ISSN 2303-4521
Periodicals of Engineering and Natural Sciences *Original Research*
Vol. 11, No. 2, March 2023, pp.78-91

Determination of combined sibship indices using 15 STR loci and a grey zone in a small local population in central Bosnia and Herzegovina

Jasmin Mušanović[1], Azra Metović[1], Adnan Fojnica[2], Jasmin Šutković[3], Tarik Ikanović[1], Perica Vasiljević[4], Damir Marjanović[5]

[1] Department of Medical Biology and Human Genetics, University of Sarajevo, Bosnia
[2] Institute of Biochemistry, Graz University of Technology, Austria
[3] Department of Genetics and Bioengineering, International University of Sarajevo, Bosnia
4 Department of Biology and Ecology, University of Niš, Serbia
[5] Institute for Anthropological Research, Croatia

ABSTRACT

This study investigates the genetic diversity and relatedness among a small local population in Bosnia and Herzegovina. In a genetic research study, a sample of 38 individuals was collected from the village of Vukotići, consisting of 21 male and 17 female subjects. The total genomic DNA was extracted using a modified Miller protocol. The QUANTIFILER DNA identification kit was used to quantify the total human DNA in the sample. The sibship relationship was assessed by computing the likelihood ratio for each of the 15 STR loci in both relatives and non-relatives. Results showed a higher homogeneity of the small local population compared to the mixed population within the larger population. Variability in peak height observed in the genetic analysis was attributed to differences in DNA concentration in the extracted samples. Probability of relatedness among participants in the Vukotići village was found to be low. Central tendency and variability measures revealed valuable insights into sample distribution and variation. The study concludes that CSI=1 and CSI=3 can be used as reliable tools to determine sibship in small local populations without a "gray zone".

Keywords: Genetic Diversity, DNA identification, 15 STR, Gray zone, Likelihood ratio

Corresponding Author:

Jasmin Mušanović
Department of Biology and Human Genetics
University of Sarajevo
Čekaluša 90, Sarajevo 71000, Bosnia
E-mail: jasmin.musanovic@mf.unsa.ba

1. Introduction

The field of population genetics aims to understand the distribution of genetic variation within and between populations [1]. This knowledge is crucial for understanding the evolution of species, the genetic basis of disease, and the development of effective strategies for genetic conservation and improvement [2]. In addition, population genetics provides the theoretical basis for the use of molecular markers such as short tandem repeats (STRs) and their application in human identification, paternity testing and sibship identification [3].

STRs are genetic markers that consist of repeating units of 2-6 base pairs that are directly adjacent to each other and can form series with a length of up to 150 base pairs [4]. These repeats vary among individuals, making them a highly discriminative tool in population genetics [4,5]. The analysis of 15 STR loci in local populations has become a standard method for this purpose [8]. In this study, fifteen autosomal STR loci (D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820, TH01, TPOX, CSF1PO, and D16S539, Penta D and Penta E) commonly used in parentage testing and sibship determination were studied [8].

However, the accuracy of this analysis depends on the frequency of the alleles in the population [8]. In populations with a high degree of genetic variation, it is possible to obtain false positive results, indicating a relationship where none exists [9]. This is particularly true for populations with a mixture of different ethnic groups, where the frequency of alleles may be different from one group to another [8, 9]. To overcome this challenge, it is possible to focus on a small, isolated population to establish clear threshold values for sibship determination, which can then be extrapolated to larger and more heterogeneous populations. In addition, this allows the identification of population-specific genetic markers that are used for a variety of purposes in forensic analysis [10-12].

In this study, the goal is to investigate the effectiveness of 15 STR loci in determining sibship relationships in the remote and isolated population of Vukotici village in Bosnia and Herzegovina (B&H). The aim is to develop standardized procedures for human identification and paternity testing in B&H by analysing the allele frequencies of the 15 STR loci. By focusing on a small, isolated population, the study aims to establish clear threshold values for sibship determination, which can then be extrapolated to larger and more heterogeneous populations in B&H. This research was undertaken as part of continuing holistic genetics investigations of populations of the rural communities in Bosnia and Herzegovina that will help to understand genetic makeup of the population as well as the migration of the local communities over course of history.
In addition, the study aims to create a "gray zone" for uncertain sibship relationships, which has not been well explored in Bosnian populations. By examining rural populations in B&H, the study hopes to construct a zone of inconclusive sibship relationships with a low rate of false sibling determination. This "gray zone" is an essential part of each forensic analysis and will contribute to a better understanding of larger and more heterogeneous populations in B&H and the region.

2. Research method

2.1. Description of the locality – Vukotici village

The village of Vukotići (Figures 1, 2, and 3) is located in Bosnia and Herzegovina, Federation of Bosnia and Herzegovina, Zenica-Doboj Canton. The village is found at 30 km from the city of Zenica and located between villages of Šerići and Jastrepac on a hill facing southeast [13, 14]. This place with 115 residential buildings has 958 inhabitants and the locals are mainly engaged in agriculture and livestock farming. The wealthier population is emigrating to urban and suburban settlements. There are no visible remains of any previous civilization [13, 14]. Village's area covers a larger area of land and is well recognized by its settlements: Kozle, Lipa, Bašča, and Donje Selo [13, 14].

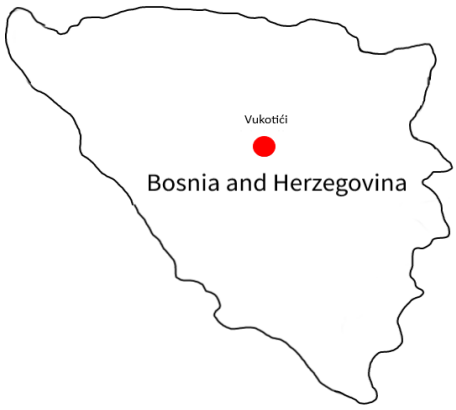

Figure 1. Schematic representation of the village of Vukotići

Figure 2. Satellite view of the village of Vukotići (taken from: Google Earth)

As for the infrastructure, due to its location and steepness (Figure 2), Vukotići is very limited in communication with the surrounding areas.

