J. Hortic. Sci.
Vol. 19(2), 2024

https://doi.org/10.24154/jhs.v19i2.2381

Original Research Paper

EST-SSR based genetic diversity analysis and their tagging to coding and non-coding regions in *Moringa oliefera*

Poornima K.N.[1*], Supriya M.[2], Raja Shankar[3], Ravishankar K.V.[4]

[1,4]Division of Basic Sciences, [2,3]Division of Vegetable Crops,
ICAR-Indian Institute of Horticultural Research, Bangalore - 560 089, India
*Corresponding author Email: poornima.n@icar.gov.in

ABSTRACT

Moringa oliefera is one of the most important vegetable tree crops with every part being important for commercial utilization with nutritional and medicinal properties. Development of genomic resources such as expressed sequence tag-based simple sequence repeat markers (EST-SSRs) are desirable due their multiallelic and polymorphic nature. In the present study, a total of 48 polymorphic EST-SSRs were mined from transcriptome data of different parts of *Moringa* plant. A set of 32 genotypes selected from the germplasm have been assessed for diversity using these markers. The mean expected heterozygosity (*HE*) and observed heterozygosity (*HO*) was 0.827 and 0.990, respectively. The PIC value ranged from 0.511 to 0.94. While assigning function to the ESTs, one of the ESTs has been deciphered to contain a CTP. According to the peptide database, ESTs with EST-SSRs had similarity to PR genes and also showed presence of PlantCirc RNA. These developed EST-SSRs were tagged with different triats which can be useful in future population genetics analysis and breeding programs in *Moringa* species.

INTRODUCTION

Moringa oliefera also known as miracle tree crop belongs to the family Moringaceae and contains 13 different species of which *M. oliefera* and *M. peregrina* are widely cultivated. It is grown worldwide due to its vast medicinal values present in each and every tissue of the tree (El-Awady et al., 2016). The leaves of Moringa are rich in nutrients like iron, calcium, phosphorous, potassium, and in antioxidants, sugars and amino acids which makes it suitable to alleviate malnutrition in many countries (Fahey, 2005). Moringa is a true diploid with 2n=28 chromosomes and said to have originated from northwestern Himalayan tract of India (Pandey et al., 2011). The wide applicability of leaf extracts is that it affects various physiological and biochemical functions and alleviates the adverse effects of environmental stresses making it more valuable. Some of the genotypes present in India are Jaffna (soft and taste fruits), Chavakacheri murungai (similar to Jaffna), Chemmurungai (red tipped fruits), Kadumurungai (small and inferior fruits) Palmurungai (bitter taste), Punamurungai (similar to Palmurungai), Kodikalmurungai (short fruit), Palmurungai, Puna Murungai and Kodikkal Murungai and wild Kadumurunga (Leone et al., 2015). The two varieties PKM-1; PKM2 have been developed at Horti Nursery Networks, Tamil Nadu, India for pod production.

Molecular markers are valuable tools for identification, characterization and assessment of genetic diversity both within and between species (Awady et al., 2016). The utilization of molecular markers in assisting conventional plant breeding gives them the strength of their applicability in genetic diversity assessment and also marker assisted plant breeding. DNA markers such as RAPD, ISSRs, AFLP, SSRs have been conventionally used in diversity assessment studies (Hassan et al., 2020).

SSR markers have been used to assess the diversity of other tree crops such as mango (Ravishankar et al., 2015) and Indian sapota (Meghala et al., 2005). EST-SSRs are easy to develop, multi-allelic, transferable across genera and are derived from expressed part of the genome, it becomes the marker of choice to tag the traits to markers (Serbessa et al., 2021). The *Moringa* diversity using EST-SSR markers has been studied (Poornima et al., 2023). The EST-SSRs have been used for study of polymorphism and genetic diversity in many crops like *Jatropha curcas* (Wen et al., 2010), soybean (Mulato., 2010), *Curcuma*

species (Sahoo et al., 2021) etc. Various genetic resources such as RAPD, SSR, ISSR, CYTP450 and other types of markers (Saini et al., 2013; Swat et al., 2020; Ravi et al., 2020) have already been reported in *Moringa* diversity analysis. The present study is emphasized on annotating the transcripts from which the EST-SSRs reported in moringa (Poornima et al., 2023) were derived. Along with genic regions the presence of non-genic regions like CTPs (chloroplast transit peptides), circular RNAs, miRNAs are also deciphered in the present study. An effort is also made to tag the markers to their putative functions and phenotypes. These tagged markers can be used in further breeding programs and population genetic studies.

MATERIALS AND METHODS

The transcriptome sequence from SRA database with accession numbers (SRX3011281-root, SRX3011280-pod, SRX3011278-leaf, SRX3011259-flower, SRX3011282-stem) were down loaded from NCBI (http://www.ncbi.nlm.nih.gov/dbest/) and then 10,000 sequences from these databases were subjected to MISA (https://webblast.ipk-gatersleben.de/misa/) for identification of simple sequence repeats. About 300 sequences showed presence of SSRs and among these about 50 sequences covering all the tissue specific transcripts were used to develop EST-SSR primers using the primer 3 software (http://frodo.wi.mit.edu/primer3). The primers were then synthesized for their validation in the germplasm of *Moringa oliefera* (Poornima et al., 2023).

DNA isolation and PCR

Thirty-two genotypes were selected from the germplasm of moringa which is maintained at ICAR-Indian Institute of Horticultural Research, Bangalore field gene bank. DNA was isolated using CTAB method (Ravishankar et al., 2000). The quality of DNA was assessed on 0.8% agarose gel and quantity was checked with a NanoDrop spectrophotometer ND 1000 (NanoDrop Products, USA). Finally, the DNA was normalized to 10 ng μL^{-1} for PCR amplification.

A 15 µL reaction volume was used for PCR amplification of genomic DNA in an eppendorf tube containing 50 ng genomic DNA, 10 pmol forward primer, 10 pmol reverse primer (tailed with M13 tags), 10 pmol "M13 tag" and 2.5 mM of dNTPs, 0.6 µl of 25mM $MgCl_2$, 0.3 U of *Taq* polymerase (Sigma, USA) and 1.5 µl of 10X PCR buffer (10 mM Tris pH 8.0, 50 mM KCl, 1.8 mM $MgCl_2$ and 0.01 mg/ml gelatin) using Thermal Cycler PCR (Applied Biosystems, Foster City, CA, USA). An initial denaturation at 95°C for 5 min followed by 35 cycles of denaturation at 95 °C for 1 min, annealing at 55-60 °C for 30 s (primer-specific) and extension at 72 °C for 1 min followed by final extension of 72°C for 10 mins and hold at 4°C was followed for the PCR. The confirmation of PCR amplification was done on 1.5% agarose gel, post-PCR multiplex sets were constructed based on fluorescence-labelled primer dyes. The results obtained from DNA Analyzer for electrophoresis for SSR amplicon size was determined with GeneMapper 4.0 software (Applied Biosystems, USA). The PCR-amplified SSR markers were scored by allele size based on the LIZ standard.

