

The Peopling of Modern Bosnia-Herzegovina: Y-chromosome Haplogroups in the Three Main Ethnic Groups

D. Marjanovic^{1,*}, S. Fornarino², S. Montagna², D. Primorac^{3,4}, R. Hadziselimovic¹, S. Vidovic⁵, N. Pojskic¹, V. Battaglia², A. Achilli², K. Drobnic⁶, S. Andjelinovic³, A. Torroni², A. S. Santachiara-Benerecetti² and O. Semino^{2,†}

¹*Institute for Genetic Engineering and Biotechnology, University of Sarajevo, Sarajevo, Bosnia and Herzegovina*

²*Dipartimento di Genetica e Microbiologia "A. Buzzati-Traverso", Università di Pavia, Pavia, Italy*

³*Medical School at Split University, Split, Croatia*

⁴*Medical School at Osijek University, Osijek, Croatia*

⁵*Faculty of Medicine, University of Banjaluka, Banjaluka, Bosnia and Herzegovina*

⁶*Forensic Laboratory and Research Center, Ministry of the Interior, Ljubljana, Slovenia*

Summary

The variation at 28 Y-chromosome biallelic markers was analysed in 256 males (90 Croats, 81 Serbs and 85 Bosniacs) from Bosnia-Herzegovina. An important shared feature between the three ethnic groups is the high frequency of the "Palaeolithic" European-specific haplogroup (Hg) I, a likely signature of a Balkan population re-expansion after the Last Glacial Maximum. This haplogroup is almost completely represented by the sub-haplogroup I-P37 whose frequency is, however, higher in the Croats (~71%) than in Bosniacs (~44%) and Serbs (~31%). Other rather frequent haplogroups are E (~15%) and J (~7%), which are considered to have arrived from the Middle East in Neolithic and post-Neolithic times, and R-M17 (~14%), which probably marked several arrivals, at different times, from eastern Eurasia. Hg E, almost exclusively represented by its subclade E-M78, is more common in the Serbs (~20%) than in Bosniacs (~13%) and Croats (~9%), and Hg J, observed in only one Croat, encompasses ~9% of the Serbs and ~12% of the Bosniacs, where it shows its highest diversification. By contrast, Hg R-M17 displays similar frequencies in all three groups. On the whole, the three main groups of Bosnia-Herzegovina, in spite of some quantitative differences, share a large fraction of the same ancient gene pool distinctive for the Balkan area.

Keywords: Y-Chromosome haplogroups, Polymorphisms, Bosnia-Herzegovina, Balkans.

Introduction

Archaeological findings in central Bosnia (~100,000 years old) and in the north of the country (~50,000

years old) indicate that the territory of modern Bosnia-Herzegovina was continuously settled since the Palaeolithic (Imamović, 1998). At the beginning of the second millennium BC this part of the Balkans was inhabited by different Illyrian tribes, which established the oldest central-western Balkan civilization (Wilkes, 1992). These local inhabitants probably completely assimilated most of the first immigrants (Avars and Slavs) who entered the region soon after the splitting of the Roman Empire. The process of the peopling of Bosnia and Herzegovina was extensively shaped with the arrival of two new tribes, the Croats and the Serbs, whose

*Address for correspondence: Dr. Damir Marjanovic, Institute for Genetic Engineering and Biotechnology, University of Sarajevo, Kemalbegova 10, 71.000 Sarajevo, Bosnia and Herzegovina. Tel. +387 33 220 926; Fax +387 33 442 891, E-mail: damirm.marjanovic@ingeb.ba

†Dr. Ornella Semino, Dipartimento di Genetica e Microbiologia, Università di Pavia, Via Ferrata, 1, 27100 Pavia, Italy. Tel +39 0382 985543; Fax +39 0382 528496, E-mail: semino@ipvgen.unipv.it

