bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Croatian Genetic Heritage: Renewed Y Chromosome Story a Decade Later

Dragan Primorac[1-11], Vedrana Škaro[12,13], Petar Projić[12,13], Saša Missoni[12], Ivana Horjan Zanki[15], Sinisa Merkaš[15], Jelena Šarac[12], Natalija Novokmet[12], Andrea Ledić[15], Adela Makar[15], Gordan Lauc[13,14], Šimun Anđelinović[16], Željana Bašić[17], Ivana Kružić[17], Marijana Neuberg[18], Martina Smolić[6,19], Robert Smolić[6,19], Irena Hrstić[7,20], Dragan Trivanović[7,20], Rijad Konjhodžić[21], Lana Salihefendić[21], Naida Babić Jordamović[22], Damir Marjanović[9,22]

[1]St. Catherine Hospital, Zagreb, Croatia

[2]School of Medicine, University of Split, Split, Croatia

[3]University Department of Forensic Sciences, University of Split, Split, Croatia

[5]Faculty of Medicine, University of Osijek, Osijek, Croatia

[6]Faculty of Dental Medicine and Health, University of Osijek, Osijek Croatia

[7]School of Medicine Rijeka, University of Rijeka, Croatia

[7]School of Medicine Rijeka, University of Rijeka, Croatia

[8]Eberly College of Science, Pennsylvania State University, University Park, PA, USA

[9]Henry C. Lee College of Criminal Justice and Forensic Sciences, University of New Haven, West Haven, CT, USA

[10]Medical School REGIOMED, Coburg, Germany

[11]The National Forensic Sciences University, Gandhinagar, Gujarat, India

[12]Molecular Anthropology Laboratory, Center for Applied Bioanthropology, Institute for Anthropological Research, Zagreb, Croatia

[13]DNA Laboratory, Genos Ltd., Zagreb, Croatia

[14]Department of Pharmacology, Faculty of Pharmacy and Biochemistry, University of Zagreb

[15]Forensic Science Centre „Ivan Vučetić“, Zagreb, Croatia

[16]University Hospital Center, Split, Croatia

[17]University Department of Forensic Sciences, University of Split, Split, Croatia

[18]University North, Varaždin, Croatia

[19]Faculty of Medicine, University of Osijek, Osijek, Croatia

[20]General Hospital Pula, Pula, Croatia

[21]Alea Genetic Center, Sarajevo, Bosnia and Herzegovina

[22]Department of Genetics and Bioengineering, International Burch University, Sarajevo, Bosnia and Herzegovina

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Correspondence to: Vedrana Škaro, Genos Ltd., DNA Laboratory, Zagreb, Croatia; vskaro@genos.hr

Abstract

Aim: To analyze additional set of Y-Chromosome genetic markers to acquire a more detailed insight into the diversity of the Croatian population.

Methods: A total number of 518 Yfiler™ Plus profiles was genotyped. Allele, haplotype frequencies and haplotype diversity, were calculated using the STRAF software package v2.0.4. Genetic distances were quantified by *R*st using AMOVA online tool from the YHRD. The evolutionary history was inferred using the neighbor-joining method of phylogenetic tree construction in MEGAX software. Whit Athey's Haplogroup Predictor v5 was used for additional comparison with selected European populations.

Results: The total of 507 haplotypes were used for genetic STR analysis. The interpopulation comparison with the original 27 Y-STR markers shows the lowest genetic diversity between Croatian and Serbian population, and the highest between Croatian and Spanish population. Interpopulation study on 17 Y-STR markers shows the lowest genetic diversity between Croatian and Bosnian-Herzegovinian population, and the highest between Croatian and Irish population. Total of 518 haplotypes were used in the determination of haplogroup diversity. Haplogroup I with its sublineage I2a expressed the highest prevalence. Haplogroup R, with its major sublineage R1a, is the second most abundant in the studied Croatian population, except for the subpopulation of Hvar, where E1b1b is the second most abundant haplogroup. Rare haplogroups also confirmed in this study are L, T and Q. G1 is detected for the very first time in Croatian population.

Conclusion: New insight into differences between examined subpopulations of Croatia and their possible (dis)similarities with neighboring abroad populations was notified.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The Y chromosome (~ 60 Mb) is relatively small and inherited from father to son unchanged (except for occasional mutations). The length of the Y chromosome is 57 Mb. Except for the small pseudoautosomal regions PAR, there is no recombination between the X and Y chromosome (1-3). This event is the reason why haplotype inheritance through the male lineage that can be tracked and analyzed (2, 4-6).

The Y chromosome mostly consists of repetitive sequences (around 50%), which are single-base substitutions, Alu elements, and LINEs. Short tandem repeats, as repetitive elements are the base of population genetic studies. Their average mutational frequency is ~0.2% per generation (7, 8).

Y haplogroup can be defined as a part of the Y chromosome family related by ancestry and determined by a specific set of Y-SNP. It is of great importance for the better understanding of the demographic processes that shaped modern populations (8-10). A relatively new approach related to Y haplogroup prediction is using a set of Y-STR markers (haplotypes) for the determination of a haplogroup of an individual.

The main aim of this research is to update information about Croatian Y chromosome diversity by using additional Y STR loci to compare new results with the previously published results generated using Y-STR and Y-SNP markers (11-14) and to restart the Croatian Y chromosome settlement scenario established in one of our previously published studies (15). Additionally, one of the goals was to analyze the subpopulation genetic structure of five Croatian regional subpopulations (with local center in Osijek, Pula, Varaždin, Split, and Hvar Island) identifying the most common haplogroups, as well as genetic differences between these five subpopulations, and analyzing potential (dis)similarity of those regional subpopulations with neighboring countries.

Materials and methods

For this study, buccal swab samples were obtained from 518 adult male individuals from five different regions of Croatia: Hvar (n=104), Osijek (n=110), Pula (n=99), Varaždin (n=100) and Split (n=105).

DNA extraction was performed on QIAsymphony instrument using the QIAsymphony DNA Investigator Kit and protocol (Qiagen, Hilden, Germany). DNA was quantified on Rotor-Gene Q real-time PCR cycler (Qiagen, Hilden, Germany) using Q-Rex Software and Investigator Quantiplex Pro RGQ Kit. Yfiler™ Plus PCR Amplification Kit (Applied Biosystems, Foster City, CA, USA) was used to simultaneously amplify 27 Y chromosome STR loci.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Amplification was carried out following the manufacturer's protocol. PCR amplification was performed on Mastercycler® nexus SX1 PCR thermal cycler (Eppendorf AG, Hamburg, Germany) according to the manufacturer's instructions. PCR-amplified products were separated and detected using standard protocols for electrophoresis on 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Allele calling was performed with GeneMapper™ ID-X Software v1.4 (Applied Biosystems, Foster City, CA, USA) using custom panel and bin sets.

A total number of 507 fully genotyped Y-STR profiles of the present study were submitted to Y Chromosome Haplotype Reference Database – YHRD with the accession numbers assigned as follows: Hvar (n=104; YA004742), Osijek (n=109; YA004743), Pula (n=94; YA004744), Varaždin (n=98; YA004746), and Split (n=102; YA004745).

Statistical analysis

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 ith haplotype frequency. Gene diversity was calculated as $1-\sum p_i^2$, where p_i is the allele frequency. The formula $\sum p_i^2$ was used to calculate match probability (MP), where p_i is the frequency of the ith haplotype. Discrimination capacity (DC) was calculated by dividing the number of haplotypes by the number of individuals in the population (16, 17). Allele and haplotype frequencies, as well as gene and haplotype diversity were calculated using the STRAF software package v2.0.4 (18, 19).

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 (20, 21). In addition, associated probability values (*P* values) with 10,000 permutations were included for the studied populations. Genetic distances were used to generate MDS plots for the comparison of population haplotype data from YHRD.

AMOVA analysis was done with two population groups. The first comparison for analysis was made using the original number of the Yfiler™ Plus kit (27 Y-STR loci). The second comparison was made with a reduced number of markers using the Yfiler™ marker set (17 Y-STR loci). The reason for the two analyses was because of the increased number of markers in

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the analyzed population and the reduced number of markers in the populations that were available for comparison. The first group of populations selected for comparison with the population of Croatia using the Yfiler™ Plus marker set included: Croatia (n = 507, present study), Slovenia (n = 194, (22)), Belgium (n = 160, (23-26)), Hungary (n = 218, (27-33)), Austria (n =392, (34-38)), Germany (n = 495, (39-49)), Italy (n = 689, (50-67)), North Macedonia (n = 295, (68-71)), Serbia (n = 183, (72-74)), Denmark (n=177, (75)), Ethiopia (n=290, (76)), French Polynesia (n=81), (77, 78)), Ghana (n=584, (79)), India (n=541, (80-98)), Lithuania (n=251, (99-101)), Mexico (n=354, (102-107)), Nigeria (n=337, (108)), Pakistan (n=280, (109-121)), Poland (n=612, (122-130)), Russian Federation (n=958, (131-143)), Saudi Arabia (n=156, (144, 145)), Spain (n=316, (146-164)), Switzerland (n=724, (19, 165, 166)), and United Kingdom (n=115, (167, 168)).

The second group of populations selected for comparison with the population of Croatia using reduced 17 Y-STR Yfiler™ marker set includes: Tiroler Unterland, Austria (n=547, (37)), Antwerpen, Belgium (n=309, (26)), Bosnia and Herzegovina (n=574, (169, 170)), Bulgaria (n=91, (31, 171)), Rostock, Germany (n=598, (25, 49)), Greece (n=191, (172)), Hungary (n=303, (30)), Italy (n=147, (65)), Warsaw, Poland (n=491, (122)), Serbia (n=567, (74)), Albania (n=315, (58, 66, 69, 173)), Czech Republic (n=109, (174)), Estonia (n=123, (99)), Ireland (n=863, (75, 168)), Lithuania (n=531, (100, 101)), North Macedonia (n=493, (68, 70, 71)), Norway (n=1555, (175, 176)), Slovenia (n=294, (22)), Sweden (n=296, (177, 178)), and Ukraine (n=212, (179)).

The evolutionary history was inferred for both sets of markers using the neighbor-joining (NJ) method of phylogenetic tree construction (180) in MEGAX (181), whereby the optimal tree is shown.

Y-chromosomal haplogroup prediction using allele frequencies on 518 Yfiler™ Plus profiles was performed using Whit Athey's Haplogroup Predictor v5 (182), an algorithm based on the Bayesian-allele-frequency approach (182, 183).

Results and discussion

A total of 518 haplotypes were detected and used for haplogroup prediction. Eleven haplotypes had potentially newly detected microvariants, which needs to be confirmed by additional

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

analysis. Therefore, 507 haplotypes (only the ones without potentially new microvariants) were used for additional population genetic STR analysis. On a sample of fully genotyped 507 Y-STR profiles, a total of 502 different haplotypes were detected in the study, with 497 unique haplotypes and 5 haplotypes appearing twice. In addition, 196 alleles at 27 Y-STR loci were detected (Table 1). Apart from the DYS385a/b double locus, the largest number of alleles was recorded on DYS481 with 14 detected alleles. Three loci had the smallest number of alleles, namely DYS393, DYS437, and Y-GATA-H4 with four alleles each.

Haplotype diversity was calculated to be 1.0000 ± 0.0014 with DC of 1.00 and MP of 0.01. The average genetic diversity for the study population was 0.656 across all loci, ranging from 0,886 for DYS481to 0,251 for DYS392. At the population level, the most common allele is allele 11 at locus DYS392 with a frequency of 0.862. This was not surprising considering that DYS392 is one of the least polymorphic loci in the studied population with six detected alleles and the lowest genetic diversity.

[Table 1. here]

In order to determine additional genetic differences within the analyzed population of Croatia, an interpopulation analysis was done between five regions: Hvar (n=104), Osijek (n=109), Pula (n=94), Split (n=102) and Varaždin (n=98). The lowest genetic diversity observed for the population of Hvar was compared to population of Split (*R*st=0.0009, *P*=0.3240). The greatest genetic diversity observed for the population of Hvar was compared to the population of Varaždin (*R*st=0.0979, *P*=0.0000), the population of Pula (*R*st=0.0284, *P*=0.0042) and the population of Osijek (*R*st=0.0210, *P*=0.0097). The lowest genetic diversity observed for the population of Osijek was compared to population of Split (*R*st=0.0063, *P*=0.1199) and the population of Pula (*R*st=0.0069, *P*=0.1138). The greatest genetic diversity observed for the population of Osijek was compared to the population of Varaždin (*R*st=0.0551, *P*=0.0000) and the population of Hvar (*R*st=0.0210, *P*=0.0097). The lowest genetic diversity observed for the population of Pula was compared to the population of Osijek (*R*st=0.0069, *P*=0.1138). The greatest genetic diversity observed for the population of Pula was compared to the population of Hvar (*R*st=0.0284, *P*=0.0042), the population of Split (*R*st=0.0180, *P*=0.0233) and the population of Varaždin (*R*st=0.0166, *P*=0.0260). The lowest genetic diversity observed for the population of Split was compared to population of Hvar (*R*st=0.0009, *P*=0.3240) and the population of Osijek (*R*st=0.0063, *P*=0.1199). The greatest genetic diversity observed for the

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

population of Split was compared to the population of Varaždin (*R*st=0.0821, *P*=0.0000) and the population of Pula (*R*st=0.0180, *P*=0.0233). The lowest genetic diversity observed for the population of Varaždin was compared to the population of Pula (*R*st=0.0166, *P*=0.0260). The greatest genetic diversity observed for the population of Varaždin was compared to the population of Hvar (*R*st=0.0979, *P*=0.0000), the population of Split (*R*st=0.0821, *P*=0.0000) and the population of Osijek (*R*st=0.0551, *P*=0.0000).

[Table 2. here]

In order to compare the studied population with a large number of worldwide published population data, interpopulation analyses were completed by comparing the analyzed population with two groups of countries. The first group of selected countries used the original set of markers 27 Y-STRs; the selection was limited since this is an expanded panel of Y-STR markers.

