Phylogeography of Y chromosomal haplogroups as reporters of Neolithic and post-Neolithic population processes in the Mediterranean area

Fiorenza Pompei[1], Fulvio Cruciani[2], Rosaria Scozzari[2], Andrea Novelletto[1]

1 Department of Biology, Università "Tor Vergata", Rome, Italy
novelletto@bio.uniroma2.it
2 Department of Genetics and Molecular Biology, Sapienza Università di Roma, Rome, Italy

ABSTRACT – *The phylogeny of the human Y chromosome as defined by unique event polymorphisms is being worked out in fine detail. The emerging picture of the geographic distribution of different branches of the evolutionary tree (haplogroups), and the possibility of genetically dating their antiquity, are important tools in the reconstruction of major peopling, population resettlement and demographic expansion events. In the last 10 000 years many such events took place, but they are so close together in time that the populations that experienced them carry Y chromosomal types which can hardly be distinguished genetically. Nevertheless, under some circumstances, one can detect departures from the model of a major dispersal of people over much of the territory, as classically claimed for the European Neolithic. The results of three studies of haplogroups relevant for Southern European populations are discussed. These analyses seem to resolve the signal of recent post-Neolithic events from the noise of the main East-to-West Palaeolithic/early Neolithic migrations. They also confirm that, provided an appropriate level of resolution is used, patterns of diversity among chromosomes which originated outside Europe may often be recognized as the result of discontinuous processes which occurred within Europe.*

IZVLEČEK – *Filogenija človeškega kromosoma Y, kot jo lahko preberemo skozi zaporedje polimorfizmov, je dobro poznana. Čedalje jasnejša slika geografskih distribucij posameznih vej evolucijskega drevesa (haploskupin) in možnosti njihovega genetskega datiranja so pomembna orodja pri preučevanju širjenja človeštva, premikov in širitev populacij. V zadnjih 10 000 letih se je zgodilo kar nekaj takih dogodkov, ki pa so si časovno tako blizu, da populacije, ki so bile vanje vpletene, nosijo tako zelo podobne Y kromosome, da jih genetsko le težko razločimo med seboj. Kljub temu je moč pod nekaterimi pogoji opaziti razlike, ki se ločijo od klasičnega modela širjenja populacij, ki velja za evropski neolitik. Predstavljamo rezultate treh študij haploskupin južnoevropskih populacij. Analize so pokazale, da je moč iz šuma glavnih paleolitskih in neolitskih migracij iz vzhoda proti zahodu razločiti nekatere po-neolitske demografske dogodke. Študija tudi potrjuje, da je mogoče – ob dovolj visoki ločljivosti – nekatere vzorce kromosomov, ki izvirajo izven Evrope, pripisati seriji prekinjenih procesov znotraj Evrope.*

KEY WORDS – *Y chromosome; Neolithic; peopling of Europe; population genetics; demographic expansions*

Introduction

The genetic characterization of human populations has long been recognized as an important and often indispensable complement to historical research for the understanding of population stratification, the reconstruction of migrations and the evaluation of gene flow. A major leap forward in this field was re-presented by the possibility of assembling and analysing genetic data into a phylogenetic perspective. Here we are concerned with the application of this approach to population processes that occurred in the Neolithic and post-Neolithic, as inferred from the current population distribution of genetic diversity

of the male-specific portion of the human Y chromosome (MSY).

The phylogenetic approach takes into account the sequential accumulation of mutations in a given stretch of DNA (in this case the MSY) over time. A mutation in a given DNA position produces a so-called derived allele at that position. Whenever this event can be considered unique, and subjects carrying the derived allele coexist in the population with subjects carrying the non-mutated (ancestral) allele, a so-called Unique Event Polymorphism (UEP) can be observed (also called biallelic polymorphisms, as typically only two alleles are observed at a given position). In this situation, each derived allele becomes a genetic marker whose origin can be located in a time when the 'parental' type already existed and can, in turn, be considered 'parental' for other mutations that appeared later. Graphs that summarize the overall process are called phylogenetic trees, and they display branches that diverge progressively, each new branch being defined by a new derived allele in any position along the MSY.

A direct extension of these concepts is that all MSY copies (each carried by a different subject) bearing the same derived allelic variant at a given position can be considered, as a first approximation, descendants of the first one in which that particular mutational event occurred (*i.e.* have a monophyletic origin). When considering more than one position on the same DNA molecule, the particular combination of allelic variants (the haplotype) thus represents a record of all the mutational events that occurred on the lineage leading to that haplotype. Alleles shared by two haplotypes testify to their common ancestry, whereas alleles which differentiate two haplotypes show that they belong to lineages that diverged some time in the past and, since then, have accumulated a different series of mutations.

The principles and methods of phylogenetic reconstruction from experimental data can be found in basic books. A general consensus has been reached on the nomenclature of lineages of the human MSY, with alternating letters and

numbers from the deepest to the terminal branches (*Y Chromosome Consortium 2002*) (Fig. 1). Each lineage defined by biallelic markers is referred to as a haplogroup, whereas the term haplotype has been restricted to a combination of alleles at Short Tandem Repeats (STRs, see below).

