International Journal of Advanced Research in Biological Sciences
ISSN: 2348-8069
www.ijarbs.com
DOI: 10.22192/ijarbs Coden: IJARQG (USA) Volume 7, Issue 8 -2020

Research Article

DOI: http://dx.doi.org/10.22192/ijarbs.2020.07.08.011

Using different dataset and different algorithms for genetic diversity analysis in sunflower broomrape

Maria Duca, Rodica Martea
Center of Functional Genetics, "Dimitrie Cantemir" State University, Chisinau, Republic of Moldova
Corresponding author: MARTEA Rodica;
E-mail: *rodica.martea@gmail.com*

Abstract

In this study, are integrated different molecular techniques and different statistical tools in order to obtain new knowledge on genetic structure and level of variation between *Orobanche cumana* Wallr. populations, which parasitizes the sunflower crops. We used two DNA-based fingerprinting techniques (ISSR and SSR), some clustering (AHC), and multivariate (PCoA) methods for describing genetic structure of thirty-nine broomrape populations collected in the Republic of Moldova. The results indicated that ISSR markers were highly polymorphic with an average of 9.40 alleles per primer. The SSR primers were less polymorphic with nearly 4.17 per primer. The polymorphism information content (PIC) have different values (0.40 for ISSR and 0.57 for SSR), underlying differences of genetic structure at the molecular level. Comparative analysis highlighted that grouping of populations within groups was not similar in ISSR and SSR, whereas ISSR and combined data are convergent, demonstrating the tendency of the geographical spread of broomrape populations and races differentiation. In conclusion, using different dataset and different algorithms for genetic diversity analysis in sunflower broomrape, are revealed that ISSR to be the better markers, as they exhibited a remarkably strong association with geographic regions of *O. cumana* populations and a weak association with racial diversity.

Keywords: Agglomerative Hierarchical Clustering (AHC), Broomrape (*Orobanche cumana* Wallr), Genetic Diversity, Inter Simple Sequence Repeats (ISSR), Principal Coordinates Analysis (PCoA), Simple Sequence Repeats (SSR)

Introduction

There are a number of studies using molecular markers techniques that have been developed and used worldwide in different systems, due to their stability, cost-effectiveness and ease of use (Sharma et al., 2012). However, only a handful of these techniques, namely RFLPs, RAPDs, AFLPs, ISSRs, SSRs and SNPs have received global acceptance. These methods are most commonly used in the estimation of genetic variability, playing an increasingly important role in genetic diversity studies in plant species. Of these, Inter Simple Sequence Repeat (ISSR) and Simple Sequence Repeats (SSR) are very important to explain genetic diversity and population structures. ISSR markers are considered as more informative. Allelic polymorphisms occur whenever the repeated sequence is missing or insertions/deletions modify the distance between repeats (Zeinab and Mohammad, 2019.) SSR or microsatellites are highly reproducible thanks to specific amplifications of short repeated sequences, which produce polymorphic patterns depending on the number of repeated units (Paiva et al., 2014). SSRs, due to their high variability and statistical power

(Brown-Guedira et al., 2000; Agrama and Tuinstra, 2003) as well as ISSR (Bornet and Branchard, 2001; Ramasetty et al. 2016; Mei et al., 2017) have been applied successfully in genetic diversity studies of populations.

In recent years, many researchers compare the utility of SSR and RAPD (Brown-Guedira et al., 2000; Agrama and Tuinstra, 2003; Mei et al., 2017). Additionally, the analysis of the effectiveness of ISSRs and RAPDs (Ramasetty et al. 2016), SSRs and RFLPs (Ramasetty et al. 2016; Smith et al., 1997), ISSRs and SSRs (Hamza et al., 2013), was done. Comparison of three molecular markers system - RAPD, ISSR and SSR (Manimekalai et al., 2006), and four molecular markers - RFLP, RAPD, SSR and AFLPs (Pejic et al., 1998), demonstrated their specificity and the utility of each.

SSR are more reproducible than Random Amplified Polymorphic DNA (Williams et al., 1990), and less expensive to use than Amplified Fragment Length Polymorphism AFLP in identification aspects (Vos et al. 1995).

Hamza H. et al (2013) demonstrated in palms that SSR markers differentiate subpopulations according to the characteristics of fruit in a more convincing way than ISSR markers (Hamza et al., 2013). Also, in accordance with polymorphism information content, which is considered as one of the most important indicators of effectiveness (Manimekalai et al., 2006; Ismail, 2019). SSRs have higher values then AFLPs. However, AFLPs were the most efficient marker system because of their capacity to reveal several bands in a single amplification (Pejic et al., 1998).
Comparative genetic diversity studies using RFLP and RAPD markers (in genus Lathyrus) indicate that RAPDs are equivalent to RFLPs in the estimation of genetic diversity. Moreover, because of their relative simplicity and lower cost, RAPDs are considered more practical than RFLPs for studies on germplasm organization and characterization (Chtourou-Ghorbel et al., 2001).

Multivariate analytical techniques, which simultaneously analyze multiple measurements on each individual had been developed for the analysis of the structure of natural populations through molecular data (Prabodh et al. 2014), Among these algorithms, cluster analysis (Odong et al., 2011), principal component analysis (Jolliffe, 2005; Yang W. et al., 2020), and principal coordinate analysis (Borg and Groenen, 2006), are most commonly employed and appear particularly useful.

