Mol Biol Rep (2012) 39:7727–7741
DOI 10.1007/s11033-012-1610-3

Croatian national reference Y-STR haplotype database

Gordan Mršić · Branka Grško vić · Andro Vrdoljak · Maja Popović · Ivica Valpotić · Šimun Anđelinović · Vlastimil Stenzl · Edvard Ehler · Ludvik Urban · Gordana Lacković · Peter Underhill · Dragan Primorac

Received: 23 February 2011 / Accepted: 7 February 2012 / Published online: 6 March 2012

Abstract A reference Y-chromosome short tandem repeat (STR) haplotype database is needed for Y-STR match interpretation as well as for national and regional characterization of populations. The aim of this study was to create a comprehensive Y-STR haplotype database of the Croatian contemporary population and to analyze substructure between the five Croatian regions. We carried out a statistical analysis of the data from previously performed genetic analyses collected during routine forensic work by the Forensic Science Centre "Ivan Vučetić". A total of 1,100 unrelated men from eastern, western, northern, southern and central Croatia were selected for the purpose of this study. Y-STRs were typed using the AmpFISTR Yfiler PCR amplification kit. Analysis of molecular variance calculated with the Y chromosome haplotype reference database online analysis tool included 16 population samples with 20,247 haplotypes. A total of 947 haplotypes were recorded, 848 of which were unique (89.5%). Haplotype diversity was 0.998, with the most frequent haplotype found in 9 of 1,100 men (0.82%). Locus diversity varied from 0.266 for DYS392 to 0.868 for DYS385. Discrimination capacity was 86.1%. Our results

Electronic supplementary material The online version of this article (doi:10.1007/s11033-012-1610-3) contains supplementary material, which is available to authorized users.

G. Mršić (✉) · B. Grško vić · A. Vrdoljak
Forensic Science Centre "Ivan Vučetić", Ministry of Interior, Ilica 335, 10000 Zagreb, Croatia
e-mail: gmrsic@unist.hr; gmrsic@mup.hr

G. Mršić · B. Grško vić · A. Vrdoljak · M. Popović · I. Valpotić · Š. Anđelinović
University Center for Forensic Sciences, University of Split, Split, Croatia

M. Popović · I. Valpotić
Department of Biology, Faculty of Veterinary Medicine, University of Zagreb, Zagreb, Croatia

Š. Anđelinović
Department of Pathology and Forensic Medicine, Split University Hospital and School of Medicine, Split, Croatia

V. Stenzl
Department of Forensic Genetics, Institute of Criminalistics, Prague, Czech Republic

E. Ehler
Department of Biology and Environmental Studies, Faculty of Education, Charles University in Prague, Prague, Czech Republic

L. Urban
Faculty of Mathematics and Physics, Charles University in Prague, Prague, Czech Republic

G. Lacković
Division of Biology, Faculty of Science, University of Zagreb, Zagreb, Croatia

P. Underhill
Department of Genetics, Stanford University, School of Medicine, Palo Alto, CA, USA

D. Primorac
Medical School, University of Split, Split, Croatia

D. Primorac
Medical School, University of Osijek, Osijek, Croatia

D. Primorac
Eberly College of Science, Penn State University, University Park, PA, USA

D. Primorac
University of New Haven, New Haven, CT, USA

suggested high level of similarity among regional subpopulations within Croatia, except for mildly different southern Croatia. Relative resemblance was found with Bosnia and Herzegovina and Serbia. Whit Atheys' Haplogroup Predictor was used to estimate the frequencies of Y-chromosome haplogroups. I2a, R1a, E1b1b and R1b haplogroups were most frequent in all Croatian regions. These results are important in forensics and contribute to the population genetics and genetic background of the contemporary Croatian population.

Keywords Y-chromosome · Short tandem repeat · Population genetics · Haplotypes · Haplogroups · Croatian population

Introduction

Croatia is located in the south-eastern Europe, sharing boundaries with Slovenia on the north-west, Hungary on the north, Serbia on the east, Montenegro on the south-east, Bosnia and Herzegovina on the south and east, and the Adriatic Sea on the west. Genetic structure of contemporary populations is interesting for a number of reasons, including genetic epidemiology, forensics, and enlightenment of the historical events. The use of Y-STRs has a central role in a number of such attempts, since it provides a good proxy of paternal lineage [1]. In forensics, reference Y-STR database within local population is needed for Y-STR match interpretation [2]. For this reason, Ministry of Interior of the Republic of Croatia created the reference Croatian Y-STR haplotype database. Variant alleles which contain incomplete repeat units are increasing the power of discrimination in DNA comparisons and can reveal phylogenic substructure within the Y-chromosome phylogeny [3].

Creation of a national Y-STR haplotype database is one of the principal steps in a detailed genetic characterization. This is interesting not only for population genetics and forensics, but also for historical analyses, since genetic material can be used to infer information on demographic events.

Several previous studies were published in the Y-STRs in Croatia, mostly focusing on regional samples or rather small number of total samples involved in the study [4–9]. Recently, we announced population data of Y-STR haplotyping in eastern, northern and central Croatia [10–12] to emphasize the importance of Y-STR variations on regional level in Croatia.

The aim of this study was to create a large and comprehensive Y-STR haplotype database of the Croatian contemporary population and to analyze intrapopulation variability between the five Croatian regions. This intrapopulation variability may be mirrored in Y-chromosomal haplogroups distribution.

Subjects and methods

Study sample

Statistical analysis of the data from previously performed genetic analyses collected during routine forensic work by the Forensic Science Centre "Ivan Vučetić" was performed in this study. Based on geographical and historical information we divided Croatia into five regions: eastern, western, southern, northern and central (Fig. 1). A total of 1,100 samples from eastern Croatia [$n = 220$] (Virovitičko-podravska, Požeško-slavonska, Brodsko-posavska, Osječko-baranjska and Vukovarsko-srijemska counties), western Croatia [$n = 220$] (Primorsko-goranska, Ličko-senjska and Istarska counties), northern Croatia [$n = 220$] (Krapinsko-zagorska, Varaždinska, Koprivničko-križevačka and Međimurska counties), southern Croatia [$n = 220$] (Zadarska, Šibensko-kninska, Splitsko-dalmatinska and Dubrovačko-neretvanska counties) and central Croatia [$n = 220$] (Zagrebačka, Sisačko-moslavačka, Karlovačka, Bjelovarsko-bilogorska counties and city of Zagreb) were collected. The donors had no known familial relationships among one another and were of sufficient quality to be included in the statistical analysis. The study was approved by the Ethics Committee of the Institute for Medical Research and Occupational Health, Zagreb, Croatia.

DNA analysis

Genomic DNA from all samples of the materials expertise was extracted from FTA cards (Whatman, Maidstone, Kent, UK) using Chelex [13]. The AmpFISTR Yfiler PCR amplification kit was used for the amplification of 17 Y-STRs: DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635 and GATAH4 according to manufacturer's instruction (Applied Biosystems Foster City, California, USA). Y-STR amplification products were analyzed on 3130xl Genetic Analyzer (Applied Biosystems). Analysis of the data was performed using Genemapper® software (version 3.2, Applied Biosystems). Amplicon sizing was performed using an internal size standard (GeneScan-500 LIZ, Applied Biosystems), and the amplicons were compared with the AmpFlSTR Yfiler allelic ladder for unambiguous allele designation.

Quality control

Blind testing and evaluation was performed according to the quality assessment scheme provided by the Y Chromosome Haplotype Reference Database (YHRD, www.yhrd.org) [14] and a certificate for successful testing was issued at November 23, 2009 for the Forensic Science Centre "Ivan

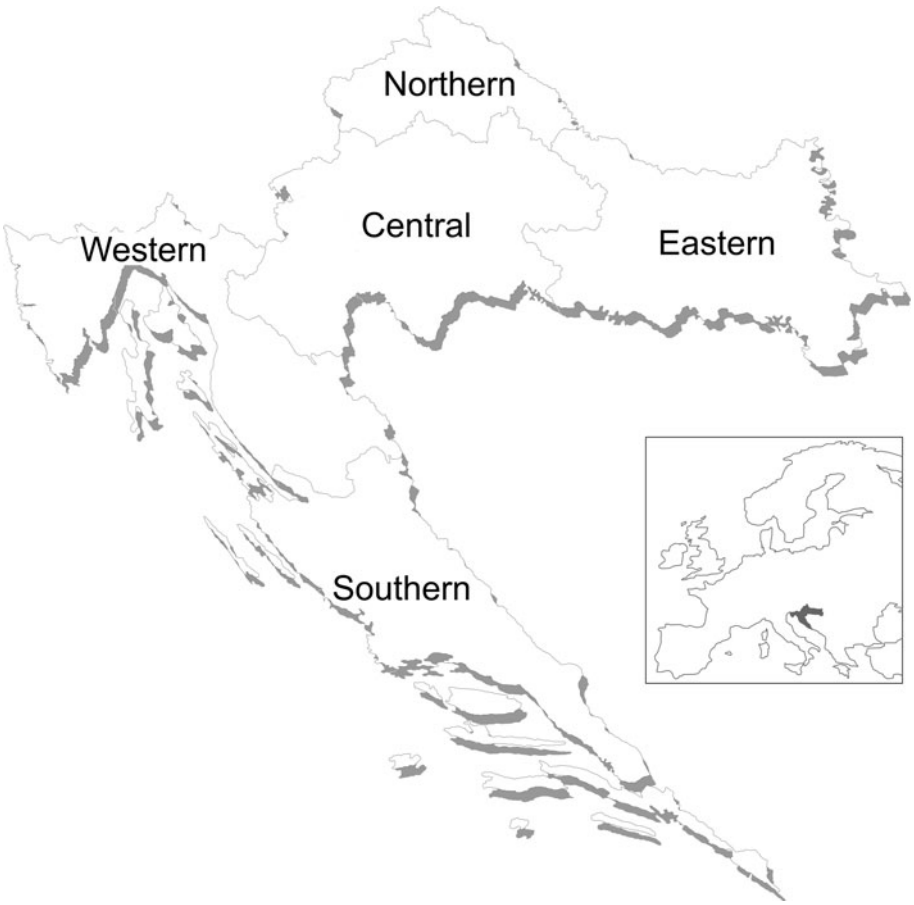

Fig. 1 Map of Croatian regions

Vučetić". The contributing Centre (YC000198) received the following accession numbers after evaluation of the data: central (YA003593), eastern (YA003594), northern (YA003595), southern (YA003596) and western (YA003601) Croatia.

