JASs Invited Reviews
Journal of Anthropological Sciences
Vol. 86 (2008), pp. 59-89

History and geography of human Y-chromosome in Europe: a SNP perspective

Paolo Francalacci & Daria Sanna

Dipartimento di Zoologia e Genetica Evoluzionistica, Università di Sassari, via Muroni 25 – 07100 Sassari, Italy
e-mail: pfrancalacci@uniss.it

Summary - *The genetic variation observed in the modern European populations can be used to reconstruct the history of the human peopling of the continent. In recent times, a great importance has been given to uniparental markers such as the Y-chromosome. This chromosome, which is passed from father to son, does not have a counterpart subject to recombination and the only possible source of variation is mutation. The nucleotide changes accumulate over time in the molecule, with no rearrangement among lineages. Lately, the D-HPLC technique, which allows the effective detection of single nucleotide polymorphisms (SNPs), was used to boost the number of available polymorphisms on the Y-chromosome. Since the year 2000, a number of studies were aimed both at the reconstruction of Y-chromosome phylogeny and the geographic distribution of Y-chromosome variation in Europe. The distribution of distinctive Y-chromosome lineages can also display a correspondence with geography, thus providing patterns of affinity and clues concerning past human movements. It is therefore possible to recognize the effect of the colonization of Europe following the Last Glacial Maximum, both from the western Iberian and the eastern Balkan refuges. Other lineages show a migratory wave from the Near East, consistent with the demic diffusion model of agriculture. A minor east-west genetic cline was proposed as a signal of an expansion from north of the Black Sea, related with the diffusion of people speaking languages of the Indo-European family.*

Keywords - *Phylogeography, Single Nucleotide Polymorphisms, Non-Recombining portion, Y-chromosome, Population genetics.*

Introduction

Since the publication of "The history and geography of human genes" by Cavalli-Sforza and collaborators (1994), a pivotal study for human evolutionary genetics focused on divergence among human populations using many autosomal inherited traits (e.g. blood types, HLA factors, proteins), an increasing attention has been given to DNA markers at uniparental transmission, such as mitochondrial DNA (mtDNA) (Richards & Macaulay, 2001), and the non recombining portion of Y-chromosome (NRY) (Jobling & Tyler-Smith, 2003). The reason of this interest resides in their uniqueness in the human genome, being the only two haploid segments of the human genome, not subject to the rearrangement of recombination, and consequently inherited linearly identical through generations. Mutation is the only possible source of variation for these two genetic systems and nucleotide changes accumulate over time in the molecule. All the mtDNAs and NRYs of living humans descend by coalescence from a single female and male respectively and mutations from the common ancestors accumulated along lineages can be used for tracking back their evolutionary history. However, the haploidy makes mtDNA and Y chromosome more sensitive to the effect of genetic drift which may significantly skew the haplogroup frequencies. This is particularly true for the Y chromosome, considering the differences between male and female effective population sizes (Jorde *et al.,* 2000). Furthermore, the NRY is far larger than the mitochondrial genome (about 60 Mb in respect to 16.569 bp), giving to this system huge potentiality for reconstructing human evolution, at least from the paternal side.

Recent advances in molecular biology, namely, the development of Denaturing-High Performance Liquid Chromatography (D-HPLC) (Underhill *et al.*, 1997), improved the knowledge of the phylogeny of the Y-chromosome at an extraordinary detail, and the number of known polymorphisms increased dramatically in the past decade from just a handful to over 600 (Karafet *et al.*, 2008). This technique made easy, fast and inexpensive the detection of new nucleotide changes in a given stretch of DNA, and it was possible to identify a large number of biallelic markers, called single nucleotide polymorphisms (SNPs) all over the molecule. The mutational event that introduces a SNP in the NRY is so rare as to be considered unique and recurrent independent mutations are practically absent (Underhill *et al.*, 2000). This feature allows a hierarchical approach in reconstructing the molecular phylogeny, so when different haplotypes share a SNP are considered derived from a common ancestor and clustered together in a haplogroup. Moreover, fast evolving multiallelic markers, such as Short Tandem Repeats (STRs), are also present in the Y-chromosome and their variability, considering an average mutation rate of about 6,9 x 10^{-4} per generation (25 years) (Zhivotovski *et al.*, 2004), can be used to date the approximate age of a haplogroup. In addition, the combined use of fast evolving STRs can also provide further helpful information to dissect the evolutionary history of lineages where no downstream SNP markers are available.

Early papers used different nomenclatures to define haplogroups (Underhill *et al.,* 2000; Semino *et al.,* 2000; Hammer *et al.,* 2001: Rosser *et al.*, 2000) and this fact, together with the contemporary existence of a STR-based classification (Jobling & Tyler-Smith, 2000), introduced unwanted source of misunderstanding. To overcome this inconvenient, a strictly cladistic nomenclature, paralleled to the one in use for mtDNA, has been proposed and now commonly accepted (Y Chromosome Consortium, 2002). The new system is based on capital letters in order to identify the broader clades and a succession of numbers and letters for lower hierarchical levels, thus flexible enough to allow the unambiguous naming of haplogroups defined by newly discovered downstream markers. However, the internal nodes are highly sensitive to changes in tree topology and to the addition of new SNPs which resolve paraphyletic into monophyletic groups. This occurrence may require the periodical update of the nomenclature (Karafet *et al.*, 2008), especially for the commonest haplogroups such as E, I or J, and creating disorder when comparing data between papers published in a different time. Moreover, a lineage showing a root SNP, but not presenting any of the derived ones, is identified by a * (star) symbol, but it should be distinguished by lineages where derived SNP were not tested, and this is obviously the case if the derived SNP were not yet discovered. For this reason, a "star" haplogroup may not have the same meaning in different papers. Moreover, to distinguish haplogroups where downstream markers were assayed but not detected, at the haplogroup name is usually added, in brackets, the symbol "x" followed by the name of the excluded haplogroups. So, to overcome possible ambiguities, it is strongly advisable (after a statement of all the SNPs assayed, including those not detected) to identify a given lineage adding to the haplogroup name the last downstream SNP observed.

The distribution of haplogroups in a given territory is not random, reflecting the spread of human beings in the past. Phylogeographic analysis confirms the African origin of *Homo sapiens* (Underhill *et al.*, 2001) and also provides clues concerning past human movements at continental and regional level. Geographical clines of Y-chromosome haplogroups in Europe have been reported (Semino *et al.,* 2000; Rosser *et al.*, 2000) and their distribution can be related to archaeological, linguistic and other non genetic data to infer the past events of the peopling of the continent. In this work we review papers dealing with Y-chromosome variability based on SNPs and focused on Europe and neighboring areas (Anatolia, the Caucasus, the Levant and North Africa). In order to homogenize the data from many studies carried out at various depth of analysis, the haplogroup frequencies for all the population studied are reported in Appendix 1 in a summarized form, tracing back into the main haplogroups the sub-clades not analyzed in the whole set of the considered papers. The complete data set is available online at the Journal of Anthropological Science website as supplementary material. Some papers (Weale *et al.*, 2002; Di Giacomo *et al.*, 2003; Dupuy *et al.*, 2006), using STRs and/or a limited set of SNPs, could not provide the necessary level of resolution to be included in the overall analysis but they were taken into account for the general discussion. Other papers focused on the phylogeny of specific haplogroups (Di Giacomo *et al.,* 2003; Rootsi *et al.,* 2004; Semino *et al.,* 2004; Cruciani *et al.,* 2004) have been used for the drawing of the frequency maps of the respective haplogroup.

History

Figure 1 shows the phylogenetic tree inferred from the data reported in Appendix 1 (see also the Supplementary Data available online). The tree does not report all the haplotypes recently identified and described at the International Society of Genetic Genealogy website. Genetic genealogy is a field in exponential growth: in the last few years, a number of commercial companies are offering tests at affordable prices and lot of people are becoming more and more interested in their ancient origin, supplying DNA samples. Their private laboratories are engaged in active research, providing a significant increase in the detection of new SNPs and, consequently, new haplogroups. However, this review takes into account only scientific papers referring to population studies and Figure 1 reports only haplogroups retrieved in these works.

Haplogroup A

Defined by the mutation M91, this haplogroup is the first branch of the evolutionary tree of mankind and it is diffused in sub-Saharan Africa with higher frequencies in Khoisan populations (Underhill *et al.,* 2001). This haplogroup can be sporadically found in southern European populations such as Greeks (Di Giacomo *et al.,* 2003; Capelli *et al.,* 2006) and Portuguese (Goncalves *et al.*, 2005), although in these studies no downstream markers were assayed and European Y-chromosomes belonging to haplogroup A may not be closely related with the southern African lineages. An early divergence originated the East African subclade A3b2-M13, (Semino *et al.*, 2002) found at high percentage in Falasha (Ethiopian Jews) (Cruciani *et al.,* 2002; Shen *et al.,* 2004) and consistently observed, although at very low frequencies, in Sardinia (Semino *et al.,* 2000; Passarino *et al.,* 2001; Contu *et al.,* 2008). However, the extreme rarity of this haplogroup does not offer clues about its possible appearance in Europe.

Haplogroup C

This clade is derivative of the CR super-haplogroup, defined by the M168 change and representing the ancestor type of the "out of Africa" event. It is characterized by the mutations M216 and RPS4Y (also known as M130). It was possibly originated shortly after M168 and was diversified in the early human migration along the southern shores of Asia, although it can be found nowadays at higher frequencies in northern Eurasia. Its occurrence in Europe, including Caucasus and Anatolia, is scarce, and chiefly limited to some southern populations (Semino *et al.,* 2000; Passarino *et al.,* 2001; Cinnioglu *et al.,* 2004; Nasidze *et al.,* 2004a; Goncalves *et al.,* 2005; Capelli *et al.,* 2006).

Haplogroup E

The *Alu* insertion known as YAP+ (or M1), together with the SNPs M145 and M203, is the common ancestral state of DE super-haplogroup, and the E clade identified by M96 has

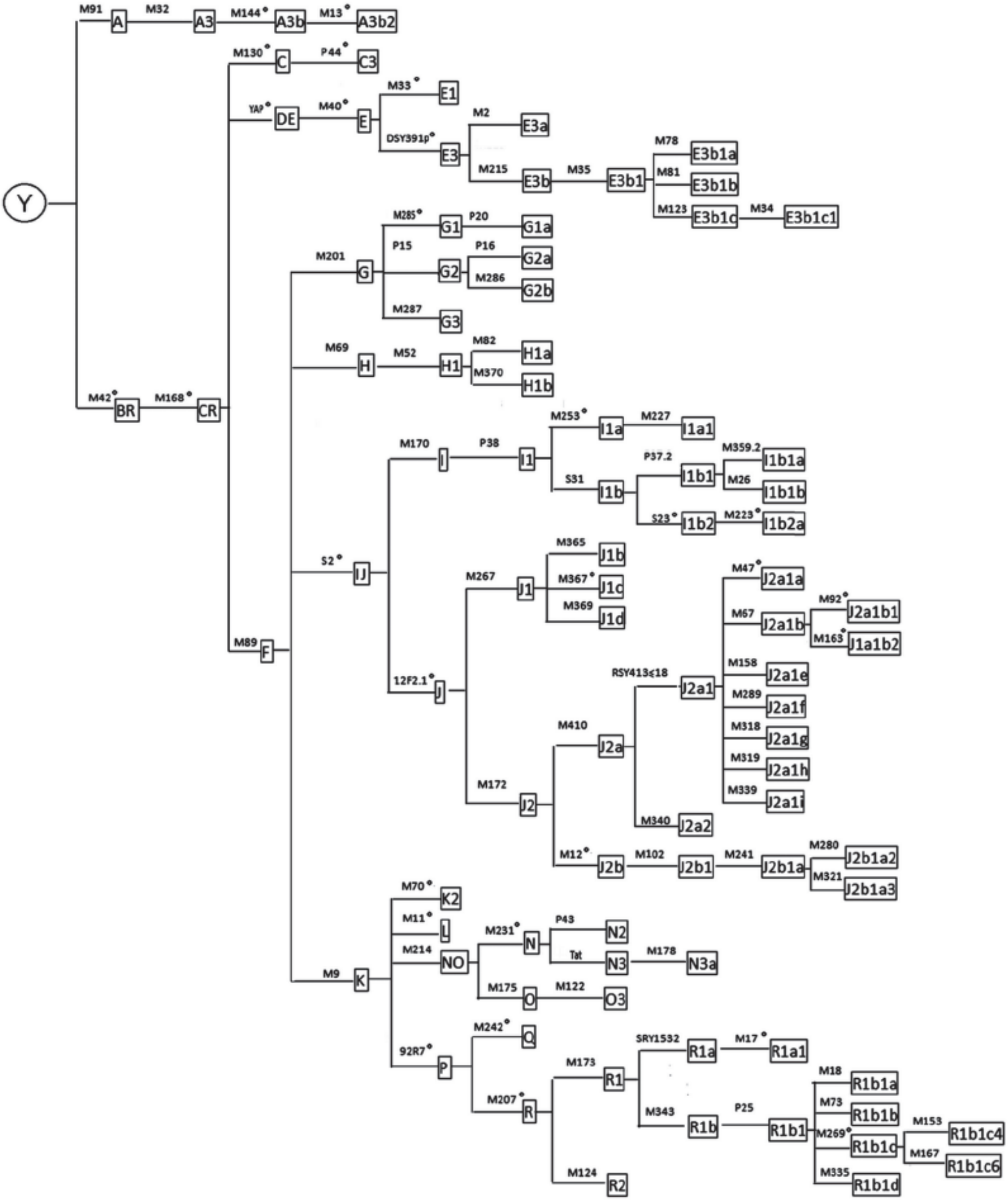

Fig. 1 - Phylogenetic tree of European haplogroups. The nodes defined by more than one marker are indicated by the symbol: °. Only the marker more commonly investigated is reported. For a complete list of the markers see: http://www.isogg.org.