Statistical analysis for the calculation of observed heterozygosity (Ho), gene diversity or expected heterozygosity (He), major allele frequency, and polymorphic information content (PIC) of EST–SSR markers was performed with Cervus (https://cervus.software.informer.com/3.0/). The genetic diversity was estimated using the software DARwin var. 6.0.14 wherein the phylogenetic tree was constructed based on the unweighted neighbour-joining method

Annotation of EST-SSRs

The PlantPepDB (http://www.nipgr.ac.in/PlantPepDB/) is a manually curated database of plant peptides having different functions like biochemical or therapeutic. The EST-SSR containing transcripts were subjected to Blast analysis in database to decipher any function encoded by the transcript and assign a genic function.

To identify non-genic functional annotation of the EST-SSR containing transcripts, the databases such as ChloroP (https://services.healthtech.dtu.dk/service.php?ChloroP-1.1) and TargetP (https://services.healthtech.dtu.dk/service.php?TargetP-1.1) were used to identify if the sequences contained any organellar targeting peptide.

The online tool PlantcircNet (http://bis.zju.edu.cn/plantcircnet/) was used to identify plant circular RNAs that may have regulatory properties in plant gene regulation.

Phenotyping

The biochemical analysis has been done using few of the genotypes. The assays include estimation of total phenols, antioxidants, flavonoids, carotenes, nitrogen, calcium magnesium and iron, manganese, zinc and copper. The mean significant values for each of the components are presented in the Table 1.

Table 1 : Biochemical analysis of *Moringa oliefera* genotypes taken as phenotypic traits for tagging SSRs

Genotype	Total phenols (mg/100 g)	FRAP antioxidant activity (mg/100 g fresh weight)	Antioxidant activity (DPPH)	Total flavonoids (mg/100 g)	Total carotenoids (mg/100 g)	N (%)	Ca (%)	Mg (%)	Fe (ppm)	Mn (ppm)	Zn (ppm)	Cu (ppm)
PKM-2	2190.67	2067.1	1234.3	1333	38.9	2.45	1.26	0.42	173.6	319.1	56.6	13.1
PKM-1	2924.25	3375.5	935.5	1942.5	64.6	2.4	0.91	0.43	190.6	201.4	79.5	11.4
IIHR-D-3	2578.17	2593.7	1217.2	1560.3	67.8	1.71	1.27	0.38	135.2	115	55.8	17.9
IIHR-D-5	2221.67	2498.6	510	1446.7	47.9	1.4	1.02	0.49	146.9	178.5	56.5	17.3
IIHR-D-6	2583	2795.5	963.2	1659	55.8	1.61	1.79	0.56	170.7	210.5	67.1	20.7
IIHR-D-7	3229.17	2828.9	1036.3	971.3	46	2.73	1.5	0.55	147.9	260.1	95.3	22.7
IIHR-D-9	2046	2158	983.2	1178	44.7	2.23	2.12	0.6	173.1	244.8	108.5	27.6
IIHR-D-10	3084.5	3756	1061.4	2118.3	59.4	2.6	2.38	0.83	174.6	196.6	91.8	24
IIHR-D-11	2629.83	3260	1112	1457	53.8	2.88	2.02	0.72	146.4	159.4	100.4	17.7
IIHR-D-15	2263	1929.4	1208.3	1384.7	59.5	1.89	3.8	0.94	222.6	195.6	129.6	28.7
Baghya	2829.75	3040	1131.2	1638	55.8	1.95	0.8	0.36	168.5	107.2	69.7	12.5
IIHR-D-17	2630.2.5	2769.1	999.1	1606.5	85	1.21	2.19	0.48	114.1	108.4	51	15.7
IIHR-D-18	2020.17	2218.8	1024.3	1415.7	44.3	1.43	2.32	0.75	108.5	99.8	57.6	16.6
IIHR-D-27	2707.33	3260	1112.3	1715.3	83	2.64	1.88	0.47	134.2	158.2	36.2	15
IIHR-D-28	2509.5	2499.4	1041.3	1533	40.7	1.47	1.56	0.6	188.3	101.1	45.5	12.4
IIHR-D-38	2614.33	2792.5	894.3	1529.3	15.7	2.76	1.95	0.64	89.8	120.8	51.1	14.6
IIHR-D-55	1689.5	1887.7	1891.4	1105.7	64	1.64	3.47	0.62	173.7	422.1	128.9	30.9
IIHR-D-62	2588.5	2718.5	1173.3	1705	65.1	2.23	2.91	0.66	422.9	203.3	40.8	17.9
IIHR-D-64	2717.67	2795.4	1123.3	764.7	53.9	1.4	2.15	0.57	123.3	207.9	45.1	15.1
IIHR-D-69	1937.5	2220	846.2	1374.3	62.9	2.2	3.43	0.69	209.8	210.3	130.2	34.7
IIHR-D-78	2691.83	3121.9	1718.18	1880.7	48.5	1.33	1.7	0.61	102.9	217.2	45.1	15.3
IIHR-D-101	2091.1	2077.32	2181.18	1744.99	80.26	2.56	1.95	0.66	142.95	34.73	45.47	12.79
IIHR-D-103	2098.08	2023.23	2507.58	1305.86	76.08	3.2	2.21	0.67	129.17	33.5	38.27	14.99
IIHR-D-107	2559.55	2758.79	3129.9	1653.61	77.06	2.8	1.44	0.58	137.89	25.56	36.61	24.13
IIHR-D-120	2143.9	1979.54	3167.12	1745.37	66.43	3.34	2.36	0.64	113.54	31.86	46.97	11.76
IIHR-D-126	2662.91	2719.61	3593.64	1850.39	73.07	2.85	1.48	0.67	100.06	22.66	31.65	13.58

Nitrogen was determined by Micro-Kjeldahl (Kjeldahl et al., 1883) method. Plant sample (0.5 g) was digested using 10 ml concentrated sulphuric acid in presence of digestion mixture ($CuSO_4$, K_2SO_4 and Se in 100:20:1 ratio) in a digestion chamber.

