Molecular Biology Reports (2022) 49:10307–10314
https://doi.org/10.1007/s11033-022-07784-5

ORIGINAL ARTICLE

De novo transcriptome analysis of *Justicia adhatoda* reveals candidate genes involved in major biosynthetic pathway

Deepthi Padmanabhan[1] · Adil Lateef[1] · Purushothaman Natarajan[1,2] · Senthilkumar Palanisamy[1]

Received: 16 January 2022 / Accepted: 8 July 2022 / Published online: 12 September 2022

Abstract

Background *Justicia adhatoda* is an important medicinal plant traditionally used in the Indian system of medicine and the absence of molecular-level studies in this plant hinders its wide use, hence the study was aimed to analyse the genes involved in its various pathways.

Methods and results The RNA isolated was subjected to Illumina sequencing. *De novo* assembly was performed using TRINITY software which produced 171,064 transcripts with 55,528 genes and N50 value of 2065 bp, followed by annotation of unigenes against NCBI, KEGG and Gene ontology databases resulted in 105,572 annotated unigenes and 40,288 non-annotated unigenes. A total of 5980 unigenes were mapped to 144 biochemical pathways, including the metabolism and biosynthesis pathways. The pathway analysis revealed the major transcripts involved in the tryptophan biosynthesis with TPM values of 6.0903, 33.6854, 11.527, 1.6959, and 8.1662 for Anthranilate synthase alpha, Anthranilate synthase beta, Arogenate/Prephenate dehydratase, Chorismate synthase and Chorismate mutase, respectively. The qRT-PCR validation of the key enzymes showed up-regulation in mid mature leaf when compared to root and young leaf tissue. A total of 16,154 SSRs were identified from the leaf transcriptome of *J. Adhatoda* ,which could be helpful in molecular breeding.

Conclusions The study aimed at identifying transcripts involved in the tryptophan biosynthesis pathway for its medicinal properties, as it acts as a precursor to the acridone alkaloid biosynthesis with major key enzymes and their validation. This is the first study that reports transcriptome assembly and annotation of *J. adhatoda* plant.

Keywords RNA sequencing · *De novo* assembly · Annotation · SALMON · SSR · qRT-PCR.

Introduction

Justicia adhatoda Linn. is a perennial shrub found in the tropical regions of Southeast Asia belonging to the *Acanthaceae* family. It is widely branched with pink, purple, or white flowers and can grow up to a height of 2.5 m. The plant has been widely used in the Ayurvedic and Unani systems of Indian medicine for treating various diseases such as bronchitis, common cold, asthma, cough, and tuberculosis [1]. *J. adhatoda* has multiple pharmacological properties that have been used to treat blood disorders, jaundice, mouth ulcers, vomiting, fever and heart-related problems [2, 3]. Phytochemical screening of *J. adhatoda* methanolic leaf extract revealed the presence of alkaloids, anthraquinones, cardiac glycosides, flavonoids, saponins, polyphenols, terpenoids and phytosterols [4]. The leaf extract of *J. adhatoda* acts as a potent anti-diabetic and anti-lipidaemic agent [5] and is also found to possess various pharmacological properties such as anti-microbial, hepatoprotective, wound healing, anti-ulcer, and anti-inflammatory activity [6]. The major alkaloids present in the leaf extracts are quinazoline alkaloids such as vasicoline, adhatodine, vasicolinone, adhavasinone, and

✉ Senthilkumar Palanisamy
mpsenthilkumar@gmail.com

Deepthi Padmanabhan
deepthipadmanabhan2010@gmail.com

Adil Lateef
lateef.adil@gmail.com

Purushothaman Natarajan
purushoth.n@gmail.com

[1] Department of Genetic Engineering, School of Bioengineering, SRM Institute of Science and Technology, 603203 Kattankulathur, India

[2] Associate Faculty Member, Department of Biology, , West Virginia State University, 225, Hamblin Hall, 25112-1000 Institute, WV, USA

vasicol [6]. GC-MS analysis of methanolic leaf extract of *J.adhatoda* revealed the presence of various bioactive compounds such as 9, 12, 15-Octadecatrienoic acid (Z, Z, Z), n-Hexadecaonoic acid, Phytol, and Amrinone phytocompounds [7]. Anisotine and vasicoline are potent inhibitors of the main protease, which is a key component for cleaving the viral polyprotein that is targeted for treating COVID-19 [8]. The development of next-generation sequencing has been useful to elucidate the secondary metabolites and its candidate genes with low labour and cost-effectiveness [9]. *De novo* sequencing plays a vital role in gene discovery where a reference genome is not available in a non-model organism, as it is a cost-effective method [10]. An assembly of short-read sequence data is used to identify the candidate genes involved in biosynthetic pathways and gene expression analysis [10]. Quantitative real-time PCR elucidates the level of gene expression that encodes the various components involved in the biosynthetic pathways. Despite the well-established role of *J. adhatoda* in the Indian medicinal system, genomic level study of the plant is not well known [11]. Thus, the present study provides a better understanding of the biosynthetic pathways and the candidate genes that are attributed to the medicinal properties of *J. adhatoda*.

