Original Article

Analysis of DNA Markers from Vietnamese Asarum L. Species

Huynh Thi Thu Hue1,2*, Le Dinh Ky3, Nguyen Huy Hoang1,2

1Institute of Genome Research, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam
2Graduate University of Science and Technology, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam
3Global KB Thien Phu Technology Company Limited, Gia Lam, Hanoi, Vietnam

Received 19 September 2022
Revised 26 October 2022; Accepted 28 November 2022

Abstract: The Asarum L. genus includes 128 species with many medicinal effects. It has been used extensively in folk prescriptions to treat several diseases such as bronchitis, asthma, rheumatism, and hepatitis. Many kinds of research showed that the Asarum L. species have many essential oils such as myristicin, methyl eugenol, myrcene, asafrol, borneol, safrol, or α-pinene, and substances such as sesquiterpenes, sterols, naringenin, or glycosyl flavonoids. Some reports in Vietnam have been conducted to analyze the secondary compounds of these species. However, little research has been done on the classification of these species through molecular markers. Indeed, DNA barcoding will help to overcome the difficulties in recognizing many plants that are endangered or rare and have high value in medicine. In the study, four DNA barcoding regions - ITS1, ITS2, matK and rpoC were analyzed from three Asarum species - A. splendens, A. yentuense and A. petelotii. The four barcodes were well amplified by PCR and sequenced by the Sanger principle. The results of sequence analysis of ITS1, ITS2 matK, and rpoC showed that ITS1 and ITS2 had the highest polymorphism. In addition, the ITS and matK markers were suitable for distinguishing the Asarum L. species. These results contribute to application of DNA barcodes in identification of Asarum species.

Keywords: Asarum, DNA barcoding, ITS, matK, rpoC.

1. Introduction

The Asarum L. genus of Aristolochiaceae includes 128 species [1] that grow diversely in large areas. The genus has many species native to Southeast Asia, China, and Taiwan, and only one species appears in Europe. By 2017, ten Asarum species were recorded in Vietnam, distributed in many provinces such as Lai Chau, Cao Bang, Vinh Phuc, Lao Cai, Quang Ninh, etc. The Asarum species in Vietnam include...
A. blumei, first described by Duch in 1864; it was named Blume [2]; in 2018, a Vietnamese species - A. yentuense was described by Nguyen Anh Tuan [3]. In recent years, Asarum species have drawn the interest of scientists due to their biological activities. This genus belongs to the group of medicinal plants most abundant in essential oils [4], in which the amount of essential oil is usually around 2% - 3% depending on the species. Many kinds of research mainly focus on a few common species, such as A. sieboldii or A. heterotropoides or A. forbesii. According to Zhang 2005, A. forbesii plant contains two main components: methyl isoeugenol (33.3%) and alpha-asarone (19.2%) in leaves, and the proportions are 58.8% and 10.3%, respectively, in roots [5].

Asarinin and sesamin compounds in Asarum L. have a pharmacological effect on the dopamine system - the most abundant catecholamine neurotransmitter in the brain. Asarinin and sesamin stimulate the dopamine biosynthesis and protect PC12 cells from neurotoxic L-DOPA and inhibit 6-hydroxydopamine-induced cytotoxicity by modulating the regulatory kinase signaling system of extracellular signaling (ERK1/2) in PC12 cells [6]. In 2008, Shao-Qing Cai et al., evaluated the cytotoxic activity of Asarum L. plants and discovered that some extracts from these plants were effective against some cancer cell lines. Notably, the 95% ethanol or water extracts of A. splendens also exhibited cytotoxicity on cancer cells [7]. In the mountainous areas of Vietnam, all parts of roots, stems and leaves of Asarum L. are used to treat many diseases such as bronchitis, asthma, typhoid, rheumatism, and hepatitis in some folk prescriptions. In A. petelotii, the composition of myristicin accounts for 33.11% of the stem oil content, 84.60% of the root oil content, and 40.15% of the total oil content of the entire plant, while the safrole in the whole plant is very low, accounting for only 0.79% [8]. Moreover, recent studies show that the secondary compounds of Asarum in Vietnam are relatively diverse, including substances with anti-inflammatory, antibacterial, antihistamine effects or potentially toxic to cancer cells [9, 10].

In modern biotechnology, the application of molecular biology using specific DNA regions called DNA barcoding to distinct and compare between species is in development. The method has been widely used for genetic identification with the advantages of fast speed and high accuracy, and has great application in the assessment and classification of biodiversity and conservation of plant genetic resources. This approach can be applied to many subjects, from animals and plants to microorganisms, based on DNA sequences in the nucleus and mitochondria or chloroplasts. In plants, the common regions known as matK, rbcL [11], rpoC1, rpoB [12], belonging to the DNA plastid, or the internal transcribed spacer (ITS) sequence in the nuclear genome, are utilized. ITS regions have a fast evolutionary rate so it expresses the diversity more clearly than DNA plastid, and is being proposed as a standard DNA marker for molecular identification of plants.