Statistical analysis

Allelic and haplotype frequencies were estimated by direct counting. Locus and haplotype diversities were calculated using Arlequin 3.1 software [15]. Furthermore, for analysis of molecular variance (AMOVA) Professor Lutz Roewer and Sascha Willuweit, MSc used the AMOVA tool provided by the YHRD [14]. Due to the limited number of markers reported in other publicly available samples, the comparative analysis was performed on a minimal European Y-STR haplotype comprising nine loci: DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 and DYS385ab [16]. A total of 16 population samples with 20,247 haplotypes were included in this analysis. The samples came from the following areas (the YHRD designation is given in parentheses): 150 haplotypes from Croatia including Zagreb, Croatia [Croatian, $n = 150$], 1,100 haplotypes from Croatia including central Croatia [Croatian, $n = 220$], east Croatia [Croatian, $n = 220$], north Croatia [Croatian, $n = 220$], south Croatia [Croatian, $n = 220$], west Croatia [Croatian, $n = 220$]; 141 haplotypes from Albania including Albania [Albanian, $n = 111$], Tirana, Albania [Albanian, $n = 30$]; 485 haplotypes from Austria including Graz [Austrian, $n = 65$], Oberöstereich [Austrian, $n = 178$], Vienna [Austrian, $n = 66$], Salzburg [Austrian, $n = 176$]; 100 haplotypes from Bosnia and Herzegovina including Doboj-Banja Luka-Bjeljina [Bosnian, $n = 31$], Mostar, [Bosnian, $n = 34$], Sarajevo [Bosnian, $n = 35$]; 122 haplotypes from Bulgaria [Bulgarian]; 2,226 haplotypes from Czech Republic including Prague [Czech, $n = 645$], Southern Bohemia [Czech, $n = 111$], Karlovy Vary [Czech, $n = 31$], Plzen [Czech, $n = 62$], Liberec [Czech, $n = 45$], Usti nad Labem [Czech, $n = 87$], Central Bohemia [Czech, $n = 455$], Hradec Kralove [Czech, $n = 49$], Pardubice [Czech, $n = 84$], Southern Moravia [Czech, $n = 216$], Moravia-Silesia [Czech, $n = 115$], Zlin [Czech, $n = 64$], Olomouc [Czech, $n = 53$], Vysocina [Czech, $n = 40$], Czech Republic [Czech, $n = 169$]; 6,119 haplotypes from Germany including Cologne [German, $n = 738$], Berlin [German, $n = 657$], Bonn [German, $n = 90$], Chemnitz [German, $n = 833$], Greifswald [German, $n = 208$], Hamburg [German, $n = 275$], Freiburg [German, $n = 433$],

Halle [German, $n = 234$], Dresden [German, $n = 88$], Leipzig [German, $n = 811$], Mainz [German, $n = 104$], Magdeburg [German, $n = 283$], Muenster [German, $n = 196$], Munich [German, $n = 281$], Rostock [German, $n = 243$], Stuttgart [German, $n = 613$], Palota, Bihor, Romania [German, $n = 32$]; 515 haplotypes from Hungary including Budapest [Hungarian, $n = 200$], Hungary [Hungarian, $n = 215$], Szeged [Hungarian, $n = 100$]; 302 haplotypes from Macedonia including Macedonia [Macedonian, $n = 250$], Skopje [Macedonian, $n = 52$]; 2,023 haplotypes from North Italy including Bologna [Italian, $n = 51$], Brescia [Italian, $n = 106$], Liguria [Italian, $n = 81$], Modena [Italian, $n = 130$], Lombardy [Italian, $n = 182$], Marche [Italian, $n = 162$], Umbria [Italian, $n = 51$], Verona [Italian, $n = 153$], Tuscany [Italian, $n = 218$], Veneto [Italian, $n = 120$], Ravenna [Italian, $n = 384$], Val Marecchia [Italian, $n = 65$], Latium [Italian, $n = 222$], Rimini [Italian, $n = 98$]; 4,131 haplotypes from Poland including Bydgoszcz [Polish, $n = 411$], Bialystok [Polish, $n = 185$], Krakow [Polish, $n = 207$], Gdansk [Polish, $n = 942$], Limanowa [Polish, $n = 53$], Nowy Sacz [Polish, $n = 114$], Lublin [Polish, $n = 246$], South Eastern Poland [Polish, $n = 161$], Szczecin [Polish, $n = 105$], Suwalki [Polish, $n = 82$], Wroclaw [Polish, $n = 715$], Warsaw [Polish, $n = 393$], Nowy Targ [Polish, $n = 52$], Zakopane [Polish, $n = 7$], Southern Poland [Polish, 380], Northern Poland [Polish, $n = 78$]; 399 haplotypes from Romania including Constanta [Romanian, $n = 36$], Romania [Romanian, $n = 104$], Moldavia [Romanian, $n = 40$], Transylvania, Romania [Romanian, $n = 14$], Ploiesti [Romanian, $n = 36$], Wallachia [Romanian, $n = 96$], Oradea, Bihor [Romanian, $n = 73$]; 215 haplotypes from Serbia including Novi Sad [Serbian, $n = 215$]; 1,108 haplotypes from Slovakia including Bratislava [Slovakian, $n = 164$], Eastern Slovakia [Slovakian, $n = 629$], Slovakia [Slovakian, $n = 80$], Sinteu, Bihor, Romania [Slovakian, $n = 35$], Novi Sad, Serbia [Slovakian, $n = 200$]; 931 haplotypes from Italy Alcamo [Italian, $n = 23$]; Caccamo [Italian, $n = 19$], Puglia [Italian, $n = 70$], Pantelleria, Italy [Italian, $n = 21$], Northern Sardinia [Italian, $n = 100$], Ragusa [Italian, $n = 29$], Mazara del Vallo [Italian, $n = 25$], Santa Ninfa [Italian, $n = 34$], Troina [Italian, $n = 21$], Trapani [Italian, $n = 32$], Sciacca [Italian, $n = 20$], Sicily [Italian, $n = 314$], Piazza Armerina [Italian, $n = 30$], Reggio di Calabria, Italy [Italian, $n = 97$], Cosenza, Italy [Italian, $n = 37$], Catanzaro [Italian, $n = 59$], and 180 haplotypes from Slovenia including Ljubljana [Slovenian, $n = 180$] [17–57]. In this study we calculated population pair-wise genetic distances (R_{ST}). This is an extension of the commonly used F_{ST} measure [58], defined as $R_{ST} = (S_b - S_W)/S_b$, where S_W is the sum over all loci of twice the weighted mean of the within-population variances V(A) and V(B), and S_b is the sum over all loci of twice the variance V(A + B) of the combined population [58]. R_{ST} values were calculated by AMOVA using on-line tools of the YHRD, P values were obtained with 10,000 permutations (significance set at $P < 0.05$).

We had estimated the Y-chromosomal haplogroups from a set of acquired Y-STR data. Regarding the fact that Y-chromosomal single-nucleotidepolymorphism (Y-SNP) data had not been available for this study all predictions of haplogroups were performed by use of the web-accessible program, Whit Atheys' Haplogroup Predictor (www.hprg.com/hapest5/), which is based on a Bayesian allele frequency approach. We considered all limitations and biases [59–61]. Therefore, we included in our study only the samples with haplogroup probability over 90%. This is very often the only approach in forensic practice to found out the supplementary information for particular unknown haplotype. The haplogroup estimation was performed for every haplotype and the haplogroup frequencies were estimated by direct counting.

Median-Joining Network (MJN) was constructed with Network 4.60 software (www.fluxus-engineering.com) [62, 63]. We utilized all 17 loci (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635 and GATAH4) for the MJN analysis. For the complete set of 1,100 Croatian samples, firstly, we reduced the amount of haplotypes in the computation by star contraction pre-processing algorithm ($\Delta = 5$). Subsequently, the reduced median algorithm ($r = 2$, frequency >1 criterion active) and median-joining algorithm ($\varepsilon = 0$) were applied. Before drawing the network, maximum parsimony post-processing procedure was performed. MJN analysis for the single haplogroups included only the reduced median (r = 2) algorithm.

