Plastid DNA Barcoding and *RtActin* cDNA Fragment Isolation of *Reutealis Trisperma*: A Promising Bioresource for Biodiesel Production

Bioinformatics and Biology Insights
Volume 17: 1–8

DOI: 10.1177/11779322231182768

Nurul Jadid[1], Nur Laili Alfina Rosidah[1], Muhammad Rifqi Nur Ramadani[1], Indah Prasetyowati[1], Noor Nailis Sa'adah[1], Aulia Febrianti Widodo[1] and Dwi Oktafitria[2]

[1]Department of Biology, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia.
[2]Department of Biology, Universitas PGRI Ronggolawe, Tuban, Indonesia.

ABSTRACT: *Reutealis trisperma* belonging to the family *Euphorbiaceae* is currently used for biodiesel production, and rapid development in plant-based biofuel production has led to its increasing demand. However, massive utilization of bio-industrial plants has led to conservation issues. Moreover, genetic information on *R trisperma* is still limited, which is crucial for developmental, physiological, and molecular studies. Studying gene expression is essential to explain plant physiological processes. Nonetheless, this technique requires sensitive and precise measurement of messenger RNA (mRNA). In addition, the presence of internal control genes is important to avoid bias. Therefore, collecting and preserving genetic data for *R trisperma* is indispensable. In this study, we aimed to evaluate the application of plastid loci, *rbcL*, and *matK*, to the DNA barcode of *R trisperma* for use in conservation programs. In addition, we isolated and cloned the *RtActin* (*RtACT*) gene fragment for use in gene expression studies. Sequence information was analyzed *in silico* by comparison with other *Euphorbiaceae* plants. For actin fragment isolation, reverse-transcription polymerase chain reaction was used. Molecular cloning of *RtActin* was performed using the pTA2 plasmid before sequencing. We successfully isolated and cloned 592 and 840 bp of *RtrbcL* and *RtmatK* fragment genes, respectively. The *RtrbcL* barcoding marker, rather than the *RtmatK* plastidial marker, provided discriminative molecular phylogenetic data for *R Trisperma*. We also isolated 986 bp of *RtACT* gene fragments. Our phylogenetic analysis demonstrated that *R trisperma* is closely related to the *Vernicia fordii Actin* gene (97% identity). Our results suggest that *RtrbcL* could be further developed and used as a barcoding marker for *R trisperma*. Moreover, the *RtACT* gene could be further investigated for use in gene expression studies of plant.

KEYWORDS: Actin, DNA barcoding, ***Euphorbiaceae***, gene expression, renewable energy, *Reutealis trisperma*

RECEIVED: November 10, 2022. **ACCEPTED:** May 30, 2023.

TYPE: Original Research Article

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors are indebted to the Directorate of Research and Community Services (DRPM) of the Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia for providing financial support under the PAKERTI scheme (1658/PKS/ITS/2022).

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Nurul Jadid, Department of Biology, Institut Teknologi Sepuluh Nopember (ITS), Surabaya 60111, Indonesia. Email: nuruljadid@bio.its.ac.id

Introduction

Perennial *Reutealis trisperma*, also known as *Kemiri sunan* or Philippine Tung, is a member of the family *Euphorbiaceae*. It is native to the Philippines; however, became prominent in Southeast Asia eventually. This woody and non-edible plant has been introduced in several Asian and South American regions, including Indonesia, Malaysia, China, India, Cuba, and Dominica.[1] Although it is reported to grow naturally in humid areas, *R trisperma* is also known to be drought-tolerant. Therefore, it also grows well in critical or unfavorable areas.[2] Together with other tropical plants,[3] *R trisperma* is known as a medicinal plant. Tigliane diterpenoids isolated from *R trisperma* have exhibited anti-HIV properties.[4] In addition, the potential of *R trisperma* as a bioresource for biodiesel production have been demonstrated.[5,6] The kernels are extracted for their oil and produce 50 to 52 wt%.[5] Furthermore, the conversion of *R trisperma* oil to biodiesel through esterification and transesterification has been reported to meet the ASTM standard D6751.[5]

Massive utilization of fossil fuels for transportation, industry, and many anthropogenic activities contributes to environmental issues, including air pollution, human health damage, and global climate change.[7,8] Therefore, the development of alternative renewable energy sources is promising. This could reduce global carbon emissions and might be sustainable as renewable bioresources.[9] Some non-edible plant-based bioenergy have been studied, including *Jatropha curcas*,[10] *Ceiba pentandra*,[11] *Calophyllum inophyllum*,[12] *Azadirachta indica*,[13] *Hevea brassiliensis*,[14] *Melia azedarach*,[15] *Nicotiana tabacum*,[16] *Pongamia pinnata*,[17] *Sapindus mukorossi*,[18] *Simmondsia chinensis*,[19] *Madhuca indica*,[20] *Datura stramonium*,[21] *Carberra odollam*,[22] and *Silybum marianum*.[23] Recently, the potential of various tropical and non-edible plants, such as *R trisperma*, for biodiesel production has been reported.[5] However, exploitation of the Philippine Tung plant as energy bio-sources risks its habitat loss. As a decline in the *R trisperma* population worldwide has been reported, it is listed as a vulnerable plant.[24] Therefore, conservation efforts are necessary. Species identification is an important step in conserving vulnerable plants and strengthening species conservation regulations.[25] A promising method to identify plant species is through DNA barcoding, which requires genetic information of the plant. As genomic information on *R trisperma* is still limited,[26] developing barcoding markers for this species is necessary. Some molecular markers

have been used for barcoding plant DNA, including *rbcL, matK, trnH-pbsA*, ITS, and ITS2. However, the Consortium for the Barcode of Life (CBOL) suggested *matK* and *rbcl* as the preferred plant barcoding region.[27] A suitable locus for plant DNA barcoding depends on the taxonomic group of the plants.[28] In addition, the discriminative power of the markers varies across plant species.[29] Therefore, in this study, we aimed to evaluate the use of both plastid *rbcL* and *matK* loci as plant DNA barcoding markers for *R trisperma.*

Despite its conservation issues, *R trisperma*, as a new source of biodiesel, has received much attention in agronomy and bioenergy-related gene expression studies. Quantification of gene expression might provide information on the activity of biodiesel-related genes or other biological activities and mechanisms, for which quantitative PCR (qPCR) is considered a reliable method. However, certain parameters should be carefully considered, including the concentration of RNA samples, PCR components, and complementary DNA (cDNA) synthesis and efficiency of the reference genes.[30] Reference genes are used to reduce quantification errors by normalizing the targeted gene expression with constitutive expression of the internal control.