Secondary nutrients *viz.*, Ca and Mg were determined with the help of atomic absorption spectrophotometer (Jackson et al., 1973) and expressed in per cent. Micronutrients (iron, manganese, zinc and copper) were determined by the atomic absorption spectrophotometer (Jackson et al., 1973) expressed in ppm (parts per million).

Vitamin C (ascorbic acid) content was estimated by visual titrimetric method using 2,6- dichlorophenol indophenols dye solution (50.0 mg dye dissolved in 42.0 mg sodium bicarbonate solution and volume made up to 200 ml). About 0.5 g leaf powder was weighed, and volume made up to 15.0ml with 3.0 % HPO_3 (meta phosphoric acid) and then filtered. 1.0 ml of filtrate was taken in a conical flask and titrated against dye to a distinct pink end point, which persisted for at least 15 seconds. Dye solution was standardized by titrating it against freshly prepared ascorbic acid solution (50.0 mg ascorbic acid in 50.0 ml of 3.0% meta- phosphoric acid solution). Then, ascorbic acid content in terms of mg/100.0 g dry weight was calculated using the following formula:

$$= \frac{\text{Titre volume of sample x dye factor x volume made up x 100}}{\text{Aliquat of extract taken for estimation x weight or volume of sample taken for estimation}}$$

Total carotenoids content of drumstick leaf powder was estimated using 100.0 mg sample added with acetone up to 25.0 ml. The solution was shaken well and kept for one whole night in dark condition to get two separate clear layers. Collected the solution from upper layer and reading were taken at 450 nm, 470 nm, 644.8 nm and 661.6 nm.

Total phenolic content in the individual genotypes leaf samples was determined using Folin-Ciocalteu (FC) assay (Singleton et al., 1999), and the values are expressed as mg gallic acid equivalent (GAE), per 100 g fresh weight. Total flavonoid content (TFC) was determined using aluminum chloride/sodium nitrite method according to Zhishen et al. (1999).

Total antioxidant potential was measured both in terms of radical scavenging activity using DPPH (1,12 - diphenyl-12-picrylhydrazyl) and ABTS radicals as well as in terms of reducing power using FRAP methods. Trolox was used as a positive control in both methods, and the DPPH radical scavenging activity (*S%*) was calculated using the following equation:

$S\% = [(A \text{ control} - A \text{ sample}) / A \text{ control}] \times 100$

Where, *A* control is the absorbance of the blank control (containing all reagents except the sample extract): *A* sample is the absorbance of the test sample.

The 50% DPPH radicals scavenging concentration (SC_{50}) of the sample as well as trolox (*i.e.* The amount of tested samples and trolox to scavenge 50% of the DPPH radicals) was calculated. Finally, DPPH values for samples were calculated using the following formula

SC_{50} of sample (μmol/100g) = (SC_{50} concentration of trolox (μmol)/weight of the sample required for SC_{50} (g) x 100

The DPPH radical scavenging activity (SC_{50}) was expressed as µmol trolox equivalent/100 g dry sample.

The FRAP assay was done according to Benzie & Strain (1996) using the reaction mixture comprising sodium acetate buffer (300 mM, pH 3.6), 10 mM TPTZ (2, 4, 6-tripyridyl-s-triazine) solution in 40 mM HCl, and 20 mM $FeCl_3 \cdot 6H_2O$ solution in the ratio of 10:1:1, and the known volume of sample was allowed to react with 2.0 mL of the reaction mixture. The intensity of color developed was measured at 593 nm after 30 minutes. Trolox served as a positive control and results were expressed in µmol TE/100 g of leaf powder.

Population structure and genetic diversity

The genetic structure of the population under study was analyzed using STRUCTURE 2.3.4 (Pritchard et al., 2000) software and the genetic diversity was estimated using the software DARwin var. 6.0.14, where in the phylogenetic tree was constructed based on the unweighted neighbour-joining method.

Tagging of EST-SSRs to Phenotype

The EST-SSRs were tagged to phenotypic trails using genomic association and prediction integrated tool (GAPIT) (Lipka et al., 2012) using R software. To carry out this analysis the SSR fragments were reclassified with (1) indicating presence of fragment and (0) indicating absence of a fragment. The quantile-quantile (QQ) plot, i.e. the QQ negative log10-transformed observed *p*-values obtained for each MTA, against their expected distribution under the null hypothesis of no genetic association.

