Diversity of Y-Short Tandem Repeats in the Representative Sample of the Population of Canton Sarajevo Residents, Bosnia and Herzegovina

Merisa Ćenanović[1], Naris Pojskić[1], Lejla Kovačević[1], Mirela Džehverović[1], Jasmina Čakar[1], Dženisa Musemić[1] and Damir Marjanović[1,2]

[1] Institute for Genetic Engineering and Biotechnology, Sarajevo University, Sarajevo, Bosnia and Herzegovina
[2] Genos. d.o.o., Zagreb, Croatia

A B S T R A C T

In our previous population study, we have used twelve Y-chromosomal short tandem repeats loci incorporated in the PowerPlex®Y System to determine Y-STR diversity in B&H human population. With intent to obtain additional verification of the previously obtained results as well as to establish specific reference for a local B&H population, we have decided to test DNA samples collected from 100 unrelated healthy male Canton Sarajevo residents (from Sarajevo region) for the same twelve Y-linked short tandem repeats loci. Qiagen DNeasy^{TM} Tissue Kit (Qiagen, GmbH, Hilden, Germany) was used for DNA extraction from buccal swabs and PowerPlex®Y System (Promega Corp., Madison, WI) has been used to simultaneously amplify Y-STR loci by PCR. PowerPlex®Y System includes 12 STR loci: DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 and DYS439. The total PCR reaction volume was 5 µL. PCR amplifications were carried out in PE GeneAmp PCR System Thermal Cycler (ABI). Electrophoresis of the amplification products was preformed on an ABI PRISM 310 genetic analyzer (ABI, Foster City, CA) according to the manufacturer's recommendations. The raw data were compiled and analyzed using the accessory software: ABI PRISM®Data Collection Software and Genemapper® version 3.2. In addition, we have compared the obtained »Sarajevo« dataset with the data previously generated for the entire Bosnian and Herzegovinian population, as well as with the available data on geographically close (neighboring) European populations. The results of this study will be used as guidelines in additional improving of research into genetic relationship among recent local B&H populations, both isolated and open, which is a long term project in our country.

Key words: *Y-STRs, population genetics, genetic diversity, PowerPlex®Y System, Canton Sarajevo, Bosnia and Herzegovina*

Introduction

Research into Y chromosome markers has experienced tremendous growth in the past several years. Y-STRs are the most used Y chromosome markers in the forensic field due to simple typing and high level of diversity. STR typing involves simple and reliable polymerase chain reaction (PCR) technique and is tolerant of degraded samples[1]. Therefore, Y-STRs are applicable to a number of human genetic testing purposes such as forensic analysis of sexual assault evidence, missing person's investigations, deficient paternity testing, elucidation of historical issues, supplementing genealogical research[2] or like in this case, to population studies.

Modern Bosnia-Herzegovina is a multi-ethnic country, with a very turbulent history. Several archeological artifacts indicate that B&H territory has continuously been populated since the Paleolithic. Historical events that marked such a long period shaped a fascinating diversity of local human populations. Because of a large number of small, more or less isolated, indigenous popu-

lations, this area is marked as a very attractive one for population-genetic surveys of different levels and approaches[3].

In our previous population studies of Bosnian and Herzegovinian human populations, we have analyzed phenotypic markers, nuclear STR and bi-allelic markers[4–6], as well as mitochondrial HV regions[7,8] to generate B&H reference database. With intent to improve our database and to obtain more specific results for local populations for variety of DNA markers, we decided to analyze additional individuals from Canton Sarajevo area (Figure 1). The Sarajevo Canton is a canton of the Federation of Bosnia and Herzegovina, with Sarajevo as a capital city. It is consisted of 9 municipalities and is also one of the most populated Canton. The Canton contains 98% of Sarajevo's population, but a much smaller percentage of the official land area. According to the estimate of Federal Office of Statistics, Federation of Bosnia and Herzegovina (data collected from the www.fzs.ba, January 25, 2010), ethnically speaking, 79.6% of the Canton population are Bosniaks, 11.2% are Serbs and another 6.7% are Croats[2]. In 1991, Bosniaks formed 45% of the population, followed by Serbs with 38%, and Croats with 7% % (data collected from the official web site http://, September 25, 2009).