By comparing the population from the present study with previously published data of the first selected group on 27 Y-STR markers for 23 populations, the lowest genetic diversity was observed between the currently analyzed population of Croatia and the previously published Serbian population (*R*st=0.0097, *P*=0.0055), and the population of Slovenia (*R*st= 0.0297, *P*=0.000). Other populations with low genetic diversity values when compared to the present results include those from Hungary (*R*st=0.0482, *P*=0.000), North Macedonia (*R*st=0.0720, *P*=0.0000), Russian Federation (*R*st=0.0779, *P*=0.0000), Pakistan (*R*st=0.0854, *P*=0.0000), Poland (*R*st=0.0905, *P*=0.0000), India (*R*st=0.0961, *P*=0.0000), and Saudi Arabia (*R*st=0.0990, *P*=0.0000). The highest genetic distance was observed when the study population was compared with the populations of Spain (*R*st= 0.3283, *P*=0.000), Ghana (*R*st=0.2787, *P*=0.0000), Nigeria (*R*st=0.2414, *P*=0.0000), Italy (*R*st=0.2043, *P*=0.0000), Belgium (*R*st=0.2031, *P*=0.0000), Switzerland (*R*st=0.1999, *P*=0.0000), Lithuania (*R*st=0.1914, *P*=0.0000), French Polynesia (*R*st=0.1878, *P*=0.0000), Mexico (*R*st=0.1785, *P*=0.0000), Ethiopia (*R*st=0.1704, *P*=0.0000), Germany (*R*st=0.1461, *P*=0.0000), Denmark (*R*st=0.1290, *P*=0.0000), Austria (*R*st=0.1255, *P*=0.0000), and United Kingdom (*R*st=0.1122, *P*=0.0000).

[Table 3. here]

The second group of populations selected for comparison with the population of Croatia using reduced 17 Y-STR set of markers included 20 populations. The lowest genetic diversity was

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

observed between the currently analyzed population of Croatia and previously published population of Bosnia and Herzegovina (R_{st}=0.0076, P=0.0002), and the population of Serbia (R_{st}= 0.0186, P=0.000). Other populations with low genetic diversity values when compared to the present results include those from Bulgaria (R_{st}=0.0144, P=0.000), Ukraine (R_{st}=0.0195, P=0.000), Slovenia (R_{st}=0.0204, P=0.000), Hungary (R_{st}=0.0238, P=0.0000), Greece (R_{st}=0.0241, P=0.0000), North Macedonia (R_{st}=0.0375, P=0.0000), Italy (R_{st}=0.0659, P=0.0000), Albania (R_{st}=0.0728, P=0.0000), Czech Republic (R_{st}=0.0767, P=0.000), and Austria (R_{st}=0.0795, P=0.0000). The highest genetic distance was observed when the study population was compared with the populations of Ireland (R_{st}= 0.3178, P=0.000), Estonia (R_{st}=0.1877, P=0.0000), Lithuania (R_{st}=0.1706, P=0.0000), Belgium (R_{st}=0.1429, P=0.0000), Norway (R_{st}=0.1270, P=0.0000) Poland (R_{st}=0.1216, P=0.0000), Sweden (R_{st}=0.1209, P=0.0000), and Germany (R_{st}=0.1036, P=0.0000)

[Table 4. here]

Genetic relationships between investigated populations are also represented through MDS plots (Figures 1-3). The results of such comparisons confirm the general trends that were observed in Tables 2-4.

[Figures 1-3 here]

To further investigate molecular evolutionary relationships between the populations of Croatia and selected populations, NJ phylogenetic trees were constructed based on *R*st values for different regions of Croatia and both compared groups (Figures 4- 6). Generally, the population comparisons point towards an overall conclusion that geographically closer populations show a higher degree of genetic relatedness.

[Figures 4-6 here]

The constructed NJ tree shows the comparison between the five regions of the analyzed population of Croatia (Figure 4a). The populations of the Hvar and Split regions are clustered together; the populations of the Osijek and Pula regions are presented in a separate cluster. The population of the Varaždin region is on a separate cluster from other regions, which could

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

indicate the potential genetic specificity (probably linked with geographical position) of this subpopulation compared with the other four examined regional subpopulations.

Additionally, the expanded NJ tree of the analyzed population from five Croatian regions together with the neighboring abroad population (Figure 4b) shows expected genetic similarity between the populations of the Hvar, Split, and Osijek regions. The populations of the Pula region are also located in the same macro-cluster. The eastern neighboring Serbian population is also included in that macro-cluster, with the closest genetic similarity with Pula, then the other three regional subpopulations. However, the population of the Varaždin region is in a relatively distanced cluster from the other four regions. The Varaždin population shows potential genetic closeness with neighboring Slovenian and Hungarian populations. These results prove again that the Y-chromosome is expected to show greater geographical clustering than other population markers (2, 15), but also could potentially mark significant migrations from the eastern neighboring countries to the Pula region especially in the second half of the 20th century.

In the NJ tree constructed for comparison of the group with the original set of 27 Y-STR markers, the Croatian population is clustered together with the populations of Serbia and Slovenia. The populations of Lithuania, Poland, and the Russian Federation are clustered together on the same branch with populations of Croatia, Serbia, Slovenia. The population of North Macedonia is positioned on the same branch as well. The populations of French Polynesia and Mexico are clustered together. The populations of Hungary and India clustered together. Populations of Belgium and Spain are clustered together on the same branch with populations of Switzerland, Italy, Germany, Austria, Denmark, and Pakistan. The populations of Ghana and Nigeria are clustered together with the populations of Ethiopia, Saudi Arabia, and the United Kingdom (Figure 5).

Compared to population of Croatia and selected populations with a reduced marker set, the NJ tree shows the population of Croatia is positioned on the same branch with the cluster of populations of Bosnia and Herzegovina and Serbia. The populations of Albania and North Macedonia are clustered together on the branch with the populations of Greece and Bulgaria. The populations of Slovenia and Ukraine are clustered together. The populations of Poland, Estonia, and Lithuania are clustered together. On a different cluster, positioned together are the populations of Germany, Belgium, and Ireland. In addition, population comparisons using

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

comparative analysis methods presented above show very low to no genetic differentiation between the currently analyzed population and the Western Balkan populations.

For the calculation of Y-chromosomal haplogroup prediction and intrapopulation variability between the five subpopulations, the total number of 518 Yfiler™ Plus profiles was used: Hvar (n=104), Varaždin (n=100), Split (n=105), Pula (n=99), and Osijek (n=110). Regarding the haplogroup diversity between these five subpopulations of Croatia, successful haplogroup assignment was obtained for all 518 Y-STR profiles (Table 5). The results of Y haplogroup prediction using Whit Athey's Haplogroup Predictor tool (182) are summarized in Figure 7. Prediction accuracy was estimated to be 100% in 492 cases. For the remaining 26 samples the prediction accuracy was 5%. Prediction accuracy varied between 63,1% and 99.58%. Out of a total of 14 detected haplogroups, the most prevalent one is I2a, which accounts for 39% of all samples, followed by R1a (24,32%) and E1b1b (10,18%). The remaining eight haplogroups were less frequent.

[Figure 7. here]

[Table 5. here]

Four of five subpopulations of Croatia showed expected results (Figure 8a-c). High frequency of haplogroup I has been reported with its known sublineage I2a in the subpopulations Hvar 52,88%, Split 50,48%, Osijek 40,91%, and Pula 31,31%. Previously published reports demonstrate similar results (11, 13, 14). However, different results were obtained in the subpopulation of Varaždin (Figure 8d). R1a was the most frequent haplogroup in this subpopulation with the frequency of 38%, while the frequency for I2a haplogroup was 18%. It is interesting to note that R1a was also the dominant haplogroup within Slovenian population, which is the closest neighboring abroad population to Varaždin county (184).

[Figure 8 here]

In summary, sublineage I2a was generally the most frequent haplogroup in the populations of Croatia, Bosnia and Serbia, according to this study, but also to all previously studied data (11, 15, 185).

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Arrival of haplogroup I in this part of Europe is approximated to around 25,000 years ago from the Middle East through Anatolia in the area of Balkan (15, 186). One scenario suggests the possibility of population expansion from one of the post-Glacial refuges into the rest of the Balkan Peninsula (169). There is also another possibility that this haplogroup could be connected with more recent population movements from the eastern Europe, but this idea still has to be examined (9). Definitively, when comparing to the other populations in Europe, the I2a haplogroup sublineage is considered the characteristic Southeast European haplogroup (187).

The R1a as a leading sublineage of haplogroup R was the second most frequent haplogroup in the studied population of Croatia, with an overall frequency of 24,32%. The prevalence of haplogroup R1a in the subpopulations of Croatia reported frequencies for Varaždin 38%, Pula 28,28%, Osijek 26,36% and Split 19,05%. In the subpopulation of Hvar a small genetic deviation in the frequency of haplogroups R1a and E1b1b was reported. The R1a haplogroup accounted for 10,58% just slightly lower than haplogroup E1b1b with frequency 11,54% (Figure 8a). This is most likely due to the founder effect which is expected for island populations. In previously reported studies on mainland population of Croatia, haplogroup R reported a frequency of 22,9% (11, 185). Migration theories of R1a origins indicates outflow of haplogroup R from West Asia to the region of Balkan as a post-Last Glacial Maximum (LGM) event, during the Mesolithic time (15, 188).

Sublineage R1b is presented in lower frequency in the studied population of Croatia. The overall frequency of the R1b sublineage for the population of Croatia accounts of 6,37%. The highest frequency of R1b haplogroup was reported in the subpopulation of Varaždin with a prevalence of 9% and in Pula with a frequency of 8,08%. Similar results were obtained in the Bosnian population based on 481 Y-STR profiles, whereby R1b accounted for 8,75% of the samples (189).

Sublineage E1b1b is the most frequent “neolithic haplogroup” for the males in this part of Europe (15). In the present study E1b1b was detected with the frequency of 10,81%. The highest prevalence of this haplogroup was reported in the subpopulation of Varaždin, with a frequency of 16%. In the other four subpopulations the frequency of haplogroup E1b1b was estimated: Hvar 11,54%, Pula 11,11%, Split 10,48% and Osijek 5,45%. In previously published results, this haplogroup was also less frequent in the population of Croatia (5.5%), Serbia (18.6%) and Bosnia and Herzegovina (14,58%) (11, 185, 189). There are two suggested theories

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

about E1b1b arrival in Europe. One theory is a post-LGM event from Asia and Africa during Neolithic period of time, while the other evidence suggests that this haplogroup is Balkan-specific, and originated around 8,000 years ago as a part of Greek colonization among the northern part of the Peninsula (15, 186). This ancient European haplogroup shows its possible dual origin from two different source populations, during recolonization of Europe from Iberia and from West Asia (15, 185).

Rare haplogroups discovered in this study were Q, T, L and G1 each present in 1,93%, 0,58%, 0,19% and 0,19% of all samples, respectively. Haplogroup L is associated with South Asia and India but is also found in low frequencies in Central Asia, Southwest Asia, and Southern Europe. With its alternative phylogenetic name K1a, haplogroup L is closely related to haplogroup T (190). Haplogroup T, also known by its phylogenetic name K1b, possibly originated in Western Asia and spread to East Africa, South Asia, and Southern Europe (191, 192). Finally, haplogroup Q represents the only Pan-American haplogroup and confirms the Asian origin of Native Americans, and provides insight into the main Asian-American migrations (193). Haplogroup G1 is detected for the very first time within Croatian population. This haplogroup is found predominantly in Iran, and is very rare in Europe. Its origin is still not clearly described.

The present analyses generally confirmed previous results and provided a more detailed insight into the genetic diversity of the subpopulations of Croatia. Analysis of 27 Y-STR loci on the currently studied population of Croatia was quantified through *R*st values and calculated based on the results of the same set of markers from other populations from the YHRD database. The results indicate that the Croatian population does not deviate significantly from the neighboring populations of Bosnia and Herzegovina, Serbia, and Slovenia. Hence, this proves that the Y chromosome genetic marker has conspicuous geographical background (2, 15) and this analysis resulted with expected geographic clustering.

A previous study (15) concluded that most of Croatian men ("owners" of HgI, R1a and R1b) are harboring the ancestral genetic impact of Old Europeans people who settled in Europe approximately 25.000 – 30.000 years ago and survived the LGM in several different refugium. Results of our studies confirmed that more than 78% of the contemporary Croatians are included in that group. The rest of the population relate to the people who arrived mostly during the Neolithization process. A small portion of examined population originated from the "owners" of rare haplogroups in the term of European genetic diversity, of which the origin is

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

still not clarified. Usage of additional Y-STR loci revealed detailed insight and supplementary information within the highly diverse and significant genetic diversity of the Croatian population.

Acknowledgment

The authors thank the sample donors for participating in this research, and colleagues at the University Hospital Center, Split, Croatia; University Department of Forensic Sciences, University of Split, Split, Croatia; University North, Varaždin, Croatia; Faculty of Medicine, University of Osijek, Osijek, Croatia; Faculty of Dental Medicine and Health, University of Osijek, Osijek Croatia; School of Medicine Rijeka, University of Rijeka, Croatia; and General Hospital Pula, Pula, Croatia for assistance with the buccal swab collection.
This collaboration was supported by the Croatian Society for Human Genetics and the International Society for Applied Biological Sciences (ISABS).

Funding Study "Analysis of Y chromosome variability in the male population of Croatia" was conducted and co-financed by the Institute of Anthropology, Zagreb, Croatia; Forensic Science Centre „Ivan Vučetić", Zagreb, Croatia; and Genos Ltd Zagreb, Croatia.

Ethical approval was given by the Ethics Committee, Institute for Anthropological Research, Zagreb, Croatia (approval number 20211053).

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1 Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature. 2003;423(6942):825-37.

2 Jobling MA, Tyler-Smith C. The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet. 2003;4(8):598-612.

3 Sun S, Heitman J. Should Y stay or should Y go: the evolution of non-recombining sex chromosomes. Bioessays. 2012;34(11):938-42.

4 Carracedo A, Butler JM, Gusmão L, Linacre A, Parson W, Roewer L, et al. Update of the guidelines for the publication of genetic population data. Forensic Sci Int Genet. 2014;10:A1-a2.

5 Butler JM. Forensic DNA Typing : Biology, Technology, and genetics of STR markers. 2nd ed. London: Academic Press; 2004. 688 p.

6 Butler JM. Advanced topics in forensic DNA typing : methodology: San Diego; 2012.

7 Ballantyne KN, Goedbloed M, Fang R, Schaap O, Lao O, Wollstein A, et al. Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. Am J Hum Genet. 2010;87(3):341-53.

8 Marjanovic D, Primorac D, Dogan S. Forensic Genetics: Theory and Application. Sarajevo, Bosnia and Herzegovina: International Burch University; 2018. 289 p.

