3 Biotech (2020) 10:271
https://doi.org/10.1007/s13205-020-02263-9

ORIGINAL ARTICLE

De novo transcriptome analysis unravels tissue-specific expression of candidate genes involved in major secondary metabolite biosynthetic pathways of *Plumbago zeylanica:* implication for pharmacological potential

Balachandran Karpaga Raja Sundari[1] · Roli Budhwar[2] · Bilikere S. Dwarakanath[1,3] · S. P. Thyagarajan[1]

Received: 30 June 2019 / Accepted: 16 May 2020 / Published online: 29 May 2020

Abstract

Key message **The present study provides comparative transcriptome analysis, besides identifying functional secondary metabolite genes of *Plumbago zeylanica* with pharmacological potential for future functional genomics, and metabolomic engineering of secondary metabolites from this plant towards diversified biomedical applications.**

Abstract *Plumbago zeylanica* is a widely used medicinal plant of the traditional Indian system of medicine with wide pharmacological potential to treat several disorders. The present study aimed to carry out comparative transcriptome analysis in leaf and root tissue of *P. zeylanica* using Illumina paired end sequencing to identify tissue-specific functional genes involved in the biosynthesis of secondary metabolites, contributing to its therapeutic efficacy. De novo sequencing assembly resulted in the identification of 62,321 "Unigenes" transcripts with an average size of 1325 bp. Functional annotation using BLAST2GO resulted in the identification of 50,301 annotated transcripts (80.71%) and GO assigned to 18,814 transcripts. KEGG pathway annotation of the "Unigenes" revealed that 2465 transcripts could be assigned to 242 KEGG pathway maps wherein the number of transcripts involved in secondary metabolism was distinct in root and leaf transcriptome. Among the secondary metabolite biosynthesis pathways, the cluster of "Unigenes" encoding enzymes of 'Phenylpropanoid biosynthesis pathway' represents the largest group (84 transcripts) followed by 'Terpenoid Backbone biosynthesis' (48 transcripts). The transcript levels of the candidate unigenes encoding key enzymes of phenylpropanoid (*PAL, TAL*) and flavanoid biosynthesis (*CHS, ANS, FLS*) pathways were up-regulated in root, while the expression levels of candidate "Unigenes" transcript for monoterpenoid (*DXS, ISPF*), diterpenoid biosynthesis (*SPS, SDS*) and indole alkaloid pathways (*STR*) were significantly higher in leaf of *P. zeylanica.* Interestingly, validation of differential gene expression profile by qRT-PCR also confirmed that candidate "Unigenes" enzymes of phenylpropanoid and flavonoid biosynthesis were highly expressed in the root, while the key regulatory enzymes of terpenoid and indole alkaloid compounds were up-regulated in the leaf, suggesting that (differences in) the levels of these functional genes could be attributed to the (differential) pharmacological activity (between root and leaf) in tissues of *P. zeylanica*.

Keywords *Plumbago zeylanica* · Secondary metabolites · Candidate unigenes · Transcriptome · Tissue-specific expression · Pharmacological potential

Introduction

Plumbago zeylanica L. (family: Plumbaginaceae), commonly known as 'Chitrak', is a widely used medicinal herb known for its white flowers, also referred as "White leadwort", that are prevalent in most parts of India and the rest of Asia. The dried roots of this plant are largely used in Ayurvedic formulations of the traditional Indian system of medicine (Satyavati and Gupta 1987), for treatment of skin

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13205-020-02263-9) contains supplementary material, which is available to authorized users.

✉ Balachandran Karpaga Raja Sundari
biokarpagam@gmail.com

Extended author information available on the last page of the article

diseases, rheumatic fever, diarrhea, and leprosy (Singh et al. 2004; Mandavkar and Jalalpure 2011). Pharmacological effects of *P. zeylanica* L. include antimicrobial, anticarcinogenic and radio-modifying properties, anti-fertility, hepatoprotective, central nervous system stimulatory activity, anti-inflammatory, anti-hyperglycemic, and anti-atherosclerotic activity (Singh et al. 2011; Chauhan 2014; Roy et al. 2017). Of all the parts of *P. zeylanica*, the root extract has been credited with tremendous pharmacological potential including anticancer (Ganesan and Gani 2013; Singh et al. 2016), imparting anti-diabetic and anti-hyperlipidemic activities (Sunil et al. 2012) and affording protection against Parkinson disease (Ittiyavirah and Ruby 2014). On the contrary, therapeutic efficacy in the leaf extract of *P. zeylanica* includes, anti-inflammatory (Vishnukanta and Rana 2008), anticancer (Hiradeve et al. 2011), anti-fertility, and antioxidant properties (Nile and Khobragade 2010).

Earlier studies on phytochemical analysis in *P. zeylanica* revealed various bioactive compounds including, alkaloids, flavonoids, naphthoquinones, (Plumbagin, chitranone, chloroplumbagin, elliptone), glycosides (Plumbagic acid glycosides), tri-terpenoids, coumarins (seselin, 5-methoxy seselin, xanthyletin, suberosin) and phenolic compounds (Roy and Bharadvaja 2017). Plumbagin, a bio-active naphthoquinone (5-hydroxy-2-methylnaphthalene-1,4-dione), is the potential compound in *P. zeylanica* that is attributed with wide spectrum of pharmacological properties, especially due to its anticancer potential (Chauhan 2014; Kawiak et al. 2017; Rondeau et al. 2018). Nevertheless, earlier phytochemical studies reported various secondary metabolite compounds in *P. zeylanica* (Roy and Bharadvaja 2017), while pharmacological properties found in its leaf and root extracts have not been analyzed to understand its therapeutic potential. Our preliminary studies with a 2-year-old *P. zeylanica* plant showed significant differences in the amount of plumbagin (1.3% w/w) in the root extracts as compared to stem (0.2% w/w) and leaf (0.00007% w/w) using HPTLC analysis (Sundari et al. 2017). Indeed, the marked antioxidant and cytotoxic potential observed in the leaf extract, with trace amount of plumbagin, suggested that leaf could contain other active phyto-constituents. Initial transcriptome analysis in the root of *P. zeylanica* provided only transcripts of a few secondary metabolite pathways, but did not bring out tissue-specific bioactive candidate genes in the absence of comparative transcriptome data (Sundari et al. 2017). Therefore, the present study was undertaken to carry out a comparative transcriptome analysis in root and leaf that could facilitate deciphering of bioactive novel genes involved in tissue-specific biosynthesis of secondary metabolites in *P. zeylanica.*

Transcriptomics approach involving next-generation sequencing (NGS) technology is predominantly used in medicinal plants for unraveling novel gene transcripts associated with the biosynthesis of secondary metabolites and to identify their expression profiles (Upadhyay et al. 2014). In the present study, Illumina NGS platform was used for RNA sequencing in leaf and root tissues of *P. zeylanica,* to identify and profile transcripts involved in secondary metabolite biosynthesis. The transcriptome assembly and functional annotation of *P. zeylanica* transcriptome resulted in the distinct identification of putative genes associated with the biosynthesis of phenylpropanoid, flavonoid, terpenoid, and indole alkaloid, both in leaf and root tissues. Further, this study also focused on differential expression profiling of the metabolic pathway genes to identify the candidate regulatory gene transcripts responsible for the biosynthesis of tissue-specific secondary metabolites, contributing to its specific pharmacological activity in the leaf and root of *P. zeylanica*.

Results

The present study focused on transcriptome sequencing only in the root and leaf tissues of *P. zeylanica* since both these tissue parts have been reported to exhibit tremendous pharmacological potential (Jain et al. 2014; Sundari et al. 2017). The type of secondary metabolite compounds synthesized could vary with the extraction process and tissue parts of medicinal plants, leading to huge differences in their pharmacological potential (Anandan et al. 2012). Hence, RNA sequencing in leaf and root tissue of *P. zeylanica* resulted in the expression profiling of functional regulatory genes in important metabolic pathways associated with the biosynthesis of tissue-specific secondary metabolites.

Transcriptome data and assembly

An average of ~ 19.58 million reads was obtained per cDNA library generating 96% reads with Phred quality score of ≥ Q20 analyzed using NGS QC Toolkit, while 93% of reads were of high quality (≥ Q30). Further, the cDNA libraries had an average of 45% GC content in both the leaf and root transcriptomes. Following RNA sequencing of both the tissues, reads having ≥ 70% of the bases with a quality score of ≥ Q30 were chosen to meet high stringency criteria for better assembly of the transcriptome. Transcriptome assembly was performed by transameliorate pipeline comprising of Trinity assembler and CD-HIT EST clustering resulted in an average of 70,000 transcripts in both root and leaf transcriptomes individually (Table 1). The total transcriptome accounted for an average of 68 Mb with a minimum transcript length of 100 bp and maximum of ~ 8000 bp, with N50 value of 1.3 kb, indicating an optimized unfragmented transcriptome assembly of the cDNA libraries. To have a single representative transcriptome for

Table 1 Statistical summary of leaf and root transcriptome assembly

Parameter	Values	Values
Sample	Leaf	Root
Number of final transcripts	71,878	101,290
Final transcriptome length (bp)	65,521,679 (~65 Mbp)	71,519,022 (~71 Mbp)
Minimum transcript length(bp)	201	201
Maximum transcript length(bp)	7600	8314
Average transcript length(bp)	911.57	706.08
N50 value	1540	1154
(G + C)%	43.47%	46.35%

both, the transcripts from their assemblies were merged and CD-HIT EST clustering was performed. This resulted in a total of 62,323 transcripts accounting for ~ 82 Mb, without much change in the minimum and maximum transcript length, and N50 value (Suppl. Fig. S2). The authenticity of the merged transcript assembly was validated by aligning the reads from the replicate cDNA libraries of both the tissues to all 62,323 transcripts using Bowtie 2. The transcript length distribution analysis of root and leaf libraries showed ~ 70% of the transcripts having an average of 1 Kb length (Suppl. Fig. S2). The raw data reads obtained from RNA sequencing in both leaf and root of *P. zeylanica* were submitted to the NCBI SRA database and an accession number was received (SRA accession number: SRP144988).

Functional annotation and biological classification

The merged transcriptome of leaf and root contained 62, 323 clustered transcripts, of which 50,253 (80.5%) transcripts could be annotated to at least one database. BLAST2GO analysis (*e* value1e −05) against Plant NR database revealed that 50,301 (80.71%) transcripts showed significant homology (< 60%) against the reference sequences, but a few of them, 5072 (10%), were hypothetical proteins and 10,556 (21%) were predicted proteins, which might be due to the lack of reference of *P. zeylanica* genome and protein sequence information for annotation. BLASTX similarity search for *P. zeylanica* transcripts showed highly significant homology with some of the known functional proteins in other plant species, including 60–70% match with glycosyl transferases (*Populus trichocarpa* and *Vitis vinifera),* 96% identity with NADPH cytochrome P450 reductase 2 family (*Arabidopsis thaliana)*, and 100% identity with NAD(P)H-quinone oxidoreductase subunit H (*Plumbago auriculata*). Earlier studies have indeed found that glycosyl transferases, cytochrome P450 reductase family, and NAD(P)H-quinone oxidoreductase subunit H have a crucial role in plant secondary metabolism including flavonoid biosynthesis, defense mechanism and further involved in the detoxification and bioactivation of certain quinones (Bowles et al. 2005; Zhang et al. 2018), suggesting that *P. zeylanica* transcripts share huge similarity with genes involved in the biosynthesis of secondary metabolites of related plant species.