been considered a "back to Africa" migration event (Cruciani *et al.,* 2002). In fact, most of its branches such as E1-M33, E2-M75 and E3a-M2 can be found in Africa and only at very low percentage in Mediterranean Europe (Cinnioglu *et al.,* 2004; Capelli *et al.,* 2006; Bosch *et al.,* 2006), particularly in the Iberian peninsula (Flores *et al.,* 2003; Flores *et al.,* 2004; Alonso *et al.,* 2005). However, this interpretation has been challenged by the discovery of a rare deep-rooting clade in Nigeria, suggesting an expansion from West Africa to Asia (Weale *et al*., 2003)

The sub-clade E3b1, defined by the mutation M35, has a wide distribution in Africa, Near East and Europe (Fig. 2a). This haplogroup probably appeared in the Horn of Africa and diffused during the Upper Paleolithic and Mesolithic periods in the Near East. The European distribution is compatible with the Neolithic demic diffusion of agriculture. In fact, two sub-clades, E3b1a-M78 and E3b1c-M123 (Fig. 2b-c) present a higher occurrence in Anatolia, the Balkans and the Italian peninsula, closely resembling to the first principal component described by Ammerman & Cavalli-Sforza (1984) and attributed to the spread of farming from the Fertile Crescent in Europe. The other clade, E3b1b-M81 (Fig. 2d) is associated with the Berber populations and can be found in regions that established a gene flow with Northern Africa, such as the Iberian peninsula (Maca-Meyer *et al.,* 2003; Flores *et al.,* 2004; Goncalves *et al.,* 2005; Alonso *et al.,* 2005; Beleza *et al.,* 2006), including the Canary Isles (Flores *et al.,* 2003), and Sicily (Semino *et al.,* 2004; Cruciani *et al.,* 2004).

Haplogroup F

This ancient haplogroup, that accounts for the majority of non African human Y-chromosomes with its derivate haplogroups from G to R, may represent a secondary migration out of Africa. The ancestral allele M89 not followed by any

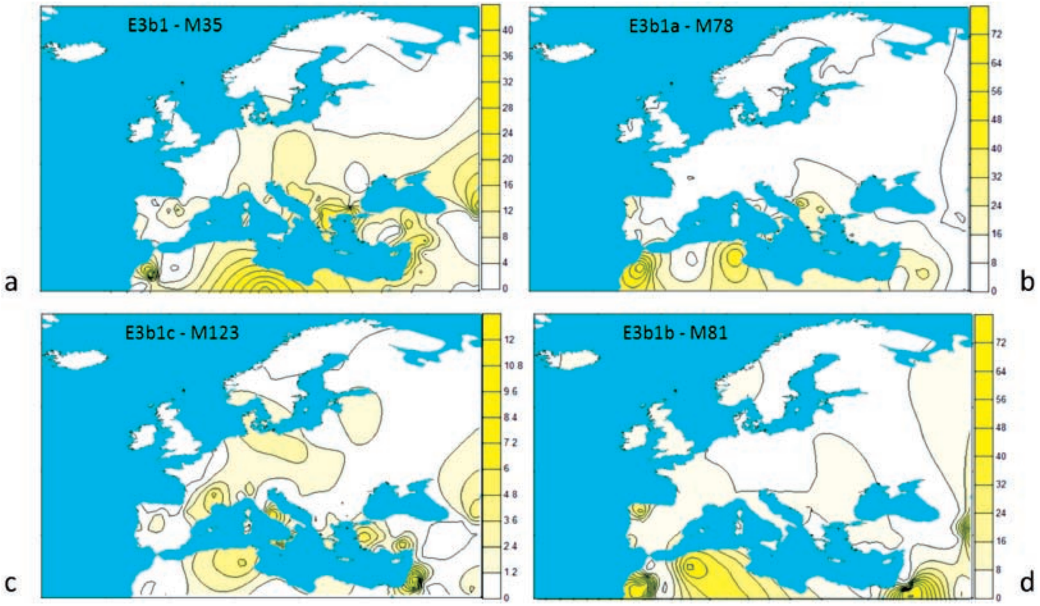

Fig. 2 - Frequency distribution maps for the haplogroup E. All the maps of this Figure, and of Figure 3 to 6, are drawn using the software Surfer 8.02, and the color code of the haplogroups follows that reported in Figure 8. Legend – a) E3b1-M35; b) E3b1a-M78; c) E3b1c-M123; d) E3b1b-M81. The colour version is available online at the JASs web site.

downstream markers can be found rather homogeneously at low frequencies in many European populations (Fig. 3a) but it has a clear peak in the Caucasus (Nasidze *et al.,* 2004a) and in the Middle East (Semino *et al.,* 2000). However, it is possible that the occurrence of F*-M89 haplogroup might be artfully enhanced by the poor analytical resolution of some papers, where not all the possible derivate markers were analyzed.

Haplogroup G

This haplogroup is defined by the SNP M201 and may be originated in Western or Southern Asia. Its occurrence in Europe has often been related with the Neolithic expansion (Semino *et al.,* 2000) but the distribution is rather uneven (Figure 3b), with very high frequency in isolated populations, such as Caucasian Georgians and Ossetians, from 30% to 70%, (Semino *et al.,* 2000; Nasidze *et al.,* 2004a,b) and northern Sardinians (about 20%) (Contu *et al.*, 2008), possibly indicating an ancient Upper Paleolithic widespread presence, now restricted to marginal areas.

Haplogroup H

This seems to be an Indian haplogroup, present in Europe almost exclusively in the Balkans, carried by the nomadic Gypsies who migrated in Europe in historic time from the Indian subcontinent, with the sub-clade H1a-M82. This mutation has been found in almost 60% of the Y-chromosomes of the Macedonian Roma people (Pericic *et al.*, 2005b). Other semi-nomadic populations of uncertain origin from the Balkans, such as the Aromuns are, however, not related to the Gypsies, and this is demonstrated by the lack of haplogroup H (Bosch *et al.,* 2006).

Haplogroup I

The haplogroup I, related with haplogroup J by common mutations, is characterized by the SNP M170. It is a primarily European haplogroup and may be arrived in Europe in Upper Paleolithic with the spread of the Gravettian culture (Semino *et al.,* 2000). The overall distribution presents a frequency cline from eastern to western Europe (Fig. 4a) but, given its ancient

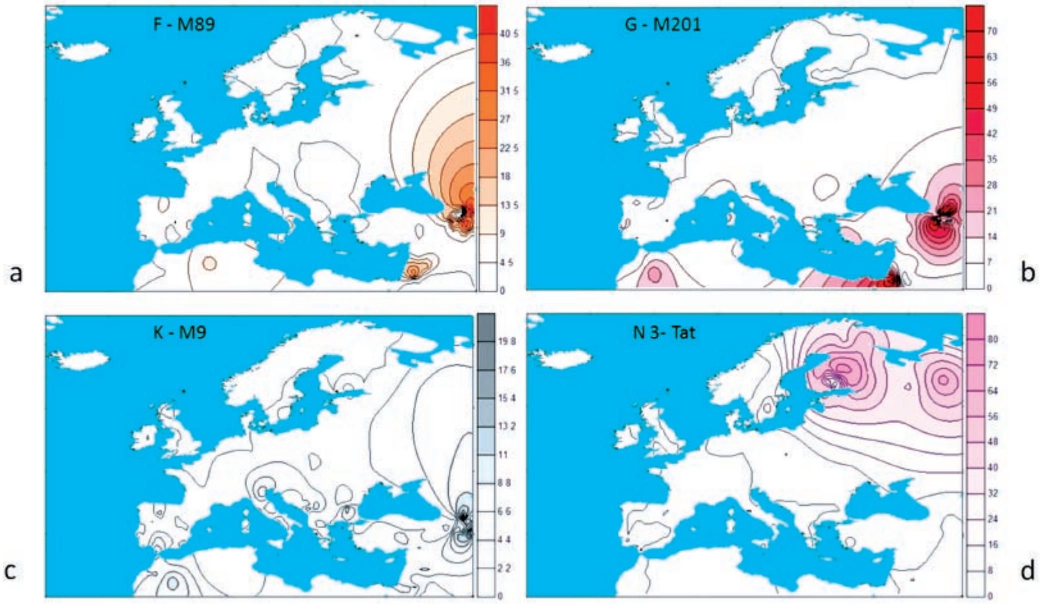

Fig. 3 - Frequency distribution maps for the haplogroups F, G, K and N. Legend – a) F-M89; b) G-M201; c) K-M9; d) N3-Tat.
The colour version is available online at the JASs web site

presence in Europe and consequent local diversification, different sub-clades present clues of independent pathways of expansion.

The sub-clade I1a-M253 reaches high frequencies in Scandinavia (Rootsi *et al.,* 2004; Karlsson *et al.,* 2006; Lappalainen *et al.,* 2006) while it is much less frequent outside northern Europe (Fig. 4ab). However, in spite of its rather scarce occurrence, this clade has a higher diversity in the south west, suggesting its origin in the Iberian refuge during the Last Glacial Maximum (LGM) and a further event of repopulation along the north-western shores, up to Fennoscandia.

The sub-clade I1b1-P37.2, largely represented in the Balkans (Fig.. 4c), with higher occurrence in Croatia (Barac *et al.,* 2003; Pericic *et al.,* 2005a; Pericic *et al.,* 2005b; Marjanovic *et al.,* 2005), could be spread westward by the post glacial migration of people coming from the Balkan refuge.

Another sub-clade, I1b2a-M223 reaches its peaks in central Europe, and it is correlated with the area of influence of Germanic people (Fig. 4c). In fact, it reaches its maximum in Germany but it is present at appreciable frequencies from France, to central-northern Italy, to eastern Europe (Rootsi *et al.,* 2004).

An apparent founder effect is shown from the I1b1b clade, defined by the mutation M26, that reaches a frequency of about 40% in the Sardinians (Semino *et al.,* 2000; Passarino *et al.,* 2001; Capelli *et al.,* 2006; Contu *et al.,* 2008) but it is quite rare elsewhere, and even virtually absent in the neighboring island of Corsica (Scozzari *et al.,* 2001; Francalacci *et al.,* 2003) (Fig. 6d). As for the I1a clade, the rare but regular presence of M26 chromosomes in the Iberian peninsula (with significant occurrence in Basques), in Morocco and in the British Isles (Capelli *et al.,* 2003) make its likely source in the Iberian refuge during the LGM.

