Molecular-genetics analysis of 15 STR loci in sibship testing in isolated rural Bosnian population – the use of the grey zone

Jasmin Mušanović[1], Azra Metović[1], Esad Pepić[2], Almira Lujinović[3], Adnan Fojnica[4], Faruk Husremović[5], Damir Šečić[2], Damir Marjanović[6,7], Jasmin Šutković[8]

[1]Department of Medical Biology and Human Genetics, [2]Department of Pathophysiology, [3]Department of Anatomy; School of Medicine, University of Sarajevo, Sarajevo, Bosnia and Herzegovina, [4]Institute of Biochemistry, Graz University of Technology, Graz, Austria, [5]Department of Urology, Cantonal Hospital Zenica, Zenica, Bosnia and Herzegovina, [6]International Burch University, Sarajevo, Bosnia and Herzegovina, [7]Institute for Anthropological Researches, Zagreb, Croatia, [8]Research and Development Centre, International University of Sarajevo, Bosnia and Herzegovina

ABSTRACT

Aim Examination of the effectiveness of STR loci in proving sibship of the Bosnian-Herzegovinian village of Orahovica and the formation of a “grey zone“.

Methods The probability of sibship was determined by calculating the likelihood ratio (LR) parameter for each of the 15 observed STR loci and for each of the pairs of relatives and non-relatives. Cumulative sibship index (CSI) was calculated for each of the pairs by multiplying the LR values of all 15 loci and obtained values are used as CSI limit for separating relatives from non-relatives. By creating a grey zone for local populations, an attempt was made to obtain a line of demarcation between siblings and non-siblings.

Results An analysis of the origin of the respondents' relatives was performed, up to the level of sibship in the third generation. The results of the CSI for pairs of relatives from the village of Orahovica showed that the highest CSI value, and therefore the sibship probability was recorded among relatives from the village of Orahovica (CSI=534211727.203;SP=99.999999812%). On the contrary, incredibly low CSI value was recorded among non-relatives, ranging from CSI=0.0000001 to 0.5261434 (SP=0.000009999% to 34.475357951%).

Conclusion For the threshold value CSI=1 and for CSI=3, this method determined sibship in 100% of pairs of relatives and the absence of biological sibship in 100% of pairs of non-relatives in the village of Orahovica. The STR system is proved to be a successful method in determining sibship or absence of sibship in small local populations.

Key words: family, genetic testing, molecular diagnostic techniques, paternity

Corresponding author:
Faruk Husremović
Department of Urology,
Cantonal Hospital Zenica
Crkvice 67, 72000 Zenica,
Bosnia and Herzegovina
Phone: +38732 447 000;
Fax:+387 32 226 576;
E-mail: farukhusremovic@gmail.com
Jasmin Mušanović ORCID ID:https://orcid.org/ 0000-0002-2900-9397

Original submission:
14 December 2022;
Revised submission:
03 January 2023;
Accepted:
08 February 2023
doi: 10.17392/1569-23

Med Glas (Zenica) 2023; 20(2):296-301

INTRODUCTION

Short tandem repeats (STRs) represent genetic markers with 2-6 base pairs (bp) that are repeated in tandems, directly adjacent to each other (1). The repeating units form series with a length of up to 150 bp (2). Number of these repeats vary among individuals, making them a highly discriminative tool used in population genetics, molecular epidemiology, forensic identifications and other applications (3-6).

For this research, STR markers were applied as an effective tool in studying genetic variation among individuals and sibship determination among relatives and nonrelatives (7,8). Genetic analysis of 15 STR loci (D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820, TH01, TPOX, CSF1PO, D16S539, Penta D and Penta E) was conducted in the Powerplex 16 system (Promega), having similar approach as previously reported (7,9).

In our previous studies we have examined the effectiveness of the 15 STR loci system in determination of sibship by using 15 STR loci and calculating different cut-off points of combined sibship indices (CSI) and distribution of sharing alleles (7,8). We found that it was very difficult to establish strict CSI cut-off values for determination of the doubtless sibship for the large population (7,8). Also, the effectiveness of 15 STR loci in determining sibship by using 15 STR loci was examined and different cut-off points of CSI and shared allele distributions were calculated in the past (10). We have found that it is very difficult to establish strict CSI cut-off values for determining/demarcating undoubted consanguinity for a large population (10).