Suitable reference genes for gene expression studies in biofuel plants have been screened. Zhang et al[31] suggested the use of *actin* and *TUB8* as reliable reference genes for gene expression studies in *Jatropha curcas*. Karuppaiya et al[32] also validated some reference genes, including *Actin*, Tubulin beta-5 chain (*TUB*), Elongation factor 1-alpha (*GAPDH1*), and Ubiquitin-conjugating enzyme E2 32 (*UBI*). Nonetheless, to date, no study has reported a search for a reference gene in *R trisperma*. Therefore, this study also aimed to isolate an *Actin* fragment gene as a candidate reference gene. Although the effectiveness of the *Actin* housekeeping genes as reference genes should be further elucidated, this study might serve as a preliminary study for further gene expression profiling in *R trisperma*.

Housekeeping genes are frequently used as reference controls. This is mainly owing to their abundance in eukaryotic cells, and their continuous expression at every stage of development and in all tissues.[33] Some of these encode actin protein, tubulin, GAPDH, and 18S ribosomal RNA (rRNA). However, specific and optimal reference genes often depend on the species, tissue type, and treatments.[34,35]

Actin protein, a globular structure with a mass of approximately 42 kDa, is found in abundance and highly conserved in almost all higher organisms.[36] These proteins play an important role in forming tissues and providing mechanical support for cells, thereby determining cell shape, cell movement, and cell division.[37,38] Actin is also important in plants for cell morphogenesis, cell elongation, as a cell wall component, and for the growth of root hairs, trichome cells, pollen tubes, and apical meristems.[39] Actin genes have been reported to be expressed stably in various tissues at all developmental stages.[40] Some gene expression studies in potatoes,[41] soybeans,[42] rice,[43] and in leaves, flowers, stems, and roots of *Carthamus tinctorius*[44] used the *actin* gene as the reference standard in qPCR analysis. Nevertheless, the *R trisperma Actin* has not been isolated and is not yet available in the database. This study was conducted to perform DNA barcoding of *R trisperma* using the *rbcL* and *matK* genes, as well as to isolate the *R trisperma Actin*-coding gene fragment. This study might serve as a preliminary study for selecting reliable reference standards for further transcriptomic studies.

Materials and Methods

Plant materials

R trisperma used in this study was provided by the Purwodadi Botanical Garden of the National Research and Innovation Agency, Indonesia. Fresh young *R trisperma* leaves were collected *in situ* from the botanical garden and stored at –20°C for genomic DNA (gDNA) isolation and total RNA extraction. All fresh samples were ground in liquid nitrogen before being subjected to gDNA and total RNA extraction.

Data collection of Actin and primer design

Actin gene sequence data were retrieved from the National Center for Biotechnology Information database (https://www.ncbi.nlm.nih.gov/genbank/). The selected members of the *Euphorbiaceae* family included *Ricinus communis* (AY360221.1), *Jatropha curcas* (HM044307.1), *Vernicia fordii* (JQ680035.1), and *Hevea brasiliensis* (JF775488.1). The sequence data of the four species were downloaded in the FASTA format, and sequence alignment was performed using MEGA11. The DNA alignment results showing conserved regions were used as the basis for designing the primers (Suppl. 1). Both forward and reverse primers contained 20 to 24 bp selected from the conserved region. The selected primers were then analyzed for secondary structure, primer dimer, melting temperature, and GC content using primer analysis and Tm Calculator tool (https://sg.idtdna.com/calc/analyzer).

Nucleic acid extraction and first-strand cDNA synthesis

gDNA extraction was performed to obtain nuclear, chloroplast, and mitochondrial DNA for *R trisperma* DNA barcoding analysis. gDNA extraction was carried out using ZR™Plant/Seed DNA MicroPrep (Zymo Research) as per the manufacturer's protocol. The samples used for gDNA extraction were obtained from 150 mg of *R trisperma* leaves, which were ground using liquid nitrogen prior to extraction.[45] Meanwhile, we used Plant Total RNA Mini Kit (Genaid) to isolate the total RNA from *R trisperma* seeds, as described by Jadid et al.[46] The plant endosperm was prepared and ground in liquid nitrogen. We measured the concentration and quality of the DNA and total RNA extracts using NanoDrop™. The A_{260}/A_{280} ratio was used to evaluate the quality of the DNA and RNA extracts.

Approximately 0.2 µg of total RNA was used as a template for cDNA synthesis using ReverTra Ace™ qPCR RT (Toyobo) according to the kit's instructions.

DNA amplification and sequencing

DNA barcoding was carried out by amplification of the extracted gDNA using universal primers and TOYOBO KOD FX Neo (TOYOBO). The primers used included *RbcL forward* (5′ ➔ 3′): ATGTCACCCACCAACAGAGACT AAAGC, *RbcL reverse* (5′ ➔ 3′): GTAAAATCAAGTCCA CCRCG, *matK forward* (5′ ➔ 3′): ACCCAGTCCATC TGGAAATCTTGGTTC, and *matK reverse* (5′ ➔ 3′): CGTACAGTACTTTTGTGTTTACGAG. The amplification parameters consisted of an initial denaturation at 95°C for 3 min, followed by 35 cycles of denaturation at 98°C for 10 s, primer annealing at 53°C for 30 s, and extension at 68°C for 45 s. A final extension was performed at 68°C for 5 min to complete the PCR process.

