Development of a New Screening Method “Allele Matching Cut Off Score (AMCOS)” for Faster Kinship Analysis in Cases of Mass Disasters: A Proof of Concept Study

Sonia Kakkar
PGIMER
Phulen Sarma
PGIMER
Inusha Pannigrihi
PGIMER
S. P. Mandal
PGIMER
R. K. Kumawat
State Forensic Science Laboratory, Jaipur
Pankaj Shrivastava (pankaj.shrivastava@rediffmail.com)
State Forensic Science Laboratory

Research Article

Keywords: Kinship, X-STR, Autosomal, Identical by state (IBS), Identical by descendent (IBD)

Posted Date: May 27th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-525011/v1

Abstract

kinship analysis in forensic is based on calculation of respective kinship indices. But calculation of the same is possible only when the subject under identification has been associated to a particular population whose gene frequency data is available for the particular set of markers used in forensic practices. In case of the mass disasters where a huge number of individuals are to be identified, gathering the population frequency data and calculating the kinship indices can be an intricate progression requiring a lot of time and huge resources. The present study is based on allele matching score values which doesn't require the use of allele frequency data to establish kinship. This method is based on the allele sharing approach which simply refers to the number of shared alleles (1 or 2) between two individuals; also known as identical by state (IBS) alleles which might have been inherited from a recent common ancestor in which the alleles are identical by descendent (IBD). In case of mass disasters this method can be used to narrow down the investigation by screening the number of related individuals which can further be confirmed with other tests if required. This method has been tested for various statistical parameters and has shown promising values which suggests the potential use of this method in forensic practice. This method has been tested on siblings and grandparent-grandchildren by using autosomal and X-STR markers both as the reference samples from parents can't always be available. The present study also compares the results shown by autosomal and X-STR markers in siblings and grandparent-grandchildren identification, thereby suggesting the better set of markers for siblings and grandparent-grandchildren identification.

1. Introduction

Sib ship analysis plays a vital role in individual identification in civil and criminal law cases and in searching for a missing person when the parents are absent or dead[1]. In situations where parentage (family trio) analysis is not feasible, DNA comparison with an alleged sibling may solve the purpose of identification. Since there are no obligatory alleles between the siblings that can help in excluding the case with absolute certainty, sibship analyses are more complicated[2]. It is not possible to eliminate sibship with confidence by using the genetic markers if only siblings are available for the study[3].

According to Mendelian genetics law, full siblings acquire the alleles from their parents. Probabilities that a full sibling will share 0, 1, or 2 alleles identical by descent is ¼, ½, and ¼ [4]. Studies have been conducted from time to time, to develop and access the validity of the sibling comparison test.

STR multiplex markers are the predominantly used technology for human identification[5]. Multiplex STR assays have been evaluated for their use in pairwise kinship analysis[6,2]. The study has been conducted on 9, 12, and 15-STR markers to develop a method for sibling identification[7].

The study of kinship requires the analysis of identical by descendent (IBD) alleles[8], and the research on genetic relatedness has always been linked to the root concept of IBD. Previous studies on sibship analysis have also been based on the idea of IBD [9–12]. By combining the IBD method with the Identical by state (IBS) information, the inference of genetic relatedness between the individuals (in pedigree and /or in large population-based studies) has also been reported by Stevens, 2011[13]. IBS term is used to describe two identical alleles at a locus between two individuals who do not share a recent common ancestry. IBS method was proposed by Chakraborty and Jin (1983) for the inference of a pairwise relationship[14]. On the other hand, IBD (identity by descendent) describes two identical alleles that share the common ancestry. Two individuals who share 1 or 2 alleles IBS at a given locus may have inherited the alleles from a recent common ancestor in which the alleles are IBD[13].

Yuan (2017) reported the application of autosomal STR loci with the IBS method, and a discriminant function algorithm has also been studied for their utility in Sibling identification. It was concluded from the study that STRs with higher discrimination power (PD) values should be selected when additional autosomal markers are required for full sibling identification and discriminant analysis with IBS was reported to be highly useful for the full sibling test [4].

To infer a biological relationship from pairwise genetic data in loci is based on population frequencies of the observed alleles shared by the pairs of individuals and on probability equations for genotype combinations[3,15]. The Likelihood ratio (LR) is calculated by using the frequency data to express the probability ratio of relatives to non-relatives. However, in some cases where the population frequencies of alleles may be unknown, or the ethnic origins may be unclear for foreign individuals, the LR based method fails to infer the relationship between two individuals[15].

In this study we have used the allele sharing approach which simply refers to the number of shared alleles (1 or 2) between two individuals; also known as IBS alleles which might have been inherited from a recent common ancestor in which the alleles are IBD[13,16] and developed a new method of "*Allele matching cut off score (AMCOS)*." The utility of this method was checked in siblings and grandparent-grandchildren (GP-GC) identification. We applied AMCOS on the sibling and grandparent-grandchildren data obtained from the most frequently used autosomal and relatively newer X-STR markers. The reason for choosing X-STR markers was their increasing popularity and their promising performance in kinship testing[17,18]. The use of X-STRs has also been suggested in certain pedigree analyses, which are reported to be indistinguishable by autosomal STR analysis[8]. Also, the options for chromosome X marker typing utilizing short amplicons and the ease of analysis over mtDNA, which is an intricate process[19,20], can be considered when degraded samples from the mass disasters are to be analyzed[21]. Autosomal STR analysis was chosen because the use of unlinked biallelic markers has always been a worldwide standard practice in forensic laboratories for the last two decades[22,23]. The present study based on AMCOS values will give us a cut-off score/value based on IBS allele matches (which could be IBD also) between the siblings and grandparents-grand children. The cut-off score/value can be used to shortlist the number of individuals to be matched for sibship and GP-GC identification in cases of any mass disaster or natural calamity. This AMCOS based method can help the analyst to save time and resources by short listing the number of individuals to be matched for kinship establishment in Disaster Victim Identification (DVI), which may further be confirmed by other analyses if required.

2. Results

Brother-Sister (B-S) kinship analysis by autosomal STRs.

To examine the kinship analysis between brother and sister analysis using autosomal STR markers, sensitivity and 1-specificity values at different allele matching scores were calculated. Independent t-Test showing a statistically significant difference between B-S (related) and non-B-S (unrelated) group based on OAM and TAM score were calculated. The average TAM score for the B-S group is 3.92 ± 1.754 (SD), and for the non-B-S group average, TAM is 0.80 ± 0.764 (SD). Whereas the average OAM score for the B-S group is 8.64 ± 1.846 and for the non-B-S group is 7.48 ± 1.531 (Fig. 1). Sensitivity and 1-specificity values at different allele matching scores calculated based on the ROC curve were shown in Table 1. The smallest OAM cutoff score value is the minimum observed test value minus 1, and the most considerable cutoff value is the maximum perceived test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values (Table 1).

Table 1
Sensitivity and 1-specificity values at different allele matching scores calculated based on the ROC curve (Coordinates of the curve)

Test Result Variable(s): One Allele Match Score		
OAM scores	Sensitivity (Y-axis coordinates)	1 – Specificity (X-axis coordinates)
3.00	1.000	1.000
4.50	1.000	0.960
5.50	1.000	0.880
6.50	0.800	0.800
7.50	0.720	0.480
8.50	0.520	0.280
9.50	0.400	0.080
10.50	0.160	0.000
11.50	0.040	0.000
13.00	0.000	0.000

The sensitivity and specificity of the test with AMCOS of 9 were found to be 52% and 72 %, respectively. The predictive values for positive and negative predictions were found to be 65% and 60%, respectively, and the overall accuracy of the test was found to be 62% (Table 2).

Table 2
Evaluation of AMCOS (OAM) in AMCOS in brother-sister kinship analysis by autosomal STR analysis:

Statistic	Value	95% CI
Sensitivity	52.00%	31.31–72.20%
Specificity	72.00 %	50.61–87.93%
Positive Predictive Value	65.00%	47.16–79.44%
Negative Predictive Value	60.00 %	48.25–70.70%
Accuracy	62.00%	47.17–75.35%

Analysis of TAM scores in Brother - Sister (B-S) identification.

Sensitivity and 1-specificity values at different allele matching scores calculated based on the ROC curve (Coordinates of the curve) were given in Table 3. The smallest TAM cutoff score value is the minimum observed test value minus 1, and the most considerable cutoff value is the maximum perceived test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values (Table 3).

Table 3
Sensitivity and 1-specificity values at different allele matching scores calculated based on the ROC curve (Coordinates of the curve)

TAM scores	Sensitivity (Y-axis coordinates)	1 – Specificity (X-axis coordinates)
-1.00	1.000	1.000
0.50	0.960	0.600
1.50	0.960	0.200
2.50	0.920	0.000
3.50	0.520	0.000
4.50	0.320	0.000
5.50	0.120	0.000
7.00	0.080	0.000
9.00	0.000	0.000

The sensitivity and specificity of the test with AMCOS of 3 were found to be 92% and 100 %, respectively. The predictive values for positive and negative predictions were found to be 100% and 93%, respectively, and the overall accuracy of the test was found to be 96% (Table 4).

Table 4
Evaluation of AMCOS in AMCOS (TAM) in brother-sister kinship analysis by autosomal STR markers.

Statistic	Value	95% CI
Sensitivity	92.00%	73.97–99.02%
Specificity	100.00 %	86.28–100.00%
Positive Predictive Value	100.00%	-
Negative Predictive Value	92.59 %	76.79–97.93%
Accuracy	96.00%	86.29–99.51%

Brother-sister (B-S) kinship analysis by X-STRs.

Independent t-Test showing a statistically significant difference between B-S (related) and non-B-S (unrelated) groups based on the OAM score were analyzed. Average OAM score for B-S group is 7.88 ± 2.075 (SD), whereas for the non-B-S group is 4.24 ± 1.363 (SD). ROC curve to assess the accuracy of the test in discriminating the true cases (B-S) from the false (Non-B-S) cases are shown in Fig. 2. The smallest OAM cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values (Table 5)

Table 5
Sensitivity and 1-specificity values at different allele matching scores calculated based on the ROC curve (Coordinates of the curve)

Test Result Variable(s): One allele (maternal) matching (OAM) score /12		
OAM score	Sensitivity (Y-axis coordinates)	1 – Specificity (X-axis coordinates)
1.00	1.000	1.000
2.50	1.000	0.880
3.50	0.880	0.760
4.50	0.840	0.320
5.50	0.800	0.240
6.50	0.680	0.040
7.50	0.560	0.000
8.50	0.520	0.000
9.50	0.360	0.000
10.50	0.160	0.000
11.50	0.080	0.000
13.00	0.000	0.000

The sensitivity and specificity of the test with AMCOS of 6, when applied to the B-S sibling group, was found to be 80% and 76 %, respectively. The predictive values for positive and negative predictions were found to be 77% and 79%, respectively, and the overall accuracy of the test was found to be 78% (Table 6).