GO annotation

Gene Ontology (GO) assignment is a widely used standard to classify the function of unigene transcripts based on sequence homology. GO annotation assigned for 18,814 transcripts were functionally categorized, based on their sequence homology, into three GO classification including biological process (BP), molecular function (MF), and cellular component (CC). The top 10 GO annotation classifications under three functional categories are represented in Fig. 1a. Under these functional classification ontology, ATP binding (GO: 0005524, 7.66%), metal ion binding (GO: 0046872, 3.93%), and nucleotide binding (GO: 0000166, 3.67%) were highly represented in the molecular function. Under biological process ontology, oxidation–reduction process (GO: 0055114, 7.97%), protein phosphorylation (GO: 0006468, 6.06%), and regulation of transcription (GO: 0006355, 5.37%) were the most enriched categories. However, among the cellular component category of GO, sequences related to the integral component of membrane (GO: 0016021, 16.64%), nucleus (GO: 0005634, 3.98%), and cytoplasm (GO: 0005737, 2.95%) were found to be highly significant.

KEGG functional classification

The assembled transcripts of *P. zeylanica* annotated with GO terms were mapped to the reference KEGG pathways in KAAS database and ECs were generated (Kanehisa et al. 2017). KEGG annotation of the unigenes revealed that 2465 transcripts could be assigned to 242 KEGG pathway maps with 412 Enzyme Commissions, while 35, 205 transcripts were annotated using the PlantCyc database. Within the KEGG pathway maps, 'Metabolism' category showed the highest number of mapped entry in transcriptome of *P. zeylanica* and highly represented KEGG pathways with maximum "Unigenes" transcripts among them being thiamine metabolism (646 transcripts) followed by aminobenzoate degradation (190), starch and sucrose metabolism (84) pyrimidine and purine metabolism (67) and oxidative phosphorylation (56). These significantly represented pathways are the major essential pathways for nucleotide, carbohydrate, and energy metabolism, necessary for basic growth of all plant species, indicating the effectiveness of annotation and assembly of our *P. zeylanica* de novo transcriptome. The comparative analysis of top 10 KEGG pathway classifications identified with distinct number of transcripts in

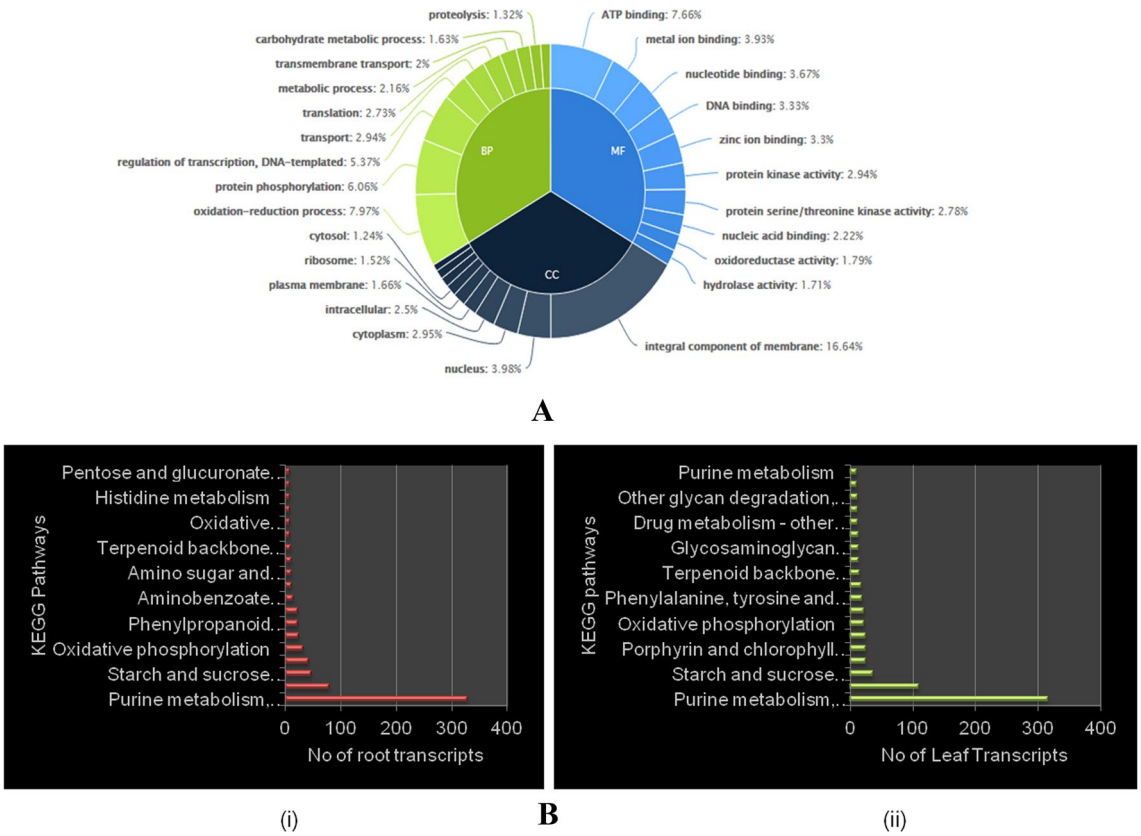

Fig. 1 **a** Functional classification of genes in *Plumbago zeylanica* based on gene ontology classification. Figure shows the top ten highly represented functional categories of unigenes grouped under three main categories of GO annotation: biological process (BP), cellular component (CC) and molecular function (MF). **b** KEGG pathway annotation for assembled unigenes in *P. zeylanica* transcriptome. The top 10 most significantly enriched KEGG pathway categories' identified in (i) leaf and (ii) root transcriptome are represented

leaf and root transcriptome are illustrated in Fig. 1b. Indeed, the highly represented pathways necessary for plant growth and metabolism were similar in both the leaf and root transcriptome of *P. zeylanica* as expected, but the number of transcripts involved in regulating secondary metabolic pathways were distinct (for example, terpenoid biosynthesis: 16 transcripts in leaf and 20 transcripts in root).

Prediction of bioactive unigenes involved in secondary metabolite biosynthesis of *P. zeylanica*

Interestingly, we could also identify a number of *P. zeylanica* unique sequences with 254 transcripts that were found to be involved in the biosynthesis of a wide range of secondary metabolites (Suppl. Table S2). Among the secondary metabolite biosynthesis pathways, the cluster of unigene encoding enzymes of 'Phenylpropanoid biosynthesis pathway' represents the largest group (84 transcripts) followed by 'Terpenoid Backbone biosynthesis' (48 transcripts). Earlier pathway analysis in medicinal plants of Polygonaceae family, one of the closest families of Plumbaginanceae, have also suggested a higher representation of phenylpropanoid biosynthesis pathway in KEGG annotation of transcripts for *Fagopyrum dibotrys* (Chen and Li 2016) and *Polygonum minus* (Rahnamaie-Tajadod et al. 2017). Out of the annotated 19 metabolic pathways, five major pathways including flavonoid, phenylpropanoid, terpenoid backbone biosynthesis, isoquinoline indole alkaloid, ubiquinone, and other terpenoid–quinone biosynthesis related to the biosynthesis of secondary metabolite compounds were well represented in our transcriptome datasets, which have already been reported as the major secondary metabolites in *P. zeylanica* (Roy and Bharadvaja 2017). The candidate unigenes in leaf and root transcriptome of *P. zeylanica* were identified from each of these secondary metabolite biosynthesis pathways through differential expression profiles.

Identification of candidate "Unigenes" involved in the phenylpropanoid biosynthesis

From the KEGG pathway and PlantCyc annotation of *P. zeylanica* transcriptome, it was found that 84 transcripts encoding unigene enzymes were assigned to the phenylpropanoid biosynthesis pathway. To decipher the key genes

involved in total phenolic compound biosynthesis, differentially expressed unigenes involved in the phenylpropanoid pathway were identified. Among these unigenes, transcripts coding for 'Lactoperoxidase' (LPO) were well represented in transcriptome leading to the phenylpropanoid biosynthesis pathway, of which 3 LPO unigenes were root-specific transcripts and 10 unigenes showed tissue-specific differential expression in *P. zeylanica*. From the digital gene expression profiling using DeSeq2, it was found that seven unigenes of lactoperoxidase [EC 1.11.1.7, LPO], gentiobiase [betaglucosidase, EC: 3.2.1.21], phenyl ammonia lyase [EC 4.3.1.24, PAL], and tyrosine ammonia lyase [EC 4.3.1.25, TAL] were all up-regulated in the root of *P. zeylanica* (Table 2) as shown in Fig. 2. On the contrary, three unigene transcripts

Table 2 Candidate "Unigenes" encoding enzymes involved in the phenylpropanoid and flavonoid biosynthesis identified from *P. zeylanica* transcriptome

Pathway	Name	Description	KO no.	EC no.	Total no. of unigenes
Phenylpropanoid biosynthesis	*LPO*	Lactoperoxidase	K12550	EC 1.11.1.7	10
	BGL	Gentiobiase (betaglucosidase)	K01188	EC: 3.2.1.21	10
	PAL	Phenylalanine ammonia lyase	K10775	EC.4.3.1.5	10
	TAL	Tyrosine ammonia lyase	K10774	EC 4.3.1.25	10
Flavanoid/polyketide biosynthesis	*FLS*	Flavonol synthase	K05278	EC.1.1.1.219	1
	CHS	Chalcone synthase	K00660	EC 2.3.1.74	9
	CHI	Chalcone isomerase	K01859	EC.5.5.1.6	9
	ANS	Anthocyanidin synthase	K05277	EC.1.14.11.19	9

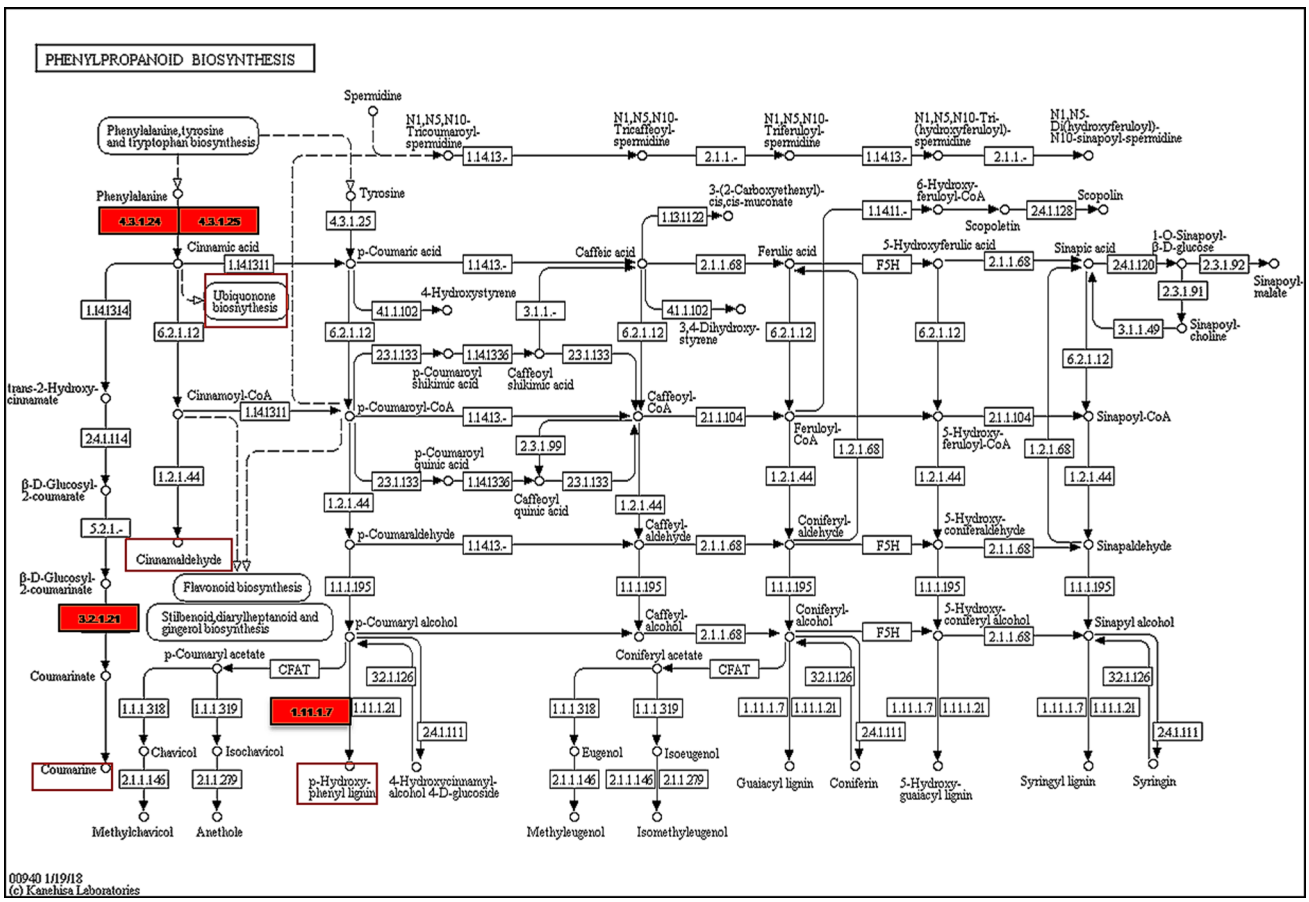

Fig. 2 KEGG pathway map showing metabolites and differentially expressed key unigenes assigned to "Phenylpropanoid biosynthesis" pathway. The fully shaded red boxes represent gene transcripts with high expression in roots of *P. zeylanica*. The red colored column also indicates the biosynthesis of specific metabolic compounds at high level. This color-coded map corresponds to map00940 in the KEGG database

