Malaysian Journal of Chemistry, 2021, Vol. 23(2), 110-114
(Special Issue: *IKM Pahang Branch Online Symposium 2021 – Chemistry for Sustainable World*)

Isolation and Cloning of Sesquiterpene Synthases (*AmGS3* and *AmGS4*) and Chalcone Synthase (*AmCHS*) from *Aquilaria malaccensis* Responsible for Agarwood Formation

Aimi Wahidah Aminan[1], Siti Zulaiha Zailani[1], Saiful Nizam Tajuddin[2], Aizi Nor Mazila Ramli[1,2*]
[1]Faculty of Industrial Science & Technology, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Kuantan, Pahang Malaysia
[2]Bio Aromatic Research Centre of Excellent, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Kuantan, Pahang Malaysia
*Corresponding author (e-mail: aizinor@ump.edu.my)

Sesquiterpene and phenylethyl chromone, two types of agarwood marker compounds, have been extensively studied. However, genetic studies of agarwood (*Aquilaria malaccensis*) are still scarce. This study describes the isolation and cloning of sesquiterpene synthase genes (*AmGS3* and *AmGS4*), and chalcone synthase gene (*AmCHS*) identified from *A. malaccensis* transcriptome data mining. The sizes of *AmGS3*, *AmGS4*, and *AmCHS* were 1162, 1466, and 1623 bp in length. The open reading frames (ORFs) of *AmGS3*, *AmGS4*, and *AmCHS* detected were 948, 1047, and 1185 bp, with a polypeptide length of 348, 315, and 394 amino acids. The full-length sequences of *AmGS3*, *AmGS4*, and *AmCHS* were successfully isolated from the infected stem of *A. malaccensis*, amplified via polymerase chain reaction (PCR), cloned into the pGEM-T Easy Vector, and transformed into prepared *Escherichia coli* DH5a competent cells. The sequencing result and BLASTn analysis revealed that the ORFs of *AmGS3* and *AmGS4* are highly homologous to putative delta-guaiene synthase from *Aquilaria sinensis*, with a similarity of 98.1% and 98.08% respectively, while the ORF of *AmCHS* is highly homologous to chalcone synthase from *A. sinensis* with a similarity of 99.24%. These results demonstrated the successful isolation of sesquiterpene synthase and chalcone synthase genes that may play important roles in forming agarwood sesquiterpene and phenylethyl chromone in *A. malaccensis*.

Key words: *Aquilaria malaccensis*; agarwood; sesquiterpene synthase; chalcone synthase

Received: November 2020; Accepted: January 2021

Aquilaria, or the *karas* tree, originating from the family of Thymelaeaceae, is known as an important source of agarwood production [1] . Several *Aquilaria* species, including *Aquilaria crassna*, *Aquilaria sinensis*, and *Aquilaria malaccensis* are known as agarwood-producing plants. Among those species, *A. malaccensis* is an industrial crop in Malaysia, a major agarwood producer. [2]. *Aquilaria* sp. has been listed in Appendix II of the Convention on International Trade in Endangered Species of Wild Fauna and Flora since 2005 due to its near extinction in the wild [3-5]. Due to depleting *Aquilaria* sources, the international import and export of agarwood products are strictly controlled [6]. This paves the way for researchers to produce artificial agarwood efficiently. Agarwood is a dark-coloured resin that accumulates in a chip wood region of the stem, branch, or root of *Aquilaria* and Gyrinops trees [7]. Agarwood emits a unique scent when burnt, and comprises aromatic plant materials and volatile essential oils [8]. The resin's unique fragrance properties are in great demand in the perfumery industry [9]. High-quality agarwood has great economic value, which could reach up to US$1,000, and is costlier than gold [7]. Nevertheless, the infrequent and extended time period for natural agarwood production in older trees cannot cope with market demands [3,10]. Due to the discovery of sesquiterpene and phenylethyl chromones as major agarwood constituents [5,11], more in-depth studies have come up with the notion of focusing on the molecular aspects for the biosynthesis of specific compounds [3]. Information on the biosynthesis of important agarwood compounds, namely sesquiterpene and chromone derivatives, is critical for the artificial synthesis of agarwood.