Formation of the "gray zone" for the Bosnian population is not well-explored area (7,10). Additional examination of urban and rural populations in Bosnia and Herzegovina would construct the zone of inconclusive sibship, with low rate of false sibling determination, as the essential part for each forensic analysis.

The aim of this research is to examine the effectiveness of STR markers in creating a grey zone as a reliable tool in the assessment of sibship within small local population.

We intentionally focused on smaller population, trying to establish clear CSI threshold values and extrapolate their application for larger and more heterogenous populations in B&H. Also, we tried to construct "grey zone" for uncertain pairs and set the basis for understanding of more heterogenous populations in B&H.

MATERIAL AND METHODS

Population and study design

The village of Orahovica, as a populated geographical area, appears in documents for the first time in the population census from 1489 (11,13). This village is located in Bosnia and Herzegovina, Federation of Bosnia and Herzegovina, Zenica-Doboj Canton and belongs to the municipality of Zenica (11,13). Village Orahovica was built with the constant pressure of the Hungarian army, once the population from Nemila fled deeper into the hills along the valley of the Orahovica river (11,13). At that time, Nemila and Orahovica each had 11 households. Today, an asphalted road leads to the village of Orahovica, which partially enables easier communication with other towns and cities within a radius of 30 km (11,13).

With the signing of the informed consent, a swab of the buccal mucosa was taken from a total of 60 people, 46 male and 14 female in 2010, in Orahovica, and further analysed at the Institute of Genetic Engineering and Biotechnology in Sarajevo. The formation of couples for the analysis of sibship was done on the basis of genealogical connection by confirming the DNA profile of relatives and parents; it was used to prove sibship. By pairing relatives of common parents, thirty pairs of relatives from the village of Orahovica were formed, while the control group for this village was formed by random pairing of non-relatives from the same samples by taking care of elimination of possible sibship. All samples of buccal mucosa were collected and analysed 12 years ago and a detailed description was previously reported (10). For the purpose of this publication, no ethical approvals were needed, as only in silica approach was used throughout the project.

Methods

For the extraction of total genomic DNA from the buccal mucosa swab back in 2010, a modified Miller protocol was used (14-16). QUANTI-

FILER DNA Identification kit (Promega Corp., Madison, WI) is designed for simultaneous quantification of total human DNA in the analysed sample and is completely based on TaqMan technology. When setting up the reaction, an internal PCR control (IPC) of size 130 bp was added to each sample, with the analysis of which potentially present PCR inhibitors can be detected. The reaction for this research was realized on the 7300 Real-Time PCR system from Applied Biosystem, 2006. The PowerPlexTM16 Kit (Promega Corp., Madison, WI), a commercial kit that includes 15 STR loci and amelogenin was used to amplify STR loci. After the amplification, the detection of the results followed, i.e. DNA profile generation. Hi Di Formamide (Applied Biosystem, Foster City, CA), which denatures nucleic acid and maintains it in a single-stranded form, and ILS 600 (Applied Biosystem, Foster City, CA) were used in this phase of the analysis; it consists of 22 artificially synthesized DNA fragments ranging in length from 60-600 base pairs. Each fragment is labelled with carboxy-x-rhodamine (CXR) and detected by Promega red, 2009 (14-16).

When it comes to the threshold value of CSI that would be used to prove sibship between two individuals, different authors use different threshold values of CSI as the threshold value when proving sibship. Various CSI limit values are mentioned in the literature, such as, for example, CSI=0.0182, CSI=0.067, CSI=1, CSI=3, CSI=10, CSI=10.3, CSI=19.0015 (10,18,19, 20-24). However, there is no consensus for the exact threshold value of CSI when proving sibship. Based on the different threshold values of CSI used in this work, reliability indicators (sensitivity, specificity, PPV and NPV) of using CSI at different levels as a method for determining sibship between two people in small local populations of the village of Orahovica were calculated. Also, an attempt was made to preliminarily calculate the threshold values of the "grey zone" - the zone of uncertain CSI values at different levels for the observed village (Table 1) in order to compare the results obtained for the investigated local populations with the results of previous research related to Bosnia and Herzegovina. Determining this zone increases the accuracy and efficiency of the performed test, i.e. methods of determining sibship (23).