9 Marjanović D, Primorac D, Mesihović S. The ancestors in us: Genetic heritage of Bosnia and Herzegovina [In Bosnian: Preci u nama: Genetičko blago Bosne i Hercegovine]. Sarajevo, Bosnia and Herzegovina: Mladinska knjiga; 2019. 302 p.

10 Marjanović D, Primorac D. Molecular forensic genetics [Molekularna forenzična genetika]. Sarajevo. Bosnia and Herzegovina: Institute for Genetic Engineering and Biotechnology; 2009. 250 p.

11 Barać L, Pericić M, Klarić IM, Rootsi S, Janićijević B, Kivisild T, et al. Y chromosomal heritage of Croatian population and its island isolates. Eur J Hum Genet. 2003;11(7):535-42.

12 Gršković B, Mršić G, Polašek O, Vrdoljak A, Merkaš S, Anđelinović S. Population data for 17 short tandem repeat loci on Y chromosome in northern Croatia. Mol Biol Rep. 2011;38(3):2203-9.

13 Mršić G, Gršković B, Vrdoljak A, Popović M, Valpotić I, Anđelinović Š, et al. Croatian national reference Y-STR haplotype database. Mol Biol Rep. 2012;39(7):7727-41.

14 Šarac J, Šarić T, Havaš Auguštin D, Novokmet N, Vekarić N, Mustać M, et al. Genetic heritage of Croatians in the Southeastern European gene pool-Y chromosome analysis of the Croatian continental and Island population. Am J Hum Biol. 2016;28(6):837-45.

15 Primorac D, Marjanović D, Rudan P, Villems R, Underhill PA. Croatian genetic heritage: Y-chromosome story. Croat Med J. 2011;52(3):225-34.

16 Nei M. Molecular Evolutionary Genetics. New York, USA: Columbia University; 1987. 512 p.

17 Nei M, Kumar S. Molecular Evolution and Phylogenetics. New York, USA: Oxford University Press; 2000. 348 p.

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

19 Zieger M, Utz S. The Y-chromosomal haplotype and haplogroup distribution of modern Switzerland still reflects the alpine divide as a geographical barrier for human migration. Forensic Sci Int Genet. 2020;48:102345.

20 Willuweit S, Roewer L. Y chromosome haplotype reference database (YHRD): update. Forensic Sci Int Genet. 2007;1(2):83-7.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

21 Roewer L, Krawczak M, Willuweit S, Nagy M, Alves C, Amorim A, et al. Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. Forensic Sci Int. 2001;118(2-3):106-13.

22 Sterlinko H, Pajnic IZ, Balazic J, Komel R. Human Y-specific STR haplotypes in a Slovenian population sample. Forensic Sci Int. 2001;120(3):226-8.

23 De Maesschalck K, Vanhoutte E, Knaepen K, Vanderheyden N, Cassiman JJ, Decorte R. 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. 2005;152(1):89-94.

24 Mertens G, Jehaes E, Leijnen G, Rand S, Jacobs W, Van Marck E. Twelve-locus Y-STR haplotypes in the Flemish population. J Forensic Sci. 2007;52(3):755-7.

25 Purps J, Siegert S, Willuweit S, Nagy M, Alves C, Salazar R, et al. A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. Forensic Sci Int Genet. 2014;12(100):12-23.

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

27 Füredi S, Woller J, Pádár Z, Angyal M. Y-STR haplotyping in two Hungarian populations. Int J Legal Med. 1999;113(1):38-42.

28 Egyed B, Füredi S, Angyal M, Boutrand L, Vandenberghe A, Woller J, et al. Analysis of eight STR loci in two Hungarian populations. Forensic Sci Int. 2000;113(1-3):25-7.

29 Beer Z, Csete K, Varga T. Y-chromosome STR haplotype in Szekely population. Forensic Sci Int. 2004;139(2-3):155-8.

30 Völgyi A, Zalán A, Szvetnik E, Pamjav H. Hungarian population data for 11 Y-STR and 49 Y-SNP markers. Forensic Sci Int Genet. 2009;3(2):e27-8.

31 Martínez-Cruz B, Mendizabal I, Harmant C, de Pablo R, Ioana M, Angelicheva D, et al. Origins, admixture and founder lineages in European Roma. Eur J Hum Genet. 2016;24(6):937-43.

32 Pamjav H, Fóthi Á, Fehér T, Fóthi E. A study of the Bodrogköz population in north-eastern Hungary by Y chromosomal haplotypes and haplogroups. Mol Genet Genomics. 2017;292(4):883-94.

33 Pamjav H, Zalán A, Béres J, Nagy M, Chang YM. Genetic structure of the paternal lineage of the Roma people. Am J Phys Anthropol. 2011;145(1):21-9.

34 Berger B, Lindinger A, Niederstätter H, Grubwieser P, Parson W. Y-STR typing of an Austrian population sample using a 17-loci multiplex PCR assay. Int J Legal Med. 2005;119(4):241-6.

35 Roewer L, Croucher PJ, Willuweit S, Lu TT, Kayser M, Lessig R, et al. Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. Hum Genet. 2005;116(4):279-91.

36 Erhart D, Berger B, Niederstätter H, Gassner C, Schennach H, Parson W. Frequency data for 17 Y-chromosomal STRs and 19 Y-chromosomal SNPs in the Tyrolean district of Reutte, Austria. Int J Legal Med. 2012;126(6):977-8.

37 Niederstätter H, Berger B, Kayser M, Parson W. 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. 2016;24:180-93.

38 Pickrahn I, Müller E, Zahrer W, Dunkelmann B, Cemper-Kiesslich J, Kreindl G, et al. Yfiler® Plus amplification kit validation and calculation of forensic parameters for two Austrian populations. Forensic Sci Int Genet. 2016;21:90-4.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

39 Junge A, Madea B. 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. 1999;101(3):195-201.
40 Lessig R, Edelmann J. Y chromosome polymorphisms and haplotypes in west Saxony (Germany). Int J Legal Med. 1998;111(4):215-8.
41 Schneider PM, Meuser S, Waiyawuth W, Seo Y, Rittner C. Tandem repeat structure of the duplicated Y-chromosomal STR locus DYS385 and frequency studies in the German and three Asian populations. Forensic Sci Int. 1998;97(1):61-70.
42 Anslinger K, Keil W, Weichhold G, Eisenmenger W. Y-chromosomal STR haplotypes in a population sample from Bavaria. Int J Legal Med. 2000;113(3):189-92.
43 Hidding M, Schmitt C. Haplotype frequencies and population data of nine Y-chromosomal STR polymorphisms in a German and a Chinese population. Forensic Sci Int. 2000;113(1-3):47-53.
44 Henke J, Henke L, Chatthopadhyay P, Kayser M, Dülmer M, Cleef S, et al. Application of Y-chromosomal STR haplotypes to forensic genetics. Croat Med J. 2001;42(3):292-7.
45 Schmidt U, Meier N, Lutz S. Y-chromosomal STR haplotypes in a population sample from southwest Germany (Freiburg area). Int J Legal Med. 2003;117(4):211-7.
46 Immel UD, Kleiber M, Klintschar M. Y chromosome polymorphisms and haplotypes in South Saxony-Anhalt (Germany). Forensic Sci Int. 2005;155(2-3):211-5.
47 Kayser M, Lao O, Anslinger K, Augustin C, Bargel G, Edelmann J, et al. Significant genetic differentiation between Poland and Germany follows present-day political borders, as revealed by Y-chromosome analysis. Hum Genet. 2005;117(5):428-43.
48 Hohoff C, Dewa K, Sibbing U, Hoppe K, Forster P, Brinkmann B. Y-chromosomal microsatellite mutation rates in a population sample from northwestern Germany. Int J Legal Med. 2007;121(5):359-63.
49 Rodig H, Grum M, Grimmecke HD. Population study and evaluation of 20 Y-chromosome STR loci in Germans. Int J Legal Med. 2007;121(1):24-7.
50 Grignani P, Peloso G, Fattorini P, Previderè C. Highly informative Y-chromosomal haplotypes by the addition of three new STRs DYS437, DYS438 and DYS439. Int J Legal Med. 2000;114(1-2):125-9.
51 Presciuttini S, Caglià A, Alù M, Asmundo A, Buscemi L, Caenazzo L, et al. Y-chromosome haplotypes in Italy: the GEFI collaborative database. Forensic Sci Int. 2001;122(2-3):184-8.
52 Ghiani ME, Vona G. Y-chromosome-specific microsatellite variation in a population sample from Sardinia (Italy). Coll Antropol. 2002;26(2):387-401.
53 Cerri N, Verzeletti A, Bandera B, De Ferrari F. Population data for 12 Y-chromosome STRs in a sample from Brescia (northern Italy). Forensic Sci Int. 2005;152(1):83-7.
54 Capelli C, Arredi B, Baldassari L, Boschi I, Brisighelli F, Caglià A, et al. A 9-loci Y chromosome haplotype in three Italian populations. Forensic Sci Int. 2006;159(1):64-70.
55 Turrina S, Atzei R, De Leo D. Y-chromosomal STR haplotypes in a Northeast Italian population sample using 17plex loci PCR assay. Int J Legal Med. 2006;120(1):56-9.
56 Robino C, Inturri S, Gino S, Torre C, Di Gaetano C, Crobu F, et al. Y-chromosomal STR haplotypes in Sicily. Forensic Sci Int. 2006;159(2-3):235-40.
57 Onofri V, Alessandrini F, Turchi C, Fraternale B, Buscemi L, Pesaresi M, et al. Y-chromosome genetic structure in sub-Apennine populations of Central Italy by SNP and STR analysis. Int J Legal Med. 2007;121(3):234-7.
58 Ferri G, Ceccardi S, Lugaresi F, Bini C, Ingravallo F, Cicognani A, et al. Male haplotypes and haplogroups differences between urban (Rimini) and rural area (Valmarecchia) in Romagna region (North Italy). Forensic Sci Int. 2008;175(2-3):250-5.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

59 Ferri G, Alù M, Corradini B, Radheshi E, Beduschi G. Slow and fast evolving markers typing in Modena males (North Italy). Forensic Sci Int Genet. 2009;3(2):e31-3.
60 Verzeletti A, Cerri N, Gasparini F, Poglio A, Mazzeo E, De Ferrari F. 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). 2009;11(1):37-40.
61 Rodríguez V, Tomàs C, Sánchez JJ, Castro JA, Ramon MM, Barbaro A, et al. Genetic sub-structure in western Mediterranean populations revealed by 12 Y-chromosome STR loci. Int J Legal Med. 2009;123(2):137-41.
62 Brisighelli F, Blanco-Verea A, Boschi I, Garagnani P, Pascali VL, Carracedo A, et al. Patterns of Y-STR variation in Italy. Forensic Sci Int Genet. 2012;6(6):834-9.
63 Piglionica M, Baldassarra SL, Giardina E, Stella A, D'Ovidio FD, Frati P, et al. Population data for 17 Y-chromosome STRs in a sample from Apulia (Southern Italy). Forensic Sci Int Genet. 2013;7(1):e3-4.
64 Robino C, Ralf A, Pasino S, De Marchi MR, Ballantyne KN, Barbaro A, et al. Development of an Italian RM Y-STR haplotype database: Results of the 2013 GEFI collaborative exercise. Forensic Sci Int Genet. 2015;15:56-63.
65 Rapone C, D'Atanasio E, Agostino A, Mariano M, Papaluca MT, Cruciani F, et al. Forensic genetic value of a 27 Y-STR loci multiplex (Yfiler(®) Plus kit) in an Italian population sample. Forensic Sci Int Genet. 2016;21:e1-5.
66 Sarno S, Tofanelli S, De Fanti S, Quagliariello A, Bortolini E, Ferri G, et al. Shared language, diverging genetic histories: high-resolution analysis of Y-chromosome variability in Calabrian and Sicilian Arbereshe. Eur J Hum Genet. 2016;24(4):600-6.
67 Lacerenza D, Aneli S, Di Gaetano C, Critelli R, Piazza A, Matullo G, et al. Investigation of extended Y chromosome STR haplotypes in Sardinia. Forensic Sci Int Genet. 2017;27:172-4.
68 Spiroski M, Arsov T, Krüger C, Willuweit S, Roewer L. Y-chromosomal STR haplotypes in Macedonian population samples. Forensic Sci Int. 2005;148(1):69-73.
69 Bosch E, Calafell F, González-Neira A, Flaiz C, Mateu E, Scheil HG, et al. Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns. Ann Hum Genet. 2006;70(Pt 4):459-87.
70 Jakovski Z, Nikolova K, Jankova-Ajanovska R, Marjanovic D, Pojskic N, Janeska B. Genetic data for 17 Y-chromosomal STR loci in Macedonians in the Republic of Macedonia. Forensic Sci Int Genet. 2011;5(4):e108-11.
71 Jankova R, Seidel M, Videtič Paska A, Willuweit S, Roewer L. Y-chromosome diversity of the three major ethno-linguistic groups in the Republic of North Macedonia. Forensic Sci Int Genet. 2019;42:165-70.
72 Veselinovic IS, Zgonjanin DM, Maletin MP, Stojkovic O, Djurendic-Brenesel M, Vukovic RM, et al. Allele frequencies and population data for 17 Y-chromosome STR loci in a Serbian population sample from Vojvodina province. Forensic Sci Int. 2008;176(2-3):e23-8.
73 Veselinovic I, Petric G, Vapa D. Genetic polymorphism of 17 Y chromosomal STRs in the Rusyn population sample from Vojvodina Province, Serbia. Int J Legal Med. 2014;128(2):273-4.
74 Zgonjanin D, Alghafri R, Antov M, Stojiljković G, Petković S, Vuković R, et al. Genetic characterization of 27 Y-STR loci with the Yfiler(®) Plus kit in the population of Serbia. Forensic Sci Int Genet. 2017;31:e48-e9.
75 Xu H, Wang CC, Shrestha R, Wang LX, Zhang M, He Y, et al. Inferring population structure and demographic history using Y-STR data from worldwide populations. Mol Genet Genomics. 2015;290(1):141-50.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

