Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Gene 851 (2023) 146981

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Whole transcriptome analysis identifies full-length genes for neoandrographolide biosynthesis from *Andrographis alata*, an alternate source for antiviral compounds

Tanuja, Madasamy Parani *

Department of Genetic Engineering, College of Engineering and Technology, Faculty of Engineering and Technology, SRM Institute of Science and Technology, SRM Nagar, Kattankulathur 603203, Kanchipuram, Chennai, TN, India

ARTICLE INFO

Edited by Takashi Gojobori

Keywords:
Andrographolide
Andrographis
Andrographis alata
Neoandrographolide
Secondary metabolite
Transcriptomics

ABSTRACT

Andrographolide and related compounds are effective against several viral diseases, including dengue, COVID-19, influenza, and chikungunya. Andrographis paniculata is the primary source for these compounds, but its availability is limited. A. alata is a potential alternative source, and neoandrographolide (NAG) is the major antiviral compound in this species. Since molecular studies in A. alata are scarce, we sequenced its leaf transcriptome to identify the full-length genes involved in neoandrographolide biosynthesis. We assembled 13.6 Gb RNA-Seq data and generated 81,361 transcripts with 1007 bp average length and 1,810 bp N50. The transcripts were categorized under biological processes (2,707), cellular components (678), and molecular functions (2,036). KEGG analysis mapped 975 transcripts to the secondary metabolite pathways. Among the 420 transcripts mapped to terpenoids and polyketides pathways, 142 transcripts were related to the biosynthesis of andrographolide and its derivatives. After a detailed analysis of these transcripts, we identified 32 full-length genes coding for all the 22 enzymes needed for andrographolide biosynthesis. Among them, 15 full-length genes were identified for the first time from Andrographis species. These full-length genes and the transcripts shall serve as an invaluable resource for the metabolic engineering of andrographolides and neoandrographolide in Andrographis and other species.

1. Introduction

Plant-derived andrographolides have potent antiviral activity against diverse viruses. The therapeutic value of these compounds was proved in several mechanistic studies reported over the past two decades. Experiments in Chikungunya virus, Influenza A virus, Zika virus, Hepatitis B virus, Dengue virus, Human immunodeficiency virus, Epstein-Barr virus, Hepatitis C virus, and Herpes simplex virus type-1 showed that these compounds could reduce the expression of viral RNA and proteins (Wintachai et al., 2015), inhibit viral DNA replication (Chen et al., 2014; Lee et al., 2014; Yuan et al., 2016), restrict the viral particle multiplication (Li et al., 2020), reduce virus-induced cell mortality (Yu et al., 2014), increase CD4 + T cell count (Calabrese et al., 2000), and inhibit the expression of viral lytic proteins (Lin et al., 2008). *In silico* studies showed that andrographolides and neoandrographolide are promising candidates for treating SARS-CoV-2 (Murugan et al., 2021; Sa-Ngiamsuntorn et al., 2021). Therefore, extracts from the plants producing these compounds are used as a general-purpose therapy

Abbreviations: UGT, UDP-glucotransferase; GO, Gene Ontology; EC, Enzyme Commission; RIN, RNA integrity number; DXS, 1-deoxy-D-xylulose-5-phosphate synthase; GGPPS, Geranylgeranyl pyrophosphate synthase; FPPS, Farnesyl-pyrophosphate synthase; HMGS, Hydroxymethylglutaryl-CoA synthase; HMGR, Hydroxymethylglutaryl-CoA reductase; HDS, 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate synthase; FTase, Farnesyltransferase; ICMT, Protein-S-isoprenylcysteine O-methyltransferase; KS, *Ent*-kaurene synthase; GGPP, Geranylgeranyl diphosphate; IPPI, Isopentenyl diphosphate isomerase; CDS, Coding sequence; MVA, Mevalonic acid; MEP, Methyl 2-C-methyl-D-erythritol-4-phosphate; DMAPP, Dimethylallyl pyrophosphate; AACT, Acetyl-CoA C-acetyl transferase; MVK, Mevalonate kinase; PMK, 5-phosphomevalonate kinase; MDD, Mevalonate diphosphate decarboxylase; DXR, 1-deoxy-D-xylulose 5-phosphate reductoisomerase; MCT, 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase; CMK, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; MCS, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; HDR, 4-hydroxy-3-methylbut-2-enyl diphosphate reductase; CPS, *ent*-copalyl diphosphatase synthetase; CPR, NADPH:cytochrome P450 reductase; RT-PCR, Reverse transcriptase-polymerase chain reaction; qRT-PCR, Quantitative Reverse transcriptase-polymerase chain reaction.

* Corresponding author.
E-mail address: paranim@srmist.edu.in (M. Parani).

https://doi.org/10.1016/j.gene.2022.146981
Received 26 April 2022; Received in revised form 28 August 2022; Accepted 13 October 2022
Available online 19 October 2022

against viral infections and fever. During viral epidemics and pandemics like the recent one due to SARS-CoV-2, the demand for these plants increases dramatically. Meeting such a demand is challenging because these plants are obtained mainly from wild collections.

The *Andrographis* genus contains 29 species, including *A. paniculata, A. alata, A. gracilis, A. affinis, A. longipedunculata*, *A. elongata*, and *A. macrobotrys* (Gnanasekaran and Murthy, 2015). Genes coding for NADPH: cytochrome P450 reductase, *ent*-copalyl diphosphate synthase, geranylgeranyl pyrophosphate synthases (GGPPS), and UDP-glucotransferase (UGT) were cloned from *A. paniculata* and functionally characterized (Lin et al., 2017; Misra et al., 2015; Wang et al., 2019; Srivastava et al., 2021). Chromosome-level genome sequence and transcriptome data are also available for *A. paniculata* (Liang et al., 2020). In other research areas also, *A. paniculata* is well-studied among the Andrographis species. The greater scientific interest in *A. paniculata* stems from the fact that this is the species predominantly used for therapeutic purposes. However, considering the frequent viral epidemics and increasing acceptance of plant-based therapy, exploring other potential sources for these compounds would be desirable. Ethanol extract from *A. alata* was reported to have antipyretic and anti-inflammatory activities in male albino rats (Balu et al., 1993; Balu and Alagesaboopathi, 1993). *A. alata* leaves are an excellent source of neoandrographolide with 33.21 mg/g dry weight (Dalawai et al., 2021). *In vitro* regenerated *A. alata* plants contained 32.21 mg/g dry weight neoandrographolide (Kadapatti and Murthy, 2021). However, research on *A. alata* is minimal, with just eight publications since 1993. Cloning the genes coding for the relevant enzymes will help increase the yield of andrographolides and related compounds in the resident species and transfer the biosynthetic pathway to new species. As of now, the GenBank database of NCBI contains only 413 protein-coding sequences from *Andrographis* species. Several genes in the terpenoid pathway in general and andrographolide biosynthesis, in particular, remain to be characterized. RNA-Seq is a promising method for large-scale gene identification in a cost-effective and high throughput manner (Kumar et al., 2012). In the current study, we have assembled the leaf transcriptome of *A. alata* and identified several genes involved in the biosynthesis of neoandrographolide.

2. Methods

2.1. Plant collection and RNA isolation

Seedlings of *A. alata* were collected from the surrounding lakes near Potheri, Chennai, Tamil Nadu. The seedlings were transplanted into pots containing garden soil, and grown in a greenhouse at 23–25 °C temperature, 16 h light, and 50 % relative humidity. A botanist taxonomically identified the plants. Mature leaves from a 2-month-old plant were harvested, frozen in liquid nitrogen, and stored at −80 °C. Total RNA was extracted using RNAiso plus reagent (TaKaRa, Japan) according to the manufacturer's protocol. DNase I (Qiagen, GmbH, Germany) was used to remove DNA contamination and the RNeasy MinElute Clean-up Kit was used to purify DNase treated RNA (Qiagen, GmbH, Germany). Intact rRNA bands without degradation or DNA contamination were observed in the purified total RNA samples. RNA concentration and quality were estimated using the Qubit RNA BR assay kit (Invitrogen, California, USA) and NanoDrop spectrophotometer. The RNA integrity was checked using Bioanalyzer 2100 (Agilent Technologies, California, USA). Total RNA (550 ng/μl) with 260/280 ratio of 1.9 and RNA integrity number (RIN) of 8.5 was used for cDNA library construction.

2.2. cDNA library construction and sequencing

The cDNA library was prepared from 1 μg of total RNA according to the Illumina TrueSeq mRNA library preparation kit protocol (https://protocols.hostmicrobe.org/legacy-protocols/truseq-mrna-library-prep). In brief, mRNA was purified from the total RNA using poly-A enriched Oligo-dT magnetic beads. The purified mRNA was fragmented, and first-strand cDNA synthesis was performed by reverse-transcription with Superscript II reverse transcriptase and random hexamers. Second strand cDNA synthesis was done in the presence of DNA polymerase 1, and cDNA size selection was performed using magnetic beads. The 3′ ends of double-stranded cDNAs were end-repaired and adenylated using an end repair and A-Tailing mix. The adapters were ligated to the fragmented cDNAs and enriched for adapter-ligated fragments by PCR amplification. Finally, validation of the library was done using bioanalyzer 2100 (Agilent Technologies, USA). The RNA-Seq library was quantified using the Qubit dsDNA BR assay kit for Qubit3.0 fluorimeter and sequenced in the Illumina NextSeq500 platform (Illumina, USA). The sequence data were submitted to NCBI Sequence Read Archive (SRA) database (Accession No. SRR18210043 and Bio project ID PRJNA811695).