Table 1. Evaluation of sensitivity (SEN), specificity (SPE), positive assumed value (PAV), negative assumed value (NAV) and threshold values of the grey zone for determining sibship for respondents from the village of Orahovica

CSI	SEN%	SPE%	PAV%	NAV%	minGZ	maxGZ
0.0182	100	86.667	88.235	100	0	7.500187505
0.067	100	93.33	93.75	100	0	14.99250375
1	100	100	100	100	0	undef.
3	100	100	100	100	0	undef.
10	93.33	100	100	93.75	0.0667	undef.
10.3	93.33	100	100	93.75	0.0667	undef.
19.0015	86.667	100	100	88.235	0.13333	undef.
Minimum	86.667	86.667	88.235	88.235	0	undef.
Maximum	100	100	100	100	0.13333	undef.

CSI, cumulative sibship index; SP (%), sibling probability; minGZ, the minimum value of the grey zone; maxGZ, the maximum value of the grey zone for determining sibship for respondents from the village of Orahovica; undefined, undefined;

Statistical analysis

The probability of sibship was determined by calculating the likelihood ratio (LR) parameter (8,9) for each of the 15 observed STR loci in both relatives and non-relatives. Cumulative or combined sibship index (CSI) (10,18) (in the literature also known as cumulative sibship index) (17-20) was calculated for each of the pairs by multiplying the LR values of all 15 loci. For each tested related and unrelated pairs, the same, standardized tabular form was applied. The form proposed by Brenner in 2006 was used as a form for calculating LR (15). To estimate difference between the means of CSI values for all groups of siblings and non-siblings, Student t-test was used. Comparative analysis of measures of central tendency and variability of CSI and SP (%): minimum (Min), maximum value (Max), range, logarithmic mean (Ls), arithmetic mean (Xs), standard deviation (SD), variance (S2) and coefficient of variation (Kv) for relatives and non-relatives from the village of Orahovica is also calculated.

RESULTS

In this study in order to test the sibship, a method of calculating the CSI with different threshold values was used. Considering we were dealing with relatively small local population it was expected to observe a genetic drift as previously reported (22). An analysis of the origin of the respondents' relatives was also performed, up to the level of sibship in the third generation.

The results of the combined sibship index - CSI (cumulative sibship index/ combined sibship index) for pairs of relatives from the village of Orahovica showed the highest value of CSI, and therefore the SP was recorded among relatives from the village of Orahovica (CSI=534211727.203; SP=99.999999812%) (Table 2).

Table 2. Calculated sibling probability (SP) for tested relatives pairs from the village of Orahovica

Pairs	SP (%)	Pairs	SP (%)
1	99.993392797	16	99.999999812
2	87.958045831	17	99.999930494
3	97.951734675	18	99.998031738
4	99.924191029	19	99.997659656
5	99.531303533	20	89.276024407
6	99.99993213	21	95.857240507
7	99.99961181	22	99.993526087
8	99.795449278	23	94.811582622
9	99.99963382	24	99.95787642
10	99.184996104	25	99.999295038
11	98.972718484	26	99.980332893
12	99.99979107	27	99.999883509
13	93.327728625	28	99.996537273
14	98.933277579	29	99.992577721
15	99.641633705	30	98.937143743

The results of the combined sibship index - CSI for unrelated pairs from the village of Orahovica showed the highest CSI value recorded among non-relatives from the village of Orahovica was CSI=0.5261434 and SP=34.475357951%, while the lowest value was CSI=0.0000001 and SP=0.000009999% (Table 3).

