Available online: www.notulaebiologicae.ro
Print ISSN 2067-3205; Electronic 2067-3264
Not Sci Biol, 2019, 11(2):222-232. DOI: 10.15835/nsb11210470
Original Article

Genetic Polymorphism of the Wild and *In vitro* Regenerated Plants of the Medicinal Grass *Cymbopogon schoenanthus* subsp. *proximus*

Asmaa M. ABDELSALAM[1,2], Kamal CHOWDHURY[2], Ahmed A. El-BAKRY[1]*

[1]*Helwan University, Faculty of Science, Ain Helwan, Cairo 11795, Egypt; asm.abdelsalam@yahoo.com; ael_bakry@yahoo.com (*corresponding author)*
[2]*Claflin University, Biology Department, 400 Magnolia St., Orangeburg, SC 29115, USA; kchowdhury@claflin.edu*

Abstract

Cymbopogon schoenanthus subsp. *proximus* grows wild in subtropical Africa, Sudan and Egypt. The species is heavily collected for its use in folk medicine and drug production. A wild population from south Egypt was used to determine preliminary genetic polymorphism within the species, using nineteen ISSR, fourteen RAPD and seven cpSSR primers. Three regeneration systems, somatic embryogenesis (SE), direct organogenesis (D), and indirect organogenesis (ID), were established from seed explants of the same population and polymorphism within regenerated plants was determined. ISSR generated a total of 222 amplified fragments for all genotypes, while RAPDs and cpSSR yielded 139 and 34 fragments, respectively. Wild plants showed an average low polymorphism for all marker types of 45.8%. Regenerated plants polymorphism was also low (SE=44.6, D=44, ID=46.2%). ISSR and cpSSR markers were more sensitive in elucidating polymorphism (51.5 and 46.87%) than RAPD (37.85%). ISSR was the most significant marker in producing unique bands, for the wild genotypes (6), SE (7), D (5) and ID (6). cpSSR followed producing 4 for wild genotypes, 6 for SE, 3 for D and 6 for ID organogenesis. Unweighted pair group with arithmetic average (UPGMA) clustering analysis and Jaccard's similarity data suggested that wild plants and those regenerated through somatic embryogenesis and direct organogenesis are more similar. The study elucidated low polymorphism within both the wild population and regenerated genotypes, with plants regenerated through somatic embryogenesis and direct organogenesis being more similar to wild plant genotypes, suggesting their future use in studies with genetic transformation and *ex-situ* conservation of the species.

Keywords: cpSSR; ISSR; organogenesis; RAPD; somatic embryogenesis

Introduction

Cymbopogon schoenanthus subsp. *proximus* is a wild aromatic herb that grows in sandy soil under arid climate (Bolous, 1999). Wild plant extract of this subspecies contains compounds that are important for folk medicine and pharmaceutical industry (Bolous, 1983). For example, sesquiterpene proximadiol, a component in the plant oil, has been found to be responsible for repulsion of the renal calculi (Locksley *et al*., 1982; El-Askary *et al*., 2003). Other terpenoids were identified in the plant extract (El Taher and Abdel-Kader, 2008). Antimicrobial and antioxidant activities of the plant extract have also been reported (Selim, 2011). NMR metabolic profiling for the species polar extract showed the presence of additional compounds of metabolic, medicinal and economic importance as trehalose, choline and the anticancer alkaloid trigonelline (Abdelsalam *et al*., 2017a).

In vitro propagation is a valuable method because of its ability to produce true-to-type genotypes in a short time (Jagesh *et al*., 2013; Agrawal *et al*., 2014). However, in some cases, some morphological, biochemical and genetic changes could occur during plant regeneration (Nehra *et al*., 1992; Peredo *et al*., 2006; El-Dougdoug *et al*., 2007).

DNA markers are very useful to reveal genetic diversity between related genotypes because results are independent of plant age, type of tissue and environmental conditions. Different types of molecular markers such as RAPD, ISSR and SSR have been used widely for the detection of genetic variation between and within *Cymbopogon* species (Sangwan *et al*., 2001; Kumar *et al*., 2009; Adhikari *et al*., 2015).

In the present study we aimed at elucidating preliminary genetic diversity within a wild population of *Cymbopogon schoenanthus* subsp. *proximus* and within plants regenerated *in vitro* through somatic embryogenesis, direct and indirect

Received: 13 Mar 2019. Received in revised form: 15 May 2019. Accepted: 16 May 2019. Published Online: 28 Jun 2019.

organogenesis. Genetic relationship among wild genotypes and regenerated plants through different *in vitro* culture systems was studied to identify the best systems that may produce genetic fidelity and genetic diversity, that can be used in future conservation of plant germplasm and for genetic transformation.

Materials and Methods

Plant material

C. schoenanthus leaves and mature inflorescences were collected from plants growing at Aswan University Botanical Garden, Egypt during spring 2014. The original source of this population is a 15 years-old transplant collected from Egypt-Sudan border, which was left to naturally grow and reproduce, with some occasional supplemental watering. The population consisted of a total of 32 individual plants, of which 15 were mature with inflorescences at different stages of maturity, the others were vegetative. Number of inflorescences/mature individual ranged from 3-30 and the number of flowers/inflorescence ranged from 58-310, of which an average of 50% producing seeds.

Inflorescences were collected from all 15 mature individuals (pale yellow). Seeds for *in vitro* cultures were collected randomly from all inflorescences.

Chemicals

Tissue culture media were purchased from PhytoTechnology Laboratories (Shawnee Mission, KS, USA). DNA extraction kits and Taq polymerase for PCR were purchased from (Promega, Madison, WI, USA) and DNA ladder from (Fisher Scientific, USA).

In vitro regeneration of plants

Seed sterilization

Mature seeds were collected from inflorescences 24 h prior to seed culture. Seeds were collected in cheesecloth, washed for 15 min. under running tap water, followed by rinsing for 5 min. in double distilled water. Seeds were then surface-sterilized by immersing in 95% ethanol for 1 min. followed by 20% Clorox solution for 20 min., with stirring. Under aseptic conditions, Clorox solution was drained and the seeds were rinsed 3 times in sterile double distilled H_2O before using as explants for all three tissue culture systems tested in this investigation.

Somatic embryogenesis

Embryogenic callus was induced from mature seeds on Murashige and Skoog medium (1962) with B5 vitamins (Gamborg *et al*., 1968) (MSB5) supplemented with 1 mg L^1 2,4-Dichlorophenoxy acetic acid (2,4 -D, Sigma, Saint Louis, MO, USA) and 0.5 mg L^{-1} 6-Benzyl adenine (BA, Sigma). First subculture was carried out on the same medium composition and growth regulator concentrations after 6 weeks. Embryogenic callus were transferred to MSB5 medium with 0.25 mg L^{-1} 2, 4-D and 0.125 mg L^{-1} BA, and embryos were germinated on medium with 0.2 mg L^{-1} BA and without 2, 4-D. Plantlets were transferred to MSB5 hormone free medium for root development. Subcultures were carried out at four weeks intervals, sucrose concentrations were 3% in induction and subculture media and 2% in the rooting medium (El-Bakry and Abdelsalam, 2012).

Direct organogenesis

Seeds were cultured on solid MSB5 medium supplemented with 7 mg L^{-1} BA and 0.05 mg L^{-1} 1-Naphthaleneacetic acid (NAA, Sigma). After 4 weeks, multiple shoots were initiated directly from the seed explant without forming callus. Root induction of these shoots was achieved by the addition of liquid MS media supplemented with 0.2 mg L^{-1} BA to the solid induction medium. Elongation of the adventitious roots was achieved by applying hormone free liquid MSB5 medium directly to the solid medium in the original vessel. Plantlets were harvested after 10 weeks from initial seed culture (Abdelsalam *et al*., 2017b).

Indirect organogenesis

Morphogenic calli (non-embryogenic) were induced on MSB5 medium supplemented with 4 mg/L (NAA) and 0.5 mg L^{-1} BA using seeds as explants. Four weeks later, morphogenic calli were subcultured on the same media. Indirect regenerated shoots were formed after 2-3 weeks from first subculture. Shoots were dissected (to individual shoots) and transferred to MSB5 hormone free medium supplemented with 6% sucrose to induce adventitious root formation. Healthy plantlets with roots were harvested after 10 weeks from initial seed culture.

For all experiments, pH of the media was adjusted to 5.8 before autoclaving. Media were solidified by adding 2 g L^{-1} phytagel (Sigma). Replications were 5 plates or magenta containers per treatment, with 5 seeds/culture vessel. Cultures were incubated at 25 °C under cool white fluorescent light (3000 lux) for 16/8 h light/dark photoperiod.

DNA profiling

Sample collection

In vitro propagated shoots from somatic embryogenesis, direct and indirect organogenesis methods were harvested on ice then lyophilized for 48 h.

DNA Extraction

DNA was extracted according to Khanuja *et al*. (1999). DNA pellet was resuspended in 200 µl of TE buffer (Tris EDTA buffer-10 mM Tris HCl, 1 mM EDTA, pH 8.0). DNA quality and concentration were determined using NanoDrop-1000® Spectrophotometer (Nano-Drop Technologies, Wilmington, DE, USA). DNA was diluted to 100 ng μl^{-1} before use in PCR reactions.

PCR reactions, RAPD, ISSR and cpSSR

Each PCR reaction of 25 µl was prepared as follows: one µl of 100 ng μl^{-1} template DNA, mixed with 5 µl 5X Taq buffer, 2.5 µl from 2 mM dNTPs, 1 U of Taq polymerase (0.2 µl), 2 µl of 25 mM MgCl2 and 11.3 µl of PCR-quality distilled water. A total of 18 DNA samples were used in ISSR and RAPD investigation. Those comprised of 3 genotypes from wild plants, five replications from each *in vitro* regeneration method. In cpSSR study, 22 genotypes

were used: 4 genotypes from each of wild plants, and 6 plants from each regeneration system. Preliminary screen of DNA markers was carried out using 50 RAPD primers, 44 ISSR primers and 49 cpSSR primers (Eurofins MWG Operon, AL, USA). Nucleotide sequences information of all ISSR and cpSSR primers is included in the supplementary data (Tables 7 and 8). Thermal cycling conditions were as follows: initial denaturation at 94 °C for 4 min. followed by 40 cycles of 1 min denaturation at 94 °C, 30 s annealing at the specified temperature for each primer, and 1 min extension at 72 °C and final extension for 10 min at 72 °C.