Haplogroup J

Haplogroup J, defined by the 12f2 polymorphism, presents an east-west gradient in Europe, indicating its arrival from the Near East. However,

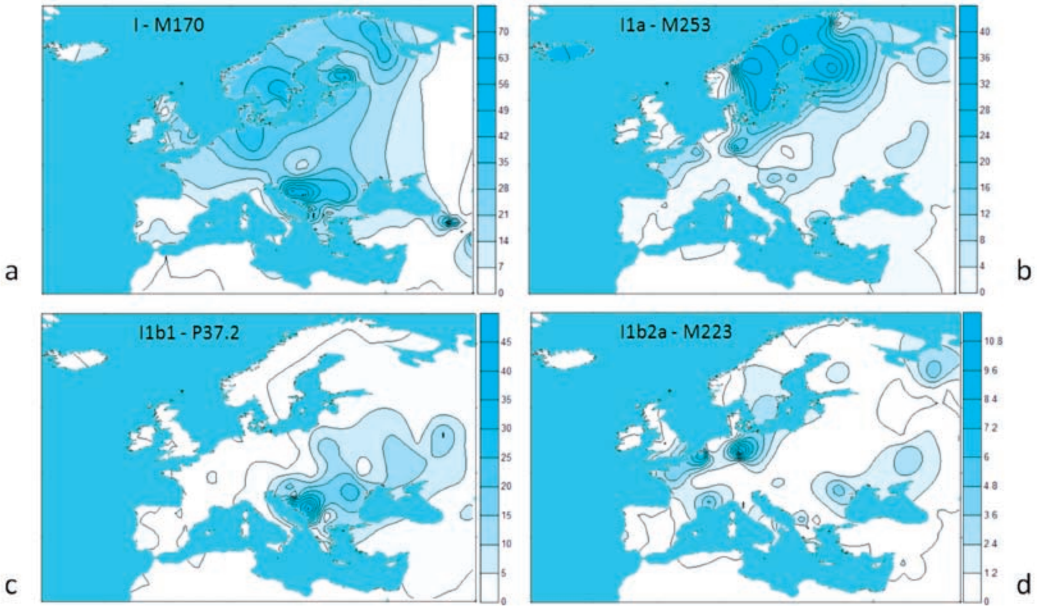

Fig. 4 - Frequency distribution maps for the haplogroup I. Legend – a) I-M170; b) I1a-M253; c) I1b1-P37.2; d) I1b2a-M223.
The colour version is available online at the JASs web site.

the two sister sub-clades, J1 and J2 have a dissimilar distribution, reflecting different peopling pathways. J1-M267 has its peaks in the Levant and in northern Africa and it is closely associated to the diffusion of Arab people, dropping abruptly outside of this area (including Anatolia and the Iberian peninsula), even if it shows an appreciable percentage in Sicily (Semino *et al.,* 2004) (Fig.5a).

The clade J2, defined by M172, has it higher frequencies in Anatolia and Mesopotamia, and decreases westwards (Semino *et al.,* 2004) (Fig. 5b). This haplogroup, particularly with its subclade J2a1b1-M92 (Fig. 5c), presents a similar pattern of diffusion of E31b1a-M78 and it is apparently correlated to the Neolithic demic expansion. Some downstream branches, such as J2b-M12 and J2a1b-M67, relatively frequent in Greece, southern Italy and western Turkey (Fig. 5d) are compatible with the influence of the Greek colonization (Di Giacomo *et al.*, 2004). The presence of J2 haplotypes in central Italy may be related to the debated hypothesis of an Anatolian origin of the Etruscan people, as suggested by mtDNA analysis on modern and ancient samples (Francalacci *et al.,* 1996; Vernesi *et al.,* 2004; Achilli *et al.,* 2007), although it could be arrived in the region much earlier following the Neolithic expansion.

Haplogroup K

It is characterized by the mutation M9 and is the ancestral haplogroup for most Y-chromosomes of northern Eurasia and Americas. The Y-chromosomes belonging to haplogroup K* are diffused at low frequencies in many European populations, with a slight prevalence in the south (Flores *et al.,* 2004; Marjanovic *et al.,* 2005), but further analysis could place them in some of the derived clades such as K2-M70. A higher frequency of this haplogroup can be observed in the Caucasus (Nasidze *et al.,* 2004a) (Fig. 3c).

Haplogroup L

This haplogroup, defined by the M11 and M20 SNPs, is predominantly Indian, but it has

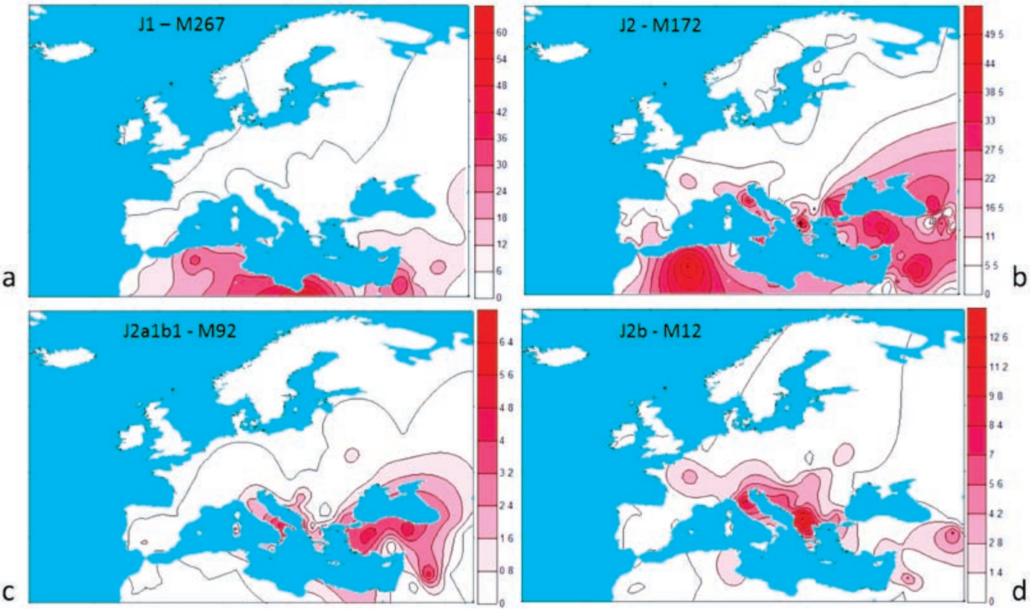

Fig. 5 - Frequency distribution maps for the haplogroup J. Legend – a) J1-M267; b) J2-M172; c) J2a1b1-M92; d) J2b-M12.
The colour version is available online at the JASs web site.

also been found at remarkable percentage in the Druze community in Lebanon (Shen *et al.,* 2004). It is also reported, at much lower frequencies, in Turkey (Cinnioglu *et al.,* 2004), and scattered in various southern and central European populations (Semino *et al.,* 2000).

Haplogroup N

This haplogroup, characterized by the mutations LLY22G and M231, is possibly originated in the Siberian region (Derenko *et al.,* 2007) and carried out in Scandinavia and northern Russia by Uralic speaking peoples (Fig. 3d). In fact, although its derived clade N3-Tat (also known as M46) is prevalent among Finnish and Saami people (Tambets *et al.,* 2004; Lappalainen *et al.,* 2006; Karlsson *et al.,* 2006), it is not so frequent in Germanic speaking Swedish and Norwegians (Weale *et al.,* 2002; Behar *et al.,* 2003; Capelli *et al.,* 2003; Karlsson *et al.,* 2006; Lappalainen *et al.,* 2006) and it is only present in one out of 181 Icelanders (Helgason *et al.*, 2000). It is interesting to notice that the Hungarians, who speak a language belonging to the Ugro-Finnic family, do not show any occurrence of Y-chromosomes attributed to haplogroup N (Semino *et al.,* 2000), pointing out to a significant admixture of the Uralic Magyar invaders with the local populations.

Haplogroup P

This rare haplogroup, defined by M45 and 92R7 among other markers, probably arose in northern Eurasia, possibly Siberia, and soon gave origin to two sister haplogroups much more successful: Q, that encompass with the Q3 sub-clade almost the totality of the Amerindian (not Na-Dene) Y-chromosomes, and R, that is the most frequent European haplogroup. Y-chromosomes belonging to haplogroup P* are sporadically observed in Europe with the only significant exception of the Dalmatian Isles of Korcula and Hvar, where it reaches a frequency of 14.3%, indicating a connection with central Asian populations such as the Avars (Barac *et al.,* 2003).

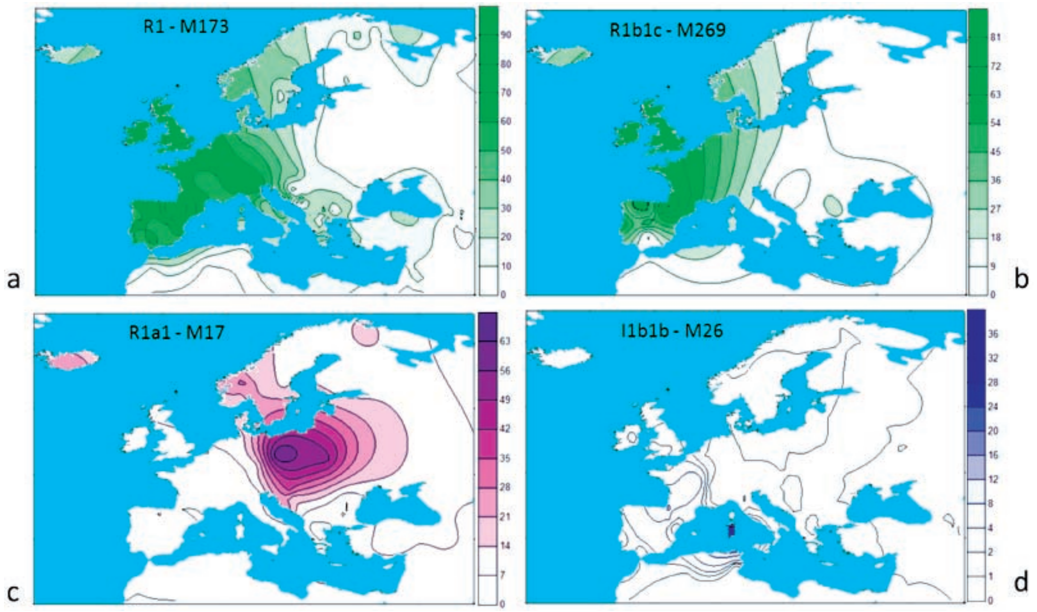

Fig. 6 - Frequency distribution maps for the haplogroups R and I. Legend – a) R1-M173; b) R1b1c-M269; c) R1a1-M17; d) I1b1b-M2.
The colour version is available online at the JASs web site.

Haplogroup R

Haplogroup R, characterized by the M207 change, was originated probably in north west Eurasia and it is divided in two sub-clades. The less frequent is R2-M124: this is mainly an Asian haplogroup, that can be found only at the edge of Europe, such as the Caucasus (Nasidze *et al.,* 2004b), Turkey (Cinnioglu *et al.,* 2004), and in the Sinte Roma (Gypsy) of known Indian origin (Wells *et al.*, 2001). Surprisingly, a few R2-M124 individuals were found in the Tyrrhenian islands of Sardinia (Semino *et al.,* 2000; Passarino *et al.,* 2001) and Corsica (Francalacci *et al.,* 2003).