of all these enzymes were found to be downregulated in the leaf of *P. zeylanica*. PAL [EC 4.3.1.24] and TAL [EC 4.3.1.25] are the two major key enzymes of the the phenylpropanoid pathway, wherein the upregulation of both PAL and TAL are known to enhance the biosynthesis of cinnamic acid along with its derivatives and coumarins (Shivashankar et al. 2015), which are known to stimulate insulin secretion necessary for exerting anti-diabetic potential (Adisakwattana 2017). Earlier reports have shown that the enzyme LPO plays an important role in the biosynthesis of lignin-derived compounds, while high expression of gentiobiase [EC: 3.2.1.21] could result in enhanced content of coumarin (Rahnamaie-Tajadod et al. 2017) that leads to scavenging of the reactive oxygen species (ROS) necessary for antioxidant and anticancer activities (Singh et al. 2016; Liu et al. 2018).

Identification of candidate unigenes involved in the flavonoid biosynthesis

Flavonoid biosynthesis is a vital part of secondary metabolisms synthesized through the phenylpropanoid pathway and is responsible for the cytotoxicity and antioxidant potential through scavenging of free radicals (Ferreyra et al. 2012). From the KEGG and PlantCyc pathway annotation, unigene transcripts encoding the key enzymes involved in the flavonoid biosynthesis pathway (map00941) could be identified in *P. zeylanica* transcriptome (Table 2). Pathway annotation suggested that one of the unigenes encoding a key enzyme dihydroflavonol-4-reductase (DFR, EC 1.1.1.219) was found to be a root-specific transcript, which is known to play a role in the biosynthesis of anthocyanidins (Ferreyra et al. 2012). Further, differential gene expression analysis using FPKM revealed that the expression of five of the key unigene encoding enzymes including chalcone synthase (CHS, EC 2.3.1.74), chalcone isomerase (CHI, EC 5.5.1.6), anthocyanidin synthase (ANS, EC: 1.14.11.19), flavonol synthase (FLS, EC.1.14.11.23), and dihydroflavonol-4-Reductase (DFR, EC 1.1.1.219) were altered in the flavonoid biosynthesis pathway as shown in Fig. 3.

CHS is a key enzyme of the flavonoid/isoflavonoid biosynthesis pathway, of which nine unigene transcripts mapped to flavonoid biosynthesis pathway, wherein three transcripts were expressed at high level in root (threefold). The validation of DEGs using qRT-PCR also confirmed that CHS is highly up-regulated in root (14-fold), while it was found to be downregulated in leaf (> -onefold). Indeed, Chalcone synthases have been shown to be the Plant type III polyketide synthase (PKSs) for producing polyketides essential for plant metabolism and defense (Dao et al. 2011; Stewart et al. 2013). Polyketides and their derivatives are also reported as vital compounds for human health, acting as antifungal agents, immunosuppressants, insecticides, anticancer drugs, and antibiotics (Ferreyra et al. 2012; Jiang et al. 2016). Similarly, comparative transcriptome analysis in *Bacopa monnieri* revealed that the expression of CHS enzymes was found to be highly up-regulated in root tissue, leading to the biosynthesis of flavonoids, anthocyanins, and isoflavonoids (Jeena et al. 2017). The up-regulation of other unigene transcripts encoding key enzymes, CHI, ANS, FLS and DFR were also found in the root as compared to leaf of *P. zeylanica*, which have been reported as the key regulatory enzymes in flavonoid biosynthesis responsible for the formation of flavanones, flavanols and anthocyanidins, and are known to be responsible for hepato-protective, anti-inflammatory and anticancer activities (Ferreyra et al. 2012; Kumar and Pandey 2013).

Prediction of the key unigene transcripts involved in terpenoid backbone biosynthesis pathway

Terpenoids, also known as isoprenoids, are synthesized through two biosynthetic pathways, including the mevalonate pathway [MD:M00095] leading to the synthesis of the precursors of sesquiterpenoids (C15) and the non-mevalonate or the MEP/DOXP pathway [MD:M00096] resulting in the synthesis of monoterpenoids (C10) and diterpenoids (C20) (Tholl 2015). KEGG and PlantCyc annotation mapped a total of 55 transcripts encoding five key differentially expressed unigene enzymes related to the terpenoid backbone biosynthesis pathway (Table 3). Two of the key enzymes, 1-deoxy-D-xylulose-5-phosphate synthase (EC:2.2.1.7, DXPS) and 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC:4.6.1.12, ISPF) were found to be involved in the monoterpenoid biosynthesis, while two of the pivotal enzymes all-trans-nonaprenyl-diphosphate synthase [geranyl-diphosphate specific] (E.C. 2.5.1.84, SPS) and all-trans-nonaprenyl-diphosphate synthase [geranylgeranyl-diphosphate specific] (E.C. 2.5.1.85, SDS) were associated with the diterpenoid biosynthesis pathway (Tholl 2006). On the other hand, 3-hydroxy-3-methylglutaryl-CoA reductase (EC 1.1.1.34: HMGCR) was found to regulate the sesquiterpenoid and triterpenoid biosynthesis pathway.

The differential expression profiles of all the four unigene transcripts including DXS, (EC:2.2.1.7), ISPF (EC:4.6.1.12), SPS (E.C. 2.5.1.84) and SDS (E.C. 2.5.1.85) were found to be highly induced in leaf of *P. zeylanica* (up to threefold) as shown in Fig. 4. Two of these enzymes, DXS and ISPF, have already been reported as regulatory enzymes for the rate limiting step of non-mevalonate pathway leading to the biosynthesis of monoterpenoid compounds that are responsible for anticancer effect in *Curcuma longa* (Annadurai et al. 2013). Moreover, the other two unigenes, SPS and SDS, were known to be involved in diterpenoid biosynthesis that perhaps cause inhibition of cell proliferation and tumor growth in different types of human cancers (Yang and Dou 2010). These results were concurrent with the abundant

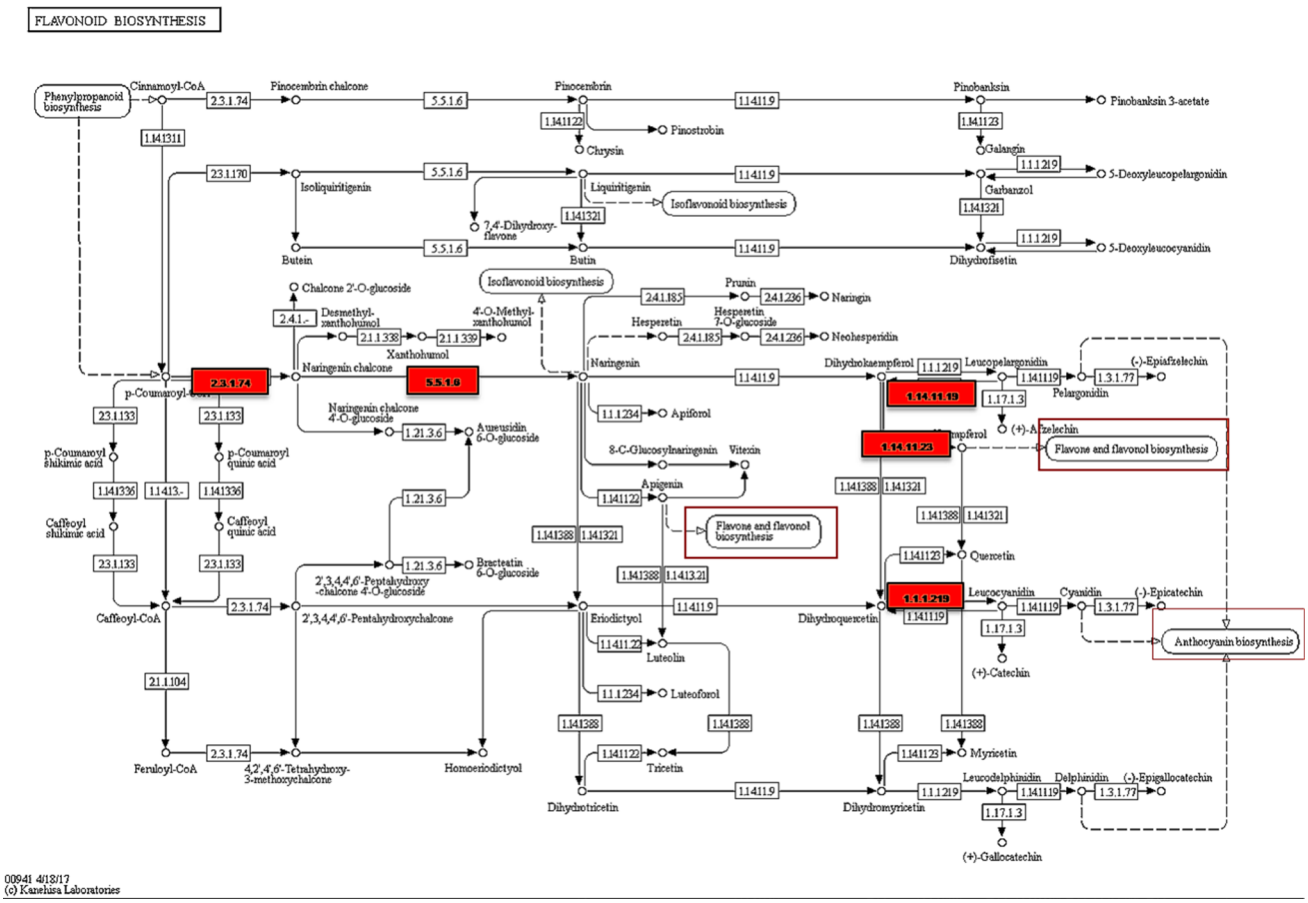

Fig. 3 KEGG pathway map showing metabolites and differentially expressed key genes identified in the "Flavonoid biosynthesis" pathway. The fully shaded red boxes represent gene transcripts with high expression in roots of *P. zeylanica*. The red-colored column also indicates the biosynthesis of specific metabolic compounds at high level. This color-coded map of DEGs corresponds to map00941 in the KEGG database