Reverse transcription-polymerase chain reaction (RT-PCR) amplification of the *Actin* gene was carried out according to the manufacturer's instructions using GoTaq® Green Master Mix (Promega). *Actin* fragment isolation was performed by using PCR. The samples were denatured initially at 95°C for 2 min, followed by 35 cycles of amplification, which comprised denaturation at 95°C for 30 s, annealing at 55°C for 30 s, and extension at 72°C for 90 s, with final extension at 72°C for 5 min. The primers used for amplification were 10ACTks (5′-GCTGGGTTT GCTGGTGATGATG-3′) and 15ACTKs (5′-GCAAGGATA GATCCTCCAATCC-3′). The PCR product was separated via DNA electrophoresis on 1% polyacrylamide gels at 20 volts/cm^3 for approximately 30 min. DNA purification was performed using Zymoclean™ Gel DNA Recovery Kit (Zymo Research). Finally, *RtrbcL, RtmatK*, and *Actin* fragments were sequenced.

Molecular gene cloning of Actin fragment

The purified *Actin* fragment was cloned into a pTA2 vector (Toyobo, TAK-101). Subsequently, the pTA-*Actin* plasmid was integrated into *Escherichia coli* DH5α competent cells (Zymo Research, T3007).[47] The transformed bacteria were then cultured in Luria Bertani (LB) medium supplemented with ampicillin. Blue-white bacterial screening was performed to select bacterial colonies harboring *Actin* fragments. Eight bacterial colonies were randomly selected for colony PCR analysis using T3 and T7 promoter primers (KOD FX Neo [Toyobo]).[48] The expected bacterial colonies were cultured in an LB liquid medium, followed by plasmid extraction.

DNA sequencing

DNA sequencing was performed using an ABI PRISM® BigDye™ Terminator Cycle Sequencing Kit (Applied Biosystems). Sequencing of *the rbcL* and *matK* genes fragments was performed using the universal primers rbcL and matK. Plasmid sequencing was carried out at using T7 promoter primer (5′-TAATACGACTCACTATAGGG-3′) and T3 promoter primer (5′-AATTAACCCTCACTAAAGGG-3′).

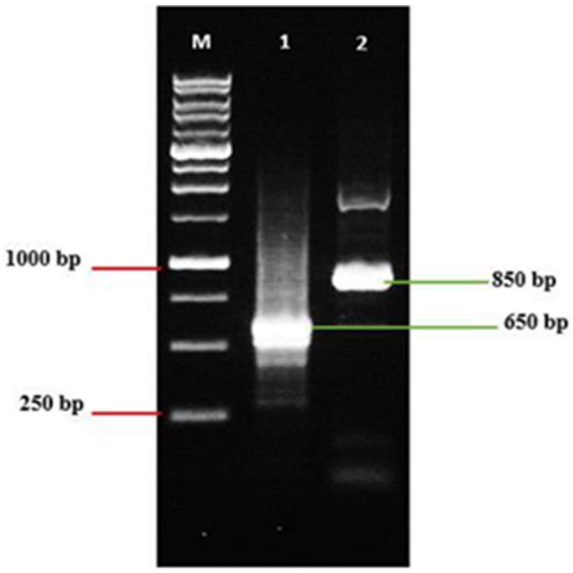

Figure 1. Electrophoregram of PCR products of *rbcL* and *matK Reutealis trisperma* gene fragments.
The PCR product was electrophoresed through 1% polyacrylamide gels.
1, indicates *rbcL* gene fragment; 2, *matK* gene fragment; M, marker; PCR, polymerase chain reaction.

Phylogenetic analysis

The nucleotide bases from the sequencing were then evaluated using nucleotide BLAST (https://blast.ncbi.nlm.nih.gov/) to determine gene similarity within the tested plant species. The nucleotide sequences obtained from BLAST analysis were downloaded and subsequently aligned with the *Actin* (*ACT) R trisperma* gene using the ClustalW Alignment algorithm (BioEdit v7.0.5). The phylogenetic tree construction was compiled using the MEGA11 application with neighbor-joining (NJ) and a 1000× bootstrapping value.

Results and Discussion

DNA barcoding using R trisperma rbcL and matK genes

The *rbcL* and *matK* gene fragments were detected as 650 and 850 bp, respectively, (see Figure 1). Subsequently, we purified the DNA via gel extraction according to the expected DNA size before performing DNA sequencing. The purification process washes DNA from remnants of salt, protein, and polyphenolic compounds.[49]

The sequencing results were analyzed for gene sequence homology using BLAST (see Table 1). BLAST analysis showed an *E*-value of 0 for some *Euphorbiaceae* members, indicating high similarity of *rbcL* gene among this plant family. The resulting query coverage value was 97% to 100%, with a similarity (Identity) value of 97% to 99%, compared with other

Table 1. BLAST analysis of *Reutealis trisperma rbcL* gene fragments compared with other *Euphorbiaceae rbcL* gene sequences.

NO.	ACCESSION NUMBER	SPECIES DESCRIPTION	QUERY COVER, %	*E*-VALUE	IDENTITY, %
1	KY628420.1	*V fordii*	99	0	99
2	FJ695500.1	*J curcas*	99	0	99
3	AY794889.1	*C aurea*	97	0	99
4	KM273035.1	*T sebifera*	99	0	97
5	KF523366.1	*C persimilis*	100	0	97
6	MF981856.1	*H brasiliensis*	99	0	92

Table 2. BLAST analysis of *Reutealis trisperma matK* gene fragments compared with other *Euphorbiaceae matK* gene sequences.

NO.	ACCESSION NUMBER	SPECIES DESCRIPTION	QUERY COVER, %	*E*-VALUE	IDENTITY, %
1	KR073223.1	*C maxima*	100	0	99
2	LC218435.1	*C depressa*	99	0	99
3	AB626776.1	*C junos*	99	0	99
4	AB626749.1	*C unshiu*	99	0	99
5	FJ716731.1	*S buxifolia*	99	0	98
6	AB762385.1	*N crenulata*	99	0	97
7	KU556679.1	*Z officinale*	90	0	99

Euphorbiaceae species. The high similarity value indicated high homology of the *R trisperma rbcL* sequence with *rbcL* from other members of the *Euphorbiaceae* family. The highest and lowest similarity values were with the *V. fordii rbcL* gene (accession number KY628420.1, 99%) and *H. brasiliensis rbcL* gene (accession number MF_981856.1, 92%), respectively.