The sub-clade R1, defined by the mutation M173, is the most common European haplotype (Fig. 6a), although its early branches can be found also elsewhere, and notably in northern Cameroon (Cruciani *et al.,* 2002), where it represents an ancient back migration to Africa from Asia. The European lineages are believed to arrive in the Upper Paleolithic, carrying the Aurignatian culture (Semino *et al.,* 2000), and they were confined in southern refuges during the LGM. After the withdrawal of the glaciers, R1, already differentiated in separate sub-clades, began to recolonize Europe. The expansion of R1b1c-M269 was originated from the Iberian refuge, and it is now typical of western Europe, reaching frequencies as high as 85% in Ireland (Moore *et al.*, 2006) (Fig. 6b). A sister clade, R1b1a-M18 is a private haplogroup of Sardinia, where it is present at low frequencies all over the island (Contu *et al.*, 2008). Another R1 subclade, R1a1-M17, has a very different distribution (Fig. 6c), having its maximum in eastern Europe with frequencies passing the 50% among Slavic people such as Sorbs, Ukrainians, Polish and Belarusians, but also in the linguistically non Slavic Hungarians (Semino *et al.,* 2000; Behar *et al.,* 2003). This haplogroup may be originated in the Ukrainian refuge during the LGM but its expansion is more likely related to the diffusion of the Kurgan culture, carried by horsemen speaking proto-Indo-European languages, as hypothesized by Semino *et al.* (2000). The R1a1-M17 Y-chromosomes reached with significant percentage Greece in the south, up to the island of Crete (Martinez *et al.*, 2007) and Scandinavia in the north (Karlsson *et al.,* 2006; Lappalainen *et al.,* 2006). The M17 mutation has been consider a marker of the Viking influence since it can be found at high frequencies in Iceland (Helgason *et al.,* 2000) and in the Northern Isles (Shetland and Orkney) (Capelli *et al.,* 2003), while its absence is indicative of a scant admixture of Norse invaders with the local populations, as happened in Ireland (McEvoy *et al.*, 2006).

Geography

The data reported in Appendix 1 have been used for describing the European Y-chromosome variation. The Principal Component Analysis has been carried out using the XLStat (version 7.5.2) package. The first two principal components, accounting for the 25,47% of the total variance, clearly separated the peripheral populations (Anatolia, the Caucasus, the Levant and North Africa) from the more homogeneous proper European ones (Fig. 7). The main discriminant contribution is given, in the first component, by the J2-M172 haplogroup in respect to I-M170, and, in the second component, by E3b1-M35 in respect to R1-M173. The first principal component (showing the Levantine and Anatolian populations on the positive portion of the axis, and the northern and western Europeans on the negative, with Italians and Balkan population in intermediate position) reflects the main demographic event of the peopling of Europe: the demic diffusion westward of Neolithic farmers. However, from the standpoint of Y-chromosome variability, this important migration wave seems to be limited to the south-eastern European areas, while the north-west keeps its ancient Paleolithic features. The second component separates the Anatolian and Caucasian regions from the Levant and north Africa, two broad areas subject to different peopling histories: the former characterized by the presence of ancient haplogroups, such as F-M89, G-M201 and K-M9 and by central Asian influence; the latter signed by the admixture between Berber haplogroups, such as E3b1b-M81, with markers of the Arab expansion, such as J1-M267. The more ancient event, but with reduced demographic impact, of the Mesolithic repeopling of Europe after the LGM, eastward from the Iberian refuge, and westward from the Balkan (and possibly Ukrainian) refuge, can be seen, less clearly, in the second component. In fact, the north-eastern populations, characterized by high frequencies of I-M170, N3-Tat and R1a1-M17, tend to be placed in the positive portion of the axis, while the north-westerns populations, showing high frequencies of the R1b1c-M269 sub-clade, are mainly located in the opposite position. An outlier of the European variation is represented by the Macedonian Roma people, whose exogenous origins is well known. The third principal component (data not shown) separates Sardinia from the rest of Europe, having I1b1b-M26 as the major contributing variable to this component.

Figure 8 shows the distribution of the main haplogroups in Europe, which will be discussed in detail below.

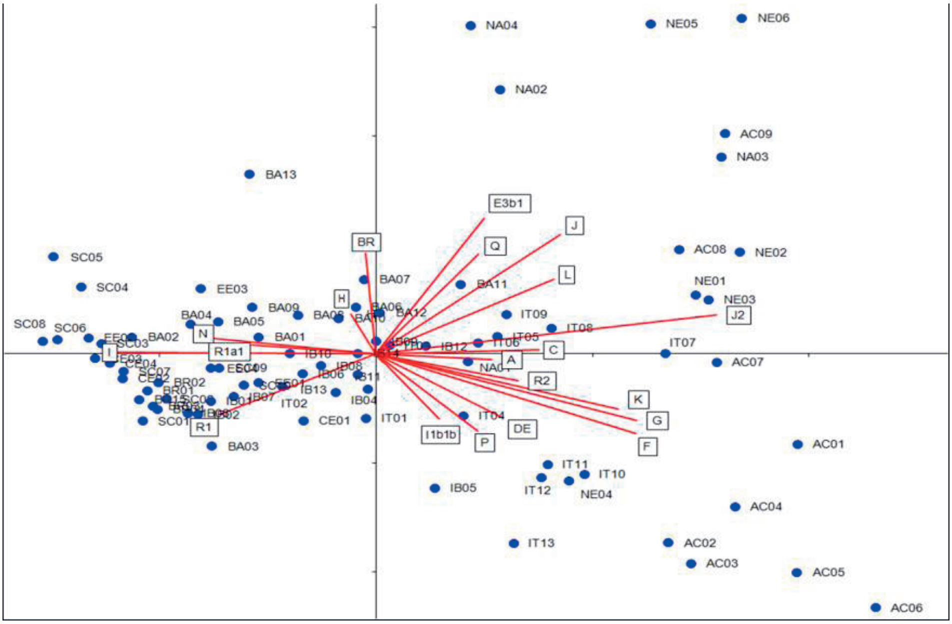

Fig. 7 - Plot of the first two principal component carried out on the data reported in Tab.1. For graphic clearness, different samples coming from the same region and analyzed by the same Authors have been pooled together. Red lines indicate the relative contribution of the variables (haplogroups).

Legend – IB=Iberian peninsula: 01-03=Basque, 04-07=North Spain, 08-10=Portugal, 11-13=South Spain, 14=Canary Isles; BR=British Isles: 01,05=Scotland, 02=England, 03-04=Ireland; CE=Central Europe: 01=France, 02=Nederland, 03-04=Germany; EE=Eastern Europe: 01=Czech and Slovakia, 02=Poland, 03=Hungary, 04=Ukraine; SC=Fennoscandia: 01-03=Scandinavia, 04-05=Finland, 06-08=Saami, 09=Iceland. – IT=Italian peninsula: 01-03=Central-North Italy, 04=Corsica, 05-07=South Italy, 08-09=Sicily, 10-13=Sardinia. – BA=Balkans: 01=Romania, 02-05=Former Yugoslavia, 06-09=Albania and Macedonia, 10-11=Greece, 12=Aromuns, 13=Macedonian Rom. AC=Anatolia and Caucasus: 01-02=Georgia, 03=Armenia, 04=Azerbaijan, 05=South Caucasus, 06=North Caucasus, 08-10=Turkey; NE=Levant: 01=Cyprus 02-03=Lebanon, 04=Syria, 05=Ashkenazi, 06=Israel; NA=North Africa: 01=Malta, 02-03=Algeria And Tunisia, 04=Berber.

The colour version is available online at the JASs web site.

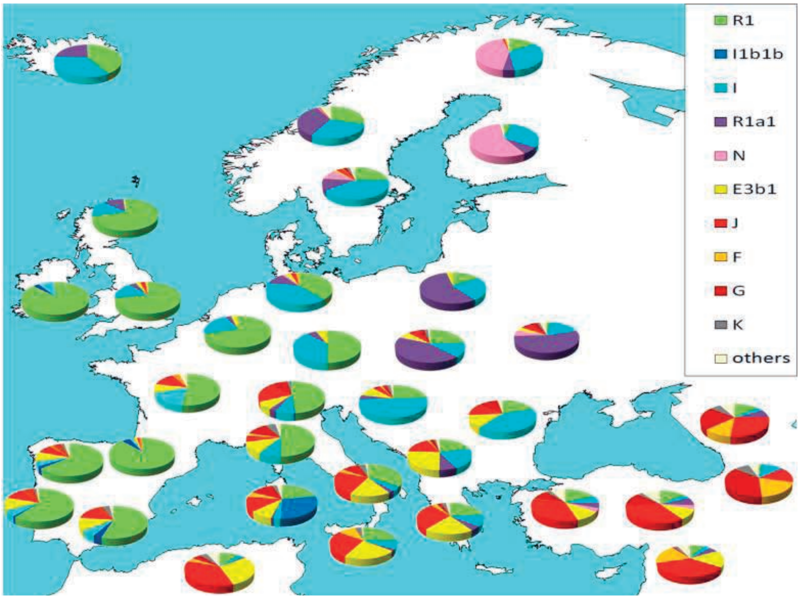

Fig. 8 - Frequency distribution of the main haplogroups in Europe. The colour version is available online at the JASs web site.

Iberian peninsula

The region was a refuge during the LGM and its Y-chromosome pool reflects features of isolation, differentiation, and subsequent expansion, being the source of the Mesolithic repeopling of the inhabited areas toward north and east. Among the haplogroups here originating, R1b1c-M269 is particularly important, accounting for the majority of western European Y-chromosomes. It encompasses the large majority of Basque variability (Alonso *et al.*, 2005) and it is generally predominant everywhere in the region. Although relatively infrequent in its territory, the Iberian peninsula also originated sub-clades belonging to the I haplogroup, such as I1b1b-M26, and possibly, I1a-M253, presently showing elevated percentages in Sardinia and Scandinavia respectively. Traces of the ancient Neolithic expansion wave can be observed mainly in the south (Bosch *et al.*, 2001) at a frequency of about 10%, while the contribution of a north African gene flow is negligible (Bosch *et al.*, 2001; Flores *et al.*, 2004) in spite of the long-time Arab rule.

British Isles

The region presents different paternal ancestries in a continuous range from the more autochthonous populations, with a clear Iberian post glacial footprint and are characterized by high frequencies of R1(xR1a1)-M173 haplotypes, to the areas more interested by foreign gene flow, such as the south-east, that denotes the Anglo-Saxon influence for the higher incidence of the I(xI1b2)-M170, and the Scottish Isles, where the Viking occupation is documented by the significant occurrence of R1a1-M17 (Capelli *et al.*, 2003). The Neolithic markers (E3b-M35 and J2-M172) are relatively rare in the south and virtually absent in Scotland. Ireland shows a homogeneous peopling of predominant Iberian origin, while the contacts with the Norse left a limited genetic trace (McEvoy *et al.*, 2006).

Central-Eastern Europe

Central Europe has been poorly investigated in comparison to other European areas. It is characterized by the admixture of the two main

Mesolithic repeopling waves: the Iberian and the Balkan, with R1-M173 and M-170 respectively as the main contributing haplogroups. A typical central European marker is I1b2a-M223 with frequencies approaching 13% in Germany (Rootsi *et al.,* 2004). Eastern Europe shows similarities with the central areas of the continent, but with a reduced presence of R1-M173, complemented by a significant increase of R1a1-M17 Y-chromosomes. This haplogroup, which reaches higher value in the Slavic area, was probably originated in the Ukrainian glacial refuge and spread westward and southward up to India (Kivisild *et al.,* 2003; Sengupta *et al.,* 2006) by the Indo-European speaking people in more recent time. A sharp genetic differentiation was observed between German and Slavic speaking populations, mainly due to the prevalence of R1-M173 in Germany and R1a1-M17 in Poland (Kayser *et al.*, 2005).

Fennoscandia

This area presents a clear distention between Germanic and Ugro-Finnic populations. Norwegian, Swedish and Icelandic samples present roughly even distribution of three main haplogroups: R1a1-M17 of eastern origin and, from the west, I1a-M253 and R1b1c-M269. The Saami and the Finnish populations are characterized by a high prevalence of the N3-Tat marker, which reaches peaks of more than 70% near the Polar circle (Lappalainen *et al.,* 2006). The Y-chromosome variability of the Saami shows no evidence of aboriginal Siberian populations, indicating a distinctive European ancestry for this population (Tambets *et al.,* 2004).

Italian peninsula

The Y-chromosome variation in the Italian peninsula is clinal, along the north-south axis, with the increasing frequency of haplogroups arrived in Italy after the arrival of agriculture in the Neolithic (Capelli, 2007; Capelli *et al.*, 2007). Northern Italy shows similar frequencies with those of Central Europe, but with the prevalence of the western R1-M173 haplogroup in respect to the eastern I-M170, while E3b1-M35 and J2-M172, have a reduced weight. On the contrary, these two haplogroups are quite common in southern Italy, whereas J2-M172 reaches its higher occurrence in Europe, indicating a significant impact of the gene flow carried out by the Neolithic farmers. R1a1-M17 is rather rare, both in the north, where it probably comes from eastern Europe, and in the south, of possible Greek provenience. A certain occurrence of J2-M172 Y-chromosomes in Tuscany has been related to the Etruscan heritage of the region (Achilli *et al.,* 2007). The three major islands show a different demographic history: the Y-chromosome variability of Sicily shares a common history with that of southern Italy, enriched by an additional Arab contribution; Corsica has similar frequencies to central-northern Italy, possibly as a consequence of a recent gene flow that hampered an ancient Sardinian connection (Francalacci *et al.,* 2003); Sardinia, an outlier in European variability, has a peopling of ancient Iberian origin and shaped by founder effect, drift, and divergence of private haplotypes (Contu *et al.*, 2008).