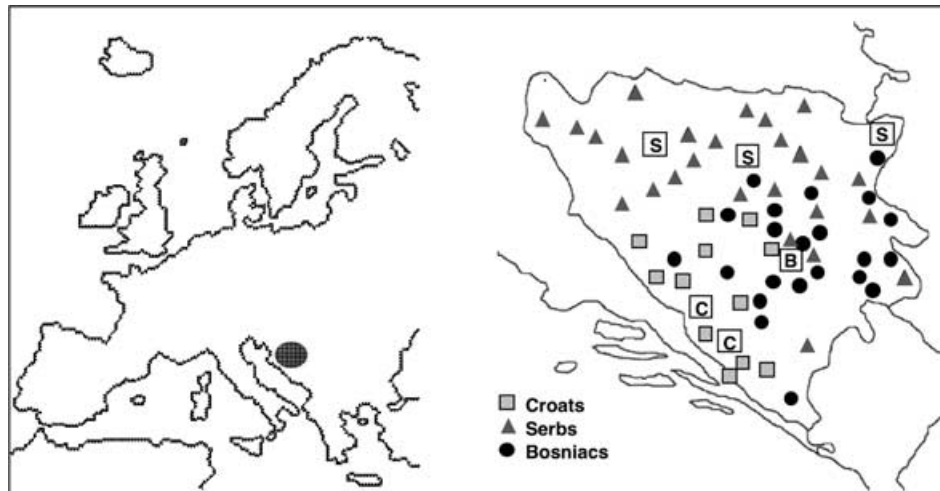


Figure 1 The location of Bosnia-Herzegovina (left) and the locations in Bosnia-Herzegovina from which the population samples were collected (right). Boxed S, B and C refer to Serb, Bosniac, and Croat collecting points. Places of sample origin: Croats – ■; Serbs – ▲; Bosniacs – ●.

origins are still under discussion. The present Bosnia-Herzegovina territory is located between the areas inhabited by Croats (westward) and Serbs (eastward). The populations from the isolated mountain villages, according to the most diffused theory, could be direct descendants of Bosnia's non-Slavic aboriginal inhabitants (Malcolm, 1994). The second important historical period for this country is represented by the expansion of the Ottoman Empire in the fifteen century that was responsible for the Islamicization of a large part of the population (Malcolm, 1994), and pointed to this area as crossroads for different cultures and populations.

Modern Bosnia-Herzegovina is a multinational and multi-religious country with a very stormy recent history. During the 20th century, political justifications for the conflicts in the area were sought through different interpretations of the origins of the three main ethnic groups (Croats, Serbs and Bosniacs) from this region.

The recent detection of a large number of informative biallelic markers in the non-recombining region of the Y chromosome (NRY) has already significantly contributed to the understanding of European pre-history and history (Semino *et al.* 1996, 2000; Underhill *et al.* 2001). Here we have investigated the Y-chromosome variation in 256 individuals from the three major ethnic groups of Bosnia-Herzegovina, with the aim of providing new clues about their origin, and about the ancient and recent events of gene flow which have influenced this area located in the heart of Europe.

Material and Methods

The Sample

The sample consists of 256 (90 Croats, 81 Serbs and 85 Bosniacs – the last previously identified as “Muslim”) unrelated males born in Bosnia-Herzegovina. Blood collection was carried out at 6 collection points, from volunteers appropriately informed about the research. A detailed analysis of the list of blood donors and their family records showed that the sample is representative of all of Bosnia-Herzegovina, since the subjects are from more than 50 different locations (Figure 1). The attribution of the subjects to the three ethnic groups was carried out according to the origin of their paternal grandfather.

Y-chromosome DNA Analysis

DNA was extracted from whole blood according to the standard phenol/chlorophorm procedure, followed by ethanol precipitation. Twenty-eight NRY biallelic markers (12f2, YAP, P37, M9, M12, M17, M26, M35, M47, M67, M68, M74, M78, M81, M89, M92, M102, M123, M170, M172, M173, M201, M207, M223, M227, M253, M267 and M269) (The Y Chromosome Consortium, 2002; Cruciani *et al.* 2002, Rootsi *et al.* 2004) were examined in hierarchical order in agreement with the Y-chromosome phylogeny (YCC, 2002). M9 was chosen as the initial marker and surveyed in all samples.