76 Iacovacci G, D'Atanasio E, Marini O, Coppa A, Sellitto D, Trombetta B, et al. Forensic data and microvariant sequence characterization of 27 Y-STR loci analyzed in four Eastern African countries. Forensic Sci Int Genet. 2017;27:123-31.
77 Zeng Z, Rowold DJ, Garcia-Bertrand R, Calderon S, Regueiro M, Li L, et al. Taiwanese aborigines: genetic heterogeneity and paternal contribution to Oceania. Gene. 2014;542(2):240-7.
78 Palencia-Madrid L, Baeta M, Villaescusa P, Nuñez C, de Pancorbo MM, Luis JR, et al. The Marquesans at the fringes of the Austronesian expansion. Eur J Hum Genet. 2019;27(5):801-10.
79 Kofi AE, Hakim HM, Khan HO, Ismail SA, Ghansah A, David AA, et al. Population data of 23 Y chromosome STR loci for the five major human subpopulations of Ghana. Int J Legal Med. 2020;134(4):1313-5.
80 Nagy M, Henke L, Henke J, Chatthopadhyay PK, Völgyi A, Zalán A, et al. Searching for the origin of Romanies: Slovakian Romani, Jats of Haryana and Jat Sikhs Y-STR data in comparison with different Romani populations. Forensic Sci Int. 2007;169(1):19-26.
81 Ghosh T, Kalpana D, Mukerjee S, Mukherjee M, Sharma AK, Nath S, et al. Genetic diversity of 17 Y-short tandem repeats in Indian population. Forensic Sci Int Genet. 2011;5(4):363-7.
82 Yadav B, Raina A, Dogra TD. Haplotype diversity of 17 Y-chromosomal STRs in Saraswat Brahmin Community of North India. Forensic Sci Int Genet. 2011;5(3):e63-70.
83 Roy S, Noor S, Haque I. Genetic analysis of 17 Y-chromosomal STRs haplotypes of three ethnic groups residing in West Bengal, India. Forensic Sci Int Genet. 2012;6(1):e5-7.
84 Gao T, Yun L, Gu Y, He W, Wang Z, Hou Y. Phylogenetic analysis and forensic characteristics of 12 populations using 23 Y-STR loci. Forensic Sci Int Genet. 2015;19:130-3.
85 Perez-Benedico D, Chennakrishnaiah S, Gayden T, Rowold DJ, Garcia-Bertrand R, Herrera RJ. Y-STR markers from Ladakh in the Himalayas. Leg Med (Tokyo). 2016;21:29-32.
86 Raghunath R, Krishnamoorthy K, Balasubramanian L, Kunka Mohanram R. Genetic portrait of Tamil non-tribal and Irula tribal population using Y chromosome STR markers. Int J Legal Med. 2016;130(2):367-9.
87 Shrivastava P, Jain T, Trivedi VB. A genetic portrait of Oraon Indian tribe drawn with 15 autosomal and 17 Y chromosomal STR markers. Int J Legal Med. 2016;130(5):1185-6.
88 Shrivastava P, Jain T, Trivedi VB. Haplotype data for 17 Y-STR loci in the population of Madhya Pradesh, India. Forensic Sci Int Genet. 2017;26:e31-e2.
89 Jain T, Shrivastava P, Trivedi VB. Genetic portrait of Majhi tribe of Chhattisgarh, India based on 15 autosomal STRs and 23 Y-STRs. Int J Legal Med. 2017;131(1):113-4.
90 Imam J, Reyaz R, Singh RS, Bapuly AK, Shrivastava P. Genomic portrait of population of Jharkhand, India, drawn with 15 autosomal STRs and 17 Y-STRs. Int J Legal Med. 2018;132(1):139-40.
91 Singh M, Sarkar A, Nandineni MR. A comprehensive portrait of Y-STR diversity of Indian populations and comparison with 129 worldwide populations. Sci Rep. 2018;8(1):15421.
92 Singh M, Sarkar A, Kumar D, Nandineni MR. The genetic affinities of Gujjar and Ladakhi populations of India. Sci Rep. 2020;10(1):2055.
93 Mohapatra BK, Chauhan K, Shrivastava P, Sharma A, Dagar S, Kaitholia K. Haplotype data for 17 Y-STR loci in the population of Himachal Pradesh, India. Int J Legal Med. 2019;133(5):1401-2.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

94 Kumar A, Kumar R, Kumawat RK, Mathur B, Shrivastava P, Chaubey G, et al. Genetic portrait study for 23 Y-STR loci in the population of Rajasthan, India. Int J Legal Med. 2020;134(5):1691-3.
95 Kumawat RK, Shrivastava P, Shrivastava D, Mathur GK. Molecular diversity of 23 Y-STR genetic markers in the population of Rajasthan, India. Meta Gene. 2020;24:100694.
96 Shrivastava P, Kushwaha K, Kumawat R, Chauhan T, Saiz M, Lorente JA, et al. Untangled the genetic structure of Kahar and Tharu, using 23 Y chromosomal paternal lineage markers. Meta Gene. 2020;26:100797.
97 Joshi B, Lemtur S, Humtsoe M, Verma K, Kumawat RK, Kushwaha P, et al. Genetic portrait of 23 Y-STR loci in the Naga tribes of Nagaland, India. Int J Legal Med. 2021.
98 Sahoo S, Samal R, Behera S, Biswas S, Dixit S, Kumawat RK, et al. Genomic insight into Y-STR diversity in the population of Odisha, India. Int J Legal Med. 2021;135(5):1771-2.
99 Lessig R, Edelmann J. Population data of Y-chromosomal STRs in Lithuanian, Latvian and Estonian males. Forensic Sci Int. 2001;120(3):223-5.
100 Ruzgaitė G, Čaplinskienė M, Baranovienė R, Jankauskienė J, Kukienė J, Savanevskytė K, et al. Forensic application of Y-chromosomal STR analysis in Lithuanian population. Biologija. 2015;61(2):60-72.
101 Jankauskiene J, Kukiene J, Ivanova V, Aleknaviciute G. Population data and forensic genetic evaluation with the Yfiler™ Plus PCR Amplification kit in the Lithuanian population. Forensic Science International: Genetics Supplement Series. 2017;6:e606-e7.
102 Gutiérrez-Alarcón AB, Moguel-Torres M, León-Jiménez AK, Cuéllar-Nevárez GE, Rangel-Villalobos H. Allele and haplotype distribution for 16 Y-STRs (AmpFlSTR Y-filer kit) in the state of Chihuahua at North Center of Mexico. Leg Med (Tokyo). 2007;9(3):154-7.
103 Luna-Vázquez A, Vilchis-Dorantes G, Aguilar-Ruiz MO, Bautista-Rivas A, Pérez-García A, Orea-Ochoa R, et al. Haplotype frequencies of the PowerPlex Y system in a Mexican-Mestizo population sample from Mexico City. Forensic Sci Int Genet. 2008;2(1):e11-3.
104 Salazar-Flores J, Dondiego-Aldape R, Rubi-Castellanos R, Anaya-Palafox M, Nuño-Arana I, Canseco-Avila LM, et al. Population structure and paternal admixture landscape on present-day Mexican-Mestizos revealed by Y-STR haplotypes. Am J Hum Biol. 2010;22(3):401-9.
105 Perez-Benedico D, La Salvia J, Zeng Z, Herrera GA, Garcia-Bertrand R, Herrera RJ. Mayans: a Y chromosome perspective. Eur J Hum Genet. 2016;24(9):1352-8.
106 Ramos-González B, Aguilar-Velázquez JA, de Lourdes Chávez-Briones M, Del Rocío Escareño-Hernández M, Alfaro-Lopez E, Rangel-Villalobos H. Genetic population data of three Y-STR genetic systems in Mexican-Mestizos from Monterrey, Nuevo León (Northeast, Mexico). Forensic Sci Int Genet. 2017;29:e21-e2.
107 López-Ramírez YL, Aguilar-Velázquez JA, López-Armenta M, Ruiz-Hernández M, Rangel-Villalobos H. Paternal lineages and forensic parameters based on 23 Y-STRs (Powerplex® Y23) in Mestizo males from Mexico City. Int J Legal Med. 2020;134(1):199-202.
108 Martinez B, Catelli L, Romero M, Okolie V, Keshinro S, Carvalho E, et al. Forensic evaluation of 27 y-str haplotypes in a population sample from nigeria. Forensic Science International: Genetics Supplement Series. 2017;6:e289-e91.
109 Ilyas M, Shahzad MS, Israr M, Jafri SS, Shafique M, Zar MS, et al. Y-Chromosomal STR Haplotype Profiling in Yousafzai's living in Swat Valley Pakistan. 22nd Congress of the IALM; July 5-8; Istanbul, Turkey: Medimond; 2012. p. 357-63.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

110 Lee EY, Shin KJ, Rakha A, Sim JE, Park MJ, Kim NY, et al. Analysis of 22 Y chromosomal STR haplotypes and Y haplogroup distribution in Pathans of Pakistan. Forensic Sci Int Genet. 2014;11:111-6.
111 Perveen R, Rahman Z, Shahzad MS, Israr M, Shafique M, Shan MA, et al. Y-STR haplotype diversity in Punjabi population of Pakistan. Forensic Sci Int Genet. 2014;9:e20-1.
112 Perveen R, Shahid AA, Shafique M, Shahzad M, Husnain T. Genetic variations of 15 autosomal and 17 Y-STR markers in Sindhi population of Pakistan. Int J Legal Med. 2017;131(5):1239-40.
113 Adnan A, Rakha A, Noor A, van Oven M, Ralf A, Kayser M. Population data of 17 Y-STRs (Yfiler) from Punjabis and Kashmiris of Pakistan. Int J Legal Med. 2018;132(1):137-8.
114 Ullah I, Olofsson JK, Margaryan A, Ilardo M, Ahmad H, Sikora M, et al. High Y-chromosomal Differentiation Among Ethnic Groups of Dir and Swat Districts, Pakistan. Ann Hum Genet. 2017;81(6):234-48.
115 Javed F, Sumbal S, Shafique M, Shahid AA, Shahzadi A, Rani N, et al. Male individualization using 12 rapidly mutating Y-STRs in Araein ethnic group and shared paternal lineage of Pakistani population. Int J Legal Med. 2018;132(6):1621-4.
116 Adnan A, Rakha A, Kasim K, Noor A, Nazir S, Hadi S, et al. Genetic characterization of Y-chromosomal STRs in Hazara ethnic group of Pakistan and confirmation of DYS448 null allele. Int J Legal Med. 2019;133(3):789-93.
117 Adnan A, Rakha A, Ameen F, Alarfaj AA, Almansob A, Wang CC, et al. Genetic structure and forensic characteristics of Saraiki population from Southern Punjab, Pakistan, revealed by 20 Y-chromosomal STRs. Int J Legal Med. 2020;134(3):977-9.
118 Aslam MA, Hussain M, Khan K, Zahra FT, Shafique M, Javeed S. Haplotype diversity of 17 Y-STRs in Sheikh population of Punjab. Int J Legal Med. 2020;134(4):1325-6.
119 Perveen R, Shahid AA, Shafique M, Shehzad M, Akram M. Kashmiris phylogenetic depictions through uniparental and biparental genetic markers. Int J Legal Med. 2020;134(4):1311-2.
120 Zahra FT, Hussain M, Khan K, Aslam MA, Shafique M, Rubab A, et al. Genetic polymorphism of Y-chromosomal STRs in Gujjar population of Punjab. Int J Legal Med. 2020;134(4):1333-4.
121 Adnan A, Rakha A, Nazir S, Alghafri R, Hassan Q, Wang CC, et al. Forensic features and genetic legacy of the Baloch population of Pakistan and the Hazara population across Durand line revealed by Y-chromosomal STRs. Int J Legal Med. 2021;135(5):1777-84.
122 Ploski R, Wozniak M, Pawlowski R, Monies DM, Branicki W, Kupiec T, et al. Homogeneity and distinctiveness of Polish paternal lineages revealed by Y chromosome microsatellite haplotype analysis. Hum Genet. 2002;110(6):592-600.
123 Pawłowski R, Dettlaff-Kakol A. Population data of nine Y-chromosomal STR loci in northern Poland. Forensic Sci Int. 2003;131(2-3):209-13.
124 Pepinski W, Niemcunowicz-Janica A, Skawronska M, Koc-Zorawska E, Janica J, Soltyszewski I. Y-chromosome STR haplotypes and alleles in the population sample of Old Believers residing in the Northeastern Poland. Forensic Sci Int. 2004;143(1):65-8.
125 Pepinski W, Niemcunowicz-Janica A, Skawronska M, Koc-Zorawska E, Janica J, Soltyszewski I. Y-chromosome STR haplotypes in a population sample of the Byelorussian minority living in the northeastern Poland. Forensic Sci Int. 2004;140(1):117-21.
126 Janica J, Pepinski W, Niemcunowicz-Janica A, Skawronska M, Aleksandrowicz-Bukin M, Ptaszynska-Sarosiek I, et al. Y-chromosome STR haplotypes and alleles in the ethnic

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

group of Polish Tatars residing in the Northeastern Poland. Forensic Sci Int. 2005;150(1):91-5.

127 Pepinski W, Skawronska M, Niemcunowicz-Janica A, Ptaszynska-Sarosiek I, Koc-Zorawska E, Janica J, et al. Population genetics of Y-chromosome STRs in a population sample of the Lithuanian minority residing in the northeastern Poland. Forensic Sci Int. 2005;153(2-3):264-8.

128 Rebała K, Szczerkowska Z. Polish population study on Y chromosome haplotypes defined by 18 STR loci. Int J Legal Med. 2005;119(5):303-5.

129 Wolańska-Nowak P, Branicki W, Parys-Proszek A, Kupiec T. 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. 2009;4(1):e43-4.

130 Spólnicka M, Dąbrowska J, Szabłowska-Gnap E, Pałeczka A, Jabłońska M, Zbieć-Piekarska R, et al. Intra- and inter-population analysis of haplotype diversity in Yfiler(®) Plus system using a wide set of representative data from Polish population. Forensic Sci Int Genet. 2017;28:e22-e5.

131 Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, et al. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. Am J Hum Genet. 2000;67(6):1526-43.

132 Nasidze I, Schädlich H, Stoneking M. Haplotypes from the Caucasus, Turkey and Iran for nine Y-STR loci. Forensic Sci Int. 2003;137(1):85-93.

133 Lessig R, Edelmann J, Kleemann WJ, Kozhemyako V. Population data of Y-chromosomal STRs in Russian males of the Primorye region population. Forensic Sci Int. 2006;159(1):71-6.

134 Pakendorf B, Novgorodov IN, Osakovskij VL, Danilova AP, Protod'jakonov AP, Stoneking M. Investigating the effects of prehistoric migrations in Siberia: genetic variation and the origins of Yakuts. Hum Genet. 2006;120(3):334-53.

