Y-chromosomal evidence of the cultural diffusion of agriculture in southeast Europe

Vincenza Battaglia1, Simona Fornarino1, Nadia Al-Zahery1, Anna Olivieri1, Maria Pala1, Natalie M Myres2, Roy J King3, Siiro Rootsi4, Damir Marjanovic5,6, Dragan Primorac7,8, Rifat Hadziselimovic5, Stojko Vidovic9, Katia Drobnic10, Naser Durmishi11, Antonio Torroni1, A Silvana Santachiara-Benerecetti1, Peter A Underhill3 and Ornella Semino*,1

1Dipartimento di Genetica e Microbiologia, Università di Pavia, Pavia, Italy; 2Sorenson Molecular Genealogy Foundation, Salt Lake City, UT, USA; 3Department of Psychiatry and Behavioral Sciences, School of Medicine, Stanford University, Stanford, CA, USA; 4Department of Evolutionary Biology, University of Tartu and Estonian Biocentre, Tartu, Estonia; 5Institute for Genetic Engineering and Biotechnology, University of Sarajevo, Sarajevo, Bosnia and Herzegovina; 6Genos doo, Zagreb, Croatia; 7Medical School, Split University, Split, Croatia; 8Medical School, Osijek University, Osijek, Croatia; 9Faculty of Medicine, University of Banja Luka, Banja Luka, Bosnia and Herzegovina; 10Forensic Laboratory and Research Center, Ministry of the Interior, Ljubljana, Slovenia; 11Medical School, University of Tetovo, Tetovo, Former Yugoslavia Republic of Macedonia

The debate concerning the mechanisms underlying the prehistoric spread of farming to southeast Europe is framed around the opposing roles of population movement and cultural diffusion. To investigate the possible involvement of local people in the Balkan’s diffusion of agriculture, we analysed patterns of Y-chromosome diversity in 1206 subjects from 17 population samples, mainly from southeast Europe. Three main Y-chromosome clades, I-M423, E-V13 and J-M241, turned out to be informative in distinguishing into southeast Europe between Holocene Mesolithic foragers and Neolithic range expansions from the eastern Sahara and the Near East. In particular, the Balkan microsatellite variation associated with J-M241 concords with Neolithic times, those related to E-V13 and I-M423 Balkan Y chromosomes are consistent with a late Mesolithic time frame. In addition, the low frequency and variance associated with I-M423 and E-V13 in Anatolia and the Middle East make very likely a south European Mesolithic origin for both. Thus, Balkan Mesolithic foragers with their own autochthonous genetic signatures were the earliest to adopt agriculture when it was introduced by farming people from the Near East. These initial local converted farmers became the principal agents spreading this economy using maritime leapfrog colonization strategies in the Adriatic and transmitting the Neolithic cultural package to other adjacent Mesolithic populations. The ensuing range expansions of J-M241, E-V13 and I-M423 parallel in space and time the diffusion of Neolithic Impressed Ware, thereby documenting a case of cultural diffusion using genetic evidence.

European Journal of Human Genetics (2008), 1–11
& 2008 Macmillan Publishers Limited All rights reserved 1018-4813/08 $32.00
www.nature.com/ejhg

Keywords: Balkan Neolithic; farming transition; peopling of Europe; Y-chromosome haplogroups

*Correspondence: Dr O Semino, Dipartimento di Genetica e Microbiologia, Università di Pavia, Pavia 27100, Italy. Tel: +39 0382 985543; Fax: +39 0382 528496; E-mail: semino@ipvgen.unipv.it

Received 2 September 2008; revised 17 November 2008; accepted 20 November 2008

Introduction

Europe was colonized by modern humans about 40000 years ago and underwent a second colonization wave during the Neolithic, with the spread of farming.1,2
relative Palaeolithic and Neolithic contribution to the current European gene pool has been widely debated and is still under discussion.3–6 Two opposing models have been cited to account for the spread of farming in Europe: the demic diffusion model, which implies a movement of people and therefore a significant Near Eastern genetic input,7 and the cultural diffusion model, which, on the contrary, considers the transition to agriculture as a cultural phenomenon, and therefore without major changes at the genetic level.5 Archaeological evidences suggest, however, that the spread of agriculture has been a complex process characterized by migrations and local admixture.9,10