In the present study, four DNA barcoding regions - ITS1, ITS2, matK and rpoC - were analysed from four samples of A. splendens, A. yentuense and A. petelotii. The compared nucleotide sequences of the markers can contribute to DNA barcoding data for the Asarum L. species of Vietnam.

2. Materials and Methods

2.1. Materials

One sample of A.splendens named LC was collected at Lai Chau province; two samples of A. yentuense were collected at Yen Tu, Quang Ninh province named YT, and a sample named BPS that grew at Bac Phong Sinh, Uong Bi, Quang Ninh province; and the last one is A. petelotii found at Tam Dao, Vinh Phuc province, named TD [2].
2.2. Methods

Total DNA from four samples was extracted using Cetyltrimethylammonium bromide (CTAB) [13]. The specific primer pairs ITS1, ITS2, matK and rpoC were designed based on sequences KM982345.1 and JQ886464.1 on GenBank (Table 1) according to Chang et al., [14].

The PCR to amplify ITS, rpoC and matK gene fragments was performed in a 25 µl reaction with 1 µl DNA template (50 ng), 1 µl of each primer (10 µM), 12.5 µl DreamTaq master mix (2X), and 10.5 µl ddH2O. The PCR program comprised an initial denaturation step at 95 °C for 3 min, followed by 35 cycles (95°C/30s; 58°C/30s; 72°C/45s) and a final extension step at 72 °C for 5 min. The amplified products were detected on 0.8% agarose gels by electrophoresis, then purified by GeneJET Gel Extraction Kit (Thermo Scientific, USA). The PCR products were sequenced by the ABI 3500 Genetic Analyzer system based on Sanger's principles, with the BigDye®Terminator v 3.1 Cycle Sequencing kit (Applied Biosystems, USA).

The DNA barcoding sequences were processed and analyzed using BioEdit software v7.0.5.9, in which comparisons with the corresponding sequences were published on GenBank. The similarity percentage parameter between the species was determined using the MEGAX software. The phylogenetic tree was performed by the Maximum-Likelihood approach of the MEGAX software with the bootstrap of 1,000 replications [15].

3. Results and Discussion

3.1. DNA Extraction

In the research, total DNA of four Asarum L. samples were extracted from 100 mg of leaf by the CTAB method as mentioned. CTAB extraction buffer was adjusted and combined with a mixture of phenol: chloroform: isoamylalcohol (25:24:1) solution to remove proteins. Then, DNA was precipitated in 100% EtOH. The genomic DNAs observed in agarose gel were a sharp, clear DNA band without breaking, and RNA has been removed (Figure 1), it means that total DNA is in good condition. Because Asarum L. leaves contain abundant secondary compounds, extraction of genomic DNA from this plant was difficult. We had to improve the extraction process by repeating the extraction step with chloroform: isoamylalcohol to get sufficient quality of the DNA. As a result, all DNA samples had the ratio A260/A280 between 1.8 - 2.0 values (Figure 1).

3.2. Amplification of DNA Markers ITS1, ITS2, MatK and RpoC

Primers used to amplify fragments of ITS1, ITS2, matK and rpoC markers of the four Asarum samples were indicated in Table 1. The results showed that the fragments of ITS1

### Table 1. List of primers used in the study

<table>
<thead>
<tr>
<th>No.</th>
<th>Primer name</th>
<th>Sequence (5'→3')</th>
<th>Estimated size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ITS1-F</td>
<td>CCTTATCAYTTAGAGGAAGGAG</td>
<td>300-350</td>
</tr>
<tr>
<td>2</td>
<td>ITS1-R</td>
<td>GCCRAGATATCCGTTGCGAG</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>ITS2-F</td>
<td>YGACTCTCGGCAACGGATA</td>
<td>450-500</td>
</tr>
<tr>
<td>4</td>
<td>ITS2-R</td>
<td>CCGCTTAKTGTATATGCTTAAA</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>matK-F</td>
<td>ACCGTACTTTGTTGACGAC</td>
<td>880</td>
</tr>
<tr>
<td>6</td>
<td>matK-R</td>
<td>TCCATCTRGAAATMTTRGTTCA</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>rpoC-F</td>
<td>GGCAAAAGGGAAGATTTCG</td>
<td>550</td>
</tr>
<tr>
<td>8</td>
<td>rpoC-R</td>
<td>CCATAAGCATATCTTGGTTGG</td>
<td></td>
</tr>
</tbody>
</table>
(0.3kb), ITS2 (0.4kb), rpoC (0.5 kb) and matK (0.8kb) were specifically amplified from all DNA templates. The products were purified by GeneJET Gel Extraction Kit (Figure 2) for sequencing by the Sanger principle.