Balkan peninsula

This region hosted a refuge during the LGM and contributed significantly to the repopulation of central-eastern Europe. In spite of the remarkable linguistic stratification of the region, the Y-chromosome variation is rather homogeneous in the main ethnic groups (Bosch *et al.,* 2006). However, a certain differentiation can be seen between the northern populations, because of the remarkable frequency shown by I1b1-P37.2 (with a peak in Croatia), and southern populations from Albania, Macedonia and Greece, where, as observed in southern Italy, the Near Eastern haplotypes J2-M172 and E3b1-M35 (especially with the sub-clade E3b1a-M78) reach high percentage, following the migration of Neolithic people in the area.

Anatolia and Caucasus

The Anatolian peninsula is a bridge between Europe and the Mesopotamian region, and it was a core area of the development of agriculture. Not surprisingly, the haplogroups linked to the Neolithic demic diffusion are quite frequent in this area, accounting cumulatively more than the 50% of the Y-chromosome variation with a prevalence of J2-M172, E3b1-M78 and E3b3-M123 clades (Cinnioglu *et al.,* 2004). The region shows also appreciable frequencies of typical European haplogroups such as I-M170, pointing out to a Balkan gene flow, and the main western European haplogroup R1b3-M269, indicative of an ancient Aurignatian substratum in the pre-Neolithic Turkey. A weak but noticeable gene flow from Central Asia is also present, enriching the Y-chromosome variability of the area. This Asiatic connection is much more apparent in the Caucasus region (Nasidze *et al.,* 2004a), where the most frequent Y-chromosomes belong to the basal hapologroups G-M201, F-M89 and K-M9, whose distribution only partially overlaps the Neolithic expansion, and they are possibly the remnant of an ancient post-glacial peopling.

North Africa and the Levant

The area presents a rather different Y-chromosome landscape, suggesting a genetic barrier between the two rims of the Mediterranean, in spite of a long history of cultural exchanges. The Berber clade E3b1b-M81 is only sporadically observed in the Iberian peninsula and the Arab clade J1-M267 has a appreciable occurrence only in Sicily (Semino *et al.,* 2004). The various Jews communities, although sharing a relative common variability, maintain traces of their historical provenience, as the Ethiopian Jews, characterized by the east African A3b2-M13 haplogroup (Shen *et al.,* 2004), and the Ashkenazy Jews who show high frequencies of the eastern European R1a1-M17 clade (Behar *et al.,* 2003).

Conclusions

The history of the European populations can be understood only by the integration of information coming from fields as linguistics, archaeology and anthropology with genetic data. In this regards, the rapid between-population divergence of unilinear markers and the ability to preserve elements of common ancestry, make the mitochondrial genome and the Y-chromosome unique tools for investigating migration events and relative degree of admixture with pre-existing populations. In particular, recent technical advancement brought the molecular phylogeny of Y-chromosome at unprecedented level of resolution, allowing a detailed reconstruction of the genetic relationships among the European regions, spanning from the Upper Paleolithic to historic time.

Info on the web

http://www.isita-org.com/jass/Contents/ContentsVol86.htm

Complete data set for Y-Chromosome aplogroup frequences in Europe

https://www3.nationalgeographic.com/genographic/

National Geographic, "Genographic Project"

http://www.isogg.org/

International Society of Genetic Genealogy

http://ycc.biosci.arizona.edu/

Y-Chromosome Consortium

http://www.ysearch.org/

Public search engine for Y-chromosome haplotypes modern humans

References

Achilli A., Olivieri A., Pala M., Metspalu E., Fornarino S., Battaglia V., Accetturo M., Kutuev I., Khusnutdinova E., Pennarun E., Cerutti N., Di G.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: 759-768.

Alonso S., Flores C., Cabrera V., Alonso A., Martin P., Albarran C., Izagirre N., de la Rua C. & Garcia O. 2005. The place of the Basques in the European Y-chromosome diversity landscape. *Eur. J. Hum. Genet.*, 13: 1293-1302.

Ammermman A.J. & Cavalli-Sforza L.L. 1984. *The Neolithic transition and the genetics of populations in Europe*. Princeton University Press, Princeton.

Barac L., Pericic M., Klaric I.M., Rootsi S., Janicijevic B., Kivisild T., Parik J., Rudan I., Villems R. & Rudan P. 2003. Y chromosomal heritage of Croatian population and its island isolates. *Eur. J. Hum. Genet.*, 11: 535-542.

Behar D.M., Thomas M.G., Skorecki K., Hammer M.F., Bulygina E., Rosengarten D., Jones A.L., Held K., Moses V., Goldstein D., Bradman N. & Weale M.E. 2003. Multiple origins of Ashkenazi Levites: Y chromosome evidence for both Near Eastern and European ancestries. *Am. J. Hum. Genet.*, 73: 768-779.

Beleza S., Gusmao L., Lopes A., Alves C., Gomes I., Giouzeli M., Calafell F., Carracedo A. & Amorim A. 2006. Micro-phylogeographic and demographic history of Portuguese male lineages. *Ann. Hum. Genet.*, 70: 181-194.

Bosch E., Calafell F., Comas D., Oefner P.J., Underhill P.A. & Bertranpetit J. 2001. High-resolution analysis of human Y-chromosome variation shows a sharp discontinuity and limited gene flow between northwestern Africa and the Iberian Peninsula. *Am. J. Hum. Genet.*, 68: 1019-1029.

Bosch E., Calafell F., Gonzalez-Neira A., Flaiz C., Mateu E., Scheil H.G., Huckenbeck W., Efremovska L., Mikerezi I., Xirotiris N., Grasa C., Schmidt H. & Comas D. 2006. Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns. *Ann. Hum. Genet.*, 70: 459-487.

Capelli C. 2007. The genetic history of Italy: a male perspective. *Journal of Anthropological Sciences*, 85:3-4.

Capelli C., Brisighelli F., Scarnicci F., Arredi B., Caglia' A., Vetrugno G., Tofanelli S., Onofri V., Tagliabracci A., Paoli G. & Pascali V.L. 2007. Y chromosome genetic variation in the Italian peninsula is clinal and supports an admixture model for the Mesolithic-Neolithic encounter. *Mol. Phylogenet. Evol.*, 44: 228-239.

Capelli C., Redhead N., Abernethy J.K., Gratrix F., Wilson J.F., Moen T., Hervig T., Richards M., Stumpf M.P., Underhill P.A., Bradshaw P., Shaha A., Thomas M.G., Bradman N. & Goldstein D.B. 2003. A Y chromosome census of the British Isles. *Curr. Biol.*, 13: 979-984.

Capelli C., Redhead N., Romano V., Cali F., Lefranc G., Delague V., Megarbane A., Felice A.E., Pascali V.L., Neophytou P.I., Poulli Z., Novelletto A., Malaspina P., Terrenato L., Berebbi A., Fellous M., Thomas M.G. & Goldstein D.B. 2006. Population structure in the Mediterranean basin: a Y chromosome perspective. *Ann. Hum. Genet.*, 70: 207-225.

Cavalli-Sforza L.L., Menozzi P. & Piazza A. 1994. *The History and Geography of Human Genes*. Princeton University Press, Princeton..

Cinnioglu C., King R., Kivisild T., Kalfoglu E., Atasoy S., Cavalleri G.L., Lillie A.S., Roseman C.C., Lin A.A., Prince K., Oefner P.J., Shen P., Semino O., Cavalli-Sforza L.L. & Underhill P.A. 2004. Excavating Y-chromosome haplotype strata in Anatolia. *Hum. Genet.*, 114: 127-148.

Contu D., Morelli L., Santoni F., Foster J.W., Francalacci P. & Cucca F. 2008. Y-chromosome based evidence for pre-Neolithic origin of the genetically homogeneous but diverse Sardinian population: inference for association scans. *PLOS One,* 3: e1430.

Cruciani F., La F.R., Santolamazza P., Sellitto D., Pascone R., Moral P., Watson E., Guida V., Colomb E.B., Zaharova B., Lavinha J., Vona G., Aman R., Cali F., Akar N., Richards M., Torroni A., Novelletto A. & Scozzari R. 2004. Phylogeographic analysis of haplogroup E3b (E-M215) y chromosomes reveals multiple migratory events within and out of Africa. *Am. J. Hum. Genet.*, 74: 1014-1022.

Cruciani F., Santolamazza P., Shen P., Macaulay V., Moral P., Olckers A., Modiano D., Holmes S., Destro-Bisol G., Coia V., Wallace D.C., Oefner P.J., Torroni A., Cavalli-Sforza L.L., Scozzari R. & Underhill P.A. 2002. A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *Am. J. Hum. Genet.*, 70: 1197-1214.

Derenko M., Malyarchuk B., Denisova G., Wozniak M., Grzybowski T., Dambueva I. & Zakharov I. 2007. Y-chromosome haplogroup N dispersals from south Siberia to Europe. *J. Hum. Genet.*, 52:763-770.

Di Giacomo F., Luca F., Anagnou N., Ciavarella G., Corbo R.M., Cresta M., Cucci F., Di S.L., Agostiano V., Giparaki M., Loutradis A., Mammi' C., Michalodimitrakis E.N., Papola F., Pedicini G., Plata E., Terrenato L., Tofanelli S., Malaspina P. & Novelletto A. 2003. Clinal patterns of human Y chromosomal diversity in continental Italy and Greece are dominated by drift and founder effects. *Mol. Phylogenet. Evol.*, 28: 387-395.

Di Giacomo F., Luca F., Popa L.O., Akar N., Anagnou N., Banyko J., Brdicka R., Barbujani G., Papola F., Ciavarella G., Cucci F., Di S.L., Gavrila L., Kerimova M.G., Kovatchev D., Kozlov A.I., Loutradis A., Mandarino V., Mammi' C., Michalodimitrakis E.N., Paoli G., Pappa K.I., Pedicini G., Terrenato L., Tofanelli S., Malaspina P. & Novelletto A. 2004. Y chromosomal haplogroup J as a signature of the post-neolithic colonization of Europe. *Hum. Genet.*, 115: 357-371.

Dupuy B.M., Stenersen M., Lu T.T. & Olaisen B. 2006. Geographical heterogeneity of Y-chromosomal lineages in Norway. *Forensic Sci. Int.*, 164: 10-19.

Flores C., Maca-Meyer N., Gonzalez A.M., Oefner P.J., Shen P., Perez J.A., Rojas A., Larruga J.M. & Underhill P.A. 2004. Reduced genetic structure of the Iberian peninsula revealed by Y-chromosome analysis: implications for population demography. *Eur. J. Hum. Genet.*, 12: 855-863.

Flores C., Maca-Meyer N., Perez J.A., Gonzalez A.M., Larruga J.M. & Cabrera V.M. 2003. A predominant European ancestry of paternal lineages from Canary Islanders. *Ann. Hum. Genet.*, 67: 138-152.

Francalacci P., Bertranpetit J., Calafell F. & Underhill P.A. 1996. 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.

Francalacci P., Morelli L., Underhill P.A., Lillie A.S., Passarino G., Useli A., Madeddu R., Paoli G., Tofanelli S., Calo C.M., Ghiani M.E., Varesi L., Memmi M., Vona G., Lin A.A., Oefner P. & Cavalli-Sforza L.L. 2003. Peopling of three Mediterranean islands (Corsica, Sardinia, and Sicily) inferred by Y-chromosome biallelic variability. *Am. J. Phys. Anthropol.*, 121: 270-279.

Goncalves R., Freitas A., Branco M., Rosa A., Fernandes A.T., Zhivotovsky L.A., Underhill P.A., Kivisild T. & Brehm A. 2005. Y-chromosome lineages from Portugal, Madeira and Acores record elements of Sephardim and Berber ancestry. *Ann. Hum. Genet.*, 69: 443-454.

