Genetics & Applications

Institute for Genetic Engineering and Biotechnology University of Sarajevo

Short communication Open access

Contemporary population genetics data for 23 Y-STR loci in the general Bosnian-Herzegovinian population.

Naida Babić Jordamović[1], Tamara Kojović[1], Serkan Dogan[1], Larisa Bešić[1], Lana Salihefendić[1,2], Rijad Konjhodžić[2], Sabahudin Ćordić[1], Vedrana Škaro[3,4], Petar Projić[3,4], Vesna Hadžiavdić[5], Adna Ašić[1,*], Damir Marjanović[1,3]

[1]*Department of Genetics and Bioengineering, International Burch University, Sarajevo, Bosnia and Herzegovina*
[2]*Alea Genetic Centre, Sarajevo, Bosnia and Herzegovina*
[3]*Institute for Anthropological Research, Centre for Applied Bioanthropology, Molecular Anthropology Laboratory, Zagreb, Croatia*
[4]*Genos Ltd, DNA Laboratory, Zagreb Croatia*
[5]*University of Tuzla, Department of Biology, Tuzla, Bosnia and Herzegovina.*

DOI: 10.31383/ga.vol5iss1pp51-63

***Correspondence**
E-mail: adna.asic@ibu.edu.ba

Received
May, 2021
Accepted
June, 2021
Published
June, 2021

Copyright: ©2021 Genetics & Applications, The Official Publication of the Institute for Genetic Engineering and Biotechnology, University of Sarajevo

Keywords

Y chromosome, Y-STR, Bosnia and Herzegovina, PowerPlex Y23 system, population genetics

Abstract

Bosnia and Herzegovina is located in the South-Eastern Europe, characterized by numerous historical influences, massive migration processes and complex population structure. For that reason, the aim of this study is to provide an accurate and precise update of the population genetics data of allele frequencies on 23 Y-STR loci in Bosnia and Herzegovina using larger sample size. For this purpose, 480 adult male individuals from the general population have been genotyped over 23 Y-STR loci contained in the PowerPlex Y23 system. Population genetics parameters have been calculated, namely allele and haplotype frequencies, gene and haplotype diversity, as well as R_{st} and P values for the assessment of interpopulation differences. The obtained results are in close agreement with previously published data for Bosnian-Herzegovinian population, as well as for local subpopulations. This study offers significantly increased resolution and information content, with 454 unique haplotypes. Population comparison reveals no statistically significant differences between the study population and 12 European populations used for comparison, as visualized through an MDS plot and neighbour-joining phylogenetic tree. This study offers representative data for local Y chromosomes that can be used for forensic applications, paternity and kinship testing, as well as for genealogical studies.

Introduction

With 57 Mb in length, the Y chromosome is one of the smallest chromosomes in the human genome (Skaletsky et al., 2003). Its main advantage for the population genetics studies is that a major part of the chromosome is inherited from father to son

unchanged, except for occasional mutations, as a lineage marker. There is no recombination between the X and Y chromosomes during meiosis, except for the small defined pseudoautosomal regions (PARs) located at the tips of the sex chromosomes (Jobling and Tyler-Smith, 2003; Sun and Heitman, 2012). This allows Y-chromosomal alleles to be inherited as a haplotype through male lineage that can be easily tracked and analysed for newly introduced mutational events (Jobling and Tyler-Smith, 2003; Butler, 2005).

Apart from being relatively small, 50% of the Y chromosome consists of repetitive sequences, including single-base substitutions, *Alu* elements and long interspersed nuclear elements (LINEs). When it comes to more mutable repetitive elements, prominent examples are short tandem repeats (STRs), with an average mutation frequency of ~0.2% per generation and the minisatellite locus *MSY1* with a mutation frequency of 6%–11% per generation (Marjanovic and Primorac, 2009; Ballantyne et al., 2010).

Bosnia and Herzegovina (B&H) is a small multi-ethnic country located in the Southeast Europe on the Balkan Peninsula. According to the results of the 2013 census, the total population size is approximately 3.8 million people (Agency for Statistics of Bosnia and Herzegovina, 2013). Just like the rest of the region, B&H is a very interesting area for the population studies, since the country is a home to several partially isolated indigenous populations, in addition to being found at the crossroads between the Middle East and Western Europe. Due to that, country-level studies on autosomal and linage markers (Marjanovic et al., 2004a; Marjanovic et al., 2005; Kovacevic et al., 2013), as well as the studies of characteristic subpopulations, such as the Sarajevo Canton (Cenanovic et al., 2010) and the Tuzla Canton (Babic et al., 2017) populations, gained the attention of the research community in the past. In addition, studies of isolated populations inhabiting high-altitude areas significantly contributed to the knowledge of the molecular diversity of genetic markers in B&H (Marjanovic et al., 2004b).

The main aim of this study is to perform a high-resolution molecular characterization of Bosnian-Herzegovinian Y chromosomes through Y-STR marker analysis. The study is conducted on an increased sample size according to the current recommended standards (Carracedo et al., 2014), therefore aiming to revise and update previously published data on the smaller study cohort, as well as to type the study chromosomes using 23 Y-STR loci in order to increase the informativeness of obtained haplotypes.

Material and methods

In this study, buccal swab samples were obtained during 2019 from 480 adult male individuals from different regions of B&H. This is a set of completely new samples that were never previously used for population genetics studies in B&H. All participants signed an informed consent form. Prior to sample collection, the approval of the Ethical Committee of the Faculty of Engineering and Natural Sciences was obtained. The study was conducted in line with Helsinki declaration. Genomic DNA was isolated using Qiagen DNeasy™ Blood and Tissue Kit (Hilden, Germany) according to manufacturer's recommendations.