Table 3 "Unigenes" encoding enzymes involved in terpenoid and indole alkaloid biosynthesis in *P. zeylanica*

Pathway	Name	Description	KO no.	EC no.	Total no of unigenes
Mono and diterpenoid biosynthesis pathway	DXPS	1-Deoxy-d-xylulose-5-phosphate synthase	K01662	E.C.2.2.1.7	6
	ISPF	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	K01770	E.C.4.6.1.12	1
Ubiquinone and terpenoid quinone biosynthesis pathway	SPS	All-trans-nonaprenyl-diphosphate synthase [geranyl-diphosphate specific]	K05356	E.C. 2.5.1.84	1
	SDS	All-trans-nonaprenyl-diphosphate synthase [geranylgeranyl-diphosphate specific]	K05356	E.C. 2.5.1.85	1
	TAT	Tyrosine aminotransferase	K00815	E.C.2.6.1.5	9
	HPD	4-Hydroxyphenylpyruvate decarboxylase	K00457	E.C. 1.13.11.27	9
Triterpene and sesquiterpenoid compounds	HMGCR	3-Hydroxy-3-methylglutaryl-CoA reductase	K00021	EC 1.1.1.34:	2
Indole alkaloid biosynthesis	STR	Strictosidine synthase	K01757	E.C. 4.3.3.2	6

expression of these similar unigene transcripts in the MEP pathway [DXS,SPS,ISPF] related to the biosynthesis of terpenoid compounds observed in leaf transcriptome of *Salvia miltiorrhiza* (Yang et al. 2013a, b). On the contrary, 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR), which catalyzes major key steps of the mevalonate pathways,

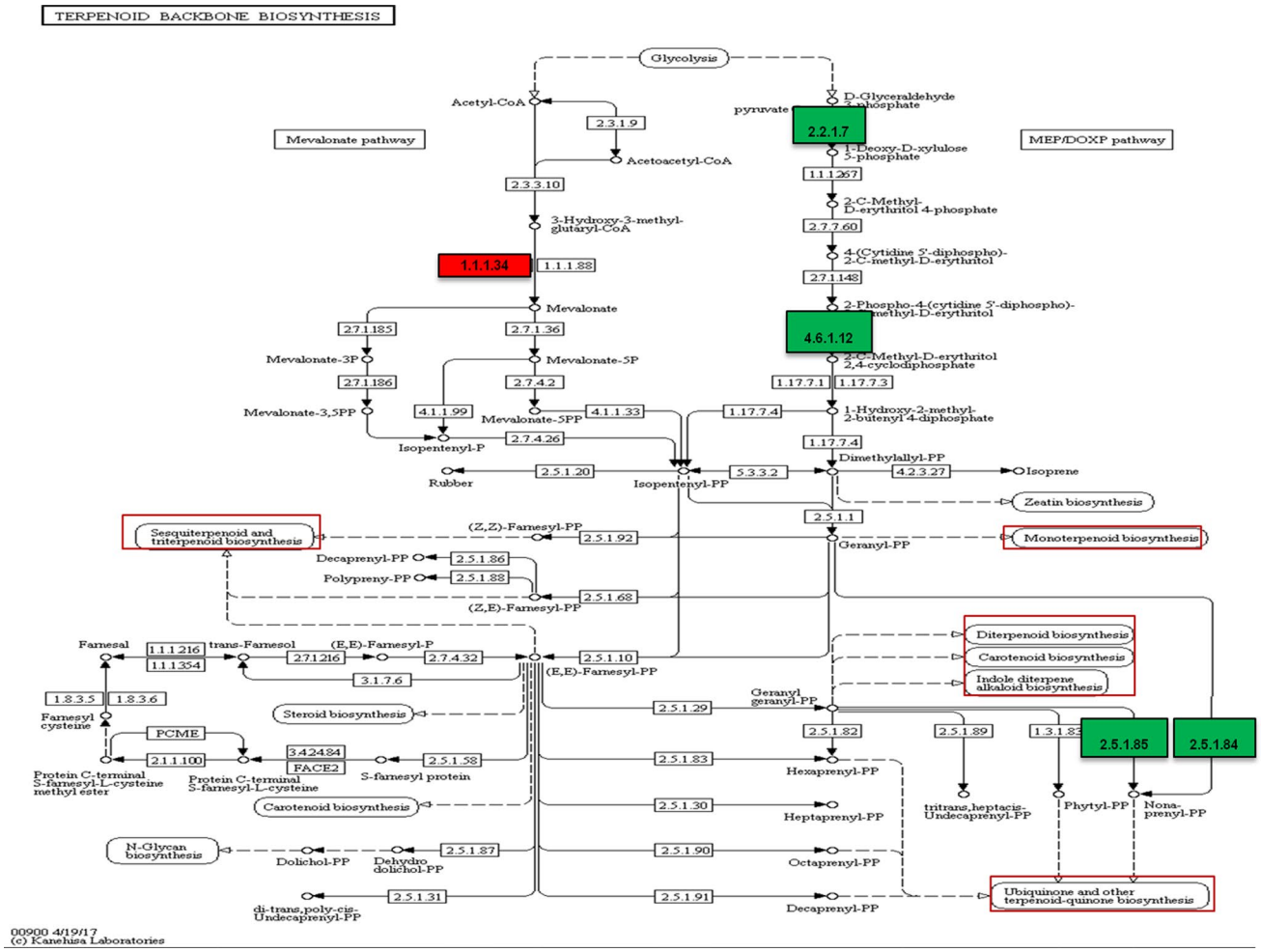

Fig. 4 KEGG pathway map showing metabolites and differentially expressed key genes identified in the "Terpenoid Backbone Biosynthesis" pathway. The fully shaded red boxes represent gene transcripts with high expression in roots of *P. zeylanica*. The green-colored boxes indicate transcripts with relatively higher expression in leaf of *P. zeylanica.* The red-colored column also indicates the biosynthesis of specific metabolic compounds at high level. This color-coded map corresponds to map00900 in the KEGG database

was found to be up-regulated in the root of *P. zeylanica* with two unigene transcripts (up to threefold). Similar results of relatively high expression of HMGR genes in the MVA pathway were found to be active in the root of *Salvia miltiorrhiza* (Yang et al. 2013a, b). Moreover, HMGCR has been reported as the crucial regulatory enzyme for triterpene and sesquiterpenoid biosynthesis in *Glycyrrhiza uralensis* (Liu et al. 2014), which imparts anti-inflammatory, antimicrobial, antitumoral, and immunomodulator effects on the immune system of target cells (Rios 2010).

Prediction of the key unigene transcripts involved in ubiquinone and other terpenoid quinone biosynthesis pathway

Plumbagin, a natural naphthoquinone, is identified as an active constituent in the plant families of Plumbaginaceae, and is responsible for a wide range of pharmacological activities (Carson and Hammer 2010; Savoia 2012). From PlantCyc annotation, we could identify nine transcripts encoding unigene enzymes involved in the ubiquinone and other terpenoid quinone biosynthesis pathways, of which two key enzymes encoding genes, TAT and HPD, showed differential expression (up to 3.5-fold) in the root of *P. zeylanica* (Fig. 5). Tyrosine aminotransferase [EC:2.6.1.5, TAT], an essential enzyme for the biosynthesis of tyrosine and 4-hydroxyphenylpyruvate decarboxylase (Ec 1.13.11.27 HPD), another key enzyme involved in the formation of homogentisic acid (Ndikuryayo et al. 2017) were highly expressed in the root with one unigene transcript. Earlier studies have shown that TAT and HPD enzymes are key regulatory enzymes involved in the biosynthesis of naphthoquinones (Widhalm and Rhodes 2016), and thus play a key role in the production of antioxidants and also exert

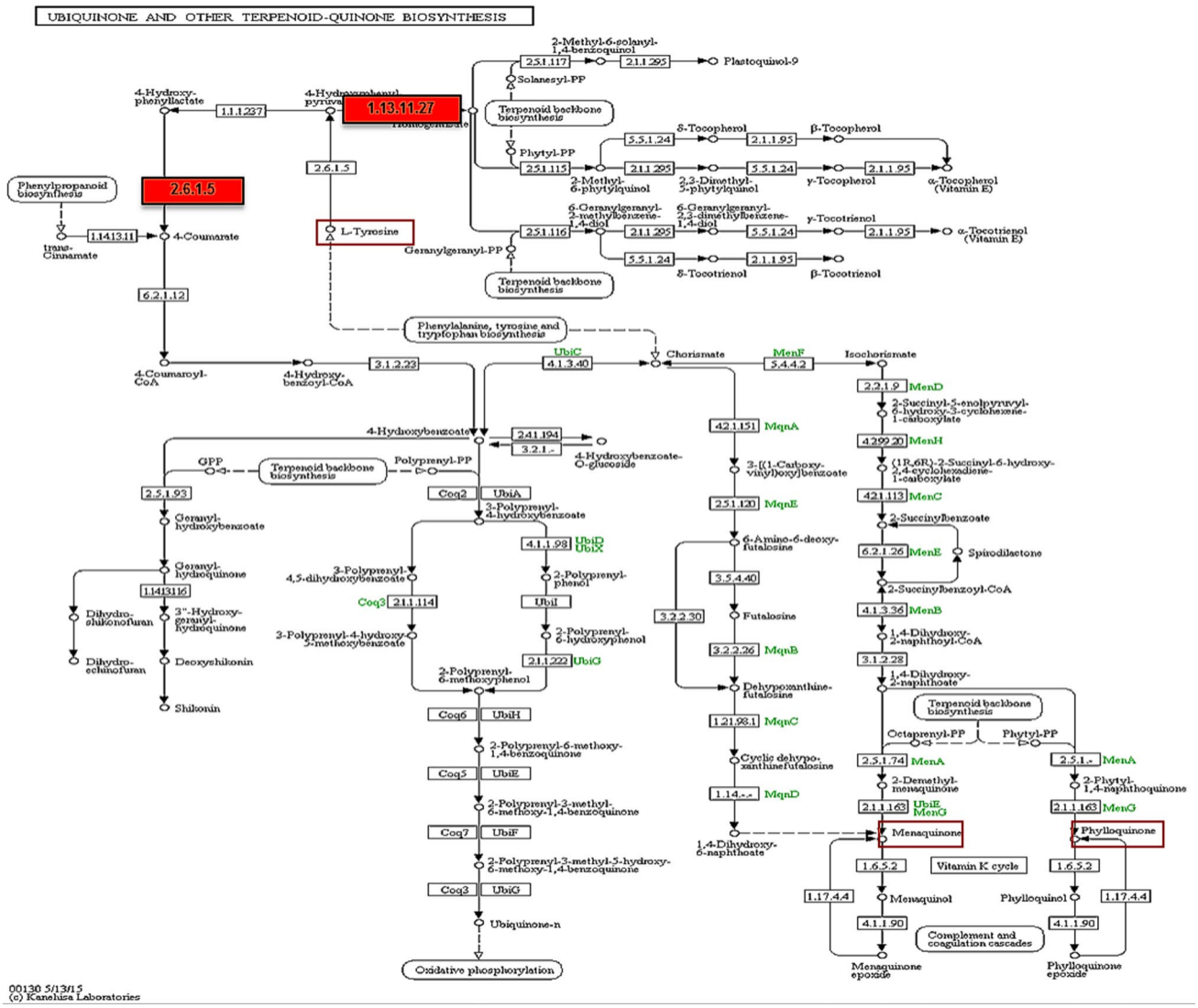

Fig. 5 Differentially expressed unigenes (DEGs) identified in the "Ubiquinone and other Terpenoid-quinone Biosynthesis" pathway. The fully shaded red boxes represent gene transcripts with high expression in roots of *P. zeylanica*. The red colored column also indicates the biosynthesis of specific metabolic compounds at high level. This color-coded map corresponds to KEGG pathway: map00130

antimicrobial, insecticidal, anti-phlogistic and wound healing properties (Raghu and Karunagaran 2014). Moreover, transcriptome analysis in *Prunella vulgaris* L has also shown a high expression of TAT encoding genes in root compared to leaves (Ru et al. 2017).