2.3. Processing of raw data and de novo assembly

The following steps were performed with the raw paired-end sequence data generated from the Illumina NextSeq500 platform. First, raw read quality checks were performed using Fastqc (Andrews, 2010). Adapters sequences and low-quality reads (<30 phred score) were removed using Trimmomatic (Bolger et al., 2014). Processed reads with bases quality score > 30 were used for transcriptome assembly using Trinity assembler (Grabherr et al., 2011). The transcripts were clustered using Cd-hit-est (Li and Godzik, 2006) to remove redundant transcripts and produce unique transcripts.

2.4. Functional annotation and pathway enrichments

Functional annotation of transcripts was conducted by sequence similarity comparison to non-redundant databases of NCBI with an E value threshold $\leq 1e^{-5}$. The resultant annotated transcripts were used for Gene Ontology (GO) annotation and Kyoto Encyclopedia of Gene and Genomes analysis. Blast2GO software (Conesa et al., 2005) was used for KEGG pathway analysis, and the retrieved GO terms were classified into biological processes, cellular components, and molecular functions. The transcripts were assigned with unique Enzyme Commission (EC) numbers and used to map KEGG pathways.

2.5. Reverse transcriptase-polymerase chain reaction (RT-PCR)

Seven genes coding for the enzymes in the neoandrographolide biosynthesis were selected for RT-PCR analysis. First-strand cDNA was synthesized using 1 μg RNA, 0.5 μM oligo-dT primer, and 200 units Superscripts II reverse transcriptase (Invitrogen, USA). Gene-specific primers for the selected genes were designed using Primer 3 plus software (https://www.primer3plus.com/) to amplify 906 to 2386 bp cDNA fragments (Supplementary Table 1). PCR reactions contained Taq 2X Master Mix RED (Ampliqon, Denmark) to 1X concentration, 0.5 μM primers, and 15 ng cDNA. PCR amplification steps comprised of initial denaturation at 95 °C for 5 min, followed by 35 cycles of denaturation at 95 °C for 30 *sec*, annealing at 52–59 °C for 30 *sec*, and extension at 72 °C for 1 min, final extension at 72 °C for 5 min, and hold at 4 °C.

2.6. Quantitative-RT-PCR (qRT-PCR)

Seven genes coding for the enzymes in the neoandrographolide biosynthesis were selected for qRT-PCR analysis. Elongation factor 1-alpha was selected as the housekeeping gene. Gene-specific primers targeting 115 to 213 bp regions in the 3′UTR of the transcripts (Supplementary Table 2) were designed using Primer 3 plus software (https://www.primer3plus.com/). First-strand cDNA was synthesized using 1 μg RNA, 0.5 μM oligo-dT primer, and 200 units Superscripts II reverse transcriptase (Invitrogen, USA). The qPCR reactions (10 μl) in triplicate containing 5 μl TB Green Premix Ex Taq II 2X mix (Takara,

Japan), 0.5 μM primers, and 15 ng of cDNA were performed in QuantStudio 5 Real-time PCR machine (Thermo Scientific, USA). The qRT-PCR amplification cycle comprised of an initial denaturation at 95 °C for 2 min, 40 cycles of amplification at 95 °C for 30 *sec*, and 57 °C for 30 *sec*, a melting curve analysis at 95 °C for 15 *sec*, 57 °C for 1 min, and 95 °C for 15 *sec*.

3. Results

3.1. Transcriptome sequencing and de novo assembly

We generated 97 million paired-end (2 x150 bp) reads from the leaves of *A. alata*. After removing low-quality reads (<Q30), 81 million high-quality reads were obtained for *de novo* transcriptome assembly. The assembled transcriptome contained 81,361 transcripts with 1,007 bp average length and 1,810 bp N50 value. The assembly statistics of the transcriptome are provided in Table 1. The transcript size varied from 200 bp to 16,682 bp. About 47 % of transcripts (38,718) were smaller than 500 bp. Transcripts of size 501–1,000 bp, 1,001–2,000 bp, and 2,001–5,000 bp were 18 % (14,680), 20 % (16,013), and 14 % (11,330), respectively. The overall size distribution of the transcripts is shown in Fig. 1.

3.2. Functional annotation

The *de novo* assembled transcripts were functionally annotated using NR, GO, KEGG, and UniProt databases. Among the assembled transcripts, 57 % (46,764) showed significant homology with other proteins, and 43 % (34,597) did not show significant similarity with any sequence in the database. About 85 % (39,728) of the annotated transcripts matched existing gene models, and 15 % (7,036) were classified under uncharacterized protein. Regarding the top hit species, the top hits were found with *Sesamum indicum* (37.2 %), followed by *Erythranthe guttata* (8 %) and *Nicotiana tabacum* (0.8 %). The top hit species distribution is shown in Fig. 2. According to the GO analysis, the transcripts were classified into three main domains: biological process, cellular components, and molecular function. GO analysis assigned 38,069 transcripts to one or more GO categories; 2,707, 679, and 2,037 categories under biological processes, cellular components, and molecular functions. The top 20 categories from each domain are represented in Fig. 3. Most of the transcripts related to the biological process belonged to the oxidation–reduction process (2,418), followed by protein phosphorylation (1,424) and transmembrane transport (1,077). Under cellular components, integral components of the membrane (7,566), cytoplasm (2,652), and membrane (2,317) were the predominant categories. Under molecular functions, ATP binding (3,826), metal ion binding (2,051), and DNA binding (1,038) were amongst the top three major classes.

3.3. Analysis of metabolic pathways

The assembled transcripts were mapped to the KEGG database to determine their involvement in various metabolic pathways. We mapped 19,169 transcripts to 150 KEGG pathways. Most transcripts were mapped to the purine metabolism pathway (1,986). The other pathways with a significantly higher number of transcripts were thiamine metabolism (1,800), starch and sucrose metabolism (714), drug metabolism (426), and glycolysis (419). Transcripts mapped to the top ten pathways are represented in Fig. 4. The KEGG pathways are classified under five groups. The highest number of transcripts were mapped to metabolism (16,053), further classified into 11 subgroups. More than 50 % of the transcripts under metabolism were mapped to the metabolism of carbohydrates (3,752), cofactors and vitamins (2,866), and nucleotides (2,173). We mapped 975 transcripts to the metabolism of terpenoids, polyketides, and other secondary metabolites. Details of the transcripts mapped to the 11 subgroups under the metabolism category are given in Fig. 5.

Table 1
Statistics of *A. alata* transcriptome sequencing and assembly.

Particulars	Numbers
Number of raw reads (Millions)	97
Number of clean reads (Millions)	81
Number of bases > Q30 (Gb)	8.9
Mapped reads (%)	95
Number of assembled transcripts	81,361
Average length (bases)	1,007
Median contig length (bases)	542
Max length (bases)	16,682
Min length (bases)	200
Contig N50 (bases)	1,810
GC content (%)	47

3.4. Terpenoid backbone biosynthesis pathway

Andrographolides and related compounds are part of the terpenoids and polyketides pathway. We mapped 420 transcripts to this pathway. Terpenoid backbone biosynthesis is the major pathway under the metabolism of terpenoids and polyketides (Fig. 6). We found 158 transcripts encoding 26 enzymes involved in synthesizing terpenoids and polyketides (Fig. 7). The highest number of transcripts (23) encoded for 1-deoxy-D-xylulose-5-phosphate synthase (DXS), followed by geranylgeranyl pyrophosphate synthase (GGPPS) and farnesyl-pyrophosphate synthase (FPPS) with 16 transcripts. Hydroxymethylglutaryl-CoA synthase (HMGS), hydroxymethylglutaryl-CoA reductase (HMGR), 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate synthase (HDS), farnesyltransferase (FTase), and Protein-S-isoprenylcysteine O-methyltransferase (ICMT) enzymes had 5 to 10 transcripts. Enzymes in the terpenoid backbone pathway and the number of transcripts coding for these enzymes are given in Table 2. We also identified 20 transcripts involved in the diterpenoid pathway, and five of them encoded Abieta-7,13-dien-18-ol hydroxylase. *Ent*-kaurene synthase (KS) and Gibberellin 3-beta-dioxygenase were encoded by three and two transcripts, respectively. The number of transcripts coding for the different enzymes in the diterpenoid backbone pathway is given in Table 3.

3.5. Full-length genes involved in neoandrographolide biosynthesis

At least 22 enzymes are required to synthesize geranylgeranyl diphosphate (GGPP) and convert it to neoandrographolide. From the leaf transcriptome of *A. alata*, we identified 130 transcripts coding for these 22 enzymes. After a detailed analysis of these transcripts, we identified full-length genes for all the enzymes required for the biosynthesis of neoandrographolide. Three full-length genes each were identified for 3-Hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR), geranylgeranyl pyrophosphate synthase (GGPPS), and 1-deoxy-D-xylulose 5-phosphate synthase (DXS) enzymes. Two genes each were identified for acetyl-CoA C-acetyl transferase (AACT), 4-Hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR), isopentenyl diphosphate isomerase (IPPI), and farnesyl pyrophosphate synthase (FPPS). Single genes encoded all other enzymes. The coding sequence (CDS) of the full-length genes ranged between 720 and 2,451 bp and encoded 240 to 817 amino acids. Sequences of all the full-length genes were submitted to the GenBank under the accession numbers ON049353 to ON049382 and ON211623 to ON211625. Details of the full-length genes involved in neoandrographolide biosynthesis identified from *A. alata* are given in Table 4.