Gel electrophoresis and DNA bands visualization

PCR products were analyzed using 1.5% agarose gel and stained using ethidium bromide (Sigma). Resulting bands were visualized and photographed by Gel Doc XR+ imager® (Bio-Rad, Hercules, California, USA). Bands sizes were determined by the aid of 1kb plus DNA ladder (Fisher Scientific, USA).

Data analysis

Data were collected from at least three genotypes of wild plants and 5-6 genotypes for *in vitro* propagated plants of each tissue culture systems. Gel images were analyzed using Quantity One Software (BioRad Laboratories, Hercules, California). Reproducible bands were scored and entered into a binary data matrix sheet (1 for presence and 0 for absence). This analysis was carried out separately for each type of markers. To analyze the combined data from different markers, the binary data from ISSR, RAPD and cpSSR were combined into one excel sheet and analyzed. The binary data matrix was used to calculate Jaccard's similarity coefficient (Jaccard, 1908) using community analysis package software (CAP) by Henderson and Seaby (1999). The dendrograms were generated using unweighted pair group method with arithmetic average (UPGMA) clustering analysis based on Jaccard's similarity coefficient. The statistical analysis of the phylogenetic tree was confirmed by using 100-bootstrap replicates.

Results and Discussion

The regeneration systems used in the study are illustrated /summarized in Fig. 1. All systems produce healthy plants, but both types of organogenic systems produce plants in a shorter time span (2-3 months), than somatic embryogenesis (5-6 months).

The results for genetic polymorphism using different primer types are given below for ISSR, RAPDs, cpSSR and a combined analysis for all data regenerated irrespective of the primer type used.

ISSR analysis

Genetic diversity among different *in vitro* regeneration systems were analyzed using ISSR markers. Initially, forty-four ISSR primers were tested with two samples per source, and 19 primers were selected depending on bands reproducibility and total amplified fragment. A total number of amplified fragments of 222 were generated with 24 monomorphic bands and 198 polymorphic bands.

Maximum number of amplified fragments was 16 with UBC842T, UBC857C and Y (TG) 7 primers. Minimum number of amplified bands was 8 with UBC846G primer (Table 1 in supplementary data). Some bands were unique to each regeneration system (Table 4). R(CA)7 primer produced higher number of unique bands (16) (Fig. 3A), of which six bands in wild type, 4 bands within plants from somatic embryogenesis, 3 from each organogenesis systems. Polymorphism percentage among wild plant genotypes was 53.3%, among SE plants was 57.9%, among indirect regenerated plants was 48.1% and among direct regenerated plants was 46.7% (Table 1).

Dendrogram generated using UPGMA showed two main clusters. The first cluster included only plants regenerated through indirect organogenesis (Fig. 2A). The second cluster was divided into two sub-clusters: wild plant population in a separate sub-cluster, while *in vitro* regenerated plants through somatic embryogenesis and direct organogenesis in the other sub-cluster.

Jaccard's similarity index (Table 4 supplementary data) showed that, the highest genetic similarity value among groups was 0.51 between embryogenic plants and direct regenerated plants, while the lowest similarity value (0.28) was recorded between embryogenic plants and indirect regenerated plants. Similarity within direct regenerated plants population was higher (0.62 - 0.86) compared to other *in vitro* regenerated systems. Among different *in vitro* propagated systems, plants regenerated through somatic embryogenesis showed higher similarity ranging 0.32 - 0.5 to wild plant, while indirect regenerated plants were less similar to wild genotypes compared to other systems recording similarity between 0.32 and 0.36.

RAPD analysis

Fifty RAPD primers were tested to study preliminary genetic diversity within the wild genotypes and *in vitro* regenerated plants, of which 14 primers were selected. A total of 139 fragments were produced by all primers, of which 24 were monomorphic and 115 were polymorphic bands. Primers OPB-03 and OPD-03 generated maximum number of amplified fragments (16), while lower number of amplified bands was 5 with OPK-07 and OPM-07 primers (Table 2 supplementary data). Some bands were unique to either regeneration system or to wild plant. Higher number of unique bands (6) was observed in indirect regenerated plants. Direct regenerated plants did not show any unique bands (Table 4). Wild genotypes showed 2 unique bands with base pair ranging from 711-700 and 1333-1294 in OPR-07 and 7N3D, respectively (Table 4). RAPD profiles of OPAC-07, OPZ07, OPD07, OPD03 primers are shown in (Fig. 3. E-H).

Polymorphism percentage within population of each plant source based on RAPD analysis was summarized in Table 2. Indirect regenerated plant population produced higher polymorphism percentage 43.5. Polymorphism percentage in direct regenerated plant population was low and close to that of wild plant population 35.3 and 34.1, respectively.

Dendrogram generated using UPGMA analysis separated the studied plants into 2 major clusters.

Table 1. ISSR analysis within W=Wild, SE= Somatic embryogenesis, D= Direct regeneration and ID = Indirect regeneration genotypes. TAF= Total amplified fragments, Mb= Monomorphic bands, Pb= Polymorphic bands

	Primer	Wild			SE			D			ID		
		TAF	**Mb**	**Pb**	**TAF**	**Mb**	**Pb**	**TAF**	**Mb**	**Pb**	**TAF**	**Mb**	**Pb**
1	UBC810	8	3	5	8	4	4	5	5	0	4	4	0
2	UBC811	9	0	9	10	1	9	8	5	3	9	0	9
3	UBC819	5	2	3	3	2	1	4	4	0	7	7	0
4	UBC842C	7	2	5	10	4	6	11	1	10	8	3	5
5	UBC842T	6	3	3	14	1	13	13	3	10	5	3	2
6	UBC846A	6	2	4	7	0	7	5	2	3	5	3	2
7	UBC846G	4	4	0	7	2	5	5	3	2	7	3	4
8	UBC849C	8	5	3	8	3	5	9	6	3	6	2	4
9	UBC849T	6	3	3	8	1	7	9	4	5	7	6	1
10	UBC856C	9	4	5	10	5	5	10	6	4	10	3	7
11	UBC856T	10	5	5	8	7	1	8	6	2	5	5	0
12	UBC857C	7	2	5	8	6	2	7	3	4	11	2	9
13	UBC857T	10	4	6	11	5	6	10	9	1	10	8	2
14	UBC873	8	5	3	11	4	7	7	4	3	6	6	0
15	R(CA)7	6	6	0	4	4	0	4	3	1	5	3	2
16	Y(TG)7	11	5	6	9	8	1	12	3	9	6	3	3
17	UBC834T	5	4	1	10	6	4	9	5	4	9	4	5
18	UBC860A	2	1	1	9	1	8	6	5	1	8	2	6
19	CCC(GT)6	8	3	5	9	5	4	8	3	5	7	3	4
	Total	**135**	**63**	**72**	**164**	**69**	**95**	**150**	**80**	**70**	**135**	**70**	**65**
	%		**46.7**	**53.3**		**42.1**	**57.9**		**53.3**	**46.7**		**51.9**	**48.1**

Table 2. RAPD analysis within W=Wild, SE= Somatic embryogenesis, D= Direct regeneration and ID = Indirect regeneration genotypes.TAF= Total amplified fragments, Mb= Monomorphic bands, Pb= Polymorphic bands

	Primer	Wild			SE			D			ID		
		TAF	**Mb**	**Pb**	**TAF**	**Mb**	**Pb**	**TAF**	**Mb**	**Pb**	**TAF**	**Mb**	**Pb**
1	OPB-03	8	5	3	8	5	3	9	4	5	8	4	4
2	OPD-03	11	10	1	13	8	5	10	8	2	12	7	5
3	OPD-07	7	2	5	11	3	8	9	4	5	11	4	7
4	OPC-07	3	3	0	4	1	3	5	2	3	5	2	3
5	OPE-07	8	4	4	5	2	3	5	4	1	5	5	0
6	OPI-07	4	3	1	8	5	3	7	1	6	9	2	7
7	OPK-07	3	2	1	4	4	0	3	3	0	4	3	1
8	OPM-07	3	3	0	3	3	0	3	3	0	5	2	3
9	OPN-07	3	2	1	8	5	3	2	2	0	3	3	0
10	OPQ-07	7	5	2	7	5	2	8	6	2	8	4	4
11	OPR-07	5	3	2	5	4	1	4	4	0	5	3	2
12	OPV-07	9	2	7	4	4	0	6	4	2	5	5	0
13	OPZ-07	6	5	1	7	4	3	8	6	2	8	6	2
14	OPAC-07	5	5	0	4	3	1	6	4	2	4	2	2
	Total	**82**	**54**	**28**	**91**	**56**	**35**	**85**	**55**	**30**	**92**	**52**	**40**
	%		**65.9**	**34.1**		**61.5**	**38.5**		**64.70**	**35.30**		**56.5**	**43.5**

Table 3. cpSSR analysis within wild plants and *in vitro* regenerated plant groups (W= Wild, SE=somatic embryogenesis, D= direct, ID= Indirect organogenesis).TAF= Total amplified fragments, Mb= Monomorphic bands, Pb= Polymorphic bands

	Primer	Wild			SE			D			ID		
		TAF	**Mb**	**Pb**	**TAF**	**Mb**	**Pb**	**TAF**	**Mb**	**Pb**	**TA**	**Mb**	**Pb**
1	VgcpSSR5-F VgcpSSR5-R	1	1	0	1	1	0	1	0	1	1	1	0
2	CCMP1-F CCMP1-R	6	3	3	4	1	3	4	2	2	5	1	4
3	CCMP2-F CCMP2-R	7	1	6	6	4	2	7	2	5	8	4	4
4	trnK intron-F trnK intron-R	2	2	0	2	1	1	2	2	0	1	1	0
5	psbK-psbI-F psbK-psbI-R	1	1	0	1	1	0	1	1	0	1	1	0
6	rpoC2-rps2-F rpoC2-rps2-R	1	1	0	1	1	0	1	1	0	0	0	0
7	atpI-atpH-F atpI-atpH-R	0	0	0	1	1	0	0	0	0	0	0	0
	Total	**18**	**9**	**9**	**16**	**10**	**6**	**16**	**8**	**8**	**16**	**8**	**8**
	%		**50**	**50**		**62.5**	**37.5**		**50**	**50**		**50**	**50**