Studies of *A. malaccensis* at the genetic level are still lacking. Although the generation of sesquiterpene has been identified in the last step of two important pathways, i.e. the mevalonic acid (MVA) and 1-deoxy-D-xylulose- 5-phosphate (DXP) pathways, information on the biosynthesis of chromone derivatives in agarwood formation is almost unknown. To date, full-length sesquiterpene synthases (*AmSesTPS1* and *AmGuaiS1*) have been reportedly isolated from callus samples of *A. malaccensis* by Siah

†Paper presented at the IKM Pahang Branch Online Symposium 2021 – Chemistry for Sustainable World

et al. (2016) [12]. However, since agarwood is found abundantly in the stem, a study of the isolation of sesquiterpene synthase from stem samples might provide new perspective on sesquiterpene synthase. Additionally, the presence of chromones in *A. malaccensis* agarwood has not been reported. Therefore, this study describes the cDNA isolation and cloning of the sesquiterpene synthases (*AmGS3* and *AmGS4*) and chalcone synthase (*AmCHS*) genes responsible for the formation of two key aromatic compounds, sesquiterpene and phenylethyl chromone of agarwood, from the stem sample of *A. malaccensis*. The results might provide a foundation for further exploration of agarwood formation in *A. malaccensis*.

MATERIALS AND METHODS

1. Plant Sample Collections

Samples of mature agarwood trees were collected from Kampung Kedaik, Rompin, Pahang. All the collected infected wood samples were tagged, wrapped in aluminium foil and stored temporarily in liquid nitrogen. The samples were taken to the laboratory for the isolation of total RNA.

2. Total RNA Isolation

Total RNA was extracted from the infected stem of *A. malaccensis* using RNaqueous extraction (Ambion, USA) and Ribospin™ (*GeneAll*®, Korea) plant kit extraction protocol. The purity and concentration of the total RNA were checked at $A^{260/280}$ and $A^{230/280}$ using the NanoDrop spectrophotometer. The RNA samples were subjected to gel electrophoresis on a 1.0% agarose gel concentration, at a voltage of 70V for 40 min.

3. First-strand cDNA synthesis

First-strand cDNA was synthesised using the GoScript™ Reverse Transcription System (Promega) kit. The RNA sample, primers, and water were mixed in a sterile PCR tube before incubation at 65 °C for 5 min and then placed on ice for 2 min. Next, the PCR master mix comprising 5× reaction buffer, 25 mM $MgCl_2$, 10 mM PCR nucleotide mix, and GoScript™ Reverse Transcriptase (Promega) was added to the mixture and incubated consecutively at 25 °C for 10 min, 42 °C for 30 min, and 85 °C for 5 min.

4. Candidate Gene Selection

The candidate gene selection was achieved by mining the *A. malaccensis* transcriptome data for transcripts related to the sesquiterpene and phenylethyl chromone biosynthetic pathways. The assembled transcripts were classified as sesquiterpene synthases and chalcone synthase, according to the homology search.