Table 6
Evaluation of AMCOS in brother-sister kinship identification by X-STR markers

Statistic	Value	95% CI
Sensitivity	80.00%	59.30–93.17%
Specificity	76.00 %	54.87–90.64%
Positive Predictive Value	76.92%	61.76–87.31%
Negative Predictive Value	79.17 %	62.73–89.56%
Accuracy	78.00%	64.04–88.47%

Brother-Brother (B-B) Kinship analysis by autosomal STRs:

Independent t-Test showing a statistically significant difference between B-B (related) and non-B-B (unrelated) group based on the TAM score were analyzed. The average TAM score for B-B is 4.85 ± 1.496 (SD), whereas, for the non-B-B

group, the average TAM is 0.60 ± 0.681 (SD). Only two allele matching (TAM) scores showed a significant difference between the two (related and unrelated) groups (Fig. 3). So only the TAM score was evaluated for its efficiency as a biomarker for brother-brother kinship analysis by autosomal STR markers. Sensitivity and 1-specificity values at different allele matching scores calculated based on the ROC curve were calculated (Table 7). The smallest TAM score cutoff value is the minimum observed test value minus 1, and the most considerable cutoff value is the maximum perceived test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.

Table 7
Sensitivity and 1-specificity values at different allele matching scores calculated based on the ROC curve (Coordinates of the curve)

Test Result Variable(s): Two Allele Match Score		
TAM scores	Sensitivity Y-axis coordinates	1 – Specificity X-axis coordinates
-1.00	1.000	1.000
0.50	1.000	0.500
1.50	1.000	0.100
2.50	0.950	0.000
3.50	0.800	0.000
4.50	0.600	0.000
5.50	0.350	0.000
6.50	0.100	0.000
7.50	0.050	0.000
9.00	0.000	0.000

The sensitivity and specificity of the test with AMCOS of 3 were found to be 95% and 100 %, respectively. The predictive values for positive and negative predictions were found to be 100% and 95%, respectively, and the overall accuracy of the test was found to be 97.5% (Table 8).

Table 8
Evaluation of AMCOS in Brother-Brother identification by autosomal STR analysis.

Statistic	**Value**	**95% CI**
Sensitivity	95.00%	75.13–99.87%
Specificity	100.00 %	83.16–100.00%
Positive Predictive Value	100.00%	-
Negative Predictive Value	95.24 %	74.75–99.27%
Accuracy	97.50%	86.84–99.94%

Brother-Brother (B-B) Kinship analysis by X STR analysis:

Independent t-Test showing a statistically significant difference between B-B and non-B-B group based on the OAM score were analyzed using X-STRs. The Average OAM score for the B-B group is 9.1 ± 2.075 (SD), whereas for the non-B-B group average OAM score is 1.85 ± 0.040 (SD). ROC curve to assess the accuracy of the test in discriminating against the true sibling (Brother-Brother) cases from false cases is shown in Fig. 4. Sensitivity and 1-specificity values at different allele matching scores calculated on the basis of the ROC curve (Coordinates of the curve) were shown in Table 9.

Table 9
Sensitivity and 1-specificity values at different allele matching scores calculated on the basis of the ROC curve (Coordinates of the curve)

Test Result Variable(s): One Allele Match Score		
OAM score	Sensitivity (Y-axis coordinates)	1 – Specificity (X-axis coordinates)
-1.00	1.000	1.000
0.50	1.000	0.950
1.50	1.000	0.600
2.50	1.000	0.200
3.50	1.000	0.100
4.50	1.000	0.000
5.50	0.950	0.000
6.50	0.900	0.000
7.50	0.800	0.000
8.50	0.550	0.000
9.50	0.450	0.000
10.50	0.250	0.000
11.50	0.200	0.000
13.00	0.000	0.000

The smallest OAM cutoff score value is the minimum observed test value minus 1, and the most considerable cutoff value is the maximum perceived test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.

The sensitivity and specificity of the test with AMCOS of 5 were found to be 100 %. The predictive values for positive and negative predictions were also found to be 100%, and the overall accuracy of the test was 100% too (Table 10).

Table 10
Evaluation of AMCOS in AMCOS in Brother-bother identification by X-STR analysis

Statistic	Value	95% CI
Sensitivity	100.00%	83.16–100.00%
Specificity	100.00%	83.16–100.00%
Positive Predictive Value	100.00%	-
Negative Predictive Value	100.00 %	-
Accuracy	100.00%	91.19%to 100.00%

Sister- Sister (S-S) kinship analysis by autosomal STRs:

Independent t-Test showing a statistically significant difference between S-S (related) and non-S-S (unrelated) groups based on the TAM score were calculated. The Average TAM score for S-S is 5.45 ± 1.63 (SD), whereas, for the non-S-S group, the average TAM is 0.95 ± 0.326 (SD). Only two allele matching (TAM) scores showed a significant difference between the two (related and unrelated) groups. So only the TAM score was evaluated for its efficiency as a biomarker for sister-sister kinship analysis by autosomal STR markers (Fig. 5). Sensitivity and 1-specificity values at different allele matching scores were calculated on the basis of the ROC curve are shown in Table 11.

Table 11
Sensitivity and 1-specificity values at different allele matching scores calculated on the basis of the ROC curve (Coordinates of the curve)

Test Result Variable(s): Two Allele Match Score		
TAM score [a]	Sensitivity (Y-axis coordinates)	1 – Specificity (X-axis coordinates)
-1.00	1.000	1.000
0.50	1.000	0.650
1.50	1.000	0.300
2.50	0.950	0.000
3.50	0.850	0.000
4.50	0.750	0.000
5.50	0.550	0.000
6.50	0.250	0.000
7.50	0.100	0.000
9.00	0.000	0.000

The smallest cutoff value is the minimum observed test value minus 1, and the most considerable cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.
The sensitivity and specificity of the test with AMCOS of 3 were found to be 95% and 100 %, respectively. The predictive values for positive and negative predictions were found to be 100% and 95%, respectively, and the overall accuracy of the test was found to be 97.5% (Table 12).

Table 12
Evaluation of AMCOS in Sister-Sister identification by autosomal STR markers.

Statistic	Value	95% CI
Sensitivity	95.00%	75.13–99.87%
Specificity	100.00 %	83.16–100.00%
Positive Predictive Value	100.00%	-
Negative Predictive Value	95.24 %	74.75–99.27%
Accuracy	97.50%	86.84–99.94%

Sister- Sister (S-S) kinship analysis by X-STR analysis:

Independent t-Test showing a statistically significant difference between S-S (related) and non-S-S (unrelated) groups based on the OAM score were analyzed. The Average OAM score for S-S is 11.85 ± 0.366 (SD), whereas, for the non-S-S group, the average OAM is 5.90 ± 2.049 (SD). ROC curve and AUC to assess the accuracy of the test in discriminating the true cases (S-S cases) from the false cases (non-S-S cases) (Fig. 6). Sensitivity and 1-specificity values at different allele matching scores calculated on the basis of the ROC curve are shown in Table 13.

Table 13
Sensitivity and 1-specificity values at different allele matching scores calculated on the basis of the ROC curve (Coordinates of the curve).

Test Result Variable(s): One Allele Match Score		
OAM score [a]	Sensitivity (Y-axis coordinates)	1 – Specificity (X-axis coordinates)
1.00	1.000	1.000
3.00	1.000	0.900
4.50	1.000	0.800
5.50	1.000	0.550
6.50	1.000	0.450
7.50	1.000	0.150
8.50	1.000	0.100
9.50	1.000	0.050
10.50	1.000	0.000
11.50	0.850	0.000
13.00	0.000	0.000

The smallest OAM cutoff score value is the minimum observed test value minus 1, and the most considerable cutoff value is the maximum perceived test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.
The sensitivity and specificity of the test with AMCOS of 11 were found to be 100%. The predictive values for positive and negative predictions were also found to be 100%, and the overall accuracy of the test was also 100% (Table 14).

Table 14
Evaluation of AMCOS in Sister-Sister identification by X-STR analysis.

Statistics	Value	95% CI
Sensitivity	100.00%	83.16–100.00%
Specificity	100.00 %	83.16–100.00%
Positive Predictive Value	100.00%	-
Negative Predictive Value	100.00 %	-
Accuracy	100.00%	91.19–100.00%

Grandparents –Grandchildren (GP-GC) by autosomal STR analysis:

Independent t-Test showing a statistically significant difference (Statistically significant difference ($p < 0.05$) between GP-GC (related) and non-GP-GC (unrelated) group based on the OAM score were analyzed. The Average OAM score for GP-GC is 11.54 ± 2.64 (SD), whereas, for the non-GP-GC group, the average OAM is 8.27 ± 2.146 (SD). The ROC curve and AUC to assess the accuracy of the test in discriminating the true cases (Grand parentage cases) from the false cases (non-Grand parentage cases) (Fig. 7). Sensitivity and 1-specificity values at different allele matching scores calculated based on the ROC curve were shown in Table 15.

Table 15
Sensitivity and 1-specificity values at different allele matching scores calculated based on the ROC curve (Coordinates of the curve):

Test Result Variable(s): OAM		
OAM scores[a]	Sensitivity (Y-axis coordinates)	1 – Specificity (X-axis coordinates)
3.00	1.000	1.000
4.50	1.000	0.962
5.50	1.000	0.923
6.50	0.962	0.769
7.50	0.962	0.615
8.50	0.923	0.462
9.50	0.852	0.350
10.50	0.654	0.154
11.50	0.577	0.038
12.50	0.385	0.038
13.50	0.192	0.000
15.00	0.000	0.000

[a]. The smallest cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.

The sensitivity and specificity of the test with AMCOS of 10 were found to be 85% and 65 %, respectively. The predictive values for positive and negative predictions were found to be 74% and 85%, respectively, and the overall accuracy of the test was found to be approx 79 % (Table 16).

Table 16
Evaluation of AMCOS in Grand-parentage identification by autosomal STR markers.

Statistic	Value	95% CI
Sensitivity	85.20%	69.85–97.55%
Specificity	65.00 %	48.21–85.67%
Positive Predictive Value	74.19%	61.37–83.88%
Negative Predictive Value	85.71 %	66.75–94.72%
Accuracy	78.85%	65.30–88.94%

Grandparents –Grandchildren (GP-GC) by X-STR analysis:

Grandparents –Grandchildren relationship are found in paternal and maternal side. Therefore in this study kinship analysis of GP-GC paternal side and maternal side studied as follows.