The aim of the present study was to investigate the discrimination power between *O. cumana* populations, using ISSR and SSR methods (i), to evaluate their comparative genetic diversity and detect different levels of variability within broomrape populations (ii) and to ascertain whether genetic relationships derived from SSR markers confirm or conflict with ISSR (iii).

Materials and Methods

Plant Material

Thirty-nine broomrape populations (*Orobanche cumana* Wallr.) were collected from different regions (North, South, Center) of the Republic of Moldova. Identification of races based on their morpho-physiological characteristics was performed and 14 populations belonging to ≤ E race, 6 – F race, 10 – G race, and 9 – H race were analyzed (Table 1). Broomrape aerial shoots from each population (approximately 50 mg), where stored in liquid nitrogen at -80°C for DNA isolation.

Table 1. Racial distribution of O. cumana populations in the Republic of Moldova

Broomrape race				Region
≤ E	F	G	H	
Donduseni – NEI	**Prepelita – NF2**	**Soroca – NG3**	**Balti – NH4**	**North 4**
Bacioi – CE1 Brinzenii Noi – CE2 Buteni – CE3 Cazanesti – CE4 Chisinau – CE5 Floreni – CE6 Fundul Galbenei – CE7 Holercani – CE8 Izbiste – CE9 Singera – CE10	Rassvet – CF11	Costuleni – CG12 Frasinesti – CG13 Verejeni – CG14	Ciocilteni – CH15 Sarata Mereseni – CH16	Center 16
Cazangic – SE1 Cimislia – SE2 Crihana Veche – SE3	Carabetovca – SF4 Grigorievca – SF5 Slobozia-Mare – SF6 Stefan-Vodă – SF7	Besalma – SG8 Chirsova – SG9 Ceadar Lunga – SG10 Gura Galbenei – SG11 Manta – SG12 Talmaza – SG13	Alexanderfield – SH14 Congaz – SH15 Corteni – SH16 Ermoclia – SH17 Svetlîi – SH18 Taraclia – SH19	South 19
14	6	10	9	39

DNA Extraction

Genomic DNA was isolated from broomrape tissues (bulk samples) using a GeneJET Plant Genomic DNA Purification Mini Kit (Thermo Scientific) following methodology described by the manufacturer. The DNA integrity and quantification were assessed in 1.0% agarose gel and then stored at −80°C.

ISSR Analysis

Fourteen ISSR markers, described by Benharrat H. (2002) were used for genotyping assays (Table 2). The reaction mixture had a final volume of 25 µl/tube: 30 ng of DNA; 265.6 µM dNTP; 7.5 pmol primer (0.4 µM); 1.3 U DNA Polymerase. The amplification reactions were performed individually using GeneAmp® PCR system 9700 thermocycler (Applied Biosystems), with the cycling conditions: 94 °C-7 min; 35 cycles of 94°C-30 s, 45°C-45 s, 72°C-2 min; 72°C-2 min.

SSR Analysis

Fifteen highly polymorphic SSR primer pairs, described by Pineda-Martos R. (2013) were used for genotyping assays (Table 2). PCR was realized with the following concentrations of components in the reaction mixture: 200µM dNTP, 2.0mM MgCl2, 1.0 unit of DreamTaq Green DNA Polymerase (Thermo Scientific), 0.4µM of each primer and 50 ng of extracted DNA.

Reaction volume was 15 µl. PCR was performed in Veriti Termocycler (Applied Biosystems) with the cycling conditions: 95 °C-3 min; 35 cycles of 95°C-30 s, 57°C-45 s, 72°C-1 min; 72°C-5 min. The reaction products were evaluated for polymorphisms in 1.5% agarose gel (ISSR) and in 8% polyacrylamide non-denaturing gel (SSR). Gels were stained with 1 µg mL-1 ethidium bromide for 30 to 60 min. and visualized by a UV-transilluminator.

Data Analysis

The bands were detected using the FotoCapt program and then were sized. Only the fragments with a high luminous intensity were considered in the statistical analysis. The number of bands produced for each primer were scored for their presence (1) or absence (0), and a binary matrix was generated into Microsoft Office Excel 2003 (file.xls).

Polymorphism Level and Polymorphism Information Content (PIC)

PIC for each set was determined as described in Smith et al. (1997). The ISSR and SSR binary data set was used for estimated population genetic similarity values and genetic distances by DendroUPGMA tool accompanied with the Jaccard similarity coefficient for binary data (Schoenberg, 1935).

Agglomerative Hierarchical Clustering (AHC)

AHC for broomrape populations was done by using Jaccard coefficient (Schoenberg, 1935) as agglomeration method by XLSTAT, version 2018.1.

Principal Coordinates Analysis (PCoA)

PCoA was calculated on pairwise genetic distance for both the ISSR and SSR, and for the combined molecular data, XLSTAT was employe, version 2018.1, and Jaccard coefficient (Schoenberg, 1935).