After a pioneering era, the search for biallelic markers exploited high-throughput methods that were first applied to samples representative of the entire world population and, later, oriented to resolve in finer detail some specific lineages.

Another important class of markers is represented by STRs. These include loci with different lengths of the basic repeat, and extensive searches for developing them as markers have been performed (*Kayser et al. 2004*). Mutation at these loci occurs by the addition/subtraction of a number of repeats that is one in the majority of cases. This latter feature fits the theoretical 'Stepwise Mutational Model', which allows us to create expectations for the rate of accumulation of diversity and the distribution of allele sizes. What matters here is that the overall amount of STR diversity observed among the carriers of a specific lineage defined by biallelic markers is a function of the time elapsed since the origin of that lineage, and this property is exploited to arrive at an evaluation of the antiquity of that lineage purely on genetics grounds.

The genetic concepts and tools described above have been used to search for the genetic signatures that
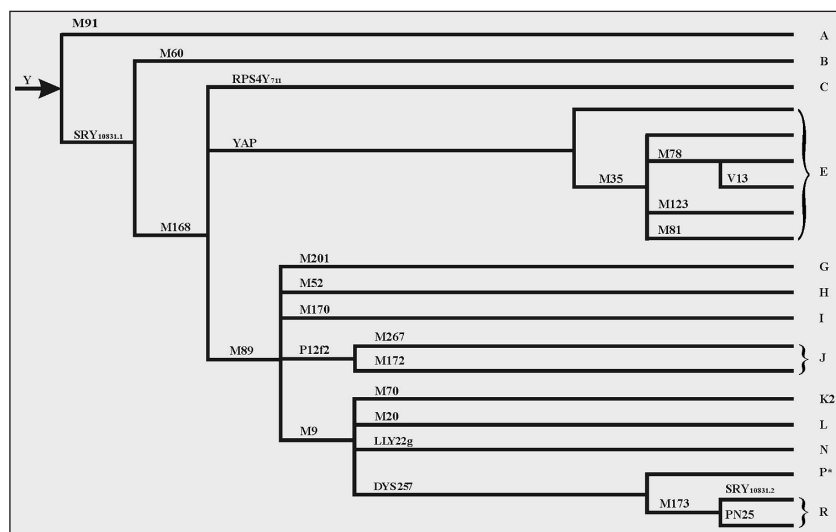

Fig. 1. Schematic representation of the human MSY phylogenetic tree. Only the main branches found in Europe are shown. Mutations that identify each branch are reported above the corresponding line. Letters used in the unified nomenclature for the main haplogroups are shown on the right. The positions of the nodes are not proportional to age estimates.

the Neolithic revolution has left in the male gene pool of populations of the Mediterranean region and other areas nearby. However, it has to be emphasized that events that occurred in the last tenth of millennia or later may have left traces that could only modify the pre-existing repertoire of genetic markers and their particular geographic distributions. These were the result of processes occurring over a much longer time preceding the Neolithic. In fact, even in the current description of the MSY phylogenetic tree, most of the markers are older than 10–15 ky BP, *i.e.* they were already present in the populations that experienced the demographic changes associated with the Neolithic revolution. In conclusion, the question for the geneticist is whether a DNA polymorphism which is able to mark a specific episode indeed exists and is known. In the phylogenetic framework, only under some circumstances one can safely assume that a particular pattern of genetic variation within a single or a group of populations can be the result of a Neolithic or post-Neolithic event. These are:

❶ a biallelic marker near to the tip of a branch of the MSY tree is dated at a time compatible or younger than the Neolithic or,

❷ no such marker is known but, within an older lineage, a subset of populations display a limited amount of STR variation, as if they had been founded at a more recent time and by a reduced number of founders.

We review and discuss here three studies (*Di Giacomo et al. 2004; Cruciani et al. 2007; Luca et al. 2007*) that found genetic evidence of demographic events which occurred after the spread of the Neolithic culture from the Levant and involved Central and South-Eastern Europe.

Post-Neolithic expansion from the Aegean detected by haplogroup J2f1–M92

Haplogroup J has been considered to represent a signature of Neolithic demic diffusion associated with the spread of agriculture (*Semino et al. 1996*). Di Giacomo et al. (*2004*) provided population data which give insights into the ways in which this haplogroup spread.

Phylogenesis. Haplogroup J can be subdivided into two major clades J1 and J2 – characterized by the markers M267 and M172, respectively – plus the rare paragroup J*(xJ1,J2). Within J2, the analysis of a multi-repeat deletion in the dinucleotide STR locus DYS413 (*Malaspina et al. 1998*) resolves a major multifurcation of six independent lineages, recently increased to 11 (*Sengupta et al. 2006*). This additional mutational step within J2 enhances the possibility of performing phylogeographic studies of the entire J2 sub-haplogroup in the Mediterranean area (Fig. 2).

