BIODIVERSITAS
Volume 25, Number 7, July 2024
Pages: 2999-3007

ISSN: 1412-033X
E-ISSN: 2085-4722
DOI: 10.13057/biodiv/d250721

Genetic comparison among some cultivars of *Catharanthus roseus* (L.) G. Don. using three intergenic spacers of the chloroplast genome

AGUS HERY SUSANTO♥, MURNI DWIATI
Faculty of Biology, Universitas Jenderal Soedirman. Jl. Dr. Suparno No. 63, Purwokerto Utara, Banyumas 53122, Central Java, Indonesia.
Tel.: +62-281-638794, Fax.: +62-281-631700, ♥email: susanto1408@unsoed.ac.id

Manuscript received: 29 November 2023. Revision accepted: 10 July 2024.

Abstract. *Susanto AH, Dwiati M. 2024. Genetic comparison among some cultivars of* Catharanthus roseus *(L.) G. Don. using three intergenic spacers of the chloroplast genome. Biodiversitas 25: 2999-3007. Catharanthus roseus* (L.) G. Don belongs to the Apocynaceae family and is recognized as both an ornamental plant and a medicinal herb. By means of both hybridization and mutation breeding, numerous cultivars of *C. roseus* have been developed over time. This encourages the necessity for molecular characterization to establish a DNA barcode for each cultivar, particularly in relation to its potential as a medicinal herb. Therefore, this study aimed to genetically compare some cultivars of *C. roseus* based on three Intergenic Spacers (IGS) of the chloroplast genome, namely *trnQ-rps16*, *rps16-trnK*, and *trnL-rpl32*. Ten samples of *C. roseus* were randomly collected from around Purwokerto City, Central Java, Indonesia, representing cultivars with various corolla and flower eye colors. Universal primers were used to amplify the three molecular markers. The results showed that some insertion-deletions (indels) with *rps16-trnK* and *trnL-rpl32* were observed, indicating some potential barcodes for *C. roseus* cultivars, following pharmacological analysis. On the other hand, very little genetic difference among samples was found with *trnQ-rps16*. It can be concluded that some other molecular markers, specifically those more closely related to the expression of flower color in *C. roseus*, should be explored to obtain candidate DNA barcodes in assisting the authentication of medicinal herbs.

Keywords: *Catharanthus roseus*, chloroplast genome, intergenic spacer, molecular characterization

INTRODUCTION

Catharanthus roseus (L.) G. Don, commonly known as Madagascar periwinkle, was formerly known as an ornamental plant but has now gained recognition as a valuable medicinal herb. More than 130 secondary metabolites are potentially utilized in the treatment of various non-infectious diseases such as diabetes, hypertension, asthma, constipation, menstruation problems, and several cancers related cases (Aruna et al. 2015; Pan et al. 2016). The dimetric alkaloids vincristine and vinblastine derived from *C. roseus*, are known to have anticancer properties, thereby widely needed in pharmaceutical industries (Das and Sharangi 2017; Zhang et al. 2018; Vu et al. 2022). These compounds are particularly used in the treatment of lymph node cancer and leukemia and are mostly found in stems and leaves of *C. roseus* (Dhyani et al. 2022).

Originating from Madagascar Island, as the name implies, *C. roseus* is distributed across numerous tropical and subtropical regions worldwide (Paarakh et al. 2019). This plant species has been reported to be found in some states of the southern US (Mishra and Verma 2017). In Spain and various other southern European countries, *C. roseus* has been extensively cultivated for commercial purposes related to their utilization as medicinal herbs (Das and Sharangi 2017).

There are two cultivars of *C. roseus*, namely one characterized by pink flowers and another with white flowers (Aruna et al. 2015). Despite the challenges posed by the plant's closed pollination system (cleistogamous), which rarely includes insect-assisted pollination, numerous other *C. roseus* cultivars have emerged through hybridization. These include varieties with Mirabella red, maroon red, and purple corollas (Chen and Yeh 2012). Another approach to creating genetic variation in *C. roseus* is by means of mutation breeding using gamma-ray irradiation on various plant parts (Noormohammadi et al. 2020). Induced somaclonal variation using gamma-ray irradiation has been proven effective in producing genetic variation in *C. roseus*, specifically in inducing the expression of some genes involved in the regulation of alkaloid biosynthesis (Noormohammadi et al. 2018).

The proliferation of numerous new cultivars of *C. roseus* underscores the importance of characterization efforts to establish barcodes corresponding to each cultivars potential as a medicinal herb. However, morphological characterization can only be conducted at the generative phase, since the differences commonly lie merely on the corolla and flower eye colors. The majority of medical values of *C. roseus* are found in the vegetative phase, particularly in the stems and leaves. Therefore, molecular approaches, which are not depending on the stages of plant growth and development, are needed (Makki et al. 2019).

Many molecular markers commonly used in the assessment of intra-specific genetic variation across various plant species, specifically in Angiosperms, are readily available. The chloroplast genome is identified as an ideal source of molecular markers for genetic variation analysis

in Angiosperm. This is because of the uniparental inheritance of the chloroplast genome, which rules out the influence of recombination on the genetic variation observed (Park et al. 2021).

The complete map of the chloroplast genome in *C. roseus* is now available, enabling its utilization in studying genetic variation, whether intra-specific or inter-specific among other species of *Catharanthus*. Four Intergenic Spacers (IGS) have been identified as the candidates of molecular markers in the study of inter-specific genetic variation in *Catharanthus* due to their high mutation rates. They are *trnT*(GGU)-*trnE*(UUC), *trnQ*(UUG)-*rps16*, *rps16*-*trnK*(UUU) and *trnL*(UAG)-*rpl32* (Ku et al. 2013). With the exception of *trnT*(GGU)-*trnE*(UUC), these molecular markers show high variabilities within some other Angiospermae species, such as *Magnolia acuminata* (L.) L., *Liriodendron tulifioera* L., *Triticum ovatum* (L.) Gren. & Godr., *Pseudotrillium rivale* (S.Watson) S.B.Farmer, *Minuartia uniflora* (Walter) Mattf., *Prunus hortulata* L.H.Bailey, *Hibiscus cannabinus* L., *Gratiola brevifolia* Raf., and *Eupatorium capillifolium* (Lam.) Small. The levels of variability even surpass those of other non-coding regions reported previously, thereby the other three molecular markers are highly suitable for studying intra-specific genetic variation in Angiosperms (Shaw et al. 2007). Therefore, in this study, the genetic differences among some cultivars of *C. roseus* were assessed using *trnQ*(UUG)-*rps16*, *rps16*-*trnK*(UUU), and *trnL*(UAG)-*rpl32* as the molecular markers. The molecular data obtained will support the phenotype-based comparative study among some *C. roseus* cultivars, particularly related to their respective potential as a medicinal herb. In other words, this will contribute to the scientific basis for *C. roseus* applications in biopharmaceuticals.