Part 1: Paternal grandparents.

Independent t-Test showing a statistically significant difference between paternal GP-GC (related) and non-GP-GC (unrelated) group based on the OAM score were examined. The average OAM score for GP-GC is 12 ± 0.00 (SD), whereas, for the non-GP-GC group, the average OAM is 6.57 ± 0.976 (SD). ROC curve and AUC to assess the accuracy of the test in discriminating the true cases (Grand parentage cases) from the false cases (non-Grand parentage cases) (Fig. 8). Sensitivity and 1-specificity values at different allele matching scores calculated on the basis of the ROC curve (Coordinates of the curve) were shown in Table 17.

Table 17
Sensitivity and 1-specificity values at different allele matching scores calculated on the basis of the ROC curve (Coordinates of the curve)

Test Result Variable(s): one allele matching score from PGM to Grand DAUGHTER		
OAM scores[a]	Sensitivity (Y-axis coordinates)	1 – Specificity (X-axis coordinates)
4.00	1.000	1.000
5.50	1.000	0.857
6.50	1.000	0.571
7.50	1.000	0.143
10.00	1.000	0.000
13.00	0.000	0.000

[a]. The smallest OAM cutoff score value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed

test values.
The sensitivity and specificity of the test with AMCOS of 10 were found to be 100 %. The predictive values for positive and negative predictions were found to be 100%, and the overall accuracy of the test was also found to be 100% (Table 18).

Table 18
Evaluation of AMCOS in paternal Grand-parentage determination by X-STR analysis.

Statistic	Value	95% CI
Sensitivity	100.00%	59.04–100.00%
Specificity	100.00 %	59.04–100.00%
Positive Predictive Value	100.00%	-
Negative Predictive Value	100.00 %	-
Accuracy	100.00%	76.84–100.00%

Part2: Maternal Grandparents.

Independent t-Test showing a statistically significant difference between maternal GP-GC (related) and non-GP-GC (unrelated) group based on the OAM score were analyzed. The average OAM score for GP-GC is 8.85 ± 2.794 (SD), whereas, for the non-GP-GC group, the average OAM is 3.15 ± 1.625 (SD). ROC curve and AUC to assess the accuracy of the test in discriminating the true cases (Grand parentage cases) from the false cases (non-Grand parentage cases) were shown in Fig. 9. Sensitivity and 1-specificity values at different allele matching scores calculated on the basis of the ROC curve (Coordinates of the curve) were shown in Table 19.

Table 19
Sensitivity and 1-specificity values at different allele matching scores calculated on the basis of the ROC curve (Coordinates of the curve).

Test Result Variable(s): ONE ALLELE MATCHING SCORE /12		
OAM score [a]	Sensitivity (Y-axis coordinates)	1 – Specificity (X-axis coordinates)
-1.00	1.000	1.000
0.50	1.000	0.923
1.50	1.000	0.846
2.50	1.000	0.769
3.50	1.000	0.308
4.50	0.846	0.231
5.50	0.846	0.077
6.50	0.769	0.000
8.00	0.692	0.000
9.50	0.462	0.000
10.50	0.385	0.000
11.50	0.154	0.000
13.00	0.000	0.000

[a]. The smallest OAM cutoff score value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.

The sensitivity and specificity of the test with AMCOS of 6 were found to be 85% and 92%, respectively. The predictive values for positive and negative predictions were found to be 92% and 86%, respectively, and the overall accuracy of the test was found to be 88% (Table 20).

Table 20
Evaluation of AMCOS in maternal Grandparent – Grandchildren identification by X-STR markers

Statistic	Value	95% CI
Sensitivity	84.62%	54.55–98.08%
Specificity	92.31%	63.97–99.81%
Positive Predictive Value	91.67%	62.26–98.65%
Negative Predictive Value	85.71%	62.42–95.59%
Accuracy	88.46%	69.55–97.55%

Over all, above mentioned results were presented in the Table 21 and Table 22.

Table 21
AMCOS for autosomal and X-STR markers in various sorts of kinship case

Type of kinship	Allele Match cut off score (AMCOS)			
	Autosomal STR		X-STR	
	TAM	OAM	OAM	TAM
B-S	3	9	6	NA
B-B	3	NS	5	NA
S-S	3	NS	11	NS
GP-GC	NA	10	6 for MGP & 10 for PGP	NA

NA: Not applicable, NS: Non- significant.

MGP: Maternal Grandparents.

PGP: Paternal Grandparents.

Table 22: Comparison table for autosomal and X-STR markers with respect to the statistical parameters calculated for various kinship analyses.

PARAMETER	B-S		B-B		S-S		GP-GC		
	X	AUTO	X	AUTO	X	AUTO	X		AUTO
							Maternal-GP	Paternal-GP	
Sensitivity	80% 59.30% to 93.17% (95% CI)	92% 73.97% to 99.02% (95% CI)	100% 83.16% to 100% (95%CI)	95% 75.13% to 99.87% (95% CI)	100% 83.16% to 100% (95% CI)	95% 75.13 % to 99.87% (95%CI)	84.62% 54.55% to 98.08% (95%CI)	100% 59.04% to 100% (95%CI)	85.20% 69.85% to 97.55% (95%CI)
Specificity	76% 54.87% to 90.64% (95% CI)	100% 86.28% to 100% (95% CI)	100% 83.16% to 100% (95%CI)	100% 83.16% to 100% (95% CI)	100% 83.16% to 100% (95% CI)	100% 83.16 % to 100% (95%CI)	92.31 % 63.97% to 99.81% (95%CI)	100% 59.04% to 100% (95%CI)	65% 48.21% to 85.67% (95%CI)
PPV	76.92% 61.76% to 87.81% (95% CI)	100%	100%	100%	100%	100%	91.67% 62.26% To 98.65% (95%CI)	100%	74.19% 61.37% to 83.88% (95%CI)
NPV	79.17% 62.73% to 89.56% (95% CI)	92.59% 76.79% to 97.93% (95% CI)	100%	95.24% 74.75% to 99.27% (95% CI)	100%	95.24% 74% to 99.27% (95%CI)	85.71 % 62.42% To 95.59% (95%CI)	100%	85.71% 66.75% to 94.72% (95%CI)
ACCURACY	78% 64.04% to 88.47% (95% CI)	96% 86.29% to 99.51% (95% CI)	100% 91.19% to 100% (95%CI)	97.50% 86.84% to 99.94% (95% CI)	100% 91.19% to 100% (95% CI)	97.50% 86.84% to 99.94% (95%CI)	88.46% 69.55% To 97.55% (95%CI)	100% 76.84% to 100% (95%CI)	78.85% 65.30% to 88.94% (95%CI)

Discussion

In cases of mass disasters the dead bodies or their mortal remains have to be identified and handed over to the grieving families to perform the last rights and for other civil matters like insurance, property and job claims .The number of sample pairs to be matched in such cases is enormous and the time frame is short. Such situation calls the need of a screening method which screens out the number of sample pairs (dead and its kin) to be analyzed for relatedness, which can further be confirmed for kinship. To avoid the wastage of time and resources the present study sets a standard “allele match cut off score (AMCOS)” for both the marker sets (Autosomal and X-STR) in various kinship analyses (Table 21) for the purpose of screening out pairs out of hundreds and thousands of individuals and the dead bodies /remains to be tested for kinship in cases of mass disasters, this method is solely based on allele matches at different loci and doesn’t require any allele frequency data. To make the evidence more comprehensible in the court of law, the forensic reports for Human Identification (HID) are presented in the form of likelihood ratios (LRs), and to calculate the LRs, allele frequencies are required for the population, which the person under-identification belongs to. In a diverse and developing country like India and many other developing countries, where the resources are scarce, the population data is rarely available. Also in cases of intra and inter-population migrations it gets difficult to obtain a population specific database. The AMCOS method has been devised to be used in such a condition.

The study uses two set markers, autosomal and X-STRs, for the same set of kinship analyses (B-S, B-B, S-S, and GP-GC). In B-S analysis by autosomal STR the significant TAM score of 3 and OAM score of 9 was found to be 92% sensitive with a specificity of 100% and accuracy of 96%. On the other hand OAM of 6 was found to have sensitivity, specificity and accuracy of 80%, 76% and 78% respectively, when B-S analysis by performed by X-STR. In B-B analysis by autosomal STR, the sensitivity, specificity and accuracy were 95%, 100% and 97.5% respectively, while by X-STR the same set of B-B cases showed a sensitivity, specificity and accuracy of 100%. Similarly S-S analysis showed a sensitivity, specificity and accuracy of 100% with X-STRs and the same showed a sensitivity, specificity and accuracy of 95%, 100% and 97.5% respectively, when autosomal STRs were used for the analysis. The X-STRs have shown better values of statistical parameters in GP-GC identification cases with a sensitivity, specificity and accuracy of 100% in paternal GP-GC and 84.6%, 92.31% and 88.46% respectively in maternal GP-GC cases.

The outcome of the study shows that X-STRs seemingly performs better relating to the statistical parameters like sensitivity, specificity in B-B, S-S, and GP-GC identification cases, while the negative predictive value (NPV), and positive predictive value (PPV) remained the same with both autosomal and X-STRs. Whereas, autosomal STR analysis showed better values of all the statistical parameters in B-S identification cases. We tried the AMCOS method for GP-GC screening by autosomal STR analysis, which otherwise is reported to be indistinguishable by the unlinked autosomal markers with LR based methods. The AMCOS method gave fairly good values of statistical parameters with GP-GC identification cases. Though, the technique needs to be validated in a larger sample size of GP-GC pairs. The present study is a proof of concept based study and needs to be confirmed in a larger sample size of siblings and GP-GC.

To the best of the author's knowledge, AMCOS based method has never been used earlier to establish the kinship. The present study shows the successful application of AMCOS method to identify siblings and GP-GC relationships. The results support the potential use of this technique in forensic settings to identify siblings and GP-GC. Besides that, the present study has compared the results of X-STR and autosomal STR analysis in the same samples concerning statistical and forensic parameters and has suggested the use of a better set of markers for the above mentioned kinship analyses in question.