Materials and methods

Plant material and RNA isolation

Mature and healthy leaves of *J.adhatoda* were collected and RNA extraction was carried out using TRIzol® Reagent (Invitrogen, USA). The total RNA extracted was treated with DNase A and purified with the RNeasy MinElute clean-up kit (Qiagen Inc., GmbH, Germany, USA). The quantity and quality of total RNA extracted were evaluated using Nanodrop 2000 spectrophotometer (Thermo Scientific, Wilmington, Delaware, USA). RNA integrity value was measured using Bioanalyzer 2100 (Agilent Technologies, Santa Clara, California, USA). The sample was further subjected to cDNA library preparation to carry out Illumina sequencing [11].

Library preparation and Illumina sequencing

The total RNA was made rRNA free using the Ribo-Zero rRNA removal kit (Illumina Inc., Singapore) and the remaining fraction was purified and eluted.The purified RNA sample was fragmented into short sequences using fragmentation buffer, and these fragments were used for first-strand cDNA synthesis using superscript II reverse transcriptase (Invitrogen, Carlsbad, Califonia, USA). The second strand was synthesized and purified. Poly-A tailing and adapter ligation was performed for paired-end library preparation. The primers based on adapter region were used for amplification of library for enrichment of the cDNA fragments. The quality of sequencing library was assessed using Caliper LabChip GX using HT DNA High Sensitivity Assay kit (Caliper Life Sciences Inc., USA) Sequencing was performed with the NextSeq 500 using TruSeq v3-HS kit to generate 100 bp paired-end reads (Illumina Inc., USA).

De novo transcriptome assembly and clustering

The quality of raw paired-end reads was assessed using FastQC v0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and the adapter sequences and low-quality bases were removed using Cutadapt v3.5 and Sickle v1.33 tools, respectively [12–14]. The bases with Phred score ≥ 30 reads were retained, and the *de novo* assembler Trinity v2.13.2 [15] was used for the transcriptome assembly, which combines three independent software modules: *Inchworm* assembles the data into full-length transcripts for a dominant isoform;*Chrysalis* clusters the contigs and constructs de Bruijn graphs for each cluster, representing the transcriptional complexity for a gene; *Butterfly* processes the graphs and reports full-length transcripts for alternatively spliced isoforms. CD-HIT v4.8.1 [16] was used for the clustering to remove the redundancy of assembled contigs.

Assessment of gene completeness

An online tool TRAPID was used to analyse functional and comparative *de novo* transcriptome data set (http://bioinformatics.psb.ugent.be/trapid_02/), which compared the unigene transcripts against PLAZA 4.5 dicots plant database with an E-value < 1E-5 for significant similarity search and annotation [17]. Considering one or more hits in the TRAPID database for full length, quasi full length, or partial length based on ORF, unigene completeness was assessed [18].

Functional annotation and gene ontology (GO) classification

The sequences of transcriptome assembly were compared against the non-redundant (nr) protein sequence database at National Centre for Biotechnology Information (NCBI) using BLASTX function from BLAST + package. The output was further analysed by Blast2GO software for retrieving

the Gene Ontology (GO) terms of assembled unigenes[19]. The Kyoto encyclopedia of genes and genomes (KEGG) was performed for pathway mapping. The KEGG Automated Annotation Server (KAAS) database was also used for mapping the biosynthetic pathways of various secondary metabolites with an E-value of 1E-05. The GO terms were classified based on three different categories including Biological process, Molecular function and Cellular component.

Simple sequence repeat (SSR) detection

Simple Sequence Repeats (SSRs) of *J. adhatoda* were identified using MISA tool (https://webblast.ipk-gatersleben.de/misa/) with default parameters to identify mono-, di-, tri-,tetra-, penta-, and hexa-nucleotide motifs as 10, 6, 5, 5, 5, and 5 repeats respectively.

Transcript quantification

The quantification of transcripts of *J. adhatoda* leaf was performed using the SALMON tool [20]. It quantifies the transcripts based on the GC content, which estimates the sensitivity and abundance of differential gene expression, and quasi mapping of reads was done, where the transcripts mapped itself fast and accurately.