Table 2. List of primers used in a genetic diversity study of O. cumana

Nr.	Primer name	Primer sequence (5′-3′)	pb
ISSR primers			
1.	807	AGAGAGAGAGAGAGAGT	17
2.	810	GAGAGAGAGAGAGAGAT	17
3.	835	AGAGAGAGAGAGAGAGYC	18
4.	841	GAGAGAGAGAGAGAGAYC	18
5.	857	ACACACACACACACACYG	18
6.	(CAA)5	CAACAACAACAACAA	15
7.	(GACA)4	GACAGACAGACAGACA	16
8.	(GATA)4	GATAGATAGATAGATA	16
9.	(CA)6RG	CACACACACACARG	14
10.	(CTC)4RC	CTCCTCCTCCTCRC	14
11.	(CAG)5	CAGCAGCAGCAGCAG	15
12.	(CT)8TC	CTCTCTCTCTCTCTCTTC	18
13.	(CA)6AC	CACACACACACAAC	14
14.	(AG)8YA	AGAGAGAGAGAGAGAGYA	18
SSR primers			
1.	Ocum-52F	CATGTCTAAGCTTTTGGCTCG	21
	Ocum-52R	CAAGACTTGGAACAAGCAAATC	22
2.	Ocum-59F	TCTTGATTTGTATATGTCTGATGCAAT	27
	Ocum-59R	ATGCTACAATAGAAATACACAACGAAC	27
3.	Ocum-70F	AAGCTGTAAACAATGCCTGAA	21
	Ocum-70R	CCTCCTCCAGTACCACTAGGC	21
4.	Ocum-74F	CCTAAAATTGAAACCTTAAGGAAA	24
	Ocum-74R	ACTTTCCGTGAGACGGAGTC	20
5.	Ocum-75F	TGTGGATAGAGTATAAGCTACCAGTTC	27
	Ocum-75R	TTCCCGTAGCTTGGAGAATG	20
6.	Ocum-81F	TTACAAGGTGAAACCACCCA	20
	Ocum-81R	CAGCTACTGTCCGCAAGAAA	20
7.	Ocum-87F	TTCTCGACAGCTTTGGGTAAA	21
	Ocum-87R	ATGCCAACTTCGAGTGATCC	20
8.	Ocum-108F	TCGTTAATAAGTGGTTCACGAAAA	24
	Ocum-108R	TGACTAAAAATAAAATGTACGGGTG	25
9.	Ocum-122F	GGAATACATCATTAAAGTAGTTGTCCG	27
	Ocum-122R	GAAGGAGTCATTAAACTCCGTGA	23
10.	Ocum-141F	CAGCAACTGTTTCTTCCATAGAG	23
	Ocum-141R	TCCAAGAAGAGGAAAAGAAGTGA	23
11.	Ocum-160F	TGAGGGTTTGTAAAGTGGGC	20
	Ocum-160R	CGTACCTTATCCCTCCGTCA	20
12.	Ocum-174F	CAACCAACAAACAAGTAGTGACG	23
	Ocum-174R	TCTTGCGGCAAAACCATT	18
13.	Ocum-196F	GTATGTGCGCCCGTCTTG	18
	Ocum-196R	GGGGATGACTGTGTTTCGAT	20
14.	Ocum-197F	AGAGACGGCATCATCAATCA	20
	Ocum-197R	GTGATCGTGCAGGCACCTA	19
15.	Ocum-206F	CCGATTGCTGTTTATGTTGTATT	23
	Ocum-206R	TGTAGGAGATGCCCAGTTCA	20

Results and Discussion

In this study, we integrated different molecular techniques and different statistical tools in order to obtain new knowledge on genetic structure and level of variation between broomrape populations. The comparison of the two types of molecular markers for their efficiency, effectiveness and informativeness in populations analysis has been carried out in *Orobanche cumana* Wallr.

The 39 broomrape populations, evaluated in this study were differentiated uniquely using ISSR and SSR molecular markers as tools for assessing genetic variation, races identification and determining the relationships among different populations from a wide range of geographical origins. The ISSR and SSR analysis conducted for *O. cumana* populations revealed the presence of an appreciable level of genetic diversity. To assess the differences between aforementioned marker indicators, including the percentage of polymorphic bands, polymorphism, PIC index, genetic distance, and similarity index were calculated.

Polymorphism (P,%)

The general level of genetic polymorphism obtained with the two type of markers where different especially, according to its PCR molecular mechanisms, underlying population specificity and thus - their utility in the analysis of genetic relationships.

Thus, all the 14 *ISSR primers*, as markers, targeted multiple genomic loci to amplify mainly the inter SSR sequences of different sizes (Zietkiewicz et al., 1994) generated 132 clear and reproducible bands (mean = 9.4/primer), 74.7% of which were polymorphic (Table 3).

Electrophoretic analysis revealed that primer 807 generated the largest number of loci (20), being the most informative, whereas primer (CA)6AC - produced the minimum number (5). Different primers showed a different level of polymorphism (P). Thus, polymorphism had values between 100%, in case of primers 807, $(GACA)_{4,}$ $(CTC)_4RC$, $(CAG)_{5,}$ $(CT)_8TC$ and 22% for primer 810, values, which according to the method of data analysis (Calderini et al., 1999), can be considered statistically significant.

The studied *SSR primers*, which detects the allelic variation by way of repeat numbers with in a locus and one pair of SSR primer deals with one locus (Weising et al., 2005) showed different level of polymorphism and produced a total number of 50 alleles, with the average 4.17 (Table 3).

Each SSR primer pair produced bands with high and low frequency. The number of detected alleles per locus in various accessions ranged from two (Ocum-141) to eight (Ocum-81).

Thus, in our research, SSR markers revealed a higher number of alleles per locus, whereas in other similar studies evaluating populations from Spain (Pineda-Martos et al., 2013) or Russia, Kazakhstan and Romania (Guchetl et al., 2014). The SSR primer was considered to be polymorphic when the most abundant allele in the population has frequency lower than 95.4% (Agrama et al., 2003). So, among the two marker types, the percent polymorphisms of ISSR markers scored higher (74.7%) than SSR (40%) (Table 3).