Estimation of genetic diversity at 12 Y-chromosomal short tandem repeats loci included in the PowerPlex® Y System would extend the existing database, and create more realistic view of the state of the genetic structure within Bosnian and Herzegovinian human population, in particular regarding the diversity among the isolated and non-isolated local populations. In addition, our intent is to estimate genetic distinctiveness of the Canton Sarajevo population within general B&H population and to compare it with populations of geographically neighboring countries.

Material and Methods

The total of 100 unrelated, healthy male individuals living in Canton Sarajevo has been sampled for the analysis. DNA was collected using buccal swabs. Genomic DNA was extracted using a Qiagen DNeasy™ Tissue Kit (Qiagen, GmbH, Hilden, Germany). 12 Y-chromosomal short tandem repeats loci (DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 and DYS439) incorporated in the PowerPlex®Y System (Promega Corp., Madison, WI), were amplified by PCR. The total volume of PCR reaction was 5 µL. The PCR amplifications have been carried out in PE GeneAmp PCR System Thermal Cycler (ABI, Foster City, CA, USA) according to the manufacturer's recommendations. Separation and detection of the amplification products was performed on an ABI PRISM® 310 Genetic Analyzer (ABI). Raw data have been compiled and analyzed using ABI PRISM®Data Collection Software and Genemapper® version 3.2 software. Within population diversity of Canton Sarajevo residents was estimated by calculating gene diversity (GD), allele fre-
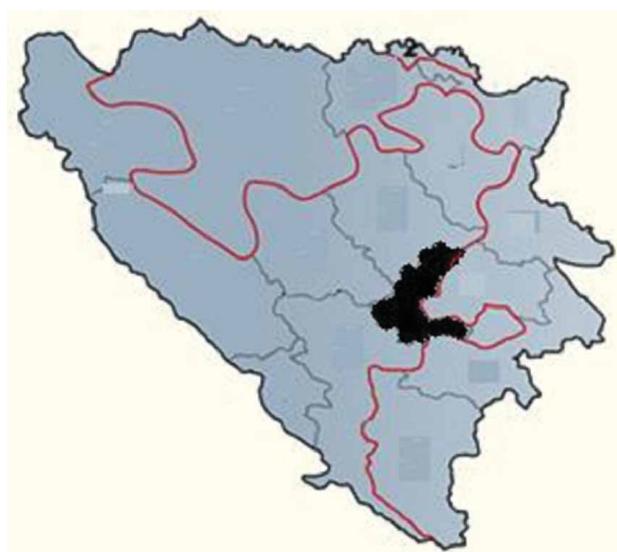

Fig. 1. Geographical position of Canton Sarajevo area.

quencies, number of alleles as well as haplotype diversity (HD)[9]. For comparison of Canton Sarajevo population with general B&H population (samples collected from various regions of country) exact test of population differentiation[10] and θ[11] have been calculated. In order to determine Canton Sarajevo population relationships with general B&H population[12,13] and populations from neighboring countries[14–18], pairwise F_{ST}[11], Reynolds et al. model of genetic distance[19], exact test of population differentiation and AMOVA[20] have been applied. Neighbor-joining dendrogram[21] based on genetic distance analysis, has been constructed. Since only 7 Y-STR loci (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393) were shared from all previously published results for neighboring populations, for purpose of appropriate pairwise F_{ST} and genetic distance analyses only above-mentioned loci have been included. Considered significance level for pairwise F_{ST} after Bonferroni's correction was $p<0.01$. The above listed calculations were made using Arlequin version 3.11[20], PowerMaker version 3.32[22] and MEGA4[23] software.