Validation by real-time PCR

To validate the transcriptome assembly as well as to understand the expression of unigenes from *J. adhatoda*, qRT-PCR analysis was performed using QuantStudio 5 Real-Time PCR system (Thermo Scientific, Wilmington, Delaware, USA) and QuantiNova SYBR Green PCR Kit (Qiagen Inc., GmbH, Germany). Control reaction without a template was included for each selected gene. The Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) from *J. adhatoda* was used as an internal reference gene for normalization and estimation of gene expression. A comparative Ct ($2^{-\Delta\Delta Ct}$) method was used to analyse the qRT-PCR data, and fold change in gene expression was calculated using ΔCt values [21]. The experiment was repeated with three technical and two biological replicates.

Table 1 Summary of paired-end sequencing and *de novo* assembly of *J. adhatoda* transcriptome

Particulars	Numbers
Number of raw reads	51,639,915
Number of clean reads	39,134,610
No. of bases (after processing)	33,247,241
Mean Phred score	37.45
Total transcripts	171,064
Percentage of successful assembly from raw reads	64.38
Average length (bp)	1383.89
Median contig length	1096
GC %	41.69
Contig N50 (bp)	2065

Results

RNA sequencing and *de novo* transcriptome assembly

RNA sequencing of leaf transcriptome from *J. adhatoda* generated 51,639,915 raw reads and the pre-processing of raw data was performed to remove the adapter sequences, 39,134,610 high quality reads were retained with GC content of 41.69%. The raw reads were deposited at National Centre for Biotechnology Information (NCBI) Short Read Archive (SRA) database under the Accession number PRJNA84216. A total of 171,064 transcripts were assembled from raw reads with maximum and minimum unigene lengths as 15,630 bp and 201 bp, respectively with an N50 value of 2065 bp for the assembled transcriptome. The transcriptome assembly details are mentioned in Table 1 and The length distribution of unigenes is given in Supplementary Fig. 1.

Assessment of gene completeness

Gene completeness analysis of *J. adhatoda* resulted in meta annotation of 30,857 (28.9%) full-length unigenes, 14,071 (13.2%) quasi full-length unigenes, 22,737 (21.3%) of partial unigenes, and 39,221 (36%) unigenes did not match to any proteins in PLAZA 4.5 dicots plant database.

Functional annotation of unigenes

The *de novo* assembled unigenes of *J. adhatoda* were annotated for sequence similarity search using BLASTX against a non-redundant protein database at NCBI with a 1E-5. The results showed 105,572 annotated unigenes and 40,288 non-annotated unigenes, of which 4772 unigenes were predicted due to inadequate genomic information in public databases. The unigene similarity search reveals that the plant has

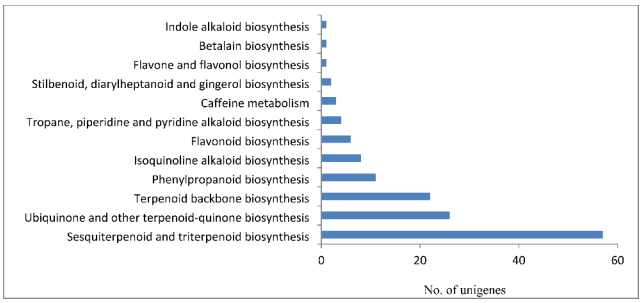

Fig. 1 KEGG pathway unigene assignments based on secondary metabolite biosynthesis

the best similarity with *Erythranthe guttata* (55%), *Coffea canephora* (6.3%), Utricularia gibba (3.4%), among others. Thus, the results indicated that *J .adhatoda* is more closely associated with *Erythranthe guttata.*

Gene Ontology (GO) classification of unigenes

GO classification was performed using the BLAST2GO tool in order to classify the genes based on gene annotation into three different categories: molecular function, cellular component, and biological process and 47 sub-categories. A total of 143,277 genes were assigned to 20 classes in the biological function, 26,732 genes in organic substance metabolic process, 26,713 genes in primary metabolic process, 21,100 genes in nitrogen compound metabolic process, 19,714 genes in the cellular metabolic process, and 11,476 genes in the biosynthetic process. In molecular function, 93,438 genes were classified into 14 sub-categories, which include the heterocyclic compound of 21,276 genes, organic cyclic compound binding with 21,276 genes, 16,105 genes with transferase activity, 13,563 genes with small-molecule binding and 12,904 genes with hydrolase activity. Additionally, 95,725 genes were assigned to cellular component category with 13 sub-categories where intracellular had the highest number of unigenes (23,055), the intracellular part 21,946 genes, 18,296 genes of intracellular organelle, and 16,066 genes of the membrane-bound organelle. (Supplementary Fig. 2).