Population Data. Data on the overall occurrence of the entire J haplogroup display an area of high frequencies (>20%) stretching from the Middle East to the central Mediterranean. A review of the frequency data on Europe, the Caucasus, Iran, Iraq and North Africa reveals that, in the Mediterranean, this haplogroup is mainly confined to coastal areas. The high frequencies in Turkey, Jewish and non-Jewish Middle Eastern populations and in the Caucasus, identify the fertile crescent and the east Mediterranean as the focal area for the westward dispersal of the haplogroup. However, the data agree in showing that this haplogroup did not leave a strong signature in the peoples of the northern Balkans and central Europe, this being the most likely route under the demic diffusion model for the entry of agriculturalists into the European continent north to the Alps. Instead, the raw frequency data from within the Iberian, Italian and Balkan peninsulas are more in line with alternative routes of westward spread, possibly maritime.

Internal J diversity. The highest UEP diversity is observed in Turkey, Egypt and three locations in southern Europe. The two most derived sub-haplogroups typed (J2f1–M92 and J2e–M12) were only found in Turkey and locations west to it, boosting the UEP internal diversity. The sub-haplogroup distribution found in Turkey is similar to that reported by Cinnioglu et al. (*2004*).

The UEP diversity within J2 is lower in the Middle East compared to both Turkey and the European locations. In conclusion, the UEP diversity of J in Turkey and southern Europe does not seem to be a simple subset of that present in the area where this haplogroup first originated. This finding, also confirmed in the data by Semino et al. (*2004*), points to Turkey and the Aegean as a relevant source for the J diversity observed throughout Europe.

The contribution of STRs for dating. When combined with the results of 5 STRs, the age returned for the entire J clade and its confidence interval fell within the range reported in previous works (39.6

–10.5 ky BP). Conversely, two of the terminal branches (J2f–M67 and J2f–M92) turned out to be much younger, with estimated ages of 4 and 2.6 ky BP, respectively (C.I. 2.4–7.7 and 1.6–4.2, respectively).

Conclusion 1

The dating estimates obtained by Di Giacomo et al. (*2004*) are in agreement with the appearance of J1 and J2 in the Levant at the time of the Neolithic agriculture revolution. Implicitly, these figures make these haplogroups of little aid in identifying splits in population that may have accompanied the westward dispersal of the entire haplogroup.

The data by Di Giacomo et al. (*2004*) and Semino et al. (*2004*) show that J2f1–M92 is predominantly

found in the northern Mediterranean, from Turkey westward. In particular, the estimates for this latter sub-haplogroup are barely compatible with its presence among the early Levantine agriculturalists. Thus the most likely explanation is the emergence of J2f1 in the Aegean area, possibly during the population expansion phase also detected by Malaspina et al. (*2001*), and coincident with the expansion of the Greek world up to the European coast of the Black sea. This scenario would agree with the clustering of J2f1–M92 chromosomes in the north-west of Turkey (*Cinnioglu et al. 2004*).

In summary, this set of data is in agreement with a major discontinuity for the peopling of southern Europe. Here, haplogroup J constitutes not only the signature of a single wave-of-advance from the Levant but, to a greater extent, also of the expansion of the Greek world, with an accompanying novel quota of genetic variation produced during its demographic growth. Recently Cadenas et al. (*2007*) described similar evidence concerning haplogroup J1–M267 as a marker of the Neolithic spread from the fertile crescent to the South Arabian peninsula.
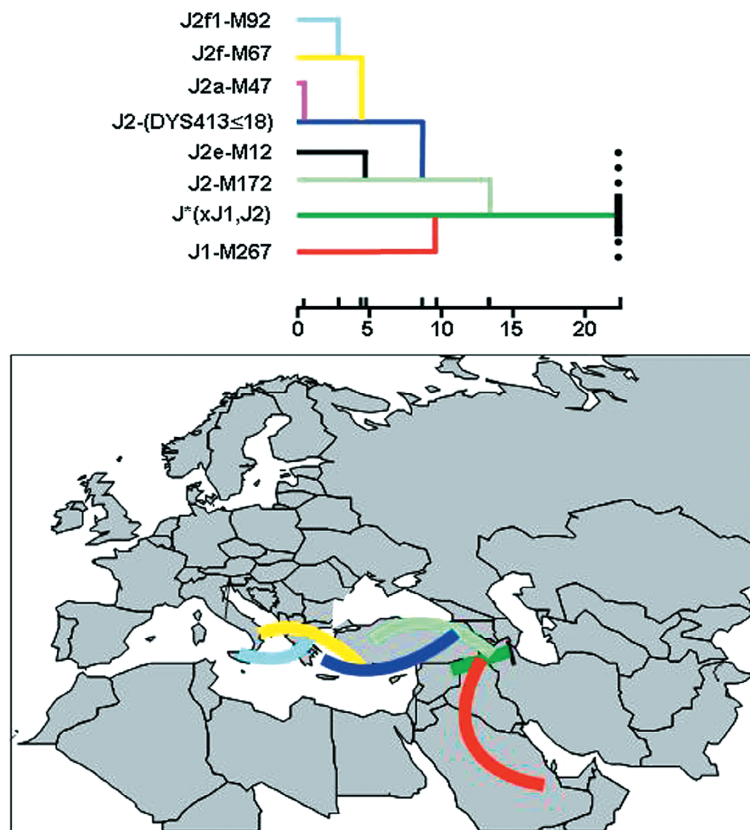
Post-Neolithic expansion from within the Balkans detected by haplogroup E–V13