Figure 1. Agarose electrophoresis of total DNAs extracted from four *Asarum* L. samples. 

Figure 2. Agarose electrophoresis of purified PCR products of nuclear and chloroplast markers. 
*Note*: a) Nuclear marker b) Chloroplast marker. 

### 3.3. Analysis of ITS1, ITS2, MatK and RpoC Sequences

Nucleotide sequences of the four markers were determined by the Sanger method. After the bad quality regions were removed from both ends of each sequence, the lengths of the remaining sequences of ITS1, ITS2, matK and rpoC were 240 bp, 339 bp, 820 bp and 474 bp respectively. All sequences were edited and aligned using ClustalW on Bioedit software combining with the DNA sequences downloaded from the GenBank database. The analysis on each marker includes 12 sequences: four sequences generated from this study, seven references in *Asarum* L., and an outgroup.

As observed with ITS1, the Lai Chau sample (*A. splendens*) had the highest nucleotide similarity with JF975949.1 (*A. splendens*), followed by JF975922.1 (*A. delavayi*). In addition, it maintained several diverse positions compared to all other species such as 96 C>T, 137 A>G, 143 A>G, 184 C>T, 214 C>T. The two samples collected at Bac Phong Sinh and Yen Tu (*A. yentuense*) share the highest homology (100%) (Table 2) and both have a unique T at 111 position that may be a distinctive feature of the samples. It was explained that the geographical proximity determined this homology. In contrast, the nucleotide sequence of the Tam Dao sample was significantly different compared with other sequences. These differences are in some positions such as 47, 208, 213, 216 and 240. Thus, there are a lot of polymorphisms in the ITS1 region of the *Asarum* species in research.

For ITS2 marker, the polymorphic regions mainly occurred from 205 to 255 positions. Both Bac Phong Sinh and Yen Tu (*A. yentuense*) samples still showed high similarities with each other (99.7%) (Table 2). The Tam Dao sample (*A. petelotii*) also had many similar positions to the reference LC530386.1 (*A. petelotii*) and differ from other sequences (208 A>T, 222 C>T, 228 C>T, addition of AG at 252, 253, and 255 C>T) interestingly, most of the differences were mutations C>T. It showed a clear difference of the *A. petelotii* compared with other species. The *A. splendens* sample had the highest sequence similarity with JF975949.1 (*A. splendens*) and JF975922.1 (*A. delavayi*) among the compared sequences. The matK region showed a higher conservative level than the ITS fragments; the homology is over 99% between all of them (Table 2). However, there were still some differences in the sequence of the studied samples. For example, the Lai Chau
sample (*A. splendens*) had some distinctive nucleotides which differ from most samples (218 C > T, 300 C > T, 453 G > T and 494 A > C) so it could be unique polymorphisms of *A. splendens* species. Likewise, the *A. yentuense* also had some unique polymorphisms (110 T>G and 786 A>C). Nevertheless, the rpoC region is the most conservative sequence. When compared with the references from the GenBank database, there was no difference in the rpoC region (Table 2). This demonstrated that the region is highly conserved and could not be a marker to differentiate species in *Asarum* L.

Table 2. The similarity percentage of the ITS1, ITS2, matK and rpoC sequences

<table>
<thead>
<tr>
<th>ITS1</th>
<th>LC</th>
<th>BPS</th>
<th>TD</th>
</tr>
</thead>
<tbody>
<tr>
<td>LC</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BPS</td>
<td>71.73</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TD</td>
<td>93.67</td>
<td>67.09</td>
<td>-</td>
</tr>
<tr>
<td>YT</td>
<td>71.73</td>
<td>100.00</td>
<td>67.09</td>
</tr>
<tr>
<td>ITS2</td>
<td>LC</td>
<td>BPS</td>
<td>TD</td>
</tr>
<tr>
<td>LC</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BPS</td>
<td>99.08</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TD</td>
<td>79.03</td>
<td>78.42</td>
<td>-</td>
</tr>
<tr>
<td>YT</td>
<td>98.78</td>
<td>99.69</td>
<td>78.12</td>
</tr>
<tr>
<td>matK</td>
<td>LC</td>
<td>BPS</td>
<td>TD</td>
</tr>
<tr>
<td>LC</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BPS</td>
<td>99.02</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TD</td>
<td>99.27</td>
<td>99.51</td>
<td>-</td>
</tr>
<tr>
<td>YT</td>
<td>99.15</td>
<td>99.88</td>
<td>99.63</td>
</tr>
<tr>
<td>rpoC</td>
<td>LC</td>
<td>BPS</td>
<td>TD</td>
</tr>
<tr>
<td>LC</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BPS</td>
<td>100.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TD</td>
<td>100.00</td>
<td>100.00</td>
<td>-</td>
</tr>
<tr>
<td>YT</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