PCR amplification of 23 Y-chromosomal short tandem repeat (STR) loci (*DYS19*, *DYS385a/b*, *DYS389I/II*, *DYS390*, *DYS391*, *DYS392*, *DYS393*, *DYS437*, *DYS438*, *DYS439*, *DYS448*, *DYS456*, *DYS458*, *DYS635*, *DYS481*, *DYS533*, *DYS549*, *DYS570*, *DYS576*, *DYS643* and *Y-GATA-H4*) incorporated in the PowerPlex® Y23 System (Promega Corporation, Madison, WI, USA) was performed according to manufacturer's instructions using SimpliAmp™ Thermal Cycler (Applied Biosystems, Foster City, CA, USA). PCR products were detected through capillary electrophoresis on ABI PRISM® 310 Genetic Analyzer (Applied Biosystems) according to manufacturer's instructions. Allelic data analysis and haplotype assignment were performed using GeneMapper™ v3.2 (Applied Biosystems). Amplified fragment analysis and Y-STR typing were carried out according to the quality assurance standards recommended by the Scientific Working Group on DNA Analysis Methods (SWGDAM, 2014).

The number of alleles and different haplotypes, allele and haplotype frequencies, and gene and

haplotype diversity were estimated in order to assess the intrapopulation diversity.

Haplotype diversity was calculated using Nei's formula: HD = $(1-\sum p_i^2)*n/(n-1)$, where n is the sample size and p_i is the i^{th} haplotype frequency. Gene diversity was calculated as $1-\sum p_i^2$, where pi is the allele frequency. The formula $\sum p_i^2$ was used to calculate match probability (MP), where pi is the frequency of the i^{th} haplotype. Discrimination capacity (DC) was calculated by dividing the number of haplotypes by the number of individuals in the population (Nei, 1987; Nei and Kumar, 2000). Allele and haplotype frequencies, as well as gene and haplotype diversity were calculated using the STRAF software package v1.0.5 (Gouy and Zieger, 2017).

Genetic distances between groups of males and between populations were quantified by R_{st} using AMOVA online tool from the Y Chromosome Haplotype Reference Database – YHRD (Willuweit and Roewer, 2007). In addition, associated probability values (P values) with 10,000 permutations were included for the studied European populations. Genetic distances were used to produce MDS plots for the comparison of population haplotype data, which were automatically generated on YHRD using the data available in this database. European populations selected for comparison with the population of B&H include: B&H (n = 480, present study), previously published data for the general Bosnian-Herzegovinian population (n = 100, Kovacevic et al., 2013), Croatia (n = 232, Grskovic et al., 2011), Slovenia (n = 97, Sterlinko et al., 2001), Belgium (n = 708, Roewer et al., 2001; Maesschalck et al., 2005; Mertens et al., 2007; Purps et al., 2014; Claerhout et al., 2020), Hungary (n = 576, Füredi et al., 1999; Egyed et al., 2000; Beer et al., 2004; Völgyi et al., 2009; Pamjav et al., 2017), Austria (n =253, Berger et al., 2005; Roewer et al., 2005; Erhart et al., 2012; Niederstätter et al., 2016; Pickrahn et al., 2016), Germany (n = 2756, Junge et al., 1997; Lessig and Edelmann, 1998; Schneider et al., 1998; Anslinger et al., 2000; Hidding et al., 2000; Henke et al., 2001; Schmidt et al., 2003; Immel et al., 2005; Kayser et al., 2005; Hohoff et al., 2007; Rodig et al., 2007), Czech Republic (n = 109, Zastera et al., 2010), Greece (n = 242, Parreira et al., 2002; Robino et al., 2004; Bosch et al., 2006; Kovatsi et al., 2009; Katsaloulis et al., 2013; Martínez et al., 2016), Italy (n = 1860, Grignani et al., 2000; Presciuttini et al., 2001; Ghiani et al., 2002; Cerri et al., 2005; Turrina et al., 2006; Robino et al., 2006; Onofri et al., 2007; Ferri et al., 2008; Ferri et al., 2009; Verzeletti et al., 2009; Rodríguez et al., 2009; Brisighelli et al., 2012; Piglionica et al., 2013; Robino et al., 2015; Rapone et al., 2016; Sarno et al., 2016; Lacerenza et al., 2017), North Macedonia (n = 96, Spiroski et al., 2005; Jakovski et al., 2019; Jankova et al., 2019), and Serbia (n = 379, Veselinovic et al., 2008; Veselinovic et al., 2014; Zgonjanin et al., 2017). These populations were selected because of the availability of high-quality population genetics studies performed on 23 Y-STR loci and on a significant number of samples that could produce meaningful results in population comparison efforts. In addition, we wanted a group of neighbouring populations to B&H, as well as a set of other European populations, in order to check the informativeness of obtained Y-STR data for interpopulation differentiation.

The evolutionary history was inferred using the neighbour-joining (NJ) method of phylogenetic tree construction (Saitou and Nei, 1987) in MEGAX (Kumar et al., 2018), whereby the optimal tree is shown.

Results and Discussion

On a sample of 480 participants, a total of 467 different haplotypes were detected in the study, with 454 unique haplotypes and 13 haplotypes appearing twice. In addition, 173 alleles at 23 Y-STR loci were detected (Table S1). Apart from *DYS385a/b* double locus, the largest number of alleles was recorded on DYS481 with 16 detected alleles. Two loci had the smallest number of alleles, namely *DYS438* and Y-GATA-H4 with five alleles each. Average genetic diversity for the study population was 0.634 across all loci, ranging from 0.344 at the locus *DYS392* to 0.884 at *DYS481*. At the population level, the most common allele is allele 11 at locus *DYS392* with frequency of 0.8021. This was not surprising considering that *DYS392* is one of the least polymorphic loci in the population with six detected alleles and lowest genetic diversity.