Cruciani et al. (*2007*) provided detailed population data on the distribution of E–M78 binary sub-haplogroups defined by ten UEPs in 81 populations mainly from Europe, western Asia and Africa. In order to obtain estimates of the internal diversity and coalescence age of E–M78 sub-haplogroups and their associated human migrations and demographic expansions, a set of eleven microsatellites was also analyzed. The same set of microsatellites was also analyzed in a sample of Y chromosomes belonging to the haplogroup J–M12. These results not only provide a refinement of previous evolutionary hypotheses based on microsatellites alone, but also well defined time frames for different migratory events that led to the disper-

*Fig. 2. Top. Phylogenetic arrangement of lineages within haplogroup J, as analysed by Di Giacomo et al. (*2004*). Other internal lineages (*Y Chromosome Consortium 2002; Sengupta et al. 2006*) are not shown. The positions of the nodes of the tree are according to age estimates (*Di Giacomo et al. 2004*) and are marked on the lower bar (0 = present). Bottom. Same phylogenetic tree as above superimposed onto geography, to show the main routes of dispersal of the different lineages. The origin of the entire J haplogroup was arbitrarily placed in the fertile crescent and only south and westward dispersals are outlined. For simplicity, J2–M12 and J2–M47 are not shown. The endpoints of each line are schematic and do not represent exclusive directions of migration (e.g. J1–M267 is found not only in the Arabian Peninsula, but also in other areas where J is present).*

sal of these haplogroups and sub-haplogroups in the Old World.

Phylogenesis. By analyzing a worldwide sample of 6501 male subjects, 517 chromosomes belonging to haplogroup E–M78 were identified, more than twice the number found in a previous study (*Cruciani et al. 2004*). These chromosomes have been further analyzed for 10 biallelic markers. Four sub-haplogroups were either rare or absent in the global sample, while the other haplogroups/paragroups were relatively common.

Population data and dating. The subdivision of E–M78 in the six common major clades revealed a pronounced geographic structuring: haplogroup E–V65 and the paragroups E–M78* and E–V12* were observed mainly in northern Africa, haplogroup E–V13 was found at high frequencies in Europe, and haplogroup E–V32 was observed at high frequencies only in eastern Africa. The only haplogroup showing a wide geographic distribution was E–V22, relatively common in north-eastern and eastern Africa, but also found in Europe, western Asia, up to southern Asia.

The peripheral geographic distribution of the most derived sub-haplogroups with respect to north-eastern Africa, as well as the results of quantitative analysis of UEP and microsatellite diversity, are strongly suggestive of a north-eastern African origin of E–M78. The evolutionary processes that determined the wide dispersal of the E–M78 lineages from north-eastern Africa to other regions can then be addressed.

Previous studies on the Y chromosome phylogeography have revealed that central and western Asia were the main sources of Palaeolithic and Neolithic migrations contributing to the peopling of Europe (*Underhill et al. 2000; Wells et al. 2001*). The molecular dissection of E–M78 contributes to the understanding of the genetic relationships between northern Africa and Europe. Several lines of evidence suggest that E–M78 sub-haplogroups E–V12, E–V22 and E–V65 were involved in trans-Mediterranean migrations directly from Africa. These haplogroups are common in northern Africa, where they probably originated, and are observed almost exclusively in Mediterranean Europe, as opposed to central and eastern Europe. Also, among the Mediterranean populations, they are more common in Iberia and south-central Europe than in the Balkans, the natural entry-point for chromosomes coming from the Levant. Such findings are hardly compatible with the south-eastern entry of E–V12, E–V22 and E–V65 haplogroups into Europe. Upper limits for the introduction of each of these haplogroups in Europe are given by their estimated ages (18.0, 13.0 and 6.2 ky BP, respectively), while lower bounds should be close to the present time, given the lack of internal geographic structuring.

Haplogroup E–V13 is the only E–M78 lineage that reaches the highest frequencies outside Africa. In fact, it represents about 85% of European E–M78 chromosomes, with a clinal pattern of frequency distribution from the southern Balkan peninsula (19.6%) to western Europe (2.5%) (Fig. 3). The same haplogroup is also present at lower frequencies in Anatolia (3.8%), the Near East (2.0%) and the Caucasus (1.8%). In Africa, haplogroup E–V13 is rare, being observed only in northern Africa at a low frequency (0.9%). The European E–V13 microsatellite haplotypes are related to each other to form a nearly perfect, star-like network, a likely consequence of rapid demographic expansion (*Jobling et al. 2004*). The age of the European E–V13 chromosomes turns out to be 4.0–4.7 ky BP. On the other hand, when only E– V13 chromosomes from western Asia are considered, the resulting network does not show such a star-like shape, and a much earlier age of 11.5 ky BP (95% C.I. 6.8–17.0) is obtained. These results present the possibility of recognizing time windows for i) population movements from the E–M78 homeland in north-eastern Africa to Eurasia, and ii) population movements from western Asia into Europe and, later, within Europe.