Material and methodology

The study was commenced after taking ethical clearance from the internal ethical committee of Post Graduate Institute of Medical Education, and Research vide letter no: INT/IEC/2016/2409, dated: Oct 4th 2016. Informed consent was taken from all subjects including parents or grandparents and/or legal guardian, if subjects are under 18 years. 1ml peripheral blood sample was withdrawn from volunteer siblings and Grandparents-Grandchildren. The study was conducted at the Department of Forensic medicine, PGIMER, Chandigarh. All methods were performed in accordance with the relevant guidelines and regulations. Total of 170 pairs, 50 B-S (25 test and 25 control), 40 B-B (20 test and 20 control), 40 S-S (20 test and 20 control) and 40 GP-GC (20 test and 20 control), were studied. The kinship was confirmed verbally from parents, siblings, and grandparents. Also, to ensure the sibship and grandparentage, we followed the certainty threshold for likelihood ratios and selected the pairs with kinship indices between 100-1000(or >1000)[24]. Since all the studied pairs hailed from the Punjab region of India, the population allele frequencies were calculated for the same population by using the GenAlEx 6.5 software[25], and distant kinship (Sib ship and GP-GC) indices were calculated by using the X-STR data and FamLinkX software[26]. X-STR data was used to confirm kinship because unlinked autosomal STR markers are not efficient enough to distinguish the pedigrees like GP-GC[8,27]. Although samples were collected from all the known families, the parentage of the mother-father-child trio was confirmed for both siblings by paternity analysis. In the case of siblings and grandchildren, the majority of the volunteers were adults, and in the case of children, written informed consent was obtained from the parents or grandparents.

1. ***Extraction:*** DNA extraction was done by using the QIAAMP DNA BLOOD MINI KIT (Qiagen, Hilden, Germany) as per the manufacturer's recommended protocol.
2. ***Quantification and amplification:*** The extracted DNA samples were quantified using QUANTIFILER HUMAN DNA QUANTIFICATION KIT (ABI, Thermo Fisher Scientific, US) as per the manufacturer's recommended protocol. The samples were then amplified by using AMPF*L*STR IDENTIFILER PLUS PCR AMPLIFICATION KIT (Thermo Fisher Scientific, USA) for 15 autosomal markers (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818, and FGA) and INVESTIGATOR ARGUS X-12 PCR AMPLIFICATION KIT (Qiagen, Hilden, Germany) for 12 X STR markers (DXS10103, DXS8378, DXS7132, DXS10134, DXS10074, DXS10101, DXS10135, DXS7423, DXS10146, DXS10079, HPRTB, and DXS10148). Amplification of 500pg (picogram) DNA was performed according to the manufacturer's recommended protocol of both the kits, except half the reaction volume was used[28].
3. ***Fragment analysis:*** Samples were run on genetic analyzer 3100 (Thermo Fisher Scientific, USA) using POP-4 with dye set G5 LIZ 500 (Thermo Fisher Scientific, USA) and BTO-550 (Qiagen, Hilden, Germany) were used as size standards for autosomal and X-STR analysis, respectively. Fragment analysis was performed as per the manufacturer's recommended protocol. Obtained profiles were obtained and analyzed using profile quality parameters (data not shown).
4. ***Data analysis***: By using GeneMapper® ID software V3.2.1 (Thermo Fisher Scientific, USA), the data were analyzed. A peak detection threshold of 50 RFUs was used for allele designation. Alleles were designated on the basis of the number of allele's repeats and in accordance with the guidelines of IFSG by the help of allelic ladders provided by the manufactures of both the kits, AMPF*L*STR IDENTIFILER PLUS PCR AMPLIFICATION KIT (Thermo Fisher Scientific, USA) and INVESTIGATOR ARGUS X-12 PCR AMPLIFICATION KIT (Qiagen, Hilden, Germany).

Methodology for the calculation of allele matching score (AMS), which is further used to calculate the AMCOS, in siblings and Grandparents-Grandchildren (GP-GC) for further statistical analysis is given in supplementary tables S1 to S8.

Statistical analysis:

1. Independent t-test was applied on test pairs (related) and corresponding control pairs (unrelated) of B-S, B-B, S-S, and GP-GC.
2. If $p<0.05$, the difference between the related and unrelated groups was considered significant, and that particular group (B-S, B-B, S-S, or GP-GC) was considered for further analysis.
3. Further statistical analysis was done by using SPSS software version 22.0. Using the same receiver-operator curve (ROC) was drawn, area under this curve (AUC) depicted the efficiency of the test (in percentage) in discriminating the true cases from false cases of relatedness amongst all the groups (B-S, B-B, S-S, and GP-GC).
4. From the coordinates of the curve (ROC), the score with the maximum value of sensitivity and specificity (i.e., the minimum value of 1-specificity) was chosen as the allele matching cut of the score (AMCOS). The AMCOS was selected for all the groups (B-S, B-B, S-S, and GP-GC) in a similar fashion.
5. The allele matching score (AMS) in each pair of siblings and GP-GC (test and control both) was calculated (mentioned in supplementary tables: S1-S8) and compared against AMCOS ($\geq$ or < AMCOS).
6. Based on this comparison, a two by two table was drawn for each group (B-S, B-B, S-S, and GP-GC) which showed the entire true positive (allele matching score $\geq$ AMCOS), true negative (allele matching score <AMCOS), false negative (related but allele matching score < AMCOS) and false-positive cases (unrelated but allele matching score $\geq$AMCOS). The method was followed in all the pairs of both test (related) and control (unrelated) groups of (B-S, B-B, S-S, and GP-GC) (2x2 tables not shown in results).
7. The obtained value of two by two table in all the groups (B-S, B-B, S-S, and GP-GC) were further subjected to statistical analysis to evaluate the AMCOS by calculating the other parameters like sensitivity, specificity, negative predictive value (NPV), positive predictive value (PPV) and accuracy of the test. The above-mentioned parameters were calculated by using software MedCalc for windows version 15.0.
8. The sensitivity of the test here is its ability to identify true positive cases of kinship, i.e., related individuals. In contrast, specificity defines the strength of the test method to detect the true negative cases of kinship, i.e., unrelated individuals alleged to be kins.
9. Positive predictive value (PPV) is the probability that the individuals who have been identified as related (kinship is established) are actually related. Similarly, negative predictive value (NPV) is the probability of identifying unrelated individuals as unrelated.

Declarations

Acknowledgments

The authors are thankful to Dr. Inusha Pannigrihi for allowing working in her laboratory facilities and all the volunteer donors who participated in the study.

Author contribution:

SK planned the study, collected the samples, did the laboratory work and prepared the manuscript, PS did the statistical analysis, RK helped in statistical analysis and manuscript preparation, IP and SPM planned the study and reviewed the manuscript, PS (corresponding author) planned the study, conducted the experiments, supervised the scientific analysis and reviewed the manuscript. All the authors read and approved the final manuscript.

Competing financial interests: The authors declare no competing financial interests.

Additional information

Supplementary information is available for this paper at----------

Correspondence and requests for materials should be addressed to P.S.

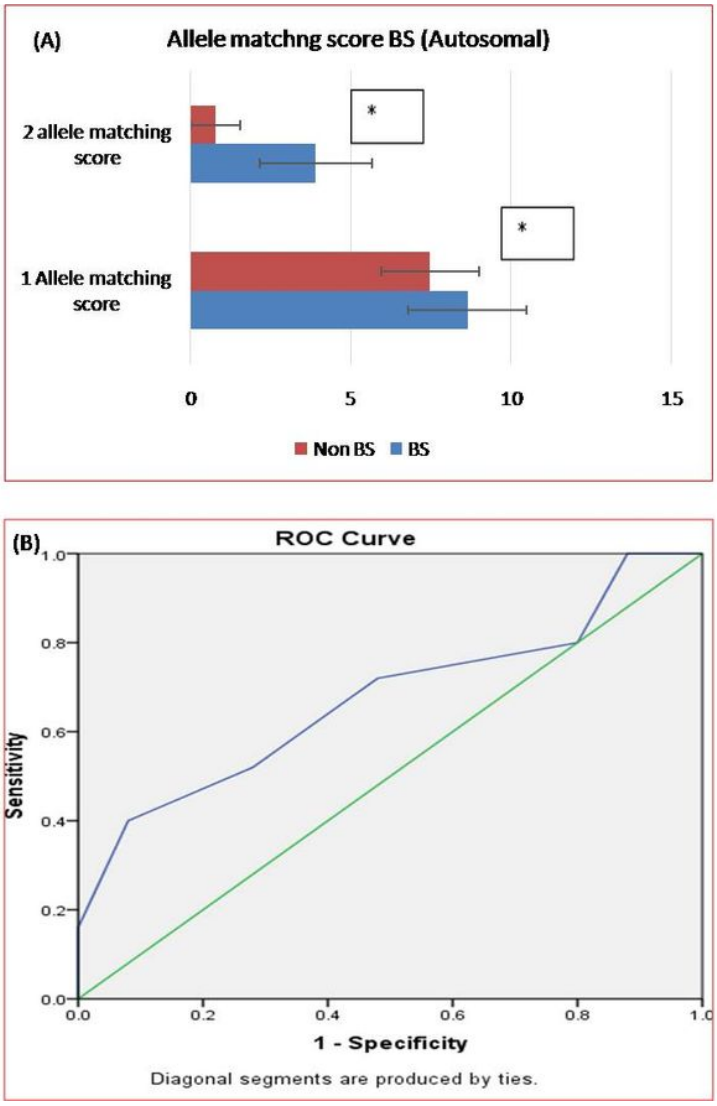
Reprints and permissions information is available at----------