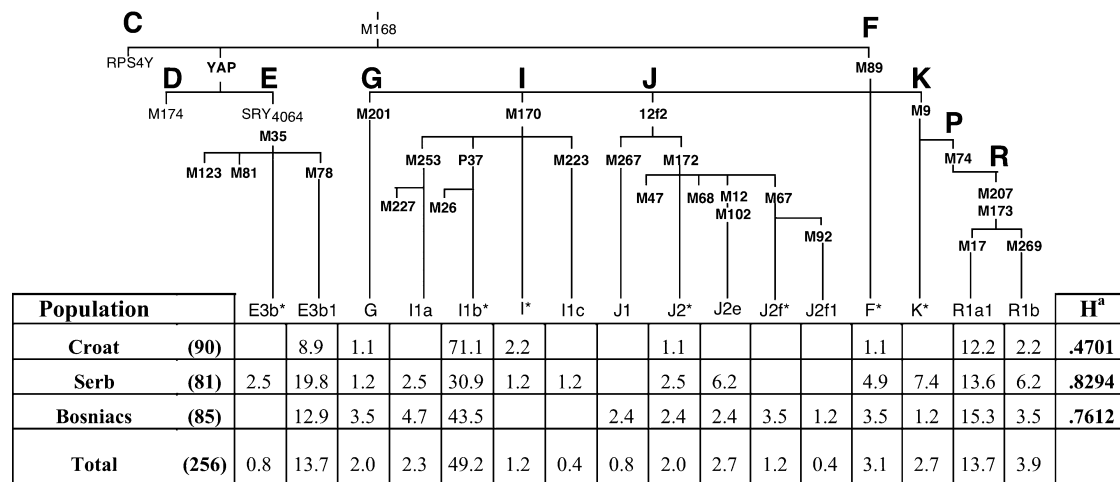


Figure 2 Phylogeny of Y-chromosome haplogroups and their frequencies in Croats, Serbs and Bosniacs of modern Bosnia-Herzegovina. Significant differences:

I-P37 Croats *vs* Serbs $P < 10^{-4}$; Croats *vs* Bosniacs $P = 2 \times 10^{-4}$

E Croats *vs* Serbs $P < 0.02$

J Croats *vs* Serbs $P_{\text{Fisher}} = 0.020$; Croats *vs* Bosniacs $P_{\text{Fisher}} = 0.003$

^aIntrapopulation haplogroup diversity.

The 12f2 and YAP polymorphisms were analyzed according to Rosser *et al.* (2000) and Hammer & Horai (1995), respectively; mutations P37, M9 and M269 were typed through PCR/RFLP assays according to Karafet *et al.* (2001) (P37) and Cruciani *et al.* (2002) (M9 and M269). All the other mutations were detected by PCR/DHPLC, as reported by Underhill *et al.* (2000). The YCC (2002) nomenclature was used for haplogroup labelling.

Statistical Analysis

Haplogroup diversity was computed using Nei's standard method (Nei & Kumar, 2000). Genetic distance tests were performed using the approach of Reynolds *et al.* (1983). Population differentiation tests (Raymond & Rousset, 1995) and AMOVA (Weir & Cockerham, 1984; Excoffier *et al.* 1992) were performed using the Arlequin package, ver. 2000 (Schneider *et al.* 2000). Principal Component (PC) analysis was performed on haplogroup frequencies by using Excel implemented by the XL-stat program.