The most parsimonious and plausible scenario is that E–V13 originated in western Asia about 11 ky BP, and its presence in northern Africa is the result of a more recent introgression. Under this hypothesis, E–V13 chromosomes sampled in western Asia and their coalescence estimate detect a likely Palaeolithic exit from Africa of E–M78 chromosomes devoid of the V13 mutation, which later occurred somewhere in the Near East/Anatolia. The refinement of location for the source area of such movements and associated chronologies attained by Cruciani et al. (*2007*) may be relevant to controversies on the spread of cultures (and languages) between Africa and Asia in the corresponding timeframes (*Bellwood 2004; Ehret et al. 2004*).

Two haplogroups support the same scenario. As to a western Asia-Europe connection, the data suggest that western Asians carrying E–V13 may have reached the Balkans anytime after 17.0 ky ago,

but expanded into Europe not earlier than 5.3 ky ago. Accordingly, the allele frequency peak is located in Europe, whereas the distribution of microsatellite allele variance shows a maximum in western Asia. Based on previously published data discussed above, Cruciani et al. (*2007*) observed that another haplogroup, J–M12, shows a frequency distribution within Europe similar to that observed for E–V13. In order to evaluate whether the present distribution of these two haplogroups can be the consequence of the same expansion/dispersal microevolutionary event, the two frequency distributions in Europe were compared. A high and statistically significant correspondence between the frequencies of the two haplogroups (r = 0.84, 95% C.I. 0.70–0.92) was observed. A similar result (r = 0.85, 95% C.I. 0.70–0.93) was obtained when the series was enlarged with the J–M12 data from Bosnia, Croatia and Serbia (*Marjanović et al. 2005*) matched with the frequencies of E–M78 cluster α (*Peričić et al. 2005*) as a proxy for haplogroup E–V13.

*Fig. 3. Map of the observed haplogroup E–V13 frequencies (*Cruciani et al. 2007*).*

Finally, tetranucleotide microsatellite data were used in order to obtain a coalescence estimate for the J–M12 haplogroup in Europe. By taking into consideration two different demographic expansion models, age estimates very close to those of E–V13 were obtained, *i.e.* 4.1 ky BP (95% C.I. 2.8-5.4 ky BP) and 4.7 ky BP (95% C.I. 3.3-6.4 ky), respectively.

The overall view was confirmed by subsequent works aimed at clarifying the peopling of Crete. According to Martinez et al. (*2007*) E–M78 cluster α chromosomes (which largely overlap E–V13) may have reached Crete as a result of gene flow from mainland Greece during and/or after the Neolithic. King et al. (*2008*) dated the expansion of E–V13 chromosomes in Crete at 3.1 ky BP, "*arguably reflecting the presence of a mainland Mycenaean population in Crete*". Also, the V13 marker is able to rule out recent genetic affinities between Crete and Egypt, where E chromosomes are mainly devoid of V13.

Conclusion 2
The congruence between frequency distributions, shape of the networks, pair-wise haplotypic differences and coalescent estimates point to a single
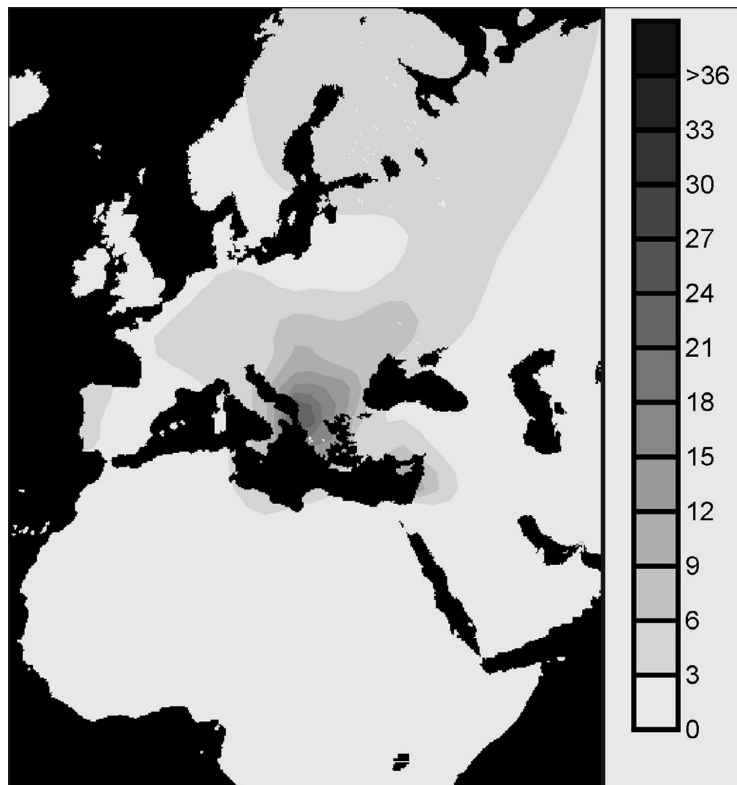
evolutionary event at the basis of the distribution of haplogroups E–V13 and J–M12 within Europe, a finding never appreciated before. These two haplogroups account for more than one fourth of the chromosomes currently found in the southern Balkans, underlining the strong demographic impact of the expansion in the area.