MATERIALS AND METHODS

Plant materials

Ten plants were randomly collected from Purwokerto City, Central Java, Indonesia, representing some *C. roseus* cultivars with different corolla and flower eye colors. Among these plants, there were four wild-type and six mutant individuals. The wild-types comprised two with pink flowers (C6 and C9) and another two with white flowers (C7 and C8), while the mutants included those of red corolla with white flower eye (C1), white corolla with white flower eye (C2), Mirabella red (C3), purple corolla with white flower eye (C4), maroon red (C5) and large white corolla (C10). Stalks were cut from each plant and grown in a pot containing soil medium. Once they exhibited sufficient growth, fresh leaves were harvested for DNA genomic extraction.

Procedures

Genomic DNA extraction

The genomic DNA extraction was carried out following the CTAB method (Doyle and Doyle 1990). Approximately 0.1 g of leaf was cut into pieces and put into a 1.5 mL microcentrifuge tube. Then, 800 µL of CTAB solution previously heated at 65°C for 30 minutes was added. The leaf pieces were gently crushed with a mini-bead beater and heated at 65°C for an hour, then cooled to room temperature. Next, 500 µL of Chloroform-Isoamyl Alcohol (CIAA) was added, mixed gently, and centrifuged at 12,000 rpm for 5 minutes. The supernatant was transferred into another 1.5 mL microcentrifuge followed by the addition of 1/10 volume of 3M sodium acetate. Subsequently, isopropanol of 2/3 of the total volume was added and mixed gently by inverting the tube several times. This mixture was kept in the freezer for 24 hours prior to centrifugation at 12,000 rpm for 10 minutes. The supernatant was removed and the DNA pellet was washed with the addition of 500 µL 70% ethanol. This solution was centrifuged at 12,000 rpm for 5 minutes and the supernatant was removed. The DNA pellet was air-dried and dissolved in 100 µL TE buffer, and the DNA solution was stored at 4°C before quantification and PCR amplification. The DNA was visualized in a 1.5% agarose gel electrophoresis using 1x TBE buffer for quality examination, while quantification was performed using gene quant.

Molecular marker amplification

The three IGS molecular markers, namely trnQ (UUG)-rps16, rps16-trnK(UUU), and trnL(UAG)-rpl32, were amplified using PCR technique, in which the extracted genomic DNAs served as the templates. The pairs of universal primers used are presented in Table 1.

Each PCR mixture was prepared in a total volume of 10 µL containing 2.5 µL 5 ng/µL template DNA, 0.25 µL 1 µM primers (forward and reverse), 5 µL Kapa Taq DNA polymerase (Roche) and 2.25 µL nuclease-free water. The PCR conditions used are as follows: pre-denaturation at 94°C for 3 minutes, followed by 33 PCR cycles consisting of denaturation at 94°C for 45 seconds, primer annealing at 55°C for 45 seconds, primer extension at 72°C for 2 minutes, and proceeded by a final extension at 72°C for 3 minutes. The PCR reaction was performed in a Bio-Rad-T100™ Thermal Cycler machine. The PCR products were visualized in a 1% agarose gel electrophoresis using 1x TBE buffer. When a clear single PCR band of the respective sample was obtained, the PCR reaction was scaled up to a total volume of 50 µL.

Table 1. The universal primers used to amplify and sequence molecular markers

Molecular marker	Universal primer (5' - 3')	Reference
trnQ(UUG)-*rps16*	F: GCGTGGCCAAGYGGTAAGGC	Shaw et al. (2007)
	R: GTTGCTTTYTACCACATCGTTT	
rps16-*trnK*(UUU)	F: AAAGTGGGTTTTTATGATCC	Shaw et al. (2007)
	R: TTAAAAGCCGAGTACTCTACC	
trnL(UAG)-*rpl32*	F: CTGCTTCCTAAGAGCAGCGT	Shaw et al. (2007)
	R: CAGTTCCAAAAAAACGTACTTC	

Not: F: Forward primer; R: Reverse primer

DNA sequencing

All the PCR products were first purified using the QIA Quick kit (Qiagen), and, sent to Firstbase Malaysia for sequencing using the automated dideoxy method (de Camargo Smidt et al. 2020).

Data analysis

The DNA sequences obtained were edited using Bioedit software version 7.0.4.1 (Hall 1999) and checked manually. The edited sequences were compared to those from GenBank® using Basic Local Alignment Search Tool (BLAST) at NCBI for validation. Then, sequence alignment was performed using ClustalW (Thompson et al. 1994) implemented in MEGA version 11 (Tamura et al. 2021). Genetic distance was calculated using the Bioedit software version 7.0.4.1, while a phylogenetic dendrogram with UPGMA was constructed using the MEGA version 11.

RESULTS AND DISCUSSION

The genomic DNAs from all ten plant samples were successfully extracted resulting in sufficiently high concentrations ranging from 722 ng/µL to 1,935 ng/µL. The A260 nm/A280 nm ratio varies from 2.065 to 2.096 indicating relatively high quality of DNAs. This means that all the extracted genomic DNAs could serve as PCR templates for molecular marker amplification.

All three IGS molecular markers, namely *trnQ* (UUG)-*rps16*, *rps16*-*trnK*(UUU), and *trnL*(UAG)-*rpl32*, were successfully amplified showing one clear band in the respective sample though some faint bands are also present as shown in Figures 1.A-1.C.

Figure 1 shows that all the PCR products with *trnQ*(UUG)-*rps16* IGS are approximately 1,700 bp in length. However, after manual editing, the sequences of only 1,263 bp long are obtained. The size of sample C2, at 1,295 bp, differs significantly from the other nine samples, and its base sequence is entirely different from that of the others. However, despite these differences, blasting the sequences against the data available in the NCBI database reveals that all sequences, including that of C2, exhibit a considerably high identity to the chloroplast genome of *C. roseus* cultivar Pacifica Punch Halo (Acc. no. NC_021423.1) as presented in Table 2. Meanwhile, all the PCR products with *rps16*-*trnK*(UUU) and *trnL*(UAG)-*rpl32* show about 900 bp and 1,200 bp long in size respectively. The sequences obtained after manual editing are 820 bp for *rps16*-*trnK*(UUU) and 1,112 bp for *trnL*(UAG)-*rpl32*. Regarding *trnQ*(UUG)-*rps16*, blasting against the NCBI database also shows a very high percentage of identity to the chloroplast genome of the *C. roseus* cultivar Pacifica Punch Halo (Table 2).

In this study, the size of the *trnQ*(UUG)-*rps16* IGS sequences obtained are much shorter in comparison to those found in some species of *Actinidia* (Actinidiaceae) ranging from 2,254 to 2,324 bp (Tang et al. 2019). This IGS is one of the divergent regions among plant species (Xie et al. 2021). Different *trnQ*(UUG)-*rps16* IGS sequences are identified between two chloroplast genomes of *Veronica nakaiana* (Plantaginaceae) (Lee et al. 2021). On the other hand, a significantly lower genetic difference with respect to *trnQ*(UUG)-*rps16* IGS sequences between *Campanula takesimana* (Campanulaceae) and its progenitor, *C. punctata*, has been reported (Cheong et al. 2020).