Software Splitstree 4.10 (www.splitstree.org) was used for constructing the Neighbor-Joining (NJ) phylogenetic tree [64]. As input data we used the matrix of R_{ST} pair-wise distances of the Croatian population focused on in this study (central, western, northern, southern and eastern) and neighboring populations as described previously, including population from Zagreb, Croatia.

Results

A total of 1,100 samples from five regions of Croatia were included in the analysis. Locus diversity varied from 0.266 for DYS392 to 0.868 for DYS385. Variant alleles have been found at DYS458 locus. Allele frequencies and locus diversity values are shown in Table 1. A total of 947 haplotypes were recorded, 848 of which were unique (89.5%). Total haplotype diversity was 0.998, with the most common recorded haplotype found in 9 of 1,100 men (0.82%). Discrimination capacity was 86.1%.

Table 1 Allele frequencies and locus diversity values at 17 Y-STR loci in Croatian population

Allele	STR locus on Y chromosome																
	DYS456	DYS389I	DYS390	DYS389II	DYS458	DYS19	DYS393	DYS391	DYS439	DYS635	DYS392	YGATAH4	DYS437	DYS438	DYS448	Genotype	DYS385[a] (%)
8								0.09	0.36							10–10	0.18
9								1.55	0.45			0.45		9.00		10–13	0.09
10								49.73	13.00		0.55	2.64		59.91		10–14	2.00
11		0.27					0.45	47.27	26.27		85.09	53.09		23.64		10–15	0.09
12	0.73	14.09			0.09	0.09	12.27	1.36	32.09		4.18	36.82	0.09	7.27		10–17	0.09
13	1.73	70.09			0.73	9.27	79.73		24.73		8.27	6.09	0.18	0.18		10–17.1	0.09
14	10.45	15.27			1.82	19.45	7.27		2.73		1.18	0.91	42.82			10–20	0.09
15	53.55	0.27			23.18	24.00	0.27		0.27		0.27		46.91			11–11	0.91
16	20.55				24.91	37.45			0.09		0.45		9.27			11–12	0.27
17	11.00				30.27	9.45							0.73		0.09	11–13	3.36
18	2.00				14.09	0.27									1.36	11–14	17.09
19					3.00					0.09					41.09	11–15	3.63
20					0.45					4.18					48.82	11–16	0.73
21			0.45							15.27					7.36	11–18	0.09
22			7.64							22.00					1.09	12–12	0.18
23			13.73							48.36					0.18	12–14	1.73
24			55.82							8.73						12–15	0.82
25			20.73							1.36						12–16	0.27
26			1.45													12–17	0.09
27			0.18	0.82												12–18	0.45
28				9.09												13–13	0.27
29				17.45												13–14	4.18
30				31.00												13–15	2.27
31				34.00												13–16	1.00
32				7.18												13–17	0.64
33				0.45												13–17.2	0.27
12.2					0.09											13–18	0.18
15.1					0.09											13–19	0.36
16.2					0.09											13–20	0.27
17.2					0.36											14–14	3.36
18.2					0.27											14–15	30.00
19.2					0.55											14–16	4.55
15,16						0.09										14–17	1.55
																14–18	0.82
																14–19	0.27
																15–15	3.09
																15–16	0.36
																15–17	2.00
																15–18	0.82
																15–19	0.91
																15–20	0.36

Table 1 continued

Allele	STR locus on Y chromosome																	
	DYS456	DYS389I	DYS390	DYS389II	DYS458	DYS19	DYS393	DYS391	DYS439	DYS635	DYS392	YGATAH4	DYS437	DYS438	DYS448	Genotype	DYS385[a] (%)
																I6–16	0.36
																I6–17	1.00
																I6–18	4.73
																I6–19	0.73
																I6–20	0.27
																I7–17	0.91
																I7–18	1.46
																I7–19	0.27
																I8–19	0.09
																I4.2–18	0.09
																I7.1–17	0.09
																I7.1–17.1	0.09
																I7.1–18	0.09
LDV	0.645	0.466	0.620	0.742	0.770	0.748	0.342	0.526	0.747	0.682	0.266	0.574	0.581	0.567	0.579		0.868

Bold values were used to emphasize the variant alleles, which were described for the first time in Croatian population

LDV locus diversity value

[a] The table shows allele frequencies for each investigated locus except DYS385 for which genotype frequencies are reported as percentages. These genotype frequencies were calculated for the combination of two alleles

Y-STR haplotypes in five Croatian regions are presented in Supplementary Table 1.

A pair-wise comparison suggested a strong similarity of Croatian subpopulations (originating from the investigated regions), with the exception of southern Croatia, which was mildly different from other regions. In terms of neighboring countries, we recorded relative similarity with the populations from Bosnia and Herzegovina as well as from Serbia. While analyzing available haplotypes within YHRD, Professor Lutz Roewer and Sascha Willuweit, MSc found a strong difference between Croatian and Slovenian samples [14]. However, in the same analysis they noted a strong delineation between Croatian subpopulations and other Slavic populations (notably those from the Czech Republic, Poland and Slovakia, i.e., western Slavs) (Table 2; Fig. 2).

Single haplogroups had slightly different extent of haplotypes diversifications except of haplogroup I2a (xI2a1), where a remarkable difference in frequencies was observed between northern and southern haplotypes. I2a haplogroup seemed to be the most homogenous haplogroup irrespectively of regional data source (Table 3). The most frequent haplogroups in all five Croatian regions are shown to be I2a, R1a, E1b1b and R1b (Table 3). Haplogroups were not distributed equally in all five Croatian subpopulations. The most different subpopulation from all the others was the southern Croatian. Frequencies of I2a and E1b1b were statistically significantly different between the southern and eastern, western, northern and central Croatia subpopulations (χ^2 test, $P < 0.05$). Southern Croatian subpopulation was also significantly different in R1b frequency from central Croatia and northern Croatia and in R1a frequency from northern Croatia.

The displayed MJN of all pooled Croatian haplotypes (Fig. 3) did not reflect simple structure based on the Y-chromosomal haplogroups. Only the uppermost cluster consisted mainly of I2a haplotypes, while other clusters are composed of haplotype mixture from different haplogroups. The inner structure of each Y-chromosomal haplogroup is displayed in Figures S1–S8. It is difficult to describe the proper magnitude of the sources of the displayed population structure in Fig. 3. Therefore, we can more easily estimate which sources were not detected in the analysis. The latest is the effect of geographic substructure that we did not detect in the MJN analysis. None of the clusters showed remarkable overrepresentation of the haplotypes from one of the five Croatian subpopulations. Thus, we had concluded that Fig. 3 reflected the genetic variability of Croatian Y-chromosomal diversity based both on fast and slow mutating Y-STR and Y-SNP polymorphisms.

Subsequently the haplotypes assigned to single haplogroups were analyzed in more detail (unreduced MJN

Table 2 Analysis of molecular variance pair-wise distances based on R_{ST} values between the Croatian subpopulations in this study and selected comparison populations

Population samples[a] Population samples[b]	Central Croatia, Croatia (Croatian)	East Croatia, Croatia (Croatian)	North Croatia, Croatia (Croatian)	South Croatia, Croatia (Croatian)	West Croatia, Croatia (Croatian)	Zagreb, Croatia (Croatian)	Albanian	Austrian	Bosnian	Bulgarian	Czech
Central Croatia, Croatia [Croatian]	–	0.8105	0.6582	0.0322	0.6352	0.9354	0.0001	0.0000	0.0446	0.0001	0.0000
East Croatia, Croatia [Croatian]	−0.0027	–	0.2100	0.1037	0.9564	0.7050	0.0001	0.0000	0.1719	0.0010	0.0000
North Croatia, Croatia [Croatian]	−0.0021	0.0018	–	0.0044	0.1296	0.5236	0.0000	0.0000	0.0098	0.0002	0.0000
South Croatia, Croatia [Croatian]	0.0084	0.0042	0.0171	–	0.0699	0.0522	0.0000	0.0000	0.0195	0.0000	0.0000
West Croatia, Croatia [Croatian]	−0.0019	–0.0036	0.0035	0.0055	–	0.6297	0.0000	0.0000	0.1218	0.0005	0.0000
Zagreb, Croatia [Croatian]	−0.0043	−0.0028	−0.0015	0.0085	−0.0023	–	0.0007	0.0000	0.0621	0.0008	0.0000
Albanian	0.0539	0.0473	0.0598	0.0904	0.0510	0.0490	–	0.0000	0.0148	0.1331	0.0000
Austrian	0.0761	0.0902	0.0682	0.1354	0.0942	0.0746	0.1088	–	0.0000	0.0000	0.0000
Bosnian	0.0125	0.0040	0.0226	0.0162	0.0061	0.0118	0.0295	0.1362	–	0.0092	0.0000
Bulgarian	0.0394	0.0312	0.0475	0.0717	0.0325	0.0383	0.0063	0.0977	0.0265	–	0.0000
Czech	0.0860	0.1066	0.0697	0.1335	0.1113	0.0877	0.1818	0.0372	0.1591	0.1678	–
German	0.1070	0.1267	0.0926	0.1623	0.1296	0.1054	0.1679	0.0105	0.1773	0.1544	0.0185
Hungarian	0.0189	0.0301	0.0120	0.0572	0.0322	0.0203	0.0796	0.0235	0.0640	0.0644	0.0289
Macedonian	0.0357	0.0257	0.0454	0.0508	0.0262	0.0341	0.0084	0.1217	0.0101	–0.0011	0.1851
North Italian	0.0938	0.1050	0.0844	0.1500	0.1090	0.0924	0.0908	0.0135	0.1364	0.0875	0.0651
Polish	0.1409	0.1652	0.1215	0.1780	0.1709	0.1454	0.2825	0.1221	0.2244	0.2655	0.0270
Romanian	0.0150	0.0113	0.0195	0.0357	0.0106	0.0154	0.0229	0.0794	0.0138	0.0019	0.1286
Serbian	0.0125	0.0061	0.0186	0.0281	0.0072	0.0125	0.0177	0.1005	0.0011	0.0046	0.1425
Slovakian	0.0461	0.0626	0.0344	0.0791	0.0666	0.0472	0.1510	0.0527	0.1074	0.1363	0.0113
Slovenian	0.0690	0.0879	0.0550	0.1193	0.0966	0.0695	0.1618	0.0533	0.1516	0.1711	0.0067
South Italian	0.0584	0.0594	0.0569	0.1002	0.0587	0.0575	0.0295	0.0460	0.0701	0.0169	0.1218