The *matK* gene sequence was analyzed using BLAST. It showed results identical to several species from the Rutaceae and Zingiberaceae families (see Table 2), indicated by an *E*-value of 0. The resulting query coverage value ranged from 90% to 100%, with a resulting similarity value of 97% to 99%. These results indicate that the *matK* nucleotide sequence of *R trisperma* is highly homologous to those of several species from the Rutaceae and Zingiberaceae families. The highest and lowest similarity values were shown to the *Citrus maxima matK* gene (accession number KR073223.1, 99%) and *Naringa cranulata matK* gene (accession number AB762385.1, 97%), respectively.

The *rbcL* and *matK* phylogenetic trees for DNA barcoding were constructed using MEGA11, which describes the evolutionary relationship between sample organisms.[50] Using the NJ algorithm, we compared the *R trisperma* sequence with other organisms from the BLAST results. The NJ method best estimates the branch length that most closely reflects the distance between sequences.[51]

Two main clusters were identified in the phylogenetic tree analysis of the *rbcL* gene (see Figure 2). The first cluster comprised species from the *Euphorbiaceae* family, namely, *RtrbcL, VfrbcL, CarbcL, TsrbcL, CprbcL, HbrbcL*, and *JcrbcL*. The second

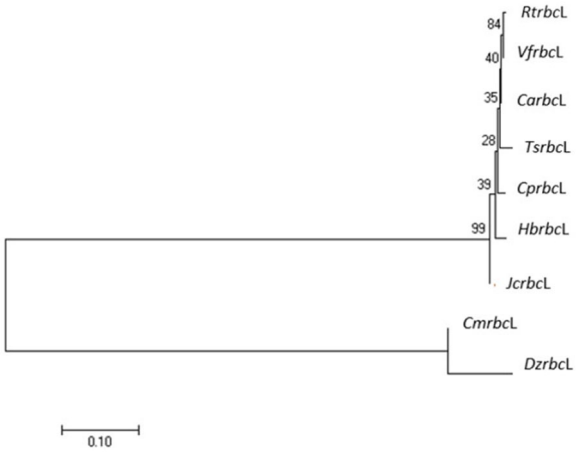

Figure 2. Phylogenetic tree of RtrbcL with several species of the family *Euphorbiaceae*.
CarbcL indicates *rbcL Cavacoa aurea; CmrbcL, rbcL Citrus maxima; CprbcL, rbcL Croton persimilis; DzrbcL, rbcL Durio zibethinus; HbrbcL, rbcL Hevea brasiliensis; JcrbcL, rbcL Jatropha curcas; TsrbcL, rbcL Triadica sebifera; VfrbcL, rbcL Vernicia fordii*.

cluster comprised *DzrbcL* and *CmrbcL*. Our data indicate that *RtrbcL* belongs to a monophyletic group with other *Euphorbiaceae* families. This result is supported by the unique characteristics of the *Euphorbiaceae* family, which contains white to red latex and has a box-shaped fruit.[1,52]

A comparison of the *R trisperma matK* gene with those of other *Euphorbiaceae* species was also displayed in the constructed phylogenetic tree based on BLAST analysis (see Table 2). To construct the tree, additional species from the *Euphorbiaceae*

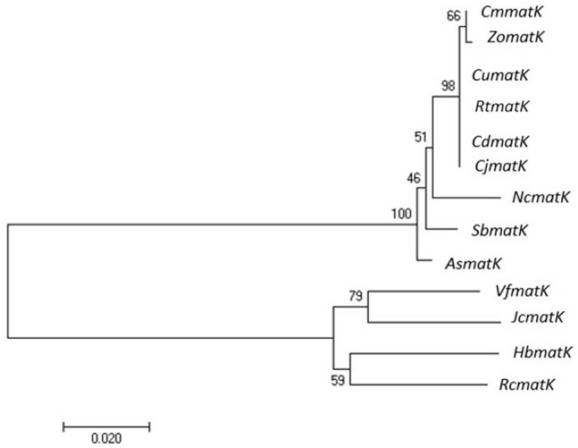

Figure 3. Phylogenetic tree *of RtmatK* with several *Euphorbiaceae* species.
AsmatK indicates matK *Atalantia* spinosa; CdmatK, matK *Citrus* depression; CjmatK, matK *Citrus* junos; CmmatK, matK *Citrus* maxima; CumatK, matK *Citrus* unshiu; HbmatK, matK *H*evea brasiliensis; JcmatK, matK *Jatropha* curcas; NcmatK, matK *Naringi* crenulata; RcmatK, matK *Ricinus* communis; SbmatK, matK *Severinia* buxifolia; VfmatK, matK *Vernicia* fordii; ZomatK, matK *Zingiber* officinale.

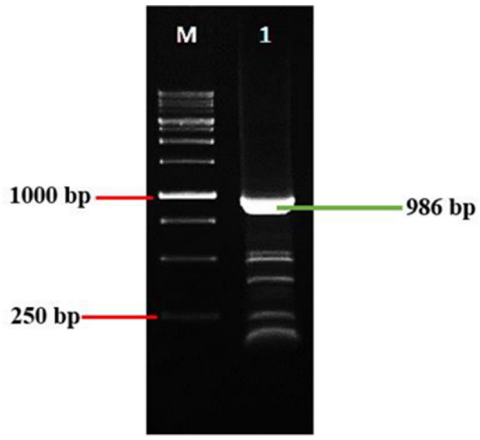

Figure 4. Amplification of *Reutealis trisperma Actin* fragment.
The *Actin* fragment was electrophoresed using 1% polyacrylamide gels. 1 indicates *R trisperma* actin-coding gene fragment; M, marker 10 kb.