Genetic studies have described allele frequency clines for different markers along the European colonization routes.11–13 These have been interpreted in favour to demic diffusion and considered a strong indication of a Neolithic contribution to the modern European gene pool.14,15 However, a number of simulation studies have demonstrated that allele frequency clines can also occur in range expansions where admixture is not present.16–19 The impact of the Neolithic in the Balkans, as archaeological records show, is regarded to be considerable,20 although Mesolithic hunter-gatherers were present in this region just before the first appearance of Neolithic pottery.21 Thus, the question to address is the following: can we detect present-day signals of interactions between indigenous Mesolithic people and agricultural colonists in the southeast European gene pool? At this regard, particularly useful are Y-chromosome markers, whose distribution is often highly correlated with their phylogeny.22

Interestingly, the Y-chromosome gene pool of southeast European populations is characterized both by ‘autochthonous’ markers, such as haplogroup (Hg) I, present in the Balkans in pre-Neolithic times,23,24 and by markers mainly belonging to Hgs E and J, which originated outside Europe, in Africa and the Middle East, respectively.12,13,25–28 In addition, new Y-chromosome polymorphisms have added further sophistication to phylogenetic relationships, especially within Hgs E and J,29–32 providing the opportunity to more fully evaluate the above issue. This possibility induced us to carry on a deeper genetic structure characterization of the Balkan area through the analysis of 80 Y-chromosome bi-allelic markers and 12 linked STR loci in 1206 subjects from 17 population samples mainly from southeast Europe.

Materials and methods

Samples

The sample consists of 1206 unrelated male individuals from 17 population samples (Figure 1). Two-hundred and thirty-five of these, namely 64 Albanians from Former Yugoslavia Republic of Macedonia, 29 Croatians from Osijek, 75 Slovenians and 67 northeast Italians (from the province of Trento), are reported here for the first time. The remaining include samples reported earlier,23,27,33 and consist of 104 Caucasians (38 Balkarians and 66 Georgians), 149 Greeks (92 from Athens and 57 from Macedonia), 55 Albanians (collected in Tirana), 89 Croatians, 99 Polish, 75 Czechs, 92 Ukrainians, 53 Hungarians and 255 Bosnia-Herzegovinians (84 Bosnians, 90 Croats and 81 Serbs). Blood samples were collected from healthy unrelated adults after obtaining informed consent. DNA was extracted from whole blood according to the standard phenol/chloroform procedure, followed by ethanol precipitation.

In addition, P37.2* samples identified from a screening of the Sorenson Molecular Genealogy Foundation collection (over 14000 Y chromosomes from more than 100 countries) were also included.

Figure 1 Geographic location of the studied samples: 1, Greeks; 2, Macedonian Greeks; 3, Albanians; 4, Albanians from Former Yugoslavia Republic of Macedonia (FYROM); 5, Bosniacs; 6, Bosnia-Croats; 7, Bosnia-Serbs; 8, Croats; 9, Croats from Osijek; 10, Slovenians; 11, northeast Italians; 12, Hungarians; 13, Czechs; 14, Poles; 15, Ukrainians; 16, Georgians; and 17, Balkarians.
Examined bi-allelic markers and microsatellites
Eighty Y-specific unique event polymorphisms (Figure 2) were examined in hierarchical order. Two new mutations were discovered: M507 (AC006376.2: g.148697 T→G) and M521 (AC012068.5: g. 5219 C→T). The first is associated with I-M253 and was discovered while typing M227, the second is associated with E-M78* and was ascertained while typing V12.

Bi-allelic markers were analysed by PCR/AFLPs (YAP44 and 12f213), by PCR/RFLPs (P37,35 M9, M269,16 V12, V13, V22, V27, V32 and V6528-30) or by PCR/DHPLC (M241, M253, M267, M285, M287, M304, M343, M365, M367–M36928, M410,37 M423, M429, M43831 and M40638 and those reported in the YCC39).