By comparing the population from the present study with previously published data for 12 European

populations, the lowest genetic diversity was observed between the currently analysed population of B&H and previously published Bosnian-Herzegovinian population (R_{st}=0.0021, P=0.2230, Kovacevic et al., 2013), as well as the population of Serbia (R_{st}= 0.0028, P=0.0647, Veselinovic et al., 2008; Veselinovic et al., 2014; Zgonjanin et al., 2017). Other populations with low genetic diversity values when compared to the present results include those from Croatia (R_{st}=0.0077, P=0.0148, Grskovic et al., 2011), Slovenia (R_{st}=0.0497, P=0.0000, Sterlinko et al., 2001), Hungary (R_{st}=0.0806, P=0.0000, Füredi et al., 1999; Egyed et al., 2000; Beer et al., 2004; Völgyi et al., 2009; Pamjav et al., 2017), Greece (R_{st}=0.0943, P=0.0000, Parreira et al., 2002; Robino et al., 2004; Bosch et al., 2006; Kovatsi et al., 2009; Katsaloulis et al., 2013; Martínez et al., 2016) and North Macedonia (R_{st}=0.1000, P=0.0000, Spiroski et al., 2005; Jakovski et al., 2019; Jankova et al., 2019). The highest genetic distance was observed when our study population was compared with the populations of Belgium (R_{st}= 0.2341, P=0.000, Roewer et al., 2001; Maesschalck et al., 2005; Mertens et al., 2007; Purps et al., 2014; Claerhout et al., 2020), Italy (R_{st}=0.2133, P=0.0000, Grignani et al., 2000; Presciuttini et al., 2001; Ghiani et al., 2002; Cerri et al., 2005; Robino et al., 2006; Turrina et al., 2006; Onofri et al., 2007; Ferri et al., 2008; Ferri et al., 2009; Rodríguez et al., 2009; Verzeletti et al., 2009; Brisighelli et al., 2012; Piglionica et al., 2013; Robino et al., 2015; Rapone et al., 2016; Sarno et al., 2016; Lacerenza et al., 2017), Germany (R_{st}=0.1804, P=0.0000, Junge et al., 1997; Lessig and Edelmann, 1998; Schneider et al., 1998; Anslinger et al., 2000; Hidding et al., 2000; Henke et al., 2001; Schmidt et al., 2003; Immel et al., 2005; Kayser et al., 2005; Hohoff et al., 2007; Rodig et al., 2007), Austria (R_{st}=0.1749, P=0.0000, Berger et al., 2005; Roewer et al., 2005; Erhart et al., 2012; Niederstätter et al., 2016; Pickrahn et al., 2016) and population of Czech Republic (R_{st}=0.1306, P=0.0000, Zastera et al., 2010) (Table S2).

Genetic relationships between investigated populations are also represented through an MDS plot (Figure 1). The results of such comparisons confirm the general trends that were observed in Table S2. To further investigate molecular evolutionary relationships between the populations of B&H and other European populations, a NJ phylogenetic tree was constructed based on R_{st} values (Figure 2). Generally, population comparisons point towards a general conclusion that geographically closer populations show higher degree of genetic relatedness. In the NJ tree (Figure 2), two Bosnian-Herzegovinian populations (including present and previously published data), Serbia, Croatia, and North Macedonia cluster together, while the populations of Slovenia, Hungary, Belgium, Italy, Austria, Czech Republic, Greece and Germany form another cluster on the phylogenetic dendrogram. Greece and Slovenia are the most distant in the second cluster, corresponding to their location in the Southern Europe and in geographic neighbourhood of the Balkan Peninsula, respectively. In addition, population comparisons using comparative analysis methods presented above show very low to no genetic differentiation between the currently analysed population and the Western Balkan populations.

Comparing our results with previously published Y-STR data for the Bosnian-Herzegovinian population, it is useful to start by observing individual loci. Loci *DYS390*, *DYS438*, *DYS437*, *DYS391* and *DYS389I* were previously identified as the least polymorphic ones in B&H (Cenanovic et al., 2010; Kovacevic et al., 2013). Current results obtained on 480 participants provide a good agreement, since all of them have six alleles per locus, except for *DYS438* with only five alleles. In addition, genetic diversity values for these loci are rather low, ranging from 0.4325 for *DYS389I* to 0.6206 for *DYS391*. Similarly, *DYS391* locus, with six alleles, was found to be the least polymorphic in the Turkish population recently settled in Sarajevo (Dogan et al., 2014).

Conversely, *DYS481* is the most polymorphic locus on 480 samples analysed here. This result is also in agreement with previous studies on neighbouring populations, but also for 100 samples of Bosnian-Herzegovinian population published in 2013 (Kovacevic et al., 2013; Babic et al., 2016; Zgonjanjin et al., 2017; Kacar et al., 2019).

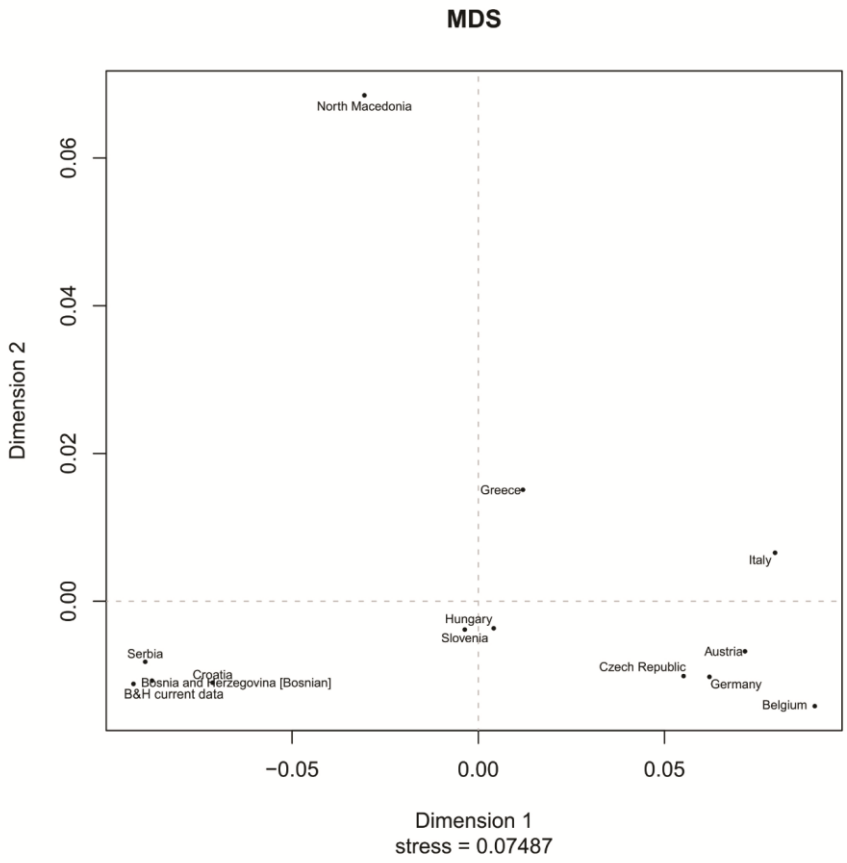


Figure 1. MDS plot showing genetic differentiation between the analysed populations in two dimensions.