Table 4. Unique bands in wild plants and *in vitro* regenerated plant groups (W= wild, SE=somatic embryogenesis, D= direct, ID= indirect organogenesis)

	Primer	Band size range(bp)	Unique bands			
			Wild	SE	D	ID
ISSR	UBC810	1712-1702	-	-	-	1
	UBC842T	1871-1803	-	1	-	-
	UBC846A	2128-2027	-	-	1	-
		1665-1643	-	-	-	1
	UBC856T	1727-1660	-	-	1	-
	UBC857C	2394-2354	-	1	-	-
	UBC857T	1761-1680	-	1	-	-
	R(CA)7	1124-118 1140-1080 1067-1024 932-894 762-707 664-656 565-562 509-506	6	4	3	3
	Y(TG)7	424-343	-	-	-	1
RAPD	OPD-03	365	-	-	-	1
	OPK-07	453-427	-	-	-	1
	OPN-07	1140-1113	-	1	-	-
	OPQ-07	1347-1277 714-668	-	-	-	2
	OPR-07	1967-1903	-	-	-	1
		711-700	1	-	-	-
	OPZ-07	1303-1220	-	-	-	1
	OPAC-07	1333-1294	1	-	-	-
cpSSR	atpI-atpH-F atpI-atpH-	425-444	-	1	-	-
	CCMP1-F	2302-2275 976-918	3	1	1	2
	CCMP1-R	879-832 812-793 778-773				
	CCMP2-F	5927-5785 2359-2263 1000-918	1	4	2	4
	CCMP2-R	906-855 767-749 2580-2519				

Table 5. Polymorphism percentage within studied plants groups (W= wild, SE=somatic embryogenesis, D= direct, ID= indirect organogenesis)

Protocol Genotypes	ISSR	RAPD	cpSSR	All markers
W	53.3	34.1	50	45.8
SE	57.9	38.5	37.5	44.6
D	46.7	35.3	50	44.0
ID	48.1	43.5	50	47.2
All genotypes	51.5	37.85	46.87	

The first cluster included wild genotypes and somatic embryogenic plants, while second cluster included *in vitro* regenerated plants from direct and indirect organogenesis (Fig. 2B).

Jaccard's similarity index is presented in Table 5, supplementary data. Wild plants and embryogenic plants showed higher similarity value among the studied groups (0.64), while indirect regenerated and wild plants showed less similarity (0.46). Similarity values within groups showed that, direct regenerated plants population produced higher within similarity (0.94).

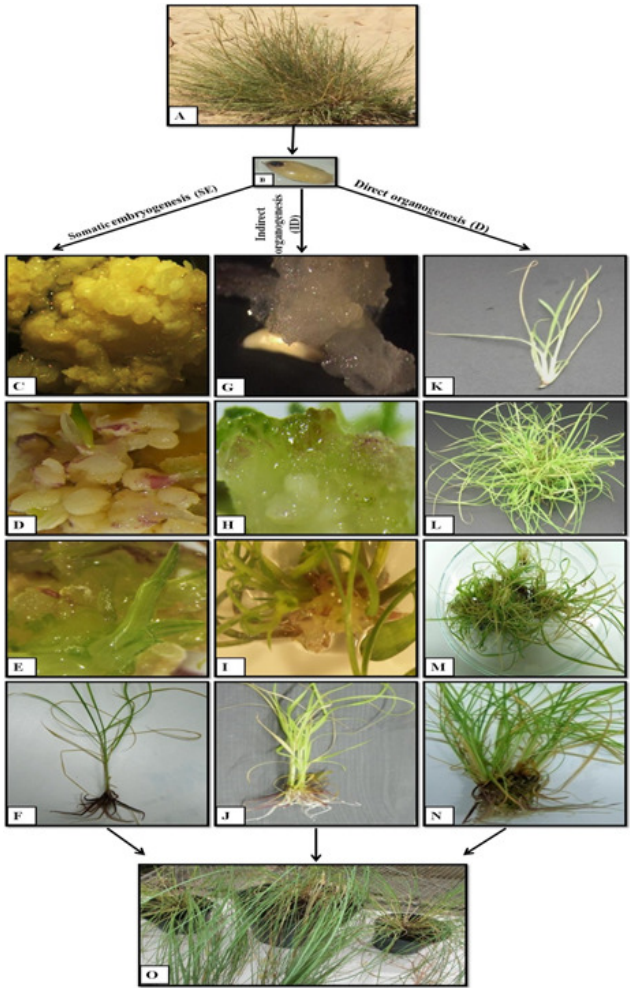

Fig. 1. *In vitro* regeneration from *C. schoenanthus* subsp. *proximus*:(A) Wild plant with mature inflorescence; (B) Mature seed under light microscope; (C) Embryogenic callus, (D) Embryogenic callus with mature somatic embryos; (E) Embryogenic callus with germinating somatic embryos; (F) Rooted plant; (G) Non embryogenic callus; (H) Organogenic callus; (I) Indirect regenerated shoots; (J) Rooted plant; (K) Direct regenerated shoots 1 week old; (L) Direct regenerated shoots 3 weeks old; (M) Direct regenerated shoots 5 weeks old; (N) Rooted plant and (O) Acclimatized plant

cpSSR analysis

Genetic similarities among and within *in vitro* regenerated plants through different tissue culture systems and wild plants were studied using cpSSR DNA markers. Out of 49 primers screened (Table 8, supplementary data), 7 primers were selected. Total number of fragments was 29 that include 23 polymorphic bands and 6 monomorphic bands among all studied groups (Table 3, supplementary data). Each system of the *in vitro* regenerated plants was characterized by unique bands (Table 4). CCMP2 primer (Fig. 3I) produced higher number of unique bands (11 bands), of which 4 bands in each of the indirect organogenic plants and somatic embryogenic plants. Wild plants were characterized by one unique band.

Polymorphism percentage within wild, direct organogenic and indirect organogenic populations were 50%. Somatic embryogenesis system showed less within population polymorphism percentage (37.5, Table 3).

Dendrogram generated using UPGMA analysis for cpSSR data separated plants regenerated through indirect organogenesis into one cluster, while the wild and *in vitro* regenerated plants through direct organogenesis and somatic embryogenesis into a second separate cluster (Fig. 2C). Wild genotypes and somatic embryogenesis genotypes were grouped into 2 sub-clusters and separate from the direct regenerated genotypes. Jaccard's similarity coefficient calculated Table 6 (supplementary data), showed that among *in vitro* regenerated plants, somatic embryogenic regenerated plants were more similar to wild plants with similarity value (0.35 - 0.67). Direct regenerated plants showed high within population similarity range (0.5 - 0.92) within the population.

Combined markers analysis

Combined ISSR, RAPD and cpSSR markers produced a total of 386 bands with 334 polymorphic bands. Jaccard's similarity coefficient among studied groups showed that, plants produced from somatic embryogenesis were more similar to wild plants when compared to plants regenerated from organogenesis. Dendrogram generated using UPGMA data analysis showed same highest similarity of plants regenerated from somatic embryogenesis to the wild (Fig. 4).

The DNA profiles generated based on ISSR, RAPD and cpSSR marker systems were found to be uniform in elucidating genetic diversity of within and among wild plants and those regenerated through different *in vitro* culture systems.

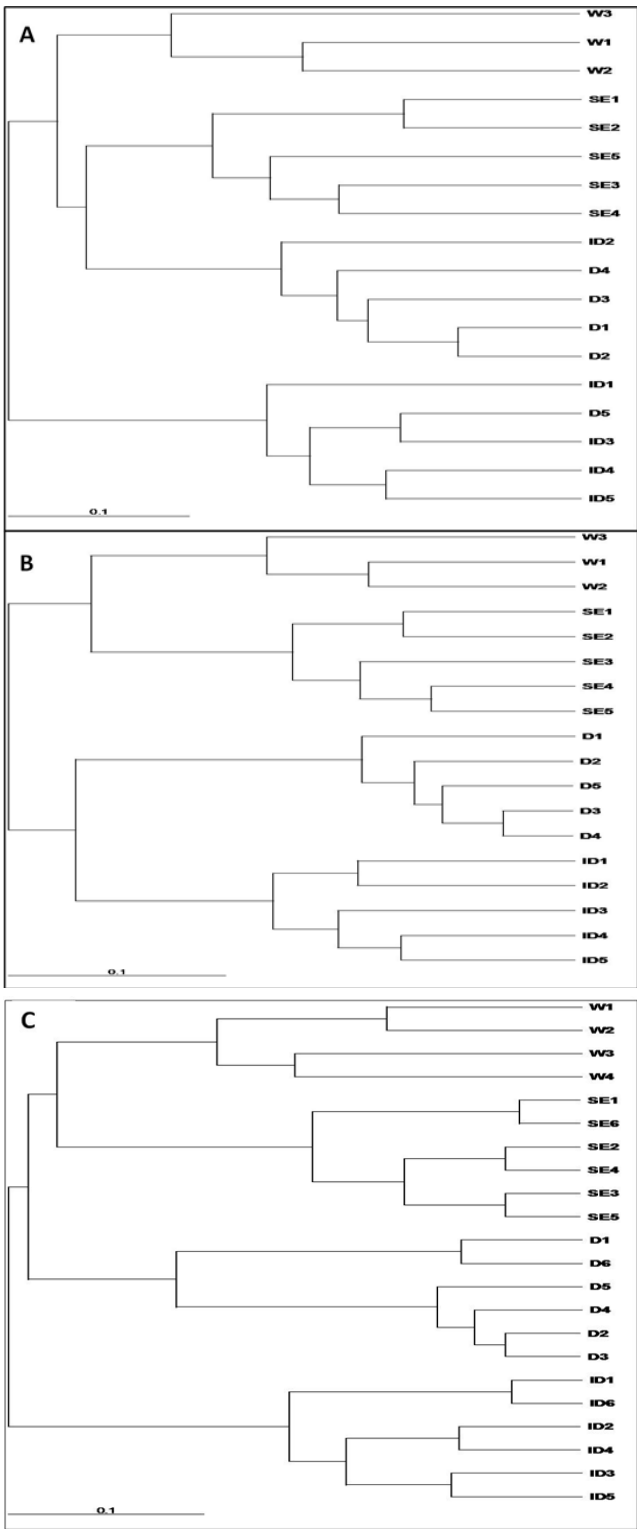

Fig. 2. Dendrogram generated using UPGMA analysis estimated by Jaccard similarity coefficient (100 bootstrap)based on different markers:(A) ISSR, (B) RAPD, (C) cpSSR, for wild plants (W) and *in vitro* regenerated plants from somatic embryogenesis (SE), direct organogenesis (D) and indirect organogenesis (ID)