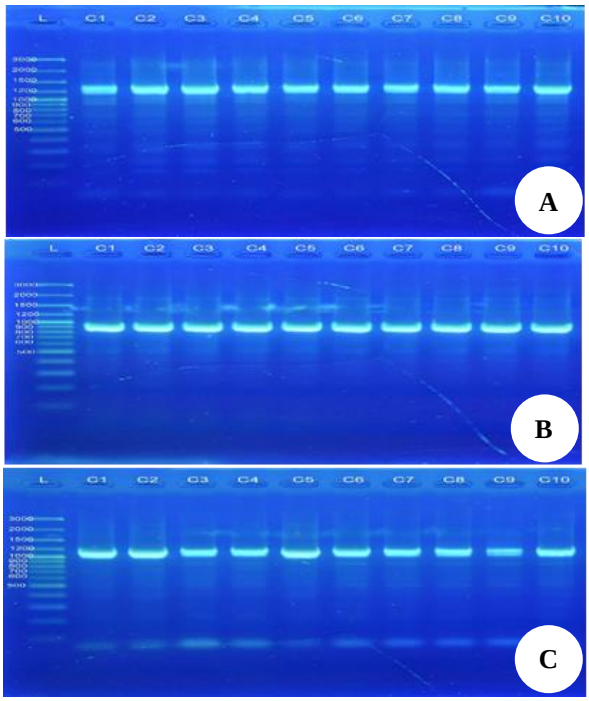

Figure 1. PCR amplification of molecular markers. A. *trnQ*(UUG)-*rps16* B. *rps16*-*trnK*(UUU) C. *trnL*(UAG)-*rpl32* L: 100 bp DNA ladder; C1: Red corolla with white flower eye; C2: White corolla with white flower eye; C3: Mirabella red; C4: Purple corolla with white flower eye; C5: Maroon red; C6: Wildtype pink 1; C7: Wildtype white 1; C8: Wildtype white 2; C9: Wildtype pink 2; C10: Large white corolla

Table 2. Blasting results to the NCBI database

Genetic marker	Maximum score	Total score	% Query cover	E-value	% Identity
trnQ(UUG)-*rps16*	2,292	2,292	99	0	99.68
trnQ(UUG)-*rps16* (sample C2)	2,265	2,265	98	0	98.67
rps16-*trnK*(UUU)	1,297	1,297	100	0	96.68
trnL(UAG)-*rpl32*	1,951	1,951	100	0	99.35

Note: The sequence referred for blasting is *Catharanthus roseus* cultivar Pacifica Punch Halo chloroplast genome (Acc. no. NC_021423.1)

In contrast to the *trnQ*(UUG)-*rps16* IGS, shorter *rps16-trnK*(UUU) IGS sequences are observed, including *rps16-trnK*(UUU) IGS of *Anethum graveolens* (Apiaceae) ranging from 747 to 779 bp. The variation within this species is mainly due to base substitution, although some indels are also reported (Sung et al. 2013). Oppositely, much longer *rps16-trnK*(UUU) IGS, namely 1,079 bp long, is obtained as a result of sequence alignment among seven genera of the subtribe Eleusininae (Poaceae) (Peterson et al. 2022b). Similarly, longer *rps16-trnK*(UUU) IGS, namely 1,044 bp long, is also found among some species of *Trisetum* and other members of the subtribe Koeleriinae (Poaceae) (Barberá et al. 2020). Meanwhile, sequences of rps16-trnK(UUU) IGS similar in size to the results in this study are found after sequence alignment among several species of the tribe Triraphideae (Poaceae), with a length of 887 bp (Peterson et al. 2022a).

A slightly longer *trnL*(UAG)-*rpl32* IGS of 1,246 bp is observed among some cultivars of *Nelumbo nucifera* (Nelumbonaceae). This sequence shows seven variable sites occupying 0.562% of the total sequence length (Long et al. 2022). Meanwhile, a 917 bp long *trnL*(UAG)-*rpl32* is observed among 24 individuals representing 15 species of *Zygophyllum* (Zygophylaceae) (Zhang et al. 2021).

Sequence alignments

When comparing the differences among samples, sequence alignment using *trnQ*(UUG)-*rps16* IGS shows some indels and base substitutions, as presented in Figure 2. All sites not shown in the figure display no polymorphism.

C5 (maroon red) shows an insertion in site 8 but deletions in sites 56, 82, and 101 in comparison to wildtypes (C6, C7, C8, C9). As well, some base substitutions in C1 (red corolla with white flower eye), C3 (Mirabella red), and C4 (purple corolla with white flower eye) are observed. However, some variations occur also between the two pink wildtypes and between the two white wildtypes, which means that *trnQ*(UUG)-*rps16* IGS is not an ideal molecular marker to distinguish *C. roseus* cultivars. In other words, this marker could not result in any potential barcode for *C. roseus* cultivars. This is in contrast to the case with eleven *Aristolochia* species (Aristolochiaceae), which can be molecularly identified using *trnQ*(UUG)-*rps16* IGS (Bai et al. 2023).

High sequence variation and intra and inter-specific divergence distinction are important in providing powerful barcodes in plant species. Some chloroplast genome markers, such as *matK*, *rbcL*, *psbA-trnH* IGS, and *trnL-trnF* IGS, have been considered for accurate discrimination of plants at the lower taxonomic levels. They have been used as standard DNA barcodes for identifying Chinese medicinal herbs (Zhu et al. 2022).

With regard to *rps16-trnK*(UUU) IGS and *trnL*(UAG)-*rpl32* IGS, relatively more variable sites are observed as shown in Figures 3 and 4 respectively. Regarding *trnQ*(UUG)-*rps16* IGS, all sites excluding those shown in the figures have no polymorphism.

Figure 3 shows slightly long indels observed in C1 to C5, which are all mutant cultivars. However, variations within wildtype cultivars, either pink or white corolla (C6 to C9), exist as well. Apart from sites 330 to 380, there are somewhat conserved sequences present within wild-type cultivars. Similar sequences are found within wildtype pink and wildtype white cultivars. On the other hand, many variations with respect to indels and base substitutions occur among mutant cultivars, indicating *rps16-trnK*(UUU) IGS as a potential source for DNA barcode in *C. roseus*. This molecular marker has been reported as one of the DNA barcodes for identifying *Tetrataenium candicans* (Apiaceae), a traditional Chinese medicinal herb species (Kang et al. 2019).

Conserved sequences of *trnL*(UAG)-*rpl32* IGS are observed within all four wildtype cultivars (C6 to C9) and three mutants (C3, C4, and C10) as shown in Figure 4. Meanwhile, some sufficiently long deletions occur at sites 275 to 304 and some shorter ones are present at sites 482 to 488 in mutant cultivars C1, C2, and C5. This suggests that *trnL*(UAG)-*rpl32* IGS could serve as a candidate for DNA barcode to identify *C. roseus* cultivars of red corolla with white eye flower, white corolla with white flower eye, and maroon red flower.