Hammer M.F., Karafet T.M., Redd A.J., Jarjanazi H., Santachiara-Benerecetti S., Soodyall H. & Zegura S.L. 2001. Hierarchical patterns of global human Y-chromosome diversity. *Mol. Biol. Evol.*, 18: 1189-1203.

Helgason A., Sigureth a.S., Nicholson J., Sykes B., Hill E.W., Bradley D.G., Bosnes V., Gulcher J.R., Ward R. & Stefansson K. 2000. Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland. *Am. J. Hum. Genet.*, 67: 697-717.

Jorde L.B., Watkins W.S., Bamshad M.J., Dixon M.E., Ricker C.E., Seielstad M.T. & Batzer M.A. 2000. The distribution of human W. genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data. *Am. J.*

Hum. Genet., 66: 979–988 .

Jobling M.A. & Tyler-Smith C. 2003. The human Y chromosome: an evolutionary marker comes of age. *Nat. Rev. Genet.*, 4: 598-612.

Jobling M.A. & Tyler-Smith C. 2000. New uses for new haplotypes the human Y chromosome, disease and selection. *Trends Genet.*, 16: 356-362.

Karafet T.M., Mendez F.L., Meilerman M.B., Underhill P.A., Zegura S.L. & Hammer M.F. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res.*, 18: 830-838.

Karlsson A.O., Wallerstrom T., Gotherstrom A. & Holmlund G. 2006. Y-chromosome diversity in Sweden - a long-time perspective. *Eur. J. Hum. Genet.*, 14: 963-970.

Kayser M., Lao O., Anslinger K., Augustin C., Bargel G., Edelmann J., Elias S., Heinrich M., Henke J., Henke L., Hohoff C., Illing A., Jonkisz A., Kuzniar P., Lebioda A., Lessig R., Lewicki S., Maciejewska A., Monies D.M., Pawłowski R., Poetsch M., Schmid D., Schmidt U., Schneider P.M., Stradmann-Bellinghausen B., Szibor R., Wegener R., Wozniak M., Zoledziewska M., Roewer L., Dobosz T. & Ploski R. 2005. Significant genetic differentiation between Poland and Germany follows present-day political borders, as revealed by Y-chromosome analysis. *Hum. Genet.*, 117:428-443.

Kivisild T., Rootsi S., Metspalu M., Mastana S., Kaldma K., Parik J., Metspalu E., Adojaan M., Tolk H.V., Stepanov V., Golge M., Usanga E., Papiha S.S., Cinnioglu C., King R., Cavalli-Sforza L., Underhill P.A. & Villems R. 2003. The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. *Am. J. Hum. Genet.*, 72: 313-332.

Lappalainen T., Koivumaki S., Salmela E., Huoponen K., Sistonen P., Savontaus M.L. & Lahermo P. 2006. Regional differences among the Finns: a Y-chromosomal perspective. *Gene*, 376: 207-215.

Maca-Meyer N., Sanchez-Velasco P., Flores C., Larruga J.M., Gonzalez A.M., Oterino A. & Leyva-Cobian F. 2003. Y chromosome and mitochondrial DNA characterization of Pasiegos, a human isolate from Cantabria (Spain). *Ann. Hum. Genet.*, 67: 329-339.

Marjanovic D., Fornarino S., Montagna S., Primorac D., Hadziselimovic R., Vidovic S., Pojskic N., Battaglia V., Achilli A., Drobnic K., Andjelinovic S., Torroni A., Santachiara-Benerecetti A.S. & Semino O. 2005. The peopling of modern Bosnia-Herzegovina: Y-chromosome haplogroups in the three main ethnic groups. *Ann. Hum. Genet.*, 69: 757-763.

Martinez L., Underhill P.A., Zhivotovsky L.A., Gayden T., Moschonas N.K., Chow C.E., Conti S., Mamolini E., Cavalli-Sforza L.L. & Herrera R.J. 2007. Paleolithic Y-haplogroup heritage predominates in a Cretan highland plateau. *Eur. J. Hum. Genet.*, 15: 485-493.

McEvoy B., Brady C., Moore L.T. & Bradley D.G. 2006. The scale and nature of Viking settlement in Ireland from Y-chromosome admixture analysis. *Eur. J. Hum. Genet.*, 14: 1288-1294.

Moore L.T., McEvoy B., Cape E., Simms K. & Bradley D.G. 2006. A Y-chromosome signature of hegemony in Gaelic Ireland. *Am. J. Hum. Genet.*, 78: 334-338.

Nasidze I., Ling E.Y., Quinque D., Dupanloup I., Cordaux R., Rychkov S., Naumova O., Zhukova O., Sarraf-Zadegan N., Naderi G.A., Asgary S., Sardas S., Farhud D.D., Sarkisian T., Asadov C., Kerimov A. & Stoneking M. 2004a. Mitochondrial DNA and Y-chromosome variation in the caucasus. *Ann. Hum. Genet.*, 68: 205-221.

Nasidze I., Quinque D., Dupanloup I., Rychkov S., Naumova O., Zhukova O. & Stoneking M. 2004b. Genetic evidence concerning the origins of South and North Ossetians. *Ann. Hum. Genet.*, 68: 588-599.

Passarino G., Underhill P.A., Cavalli-Sforza L.L., Semino O., Pes G.M., Carru C., Ferrucci L., Bonafe M., Franceschi C., Deiana L., Baggio G. & De B.G. 2001. Y chromosome binary markers to study the high prevalence of males in Sardinian centenarians and the genetic structure of the Sardinian population. *Hum.*

Hered., 52: 136-139.

Pericic M., Barac L.L., Martinovic K., I, Janicijevic B. & Rudan P. 2005a. Review of Croatian genetic heritage as revealed by mitochondrial DNA and Y chromosomal lineages. *Croat. Med. J.*, 46: 502-513.

Pericic M., Lauc L.B., Klaric I.M., Rootsi S., Janicijevic B., Rudan I., Terzic R., Colak I., Kvesic A., Popovic D., Sijacki A., Behluli I., Dordevic D., Efremovska L., Bajec D.D., Stefanovic B.D., Villems R. & Rudan P. 2005b. High-resolution phylogenetic analysis of southeastern Europe traces major episodes of paternal gene flow among Slavic populations. *Mol. Biol. Evol.*, 22: 1964-1975.

Richards M. & Macaulay V. 2001. The mitochondrial gene tree comes of age. *Am. J. Hum. Genet.*, 68: 1315-1320.

Rootsi S., Magri C., Kivisild T., Benuzzi G., Help H., Bermisheva M., Kutuev I., Barac L., Pericic M., Balanovsky O., Pshenichnov A., Dion D., Grobei M., Zhivotovsky L.A., Battaglia V., Achilli A., Al-Zahery N., Parik J., King R., Cinnioglu C., Khusnutdinova E., Rudan P., Balanovska E., Scheffrahn W., Simonescu M., Brehm A., Goncalves R., Rosa A., Moisan J.P., Chaventre A., Ferak V., Furedi S., Oefner P.J., Shen P., Beckman L., Mikerezi I., Terzic R., Primorac D., Cambon-Thomsen A., Krumina A., Torroni A., Underhill P.A., Santachiara-Benerecetti A.S., Villems R. & Semino O. 2004. Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am. J. Hum. Genet.*, 75: 128-137.

Rosser Z.H., Zerjal T., Hurles M.E., Adojaan A., Alavantic A., Amorim A., Amos W., Armenteros M., Arroyo E., Barbujani G., Beckman G., Beckman L., Bertranpetit J., Bosch E., Bradley D.G., Brede G., Cooper G., Côrte-Real H.B.S.M., de Knijff P., Decorte R., Dubrova Y.E., Evgrafov O., Gilissen A., Glisic S., Gölge M., Hill E.W., Jeziorowska A., Kalaydjieva L., Kayser M., Kivisild T., Kravchenko S.A., Krumina A., Kučinskas V., Lavinha J., Livshits L.A., Malaspina P., Maria S., McElreavey K., Meitinger T.A., Mikelsaar A.V., Mitchell R.J., Nafa K., Nicholson J., Nørby S., Pandya A., Parik J., Patsalis P.C., Pereira L., Peterlin B., Pielberg G., Prata M.J., Prevederé C., Roewer L., Rootsi S., Rubinsztein D.C., Saillard J., Santos F.R., Stefanescu G., Sykes B.C., Tolun A., Villems R., Tyler-Smith C. & Jobling M.A. 2000. Y-Chromosomal Diversity in Europe Is Clinal and Influenced Primarily by Geography, Rather than by Language. *Am. J. Hum. Genet.*, 67: 1526-15437.

Scozzari R., Cruciani F., Pangrazio A., Santolamazza P., Vona G., Moral P., Latini V., Varesi L., Memmi M.M., Romano V., De Leo G., Gennarelli M., Jaruzelska J., Villems R., Parik J., Macaulay V. & Torroni A. 2001. Human Y-chromosome variation in the western Mediterranean area: implications for the peopling of the region. *Hum. Immunol.*, 62: 871-884.

Semino O., Magri C., Benuzzi G., Lin A.A., Al-Zahery N., Battaglia V., Maccioni L., Triantaphyllidis C., Shen P., Oefner P.J., Zhivotovsky L.A., King R., Torroni A., Cavalli-Sforza L.L., Underhill P.A. & Santachiara-Benerecetti A.S. 2004. Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. *Am. J. Hum. Genet.*, 74: 1023-1034.

Semino O., Passarino G., Oefner P.J., Lin A.A., Arbuzova S., Beckman L.E., De Benedictis G., Francalacci P., Kouvatsi A., Limborska S., Marcikiae M., Mika A., Mika B., Primorac D., Santachiara-Benerecetti A.S., Cavalli-Sforza L.L. & Underhill P.A. 2000. The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science*, 290: 1155-1159.

Semino O., Santachiara-Benerecetti A.S., Falaschi F., Cavalli-Sforza L.L. & Underhill P.A. 2002. Ethiopians and Khoisan share the deepest clades of the human Y-chromosome phylogeny. *Am. J. Hum. Genet.*, 70: 265-268.

Sengupta S., Zhivotovsky L.A., King R., Mehdi S.Q., Edmonds C.A., Chow C.E., Lin A.A., Mitra M., Sil S.K., Ramesh A., Usha Rani M.V., Thakur C.M., Cavalli-Sforza L.L., Majumder P.P. & Underhill P.A. 2006.

Polarity and temporality of high-resolution y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. *Am. J. Hum. Genet.*, 78: 202-221.

Shen P., Lavi T., Kivisild T., Chou V., Sengun D., Gefel D., Shpirer I., Woolf E., Hillel J., Feldman M.W. & Oefner P.J. 2004. Reconstruction of patrilineages and matrilineages of Samaritans and other Israeli populations from Y-chromosome and mitochondrial DNA sequence variation. *Hum. Mutat.*, 24: 248-260.

Tambets K., Rootsi S., Kivisild T., Help H., Serk P., Loogvali E.L., Tolk H.V., Reidla M., Metspalu E., Pliss L., Balanovsky O., Pshenichnov A., Balanovska E., Gubina M., Zhadanov S., Osipova L., Damba L., Voevoda M., Kutuev I., Bermisheva M., Khusnutdinova E., Gusar V., Grechanina E., Parik J., Pennarun E., Richard C., Chaventre A., Moisan J.P., Barac L., Pericic M., Rudan P., Terzic R., Mikerezi I., Krumina A., Baumanis V., Koziel S., Rickards O., De Stefano G.F., Anagnou N., Pappa K.I., Michalodimitrakis E., Ferak V., Furedi S., Komel R., Beckman L. & Villems R. 2004. The western and eastern roots of the Saami--the story of genetic “outliers” told by mitochondrial DNA and Y chromosomes. *Am. J. Hum. Genet.*, 74: 661-682.

Underhill P.A., Jin L., Lin A.A., Mehdi S.Q., Jenkins T., Vollrath D., Davis R.W., Cavalli-Sforza L.L. & Oefner P.J. 1997. Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res.*, 7: 996-1005.