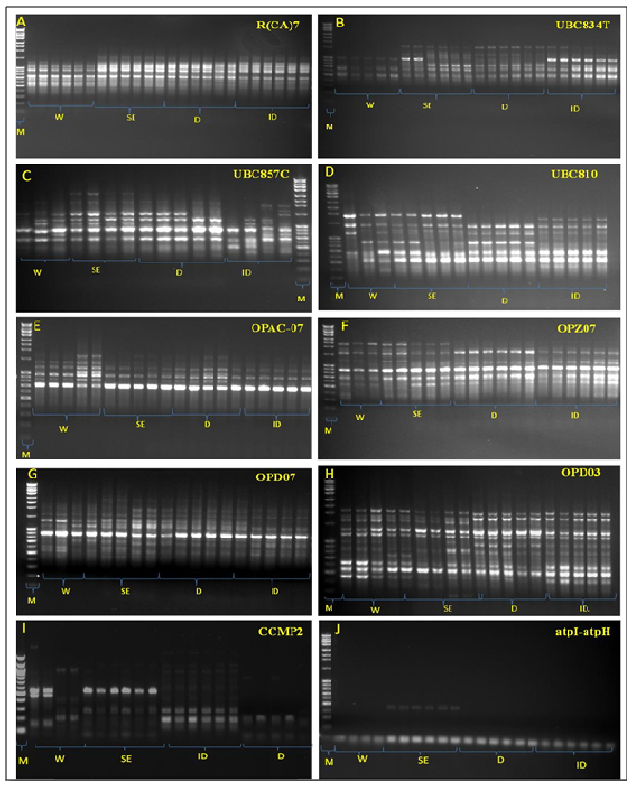

Fig. 3. Gel profiling of wild= W, *in vitro* regenerated plants through somatic embryogenesis (SE), direct organogenesis(D), indirect organogenesis (ID) using ISSR analysis (A-D), RAPD (E-H),cpSSR I,J. M= DNA marker, 100-10000 bp

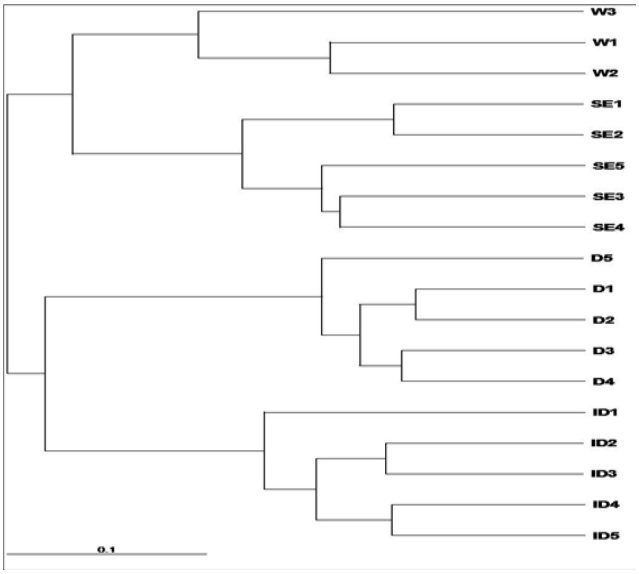

Fig. 4. UPGMA dendrogram generated using Jaccard similarity coefficient and 100 bootstrap for the genetic fidelity of *in vitro* regenerated shoots: SE=somatic embryogenesis, D=direct organogenesis, ID=indirect organogenesis) in comparison to W=wild plants, based on the combined data analysis of ISSR, RAPD and cpSSR primers

Table 6. Jaccard's similarity coefficient range within studied plant groups (W= wild, SE=somatic embryogenesis, D= direct, ID= indirect organogenesis)

Protocol/Genotypes	ISSR		RAPD		cpSSR	
	Range	Average	Range	Average	Range	Average
W	0.52-0.69	0.59	0.70-0.80	0.74	0.55-0.80	0.67
SE	0.52 -0.8	0.64	0.69-0.87	0.78	0.65-0.94	0.79
D	0.62-0.86	0.72	0.79-0.94	0.85	0.61-0.93	0.75
ID	0.60-0.80	0.70	0.69-0.84	0.75	0.50-0.88	0.72

Table 7. Jaccard's similarity coefficient range between wild plants (W) and *in vitro* regenerated plants SE=somatic embryogenesis, D= direct, ID= indirect organogenesis)

	ISSR		RAPD		cpSSR	
	Range	Average	Range	Average	Range	Average
W vs SE	0.36 -0.5	0.45	0.49-0.64	0.55	0.35-0.67	0.46
W vs D	0.35-0.44	0.39	0.41-0.46	0.43	0.38-0.53	0.43
W vs ID	0.32-0.36	0.34	0.42-0.55	0.5	0.29-0.45	0.36

Table 8. Jaccard's similarity coefficient based on combined data analysis ofISSR, RAPD, cpSSR for wild plants (W), *in vitro* regenerated plants through somatic embryogenesis(SE), direct organogenesis(D) and indirect organogenesis (ID)

0	W 1	W 2	W 3	SE 1	SE 2	SE 3	SE 4	SE 5	D 1	D 2	D 3	D 4	D 5	ID 1	ID 2	ID 3	ID 4	ID 5
W1	1																	
W2	0.75	1																
W3	0.60	0.63	1															
SE 1	0.49	0.53	0.52	1														
SE 2	0.51	0.52	0.49	0.81	1													
SE 3	0.49	0.50	0.45	0.68	0.71	1												
SE 4	0.49	0.49	0.43	0.64	0.68	0.75	1											
SE 5	0.47	0.49	0.43	0.59	0.63	0.74	0.73	1										
D1	0.41	0.42	0.39	0.42	0.43	0.46	0.46	0.46	1									
D2	0.41	0.43	0.40	0.41	0.42	0.44	0.46	0.44	0.82	1								
D3	0.41	0.43	0.41	0.43	0.43	0.48	0.45	0.46	0.75	0.83	1							
D4	0.42	0.44	0.42	0.42	0.45	0.46	0.45	0.45	0.73	0.79	0.82	1						
D5	0.41	0.42	0.39	0.40	0.41	0.43	0.45	0.43	0.67	0.74	0.75	0.78	1					
ID1	0.39	0.42	0.40	0.42	0.43	0.44	0.49	0.48	0.51	0.46	0.46	0.48	0.48	1				
ID2	0.40	0.42	0.39	0.41	0.39	0.42	0.44	0.47	0.44	0.41	0.44	0.45	0.45	0.75	1			
ID3	0.40	0.40	0.41	0.40	0.37	0.41	0.43	0.45	0.45	0.42	0.44	0.47	0.47	0.68	0.79	1		
ID4	0.40	0.40	0.42	0.40	0.38	0.40	0.43	0.43	0.48	0.45	0.46	0.48	0.48	0.65	0.70	0.77	1	
ID5	0.39	0.40	0.40	0.39	0.38	0.40	0.44	0.44	0.47	0.45	0.43	0.46	0.46	0.66	0.70	0.74	0.80	1

Discussion

Although the studied wild population was limited in number and origin, but due to the high medicinal value of the species and its potential use in the future, the present work was carried out to explore the genetic diversity using different markers, that is expected to be the basis for future genetic analysis of genotypes from different geographical areas, where the species is known to grow. Also, because the information on the biotechnology in the genus *Cymbopogon* is quite rich (Kumar *et al*., 2009; Quiala *et al*., 2016; Baruah *et al*., 2017), but poor on *C. schoenanthus*, both the different *in vitro* culture regeneration systems (El-Bakry and Abdelsalam, 2012; Abdelsalam *et al*., 2015; Abdelsalam *et al*., 2017b), their genetic analysis and their correlation with the wild genotypes would lay the basis for future genetic improvement of the species for producing high-value genotypes using different avenues of plant biotechnology.

Micro propagation using different *in vitro* culture protocols is an efficient technology for the rapid and continuous supply of the important plants using a limited space under controlled nutrients and environmental conditions (Debnath *et al*., 2006). In some cases, genetic variation occurs in the *in vitro* regenerated plants (somaclonal variation) due to use of supra-optimum concentration of growth regulators and continuous sub-culturing (Nehra *et al*., 1992; Martins *et al*., 2006).

Epigenetic variation plays a role in trans-generational stress-adaption of plants (Weinhold, 2018). A number of epigenetic variations have been reported in family Gramineae including rice and wheat (Chen and Zhou, 2013; Venetsky *et al*., 2015; Lanciano and Mirouze, 2017). Some studies reported that, the *in vitro* regenerated plants showing genetic dissimilarity as a result of genetic and/or epigenetic variation (Smulders and Klerk, 2011; Machczyńska *et al*., 2015).

Molecular markers have been considered as a fast and reliable method for determining genetic fidelity and distinguishing related populations (Semagn *et al*., 2006; Thorat *et al*., 2017).

RAPD and ISSR molecular markers have been used to evaluate genetic diversity in the genus *Cymbopogon* (Sangwan *et al*., 2001; Khanuja *et al*., 2006; Kumar *et al*., 2009).