Table 3. Calculated sibling probability (SP) for tested pairs of non-relatives (SP) from the village of Orah ovica

Pairs	SP (%)	Pairs	SP (%)
1	1.225368446	16	0.138577695
2	0.1331026	17	2.846575895
3	0.254042977	18	34.475357951
4	0.009999	19	0.261444675
5	9.659936512	20	0.086275501
6	1.085644401	21	1.120042937
7	0.014397926	22	0.000469997
8	0.019306271	23	0.047547381
9	0.007189483	24	0.567777874
10	1.403413946	25	0.089120504
11	0.000729994	26	2.409604456
12	0.000009999	27	0.025813334
13	0.015637554	28	0.081303843
14	0.000349998	29	0.024164159
15	0.212935616	30	0.002389942

Based on the value of the t-test for the compared arithmetic means in the groups of relatives and non-relatives from the village of Orahovica higher CSI in relatives compared to non-relatives was found with a statistically significant difference (p=4.19097E-59; p<0.0001) (Table 4). This indicates that these two groups (relatives and non-relatives) can be statistically clearly distinguished based on the obtained results of sibship testing using DNA analysis.

Table 4. Comparative analysis of measures of central tendency and variability of cumulative sibship index (CSI) and sibling probability (SP) for the analysed groups

CSI	Relatives Orahovica	NoN-relatives Orahovica
Min. (Min.%)	7.3043 (87.95804583)	0.0000001 (0.000009999)
Max. (Max.%)	534211727.2030 (99.999999812)	0.5261434 (34.47535795)
Range (range %)	534211719.8987 (12.041953982)	0.5261433 (34.475347951)
Ls	17977007.4674	0.025200133
Xs (%)	98.46703608	1.873951029
SD (SD %)	97501964.52 (3.14420541)	0.096729177 (6.424337686)
S2 (S2%)	9.50663E+15 (9.886027661)	0.009356534 (41.27211471)
Kv (Kv %)	542.37038448592 (3.19315532910473)	383.8439146 (342.8231361)

Min, minimum; Max, maximum values; Ls, logarithmic mean, Xs, arithmetic mean; SD, standard deviation; S2, variance; Kv, coefficient of variation

For values of CSI>3 for the same population (CSI=10, CSI=10.3, CSI=19.0015), the percentage of proven sibship among relatives decreases, while the percentage of proven absence of sibship among non-relatives remains the same as for CSI=1 and CSI=3. For these CSI values, only the lower limit of the “grey zone” that reduces the percentage of proven absence of sibship in non-relatives could be determined, while the upper limit could not be statistically defined. The results obtained in this way suggest that the creation of a “grey zone” does not make sense in this case either, and that these values are less reliable in proving sibship than CSI=1 and CSI=3. Consequently, it can be noted that the CSI is the threshold value that can be taken as clear and reliable in proving sibship for the examined population in the village of Orahovica, with the threshold value of CSI=1 (Figure 1).

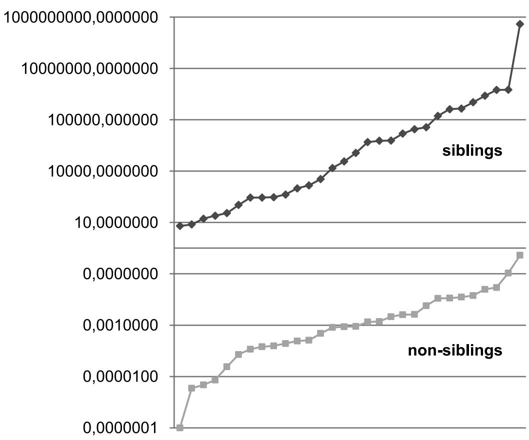

Figure 1. Representation of cumulative sibship index (CSI) values for relatives and non-relatives from the village of Orahovica

DISCUSSION

In previous publications, the importance of "grey zone" formation has been stressed and significance in the analysis of uncertain results (e.g. results that could be falsely characterized as relatives or unrelated by their value) was outlined (10,14). The importance of determining the sensitivity, specificity, positive assumed value (PAV) and negative assumed value (NAV) of the selected methods of determining sibship, lies in reducing the number of false positive and false negative findings. In the forensic science, characterizing someone as a false relative or false unrelated has a direct effect on the lives of people whose sibship is being tested (10,12).

The most commonly used limit value in the literature was CSI=1.We applied the very same method also in this case of Orahovica and determined sibship (CSI ≥1) in 100% of pairs of relatives and the absence of biological sibship (CSI<1) in 100% of pairs of non-relatives.