5. Isolation of Full-Length *AmGS3*, *AmGS4*, and *AmCHS*

The predicted open reading frames (ORFs) for *AmGS3*, *AmGS4*, and *AmCHS* were amplified by polymerase chain reaction (PCR) using a GoTaq® Flexi DNA Polymerase (Promega, USA) kit. The cDNA gene-specific primers were designed according to the ORF regions of the gene sequences retrieved from raw transcriptomic data of *A. malaccensis*. The Primer3 software (http://www.bioinformatics.nl/cgibin/primer3/primer3.cgi/) was used to check the suitability of the selected sequence. The Oligo calculate software (http://www.basic.northwestern.edu/biotools/oligocalc.html) was used to analyse all suitable primer pairs designed for the hairpin, palindromes, dimmers, and annealing temperature (Tm). The primer pairs used were *AmGS3*_F(5'ATGTTGCAAGCTTTACACCAACAGTG3'); *AmGS3*_R(5'TTAGATTTCAATAGCATGACGC3'); *AmGS4*_F(5'ATGCAAAGGCTGGAAGCAAGG3'); *AmGS4*_R(5'TCATATAGTAATTGGATGGACCAGC3'); *AmCHS*_F(5'ATGGCGGCCAAAGTGGAGGAGATCC3') and *AmCHS*_R(5'TCAATGAGCCGACTCGGTTGCTACAC3'). The PCR reaction mixture contained 1× reaction buffer, 2 mM $MgCl_2$, 0.2 mM dNTPs, 5 units of Taq polymerase, 0.5 µM of forward and reverse primers, and 20 ng of template cDNA. The reaction was performed under the following conditions: pre-denaturation at 98 °C for 30 s, followed by 32 cycles of 98 °C for 10 s, 60 °C for 10 s, and 72 °C for 20 s, with a final extension at 72 °C for 2 min.

6. Cloning of Recombinant *AmGS3*, *AmGS4*, and *AmCHS*

The amplicons were double-digested with SalI before being cloned into the pGEM-T Easy cloning vector (Promega). The ligation mixture was transformed into *Escherichia coli* DH5a competent cells. The positive transformants were screened on LB agar supplemented with 100 µg/mL ampicillin. The positive transformants were confirmed by colony PCR, and the gene sequences were verified via DNA sequencing (First BASE Laboratories, Seri Kembangan, Selangor, Malaysia).

7. Full-Length cDNA Sequence Analysis

The chromatogram file was checked and analysed via Chromas Lite software from the First BASE Company. The ORF for *AmGS3*, *AmGS4*, and *AmCHS* were predicted using the ORF Finder program (http://www.ncbi.nlm.nih.gov). Subsequently, the ORF sequences of *AmGS3*, *AmGS4*, and *AmCHS* were analysed with the basic local alignment search tool (BLASTn and BLASTx) to check for the description, query coverage, E-value, and max identity of the genes.

Table 1. The characteristics of the *A. malaccensis AmGS3*, *AmGS4*, and *AmCHS* genes obtained from BLASTn

Gene name	Description of the highest score hit	Score (bits)	E-value	Identity	Accession Number
AmGS3	*Aquilaria sinensis* delta-guaiene synthase (*ASS2*)	1648	0.0	98.10%	JQ712683.1
AmGS4	*Aquilaria sinensis*, putative delta-guaiene synthase (*SesTPS1*)	1810	0.0	98.08%	KM881472.1
AmCHS	*Aquilaria sinensis*, chalcone synthase (*CHS2*)	2130	0.0	99.24%	EF103197.1

RESULTS AND DISCUSSIONS

The purpose of isolating the *AmGS3*, *AmGS4*, and *AmCHS* genes from *A. malaccensis* was to verify and sequence their presence in the plant as the basis for further expression analysis. In this study, the cDNAs of *AmGS3*, *AmGS4*, and *AmCHS* were successfully amplified via PCR and cloned into pGEM-T Easy cloning vectors. The purpose of cDNA cloning is to improve the quality of DNA sequences to be applied in subsequent expression studies. All targeted gene sequences retrieved from sequencing services were analysed using the Basic Local Alignment Search Tool server (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to determine the identity of each isolated gene. According to nucleotide BLAST analysis, *AmGS3* gene has high similarity (98.10%) with the *A. sinensis* delta-guaiene synthase (*ASS2*) gene from the GenBank database, based on the pairwise alignment done with 0.0 Expect value. This E-value is a statistically significant BLAST analysis parameter, widely accepted as an indicator for assessing the potential biological relationship of the query sequence with the matched sequence [13]. Lower E-valuesdenote more significant hits.

The second targeted gene, *AmGS4*, is highly homologous to the *A. sinensis* putative delta-guaiene synthase (*SesTPS1*) gene, with similarities of 98.08%. The zero E-value indicates that the *AmGS3* and *AmGS4* genes belong to the terpene synthase (sesquiterpene) family. Sesquiterpene synthase genes are responsible for producing the agarwood sesquiterpene compound in agarwood-producing plants [14]. Production of these genes could be associated with the plant defence mechanism's response towards an infection.