Results

The analysis of 28 Y-chromosome biallelic markers has allowed the classification of the 256 Bosnian samples

into haplogroups E, F*, G, I, J, K* and R. The phylogenetic relationships of these haplogroups, their frequencies in the three studied sub-groups (Croats, Serbs and Bosniacs) and in the pooled sample are illustrated in Figure 2 with the P values reported for significantly different frequencies. Haplogroup I is the most commonly represented haplogroup, accounting for more than 50% of the Y chromosomes, and it is almost exclusively represented by its sub-haplogroup I-P37. Other subsets of Hg I such as I-M253 and I-M223 were scarce, while I-M26 was not observed. Additional haplogroups with overall frequencies higher than 5% are haplogroups E (14.5%), R-M17 (13.7%) and J (7.1%). Their distributions, however, differ in the three groups. In particular, the Croats show the highest I-P37 frequency (71.1% *vs* 30.9% in Serbs and 43.5% in Bosniacs) and the lowest frequencies for both Hg E (8.9% *vs* 22.3% in Serbs and 12.9% in Bosniacs) and Hg J (1.1% *vs* 8.7% in Serbs and 11.9% in Bosniacs). The particular haplogroup distribution in the Croats is mirrored in their intrapopulation diversity value (Figure 2). Indeed, the Croats are the most homogeneous group ($H = 0.47$), followed by the Bosniacs ($H = 0.76$) and the Serbs ($H = 0.83$).

On the whole, the haplogroup distribution, which accounts for 5.8% of the variation among populations, significantly differs between Croats and Bosniacs ($P < 10^{-5}$), as well between Croats and Serbs ($P < 10^{-5}$),

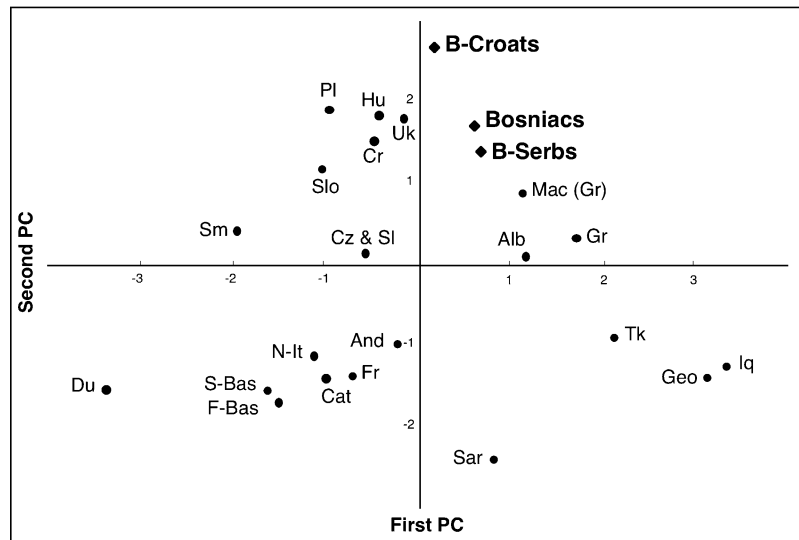


Figure 3 PC analysis performed using the frequencies of the Y-chromosome haplogroups E-M35, J1, J2, G, I*, I-P37*, I-M26, I-M223, I-M253, R-M17 and R-M269 in the three Bosnian groups (present paper), and in other European populations. The data from the additional European populations (Alb = Albanians, And = Andalusians, F-Bas = French Basques, S-Bas = Spanish Basques, Cat = Catalans, Cr = Croats of Croatia, Cz & Sl = Czechoslovaks, Du = Dutch, Fr = French, Hu = Hungarians, Geo = Georgians, Gr = Greeks, Mac (Gr) = Macedonian Greeks, Iq = Iraqis, N-It = North Italians, PI = Poles, Sar = Sardinians, Slo = Slovenians, Sm = Saami, Tk = Turks, Uk = Ukrainians) are from Semino *et al.* (2000), Rootsi *et al.* (2004), and unpublished data. On the whole, 41% of the total variance is represented: 22% by the first PC and 19% by the second PC.

but not between Serbs and Bosniacs. The relationships among the three groups in the overall European scenario are summarized by the plot of the first and second principal components illustrated in Figure 3. Croats and Serbs, according to the calculated genetic distances, are the most divergent from each other ($D = 0.1234$), followed by Croats *vs* Bosniacs ($D = 0.0550$) and Serbs *vs* Bosniacs ($D = 0.0082$).