2.2. Participants and sample collection

In a genetic research study, a sample of 38 individuals was collected from the village of Vukotići, consisting of 21 male and 17 female subjects. The selection of participants was based on their voluntary participation and the existence of a sibling who was also willing to provide their genetic material. The collection of samples was carried out ethically, with informed consent obtained from each subject. The formation of related individuals for sibship analysis was based on genealogical connections, resulting in 30 related pairs from the village of Vukotići. To serve as a control group, a random pairing of non-relatives from the same cohort was performed, resulting in 30 unrelated pairs for comparison of statistical parameters.

The collection of buccal mucosal cells was performed using a cotton swab, which was rubbed gently against the inner surface of the cheek ten times on each side. The person collecting the samples took precautions to avoid contamination of the samples, including wearing gloves, a face mask, and a cap. After collection, the cotton tip of the swab was removed and placed in a sterile 1.5 ml plastic tube, which was stored in a refrigerator at -20°C until the extraction process. The results of this study were compared with previous research to obtain a purified population sample [12, 13].

2.3. Methods for DNA Extraction and Analysis

The total genomic DNA was extracted from the buccal mucosa swab in 2010 using a modified Miller protocol [19]. The QUANTIFILER DNA Identification kit (Promega Corp.) was used to quantify the total human DNA in the sample [10]. The kit is based on TaqMan technology and has an internal PCR control (IPC) of 130 bp that can detect potentially present PCR inhibitors. The reaction was performed on the 7300 Real-Time PCR

system from Applied Biosystem. The PowerPlexTM16 Kit (Promega Corp.), which includes 15 STR loci and amelogenin, was used to amplify the STR loci [10]. The results were then detected and the DNA profile was generated. Hi Di Formamide and ILS 600 from Applied Biosystem were used in this phase of the analysis. The ILS 600 consists of 22 artificially synthesized DNA fragments ranging from 60-600 base pairs, each labelled with carboxy-x-rhodamine and detected by Promega red.

2.4. Threshold value for CSI in sibship determination

Different authors use different threshold values of CSI to prove sibship between two individuals. The values mentioned in the literature range from CSI=0.0182 to CSI=19.0015 [16-20]. There is no consensus on the exact threshold value of CSI for sibship determination. In this work, reliability indicators (sensitivity, specificity, PPV and NPV) of using different levels of CSI as a method for determining sibship in small local populations of the village of Vukotici were calculated. An attempt was made to calculate the threshold values of the "gray zone" - the zone of uncertain CSI values - in order to compare the results with previous research related to Bosnia and Herzegovina. This increases the accuracy and efficiency of the sibship determination method [20].

2.5. Statistics

The sibship relationship was assessed by computing the likelihood ratio (LR) for each of the 15 STR loci in both relatives and non-relatives. The cumulative sibship index (CSI), also referred to as the combined sibship index in literature, was calculated by multiplying the LR values of all 15 loci for each pair [20]. A standardized tabular form was used for calculating LR for each pair of related and unrelated individuals, using the form proposed by Brenner in 2006 [21].

3. Results

The genetic analysis of buccal samples collected from the village of Vukotici was found to be successful, as all 38 samples yielded usable DNA profiles for 16 loci (15 STR loci and amelogenin). The range of peak height variation observed in these profiles ranged from 1200 to 9000 RFU, as depicted in Figure 4 and 5.