The *A. malaccensis AmCHS* gene is highly similar to the *A. sinensis* chalcone synthase (*CHS2*) gene, with similarities of 99.24% from BLAST analysis. The 0.0 E-value shows that *AmCHS* could be identified as chalcone synthase gene. Previous studies showed that the *CHS* gene may regulate phenylethyl chromone biosynthesis via a phenylalanine metabolism pathway [15]. Moreover, *CHS* gene expression induced chromone accumulation with fungal induction in *Aquilaria* [11].

The nucleotide sequence analyses of *AmGS3*, *AmGS4* and *AmCHS* genes were then translated into

```
  1 ATGTTGCAAGCTTTACACCAAAGTGAACTAAGAGAAGCTTCTCGTTGGTGGAAGGAATTTGATTTTCCATCCAAGCTTCCTTATGCAAGA
     M  L  Q  A  L  H  Q  S  E  L  R  E  A  S  R  W  W  K  E  F  D  F  P  S  K  L  P  Y  A  R
 91 GACAGAATTGCTGAAGGCTACTACTGGATGATGGGTGCCCATTTTGAGCCTAAATTCTCTCTTAGTAGAAAATTTCTCAATAGAATAATT
     D  R  I  A  E  G  Y  Y  W  M  M  G  A  H  F  E  P  K  F  S  L  S  R  K  F  L  N  R  I  I
181 GGGATTACTTCTCTGATCGATGACACATATGATGTTTATGGCACATTGGAAGAAGTTACATTGTTCACTGAAGCAGTCGAGAGGTGGGAT
     G  I  T  S  L  I  D  D  T  Y  D  V  Y  G  T  L  E  E  V  T  L  F  T  E  A  V  E  R  W  D
271 ATTGAAGCTGTACAAGATATTCCTAAATACATGCAAGTAATCTACACTGGTATGTTGGGCATTTTTGAAGATTTCAAGGACAATCTGATC
     I  E  A  V  Q  D  I  P  K  Y  M  Q  V  I  Y  T  G  M  L  G  I  F  E  D  F  K  D  N  L  I
361 AATGCAAGAGGGAAAGACTATTGCATTGATTATGCGATAGAAGTGTTTAAGGAGATTGTCAGATCTTACCAAAGAGAAGCAGAATATTTC
     N  A  R  G  K  D  Y  C  I  D  Y  A  I  E  V  F  K  E  I  V  R  S  Y  Q  R  E  A  E  Y  F
451 CACACTGGATATGTGCCTAGTTATGATGAGTACATGGAGAACTCCATAATAAGTGGTGGGTACAAGATGTTCATTATTCTGATGTTGATC
     H  T  G  Y  V  P  S  Y  D  E  Y  M  E  N  S  I  I  S  G  G  Y  K  M  F  I  I  L  M  L  I
541 GGAAGGGGAGAGTTTGAACTCAAGGAAACTCTAGATTGGGCTTCAACAATCCCAGAAATGGTCAAAGCTTCTTCACTTATTGCTCGTTAT
     G  R  G  E  F  E  L  K  E  T  L  D  W  A  S  T  I  P  E  M  V  K  A  S  S  L  I  A  R  Y
631 ATTGATGACCTTCAGACATACAAGGCTGAAGAAAAAAGAGGGGAAACTGTTTCAGCGGTGCGGTGTTACATGAGGGAGTACGGTGTTTCA
     I  D  D  L  Q  T  Y  K  A  E  E  K  R  G  E  T  V  S  A  V  R  C  Y  M  R  E  Y  G  V  S
721 GAAGAAGAGTCATGCAAGAAGATGAGGGAGATGATTGAGATCGAGTGGAAGAGACTGAACAAGACGACCCTAGAAGCGAATGAAATCTCT
     E  E  E  S  C  K  K  M  R  E  M  I  E  I  E  W  K  R  L  N  K  T  T  L  E  A  N  E  I  S
811 TCGTCAGTTGTGATCCCGTCCCTGAATTTCACGCGAGTGTTGGAGGTGATGTACGATAAAGGTGATGGATATAGCGATTCCCAAGGTGTG
     S  S  V  V  I  P  S  L  N  F  T  R  V  L  E  V  M  Y  D  K  G  D  G  Y  S  D  S  Q  G  V
901 ACCAAAGATAGAATTGCCGCCTTGTTGCGTCATGCTATTGAAATCTAA
     T  K  D  R  I  A  A  L  L  R  H  A  I  E  I  *
```