3.6. Reverse-transcriptase validation

The genes coding for the MCS, IPPI2, MCT, MVK, FPPS2, HDR2, HMGR1A, and DXS2A enzymes in the neoandrographolide biosynthesis

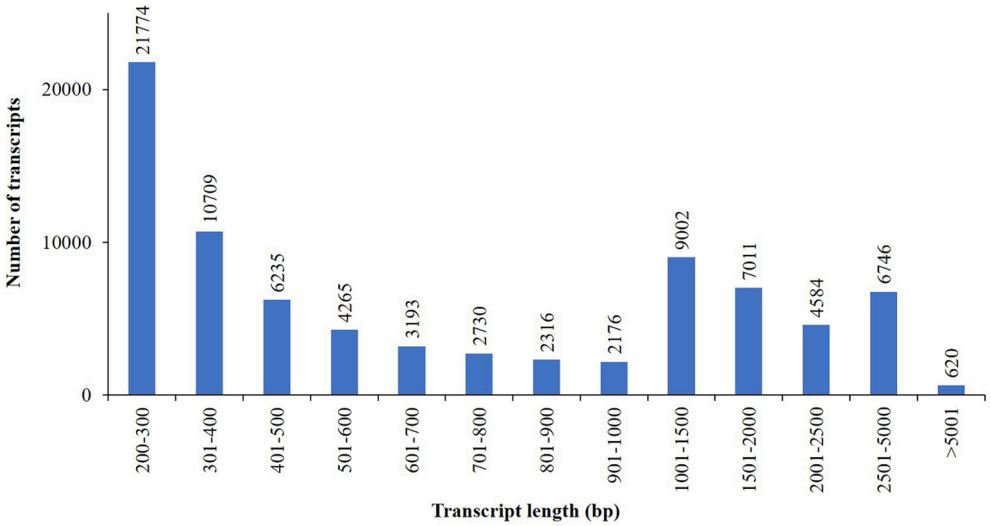

Fig. 1. Size distribution of the transcripts assembled from *A. alata*.

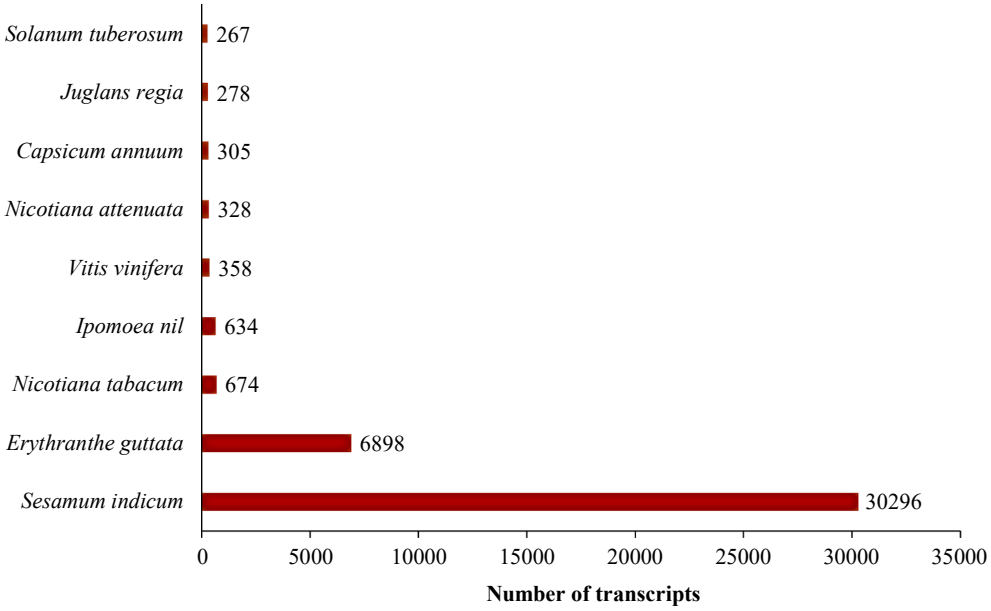

Fig. 2. Top hit species based on the similarity search to the non-redundant database of NCBI.

pathway were selected for validating the transcript assembly and gene expression. RT-PCR amplification of full-length genes yielded the cDNA fragments of the expected size ranging from 906 bp to 2368 bp (Fig. 8) indicating correct *de novo* transcript assembly. In qRT-PCR, all the genes selected from the neoandrographolide biosynthesis pathway showed a high level of expression as indicated by the Cq value ranging between 20.9 and 25.5 (Fig. 9).

4. Discussion

The antiviral property of Andrographis species is mainly due to andrographolide and related diterpenoid compounds (Wintachai et al., 2015; Chen et al., 2009). Although *Andrographis paniculata* is the primary source for these compounds, increasing demand for this species necessitates searching for alternate sources. Another species of *Andrographis*, *A. alata*, is rich in neoandrographolide and shows antipyretic, anti-inflammatory, and antiviral properties (Balu et al., 1993; Balu and Alagesaboopathi, 1993; Dalawai et al., 2021). However, compared to *A. paniculata*, research on *A. alata* is extremely limited. The present study focused on sequencing the whole transcriptome from *A. alata* leaves to generate a large-scale genomics resource and to identify the full-length gene involved in neoandrographolide biosynthesis. *De novo* transcriptome assembly using next-generation sequencing is a fast and economical tool for generating expressed sequence data for non-model plants, particularly for those species which do not have genome sequences yet (Tanuja et al., 2021; Bhambhani et al., 2017). This is the first time a transcriptome is assembled for *A. alata*. It had a higher number of transcripts and a greater N50 value than the leaf and root transcriptomes of *A. paniculata* reported earlier (Garg et al., 2015; Patel et al., 2020). The assembled transcriptome showed 95 % alignment with the quality-filtered reads. Based on the top-hit species distribution, most of the *A. alata* transcripts matched with *Sesamum indicum*. This may be because the GenBank database contains a significantly higher number of gene sequences from *Sesamum* (66,404 entries) than *Andrographis* (447 entries), and both species belong to the same order, Lamiales.

Secondary metabolites from plants are significantly used in the pharmaceutical industries and traditional medicine. Understanding the biosynthetic pathways, regulation, and the genes involved will help

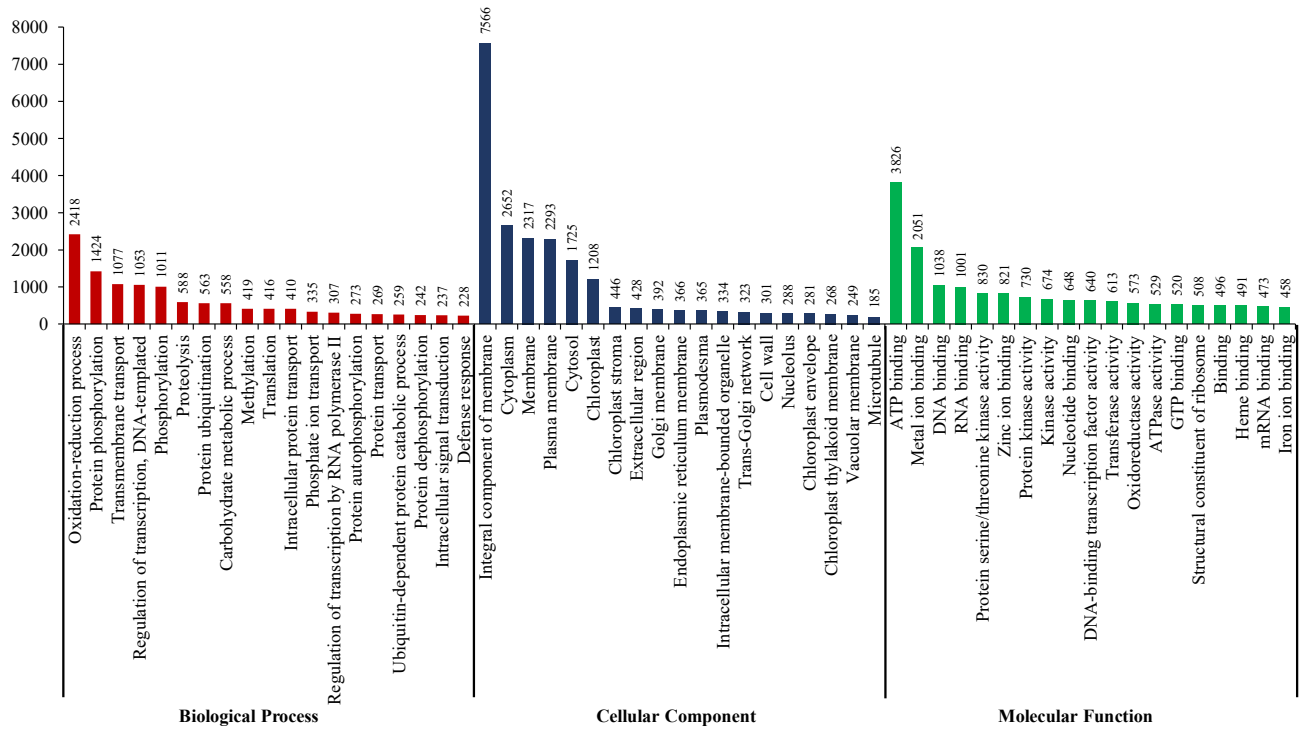

Fig. 3. Histogram showing GO term assigned to *A. alata* transcripts under biological process, cellular component, and molecular function.