Population samples[a] Population samples[b]	German	Hungarian	Macedonian	North Italian	Polish	Romanian	Serbian	Slovakian	Slovenian	South Italian
Central Croatia, Croatia [Croatian]	0.0000	0.0003	0.0000	0.0000	0.0000	0.0026	0.0154	0.0000	0.0000	0.0000
East Croatia, Croatia [Croatian]	0.0000	0.0000	0.0001	0.0000	0.0000	0.0068	0.0616	0.0000	0.0000	0.0000
North Croatia, Croatia [Croatian]	0.0000	0.0043	0.0000	0.0000	0.0000	0.0013	0.0037	0.0000	0.0000	0.0000
South Croatia, Croatia [Croatian]	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000
West Croatia, Croatia [Croatian]	0.0000	0.0000	0.0004	0.0000	0.0000	0.0072	0.0502	0.0000	0.0000	0.0000
Zagreb, Croatia [Croatian]	0.0000	0.0008	0.0000	0.0000	0.0000	0.0060	0.0235	0.0000	0.0000	0.0000
Albanian	0.0000	0.0000	0.0591	0.0000	0.0000	0.0022	0.0131	0.0000	0.0000	0.0004
Austrian	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Bosnian	0.0000	0.0000	0.0602	0.0000	0.0000	0.0223	0.2920	0.0000	0.0000	0.0000
Bulgarian	0.0000	0.0000	0.4661	0.0000	0.0000	0.2156	0.1396	0.0000	0.0000	0.0043

Table 2 continued

Population samples[a] Population samples[b]	German	Hungarian	Macedonian	North Italian	Polish	Romanian	Serbian	Slovakian	Slovenian	South Italian
Czech	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0171	0.0000
German	–	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Hungarian	0.0384	–	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000
Macedonian	0.1836	0.0765	–	0.0000	0.0000	0.0201	0.1624	0.0000	0.0000	0.0000
North Italian	0.0275	0.0433	0.1172	–	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Polish	0.0828	0.0863	0.2741	0.1619	–	0.0000	0.0000	0.0000	0.0000	0.0000
Romanian	0.1291	0.0378	0.0074	0.0824	0.2119	–	0.3228	0.0000	0.0000	0.0000
Serbian	0.1494	0.0476	0.0024	0.1009	0.2209	0.0003	–	0.0000	0.0000	0.0000
Slovakian	0.0457	0.0174	0.1437	0.0859	0.0278	0.0949	0.1019	–	0.2328	0.0000
Slovenian	0.0368	0.0265	0.1649	0.0873	0.0213	0.1180	0.1296	0.0010	–	0.0000
South Italian	0.0956	0.0469	0.0391	0.0360	0.2264	0.0242	0.0379	0.1134	0.1216	–

[a] *P* values are shown above the diagonal and R_{ST} values below

[b] Names in parentheses refer to the YHRD designations (www.yhrd.org)

graphs) to found out intra haplogroup diversity. Some very rare European haplogroups (Q,T) were detected in the sample that are very useful for forensic application.

To examine the substructure of Croatian population and further investigate its connections with European populations, we constructed the phylogenetic tree based on R_{ST} pair-wise genetic distances with NJM. The topology of NJ tree was in concordance with previous genetic studies and known historic layout of European populations.

Length of the branches correlated with genetic distances between particular populations (Fig. 4).

The mutual relationships between populations were displayed without the dependence on the genetic distances in the cladogram (Fig. 5).

The topology of NJ tree also confirmed the substructure in Y-chromosomal variation in Croatian population. The major dichotomy in the tree divided the Croatian subpopulations into southern and central European clusters. The second cluster was further divided into Germanic populations (German, Austrian, and North Italian), Slavic populations (Czech, Polish, Slovenian and Slovak) and Hungarian. Interestingly, subpopulation from northern Croatia was, regarding to the R_{ST} genetic distance, closer to this cluster than to the rest of Croatia.

Southern European cluster consisted mostly of the populations with putative Illyrian (Albanians, Serbians, Bosnians and disputably Macedonians) and Roman (Latin) roots (Southern Italians and Romanians, although here is the influence of Dacians probably more prevalent) [65, 66]. Subpopulations from eastern, western and southern Croatia were close to this cluster, but at the same time exhibited a degree of isolation from it.

Remarkable is the position of central Croatian subpopulation as a part of the deepest dichotomy of the NJ tree, and although it lay closer to the northern populations cluster, it was isolated both from northern as well as from eastern, southern and western Croatian subpopulations. It is positioned exactly on the boundary between the northern Croatian subpopulation with its affinity to Slavic/Germanic cluster and the rest of the Croatian subpopulations with their affinity to Balkan/southern European cluster.

Discussion

The results of this study are based on the first attempt to create a national reference Y-STR haplotype database on a sufficient number of samples for both national and regional characterization.

We do acknowledge the previous studies in this light and should also state that the results of this study show a strong similarity with the existing data, suggesting that the results are likely valid. This paper follows the

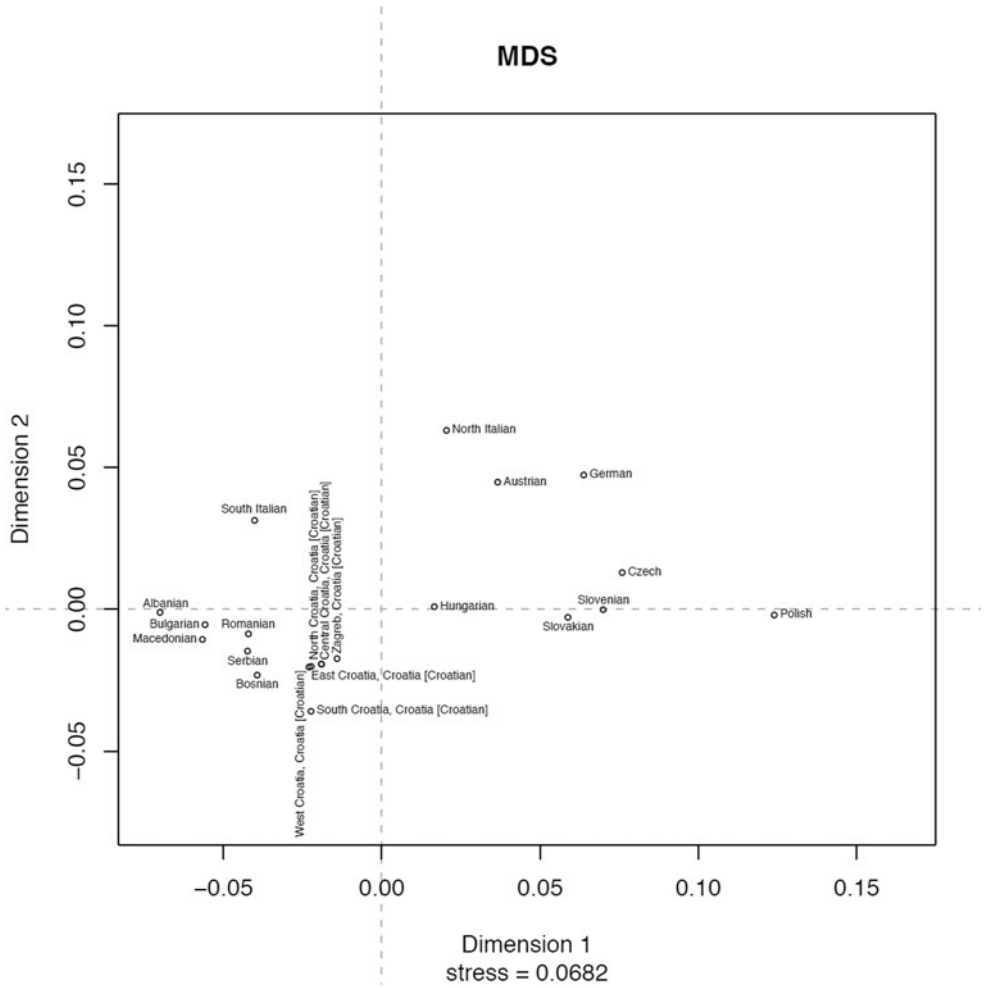

Fig. 2 Multidimensional scaling (MDS) based on pair-wise R_{ST} genetic distances between Croatia and neighboring countries

recommendations of the ISFG on the use of Y-STRs in forensic analysis [67]. Y-STR population data have been submitted to the YHRD as stated in "Subjects and methods" section.