Discussion

The analysis of Y-chromosome variation in the three main ethnic groups of modern Bosnia-Herzegovina reveals that Bosnian Croats, Bosnian Serbs and Bosniacs harbour haplogroups which they share with many other Europeans (Semino *et al.* 2000). Indeed the people of Bosnia-Herzegovina display European-specific haplogroups that most likely arose in different glacial refuge areas of Europe (I-M170, R-M17 and R-M269 from Balkan, Ukrainian and Franco-Cantabrian refuges, respectively), and haplogroups considered to

have originated in Africa (E-SRY₄₀₆₄) and the Middle East (J-12f2) and to have arrived in Europe through the prolonged gene flow from the Middle East (Cruciani *et al.* 2004, Semino *et al.* 2004). However, the Y chromosomes in Bosnia-Herzegovina differ from those of most other European regions because more than 50% belong to I-P37, a specific sub-haplogroup of I.

The frequency of sub-haplogroup I-P37 observed in the Croats is particularly high (71.1%) and could be partially attributed to genetic drift, but the high frequencies observed also in the Bosniacs (43.5%) and Serbs (30.9%) show that the three population groups share a major subset of their gene pool, and that this ancestral gene pool was affected by a major demographic event. Taking into account that a Palaeolithic origin of the P37 mutation in this Balkan district has been previously suggested (Rootsi *et al.* 2004), it is possible that the post-LGM expansion of a population with a high frequency of I-P37 from one of the refuges present in the Balkans played a major role in the peopling of Bosnia-Herzegovina and surrounding areas.

The second most frequent haplogroup in Bosnia-Herzegovina is Hg E (14.5%), whose presence in Europe has been attributed to multiple migrations from the Middle East and North Africa during and after the Neolithic (Cruciani *et al.* 2004, Semino *et al.* 2004). However, Hg E is almost exclusively represented by the sub-clade E-M78. This accounts for 13.7% of the pooled samples, but is less common in Croats (8.9%) and Bosniacs (12.9%) than in Serbs (19.8%). The frequency observed in the Serbs is significantly different from that in Croats ($P < 0.05$) but it is similar to that in other southern Balkan populations (20% in Greeks and 25% in Albanians) (Semino *et al.* 2004), where the highest European frequencies of E-M78 are observed. It is worth remembering that the clinal distribution in Europe of E-M78 and its internal microsatellite variance have been attributed to dispersals, in Neolithic and post-Neolithic times, from the Balkans to all directions, as far as Iberia to the west and, most likely, also to Turkey in the southeast (Cinnioğlu *et al.* 2004; Cruciani *et al.* 2004; Semino *et al.* 2004). In this frame, our data suggest that the expansion(s) of E-M78 would have affected, to different extents, the ancestral Balkan populations.

Haplogroup J is another haplogroup that arrived in Europe from the Middle East and its sub-clades probably marked complex migration processes during and after Neolithic time (Cinnioğlu *et al.* 2004, Di Giacomo *et al.* 2004, Semino *et al.* 2004). In Bosnia-Herzegovina Y chromosomes belonging to J are found mainly in the Bosniacs who, interestingly, harbour almost all the known J sub-clades (Semino *et al.* 2004). These are: J-M267, which has been associated with the Arab expansions; J-M92, which suggests genetic links between Anatolia and southern Italy; J-M67, which is frequent in the Caucasus; and finally J-M102, which shows frequency peaks in the southern Balkans and central-southern Italy. The last, however, appears to be more represented in the Serbs (6.2%). Thus, overall a higher extent of gene flow could have occurred in the Bosniacs, while the Croats, in whom a single undifferentiated J-M172 Y chromosome was encountered, were probably the group in which genetic drift and founder events played the most important role, as already suggested by the extremely high frequency of I-P37.