Nomenclature used for Hg labelling is in agreement with YCC39 conventions and recent updating.31,32,38 Subsets of samples belonging to E-V13, J-M241 and I-M423 Hgs were also analysed at 12 STR loci: DYS19, YCAIIa/b, DYS388, DYS389/II, DYS390, DYS391, DYS392, DYS393, DYS439 and DYS460 by using multiplex reactions according to STRbase information (http://www.cstl.nist.gov/biotech/strbase/y20prim.htm) and ABI PRISM® 3100 DNA Sequencer, internal size standard and GeneScan fragment analysis software. DYS44538 was analysed separately in J-M410 samples.

Statistical analysis
Hg diversity (H) was computed using the standard method of Nei.40 Principal components (PC) analysis was performed on Hg frequencies using Excel software implemented by Xlstat. The relative amount of accumulated diversity, as a function of geography, was evaluated through the mean microsatellite variance estimated for each population with a sample size of at least five individuals. Hg frequency and variance maps were generated by using Surfer Golden software following the Kriging procedure.42 Within specific Hgs, Median-Joining (MJ) networks42 were constructed using Network 4.5.0.0 program (Fluxus Engineering, http://www.fluxus.engineering.com). Networks were calculated by the MJ method, where ε = 0 and microsatellites loci were weighted proportionally to the inverse of the repeat variance observed in each Hg and after having processed the data with the reduced-median method. The age of microsatellite variation within Hgs was evaluated employing the methodology described by Zhivotovsky et a43 as modified according to Sengupta et al.37 A microsatellite evolutionary effective mutation rate of 6.9 × 10−4 was chosen as it is suitable for use in situations where the elapsed time frame is ≥1000 years or ~40 generations, clearly appropriate given the prehistoric time depths being explored in this study.44

It is worth mentioning that ambiguities related to past episodes of population history (eg, size fluctuations, bottlenecks, etc.) create inherent uncertainties in the calibration of the YSTR molecular clock; thus the estimated ages of microsatellite variation should be considered with caution.

Results
Figure 2 illustrates the phylogenetic relationships of Y-chromosome Hgs and their distribution in the examined southeast European populations. The main Hgs observed in Europe45 (E, I, J, R1a and R1b) contribute differently to the gene pool of the various East European areas, Hg I and Hg R being the most represented always and Hg E and Hg J being mainly frequent in the southern Balkan population.

Hg I is restricted to western Eurasia23 and is particularly frequent in the Balkans where it characterizes 36.3% of the total Y chromosomes. Two of its branches, I-M223 and I-M253, are scarcely represented in southeast Europe, the first being only sporadically observed and the second showing frequencies around 5%, with higher values (around 9%) in Macedonian Greeks and Croats. Differently, the recently described M423 SNP,44 which characterizes the previously paraphyletic P37.2 clade, accounts for the majority (77.2%) of the East European Hg I chromosomes. Its diffusion seems not to have affected the neighbouring North Italian populations, where low incidences (0–2%) are observed. The I-M423 sub-clade is characterized by a frequency distribution with high Central Balkan values (>70%) in Bosnia-Herzegovina and decreasing frequencies moving from the southern Dinaric Alps to northern Croatia. Although I-M423 comprises virtually all the I-P37.2 Balkan-related chromosomes reported earlier,31 we have also detected one I-P37.2* Albanian subject and, from a screening of previously identified P37.2 chromosomes (Rootsi et al23 and SMGF collection), 30 further P37.2* subjects, two from Moldavia23 and 28 of either documented or presumed western European ancestry, were identified.

Hg R1 is common throughout western Eurasia22 and accounts for more than 30% of the Balkan Y-chromosome pool. With the exception of one R1a*-SRY10831.2 and five R1b-M343* individuals, all the remaining R1 lineages belong to R1a1-M17* and R1b3-M269. These two sub-clades, which show in Europe opposite-frequency gradients with maximum incidences in eastern and western regions, respectively, still display high values in northern Balkans and sensibly decrease southward. R-M269 chromosomes are common in the Balkans and Anatolia and, according to the observed internal divergence of their 49a,f branches,35 most likely predated the origin of agriculture. However, the current lack of informative Hg sub-division within these populations (Figure 2) does not allow, at this time, to evaluate the role of R-M269 chromosomes during the transition to agriculture.

