the compared populations. Longobards, the largest group, show similarity (D_A<0.03) with northern and north-eastern European populations and, in Italy, with Bologna and Molise region. According to the Nei's D_A, among all the modern populations, Val di Susa is on average the nearest to the past Piedmont. Merovingians are the only ancient population that represents an exception to this trend.

Discussion

Our analyses represent the first contribution to understand the population history of Piedmont. The results show that Protohistorical group has a gene diversity equal to 1, and a MPWD more similar to modern populations (4,53 on average) than to ancient populations (2,54 on average). This could indicate that the protohistorical Piedmont was inhabited by a genetically variable population. But the small number of individuals does not permit to make strong this observation. It's interesting to find how the modern Val di Susa population, supposedly isolated, shows a particolar affinity to the protohistorical group. On the contrary, Roman, Ostrogots and Merovingians show an internal diversity similar to the other ancient population so far studied. Merovingians has the lowest gene diversity value (0,73) compared to all the ancient and modern populations examined (respectively, 0,90 and 0,96 on average); therefore, they appear as an homogeneous population with generally common characteristics (high HS values with several populations compared). Despite this high HS, Merovingian are distant from all the populations according to Fst indexes. Also Nei's D_A index shows that Merovingians are the only ancient Piedmont group to be far from Val di Susa population.

This is probably due to the presence of a private haplotyte in Merovingian group that influence genetic distances but not HS calculation. Late-Medieval group appears, instead, heterogeneous both in the M-J Network and as gene diversity and MPWD (7,07) indicate.

It seems to represent a composite population, but we need a greater number of samples to confirm this first observation. Broader considerations are possible for the Longobard group, numerically consistent. Their internal variability is close to the average values found for compared populations. HS and Nei's distances indicate their proximity to North, Center and East Europe, and to Bologna, Vercelli and Molise in Italy, the latter places known to be affected by the Longobard settlements. Therefore, from these data, it would seem that the new cultural elements appeared in Piedmont in the Early Middle Age could be related to migrations from North, Centre and East Europe, rather than to acculturation phenomena.

This seems to be the case of Longobards, at least. Different Early Middle Age cultures could be related to different populations, but we need more data to assert this. There could be some discontinuity in time, for example between the Early Middle Age and the Late Middle Age, when groups with different genetic variability characteristics have given way to a probably composite population. We find also some cases of genealogical continuity in Piedmont: between the protohistorical samples and the modern Val di Susa inhabitants, and between Longobards and today's Vercelli province.

From this first description of genetic variability of ancient Piedmont, as future project, we will hypothesize demographic dynamics to test by coalescent simulations and Bayesian Statistics.,

Acknowledgments

This study was supported by Fondazione Compagnia di San Paolo.

References

Achilli A., Olivieri A., Pala M., Metspalu E., Fornarino S., Battaglia V., Accetturo M., Kutuev I., Khusnutdinova E., Pennarun E., Cerutti N., Di Gaetano C., Crobu F., Palli D., Matullo G., Santachiara-Benerecetti A.S., Cavalli-Sforza L.L., Semino O., Villems R., Bandelt H-J, Piazza A., Torroni A. 2007. Mitochondrial DNA variation of modern Tuscans supports the near eastern origin of Etruscans. *Am. J. Hum. Genet.*, 80(4): 759-768.

Bandelt H.J., Forster P., Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.*, 16(1): 37- 48.

Bandelt H. J., Macaulay V., Richards M. 2000. Median Networks: speedy construction and greedy reduction, one simulation, and two case studies from human mtDNA. *Mol. Phylogenet. Evol.*, 16: 8-28.

Barbujani G., Sokal R.R. 1991a. Genetic Population-Structure of Italy.1. Geographic Patterns of Gene-Frequencies. *Hum. Biol.*, 63(3): 253-272.

Barbujani G., Sokal, R.R. 1991b. Genetic Population-Structure of Italy.2. Physical and Cultural Barriers to Gene Flow. *Am. J. Hum. Genet.*, 48(2): 398- 411.

Belle E.M.S., Ramakrishnan U., Mountain J.L., Barbujani G. 2006. Serial coalescent simulations suggest a weak genealogical relationship between Etruscans and modern Tuscans. *P. Natl. Acad. Sci. USA*, 103(21): 8012-8017.

Caramelli D., Milani L., Vai S., Modi A., Pecchioli E., Girardi M., Pilli E., Lari M., Lippi B., Ronchitelli A., Mallegni F., Casoli A., Bertorelle G., Barbujani G. 2008. A 28,000 years old Cro-Magnon mtDNA sequence differs from all potentially contaminating modern sequences. *PLoS ONE*, 3:(7).

Excoffier L.G., Schneider S. 2005. Arlequin ver. 3.0: An integrated sofware package for population genetics data analysis. *Evol. Bioinform.*, 1: 47-50.

Ghirotto S., Mona S., Benazzo A., Paparazzo F., Caramelli D., Barbujani G. 2010. Inferring Genealogical Processes from Patterns of Bronze-Age and Modern DNA Variation in Sardinia. *Mol. Biol. Evol.*, 27(4): 875 - 886.

Guimaraes S., Ghirotto S., Benazzo A., Milani L., Lari M., Pilli E., Pecchioli E., Mallegni F., Lippi B., Bertoldi F., Gelichi S., Casoli E., Belle M.S., Caramelli D., Barbujani G., 2009. Genealogical discontinuities among Etruscan, Medieval and contemporary Tuscans. *Mol. Biol. Evol.*, 26(9): 2157-2166.

Jin L., Nei M. 1990. Limitations of the evolutionary parsimony method of phylogenetic analysis. *Mol. Biol. Evol.*, 7(1): 82-102.

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.*, 16(2): 111-120.

Macaulay V.A., Richards M. B., Forster P., Bendall K. E., Watson E., Sykes B., Bandelt H. J. 1997. mtDNA mutation rates--no need to panic. *Am. J. Hum. Genet.*, 61: 983-990.

Piazza A., Cappello N., Olivetti E., Rendine S. 1988. A genetic history of Italy. *Ann. Hum. Genet.*, 52(Pt 3): 203-213.

R-DevelopmentCoreTeam. 2006. R: A language and Environment for Statistical Computing.

Richards M.B., Macaulay V.A., Bandelt H.J., Sykes B.C. 1998. Phylogeography of mitochondrial DNA in western Europe. *Ann. Hum. Genet.*, 62: 241-260.

Vai S. 2011. *Analisi genetica diacronica e sincronica di antiche popolazioni piemontesi.* Tesi di Dottorato, Università degli Studi di Firenze, Firenze.