Figure 1. The *AmGS3* Open Reading Frame (315 aa) detected by NCBI Open Reading Frame Finder. The codon in black indicates the ATG initiation codon and (*) indicates stop codon.

```
   1 ATGCAAAGGCTGGAAGCAAGGCAGTACATCCCCATTTACGAAGCAGACATGACCAAGAACATTTCTCTGCTGCATTTTGCAAAACTGGAT
     M  Q  R  L  E  A  R  Q  Y  I  P  I  Y  E  A  D  M  T  K  N  I  S  L  L  H  F  A  K  L  D
  91 TTTAATTTGCTACAAGCATTACATCAAAGTGAAATCAGAGAGATCACAAGATGGTGGAAGGATCTTGATTTCAAAACAAGGCTCCCATAT
     F  N  L  L  Q  A  L  H  Q  S  E  I  R  E  I  T  R  W  W  K  D  L  D  F  K  T  R  L  P  Y
 181 GCTAGAGATAGATTGGTGGAGTGTTACTTTTGGATTTTGGGAGTCCAATATGAGCCCCAATACTCCATGAGCAGATTGTTTCTCACAAAA
     A  R  D  R  L  V  E  C  Y  F  W  I  L  G  V  Q  Y  E  P  Q  Y  S  M  S  R  L  F  L  T  K
 271 GTTATATCTTTGGCATCTGTATTTGATGACACCTACGACATTTATGGAACTTTTGAAGAACTTAAGCTACTCACAGATGCAATTGAGAGG
     V  I  S  L  A  S  V  F  D  D  T  Y  D  I  Y  G  T  F  E  E  L  K  L  L  T  D  A  I  E  R
 361 TGGGAAATTGAAGCGACGGACAGTCTTCCTAGTTATATGCAAATTCTATATCGTGCCCTTTTGGATGTTTTTGATGAGTACAAGGACAAG
     W  E  I  E  A  T  D  S  L  P  S  Y  M  Q  I  L  Y  R  A  L  L  D  V  F  D  E  Y  K  D  K
 451 CTGATTAATGCAGAAGGCAAAGACTATTGTCTCTACTATGGAAAAGAAGCAATGAAAGGGCTCATTAGGTCTTACCACACTGAAGCAGTG
     L  I  N  A  E  G  K  D  Y  C  L  Y  Y  G  K  E  A  M  K  G  L  I  R  S  Y  H  T  E  A  V
 541 TCTTTCCACACTGGATATGTACAGAACTTCGAAGAATATTTGGACAACTCTGCAGTGAGCAGTGGCTACCCAATGCTCACGGTAGAAGCT
     S  F  H  T  G  Y  V  Q  N  F  E  E  Y  L  D  N  S  A  V  S  S  G  Y  P  M  L  T  V  E  A
 631 CTCATAGGCATGGGGCATCCATATGCAACCAAAGAAGCTCTTGATTGGGCTCTAAAAGTGCCAAGAGTAATCAAAGCTTCTTCTGATATT
     L  I  G  M  G  H  P  Y  A  T  K  E  A  L  D  W  A  L  K  V  P  R  V  I  K  A  S  S  D  I
 721 TGTCGTCTAGTGGATGATCTTCGAACGTACAAGGTTGAGGAAGAGAGAGGAGATGCTCCATCAGGTGTACACTGCTACATGAGGGATTAT
     C  R  L  V  D  D  L  R  T  Y  K  V  E  E  E  R  G  D  A  P  S  G  V  H  C  Y  M  R  D  Y
 811 AATGTTCCAGAAGAAGAGGCATGCTCCAAAATAGAGGAGATGATAGATTTAGCTTGGAAGGCAATTAATGAAGAAATGCAAAAGCCAGGA
     N  V  P  E  E  E  A  C  S  K  I  E  E  M  I  D  L  A  W  K  A  I  N  E  E  M  Q  K  P  G
 901 CATCTCCCACTACCAATTCTCTTACCGGCTTTGAATTTTACACGAATGATGGAGGCATTGTACCAGAACATAGATGGTTANTCCAACTCA
     H  L  P  L  P  I  L  L  P  A  L  N  F  T  R  M  M  E  A  L  Y  Q  N  I  D  G  X  S  N  S
 991 GGAGGGAGAACCAAGGACAGAATAACTTCATTGCCGGGTCCATCCAATTGA
     G  G  R  T  K  D  R  I  T  S  L  P  G  P  S  N  *
```