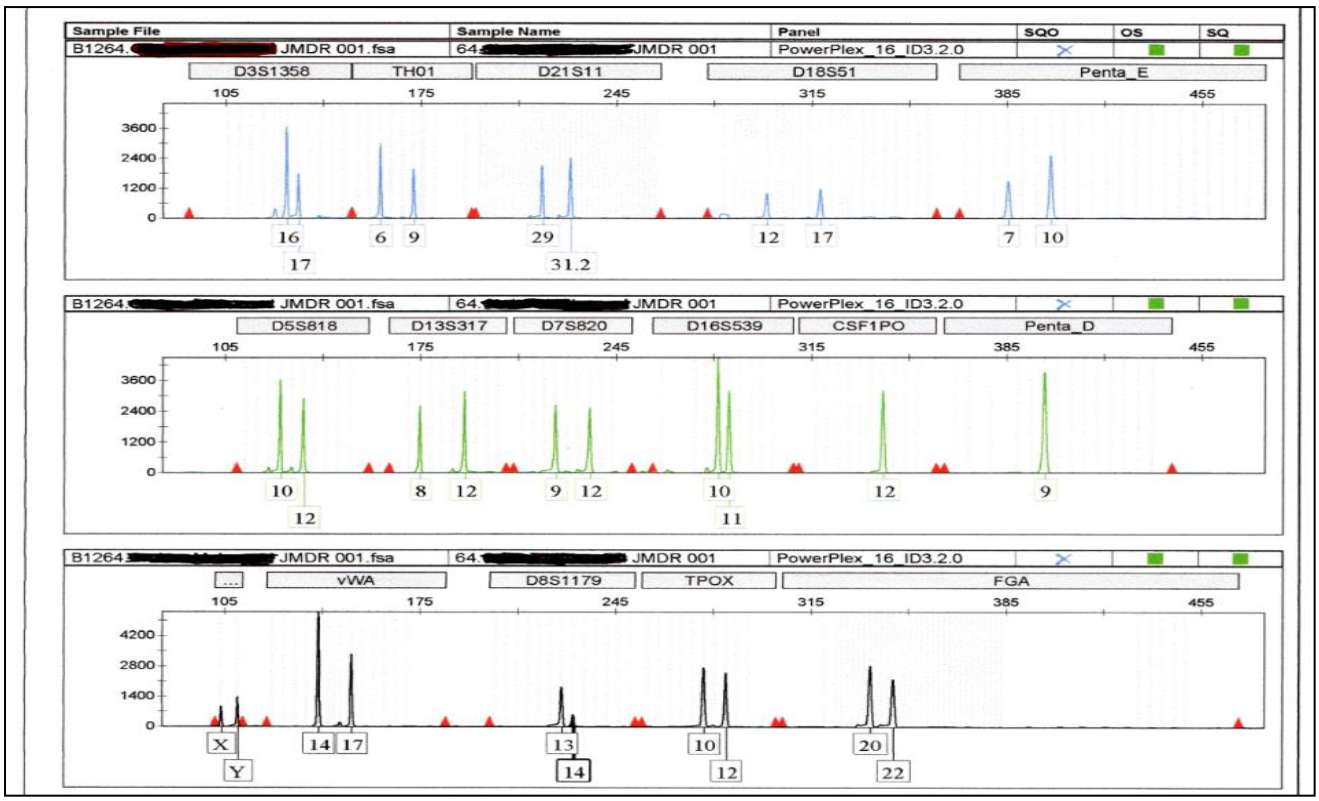

Figure 4. DNA profile with lowest detected peaks

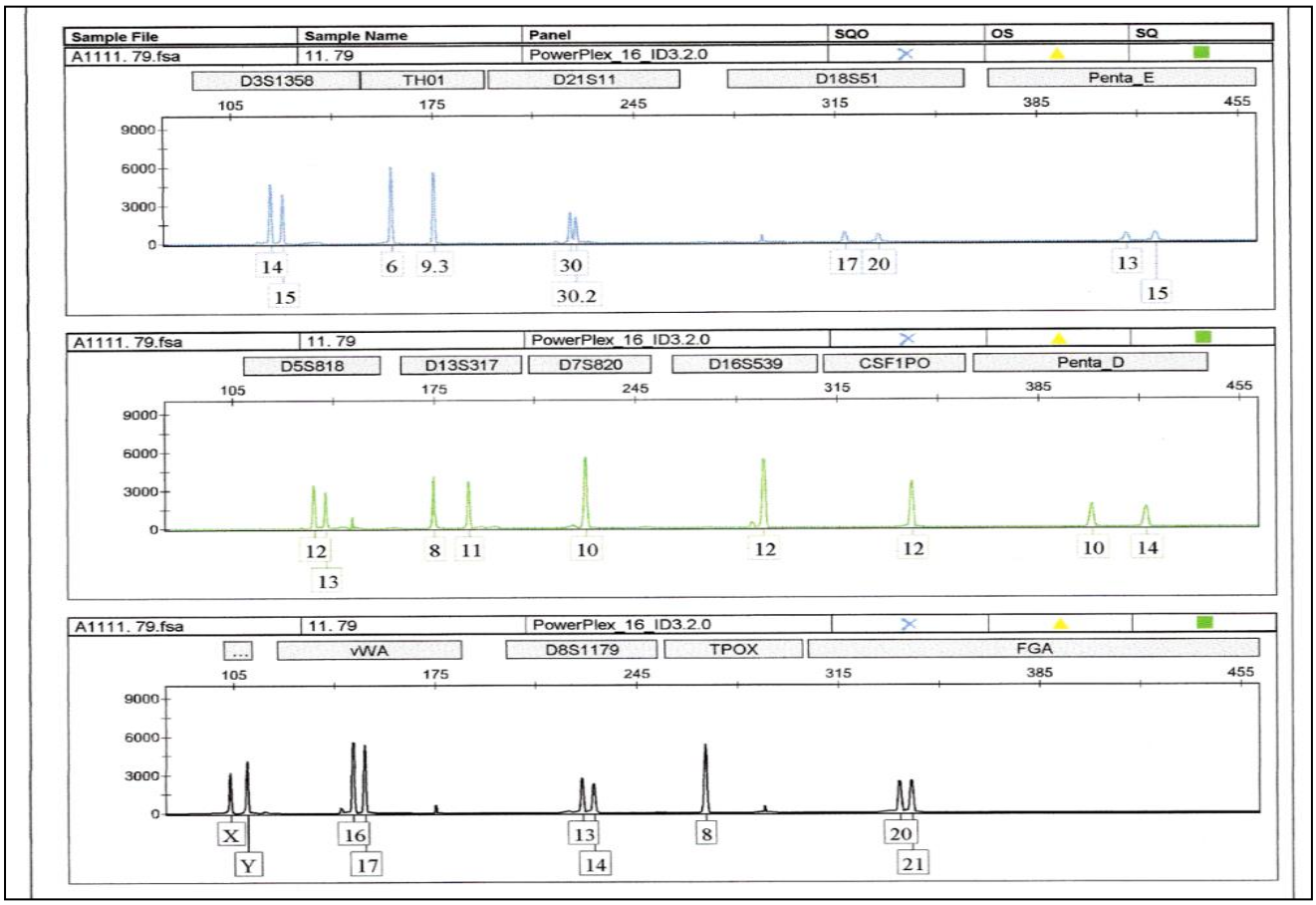

Figure 5. DNA profile with highest detected peaks

The detected range of peak height variation in the profiles from both villages exceeded the suggested ideal range of 400 to 2000 RFU for buccal mucosa samples, as well as the range obtained in a previous study using the AB377 DNA sequencer [12, 22]. However, it should be noted that these values were obtained using the AB310 genetic analyser, which has a different performance profile. Sometimes even within the same model of genetic analyser, results vary [10, 12]. Our laboratory's experience suggests that peak heights ranging from 1000 to 3000 RFU can be obtained using this particular instrument [12, 22].

Calculated values of the Combined Sibship Index (CSI) and Sibship Probability (SP) for the observed samples from the village of Vukotići.

We employed two methods to examine the relatedness of the subjects. Firstly, we calculated the CSI with varying threshold values, and secondly, we examined the distribution of allele sharing across loci. Notably, we conducted an analysis of the origin of the participants' relatives up to the level of third-generation relatives, recognizing that the relatively small local populations in this area may be subject to genetic drift, as previously observed [23].