Thus, ISSR is a better marker to evaluate the genetic diversity in *O. cumana* populations. The higher level of polymorphism associated with ISSR markers may be a function of the unique replication slippage mechanism responsible for generating SSR allelic diversity (Pejic et al.,1998).

Polymorphism Information Content (PIC)

The PIC provides an estimate of the discriminating power of markers. ISSR are bi-allelic markers and PIC can have maximum value of 0.50.

SSR are multi-allelic markers and their PIC values will be ranged from 0, if it is monomorphic, to 1, if it is highly discriminative. The comparative analysis of PIC index revealed the values ranging from 0.27 to 0.77 with a mean value 0.57 and 0.23 to 0.48 with mean value 0.40, for SSR and ISSR markers, respectively (Table 3).

Table 3. Comparison of ISSR and SSR markers in evaluating genetic diversity of O. cumana populations

MOLECULAR MARKERS	TOTAL	
	ISSR	SSR
No of populations	39	39
No of primers	14	15
No of primers excluded from this study	2	3
Polymorphism level		
Total number of bands	132	50
Polymorphic bands	132	50
No of detected alleles per locus	5 – 20	2 – 8
Polymorphism, %	22 – 100	25 - 45
Polymorphism (%), average	74.7	40
Bands average	9.40	4.17
Polymorphism Information Content		
PIC	0.23 – 0.46	0.27 – 0.77
PIC, average	0.4	0.57
Average Marker Index	0.49	0.45
Correlation PIC / no of alleles per locus	0.84	0.46
Genetic similarity		
Genetic distance range	0.16 – 0.74	0.17 – 0.81
Genetic distance, average	0.45	0.49
Similarity indices	0.26 – 0.84	0.19 – 0.83
Similarity indices, average	0.55	0.51

Overall, SSR markers were more polymorphic (0.57 PIC, average) than ISSR (0.4 PIC, average) however, the number of polymorphic bands per assay unit was higher in ISSR (9.4) as compared to SSR (4.17).

A similar edge of ISSR over SSR in terms of discriminative capability for a given set of genotypes has been observed in certain other studies (Sethi et al., 2016).

The analysis of medium correlation between PIC value and the number of alleles per locus (Spearman`s rank correlation coefficient) of ISSR markers scored higher (0.84) than SSR (0.46) (Table 3).

Genetic Similarity (GS)

GS for *O. cumana* populations show an average of 0.55 (ISSR) and 0.51 (SSR). GS was obtained in the range of 0.26-0.84 for ISSR markers, 0.19-0.83 for SSR markers (Table 3) Similar observation was also made by Sethi K. (2016) to cotton. The results showed that the two methods were effective for genetic diversity investigation of broomrape. By comparing the two marker systems, ISSR where more efficient based on PIC value, percentage of polymorphism and better reproducibility. However, for race identification purposes SSRs are better tools, since they provide a higher level of polymorphism and reproducibility.

Agglomerative Hierarchical Clustering (AHC)

The AHC method of the two molecular marker systems was used for geographic and racial origin discrimination. Thus, the results obtained in the case of the *ISSR molecular technique*, reveal that 35 of the populations were grouped in two clusters - C_1 and C_2 and 4 populations were found secluded, demonstrating a specific genetic structure. These populations involve three accessions from the South (*SG8, SH14, SH15*) and one – from the North (*NG3*), representing two more evolved and more aggressive races - race G and race H (Figure 1, A).

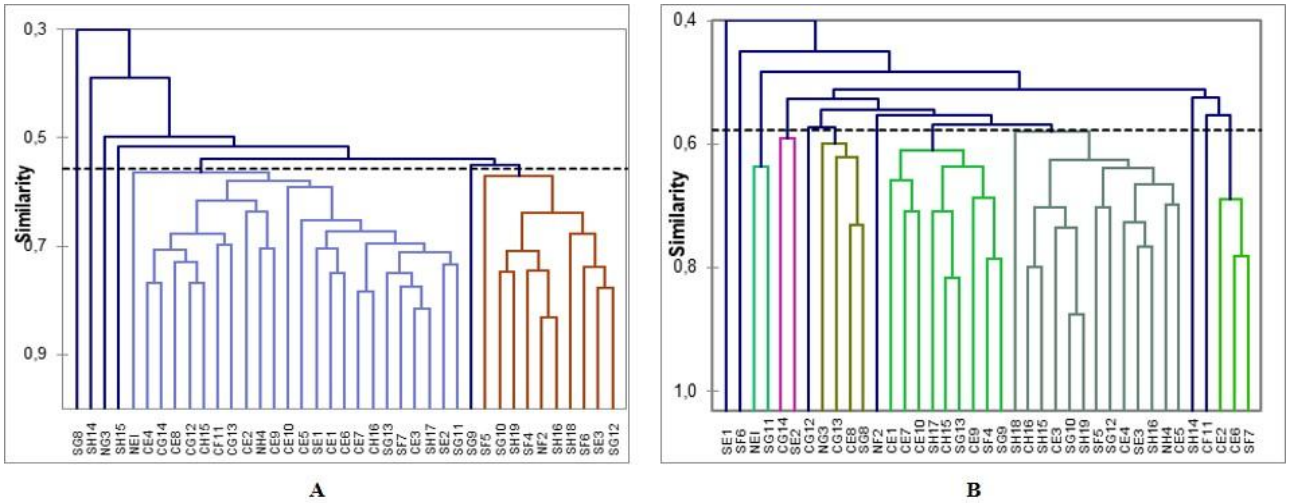

Figure 1. Hierarchical cluster analysis of ISSR data (A) and SSR data (B).