Figure 2. The *AmGS4* Open Reading Frame (346 aa) detected by NCBI Open Reading Frame Finder. The codon in black indicates the ATG initiation codon and (*) indicates stop codon.

corresponding protein sequences using the NCBI ORF Finder server. The ORF is a region in a gene sequence that is potentially translatable into proteins. Thus, it is crucial to verify the cloned sequences that comprise this region. Figure 1 shows the longest ORF for the *AmGS3* gene, with codon identities to respective nucleotides detected by the ORF Finder server. According to the result, the ORF of *AmGS3* is detected as 948 bp nucleotides, encoding 315 amino acids. Meanwhile, the ORF for the second isolated sesquiterpene synthase gene, *AmGS4*, yielded 346 amino acids encoded by 1047 bp nucleotides, as shown in Figure 2.

The ORF for *A. malaccensis* chalcone synthase gene, *AmCHS*, as shown in Figure 3, contains 1185 bp of nucleotides, encoding protein sequences comprising 394 amino acids. These results suggest that all cloned *AmGS3*, *AmGS4*, and *AmCHS* sequences are translatable into protein sequences and suitable for further expression analysis. The generation of clones producing only protein-coding regions becomes a clue for the comprehensive research of protein expression and function in different systems. Additionally, gene prediction quality greatly depends on the cDNA sequence information, as gene identification is mostly based on cDNA sequences.

```
   1 ATGACCGAGCTCAAGGAGAAGTTCAAGCGCATGTGTGACAAATCGATGATCAAGAAGAGGTACATGCACGTGACGGAGGAGGTTCTGAAG
     M  T  E  L  K  E  K  F  K  R  M  C  D  K  S  M  I  K  K  R  Y  M  H  V  T  E  E  V  L  K
  91 GAGAACCCCTCCATGGCCGACTACTGGTCGCCGTCGCTGGACGCCCGCCAGGATATCGTGGTGGTGGAGATCCCGAAGTTGGGCAAAGAA
     E  N  P  S  M  A  D  Y  W  S  P  S  L  D  A  R  Q  D  I  V  V  V  E  I  P  K  L  G  K  E
 181 GCTGCCCAGAAGGCCATCAAGGAGTGGGGCCAGCCCAAGTCCAAGATCACACACGTCGTCTTCTGCACCACTTCCGGCGTCGACATGCCC
     A  A  Q  K  A  I  K  E  W  G  Q  P  K  S  K  I  T  H  V  V  F  C  T  T  S  G  V  D  M  P
 271 GGCGCCGACTACCAGCTCACCAAGCTCCTCGGCCTCCGCCCCTCCGTCAAGCGCCTCATGATGTACCAGCAGGGCTGCTTTGCCGGCGGC
     G  A  D  Y  Q  L  T  K  L  L  G  L  R  P  S  V  K  R  L  M  M  Y  Q  Q  G  C  F  A  G  G
 361 ACTGTCCTCCGTCTTGCCAAGGACCTCGCTGAGAACAACAAGGGCGCGCGCGTGCTGGTCGTCTGCTCTGAAATCACCGCCGTCACCTTC
     T  V  L  R  L  A  K  D  L  A  E  N  N  K  G  A  R  V  L  V  V  C  S  E  I  T  A  V  T  F
 451 CGCGGCCCTTCAGAGACCCACCTTGACTCGATGGTTGGCCAGGCACTTTTCGGCGACGGCGCAGCTGCTATAATCGTCGGCTCCGACCCT