When determining the threshold values of the "grey zone" for CSI=1, it turned out that they cannot be defined statistically. Based on the previous results (10) of the detected presence of sibship in relatives or the absence of sibship in non-relatives, and the impossibility of creating a "grey zone", it can be concluded that CSI=1 is a clear, reliable and sufficient threshold value for proving sibship in this population. The stated results and conclusions for CSI=1, were also obtained for CSI=3 for the same population. For the examined values of CSI<1 (CSI=0.0182, CSI=0.067), the percentage of proven sibship remained the same as for CSI=1 and CSI=3, while the percentage of proven absence of sibship among non-relatives decreased (86.667% and 93.33%).

By trying to determine the limits of the "grey zone" for values obtained for all pairs of non-relatives in our population enter uncertain, suspicious values, and the percentage of really positive results also decreases, i.e. percentage of proven sibship, so the "grey zone" created in this way does not improve the existing result, and it can be concluded that its formulation in this case does not make sense and that these values are less reliable than the CSI=1 and CSI=3 values.

In the literature, there are many reported disagreements (10,18,19,25) regarding the evaluation of the optimal and reliable cut-off value for any of sibship determination methods (10,20,23,25). Since applications of these are methods are very delicate and sensitive, correct determination could seriously impact lives of the examiners, especially in the forensic practice. One of the studies (22) determined a CSI cut-off value of 1 as a clear cut off between relatives and non-relatives using 15 STR loci, but with known profiles of the subjects' parents. Another study (26) states the level of difficulty in the sibship determination between two individuals, once maximum of 9 STR loci are used, without knowing DNA profiles of the parents; they suggest the analysis of a larger number of STR loci in proving sibship in such individuals (26). The results obtained in this research indicate the possibility of determining sibship in the examined small local populations, using 15 STR loci and CSI threshold values of 1 and 3.

By creating a "grey zone", the most reliable limit for determining relatives and non-relatives turned out to be 0.067, but with the existence of still uncertain pairs that require additional testing (10,24,25). However, it seems that the formulation of the "grey zone" in the case of small local populations is not needed, as there are clear cut-off values separating relatives from non-relatives. Also, there are no uncertain pairs for CSI values of 1 and 3. The reason for this appearance of the possibility for CSI boundary formulation in the sample from a small local population, may arise from lower genetic variability, demonstrated earlier in these samples once compared to the samples of a mixed population (10).

In conclusion, CSI-1 and CSI-3 can be used as a good tool in order to determine sibship in 100% of pairs of relatives and the absence of biological sibship, also in 100% of pairs of non-relatives. However, limit values of the "grey zone" for CSI=1 and CSI=3 could not be defined statistically. For the village Orahovica, CSI=1 and CSI=3 values are clear, reliable and sufficient threshold values for proving sibship in this population but not for the formation of the grey zone.

FUNDING

No specific funding was received for this study.

TRANSPARENCY DECLARATION

Competing interests: None to declare.