E-M35 is the only branch of Hg E observed in this survey. It is prevalently represented by E-M78 chromosomes,
Transition to farming in the Balkans

V Battaglia et al
almost completely (>90%) belonging to the recently described E-V13 sub-clade. Only four E-M78*, which do not belong to any already described sub-clade, have been observed in the southern Balkans. Two of them (from Greece) turned out to be characterized by the mutation M521 and therefore represent a new M78 lineage.

The majority of the Balkan Hg J Y chromosomes belong to the J-M172 sub-Hg and range from 2% to 20%. Both its main branches, J-M410 and J-M12/M102*, were observed; although the first is scattered in different sub-clades (J-M67, J-M92 and J-DYS445-6) with distinct local patterns, the second is most represented by J-M241.

The PC analysis, from the perspective of population Hg frequencies (Figure 3), reveals a tight cluster of populations not comprising southern Balkan and Caucasian groups. Common to this cluster are lower frequencies of Hgs, G-M201 and J-M410, and higher frequencies of Hgs, I-M423, E-V13 and J-M241. Whereas the first two are primarily Middle Eastern Hgs and have been shown to be associated with the early Neolithic colonization of Crete, Italy, and southern Caucasus, I-M423, E-V13 and J-M241, in spite of parallel Balkan patterns of distribution, have clearly different origins. Their comparison can therefore provide insights into the
complex interaction between the European Mesolithic foragers and the Middle Eastern Neolithic farmers during the transition to farming society in the Balkans. These Hgs, although characterized by different distribution patterns of frequency and variance (Figure 4), display networks of microsatellite haplotypic variation (Supplementary Figures S1–S3), all consistent with a Balkan expansion.

Figure 4  Frequency (left) and variance (right) distributions of the main Y-chromosome haplogroups, I-M423, E-V13 and J-M241, observed in this survey. Frequency data are reported in Figure 2, variance data are relative to the examined microsatellite reported in the Supplementary Table S2. We acknowledge that interpolated spatial frequency surfaces should be viewed with caution because of sample size.13,23,27,28,36,45,49–54 Data from this study. Frequency and variance values were assigned to sample-collection places (dots). Population samples (geographically close) with less than five observations were pooled and the corresponding variance assigned to a middle position of the pooled sample locations. + Data from the literature.13,23,27,28,36,45,49–54

European Journal of Human Genetics
Discussion
Various episodes of population movement have affected southeast Europe, and the role of the Balkans as a long-standing gateway to Europe from the Near East is illustrated by the phylogenetic unification of Hgs I and J by the basal M429 mutation.1 This evidence of common ancestry suggests that ancestral JJ-M429* Y chromosomes probably entered Europe through the Balkan route sometime before the Last Glacial Maximum. They subsequently evolved into Hg J in the Middle East and Hg I in Europe in a typical disjunctive phylogeographic pattern. Such a geographic corridor is likely to have experienced additional subsequent gene flows, including the migration of agricultural colonists from the Middle East. Pottery is a useful proxy for the spread of farming both spatially and temporally. The first appearance of pottery in the Adriatic region was in Corfu at 6500 BC and reached the northern most Adriatic ~1000 years later.21 Its dispersal provides a comparative template for spatial and temporal patterns of Y chromosome Hg diversity observed in this area.

Hg J is most common (~50%) in the Middle East and Anatolia,27,29,47 with a spread zone spanning from northwest Africa to India.2,55 It has been related to different Middle Eastern migrations.12,56 In addition to Hg J-M410, Hg G-P15 chromosomes, which are also common in Anatolia,29 have been implicated in the colonization and subsequent expansion of early farmers in Crete, the Aegean and Italy.36,45,46 Earlier studies have concluded that the J-M410 sub-clades, J-DYS445-6 and J-M67, are linked to the spread of farming in the Mediterranean Basin,38,47 with a likely origin in Anatolia.29 Interestingly, J-DYS445-6 and J-M92 (a sub-lineage of M67), both have expansion times between 7000 and 8000 years ago (Table 1), consistent with the dating of the arrival of the first farmers to the Balkans.