Underhill P.A., Passarino G., Lin A.A., Shen P., Mirazon L.M., Foley R.A., Oefner P.J. & Cavalli-Sforza L.L. 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann. Hum. Genet.*, 65: 43-62.

Underhill P.A., Shen P., Lin A.A., Jin L., Passarino G., Yang W.H., Kauffman E., Bonne-Tamir B., Bertranpetit J., Francalacci P., Ibrahim M., Jenkins T., Kidd J.R., Mehdi S.Q., Seielstad M.T., Wells R.S., Piazza A., Davis R.W., Feldman M.W., Cavalli-Sforza L.L. & Oefner P.J. 2000. Y chromosome sequence variation and the history of human populations. *Nat. Genet.*, 26: 358-361.

Vernesi C., Caramelli D., Dupanloup I., Bertorelle G., Lari M., Cappellini E., Moggi-Cecchi J., Chiarelli B., Castri L., Casoli A., Mallegni F., Lalueza-Fox C. & Barbujani G. 2004. The Etruscans: a population-genetic study. *Am. J. Hum. Genet.*, 74: 694-704.

Weale M.E., Shah T., Jones A.L., Greenhalgh J., Wilson J.F., Nymadawa P., Zeitlin D., Connell B.A., Bradman N. & Thomas M.G. 2003. Rare deep-rooting Y chromosome lineages in humans: lessons for phylogeography. *Genetics*, 165:229-234.

Weale M.E., Weiss D.A., Jager R.F., Bradman N. & Thomas M.G. 2002. Y chromosome evidence for Anglo-Saxon mass migration. *Mol. Biol. Evol.*, 19: 1008-1021.

Wells R.S., Yuldasheva N., Ruzibakiev R., Underhill P.A., Evseeva I., Blue-Smith J., Jin L., Su B., Pitchappan R., Shanmugalakshmi S., Balakrishnan K., Read M., Pearson N.M., Zerjal T., Webster M.T., Zholoshvili I., Jamarjashvili E., Gambarov S., Nikbin B., Dostiev A., Aknazarov O., Zalloua P., Tsoy I., Kitaev M., Mirrakhimov M., Chariev A. & Bodmer W.F. 2001. The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc. Natl. Acad. Sci. U.S.A*, 98: 10244-10249.

Y Chromosome Consortium 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res.*, 12: 339-348.

Zhivotovsky L.A., Underhill P.A., Cinnioğlu C., Kayser M., Morar B., Kivisild T., Scozzari R., Cruciani F., Destro-Bisol G., Spedini G., Chambers G.K., Herrera R.J., Yong K.K., Gresham D., Tournev I., Feldman M.W. & Kalaydjieva L. 2004. The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *Am. J. Hum. Genet.*, 74: 50-61.

Editor, Giovanni Destro-Bisol

Appendix 1 - Y-chromosome haplogroup frequencies in European populations. The haplogroups are defined in Figure 1. More detailed information regarding haplogroup distribution is available online in the supplementary materials.

			M13	SRY1532	RPS4Y	YAP	M35	M89	M201
Population	**ref**	**N**	**A**	**BR**	**C**	**DE**	**E3b1**	**F**	**G**
Anatolia - Caucasus									
Abazinia	24	14			7			29	29
Aegean Turkey	8	30	6,7				6,7		20
Alagir	23	24							75
Anatolia C	8	90			1,1		8,9		12,2
Ardon	23	28						4	21
Armenia	24	10				6		18	11
Azerbaijan	24	72				6		11	18
Black Sea C	8	29			3,4		3,4		10,3
Black Sea E	8	83			1,2		8,4		15,7
Chechenia	24	19			5			32	5
Darginia	24	26				4		27	4
Digora	23	31						3	74
Georgia	24	77				3		14	31
Georgia	2	63						3,2	30,1
Ingushia	24	22						27	27
Istambul	8	81			4,9	3,7	12,3		7,4
Kabardinia	24	59						24	29
Lezgi (Dagestan)	24	25						32	36
Oseetia (Digora)	24	31						3	74
Ossetia (Ardon)	24	28						4	21
Ossetia S.	23	17				18		41	
Rutulia	24	24						58	38
Trace	8	52					15,4		3,8
Turkey NE	8	82					14,6		8,5
Turkey S	8	33					12,1		9,1
Turkey SE	8	43					9,3		14
Turkey	2	30				3,3	13,3	3,3	6,6
Zamankul	23	23						8,7	60,9
Zil'ga	23	23						4,3	56,5
Balkan peninsula									
Albania	4	30					23,3		3,3
Albania	2	51				2	21,6	4	2
Aromuns (Andon Poci – Albania)	4	19					15,8		
Aromuns (Dukasi - Albania)	4	39					17,9		10,3

M82	M170	M26	12f2.1	M172	M9	M11	Tat	92R7	M242	M173	M17	M124
H	I	I1b1b	J	J2	K	L	N	P	Q	R1	R1a1	R2
				7	14						14	
	6,7		3,3	13,3	3,3	3,3	10		3,3	20	3,3	
				4,2	8,3					4,2		8,3
1,1	4,4		5,6	28,9	3,3	3,3	5,6		3,3	15,6	5,6	1,1
	32			29	7			4			4	
	5			24	7			4		19	6	
	3			31	11					11	7	
	10,3		6,9	27,6			3,4			31		3,4
	2,4		9,6	19,3	2,4	12	7,2		1,2	14,5	4,8	
				26	11			16			5	
	58			4						4		
	13			3				6				
	4			21	3			4		10	10	
				33,3	3,2	1,6				14,3	7,9	6,3
	5			32	5			5				
	9,9		7,4	19,8	3,7	4,9	2,5			14,8	8,6	
	10			12	15			7		2	2	
					28					4		
	13			3				6				
	32			29	7			4			4	
				24						12	6	
				4								
1,9	11,5		15,4	30,8		1,9				13,5	3,8	1,9
1,2			12,2	26,8	1,2	3,6	1,2		6,1	13,4	11	
	3		15,2	24,2	3		3			21,2	9,1	
	4,7		7	25,6	4,7		2,3			16,3	11,6	4,7
	3,3			40	6,6	3,3	3,3	3,3		6,6	6,6	
				8,7	21,7							
				26,1	13							
6,7	16,7		3,3	16,7	3,3					26,6		
	19,6			23,5						17,6	9,8	
	42,1			5,3						36,8		
	17,9		2,6	46,2						5,2		

Appendix 1 (continued)

			M13	SRY1532	RPS4Y	YAP	M35	M89	M201
Population	**ref**	**N**	**A**	**BR**	**C**	**DE**	**E3b1**	**F**	**G**
Aromuns (Krusevo - Macedonia)	4	43					20,9		7
Aromuns (Romania)	4	42			4,8		7,1		
Aromuns (Stip – Macedonia)	4	65					18,5		
Bosnia	18	69					10,1		4,4
Bosnia	19	85					12,9	3,5	3,5
Braè	22	47					4,3	2,1	6,4
Croatia Mainland	22	108					5,6	1,9	0,9
Croatia	2	58					6,9		1,7
Croatia	18	108					5,6	1,9	0,9
Croatia	19	90					8,9	1,1	1,1
Cyprus	6	65	3,1				20	9,2	
Greece	2	76			1,3	1,3	22,4	1,3	2,6
Greece	4	41					19,5		4,9
Herzegovina	18	141					8,5		
Hvar	22	91					4,4	1,1	1,1
Korcula	22	132					3,8	1,5	10,6
Kosovo	18	114					47,4		
Krk	22	73					6,8		
Macedonia	4	52					23,1		3,8
Macedonia	18	79					24	1,3	5,1
Macedonia	2	20					15	5	
Rom (Macedonia)	18	57					29,8		
Romania (Constanza)	4	31				3,2	9,7		12,9
Romania (Ploiesti)	4	36					16,7		8,3
Serbia	18	113					22,3		
Serbia	19	81					22,3	4,9	1,2
British Isles									
Castlerea	5	43							
Channel Islands	5	128					4	2	
Chippenham	5	51						2	
Cornall	5	52							
Dorchester	5	73					4	1	
Durness	5	51							
Estern Isles	5	88							
Faversham	5	55					4		
Haverfordest	5	59					3		
Ireland	26	796					0,9	0,3	

M82	M170	M26	12f2.1	M172	M9	M11	Tat	92R7	M242	M173	M17	M124
H	**I**	**I1b1b**	**J**	**J2**	**K**	**L**	**N**	**P**	**Q**	**R1**	**R1a1**	**R2**
	20,9			11,6						39,5		
	19			33,3	9,5					26,2		
	16,9			20						44,6		
	53,7									26		
	48,2		2,4	9,5	1,2					3,5	15,3	
	53,2				2,1					34,1		
	37			1,9	1,9			1,9		50		
	44,8			5,2	1,7					10,3	29,3	
	37			1,9	0,9			1,9		50		
	73,3			1,1						2,2	12,2	
	7,7		6,2	36,9	4,6					9,2	3,1	
	7,9			21	1,3	1,3				27,6	11,8	
	19,5			19,5	2,4					34,2		
	70,9			1,4	2,8				0,7	15,6		
	65,9			3,3	1,1			14,3		15,2		
	53			2,3	1,5			6,1		21,6		
	8			16,7				1,8		25,5		
	27,3			10,9						54,8		
	28,8			11,5	3,8					27	1,9	
	34,2			12,6	1,3					20,3		
	20			15						10	35	
59,6	5,3			1,8						3,5		
	41,9			6,5						25,8		
2,8	38,9		2,8	16,7						13,9		
0,9	36,3			8	7,1					26,6		
	35,8			8,7	7,4					6,2	13,6	
	7	2								90		
	21	3	1	1						66	3	
	20	2		4						65	8	
	12			2						79	8	
	14		1	3						73	4	
	14									80	6	
	25									66	9	
	11			5						76	4	
	2	2								91	2	
	10,9	1,1			0,4		0,1	0,5		85,9		

Appendix 1 (continued)

Population	**ref**	**N**	**M13** **A**	**SRY1532** **BR**	**RPS4Y** **C**	**YAP** **DE**	**M35** **E3b1**	**M89** **F**	**M201** **G**
Isle of Man	5	62					2		
Llangefni	5	80					4		
Llanidloes	5	57					5	4	
Midhurst	5	80					1	1	
Morpeth	5	95						2	
Norfolk	5	121					3		
Oban	5	42						2	
Orkney	5	121							
Penrith	5	90					3	1	
Pitlochry	5	41							
Rush	5	76							
Shetland	5	63							
Southell	5	70					6		
Stonehaven	5	44						2	
Uttoxeter	5	84					4	1	
York	5	46					4	2	
Central Europe									
Dutch	2	27					3,7		
French	2	23					8,7	4,3	
German	2	16					6,2		
Germany/Denmark	5	190					3	2	
Eastern Europe									
Czech and Slovakia	2	45					2,2		4,4
Hungary	2	45					8,9		2,2
Poland	2	55					3,6		
Ukraine	2	50					4		4
Fennoscandia									
Blekinge/Kristianstad	12	41						2,4	4,9
Finland S	11	50				4			
Finnish Saami	21	69							
Gotland	12	40					2,5		
Häme	11	49							
Iceland	28	181							
Karelia N	11	22							
Karelia S	11	48							

M82 H	M170 I	M26 I1b1b	12f2.1 J	M172 J2	M9 K	M11 L	Tat N	92R7 P	M242 Q	M173 R1	M17 R1a1	M124 R2
	16									70	13	
	4			1	1					89	1	
	19			2						66	4	
	15	3	1	4						74	1	
	18		1	3						73	3	
	31			2						60	4	
	7									86	4	
	14	1						2		64	19	
	18			2						68	8	
	1			7						80	2	
	8	3								86	4	
	10									66	23	
	18			6						64	5	
	13									79	5	
	18			4						71	2	
	32									57	4	
	22,2									70,4	3,7	
	17,4			13				4,3		52,2		
	37,5									50	6,2	
	39			3			2			39	12	
	15,6			8,9	4,4		2,2			35,6	26,7	
	11,1			2,2		2,2				13,3	60	
	23,6									16,4	56,4	
2	18			6	2		6	2		2	54	
	46,3				2,4		2,4	2,4		22	17,1	
	30						60			4	2	
	40,6						55,1			4,3		
	50				5		10	5		15	12,5	
	36,7						55,1			6,1	2	
	34,2						0,6			41,4	23,8	
	18,2						68,2		4,6	4,6	4,6	
	18,7						70,8	2,1			8,3	