In the present study preliminary genetic polymorphism within wild plant population genotypes using ISSR primer analysis was 53.3%, using RAPD was 31.4% and in using cpSSR was 16.1%. Adhikari *et al*. (2015) detected a wide range of genetic polymorphism (81.81, 36.36, 5.19) in 3 species of the genus *Cymbopogon* (*C. martinii*, *C. flexuosus* and *C. winterianus*) based on ISSR analysis. Debajit *et al*. (2015) reported higher polymorphism percentage (82.7) in *C. flexuosus*.

Our data showed that, genetic polymorphism among *in vitro* regenerated plants using ISSR markers was higher when compared to RAPD primers. These data agree with Bishoyi *et al*. (2016), where they reported that ISSR primers produced higher polymorphism values in comparison with RAPD primers in nine varieties of genus *Cymbopogon* due to the abundant microsatellite diversity in the genome compared to random primer sequences. Metabolic profiling, as signified by NMR spectroscopy (Abdelsalam *et al*., 2017a), showed that shoots regenerated through somatic embryogenesis were the closest in their metabolic profile to the wild plant population than direct and indirect regenerated plants, in agreement with the present DNA polymorphism, where all analysis grouped the wild and somatic embryogenesis regenerated plants in the same cluster or sub-cluster.

In vitro regenerated plants are genetically similar to wild type and polymorphism exhibited is a mere reflection of variations present within the wild type populations. Bhattacharya *et al*. (2009) evaluated the genetic fidelity of *in vitro* propagated *Cymbopogon martini* plants from rhizome culture using ISSR primers; regenerated plants were genetically similar to donor plants. *In vitro* regeneration of genetically stable *Simmondsia chinensis*, *Alpinia galangal, Musa* spp. and *Artemisia nilagirica* plants has been carried out in different studies (Kumar *et al*., 2011; Parida *et al*., 2011; Agrawal *et al.* 2014; Shinde *et al*., 2016).Genetic polymorphism among wild plant population genotypes was close to genetic polymorphism in *in vitro* regenerated systems using ISSR, RAPD, cpSSR primers, but similarity coefficients were quite low indicating low levels of genetic fidelity between the wild and *in vitro* regenerated plants in our study.

Chloroplast microsatellite (cpSSR) markers have been considered a very useful tool to detect and analyse various forms of genetic diversity such as differentiation and spatial structure among and within populations (Walter and Epperson, 2004).cpSSR makers were used to analyse genetic diversity among and within a range of plant populations, Phumichai *et al*. (2015) studied the genetic diversity of eight populations of *Hevea* rubber genotypes collected from different countries. Genetic variation of two species in the sea grass genus Zostera has been examined by Provan *et al*. (2008). In this study, cpSSR analysis showed high polymorphism ratio within wild plant (50%) and *in vitro* regenerated plant through direct and indirect regenerated plant (50%). Cesare *et al*. (2010) reported that, cpSSR markers were highly polymorphic by analysing different genotypes from *Miscanthus* and related species belong to subfamily Panicoideae including *Cymbopogon citratus*. Our data showed that, UPGMA dendrogram generated using cpSSR data separated the studied plants to two main clusters, wild, embryogenic and direct regenerated plants in one cluster and the indirect regenerated plants in the other. Salvi *et al*. (2001) and Gunathilake *et al*. (2008) described the plants regenerated from direct organogenesis as a safe method for avoiding somaclonal variation, while indirect organogenesis showed higher probability to somaclonal variation (Ramírez-Mosqueda and Iglesias-Andreu, 2015).

Our data showed that, cluster analysis based on ISSR, RAPD and cpSSR separately or using a combined data analysis, plants from somatic embryogenesis and organogenesis were separated into different clusters and wild plants were grouped with the plants regenerated through somatic embryogenesis.

Conclusions

The present work represents a preliminary study of genetic polymorphism in *C. schoenanthus* subsp. *proximus,* collected from Africa. The different DNA marker methods used showed moderately low polymorphism within both the wild genotypes and within the plants regenerated through the different *in vitro* culture systems. UPGAMA analyses and Jaccard's similarity data suggested that plants regenerated through somatic embryogenesis and the wild plants are more similar compared to the plants regenerated through the organogenic systems, indicating that somatic embryogenesis system may be a better choice for future studies towards achieving higher level of clonal fidelity and genetic transformation of the species. The high similarity coefficient manifested within genotypes of plants regenerated through direct organogenesis, and its relatively close similarity to the wild genotypes, suggests its future use for *ex-situ* conservation. The relatively higher genetic polymorphism within plants regenerated through indirect *de novo* organogenesis, suggests their future use in generating and studying genetic variation for the production of high value genotypes of this medicinal plants species. Also, further genetic analysis of other populations of the species in Africa would be both necessary and valuable for the evaluation of the species gene pool, its conservation status and future economic use.

Acknowledgements

This research work was financially supported by the Culture Affairs and Missions Sector, Ministry of Higher Education, Egypt.

Conflicts of interest

The authors declare that there are no conflicts of interest related to this article.

References

Abdelsalam AM, Chowdhury K, El-Bakry AA (2015). Effect of sugar types, culture age, concentrations of 2, 4-D and sucrose on somatic embryogenesis of *Cymbopogon schoenanthus* subsp. *proximus*. Plant Tissue Culture & Biotechnology 25(1):51-62.

Abdelsalam A, Mahran E, Chowdhury K, Boroujerdi A, El-Bakry A (2017a). NMR-based metabolomic analysis of wild, greenhouse, and *in vitro* regenerated shoots of *Cymbopogon schoenanthus*. subsp. *proximus* with GC–MS assessment of proximadiol. Physiology and Molecular Biology of Plants 23(2):369-383.

Abdelsalam A, Chowdhury K, El-Bakry A (2017b). Micropropagation through *in vitro* tillering from seed cultures of the medicinal plant *Cymbopogon schoenanthus* subsp. *proximus*. Asian Journal of Applied Science 5(1):31-40.

Adhikari S, Saha S, Bandyopadhyay TC, Ghosh P (2015). Efficiency of ISSR marker for characterization of *Cymbopogon* germplasms and their suitability in molecular barcoding. Plant Systematic and Evolution 301(1):439-450.

Agrawal A, Sanayaima R, Singh R, Tandon R, Verma S, Tyagi RK (2014). Phenotypic and molecular studies for genetic stability assessment of cryopreserved banana meristem derived from field and *in vitro* explant sources. *In Vitro* Cellular and & Developmental Biology-Plant 50(3):345-356.

Baruah J, Gogoi B, Das K, Ahmed NM, Sarmah DK, La M, Bhau BS (2017). Genetic diversity study amongst *Cymbopogon* species from NE-India using RAPD and ISSR markers. Industrial Crops and Products 95:235-243.

Bhattacharya S, Bandopadhtay TK, Ghosh PD (2009). Somatic embryogenesis in *Cymbopogon pendulus* and evaluation of clonal fidelity of regenerates using ISSR marker. Scientia Horticulturae 123(4):505-513.

Bishoyi AK, Sharma A, Kavane A, Geetha KA (2016). Varietal discrimination and genetic variability analysis of *Cymbopogon* using RAPD and ISSR markers analysis. Applied Biochemistry and Biotechnology 179(4):659-670.

Boulos L (1983). Medicinal plants of North Africa. Reference Publication Inc. Michigan, USA.

Boulos L (1999). Flora of Egypt. Al Hadara publishing, Cairo, Egypt.

Cesare M, Hodkinson TR, Barth S (2010). Chloroplast DNA markers (cpSSRs, SNPs) for Miscanthus, Saccharum and related grasses (Panicoideae, Poaceae). Molecular Breeding 26(3):539-544.

Chen X, Zhou DX (2013). Rice epigenomics and epigenetics: challenges and opportunities. Current Opinion in Plant Biology 16(2):164-169.

Debajit S, Sukriti D, Sneha G, Mohan L, Singh BB (2015). RAPD and ISSR based Intra-specific molecular genetic diversity analysis of *Cymbopogon flexuosus* L. Stapf with a distinct correlation of morpho-chemical observations. Research Journal of Biotechnology 10(7):105-113.

Debnath M, Malik CP, Bisen PS (2006). Micropropagation: A tool for the production of high quality plant-based medicines. Current Pharmaceutical Biotechnology 7(1):33-49.

El Taher K, Abdel-Kader MS (2008). Chemical and pharmacological study of *Cymbopogon proximus* volatile oil. Research Journal of Medicinal Plants 2(2):53-60.

El-Askary HI, Meselhy MR, Galal AM (2003). Sesquiterpenes from *Cymbopogon proximus*. Molecules 8(9):670-677.

El-Bakry AA, Abelsalam AM (2012). Regeneration from embryogenic callus and suspension cultures of the wild medicinal plant *Cymbopogon schoenanthus*. African Journal of Biotechnology 11(43):10098-10107.

El-Dougdoug KA, El Harthi HMS, Korkar HM, Taha RM (2007). Detection of somaclonal variations in banana tissue culture using isozyme and DNA fingerprinting analysis. Journal of Applied Sciences Research 3(7):622-627.

Gamborg OL, Miller RA, Ojima K (1968). Nutrient requirements of suspension culture of soybean root cells. Experimental Cell Research 50(1):151-158.

Gunathilake PMPC, Abayagunawardhana N, Wilamasiri S, Sangakkara UR, Eeswara JP (2008). Direct shoot organogenesis from nodal and leaf explants of *Munronia pinnata* (Wall.) Theob: a valuable medicinal plant. Tropical Agriculture Research 20:213-225.

Jaccard P (1908). Nouvelles recherchés sur la distribution florale. Bulletin de la Societe Vaudoise des Sciences Naturelles 44:223-270.

Khanuja SPS, Shasany AK, Darokar MP, Kumar S (1999). Rapid isolation of DNA from dry and fresh samples of plants producing large amounts of secondary metabolites and essential oils. Plant Molecular Biology Reporter 17(1):1-7.

Khanuja SPS, Shasany AK, Pawar A, Lal RK, Darokar MP, Naqvi AA, ... Kumar S (2006). Essential oil constituents and RAPD markers to establish species relationships in *Cymbopogon* Spreng. (Poaceae). Biochemical Systematic & Ecology 33(2):171-186.