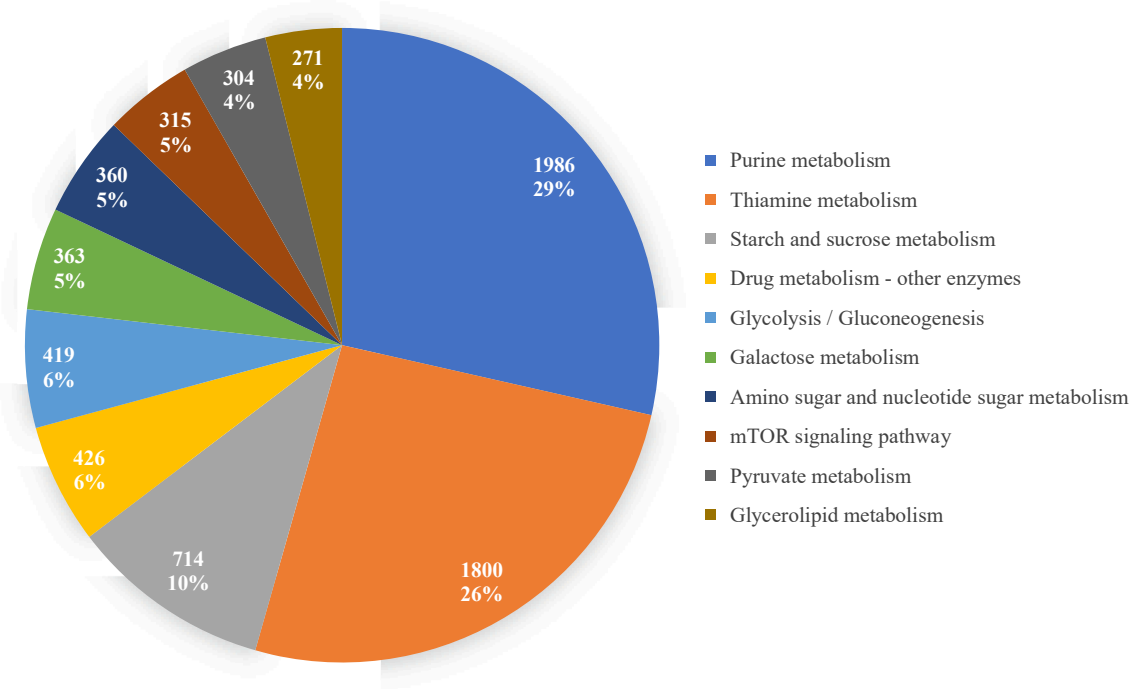

Fig. 4. The distribution of *A. alata* transcripts among the top ten pathways in KEGG pathway analysis.

improve yield and metabolic engineering. Terpenoids constitute the largest class of structurally diverse secondary metabolites and are categorized based on the number of 5-carbon isoprene units: monoterpenes (C10), sesquiterpenes (C15), diterpenes (C20), triterpenes (C30), tetraterpenes (C40) and polyterpenes (>C40). Neoandrographolide (NAG), a diterpenoid compound, is the major secondary metabolite responsible for the antiviral property of *A. alata* (Murugan et al., 2021). Geranylgeranyl pyrophosphate (GGPP) is the precursor from which several important diterpenoids, including NAG, are synthesized. GGPP is synthesized from the mevalonic acid (MVA) pathway and methyl 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway in the cytosol and plastids, respectively (Srivastava and Akhila, 2010). These two pathways are divergent till the synthesis of dimethylallyl pyrophosphate (DMAPP), which is converted to GGPP through a common pathway. In the MVA pathway, the synthesis of DMAPP from acetyl-CoA involves seven enzymes; acetyl-CoA C-acetyl transferase

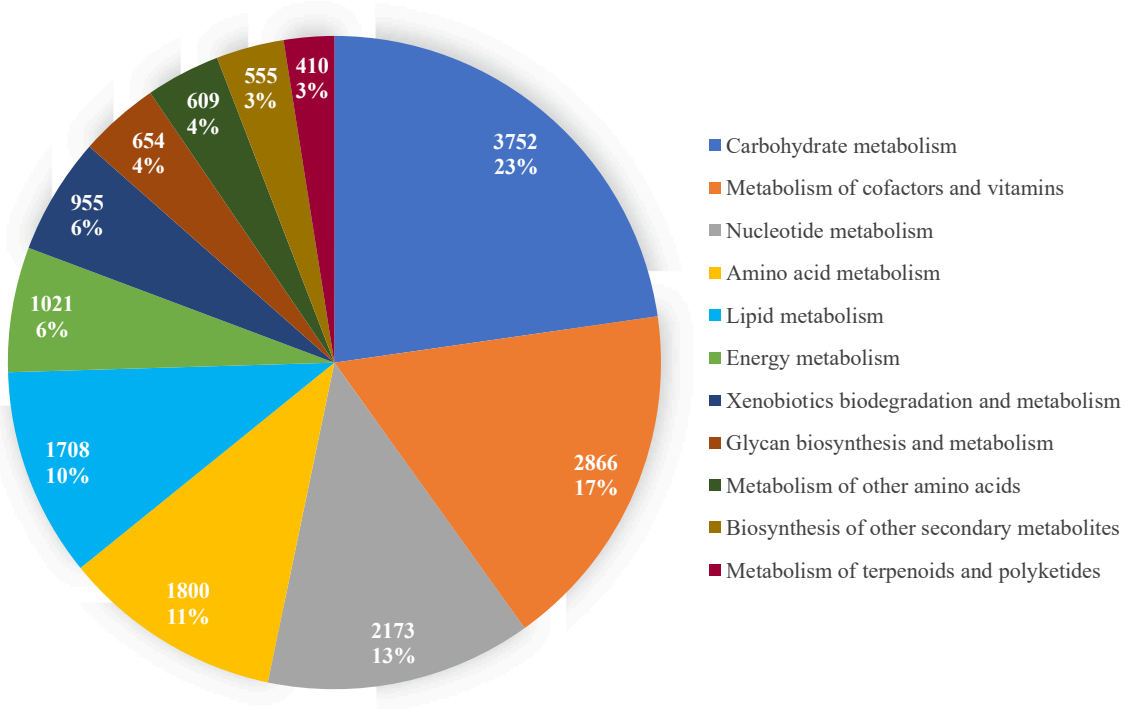

Fig. 5. The distribution of *A. alata* transcripts among the 11 metabolic pathways in KEGG pathway analysis.

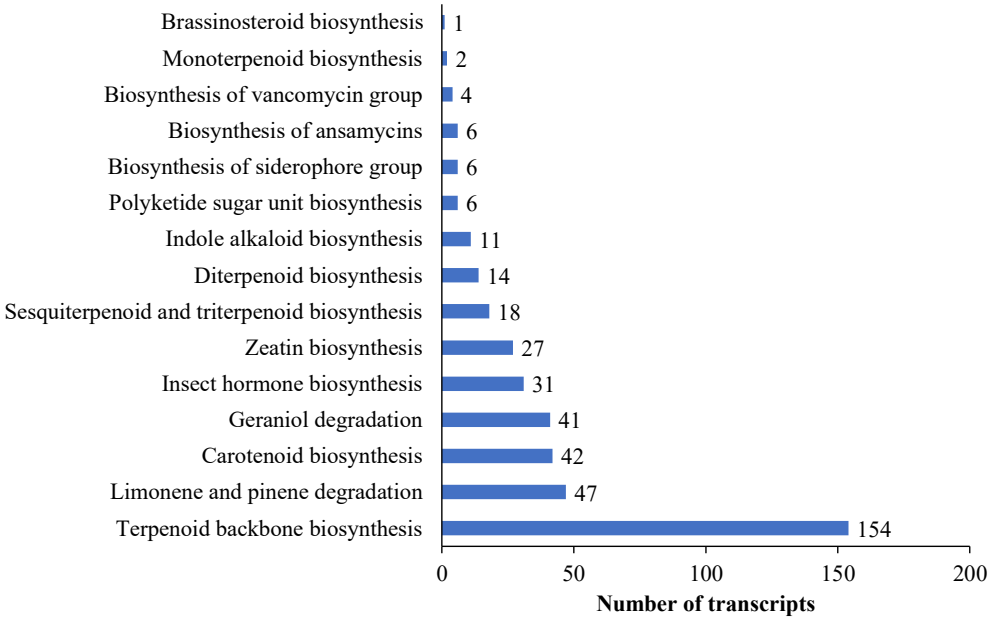

Fig. 6. The distribution of *A. alata* transcripts among the metabolic pathways of different terpenoids and polyketides in KEGG pathway analysis.