Identification of candidate "Unigenes" involved in indole alkaloid biosynthesis pathway

In various plant species, most alkaloids are synthesized through the terpenoid indole alkaloid (TIA) pathway that generates a strictosidine backbone for downstream products of their biosynthesis (Guo et al. 2013; Miralpeix et al. 2014). In the indole alkaloid biosynthesis pathway of *P. zeylanica*, we could identify strictosidine synthase (STR; EC 4.3.3.2) as one of the key enzymes encoded by "Unigenes" with six transcripts (Table 3). Interestingly, the differential expression analysis revealed that two unigene transcripts encoding strictosidine synthase were abundantly expressed in leaf of *P. zeylanica* (Suppl. Fig. S3). STR has been reported as a crucial regulatory enzyme involved in the biosynthesis of monoterpenoid indole alkaloids of *Rauvolfia serpentine* and also responsible for the production of 3-alpha-(S)-strictosidine (Stockigta et al. 2008). Indeed, the expression profile analysis of STR is concordant with its similar abundant expression observed in leaves of *Catharanthus roseus* (Zhu et al. 2015). In addition to the differential expression analysis, our qRT-PCR validation studies also showed five fold up-regulations in the expression of STR gene transcript in leaves as compared to the root. The highly induced

expression of this STR gene in leaf suggests that this may be one of the key enzymes regulating the biosynthesis of indole alkaloids and might play a role in the accumulation of alkaloids in the leaf of *P. zeylanica,* which needs to be confirmed by functional studies.

Expression level of the transcripts and differentially expressed gene (DEG) analysis

To quantify the expressed transcripts, the high-quality mapped reads were subjected to FPKM analysis for calculating fold change induction between leaf and root tissue of *P. zeylanica* (< twofold) (Suppl. Fig. S4). The differentially expressed transcripts were identified based on normalized FPKM values (*p* value cut-off ≥ 0.05). Comparative transcript abundance level revealed 7683 significant differentially expressed unigenes between the root and leaf libraries. In the root, 4198 transcripts were up-regulated (54.64% of significant unigenes) and 3485 down-regulated transcripts (45.35%), relative to their expression in the leaf (log2-fold change ≥ 2 for up-regulation and log2-fold value ≤ -2 down-regulation; $P < 0.05$) as shown in Fig. 6a. Further, we also identified some of the reads showing tissue-specific expression that were uniquely expressed either in the root or leaf. The total number of tissue-specific transcripts identified was 8670, of which 946 unigenes were leaf-specific transcripts and 7724 unigenes were exclusively expressed in root. Among the differentially expressed transcripts, gene transcripts that encode enzymes involved in chlorophyll biosynthesis, carbohydrate metabolism, and fatty acid metabolism were highly abundant in leaves and down-regulated in root of *P. zeylanica*, which substantiates the functional role of leaves in photosynthesis and energy production during plant growth, as observed in the transcriptome of *S. miltiorrhiza* (Yang et al. 2013a, b). However, the transcripts encoding enzymes for disease resistance and metabolite transport

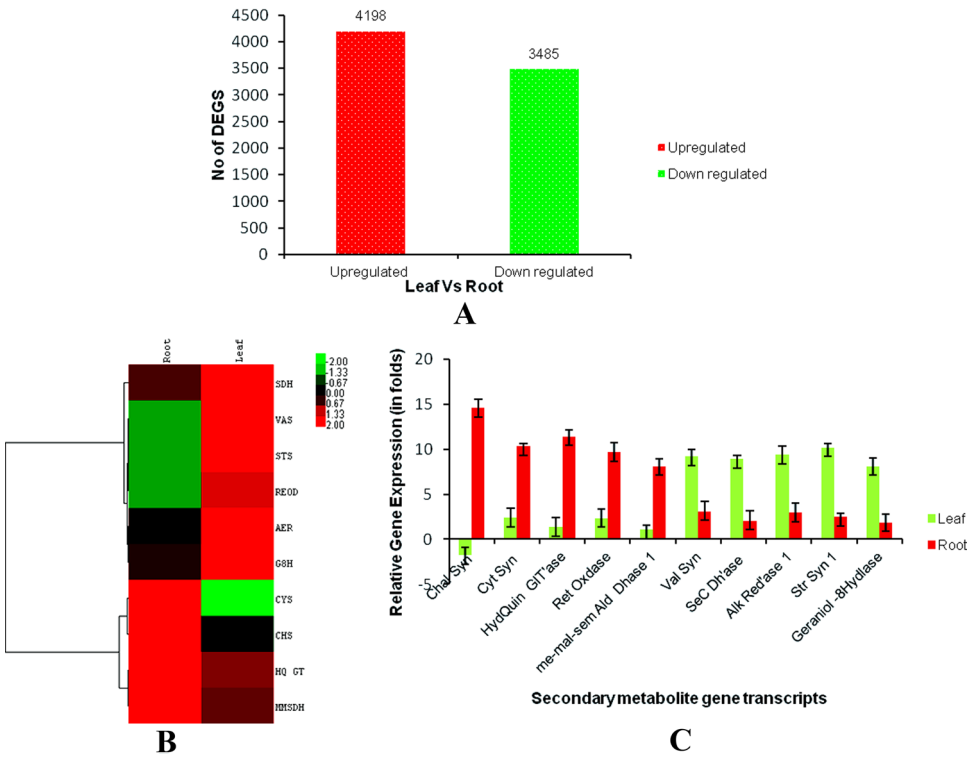

Fig. 6 **a** Comparison of differentially expressed unigenes in transcriptome of *P. zeylanica* obtained through FPKM analysis. We used a cut-off with an FPKM ≥ 1 for considering transcripts to be expressed. No of DEGs was accurately quantified based on the abundance of transcript-level sequences between leaf and root of *P. zeylanica*. Red and green bars represent the up-regulated and down-regulated unigenes in leaf compared with those in root. **b** Heat map of differentially expressed unigenes encoding enzymes of secondary metabolite biosynthetic pathways. Red and green colors represent the expression pattern of up- and down-regulated transcripts in leaf and root of *P. zeylanica*, respectively. Secoisolariciresinol dehydrogenase-like (SDH), valencene synthase (VAS), strictosidine synthase 3 (STR), reticuline oxidase (REO), 2-alkenal reductase (NADP(+)-dependent) (AER), geraniol-8-hydrolase (G8H), cytochrome P450 family synthase (CYS), chalcone synthase 1 (CHS), methylmalonate-semialdehyde dehydrogenase family (MMSDH) and hydroquinone glucosyltransferase (HQGT). **c** Validation of differentially expressed unigenes by qRT-PCR. Top 10 significant genes involved in secondary metabolism biosynthesis pathway were validated by q-PCR. The relative gene expression levels are represented as log2-fold change with reference to internal control Actin gene. Mean values and standard errors (bars) were obtained from three independent experiments for each gene transcript

proteins were highly expressed in root, indicating their significant role in root developmental function. Moreover, we found that 54 differentially expressed unigenes in the leaf and roots are known to be involved in secondary metabolism from GO annotation and BLASTx analysis. Some of the unigene transcripts with significant FPKM values belonging to the major secondary metabolic pathway, which showed significant tissue-specific expression (Fig. 6b), were selected for validation by qRT-PCR. Interestingly, these unigenes are also reported as the highly expressed unitranscripts involved in secondary metabolite biosynthesis of *Phyllanthus amarus* leaves (Mazumdar and Chattopadhyay 2015).

Validation of selected secondary metabolite pathway gene transcripts by qRT-PCR

To ascertain whether annotated transcripts from the de novo sequence assembly were actually expressed and also to characterize their expression profiles in *P. zeylanica*, a few of the differentially expressed transcripts cognate to secondary metabolite biosynthesis (each from flavonoid, quinones, terpenoids, and indole alkaloid biosynthetic pathways) were selected for qRT-PCR analysis. The ten transcripts selected for validation include Secoisolariciresinol dehydrogenase-like (SDH), valencene synthase (VAS), strictosidine synthase 3 (STR), reticuline oxidase (REO), cytochrome P450 family, chalcone synthase 1 (CHS), methylmalonate-semialdehyde dehydrogenase family (MMSDH), hydroquinone glucosyltransferase (HQGT), 2-alkenal reductase (NADP(+) dependent) and geraniol-8-hydrolase, a cytochrome P450 monooxygenase enzyme. The relative expression of the quantified ten unigenes showed that five of the transcripts related to the secondary metabolic pathways were highly expressed in root including, chalcone synthase (14 fold), cytochrome P450 family synthase (10 fold), hydroquinone glucosyltransferase (11 fold), reticuline oxidase (9 fold) and methylmalonate-semialdehyde dehydrogenase family (8 fold), when compared to leaf of *P. zeylanica* as represented in Fig. 6c. Among these five transcripts, *CHS, HQGT, CYS,* and *MMSDH* were found to be the key enzymes involved in polyketide/flavonoid biosynthesis pathway of *Phyllanthus emblica* (Kumar et al. 2016) and *Withania somnifera* (Jadhav et al. 2012). The other five transcripts include, VAS (ninefold), precursor enzymes responsible for terpenoid indole alkaloid biosynthesis, SDH (8.8fold) involved in monoterpenoid biosynthesis pathway, 2-alkenal reductase (9.3 fold), geraniol-8-hydrolase (8 fold), and STR 3 (10 fold), key regulatory enzymes of indole alkaloid biosynthesis pathway were enriched in leaf tissue of *P. zeylanica*. Similarly, these key enzymes were found to be abundant in leaf of *Catharanthus roseus* (Zhu et al. 2015). Thus, we could observe that the mRNA levels of ten unigenes, possibly involved in secondary metabolite biosynthesis validated by qRT-PCR, demonstrated similar tissue-specific expression patterns as predicted from FPKM analysis confirming its reliability.

Table 4 Statistics of SSRs identified from *P. zeylanica* transcriptome

Assembly statistics	No. of transcripts
Total number of sequences examined	62,321
Total number of identified SSRs	31,906
Number of SSRs containing sequences	23,972
Number of sequences containing more than 1 SSR	5921
Number of SSRs present in compound formation	2542

Table 5 Distribution of SSRs in different repeat types identified in *P. zeylanica*

Unit size	Number of SSRs (%)
Mononucleotide	8627
Dinucleotide	6894
Trinucleotide	5319
Tetranucleotide	334
Pentanucleotide	103

Identification of SSR

In silico MISA tool predicted that 31,906 SSRs were found in 23,972 transcripts of *P. zeylanica*, wherein 5921 unigenes contained more than one SSR loci. SSR motif analysis revealed that various classes of repeat motifs were not distributed uniformly (Table 4). Among the SSR repeats, mononucleotide motif was highly represented (41%, 8627 transcripts), followed by the di-nucleotide repeats (28.75%, 6894), trinucleotide SSR (22.18%, 5319), tetranucleotide (1.3%, 334,) and pentanucleotides (0.42%, 103), as listed in Table 5 and represented in Fig. 7a. Among the dinucleotide repeat, $(TA)_6$ appears frequently (8%), while $(AAT)_5$ was the most abundant among trinucleotide repeats. The tetranucleotide (1.3%), pentanucleotide (0.42%), and hexanucleotide (0.6%) repeats occurred in low numbers in the transcriptome of *P. zeylanica* (Fig. 7a), indicating that they possess highly complex base pattern as observed in *Physalis alkekengi* and *P. peruviana* (Fukushima et al. 2016). Indeed, the dinucleotide repeats were highly represented within the unigene associated with enzymes encoding squalene synthase, dioxygenase and 3-hydroxy-3-methylglutaryl-coenzyme A reductase, which are all found to be associated with terpenoid biosynthesis pathway (Lu et al. 2016). However, transcripts encoding phenyl ammonia lyase and peroxidases involved in phenylpropanoid biosynthesis pathway showed the presence of mono-, di-, tri-nucleotide motif and compound SSRs.