Results and Discussion

Y-STR haplotypes were generated for a sample of 100 unrelated, healthy male individuals living in Canton Sarajevo (Bosnia and Herzegovina) using PowerPlex®Y System kit (Promega Corp., Madison, WI) that includes 12 loci (DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 and DYS439). Within this pool, the totals of 81 different haplotypes were detected with 71 of them unique. Absolute frequencies of the remaining 10 haplotypes were two for six haplotypes, three for two haplotypes, five for one haplotype and six for one haplotype. Estimated HD was extremely high (1.00 ± 0.0014, Table 1). Allele frequencies and GD for each Y-STR locus analyzed in this study are

TABLE 1
POWERPLEX Y HAPLOTYPES WITH FREQUENCY HIGHER
THAN 1 IN B&H POPULATION AND THE OBSERVED
HAPLOTYPE DIVERSITY

PowerPlex® Y haplotypes	Repetitions
11-13-12-31-10-15-16-11-13-24-14-15	6
11-13-13-31-10-15-16-11-13-24-14-15	5
10-13-12-30-10-14-13-11-13-24-16-18	3
10-12-11-28-10-16-15-12-13-22-13-14	3
10-13-12-30-10-14-13-11-13-24-16-17	2
11-13-13-31-10-15-17-11-13-24-14-15	2
11-13-14-31-10-15-15-11-13-24-14-15	2
10-13-12-31-10-14-13-11-13-24-16-18	2
11-13-31-10-15-15-11-13-24-14-15	2
10-13-13-31-10-15-16-11-13-24-14-15	2
Haplotype diversity (h)	1.0000±0.0014

TABLE 2
GENE DIVERSITY, NUMBER OF ALLELES AND MAJOR ALLELE
FREQUENCY FOR OBSERVED Y STR LOCI IN CANTON
SARAJEVO POPULATION

Marker	Major Allele Frequency	Number of Alleles	Gene Diversity
DYS391	0.5500	3.0000	0.5038
DYS389I	0.7300	3.0000	0.4282
DYS439	0.4400	6.0000	0.6820
DYS389II	0.4600	5.0000	0.6954
DYS438	0.8100	5.0000	0.3306
DYS437	0.4900	3.0000	0.6154
DYS19	0.3600	5.0000	0.7410
DYS392	0.8400	5.0000	0.2842
DYS393	0.8200	4.0000	0.3126
DYS390	0.7100	5.0000	0.4694
DYS385a	0.4700	8.0000	0.7240
DYS385b	0.4600	8.0000	0.7282
Mean	0.5950	5.0000	0.5429

shown in Table 2. Estimated allele frequency ranged from 0.3600 (DYS19) to 0.8400 (DYS392). Highest number of alleles was detected at DYS385a and DYS385b loci (8) and the lowest at DYS391, DYS389I, DYS437 loci (3). GD of the loci ranged from 0.2842 (DYS392) to 0.7410 (DYS19). Mean GD value considering all 12 Y-STR loci was 0.5429, which indicated modest diversity of the detected haplotypes (Table 2).

We have compared Canton Sarajevo population data with general B&H population data[12,13]. The total number of different haplotypes was equal in the two analyzed populations (81). Also, similar number of unique ha-

plotypes was observed (69 for general B&H, 71 for Canton Sarajevo population). Most frequent haplotype in Canton Sarajevo population (11-13-12-31-10-15-16-11-13- -24- -14-15) was detected in six individuals while the same haplotype occurred only twice in general B&H population. Most frequent haplotype in general B&H population (10-13-12-30-10-14-13-11-13-24-16-18) was detected in five individuals while it occurred three times in Canton Sarajevo population. The lowest GD was detected at DYS392 Y-STR locus in both populations. The highest GD in local Canton Sarajevo population was detected at DYS19 locus. In general B&H population, the highest GD was detected at DYS385 locus, but in previous study of general B&H population, analysis collated results for both DYS385a and DYS385b which may have caused certain differences. If we exclude the collated data for DYS385 locus, the highest GD in B&H population was also detected at DYS19 locus. GD means for the observed populations were almost equal (0.5419 for general B&H, 0.5429 for Canton Sarajevo population). There was no significant difference in allele frequencies across Y-STR loci between Canton Sarajevo and general B&H populations (p>0.05).

Concurrence between the obtained results suggests that the local population of Canton Sarajevo, with respect to the detected haplotype and gene diversity, may be considered as a projection of general B&H population. Since this population represents the largest regional population in Bosnia and Herzegovina with emphasized migrational influx this is quite logical outcome.