In population genetics the sample size does not confirm the quality of population sampling [14, 17, 25, 68]. Therefore, the important aspect of population sampling is the good representation of different population specific haplotypes. On the basis of these facts we created our database. The high value of haplotype diversity (0.998) confirms the sampling quality. Moreover, large data sets typically identify a larger number of rare alleles as more individuals in a population are included in the analysis, as was shown in our case for DYS458 locus. At this locus we reported for the first time variant alleles 12.2, 16.2, 17.2 and 19.2 in eastern [10], 17.2, 18.2 and 19.2 in central [12] and 15.1 in southern Croatia.

Barać et al. [6] published population genetic study, where eight Y-STRs were analyzed in the sample of 457 participants; 109 from mainland Croatia and the rest of them from four Adriatic islands; Krk, Brač, Hvar and Korčula. Our results are in concordance to those published by Barać et al. concerning the allelic frequency distribution in Croatia, when taking into account the identical ones, except DYS391. Moreover, they also concluded that majority of their loci showed the similar allelic frequency distribution to those published for other European populations [16]. The exception was found at DYS19 and DYS389II loci. Allele 16 at DYS19 locus was the most frequent in Barać et al. as well as in our study, opposite to allele 14 in Europe [16]. In addition, both Croatian studies showed alleles 30 and 31 as the most common at DYS389II locus, opposite to allele 29 at the same locus in Europe.

The greatest publicly available YHRD is very important in forensic community, because it provides information on Y-chromosome haplotypes from world-wide distributed populations [14]. Enlargement of Y-STR databases sample size allows detection of variant alleles that are informative for Y-chromosome diversity [69]. European minimal haplotype set consisting of nine loci with high levels of variability in worldwide populations was used for analysis of molecular variance pair-wise distances based on R_{ST} values

Table 3 Haplogroup composition in single Croatian regional dataset

Haplogroup	Central Croatia Number of haplotypes	North Croatia Number of haplotypes	East Croatia Number of haplotypes	South Croatia Number of haplotypes	West Croatia Number of haplotypes	Pooled frequency (%)
I2a(×I2a1)	70	56	88	120	81	37.7
R1a	52	64	41	42	44	22.1
E1b1b	26	24	25	14	28	10.6
R1b	23	23	18	10	13	7.9
I1	11	9	13	12	19	5.8
J2b	11	11	6	6	7	3.7
G2a	8	7	4	4	7	2.7
H	3	11	2	1	4	1.8
J2a1h	1	1	5	2	4	1.2
J1	5		6		1	1.1
J2a1b	2	1	3	1	4	1
E1b1a		1				<1
G2c					1	<1
I2a1	1					<1
I2b1	2	5	1	1	1	<1
I2b(×I2b1)	2	2				<1
J2a1-bh			1	2	1	<1
L	1	1		1		<1
N	2	1	2		2	<1
Q			4	3	1	<1
T		3	1	1	2	<1

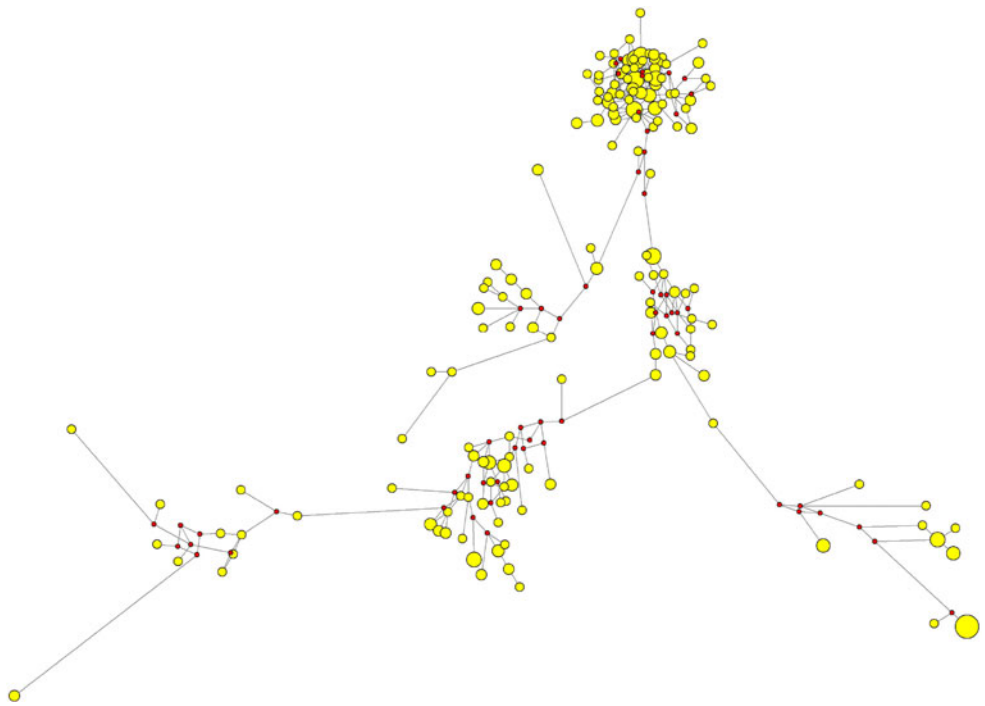

Fig. 3 Median-Joining network of the pooled Y-chromosomal haplogroups in Croatian population ($n = 1{,}100$)

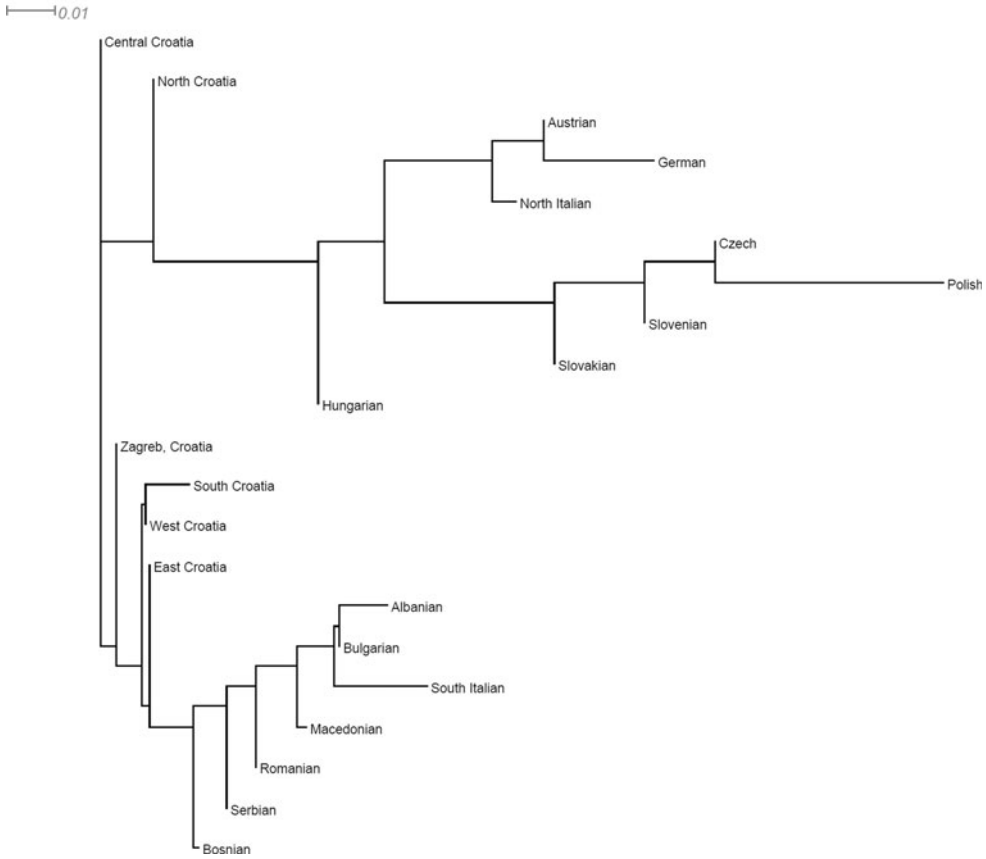

Fig. 4 Phylogram of Neighbor-Joining tree of the examined populations

between the Croatian subpopulations and neighboring countries. We compared a total of 16 population samples with 20,247 haplotypes. Only small genetic distances were documented between the Croatian regions, except for southern Croatia, indicating that Croatian population is quite homogeneous in terms of paternal genetic contribution. The relative genetic similarity was evident between the Croatian subpopulations when compared to Bosnian and Serbian populations. This similarity was less pronounced with Hungarian population. In this comprehensive study, we also showed that Croatian subpopulations are genetically distant from Czech, Polish and Slovakian populations. Strong genetic difference was observed from Slovenia concerning the paternal lineage. This result is very interesting in several historical dimensions. Croatian history is vivid and marked with strong dynamics. Initial historical records indicate the presence of old Illyrian tribes in the region, who were slowly or abruptly being assimilated into numerous newly arrived groups which include at least some Greek colonies in the coastal parts, Romans across entire territory, Slavic tribes in VIth century and later on, followed by, intrusions of Franks, and other Western-European groups, the long-lasting occupations by Ottoman Empire, agglomeration within Austrian-Hungarian empire and lastly substantial population changes imposed by the World War I and II, and most recently changes during the Croatian Homeland War in 1991–1995 [70].