Four indels are identified among *trnL*(UAG)-*rpl32* IGS of six Atractylodes species (Asteraceae), providing the molecular marker as a potential DNA barcode for the six species (Wang et al. 2021). The *trnL*(UAG)-*rpl32* IGS is found as one of the most variable molecular markers among six species of Neotropical Cranichideae (Orchidaceae) (de Camargo Smidt et al. 2020). This marker, along with five others, has been effectively used as a DNA barcode to distinguish *Styrax species* (Styracaceae), which are plants of medicinal, ornamental, and economic significance (Song et al. 2022). Three herbarium specimens and another two samples of *Carex microcarpa* (Cyperaceae) have been successfully identified using *trnL*(UAG)-*rpl32* IGS. The morphological-based examination previously performed results in uncertain taxonomic identity (Miguez et al. 2022).

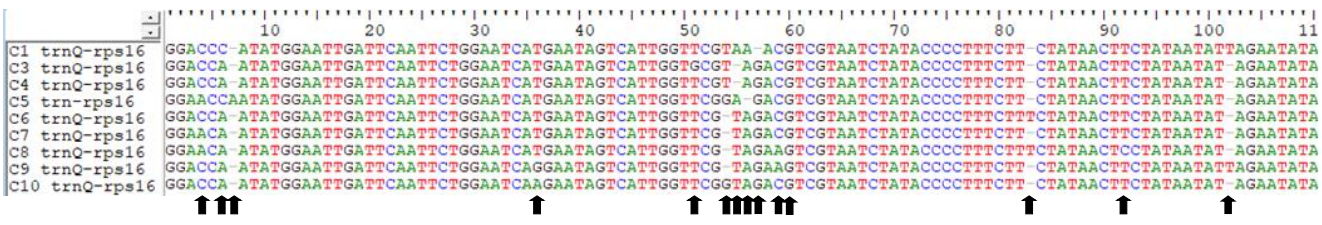

Figure 2. Sequence alignment among samples using *trnQ*(UUG)-*rps16* IGS. Black arrows indicate polymorphic sites

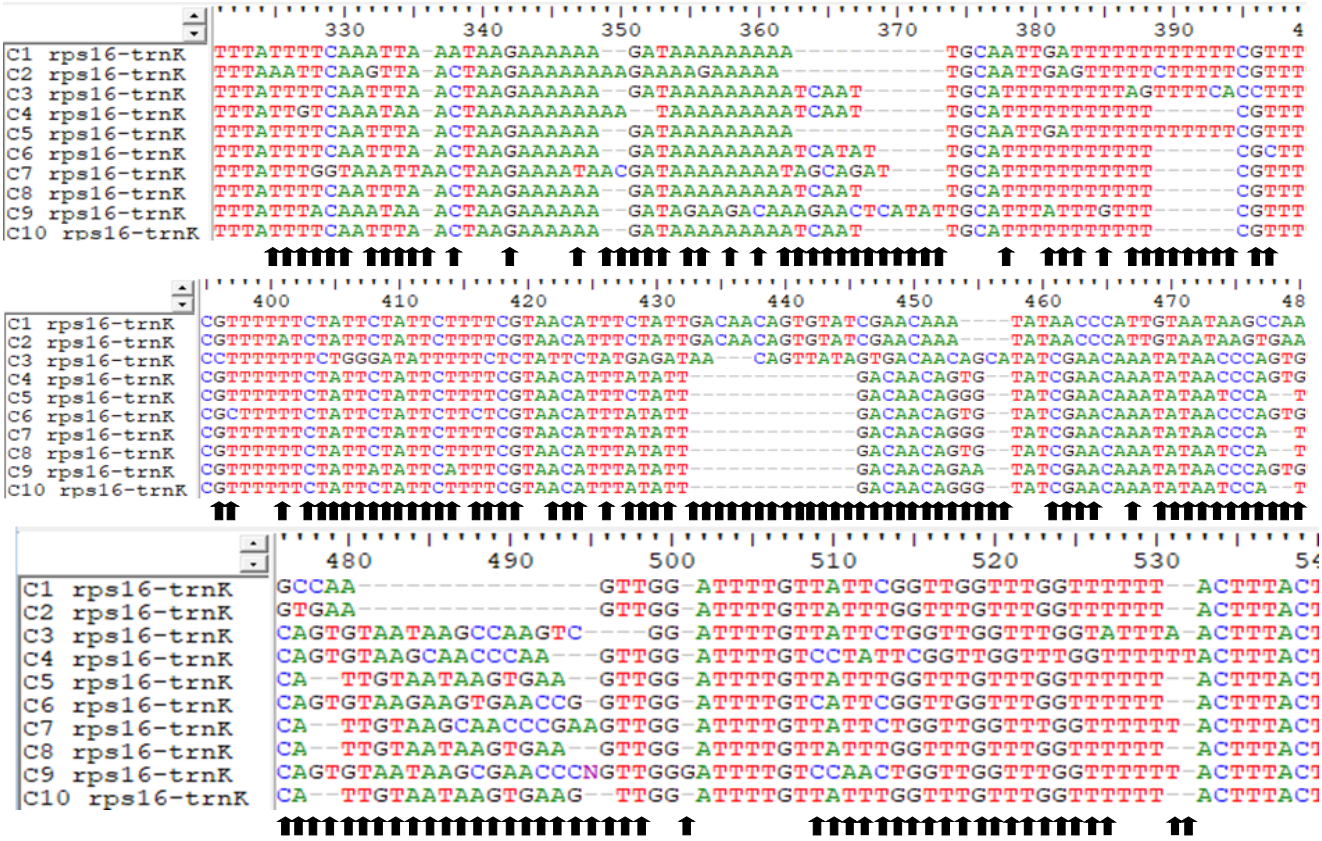

Figure 3. Sequence alignment among samples using *rps16- trnK*(UUU) IGS. Black arrows indicate polymorphic sites