Whereas, the results obtained based on the *SSR technique* reveal a less structured association. Thus, 35 of the populations were grouped in two clusters (C^1 and C^2), of which C^2, shows a distribution of populations in two other groups C^2A and C^2B, the latter being in turn differentiated into 2 groups C^2B^1 and C^2B^2 (Figure 1, B).

At the same time, three populations from the southern part of the country (*SE1, SF6 and SG11)* as well as Donduseni (*NE1)* representing race E, F and G, reveal maximum genetic distance (0.70 - 0.74) from all analyzed populations, which explains their positioning outside the formed clusters (Figure 1, B). Thus, each molecular technique removed four, but each different, solitary population from the clusters, which were not found in the formed groups. Thus, according to geographical origin, the obtained data reflect a higher capacity for clustering and grouping of populations in the case of ISSR markers compared to SSR.

The clustering in the frame of ISSR amplification products, groups all 16 populations from the Center, regardless of race, in the largest numerically cluster - C_1, which includes 23 populations.

Nine populations from the Center are found in C_1A, and another 7 populations in C_1B sub-clusters. C_1A is more uniform (9 populations from the Center and 2 - from the North), while the C_1B sub-cluster includes almost equal populations from the Center (6) and the South (6) of the country, representing a direct transition to the C_2 cluster. This cluster (C_2) includes 10 populations, out of the 19 collected from the South of the Republic of Moldova and is practically uniform according to the origin.

At the same time, SSR data reveals a less structured association according to the collected site, populations from one region being distributed in all clusters and sub-clusters. Thus, 13 populations from the Center of the country are grouped in cluster one (C^1), including four - below cluster C^1A and 9 - in C^1B and three populations from the Center are positioned in - cluster C^2. The same type of grouping was revealed for the populations collected in the South of the country. It is important to mentioned that the broomrape populations were distributed in all clusters, including C^1 - 14 (C^1A – 2 and C^1B - 12) and C^2 – 2.

In an effort to realize an analysis of the hierarchy distribution of broomrape populations in accordance with the races, it was observed that in the case of *ISSR markers* only 13 of the 14 populations belonging to E race were located in cluster C_1 (5 - C_1A and 8 - C_1B), while representatives of the F, G and H races where distributed in each of the formed clusters (C_1, C_2) and sub-clusters.

In the case of SSR markers, a more specific association with the racial structure of the populations was highlighted. Thus, the representatives of the most aggressive H-race (8 out of 9) practically grouped in the C^1B sub-cluster, and the populations belonging to the G-race (9 out of 10) - in the C^1 cluster.

Therefore, our results are in accordance with Ramasetty (2016), Hamza (2013), Zhang (2011) and revealed that ISSR and SSR markers provide similar, but not identical information and may differ for their level of informativeness.

Thus, there are four populations from the southern part of the country (*SE3, SG12, SH18, SF5*), which were grouped together in the same cluster regardless of the applied method (ISSR in C_2; SSR in C^1B^2). Six populations collected from the Center are similarly grouped, including three (*CG14, CE8, CG13*) in C_1A and C^1A, other three (*CE1, CE10, SG13*) in C_1B and C^1B. In eight cases, two by two populations were grouped together in common clusters according to both used methods. This grouping pattern is in accordance with the specificity of genetic structure, since the same populations were grouped in the same clusters based on ISSR and SSR data.

Generalizing the data obtained in this study, it is important to mention that although grouping of populations within groups was not similar in ISSR and SSR derived hierarchical clustering, whereas the pattern of a grouping of the populations remained more or less the same. So, in 61.5% of cases the populations were grouped together in common clusters due to a smaller genetic distance between them, 0.28 – 0.42 for ISSR and 0.33 – 0.58 for SSR, demonstrating a similar genetic structure and a high capacity to discriminate between the two methods used. Even though AH clustering analysis grouped the accessions into different groups, according to ISSR and SSR data sets, there are some accessions that combine similar genetic backgrounds.

Principal Coordinate Analysis (PCoA)

In order to support AHC and determine the consistency of differentiation among populations defined by the cluster analysis, as well as to understand the population structure of sunflower broomrape populations in a two-dimensional graph, PCoA was performed (Figure 2 and Figure 3).

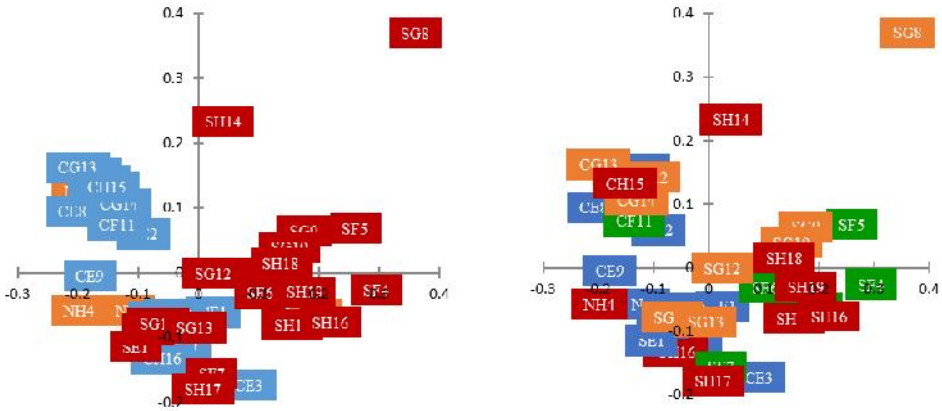

Figure 2. Principal Coordinates Analysis ordination of broomrape populations based on ISSR data by geographic origin (A) and races (B).