The first detection of milk residue in ceramic pottery occurs in sites from northwest Anatolia 7000–8500 years ago,58 an age that approximates the Hg-expansion times. Although Hg J-M241 shows high variance in India,37 its place of origin is still uncertain. As J-M241 has older expansion times in Sicily, Apulia and Turkey (Table 2), it may have arrived in the Balkans from elsewhere.

On the other hand, the expansion times of Hg V13 (Table 3) are consistent with a late Mesolithic time frame. The Greek Mesolithic, although different in its material culture from the Natufian Mesolithic of the Levant, bears some resemblance to the Mesolithic of southern Anatolia.60 This archaeological congruence between the Mesolithic of the Balkans and southern Anatolia may mirror the similar E-V13 expansion times observed for Konya, Franchthi Cave and Macedonian Greece, all approximately 9000 years ago. Moreover, E-V13 YSTR-related data from Bulgaria and Macedonia,28 both with variances of 0.28, suggest an expansion time of approximately 10000 years ago. It is likely that the origin of V13 occurred somewhere within the zone of these sample collections. In addition, it is also worth noting that in the Anatolian region of supposed Einkorn wheat origin61 (region 5 of

### Table 1 Ages of microsatellite variation and mean variance of microsatellite loci within haplogroups

<table>
<thead>
<tr>
<th>Haplogroup (N)</th>
<th>Age of microsatellite variation(a) ± SE(b) (ky)</th>
<th>Mean variance(a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-P37*+DYS388-15 (28)</td>
<td>4.0 ± 1.7</td>
<td>0.107</td>
</tr>
<tr>
<td>I-M423 (219)</td>
<td>7.8 ± 2.3</td>
<td>0.198</td>
</tr>
<tr>
<td>E-V13 (92)</td>
<td>5.9 ± 1.7</td>
<td>0.155</td>
</tr>
<tr>
<td>J-M410* (23)</td>
<td>15.8 ± 3.8</td>
<td>0.410</td>
</tr>
<tr>
<td>J-M67 (42)</td>
<td>11.0 ± 2.4</td>
<td>0.287</td>
</tr>
<tr>
<td>J-M67* (32)</td>
<td>11.2 ± 2.9</td>
<td>0.298</td>
</tr>
<tr>
<td>J-M92 (10)</td>
<td>7.0 ± 1.9</td>
<td>0.187</td>
</tr>
<tr>
<td>J-DYS445-6(10)</td>
<td>7.7 ± 2.6</td>
<td>0.164</td>
</tr>
<tr>
<td>J-M12 (51)</td>
<td>7.4 ± 1.4</td>
<td>0.191</td>
</tr>
<tr>
<td>J-M12/102* (15)</td>
<td>12.3 ± 2.7</td>
<td>0.332</td>
</tr>
<tr>
<td>J-M241 (34)</td>
<td>4.8 ± 1.3</td>
<td>0.121</td>
</tr>
</tbody>
</table>

\(a\)Values computed over the following microsatellite loci: DYS19, DYS388, DYS389I, DYS390, DYS391, DYS392, DYS393, DYS439 and DYS460 (values of YCAa,b were excluded from the calculation because of the impossibility of assigning each allele to one or the other locus of this ambiguous system). Ages of microsatellite variation were estimated as reported in Sengupta et al.37 Owing to the small sample sizes, the estimates of the expansion time should be considered indicative.

\(b\)SE was computed over loci.

### Table 2 Ages of microsatellite variation and mean variance of microsatellite loci within J-M241 haplogroup in Turkey, the Balkans and Italy

<table>
<thead>
<tr>
<th>Region (N)</th>
<th>Age of microsatellite variation(a) ± SE(b) (ky)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turkey(c) (5)</td>
<td>10.1 ± 3.4</td>
</tr>
<tr>
<td>Albania(c) (8)</td>
<td>5.4 ± 2.3</td>
</tr>
<tr>
<td>FYROM(c) (9)</td>
<td>2.4 ± 1.2</td>
</tr>
<tr>
<td>Greece(c) (5)</td>
<td>2.9 ± 1.2</td>
</tr>
<tr>
<td>Central Italy(c) (5)</td>
<td>5.8 ± 1.4</td>
</tr>
<tr>
<td>Apulia(c) (5)</td>
<td>9.6 ± 3.4</td>
</tr>
<tr>
<td>Sicily(c) (11)</td>
<td>10.0 ± 5.4</td>
</tr>
</tbody>
</table>

\(a\)Values computed over the following microsatellite loci: DYS19, DYS388, DYS389I, DYS390, DYS391, DYS392, DYS393, DYS439 and DYS460 (values of YCAa,b were excluded from the calculation because of the impossibility of assigning each allele to one or the other locus of this ambiguous system). Ages of microsatellite variation were estimated as reported in Sengupta et al.37 Owing to the small sample sizes, the estimates of the expansion time should be considered indicative.