Appendix 1 (continued)

Population	ref	N	M13 A	SRY1532 BR	RPS4Y C	YAP DE	M35 E3b1	M89 F	M201 G
Kola Saami	21	23				8,7			
Mari	2	46						6,5	
Norway	5	201						0,5	
Osterbotten	12	40							5
Ostergotland/Jonkoping	12	41					2,4		
Ostrobothnia N	11	129		0,8					
Ostrobothnia Swedish	11	58		1,7					
Ostrobothnia Swedish	11	25		4					
Saami	2	24							
Satakunta	11	48		2,1					
Savo N	11	107							
Skaraborg	12	45					2,2		
Swedish Saami	12	38							
Swedish Saami	21	35						5,7	
Udmurt	2	43					4,7	4,7	
Uppsala	12	55							
Varmland	12	42							2,4
Vasterbotten	12	32					3,1	3,1	6,3
Iberian peninsula									
Açores	15	121				0,8	9,1		8,3
Alentejo	10	65				1,5	18,3		7,7
Algarve	10	21							4,8
Andalusia	2	29					10,3	3,4	
Asturie	14	90				11,1	2,2	11,1	
Basque	9	45					2,2		
Basque Fr	2	22							
Basque Sp	2	45					2,2	2,2	
Basque	3	168				0,6	3,6	1,8	0,6
Basque	5	42					2	2	
Beira Interior	10	58					12,3	1,8	8,8
Beira Litoral	10	116					11,3		2,6
Cadiz	9	28					3,6		
Cantabria	9	70				4,3	8,6		8,6
Castile	9	21					4,8		
Catalonia	9	16						6,2	6,2
Catalonia	2	24					4,2		8,3
Cordoba	9	27					11,1		

M82	M170	M26	12f2.1	M172	M9	M11	Tat	92R7	M242	M173	M17	M124
H	I	I1b1b	J	J2	K	L	N	P	Q	R1	R1a1	R2
	17,4		4,3				39,1			30,4		
	4,3				6,5		69,5				13	
	28			0,5			1	4		30	34	
	20						65			2,5	7,5	
	31,7		9,8				7,3	2,4		36,6	9,8	
	23,3						65,1			4,7	6,2	
	46,6						29,3			3,5	19	
	36						40			8	12	
	41,7						41,7			8,3	8,3	
	54,1				2,1		27,1			6,3	8,3	
	17						78,5			0,9	4,7	
	37,8		4,4				2,2	6,7		37,8	8,9	
	31,6						44,7			7,9	15,8	
	31,4						37,1			25,7		
	7				4,6		30,2			11,6	37,2	
	38,2		5,5		1,8		14,5	1,8		27,3	10,9	
	50		4,8				9,5	2,4		16,7	14,3	
	25				3,1		25	3,1		18,8	12,5	
	6,6		2,5	10,7	1,7					60,6		
	7,7	3,1	6,2	9,2	1,5					44,6		
	4,8		4,8	9,5						76,2		
	3,4			6,9	6,9	3,4				65,5		
			3,3	6,7	1,1					64,4		
	4,4	4,4								89		
		9,1		4,5						86,4		
	2,2	4,4								88,9		
	1,8	1,2	1,2	0,6	0,6					88,1		
	2	5								89		
	3,5		1,8	3,5	1,8					67,6		
	3,4	0,9	6	6	4,3	0,9				64,6		
	3,6	10,7	3,6	14,3	10,7					53,5		
	4,3	1,4	2,9	2,9	4,3					62,7		
	14,3	19		9,5						52,4		
	6,2		6,2							75,1		
	4,2			4,2						79,2		
	11	3,7		14,7	3,7					55,8		

Appendix 1 (continued)

			M13	SRY1532	RPS4Y	YAP	M35	M89	M201
Population	**ref**	**N**	**A**	**BR**	**C**	**DE**	**E3b1**	**F**	**G**
El Hierro	9	47				4,3	10,6		4,3
Entre Douro	10	228				0,8	11,4		6,1
Extremadura	10	43					16,2		7
Fuerteventura	27	75				2,7	16		2,7
Galicia	9	19				5,3	26,3		
Gran Canaria	27	78				1,3	18		3,8
Huelva	9	22				4,5	9		
Iberia	3	662		0,2		1,8	10,6	0,3	4,5
La Gomera	27	92					10,8		5,4
La Palma	27	85				2,4	11,8		2,4
Lanzarote	27	97				2	9,3		5,2
Lebaniegos	13	37				5			11
Leon	9	60					10	1,7	6,7
Lisboa	10	62				1,6	12,9		4,8
Madeira	15	129	0,8				14		3,2
Malaga	9	26				7,6	23		3,8
Other Cantabrians	13	36				3	17		5
Pasiegos	13	45					18	2	7
Portugal C	15	102	1			1	14,7		7,8
Portugal N	15	101			1	3	20,8	1	5
Portugal N	9	109		0,9		0,9	14,6	0,9	7,3
Portugal S	15	10	1		1	1	8	1	7
Seville	9	155				0,6	6,4		4
Tenerife	27	178				0,6	18		3,9
Tra Os Montes	10	64					7,8		3,1
Valencia	9	31				3,2	13		
Italian Peninsula									
Alto Adige	7	34					6		3
Apennines C-N	7	27					14		15
Apulia N	7	46					4		13
Apulia S	7	71					23	4	11
Calabria	7	57					16		11
Calabria	2	37				2,7	13,5	10,8	8
Campania	7	84				1	18		10
Corsica	16	34					11,8	2,9	11,8
Elba Island	7	95					7		11
Italy C-N	2	50					2		10

M82 H	M170 I	M26 I1b1b	12f2.1 J	M172 J2	M9 K	M11 L	Tat N	92R7 P	M242 Q	M173 R1	M17 R1a1	M124 R2
	2,1		6,4	12,8						59,5		
	7,5	0,9	3,5	8,3	2,6	0,4				58,9		
	9,3	4,7		4,7						58,1		
	2,7	2,7	4	14,7	1,3					53,2		
			5,3							63,1		
	5,1	1,3	3,8	5,1	5,1					56,5		
	9,2			13,7	4,5					59,1		
	6	3	2,6	8,8	4,1		0,2	0,5		57,6		
	12	8,7	7,6	10,9	1,1					43,5		
	4,7	4,7	7	9,4	3,5		1,2			52,9		
	7,2	6,2	5,2	10,3	1					53,6		
	5		8	3	8					60		
	1,7	1,7	1,7	5	3,3					68,2		
	1,6		1,6	6,5						70,9		
	7		0,8	10,9	3,2	1,6	2,3			56,6		
				15,4						50,2		
	3	3		3						66		
	2			4						67		
	4,9		4,9	2,9						62,8		
	6,9		1	6,9						54,5		
	3,7	1,8	3,7	8,3	1,8			2,8		53,3		
	7		7	11					1	55		
	9,7	2,6	3,2	7,8	4,5		0,6			60,6		
	5,6	1,7	2,2	6,2	6,2					55,6		
	9,4	4,7		6,3	4,7					64,1		
	9,7	3,2		9,7	3,2					58		
	6			9	9					68		
	4			22	4					33	7	
			2	17	4					52	7	
	3		1	24	6					27	1	
	2		4	35						32	2	
				21,6	2,7	5,4		2,7		32,4		
	8	2	5	17	8					29	2	
	8,8		2,9	2,9	5,9					50		2,9
	16	4		8						53	1	
	8			14						62	4	

Appendix 1 (continued)

Population	ref	N	M13 A	SRY1532 BR	RPS4Y C	YAP DE	M35 E3b1	M89 F	M201 G
Italy S	6	68					26,5	14,7	
Latium NE	7	55					26		13
Latium S	7	51					6		12
Marche C	7	59					5		7
Sardinia	17	116	0,9		0,9		9,5	6,9	14,7
Sardinia	25	930				1	8	1	13
Sardinia C	25	103					7,8	1	11,7
Sardinia N	25	86					12,9	3,5	20,9
Sardinia S	25	187	0,5			0,5	10,6	1,1	13,9
Sardinia	2	77	1,3		1,3		10,4	5,2	14,2
Sardinia	6	81	1,2				9,9	18,5	
Sicily	16	51					2	3,9	11,8
Sicily E	6	87					28,7	4,6	
Sicily N	6	70			1,4		21,4	11,4	
Sicily S	6	55					16,4	14,5	
Tuscany C	7	41					9		7
Tuscany S	7	79					12		15
North Africa – The Levant									
Algeria	6	19					5,3	10,5	
Ashkenazi Jews	1	442				0,7	19,7	0,9	9,7
Ashkenazi Jews	20	20					20		5
Berber	3	75		1,3		2,7	76		2,7
Druze	20	20					15		
Lebanon	2	31			3,2		25,8	16,1	3,2
Lebanon	20	20					30		10
Lebanon (christians)	6	43					16,3	18,6	
Lebanon (muslims)	6	39					17,9	12,8	
Malta	6	90				2,2	8,9	6,7	
Morocco	20	19					21,1		26,3
Palestine	20	20		5			15		75
Samaritan Jews	20	12					16,7		
Syria	2	20				10	10	30	
Tunis	6	39					51,3	5,1	

M82	M170	M26	12f2.1	M172	M9	M11	Tat	92R7	M242	M173	M17	M124
H	I	I1b1b	J	J2	K	L	N	P	Q	R1	R1a1	R2
	4,4	1,5	4,4	16,2	2,9			1,5		25	2,9	
	5			15	2					38	2	
	6	2		25	8					37	4	
	8		5	36						37	2	
	3,4	35,3		5,2	0,9					21,5		0,9
	2	37	2	10	2					21	1	
		36,9	2,9	5,8	1,9					291	2,9	
		27,9		8,2	7					19,8		
	3,2	31,6	2,7	11,8	1,1					22,4	0,5	
	2,6	35,1		5,2	1,3					22,1		1,3
	4,9	25,9	4,9	9,9	1,2			2,5		21		
	3,9		11,8	31,4	3,9					27,5	3,9	
	3,4	1,1	6,9	28,7	4,6					19,5	2,3	
	15,7		7,1	11,4	1,4			1,4		25,7	2,9	
	5,5			27,3	5,5					29,1	1,8	
	5		7	17	2					46	5	
	5			19	4					41	5	
			31,6	52,6								
	4,1		19	19	2	0,2	0,2	0,5	5,2	11,4	7,5	
			25	15					5	10	20	
				6,7	4					6,7		
			5	25	5	35				10	5	
	3,2			29		3,2				6,4	9,7	
			10	40						5	5	
	4,7		9,3	34,9	4,7			7		4,7		
			30,8	25,6	5,1					5,1	2,6	
	12,2		7,8	21,1	4,4			1,1		32,2	3,3	
			10,5	21	10,5					10,5		
					5							
			66,7	16,7								
5	5			15						15	10	
	2,6		23,1	10,3						5,1	2,6	

Data from: 1) Behar *et al.*, 2003; 2) Semino *et al.*, 2000; 3) Alonso *et al.*, 2005; 4) Bosch *et al.*, 2006; 5) Capelli *et al.*, 2003; 6) Capelli *et al.*, 2006; 7) Capelli *et al.*, 2007; 8) Cinnioglu *et al.*, 2004; 9) Flores *et al.*, 2004; 10) Beleza *et al.*, 2006; 11) Lappalainen *et al.*, 2006; 12) Karlsson *et al.*, 2006; 13) Maca-Meyer *et al.*, 2003; 14) Scozzari *et al.*, 2001; 15) Goncalves *et al.*, 2005; 16) Francalacci *et al.,* 2003; 17) Passarino *et al.*, 2001; 18) Pericic *et al.*, 2005b; 19) Marjanovic *et al.*, 2005; 20) Shen *et al.*, 2004; 21) Tambets *et al.*, 2004; 22) Barac *et al.*, 2003; 23) Nasidze *et al.*, 2004b; 24) Nasidze *et al.*, 2004a; 25) Contu *et al.*, 2008; 26) Moore *et al.*, 2006; 27) Flores *et al.*, 2003; 28) Helgason *et al.*, 2000.