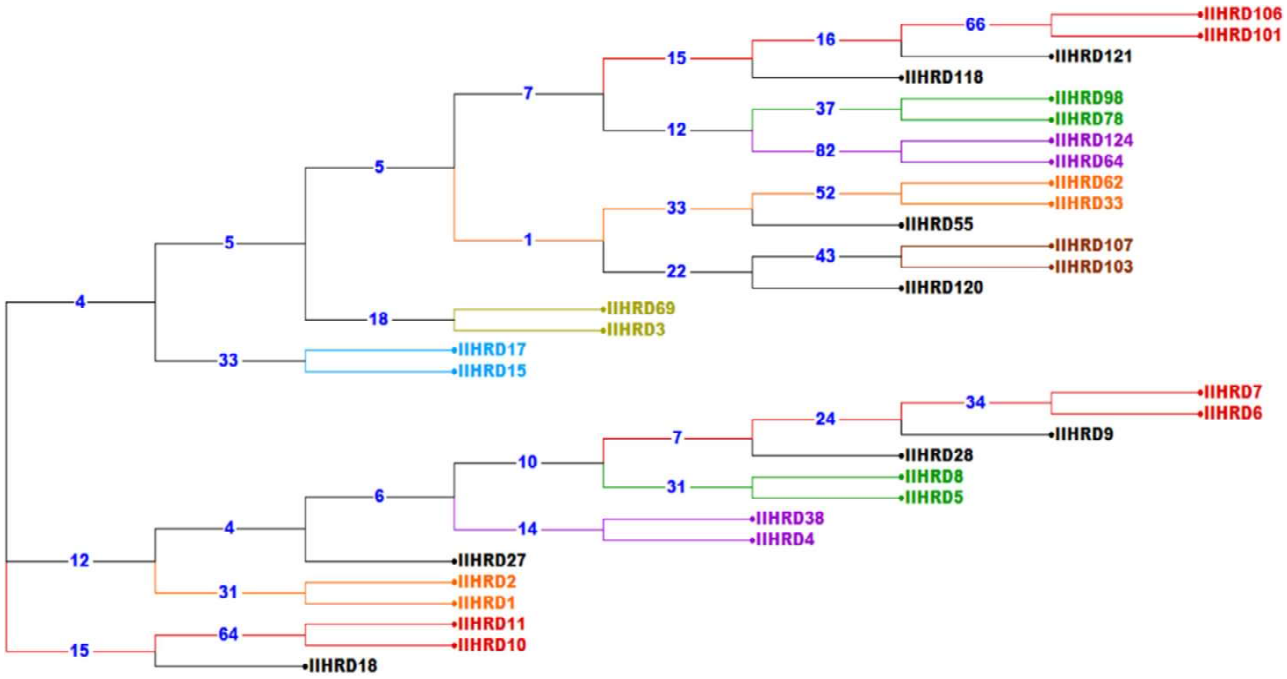

Fig. 1 : Neighbour-joining phylogenetic tree of *M. oliefera* genotypes under study

RESULTS AND DISCUSSION

The EST-SSRs reported in the study (Poornima et al., 2023) were composed mainly of mononucleotide repeats followed by trinucleotide repeats. The mean PIC value of the SSR markers was found to be 0.78 and the expected heterozygosity was 0.82. The diversity analyzed using these EST-SSR markers placed the 32 genotypes into three clusters (Fig. 1). Structure analysis was performed by STRUCTURE and prediction of the number of subpopulations was done through estimation of delta K. The delta K value was found to be four upon 10000 burn-in and 100000 MCMC with 10 iterations. Thus, it indicated presence of four subpopulations within the association panel (Fig. 2).

The blast results of the SSR containing expressed sequences with the PlantPeptide database are given in Table 2. Most of these genes are related to plant pathogenesis especially the thaumatin family of protein. The sequence from which the EST-SSR primer 4 is derived encodes for a lipid transfer protein. The e-values of these predicted genes ranged from 0.79 to 10 which are quite significant.

The ChloroP-1.1 predicts the presence of chloroplast transit peptides (cTP) in protein sequences and the location of potential cTP cleavage sites. Results showed that 13 of the EST-SSR containing sequences encode for chloroplast transit peptide with highest CS-score of 8.834 being given to Primer 2 containing sequence (Table 3). The total length of this CTP was found to be 24 bases which are also significant. TargetP-2.0 predicts the presence of N-terminal presequences: signal peptide (SP), mitochondrial transit peptide (mTP), chloroplast transit peptide (cTP) or thylakoid luminal transit peptide (lTP). For the sequences predicted to contain an N-terminal presequence a potential cleavage site is also predicted (Table 4).

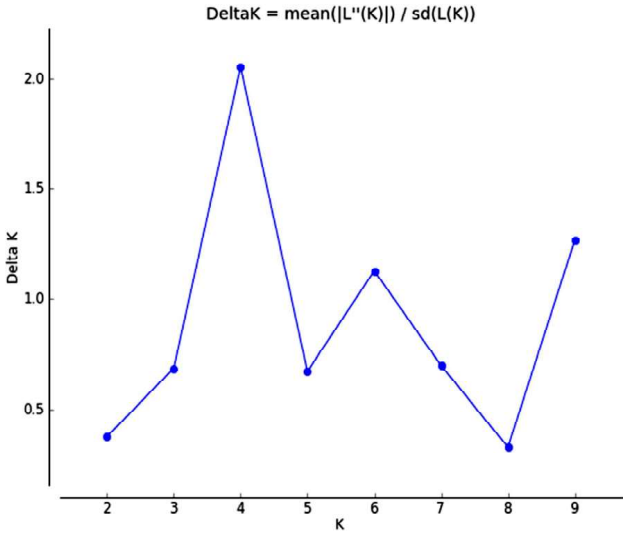

Fig. 2 : Population structure of the 32 genotypes of *M. oliefera* showing highest K value=4

Table 2 : EST annotation using plant peptide database

Primer	Plant	Peptide DB annotation	Peptide Function	E-value
P1	*Medicago truncatula*	Pathogenesis-related thaumatin family protein	Antimicrobial	4.4
P2	*Pinus monticola*	Thaumatin-like protein L4	Antimicrobial	10
P3	*Hordeum vulgare*	Hordothionin omega	Protein-translation-inhibitor	2.2
P4	*Triticum aestivum*	Ta-LTP3	Lipid-transfer protein	8.7
P10	*Beta vulgaris*	AX1	Antibacterial, Antifungal	3.4
P11	*Gloeospermum pauciflorum*	Glopa B	Enzymatic-digestion	6.4
P13	*Solanum lycopersicum*	pathogenesis-related protein 5	Antimicrobial	8.7
P15	*Zea mays*	Zeamatin precursor	Antimicrobial	5.7
P23	*Viola tricolor*	tricyclon B	Hemolytic, Antimicrobial	4.4
P24	*Fragaria vesca*	thaumatin-like protein-like	Antimicrobial	3.1
P26	*Viola tricolor*	Cyclotide vitri peptide 95	Antimicrobial	1.4
P29	*Capsicum annuum*	Proteinase inhibitor PSI-1.2, HyPep	Antifungal, Serine-protease-inhibitor	2.2
P31	*Solanum lycopersicum*	pathogenesis-related protein 5	Antimicrobial	8.7
P32	*Zea mays*	Zeamatin precursor	Antimicrobial	5.7
P40	*Chenopodium quinoa*	Osmotin-like protein 4	Antimicrobial	7.4
P42	*Vitis vinifera*	Chitinase basic class I [21-60]	Antimicrobial, Antifungal	0.79
P43	*Arabidopsis thaliana*	At2g28790 (Pathogenesis-related thaumatin-like protein)	Antimicrobial	8.4
P45	*Secale cereale*	Basic endochitinase C	Antifungal	7.2

PlantCircNet is a database that provides information on plant circRNA-miRNA-gene regulatory networks. Among the 48 EST-SSR containing sequences, two of them showed the presence of circular RNA when a sequence BLAST against the PlantCircNet database was performed using EST-SSR containing sequences. The two primers, primer 4 and primer 16 can be tagged to the circular RNAs which have gene functions related to *Oryza sativa* and *Glycine max* genes. The sequence tagged by primer 4 has match to Os_ciR5493 circular RNA and primer 16 has match to GMA_circ00564 (Fig. 1)

The association between markers and traits was predicted using GAPIT software. The phenotyping (Table 1) and genotyping data (Supplementary Table 1) are combined to derive at the associations based on mixed linear model (MLM) which is a statistical model used to detect significant marker-trait associations while controlling for population structure and kinship. The primers associated with the phenotype are presented in Supplementary Table 2, where the significant P-values of <0.05 was considered for associating the markers to phenotypic traits.