REFERENCES

1. Hannan AJ. Tandem repeats mediating genetic plasticity in health and disease. Nat Rev Genet2018; 19:286-98.
2. Fang L, Liu Q, Monteys AM. Deep Repeat: direct quantification of short tandem repeats on signal data from nanopore sequencing. Genome Biol 2022; 23:108.
3. Wyner N, Barash M, McNevin D. Forensic autosomal short tandem repeats and their potential association with phenotype. Front Genet2020; 11:884.
4. Roewer L. Y-chromosome short tandem repeats in forensics—sexing, profiling, and matching male DNA. WIREs Forensic Sci2019; 1:1336.
5. Fan H, Chu JY. A brief review of short tandem repeat mutation. Genomics Proteomics Bioinformatics 2007; 5:7-14.
6. Calafell F. Short tandem repeat polymorphism evolution in humans. Eur J Hum Genet 1998; 6:38–40.
7. Mušanović J, Metović A, Šečić D, Pepić E, Loga-Zec S, Ramić N, Šljuka S, Ećo I. Genetic Polymorphism of 15 STR loci in the population of two Bosnian and Herzegovinian villages. EJBPS 2022; 9:37-41.
8. Kakkar S, Sarma P, Panigrahi I, Mandal SP, Shrivastava P, Kumawat RK. Development of a new screening method for faster kinship analyses in mass disasters: a proof of a concept study. Sci Rep 2022; 27:20372.
9. Inoue H, Manabe S, Fujii K, Iwashima Y, Miyama S, Tanaka A, Saitoh H, Iwase H, Tamaki K, Sekiguchi K. Sibling assessment based on likelihood ratio and total number of shared alleles using 21 short tandem repeat loci included in the GlobalFiler™ kit. Leg Med 2016; 19:122-6.
10. Musanovic J, Filipovska-Musanovic M, Kovacevic L. et al. Determination of combined sibship indices "gray zone" using 15 STR loci for central Bosnian human population. Mol Biol Rep 2012; 39:5195–200.
11. Wikipedia. Orahovica, Zenica. https://en.wikipedia.org/wiki/Orahovica (22 November 2022).
12. Mušanović J, Metović A, Pepić E, Imamović B, Džubur- Alić A, Filipovska- Mušanović M, Marjanović D. Division of alleles in the determination of kinship in small localpopulations. EJBPS 2017; 4:128-32.
13. Kozlić S, Valentić OA. Orahovica u vremenu i prostoru (Orahovica in time and space) [in Bosnian]. In: Monografija sela od 1489.do danas. Zenica: UG Gajret,2007.
14. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988; 16:1215.
15. Ensenberger, MG, Thompson J, Hill B, Homick K, Kearney V, Mayntz-Press KA, Krenke BE. Developmental validation of the PowerPlex® 16 HS System: an improved 16-locus fluorescent STR multiplex. Forensic Sci Int Genet 2010; 4:257-64.
16. Butler JM. Forensic DNA Typing, Biology, Technology and Genetics of STR markers. 2nd ed. USA: Elsevier, 2005.
17. Marjanović D, Primorac D. Molekularna forenzična genetika (Molecular forensic genetics) [in Bosnian]. Sarajevo: Institut za genetičko inženjerstvo i biotehnologiju, 2009.
18. En Pu Ch, Linacre A, Systematic evaluation of sensitivity and specificity of sibship determination by using 15 STR loci. Forensic Leg Med 2008; 15: 329–34.
19. Tzeng CH, Lyou JY, Chen YR, Hu HY, Lin JS, Wang SY, Lee JC. Determination of sibship by PCR-amplified short tandem repeat analysis in Taiwan. Transfusion 2000; 40:840-5.
20. Giroti IR, Verma S, Singh K, Rohit M, Talwar I. A gray zone approach for evaluation of 15 short tandem repeat loci in sibship analysis: A pilot study in Indian subjects. J Forensic Leg Med 2007; 14:261–5.
21. Marjanović D, Durmić-Pašić A, Bakal N, Haverić S, Kalamujić B, Kovačević L, Primorac D. DNA identification of skeletal remains from World War II mass graves uncovered in Slovenia. Croat Med J 2007; 48:513.
22. Marjanovic D, Kapur L, Drobnic K, Budowle B, Pojskic N, Hadziselimovic R. Comparative study of genetic variation at 15 STR loci in three isolated populations of the Bosnian mountain area. Hum Biol 2004; 76:15-31.
23. Reid TM, Wolf CA, Kraemer CM, Lee SC, Baird ML, Lee RF. Specificity of sibship determination using the ABI identifiler multiplex system. J Forensic Sci 2004; 49:1–3.
24. Mušanović J. Analiza učinkovitosti primjene kratkih repetitivnih sekvenci u utvrđivanju srodstva među potomcima jednog roditeljskog para (Efficiency analysis of short repetitive sequence usage in determining sibship among siblings) [in Bosnian]. Sarajevo: Prirodno-matematicki univerzitet, 2007**.**
25. Coste J, Pouchot J. A grey zone for quantitative diagnostic and screening tests. Int J Epidemiol 2003; 32:304–13.
26. Gaytmenn R, Hildebrand DP, Sweet D, Pretty IA. Determination of the sensitivity and specificity of sibship calculation using AmpFlSTR Profiler Plus. Int J Legal Med 2002; 116:161–4.