135 Woźniak M, Derenko M, Malyarchuk B, Dambueva I, Grzybowski T, Miścicka-Sliwka D. Allelic and haplotypic frequencies at 11 Y-STR loci in Buryats from South-East Siberia. Forensic Sci Int. 2006;164(2-3):271-5.

136 Pakendorf B, Novgorodov IN, Osakovskij VL, Stoneking M. Mating patterns amongst Siberian reindeer herders: inferences from mtDNA and Y-chromosomal analyses. Am J Phys Anthropol. 2007;133(3):1013-27.

137 Roewer L, Krüger C, Willuweit S, Nagy M, Rodig H, Kokshunova L, et al. Y-chromosomal STR haplotypes in Kalmyk population samples. Forensic Sci Int. 2007;173(2-3):204-9.

138. Roewer L, Willuweit S, Krüger C, Nagy M, Rychkov S, Morozowa I, et al. Analysis of Y chromosome STR haplotypes in the European part of Russia reveals high diversities but non-significant genetic distances between populations. Int J Legal Med. 2008;122(3):219-23.

139 Pimenoff VN, Comas D, Palo JU, Vershubsky G, Kozlov A, Sajantila A. Northwest Siberian Khanty and Mansi in the junction of West and East Eurasian gene pools as revealed by uniparental markers. Eur J Hum Genet. 2008;16(10):1254-64.

140 Thèves C, Balaresque P, Evdokimova LE, Timofeev IV, Alekseev AN, Sevin A, et al. Population genetics of 17 Y-chromosomal STR loci in Yakutia. Forensic Sci Int Genet. 2010;4(5):e129-30.

141 Trynova EG, Tsitovich TN, Vylegzhanina EY, Bandurenko NA, Parson W. Presentation of 17 Y-chromosomal STRs in the population of the Sverdlovsk region. Forensic Sci Int Genet. 2011;5(3):e101-4.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

142 Dudás E, Vágó-Zalán A, Vándor A, Saypasheva A, Pomozi P, Pamjav H. Genetic history of Bashkirian Mari and Southern Mansi ethnic groups in the Ural region. Mol Genet Genomics. 2019;294(4):919-30.
143 Semikhodskii A, Krassotkin Y, Makarova T, Zavarin V, Ilina V, Sutyagina D. Population genetic data and forensic parameters of the 27 Y-STR panel Yfiler(®) Plus in Russian population. Int J Legal Med. 2021;135(5):1785-7.
144 Alshamali F, Pereira L, Budowle B, Poloni ES, Currat M. Local population structure in Arabian Peninsula revealed by Y-STR diversity. Hum Hered. 2009;68(1):45-54.
145 Khubrani YM, Wetton JH, Jobling MA. Extensive geographical and social structure in the paternal lineages of Saudi Arabia revealed by analysis of 27 Y-STRs. Forensic Sci Int Genet. 2018;33:98-105.
146 Gené M, Borrego N, Xifró A, Piqué E, Moreno P, Huguet E. Haplotype frequencies of eight Y-chromosome STR loci in Barcelona (North-East Spain). Int J Legal Med. 1999;112(6):403-5.
147 Hurles ME, Veitia R, Arroyo E, Armenteros M, Bertranpetit J, Pérez-Lezaun A, et al. Recent male-mediated gene flow over a linguistic barrier in Iberia, suggested by analysis of a Y-chromosomal DNA polymorphism. Am J Hum Genet. 1999;65(5):1437-48.
148 Pestoni C, Cal ML, Lareu MV, Rodríguez-Calvo MS, Carracedo A. Y chromosome STR haplotypes: genetic and sequencing data of the Galician population (NW Spain). Int J Legal Med. 1998;112(1):15-21.
149 Aler M, Salas A, Sánchez-Diz P, Murcia E, Carracedo A, Gisbert M. Y-chromosome STR haplotypes from a Western Mediterranean population sample. Forensic Sci Int. 2001;119(2):254-7.
150 Kayser M, Krawczak M, Excoffier L, Dieltjes P, Corach D, Pascali V, et al. An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations. Am J Hum Genet. 2001;68(4):990-1018.
151 Zarrabeitia MT, Riancho JA, Sánchez-Diz P, Sánchez-Velasco P. 7-Locus Y chromosome haplotype profiling in a northern Spain population. Forensic Sci Int. 2001;123(1):78-80.
152 Gamero JJ, Romero JL, González JL, Carvalho M, Anjos MJ, Real FC, et al. Y-chromosome STR haplotypes in a southwest Spain population sample. Forensic Sci Int. 2002;125(1):86-9.
153 Martinez-Jarreta B, Nievas P, Abecia E, Hinojal R, Budowle B. Haplotype distribution of nine Y-chromosome STR-loci in two northern Spanish populations (Asturias and Aragon). J Forensic Sci. 2003;48(1):204-5.
154 Martín P, García-Hirschfeld J, García O, Gusmão L, García P, Albarrán C, et al. A Spanish population study of 17 Y-chromosome STR loci. Forensic Sci Int. 2004;139(2-3):231-5.
155 Tomàs C, Jiménez G, Picornell A, Castro JA, Ramon MM. Differential maternal and paternal contributions to the genetic pool of Ibiza Island, Balearic Archipelago. Am J Phys Anthropol. 2006;129(2):268-78.
156 Gaibar M, Esteban E, Moral P, Gómez-Gallego F, Santiago C, Bandrés F, et al. STR genetic diversity in a Mediterranean population from the south of the Iberian Peninsula. Ann Hum Biol. 2010;37(2):253-66.
157 Ambrosio B, Novelletto A, Hernandez C, Dugoujon JM, Fortes-Lima C, Rodriguez JN, et al. Y-STR genetic diversity in autochthonous Andalusians from Huelva and Granada provinces (Spain). Forensic Sci Int Genet. 2012;6(2):e66-71.
158 Valverde L, Köhnemann S, Rosique M, Cardoso S, Zarrabeitia M, Pfeiffer H, et al. 17 Y-STR haplotype data for a population sample of Residents in the Basque Country. Forensic Sci Int Genet. 2012;6(4):e109-11.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

159 Nuñez C, Baeta M, Fernández M, Zarrabeitia M, Martinez-Jarreta B, de Pancorbo MM. Highly discriminatory capacity of the PowerPlex(®) Y23 System for the study of isolated populations. Forensic Sci Int Genet. 2015;17:104-7.
160 Martinez-Cadenas C, Blanco-Verea A, Hernando B, Busby GB, Brion M, Carracedo A, et al. The relationship between surname frequency and Y chromosome variation in Spain. Eur J Hum Genet. 2016;24(1):120-8.
161 García O, Yurrebaso I, Mancisidor ID, López S, Alonso S, Gusmão L. Data for 27 Y-chromosome STR loci in the Basque Country autochthonous population. Forensic Sci Int Genet. 2016;20:e10-e2.
162 Saiz M, Alvarez-Cubero MJ, Lorente JA, Alvarez JC, Martinez-Gonzalez LJ. Genetic structure in the paternal lineages of South East Spain revealed by the analysis of 17 Y-STRs. Sci Rep. 2019;9(1):5234.
163 Ferragut JF, Ramon C, Castro JA, Amorim A, Alvarez L, Picornell A. Middle eastern genetic legacy in the paternal and maternal gene pools of Chuetas. Sci Rep. 2020;10(1):21428.
164 Luis JR, Palencia-Madrid L, Mendoza VC, Garcia-Bertrand R, de Pancorbo MM, Herrera RJ. The Y chromosome of autochthonous Basque populations and the Bronze Age replacement. Sci Rep. 2021;11(1):5607.
165 Gehrig C, Hochmeister M, Budowle B. Swiss allele frequencies and haplotypes of 7 Y-specific STRs. J Forensic Sci. 2000;45(2):436-9.
166 Haas C, Wangensteen T, Giezendanner N, Kratzer A, Bär W. Y-chromosome STR haplotypes in a population sample from Switzerland (Zurich area). Forensic Sci Int. 2006;158(2-3):213-8.
167 Ali N, Coulson-Thomas YM, Norton AL, Dixon RA, Williams DR. Announcement of population data: genetic data for 17 Y-STR AmpFℓSTR® Yfiler™ markers from an immigrant Pakistani population in the UK (British Pakistanis). Forensic Sci Int Genet. 2013;7(2):e40-2.
168 Aliferi A, Thomson J, McDonald A, Paynter VM, Ferguson S, Vanhinsbergh D, et al. UK and Irish Y-STR population data-A catalogue of variant alleles. Forensic Sci Int Genet. 2018;34:e1-e6.
169 Marjanovic D, Bakal N, Pojskic N, Kapur L, Drobnic K, Primorac D, et al. Allele frequencies for 15 short tandem repeat loci in a representative sample of Bosnians and Herzegovinians. Forensic Sci Int. 2006;156(1):79-81.
170 Kovacevic L, Tambets K, Ilumäe AM, Kushniarevich A, Yunusbayev B, Solnik A, et al. Standing at the gateway to Europe--the genetic structure of Western balkan populations based on autosomal and haploid markers. PLoS One. 2014;9(8):e105090.
171 Zaharova B, Andonova S, Gilissen A, Cassiman JJ, Decorte R, Kremensky I. Y-chromosomal STR haplotypes in three major population groups in Bulgaria. Forensic Sci Int. 2001;124(2-3):182-6.
172 Katsaloulis P, Tsekoura K, Vouropoulou M, Miniati P. Genetic population study of 11 Y chromosome STR loci in Greece. Forensic Sci Int Genet. 2013;7(3):e56-8.
173 Robino C, Gino S, Ricci U, Grignani P, Previdere C, Torre C. Y-chromosomal STR haplotypes in an Albanian population sample. Forensic Sci Int. 2002;129(2):128-30.
174 Zastera J, Roewer L, Willuweit S, Sekerka P, Benesova L, Minarik M. Assembly of a large Y-STR haplotype database for the Czech population and investigation of its substructure. Forensic Sci Int Genet. 2010;4(3):e75-8.
175 Dupuy BM, Andreassen R, Flønes AG, Tomassen K, Egeland T, Brion M, et al. Y-chromosome variation in a Norwegian population sample. Forensic Sci Int. 2001;117(3):163-73.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

176 Dupuy BM, Stenersen M, Lu TT, Olaisen B. Geographical heterogeneity of Y-chromosomal lineages in Norway. Forensic Sci Int. 2006;164(1):10-9.
177 Holmlund G, Nilsson H, Karlsson A, Lindblom B. Y-chromosome STR haplotypes in Sweden. Forensic Sci Int. 2006;160(1):66-79.
178 Karlsson AO, Wallerström T, Götherström A, Holmlund G. Y-chromosome diversity in Sweden - a long-time perspective. Eur J Hum Genet. 2006;14(8):963-70.
179 Mielnik-Sikorska M, Daca P, Woźniak M, Malyarchuk BA, Bednarek J, Dobosz T, et al. Genetic data from Y chromosome STR and SNP loci in Ukrainian population. Forensic Sci Int Genet. 2013;7(1):200-3.
180 Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4(4):406-25.
181 Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol. 2018;35(6):1547-9.
182 Athey TW. With Athey's Haplogroup Predictor [updated September 2021; cited 2021 September]. Available from: http://www.hprg.com/hapest5/.
183 Athey T. Haplogroup Prediction from Y-STR Values Using an Allele Frequency Approach. Journal of Genetic Genealogy. 2005;1:1-7.
184 Maisano Delser P, Ravnik-Glavač M, Gasparini P, Glavač D, Mezzavilla M. Genetic Landscape of Slovenians: Past Admixture and Natural Selection Pattern. Front Genet. 2018;9:551.
185 Peričić M, Lauc LB, Klarić IM, Rootsi S, Janićijevic B, Rudan I, et al. High-resolution phylogenetic analysis of southeastern Europe traces major episodes of paternal gene flow among Slavic populations. Mol Biol Evol. 2005;22(10):1964-75.
186 Battaglia V, Fornarino S, Al-Zahery N, Olivieri A, Pala M, Myres NM, et al. Y-chromosomal evidence of the cultural diffusion of agriculture in Southeast Europe. Eur J Hum Genet. 2009;17(6):820-30.
187 Kushniarevich A, Utevska O, Chuhryaeva M, Agdzhoyan A, Dibirova K, Uktveryte I, et al. Genetic Heritage of the Balto-Slavic Speaking Populations: A Synthesis of Autosomal, Mitochondrial and Y-Chromosomal Data. PLoS One. 2015;10(9):e0135820.
188 Myres NM, Rootsi S, Lin AA, Järve M, King RJ, Kutuev I, et al. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. Eur J Hum Genet. 2011;19(1):95-101.
189 Babić Jordamović N, Kojović T, Dogan S, Bešić L, Salihefendić L, Konjhodžić R, et al. Haplogroup Prediction Using Y-Chromosomal Short Tandem Repeats in the General Population of Bosnia and Herzegovina. Front Genet. 2021;12:671467.
190 Mahal DG, Matsoukas IG. The Geographic Origins of Ethnic Groups in the Indian Subcontinent: Exploring Ancient Footprints with Y-DNA Haplogroups. Front Genet. 2018;9:4.
191 Mathieson I, Alpaslan-Roodenberg S, Posth C, Szécsényi-Nagy A, Rohland N, Mallick S, et al. The genomic history of southeastern Europe. Nature. 2018;555(7695):197-203.
192 Mendez FL, Karafet TM, Krahn T, Ostrer H, Soodyall H, Hammer MF. Increased resolution of Y chromosome haplogroup T defines relationships among populations of the Near East, Europe, and Africa. Hum Biol. 2011;83(1):39-53.
193 Grugni V, Raveane A, Ongaro L, Battaglia V, Trombetta B, Colombo G, et al. Analysis of the human Y-chromosome haplogroup Q characterizes ancient population movements in Eurasia and the Americas. BMC Biol. 2019;17(1):3.

TABLE 1. Allele frequencies for the 27 Y-STR loci in the population of Croatia.