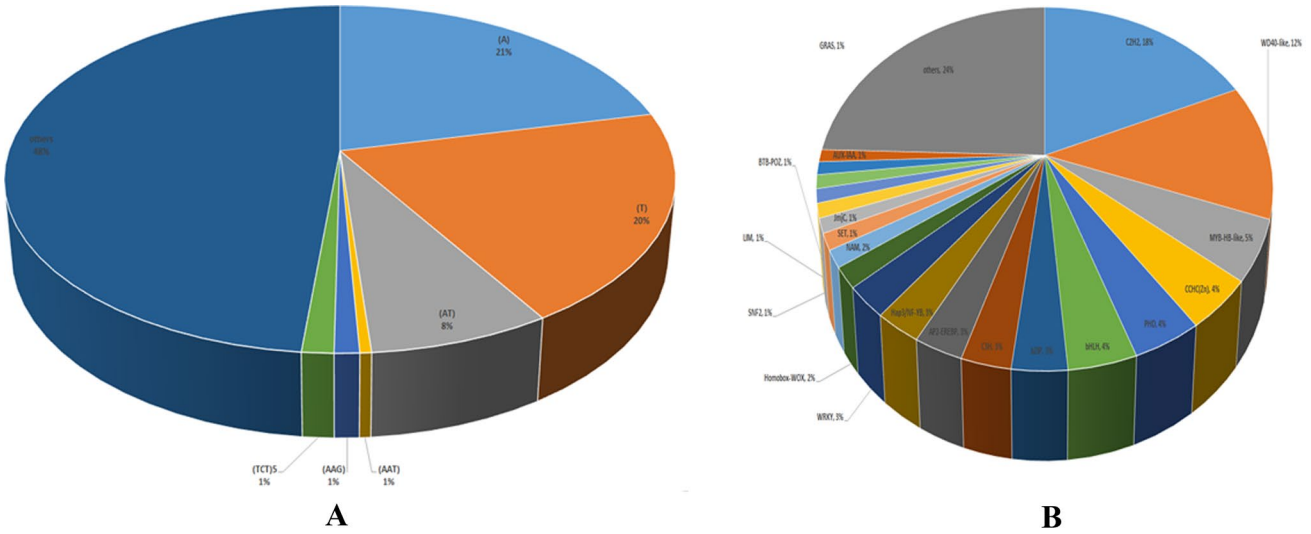

Fig. 7 **a** Frequency of most abundant SSR motifs identified in *P. zeylanica* merged transcriptome. Figure represents percentage frequency of SSRs classified into mono-, di- and tri-nucleotide repeat types of leaf and root unigene transcripts. **b** Distribution of transcription factor families in *P. zeylanica* merged transcriptome. Classes of transcription factors represented as % value that are identified from leaf and root annotated unitranscripts of *P. zeylanica*

Identification of transcription factors associated with regulation of secondary metabolite biosynthesis

From the transcriptomic data, we identified 5331 transcripts encoding putative TFs belonging to 97 TF families. Members of the C2H2 transcription factor family were the most abundant (982, 18%), followed by WD40-like (647, 12%), MYB-HB like (286, 5%), CCHC (234, 4%), and PHD (224, 4%) (Fig. 7b). Previous studies on TFs have identified many TF families, including WRKY, MYB, MYB-related, and basic helix-loop-helix (bHLH), that regulate secondary metabolism biosynthesis in plants (Liu et al. 2017). Out of these TFs, 23 unigenes encoding bHLH TFs were identified in the transcriptome of *P. zeylanica*. Results of differential gene expression analysis also predicted that all the unigenes encoding bHLH TF families were up-regulated in the root of *P. zeylanica*, of which two of the unigenes belonged to MYC-2 TFs. Earlier studies have revealed that the homologues of MYC2 TF impart an essential role in regulating the genes involved in sesquiterpene biosynthesis in *Catharanthus roseus*, *Artemisia annua*, and *Solanum lycopersicum* (Zhang et al. 2011; Ji et al. 2014). Further, our differential expression analysis confirmed that the candidate unigene transcripts encoding key enzyme, HMGCR involved in sesquiterpenoid and triterpenoid biosynthesis were also highly up-regulated in root transcriptome of *P. zeylanica*.

Discussion

Despite the wide pharmacological importance in *P. zeylanica*, only limited information is available on the molecular details and are restricted to 1842 expressed sequence tags (ESTs), as of May 2020 in NCBI GenBank for *P. zeylanica*. Studies involving transcript profiling approaches including cDNA–AFLP technique and subtractive hybridization-based specific expression subset analysis (SESA) have been employed for identifying the subset of expressed genes from active ingredient, i.e., root bark of *P. zeylanica* through root cDNA libraries, which resulted in the identification of 62 ESTs from the root of *P. zeylanica* (Shukla et al. 2013). Additionally, DNA barcoding studies using *rbcL* gene candidate have been carried out to solve intra-specific similarity and variation among three different Plumbago species, viz., *Plumbago zeylanica* Linn., *P. auriculata* Lam. and *P. rosea*, which revealed seven sequence accessions from *rbcl* region of *P. zeylanica* (Malar 2013). In recent years, functional genomics approach supported by transcriptome profiling has been widely used for elucidating bioactive genes in specific metabolic pathways of phytomedicinals (Misra 2014). This study generated ~ 82 Mbp of transcriptome, which could enable molecular studies in *P. zeylanica* and also on its related species of Plumbaginaceae including *P. indicia* and *P. rosea.* The precision of assembled sequences were validated by BLAST alignment with some of the *P. zeylanica* coding domain sequence (CDS) available at NCBI

database, which showed significant similarity of 99% to the already available mRNA sequences of *P. zeylanica*.

In plants, phenylpropanoids and flavonoids are the most essential phenolic compounds exhibiting antioxidant properties necessary for plant defense (Cuong et al. 2018) and dietary antioxidants for human health and disease (Lin et al. 2016). Accumulation of polyphenolic compounds via the phenylpropanoid pathway may be one of the mechanisms for plant defense and also for imparting pharmacological properties including, anticancer (Hemaiswarya and Doble 2013) and anti-inflammatory activity (Pastore et al. 2007). Our earlier findings had shown that the total phenolic compounds in root of *P. zeylanica* were higher (17.5% w/w) than in the leaf (2.09% w/w) and also exhibited high antioxidant and cytotoxic potential (Sundari et al. 2017). Interestingly, results of our DGE analysis suggests that these four unigenes (PAL, TAL, LPO, gentiobiase) encoding enzymes of phenylpropanoid pathway with the highest expression in root could be the key regulatory enzymes responsible for production of phenolic compounds in root, which might play an essential role in antioxidant defense. Similarly, the total flavonoid content in root extract was maximum with 45.5 ± 5.2 CE/g in comparison to leaf (2.41 ± 0.021 CE/g), which also demonstrated that the antioxidant activity is due to the flavonoids and phenolic compounds present in the root of *P. zeylanica* (Nile and Khobragade 2010; Chhetri and Khatri 2017). Interestingly, we could observe that the transcripts expressing all the key enzymes of phenylpropanoid and flavonoid pathways for the synthesis of polyphenolic compounds and flavonoids were maximum in *P. zeylanica* root from transcriptome analysis, suggesting that the accumulation of all these polyphenolic amalgam could be one of the contributing factors to its high antioxidant and cytotoxic potential. Similar observation of highly abundant unigene enzymes related to the phenylpropanoid and flavonoid biosynthesis pathways (CHI, DFR, CHS) were found in root (compared to shoot) of *Bacopa moneri* leading to the biosynthesis of flavonoids, isoflavonoids, and anthocyanins (Jeena et al. 2017).

Earlier reports on preliminary phytochemical analysis in root of *P. zeylanica* revealed the presence of steroids, flavonoids, alkaloids, quinones, phenols, tannins, and terpenoids in roots without any supportive quantification analysis (Rajakrishnan et al. 2017). From the present comparative transcriptome analysis, it appears that these unigene enzymes, TAT, HPD and HMGCR, expressed abundance in root suggesting that they might be the candidate unigenes to regulate naphthoquinone and triterpenoid biosynthesis pathway in root transcriptome of *P. zeylanica,* which needs further characterization through functional studies. Moreover, it is also reported that the plant leaves and stem of *P. zeylanica* contain higher amounts of triterpenoids (Hiradeve et al. 2011). However, candidate unigene enzymes (HMGCR) encoding triterpenoid biosynthesis were highly abundant in the root, which suggests the possibilities of microRNA role in regulation of triterpenoid biosynthesis in the root. Indeed, a crucial role of micro-RNA has been found in the regulation of tissue-specific terpenoid biosynthesis in *Withania somnifera* (Srivastava et al. 2018). Alternatively, triterpenoids might be synthesized in the root of *P. zeylanica* but transported to the leaves, which needs further investigation.

In the present study, tissue-specific gene profiling associated with monoterpenoid and indole alkaloid biosynthesis pathway of *P. zeylanica* revealed that the expression profiles of a few candidate unigene transcripts encoding DXPS, ISPF, SPS, and STR (Table 3) were highly abundant in leaf transcriptome compared to the root. Hence, our findings from the pathways analysis of the leaf transcriptome suggest that the leaf extract of *P. zeylanica* might contain high quantity of terpenoid and alkaloid compounds, which could be responsible for its pharmacological properties. This finding could be substantiated by our earlier observation of phytochemical analysis and pharmacological studies demonstrating that leaves contain less plumbagin but exhibit marked cytotoxicty and antioxidant potential (Sundari et al. 2017), which suggest that there might be high terpenoids and alkaloids in the leaf extract, contributing to its therapeutic efficacy. Further, the pathway annotation in *P. zeylanica* transcriptome also revealed some of the antimicrobial-related pathways including "polyketide sugar unit biosynthesis", "Biosynthesis of vancomycin group antibiotics", "streptomycin biosynthesis", "biosynthesis of ansamycins", "novobiocin biosynthesis" (Suppl. Table S2), wherein a few of these pathways were also reported in transcriptome of other important medicinal plant, *Phyllanthus amarus,* known for its huge pharmacological properties (Mazumdar and Chattopadhyay 2015). Hence, these pathway elucidations from leaf and root transcriptome of *P. zeylanica* could also help in subsequent engineering of key secondary metabolites for enhancing their production and reconstruction of metabolic networks.

The differential expression profile analysis validated and confirmed that most of the candidate "Unigene" transcripts related to phenylpropanoid and flavonoid biosynthesis pathways were highly expressed in root of *P. zeylanica,* whereas key unigenes belonging to terpenoid and indole alkaloid biosynthesis were abundant in leaf transcriptome of *P. zeylanica.* These results also substantiate earlier studies (Sundari et al. 2017) on the phytochemical and pharmacological potential differences in leaf and root of *P. zeylanica,* which could be attributed to their tissue-specific expression of unigene transcripts. However, the regulation of these transcripts leading to the formation of metabolites has to be investigated further though miRNA functional studies. Thus, the findings from the present study suggest that de novo assembly and annotation of *P. zeylancica* transcriptome could be further used for functional genomics analysis

and metabolic engineering studies, in future applications. Further, the nucleotide repeat distribution in SSR analysis suggest that SSRs were indeed abundantly distributed in leaf and root transcripts of *P. zeylanica*, which would enable genetic diversity studies and gene mapping for further genomics studies in *P. zeylanica* and its related species.