Kumar J, Verma V, Goyal A, Shahi AK, Sparoo R, Sangwan RS, Qazi GN (2009). Genetic diversity analysis in *Cymbopogon* species using DNA markers. Plant Omics 2(1):20-29.

Kumar S, Mangal M, Dhawan AK, Singh N (2011). Assessment of genetic fidelity of micropropagated plants of *Simmondsia chinensis* (Link) Schneider using RAPD and ISSR markers. Acta Physiologiae Plantarum 33(6):2541-2545.

Lanciano S, Mirouze M (2017). DNA Methylation in rice and relevance for breeding. Epigenomes 1(2):1-13.

Locksley HD, Fayez MBE, Radwan AS, Chari VM, Cordell GA, Wagner H (1982). Constituents of local plants XXV, Constitution of the antispasmodic principle of *Cymbopogon proximus*. Planta Medica 45(2):20-22.

Machczyńska J, Orłowska R, Zimny J, Bednarek PT (2014). Extended metAFLP approach in studies of tissue culture induced variation (TCIV) in triticale. Molecular Breeding 34(3):845-854.

Martins SR, Vences FJ, Miera LES, Barroso MR, Carnide V (2006). RAPD analysis of genetic diversity among and within Portuguese landraces of common white bean (*Phaseolus vulgaris* L.). Scientia Horticulturae 108(2):133-142.

Murashige T, Skoog F (1962). A revised medium for rapid growth and bioassay with tobacco tissue culture. Plant Physiology 15(3):473-497.

Nehra NS, Kartha KK, Stushnoff C, Giles KL (1992). The influence of plant growth regulator concentrations and callus age on somaclonal variation in callus culture regenerates of strawberry. Plant Cell, Tissue & Organ Culture 29(3):257-268.

Peredo EL, Revilla MA, Arroyo-Garcı´a R (2006). Assessment of genetic and epigenetic variation in hop plants regenerated from sequential subcultures of organogenic calli. Journal of Plant Physiology 163(10):1071-1079.

Phumichai C, Phumichai T, Wongkaew A (2015). Novel chloroplast microsatellite (cpSSR) markers for genetic diversity assessment of cultivated and wild *Hevea Rubber*. Plant Molecular Biology Reporter

33(5):1486-1498.

Quiala E, Barbón R, Capote A, Pérez N, Jiménez E (2016). *In vitro* mass propagation of *Cymbopogon citratus* Stapf., a medicinal Gramineae. In: Jain S (Ed). Protocols for *in vitro* cultures and secondary metabolite analysis of aromatic and medicinal plants, Second Edition. Methods in Molecular Biology. Humana Press, New York, NY.

Ramírez-Mosqueda MA, Iglesias-Andreu LG (2015). Indirect organogenesis and assessment of somaclonal variation in plantlets of *Vanilla planifolia* Jacks. Plant Cell, Tissue and Organ Culture 123(3):657-664.

Salvi ND, George L, Eapen S (2001). Plant regeneration from leaf base callus of turmeric and random amplified polymorphic DNA analysis of regenerated plants. Plant Cell, Tissue & Organ Culture 66(2):113-119.

Sangwan NS, Yadav U, Sangwan RS (2001a). Molecular analysis of genetic diversity in elite Indian cultivars of essential oil trade types of aromatic grasses (*Cymbopogon* species). Plant Cell Reports 20(5):437-444.

Selim SA (2011). Chemical composition, antioxidant and antimicrobial activity of the essential oil and methanol extract of the Egyptian lemongrass *Cymbopogon proximus* Stapf. Grasas y Aceites 62(1):55-61.

Semagn K, Bjornstad A, Ndjiondjop MN (2006). An overview of molecular marker methods for plants. African Journal of Biotechnology 5(25):2540-2568.

Shinde S, Sebastian JK, Jain JR, Hanamanthagouda MS, Murthy HN (2016) Efficient *in vitro* propagation of *Artemisia nilagirica* var. *nilagirica* (Indian wormwood) and assessment of genetic fidelity of micropropagated plants. Physiology and Molecular Biology of Plants 22(4):595-603.

Smulders MJM, de Klerk GJ (2011). Epigenetics in plant tissue culture. Plant Growth Regulations 63(2):137-146.

Thorat AS, Sonone NA, Choudhari VV, Devarumath RM, Babu KH (2017). Plant regeneration from cell suspension culture in *Saccharum officinarum* L. and ascertaining of genetic fidelity through RAPD and ISSR markers. 3Biotech 7(1):16.

Tiwari JK, Chandel P, Gupta S, Gopal J, Singh BP, Bhardwaj V (2013). Analysis of genetic stability of *in vitro* propagated potato micro-tubers using DNA markers. Physiology & Molecular Biology of Plant 19(4):587-595.

Venetsky AL, Khasdan V, Domb V, Kashkush K (2015). Structure and extent of DNA methylation based epigenetic variation in wild emmer wheat (*T. turgidum* ssp. *dicoccoides*) populations. BMC Plant Biology 15(1):1-10.

Walter R, Epperson BK (2004). Microsatellite analysis of spatial structure among seedlings in populations of *Pinusstrobus* (Pinaceae). American Journal of Botany 91(4):549-557.

Weinhold A (2018). Tran generational stress-adaption: an opportunity for ecological epigenetics. Plant Cell Reports 37(1):3-9.

Supplementary Files

Table 1. Primer score of ISSR analysis for W=wild, SE= Somatic embryogenesis, D= Direct regeneration and ID = Indirect regenerated plants. TAF= Total amplified fragments, Mb= Monomorphic bands, Pb= Polymorphic bands

Primers	Nucleotide sequence 5'-3'	Tm	Band size (bp)	TAF	Mb	Pb
UBC810	GAGAGAGAGAGAGAGAT	54.8	1921-549	10	2	8
UBC811	GAGAGAGAGAGAGAGAC	57.2	1338-510	11	0	11
UBC819	GTGTGTGTGTGTGTGTA	54.8	2252-568	9	1	8
UBC842C	GAGAGAGAGAGAGAGACG	59.9	1883-441	13	0	13
UBC842T	GAGAGAGAGAGAGAGATG	57.6	2376-451	16	1	15
UBC846A	CACACACACACACACAAT	55.3	2128-2027	10	0	10
UBC846G	CACACACACACACACAGT	57.6	1425-441	8	2	6
UBC849C	GTGTGTGTGTGTGTGTCA	57.6	2000-609	10	1	9
UBC849T	GTGTGTGTGTGTGTGTTA	55.3	1985-553	12	1	11
UBC856C	ACACACACACACACACCA	57.6	1682-491	12	1	11
UBC856T	ACACACACACACACACTA	55.3	1827-429	11	2	9
UBC857C	ACACACACACACACACCG	59.9	3446-430	16	2	14
UBC857T	ACACACACACACACACTG	57.6	1761-445	15	2	13
UBC873	GACAGACAGACAGACA	54.2	1554-495	11	2	9
R(CA)7	GCACACACACACACA	53.4	1224-506	9	2	7
Y(TG)7	CTGTGTGTGTGTGTG	53.4	2257-343	16	1	15
UBC834T	AGAGAGAGAGAGAGAGTT	55.3	908-200	11	2	9
UBC860A	TGTGTGTGTGTGTGTGAA	55.3	1738-507	10	0	10
CCC(GT)6	CCCGTGTGTGTGTGT	56.2	1849-571	12	2	9
Total				**222**	**24**	**198**

Table 2. RAPD analysis of wild plant and *in vitro* regenerated plants (SE= Somatic embryogenesis, D= Direct regeneration and ID= Indirect regeneration).TAF= Total amplified fragments, Mb= Monomorphic bands, Pb= Polymorphic bands

Primer name	Nucleotide sequence 5'-3'	Tm	Band size range (bp)	TAF	Mb	Pb
OPB-03	CATCCCCCTG	43.6	1836-324	16	0	16
OPD-03	GTCGCCGTCA	43.6	2169-365	16	3	13
OPD-07	TTGGCACGGG	43.6	1421-373	15	1	14
OPC-07	GTCCCGACGA	43.6	1297-441	7	1	6
OPE-07	AGATGCAGCC	39.5	2438-807	10	2	8
OPI-07	CAGCGACAAG	39.5	2320-436	9	1	8
OPK-07	AGCGAGCAAG	39.5	1146-427	5	1	4
OPM-07	CCGTGACTCA	39.5	1377-278	5	2	3
OPN-07	CAGCCCAGAG	43.6	1140-362	9	1	8
OPQ-07	CCCCGATGGT	43.6	1674-444	11	1	10
OPR-07	ACTGGCCTGA	39.5	1907-700	8	3	5
OPV-07	GAAGCCAGCC	43.6	1984-500	9	2	7
OPZ-07	CCAGGAGGAC	43.6	2275-523	12	4	8
OPAC-07	GTGGCCGATG	43.6	1018-625	7	2	5
Total				139	24	115

Table 3. cpSSR marker analysis of wild plant and *in vitro* regenerated plants (SE= Somatic embryogenesis, D= Direct regeneration and ID= Indirect regeneration).TAF= Total amplified fragments, Mb= Monomorphic bands, Pb=Polymorphic bands

Primer name	Nucleotide sequence 5'-3'	Tm	Band size range (bp)	TAF	Mb	Pb
VgcpSSR5-F	AGCCCACTTTTCCGTAGGTT	60.4	967-902	1	1	0
VgcpSSR5-R	CTTTTCCTTGCCATAATGGTT					
CCMP1-F	CAGGTAAACTTCTCAACGGA	58.4	8486-773	11	0	11
CCMP1-R	CCGAAGTCAAAAGAGCGATT					
CCMP2-F	GATCCCGGACGTAATCCTG	62.3	8206-764	11	1	10
CCMP2-R	ATCGTACCGAGGGTTCGAAT					
trnK intron-F	ATACAGTCTCTTTATCAATATACTG	56.4	154-37	2	1	1
trnK intron-R	GACGTTAAAAATAGATTAGTGCC					
psbK-psbI-F	GGAAAAAYKGGTAATCTATTCC	57.1	127-153	1	1	0
psbK-psbI-R	GAAACAAAGAATATCACTACTG					
rpoC2-rps2-F	TTATTTATTTCAAGCTATTTCGG	53.9	134-123	1	1	0
rpoC2-rps2-R	AATATCTTCTTGTCATTTTTTCC					
atpI-atpH-F	TGGTTGATTGTATCCTTAACC	56.7	445-58	2	1	1
atpI-atpH-R	GAAGCAGCAGCAATTAGTGG					
Total				29	6	23