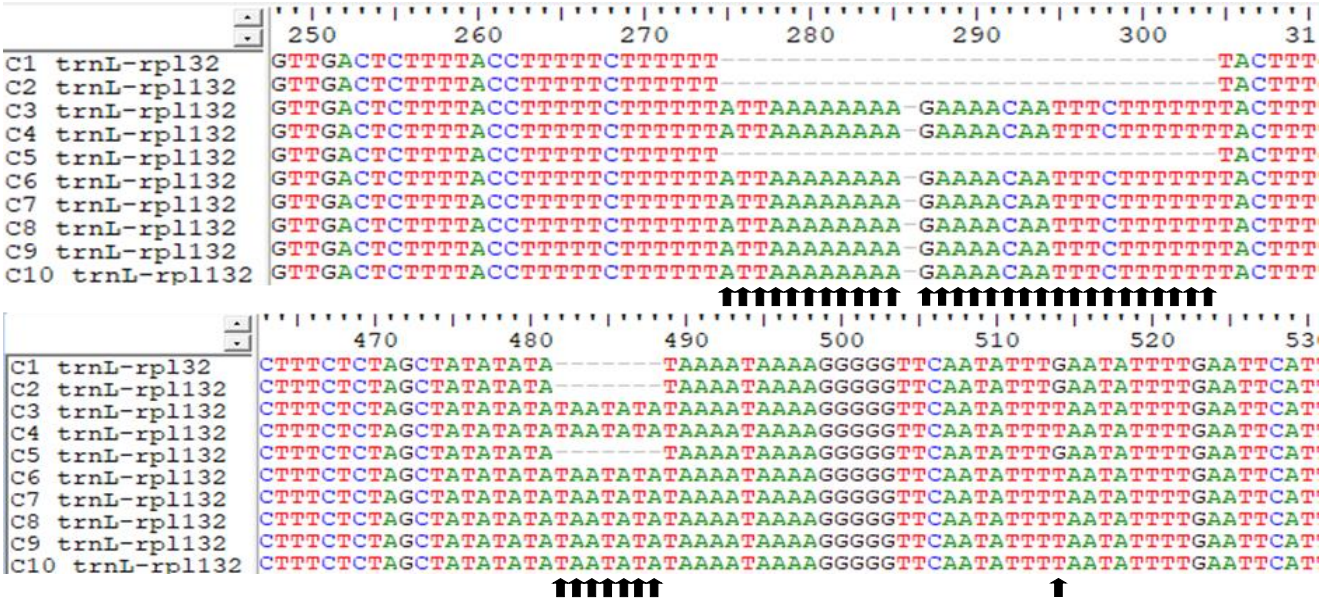

Figure 4. Sequence alignment among samples using *trnL*(UAG)-*rpl32* IGS. Black arrows indicate polymorphic sites

Genetic distances

Without involving C2 in the molecular comparison of *C. roseus* using *trnQ*(UUG)-*rps16* IGS, the pairwise genetic distances of the remaining nine samples are shown in Table 3. This table shows that the genetic distances of nine *C. roseus* samples using *trnQ*(UUG)-*rps16* IGS range from 0.0024 to 0.0138. The lowest genetic distance is observed between C6 and C9, while the highest is between C5 and C8. This is a reasonable sign that the lowest genetic distance between C6 and C9 is obtained since both are wild-type pink cultivars. Similarly, relatively low pairwise genetic distances of other wildtype cultivars are also observed. Using *trnQ*(UUG)-*rps16* IGS, a wider range of genetic distance is obtained within the samples of

Heracleum mantegazzianum (Apiaceae), namely 0.0000 - 0.0500 (Shadrin et al. 2023).

The pairwise genetic distances of *C. roseus* using *rps16-trnK*(UUU) IGS and *trnL*(UAG)-*rpl32* IGS are presented in Tables 4 and 5 respectively. These based on *rps16-trnK*(UUU) IGS varies from 0.0234 to 0.1344, and the lowest genetic distance is found between C8 and C10 samples. Low pairwise genetic distances can also be seen within wild-type cultivars, such as those between C6 and C7. On the other hand, higher genetic distances are observed between mutant and wild-type cultivars. In comparison, genetic grouping by the use of *rps16-trnK*(UUU) IGS is successfully performed among 58 *Vriesea* species (Bromeliaceae), a tropical plant genus harboring a wide variety of pollinator insects (Neves et al. 2021).

Table 3. Genetic distances of *Catharanthus roseus* samples based on *trnQ*(UUG)-*rps16* IGS

	C1	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1									
C3	0.0105								
C4	0.0113	0.0072							
C5	0.0048	0.0105	0.0113						
C6	0.0072	0.0040	0.0048	0.0097					
C7	0.0121	0.0072	0.0048	0.0130	0.0048				
C8	0.0130	0.0097	0.0072	0.0138	0.0056	0.0056			
C9	0.0097	0.0064	0.0072	0.0122	0.0024	0.0072	0.0048		
C10	0.0113	0.0080	0.0088	0.0130	0.0048	0.0097	0.0105	0.0064	

Note: C1: Red corolla with white eye flower; C3: Mirabella red; C4: Purple corolla with white flower eye; C5: Maroon red; C6: Wildtype pink 1; C7: Wildtype white 1; C8: Wildtype white 2; C9: Wildtype pink 2; C10: Large white corolla. Values higher than 0.0070 indicates different clades

Table 4. Genetic distances of *Catharanthus roseus* samples based on *rps16-trnK*(UUU) IGS

	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1										
C2	0.0446									
C3	0.1344	0.1216								
C4	0.1254	0.1109	0.0980							
C5	0.0754	0.0612	0.1070	0.0903						
C6	0.0821	0.0708	0.0738	0.0611	0.0494					
C7	0.0948	0.0844	0.0888	0.0752	0.0573	0.0520				
C8	0.0862	0.0716	0.0873	0.0724	0.0307	0.0388	0.0494			
C9	0.1105	0.0967	0.0825	0.0624	0.0783	0.0486	0.0605	0.0702		
C10	0.0946	0.0813	0.0951	0.0818	0.0335	0.0475	0.0584	0.0234	0.0762	

Note: C1: Red corolla with white eye flower; C2: White corolla with white eye flower; C3: Mirabella red; C4: Purple corolla with white flower eye; C5: Maroon red; C6: Wildtype pink 1; C7: Wildtype white 1; C8: Wildtype white 2; C9: Wildtype pink 2; C10: Large white corolla. Values higher than 0.0070 indicates different clades

Table 5. Genetic distances of *Catharanthus roseus* samples based on *trnL*(UAG)-*rpl32* IGS

	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1										
C2	0.0037									
C3	0.0151	0.0190								
C4	0.0161	0.0199	0.0027							
C5	0.0037	0.0037	0.0170	0.0180						
C6	0.0161	0.0200	0.0009	0.0018	0.0180					
C7	0.0141	0.0180	0.0027	0.0018	0.0180	0.0018				
C8	0.0151	0.0190	0.0018	0.0027	0.0170	0.0027	0.0027			
C9	0.0142	0.0180	0.0009	0.0018	0.0161	0.0018	0.0018	0.0009		
C10	0.0132	0.0170	0.0018	0.0027	0.0170	0.0027	0.0009	0.0018	0.0009	

Note: C1: Red corolla with white flower eye; C2: White corolla with white flower eye; C3: Mirabella red; C4: Purple corolla with white flower eye; C5: Maroon red; C6: Wildtype pink 1; C7: Wildtype white 1; C8: Wildtype white 2; C9: Wildtype pink 2; C10: Large white corolla. Values higher than 0.0070 indicates different clades

The pairwise genetic distances based on *trnL*(UAG)-*rpl32* IGS range from 0.0009 to 0.0200 as shown in Table 5. Lower genetic distances can be seen within wildtype cultivars, such as those between C8 and C9, while higher genetic distances are observed between mutants and wildtypes, such as those between C2 and C6. This genetic grouping provides strong support for the potential of *trnL*(UAG)-*rpl32* IGS as a candidate DNA barcode for identifying certain *C. roseus* cultivars. Similar results using the molecular marker have been reported in the analysis of genetic differences among *Physalis ixocarpa* (Solanaceae) populations in Michoacan State, Mexico (Martínez-Vega et al. 2022).