Biological pathway analysis

The biochemical pathways of *J. adhatoda* were identified by mapping the unigenes to the KEGG pathway database with the help of the KAAS and BLAST2GO software with 5981 unigenes annotated to 144 biochemical pathways. A total of 6409 unigenes were assigned to the metabolic pathway, including nucleotide metabolism, carbohydrate

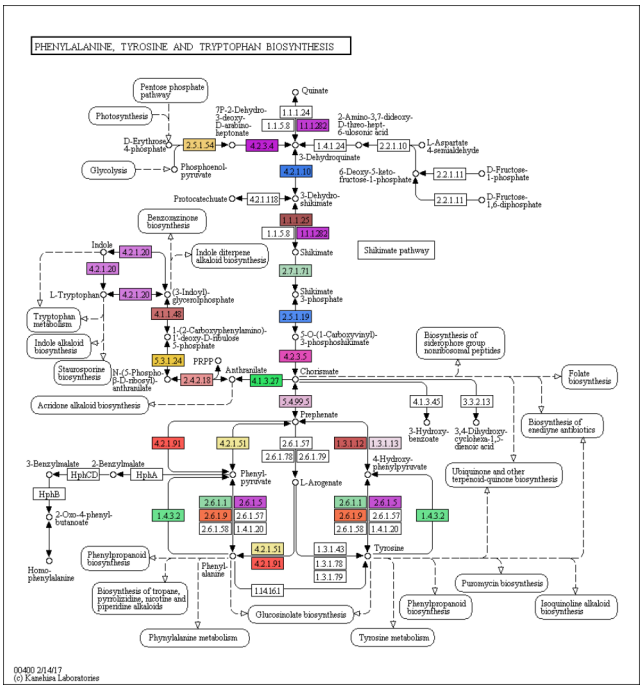

Fig. 2 Tryptophan biosynthesis pathway based on KEGG mapping represents the various enzymes identified from transcriptome data concerning *Justicia adhatoda* in each color along with its enzyme code (EC number). The KEGG pathway map is adapted from map00400 https://www.kegg.jp/keggbin/highlight_pathway?scale=1.0&map=map00400&keyword=tryptophan

metabolism, and amino acid metabolism having 2539, 347, and 477 unigenes, respectively (Supplementry Fig. 3). The KEGG pathway analysis of secondary metabolite biosynthesis was divided into 12 sub-categories, where the highest number of unigenes were found in sesquiterpenoid and triterpenoid biosynthesis (57 unigenes), followed by ubiquinone and another terpenoid-quinone biosynthesis (26 unigenes), terpenoid backbone biosynthesis (22 unigenes), and phenylpropanoid biosynthesis (11 unigenes) (Fig. 1).

Tryptophan biosynthesis

The major genes involved in the tryptophan biosynthesis pathway from the KEGG database are presented in Supplemnetry Table 1 and the genes identified in this pathway from transcriptome data have been depicted in Table 2. KEGG pathway map representing the biosynthesis of Tryptophan has been depicted in Fig. 2.

Table 2 Major genes involved in tryptophan biosynthesis pathway from the *J.adhatoda* leaf transcriptome

Gene name	E C Number	Unigene ID	Unigene length	TPM value	No. Of Unigene
Anthranilte synthase alpha	4.1.3.27	TRINITY_DN13973_c0_g1_i2	524	6.0903	1
Anthranilate synthase beta	4.1.3.27	TRINITY_DN9855_c0_g1_i4	2058	33.6854	1
Arogenate/ Prephenate dehydratase	4.1.1.48	TRINITY_DN8030_c0_g1_i1	1605	6.9556	7
		TRINITY_DN11169_c0_g1_i1	2648	0.8030	
		TRINITY_DN11169_c0_g1_i2	3676	2.3483	
		TRINITY_DN11169_c0_g1_i3	3139	8.1132	
		TRINITY_DN9715_c0_g1_i2	1765	11.527	
		TRINITY_DN9715_c0_g1_i3	666	6.4298	
		TRINITY_DN13484_c0_g1_i1	769	5.9688	
Chorismate synthase	4.2.3.5	TRINITY_DN12680_c0_g1_i3	2687	1.6959	2
		TRINITY_DN12680_c0_g1_i5	1132	0.8975	
Chorismate mutase	5.4.99.5	TRINITY_DN11294_c0_g1_i1	2184	8.1662	5
		TRINITY_DN11294_c0_g1_i5	1293	2.2369	
		TRINITY_DN11294_c0_g1_i6	2287	2.1131	
		TRINITY_DN11294_c0_g1_i7	2016	1.0050	
		TRINITY_DN11295_c0_g1_i1	1170	4.4995	