Table 1 presents the results of the combined sibling index - CSI (Cumulative Sibship Index/Combined Sibship Index) for pairs of relatives from the village of Vukotici, while Table 2 displays the results expressed as Sibship probability - SP. The lowest CSI value and correspondingly the low probability of relatedness was obtained for this village, having values of CSI = 5.4853 and SP = 84.58027509%. In our previous study, we saw relatively high probability of relatedness with the values of CSI = 534211727.203 and SP = 99.999999812% [23].

Table 1. Calculated likelihood ratio/cumulative sibship index (LR/CSI) for tested relative pairs in a village of Vukotići

PAIRS	LR (CSI)	PAIRS	LR (CSI)
1	406,5914	16	58,6403
2	692,7274	17	1044,8422
3	12,5673	18	1708,7137
4	809,9493	19	346206,7605
5	117644,1385	20	992,2008
6	4191864,2226	21	2388,7079
7	388,6584	22	366913,9270
8	695370,4049	23	59003,8814
9	357755,3632	24	165734,1246
10	203,2614	25	5,4852
11	27284111,1987	26	1403689,5656
12	91040709,7791	27	70622,6978
13	216797,3789	28	5232550,9294
14	376,3690	29	708,4484
15	3224,7574	30	1910534,0937

Table 2. Calculated sibiling probability (SP) for tested relative pairs in a village of Vukotići

PAIRS	SP sibiling probability (%)	PAIRS	SP sibiling probability (%)
1	99,754656256	16	98,323281405
2	99,855851159	17	99,908204438
3	92,629336714	18	99,941510675
4	99,876687728	19	99,999711156
5	99,999149986	20	99,899315425
6	99,999976063	21	99,958153881
7	99,743364957	22	99,999727457
8	99,999856191	23	99,998305224
9	99,99972048	24	99,999396627
10	99,510437712	25	84,580275087
11	99,999996334	26	99,999928759
12	99,999998910	27	99,998584044
13	99,999538741	28	99,999980888
14	99,73500738	29	99,859045421
15	99,968999528	30	99,999947658

Table 3 shows the results of the combined sibling index (CSI) or the cumulative sibship index for pairs of unrelated individuals from the village of Vukotici. The values are expressed as Sibship probability (SP), which is presented in Table 2. It is noteworthy that the highest value of CSI, indicating the highest probability of relatedness, was observed in pairs of unrelated individuals from the village of Orahovica [23]. The recorded CSI value was 0.5261434, with a corresponding SP value of 34.475357951%. Conversely, the lowest CSI value was observed in pairs of unrelated individuals, also from the village of Orahovica. The recorded CSI value was 0.0000001, and the corresponding SP value was 0.000009999%.

Table 3. Calculated likelihood ratio/cumulative sibship index (LR/CSI) for tested unrelated pairs in a village of Vukotići

PAIRS	LR (CSI)	PAIRS	LR (CSI)
1	406,5914	16	58,6403
2	692,7274	17	1044,8422
3	12,5673	18	1708,7137
4	809,9493	19	346206,7605
5	117644,1385	20	992,2008
6	4191864,2226	21	2388,7079
7	388,6584	22	366913,9270
8	695370,4049	23	59003,8814
9	357755,3632	24	165734,1246
10	203,2614	25	5,4852
11	27284111,1987	26	1403689,5656
12	91040709,7791	27	70622,6978
13	216797,3789	28	5232550,9294
14	376,3690	29	708,4484
15	3224,7574	30	1910534,0937

Table 4. Calculated sibiling probability (SP) for tested unrelated pairs in a village of Vukotići

PAIRS	*LR (CSI)*	*PAIRS*	*LR (CSI)*
1	0,015277665	16	0,161259533
2	0,252013288	17	0,009079175
3	0,0013808905	18	0,551204891
4	0,296885962	19	0,084997692
5	0,000519997	20	0,001289983
6	0,041422834	21	0,000099999
7	0,093402677	22	0,564969953
8	0,056977517	23	0,012198511
9	1,62236525	24	0,366283432
10	0,048206749	25	0,023394525
11	0,4548021	26	1,049077678
12	0,001529976	27	0,005629683
13	0,009329129	28	0,78198801
14	0,048086865	29	0,036436718
15	0,000609996	30	1,393915334

Display, comparison, and discussion of measures of central tendency and variability for the observed samples

We present here a thorough examination and discussion of the measures of central tendency and variability observed in the samples. A comparative analysis is conducted on the measures of CSI and SP (%), which includes the assessment of minimum (Min) and maximum (Max) ranges, logarithmic mean (LM), arithmetic mean (AM), standard deviation (SD), variance (S2), and coefficient of variation (CV) for both related and unrelated individuals residing in the village of Vukotici. The findings of this study are displayed in Table 5.

Table 5. Comparative analysis of measures of central tendency and variability CSI and SP (%), including minimum (Min), maximum (Max), range (R), logarithmic mean (LM), arithmetic mean (AM), standard deviation (SD), variance (S2), and coefficient of variation (CV) for the village of Vukotici.

	Related individuals in a village of Vukotići	**Unrelated individuals in a village Vukotići**
Min **CSI %**	5,4852 84,58027509	0,000001 0,000099999
Max **CSI %**	91040709,7791 99,99999891	0,0164912 1,62236525
LM **CSI %**	91040704,2938 15,41972382	0,0164902 1,622265251
AM **CSI %**	4449084,3462 99,11793154	0,002904143 0,266154534
SD **CSI %**	17110732,29 3,063022734	0,004569885 0,428778786
S^2 **CSI %**	2,92777E+14 9,38210827	2,08838E-05 0,183851247
CV **CSI %**	384,58997309445 3,0902811291657	157,3574373 161,1014397

The t-test results for the comparison of arithmetic means in the related and unrelated groups from Vukotići village ($p = 1.07117E-80$; $p < 0.0001$) indicate a high level of statistical significance and suggest a significantly higher CSI in the related individuals compared to the unrelated ones. This finding suggests that the related and unrelated groups can be distinguished statistically based on the results of DNA analysis testing for relatedness in both villages.