Considering the fact that the extent of Y-STR variability may be influenced by the haplogroup composition, we firstly analyzed raw haplogroup structure in the large Croatian sample set diversified regionally. As it was mentioned earlier, by use the web-accessible program, Whit Atheys' Haplogroup Predictor we estimated the frequencies of Y-chromosome haplogroups from a set of acquired Y-STR data. The results of our study identified four main haplogroups in all five Croatian regions: I2a (37.7%), R1a (22.1%), E1b1b (10.6%) and R1b (7.9%). Each of the detected haplogroup had different level of haplotype diversification except R1a and E1b1: I2a (0.31), R1a (0.382), E1b1(0.383) and R1b (0.402). The lowest haplotypic diversity was measured in haplogroup I2a, which seemed to be the most homogenous haplogroup in all Croatian regions. On the other hand, differences in haplogroup frequencies between northern and southern samples were statistically significant what could be important in terms of forensic applications.

MJN graphs for the eight most frequent haplogroups were performed. This form of visual displays enabled to study remote relationships between the single haplotypes

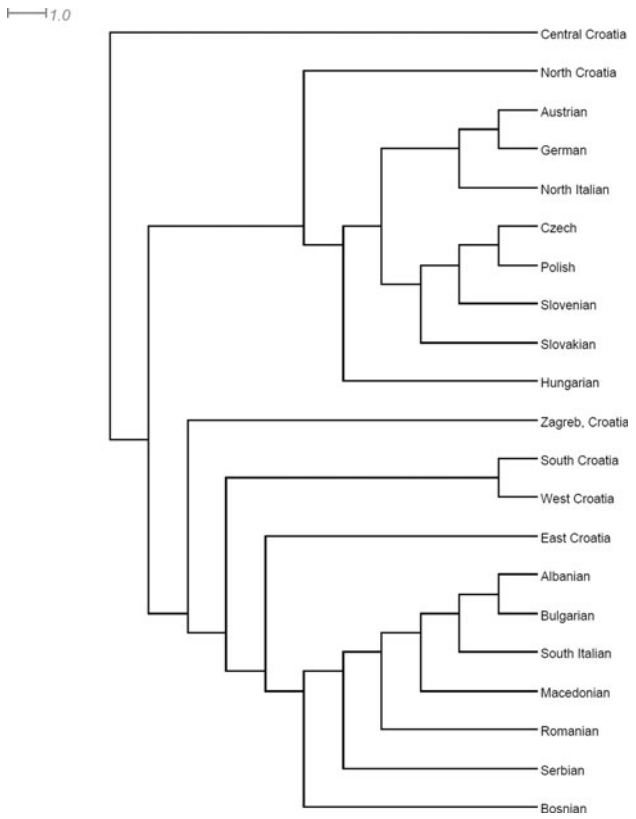

Fig. 5 Cladogram of Neighbor-Joining tree of the examined populations

inside particular haplogroup and to explore mutation events that might cause the present state of this affinity. In analyzed population sample 10 rare subhaplogroups were found (E1b1a, G2c, I2a1, I2b1, I2b(xI2b1), J2a1-bh, L, N, Q and T). The haplogroup rarity in the population sample had some benefits for the forensic applications, because in rare haplogroups the probability of finding haplotypes identical by descent is higher then finding haplotypes identical by state. This fact can be applicable in methods of familial searching where paternal lineages of special interest could be determined.

The high frequency of haplogroup I2a was observed in Bosnia and Herzegovina (70%), and decreasing frequencies were spreading from the southern Dinaric Alps to northern Croatia [71]. The same authors reported lower I2a1 haplogroup frequency in Croatia (32%).

We would like to emphasize the previously published studies based on haplogroup analysis using Y-SNPs in Croatian population [8, 9]. Barać et al. showed that the most prevalent haplogroups in Croatia are I (49%), R1a (27%) and R1b (7%), while haplogroups E, G, J, F, K and P are represented with 20% [9]. In addition, Peričić et al. analyzed 451 individuals, 348 from four Adriatic islands (Krk, Hvar, Brač, Korčula) and 109 from Croatia mainland (Osijek, Zabok, Zagreb, Delnice, Pazin and Dubrovnik) [9]. They showed that the most frequent Y-chromosomal haplogroups in the overall Croatian population are I1b-P37 (41.7%) and R1a-SRY1532 (25%) indicating the Slavic gene pool component. Moreover, Peričić et al. stress out that R1b lineage is present at relatively low frequency (8.5%) in southeastern Europe and Croatia (7.4%). Moreover, Semino et al. [72] data indicated that Croatian men can be subdivided into three major haplogroups, I (45%), R1a1 (30%) and R1b (10%). However, haplogroups E3b, J2a, G, LT are included all together with 15%. Although our haplogroup analysis was based on prediction; the frequencies of I2a (37.7%), R1a (22.1%) and R1b (7.9%) haplogroups were in concordance to those previously published for Croatian population [8, 9, 71, 72].

The main goals of our project were to create a national reference Y-STR haplotype database, as well as to investigate the Croatian intrapopulation Y-STR variability as well as genetic contribution from neighboring countries. These results may serve several important tasks, including that of forensics [73], but also contribute to our knowledge and understanding of the genetic background of contemporary Croatian populations, in terms of paternal lineages.

Future lofty goal will include analysis of informative biallelic markers in non-recombining region of Y-chromosome on this representative sample to further explore Croatian genetic heritage and to write additional chapter in the book of past European migration processes [71, 72, 74–77].

Acknowledgments This study was funded by the Ministry of Interior of Croatia. The authors thank Sunčica Papić, BSc, Marija Valić, BSc and Siniša Merkaš, BSc for technical support. Special thanks to Professor Lutz Roewer and Sascha Willuweit, MSc as well as Assistant Professor Ozren Polašek, MD, PhD for statistical analysis and support in the creation of this manuscript.

References

1. Butler JM (2003) Recent developments in Y-short tandem repeat and Y-single nucleotide polymorphism analysis. Forensic Sci Rev 15:91–111
2. Roewer L (2009) Y chromosome STR typing in crime casework. Forensic Sci Med Pathol 5:77–84. doi:10.1007/s12024-009-9089-5
3. Allor C, Einum DD, Scarpetta M (2005) Identification and characterization of variant alleles at CODIS STR loci. J Forensic Sci 50:1128–1133. doi:10.1520/JFS2005024
4. Lovrečić L, Ristić S, Brajenović B, Kapović M, Peterlin B (2005) Human Y-specific STR haplotypes in the Western Croatian population sample. Forensic Sci Int 149(2–3):257–261. doi:10.1016/j.forsciint.2004.06.026
5. Ljubković J, Stipisić A, Sutlović D, Definis-Gojanović M, Bucan K, Anđelinović Š (2008) Y-chromosomal short tandem repeat haplotypes in southern Croatian male population defined by 17 loci. Croat Med J 49(2):201–206. doi:10.3325/cmj.2008.2.201