Identification of simple sequence repeats (SSRs)

A total of 25,978 SSRs were identified from 106,886 sequences with 4374 sequences having more than 1 SSR, and 1951 SSRs were present in compound formation. The number of SSR loci with 8530 di-nucleotide repeats, 6802 tri-nucleotide repeats, 665 tetra-nucleotide repeats, 96 penta nucleotide repeats, and 61 hexa nucleotide repeats are represented in Supplementry Table 2. SSRs with five tandem repeats (3984) were the most common in *J. adhatoda*, followed by six tandem repeats (3902), seven tandem repeats (2289), nine tandem repeats (2087), eight tandem repeats (1622), and ten tandem repeats (589). Among di-nucleotide repeats, AT/AT was found to be the highest with 2968 repeats, followed by AG/CT with 2294 repeats, and tri-nucleotide repeats AAG/CTT has the highest frequency of 1737 and AAT/ATT with 1669 repeats and other motifs distributed uniformly (Supplementry Fig. 4).

Transcript quantification

The expression levels of *de novo* assembled unigenes of the *J.adhatoda* leaf transcriptome were calculated based on TPM values with the help of the SALMON tool. In the tryptophan biosynthesis pathway, the TPM values of key enzymes involved are 6.0903, 33.6854, 11.527, 1.6959, and 8.1662 for Anthranilate synthase alpha, Anthranilate synthase beta, Arogenate/ Prephenate dehydratase, Chorismate synthase, and Chorismate mutase, respectively (Table 2). The top 10 most abundant unigenes in the *J.adhatoda* leaf transcriptome are represented in Supplementry Table 3.

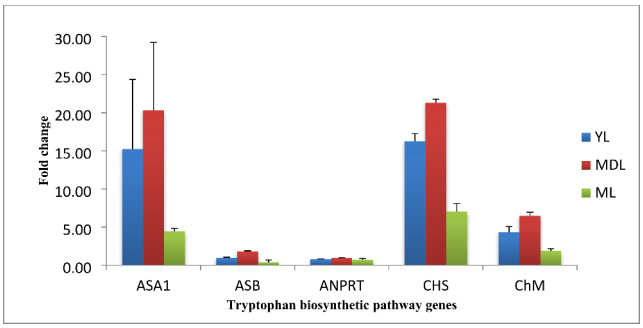

Fig. 3 qRT-PCR validation of Anthranilate synthase alpha 1 (ASA1), Anthranilate synthase beta (ASB), Arogenate/prephenate dehydratase (ANPRT), Chorismate synthase and Chorismate mutase in a young leaf (YL), mid-mature leaf (MDL), and root (ML) of *Justicia adhatoda*

Validation by real-time PCR

The qRT-PCR analysis was performed to analyse the expression pattern of the selected tryptophan biosynthesis genes and to validate the transcriptome assembly. The genes selected are Anthranilate synthase alpha (EC: 4.1.3.27), Anthranilate synthase beta (EC: 4.1.3.27), Arogenate/ Prephenate dehydratase (EC: 4.1.1.48), Chorismate synthase (EC: 4.2.3.5) and Chorismate mutase (EC: 5.4.99.5). Anthranilate synthase alpha, Anthranilate synthase beta, Chorismate synthase and Chorismate mutase showed significant up-regulation in mature leaf when compared to young leaf and root. Arogenate/prephenate dehydratase was down-regulated in mature leaf tissue. GAPDH was used as a housekeeping gene for the gene expression analysis. Anthranilate synthase beta and Arogenate/prephenate dehydratase though with low TPM values were chosen for qRT-PCR analysis, as they were also involved in the tryptophan biosynthesis pathway (Fig. 3).