Molecular breeding and molecular biology of *Moringa oliefera* has been assisted by the development of different genetic resources such as RAPD, SSR, ISSR, CYTP450 and other types of markers (Saini et al., 2013; Swat et al., 2020; Ravi et al., 2020). An addition to this repertoire of genomic resources is the EST-SSRs thus developed by our group at ICAR-IIHR. The mean expected heterozygosity (*HE*) ranged from 0.605 to 0.958 and observed heterozygosity (*HO*) ranged from 0.938 to 1.00 and PIC value ranged from 0.511 to 0.94. In a previous study by Natarajan et al. (2015) SSR markers in moringa, the SSR MO 10 recorded the highest PIC value (0.84), followed by SSR MO 58 (0.75). The lowest PIC value (0.39) was recorded for SSR MO 18. The mean PIC value was

Table 3 : chlorop v1.1 prediction results where CTP along with cleavage site is represented

Primer	Length	Score	cTP	CS- score	cTP- length
Primer1	320	0.461		-0.789	34
Primer2	421	0.516	Y	8.834	24
Primer3	284	0.479		-1.408	52
Primer4	349	0.431		2.163	20
Primer8	599	0.561	Y	-0.382	34
Primer10	1150	0.44		-0.444	7
Primer11	191	0.463		0.579	55
Primer14	432	0.54	Y	-2.637	37
Primer15	1807	0.48		-1.125	49
Primer16	247	0.434		-0.752	22
Primer21	536	0.545	Y	2.634	36
Primer22	325	0.501	Y	-0.476	63
Primer23	294	0.481		8.206	22
Primer24	471	0.439		-1.683	5
Primer26	614	0.422		3.862	91
Primer27	911	0.511	Y	-0.58	82
Primer28	450	0.562	Y	1.46	43
Primer29	632	0.453		2.005	42
Primer32	432	0.54	Y	-2.637	37
Primer33	1807	0.48		-1.125	49
Primer34	247	0.434		-0.752	22
Primer35	197	0.469		4.726	80
Primer40	302	0.544	Y	2.383	50
Primer41	295	0.425		11.007	91
Primer42	279	0.534	Y	3.605	39
Primer43	448	0.566	Y	4.374	86
Primer44	295	0.585	Y	2.411	64
Primer45	565	0.513	Y	3.805	63

0.52 which is almost same as obtained in our study (Table 5).

A greater meaning is given to these markers when they are tagged to certain genic and non-genic functions. The PlantPeptide database is the first database comprising detailed analysis and comprehensive information of phyto-peptides from a broad functional range which will be useful for peptide-based applied research. The EST-SSRs were annotated to comprise of genes involved in plant defense response. These sequences were found to be homologous to pathogenesis genes found in plants like Medicago, Pinus, Hordeum, Capsicum and many other species.

Chloroplast transit peptide (CTP) plays an important role in the plant developmental process. These peptide molecules carry different proteins into the chloroplast and also specifically into mitochondria. In one of the studies by Christian et al. (2020), the origin of plastid transit peptides, multi-genome and pan-genomic assessment of CTP evolution has been discussed. Xu et al. (2018) have cloned and functionally validated the Rubisco small subunit chloroplast transit peptide in tomato. These studies show that identification of CTP is important in genome analysis and an SSR marker tagged to such an important element of the genome may be very useful for analyzing the gene functions. Thus, the databases such as chloroP, TargetP have been used in the present study to identify the different transit peptides.

Plant circular RNA are a group of non-coding RNAs with a closed-loop structure that are produced via alternative splicing. They are proposed to have

Table 4 : Targetp v1.1 prediction results with cleavage site predictions for chloroplastic, mitochondrial or thylakoid localization

Name	Length	cTP	mTP	SP	other	Loc	RC	TPlen
primer1	320	0.079	0.021	0.753	0.11	S	2	17
primer2	421	0.124	0.818	0.006	0.053	M	2	13
primer3	284	0.039	0.233	0.014	0.432	_	5	-
primer4	349	0.077	0.277	0.044	0.511	_	4	-
primer8	599	0.135	0.641	0.016	0.091	M	3	94
primer10	1150	0.003	0.136	0.923	0.152	S	2	33
primer11	191	0.025	0.023	0.385	0.162	S	4	30
primer14	432	0.057	0.291	0.49	0.207	S	5	42
primer15	1807	0.009	0.461	0.184	0.048	M	4	15
primer16	247	0.086	0.476	0.069	0.665	_	5	-
primer21	536	0.108	0.03	0.116	0.722	_	2	-
primer22	325	0.61	0.057	0.014	0.588	C	5	63
primer23	294	0.071	0.434	0.214	0.292	M	5	21
primer24	471	0.034	0.243	0.15	0.812	_	3	-
primer26	614	0.146	0.263	0.007	0.84	_	3	-
primer27	911	0.33	0.521	0.023	0.087	M	5	20
primer28	450	0.079	0.016	0.373	0.008	S	4	16
primer29	632	0.053	0.054	0.761	0.157	S	2	18
primer32	432	0.057	0.291	0.49	0.207	S	5	42
primer33	1807	0.009	0.461	0.184	0.048	M	4	15
primer34	247	0.086	0.476	0.069	0.665	_	5	-
primer35	197	0.008	0.497	0.04	0.445	M	5	66
primer40	302	0.263	0.732	0.055	0.006	M	3	24
primer41	295	0.013	0.031	0.08	0.75	_	2	-
primer42	279	0.748	0.116	0.02	0.331	C	3	39
primer43	448	0.714	0.12	0.002	0.17	C	3	86
primer44	295	0.9	0.105	0.016	0.009	C	2	64
primer45	565	0.283	0.247	0.016	0.3	_	5	-
cutoff		0	0	0	0			

important biological roles such as plant growth, development and stress responses (Moazzam-Jazi et al., 2021). Han et al. (2020) have identified, characterized and functionally predicted circular RNAs in maize. Pan et al. (2017) reported that *Arabidopsis* seedlings presented 1583 heat-specific circRNAs, and that heat stress could increase the quantity, length, and alternative circularization events of circRNAs. Knowing the importance of circular RNAs, the EST-SSR containing sequences were subjected to prediction of circRNAs and two sequences were found homologous to circRNAs from rice and soyabean (Fig. 3).