The phylogeny tree generated from nuclear marker (ITS1 and ITS2) by the Maximum-Likelihood algorithm showed the close relationship between two of *A. yentuense*, as well as with *A. hongkongense*, *A. sagittarioides*, and *A. longerhizomatus* species. The clearly separate *A. petelotii* are based on ITS1 and ITS2 with a reliable bootstrap index. It means that there is an advantageous distinction between the species when using nuclear markers of ITS1 and ITS2 (Figure 3, 4). Among the two chloroplast markers, matK also showed a clear separation between samples. However, the relationship of species in *Asarum* L. has not been clearly separated, as in *A. yentuense* with *A. longerhizomatus*. The *A. splendens* belonged to the same clade as *A. petelotii*, *A. sagittarioides* and *A. hongkongense*. Similar to the nuclear marker, the close relationship of the *A. splendens* and LC530481.1 (the *A. splendens* sequence from GenBank or *A. chingchense*) had the highest bootstrap value in the matK taxonomic tree. However, the rpoC marker was highly conserved so the phylogeny tree could not distinguish the species (Figure 5, 6). As a result, the two nuclear markers and plastic- matK marker were absolutely suitable for distinguishing the species.

In the study, using ITS1, ITS2 and matK regions clearly distinguished between *A. yentuense*, *A. splendens* and *A. petelotii* as well as from other species. Because of convenience in short length and easy amplification, ITSs were widely applied in phylogeny and taxonomy analysis both in fungi, monocot or dicot [11, 16, 17]. The result showed that ITS1, ITS2 are more suitable for the *Asarum* research. This marker was also useful in plant DNA barcodes such as in *Physalis* of Solanaceae [18]. Similarly, many researchers used the ITS for phylogeny analysis. By using this marker, the Ehreatiaceae was divided to four groups, and *E. asperula* belong to Ehretia I group [19]. Other examples, *Panax vietnamensis* and *Panax stipuleanatus* were clearly distinguished by using ITS regions [20]. Moreover, ITS and 5.8 S regions were believed to have great significance in evolutionary and phylogenetic, and matK is also effectively used in taxonomy research [21].

In recent years, for the identification of herbs, many researches combined ITS2 and other marker for instance *psbA-trnH* to
determine raw materials for plant-originated medicine [22, 23] or other plant as Duckweed [24]. In the report of Wu et al., twenty-eight *Asarum* samples were collected from herbal markets which were assessed by ITS2 and *psbA-trnH* barcodes. The result showed that only twenty samples were *Asarum* species, the rest belonged to *Stephania tetranda*, *Aristolochia fangchi* and *Aristolochia tuberosa* [25]. The DNA marker application significantly helped to increase the accurate usage of herbs in medical treatment. Notably, due to its importance for the accurate use of the herbs, the assessment using DNA barcoding is extremely necessary specifically for the *Asarum* species.

**Figure 3.** Phylogeny tree generated from ITS1. J428635.1 (A. hongkongense), FJ428634.1 (A. longerhizomatosum), LC530386.1 (A. petelotii), FJ428633.1 (A. sagittarioides), FJ980374.1 (A. maximum), JF975922.1 (A. delavayi) JF975949.1 (A. splendens), MH711018.1 (Saruma henryi).

**Figure 4.** Phylogeny tree generated from ITS2. FJ428635.1 (A. hongkongense), FJ428634.1 (A. longerhizomatosum), LC530386.1 (A. petelotii), FJ428633.1 (A. sagittarioides), FJ980374.1 (A. maximum), JF975922.1 (A. delavayi) JF975949.1 (A. splendens), MH711018.1 (Saruma henryi).

**Figure 5.** Phylogeny tree generated from matK. FJ428687.1 (A. hongkongense), FJ428693.1 (A. longerhizomatosum), LC530482.1 (A. petelotii), FJ428690.1 (A. sagittarioides), LC530438.1 (A. maximum), MG554427.1 (A. delavayi), LC530481.1 (A. splendens), LC068520.1 (Saruma henryi).

**Figure 6.** Phylogeny tree generated from *rpoC*. GQ436064.1 (A. maximum), MG554429.1 (A. delavayi), MK577409.1 (A. heterotropoides), MG554458.1 (A. megacalyx), MG554436.1 (A. sieboldii), MG554425.1 (A. minus), LC529900.1 (A. forbesii), NC_039933.1 (Saruma henryi).
4. Conclusion

We successfully sequenced four DNA barcodes - ITS1, ITS2, matK and rpoC - from samples belonging to A. splendens, A. yentuense and A. petelotii. ITS1 and ITS2s showed the highest polymorphism. ITSs and matK markers were suitable for distinguishing the *Asarum* L. species.

References

https://doi.org/10.1371/journal.pone.0004598.