At least four major demographic events have been envisioned for this geographic area, *i.e.* the post-Last Glacial Maximum expansion (about 20 ky BP) (*Taberlet et al. 1988; Hewitt 2000*), the Younger Dryas-Holocene re-expansion (about 12 ky BP), the population growth associated with the introduction of agricultural practices (about 8 ky BP) and the development of Bronze technology (about 5 ky BP). Though large, the confidence intervals for the coalescence of both haplogroups E–V13 and J–M12 in Europe exclude the expansions following the Last Glacial Maximum, or the Younger Dryas. The estimated coalescence age of about 4.5 ky BP for haplogroups E–V13 and J–M12 in Europe (and their C.I.s) would also exclude a demographic expansion associated with the introduction of agriculture from Anatolia and would place this event at the beginning of the Balkan Bronze Age, a period that saw strong demographic changes as clearly seen in the archaeological record. The arrangement of E–V13 and J–M12

frequency surfaces appears to fit the expectations for a range expansion in an already populated territory. Moreover, similarly to what Peričić et al. (*2005*) found for the E–M78 network, the dispersion of E–V13 and J–M12 haplogroups seems to have mainly followed the rivers connecting the southern Balkans to north-central Europe, a route that had already hastened by a factor of 4–6 the spread of the Neolithic to the rest of the continent (*Davison et al. 2006*).

Post-Neolithic expansion within Central Europe detected by three haplogroups

Luca et al. (*2007*) explored the MSY diversity in five, closely spaced Czech population samples. The haplogroups P–DYS257*(xR1a) and R1a–SRY$_{10831}$ establish a major divide across central Europe, initially identified with a line roughly extending from the Adriatic to the Baltic (*Malaspina et al. 2000*). This line separates high frequencies of R1a–SRY$_{10831}$ to the East from low frequencies to the West, with an opposite trend for P–DYS257*(xR1a). Kayser et al. (*2005*) found this sharp genetic boundary to coincide with the German-Polish border, and interpreted it as the result of massive population movements associated with World War II, superimposed on pre-existing continent-wide clines. The Czech Republic appears to be affected by a much smoother frequency shift, if any, supporting the interpretation of a very recent origin of the German-Polish discrepancy.

Overall, the haplogroup frequencies identify the Czech population as one influenced to a very moderate extent by genetic inputs from outside Europe in the post-Neolithic and historical times. It thus may represent an ideal population to draw inferences on geographically confined processes that might also have occurred in other parts of central Europe.

Inferences based on STR variation in the three most common haplogroups obtained with coalescent methods deserve careful evaluation. First, even though sampling was carried out in a limited geographic area, it returned age estimates for I–M170, P–DYS257*(xR1a) and R1a–SRY$_{10831}$ similar to those obtained in reports with a wider geographical coverage (approximately 500, 400 and 350 generations ago, respectively). Conservatively, one can simply conclude that the Czech population harbours a large part of the STR variation generated in each haplogroup. The ages of the three most common haplogroups turned out to be largely overlapping, and compatible with their presence during or soon after the Last Glacial Maximum.

However, a local signal emerged from the distribution of this diversity, *i.e.* that of a fast and recent population growth, which persists even after relaxing the prior assumptions of the dating method and is similar for the three haplogroups. This is summarized by the parameters alpha (rate of population growth, 0.023, 0.031 and 0.032 for I–M170, P–DYS257*(xR1a) and R1a–SRY$_{10831}$, respectively) and beta (beginning of population growth, 97, 150 and 125 generations ago, respectively) and their relatively narrow confidence intervals (up to 1.5 fold the average). Estimation of the beta parameter most likely locates the beginning of this process in the 1st millennium BC, with confidence intervals that are barely compatible with the archaeologically documented introduction of Neolithic technology in this area (*Haak et al. 2005*). At least for the female lineage, these authors found a little genetic contribution to the present European gene pool from the first farmers settled in the area. Independently of the relevance of these data for reconstructing the genetics of Europe in the early Neolithic (*Barbujani and Chikhi 2006*), the central value for population growth coincides with a later period of repeated changes in the material cultures in this geographic region, driven by the development of metal technologies and the associated social and trade organization.

Conclusion 3
The combined use of UEP and STR markers allowed the exploration of different time horizons for the age of molecules and for the process of population growth (*Torroni et al. 2006*). In fact, the data for the Czech population favour a model in which the age of the most common MSY molecules could be separated from consistent population growth. Similar results have been obtained for Lithuania (*Kasperaviciute et al. 2004*). Both regions lie at the north-western and northern edge, respectively, of the putative homeland (central and southeastern Europe) of an aboriginal quota of the molecular MSY diversity. This offers an unprecedented opportunity to test alternative models for a continental pattern of diversity which is arranged along the southeast-to-northwest axis. The question of whether this could be the result not only of a single demic diffusion, but also of the demographic increases affecting pre-existing local gene pools is still open. Examples of the recent growth of pre-existing gene pools that add complexity to the simple demic diffusion models, are provided by mtDNA haplogroup HV and H1 (*Achilli et al. 2004*), as well as Y chromosomal haplogroup R–SRY$_{2627}$ (*Hurles et al. 1999*).