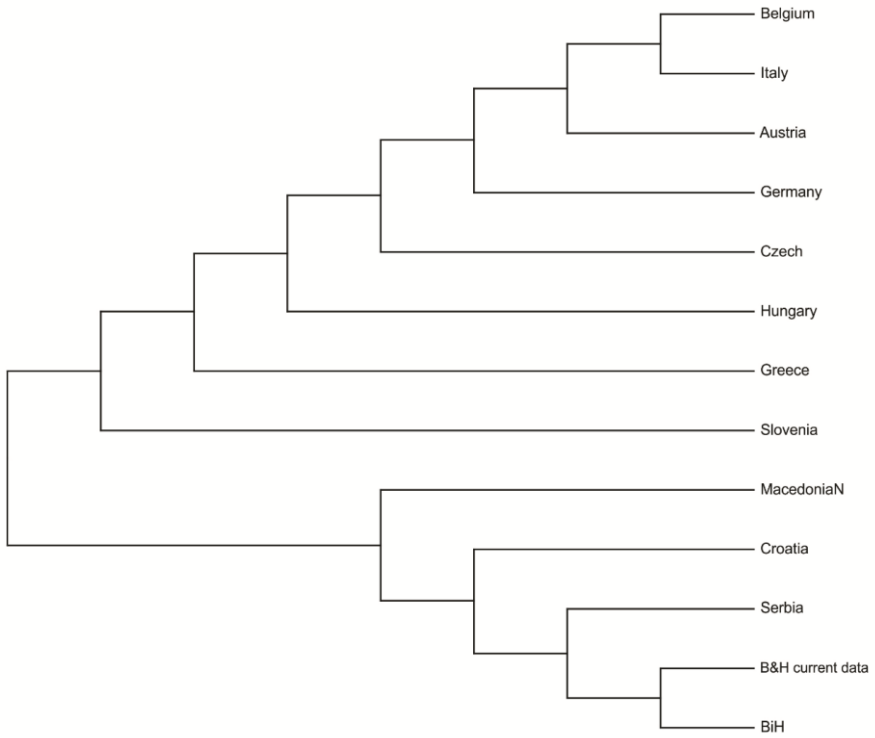


Figure 2. Neighbour-joining phylogenetic tree showing genetics relationships and clustering between the studied populations.

Allele 22 at the *DYS481* locus was found to be the most frequent allele in a recent study of Serbian population (Kacar et al., 2019), which is also the case in our study when observing this same locus.

Improved data resolution in this study is achieved not only through increased sample size and samples being collected from different parts of the country. Using 23 loci significantly improved information content of obtained haplotypes. More precisely, in a sample of 480 Y chromosomes, 454 haplotypes were unique. In a previous study of the local Tuzla Canton (B&H) population, all 100 haplotypes were unique (Babic et al., 2017), while in the local Sarajevo Canton (B&H) population, 98 samples had unique haplotypes, while one haplotype was repeated twice (Kovacevic et al., 2013), with both studies performed using the same Y-STR loci that were used in our study. For comparison, a study of the general Bosnian-Herzegovinian population on 100 typed over 12 Y-STR loci produced only 81 different haplotypes, including 71 unique (Cenanovic et al., 2010).

This study aims to provide an update of the current literature regarding Y-STR population data in B&H. Our study was prepared according to the latest guidelines on publishing forensic and population genetics data for Y-STRs (Carracedo et al., 2014), which state that a minimum of 17 Y-STR loci and 200 samples should be used. Autosomal STR data have already been updated in previous years, firstly by publishing data on 1000 samples analysed for 15 STR loci contained in the PowerPlex 16 system (Promega Corporation; Pilav et al., 2017), followed by publishing data on 22 STR loci contained in the PowerPlex Fusion system (Promega Corporation) on a sample of 600 Bosnian-Herzegovinians (Pilav et al., 2020). The need for an increased sample size and improved data resolution for Y-chromosome studies is addressed in the present study. The Y chromosomal STR data of the present study was submitted to Y-STR Haplotype Reference Database (YHRD) (http://www.yhrd.org) and accession number YA003787 was assigned [www.yhrd.org].

Conclusion

Our extended dataset of 480 Y-chromosomal haplotypes collected from different parts of the country and produced based on 23 Y-STR loci gives a more detailed insight into the allele frequency distribution on these loci. By increasing the sample size from 100 to 480 individuals, we are offering precise results on allele and haplotype frequencies, as well as population measures that can be used not only for population genetics studies, but also for forensic applications, paternity testing, kinship analysis, and missing person identification.

Conflict of Interest

The authors declare no conflict of interest.

References

Agency for Statistics of Bosnia and Herzegovina (2013) Census of population, households and dwellings in Bosnia and Herzegovina, 2013: final results. Sarajevo: Agency for Statistics of Bosnia and Herzegovina.

Anslinger K, Keil W, Weichhold G, Eisenmenger W (2000) Y-chromosomal STR haplotypes in a population sample from Bavaria. Int J Legal Med 113: 189-92.

Applied Biosystems. SimpliAmp Thermal Cycler. Foster City, CA, USA.

Applied Biosystems. ABI PRISM 310 Genetic Analyzer. Waltham, Massachusetts, USA.

Applied Biosystems. GeneMapper v3.2. Waltham, Massachusetts, USA.