References

1. Reid, T. M., Baird, M. L., Reid, J. P., Lee, S. C. & Lee, R. F. Use of sibling pairs to determine the familial searching efficiency of forensic databases. *Forensic Sci. Int. Genet.* **2**, 340–342 (2008).
2. Gaytmenn, R., Hildebrand, D. P., Sweet, D. & Pretty, I. A. Determination of the sensitivity and specificity of sibship calculations using AmpF/STR profiler plus. *Int. J. Legal Med.* **116**, 161–164 (2002).
3. Wenk, R. E., Traver, M. & Chiafari, F. A. Determination of sibship in any two persons. *Transfusion* **36**, 259–262 (1996).
4. Yuan, L. *et al.* Study of autosomal STR loci with IBS method in full sibling identification. *Leg. Med.* **26**, 14–17 (2017).
5. Kimpton, C. P. *et al.* Validation of highly discriminating multiplex short tandem repeat amplification systems for individual identification. *Electrophoresis* **17**, 1283–1293 (1996).
6. Tamura, T., Osawa, M., Ochiai, E., Suzuki, T. & Nakamura, T. Evaluation of advanced multiplex short tandem repeat systems in pairwise kinship analysis. *Leg. Med.* **17**, 320–325 (2015).
7. Tzeng, C. H. *et al.* Determination of sibship by PCR-amplified short tandem repeat analysis in Taiwan. *Transfusion* **40**, 840–845 (2000).
8. Pinto, N., Gusmão, L. & Amorim, A. X-chromosome markers in kinship testing: A generalisation of the IBD approach identifying situations where their contribution is crucial. *Forensic Sci. Int. Genet.* **5**, 27–32 (2011).
9. Zhao, S.-M. *et al.* [Establishment of universal algorithms for commonly used kinship indices between two individuals]. *Fa Yi Xue Za Zhi* **27**, 330–3 (2011).
10. Pu, C. E. & Linacre, A. Systematic evaluation of sensitivity and specificity of sibship determination by using 15 STR loci. *J. Forensic Leg. Med.* **15**, 329–334 (2008).
11. García, D., Carleos, C., Parra, D. & Cañón, J. Sib-parentage testing using molecular markers when parents are unknown. *Anim. Genet.* **33**, 364–71 (2002).
12. Browning, S. R. & Browning, B. L. Identity by Descent Between Distant Relatives: Detection and Applications. *Annu. Rev. Genet.* **46**, 617–633 (2012).
13. Stevens, E. L. *et al.* Inference of relationships in population data using identity-by-descent and identity-by-state. *PLoS Genet.* **7**, (2011).
14. CHAKRABORTY, R. & JIN, L. Determination of Relatedness between Individuals Using DNA Fingerprinting. *Human Biology* **65**, 875–895
15. Tamura, T. *et al.* Evaluation of the allele-sharing approach, known as the IBS method, in kinship analysis. *J. Forensic Leg. Med.* **20**, 112–116 (2013).
16. Aoki, Y., Nakayama, Y., Saigusa, K., Nata, M. & Hashiyada, M. Comparison of the likelihood ratio and identity-by-state scoring methods for analyzing sib-pair test cases: a study using computer simulation. *Tohoku J. Exp. Med.* **194**, 241–50 (2001).
17. Szibor, R. *et al.* Use of X-linked markers for forensic purposes. *Int. J. Legal Med.* **117**, 67–74 (2003).

18. Szibor, R. X-chromosomal markers: past, present and future. *Forensic Sci. Int. Genet.* **1**, 93–99 (2007).
19. Parson, W., Parsons, T. J., Scheithauer, R. & Holland, M. M. Population data for 101 Austrian Caucasian mitochondrial DNA d-loop sequences: application of mtDNA sequence analysis to a forensic case. *Int. J. Legal Med.* **111**, 124–132 (1998).
20. Budowle, B., Allard, M. W., Wilson, M. R. & Chakraborty, R. Forensics and mitochondrial DNA: applications, debates, and foundations. *Annu. Rev. Genomics Hum. Genet.* **4**, 119–141 (2003).
21. Asamura, H., Sakai, H., Kobayashi, K., Ota, M. & Fukushima, H. MiniX-STR multiplex system population study in Japan and application to degraded DNA analysis. *Int. J. Legal Med.* **120**, 174–181 (2006).
22. Divne, A. M., Edlund, H. & Allen, M. Forensic analysis of autosomal STR markers using Pyrosequencing. *Forensic Sci. Int. Genet.* **4**, 122–129 (2010).
23. Schneider, P. M. Beyond STRs: The Role of Diallelic Markers in Forensic Genetics. *Transfus. Med. Hemother.* **39**, 176–180 (2012).
24. Evett, I. Strength of evidence in likelihood ratios. in *The use of statistics in forensic science* (Ellis Horwood Limited, 1991).
25. Peakall, R. & Smouse, P. E. GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* (2012). doi:10.1093/bioinformatics/bts460
26. Kling, D., Dell'Amico, B. & Tillmar, A. O. FamLinkX - Implementation of a general model for likelihood computations for X-chromosomal marker data. *Forensic Sci. Int. Genet.* **17**, 1–7 (2015).
27. Pinto, N., Gusmão, L. & Amorim, A. X-chromosome markers in kinship testing: A generalisation of the IBD approach identifying situations where their contribution is crucial. *Forensic Sci. Int. Genet.* **5**, 27–32 (2011).
28. Shrivastava, P., Jain, T., Gupta, U. & Trivedi, V. Ben. Genetic polymorphism study on 12 X STR loci of investigator Argus X STR kit in Bhil tribal population of Madhya Pradesh, India. *Leg. Med.* **17**, 214–217 (2015).

Figures



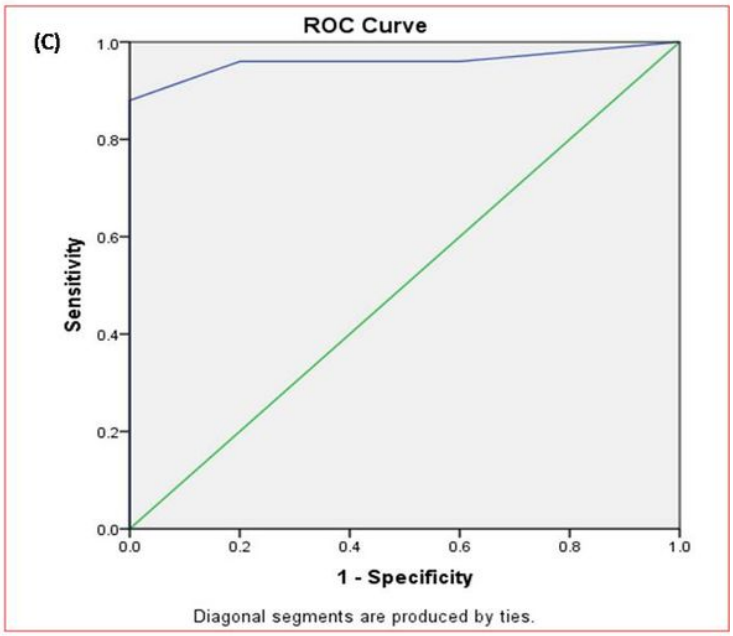
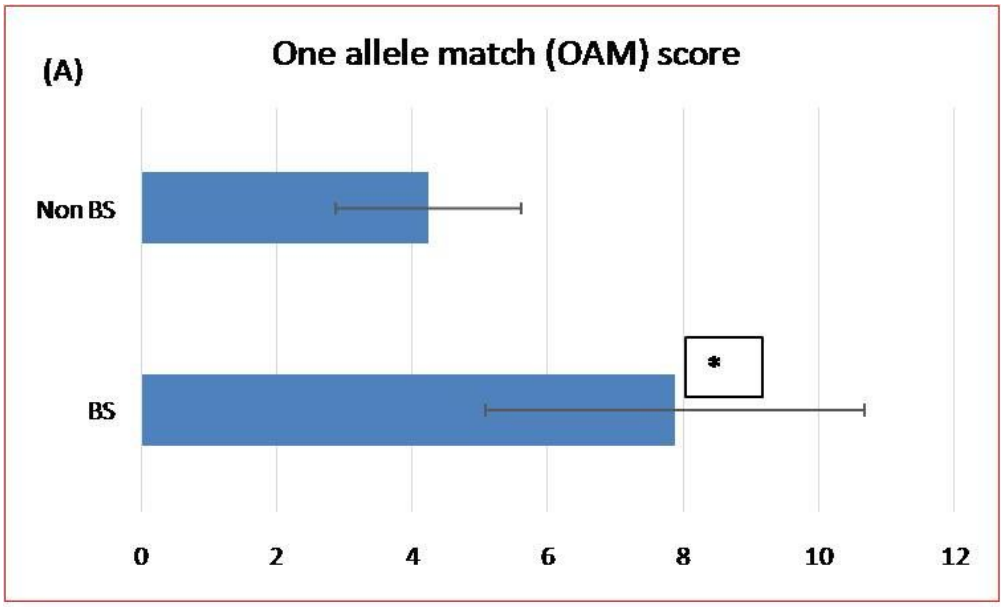


Figure 1

Independent t-Test showing a statistically significant difference between B-S (related) and non-B-S (unrelated) group based on OAM and TAM score were calculated. The average TAM score for the B-S group is 3.92±1.754 (SD), and for the non-B-S group average, TAM is 0.80±0.764 (SD). Whereas the average OAM score for the B-S group is 8.64 ± 1.846 and for the non-B-S group is 7.48±1.531



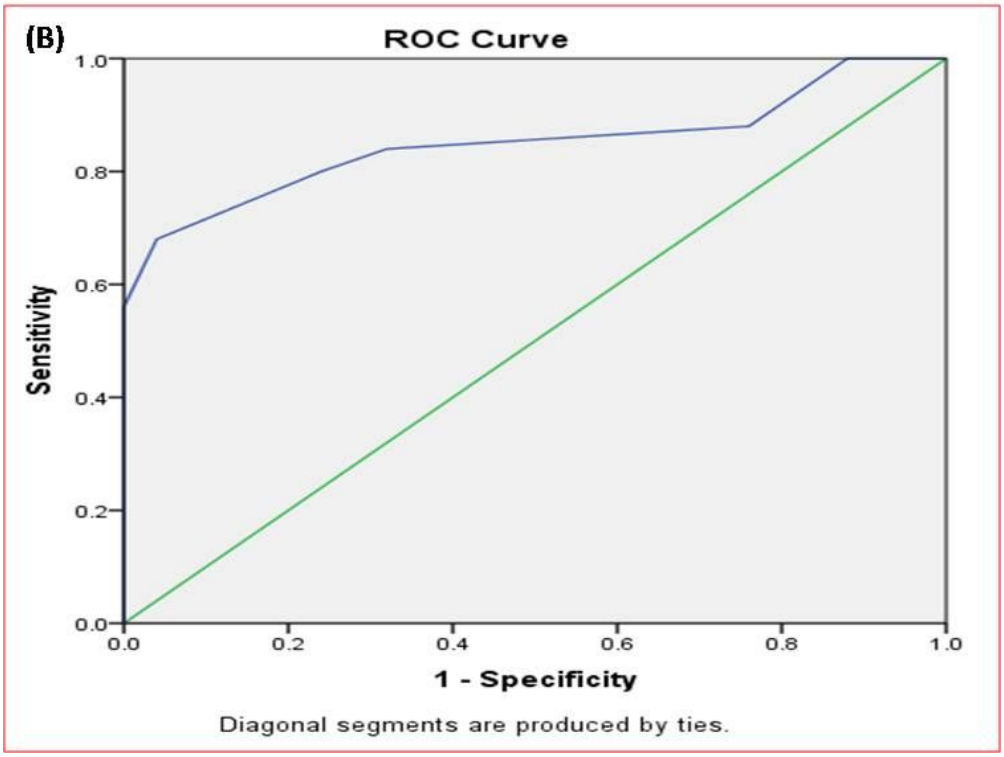
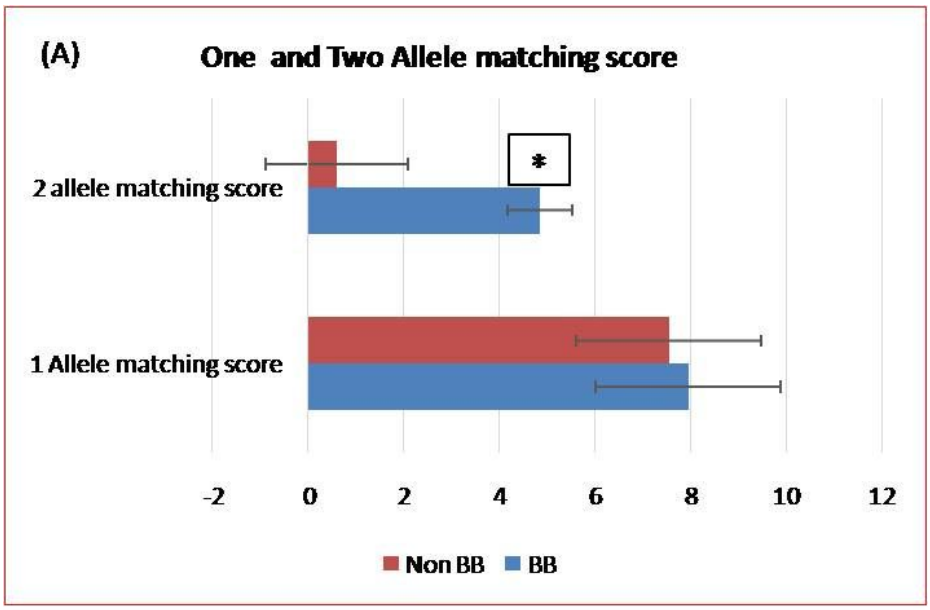