6. Barać L, Peričić M, Klarić IM, Janićijević B, Parik J, Rootsi S, Rudan P (2003) Y chromosome STRs in Croatians. Forensic Sci Int 138(1–3):127–133. doi:10.1016/j.forsciint.2003.09.004
7. Haliti N, Carapina M, Masić M, Strinović D, Klarić IM, Kubat M (2009) Evaluation of population variation at 17 autosomal STR and 16 Y-STR haplotype loci in Croatians. Forensic Sci Int Genet 3(4):e137–e138. doi:10.1016/j.fsigen.2008.11.004
8. Barać L, Peričić M, Klarić IM, Rootsi S, Janićijević B, Kivisild T, Parik J, Rudan I, Villems R, Rudan P (2003) Y chromosomal heritage of Croatian population and its island isolates. Eur J Hum Genet 11(7):535–542. doi:10.1038/sj.ejhg.5200992
9. Peričić M, Barać Lauc L, Martinović Klarić I, Janićijević B, Rudan P (2005) Review of Croatian genetic heritage as revealed by mitochondrial DNA and Y chromosomal lineages. Croat Med J 46(4):502–513
10. Gršković B, Mršić G, Vrdoljak A, Merkaš S, Anđelinović Š (2010) Population genetic analysis of haplotypes based on 17 short tandem repeat loci on Y chromosome in population sample from eastern Croatia. Croat Med J 51(3):202–208. doi:10.3325/cmj.2010.51.202
11. Gršković B, Mršić G, Polašek O, Vrdoljak A, Merkaš S, Anđelinović Š (2011) Population data for 17 short tandem repeat loci on Y chromosome in northern Croatia. Mol Biol Rep 38(3):2203–2209. doi:10.1007/s11033-010-0349-y
12. Gršković B, Mršić G, Polašek O, Vrdoljak A, Merkaš S, Anđelinović Š (2011) Genetic polymorphisms of 17 short tandem repeat loci on Y chromosome in central Croatian population. Forensic Sci Med Pathol 7(2):155–161. doi:10.1007/s12024-010-9216-3
13. Walsh PS, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques 10:506–513
14. Willuweit S, Roewer L (2007) Y chromosome haplotype reference database (YHRD): update. Forensic Sci Int Genet 1(21): 83–87. doi:10.1016/j.fsigen.2007.01.017
15. Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50
16. Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Roewer L et al (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 10(3):125–133. doi:10.1007/s004140050051 141–149
17. Roewer L, Croucher PJP, Willuweit S, Lu TT, Kayser M, Lessig R, De Knijff P, Jobling MA, Tyler Smith C, Krawczak M (2005) Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. Hum Genet 116(4):279–291. doi:10.1007/s00439-004-1201-z
18. Robino C, Gino S, Ricci U, Grignani P, Previdere C, Torre C (2002) Y-chromosomal STR haplotypes in an Albanian population sample. Forensic Sci Int 129(2):128–130. doi:10.1016/S0531-5131(02)00550-2
19. Bosch E, Calafell F, Gonzalez Neira A, Flaiz C, Mateu E, Scheil HG, Huckenbeck W, Efremovska L, Mikerezi I, Xirotiris N, Grasa C, Schmidt H, Comas D (2006) Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns. Ann Hum Genet 70(Pt 4):459–487. doi:10.1111/j.1469-1809.2005.00251.x
20. Marjanovic D, Fornarino S, Montagna S, Primorac D, Hadziselimovic R, Vidovic S, Pojskic N, Battaglia V, Achilli A, Drobnic K, Andjelinovic S, Torroni A, Santachiara Benerecetti AS, Semino O (2005) The peopling of modern Bosnia-Herzegovina: Y-chromosome haplogroups in the three main ethnic groups. Ann Hum Genet 69(Pt 6):757–763. doi:10.1111/j.1529-8817.2005.00190.x
21. Zaharova B, Andonova S, Gilissen A, Cassiman JJ, Decorte R, Kremensky I (2001) Y-chromosomal STR haplotypes in three major population groups in Bulgaria. Forensic Sci Int 124(2–3): 182–186. doi:10.1016/S0379-0738(01)00597-7
22. Zastera J, Roewer L, Willuweit S, Sekerka P, Benesova L, Minarik M (2010) Assembly of a large Y-STR haplotype database for the Czech population and investigation of its substructure. Forensic Sci Int Genet 4(3):e75–e78. doi:10.1016/j.fsigen.2009.06.005
23. Henke J, Henke L, Chatthopadhyay P, Kayser M, Dulmer M, Cleef S, Poche H, Felske Zech H (2001) Application of Y-chromosomal STR haplotypes to forensic genetics. Croat Med J 42(3):292–297
24. Hidding M, Schmitt C (2000) Haplotype frequencies and population data of nine Y-chromosomal STR polymorphisms in a German and a Chinese population. Forensic Sci Int 113(1–3): 47–53. doi:10.1016/S0379-0738(00)00261-9
25. Roewer L, Krawczak M, Willuweit S, Nagy M, Alves C, Amorim A, Anslinger K, Augustin C, Betz A, Bosch E, Caglia A, Carracedo A, Corach D, Dekairelle AF, Dobosz T, Dupuy BM, Furedi S, Gehrig C, Gusmao L, Henke J, Henke L, Hidding M, Hohoff C, Hoste B, Jobling MA, Kargel HJ, De Knijff P, Lessig R, Liebeherr E, Lorente M, Martinez Jarreta B, Nievas P, Nowak M, Parson W, Pascali VL, Penacino G, Ploski R, Rolf B, Sala A, Schmidt U, Schmitt C, Schneider PM, Szibor R, Teifel Greding J, Kayser M (2001) Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. Forensic Sci Int 118(2–3):106–113. doi:10.1016/S0379-0738(00)00478-3
26. Junge A, Madea B (1999) Population studies of the Y-chromosome specific polymorphisms DYS19, DYS389 I + II, DYS390 and DYS393 in a western German population (Bonn area). Forensic Sci Int 101(3):195–201. doi:10.1016/S0379-0738(99)00029-8
27. Schmidt U, Meier N, Lutz S (2003) Y-chromosomal STR haplotypes in a population sample from southwest Germany (Freiburg area). Int J Legal Med 117(4):211–217. doi:10.1007/s00414-003-0373-8
28. Immel UD, Kleiber M, Klintschar M (2005) Y chromosome polymorphisms and haplotypes in South Saxony-Anhalt (Germany). Forensic Sci Int 155(2–3):211–215. doi:10.1016/j.forsciint.2005.01.004
29. Rodig H, Grum M, Grimmecke HD (2007) Population study and evaluation of 20 Y-chromosome STR loci in Germans. Int J Legal Med 121(1):24–27. doi:10.1007/s00414-005-0075-5
30. Lessig R, Edelmann J (1998) Y chromosome polymorphisms and haplotypes in west Saxony (Germany). Int J Legal Med 111(4): 215–218. doi:10.1007/s004140050155
31. Schneider PM, Meuser S, Waiyawuth W, Seo Y, Rittner C (1998) Tandem repeat structure of the duplicated Y-chromosomal STR locus DYS385 and frequency studies in the German and three Asian populations. Forensic Sci Int 97(1):61–70. doi:10.1016/S0379-0738(98)00146-7
32. Kayser M, Lao O, Anslinger K, Augustin C, Bargel G, Edelmann J, Elias S, Heinrich M, Henke J, Henke L, Hohoff C, Illing A, Jonkisz A, Kuzniar P, Lebioda A, Lessig R, Lewicki S, Maciejewska A, Monies DM, Pawłowski R, Poetsch M, Schmid D, Schmidt U, Schneider PM, Stradmann Bellinghausen B, Szibor R, Wegener R, Wozniak M, Zoledziewska M, Roewer L, Dobosz T, Ploski R (2005) Significant genetic differentiation between Poland and Germany follows present-day political borders, as revealed by Y-chromosome analysis. Hum Genet 117(5): 428–443. doi:10.1007/s00439-005-1333-9
33. Hohoff C, Dewa K, Sibbing U, Hoppe K, Forster P, Brinkmann B (2007) Y-chromosomal microsatellite mutation rates in a population sample from northwestern Germany. Int J Legal Med 121(5):359–363. doi:10.1007/s00414-006-0123-9