R-M17 is the prevalent sub-haplogroup of R as previously observed in other eastern European populations (Semino *et al.* 2000, Passarino *et al.* 2001, Wells *et al.* 2001). Its frequency is very similar in the three groups, ranging from 12.2% in the Croats to 15.3% in the Bosniacs, and perfectly fits the expected distribution of R-M17 which is found almost exclusively in eastern Europe with a decreasing gradient from north-east to south-west. This gradient, initially attributed to expansion(s) from a Ukrainian glacial refuge (Semino *et al.* 2000), could also be due to infiltrations of Indo-European speaking peoples from southern Russia about 2,000 years ago (Jovanović 1979; Barać *et al.* 2003), as well as to the arrival of the Slav clans during the 6th and 7th centuries. However, to evaluate this latter scenario, larger samples of R-M17 Y chromosomes from the area and detailed analyses of their STR diversity are required.

Of interest is the presence of the R-M269 Y chromosomes in modern Bosnia-Herzegovina. Despite their relatively low frequency, ranging from 2.2% in the Croats to 6.2% in the Serbs, they indicate that the gene pool of the ancestral population(s) of the Franco-Cantabrian refuge area also contributed to some extent to this region of the Balkans, as additionally attested by the finding in the same region of the mtDNA haplogroup H1 (Achilli *et al.* 2004).

In the PC analysis (Figure 3), the first PC shows that the three populations are genetically extremely close to each other, and closely related to other populations of the Balkans. However, the second PC tends to separate the Croat group not only from both Serbs and Bosniacs, but also from the Croats of Croatia.

In conclusion our data suggest that a post-glacial expansion – possibly from a LGM refuge area in the Balkans – probably gave rise to the frequency peak of haplogroup I-P37 in the region, with decreasing consequences for the Croats, to the Bosniacs and then to the Serbs. In contrast, Neolithic and post-Neolithic gene flow appears to have played an overall less important role in all three populations. It only marginally influenced the Bosnian Croats, but provided some “Southern Balkan” (Hg E-M78) and “Anatolian” (different clades of Hg J) contributions to gene pools of the Serbs and Bosniacs, respectively. On the contrary, the migratory processes from Central Asia and Eastern Europe – marked by haplogroup R-M17 – seem to have similarly

influenced the three major ethnic groups of modern Bosnia-Herzegovina.

Acknowledgements

We are grateful to all the donors for providing the blood samples and to all the people and institutions that contributed to their collection. These include the Institute for Transfusion and the Institute for Clinical Biochemistry of Sarajevo, the Cantonal Hospital and the Primary Care Institution of Mostar, the Clinical Hospital Centre and the Institute for Transfusion of Banja Luka, the Regional Hospitals of Doboj and Bijeljina. We wish to thank the reviewers, whose comments and suggestions helped us to improve the quality of the manuscript. This research was supported by Progetto Finalizzato C.N.R. "Beni Culturali", Fondo d'Ateneo per la Ricerca dell'Università di Pavia, the Italian Ministry of the University (Progetti Ricerca Interesse Nazionale 2002, 2003), and the Ministry of Science and Education of the Sarajevo Canton, Bosnia-Herzegovina.