Babic N, Dogan S, Cakar J, Pilav A, Marjanovic D, Hadziavdic V (2017) Molecular diversity of 23 Y-chromosome short tandem repeat loci in the population of Tuzla Canton, Bosnia and Herzegovina. Ann Hum Biol 44: 419-26.

Ballantyne KN, Goedbloed M, Fang R, Schaap O, Lao O, et al (2010) Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. Am J Hum Genet 87: 341-53.

Beer Z, Csete K, Varga T (2004) Y-chromosome STR haplotype in Szekely population. Forensic Sci Int 139: 155-8.

Berger B, Lindinger A, Niederstatter H, Grubwieser P, Parson W (2005) Y-STR typing of an Austrian population sample using a 17-loci multiplex PCR assay. Int J Legal Med 119: 241-6.

Bosch E, Calafell F, Gonzalez-Neira A, Flaiz C, Mateu E, et al (2006) Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns. Ann Hum Genet 70: 459-87.

Brisighelli F, Blanco-Verea A, Boschi I, Garagnani P, Pascali VL, et al (2012) Patterns of Y-STR variation in Italy. Forensic Sci Int Genet 6: 834-9.

Butler J (2005) Forensic DNA typing. Amsetrdam: Elsevier Academic Press.

Carracedo A, Butler JM, Gusmao L, Linacre A, Parson W, et al (2014) Update of the guidelines for the publication of genetic population data. Forensic Sci Int Genet 10: A1-2.

Cenanovic M, Pojskic N, Kovacevic L, Dzehverovic M, Cakar J, et al (2010) Diversity of Y-short tandem repeats in the representative sample of the population of Canton Sarajevo residents, Bosnia and Herzegovina. Coll Antropol 34: 545-50.

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: 83-7.

Claerhout S, Roelens J, Van der Haegen M, Verstraete P, Larmuseau MHD, Decorte R (2020) Ysurnames? The patrilineal Y-chromosome and surname correlation for DNA kinship research. Forensic Sci Int Genet 44: 102204.

De Maesschalck K, Vanhoutte E, Knaepen K, Vanderheyden N, Cassiman JJ, Decorte R (2005) Y-chromosomal STR haplotypes in a Belgian population sample and identification of a micro-variant with a flanking site mutation at DYS19. Forensic Sci Int 152: 89-94.

Dogan S, Primorac D, Marjanovic D (2014) Genetic analysis of haplotype data for 23 Y-chromosome short tandem repeat loci in the Turkish population recently settled in Sarajevo, Bosnia and Herzegovina. Croat Med J 55: 530-6.

Egyed B, Furedi S, Angyal M, Boutrand L, Vandenberghe A, et al (2000) Analysis of eight STR loci in two Hungarian populations. Int J Legal Med 113: 272-5.

Erhart D, Berger B, Niederstatter H, Gassner C, Schennach H, Parson W (2012) Frequency data for 17 Y-chromosomal STRs and 19 Y-chromosomal SNPs in the Tyrolean district of Reutte, Austria. Int J Legal Med 126: 977-8.

Ferri G, Ceccardi S, Lugaresi F, Bini C, Ingravallo F, et al (2008) Male haplotypes and haplogroups differences between urban (Rimini) and rural area (Valmarecchia) in Romagna region (North Italy). Forensic Sci Int 175: 250-5.

Ferri G, Alu M, Corradini B, Radheshi E, Beduschi G (2009) Slow and fast evolving markers typing in Modena males (North Italy). Forensic Sci Int Genet 3: e31-3.

Furedi S, Woller J, Padar Z, Angyal M (1999) Y-STR haplotyping in two Hungarian populations. Int J Legal Med 113: 38-42.

Ghiani ME, Vona G (2002) Y-chromosome-specific microsatellite variation in a population sample from Sardinia (Italy). Coll Antropol 26: 387-401.

Gouy A, Zieger M (2017) STRAF-A convenient online tool for STR data evaluation in forensic genetics. Forensic Sci Int Genet 30: 148-51.

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: 125-9.

Grskovic B, Mrsic G, Polasek O, Vrdoljak A, Merkas S, Andelinovic S (2011) Population data for 17 short tandem repeat loci on Y chromosome in northern Croatia. Mol Biol Rep 38: 2203-9.

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: 47-53.

Henke J, Henke L, Chatthopadhyay P, Kayser M, Dulmer M, et al (2001) Application of Y-chromosomal STR haplotypes to forensic genetics. Croat Med J 42: 292-7.

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: 359-63.

Immel UD, Kleiber M, Klintschar M (2005) Y chromosome polymorphisms and haplotypes in South Saxony-Anhalt (Germany). Forensic Sci Int 155: 211-5.

Jakovski Z, Nikolova K, Jankova-Ajanovska R, Marjanovic D, Pojskic N, Janeska B (2011) Genetic data for 17 Y-chromosomal STR loci in Macedonians in the Republic of Macedonia. Forensic Sci Int Genet 5: e108-11.

Jankova R, Seidel M, Videtic Paska A, Willuweit S, Roewer L (2019) Y-chromosome diversity of the three major ethno-linguistic groups in the Republic of North Macedonia. Forensic Sci Int Genet 42: 165-70.

Jobling MA, Tyler-Smith C (2003) The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet 4: 598-612.

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: 195-201.

Kacar T, Stamenkovic G, Blagojevic J, Krtinic J, Mijovic D, Marjanovic D (2019) Y chromosome genetic data defined by 23 short tandem repeats in a Serbian population on the Balkan Peninsula. Ann Hum Biol 46: 77-83.

Katsaloulis P, Tsekoura K, Vouropoulou M, Miniati P (2013) Genetic population study of 11 Y chromosome STR loci in Greece. Forensic Sci Int Genet 7: e56-8.

Kayser M, Lao O, Anslinger K, Augustin C, Bargel G, et al (2005) Significant genetic differentiation between Poland and Germany follows present-day political borders, as revealed by Y-chromosome analysis. Hum Genet 117: 428-43.

Kovacevic L, Fatur-Ceric V, Hadzic N, Cakar J, Primorac D, Marjanovic D (2013) Haplotype data for 23 Y-chromosome markers in a reference sample from Bosnia and Herzegovina. Croat Med J 54: 286-90.