We have identified 146 unigenes encoding WRKY TFs, in which 23 unigene transcripts were up-regulated in the leaf, which indicates that these TFs might regulate monoterpenoid indole alkaloid biosynthesis. The role of WRKYs responsible for regulating the three major classes of plant metabolites, including phenylpropanoids, alkaloids, and terpenes, have been identified (Schluttenhofer and Yuan 2015), wherein its more specific role in regulating terpene indole alkaloid (TIA) production has also been reported in *Catharanthus* spp. (Suttipanta et al. 2011). The differential regulation of MYB transcription factors has been shown to control the genes encoding enzymes of CHS and CHI involved in the flavonoid biosynthesis in *Arabidopsis thaliana* (Stracke et al. 2007). Indeed, up-regulation of the MYB transcription factors and the unigene enzymes of CHS and CHI in the root of *P. zeylanica* substantiates that these TFs and their genes might play a crucial regulatory role in the tissue-specific expression of flavonoid biosynthesis. The prediction of large set of TFs from transcriptome of *P. zeylanica*, together with their tissue-specific expression profiles, provides a valuable resource for further characterization of specific TFs in various secondary metabolite biosynthesis pathways.

Transcriptome annotation in *P. zeylanica* provided molecular insight into the identification of tissue-specific bioactive unigene transcripts, transcription factors, and SSRs involved in various biological processes, metabolic pathways and their differential expression profiles. De novo transcriptome assembly and annotation in *P. zeylanica* generated a total of 62,323 unitranscripts. Further, tissue-specific expression profile revealed the presence of candidate unigenes involved in regulating the biosynthesis of phenolic compounds and flavonoids in the root, while terpenoids and indole alkaloids in leaf, together contributing to the pharmacological activity of *P. zeylanica*. This study provides valuable information on the functional genes involved in the biosynthesis of bioactive compounds that could lead to further metabolome engineering and functional genomics studies to facilitate phytomedicinal research in *P. zeylanica*.

Materials and methods

Plant material collection and authentication

Two-year-old *P. zeylanica* plants raised thorough clonal propagation method were collected from the medicinal plant garden of Irula Tribe Women's Welfare Society (ITWWS) Thandarai, Chengalpet, and authenticated by Prof. Jayaraman, Plant Anatomist Botanists, Plant Anatomy Research Centre, Tambaram (PARC/2017/3562). A voucher specimen was deposited at the Herbal and Indian Medicine Research Laboratory, Sri Ramachandra Medical College and Research Institute, Chennai. The leaf and root tissue samples were harvested from healthy clonal plants in the month of September 2017 and immediately snap frozen in liquid nitrogen and preserved at − 80 °C until use (Suppl. Fig. S1).

RNA sequencing and quantification

RNA was extracted from the root of *P. zeylanica* using Spectrum Plant Total RNA kit (Sigma-Aldrich, USA) according to the manufacturer's instructions and an in-house-patented nucleic acid extraction protocol of Institute of Forest Genetics and Tree Breeding, Coimbatore (Indian Patent No. 272765) was followed for RNA isolation from young leaf tissues. The quality and amount of RNA were analyzed using Nanodrop 2000 (Thermo Fisher Scientific, Wilmington, DE) and Agilent Technologies 2100 Bioanalyzer (Agilent Technologies, USA). Total RNA from three individual plants were pooled and cDNA library was prepared from 1 µg of total RNA using TruSeq RNA Sample Preparation v2 Guide (Illumina, USA). Leaf and root cDNA library samples (PL and PR) were quantified using Qubit 3.0 (Thermo Scientific, USA) and size distribution was verified using DNA 1000 chip on an Agilent Technologies 2100 Bioanalyzer. Paired end sequencing (2 × 100 bp) of PL and PR library was carried out in Bionivid Technologies, Bangalore, using the HiSeq 2000 sequencing platform (Illumina, USA). The raw (average number of ~ 30 M) reads obtained from each library were quality filtered to eliminate adapter sequences. The reads were filtered for low-quality reads using NGSQC Toolkit and SQIT tool with a phred score value of Q20 and in addition with Q30 score to obtain high-quality reads for better assembly.

Transcriptome assembly and clustering of high-quality reads

Stringent criteria of 70:30 was used to filter high-quality reads, wherein more than 70% high-quality bases (phred scores > 30) were considered for building the transcriptome assembly. De novo primary assembly of the high-quality reads was carried out in Trinity Assembler (Haas et al. 2013) using De bruijn graph-based method, while the longest transcript was considered as a uni-gene and used as reference sequence for coding sequence prediction. The final assembly was carried out by clustering the ameliorated transcripts using CD-HIT-EST (v4.6.1) to remove redundant transcripts (Bankar et al. 2015) and high-quality paired end reads were

mapped to clustered non-redundant transcripts using RSEM tool (Li and Dewey 2011).

Functional annotation and biological classification

Annotation for all the unique transcripts (> 200 bp) was done using BLASTx homology search against NCBI non-redundant (nr) protein Database (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) and pigeon pea protein database. BLAST hits with *e* value scores ≤ 0.001 and query coverage above 70% were considered as annotated homologous proteins; AWK script was used for filtering the reciprocal best hits. BLAST hits were used to obtain associated Gene Ontology (GO) terms describing biological processes, molecular functions, and cellular components. GO terms were integrated with particular gene identifiers and accession numbers. Expression levels of all the transcripts in the individual libraries in replicates were assessed by mapping high-quality (HQ) filtered reads using BOWTIE2 (Langmead and Salzberg 2012). Mapped reads were further normalized using Fragments Per Kilo base Per Millions (FPKM) method. Transcripts annotated in both tissue types were plotted as a pie-donut chart for ten major GO terms and pathways using online server (https://www.highcharts.com).

Quantification and expression profiling in *P. zeylanica* transcriptome

Differential gene expression analysis of the expressed transcripts was performed using DESeq (Love et al. 2014) software based on R programming environment. Transcripts that were two fold and above, which were up- or down-regulated with a *P* value (derived by Student's *t* test) of $P <= 0.05$, indicative of FDR, were considered as differentially expressed. Unsupervised hierarchical clustering of up- and down-regulated genes were performed using Cluster 3.0 software (de Hoon et al. 2004) by applying Pearson Uncentered Algorithm with Average Linkage Rule. Further, the clusters of transcripts were visualized using Java Tree View software (Saldanha 2004) to identify patterns of up- and down-regulated transcripts. Biological analysis of differentially expressed genes was done based on GO annotations obtained from EBI-GOA database (Camon 2004) and KEGG Pathway database (Kanehisa et al. 2017).

qRT-PCR validation of selected transcripts encoding secondary metabolite biosynthetic pathway genes

The quantitative expression of unitranscripts, found to be associated with the biosynthesis of secondary metabolite compounds, was validated by qRT-PCR. Ten genes were chosen for qRT-PCR for which primers were designed from the final assembled unigene sequences [log2-fold change value (≤ -3), *P* value (1*e*−10)] of selected secondary metabolic pathways (Suppl. Table S1). Two housekeeping genes, *Actin* and *Histone H3*, were used as internal reference genes for normalization. The relative expression of the transcripts were analyzed in three biological replicates by qRT-PCR using KAPA SYBR FAST qPCR Master Mix Kit (2X) (Sigma-Aldrich, USA) in ABI 7900HT Fast Real-Time PCR System (Applied Biosystems, USA) and melting curve analysis was carried out to confirm the specificity. The expression levels for each transcript relative to the reference genes were determined by $2^{-\Delta\Delta CT}$ values (Livak and Schmittgen 2001). Triplicate data were analyzed statistically (± standard deviation).

Identification of SSRs and transcription factors

Microsatellites or simple sequence repeats (SSR) specific to the unigene transcripts in the final transcriptome were identified using MISA—Microsatellite Search Tool (Thiel et al. 2003). Transcription factor families present in the final transcriptome were identified by mapping the unigene transcripts against Plant TFDB v4.0, (https://planttfdb.cbi.pku.edu.cn/) and subjected to BLASTX analysis (*e* value cut off 1E−06).

Acknowledgements The author Karpaga Raja Sundari acknowledges DST-National Post-Doctoral Fellowship for funding this research work. KRS also acknowledges the support of Dr. Modhumita Dasgupta, Scientist F, Institute of Forest Genetics and Tree Breeding, Coimbatore, for providing the guidance and facilities to carry out RNA extraction from tissues of *P. zeylanica*. The study was funded by Science and Engineering Research Board(SERB) (Grant no. DST/NPDF/2015/000524).

Author contributions The author BKRS and SPT conceptualized and designed the experiments for RNA sequencing. RB contributed to RNA sequencing and acquisition of data while BKRS and RB analyzed the transcriptome data. BKRS conducted experimental validation by qRT-PCR for the obtained transcripts. BKRS and BSD interpreted the data and drafted the article. BSD and SPT critically revised the article and made the final approval of the manuscript.

Compliance with ethical standards

Conflict of interest The corresponding author, Karpaga Raja Sundari has received grant from Department of science and Technology, Govt. of India. All the authors declare that they have no competing financial and non-financial interests.

Data archiving statement *P. zeylanica* leaf and root transcriptome raw data can be accessed at NCBI SRA database under SRA ID SRP144988.

References

Adisakwattana S (2017) Cinnamic acid and its derivatives: mechanisms for prevention and management of diabetes and its complications. Nutrients 9:163

Anandan A, Eswaran R, Doss A, Sangeetha G, Anand SP (2012) Chemical compounds investigation of *Lucas Aspera* leaves—a