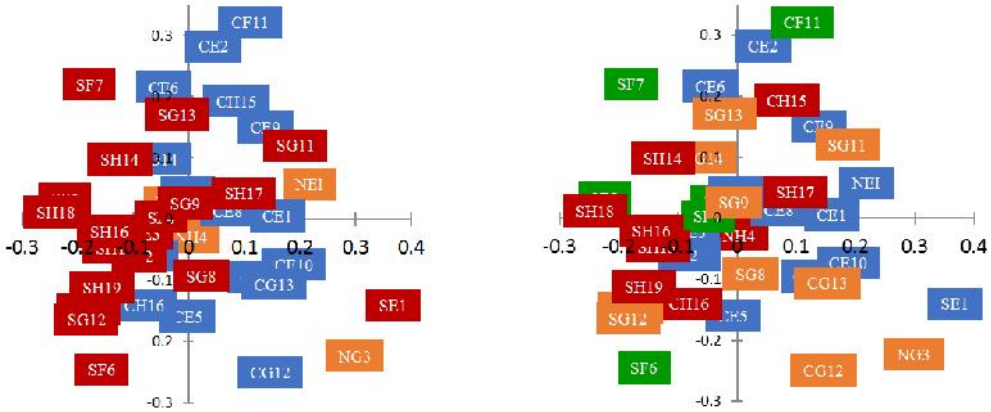

Figure 3. Principal Coordinates Analysis ordination broomrape populations based on ISSR data by geographic origin (A) and races (B).

This method has been widely used for image representation and to extract relevant information from dimensional data sets (Odong et al., 2011; Borg and Groenen, 2006).

The PCoA ordination of molecular results from this study produced similar groups, confirming the results of the aforementioned AHC (Figure 1) and the UPGMA for *ISSR* (Duca et al., 2020) and *SSR* (Duca et al., 2017) data. The PCoA of combined ISSR and SSR data was employed to examine further the genetic relationships among populations. All at once, in this simultaneous analysis of genetic diversity using both datasets provides more information and clearer discrimination of populations.

Relevant association of populations with origin of collecting site were detected, similar to hierarchical analysis provided by *ISSR* (Figure 1, A) demonstrating the higher share of ISSR data in the structural organization of populations according to the genetic structure.

Thus the 39 accessions where separated into three more or less numerically equal groups (Figure 4, A). The majority (88%) populations collected from central part of the country are located on the left side of the graph, repeating the accessions included in C_1A (Figure 1, A).

The majority (85%) of populations from the South are located on the right side of the graph 100% repeating the cluster C_2 (Figure 1, A). a more amorphous separated group, located at the bottom of the horizontal axis, was formed by the populations clustered by AHC in C_1B (53%).

Therefore, the grouping pattern observed in the PCoA is in accordance with the collected site (Figure 4, A) similar to AHC based on ISSR data. However, as solitary populations, where separated *SH14* and *SG8* as in the ISSR clustering and *SE1* and *SG 11* as in the SSR clustering.

Principal Coordinate Analysis (Combined Data)

The PCoA of combined molecular data, performed in accordance with races, showed a close position of the populations belonging to the same race (Figure 4, B). Thereby, 11 of the 14 populations representing the E race form a common group. It reveals the proximity of the four (CE8, CE4, CE2, CE5) and five (CE6, CE1, CE7, CE9, CE3) populations, which form two subgroups, all of which are collected from localities situated at a very small geographical distance (around 10-15 km). Separately there are only three populations representing the race E - NE1, SE1, SE3, the first two being located solitary in the case of hierarchical analysis of SSR data (Figure 1, B).

Four of the six populations representing race F, collected from the southern part (SF4, SF5, SF6, SF7) and one population from the northern part (NF2) were grouped in a common group (Figure 4, B).

Populations representing race G (10) were grouped into two separate groups. Thus, the three populations from the Center (CG12, CG13, CG14), as well as the population from Soroca (NG3), were positioned in a single group.
At the same time, the rest of the populations from the South (SG9, SG10, SG11, SG12, SG13) formed the second group (Figure 4, B).

Six of the nine populations representing the H race form a common group. It shows the proximity between the populations from Congaz, Corteni, Ermoclia, Svetlîi, Taraclia and Sarata Mereseni (Figure 4, B).