(AACT), 3-hydroxy-3-methylglutaryl coenzyme A synthase (HMGS), 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR), mevalonate kinase (MVK), 5-phosphomevalonate kinase (PMK), mevalonate diphosphate decarboxylase (MDD), and isopentenyl diphosphate isomerase (IPPI). We identified 12 full-length genes coding for these seven enzymes. HMGR is a rate-limiting enzyme in the MVA pathway, and its expression levels in the leaves of *A. paniculata* correlated with andrographolide content. ApHMGR gene coding for 595 amino acids was cloned from *A. paniculata* and expressed in *Escherichia coli* (Srinath et al., 2019). We identified the same gene from *A. alata*, annotated as AaHMGR1, coding for 595 amino acids with 99 % identity to ApHMGR. In addition, we identified two other full-length HMGR genes, AaHMGR1a (559 aa) and AaHMGR2 (556 aa). HMGR1 had 68 and 70 % identity with HMGR1a and HMGR2. Three splice variants of the HMGR2 gene were identified. Two genes each for AACT and IPPI and one gene each for HMGS, MVK, PMK, and MDD were also identified for the first time from *A. alata*.

In the MEP pathway, the synthesis of DMAPP from glyceraldehyde 3-phosphate involves seven enzymes, which are 1-deoxy-D-xylulose 5-phosphate synthase (DXS), 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR), 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (MCT), 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase

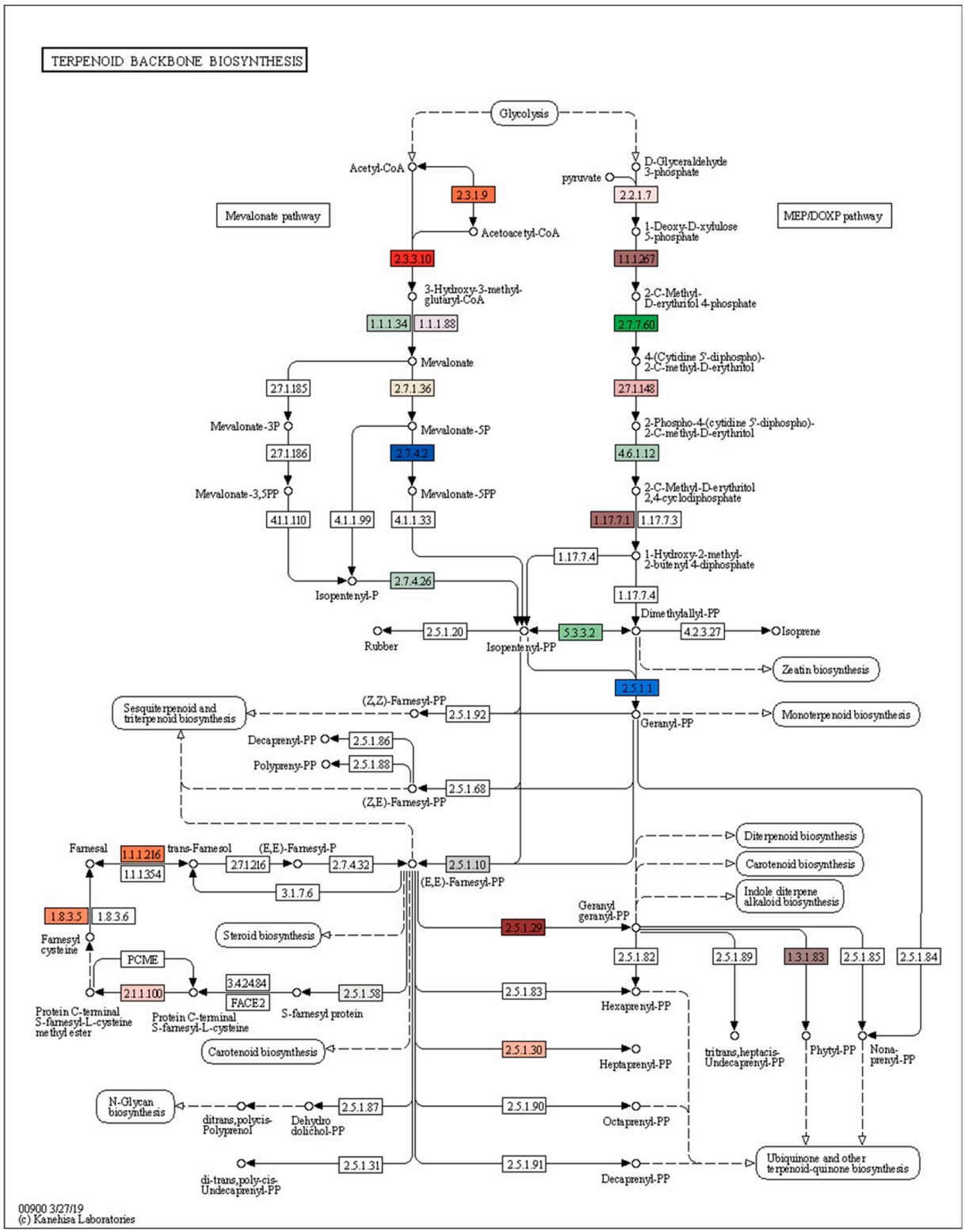

Fig. 7. KEGG analysis of the *A. alata* transcriptome showing the transcripts coding for the enzymes of the terpenoid biosynthesis pathway. Enzyme Commission number of the enzymes are given inside the boxes.

(CMK), 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MCS), 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate synthase (HDS), and 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR). In this study, we identified 11 genes coding for these seven enzymes. DXS controls flux to the MVP pathway, and it is a rate-limiting enzyme. Expression levels of the DXS gene correlated with accumulation of andrographolide. DXS gene was cloned from *A. paniculata* (ApDXS) and expressed in *E. coli* (Srinath et al., 2021). We identified a full-length AaDXS1 gene that showed 99 % identity with ApDXS. However, AaDXS1 (GenBank Acc. No. ON049379) codes for 719 amino acids, whereas ApDXS codes for 691 amino acids. Further analysis showed that the ApDXS gene (GenBank Acc. No. MG271749.1) has a single base pair

Table 2
Enzyme involved in terpenoid biosynthesis and the number of transcripts identified from *A. alata* for each enzyme.

S. No.	Enzyme code	Name of the enzyme	No. of transcripts
1	ec:2.2.1.7	1-deoxy-D-xylulose-5-phosphate synthase (DXS)	23
2	ec:2.5.1.29	Geranylgeranyl diphosphate synthase (GGPPS)	19
3	ec:2.5.1.10	Farnesyl-diphosphate synthase (FPPS)	19
4	ec:2.5.1.1	Dimethylallyltranstransferase (DMATT)	16
5	ec:2.3.3.10	Hydroxymethylglutaryl-CoA synthase (HMGS)	10
6	ec:1.17.7.1	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (Ferredoxin) (HDS)	7
7	ec:1.1.1.34	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (HDS)	8
8	ec:1.1.1.88	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (HDS)	8
9	ec:2.5.1.58	Protein farnesyltransferase (PFT)	6
10	ec:2.1.1.100	Protein-S-isoprenylcysteine O-methyltransferase.	6
11	ec:2.7.4.26	Isopentenyl phosphate kinase (IPK)	5
12	ec:2.3.1.9	Acetyl-CoA C-acetyltransferase (AACT)	4
13	ec:1.14.14.145	Abietadienol hydroxylase	4
14	ec:5.3.3.2	Isopentenyl-diphosphate Delta-isomerase (IPPI)	3
15	ec:1.1.1.267	1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR)	3
16	ec:1.3.1.83	Geranylgeranyl diphosphate reductase. (GPPS)	3
17	ec:2.7.4.2	Phosphomevalonate kinase (PMK)	2
18	ec:4.1.1.33	Diphosphomevalonate decarboxylase (DMD)	2
19	ec:2.5.1.30	Heptaprenyl diphosphate synthase (HepPPS)	2
20	ec:1.17.7.4	4-hydroxy-3-methylbut-2-enyl-diphosphate reductase (HDR)	2
21	ec:1.8.3.5	Prenylcysteine oxidase.	1
22	ec:2.7.1.36	Mevalonate kinase (MVK)	1
23	ec:2.7.1.148	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (CMK)	1
24	ec:2.7.7.60	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (MCT).	1
25	ec:4.6.1.12	2-C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase (MCS)	1
26	ec:1.1.1.216	Farnesol dehydrogenase (FDL)	1
		TOTAL	158

Table 3
-. Enzyme involved in diterpenoid biosynthesis and the number of transcripts identified from *A. alata* for each enzyme.

S. No.	Enzyme code	Name of the enzyme	No. of transcripts
1	ec:1.14.11.15	Gibberellin 3-beta-dioxygenase (GA3OX)	3
2	ec:1.14.11.13	Gibberellin 2-beta-dioxygenase (GA2OX)	2
3	ec:1.14.14.145	Abieta-7,13-dien-18-ol hydroxylase	5
4	ec:4.2.3.19	*Ent*-kaurene synthase (KS)	4
5	ec:5.5.1.13	*Ent*-copalyl diphosphate synthase (CPS)	1
6	ec:1.6.2.4	NADPH: Cytochrome p450 reductase (CPR)	3
7		UDP-glycosyltransferase 73AUI (UGT73AUI)	1
8		UDP-glycosyltransferase 86CII (UGT86CII)	1
		Total	20

deletion at 410th position, which changed the longest open reading frame to start from 323rd position (coding for 691 amino acids) instead of 240th position (coding for 719 amino acids). Additionally, we identified two more full-length DXS genes; one is cytosolic (AaDXS2,

Table 4
Details of the genes identified from A. alata involved neoandrographolide biosynthesis.