Testing the threshold values of CSI to prove relatedness in samples from small local populations

In the realm of determining sibship through CSI analysis, there is no universal consensus on the appropriate threshold value to be used for this purpose, as various authors have reported divergent threshold values in the literature. These threshold values include CSI values of 0.0182, 0.067, 1, 3, 10, 10.3, and 19.0015, as reported by different studies [10, 12, 14]. Consequently, there remains an ongoing debate and uncertainty regarding the precise CSI threshold value that should be employed to establish sibship. Figure 6 shows a visual representation of the CSI values for both related and unrelated individuals from the village of Vukotići, shedding light on the empirical distribution of CSI values in this particular context.

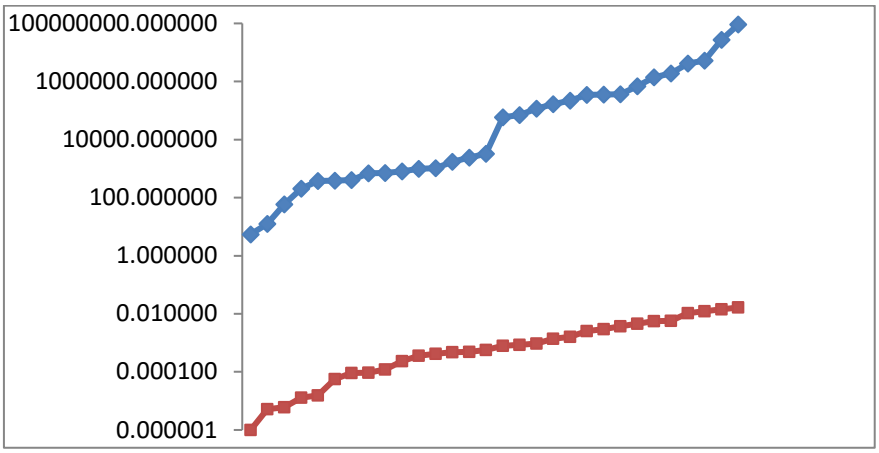

Figure 6. CSI values for the related and unrelated participants from the village of Vukotici (Blue – siblings, Red – nonsiblings)

Based on the different CSI threshold values used in this study, reliability indicators (sensitivity, specificity, PPV, and NPV) of using CSI at different levels as a method for determining sibship between two individuals in small local populations in the village of Vukotići were calculated. Additionally, attempts were made to preliminarily calculate "gray zone" threshold values - the zone of uncertain CSI values at different levels for the observed village (Table 6) in order to compare the obtained results for the examined 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 or sibship determination method [25]. In previous research in this field [10, 12] as well as in similar studies (26), the effect of creating a "gray zone" was highlighted in the analysis of uncertain results, i.e., results that could be falsely characterized as relatives or non-relatives based on their values. The importance of determining sensitivity, specificity, PPV, and NPV in order to determine the accuracy of the chosen sibship determination methods lies in reducing the number of false positive and false negative findings, because in a forensic sense, characterizing someone as a false relative or false non-relative has a direct effect on the lives of people whose sibship is being examined.

Table 6. Evaluation of sensitivity, specificity, PPV, NPV, and "gray zone" threshold values for sibship determination for participants from the village of Vukotići. SEN% - sensitivity; SPE% - specificity; PPV% - positive predictive value; NPV% - negative predictive value minSZ and maxSZ - minimum and maximum values of the "gray zone" threshold; ND – not determined

CSI	SEN%	SPE%	PPV%	NPV%	minSZ	maxSZ
0,0182	100	100	100	100	0	ND
00,067	100	100	100	100	0	ND
1	100	100	100	100	0	ND
3	100	100	100	100	0	ND
10	96,667	100	100	96,774	0,03333	ND
10,3	96,667	100	100	96,774	0,03333	ND
19,0015	93,333	100	100	93,75	0,06667	ND
Minimum	93,333	100	100	93,75	0	ND
Maximum	100	100	100	100	0,06667	ND

When it comes to the village of Vukotići, CSI value of 1 was used to establish sibship. Values of CSI≥1 were used 100% successfully for determination of related pairs and the absence of biological sibship (CSI<1) in 100% of unrelated pairs. When determining the "gray zone" boundary values for CSI=1, it was shown that they cannot be statistically defined even for this population. Based on the previously mentioned results of detected presence of sibship among relatives and absence of sibship among non-relatives, and the impossibility of creating a "gray zone", the same conclusion can be drawn as for the village of Orahovica [10] namely that CSI=1 is a clear, reliable, and sufficient boundary value for proving sibship in this population. The same results as for CSI=1 for this population were obtained for CSI=0.0182, CSI=0.067, and CSI=3. For CSI values greater than 3 for the same population (CSI=10, CSI=10.3, CSI=19.0015), the percentage of proven sibship and the percentage of proven absence of sibship among non-relatives behave as in the previously described sample from the village of Orahovica, as well as the creation of a "gray zone" and its meaning, so it can also be considered that these values are less reliable in proving sibship compared to CSI=0.0182, CSI=0.067, CSI=1, and CSI=3 for this population. From all the results for both villages, it can be observed that the CSI boundary value that can be taken as clear and reliable for proving sibship in the examined population from the village of Vukotići is much lower, and it amounts to CSI=0.0182, while for the population from the village of Orahovica, the lowest clear and reliable boundary value is CSI=1.

Comparing the results of a pooled sample of small local populations and a sample of a larger mixed population

Since one of the goals of this research was to compare the results obtained from the analysis of the effectiveness of STR loci between small local populations with the results obtained from the previously analysed sample of a larger, mixed (heterogeneous) population [10, 12], the participants from both villages were combined into one sample (Figure 7). The combination of samples from the two previously described villages into one larger sample could be done based on the proven absence of a significant statistical difference in the average values of the examined parameters of relatedness, which speaks to the similarity of these two villages. The newly obtained sample in terms of the number of participants is approximately equivalent to the size of the sample of the mixed (heterogeneous) population.