34. Anslinger K, Keil W, Weichhold G, Eisenmenger W (2000) Y-chromosomal STR haplotypes in a population sample from Bavaria. Int J Legal Med 113(3):189–192. doi:10.1007/s004140050296
35. Furedi S, Woller J, Padar Z, Angyal M (1999) Y-STR haplotyping in two Hungarian populations. Int J Legal Med 113(1):38–42. doi:10.1007/s004140050276
36. Völgyi A, Zalán A, Szvetnik E, Pamjav H (2009) Hungarian population data for 11 Y-STR and 49 Y-SNP markers. Forensic Sci Int Genet 3(2):e27–e28. doi:10.1016/j.fsigen.2008.04.006
37. Beer Z, Csete K, Varga T (2004) Y-chromosome STR haplotype in Szekely population. Forensic Sci Int 139(2–3):155–158. doi:10.1016/j.forsciint.2003.10.010
38. Spiroski M, Arsov T, Kruger C, Willuweit S, Roewer L (2005) Y-chromosomal STR haplotypes in Macedonian population samples. Forensic Sci Int 148(1):69–73. doi:10.1016/j.forsciint.2004.04.067
39. Bosch E, Calafell F, Gonzalez Neira A, Flaiz C, Mateu E, Scheil HG, Huckenbeck W, Efremovska L, Mikerezi I, Xirotiris N, Grasa C, Schmidt H, Comas D (2006) Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns. Ann Hum Genet 70(Pt 4):459–487. doi:10.1111/j.1469-1809.2005.00251.x
40. Ferri G, Ceccardi S, Lugaresi F, Bini C, Ingravallo F, Cicognani A, Falconi M, Pelotti S (2008) Male haplotypes and haplogroups differences between urban (Rimini) and rural area (Valmarecchia) in Romagna region (North Italy). Forensic Sci Int 175(2–3):250–255. doi:10.1016/j.forsciint.2007.06.007
41. Cerri N, Verzeletti A, Bandera B, De Ferrari F (2005) Population data for 12 Y-chromosome STRs in a sample from Brescia (northern Italy). Forensic Sci Int 152(1):83–87. doi:10.1016/j.forsciint.2005.02.006
42. Presciuttini S, Caglia A, Alu M, Asmundo A, Buscemi L, Caenazzo L, Carnevali E, Carra E, De Battisti Z, De Stefano F, Domenici R, Piccinini A, Resta N, Ricci U, Pascali VL (2001) Y-chromosome haplotypes in Italy: the GEFI collaborative database. Forensic Sci Int 122(2–3):184–188. doi:10.1016/S0379-0738(01)00500-X
43. Ferri G, Alú M, Corradini B, Radheshi E, Beduschi G (2008) Slow and fast evolving markers typing in Modena males (North Italy). Forensic Sci Int Genet 3(2):e31–e33. doi:10.1016/j.fsigen.2008.05.004
44. Grignani P, Peloso G, Fattorini P, Previdere C (2000) Highly informative Y-chromosomal haplotypes by the addition of three new STRs DYS437, DYS438 and DYS439. Int J Legal Med 114(1–2):125–129. doi:10.1007/s004140000153
45. Onofri V, Alessandrini F, Turchi C, Fraternale B, Buscemi L, Pesaresi M, Tagliabracci A (2007) Y-chromosome genetic structure in sub-Apennine populations of Central Italy by SNP and STR analysis. Int J Legal Med 121(3):234–237. doi:10.1007/s00414-007-0153-y
46. Turrina S, Atzei R, De Leo D (2006) Y-chromosomal STR haplotypes in a Northeast Italian population sample using 17plex loci PCR assay. Int J Legal Med 120(1):56–59. doi:10.1007/s00414-005-0054-x
47. Pepinski W, Niemcunowicz Janica A, Ptaszynska Sarosiek I, Skawronska M, Koc Zorawska E, Janica J, Soltyszewski I (2004) Population genetics of Y-chromosome STRs in a population of Podlasie, Northeastern Poland. Forensic Sci Int 144(1):77–82. doi:10.1016/j.forsciint.2004.02.024
48. Ploski R, Wozniak M, Pawlowski R, Monies DM, Branicki W, Kupiec T, Kloosterman A, Dobosz T, Bosch E, Nowak M, Lessig R, Jobling MA, Roewer L, Kayser M (2002) Homogeneity and distinctiveness of Polish paternal lineages revealed by Y chromosome microsatellite haplotype analysis. Hum Genet 110(6):592–600. doi:10.1007/s00439-002-0728-0
49. Pawlowski R, Dettlaff Kakol A (2003) Population data of nine Y-chromosomal STR loci in northern Poland. Forensic Sci Int 131(2–3):209–213. doi:10.1016/S0379-0738(02)00415-2
50. Rebala K, Szczerkowska Z (2005) Polish population study on Y chromosome haplotypes defined by 18 STR loci. Int J Legal Med 119(5):303–305. doi:10.1007/s00414-005-0547-7
51. Wolańska Nowak P, Branicki W, Parys Proszek A, Kupiec T (2009) A population data for 17 Y-chromosome STR loci in South Poland population sample—some DYS458.2 variants uncovered and sequenced. Forensic Sci Int Genet 4(1):e43–e44. doi:10.1016/j.fsigen.2009.04.009
52. Barbarii LE, Rolf B, Dermengiu D (2003) Y-chromosomal STR haplotypes in a Romanian population sample. Int J Legal Med 117(5):312–315. doi:10.1007/s00414-003-0397-0
53. Veselinovic IS, Zgonjanin DM, Maletin MP, Stojkovic O, Djurendic Brenesel M, Vukovic RM, Tasic MM (2008) Allele frequencies and population data for 17 Y-chromosome STR loci in a Serbian population sample from Vojvodina province. Forensic Sci Int 176(2–3):e23–e28. doi:10.1016/j.forsciint.2007.04.003
54. Rebala K, Mikulich AI, Tsybovsky IS, Sivakova D, Dzupinkova Z, Szczerkowska Dobosz A, Szczerkowska Z (2007) Y-STR variation among Slavs: evidence for the Slavic homeland in the middle Dnieper basin. J Hum Genet 52(5):406–414. doi:10.1007/s10038-007-0125-6
55. Ghiani ME, Vona G (2002) Y-chromosome-specific microsatellite variation in a population sample from Sardinia (Italy). Coll Antropol 26(2):387–401
56. Robino C, Inturri S, Gino S, Torre C, Di Gaetano C, Crobu F, Romano V, Matullo G, Piazza A (2006) Y-chromosomal STR haplotypes in Sicily. Forensic Sci Int 159(2–3):235–240. doi:10.1016/j.forsciint.2005.05.015
57. Sterlinko H, Pajnic IZ, Balazic J, Komel R (2001) Human Y-specific STR haplotypes in a Slovenian population sample. Forensic Sci Int 120(3):226–228. doi:10.1016/S0379-0738(01)00390-5
58. Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462
59. Athey TW (2005) Haplogroup prediction from Y-STR values using an allele-frequency approach. J Genet Geneal 1:1–7
60. Athey TW (2006) Haplogroup prediction from Y-STR values using a Bayesian-allele frequency approach. J Genet Geneal 2:34–39
61. Muzzio M, Raallo V, Motti JMB, Santos MR, Camelo JL, Bailliet G (2011) Software for Y-haplogroup predictions: a word of caution. Int J Legal Med 125:143–147. doi:10.1007/s00414-009-0404-1
62. Bandelt HJ, Forster P, Sykes BC, Richards MB (1995) Mitochondrial portraits of human populations using median networks. Genetics 141:743–753
63. Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48
64. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23(2):254–267. doi:10.1093/molbev/msj030
65. Wilkes J (1995) The Illyrians. Blackwell, London
66. Bosch E, Calafell F, González-Neira A, Flaiz C, Mateu E, Schneil HG, Huckenbeck W, Efremovska L, Mikerezi I, Xirotiris N, Grasa C, Schmidt H, Comas D (2006) Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns. Ann Hum Genet 70(Pt 4):459–487. doi:10.1111/j.1469-1809.2005.00251.x
67. Gusmão L, Butler JM, Carracedo A, Gill P, Kayser M, Mayr WR, Morling N, Prinz M, Roewer L, Tyler-Smith C, Schneider PM, DNA Commission of the International Society of Forensic Genetics (2006) DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on

the use of Y-STRs in forensic analysis. Forensic Sci Int 157(2–3):187–197. doi:10.1016/j.forsciint.2005.04.002
68. Roewer L (2003) The Y-Short Tandem Repeat Haplotype Reference database (YHRD) and male population stratification in Europe—impact on forensic genetics. Forensic Sci Rev 15(2): 164–170
69. Myres NM, Ritchie KH, Lin AA, Hughes RH, Woodward SR, Underhill PA (2009) Y-chromosome short tandem repeat intermediate variant alleles DYS392.2, DYS449.2, and DYS385.2 delineate new phylogenetic substructure in human Y-chromosome haplogroup tree. Croat Med J 50(3):239–249. doi:10.3325/cmj.2009.50.239
70. Klaić N (1971) Povijest Hrvata u ranom srednjem vijeku. Školska knjiga, Zagreb
71. Battaglia V, Fornarino S, Al-Zahery N, Olivieri A, Pala M, Myres NM, King RJ, Rootsi S, Marjanovic D, Primorac D, Hadziselimovic R, Vidovic S, Drobnic K, Durmishi N, Torroni A, Santachiara-Benerecetti AS, Underhill PA, Semino O (2009) Y-chromosomal evidence of the cultural diffusion of agriculture in Southeast Europe. Eur J Hum Genet 17(6):820–830. doi:10.1038/ejhg.2008.249
72. Semino O, Magri C, Benuzzi G, Lin AA, Al-Zahery N, Battaglia V, Maccioni L, Triantaphyllidis C, Shen P, Oefner PJ, Zhivotovsky LA, King R, Torroni A, Cavalli-Sforza LL, Underhill PA, Santachiara-Benerecetti AS (2004) Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. Am J Hum Genet 74(5):1023–1034. doi:10.1086/386295
73. Gršković B, Mršić G (2010) Y chromosome: from evolution to forensics- an overwiev. Acta Med Croatica 64:33–40
74. Primorac D, Marjanović D, Pavao Rudan, Villems R, Underhill P (2011) Croatian genetic heritage: Y chromosome story. Croat Med J 52(3):225–234. doi:10.3325/cmj.2011.52.225
75. Underhill PA, Myres NM, Rootsi S, Metspalu M, Zhivotovsky LA, King RJ, Lin AA, Chow CE, Semino O, Battaglia V, Kutuev I, Järve M, Chaubey G, Ayub Q, Mohyuddin A, Mehdi SQ, Sengupta S, Rogaev EI, Khusnutdinova EK, Pshenichnov A, Balanovsky O, Balanovska E, Jeran N, Augustin DH, Baldovic M, Herrera RJ, Thangaraj K, Singh V, Singh L, Majumder P, Rudan P, Primorac D, Villems R, Kivisild T (2010) Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a. Eur J Hum Genet 18(4):479–484. doi:10.1038/ejhg.2009.194
76. 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 LA, 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 JP, Chaventre A, Ferak V, Furedi S, Oefner PJ, Shen P, Beckman L, Mikerezi I, Terzic R, Primorac D, Cambon-Thomsen A, Krumina A, Torroni A, Underhill PA, Santachiara-Benerecetti AS, Villems R, Semino O (2004) Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. Am J Hum Genet 75(1):128–137. doi:10.1086/422196
77. Semino O, Passarino G, Oefner JP, Lin AA, Arbuzova S, Beckman EL, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S, Marcikic M, Mika A, Primorac D, Santachiara-Benerecetti AS, Cavalli-Sforza LL, Underhill AP (2000) The genetic legacy of paleolithic homo sapiens sapiens in extant Europeans: a Y chromosome perspective. Science 290:1155-1159. doi:10.1126/science.290.5494.1155