Phylogenetic relationships

As shown in Table 2, the *trnQ*(UUG)-*rps16* IGS sequence in sample C2 differs considerably from those of the other nine samples, despite its significant similarity to the chloroplast genome of *C. roseus* cultivar Pacifica Punch Halo chloroplast genome available in the NCBI database (Acc. no. NC_021423.1). Therefore, it can serve as an outgroup in the phylogenetic tree among *C. roseus* samples as shown in Figure 5.A. Without involving C2 from the phylogenetic tree (Figure 5.B), both wildtype pinks (C6 and C9) look to be in the same clade. This is also the case with the two wild-type whites (C7 and C8). However, confusion occurs as C7 seems to be closer to C4 (purple mutant) in comparison to C8.

In the absence of an outgroup, the two wildtype white cultivars (C7 and C8) appear to be separated into two different clades, while the two wildtype pinks (C6 and C9) belong to the same clade in the phylogenetic tree constructed using *rps16-trnK*(UUU) IGS. However, C6 seems to be closer to C4 (purple mutant) compared to C9, which is somewhat confusing. Similarly, C7 and C8 are separated into two different clades when the phylogenetic tree is constructed using *trnL*(UAG)-*rpl32* IGS. This is observed in the case of C6 and C9, although there is a relatively low genetic distance between them, namely 0.0018, is observed (Table 5).

It appears that genetic grouping based on phylogenetic relationships and genetic distance calculation may not directly correlate with the effectiveness of a specific molecular marker as a candidate DNA barcode. However, DNA barcode is a useful tool not only for species identification but also for phylogenetic construction (Kang et al. 2017). DNA barcoding is found suitable and essential for the characterization of genetic relationships among *Medicago sativa* cultivars (Fabaceae) (El-Sherif and Ibrahim 2020).

In the case of *C. roseus* cultivars, DNA barcoding plays an important in supporting molecular identification, specifically for cultivars possessing high contents of specific secondary metabolites. Therefore, it is essential to conduct studies on the correlation between the DNA barcode of *C. roseus* cultivars and their potential as medicinal herbs. For reference, about 35 selected local medicinal plants commonly used in Malaysia have been analyzed for correlation between the DNA barcodes using three molecular markers and their metabolite profiling (Wei 2022). The correlation between DNA barcodes of *C. roseus* and some other members of the Apocynaceae family and their potential as medicinal plants has been studied for species authentication. In a particular study, 122 plant samples belonging to 31 species and 19 genera of Apocynaceae were successfully identified using two molecular markers, namely ITS2 and *psbA-trnH* IGS. The results could support the improvement of DNA barcoding reference databases for herbal drugs and other herbal raw materials (Lv et al. 2020).

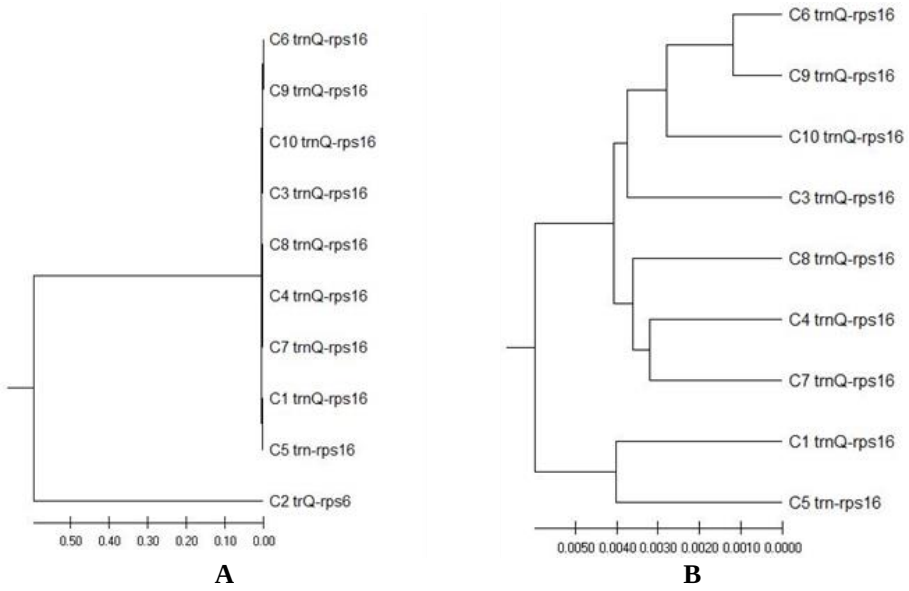

Figure 5. Phylogenetic tree among *Catharanthus roseus* cultivars using *trnQ*(UUG)-*rps16* IGS. A. Using C2 as the outgroup; B. Without outgroup. The phylogenetic trees using *rps16-trnK*(UUU) IGS and *trnL*(UAG)-*rpl32* IGS are presented in Figures 6 and 7 respectively

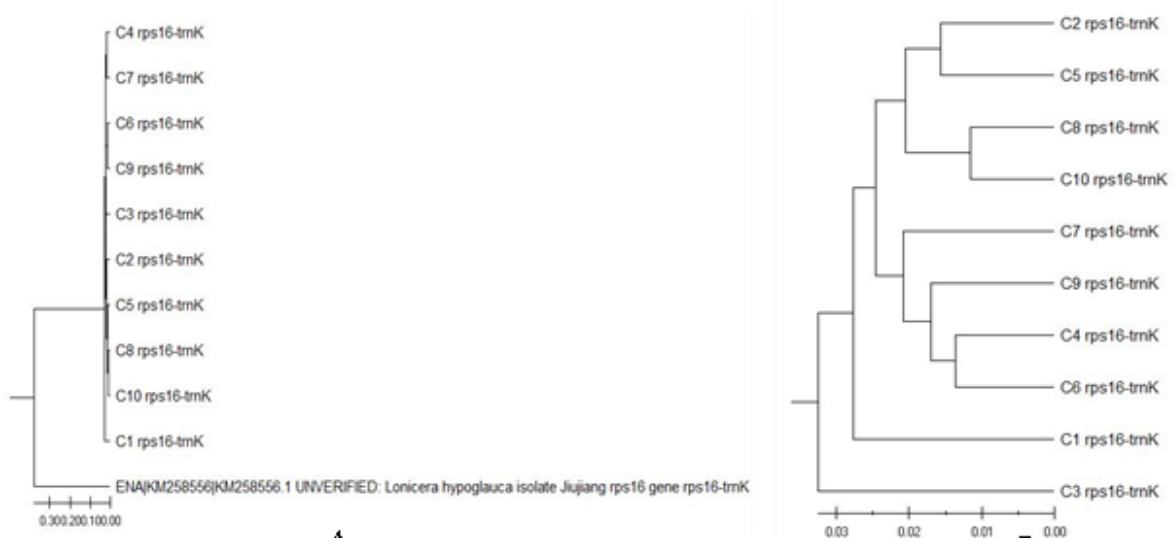

Figure 6. Phylogenetic tree among *Catharanthus roseus* cultivars using *rps16-trnK*(UUU) IGS. A. With outgroup; B. Without outgroup

Figure 7. Phylogenetic tree among *Catharanthus roseus* cultivars using *trnL***(**UAG)-*rpl32* IGS. A. With outgroup; B. Without outgroup

In conclusion, molecular comparison among various *C. roseus* cultivars to identify candidate DNA barcodes is of importance to assist in their authentication as medicinal herbs. Furthermore, exploration of other molecular markers showing a higher correlation with the expression of flower color in *C. roseus* should be performed.