S. No.	Name of the gene	mRNA size (bp)	ORF size (bp)	ORF size (aa)	GenBank Accession number
1	Acetyl- CoA C- acetyl transferase 2 (AACT 2)	1728	1248	415	ON049353
2	Acetyl- CoA C- acetyl transferase 1 (AACT 1)	1654	1209	402	ON049354
3	3-hydroxy-3-methylglutaryl coenzyme A synthase (HMGS)	1874	1383	460	ON049355
4	3-hydroxy-3-methylglutaryl-coenzyme A reductase 2 (HMGR2)	3757	1671	556	ON049356
5	3-hydroxy-3-methylglutaryl-coenzyme A reductase 1 (HMGR1)	2107	1788	595	ON049357
6	3-hydroxy-3-methylglutaryl-coenzyme A reductase 1A (HMGR1a)	2420	1680	559	ON049358
7	Mevalonate kinase (MVK)	1629	1161	386	ON049359
8	5-phosphomevalonate kinase (PMK)	2056	1521	506	ON049360
9	Mevalonate diphosphate decarboxylase (MDD)	1585	1260	419	ON049361
10	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (CMK)	1583	1197	398	ON049362
11	1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR)	1953	1437	478	ON049363
12	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (MCT)	1307	921	306	ON049364
13	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MCS)	1025	720	239	ON049365
14	1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate synthase (HDS)	2808	2238	745	ON049366
15	4-hydroxy-3-methylbut-2-enyl diphosphate reductase 1 (HDR1)	1729	1380	459	ON049367
16	4-hydroxy-3-methylbut-2-enyl diphosphate reductase 2 (HDR2)	1932	1392	463	ON049368
17	Isopentenyl diphosphate isomerase I (IPPI 1)	1092	894	297	ON049369
18	Isopentenyl diphosphate isomerase II (IPPI 2)	1283	864	287	ON049370
19	Geranyl pyrophosphate synthase (GPS)	1976	1260	419	ON049371
20	Farnesyl pyrophosphate synthase 1 (FPPS 1)	1531	1047	348	ON049372
21	Farnesyl pyrophosphate synthase 2 (FPPS 2)	1436	1029	342	ON049373
22	Geranylgeranyl pyrophosphate synthase 1 (GGPPS 1)	1521	1107	368	ON049374
23	Geranylgeranyl pyrophosphate synthase 2 (GGPPS 2)	1436	1044	347	ON049375
24	Geranylgeranyl pyrophosphate synthase 3 (GGPPS 3)	2365	969	322	ON049376
25	Copalyl/labdadienyl diphosphate synthase 2 (CPS2)	2822	2451	816	ON049377
26	Kaurene synthase 1 (KS1)	2745	2391	796	ON049378
27	1-deoxy-D-xylulose 5-phosphate synthase 1 (DXS 1)	2909	2160	719	ON049379
28	1-deoxyxylulose 5-phosphate synthase Chloroplastic 2A (DXS 2A)	2638	2163	720	ON049380
29		2624	2208	735	ON049381

(*continued on next page*)

Table 4 (*continued*)

S. No.	Name of the gene	mRNA size (bp)	ORF size (bp)	ORF size (aa)	GenBank Accession number
	1-deoxy-D-xylulose-5-phosphate synthase 2 (DXS 2)				
30	UDP-glycosyltransferase 73 (UGT73)	1809	1365	454	ON049382
31	NADPH-cytochrome P450 reductase 2 (CPR2)	2637	2136	711	ON211623
32	Cytochrome P450 76A1 (CYP76A1)	1817	1647	548	ON211624
33	UDP-glycosyltransferase 86 (UGT86) (partial)	1301	1247	416	ON211625

GenBank Acc. No. ON049381), and the other is chloroplastic (AaDXS2a, GenBank Acc. No. ON049380). We also got 16 splice variants for the AaDXS1 gene. HDR gene was cloned from *A. paniculata* (ApHDR), and its expression was 5.5 to 16-fold higher in the leaves than in the roots (Shailaja et al., 2021). We identified two full-length HDR genes, AaHDR1 (459 amino acids) and AaHDR2 (463 amino acids), with 88 % and 82 % identity to ApHDR. Besides, for the first time, we identified full-length genes for DXR, MCT, CMK, MCS, and HDS from Andrographis.

In a common pathway, DMAPP synthesized from MVA and MEP pathways is converted into GGPP by geranyl pyrophosphate synthase (GPS), farnesyl pyrophosphate synthase (FPPS), and geranylgeranyl pyrophosphate synthase (GGPPS). For the first time, we identified one full-length gene for GPS and two for FPPS from Andrographis, along with four splice variants of FPPS2. There are at least four GGPPS genes in *A. paniculata,* and only ApGGPPS2 was reported to be involved in andrographolide biosynthesis (Wang et al., 2019). We also identified a full-length GGPPS2 gene from A. alata (AaGGPPS2), showing 99 % identity with ApGGPPS2.

GGPP is the common precursor for several diterpenoids compounds, including andrographolides. The committing enzyme that diverts GGPP to *ent*-labdane-related diterpenoids is *ent*-copalyl diphosphatase synthetase (*ent*-CPS). Misra et al. (Misra et al., 2015) cloned *ent*-CPS2 gene from *A. paniculata* (Ap-*ent*-CPS2), coding for 816 amino acids that converts GGPP to *ent*-copalyl diphosphate (*ent*-CPP) (Misra et al., 2015). The *ent*-CPS2 gene from *A. alata* (Aa-*ent*-CPS2) also encoded 816 amino acids and showed 99 % identity with ApCPS2. Shen et al. (2015) have cloned and functionally characterized another CPS2 gene with 832 aa and 57 % identity with Aa-*ent*-CPS2, but we did not find this gene in our study (Shen et al., 2016). The enzyme *ent*-kaurene synthase (KS1) converts *ent*-CPP to *ent*-copalol [8(17),13-*ent*-labdadiene-15,16,19-triol]. In *E. coli*, co-expression of *ent*-CPS2 from *A. paniculata* and KS1 from rice and wheat produced *ent*-copalol (Shen et al., 2016). In our study, we got one KS1 gene from *A. alata* (AaKS1), which coded for 796 amino acids with 99 % similarity to the KS1 from *A. paniculata* (ApKS1, Accession No. KY979588.1). Cytochrome P450 76 (CYP76), along with NADPH: cytochrome P450 reductase 2 (ApCPR2), converts *ent*-copalol to andrograpanin (Lin et al., 2017). We identified full-length CPR2 and CYP76 genes from *A. alata* (AaCPR2 and AaCYP76). In the final step, UDP-glycotransferase (UGT) enzyme glycosylates andrograpanin to produce neoandrographolide. Sun et al. (2019) functionally characterized recombinant UGT73AU1, UGT71, UGT76, UGT85, and UGT88 from *A. paniculata* and reported that ApUGT37AUI could convert

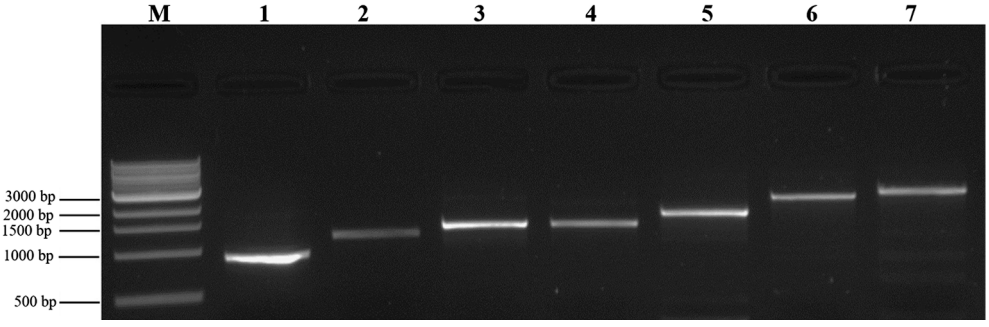

Fig. 8. Agarose gel electrophoresis of the RT-PCR amplified cDNA fragments of 1) 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MCS), 2) 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (MCT), 3) Mevalonate kinase (MVK), 4) Farnesyl pyrophosphate synthase 2 (FPPS2), 5) 4-hydroxy-3-methylbut-2-enyl diphosphate reductase 2 (HDR2), 6) 3-hydroxy-3-methylglutaryl-coenzyme A reductase 1A (HMGR1a), and 7) 1-deoxyxylulose 5-phosphate synthase Chloroplastic 2A (DXS2A). M denotes 1 Kb ladder (New England Biolabs, USA).