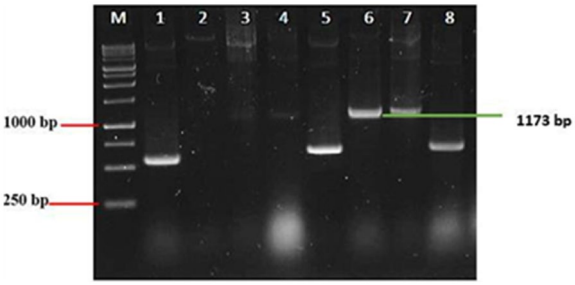

Figure 5. Validation of pTA-ACT plasmid using PCR colony.
ACT indicates *actin*; Lane 1-8, results of colony PCR from eight randomly selected bacterial colonies; PCR, polymerase chain reaction.

family were added, namely, *V fordii* (accession number NC_034803.1), *Jatropha curcas* (accession number NC_012224.1), *H brasiliensis* (accession number HQ606140.1), and *R communis* (accession number AB233767.1). The *matK* phylogenetic tree analysis (see Figure 3) revealed two main clusters. The first cluster comprised *RtmatK, CmmatK, ZomatK, CumatK, CdmatK, CjmatK, NcmatK, SbmatK*, and *AsmatK*. In contrast, the second cluster comprised *matK* from species groups of the other *Euphorbiaceae* family— *VfmatK, JcmatK, HbmatK, and RcmatK*. *RtmatK* is in a cluster with several species from the *Rutaceae* and *Zingiberaceae* families.

Some DNA barcoding studies have used multi-locus DNA barcoding markers, such as intergenic regions (*trnH-pbsA*), nuclear ribosomal ITS, or other chloroplast genomes (*rpoC1* and *rpoB*). However, CBOL has described *matK* and *rbcl* as suitable plant barcoding markers.[27] The above results revealed that the *rbcL* gene provides a more informative description of the molecular phylogenetics of *R trisperma*. The *rbcL* gene sequence has a better discrimination ability than the *matK* gene sequence, favoring molecular identification. Similar results were also described in the plant DNA barcoding study performed by Maloukh et al,[53] where *matK* was also unable to discriminate against plant species native to the United Arab Emirates. Furthermore, Barthet and Hilu[54] stated that the *matK* gene underwent nucleotide substitution three times more often than the *rbcL* gene.

Isolation and molecular cloning of Actin gene from R trisperma

We successfully isolated *the Actin* fragment from *R trisperma* using RT-PCR. The isolated fragment was 986 bp in size (see Figure 4). The PCR product was gel extracted and purified. The purpose of purification was to remove excess primers, dNTPs, enzymes, and short DNA amplicons resulting from incomplete polymerization, dyes, and salt residues in the PCR reaction.[55] The purification results were then used as DNA inserts in the ligation process. Ligation of the *Actin* gene was performed by combining the DNA fragments with a pTA2 cloning vector, resulting in a recombinant plasmid of 3967 bp. Further steps comprised transforming the pTA2 harboring the *Actin* fragment into DH5α *E. coli* competent cells, followed by inoculation of the transforming bacteria on a solid LB medium supplemented with ampicillin.

Positive transformants were selected using the blue-white screening method on LB agar supplemented with ampicillin, isopropyl β-D-1-thiogalactopyranoside (IPTG), and 5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside (X-gal). *E. coli* harboring the recombinant plasmid was indicated by white colonies.[56] These white colonies were further confirmed using a colony PCR. Eight white colonies were selected for colony PCR using the T3 promoter and T7 promoter primers (see Figure 5). Colony 6 had a matching target size (1173 bp). We conducted further plasmid extraction of colony 6 bacteria, followed by DNA sequencing. Our sequencing data confirmed the 986 bp amplicon size of *Actin* fragment. This size is sufficient to be used as a basis for quantitative real-time PCR (qRT-PCR) primers. As previously reported, the amplicon length for qRT-PCR is generally recommended to be below

Table 3. BLAST analysis of *Reutealis trisperma ACT* gene fragments compared with other *Euphorbiaceae ACT* gene sequences.

NO	ACCESSION NUMBER	SPECIES DESCRIPTION	QUERY COVER, %	*E*-VALUE	IDENTITY, %
1	JQ680035.1	*V fordii*	99	0	97
2	XM_021835715.1	*H brasiliensis*	99	0	94
3	KY656700.1	*T sebifera*	99	0	94
4	NM_001308728.1	*J curcas*	99	0	94
5	XM_021777092.1	*M esculenta*	99	0	94
6	XM_002522148.3	*R communis*	99	0	93
7	KC182753.1	*E lathyris*	99	0	92
8	XM_024058244.1	*Q suber*	99	0	91
9	XM_022883610.1	*D zibethinus*	99	0	90
10	NM_001327089.1	*G hirsutum*	99	0	90
11	KY382441.1	*H cannabinus*	99	0	89

Abbreviation: ACT, *actin*.

150 to 200 bp.[57] Some housekeeping genes have been used as reference genes or internal standards for gene expression analysis.[58] Among those genes, previous studies have demonstrated that *Actin* (*ACT*) might serve as the most suitable candidate. A study on the effect of abiotic stress in *Daucus carota* L. showed that *Actin* demonstrated good and stable performance for qPCR analysis in heat, cold, and methyl jasmonate treatment.[59] Moreover, Jadid et al[60] used *Actin* as the internal standard during gene expression profiling in different vegetative organs of *Jatropha curcas*. Similar results were reported by Feng et al,[61] who evaluated housekeeping genes in response to abiotic stress in celery (*Apium graveolens* L.). Taken together, *RtActin* may be used in *R trisperma* species. Further experiments are required to evaluate and validate the performance of the *Actin* gene.

Phylogenetic analysis of R trisperma Actin among Euphorbiaceae species

Gene sequence homology analysis was carried out using the previously obtained sequence, which was then processed using BLAST.[26] The *R trisperma Actin*-coding gene sequence has a query coverage value of 99% nucleotides, with a similarity value of 92% to 97% to different *Euphorbiaceae* species tested in this study (see Table 3). This similarity value indicates gene sequence homology.[62] The *R trsiperma Actin*-coding gene fragment showed the highest and lowest similarity values to the *V fordii Actin*-coding gene (accession number JQ680035.1, 97%) and *Euphorbia lathyris Actin* gene (accession number KC182753.1, 92%), respectively. Species numbers 1 to 7 belong to the *Euphorbiaceae* family. Species number 8 belongs to the family *Fagaceae*, while species 9 to 11 belong to the *Malvaceae* family. According to information from the *actin*-coding genes, these results indicate that *Fagaceae* and *Malvaceae* are closest to *Euphorbiaceae*.