Kovatsi L, Saunier JL, Irwin JA (2009) Population genetics of Y-chromosome STRs in a population of Northern Greeks. Forensic Sci Int Genet 4: e21-2.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol 35: 1547-49.

Lacerenza D, Aneli S, Di Gaetano C, Critelli R, Piazza A, et al (2017) Investigation of extended Y chromosome STR haplotypes in Sardinia. Forensic Sci Int Genet 27: 172-74.

Lessig R, Edelmann J (1998) Y chromosome polymorphisms and haplotypes in west Saxony (Germany). Int J Legal Med 111: 215-8.

Marjanovic D (2004a) Polimorfizam kompletno Y-vezanih molekularnih markera u referentnom uzorku bosansko hercegovačke populacije [Polymorphism of Y-linked molecular markers in a reference sample of the Bosnian human population] [dissertation]. Faculty of Natural Sciences (BiH): University of Sarajevo. Bosnian.

Marjanovic D, Kapur L, Drobnic K, Budowle B, Pojskic N, Hadziselimovic R (2004b) Comparative study of genetic variation at 15 STR loci in three isolated populations of the Bosnian mountain area. Hum Biol 76: 15-31.

Marjanovic D, Bakal N, Pojskic N, Kapur L, Drobnic K, et al (2005) Population data for the twelve Y-chromosome short tandem repeat loci from the sample of multinational population in Bosnia and Herzegovina. J Forensic Sci 50: 223-5.

Marjanovic D, Primorac D (2009) Molekularna forenzična genetika [Molecular forensic genetics]. Sarajevo: Institute for Genetic Engineering and Biotechnology. Bosnian.

Martinez-Cruz B, Mendizabal I, Harmant C, de Pablo R, Ioana M, et al (2016) Origins, admixture and founder lineages in European Roma. Eur J Hum Genet 24: 937-43.

Mertens G, Jehaes E, Leijnen G, Rand S, Jacobs W, Van Marck E (2007) Twelve-locus Y-STR haplotypes in the Flemish population. J Forensic Sci 52: 755-7.

Nei M (1987) Molecular evolutionary genetics. New York: Columbia University Press.

Nei M, Kumar S (2000) Molecular Evolution and Phylogenetics. New York: Oxford University Press.

Niederstatter H, Berger B, Kayser M, Parson W (2016) Differences in urbanization degree and consequences on the diversity of conventional vs. rapidly mutating Y-STRs in five municipalities from a small region of the Tyrolean Alps in Austria. Forensic Sci Int Genet 24: 180-93.

Onofri V, Alessandrini F, Turchi C, Fraternale B, Buscemi L, et al (2007) Y-chromosome genetic

structure in sub-Apennine populations of Central Italy by SNP and STR analysis. Int J Legal Med 121: 234-7.

Promega (2012) PowerPlex Y23 System - Technical Manual. Madison: Promega Corporation.

Qiagen (2005) QIAGEN Genomic DNA Handbook. Vienna: Qiagen Companies.

Pamjav H, Fothi A, Feher T, Fothi E (2017) A study of the Bodrogkoz population in north-eastern Hungary by Y chromosomal haplotypes and haplogroups. Mol Genet Genomics 292: 883-94.

Parreira KS, Lareu MV, Sanchez-Diz P, Skitsa I, Carracedo A (2002) DNA typing of short tandem repeat loci on Y-chromosome of Greek population. Forensic Sci Int 126: 261-4.

Pickrahn I, Muller E, Zahrer W, Dunkelmann B, Cemper-Kiesslich J, et al (2016) Yfiler((R)) Plus amplification kit validation and calculation of forensic parameters for two Austrian populations. Forensic Sci Int Genet 21: 90-4.

Pilav A, Pojskic N, Ahatovic A, Dzehverovic M, Cakar J, Marjanovic D (2017) Allele frequencies of 15 STR loci in Bosnian and Herzegovinian population. Croat Med J 58: 250-56.

Pilav A, Pojskic N, Kalajdzic A, Ahatovic A, Dzehverovic M, Cakar J (2020) Analysis of forensic genetic parameters of 22 autosomal STR markers (PowerPlex(R) Fusion System) in a population sample from Bosnia and Herzegovina. Ann Hum Biol 47: 273-83.

Piglionica M, Baldassarra SL, Giardina E, Stella A, D'Ovidio FD, et al (2013) Population data for 17 Y-chromosome STRs in a sample from Apulia (Southern Italy). Forensic Sci Int Genet 7: e3-4.

Presciuttini S, Caglia A, Alu M, Asmundo A, Buscemi L, et al (2001) Y-chromosome haplotypes in Italy: the GEFI collaborative database. Forensic Sci Int 122: 184-8.

Purps J, Siegert S, Willuweit S, Nagy M, Alves C, et al (2014) A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. Forensic Sci Int Genet 12: 12-23.

Rapone C, D'Atanasio E, Agostino A, Mariano M, Papaluca MT, et al (2016) Forensic genetic value of a 27 Y-STR loci multiplex (Yfiler((R)) Plus kit) in an Italian population sample. Forensic Sci Int Genet 21: e1-5.

Robino C, Varacalli S, Gino S, Chatzikyriakidou A, Kouvatsi A, et al (2004) Y-chromosomal STR haplotypes in a population sample from continental Greece, and the islands of Crete and Chios. Forensic Sci Int 145: 61-4.

Robino C, Inturri S, Gino S, Torre C, Di Gaetano C, et al (2006) Y-chromosomal STR haplotypes in Sicily. Forensic Sci Int 159: 235-40.

Robino C, Ralf A, Pasino S, De Marchi MR, Ballantyne KN, et al (2015) Development of an Italian RM Y-STR haplotype database: Results of the 2013 GEFI collaborative exercise. Forensic Sci Int Genet 15: 56-63.

Rodig H, Grum M, Grimmecke HD (2007) Population study and evaluation of 20 Y-chromosome STR loci in Germans. Int J Legal Med 121: 24-7.