Discussion

De novo transcriptome assembly of *J. adhatoda* revealed various enzymes involved in the secondary metabolite biosynthesis pathway that play a vital role in exhibiting various pharmacological properties. These pathways were identified using KEGG database where 144 biochemical pathways were identified using the assembled transcripts. Tryptophan is an essential amino acid, and a substrate, required for synthesis of serotonin, a neurotransmitter produced in the brain. The major genes identified in tryptophan biosynthesis were Anthranilate synthase alpha (EC: 4.1.3.27) and Anthranilate synthase beta (EC: 4.1.3.27). They are involved in the conversion of Chorismate into Anthranilate; Arogenate/prephenate dehydratase (EC: 4.1.1.48) converts 1-(2-Carboxyphenylamino)-1'-deoxy-D-ribulose 5-phosphate into (3-Indoyl)-glycerolphosphate, Chorismate synthase (EC:4.2.3.5) converts Chorismate into 5-O-(1-Carboxyvinyl)-3-phosphoshikimate, and Chorismate mutase (EC: 5.4.99.5) converts Prephenate into Chorismate. Anthranilate synthase is a rate-limiting enzyme involved in the terpenoid indole alkaloid pathway (TIA) [22]. It is a holoenzyme with two heterotetramers consisting of two alpha and beta subunits. The alpha subunits catalyse the formation of Chorismate to Anthranilate, which plays a vital role in the TIA pathway along with the binding site of tryptophan involved in feedback inhibition. The aminotransferases activity of the beta subunit plays a vital role in the transfer of amino groups from Glutamine to the alpha subunit of Anthrnilate synthase [23, 24]. Chorismate is the final compound of the shikimate pathway that is formed by the enzyme chorismate synthase. This enzyme catalyses the trans-1,4 elimination of phosphate from 5-enolpyruvylshikiukimate 3-phosphate, and it is of two types: fungal-type bifunctional chorismate synthase with NADPH-dependent flavin reductase and bacterial and plant-type monofunctional chorismate synthase [25, 26]. Prephenate aminotransferase and arogenate dehydratase catalyses the final step of phenylalanine production [27–29]. Anthranilate synthase is a rate-limiting enzyme that catalyses chorismate to anthranilate in the Indole pathway biosynthesis [22, 23]. Anthranilate synthase consists of two subunits in its holoenzyme such as alpha and beta subunits. Alpha subunit plays a vital role in catalysing chorismate to anthranilate by feedback inhibition. The aminotransferase activity of the beta subunit of anthranilate synthasetransfers the amino group from glutamine to the alpha subunit [24]. Overexpression of anthranilate synthase increased the level of tryptophan that acts as a precursor for various biosynthetic pathways [30]. The production of vasicinone and vasicine was enhanced by stimulating anthranilate synthase activity by increasing the tryptophan and sorbitol in the culture [31]. Chorismate mutase is a key enzyme that catalyses the formation of prephenate from chorismate for the biosynthesis of aromatic amino acids from the shikimate pathway [32]. In a study, Chorismate mutase was found to be a putative enzyme that is bifunctional and involved in the biosynthesis of phenylalanine, osmotic, and antibiotic tolerance [33, 34]. Acridone is a major alkaloidal compound present in *J. adhatoda*, which acts as an anti-cancer agent [35]. It is aheterocyclic alkaloid that contains tri-cyclic rings with a carbonyl group at the ninth position and nitrogen at the tenth position [36]. There are various acridone derivatives such as glyforine, thioacridones, and acronycine, all of which exhibit pharmacological properties such as antimicrobial, antipsoriatic, anti-malarial, and anti-cancer activity [37–40]. 25,978 Simple Sequence Repeats (SSRs) were identified. Many PCR based markers like RAPDs and AFLPs were usually being developed for genetic variation studies but Microsatellites with unique DNA sequences and short repeatative flanking traits can be used for polymorphism studies. The traditional techniques are quite laborious compared to the latest technology. [41, 42]. Transcripts were quantified using Salmon tool and their Transcript per million values was obtained and alternatively RSEM tool can also be used to quantify the transcripts.

Conclusion

This is the first report on *de novo* transcriptome assembly and annotation of the *J. adhatoda* and the identified genes involved in the tryptophan biosynthesis pathway related to medicinal properties. The tryptophan biosynthesis acts as a precursor for acridone alkaloid biosynthesis. These results serve as an important resource for the molecular studies on biosynthesis of various medicinally important compounds from this plant.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11033-022-07784-5.

Acknowledgements We acknowledge the SRM-DBT platform and the High-Performance Computing Cluster Facility (Genome Server) at SRM Institute of Science Technology (SRMIST). All the authors are thankful to Dr.M.Parani, Dept. of Genetic Engineering for his support and guidance through out the study.

Author contributions Senthilkumar Palanisamy conceived the experimental study, supervised the data analysis and reviewed the manuscript; Purushothaman Natarajan performed the omics box analysis; Deepthi Padmanabhan analysed the data, prepared figures and written the manuscript; Adil Lateef supported DP in RTPCR experiment. All authors contributed to the manuscript at various stages.