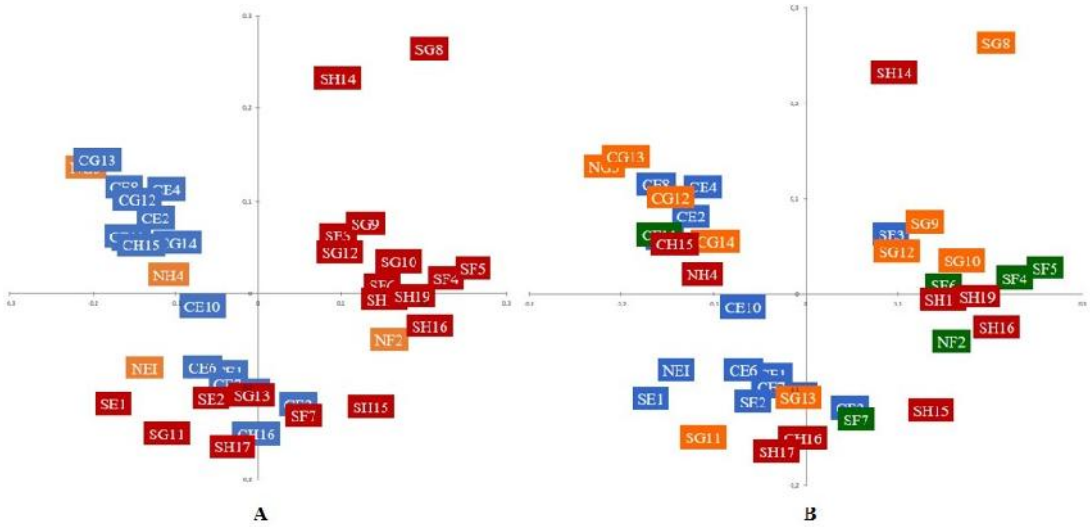

Figure 4. Principal Coordinates Analysis based on combined ISSR and SSR data according to geographic origin (A) *(orange-North, red -South, blue-Centre)* and to races (B) *(blue-E race; orange-G race; red -H race; green-F race).*

Three populations (SH14, CH15 and NH4) are found separately, the last two - similar to the results obtained for ISSR.

Both, AHC and PCoA, supported the grouping of all 39 populations into three groups. Therefore, comparative analysis of genetic structure of races and geographic origin, using ISSR, SSR and combined data. As well as two methods of grouping showed clear geographic trends among natural populations, and has demonstrated that different populations from one race are classified in the same cluster, indicating low genetic distance and high genetic similarity, only if they belonged to the same region.

Our results indicated the genetic diversity of broomrape populations were appreciable both at the population and race level. Low level race differentiation in both techniques was found, which may be attributed to limited geographic distance and abundant gene flow. The result for broomrape races supports the idea that the species with small geographic ranges tend to maintain less genetic diversity than geographically widespread species (Liu et al., 2013).
In summary, our results indicated that genetic diversity is due to race specificity and gene flow between geographically close populations and environmental conditions. These results highlight the importance of various climatic conditions on race and population specificity.

Conclusions

A specific or a combination of molecular (ISSR and SSR) methods are important for the establishment of genetic variation and relationships among *O. cumana* populations. Taken together, the results obtained in this study provide a good molecular basis for understanding genetic diversity.

This study shows the importance of exploring different methods of analysis to confirm the veracity of the results. The difference in the grouping pattern of two data sets and two methods of clustering may be attributed to the different principles adopted in each technique and to the difference in statistical proprieties between them.

The use of multivariate statistical algorithms, as an important strategy for classifying germplasm, revealed that the pattern of a grouping of populations remained more or less the same in hierarchical clustering method and PCoA.

Comparative analysis of derived dendrograms highlighted, that grouping of populations within groups was not similar in ISSR and SSR, whereas ISSR and combined data of ISSR and SSR markers are convergent, demonstrating the tendency of the geographical spread of broomrape populations and races differentiation.

The comparative analysis of ISSR and SSR marker systems revealed ISSR to be the better markers, as they generated the highest percentage of polymorphic bands, polymorphism, PIC index, genetic distance, similarity index. The ISSR markers exhibited a remarkably strong association with geographic regions of broomrape populations and a weak association with racial diversity. The SSR variation showed more clear association with genetic origin and race.

Acknowledgments

This research was undertaken in the framework of the project 15.817.05.03F "Sunflower (*Helianthus annuus* L.) resistance to broomrape (*Orobanche cumana* Wallr.): genetic, molecular and physiological mechanisms", supported by the National Agency for Research and Development (Government of Republic of Moldova).

References

Agrama, H. A., Tuinstra, M. R. 2003. Phylogenetic diversity and relationships among sorghum accessions using SSRs and RAPDs. African Journal of Biotechnology. 2(10):334-340.

Benharrat, H., Veronesi, C., Theodet, C., Thalouam, P. 2002. *Orobanche* species and population discrimination using inter simple sequence repeat (ISSR). Weed Res. 42:470-474.

Borg, I., Groenen, P. J. F. 2006. Modern Multidimensional Scaling: Theory and Applications. Journal of Educational Measurement. 40(3):277-280.

Bornet, B., Branchard, M. 2001. Nonanchored inter simple sequence repeat (ISSR) markers: reproducible and specific tool for genome fingerprinting. Plant. MolBiol. Rep. 19:209-215.

Brown-Guedira, G. L. et al. 2000. Evaluation of genetic diversity of soybean introductions and north American ancestors using RAPD and SSR markers. Crop Science. 40:815-823.

Calderini, D. F., Abeledo, L. G., Savin, R., Slafer, G. A. 1999. Carpel size and temperature in preanthesis modify potential grain weight in wheat. J Agric Sci. 132:453-460.

Chtourou-Ghorbel, N., Lauga, B., Combes, D., Marrakchi, M. 2001. Comparative genetic diversity studies in the genus *Lathyrus* using RFLP and RAPD markers. Lathyrus Lathyrism Newsletter. 2: 62.68.

Duca, M., Joiţa-Păcureanu, M., Port, A., Martea, R., Boicu, A., Rîşnoveanu, L., Clapco S. 2020. Genetic diversity analysis of sunflower broomrape populations from Republic of Moldova using ISSR markers. Romanian Agricultural Research. 37:3-11. DII 2067-5720 RAR 2020-8.