Table 5 : Estimates of population genetic parameters based on frequency of the alleles at 36 EST–SSRs marker loci in a set of 32 *M. oliefera* germplasm accessions

SSR locus	Total number of alleles	Effective number of alleles	HObs	HExp	UGD	PIC
MO13	7	3.401	1	0.706	0.720	0.646
MO14	10	4.717	1	0.788	0.804	0.742
MO15	14	6.250	0.969	0.84	0.857	0.807
MO16	16	8.696	0.969	0.885	0.903	0.859
MO17	7	3.546	1	0.718	0.732	0.66
MO18	5	3.322	1	0.699	0.713	0.634
MO19	17	7.246	0.969	0.862	0.879	0.833
MO20	21	15.625	0.969	0.936	0.955	0.916
MO21	6	4.854	1	0.794	0.810	0.751
MO22	7	3.690	1	0.729	0.744	0.669
MO23	14	5.814	0.969	0.828	0.845	0.792
MO24	17	10.753	1	0.907	0.925	0.885
MO25	7	3.115	1	0.679	0.693	0.612
MO26	10	6.098	1	0.836	0.853	0.799
MO27	21	9.524	0.969	0.895	0.913	0.873
MO28	24	19.608	1	0.949	0.968	0.93
MO29	8	5.208	1	0.808	0.824	0.763
MO30	8	4.167	1	0.76	0.775	0.71
MO31	20	10.204	0.938	0.902	0.920	0.88
MO32	30	23.810	1	0.958	0.977	0.94
MO33	7	4.651	1	0.785	0.801	0.742
MO34	6	4.115	1	0.757	0.772	0.705
MO35	20	12.195	0.938	0.918	0.936	0.897
MO36	18	12.500	0.969	0.92	0.938	0.899
MO37	10	4.545	0.969	0.78	0.796	0.74
MO38	6	4.000	1	0.75	0.765	0.694
MO39	20	11.236	1	0.911	0.929	0.89
MO40	20	15.152	1	0.934	0.953	0.914
MO41	6	4.000	1	0.75	0.765	0.693
MO42	4	2.532	1	0.605	0.617	0.511
MO43	22	7.092	1	0.859	0.876	0.836
MO44	21	13.889	1	0.928	0.947	0.908
MO45	9	4.808	1	0.792	0.808	0.749
MO46	6	3.717	1	0.731	0.746	0.677
MO47	20	13.158	1	0.924	0.942	0.903
MO48	21	14.925	1	0.933	0.952	0.913
Total	485	292.163				
Average	13.472	8.116	0.990	0.827	0.843	0.788

Hobs: observed heterozygosity, Hexp: expected heterozygosity, UGD: unbiased genetic diversity PIC: polymorphism information content

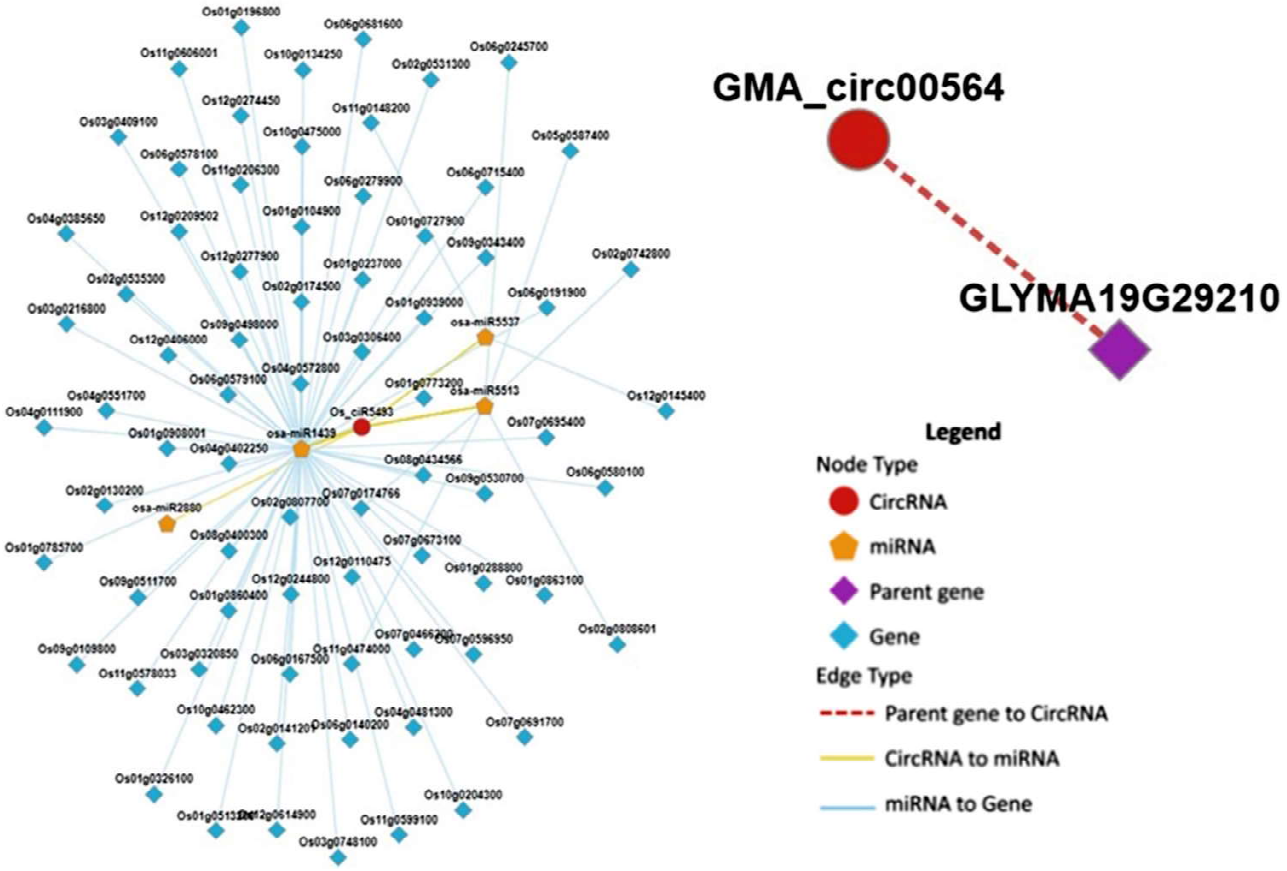

Fig. 3 : The plant circular RNA prediction and their interaction with miRNAs and other functional genes using PlantCircNet database

Tagging of markers to traits is a powerful way to investigating multiple or complex traits related to any single/multiple stress. Such associations in various plants/crops have identified novel gene candidates, or genes or quantitative trait loci, responsible for abiotic stress, biotic stress and many other traits. The structured association analysis between twelve phenotypic traits and 48 EST-SSR marker genotype data of 32 genotypes was performed using R package program GAPIT Supplementary Table 2. About nine marker trait associations were obtained and the QQ plots from single locus and multi-locus models revealed them being fitted properly to the datasets (Fig. 4).