References

- Achilli, A., Rengo, C., Magri, C., Battaglia, V., Olivieri, A., Scozzari, R., Cruciani, F., Zeviani, M., Briem, E., Carelli, V., Moral, P., Dugoujon, J. M., Roostalu, U., Loogvali, E. L., Kivisild, T., Bandelt, H. J., Richards, M., Villems, R., Santachiara-Benerecetti, A. S., Semino, O. & Torroni, A. (2004) The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *Am J Hum Genet* **75**, 910–918.
- Barać, L., Pericic, M., Klaric, I. M., Janicijevic, B., Parik, J., Rootsi, S. & Rudan, P. (2003) Y-chromosomal heritage of Croatian population and its islands isolates. *Eur J Hum Genet* **11**, 535–542.
- 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.
- 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* **72**, 1197–1214.
- Cruciani, F., La Fratta, 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.
- 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 Stasi, L., Gavrilu, L., Kerimova, M. G., Kovatchev, D., Kozlov, A. I., Loutradis, A., Mandarino, V., Mammi, C., Michalodimitrakakis, 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.
- Excoffier, L., Smouse, P. & Quattro, J. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131**, 343–359.
- Hammer, M. F. & Horai, S. (1995) Y chromosomal DNA variation and the peopling of Japan. *Am J Hum Genet* **56**, 951–962.
- Imamović, E., Lovranović, D., Nilević, B., Sunjic, M., Zlatar, B., Pelidija, E., Tepic, I., Hadzicengic, I., Imamovic, M., Isek, T., Brkljaca, S., Pelisic, M., Kamberovic, H. (1998) *Bosnia and Herzegovina – from the oldest age until Second World War*. Sarajevo: Bosanski kulturni centar.
- Jovanović, B. (1979) *Prehistory of south Slavic countries*. (ed. N. Tasic) vol 3. Sarajevo: Svjetlost.
- Karafet, T., Xu, L., Du, R., Wang, W., Feng, S., Wells, R. S., Redd, A. J., Zegura, S. L. & Hammer, M. F. (2001) Paternal population history of East Asia: sources, patterns and microevolutionary processes. *Am J Hum Genet* **69**, 615–628.
- Malcolm, N. (1994) *Bosnia. A Short History*. London: Macmillan London Limited Copyright.
- Nei, M. & Kumar, S. (2000) *Molecular Evolution and Phylogenetics*. Oxford, New York: Oxford University Press.
- Passarino, G., Semino, O., Magri, C., Al-Zahery, N., Benuzzi, G., Quintana-Murci, L., Andellnovic, S., Bullc-Jakus, F., Liu, A., Arslan, A. & Santachiara-Benerecetti, A. S. (2001) The 49a,f haplotype 11 is a new marker of the EU19 lineage that traces migrations from northern regions of the Black Sea. *Hum Immunol* **62**, 922–932.
- Raymond, M. & Rousset, F. (1995) An exact test for population differentiation. *Evolution* **49**, 1280–1283.
- Reynolds, J. B., Weir, B. S., Cockerham, C. C. (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* **105**, 767–779.
- 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 Europa. *Am J Hum Genet* **75**, 128–137.
- Rosser, Z. H., Zerjal, T., Hurler, M. E., Adojaan, M., Alavantic, D., 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., Corte-Real, H. B., de Knijff, P., Decorte, R., Dubrova, Y. E., Evgrafov, O., Gilissen, A., Glisic, S., Golge, M., Hill, E. W., Jeziorowska, A., Kalaydjieva, L., Kayser, M., Kivisild, T., Kravchenko, S. A., Krumina, A., Kucinskas, 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., Norby, S., Pandya, A., Parik, J., Patsalis, P. C., Pereira, L., Peterlin, B., Pielberg, G., Prata, M. J., Previdere, 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**, 1526–1543.
- Schneider, S., Roessli, D. & Excoffier, L. (2000) *Arlequin ver. 2000: The software for population genetics data analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- 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. *Am J Hum Genet* **59**, 964–968.
- 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* in extant Europeans: a Y chromosome perspective. *Science* **290**, 1155–1159.
- 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 late migratory events in the Mediterranean area. *Am J Hum Genet* **74**, 1023–1034.
- The Y Chromosome Consortium (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* **12**, 339–348.
- 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.
- Underhill, P. A., Passarino, G., Lin, A. A., Shen, P., Mirazon, Lahr, 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.
- Weir, B. S. & Cockerham, C. C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.
- 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 hearthland: a continental perspective on Y-chromosome diversity. *Proc Natl Acad Sci USA* **98**, 339–348.
- Wilkes, J. (1992) *The Illyrians*. Oxford (UK): Blackwell Publisher.

Received: 29 November 2004

Accepted: 15 February 2005