Figure 2

Independent t-Test showing a statistically significant difference between B-S (related) and non-B-S (unrelated) groups based on the OAM score were analyzed. Average OAM score for B-S group is 7.88±2.075 (SD), whereas for the non-B-S group is 4.24±1.363 (SD).



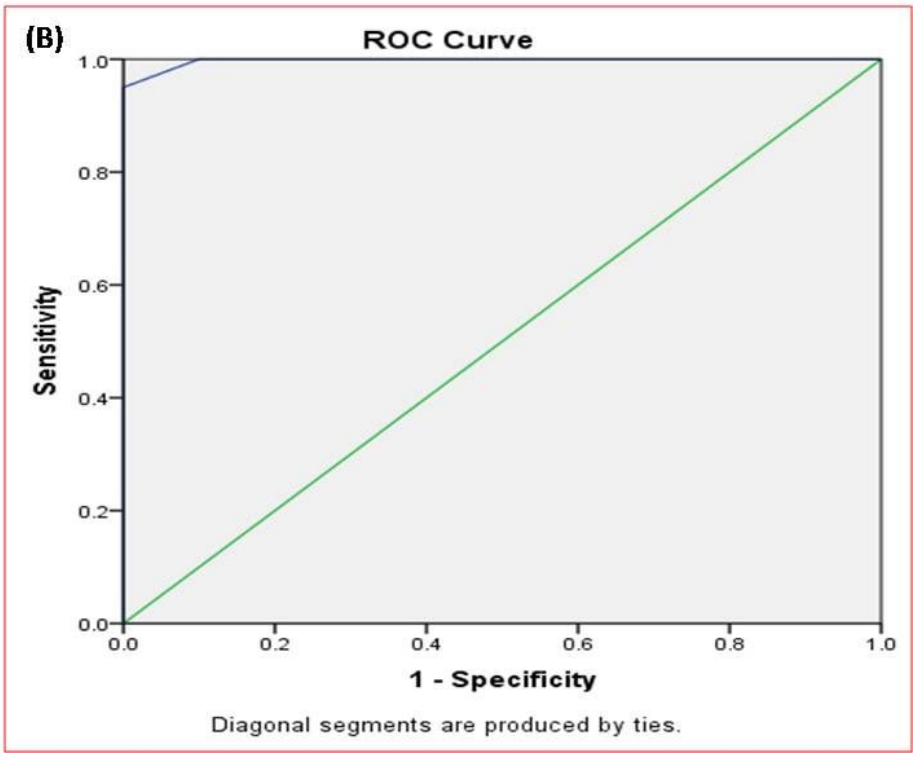


Figure 3

Independent t-Test showing a statistically significant difference between B-B (related) and non-B-B (unrelated) group based on the TAM score were analyzed. The average TAM score for B-B is 4.85±1.496 (SD), whereas, for the non-B-B group, the average TAM is 0.60±0.681 (SD).

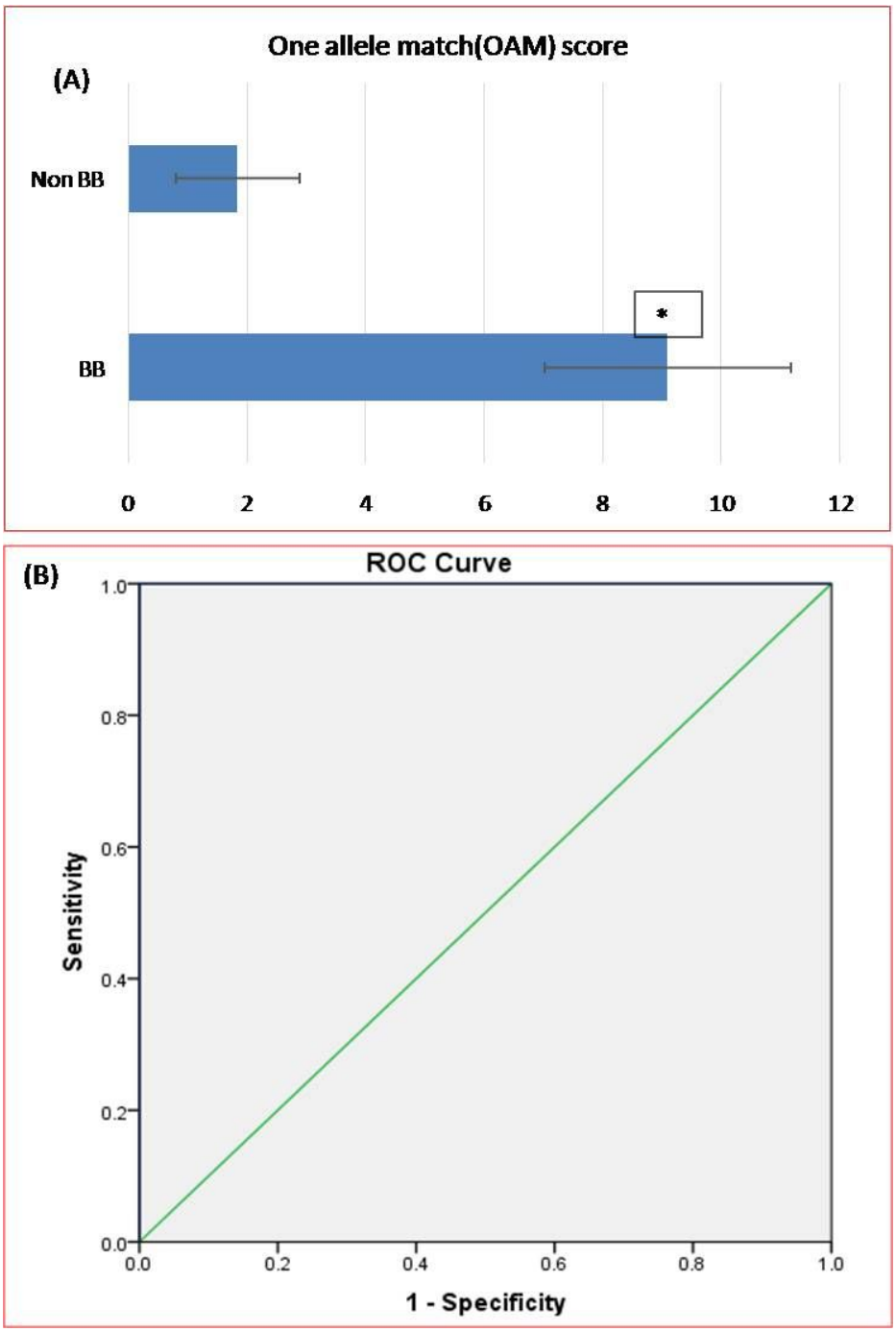


Figure 4

Independent t-Test showing a statistically significant difference between B-B and non-B-B group based on the OAM score were analyzed using X-STRs. The Average OAM score for the B-B group is 9.1±2.075 (SD), whereas for the non-B-B group average OAM score is 1.85±0.040 (SD).