\(b\)SE was computed over loci.

\(c\)Sengupta et al.37 Owing to the small sample sizes, the estimates of the expansion time should be considered indicative.
Cinnioglu et al.\(^{29}\), only one V13 chromosome out of 43 is found (PA Underhill, unpublished data). Therefore, as no evidence at present supports the association of E-V13 Hg with the attested origin of farming in southeast Anatolia, the possibility of farming adoption by Balkan E-V13-associated people is plausible. The low E-V13 frequency in this region (Figure 4) indicates that if the original source of the E-M78 may have been the Levant and ultimately onwards to Asia Minor and Europe, where they each eventually differentiated into their regionally distinctive branches.

Differently from the earlier discussed Hgs, I-M423 represents the southeast European autochthonous clade of I-P37. Its distribution reaches Anatolia, where, however, it is only sporadically observed (2.6%, updated from Rootsi et al.\(^{37}\)). Owing to the small sample sizes, the estimates of the expansion time should be considered indicative.

Transition to farming in the Balkans

V Battaglia et al

Table 3 Ages of microsatellite variation and mean variance of microsatellite loci within E-V13 haplogroup in Turkey and Greece

<table>
<thead>
<tr>
<th>Region (N)</th>
<th>Age of microsatellite variation(^a) ± SE(^b) (ky)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Konya-Turkey(^c) (10)</td>
<td>9.4 ± 2.9</td>
</tr>
<tr>
<td>Macedonian Greece(^c) (8)</td>
<td>8.5 ± 4.5</td>
</tr>
<tr>
<td>Greece(^c) (8)</td>
<td>4.8 ± 1.6</td>
</tr>
<tr>
<td>Nea Nikomedeia (North Greece)(^d) (6)</td>
<td>8.6 ± 4.0</td>
</tr>
<tr>
<td>Seklo and Dimini (Central Greece)(^d) (20)</td>
<td>4.3 ± 1.8</td>
</tr>
<tr>
<td>Lerna and Franchthi Cave (South Greece)(^d) (20)</td>
<td>9.2 ± 4.3</td>
</tr>
</tbody>
</table>

\(^a\)Values computed over the following microsatellite loci: DYS19, DYS388, DYS389I, DYS390, DYS391, DYS392, DYS393, DYS439 and DYS460 (values of YCA\(^a\),b were excluded from the calculation because of the impossibility of assigning each allele to one or the other locus of this ambiguous system). Ages of microsatellite variation were estimated as reported in Sengupta et al.\(^{37}\). Owing to the small sample sizes, the estimates of the expansion time should be considered indicative.

\(^b\)SE was computed over loci.\(^{43,57}\)

\(^c\)Present paper.

\(^d\)King et al.\(^{38}\)

or not E-V13 and J-M241 participated in the same demography remains uncertain.

The presence of E-M78* Y chromosomes in the Balkans (two Albanians), previously described virtually only in northeast Africa, upper Nile,\(^{28,63}\) gives rise to the question of what the original source of the E-M78 may have been. Correlations between human-occupation sites and radiocarbon-dated climatic fluctuations in the eastern Sahara and Nile Valley during the Holocene\(^{64}\) provide a framework for interpreting the main southeast European centric distribution of E-V13. A recent archaeological study reveals that during a desiccation period in north Africa, while the eastern Sahara was depopulated, a refugium existed on the border of present-day Sudan and Egypt, near Lake Nubia, until the onset of a humid phase around 8500 BC (radiocarbon-calibrated date). The rapid arrival of wet conditions during this Early Holocene period provided an impetus for population movement into habitat that was quickly settled afterwards.\(^{64}\) Hg E-M78* representatives, although rare overall, still occur in Egypt, which is a hub for the distribution of the various geographically localized M78-related sub-clades.\(^{28}\) The northward-moving rainfall belts during this period could have also spurred a rapid migration of Mesolithic foragers northwards in Africa, the Levant and ultimately onwards to Asia Minor and Europe, where they each eventually differentiated into their regionally distinctive branches.