ACKNOWLEDGEMENTS

The authors are very grateful to the Institution of Research and Public Service, Universitas Jenderal Soedirman, Banyumas, Indonesia for funding this project through the scheme of *Riset Fasilitasi Tugas Khusus Guru Besar* financial year 2023 with contract number 27.29/UN23.37/PT.01.03/II/2023. High appreciation is also addressed to Dr.rer.nat. Saefuddin Aziz, for data analysis assistance.

REFERENCES

Aruna MS, Prabha MS, Priya NS, Nadendla R. 2015. *Catharanthus roseus*: Ornamental plant is now a medicinal boutique. J Drug Deliv Ther 5 (3): 1-4. DOI: 10.22270/jddt.v5i3.1095.

Bai X, Wang G, Ren Y, Su Y, Han J. 2023. Insights into the taxonomy and phylogenetic relationships of eleven *Aristolochia* species based on chloroplast genome. Front Plant Sci 14: 1119041. DOI: 10.3389/fpls.2023.1119041.

Barberá P, Soreng RJ, Peterson PM, Romaschenko K, Quintanar A, Aedo C. 2020. Molecular phylogenetic analysis resolves *Trisetum* (Poaceae: Pooideae: Koeleriinae) polyphyletic: Evidence for a new genus, *Sibirotrisetum* and resurrection of *Acrospelion*. J Syst Evol 58 (4): 517-526. DOI: 10.1111/jse.12523.

Chen C-M, Yeh D-M. 2012. "Taoyuan No. 1 Rose Girl": A double-flowered periwinkle, *Catharanthus roseus*. HortScience 47 (8): 1175-1176. DOI: 10.21273/hortsci.47.8.1175.

Cheong WY, Kim S-H, Yang JY, Lee W, Pak J-H, Kim S-C. 2020. Insights from chloroplast DNA into the progenitor-derivative relationship between *Campanula punctata* and *C. takesimana* (Campanulaceae) in Korea. J Plant Biol 63: 431-444. DOI: 10.1007/s12374-020-09281-3.

Das S, Sharangi AB. 2017. Madagascar periwinkle (*Catharanthus roseus* L.): Diverse medicinal and therapeutic benefits to humankind. J Pharm Phytochem 6 (5): 1695-1701.

de Camargo Smidt E, Páez MZ, do Nascimento Vieira L, Viruel J, de Baura VA, Balsanelli E, de Souza EM, Chase MW. 2020. Characterization of sequence variability hotspots in Cranichideae plastomes (Orchidaceae, Orchidoideae). PLoS One 15 (1): e0227991. DOI: 10.1371/journal.pone.0227991.

Dhyani P, Quispe C, Sharma E, Bahukhandi A, Sati P, Attri DC, Szopa A, Sharifi-Rad J, Docea AO, Mardare I, Calina D, Cho WC. 2022. Anticancer potential of alkaloids: A key emphasis to colchicine, vinblastine, vincristine, vindesine, vinorelbine and vincamine. Cancer Cell Intl 22 (1): 206. DOI: 10.1186/s12935-022-02624-9.

Doyle JJ, Doyle JL. 1990. A rapid total DNA preparation procedure for fresh plant tissue. Focus 12 (1): 13-15.

El-Sherif N, Ibrahim M. 2020. Implications of rbcl and rpoC1 DNA barcoding in phylogenetic relationships of some Egyptian *Medicago sativa* L. cultivars. Egypt J Bot 60 (2): 451-460. DOI: 10.21608/ejbo.2020.20028.1399.

Hall TA. 1999. BIOEDIT: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/ NT. Nucleic Acids Symp Ser 41: 95-98.

Kang L, Xie D, Xiao Q, Peng C, Yu Y, He X. 2019. Sequencing and analyses on chloroplast genomes of *Tetrataenium candicans* and two allies give new insights on structural variants, DNA barcoding and phylogeny in Apiaceae subfamily Apioideae. PeerJ 7: e8063. DOI: 10.7717/peerj.8063.

Kang Y, Deng Z, Zang R, Long W. 2017. DNA barcoding analysis and phylogenetic relationships of tree species in tropical cloud forests. Sci Rep 7: 12564. DOI: 10.1038/s41598-017-13057-0.

Ku C, Chung W-C, Chen L-L, Kuo C-H. 2013. The complete plastid genome sequence of Madagascar periwinkle *Catharanthus roseus* (L.) G. Don: Plastid genome evolution, molecular marker identification, and phylogenetic implications in Asterids. PLoS One 8 (6): e68518. DOI: 10.1371/journal.pone.0068518.

Lee Y-E, Lee Y, Kim S. 2021. A report of the second chloroplast genome sequence in *Veronica nakaiana* (Plantaginaceae), an endemic species in Korea. Korean J Plant Taxon 51 (1): 109-114. DOI: 10.11110/kjpt.2021.51.1.109.

Long DT, Hong HTK, Tram LLT. 2022. Analysis genetic variation and phylogenetic relationships of local lotus collected in Thua Thien hue province, Vietnam by DNA barcoding. Res J Biotechnol 17 (2): 48-56. DOI: 10.25303/1702rjbt4856.

Lv Y-N, Yang C-Y, Shi L-C, Zhang Z-L, Xu A-S, Zhang L-X, Li X-L, Li H-T. 2020. Identification of medicinal plants within the Apocynaceae family using ITS2 and psbA-trnH barcodes. Chin J Nat Med 18 (8): 594-605. DOI: 10.1016/S1875-5364(20)30071-6.

Makki RM, Saeedi AA, Khan TK, Ali HM, Ramadan AM. 2019. Single nucleotide polymorphism analysis in plastomes of eight *Catharanthus roseus* cultivars. Biotechnol Biotechnol Equip 33 (1): 419-428. DOI: 10.1080/13102818.2019.1579671.

Martínez-Vega AL, Oregel-Zamudio E, García-Ruíz I, Villapando-Arteaga EV, Torres-García JR. 2022. Genetic and metabolomic differentiation of *Physalis ixocarpa* Brot. ex Hornem. populations in Michoacan State, Mexico. Genet Resour Crop Evol 69: 1867-1877. DOI: 10.1007/s10722-022-01347-0.