Rodriguez V, Tomas C, Sanchez JJ, Castro JA, Ramon MM, et al (2009) Genetic sub-structure in western Mediterranean populations revealed by 12 Y-chromosome STR loci. Int J Legal Med 123: 137-41.

Roewer L, Krawczak M, Willuweit S, Nagy M, Alves C, et al (2001) Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. Forensic Sci Int 118: 106-13.

Roewer L, Croucher PJ, Willuweit S, Lu TT, Kayser M, et al (2005) Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. Hum Genet 116: 279-91.

Sarno S, Tofanelli S, De Fanti S, Quagliariello A, Bortolini E, et al (2016) Shared language, diverging genetic histories: high-resolution analysis of Y-chromosome variability in Calabrian and Sicilian Arbereshe. Eur J Hum Genet 24: 600-6.

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406-25.

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, et al (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423: 825-37.

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: 211-7.

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: 61-70.

Spiroski M, Arsov T, Kruger C, Willuweit S, Roewer L (2005) Y-chromosomal STR haplotypes in Macedonian population samples. Forensic Sci Int 148: 69-73.

Sterlinko H, Pajnic IZ, Balazic J, Komel R (2001) Human Y-specific STR haplotypes in a Slovenian population sample. Forensic Sci Int 120: 226-8.

Sun S, Heitman J (2012) Should Y stay or should Y go: the evolution of non-recombining sex chromosomes. Bioessays 34: 938-42.

[SWAGDAM] Scientific Working Group on DNA Analysis Mehods (2014) Interpretation Guidelines for Y-Chromosome STR Typing. Uniated States and Canada. Approved 01/09/14.

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: 56-9.

Veselinovic I, Petric G, Vapa D (2014) Genetic polymorphism of 17 Y chromosomal STRs in the Rusyn population sample from Vojvodina Province, Serbia. Int J Legal Med 128: 273-4.

Veselinovic IS, Zgonjanin DM, Maletin MP, Stojkovic O, Djurendic-Brenesel M, et al (2008) Allele frequencies and population data for 17 Y-chromosome STR loci in a Serbian population sample from Vojvodina province. Forensic Sci Int 176: e23-8.

Verzeletti A, Cerri N, Gasparini F, Poglio A, Mazzeo E, De Ferrari F (2009) Population data for 15 autosomal STRs loci and 12 Y chromosome STRs loci in a population sample from the Sardinia island (Italy). Leg Med (Tokyo) 11: 37-40.

Volgyi A, Zalan A, Szvetnik E, Pamjav H (2009) Hungarian population data for 11 Y-STR and 49 Y-SNP markers. Forensic Sci Int Genet 3:e27-8.

Willuweit S, Roewer L, International Forensic YCUG (2007) Y chromosome haplotype reference database (YHRD): update. Forensic Sci Int Genet 1: 83-7.

Xu H, Wang CC, Shrestha R, Wang LX, Zhang M, et al (2015) Inferring population structure and demographic history using Y-STR data from worldwide populations. Mol Genet Genomics 290:141-50.

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: e75-8.

Zgonjanin D, Alghafri R, Antov M, Stojiljkovic G, Petković S, Vuković R, Drašković D (2017) Genetic characterization of 27 Y-STR loci with the Yfiler ® Plus kit in the population of Serbia. Forensic Science International: Genetics. 31. 10.1016/j.fsigen.2017.07.013.

Table S1. Allele frequencies and statistical measures for the 23 Y-STR loci in Bosnian-Herzegovinian population

DYS576		DYS3891		DYS448		DYS389II		DYS19		DYS391		DYS481		DYS549	
GD=0,7132		**GD=0,4325**		**GD=0,5771**		**GD=0,7470**		**GD=0,7568**		**GD=0,6206**		**GD=0,8836**		**GD=0,6596**	
PIC=0,6673		**PIC=0,3950**		**PIC=0,4851**		**PIC=0,7052**		**PIC=0,7192**		**PIC=0,5463**		**PIC=0,8704**		**PIC=0,5941**	
PM=0,2882		**PM=0,5684**		**PM=0,4241**		**PM=0,2546**		**PM=0,2447**		**PM=0,3807**		**PM=0,1183**		**PM=0,3418**	
PD=0,7118		**PD=0,4316**		**PD=0,5759**		**PD=0,7454**		**PD=0,7553**		**PD=0,6193**		**PD=0,8817**		**PD=0,6582**	
Al.	Freq.	**Al.**	Freq.	**Al.**	Freq.	**Al.**	Freq.	**Al.**	Freq.	**Al.**	Freq.	**Al.**	Freq.	**Al.**	Freq.
14	0,0063	10	0,0021	17	0,0063	13	0,0021	12	0,0021	9	0,0188	13	0,0021	10	0,0104
5	0,0125	11	0,0021	18	0,0063	27	0,0083	13	0,1479	10	0,3563	19	0,0021	11	0,4104
16	0,0875	12	0,1479	19	0,4833	28	0,0917	14	0,1688	11	0,4854	20	0,0125	12	0,3938
17	0,2875	13	0,7313	20	0,4313	29	0,1854	15	0,2063	12	0,1333	21	0,0479	13	0,1250
18	0,4208	15	0,1083	21	0,0667	30	0,2688	**16**	0,3792	13	0,0042	22	0,1729	14	0,0500
19	0,1354	16	0,0083	22	0,0063	31	0,3667	**17**	0,0896	16	0,0021	23	0,1563	15	0,0083
20	0,0479					32	0,0708	18	0,0063			24	0,0583	16	0,0021
21	0,0021					33	0,0042					25	0,0563		
						34	0,0021					26	0,0375		
												27	0,0313		
												28	0,0271		
												29	0,0604		
												30	0,1521		
												31	0,1563		
												32	0,0250		
												33	0,0021		