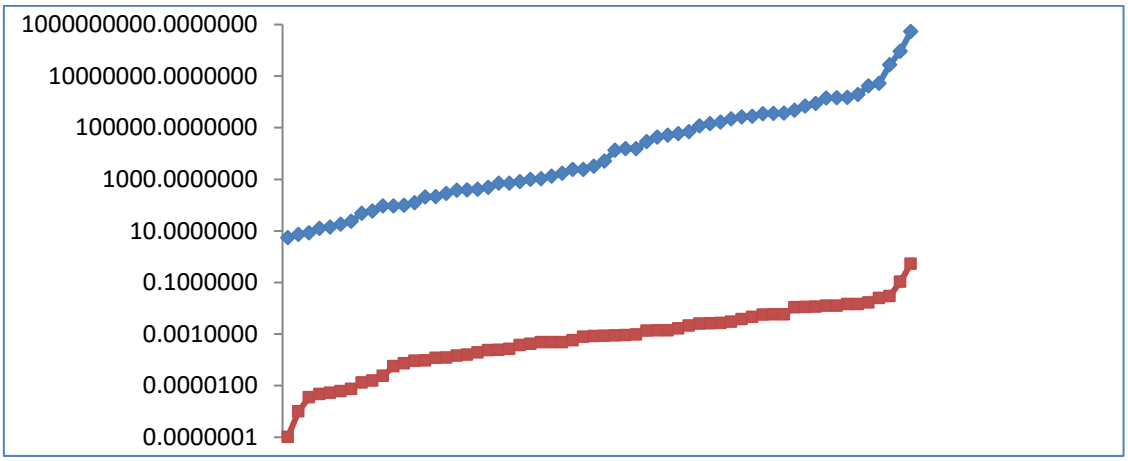

Figure 7. Representation of CSI values for relatives and non-relatives from both villages (Blue – siblings, Red – nonsiblings)

The t-test analysis revealed that there was no significant difference in the SP (%) values between related groups from observed populations (p=0.338949133) and between unrelated groups of both populations (p=0.30420467) However, when comparing the measures of central tendency and variability between the sample of small local populations and the mixed population, there were notable differences. The minimum CSI value and SP (%) for

related individuals in the mixed population were significantly lower than those in the sample of small local populations (CSI=0.0536 and SP=5.087319665 for mixed population vs. CSI=5.4852 and SP=84.58027509 for small local population). Additionally, the maximum value for unrelated individuals in the sample of small local populations was lower than that in the mixed population (CSI: 0.5261434 vs. 62.3941 and SP: 34.47535795 vs. 98.42256614).

Of particular note is the fact that the variability expressed through standard deviation and variance was much higher for both related and unrelated individuals in the mixed population compared to the sample of small local populations (SD: 14.31788783 vs. 3.094913584 for related individuals and 15.0150278 vs. 4.586269405 for unrelated individuals).

Moreover, the analysis of variance (ANOVA) for these samples demonstrated a statistically highly significant difference ($p<0.0001$) in the variability, with significantly higher variability observed in the mixed population (Figure 8).

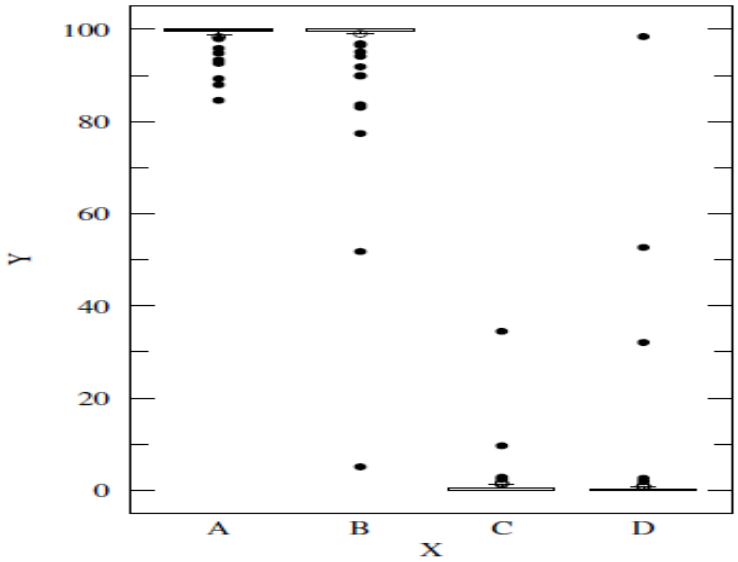

Figure 8. Distribution of the SP (%) values for the sample of small local population and the mixed population.
A- Related individuals from the small local population (Orahovica + Vukotići)
B- Related individuals from the mixed population
C- Unrelated individuals from the small local population (Orahovica + Vukotići)
D- Unrelated individuals from the mixed population

The obtained results suggest a higher homogeneity of small local populations compared to the mixed population within the population of Bosnia and Herzegovina.

4. Discussion

The genetic analysis of buccal samples collected from the village of Vukotici was found to be successful. This study is supportive of the notion that STR markers may be useful for genetic differentiation between larger and geographically more distant regions. The observed variability in peak height can be attributed to differences in DNA concentration in the extracted samples. This is likely due to the higher amount of biological material collected by the method of buccal mucosa scraping. Previous research has shown that higher DNA concentrations lead to the appearance of higher peaks, as demonstrated by Katsuya and his colleagues [31]. Excessive amounts of DNA can inhibit PCR reaction and causes the inability to detect STR profiles, this did not occur in our study. All profiles were fully detected for all 15 loci, and their "readability" was extremely high. It is common in forensic practice to yield profiles with peaks of average height, less than 50 measurement units, due to small amounts of degraded nuclear DNA. Most laboratories consider such profiles as positive

findings, provided that it is certain that the detected peak is the result of the presence of a certain allele variant, rather than an artefact. Study conducted by Wang in 2017 also revealed that buccal swab samples frequently generate DNA profiles with varying peak heights, which may be due to dissimilarities in DNA concentration in the recovered samples [30]. It is also noteworthy that the total volume of PCR in our study, as well as in the previous one, was five times lower than recommended in standard manufacturer protocols. However, laboratory experience has demonstrated that this reduced volume of PCR, along with a proportional reduction in the volume of the starting sample, does not significantly affect the ability to generate a usable DNA profile. Our findings confirm that this reduced volume of PCR is reliable for this type of research.