Duca, M., Port, A., Boicu, A., Şestacova, T. 2017. Molecular characterization of broomrape populations from republic of Moldova using SSR markers. Helia. 40(66):47-59.

Guchetl, S. et al. 2014. Interpopulation genetic differentiation *Orobanche cumana* Wallr. from Russia, Kazakhstan and Romania using molecular genetic markers. Helia. 37(61):181-191.

Hamza, H. et al. 2013. Comparison of the effectiveness of ISSR and SSR markers in determination of date palm (Phoenix dactylifera L.) agronomic traits. AJCS. 7(6):763-769.

Ismail, N. A. 2019. Genetic diversity of torch ginger (*Etlingera elatior*) germplasm revealed by ISSR and SSR markers. Hindawi BioMed Research International. 1-14, ID 5904804.

Jolliffe, I. T. 2005. Principal component analysis. Weather. 98(10):1-13.

Liu, J., Shi, S., Chang, E., Yang, W., Jiang, Z. 2013. Genetic diversity of the critically endangered *Thuja sutchuenensis* revealed by ISSR markers and the implications for conservation. International Journal of Molecular Sciences. 14(7):14860-14871.

Manimekalai, R., Nagaarajan, P., Kumaran, P. M. 2006. Comparison of effectiveness of RAPD, ISSR and SSR Markers for analysis of coconut (*Cocos nucifera* L.) germplasm accessions. Tropical Agricultural Research. 18:1-10.

Mei, Z. et al. 2017. Genetic analysis of *Penthorum chinense* Pursh. by improved-RAPD and ISSR in China. Electron J Biotechnol. 30:6-11.

Odong, T. L., van Heerwaarden, J., Jansen, J., van Hintum, T. J. L., van Eeuwijk, F. A. 2011. Determination of genetic structure of germplasm collections: are traditional hierarchical clustering methods appropriate for molecular marker data? Theor. Appl. Genet. 123(2):195-205.

Paiva, C. L., Viana, A. P., Santos, E. A., Freitas, J. C. O. et al. 2014. Genetic variability assessment in the genus *Passiflora* by SSR marker. Chil. J. Agric. Res., 74:355-360.

Pejic, I. et al. 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theoretical and Applied Genetics. 97(8):1248-1255.

Pineda-Martos, R., Velasco, L., Fernández-Escobar, J., Fernández-Martínez, J. M., Pérez-Vich, B. 2013. Genetic diversity of *Orobanche cumana* populations from Spain assessed using SSR markers. Weed Research. 53: 279-289.

Prabodh, K. B., et al. 2014. Structure and genetic diversity of natural populations of *Morus alba* in the trans-Himalayan Ladakh Region. Biochem Genet. 52(3-4):137-152.

Ramasetty, B. T., et al. 2016. Identification and genetic diversity analysis of *Memecylon* species using ISSR, RAPD and Gene-based DNA barcoding tools. Electronic Journal of Biotechnology. 24 (2016):1-8.

Schoenberg, I. J. 1935. Remarks to Maurice Frchet's article "Sur la dfinition axiomatique d'une elasse d'espaces distancibs vectoriellement applicable sur l'espace de Hilbert. Ann Math. 38(3):724-732.

Sethi, K., Siwach, P., Verma, S. K. 2016 Simple sequence repeats (SSR) and interspersed sequence repeats (ISSR) markers for genetic diversity analysis among selected genotypes of *Gossypium arboreumrace*. Afr. J. Biotechnol. 15(1):7-19.

Sharma, R., Joshi, A., Maloo, S. R., et al. 2012. Assessment of genetic finger printing using molecular marker in plants: a review. Sci Res Impact. 1:29-36.

Smith, J. S. C., Chin, E. C. L., Shu, H. 1997. An evaluation the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparison with data from RFLP and pedigree. Theor. Appl. Genet. 95:163-173.

Vos, P., et al. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23(21):4407-4414.

Weising K., et al. 2005. *Microsatellites*. In: DNA Fingerprinting in Plants: Principles, Methods, and Applications, 2nd Edition, CRC Press Reference, Boca Raton, Florida, p. 41-45.

Williams, J. G., Kubelik, A. R., Livak, K. J. et al. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18(22):6531-6535.

Yang, W. et al. 2020. Using Principal Components Analysis and IDW Interpolation to Determine Spatial and Temporal Changes of Surface Water Quality of Xin'anjiang River in Huangshan, China". Int. J. Environ. Res. Public Health. 17(2942):1-20.

Zeinab, M., Mohammad, A. 2019. Usage of Inter Simple Sequence Repeat (ISSR) marker in assessing the genetic variation of six *Parrotfish* species from the Egyptian Red Sea. Egyptian Academic Journal of Biological Sciences. 11(3):109-116.

Zhang, P., Li, J., Li, X., Liu, X., Zhao, X., Lu, Y. 2011. Population structure and genetic diversity in a rice core collection (*Oryza sativa* L.) investigated with SSR markers". PLoS One. 6(12):1-13, e27565.

Zietkiewicz, E., Rafalski, A., Labuda, D. 1994. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics. 20:176-183.

Access this Article in Online

Website: www.ijarbs.com

Subject: Genetics

Quick Response Code

DOI:10.22192/ijarbs.2020.07.08.011

How to cite this article:
Maria Duca, Rodica Martea. (2020). Using different dataset and different algorithms for genetic diversity analysis in sunflower broomrape. Int. J. Adv. Res. Biol. Sci. 7(8): 99-110.
DOI: http://dx.doi.org/10.22192/ijarbs.2020.07.08.011