Concluding remarks

The build-up of present day male-specific Y chromosome (MSY) diversity can be viewed as an increase in complexity, due to the repeated addition of new variation to the pre-existing background by two main mechanisms: the immigration of differentiated MSY copies from outer regions, and the accumulation of novel MSY variants generated by new mutations *in loco*. Recently, Sengupta et al. (*2006*) pointed out that combining highly resolved phylogenetic hierarchy, haplogroup internal diversification, geography and expansion time estimates can lead to the appropriate diachronic partition of the MSY pool. The DNA content of the MSY ensures that abundant diversity exists to proceed a long way in this process of phylogeographic refinement, eventually leading to a level of resolution for human history comparable with, or even greater than, that achieved by mitochondrial DNA (*Torroni et al. 2006*).

In addition, environmental or cultural transitions are usually considered to be the basis of dramatic changes in the size of human populations. These changes, too, are expected to leave a distinct signature in the genetic pools of the populations that experienced them. Even in the absence of known markers that are able to qualitatively mark these episodes, quantitative analysis is feasible and can sometimes lead to robust inferences.

Here we show that a growing body of work converges in disclosing a further level of complexity in the genetic landscape of central and south-eastern Europe. This appears to be, to a large extent, the consequence of a recent population increase *in situ*, rather than the result of a mere flow of western Asian migrants during the early Neolithic.

∴

REFERENCES

ACHILLI A., RENGO C., MAGRI C., BATTAGLIA V., OLIVIERI A., SCOZZARI R., CRUCIANI F., ZEVIANI M., BRIEM E., CARELLI V. et al. 2004. The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *American Journal of Human Genetics 75:910–8*.

BARBUJANI G., CHIKHI L. 2006. Population genetics: DNAs from the European Neolithic. *Heredity 97: 84–5*.

BELLWOOD P. 2004. The origins of Afroasiatic. *Science 306: 1681*.

CADENAS A. M., ZHIVOTOVSKY L. A., CAVALLI-SFORZA L. L., UNDERHILL P. A., HERRERA R. J. 2007. Y-chromosome diversity characterizes the Gulf of Oman. *European Journal of Human Genetics 13: 374–86*.

CINNIOGLU C., KING R., KIVISILD T., KALFOGLU E., ATASOY S., CAVALLERI G. L., LILLIE A. S., ROSEMAN C. C., LIN A. A., PRINCE K. et al. 2004. Excavating Y-chromosome haplotype strata in Anatolia. *Human Genetics 114: 127–48*.

CRUCIANI F., LA FRATTA R., SANTOLAMAZZA P., SELLITTO D., PASCONE R., MORAL P., WATSON E., GUIDA V., COLOMB E. B., ZAHAROVA B. et al. 2004. Phylogeographic analysis of haplogroup E3b (E–M215) y chromosomes reveals multiple migratory events within and out of Africa. *American Journal of Human Genetics 74: 1014–22*.

CRUCIANI F., LA FRATTA R., TROMBETTA B., SANTOLAMAZZA P., SELLITTO D., COLOMB E. B., DUGOUJON J. M., CRIVELLARO F., BENINCASA T., PASCONE R. et al. 2007. Tracing past human male movements in northern/eastern Africa and western Eurasia: new clues from Y-chromosomal haplogroups E–M78 and J–M12. *Molecular Biology and Evolution 24: 1300–11*.

DAVISON K., DOLUKHANOV P., SARSON G. R., SHUKUROV A. 2006. The role of waterways in the spread of the Neolithic. *Journal of Archaeological Science 33: 641–52*.

DI GIACOMO F., LUCA F., POPA L. O., AKAR N., ANAGNOU N., BANYKO J., BRDICKA R., BARBUJANI G., PAPOLA F., CIAVARELLA G. et al. 2004. Y chromosomal haplogroup J as a signature of the post-neolithic colonization of Europe. *Human Genetics 115: 357–71*.

EHRET C., KEITA S. O. Y., NEWMAN P. 2004. The origins of Afroasiatic. *Science 306: 1680–81*.

HAAK W., FORSTER P., BRAMANTI B., MATSUMURA S., BRANDT G., TANZER M., VILLEMS R., RENFREW C., GRONENBORN D., ALT K. W. et al. 2005. Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science 310: 1016–8*.

HEWITT G. 2000. The genetic legacy of the Quaternary ice ages. *Nature 405: 907–13*.

HURLES M. E., VEITIA R., ARROYO E., ARMENTEROS M., BERTRANPETIT J., PEREZ-LEZAUN A., BOSCH E., SHLUMUKOVA M., CAMBON-THOMSEN A., MCELREAVEY K. et al. 1999. Recent male-mediated gene flow over a linguistic barrier in Iberia, suggested by analysis of a Y-chromosomal DNA polymorphism. *American Journal of Human Genetics 65: 1437–48*.

JOBLING M. A., HURLES M. E., TYLER-SMITH C. 2004. *Human evolutionary genetics*. New York: Garland Science.

KASPERAVICIUTE D., KUCINSKAS V., STONEKING M. 2004. Y chromosome and mitochondrial DNA variation in Lithuanians. *Annals of Human Genetics 68: 438–52*.