The results for the CSI and SP values indicate low probability of relatedness among the participants from the village of Vukotici. Interestingly, our recent study in the neighbouring village found a relatively high probability of relatedness among participants [10]. This may be due to the fact that the two villages have different migration histories and levels of genetic drift. It is known that small populations, such as those in isolated villages, can be subject to genetic drift, which can result in increased levels of relatedness among individuals [32, 33]. It is important to stress here that the CSI values in this study were calculated based on different threshold values. In general SP values provide a more intuitive and easily interpretable measure of the likelihood of relatedness [33]. In addition to the CSI and SP calculations, we also examined the distribution of allele sharing across loci to assess the degree of relatedness among the participants. This approach is useful for detecting more distant relationships as it reveals patterns of allele sharing that are indicative of shared ancestry [35].

The central tendency and variability measures presented in Table 5 offer valuable insights into the distribution and variation of the observed samples in Vukotici village. The results indicate that related individuals have significantly higher CSI compared to unrelated ones, which is consistent with prior studies demonstrating higher relatedness among individuals in small and isolated communities [36]. Additionally, the standard deviation (SD) and coefficient of variation (CV) measures reveal a slightly higher degree of variation among related individuals, potentially due to the prevalence of specific genetic variants in this population. It is important to note that the findings are limited to the Vukotici village and may not apply to other populations. The small sample size and geographic limitations of the study are potential factors that may impact the generalizability of the results. Further research involving larger, more diverse populations is necessary to confirm these findings.

In previous publications, the importance of the "gray zone" in the analysis of uncertain results has been highlighted [7, 10, 27]. Determining the sensitivity, specificity, positive and negative assumed values of selected methods is crucial in reducing false positive and false negative findings. False characterization of someone as a relative or non-relative can have a direct impact on the lives of those being tested, especially in forensic science. Using the CSI=1 method, sibship was determined in 100% of pairs of relatives and the absence of biological sibship was determined in 100% of pairs of non-relatives. Attempts to define the limits of the "gray zone" for CSI=1 were unsuccessful. CSI=1 was found to be a clear, reliable, and sufficient threshold value for proving sibship in this population. Similar results were obtained for CSI=3. The percentage of proven sibship remained the same for CSI<1 values, while the percentage of proven absence of sibship among non-relatives decreased. Our results indicate that attempts to define a "gray zone" were unsuccessful that CSI=1 and CSI=3 can be used as a reliable tool to determine sibship in small local populations, without the need for a "gray zone". Disagreements have been reported regarding the evaluation of the optimal and reliable cut-off value for sibship determination methods [7, 10, 28]. The correct determination of these methods is crucial as it can impact the lives of examiners, especially in forensic practice. One of the studies determined CSI boundary 1 as a clear boundary between relatives and non-relatives using 15 STR loci in their research but with known profiles of the examinees' parents [26]. Another study states difficulty to determine sibship between two individuals with absolute certainty using a maximum of 9 STR loci when the DNA profiles of the parents are not known, and thus neither are the mandatory alleles, and it indicated the analysis of a larger number of STR loci in proving sibship in such individuals.

5. Conclusions

In the conclusion, this investigation of genetic diversity based on short tandem repeat polymorphisms revealed genetic homogeneity among village populations. Values of 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.

Declaration of competing interest

The authors declare that they have no known financial or non-financial competing interests in any material discussed in this paper.

Funding information

No funding was received from any financial organization to conduct this research.

References

[1] R. E. C. R., *Encyclopedia of genetics*, 1st ed. New York, NY: Routledge Member of the Taylor and Francis Group, 2001.

[2] D. S. Woodruff, "Populations, Species, and Conservation Genetics," in *Encyclopedia of Biodiversity*, Elsevier, 2001, pp. 811–829.

[3] A. J. Hannan, "Tandem repeats mediating genetic plasticity in health and disease," *Nat. Rev. Genet.*, vol. 19, no. 5, pp. 286–298, 2018.

[4] J. A. Siegel and P. J. Saukko, *Encyclopedia of forensic sciences*. Academic Press, 2012.

[5] H. Fan and J.-Y. Chu, "A brief review of short tandem repeat mutation," *Genomics Proteomics Bioinformatics*, vol. 5, no. 1, pp. 7–14, 2007.

[6] M. Lancia, A. Coletti, G. Margiotta, L. Lottanti, E. Carnevali, and M. Bacci, "Allele frequencies of 15 STR loci in an Italian population," *Int. Congr. Ser.*, vol. 1288, pp. 340–342, 2006.

[7] J. Musanovic *et al.*, "Determination of combined sibship indices 'gray zone' using 15 STR loci for central Bosnian human population," *Mol. Biol. Rep.*, vol. 39, no. 5, pp. 5195–5200, 2012.

[8] R. V. Rohlfs, S. M. Fullerton, and B. S. Weir, "Familial identification: population structure and relationship distinguishability," *PLoS Genet.*, vol. 8, no. 2, p. e1002469, 2012.

[9] A. Pilav, N. Pojskić, A. Ahatović, M. Džehverović, J. Čakar, and D. Marjanović, "Allele frequencies of 15 STR loci in Bosnian and Herzegovinian population," *Croat. Med. J.*, vol. 58, no. 3, pp. 250–256, 2017.

[10] *Genetic Polymorphism of 15 STR loci in the population of two Bosnian and Herzegovinian villages*. .

