IECAG 2025 Conference
The 5th International Electronic Conference on Agronomy
15–18 December 2025 | Online
MDPI

Characterization of EST-SSR markers associated with oil biosynthesis in Castor (*Ricinus communis* L.)

Pranjal Patil[1], Rajitha Nair[1], Harshvardhan Zala[1], Ankit M. Patel[2], K. N. Prajapati[1], and S. D. Solanki[1]

[1]Department of Genetics and Plant Breeding, C. P. College of Agriculture, S. D. Agricultural University, Sardarkrushinagar-385 506, Gujarat, India

[2]Center for Oilseeds Research, S. D. Agricultural University, Sardarkrushinagar-385 506, Gujarat, India

INTRODUCTION

- The monotypic species Castor (*Ricinus communis* L.) is a member of the Euphorbiaceae family (2n = 20).
- Grown extensively in dry and semi-arid areas, it is a significant non-edible oilseed crop.
- Among oilseed crops, castor seeds have the highest oil content (40–55%).
- It is the most valuable natural industrial oil in terms of commercial value because it contains more than 85% ricinoleic acid.
- Castor has employed DNA markers such as RAPD, AFLP, RFLP, and SSR; nonetheless, SSR markers tend to be preferred because:
  - Co-dominance
  - High polymorphism
  - Genome-wide distribution
  - Genome-wide distribution
  - High reproducibility
- Expressed Sequence Tags, or ESTs, are great resources for gene-based SSR markers since they reflect expressed genes.

MATERIAL AND METHODS

1. Plant Material & DNA Extraction

•Young leaves of **24 diverse castor genotypes** were collected.

•Genomic DNA was isolated using the **CTAB method (Doyle & Doyle, 1990)**.

•DNA quality was checked on **0.8% agarose gel** and quantified using a **BioSpectrometer**.

•DNA samples were diluted to a working concentration of **20 ng/µL**.

2. EST Data Mining & Processing

•**62,105 ESTs** of castor were downloaded from the **NCBI database**.

•Sequences were cleaned by removing **low-complexity regions**, **poly-A/T tails**, and **low-quality (<100 bp)** ends.

•**Vector sequences** were removed using **UniVec + CrossMatch** (EGassembler).

3. Functional Annotation

•Contigs and singletons were annotated using **Blast2GO**.

•Sequences were mapped and assigned **GO terms** for molecular function, biological process, and cellular component.

4. EST–SSR Identification

•SSRs were mined using **MISA** with the following repeat criteria:

•**Di–hexa nucleotide repeats:** (2–6), (3–5), (4–4), (5–3), (6–3)

•Maximum interruption between repeats: **100 bp**.

5. Primer Design (EST-SSR Development)

Primers were designed using BatchPrimer3 with the following parameters: product size 100–250 bp, primer length 18–27 bp, melting temperature 57–63°C, GC content 40–60%, and a maximum Tm difference of 1.5°C between primer pairs.

6. PCR Amplification & Gel Analysis

PCR was carried out in a 10 µL reaction containing 1× buffer, 2.5 mM Mg^{2+}, 0.2 mM dNTPs, 10 pmol primers, 1 U Taq polymerase, and 20 ng DNA, using a touchdown program (94°C for 3 min; 5 cycles of 94°C 30 s, 65→61°C 30 s, 72°C 1 min; 30 cycles at 60°C annealing; final extension at 72°C for 5 min), and the products were separated on a 3% agarose gel to assess polymorphis.

SCIENTIFIC RECOMMENDATION

It is recommended to scientific community involved in castor improvement is utilize newly developed EST-SSR markers associated with oil synthesis/fatty acids biosynthesis related traits for genetic improvement/ marker assisted breeding of castor genotypes.

RESULT AND DISCUSSION

Non-redundant sequence assembly

- From 62105 ESTs, 13811 non-redundant unigene sequences were identified. Out of 13811 unigene, 1955 (14.15%) sequences harbored SSR motifs. A total of 2425 SSRs were identified and 374 sequences harbored more than one SSR.