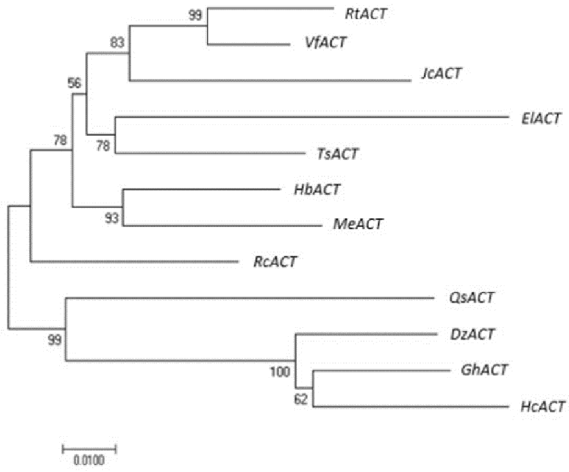

Figure 6. *RtACT* phylogenetic tree among other Euphorbiaceae species. DzACT indicates Durio zibethinus actin; ElACT, Euphorbia lathyris actin; GhACT, Gossypium hirsutum actin; HbACT, Hevea brasiliensis actin; HcACT, Hibiscus cannabinus actin; JcACT, Jatropha curcas actin; MeACT, Manihot esculenta actin; QsACT, Quercus source actin; RcACT, Ricinus communis actin; TsACT, Triadica sebifera Actin; VfACT, Vernicia fordii Actin.

The *R trisperma Actin*-coding gene was then compared with other *Actin*-coding genes of *Euphorbiaceae* species using phylogenetic analysis. Four species outside the *Euphorbiaceae* family, including *Quercus suber* (XM_ 024058244.1; Fagaceae), *Durio zibethinus* (XM_022883610.1; Malvaceae), *Gossypium hirsutum* (NM_001327089.1; Malvaceae), and *Hibiscus cannabinus* (KY382441.1; Malvaceae) were selected as outgroup. Outgroups were included in the analysis to obtain more credible information from related gene sequences. Outgroups were needed to provide polarization of characters or traits, namely, apomorphic and plesiomorphic.[63] Based on the phylogenetic tree (see Figure 6), the 11 species were separated into two main clusters. The first cluster comprised *actin*-coding genes from *Euphorbiaceae* species, including *RtACT, VfACT, JcACT, TsACT,*

ElACT, HbACT, MeACT, and *RcACT*. The second cluster comprised *QsACT, DzACT, GhACT*, and *HcACT*. This result indicates that *RtACT* is in a monophyletic group with other *Euphorbiaceae* families. *RtACT* displayed similarity to *VfACT*, supported by a bootstrap value of 99%. According to Hall,[64] a clade can be trusted with bootstrap values above 90% and cannot be trusted if bootstrap values are below 25%. These results indicate that our phylogenetic tree construction is reliable, and the *Actin* gene can provide an informative description of the molecular phylogeny of *R trisperma.*

Conclusions

This study used two plastid markers (*rbcL* and *matK*) for the DNA barcode of *R. trisperma*, a promising biosource for biodiesel production. We isolated 650 and 850 bp, corresponding to *rbcL* and *matK* fragments, respectively. However, we observed that the *rbcL* marker is more discriminatory for the molecular identification of *R trisperma* plants than the *matK marker*. Our phylogenetic tree of the *rbcL* plastid marker showed that *R trisperma* was clustered in the same group as other *Euphorbiaceae* plants and displayed a high similarity to *V fordii.* Nevertheless, other multi-locus DNA barcoding markers, such as intergenic regions (*trnH-pbsA*), nuclear ribosomal ITS, or other chloroplast genomes (*rpoC1* and *rpoB*), could be used to compare the discrimination power of the markers. Furthermore, we successfully isolated 986 bp of *the Actin*-coding gene, which might help in the development of gene expression analyses of *R trisperma*. The reference gene should be stably expressed in all plant organs because it maintains the essential cellular functions. Therefore, the actin fragment sequence obtained for differential gene expression studies of *R trisperma* organs should be conducted. This could provide additional information on *RtACT* stability as a potential reference gene. BLAST analysis showed that the gene was homologous to other *Actin*-coding genes of the *Euphorbiaceae* family and showed the highest degree of kinship with the *Actin*-coding genes of *V fordii.* Taken together, our DNA barcoding results could be used as a reference for *R trisperma* conservation programs. Meanwhile, validation of *Actin* stability and effectiveness as a reference gene for *R trisperma* transcriptomic analysis still needs to be further validated.

Author Contributions

NJ, NNS, and DO conceived and designed the study. NLAR, IP, MRNR, and AFW performed bioinformatics analysis. NJ and NLAR performed molecular biology experiments. NJ, NNS, and DO supervised the experiments and acquired the funding. NJ, NLAR, and MRNR drafted the article. NJ, NNS, and DO reviewed and edited the article. All authors have read and approved the final article.