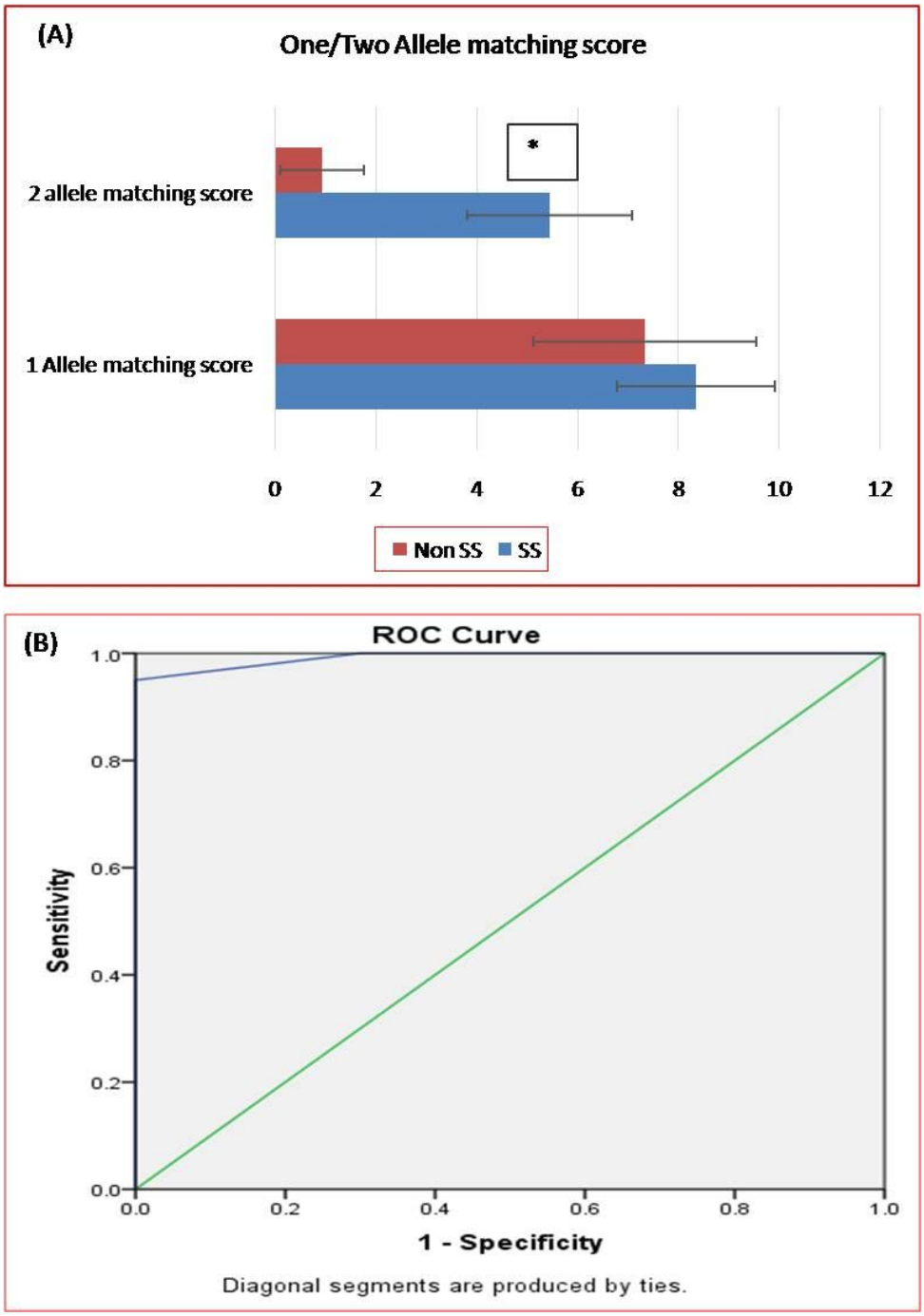
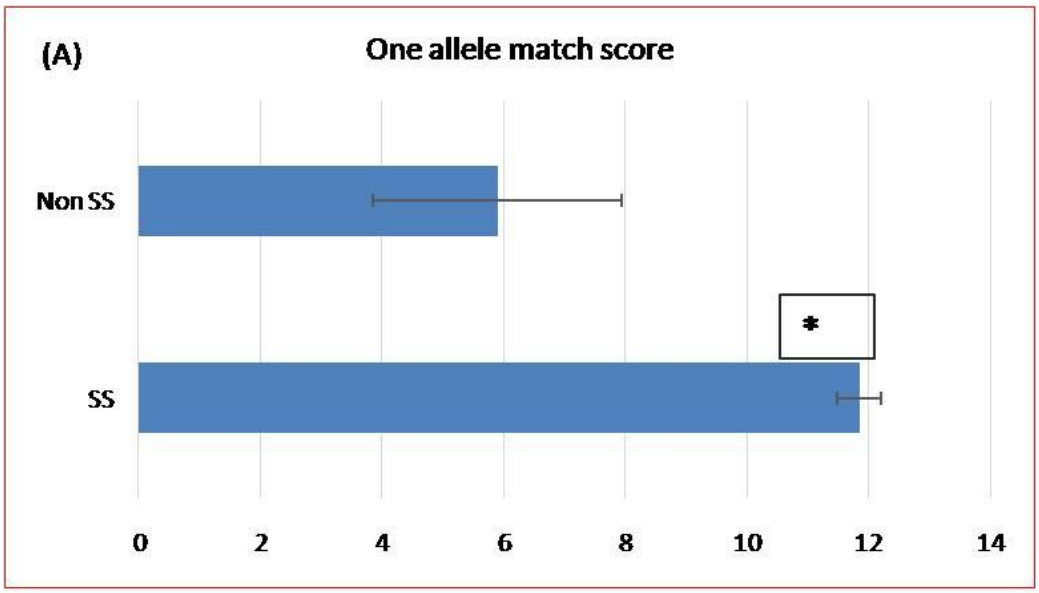


Figure 5

Independent t-Test showing a statistically significant difference between S-S (related) and non-S-S (unrelated) groups based on the TAM score were calculated. The Average TAM score for S-S is 5.45 ± 1.63 (SD), whereas, for the non-S-S group, the average TAM is 0.95 ± 0.326 (SD). Only two allele matching (TAM) scores showed a significant difference between the two (related and unrelated) groups. So only the TAM score was evaluated for its efficiency as a biomarker for sister-sister kinship analysis by autosomal STR markers.



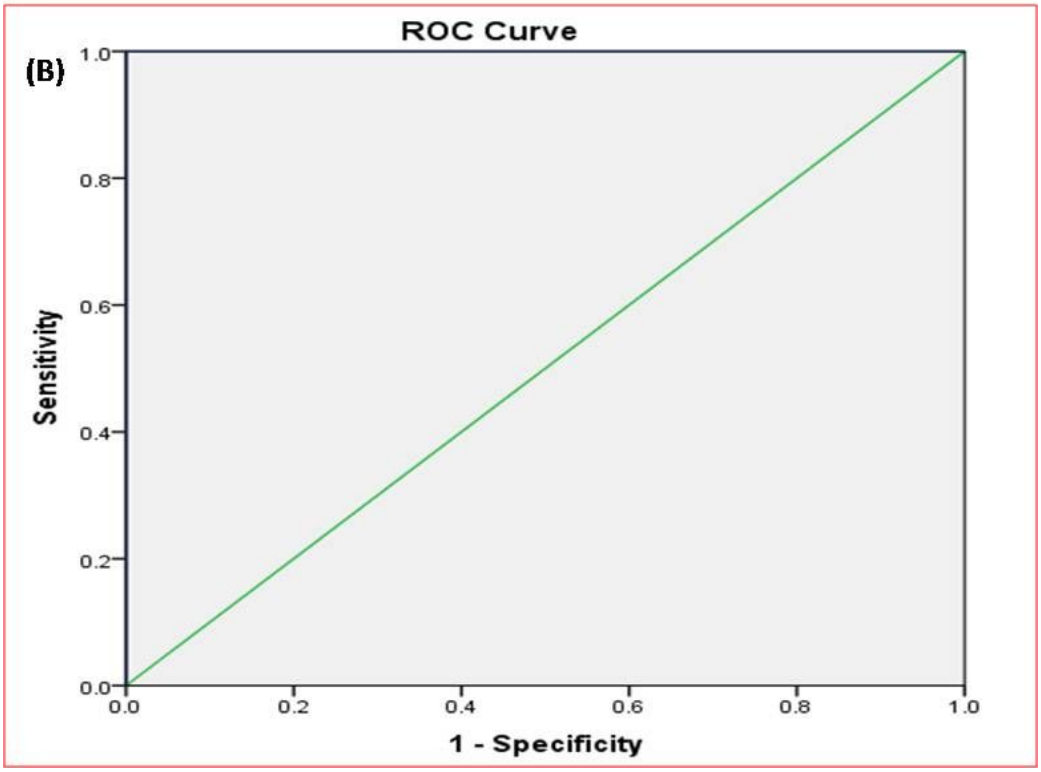
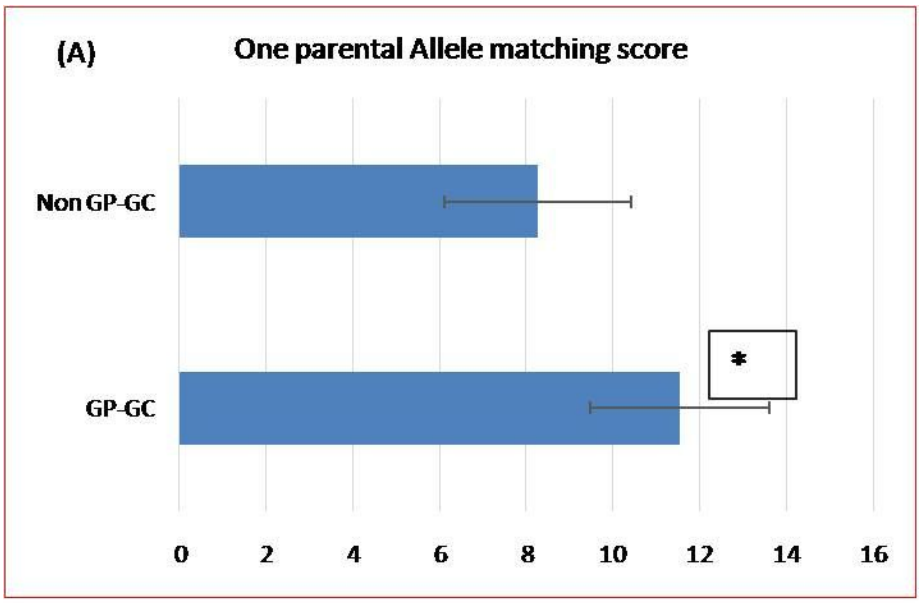


Figure 6

Independent t-Test showing a statistically significant difference between S-S (related) and non-S-S (unrelated) groups based on the OAM score were analyzed. The Average OAM score for S-S is 11.85 ± 0.366 (SD), whereas, for the non-S-S group, the average OAM is 5.90 ± 2.049 (SD). ROC curve and AUC to assess the accuracy of the test in discriminating the true cases (S-S cases) from the false cases (non-S-S cases).