The marker MO 14 showed significant association with Ca and Mg content. The marker MO15 was found to be associated with the total phenols and antioxidant traits significantly with P value 0.03 and 0.04, respectively which is less than the highest probability 0.05. the marker MO44 was found

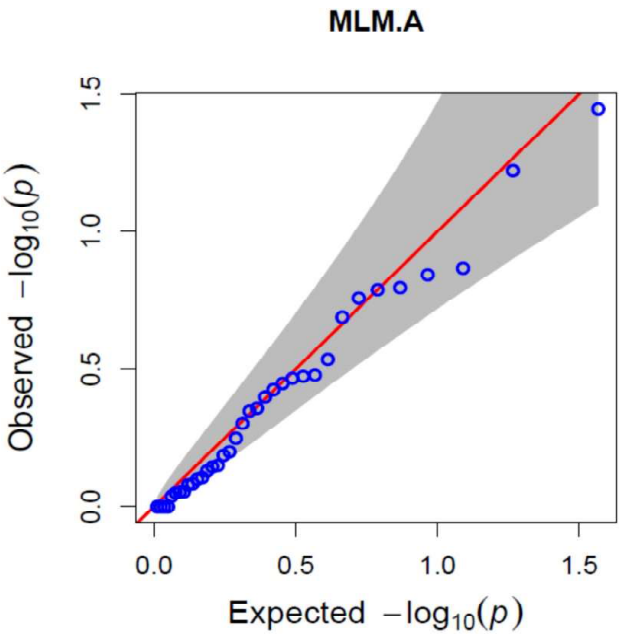

Fig. 4 : The QQ plot obtained for markers associated with phenotypic trait of total phenols

significantly associated with multiple traits like the total Fe, Mn, Zn and Cu content with a significant P-value of 0.02. The Supplementary Table 2 shows other associations and their significant P –values and R^2 values. However, a stronger association of specific SSR markers were detected for TPC, total flavonoids, SOD, chlorophyll a, Chlorophyll b, total chlorophyll, carotenoids, starch, amylose, total anthocyanin, gamma-oryzanol and total phenolics with the markers showing > 0.10 R^2 value in rice (Sahoo et al., 2020).

CONCLUSION

In the present study, 48 EST-SSR markers were used for studying the genetic diversity and the ESTs from which the SSRs were developed have been associated to certain traits and functions are annotated. Though, this study was done using 32 genotypes and 48 EST-SSR markers, the study can be extended further by the use of many more genotypes and markers.

ACKNOWLEDGEMENT

Authors are thankful to the ICAR-Indian Institute of Horticultural Research, Bangalore-560089, Karnataka, for the financial support for conducting this study.

REFERENCES

Benzie, I. F., & Strain, J. J. (1996). The ferric reducing ability of plasma (FRAP) as a measure of "antioxidant power": the FRAP assay. *Analytical biochemistry*, *239*(1), 70-76. https://doi.org/10.1006/abio.1996.0292

Christian, R. W., Hewitt, S. L., Nelson, G., Roalson, E. H., & Dhingra, A. (2020). Plastid transit peptides-where do they come from and where do they all belong? Multi-genome and pan-genomic assessment of chloroplast transit peptide evolution. *PeerJ*, *8*, e9772. Portico. https://doi.org/10.7717/peerj.9772

El-Awady, M. A., Hassan, M. M., El-Sayed, S. A. H., & Gaber, A. (2016). Comparison of the antioxidant activities, phenolic and flavonoids contents of the leaves-crud extracts of *Moringa peregrine* and *Moringa oleifera*. *International Journal of Biosciences*, *8*, 55-62. https://doi.org/10.12692/ijb/8.1.55-62

Fahey, J. W. (2005). *Moringa oleifera*: a review of the medical evidence for its nutritional, therapeutic, and prophylactic properties. Part 1. *Trees for Life Journal*, *1*(5), 1-15.

Han, Y., Li, X., Yan, Y., Duan, M. H., & Xu, J. H. (2020). Identification, characterization, and functional prediction of circular RNAs in maize. *Molecular Genetics and Genomics*, *295*, 491-503. https://doi.org/10.1007/s00438-019-01638-9

Hassan, F. A., Ismail, I. A., Mazrou, R., & Hassan, M. (2020). Applicability of inter-simple sequence repeat (ISSR), start codon targeted (SCoT) markers and ITS2 gene sequencing for genetic diversity assessment in *Moringa oleifera* Lam. *Journal of Applied Research on Medicinal and Aromatic Plants*, *18*, 100256. https://doi.org/10.1016/j.jarmap.2020.100256

Jackson, M. C., & Young, W. A. P. (1973). A capacitive integration system for the precise measurement of isotopic ratios in a mass spectrometer. *Review of Scientific Instruments*, *44*(1), 32-34. https://doi.org/10.1063/1.1685949

Jackson, K. W., Aldous, K. M., & Mitchell, D. G. (1973). A 500-channel silicon-target vidicon tube as a photodetector for atomic absorption spectrometry. *Spectroscopy Letters*, *6*(6), 315-321. https://doi.org/10.1080/00387017308065467

Kjeldahl, J. G. C. T. (1883). A new method for the estimation of nitrogen in organic compounds. *Zeitschrift für Analytische Chemie*, *22*(1), 366-382. https://doi.org/10.1007/bf01338151

Leone, A., Spada, A., Battezzati, A., Schiraldi, A., Aristil, J., & Bertoli, S. (2015). Cultivation, genetic, ethnopharmacology, phytochemistry and pharmacology of *Moringa oleifera* leaves: An overview. *International Journal of Molecular Sciences*, *16*(6), 12791-12835. https://doi.org/10.3390/ijms160612791

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., ... & Zhang, Z. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics*, *28*(18), 2397-2399. https://doi.org/10.1093/bioinformatics/bts444

Meghala, R., Ravishankar, K. V., Anand, L., & Rekha, A. (2005). Genetic diversity of Indian sapota

(*Manilkara zapota*) cultivars characterized by RAPD markers. *Plant Genetic Resources Newsletter (IPGRI/FAO)*, (142).