KAYSER M., KITTLER R., ERLER A., HEDMAN M., LEE A. C., MOHYUDDIN A., MEHDI S. Q., ROSSER Z., STONEKING M., JOBLING M. A. et al. 2004. A comprehensive survey of human Y-chromosomal microsatellites. *American Journal of Human Genetics 74: 1183–97*.

KAYSER M., LAO O., ANSLINGER K., AUGUSTIN C., BARGEL G., EDELMANN J., ELIAS S., HEINRICH M., HENKE J., HENKE L. et al. 2005. Significant genetic differentiation between Poland and Germany follows present-day political borders, as revealed by Y-chromosome analysis. *Human Genetics 117: 428–43*.

KING R. J., OZCAN S. S., CARTER T., KALFOGLU E., ATASOY S., TRIANTAPHYLLIDIS C., KOUVATSI A., LIN A. A., CHOW C.-E. T., ZHIVOTOVSKY L. A. et al. 2008. Differential Y-chromosome Anatolian influences on the Greek and Cretan Neolithic. *Annals of Human Genetics 72: 205–14*.

LUCA F., DI GIACOMO F., BENINCASA T., POPA L. O., BANYKO J., KRACMAROVA A., MALASPINA P., NOVELLETTO A., BRDICKA R. 2007. Y-chromosomal variation in the Czech Republic. *American Journal of Physical Anthropology 132: 132–39*.

MALASPINA P., CRUCIANI F., CIMINELLI B. M., TERRENATO L., SANTOLAMAZZA P., ALONSO A., BANYKO J., BRDICKA R., GARCIA O., GAUDIANO C. et al. 1998. Network analyses of Y-chromosomal types in Europe, northern Africa, and western Asia reveal specific patterns of geographic distribution. *American Journal of Human Genetics 63: 847–60*.

MALASPINA P., CRUCIANI F., SANTOLAMAZZA P., TORRONI A., PANGRAZIO A., AKAR N., BAKALLI V., BRDICKA R., JARUZELSKA J., KOZLOV A. et al. 2000. Patterns of male-specific inter-population divergence in Europe, West Asia and North Africa. *Annals of Human Genetics 64: 395–412*.

MALASPINA P., TSOPANOMICHALOU M., DUMAN T., STEFAN M., SILVESTRI A., RINALDI B., GARCIA O., GIPARAKI M., PLATA E., KOZLOV A. I. et al. 2001. A multistep process for the dispersal of a Y chromosomal lineage in the Mediterranean area. *Annals of Human Genetics 65: 339–49*.

MARJANOVIĆ D., FORNARINO S., MONTAGNA S., PRIMORAC D., HADŽISELIMOVIĆ R., VIDOVIĆ S., POJSKIĆ N., BATTAGLIA V., ACHILLI A., DROBNIČ K. et al. 2005. The peopling of modern Bosnia-Herzegovina: Y-chromosome haplogroups in the three main ethnic groups. *Annals of Human Genetics 69: 757–63*.

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. *European Journal of Human Genetics 15: 485–93*.

PERIČIĆ M., LAUC L. B., KLARIĆ I. M., ROOTSI S., JANIČIJEVIĆ B., RUDAN I., TERZIĆ R., COLAK I., KVESIĆ A., POPOVIĆ D. et al. 2005. High-resolution phylogenetic analysis of southeastern Europe traces major episodes of paternal gene flow among Slavic populations. *Molecular Biology and Evolution 22: 1964–75*.

SEMINO O., PASSARINO G., BREGA A., FELLOUS M., SANTACHIARA-BENERECETTI A. S. 1996. A view of the neolithic demic diffusion in Europe through two Y chromosome-specific markers. *American Journal of Human Genetics 59: 964–8*.

SEMINO O., MAGRI C., BENUZZI G., LIN A. A., AL-ZAHERY N., BATTAGLIA V., MACCIONI L., TRIANTAPHYLLIDIS C., SHEN P., OEFNER P. J. et al. 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. *American Journal of Human Genetics 74: 1023–34*.

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. et al. 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. *American Journal of Human Genetics 78: 202–21*.

TABERLET P., FUMAGALLI L., WUST-SAUCY A.-G. J.-F. C. 1988. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology 7: 453–64.*

TORRONI A., ACHILLI A., MACAULAY V., RICHARDS M., BANDELT H. J. 2006. Harvesting the fruit of the human mtDNA tree. *Trends in Genetics 22: 339–45.*

UNDERHILL P. A., SHEN P., LIN A. A., JIN L., PASSARINO G., YANG W. H., KAUFFMAN E., BONNE-TAMIR B., BERTRANPETIT J., FRANCALACCI P. et al. 2000. Y chromosome sequence variation and the history of human populations. *Nature Genetics 26: 358–61.*

WELLS R. S., YULDASHEVA N., RUZIBAKIEV R., UNDERHILL P. A., EVSEEVA I., BLUE-SMITH J., JIN L., SU B., PITCHAPPAN R., SHANMUGALAKSHMI S. et al. 2001. The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proceedings of the National Academy of Sciences USA 98: 10244–9.*

Y CHROMOSOME CONSORTIUM. 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Research 12: 339–48.*