DYS533		DYS438		DYS437		DYS456		DYS643		DYS393		DYS458		DYS385a	
GD=0,6333		**GD**=0,5259		**GD**=0,5736		**GD**=0,6586		**GD**=0,6203		**GD**=0,3491		**GD**=0,7652		**GD**=0,7675	
PIC=0,5693		**PIC**=0,4877		**PIC**=0,4899		**PIC**=0,6185		**PIC**=0,5856		**PIC**=0,3210		**PIC**=0,7276		**PIC**=0,7387	
PM=0,3681		**PM**=0,4752		**PM**=0,4276		**PM**=0,3428		**PM**=0,3810		**PM**=0,6516		**PM**=0,2364		**PM**=0,2341	
PD=0,6319		**PD**=0,5248		**PD**=0,5724		**PD**=0,6572		**PD**=0,6190		**PD**=0,3484		**PD**=0,7636		**PD**=0,7659	
Al.	Freq.	**Al.**	Freq.	**Al.**	Freq.	**Al.**	Freq.	**Al.**	Freq.	**Al.**	Freq.	**Al.**	Freq.	**Al.**	Freq.
9	0,0083	9	0,1021	10	0,0021	12	0,0021	8	0,0208	8	0,0021	11	0,0021	9	0,0021
10	0,0125	10	0,6583	13	0,0021	13	0,0229	9	0,0729	11	0,0083	13	0,0146	10	0,0146
11	0,1646	11	0,1583	14	0,3604	14	0,1083	10	0,5771	12	0,1354	14	0,0208	11	0,1792
12	0,5042	12	0,0792	15	0,5375	15	0,5271	11	0,1479	13	0,7938	15	0,2083	12	0,0771
13	0,2938	13	0,0021	16	0,0938	16	0,1875	12	0,1375	14	0,0563	16	0,2000	13	0,1250

14	0,0167			0,0042	17	0,1313	13	0,0375	15	0,0042	17	0,3542	15	0,4021	
					18	0,0188	14	0,0063			18	0,1604	15	0,0563	
					19	0,0021					19	0,0333	16	0,1229	
											20	0,0063	17	0,0167	
													18	0,0042	

DYS385b		DYS570		DYS635		DYS390		DYS439		DYS392		GATA H4	
GD=0,7468		**GD**=0,7424		**GD**=0,6507		**GD**=0,5583		**GD**=0,7303		**GD**=0,3442		**GD**=0,5276	
PIC=0,7187		**PIC**=0,7056		**PIC**=0,6003		**PIC**=0,5186		**PIC**=0,6838		**PIC**=0,3265		**PIC**=0,4307	
PM=0,2547		**PM**=0,2592		**PM**=0,3507		**PM**=0,4429		**PM**=0,2713		**PM**=0,6565		**PM**=0,4735	
PD=0,7453		**PD**=0,7408		**PD**=0,6493		**PD**=0,5571		**PD**=0,7288		**PD**=0,3435		**PD**=0,5265	
Al.	**Freq.**	**Al.**	**Freq.**	**Al.**	**Freq.**	**Al.**	**Freq.**	**Al.**	**Freq.**	**Al.**	**Freq.**	**Al.**	**Freq.**
11	0,0208	14	0,0063	13	0,0021	21	0,0083	9	0,0063	7	0,0021	10	0,0083
12	0,0104	15	0,0042	19	0,0021	22	0,0604	10	0,0938	10	0,0625	11	0,5729
13	0,0521	16	0,0396	20	0,0125	23	0,1396	11	0,2188	11	0,8021	12	0,3792
14	0,1958	17	0,1708	21	0,1583	24	0,6292	12	0,3854	12	0,0333	13	0,0375
15	0,4375	18	0,4125	22	0,2375	25	0,1542	13	0,2542	13	0,0896	15	0,0021
16	0,0792	19	0,2125	23	0,5146	26	0,0083	14	0,0375	14	0,0104		
17	0,0646	20	0,1083	24	0,0646			15	0,0042				
18	0,1000	21	0,0354	25	0,0083								
19	0,0354	22	0,0083										
20	0,0021	23	0,0021										
22	0,0021												

*Al. = Allele, Freq. = Frequency

Table S2. Interpopulation comparison of the current data with 12 previously published European populations using genetic distance R_{st} values and P values. Data was obtained from the relevant publications and accessed through the YHRD database

Population	BIH	BIH1	AU	BLG	CRO	CZH	GER	GRE	HUN	ITA	MAC	SRB	SLO
Current study (BIH)	*	0.2230	0.0000	0.0000	0.0137	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0678	0.0001
Bosnia and Herzegovina (BIH1)	0.0021	*	0.0000	0.0000	0.0203	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1479	0.0000
Austria (AU)	0.1749	0.2305	*	0.0000	0.0000	0.0025	0.0015	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Belgium (BLG)	0.2341	0.3079	0.0114	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Croatia (CRO)	0.0077	0.0168	0.1465	0.2014	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0008	0.0020
Czech Republic (CZH)	0.1306	0.1767	0.0162	0.0324	0.0939	*	0.1386	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005
Germany (GER)	0.1804	0.2252	0.0059	0.0140	0.1365	0.0024	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Greece (GRE)	0.0943	0.1259	0.0473	0.0998	0.0840	0.0432	0.0547	*	0.0000	0.0000	0.0000	0.0000	0.0037
Hungary (HUN)	0.0806	0.1097	0.0461	0.0876	0.0602	0.0302	0.0435	0.0169	*	0.0000	0.0000	0.0000	0.1445
Italy (ITA)	0.2133	0.2614	0.0102	0.0280	0.1855	0.0400	0.0267	0.0421	0.0592	*	0.0000	0.0000	0.0000
North Macedonia (MAC)	0.1000	0.1151	0.1459	0.2187	0.1118	0.1312	0.1556	0.0364	0.0813	0.1347	*	0.0000	0.0000
Serbia (SRB)	0.0028	0.0041	0.1862	0.2521	0.0178	0.1447	0.1880	0.0916	0.0864	0.2159	0.0880	*	0.0000
Slovenia (SLO)	0.0497	0.0756	0.0661	0.1179	0.0297	0.0320	0.0549	0.0187	0.0031	0.0826	0.0729	0.0568	*

**P* values are listed above the diagonal and R_{st} values below it