Moazzam-Jazi, M., Hedayati, V., & Moradi, S. (2021). Circular RNA in Rice (*Oryza sativa*). *Bioinformatics in Rice Research: Theories and Techniques*, 507-526. https://doi.org/10.1007/978-981-16-3993-7_22

Mulato, B. M., Möller, M., Zucchi, M. I., Quecini, V., & Pinheiro, J. B. (2010). Genetic diversity in soybean germplasm identified by SSR and EST-SSR markers. *Pesquisa Agropecuária Brasileira*, *45*, 276-283. https://doi.org/10.1590/s0100-204x2010000300007

Natarajan, S., & Joshi, J. A. (2015). Characterisation of moringa (*Moringa oleifera* Lam.) genotypes for growth, pod and seed characters and seed oil using morphological and molecular markers. https://doi.org/10.5958/2229-4473.2015.00039.7

Pan, T., Sun, X., Liu, Y., Li, H., Deng, G., Lin, H., & Wang, S. (2018). Heat stress alters genome-wide profiles of circular RNAs in Arabidopsis. *Plant molecular biology*, *96*, 217-229. https://doi.org/10.1007/s11103-017-0684-7

Pandey, A., Pradheep, K., Gupta, R., Nayar, E. R., & Bhandari, D. C. (2011). Drumstick tree (*Moringa oleifera* Lam.): a multipurpose potential species in India. *Genetic Resources and Crop Evolution*, *58*, 453-460. https://doi.org/10.1007/s10722-010-9629-6

Poornima, K. N., Shankar, R., Ramesh, S., & Ravishankar, K. V. (2023). De-novo development and validation of EST-SSRs in *Moringa oliefera*. *Journal of Plant Biochemistry and Biotechnology*, *32*(2), 319-327. https://doi.org/10.1007/s13562-022-00807-z

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945-959. https://doi.org/10.1093/genetics/155.2.945

Ravi, R. D., Siril, E. A., & Nair, B. R. (2020). The efficiency of Cytochrome P450 gene-based markers in accessing genetic variability of drumstick (*Moringa oleifera* Lam.) accessions. *Molecular Biology Reports*, *47*, 2929-2939. https://doi.org/10.1007/s11033-020-05391-w

Ravishankar, K. V., Anand, L., & Dinesh, M. R. (2000). Assessment of genetic relatedness among mango cultivars of India using RAPD markers. *The Journal of Horticultural Science and Biotechnology*, *75*(2), 198-201. https://doi.org/10.1080/14620316.2000.11511223

Ravishankar, K. V., Bommisetty, P., Bajpai, A., Srivastava, N., Mani, B. H., Vasugi, C., ... & Dinesh, M. R. (2015). Genetic diversity and population structure analysis of mango (*Mangifera indica*) cultivars assessed by microsatellite markers. *Trees*, *29*, 775-783. https://doi.org/10.1007/s00468-015-1155-x

Sahoo, A., Behura, S., Singh, S., Jena, S., Ray, A., Dash, B., ... & Nayak, S. (2021). EST-SSR marker-based genetic diversity and population structure analysis of Indian *Curcuma* species: significance for conservation. *Brazilian Journal of Botany*, *44*, 411-428. https://doi.org/10.1007/s40415-021-00711-1

Sahoo, S., Sanghamitra, P., Nanda, N., Pawar, S., Pandit, E., Bastia, R., ... & Pradhan, S. K. (2020). Association of molecular markers with physio-biochemical traits related to seed vigour in rice. *Physiology and Molecular Biology of Plants*, *26*, 1989-2003. https://doi.org/10.1007/s12298-020-00879-y

Saini, R. K., Saad, K. R., Ravishankar, G. A., Giridhar, P., & Shetty, N. P. (2013). Genetic diversity of commercially grown *Moringa oleifera* Lam. cultivars from India by RAPD, ISSR and cytochrome P 450-based markers. *Plant Systematics and Evolution*, *299*, 1205-1213. https://doi.org/10.1007/s00606-013-0789-7

Serbessa Tolera, B., Dagne Woldegebriel, K., Teshome Gari, A., Geleta Dida, M., & Tesfaye Geletu, K. (2021). Analyses of genetic diversity and population structure of anchote (*Coccinia abyssinica* (Lam.) Cogn.) using newly developed EST-SSR markers. *Genetic Resources and Crop Evolution*, *68*, 2337-2350. https://doi.org/10.1007/s10722-021-01132-5

Swat, Bhardwaj, A., Attri, C., & Kulshrestha, S. (2020). Evaluating the genetic diversity of *Moringa oleifera* obtained from south-central and northern states of India using RAPD markers technique, *Asian Journal of Plant Sciences*, *19*(1), 1-7. https://doi.org/10.3923/ajps.2020.1.7

Singleton, V. l. (1999). Analysis of total phenols and other oxidation substrates and antioxidants by means of Folin-Ciocalteu reagent. *Methods in Enzymology*, *299*, 152-178. https://doi.org/10.1016/s0076-6879(99)99017-1

Wen, M., Wang, H., Xia, Z., Zou, M., Lu, C., & Wang, W. (2010). Development of EST-SSR and genomic-SSR markers to assess genetic diversity in *Jatropha Curcas* L. *BMC research notes*, *3*, 1-8. https://doi.org/10.1186/1756-0500-3-42

Xu HuaWei, X. H., & Hou DianYun, H. D. (2018). Cloning and functional verification of tomato Rubisco small subunit chloroplast transit peptide. *Genomics and Applied Biology*, *37*(4), 1570-1575.

Zhishen, J., Mengcheng, T., & Jianming, W. (1999). The determination of flavonoid contents in mulberry and their scavenging effects on superoxide radicals. *Food Chemistry*, *64*(4), 555-559. https://doi.org/10.1016/s0308-8146(98)00102-2

(***Received : 20.1.2024; Revised : 20.7.2024; Accepted : 23.7.2024)***