In comparative analysis of Canton Sarajevo and general B&H populations with the populations of neighboring countries Reynolds et al. model of genetic distance[19] has been used, as well as pairwise F_{ST} and AMOVA. As expected, the Canton Sarajevo population has the lowest pairwise F_{ST} with general B&H population (–0.00189) as shown in Table 3. In comparisons with other observed populations, Canton Sarajevo population has low differentiation level, but significant (p<0.01) and the highest value of pF_{ST} was estimated against Kosovo population[14] (0.0327) When we exclude Canton Sarajevo population, the lowest pairwise F_{ST} was estimated between Macedonia[*] and Bulgaria[18] populations (0.00156) and the highest between populations of Croatia and Kosovo (0.03474, Table 3). It is indicative that Kosovo and Croatian populations[17] have all statistical significant pairwise F_{ST} values (p<0.01, Table 3). AMOVA (analysis of molecular variance) at 12 observed Y-STR loci, showed that 1.49% was attributed to the variation between the groups, –0.19% among populations within groups (Canton Sarajevo and general B&H populations), while 98.7% was within population variance. Estimation of genetic distance based on the data available for 7 (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393) Y-STR loci, as expected, showed the lowest genetic distance between local Canton Sarajevo and general B&H popula-

---

* unpublished data kindly provided by Zlatko Jakovski

TABLE 3
PAIRWISE F$_{ST}$ (BELOW DIAGONAL) AND ITS P-VALUE (ABOVE DIAGONAL) BETWEEN PAIRS OF THE OBSERVED POPULATIONS

POP	CS	BH	CRO	SRB	MAC	BULG	KOS
CS		0.83594	0.00000*	0.00195*	0.00684*	0.00195*	0.00000*
BH	–0.00189		0.00000*	0.05859	0.02734	0.04490	0.00000*
CRO	0.02030	0.01458		0.00000*	0.00000*	0.00000	0.00000*
SRB	0.00490	0.00168	0.00696		0.00000	0.02344	0.00000*
MAC	0.00819	0.00495	0.01969	0.00895		0.14844	0.00000*
BULG	0.00731	0.00287	0.01031	0.00200	0.00156		0.00000*
KOS	0.03275	0.02757	0.03474	0.02521	0.01878	0.01887	

*p<0.01, CS – Canton Sarajevo, BH – Bosnia and Herzegovina, CRO – Croatia, SRB – Serbia, MAC – Macedonia, BULG – Bulgaria, KOS – Kosovo

TABLE 4
GENETIC DISTANCE BASED ON 7 OF THE OBSERVED Y-STR LOCI

POP	BH	BULG	CRO	KOS	CS	MAC	SRB
BH							
BULG	0.0312						
CRO	0.0759	0.0866					
KOS	0.0842	0.0517	0.1320				
CS	0.0115	0.0362	0.0876	0.0886			
MAC	0.0332	0.0147	0.0924	0.0335	0.0337		
SRB	0.0181	0.0131	0.0605	0.0601	0.0224	0.0235	

CS – Canton Sarajevo, BH – Bosnia and Herzegovina, CRO – Croatia, SRB – Serbia, MAC – Macedonia, BULG – Bulgaria, KOS – Kosovo

tion (0.0115). The greatest genetic distance Canton Sarajevo has against Kosovo population (0.0886). When we exclude Canton Sarajevo population results of genetic distance analyses showed the lowest genetic distance between Bulgaria[18] and Serbia populations[15,16] (0.0131) and the highest between Croatia and Macedonia (0.0924) as shown in Table 4. Neighbor-joining tree constructed based on results of genetic distance analysis shows cluster of general BiH and Canton Sarajevo populations, since Macedonian, Kosovo and Bulgarian population constitute second and Croatian and Serbian populations third group (Figure 2).

Comparative analysis of the obtained results for Canton Sarajevo population, previous studies of general BiH population[12,13] and available data for neighboring countries populations[14–18] revealed no genetic differentiation between local Canton Sarajevo and general BiH population. It was not detected pronounced genetic differentiation between geographically close populations with the exception of Kosovo. Although, in most pairwise comparisons differentiation level was low, but significant (p<0.01). Similar results were previously obtained in the analysis of nuclear STR markers[13] and bi-allelic markers[6]. The observed differences could be attributed both to different demographic factors, such as different intensities, levels and directions of human migrations in the region, as well as to social aspects.