[11] D. Marjanović *et al.*, "DNA identification of skeletal remains from the World War II mass graves uncovered in Slovenia," *Croat. Med. J.*, vol. 48, no. 4, pp. 513–519, 2007.

[12] Musanovic *et al.*, "Determination of combined sibship indices 'gray zone' using 15 STR loci for central Bosnian human population," *Mol. Biol. Rep.*, vol. 39, no. 5, pp. 5195–5200, 2012.

[13] A. Kovačević, *Iskre svjetlosti: istinite priče o šehidima*. Bosna i Hercegovina, 2005.

[14] S. Izdanje Državnog zavoda za statistiku Republike Bosne i Hercegovine, *"Nacionalni sastav stanovništva - Rezultati za Republiku po opštinama i naseljenim mjestima 1991." Statistički bilten br. 234*. Sarajevo: Državni zavod za statistiku Republike Bosne i Hercegovine, 1991.

[15] D. Marjanović, D. Primorac, and K. Drobnič, *Molekularna forenzična genetika. Institut za genetičko inženjerstvo i biotehnologiju*. 2009.

[16] C. E. Pu and A. Linacre, "Systematic evaluation of sensitivity and specificity of sibship determination by using 15 STR loci," *J. Forensic Leg. Med.*, vol. 15, no. 5, pp. 329–334, 2008.

[17] R. E. Wenk, M. Traver, and F. A. Chiafari, "Determination of sibship in any two persons," *Transfusion*, vol. 36, no. 3, pp. 259–262, 1996.

[18] R. I. Giroti, S. Verma, K. Singh, R. Malik, and I. Talwar, "A grey zone approach for evaluation of 15 short tandem repeat loci in sibship analysis: a pilot study in Indian subjects," *J. Forensic Leg. Med.*, vol. 14, no. 5, pp. 261–265, 2007.

[19] S. A. Miller, D. D. Dykes, and H. F. Polesky, "A simple salting out procedure for extracting DNA from human nucleated cells," *Nucleic Acids Res.*, vol. 16, no. 3, p. 1215, 1988.

[20] M. G. Ensenberger, "Developmental validation of the PowerPlex® 16 HS System: an improved 16-locus fluorescent STR multiplex," *Forensic Science International: Genetics*, vol. 4, pp. 257–264, 2010.

[21] C. H. Brenner, *DNA-VIEW 2006 user's manual*. Oakland (CA: Copyright CH Brenner, 2006.

[22] D. Marjanovic, L. Kapur, N. Pojskic, and R. Hadziselimovic, "DNA diversity in the studies of genetic distance among isolated human populations in bosnia," *Hum. Evol.*, vol. 20, no. 2–3, pp. 157–166, 2005.

[23] D. Marjanovic, L. Kapur, K. Drobnic, B. Budowle, N. Pojskic, and R. Hadziselimovic, "Comparative study of genetic variation at 15 STR loci in three isolated populations of the Bosnian mountain area," *Hum. Biol.*, vol. 76, no. 1, pp. 15–31, 2004.

[24] C. H. Tzeng *et al.*, "Determination of sibship by PCR-amplified short tandem repeat analysis in Taiwan," *Transfusion*, vol. 40, no. 7, pp. 840–845, 2000.

[25] J. Coste and J. Pouchot, "A grey zone for quantitative diagnostic and screenig tests," *Int. J. Epidemiol.*, vol. 32, no. 2, pp. 304–313, 2003.

[26] T. M. Reid, C. A. Wolf, C. M. Kraemer, S. C. Lee, M. L. Baird, and R. F. Lee, "Specificity of sibship determination using the ABI Identifiler multiplex system," *J. Forensic Sci.*, vol. 49, no. 6, pp. 1262–1264, 2004.

[27] J. Mušanović, A. Metović , E. Pepić, B, Imamović , A. Džubur- Alić, M. Filipovska- Mušanović , D. Marjanović , "Division of alleles in the determination of sibship in small local populations" *European Journal of Biomedical and Pharmaceutical Sciences*, vol 128-3, no. 4, 2017.

[28] Pu, C. E., & Linacre, A, *Systematic evaluation of sensitivity and specificity of sibship determination by using 15 STR loci*. Journal of Forensic and Legal Medicine, 2008.

[29] H. Brown *et al.*, "Development and validation of a novel multiplexed DNA analysis system, InnoTyper® 21," *Forensic Sci. Int. Genet.*, vol. 29, pp. 80–99, 2017.

[30] S. Wang, G. Shanthan, M. M. Bouzga, H. M. Thi Dinh, C. Haas, and A. E. Fonneløp, "Evaluating the performance of five up-to-date DNA/RNA co-extraction methods for forensic application," *Forensic Sci. Int.*, vol. 328, no. 110996, p. 110996, 2021.

[31] F. Takeuchi *et al.*, "The fine-scale genetic structure and evolution of the Japanese population," *PLoS One*, vol. 12, no. 11, p. e0185487, 2017.

[32] J. Wang, "Sibship reconstruction from genetic data with typing errors," *Genetics*, vol. 166, no. 4, pp. 1963–1979, 2004.

[33] J. K. Pritchard, M. Stephens, and P. Donnelly, "Inference of population structure using multilocus genotype data," *Genetics*, vol. 155, no. 2, pp. 945–959, 2000.

[34] D. Carmelli and T. Reed, "Stability and change in genetic and environmental influences on hand-grip strength in older male twins," *J. Appl. Physiol.*, vol. 89, no. 5, pp. 1879–1883, 2000.

[35] M. C. Keller and G. Miller, "Resolving the paradox of common, harmful, heritable mental disorders: which evolutionary genetic models work best?," *Behav. Brain Sci.*, vol. 29, no. 4, pp. 385–404; discussion 405-52, 2006.

[36] R. Gaytmenn, D. P. Hildebrand, D. Sweet, and I. A. Pretty, "Determination of the sensitivity and specificity of sibship calculations using AmpF lSTR Profiler Plus," *Int. J. Legal Med.*, vol. 116, no. 3, pp. 161–164, 2002.