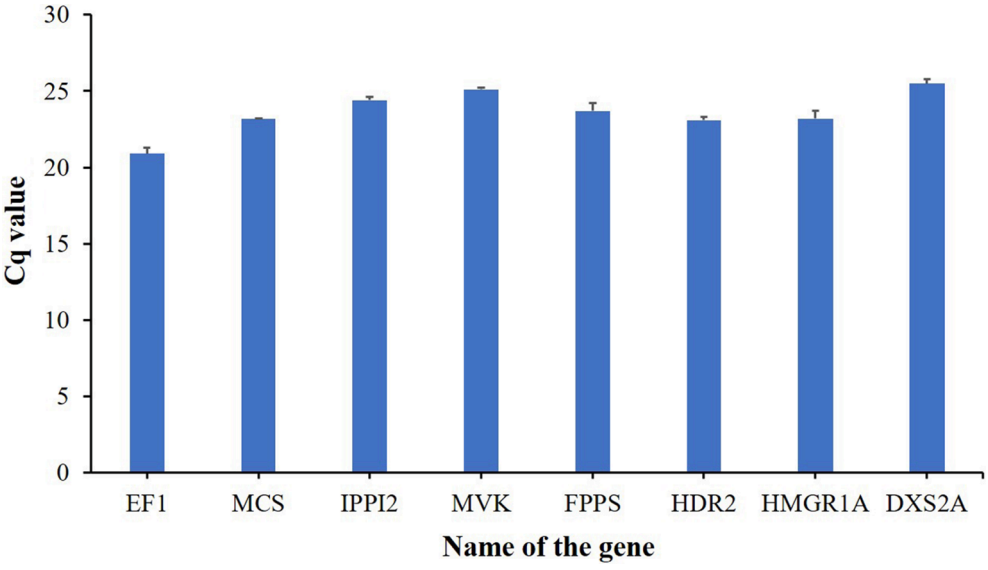

Fig. 9. Analysis of the expression of 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MCS), Isopentenyl diphosphate isomerase II (IPPI 2), Mevalonate kinase (MVK), Farnesyl pyrophosphate synthase 2 (FPPS2), 4-hydroxy-3-methylbut-2-enyl diphosphate reductase 2 (HDR2), 3-hydroxy-3-methylglutaryl-coenzyme A reductase 1A (HMGR1a), and 1-deoxyxylulose 5-phosphate synthase Chloroplastic 2A (DXS2A) genes involved in the neoandrographolide biosynthesis by qRT-PCR. The Elongation Factor 1 (EF1) is a housekeeping gene, which was used as a control. All the genes were analyzed in triplicate.

andrograpanin to neoandrographolide (Sun et al., 2019). However, an extensive functional characterization of 23 UGTs from *A. paniculata* revealed that UGT86CII could be the most probable enzyme, and it showed relatively higher catalytic activity than UGT73AUI (Srivastava et al., 2021). *A. alata* transcriptome contained 103 transcripts coding for 21 UGTs, including a full-length gene for UGT73AUI (AaUGT73AUI) and a partial gene for UGT86CII (AaUGT86CII). The proposed pathway of neoandrographolide biosynthesis based on the previous reports and the current study is given in Fig. 10.

NAG is a potential antiviral compound, and *A. alata* is the primary source of this compound. In this study, we identified one to three full-length genes for all the 22 enzymes involved in the synthesis of NAG, starting from acetyl CoA and D-glyceraldehyde 3 phosphate. Among the 32 full-length genes reported in this study, 15 were identified for the first time from *Andrographis* species. Interestingly only three genes showed splice variants. We also identified several transcripts related to the biosynthesis of other secondary metabolites. The full-length genes and transcripts identified from this study will provide valuable information for understanding the biosynthesis of NAG and other metabolites in *A. alata* and serve as an excellent genomics resource that will help genetic manipulation and metabolic engineering.

Funding

This study was financially supported by SRM-DBT Partnership Platform for Con-temporary Research Services and Skill Development in Advanced Life Sciences Technologies (Order No. BT/PR12987/INF/22/205/2015.

Data Availability Statement: The data was submitted to NCBI SRA database with Accession Number SRR18210043. Full length gene was

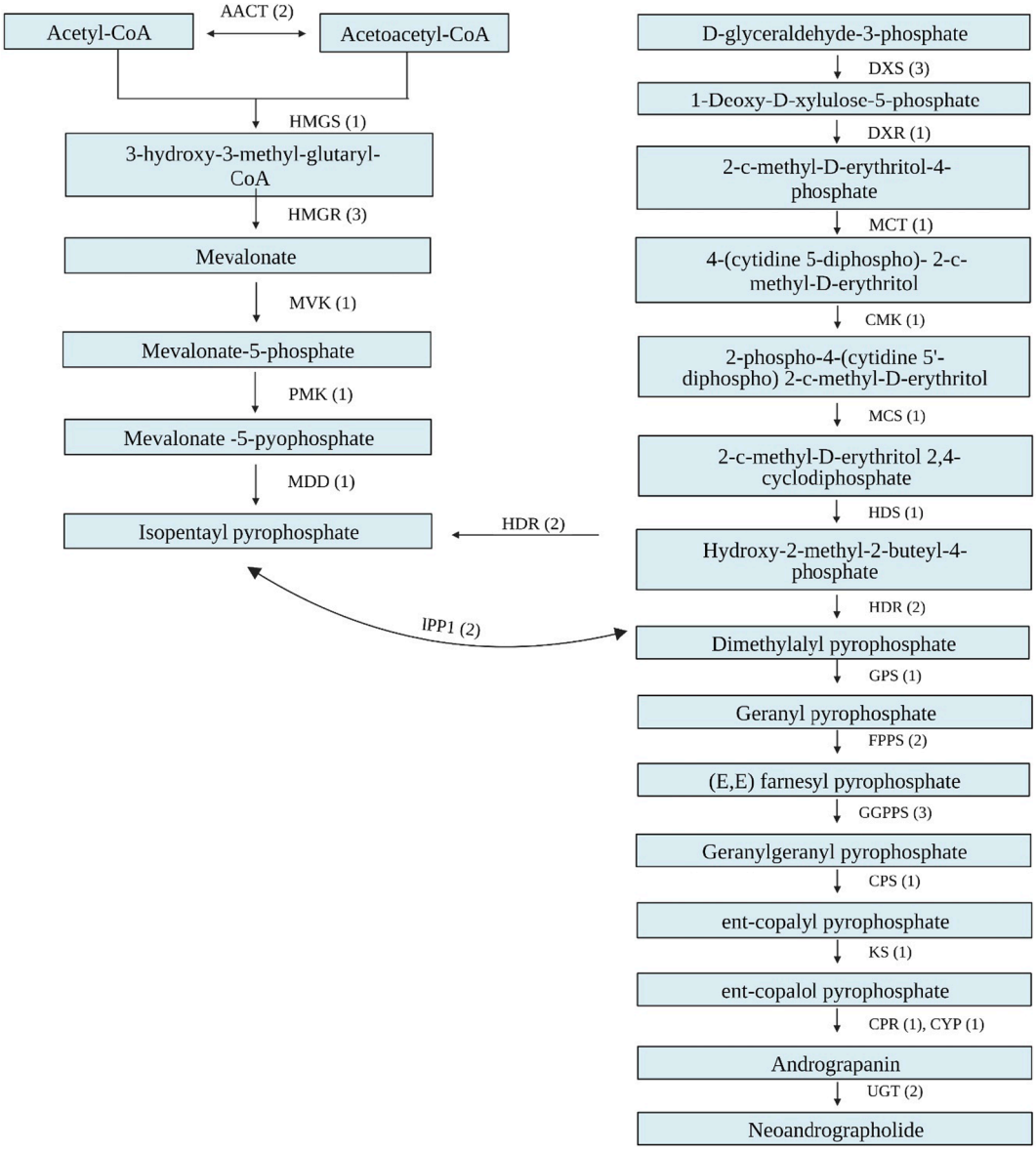

Fig. 10. A schematic representation of proposed pathways for the biosynthesis of neoandrographolide. The numbers given in the bracket indicate the number of genes identified for each enzyme. Abbreviations: AACT: Acetyl- CoA C- acetyl transferase; HMGS: 3-hydroxy-3-methylglutaryl coenzyme A synthase; HMGR 3-hydroxy-3-methylglutaryl coenzyme A reductase; MVK: Mevalonate kinase; PMK: 5-phosphomevalonate kinase; MDD: Mevalonate diphosphate decarboxylase; IPPI: Isopentenyl diphosphate isomerase; DXS: 1-deoxy-D-xylulose-5-phosphate synthase; DXR: 1-deoxy-D-xylulose 5-phosphate reductoisomerase; MCT: 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase; CMK: 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; MCS: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; HDS: 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate synthase; HDR: 4-hydroxy-3-methylbut-2-enyl diphosphate reductase; GPS: Geranyl pyrophosphate synthase; FPPS: Farnesyl pyrophosphate synthase; GGPPS: Geranylgeranyl pyrophosphate synthase; CPS: Copalyl/labdadienyl diphosphate synthase; KS: Kaurene synthase; CPR: NADPH-cytochrome P450 reductase; CYP: Cytochrome P450; UGT: UDP-glycosyltransferase.

submitted under the GenBank accession numbers ON049353 to ON049382 and ON211623 to ON211625.

CRediT authorship contribution statement

Tanuja: Conceptualization, Methodology, Investigation, Writing – original draft. **Madasamy Parani:** Conceptualization, Writing – review & editing, Project administration, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors are unable or have chosen not to specify which data has been used.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2022.146981.

References

Andrews, Simon, 2010. FastQC: a quality control tool for high throughput sequence data 2010.

Balu, S., Alagesaboopathi, C., 1993. Anti-inflammatory activities of some species of andrographis wall. (ACANTHACEAE) 1–2, 180–184.

Balu, S., Boopathi, C.A., Elango, V., Nadu, T., 1993. Antipyretic activities of some species of andrographis wall. Anc Sci Life 3–4, 399–402.