Funding This research did not receive any specific grant from funding agencies in the public, commercial or non-profit sectors. RNA Seq service was provided by SRM DBT Platform for Contemporary Research

in Life Sciences.

Declarations

Competing interests The authors have no relavent financial or non-financial interests to disclose.

Ethical approval This is an observational study, hence no ethical approval is required.

References

1. Wangkhem Sandhya Chanu & Kananbala Sarangthem (2014) Phytochemical constituents of *Justicia Adhatoda* Linn. Found in Manipur. Indian J Plant Sci 3(2):19–21
2. Hossain MT, Hoq MO (2016) Therapeutic use of *Adhatoda vasica*. Asian J Med Biol Res 2(2):156–163
3. Chowdhury II, Atiar Rahman Md, Mohammad A, Hashem M, Mosharef Hossain Bhuiyan D, Hajjar W, Alelwani, Arwa A, Makki et al (2020) Supplements of aqueous combination of *Justicia adhatoda* and *Ocimum tenuiflorum* boost antioxidative effects and impede hyperlipidemia. Anim Models Exp Med 3:140–115
4. Jayapriya G, Gricilda Shoba F (2015) GC-MS analysis of bioactive compounds in methanolic leaf extracts of *Justicia adhatoa* (Linn). J pharmacogn phytochem 4(1):113–117
5. Muhammad Gulfraz A, Ahmad MJ, Asad, Aliyasadiq U, Afzal M, Imran P, Anwar A, Zeenat KS, Abbasi S, Maqsood, Rahmat Ullah Qureshi (2011) Antidiabetic activities of leaves and root extracts of *Justicia adhatoda* Linn against alloxan induced diabetes in rats. Afr J Biotechnol 10(32):6101–6106
6. Alka Sharma & Ashwani Kumar (2016) Pharmacognostic studies on medicinal plants. Justicia adhatoda World J Pharm Res 5(7):1674–1704
7. Jayapriya G, Gricilda Shoba F (2015) GC-MS analysis of bioactive compounds in methanolic leaf extracts of *Justicia adhatoa* (Linn). J pharmacogn phytochem 4(1):113–117
8. Subha Bose Banerjee & Kuntal Gupta (2021) *Justicia adhatoda* having multifunctional pharmacological properties can prevent COVID-19 disease. World J pharm med 7(9):34–38
9. Ercan S, Unlu, OzgeKaya IE, Ekrem Gurel (2021) Sequencing, *de novo* assembly and annotation of *Digitalis ferruginea subsp. Schischkinii* transcriptome. Mol Biol Rep 48:127–137
10. Shikha Kalra BL, Puniya D, Kulshreshtha S, Kumar J, Kaur S, Ramachandran, Kashmir Singh (2013) & *De novo* Transcriptome sequencing reveals important molecular networks and metabolic pathways of the plant, *Chlorophytum borivilianum*.PLoS One8(12)
11. Adil Lateef, Sudheesh K, Prabhudas & Purushothaman Natarajan (2018) RNA sequencing and *de novo* assembly of *Solanum trilobatum* leaf transcriptome to identify putative transcripts for major metabolic pathways.Sci. Rep.8
12. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMB net Journal 17(1):10–12
13. Martin JA, Wang Z (2011) Next-generation transcriptome assembly. Nat Rev Genet 12:671–682
14. Joshi NA, Fass JN (2011) Sickle: A sliding-window, adaptive. quality-based trimming tool for FastQ files
15. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA et al (2011) Full length trasnscriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644–652
16. Li W, Godzik A (2006) CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659
17. Proost S et al (2009) PLAZA: a comparative genomics resource to study gene and genome evolution in plants. Plant Cell 21(12):3718–3731
18. Van Bel M et al (2013) TRAPID: an efficient online tool for the functional and comparative analysis of *de novo* RNA-Seq transcriptomes. Genome Biol 14(12):134
19. Ana Conesa & Stefan Gotz (2008) Blast2GO: A Comprehensive suite for functional Analysis in Plant Genomics. Int J Plant Genomics.12.
20. Rob Patro G, Duggal MI, Love RAI, Carl Kingsford (2017) Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14:417–419
21. Rao X, Huang X, Zhou Z, Lin X (2013) An improvement of the 2 (−delta delta CT) methods for quantitative real-time polymerase chain reaction data analysis. Biostat Bioinforma Biomath 3:71–85
22. Jiayi S, Manmathan H, Cheng S, Chritie AM (2016) Examining the transcriptional response of overexpressing anthranilate synthase in the hairy roots of an important medicinal plant *Cathranthus roseus* by RNA-seq. BMC Plant Biol 6(1):108
23. Hughes EH, Hong SB, Gibson SI, Shanks JV, San KY (2004) Metabolite Engineering of the indole pathway in *Catharanthus roseus* hairy roots and increased accumulation of tryptamine and serpentine. Metabolite Eng 6(4):268–276
24. Peebles CA, Hong SB, Gibson SI, Shanks JV, San KY (2005) Transient effects of overexpressing anthranilate synthase alpha and beta subunits in *Catharanthus roseus* hairy roots. Biotechnol Programme 21(5):1572–1576
25. Schaller A, Vanafferden M, Windhofer V, Bulow S, Abel G, Schmid J et al (1991) Purification and characterization of chorismate synthase from *Eugelena gracius*-comparision with chorismate synthases of plant and microbial origin. Plant Physiol 97:1271–1279
26. Maeda H, Dudareva N (2012) the shikimate pathway and aromatic amino acid biosynthesis in plants. Annu Rev Plant Biol 63:73–105
27. Cho MH, Corea ORA, Yang H, Bedgar DL, Laskar DD, Anterola AM et al (2007) Phenylalanine biosynthesis in *Arabidropsis thaliana*- identification and characterization of arogenate dehydratases. J Biol Chem 282:30827–30835
28. Huang T, Tohge T, Lytovchenko A, Fernie AR, Jander G (2010) Pleiotropic physiological consequences of feedback insensitive phenylalanine biosynthesis in *Arabidopsis thaliana*. Plant J 63:823–835
29. Takayuki Tohge M, Watanbe R, Hoefgen, Alisdair R Fernie(2013) Shikimate and phenylalanine biosynthesis in the green lineage. Front. Plant Sci12
30. Ishihara A, Matsuda F, Miyagawa H, Wakasa K (2007) Metabolomics for metabolically manipulated plants: effects of tryptophan overproduction. Metabolomics 3:319–334
31. Bharat Singh, Pooran M, Sahu, Ram A, Sharma (2017) Effect of elicitors on the production of pyrroloquinazoline alkaloids by stimulating anthranilate synthase activity in *Adhatoda vasica* Nees cell cultures. Planta 246:1125–1137
32. Kast P, Grisostomic C, Chen IA, Li SL, Krengel U, Xue YF(2000) A strategically positioned cation is crucial for efficient catalysis by chorismate mutase.Journal of Biology Chemistry.275
33. Khanapur M, Alvala M, Prabhakar M, Kumar KS, Edwin RK, Saranya (2017) *Mycobacterium tuberculosis* chorismate mutase: A potential target for TB. Bioorg Med Chem 25:1725–1736
34. Monti SM, De Simone G, D'ambrosio K (2016) L-histidinol dehydrogenase as a new target for old diseases. Curr Top Med Chem 16:2369–2378
35. Rajkumar Veligeti RB, Madhu J, Anireddy VR, Pasupuleti VKumarR, Aluva et al (2020) Synthesis of novel cytotoxic tetracyclic acridone derivatives and study of their molecular docking,