Miguez M, Bartolucci F, Jiménez-Mejías P, Martín-Bravo S. 2022. Re-evaluating the presence of *Carex microcarpa* (Cyperaceae) in Italy based on herbarium material and DNA barcoding. Plant Biosyst 156 (3): 628-634. DOI: 10.1080/11263504.2021.1897703.

Mishra JN, Verma NK. 2017. A brief study on *Catharanthus roseus*: A review. Intl J Res Pharm Pharm Sci 2 (2): 20-23.

Neves B, Kessous IM, Moura RL, Couto DR, Zanella CM, Antonelli A, Bacon CD, Salgueiro F, Costa AF. 2021. Pollinators drive floral evolution in an Atlantic Forest genus. AoB Plants 12 (5): plaa046. DOI: 10.1093/aobpla/plaa046.

Noormohammadi Z, Asghari-Mooneghi N, Farahani F. 2020. Effect of gamma radiation on morphological and genetic variation in regenerated plantlets *Catharanthus roseus* (L.) G Don. Genetika 52 (1): 15-28. DOI: 10.2298/gensr2001015n.

Noormohammadi Z, Taban M, Farahani F. 2018. Short Communication: The impact of gamma radiation on *tdc* and *str* gene expressions in *Catharanthus roseus* regenerated plantlets. Biodiversitas 19 (5): 1805-1810. DOI: 10.13057/biodiv/d190530.

Paarakh MP, Swathi S, Taj T, Tejashwini V, Tejashwini B. 2019. *Catharanthus roseus* Linn - A review. Acta Sci Pharm Sci 3 (10): 19-24. DOI: 10.31080/asps.2019.03.0393.

Pan Q, Mustafa NR, Tang K, Choi YH, Verpoorte R. 2016. Monoterpenoid indole alkaloids biosynthesis and its regulation in *Catharanthus roseus*: A literature review from genes to metabolites. Phytochem Rev 15: 221-250. DOI: 10.1007/s11101-015-9406-4.

Park H-S, Lee WK, Lee S-C, Lee HO, Joh HJ, Park JY, Kim S, Song K, Yang T-J. 2021. Inheritance of chloroplast and mitochondrial genomes in cucumber revealed by four reciprocal F1 hybrid combinations. Sci Rep 11 (1): 2506. DOI: 10.1038/s41598-021-81988-w.

Peterson PM, Romaschenko K, Arrieta YH. 2022a. A phylogeny of the Triraphideae including *Habrochloa* and *Nematopoa* (Poaceae, Chloridoideae). PhytoKeys 194: 123-133. DOI: 10.3897/phytokeys.194.80967.

Peterson PM, Romaschenko K, Herrera-Arrieta Y, Vorontsova MS. 2022b. Phylogeny, classification, and biogeography of *Afrotrichloris*, *Apochiton*, *Coelachyrum*, *Dinebra*, *Eleusine*, *Leptochloa*, *Schoenefeldia*, and a new genus, *Schoenefeldiella* (Poaceae: Chloridoideae: Cynodonteae: Eleusininae). J Syst Evol 60 (3): 630-639. DOI: 10.1111/jse.12803.

Shadrin D, Dalke I, Zakhozhiy I, Shilnikov D, Kozhin M, Chadin I. 2023. *Heracleum sosnowskyi* or *Heracleum mantegazzianum*? DNA-based identification of invasive hogweeds (Apiaceae) in two key regions of the species' invasion history in the territory of the former Soviet Union. Res Sq 2023: 1-27. DOI: 10.21203/rs.3.rs-3296382/v1.

Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. Am J Bot 94 (3): 275-288. DOI: 10.3732/ajb.94.3.275.

Song Y, Zhao W, Xu J, Li MF, Zhang Y. 2022. Chloroplast genome evolution and species identification of *Styrax* (Styracaceae). BioMed Res Intl 2022: 5364094. DOI: 10.1155/2022/5364094.

Sung J-S, Chung J-W, Lee G-A, Kang M-J, Lee S-Y, Huh M-K. 2013. Genetic diversity and phenetic relationship of dill (*Anethum graveolens* L.) by *rps16-trnK* DNA sequences. J Life Sci 23 (11): 1305-1310. DOI: 10.5352/jls.2013.23.11.1305.

Tamura K, Stecher G, Kumar S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol 38 (7): 3022-3027. DOI: 10.1093/molbev/msab120.

Tang P, Xu Q, Shen R, Yao X. 2019. Phylogenetic relationship in *Actinidia* (Actinidiaceae) based on four noncoding chloroplast DNA sequences. Plant Syst Evol 305: 787-796. DOI: 10.1007/s00606-019-01607-0.

Thompson JD, Higgins DG, Gibson TJ. 1994. ClustalW: Improving the sensitivity of progressive multiple sequence aligment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res 22 (22): 4673-4680. DOI: 10.1093/nar/22.22.4673.

Vu PTB, Cao DM, Bui AL, Nguyen NN, Van Bui L, Quach PND. 2022. In vitro growth and content of vincristine and vinblastine of *Catharanthus roseus* L. hairy roots in response to precursors and elicitors. Plant Sci Today 9 (1): 21-28. DOI: 10.14719/pst.1337.

Wang Y, Wang S, Liu Y, Yuan Q, Sun J, Guo L. 2021. Chloroplast genome variation and phylogenetic relationships of *Atractylodes* species. BMC Genomics 22 (1): 103. DOI: 10.1186/s12864-021-07394-8.

Wei LM. 2022. Metabolite profiling and DNA barcoding analysis of 35 Malaysian medicinal plants. [Dissertation]. Faculty of Medicine and Health Sciences Universiti Tunku Abdul Rahman, Malaysia.

Xie X, Huang R, Li F, Tian E, Li C, Chao Z. 2021. Phylogenetic position of *Bupleurum sikangense* inferred from the complete chloroplast genome sequence. Gene 798: 145801. DOI: 10.1016/j.gene.2021.145801.

Zhang L, Wang S, Su C, Harris AJ, Zhao L, Su N, Wang J-R, Duan L, Chang Z-Y. 2021. Comparative chloroplast genomics and phylogenetic analysis of *Zygophyllum* (Zygophyllaceae) of China. Front Plant Sci 12: 723622. DOI: 10.3389/fpls.2021.723622.

Zhang Y-W, Kong X-Y, Wang J-H, Du G-H. 2018. Vinblastine and vincristine. In: Du G-H (eds). Natural Small Molecule Drugs from Plants. Springer, Singapore. DOI: 10.1007/978-981-10-8022-7_91.

Zhu S, Liu Q, Qiu S, Dai J, Gao X. 2022. DNA barcoding: An efficient technology to authenticate plant species of traditional Chinese medicine and recent advances. Chin Med 17 (1): 112. DOI: 10.1186/s13020-022-00655-y.