Bhambhani, S., Lakhwani, D., Gupta, P., Pandey, A., Dhar, Y.V., Kumar Bag, S., Asif, M. H., Kumar Trivedi, P., 2017. Transcriptome and metabolite analyses in Azadirachta indica: identification of genes involved in biosynthesis of bioactive triterpenoids. Sci. Rep. 1–12.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 30, 2114–2120.

Calabrese, C., Berman, S.H., Babish, J.G., Ma, X., Shinto, L., Dorr, M., Wells, K., Wenner, C.A., Standish, L.J., 2000. A phase I trial of andrographolide in HIV positive patients and normal volunteers. Phytother Res 14, 333–338.

Chen, H., Ma, Y.B., Huang, X.Y., Geng, C.A., Zhao, Y., Wang, L.J., Guo, R.H., Liang, W.J., Zhang, X.M., Chen, J.J., 2014. Synthesis, structure-activity relationships and biological evaluation of dehydroandrographolide and andrographolide derivatives as novel anti-hepatitis B virus agents. Bioorganic Med. Chem. Lett. 24, 2353–2359.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21, 3674–3676.

Dalawai, D., Aware, C., Jadhav, J.P., Murthy, H.N., 2021. RP-HPLC analysis of diterpene lactones in leaves and stem of different species of andrographis. Nat. Prod. Res. 35, 2239–2242.

Garg, A., Agrawal, L., Misra, R.C., Sharma, S., Ghosh, S., 2015. Andrographis paniculata transcriptome provides molecular insights into tissue-specific accumulation of medicinal diterpenes. BMC Genom. 16, 1–16.

Gnanasekaran, G., Murthy, G.V.S., 2015. Epitypification of Andrographis Alata (Acanthaceae). Telopea (Syd) 18, 165–166.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.

Jian-Xin Chen,*,a,b Hui-Jun Xue,a Wen-Cai Ye, b Bing-Hu Fang,a Ya-Hong Liu,a Shao-Hua Yuan,c Pei Yu,*, band Yu-Qiang Wangb, 2009. Activity of andrographolide and its derivatives against influenza virus in vivoand in vitro. Biol. Pharm. Bull. (2009) 1385–1391.

Kadapatti, S.S., Murthy, H.N., 2021. Rapid plant regeneration, analysis of genetic fidelity, and neoandrographolide content of micropropagated plants of andrographis alata (Vahl) Nees. J. Genet. Eng. Biotechnol. 19, 1–8.

Kumar, R., Ichihashi, Y., Kimura, S., Chitwood, D.H., Headland, L.R., Peng, J., Maloof, J. N., Sinha, N.R., 2012. A high-throughput method for illumina RNA-Seq library preparation. Front. Plant Sci. 3, 202.

Lee, J.C., Tseng, C.K., Young, K.C., Sun, H.Y., Wang, S.W., Chen, W.C., Lin, C.K., Wu, Y. H., 2014. Andrographolide exerts anti-hepatitis C virus activity by up-regulating haeme oxygenase-1 via the P38 MAPK/Nrf2 pathway in human hepatoma cells. Br. J. Pharmacol. 171, 237–252.

Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22 (13), 1658–1659.

Li, F., Khanom, W., Sun, X., Paemanee, A., Roytrakul, S., Wang, D., Smith, D.R., Zhou, G. C., 2020. Andrographolide and Its 14-aryloxy analogues inhibit zika and dengue virus infection. Molecules 25, 5037.

Liang, Y., Chen, S., Wei, K., Yang, Z., Duan, S., Du, Y., Qu, P., Miao, J., Chen, W., Dong, Y., 2020. Chromosome level genome assembly of andrographis paniculata. Front. Genet. 11, 701.

Lin, T.-P., Chen, S.-Y., Duh, P.-D., Chang, L.-K., Liu, Y.-N., 2008. Inhibition of the epstein-barr virus lytic cycle by andrographolide. Biol. Pharm. Bull 31, 2018–2023.

Lin, H., Wang, J., Qi, M., Guo, J., Rong, Q., Tang, J., Wu, Y., Ma, X., Huang, L., 2017. Molecular cloning and functional characterization of multiple NADPH-cytochrome P450 reductases from andrographis paniculata. Int. J. Biol. Macromol 102, 208–217.

Misra, R.C., Garg, A., Roy, S., Chanotiya, C.S., Vasudev, P.G., Ghosh, S., 2015. Involvement of an ent-copalyl diphosphate synthase in tissue-specific accumulation of specialized diterpenes in andrographis paniculata. Plant Sci. 240, 50–64.

Murugan, N.A., Pandian, C.J., Jeyakanthan, J., 2021. Computational investigation on andrographis paniculata phytochemicals to evaluate their potency against SARS-CoV-2 in comparison to known antiviral compounds in drug trials. J. Biomol 39, 4415–4426.

Patel, A.A., Shukla, Y.M., Kumar, S., Sakure, A.A., Parekh, M.J., Zala, H.N., 2020. Transcriptome analysis for molecular landscaping of genes controlling diterpene andrographolide biosynthesis in Andrographis paniculata (Burm. f.) Nees. 3 Biotech 10 (12).

Sa-Ngiamsuntorn, K., Suksatu, A., Pewkliang, Y., Thongsri, P., Kanjanasirirat, P., Manopwisedjaroen, S., Charoensutthivarakul, S., Wongtrakoongate, P., Pitiporn, S., Chaopreecha, J., et al., 2021. Anti-SARS-CoV-2 activity of andrographis paniculata extract and its major component andrographolide in human lung epithelial cells and cytotoxicity evaluation in major organ cell representatives. J. Nat. Prod 84, 1261–1270.

Shailaja, A., Srinath, M., Bindu, B.V.B., Giri, C.C., 2021. Isolation of 4-hydroxy 3-methyl 2-butenyl 4-diphosphate reductase (ApHDR) gene of methyl erythritol diphosphate (MEP) pathway, in silico analysis and differential tissue specific ApHDR expression in Andrographis paniculata (Burm. f) Nees. Physiol. Mol. Biol. Plants 27 (2), 223–235.

Shen, Q., Li, L., Jiang, Y., Wang, Q., 2016. Functional characterization of ent-copalyl diphosphate synthase from andrographis paniculata with putative involvement in andrographolides biosynthesis. Biotechnol 38, 131–137.

Srinath, M., Bindu, B.B.V., Shailaja, A., Giri, C.C., 2019. Isolation, Characterization and in Silico Analysis of 3-Hydroxy-3-Methylglutaryl-Coenzyme A Reductase (HMGR) Gene from Andrographis Paniculata (Burm. f) Nees, Mol. Biol. Rep. 47 (2019) 639–654.

Srinath, M., Shailaja, A., Bindu, B.B.V., Giri, C.C., 2021. Molecular cloning and differential gene expression analysis of 1-deoxy-D-xylulose 5-phosphate synthase (DXS) in andrographis paniculata (Burm. f) Nees. Mol. Biotechnol. 2021, 63, 109–124.

Srivastava, N., Akhila, A., 2010. Biosynthesis of andrographolide in andrographis paniculata. Phytochemistry 71 (11-12), 1298–1304.

Srivastava, P., Garg, A., Misra, R.C., Chanotiya, C.S., Ghosh, S., 2021. UGT86C11 Is a novel plant UDP-glycosyltransferase involved in labdane diterpene biosynthesis. J. Biol. Chem. 297, 101045.

Sun, W., Leng, L., Yin, Q., Xu, M.M., Huang, M., Xu, Z., Zhang, Y., Yao, H., Wang, C., Xiong, C., et al., 2019. The genome of the medicinal plant andrographis paniculata provides insight into the biosynthesis of the bioactive diterpenoid neoandrographolide. Plant J. 97, 841–857.

Tanuja Parasar N.R., Kumar R., Natarajan, P., Parani, M., 2021. De Novo Assembly, Annotation and Molecular Marker Identification from the Leaf Transcriptome of Ocimum Gratissimum L. Plant Genet. 2021, 1–8.29.

Wang, J., Lin, H.X., Su, P., Chen, T., Guo, J., Gao, W., Huang, L.Q., 2019. Molecular cloning and functional characterization of multiple geranylgeranyl pyrophosphate synthases (ApGGPPS) from Andrographis Paniculata. Plant Cell Rep. 38, 117–128.

Wintachai, P., Kaur, P., Lee, R.C.H., Ramphan, S., Kuadkitkan, A., Wikan, N., Ubol, S., Roytrakul, S., Chu, J.J.H., Smith, D.R., 2015. Activity of andrographolide against chikungunya virus infection. Sci. Rep 5, 1–14.

Yu, B., Dai, C.Q., Jiang, Z.Y., Li, E.Q., Chen, C., Wu, X.L., Chen, J., Liu, Q., Zhao, C.L., He, J.X., et al., 2014. Andrographolide as an anti-H1N1 drug and the mechanism related to retinoic acid-inducible Gene-I-like receptors signaling pathway. Chin. J. Integr. Med. 20, 540–545.

Yuan, L., Zhang, C., Sun, H., Liu, Q., Huang, J., Sheng, L., Lin, B., Wang, J., Chen, L., 2016. The semi-synthesis of novel andrographolide analogues and anti-influenza virus activity evaluation of their derivatives. Bioorganic Med. Chem. Lett. 26, 769–773.