Allele/Locus	DYF387S1a	DYF387S1b	DYS19	DYS385a	DYS385b	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	DYS448	DYS449	DYS456	DYS458	DYS460	DYS481	DYS518	DYS533	DYS570	DYS576	DYS627	DYS635	YGATAH4
7													0,002														
8																			0,006								
9			0,002						0,020				0,079	0,006					0,097			0,012					
10				0,028					0,485				0,602	0,162					0,584			0,014					0,039
11				0,270	0,006	0,002			0,483	0,862			0,256	0,235					0,262			0,124					0,531
12			0,002	0,028	0,002	0,144			0,010	0,041	0,081		0,057	0,333			0,006		0,051			0,517					0,361
13			0,126	0,114	0,036	0,704			0,002	0,065	0,854	0,002	0,004	0,235			0,020	0,008				0,321					0,069
14			0,134	0,394	0,320	0,146				0,012	0,059	0,432		0,030			0,077	0,032				0,012					
15			0,249	0,059	0,406	0,002				0,018	0,006	0,465					0,509	0,249					0,006	0,022	0,014		
16			0,400	0,085	0,049	0,002				0,002		0,101					0,223	0,231					0,041	0,095	0,077		
17			0,083	0,022	0,075										0,002		0,150	0,331					0,178	0,280	0,136		
18			0,004		0,081										0,008		0,016	0,112					0,391	0,416	0,063		
19					0,018										0,357			0,036					0,217	0,128	0,105		
20					0,008										0,556			0,002		0,004			0,110	0,045	0,306	0,024	
21								0,004							0,073					0,083			0,041	0,010	0,187	0,116	
22								0,069							0,002					0,150			0,016	0,004	0,073	0,235	
23								0,148							0,002					0,178					0,032	0,499	
24								0,527												0,077					0,008	0,103	
25								0,237								0,002				0,063						0,022	
26								0,016								0,002				0,016						0,002	
27							0,004									0,014				0,037							
28							0,075									0,073				0,045							
29							0,178									0,144				0,059							
30							0,329									0,168				0,164							
31							0,321									0,233				0,103							
32							0,087									0,191				0,018							
33							0,004									0,128				0,004							
34	0,018						0,002									0,043											
35	0,144	0,034														0,002					0,002						
36	0,150	0,049																			0,018						
37	0,274	0,089																			0,051						
38	0,353	0,363																			0,124						
39	0,059	0,391																			0,292						
40	0,002	0,063																			0,254						
41		0,012																			0,160						
42																					0,067						
43																					0,028						
44																					0,004						

TABLE 2. Interpopulation comparison of the current data within five different regions genetic distance *R*st values and *P* values. Data was obtained from the relevant publications and accessed through the YHRD database.

Population	Hvar	Osijek	Pula	Split	Varaždin
Hvar	*	0.0097	0.0042	0.3240	0.0000
Osijek	0.0210	*	0.1138	0.1199	0.0000
Pula	0.0284	0.0069	*	0.0233	0.0260
Split	0.0009	0.0063	0.0180	*	0.0000
Varaždin	**0.0979**	**0.0551**	0.0166	**0.0821**	*

TABLE 3. Interpopulation comparison of the current data with 23 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	Croatia	Austria	Belgium	Denmark	Ethiopia	French Polynesia	Germany	Ghana	Hungary	India	Italy	Lithuania	Mexico	Nigeria	North Macedonia	Pakistan	Poland	Russian Federation	Saudi Arabia	Serbia	Slovenia	Spain	Switzerland	United Kingdom
Croatia	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0055	0.0000	0.0000	0.0000	0.0000
Austria	0.1255	-	0.0000	0.0024	0.0000	0.0000	0.0052	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Belgium	0.2031	0.0197	-	0.0000	0.0000	0.0000	0.0027	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.4444	0.0000
Denmark	0.1290	0.0099	0.0412	-	0.0000	0.0000	0.0013	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Ethiopia	0.1704	0.2165	0.2801	0.2495	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
French Polynesia	0.1878	0.1419	0.2024	0.1934	0.2763	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Germany	0.1461	0.0045	0.0104	0.0102	0.2501	0.1709	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Ghana	0.2787	0.3710	0.4691	0.4133	0.2459	0.3321	0.4094	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0012	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Hungary	0.0482	0.0757	0.1484	0.0879	0.1303	0.1706	0.0947	0.3044	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0000	0.0000	0.0000
India	0.0961	0.1023	0.1642	0.1154	0.1454	0.1681	0.1217	0.2779	0.0201	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Italy	0.2043	0.0249	0.0159	0.0575	0.2656	0.1578	0.0281	0.4213	0.1339	0.1529	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lithuania	0.1914	0.1028	0.1330	0.0847	0.2836	0.2703	0.0922	0.4656	0.1106	0.1226	0.1454	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mexico	0.1785	0.0949	0.1173	0.1287	0.2192	0.1469	0.1094	0.3530	0.1398	0.1410	0.1291	0.2279	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Nigeria	0.2414	0.3309	0.4230	0.3709	0.2023	0.2829	0.3721	0.0054	0.2563	0.2394	0.3853	0.4213	0.3107	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
North Macedonia	0.0720	0.0374	0.0848	0.0627	0.1212	0.1627	0.0548	0.3156	0.0348	0.0656	0.0780	0.1377	0.1012	0.2687	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Pakistan	0.0854	0.0531	0.1030	0.0519	0.2216	0.1213	0.0644	0.3449	0.0381	0.0503	0.1003	0.0963	0.1256	0.2998	0.0718	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Poland	0.0905	0.1294	0.1886	0.1076	0.2569	0.2681	0.1225	0.4045	0.0687	0.0972	0.2007	0.0684	0.2363	0.3703	0.1198	0.0762	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Russian Federation	0.0779	0.0677	0.1188	0.0648	0.1586	0.1833	0.0794	0.2947	0.0303	0.0593	0.1229	0.0704	0.1350	0.2609	0.0471	0.0689	0.0635	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Saudi Arabia	0.0990	0.1150	0.1891	0.1464	0.0732	0.1506	0.1523	0.2084	0.0562	0.0604	0.1695	0.2070	0.0956	0.1639	0.0577	0.1029	0.1896	0.0794	-	0.0000	0.0000	0.0000	0.0000	0.0002
Serbia	0.0097	0.1259	0.2101	0.1415	0.1231	0.2006	0.1525	0.2604	0.0449	0.0847	0.2057	0.2023	0.1647	0.2159	0.0517	0.1014	0.1189	0.0743	0.0669	-	0.0000	0.0000	0.0000	0.0000
Slovenia	0.0297	0.0671	0.1355	0.0615	0.1703	0.1821	0.0781	0.3378	0.0142	0.0514	0.1387	0.0932	0.1533	0.2887	0.0418	0.0334	0.0310	0.0275	0.0899	0.0423	-	0.0000	0.0000	0.0000
Spain	0.3283	0.1183	0.0490	0.1695	0.4077	0.3475	0.0887	0.5812	0.2931	0.2865	0.0760	0.2378	0.2156	0.5509	0.2032	0.2321	0.2990	0.2268	0.3483	0.3590	0.2795	-	0.0000	0.0000
Switzerland	0.1999	0.0138	-0.0003	0.0299	0.2845	0.1833	0.0075	0.4310	0.1421	0.1632	0.0180	0.1283	0.1143	0.3997	0.0840	0.1026	0.1816	0.1170	0.1850	0.2052	0.1314	0.0581	-	0.0000
United Kingdom	0.1122	0.0813	0.1442	0.1122	0.0780	0.1429	0.1167	0.2236	0.0562	0.0558	0.1344	0.1718	0.0994	0.1752	0.0438	0.0865	0.1744	0.0756	0.0181	0.0808	0.0809	0.3041	0.1438	-

TABLE 4. Interpopulation comparison of the current data with 20 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	Croatia	Austria	Belgium	Bosnia and Herzegovina	Bulgaria	Germany	Greece	Hungary	Italy	Poland	Serbia	Albania	Czech Republic	Estonia	Ireland	Lithuania	North Macedonia	Norway	Slovenia	Sweden	Ukraine
Croatia	*	0.0000	0.0000	0.0002	0.0089	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
Austria	0.0795	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0035	0.0000	0.0000	0.0000	0.0079	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Belgium	0.1429	0.0188	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0007	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Bosnia and Herzegovina	0.0076	0.0988	0.1671	*	0.0034	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Bulgaria	0.0144	0.0649	0.1467	0.0186	*	0.0000	0.1307	0.0001	0.0000	0.0000	0.0095	0.0041	0.0000	0.0000	0.0000	0.0000	0.0706	0.0000	0.0001	0.0000	0.0003
Germany	0.1036	0.0214	0.0173	0.1382	0.1206	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.2638	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Greece	0.0241	0.0519	0.1187	0.0269	0.0048	0.1025	*	0.0001	0.0008	0.0000	0.0001	0.0024	0.0000	0.0000	0.0000	0.0000	0.0785	0.0000	0.0000	0.0000	0.0000
Hungary	0.0238	0.0252	0.0623	0.0451	0.0354	0.0321	0.0283	*	0.0001	0.0000	0.0000	0.0000	0.0024	0.0000	0.0000	0.0000	0.0000	0.0000	0.0973	0.0000	0.0130
Italy	0.0659	0.0144	0.0441	0.0744	0.0465	0.0546	0.0223	0.0273	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Poland	0.1216	0.1049	0.1305	0.1687	0.1603	0.0580	0.1555	0.0679	0.1542	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Serbia	0.0186	0.1087	0.1870	0.0068	0.0145	0.1612	0.0182	0.0612	0.0735	0.2021	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Albania	0.0728	0.0960	0.1714	0.0585	0.0258	0.1723	0.0159	0.0865	0.0527	0.2372	0.0336	*	0.0000	0.0000	0.0000	0.0000	0.0112	0.0000	0.0000	0.0000	0.0000
Czech Republic	0.0767	0.0156	0.0252	0.1086	0.0875	0.0012	0.0704	0.0176	0.0390	0.0519	0.1330	0.1324	*	0.0000	0.0000	0.0000	0.0000	0.0018	0.0002	0.0000	0.0000
Estonia	0.1877	0.1038	0.0940	0.2198	0.2054	0.0623	0.1630	0.1043	0.1226	0.1052	0.2477	0.2263	0.0693	*	0.0000	0.0080	0.0000	0.0000	0.0000	0.0000	0.0000
Ireland	0.3178	0.1360	0.0852	0.3403	0.3446	0.1221	0.3169	0.2276	0.2137	0.2475	0.3677	0.3590	0.1671	0.2125	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lithuania	0.1706	0.1162	0.1109	0.2114	0.2008	0.0615	0.1754	0.0977	0.1484	0.0555	0.2415	0.2548	0.0652	0.0121	0.2139	*	0.0000	0.0000	0.0000	0.0000	0.0000
North Macedonia	0.0375	0.0747	0.1458	0.0312	0.0067	0.1358	0.0036	0.0528	0.0419	0.1887	0.0165	0.0067	0.1028	0.2029	0.3271	0.2178	*	0.0000	0.0000	0.0000	0.0000
Norway	0.1270	0.0391	0.0394	0.1616	0.1473	0.0160	0.1183	0.0500	0.0715	0.0862	0.1797	0.1913	0.0178	0.0553	0.1700	0.0670	0.1535	*	0.0000	0.0000	0.0000
Slovenia	0.0204	0.0376	0.0887	0.0463	0.0297	0.0466	0.0284	0.0025	0.0439	0.0639	0.0606	0.0868	0.0269	0.1196	0.2641	0.1046	0.0511	0.0589	*	0.0000	0.0281
Sweden	0.1209	0.0427	0.0447	0.1437	0.1301	0.0400	0.0894	0.0545	0.0449	0.1360	0.1546	0.1406	0.0363	0.0646	0.2141	0.0956	0.1176	0.0147	0.0674	*	0.0000
Ukraine	0.0195	0.0603	0.1135	0.0461	0.0405	0.0599	0.0384	0.0079	0.0619	0.0530	0.0640	0.1005	0.0396	0.1254	0.2892	0.1021	0.0632	0.0845	0.0060	0.0950	*

TABLE 5. Haplogroup composition in the five regions of Croatia.

	Haplogroup/number of haplotypes													
Region	I1	I2a	I2b	J1	J2a	J2b	R1a	R1b	G1	G2a	E1b1b	Q	T	L
Hvar	3	55			4	1	11	4		6	12	8		
Varaždin	8	18	3	1	4	3	38	9			16			
Split	5	53	2		2	3	20	5	1		11	1	2	
Pula	9	31	1	1	3	2	28	8		3	11		1	1
Osijek	12	45	1	1	5	2	29	7		1	6	1		
Total number of haplotypes	37	202	7	3	18	11	126	33	1	10	56	10	3	1

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

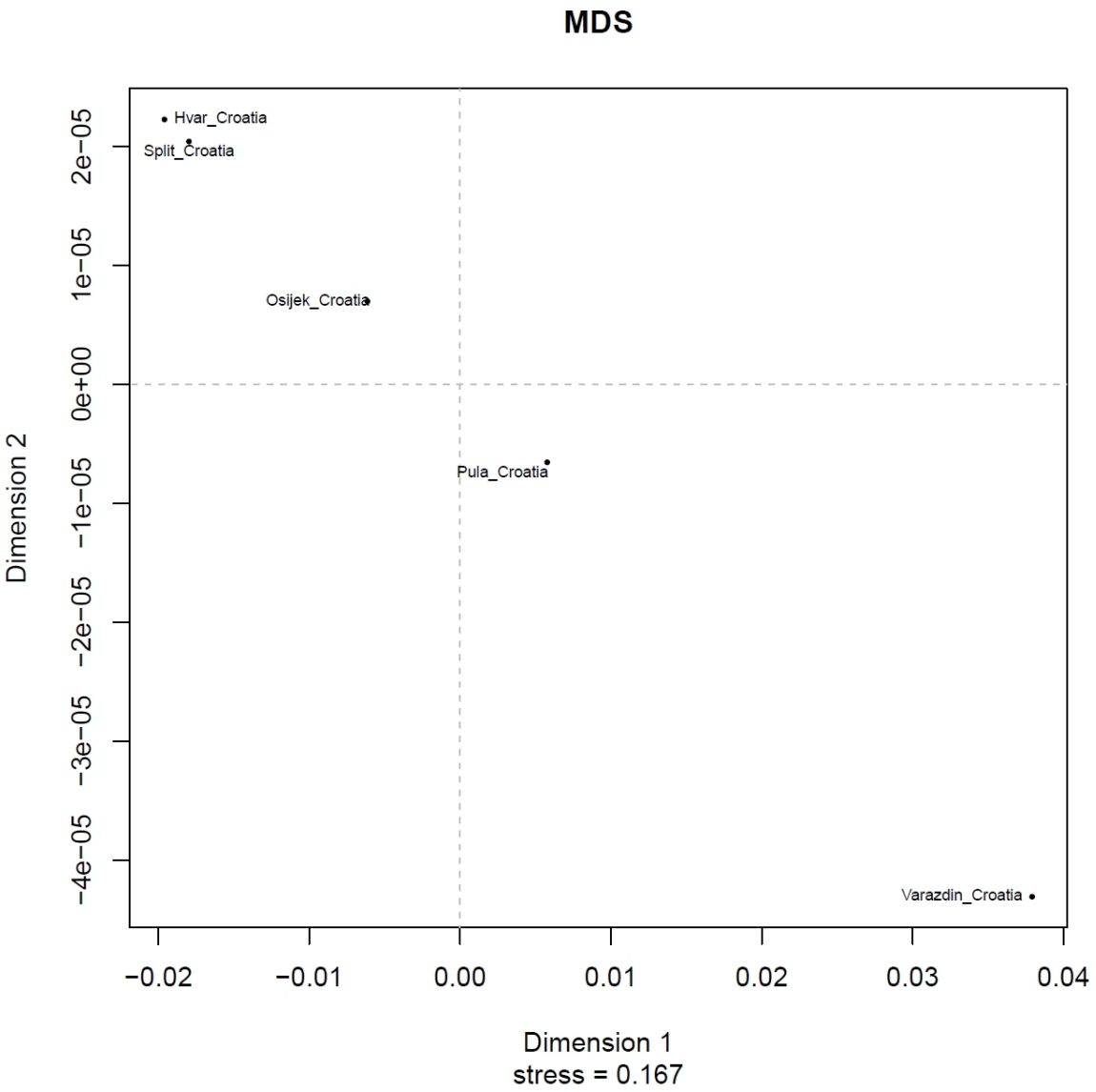

FIGURE 1. MDS plot showing genetic differentiation between five different regions of the analyzed populations of Croatia in two dimensions with original 27 Y-STR markers.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