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

REFERENCES

1. Wulandari WS, Darusman D, Cecep Kusmana W. Land suitability analysis of biodiesel crop Kemiri Sunan (Reutealis trisperma (Blanco) Airy Shaw) in the Province of West Java, Indonesia. *J Environ Earth Sci*. 2014;4:27-37.
2. Herman M, Syakir M, Pranowo D, Saefudin, dan Sumanto, Kemiri S. *(Reutealis Trisperma (Blanco) Airy Shaw) Tanaman Penghasil Minyak Nabati dan Konservasi Lahan*. IAARD Press; 2013.
3. Jadid N, Rachman RY, Hartanti SR, Abdulgani N, Wikanta W, Muslihatin W. Methanol extract of piper retrofractum vahl. potentially mediates mast cell stabilization. *Int J Pharm Bio Sci*. 2016;7:379-383.
4. Lu Y, Huang YS, Chen CH, et al. Anti-HIV tigliane diterpenoids from Reutealis trisperma. *Phytochemistry*. 2020;174:112360.
5. Holilah H, Prasetyoko D, Oetami TP, et al. The potential of Reutealis trisperma seed as a new non-edible source for biodiesel production. *Biomass Convers Biorefin*. 2015;5:347-353.
6. Riayatsyah TM, Ong HC, Chong WT, Aditya L, Hermansyah H, Mahlia TM. Life cycle cost and sensitivity analysis of Reutealis trisperma as non-edible feedstock for future biodiesel production. *Energies*. 2017;10:877.
7. Kotcher J, Maibach E, Choi WT. Fossil fuels are harming our brains: identifying key messages about the health effects of air pollution from fossil fuels. *BMC Public Health*. 2019;19:1079.
8. Lelieveld J, Klingmüller K, Pozzer A, Burnett RT, Haines A, Ramanathan V. Effects of fossil fuel and total anthropogenic emission removal on public health and climate. *Proc Nat Acad Sci*. 2019;116:7192-7197.
9. Gaurav N, Sivasankari S, Kiran GS, Ninawe A, Selvin J. Utilization of bioresources for sustainable biofuels: a review. *Renew Sust Energ Rev*. 2017;73:205-214.
10. Riayatsyah TM, Sebayang AH, Silitonga AS, et al. Current progress of jatropha curcas commoditisation as biodiesel feedstock: a comprehensive review. *Front Energ Res*. 2022;9:815416.
11. Pooja S, Anbarasan B, Ponnusami V, Arumugam A. Efficient production and optimization of biodiesel from kapok (Ceiba pentandra) oil by lipase transesterification process: addressing positive environmental impact. *Renew Energ*. 2021;165:619-631.
12. Arumugam A, Ponnusami VJ. Biodiesel production from Calophyllum inophyllum oil a potential non-edible feedstock: an overview. *Renew Energ*. 2019;131:459-471.
13. Hamadou B, Falama RZ, Delattre C, Pierre G, Dubessay P, Michaud P. Influence of physicochemical characteristics of Neem seeds (*Azadirachta indica A.* Juss) on biodiesel production. *Biomolecules*. 2020;10:616.
14. Onoji SE, Iyuke SE, Igbafe AI, Daramola MO. Rubber seed (Hevea brasiliensis) oil biodiesel emission profiles and engine performance characteristics using a TD202 diesel test engine. *Biofuels*. 2022;13:423-430.
15. Awais M, Musmar SA, Kabir F, et al. Biodiesel production from Melia azedarach and ricinus communis oil by transesterification process. *Catalysts*. 2020;10:427.
16. Wu S, Gao C, Pan H, et al. Advancements in tobacco (Nicotiana tabacum L.) seed oils for biodiesel production. *Front Chem*. 2021;9:834936.
17. Alhassan Y, Kumar N. Single step biodiesel production from Pongamia pinnata (Karanja) seed oil using deep eutectic solvent (DESs) catalysts. *Waste Biomass Valor*. 2016;7:1055-1065.
18. Khan IU, Chen H, Yan Z, Chen J. Extraction and quality evaluation of biodiesel from six familiar non-edible plants seeds. *Processes*. 2021;9:840.
19. Shah SN, Sharma BK, Moser BR, Erhan SZ. Preparation and evaluation of jojoba oil methyl esters as biodiesel and as a blend component in ultra-low sulfur diesel fuel. *Bioenergy Res*. 2010;3:214-223.
20. Choksi H, Pandian S, Gandhi YH, Deepalakshmi S. Studies on production of biodiesel from Madhuca indica oil using a catalyst derived from cotton stalk. *Energ Source A*. 2021;43:3424-3433.
21. Wang R, Zhou WW, Hanna MA, et al. Biodiesel preparation, optimization, and fuel properties from non-edible feedstock, Datura stramonium L. *Fuel*. 2012;91:182-186.
22. Kansedo J, Lee KT. Process optimization and kinetic study for biodiesel production from non-edible sea mango (Cerbera odollam) oil using response surface methodology. *Chem Eng J*. 2013;214:157-164.
23. Fadhil AB, Ahmed KM, Dheyab MM. Silybum marianum L. seed oil: a novel feedstock for biodiesel production. *Arab J Chem*. 2017;10:S683-S690.
24. International Union for Conservation of Nature and Natural Resources (IUCN). The IUCN red list of threatened species. http://www.iucnredlist.org.
25. Hartvig I, Czako M, Kjær ED, Nielsen LR, Theilade I. The use of DNA barcoding in identification and conservation of rosewood (Dalbergia spp.). *PLoS ONE*. 2015;10:e0138231.
26. Jadid N, Prasetyowati I, Rosidah NLA, et al. In silico analysis of partial fatty acid desaturase 2 CDNA from Reutealis trisperma (blanco) airy Shaw. *Bioinform Biol Insights*. 2021;15:11779322211005747-11779322211005748. doi:10.1177/11779322211005747
27. CBOL Plant Working Group. A DNA barcode for land plants. *Proc Natl Acad Sci USA*. 2009;106:12794-12797. doi:10.1073/pnas.0905845106