     R  G  P  S  E  T  H  L  D  S  M  V  G  Q  A  L  F  G  D  G  A  A  A  I  I  V  G  S  D  P
 541 GACACCAAGATCGAGCGTCCACTCTTCGAGCTGATCTCGGCAGCCCAGACCATCCTCCCCGACTCCGACGGCGCTATTGACGGCCACCTC
     D  T  K  I  E  R  P  L  F  E  L  I  S  A  A  Q  T  I  L  P  D  S  D  G  A  I  D  G  H  L
 631 CGTGAAGTGGGTCTCACCTTCCATCTTCTGAAGGACGTTCCCGGGCTGATCTCGAAGAACATCGAGAAAAGCTTGGTGGAAGCCTTTACC
     R  E  V  G  L  T  F  H  L  L  K  D  V  P  G  L  I  S  K  N  I  E  K  S  L  V  E  A  F  T
 721 CCGATCGGCATCAGCGACTGGAACTCCATCTACTGGATCGCTCACCCGGGTGGTCCTGCCATTCTCGACCAGGTTGAACAGAAACTCGGT
     P  I  G  I  S  D  W  N  S  I  Y  W  I  A  H  P  G  G  P  A  I  L  D  Q  V  E  Q  K  L  G
 811 CTAAAACAGGAGAAACTGAGGGCGACTCGCCACATACTCAGTGAGTACGGGAACATGTCCAGCGCGTGTGTCTTGTTTATCTTGGATGAA
     L  K  Q  E  K  L  R  A  T  R  H  I  L  S  E  Y  G  N  M  S  S  A  C  V  L  F  I  L  D  E
 901 ATGAGGAAGAAGTCGCTGGAGGAAGGGAAGGCCACCACTGGAGAAGGGTTGGAGTGGGGAGTTCTGTTCGGGTTCGGGCCGGGTCTGACG
     M  R  K  K  S  L  E  E  G  K  A  T  T  G  E  G  L  E  W  G  V  L  F  G  F  G  P  G  L  T
 991 GTGGAGACAGTGGTGCTGCACAGTGTAGCAACCGAGTCGGCTCATTGA
     V  E  T  V  V  L  H  S  V  A  T  E  S  A  H  *
```

Figure 3. The *AmCHS* Open Reading Frame (394 aa) detected by NCBI Open Reading Frame Finder. The codon in black indicates the ATG initiation codon and (*) indicates stop codon.

CONCLUSION

In summary, two new sesquiterpene synthases *AmGS3* and *AmGS4*, and a chalcone synthase, *AmCHS*, identified from the stem of *A. malaccensis* were successfully isolated. The full-length ORFs of *AmGS3*, *AmGS4*, and *AmCHS* were successfully cloned into the *E. coli* cloning system. This result sheds some light on the expression of sesquiterpene synthase and chalcone synthase from *A. malaccensis*, which might play important roles in the formation of agarwood.

AVAILABILITY OF DATA AND MATERIALS

All data are available from the corresponding author. Representative sequences were submitted to the GenBank database under the following accession numbers. Sesquiterpene synthase sequences: AmGS3 (QRI93561.1), AmGS4 (QRI93562.1). Chalcone synthase sequence: AmCHS (QRI93563.1).

ACKNOWLEDGMENT

This research was supported by Universiti Malaysia Pahang through the RDU182207-1 research grant.

REFERENCES

1. Chong, S. P., Osman, M. F., Bahari, N., Nuri, E. A., Zakaria, R. & Abdul-Rahim, K. (2015) Agarwood Inducement Technology : A Method for Producing Oil Grade Agarwood in Cultivated *Aquilaria malaccensis* Lamk., *Journal of Agrobiotechnology*, **6**.