Conclusion

The analysis of 12 observed Y-STR loci suggests that Canton Sarajevo human population is not Y-STR specific, compared with B&H general population. This conclusion is expected since Canton Sarajevo population is large urban population that, as capitol city, consists of residents originated from all parts of B&H. In comparison with other neighboring countries (geographically closer populations), Canton Sarajevo population has the same position as general B&H population with no major Y-STR difference. The samples observed in this study can be included into B&H general population for further human population and forensic genetics studies.
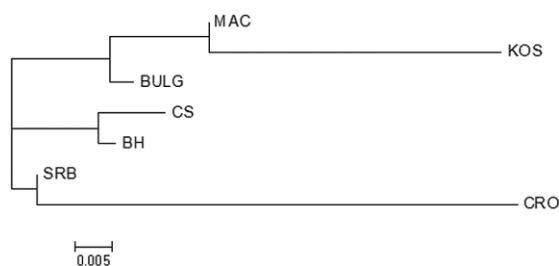

Fig. 2. NJ dendrogram based on Reynolds standard genetic distance calculated using the available data for 7 Y-STR loci.

Once again, within Bosnia and Herzegovina and other countries in Balkans region it has been noted very high level of human population diversity with quite similar genetic structure. These could be shaped by complex social and historical processes in a specific natural-geographical environment. Such environment eventually caused development of very particular genetic entities. In those terms, Bosnia and Herzegovina represents a unique region and source of genetic information discovered in previous but also, yet to be discovered in future studies.

R E F E R E N C E S

1. GUSMAO L, CARRACEDO A, Profiles in DNA, 6 (2003) 3. — 2. BUTLER JM, Forensics Sci Rev, 15 (2003) 91. — 3. MARJANOVIĆ D, POJSKIĆ N, KAPUR L, HAVERIČ S, DURMIĆ-PAŠIĆ A, BAJROVIĆ K, HADŽISELIMOVIĆ R, Coll Antrop, 32 (2008) 315. — 4. MARJANOVIĆ D, KAPUR L, DROBNIČ K, BUDOWLE B, POJSKIĆ N, HADŽISELIMOVIĆ R, Hum Biol, 78 (2004) 15. — 5. MARJANOVIĆ D, BAKAL N, POJSKIĆ N, J Forensic Sci, 50 (2005) 223. — 6. MARJANOVIĆ D, FOMARINO S, BAKAL N, MONTAGNA S, HADŽISELIMOVIĆ R, PRIMORAC D, VIDOVIĆ S, POJSKIĆ N, BATTAGLIA V, ACHILLI A, TORRONI A, ANDJELOVIĆ S, DROBNIČ K, SANTACHIARA-BENERECCETI S, SEMINO O, Ann Hum Genet, 69 (2005) 757. — 7. KAPUR L, POJSKIĆ N, MARJANOVIĆ D, DURMIĆ A, BAJROVIĆ K, HADŽISELIMOVIĆ R, MtDNA diversity in local human populations after war resettlement. In: Proceedings (2nd International Symposium, Paris, 2003). — 8. POJSKIĆ N, MtDNA hypervariable regions in determination of local human populations diversity after war ressettlement. In: Proceedings (European Human Genetics Conference, Prague, 2005). — 9. NEI M, Molecular Evolutionary genetics (Columbia University Presss, New York, 1987). — 10. RAYMOND M, ROUSSET F, Evolution, 49 (1995) 1280. — 11. WEIR BS, Genetic data analysis II (Sinauer Associates Inc., Sunderland, 1996). — 12. MARJANOVIĆ D, BAKAL N, POJSKIĆ N, KAPUR L, DROBNIĆ K, PRIMORAC D, BAJROVIĆ K, HADŽISELIMOVIĆ R, J Forensic Sci, 50 (2005) 1. — 13. MARJANOVIĆ D, POJSKIĆ N, BAKAL N, DROBNIĆ K, PRIMORAC D, BAJROVIĆ K, HADŽISELIMOVIĆ R, Progress in Forensic Genetics, 11 (2006) 243. — 14. PERIČIĆ M, BARAĆ LAUC L, MARTINOVIĆ KLARIĆ I, JANIĆIJEVIĆ B, BEHLULI I, RUDAN P, Forensic Sci Int, 146 (2004) 61. — 15. STEVANOVIĆ M, DOBRIČIĆ V, KECKAREVIĆ D, PEROVIĆ A, SAVIĆ-PAVIĆEVIĆ D, KECKAREVIĆ-MARKOVIĆ M, JOVANOVIĆ A, ROMAC S, Forensic Sci Int, 171 (2007) 216. — 16. BARAĆ LAUC L, PERIČIĆ M, MARTINOVIĆ KLARIĆ I, ŠIJACKI A, POPOVIĆ D, JANIĆIJEVIĆ B, RUDAN P, Forensic Sci Int, 150 (2005) 97. — 17. BARAĆ LAUC L, PERIČIĆ M, MARTINOVIĆ KLARIĆ I, JANIĆIJEVIĆ B, PARIK J, ROOTSI S, RUDAN P, Forensic Sci Int, 138 (2003) 127. — 18. ZAHAROVA B, ANDONOVA S, GILISEN A, CASSIMAN J, DECORTE R, KREMENSKY I, Forensic Sci Int, 124 (2001) 182. — 19. REYNOLDS JB, WEIR BS, COCKERHAM CC, Genetics, 105 (1983) 767. — 20. EXCOFFIER L, LAVAL LG, SCHNEIDER S, Evol Bioinform Online, 1 (2005) 47. — 21. SAITOU N, NEI M, Mol Biol Evol, 4 (1987) 406. — 22. LIU K, MUSE S, Bioinformatics, 21 (2005) 2129. — 23. TAMURA K, DUDLEY J, NEI M, KUMAR S, Mol Biol Evol, 24 (2007) 1596.