28. Ho VT, Tran TKP, Vu TTT, Widiarsih S. Comparison of matK and rbcL DNA barcodes for genetic classification of jewel orchid accessions in Vietnam. *J Genet Eng Biotechnol*. 2021;19:93. doi:10.1186/s43141-021-00188-1
29. Worthy SJ, Bucalo K, Perry E, et al. Ability of rbcL and matK DNA barcodes to discriminate between montane forest orchids. *Plant Syst Evol*. 2022;308:19. doi:10.1007/s00606-022-01809-z
30. Paolacci AR, Tanzarella OA, Porceddu E, Ciaffi M. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC Mol Biol*. 2009;10:11. doi:10.1186/1471-2199-10-11
31. Zhang L, He LL, Fu QT, Xu ZF. Selection of reliable reference genes for gene expression studies in the biofuel plant Jatropha curcas using real-time quantitative PCR. *Int J Mol Sci*. 2013;14:24338-24354.
32. Karuppaiya P, Yan X-X, Liao W, Wu J, Chen F, Tang L. Identification and validation of superior reference gene for gene expression normalization via RT-qPCR in staminate and pistillate flowers of Jatropha curcas—a biodiesel plant. *PLoS ONE*. 2017;12:e0172460.
33. Hannum S, Akashi K, Suharsono UW, Hartana A, Yokota A. Isolasi fragmen cDNA dari gen penyandi aktin dari Melastoma malabathricum. *Makara J Sci*. 2011;14:163-167.
34. Chervoneva I, Li Y, Schulz S, et al. Selection of optimal reference genes for normalization in quantitative RT-PCR. *BMC Bioinformatics*. 2010;11:253. doi:10.1186/1471-2105-11-253
35. Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques*. 2004;37:112-119.
36. Luo C, He XH, Chen H, Hu Y, Ou SJ. Molecular cloning and expression analysis of four actin genes (MiACT) from mango. *Biol Plant*. 2013;57:238-244.
37. Vantard M, Blanchoin L. Actin polymerization processes in plant cells. *Curr Opin Plant Biol*. 2002;5:502-506.
38. Blessing CA, Ugrinova GT, Goodson HV. Actin and ARPs: action in the nucleus. *Trends Cell Biol*. 2004;14:435-442.
39. Gilliland LU, Kandasamy MK, Pawloski LC, Meagher RB. Both vegetative and reproductive actin isovariants complement the stunted root hair phenotype of the Arabidopsis act2-1 mutation. *Plant Physiol*. 2002;130:2199-2209.
40. Podevin N, Krauss A, Henry I, Swennen R, Remy S. Selection and validation of reference genes for quantitative RT-PCR expression studies of the non-model crop Musa. *Mol Breed*. 2012;30:1237-1252.
41. Nicot N, Hausman JF, Hoffmann L, Evers D. Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *J Exp Bot*. 2005;56:2907-2914.
42. Jian B, Liu B, Bi Y, Hou W, Wu C, Han T. Validation of internal control for gene expression study in soybean by quantitative real-time PCR. *BMC Mol Biol*. 2008;9:59.
43. Zhang J, Liu D, Wang Z, et al. Expression pattern of GS3 during panicle development in rice under drought stress: quantification normalized against selected housekeeping genes in Real-Time PCR. *Asian J Plant Sci*. 2009;8:285-292.
44. Liu F, Guo DD, Tu YH, Xue YR, Gao Y, Guo ML. Identification of reference genes for gene expression normalization in safflower (Carthamus tinctorius). *Rev Bras Farmacogn*. 2016;26:564-570.
45. Andriyani Jadid N. Genetic diversity and morphological responses of Capsicum annuum varieties under alumunium stress. *Biodiversitas*. 2021;22:2576-2582. doi:10.13057/biodiv/d220516
46. Jadid N, Mardika RK, Nurhidayati T, Irawan MI. Reverse transcription-PCR analysis of geranylgeranyl diphosphate synthase (JcGGPPS) in Jatropha curcas L. and in silico analysis of casbene synthase (JcCS) among Euphorbiaceae. *AIP Conf Proc*. 2016;1744:020042. doi:10.1063/1.4953516
47. Green MR, Sambrook J. *Molecular Cloning: A Laboratory Manual*. 4th ed. Cold Spring Harbor Laboratory Press; 2012.
48. Jacobus AP, Gross J. Optimal cloning of PCR fragments by homologous recombination in Escherichia coli. *PLoS ONE*. 2015;10:e0119221.
49. Cho M, Xiao Y, Nie J, et al. Quantitative selection of DNA aptamers through microfluidic selection and high-throughput sequencing. *Proc Natl Acad Sci*. 2010;107:15373-15378.
50. Horiike T. An introduction to molecular phylogenetic analysis. *Rev Agric Sci*. 2016;4:36-45.
51. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987;4:406-425.
52. Stuppy W, Welzen PC, Klinratana P, Posa MC. Revision of the genera Aleurites, Reutealis and Vernicia (Euphorbiaceae). *Blumea*. 1999;44:73-98.
53. Maloukh L, Kumarappan A, Jarrar M, Salehi J, El-Wakil H, Rajya Lakshmi TV. Discriminatory power of rbcL barcode locus for authentication of some of United Arab Emirates (UAE) native plants. *3 Biotech*. 2017;7:144.
54. Barthet MM, Hilu KW. Expression of matK: functional and evolutionary implications. *Am J Bot*. 2007;94:1402-1412.
55. Snustad DP, Simmons MJ. *Principles of Genetics*. John Wiley & Sons; 2015.
56. Bergmans HE, van Die IM, Hoekstra WP. Transformation in Escherichia coli: stages in the process. *J Bacteriol*. 1981;146:564-570.
57. Van Holm W, Ghesquière J, Boon N, et al. A viability quantitative PCR dilemma: are longer amplicons better? *Appl Environ Microbiol*. 2021;87: e0265320.
58. Li MY, Song X, Wang F, Xiong AS. Suitable reference genes for accurate gene expression analysis in parsley (Petroselinum crispum) for abiotic stresses and hormone stimuli. *Front Plant Sci*. 2016;7:1481.
59. Tian C, Jiang Q, Wang F, Wang GL, Xu ZS, Xiong AS. Selection of suitable reference genes for qPCR normalization under abiotic stresses and hormone stimuli in carrot leaves. *PLoS ONE*. 2015;10:e0117569. doi:10.1371/journal.pone.0117569
60. Jadid N, Mardika RK, Purwani KI, Permatasari EV, Prasetyowati I, Irawan MI. Transcription profile data of phorbol esters biosynthetic genes during developmental stages in Jatropha curcas. *Data Brief*. 2018;18:700-705.
61. Feng K, Liu JX, Xing GM, et al. Selection of appropriate reference genes for RT-qPCR analysis under abiotic stress and hormone treatment in celery. *PeerJ*. 2019;7:e7925.
62. Claverie JM, Notredame C. *Bioinformatics for Dummies*. John Wiley & Sons; 2006.
63. Wheeler WC. Nucleic acid sequence phylogeny and random outgroups. *Cladistics*. 1990;6:363-367.
64. Hall BG. *Phylogenetic Trees Made Easy: A How to Manual for Molecular Biologists*. Sinauer Associates; 2001.