Supplementary Files

Table 4. Jaccard's similarity coefficient based on ISSR analysis of the *in vitro* regenerated plants through SE=Somatic embryogenesis, D=Direct organogenesis, ID=Indirect organogenesis and W=wild

0	W1	W2	W3	SE 1	SE 2	SE 3	SE 4	SE 5	D1	D2	D3	D4	D5	ID1	ID2	ID3	ID4	ID5
W1	1																	
W2	0.69	1																
W3	0.50	0.57	1															
SE 1	0.46	0.50	0.45	1														
SE 2	0.47	0.49	0.42	0.80	1													
SE 3	0.47	0.46	0.40	0.62	0.63	1												
SE 4	0.47	0.47	0.36	0.59	0.63	0.73	1											
SE 5	0.47	0.45	0.39	0.52	0.56	0.67	0.64	1										
D1	0.39	0.43	0.36	0.43	0.44	0.47	0.47	0.46	1									
D2	0.39	0.44	0.37	0.44	0.46	0.48	0.49	0.45	0.86	1								
D3	0.39	0.41	0.38	0.44	0.45	0.51	0.47	0.48	0.73	0.79	1							
D4	0.38	0.42	0.38	0.43	0.45	0.47	0.44	0.46	0.70	0.75	0.74	1						
D5	0.39	0.42	0.35	0.40	0.40	0.42	0.45	0.42	0.61	0.66	0.66	0.70	1					
ID1	0.33	0.36	0.33	0.36	0.33	0.36	0.41	0.38	0.47	0.46	0.46	0.48	0.48	1				
ID2	0.34	0.36	0.32	0.33	0.30	0.33	0.35	0.37	0.38	0.36	0.39	0.39	0.40	0.72	1			
ID3	0.35	0.34	0.35	0.31	0.28	0.34	0.34	0.36	0.38	0.37	0.39	0.42	0.42	0.66	0.80	1		
ID4	0.34	0.33	0.331	0.30	0.30	0.33	0.34	0.34	0.41	0.40	0.39	0.43	0.43	0.60	0.67	0.75	1	
ID5	0.34	0.35	0.321	0.33	0.32	0.33	0.35	0.36	0.40	0.40	0.37	0.40	0.42	0.60	0.68	0.70	0.78	1

Table 5. Jaccard's similarity coefficient based RAPD data of the *in vitro* propagated shoots through SE=Somatic embryogenesis, D=Direct organogenesis, ID=Indirect organogenesis and W=wild plant

0	W1	W2	W3	SE 1	SE 2	SE 3	SE 4	SE 5	D1	D2	D3	D4	D5	ID1	ID2	ID3	ID4	ID5
W1	1																	
W2	0.81	1																
W3	0.70	0.72	1															
SE1	0.52	0.54	0.64	1														
SE2	0.58	0.59	0.62	0.84	1													
SE3	0.52	0.56	0.54	0.77	0.83	1												
SE4	0.53	0.56	0.57	0.71	0.73	0.78	1											
SE5	0.49	0.55	0.52	0.69	0.72	0.82	0.87	1										
D1	0.45	0.42	0.41	0.39	0.43	0.45	0.49	0.48	1									
D2	0.44	0.41	0.43	0.36	0.36	0.39	0.42	0.42	0.81	1								
D3	0.44	0.44	0.44	0.40	0.40	0.43	0.43	0.44	0.80	0.89	1							
D4	0.46	0.46	0.45	0.39	0.42	0.44	0.46	0.45	0.815	0.85	0.94	1						
D5	0.43	0.43	0.42	0.38	0.40	0.42	0.44	0.43	0.79	0.83	0.86	0.90	1					
ID1	0.50	0.53	0.50	0.54	0.59	0.56	0.63	0.63	0.56	0.48	0.49	0.50	0.53	1				
ID2	0.48	0.53	0.49	0.53	0.53	0.54	0.59	0.63	0.53	0.50	0.52	0.53	0.54	0.80	1			
ID3	0.46	0.49	0.48	0.52	0.51	0.52	0.56	0.60	0.56	0.50	0.53	0.54	0.55	0.71	0.78	1		
ID4	0.49	0.52	0.55	0.54	0.52	0.52	0.58	0.60	0.60	0.55	0.57	0.58	0.57	0.71	0.72	0.79	1	
ID5	0.47	0.50	0.51	0.47	0.48	0.50	0.57	0.56	0.59	0.57	0.54	0.55	0.55	0.72	0.70	0.77	0.84	1

Supplementary Files

Table 6. Jaccard's similarity coefficient based on cpSSR dataof the *in vitro* propagated shoots through SE=Somatic embryogenesis, D=Direct organogenesis, ID=Indirect organogenesis and W=wild plant

0	W1	W2	W3	W4	SE 1	SE 2	SE 3	SE 4	SE 5	SE 6	ID1	ID2	ID3	ID4	ID5	ID6	D1	D2	D3	D4	D5	D6
W1	1																					
W2	0.80	1																				
W3	0.69	0.55	1																			
W4	0.56	0.71	0.71	1																		
SE 1	0.61	0.67	0.50	0.55	1																	
SE 2	0.44	0.42	0.42	0.40	0.71	1																
SE 3	0.53	0.50	0.42	0.40	0.81	0.86	1															
SE 4	0.39	0.37	0.37	0.35	0.65	0.92	0.79	1														
SE 5	0.47	0.44	0.37	0.35	0.75	0.77	0.92	0.85	1													
SE 6	0.56	0.61	0.45	0.50	0.94	0.65	0.75	0.69	0.80	1												
ID1	0.30	0.29	0.35	0.33	0.38	0.44	0.44	0.47	0.47	0.40	1											
ID2	0.35	0.33	0.40	0.38	0.43	0.42	0.42	0.44	0.44	0.45	0.69	1										
ID3	0.40	0.38	0.45	0.43	0.48	0.40	0.40	0.42	0.42	0.50	0.65	0.71	1									
ID4	0.38	0.36	0.43	0.41	0.46	0.38	0.38	0.40	0.40	0.48	0.61	0.88	0.82	1								
ID5	0.37	0.35	0.42	0.40	0.45	0.44	0.44	0.47	0.47	0.47	0.73	0.80	0.87	0.71	1							
ID6	0.29	0.27	0.33	0.32	0.36	0.42	0.42	0.44	0.44	0.38	0.93	0.75	0.71	0.67	0.80	1						
D1	0.40	0.38	0.53	0.43	0.41	0.40	0.40	0.35	0.35	0.36	0.56	0.45	0.43	0.48	0.40	0.53	1					
D2	0.44	0.42	0.42	0.40	0.45	0.44	0.44	0.39	0.39	0.40	0.30	0.35	0.40	0.38	0.37	0.35	0.56	1				
D3	0.47	0.44	0.44	0.42	0.47	0.47	0.47	0.41	0.41	0.42	0.32	0.37	0.42	0.40	0.39	0.37	0.59	0.92	1			
D4	0.44	0.42	0.50	0.47	0.53	0.53	0.53	0.47	0.47	0.47	0.37	0.42	0.47	0.45	0.44	0.42	0.65	0.86	0.92	1		
D5	0.47	0.44	0.44	0.42	0.47	0.47	0.47	0.41	0.41	0.42	0.32	0.37	0.42	0.40	0.39	0.37	0.50	0.92	0.85	0.79	1	
D6	0.40	0.38	0.53	0.43	0.41	0.40	0.40	0.35	0.35	0.36	0.56	0.53	0.50	0.55	0.47	0.61	0.88	0.56	0.59	0.65	0.59	1

Table 7. List of ISSR primers used in the present study with their nucleotide sequence and references

Primer ID	Nucleotide sequence 5'-3'	Reference
UBC810	GAGAGAGAGAGAGAGAT	Melo et al., 2011
UBC811	GAGAGAGAGAGAGAGAC	Zhu et al., 2011
UBC819	GTGTGTGTGTGTGTGTA	Zhu et al., 2011
UBC842C	GAGAGAGAGAGAGAGACG	Zhu et al., 2011
UBC842T	GAGAGAGAGAGAGAGATG	Zhu et al., 2011
UBC846A	CACACACACACACACAAT	Zhu et al., 2011
UBC846G	CACACACACACACACAGT	Zhu et al., 2011
UBC849C	GTGTGTGTGTGTGTGTCA	Zhu et al., 2011
UBC849T	GTGTGTGTGTGTGTGTTA	Zhu et al., 2011
UBC856C	ACACACACACACACACCA	Zhu et al., 2011
UBC856T	ACACACACACACACACTA	Zhu et al., 2011
UBC857C	ACACACACACACACACCG	Zhu et al., 2011
UBC857T	ACACACACACACACACTG	Zhu et al., 2011
UBC861	ACCACCACCACCACCACC	Melo et al., 2011
UBC862	AGCAGCAGCAGCAGCAGC	Melo et al., 2011
UBC873	GACAGACAGACAGACA	Melo et al., 2011
R(CA)7	GCACACACACACACA	Melo et al., 2011
Y(TG)7	CTGTGTGTGTGTGTG	Melo et al., 2011
TA(CAG)4	TACAGCAGCAGCAG	Melo et al., 2011
CRR(ATT)4	CAGATTATTATTATT	Melo et al., 2011
RA(GCT)6	AAGCTGCTGCTGCTGCTGCT	Melo et al., 2011
UBC827	ACACACACACACACACG	Morales et al., 2011
UBC834C	AGAGAGAGAGAGAGAGCT	Morales et al., 2011
UBC834T	AGAGAGAGAGAGAGAGTT	Morales et al., 2011
UBC845A	CTCTCTCTCTCTCTCTAG	Morales et al., 2011
UBC845G	CTCTCTCTCTCTCTCTGG	Morales et al., 2011
UBC848A	CACACACACACACACAAG	Morales et al., 2011
UBC848G	CACACACACACACACAGG	Morales et al., 2011
UBC860A	TGTGTGTGTGTGTGTGAA	Morales et al., 2011
UBC860G	TGTGTGTGTGTGTGTGGA	Morales et al., 2011
(AC)8TG	ACACACACACACACACTC	Li and Ge, 2001
(AC)8TC	ACACACACACACACACCC	Li and Ge, 2001
CCC(GT)6	CCCGTGTGTGTGTGT	Li and Ge, 2001
GCG(AC)6A	GCGACACACACACACA	Li and Ge, 2001
CA(GA)8	CAGAGAGAGAGAGAGAGA	Costa et al., 2011
GC(GA)8	GCGAGAGAGAGAGAGAGA	Costa et al., 2011
(GCT)5C	GCTGCTGCTGCTGCTC	Costa et al., 2011
(GCT)5T	GCTGCTGCTGCTGCTT	Costa et al., 2011
(AGC)5GR	AGCAGCAGCAGCAGCA	Costa et al., 2011
(AGC)5GR	AGCAGCAGCAGCAGCG	Costa et al., 2011
(GAC)5	GACGACGACGACGAC	Shi et al., 2010
(GTG)5	GTGGTGGTGGTGGTG	Shi et al., 2010
(TGTC)5	TGTCTGTCTGTCTGTCTGTC	Shi et al., 2010
(GTGTGG)3	GTGTGGGTGTGGGTGTGG	Shi et al., 2010