ADMET, QSAR, bioactivity and protein binding properties. Sci Rep 10:20720
36. Heba AEEG, Farghaly AO, Mohammed AH (2018) Acridone-based acetyl-cholinesterase inhibitors; synthesis, antioxidant activity and molecular modeling. Int J Drug Deliv 10:25–35
37. Bernaridone AMR, Azevedo AR, Pinheiro LCS, Rodrigues CR, Castro HC, Frugulhetti ICPP, Albuquerque MG (2008) SAR series of anti-HSV-1-acridone derivatives and a rational acridone-based design of a new anti-HSV-1 3H-benzo[b]pyrazolo[3,4-h]-1,6-napthyridine series. Bioorg Med Chem 16:313–321
38. Goodell JR, Madhok AA, Hiasa H, Ferguson DM (2006) Synthesis and evaluation of acridine-and acridone-based anti-hepes agents with topisomerase activity. Bioorg Med Chem 14:5467–5480
39. Michael JP (2008) Quinoline, quinazoline and acridone alkaloids. Nat Prod Rep 25:166–187
40. Putic A, Stecher L, Prinz H, Muller K (2010) Structural-activity relationship studies of acridones as potential antisporatic agents.2.Synthesis and antiproliferative activity of 10-substituted hydroxyl-10H-acridin-9-ones against human keratinocyte growth. Eur J Med Chem 45:5345–5352
41. Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Amr J Hum Genet 44:397–401
42. Smith DN, Devey ME (1994) Occurrence and inheritance of microsatellites in *Pinus radia*. Genome 37:977–983