M. Ćenanović

Institute for Genetic Engineering and Biotechnology, Sarajevo University, Gajev trg 4., 71000 Sarajevo, Bosnia and Herzegovina,
e-mail address: merisa110@bih.net.ba

DIVERZITET Y-VEZANIH KRATKIH TANDEMSKIH PONAVLJANJA (STRs) U REPREZENTATIVNOM UZORKU STANOVNIŠTVA KANTONA SARAJEVO, BOSNA I HERZEGOVINA

S A Ž E T A K

Prethodnim populaciono-genetičkim studijama analizirano je 12 STR lokusa Y hromosoma uključenih u *PowerPlex® Y System* sustav u svrhu ustanovljavanja raznolikosti kratkih tandemskih ponavljanja Y hromosoma unutar bosansko-hercegovačke humane populacije. U cilju dodatne potvrde ranije dobivenih rezultata, kao i uspostavljanja »specifičnoga uzorka« za lokalnu bh populaciju, dodatno je testirano 100 uzorke prikupljenih od zdravih muških osoba, stanovnika Kantona Sarajevo. Za izdvajanje DNA iz briseva bukalne sluznice korišten je *Qiagen DNeasy™ Tissue* kit (*Qiagen GmbH, Hilden, Germany*), te *PowerPlex® Y System* kit (Promega Corp., Madison, WI) za istovremeno umnažanje Y-STR lokusa metodom polimerazne lančane reakcije (PCR). PowerPlex® Y System obuhvata 12 STR lokusa: *DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390,DYS391, DYS392, DYS393, DYS437, DYS438 and DYS439*. Ukupni PCR-volumen iznosio je 5ìl. PCR amplifikacija izvršena je korištenjem *PE GeneAmp PCR System* instrumenta (ABI). Elektoforetska analiza umnoženih genetičkih biljega realizirana je uporabom *ABI PRISM® 310 Genetic Analyzer* instrumenta (ABI, Foster City, CA) u skladu sa preporukama proizvođača, dok je za analiza primarnih podataka korišten *ABI PRISM® Data Collection v. 3.2 i GeneMapper© v. 3.2* softver. U sklopu ove studije upoređeni su rezultati dobiveni za »sarajevsku

populaciju« sa prethodno ustanovljenim podacima za ukupnu bosanskohercegovačku populaciju, ali i dostupnim podacima za susjedne europske populacije. Rezultati ovog istraživanja će se koristiti kao smjernice u daljem istraživanju genetičke povezanosti među recentnim lokalnim populacijama u BiH, bilo izoliranim ili otvorenim, što predstavlja dugoročni projekt u našoj zemlji.