Supplementary Files

Table 8. List of cpSSR primers screened with their nucleotide sequence and references

Primer ID	Primer Nucleotide Sequences (5'-3')	References
VgcpSSR1-F	GGTGGATGTTTATACCCAATCG	Pan et al., 2014
VgcpSSR1-R	TCTTTCTGCGATACAAACAAGAA	
VgcpSSR2-F	TTTTCTATGTATGGCGCAACC	Pan et al., 2014
VgcpSSR2-R	CGGGGATAAAGCTGCCTATT	
VgcpSSR3-F	AAACCACTCGAATATTATGGAAA	Pan et al., 2014
VgcpSSR3-R	CCAGTTCAAATCTGGTTCCTG	
VgcpSSR4-F	GAAAAGAACAAGCAAATCCACA	Pan et al., 2014
VgcpSSR4-R	TGATCCTTACGATGCTTCCTTT	
VgcpSSR5-F	AGCCCACTTTTCCGTAGGTT	Pan et al., 2014
VgcpSSR5-R	CTTTTCCTTGCCATAATGGTT	
VgcpSSR7-F	TCAACCATTTCCCAACACCT	Pan et al., 2014
VgcpSSR7-R	CATCGAGTTCATGGATTTGC	
VgcpSSR9-F	TGAAATTTGAAAAACGGGGTA	Pan et al., 2014
VgcpSSR9-R	AAGCGATACGGATAGATTCCT	
VgcpSSR10-F	GGGCTCATTGGCTGTAGAAA	Pan et al., 2014
VgcpSSR10-R	CCATCTCTCCCCAATTGAAA	
VgcpSSR11-F	TTTGAGAAGGTTCAATTGTTCG	Pan et al., 2014
VgcpSSR11-R	TCGGACTCTAGGAAAGGACAA	
VgcpSSR12-F	GGCCATTTATCCCACTTTCC	Pan et al., 2014
VgcpSSR12-R	CCAGTCTCTACTGGGGGTTA	
VgcpSSR13-F	TATTGGTTTTGCACCAATCG	Pan et al., 2014
VgcpSSR13-R	ACCAGGGTGTATGTGCGACT	
VgcpSSR14-F	TGGATCATAATCCTTGAACATCA	Pan et al., 2014
VgcpSSR14-R	TGCGAAAACAAAGATAAGAAATCA	
CCMP1-F	CAGGTAAACTTCTCAACGGA	Weis and Gard, 1999
CCMP1-R	CCGAAGTCAAAAGAGCGATT	
CCMP2-F	GATCCCGGACGTAATCCTG	Weis and Gard, 1999
CCMP2-R	ATCGTACCGAGGGTTCGAAT	
CCMP3-F	CAGACCAAAAGCTGACATAG	Weis and Gard, 1999
CCMP3-R	GTTTCATTCGGCTCCTTTAT	
CCMP4-F	AATGCTGAATCGAYGACCTA	Weis and Gard, 1999
CCMP4-R	CCAAAATATTBGGAGGACTCT	
CCMP5-F	TGTTCCAATATCTTCTTGTCATTT	Weis and Gard, 1999
CCMP5-R	AGGTTCCATCGGAACAATTAT	
CCMP6-F	CGATGCATATGTAGAAAGCC	Weis and Gard, 1999
CCMP6-R	CATTACGTGCGACTATCTCC	
CCMP7-F	CAACATATACCACTGTCAAG	Weis and Gard, 1999
CCMP7-R	ACATCATTATTGTATACTCTTTC	
CCMP8-F	TTGGCTACTCTAACCTTCCC	Weis and Gard, 1999
CCMP8-R	TTCTTTCTTATTTCGCAGDGAA	
CCMP9-F	GGATTTGTACATATAGGACA	Weis and Gard, 1999
CCMP9-R	CTCAACTCTAAGAAATACTTG	
CCMP10-F	TTTTTTTTTAGTGAACGTGTCA	Weis and Gard, 1999
CCMP10-R	TTCGTCGDCGTAGTAAATAG	
CSU01-F	TTCCCGATTCTACTAGCACTC	Peng and Chen, 2011
CSU01-R	ATTATTATCGCTGGTGCAGAG	
CSU03-F	AAAGTATTCCTGACCCAATCG	Peng and Chen, 2011
CSU03-R	ACTAGGACTTATCTTTATCGC	
CSU05-F	TGTTCGATAGCAAGTTGATTG	Peng and Chen, 2011
CSU05-R	GAGTTAGTTGAACTTATCACTC	
CSU07-F	GACTTTCTACTTACAAATCCTG	Peng and Chen, 2011
CSU07-R	ATAATTCACTGATCCACCATG	
trnK intron-F	ATACAGTCTCTTTATCAATATACTG	Provan et al., 2004
trnK intron-R	GACGTTAAAAATAGATTAGTGCC	
psbK-psbI-F	GGAAAAAYKGGTAATCTATTCC	Provan et al., 2004
psbK-psbI-R	GAAACAAAGAATATCACTACTG	
rpoC2-rps2-F	TTATTTATTTCAAGCTATTTCGG	Provan et al., 2004
rpoC2-rps2-R	AATATCTTCTTGTCATTTTTTCC	
atpI-atpH-F	TGGTTGATTGTATCCTTAACC	Provan et al., 2004
atpI-atpH-R	GAAGCAGCAGCAATTAGTGG	

Supplementary Files

atpB-rbcL-F	GATTGGTTCTCATAATTATCAC	Provan et al., 2004
atpB-rbcL-R	TATTGAATTAACTAATTCATTTCC	
NTCP T10-F	CTTCCAAGCTAACGATGC	Provan et al., 2004
NTCP T10-R	CTGTCCTATCCATTAGACAATG	
NTCP T12-F	CCTCCATCATCTCTTCCAA	Provan et al., 2004
NTCP T12-R	ATTTATTTCAGTTCAGGGTTCC	
NTCP T18-F	CTGTTCTTTCCATGACCCCTC	Provan et al., 2004
NTCP T18-R	CCACCTAGCCAAGCCAGA	
NTCP T40-F	GATGTAGCCAAGTGGATCA	Provan et al., 2004
NTCP T40-R	TAATTTGATTCTTCGTCGC	
Rc3-F	TAGGCATAATTCCCAACCCA	Provan et al., 2004
Rc3-R	CTTATCCATTTGGAGCATAGGG	
Rc5-F	ATTTGGAATTTGGACATTTTGG	Provan et al., 2004
Rc5-R	ACTGATTCGTAGGCGTGGAC	
Rc6-F	GAATTTTAGAACTTTGAATTTTTTACCC	Provan et al., 2004
Rc6-R	AAGCGTACCGAAGACTCGAA	
Rc9-F	ATAAGGTTATTCCCCGCTTACC	Provan et al., 2004
Rc9-R	AAATTGGGGGAATTCGTACC	
ARCP1-F	GAACGACGGGAATTGAACC	Cheng et al., 2006
ARCP1-R	GGTGGAATTTGCTACCTTTTT	
ARCP2-F	TGGAGAAGGTTCTTTTTCAAGC	Cheng et al., 2006
ARCP2-R	CGAACCCTCGGTACGATTAA	
ARCP4-F	CAATTCGGGATTTTCCTTGA	Cheng et al., 2006
ARCP4-R	GAGCGAAGGGGTACGAAATA	
ARCP5-F	GGCCATAGGCTGGAAAGTCT	Cheng et al., 2006
ARCP5-R	GTTTATGCATGGCGAAAAGG	
ARCP7-F	TTTACCGAGCAGGTCTACG	Cheng et al., 2006
ARCP7-R	TGAACGATCCCCAGGACTTA	
ARCP9-F	GAAAAATGCAAGCACGGTTT	Cheng et al., 2006
ARCP9-R	TACGATCCGTAGTGGGTTGC	
ARCP11-F	GAGCGAAGGGGTACGAAATA	Cheng et al., 2006
ARCP11-R	CAATTCGGGATTTTCCTTGA	
ccSSR5-F	GGTGCCATTTTAGGATTCCA	Cheng et al., 2006
ccSSR5-R	GAGAAGGTTCCATCGGAACAA	
ccSSR12-F	GCTTTGGTATCTTTCGCCTCT	Cheng et al., 2006
ccSSR12-R	TCCATAGATTCGATCGTGGTT	
ccSSR21-F	AGCCATTTCATTTCGGGTTA	Cheng et al., 2006
ccSSR21-R	ACGCCAGGATGATAAAAAGC	