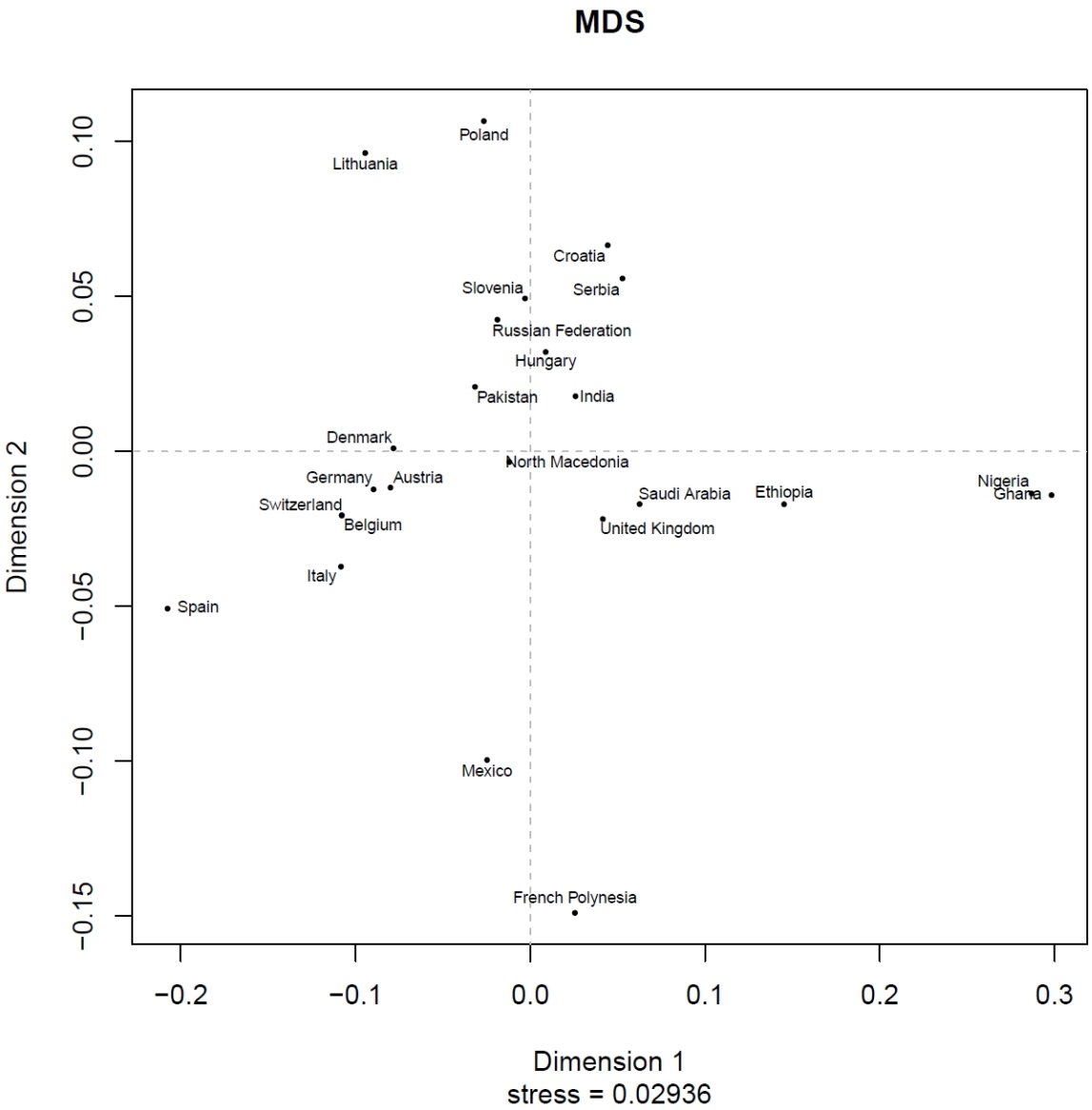

FIGURE 2. MDS plot showing genetic differentiation between 24 analysed populations in two dimensions with original 27 Y-STR markers.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

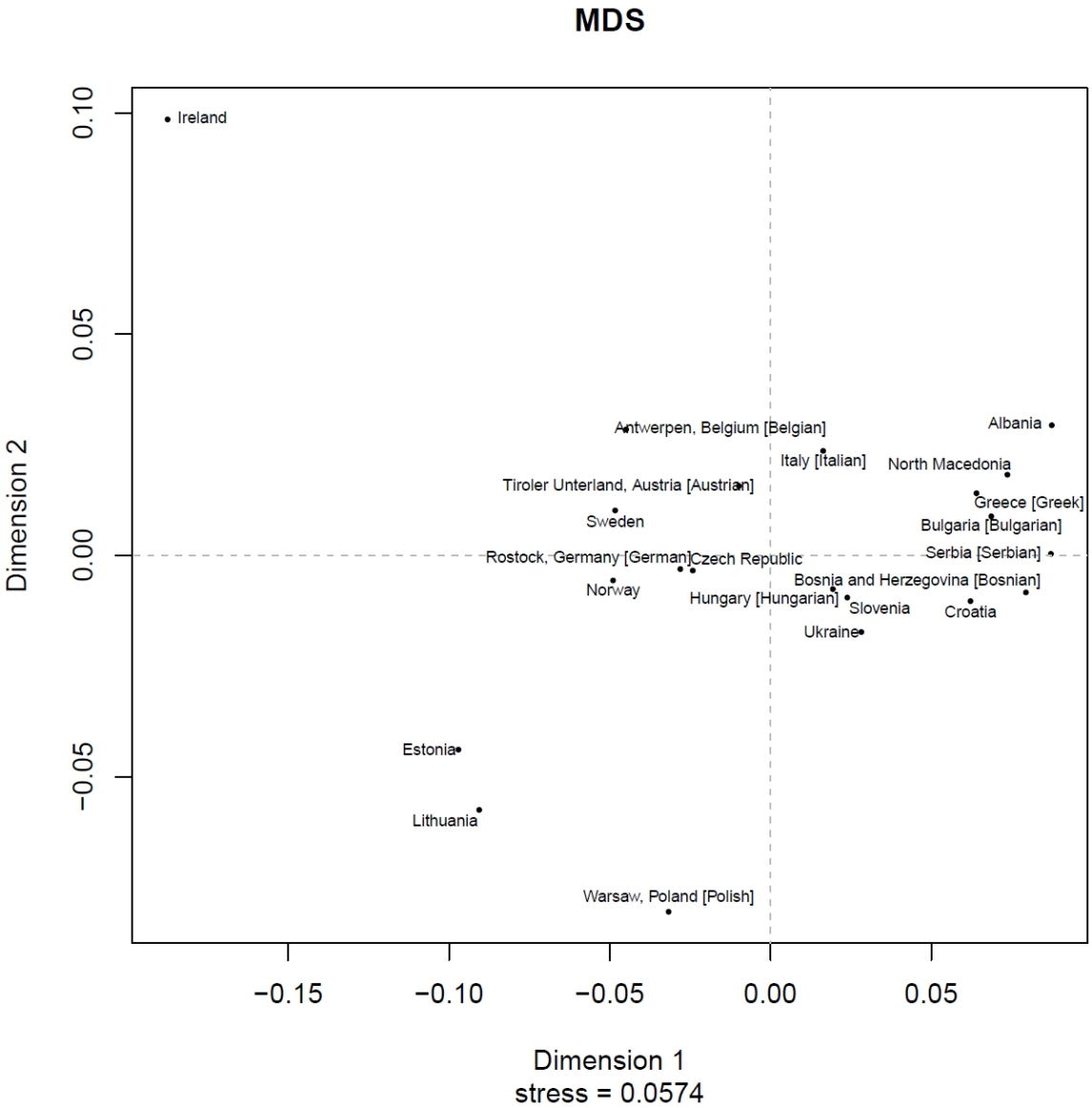

FIGURE 3. MDS plot showing genetic differentiation between the 21 analyzed populations in two dimensions with reduced number of 17 Y-STR markers.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

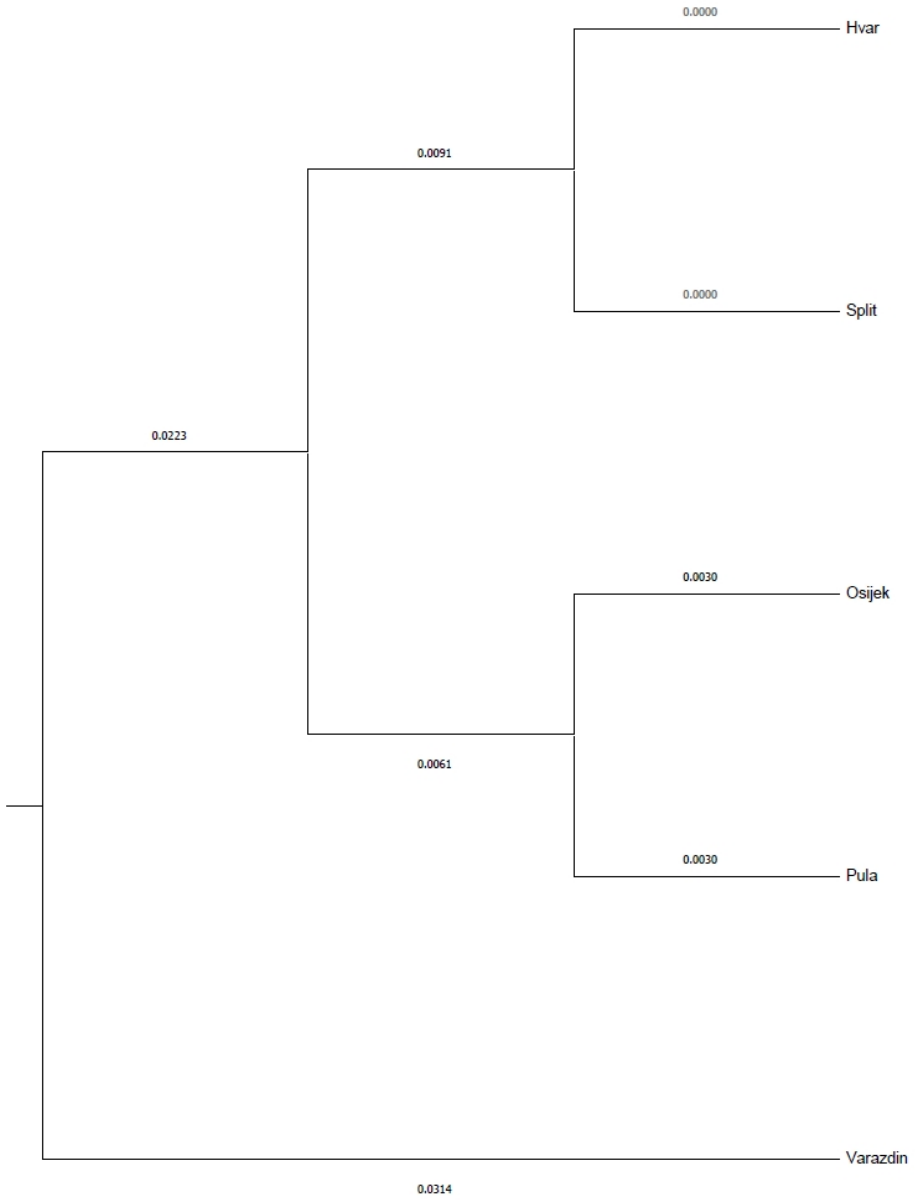

FIGURE 4a. Neighbour-joining phylogenetic tree showing genetics relationships and clustering between five different regions of the studied population of Croatia with original 27 Y-STR markers.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