Distribution of EST-SSR repeat types

- Depending upon the number of nucleotides per repeat unit, SSRs were classified as di-, tri, tetra, penta- or hexanucleotides.
- The tri-nucleotide repeat motif was found to be the most abundant (1319) and accounted for 54.4% of SSRs, followed by di- (1018, 42%) repeats (Table 1).
- The AG/CT (53.7%) di-nucleotide repeat was the most abundant motif type detected followed AC/GT (23.1%), AT/AT and CG/CG. The AGC/CTG (34.5%) tri-nucleotide repeat was the most abundant motif type detected followed AGG/CCT (25.2%), ACC/GGT (13.8%) and ACG/CGT (12.4%) motif types.
- Out of the total 1955 sequences harbored SSR motifs subjected to BlastX with non-redundant database, 1138 sequences were annotated.

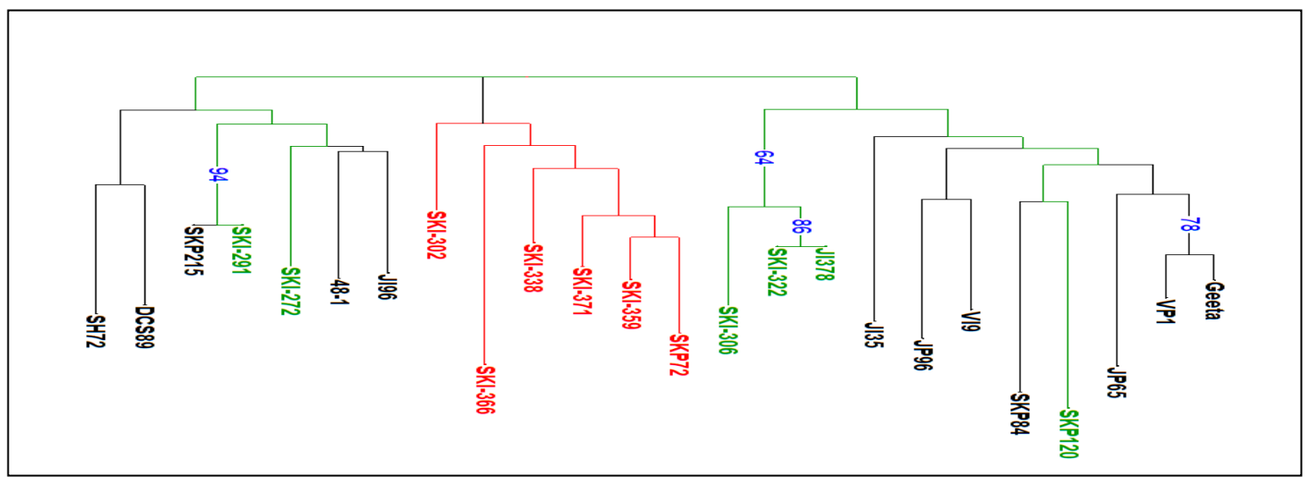
Table 1: Distribution and frequencies of SSR repeat types with repeat numbers

Motif length	Repeats number								
	5	6	7	8	9	10	>10	Total	%
Di-nucleotide	-	270	170	127	92	72	287	1018	42.0%
Tri-nucleotide	665	306	167	90	48	18	25	1319	54.4%
Tetra-nucleotide	36	10	4	5	1	-	1	57	2.4%
Penta-nucleotide	4	1	1	-	1	-	-	7	0.3%
Hexa-nucleotide	14	8	2	-	-	-	-	24	1.0%
Total	719	595	344	222	142	90	313	2425	-
%	29.6	24.5	14.2	9.2	5.9	3.7	12.9	-	-

Validation of EST-SSRs

- A total of 30 SSR markers were validated on 12 parental genotypes with six high and six low oil content genotypes.
- Out of that 15 displayed desired amplification of defined product size.
- Ten markers were polymorphic, generating a total of 25 polymorphic and 5 monomorphic loci, with an average of 2.5 alleles per locus.
- PIC values of polymorphic markers ranged from 0.239 to 0.454, indicating moderate informativeness.
- A total of 25 polymorphic loci and 5 monomorphic loci with 2.5 average numbers of alleles were detected.

Fig 1: Dendrogram depicting the genetic relationship among 24 castor genotypes based on EST-SSR markers



CONCLUSION

- The newly developed EST derived SSR markers based on functional annotation of sequences have added to the repository of molecular markers for Castor.
- Total 10 primers showed polymorphic patterns, and revealed genetic relationships among 24 castor genotypes.
- These markers can be used for genetic improvement and diversity study associated with oil content traits in castor genotypes.