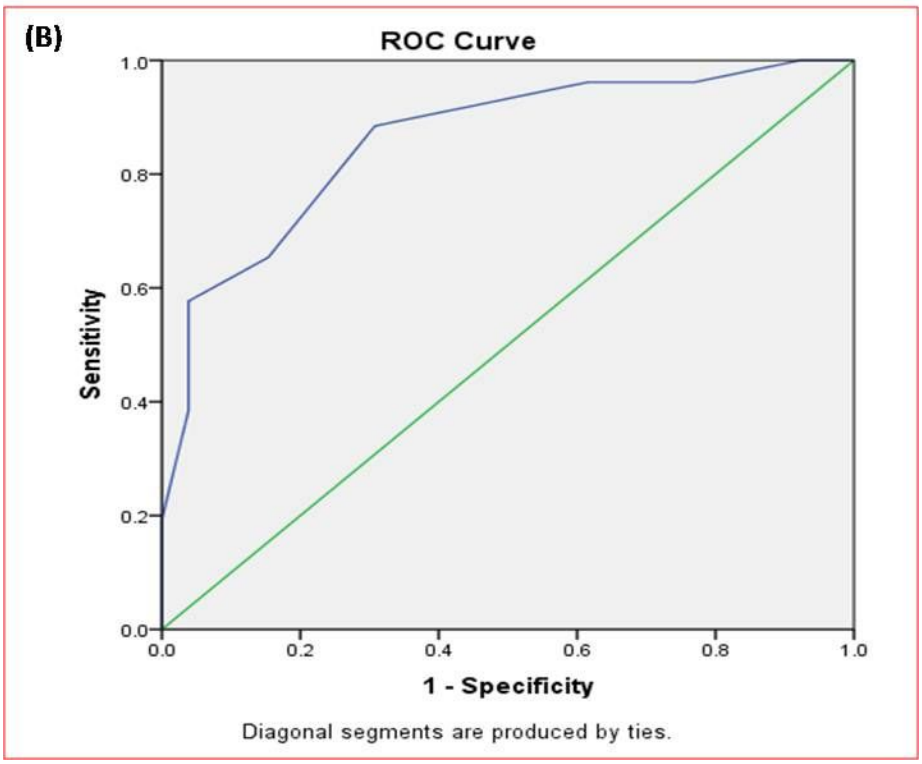
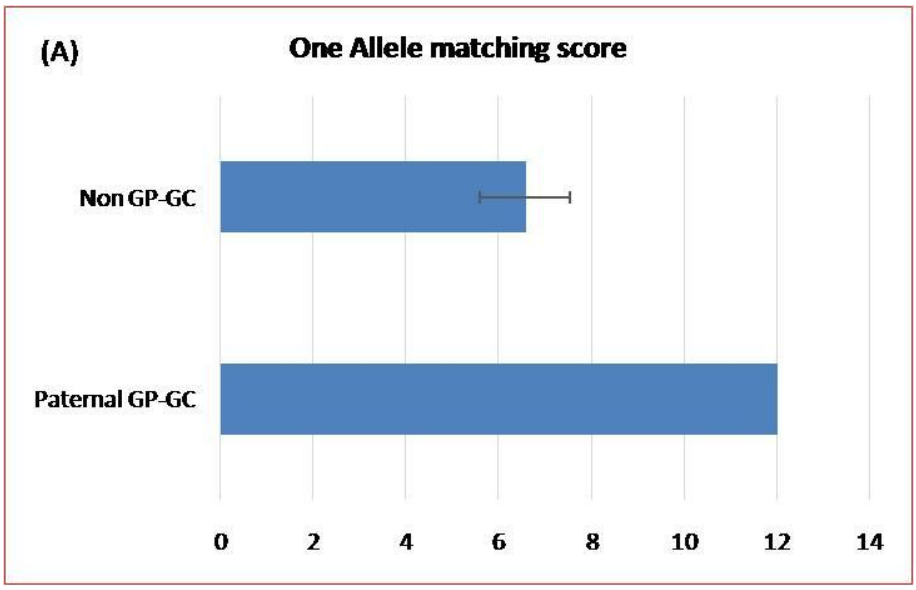


Figure 7

Independent t-Test showing a statistically significant difference (Statistically significant difference ($p<0.05$) between GP-GC (related) and non-GP-GC (unrelated) group based on the OAM score were analyzed. The Average OAM score for GP-GC is 11.54 ± 2.64 (SD), whereas, for the non-GP-GC group, the average OAM is 8.27 ± 2.146 (SD). The ROC curve and AUC to assess the accuracy of the test in discriminating the true cases (Grand parentage cases) from the false cases (non-Grand parentage cases)



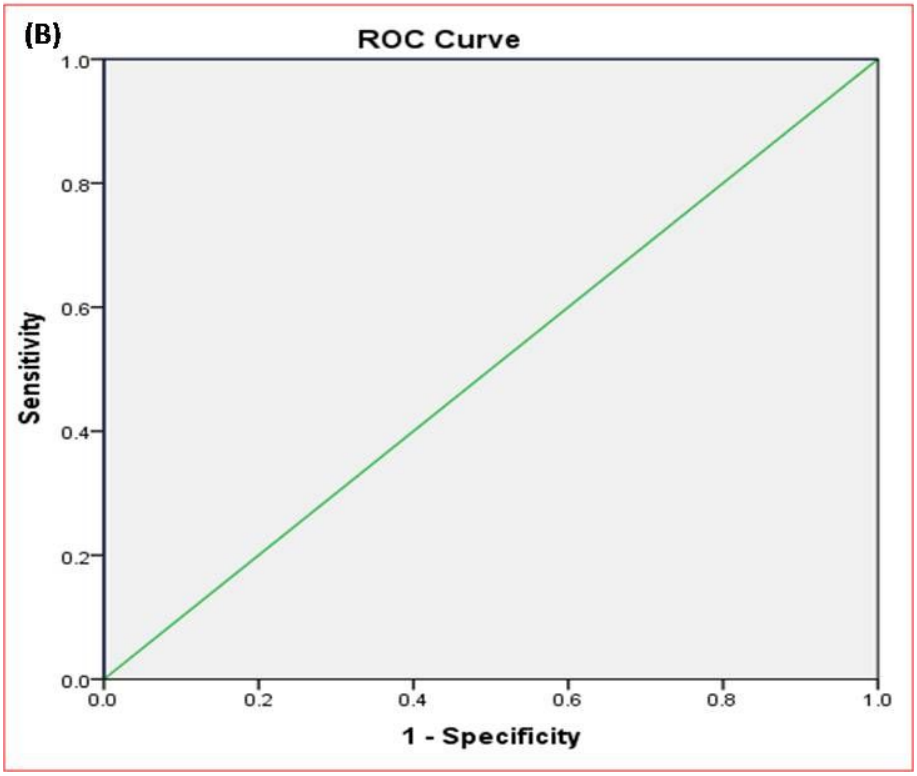


Figure 8

Independent t-Test showing a statistically significant difference between paternal GP-GC (related) and non-GP-GC (unrelated) group based on the OAM score were examined. The average OAM score for GP-GC is 12 ± 0.00 (SD), whereas, for the non-GP-GC group, the average OAM is 6.57 ± 0.976 (SD) . ROC curve and AUC to assess the accuracy of the test in discriminating the true cases (Grand parentage cases) from the false cases (non-Grand parentage cases)

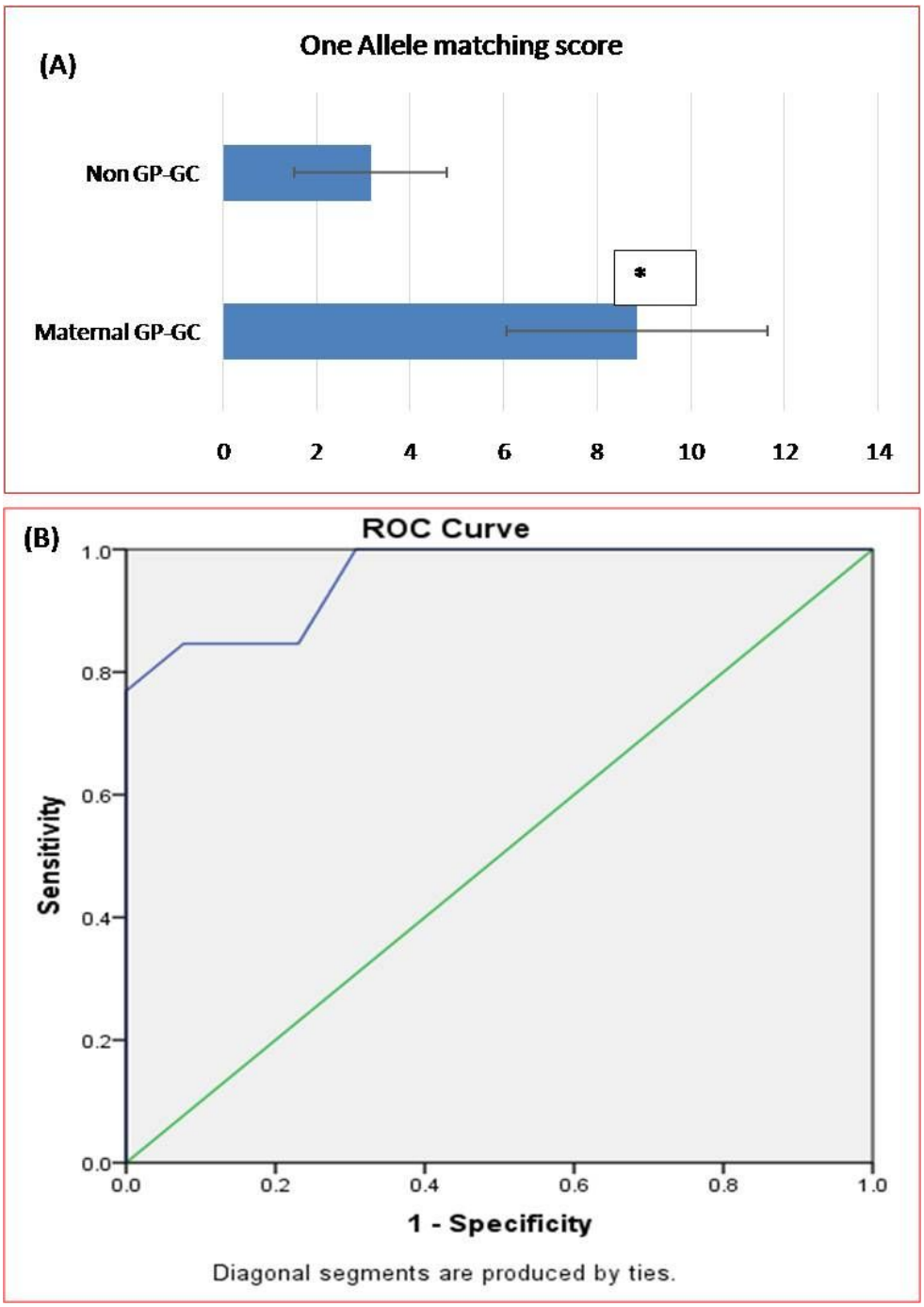


Figure 9

Independent t-Test showing a statistically significant difference between maternal GP-GC (related) and non-GP-GC (unrelated) group based on the OAM score were analyzed. The average OAM score for GP-GC is 8.85 ± 2.794 (SD), whereas, for the non-GP-GC group, the average OAM is 3.15 ± 1.625 (SD)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementrymaterialTableS1S8edited.docx