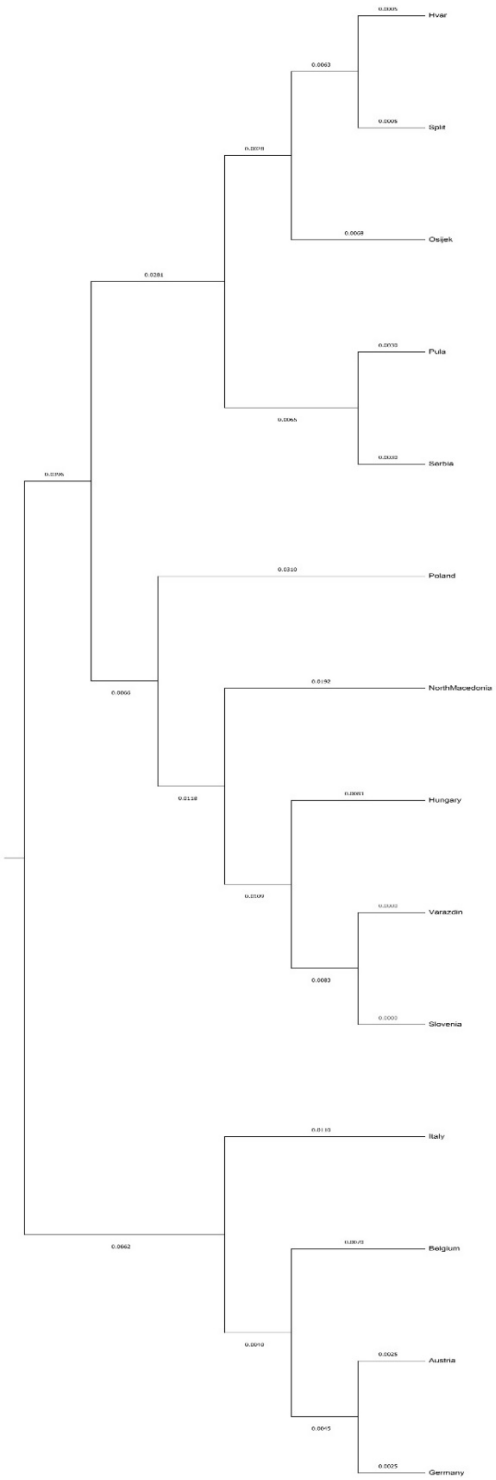

FIGURE 4b. Neighbor-joining phylogenetic tree showing genetics relationships and clustering between five different regions of the studied population of Croatia with original 27 Y-STR markers and selected neighboring countries.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

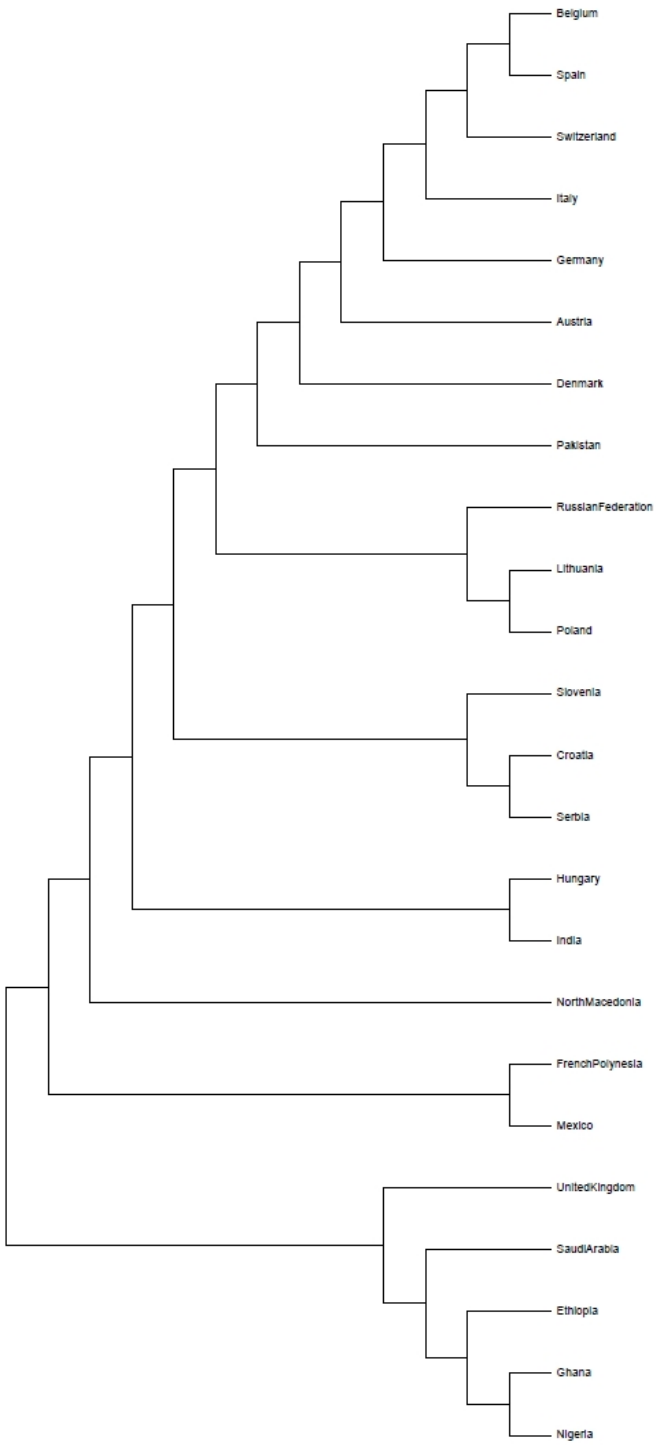

FIGURE 5. Neighbor-joining phylogenetic tree showing genetics relationships and clustering between the 24 studied populations with original 27 Y-STR markers.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

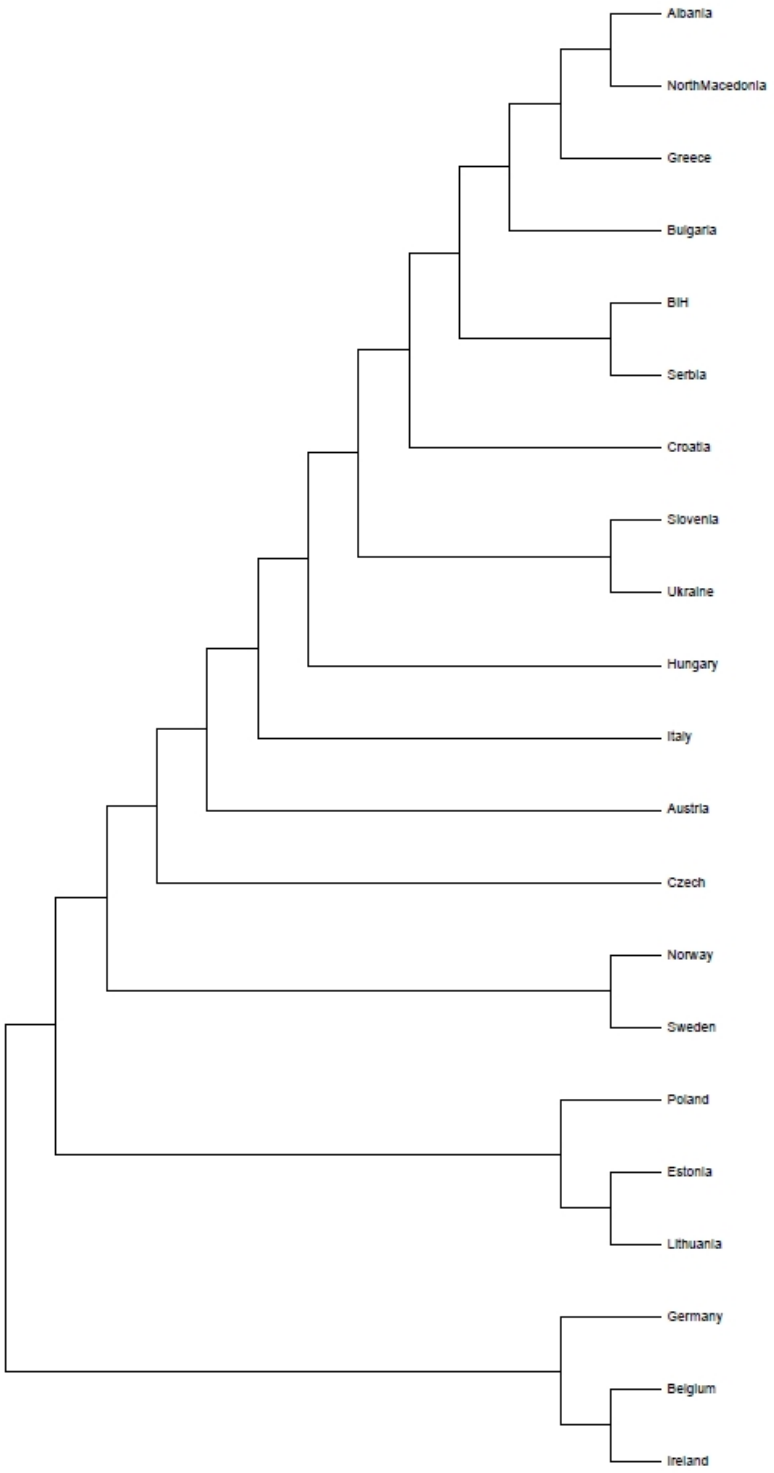

FIGURE 6. Neighbor-joining phylogenetic tree showing genetics relationships and clustering between the 21 studied populations with reduced number of 17 Y-STR markers.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

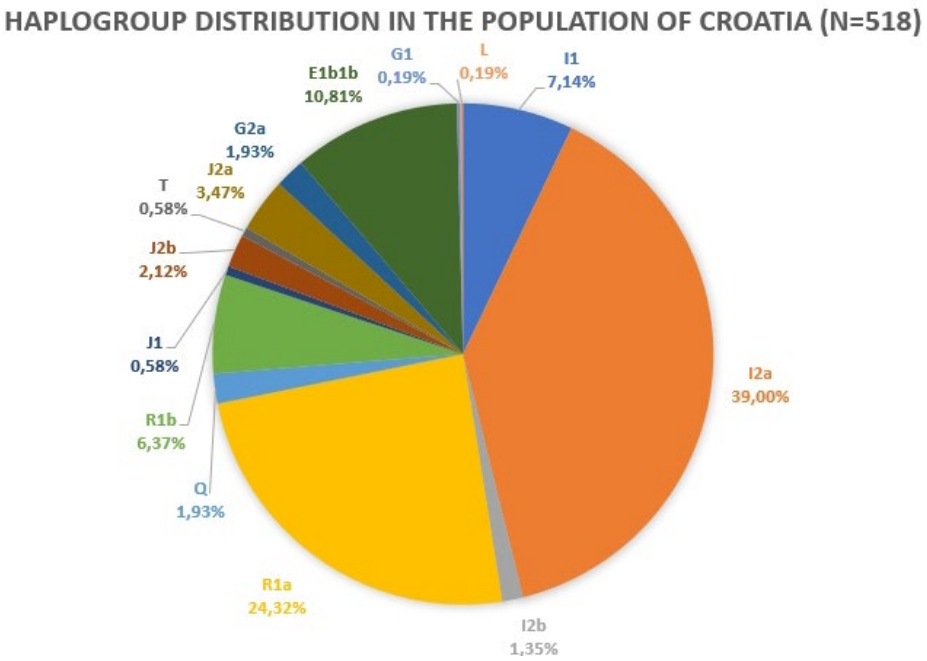

FIGURE 7. The results of Y haplogroup prediction in Croatian population (n=518).

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.21.485134; this version posted March 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

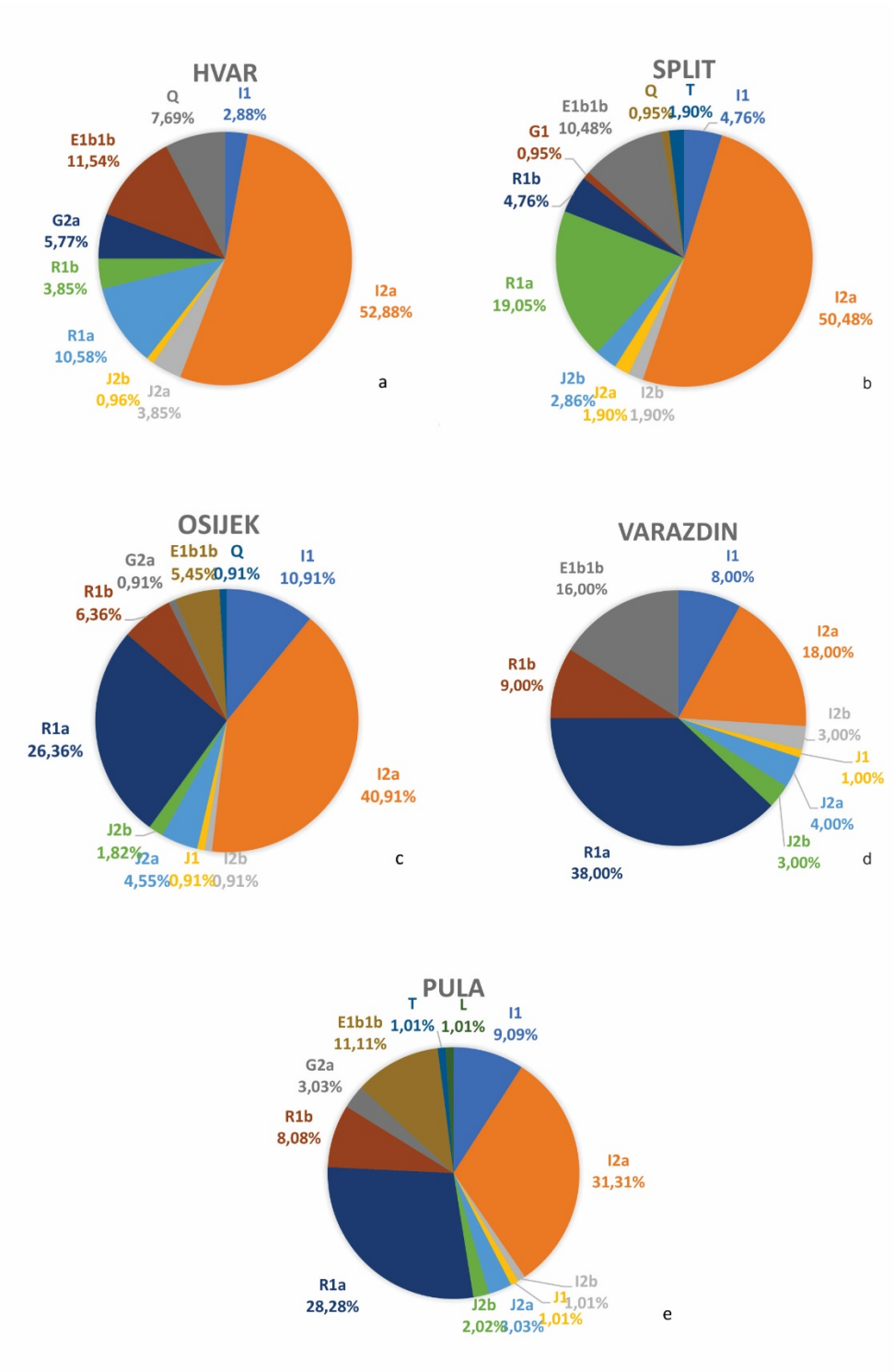

FIGURE 8. Y haplogroup frequency in five Croatian subpopulations: a. Hvar, b. Split, c. Osijek, d. Varaždin, e. Pula