As reported earlier, both J-M12 and E-V13 radiation patterns of Cruciani et al.\(^{26}\) overlap geographically in the Balkans (Figure 4). Although J-M12 chromosomes were not genotyped for M241 by Cruciani et al.\(^{26}\) the low YSTR diversity observed suggests that these are predominantly M241 derivatives. The difference between E-V13 and J-M241 (Table 1) indicates that both E-V13 frequency and haplotype diversity would have been greater than J-M241 during either the early Neolithic or Bronze Age and the age of microsatellite variation observed dates to the Bronze Age, possibly reflect a more recent gene flow. The age of accumulated microsatellite variation associated with Hg I-M423 (Table 1) dates to around 8000 years ago (Early Holocene). Thus, although Hgs G and J mark the successful colonization and subsequent demic expansions of Neolithic pioneers to
these regions, consistent with a wave of advance,19,65 the widespread adoption of farming by Mesolithic hunter-gatherers in the Balkans and Central Europe is recorded in the autochthonous Hg I-M423.

These data indicate the complex interactions between farmers and foragers rather than the large-scale replacement of hunter-gatherers by pioneering agriculturalists during the spread from the Neolithic to the southeast Europe. The data also indicate that I-M423 and probably also E-V13 representatives would have been well established in the Balkans before the arrival of a nucleus of pioneering agriculturalists.

Thus, unlike Crete, southern and central Italy and the southern Caucasus, the cultural transmission of the Neolithic package played an important role. Either the initial G and J2 Hg agriculturalists who colonized the Balkans at first flourished but later diminished in a similar manner to that proposed regarding the Linearbandkeramik in central Europe66 or the package was rapidly and robustly adopted by local Mesolithic people in the southern Balkans (plausibly characterized by E-V13), who underwent a demic expansion and a subsequent range expansion to the eastern Adriatic. These former foragers who had recently acquired the Neolithic tradition participated in ‘leapfrog’ colonizations up the Adriatic, where they eventually transmitted agricultural practices to resident Mesolithic populations represented by I-M423 Hg chromosomes.

Interestingly, the derived Y-chromosome scenario strongly recalls the fourth PC synthetic map of Europe calculated on gene frequencies at 95 nuclear loci,11 which displays a centre in the southern Balkans and a large surrounding area that terminates with a ‘propagule’ to the northeast of the central Balkans. On the basis of this observation, our assumptions could provide a possible interpretation of the described expansion centred in the southern Balkans.

Conclusion
This study provides a model that elevates the role of migratory foragers with remote eastern Saharan ancestry who, once established in Asia Minor with their own derived genetic signature, were destined to become the earliest converts to farming and the adherents of its further spread into Europe. Such an interpretation finds support in the ‘dispersal model’ of impressed Ware in which the ‘Neolithic package’ was acquired by native groups and subsequently diffused by interactions between farmers and foragers,67 although southeast Europe shows considerable archaeological evidence of the Neolithic transition.20 Our Y-chromosome results provide biological evidence of complexity21 in the transition to farming in terms of the contrasting influences of pioneering agriculturalists and Mesolithic foragers.

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, from Bosnia-Herzegovina, and Professor Simun Andjeljmovic and his team from the University Hospital of Split, from Croatia. We are also grateful to Franco Lo Curto for having provided us with the Italian samples. We thank the three anonymous reviewers for useful comments and suggestions. This research was supported by the Italian Ministry of the Foreign Affairs (to OS), Compagnia di San Paolo (to AT and OS), the Italian Ministry of the University: Progetti Ricerca Interesse Nazionale 2007 (to AT and OS) and Estonian Science Foundation Grant no. 7445 (to SR).

References
Transcription to farming in the Balkans
V Battaglia et al


Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)