2. Elias, M. F., Ibrahim, H. & Mahamod, W. R. W. (2017) A Review on the Malaysian *Aquilaria* species in Karas Plantation and Agarwood Production. *International Journal of Academic Research in Business and Social Sciences*, **7(4)**, 1021–1029.

3. Tan, C. S., Isa, N., Ismail, I. & Zainal, Z. (2019) Agarwood Induction: Current Developments and Future Perspectives. *Frontiers in Plant Science*, **10(122)**.

4. Lee, S. Y. & Mohamed, R. (2016) The origin and domestication of Aquilaria, an important agarwood-producing genus, in *Agarwood: Science Behind the Fragrance*, ed. R. Mohamed. (Berlin: Springer Singapore), 1–20.

5. Naef, R. (2011) The volatile and semi-volatile constituents of agarwood, the infected heartwood of *Aquilaria* species: A review. *Flavour and Fragrance Journal*, **26**, 73–89.

6. Gao, Z., Yang, Y., Zhang, Z., Zhao, W., Meng, H., Jin, Y., Wei, J. (2014) Profiling of MicroRNAs under Wound Treatment in *Aquilaria* sinensis to Identify Possible MicroRNAs Involved in Agarwood Formation. *International Journal of Biological Sciences*, **10(5)**, 500–510.

7. Xu, Y., Zhang, Z., Wang, M., Wei, J., Chen, H., Gao, Z., Li., W. (2013) Identification of genes related to agarwood formation: transcriptome analysis of healthy and wounded tissues of *Aquilaria* sinensis. *BMC Genomics*, **14(227)**.

8. Chhipa, H. & Kaushik, N. (2017) Fungal and Bacterial Diversity Isolated from *Aquilaria* malaccensis Tree and Soil, Induces Agarospirol Formation within 3 Months after Artificial Infection. *Frontiers in Microbiology*, **8(1286)**.

9. Jayachandran, K., Sekar, I., Parthiban, K. T., Amirtham, D. & Suresh, K. K. (2014) Analysis of different grades of Agarwood (*Aquilaria* malaccensis Lamk.) oil through GC-MS. *Indian Journal of Natrual Products and Resources*, **5(1)**, 44–47.

10. Liu, Y., Chen, H., Yang, Y., Zhang, Z., Wei, J., Meng, H., Chen, H. (2013) Whole-tree Agarwood-Inducing Technique: An Efficient Novel Technique for Producing High-Quality Agarwood in Cultivated *Aquilaria* sinensis Trees. *Molecules*, **18**, 3086–3106.

11. Chen, X., Zhu, X., Feng, M., Zhong, Z., Zhou, X., Chen, X., Gao, X. (2017) Relationship between expression of chalcone synthase genes and chromones in artificial agarwood induced by formic acid stimulation. *Molecules.* **22(686)**, 1–14.

12. Siah, C. H., Namasivayam, P. & Mohamed, R. (2016) Transcriptome reveals senescing callus tissue of Aquilaria malaccensis, an endangered tropical tree, triggers similar response as wounding with respect to terpenoid biosynthesis. *Tree Genetics and Genomes*, **12(2)**.

13. Joshi, T. & Xu, D. (2007) Quantitative assessment of relationship between sequence similarity and function similarity. *BMC Genomics*, **8**, 1–10.

14. Ye, W., He, X., Wu, H., Wang, L., Zhang, W., Fan, Y., Gao, X. (2018) Identification and characterization of a novel sesquiterpene synthase from Aquilaria sinensis: An important gene for agarwood formation. *International Journal of Biological Macromolecules*, **108**, 884–892.

15. Wang, X., Gao, B., Liu, X., Dong, X., Zhang, Z., Fan, H., Wang, J. (2016) Salinity stress induces the production of 2- (2-phenylethyl) chromones and regulates novel classes of responsive genes involved in signal transduction in *Aquilaria* sinensis calli. *BMC Plant Biology*, **16(119)**, 1–20.