potential folklore medicinal plant. Asian J Pharm Clin Res 5:86–88
Annadurai RS, Neethiraj R, Jayakumar V, Damodaran AC, Rao SN, Katta MA, Gopinathan S, Sarma SP, Senthilkumar V, Niranjan V (2013) De novo transcriptome assembly (NGS) of *Curcuma longa* L. rhizome reveals novel transcripts related to anticancer and antimalarial terpenoids. PLoS ONE 8:56217. https://doi.org/10.1371/journal.pone.0056217
Bankar KG, Todur VN, Shukla RN, Vasudevan M (2015) Ameliorated de novo transcriptome assembly using Illumina paired end sequence data with trinity assembler. Genom Data 5:352–359
Bowles DJ, Isayenkova J, Lim EK, Poppenberger B (2005) Glycosyltransferases: managers of small molecules. Curr Opin Plant Biol 8:254–263
Camon E, Magrane M, Barrell D, Lee V, Dimmer E, Maslen J, Binns D, Harte N, Lopez R, Apweiler R (2004) The gene ontology annotation (GOA) database: sharing knowledge in uniprot with gene ontology. Nucleic Acids Res 32:262–266
Carson CF, Hammer KA (2010) Chemistry and bioactivity of essential oils. In: Thormar H (ed) Lipids and essential oils as antimicrobial agents. John Wiley and Sons, New York, pp 203–238
Chauhan M (2014) A review on morphology, phytochemistry and pharmacological activities of medicinal herb *Plumbago Zeylanica* Linn. J Pharmacogn Phytochem 3(2):95–118
Chen C, Li A (2016) Transcriptome analysis of differentially expressed genes involved in proanthocyanidin accumulation in the rhizomes of *Fagopyrum dibotrys* and an irradiation-induced mutant. Front Physiol 7:100. https://doi.org/10.3389/fphys.2016.00100
Chhetri SBB, Khatri D (2017) Phytochemical screening, total phenolic and flavonoid content and antioxidant activity of selected nepalese plants. World J Pharm Pharm Sci 6:951–968
Cuong DM, Kwon SJ, Jeon J, Park YJ, Park JS, Park SU (2018) Identification and characterization of phenyl propanoid biosynthetic genes and their accumulation in bitter melon (*Momordica charantia*). Molecules 23:469. https://doi.org/10.3390/molecules23020469
Dao TTH, Linthorst HJM, Verpoorte R (2011) Chalcone synthase and its functions in plant resistance. Phytochem Rev 10:397–412
de Hoon MJ, Imoto S, Nolan J, Miyano S (2004) Open source clustering software. Bioinformatics 20:1453–1454
Ferreyra MLF, Rius SP, Casati P (2012) Flavonoids: biosynthesis, biological functions, and biotechnological applications. Front Plant Sci 3:222. https://doi.org/10.3389/fpls.2012.00222
Fukushima A, Nakamura M, Suzuki H, Yamazaki M, Knoch E, Mori T, Umemoto N, Morita M, Hirai G, Sodeoka M, Saito K (2016) Comparative characterization of the leaf tissue of *Physalis alkekengi* and *Physalis peruviana* using RNA-seq and metabolite profiling. Front Plant Sci 7:1883. https://doi.org/10.3389/fpls.2016.01883
Ganesan K, Gani SB (2013) Ethnomedical and pharmacological potentials of *Plumbago zeylanica* L—a review. Am J Phytomed Clin Ther 1(3):313–337
Guo X, Li Y, Li C, Luo H, Wang L, Qian J, Luo X, Xiang L, Song J, Sun C, Xu H, Yao H, Chen S (2013) Analysis of the *Dendrobium officinale* transcriptome reveals putative alkaloid biosynthetic genes and genetic markers. Gene 527:131–138
Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J et al (2013) De novo transcript sequence reconstruction from RNA-seq using the trinity platform for reference generation and analysis. Nat Protoc 8:1494–1512
Hemaiswarya S, Doble M (2013) Combination of phenyl propanoids with 5-fluorouracil as anti-cancer agents against human cervical cancer (HeLa) cell line. Phytomedicine 20:151–158
Hiradeve S, Danao K, Kharabe V, Mendhe B (2011) Evaluation of anticancer activity of *Plumbago zeylanica* Linn. leaf extract. Int J Biomed Sci 1:01–09
Ittiyavirah SP, Ruby R (2014) Effect of hydro-alcoholic root extract of *Plumbago zeylanica* alone and its combination with aqueous leaf extract of *Camellia sinensis* on haloperidol induced parkinsonism in Wistar rats. Ann Neurosci 21:47–50
Jadhav SKR, Patel KA, Dholakia BB, Khan BM (2012) Structural characterization of a flavonoid glycosyltransferase from *Withania somnifera*. Bioinformation 8:943–949
Jain P, Sharma HP, Basri F, Baraik B, Kumari S, Pathak C (2014) Pharmacological profiles of ethno-medicinal plant: *Plumbago zeylanica* l.—a review. Int J Pharm Sci Rev Res 24:157–163
Jeena GS, Fatima S, Tripathi P, Upadhyay S, Shukla RK (2017) Comparative transcriptome analysis of shoot and root tissue of *Bacopa monnieri* identifies potential genes related to triterpenoid saponin biosynthesis. BMC Genom 18:490. https://doi.org/10.1186/s12864-017-3865-5
Ji Y, Xiao J, Shen Y, Ma D, Li Z, Pu G, Li X, Huang L, Liu B, Ye H (2014) Cloning and characterization of AabHLH1, a bHLH transcription factor that positively regulates artemisinin biosynthesis in *Artemisia annua*. Plant Cell Physiol 55:1592–1604
Jiang N, Doseff AI, Grotewold E (2016) Flavones: from biosynthesis to health benefits. Plants 5:27. https://doi.org/10.3390/plants5020027
Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 45:353–361
Kawiak A, Domachowska A, Jaworska A, Lojkowska E (2017) Plumbagin sensitizes breast cancer cells to tamoxifen-induced cell death through GRP78 inhibition and Bik upregulation. Sci Rep. https://doi.org/10.1038/srep43781
Kumar S, Pandey AK (2013) Chemistry and biological activities of flavonoids: an overview. Sci World J. https://doi.org/10.1155/2013/162750
Kumar A, Kumar S, Bains S, Vaidya V, Singh B, Kaur R, Kaur J, Singh K (2016) De novo transcriptome analysis revealed genes involved in flavonoid and vitamin C biosynthesis in *Phyllanthus emblica* (L). Front Plant Sci 7:1610. https://doi.org/10.3389/fpls.2016.01610
Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359
Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinform 12:323
Lin D, Xiao M, Zhao J, Li Z, Xing B, Li X, Kong M, Li L, Zhang Q, Liu Y, Chen H, Qin W, Wu H, Chen S (2016) An overview of plant phenolic compounds and their importance in human nutrition and management of type 2 diabetes. Molecules 21:1374. https://doi.org/10.3390/molecules21101374
Liu Y, Zhu X, Li W, Wen H, Gao Y, Liu Y, Liu C (2014) Enhancing production of ergosterol in *Pichia pastoris* GS115 by overexpression of 3-hydroxy-3-methylglutaryl CoA reductase from *Glycyrrhiza uralensis*. Acta Pharm Sin B 4:161–166
Liu Y, Wang Y, Guo F, Zhan L, Mohr T, Cheng P, Huo N, Gu R, Pei D, Sun J, Tang L, Long C, Huang L, Gu YQ (2017) Deep sequencing and transcriptome analyses to identify genes involved in secoiridoid biosynthesis in the Tibetan medicinal plant *Swertia mussotii*. Sci Rep 7:43108. https://doi.org/10.1038/srep43108
Liu QQ, Luo L, Zheng L (2018) Lignins: biosynthesis and biological functions in plants. Int J Mol Sci. https://doi.org/10.3390/ijms19020335
Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25:402–408
Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550

Lu X, Tang K, Li P (2016) Plant metabolic engineering strategies for the production of pharmaceutical terpenoids. Front Plant Sci. https://doi.org/10.3389/fpls.2016.01647

Malar RJJT (2013) Studies on interspecific and intraspecific variation in the genus: Plumbago (Plumbaginaceae) from south India. Dissertation, Manonmaniam Sundaranar University, Tamil Nadu.

Mandavkar YD, Jalalpure SS (2011) A comprehensive review on *Plumbago zeylanica* Linn. Afr J Pharm Pharmacol 5:2738–2747

Mazumdar AB, Chattopadhyay S (2015) Sequencing, de novo assembly, functional annotation and analysis of *Phyllanthus amarus* leaf transcriptome using the Illumina platform s. Front Plant Sci 6:1199. https://doi.org/10.3389/fpls.2015.01199

Miralpeix B, Sabalza M, Twyman RM, Capell T, Christou P (2014) Strategic patent analysis in plant biotechnology: terpenoid indole alkaloid metabolic engineering as a case study. Plant Biotechnol J 12:117–134

Misra BB (2014) An updated snapshot of recent advances in transcriptomics and genomics of phytomedicinals. J Postdoct Res 2:1–15

Ndikuryayo F, Moosavi B, Yang W-C, Yang G-F (2017) 4-hydroxyphenyl pyruvate dioxygenase inhibitors: from chemical biology to agrochemicals. J Agric Food Chem 65:8523–8537

Nile SH, Khobragade CN (2010) Antioxidant activity and flavonoid derivatives of *Plumbago zeylanica*. J Nat Prod 3:130–133

Pastore S, Mariani V, Cesareo E, Korkina L (2007) Anti-inflammatory effects of phenyl propanoids on human keratinocytes. Free Radic Res 41:46

Raghu D, Karunagaran D (2014) Plumbagin down regulates Wnt signaling independent of p53 in human colorectal cancer cells. J Nat Prod 77:1130–1134

Rahnamaie-Tajadod R, Loke KK, Goh HH, Noor NM (2017) Differential gene expression analysis in *Polygonum minus* leaf upon 24 h of methyl jasmonate elicitation. Front Plant Sci. https://doi.org/10.3389/fpls.2017.00109

Rajakrishnan R, Lekshmi R, Benil PB, Thomas J, AlFarhan AH, Rakesh V, Khalaf S (2017) Phytochemical evaluation of roots of *Plumbago zeylanica* L. and assessment of its potential as a nephroprotective agent. Saudi J Biol Sci 24:760–766

Rios JL (2010) Effects of triterpenes on the immune system. J Ethnopharmacol 128:1–14

Rondeau G, Abedinpour P, Chrastina A, Pelayo J, Borgstrom P, Welsh J (2018) Differential gene expression induced by anti-cancer agent plumbagin is mediated by androgen receptor in prostate cancer cells. Sci Rep 8:2694. https://doi.org/10.1038/s41598-018-20451-9

Roy A, Bharadvaja N (2017) A review on pharmaceutically important medical plant: *Plumbago zeylanica*. J Ayu Herb Med 3:225–228

Roy A, Attre T, Bharadvaja N (2017) Anticancer agent from medicinal plants: a review. In: Tiezzi A (ed) New aspects in medicinal plants and pharmacognosy. JB Books Publisher, Poland

Ru M, Wang K, Bai Z, Peng L, He S, Wang Y, Liang Z (2017) A tyrosine aminotransferase involved in rosmarinic acid biosynthesis in *Prunella vulgaris* L. Sci Rep. https://doi.org/10.1038/s41598-017-05290-4

Saldanha AJ (2004) Java Treeview—extensible visualization of microarray data. Bioinformatics 20:3246–3248

Satyavati VG, Gupta KA (1987) Medicinal plants of India. Indian Council Med Res 2:471–477

Savoia D (2012) Plant-derived antimicrobial compounds: alternatives to antibiotics. Fut Microbiol 7:979–990

Schluttenhofer C, Yuan L (2015) Regulation of specialized metabolism by WRKY transcription factors. Plant Physiol. https://doi.org/10.1104/pp.114.251769

Shivashankar S, Sumathi M, Krishnakumar NK, Rao VK (2015) Role of phenolic acids and enzymes of phenylpropanoid pathway in resistance of chayote fruit (*Sechium edule*) against infestation by melon fly *Bactrocera cucurbitae*. Ann Appl Biol 166:420–433

Shukla AK, Mall M, Rai SK, Singh S, Nair P, Parashar G, Shasany AK, Singh SC, Joshi VK, Khanuja SPS (2013) A transcriptomic approach for exploring the molecular basis for dosha-balancing property-based classification of plants in Ayurveda. Mol Biol Rep 40:3255–3262

Singh J, Mishra NP, Jain SP, Singh SC, Sharma A, Khanuja SPS (2004) Traditional uses of *Plumbago zeylanica* (Chitrak). J Med Arom Plant Sci 26:795–800

Singh M, Nagori K, Iyer S, Khare G, Sharwan G, Tripathi DK (2011) Ethano-medicinal, traditional and pharmacological aspects of *Plumbago zeylanica*. Pharmacol Online Newslett 3:684–700

Singh S, Sharma B, Kanwar SS, Kumar A (2016) Lead phytochemicals for anticancer drug development. Front Plant Sci 7:1667. https://doi.org/10.3389/fpls.2016.01667

Srivastava S, Sanchita SR, Srivastava G, Sharma A (2018) Comparative study of withanolide biosynthesis-related miRNAs in root and leaf tissues of *Withania somnifera*. Appl Biochem Biotechnol. https://doi.org/10.1007/s12010-018-2702-x

Stewart CJR, Vickery CR, Burkart MD, Noel JP (2013) Confluence of structural and chemical biology: plant polyketide synthases as biocatalysts for a bio-based future. Curr Opin Plant Biol 16:365–372

Stockigta J, Barleben L, Panjikar S, Lorisa EA (2008) 3D-Structure and function of strictosidine synthase—the key enzyme of monoterpenoid indole alkaloid biosynthesis. Plant Physiol Biochem 46:340–355

Stracke R, Ishihara H, Barsch GHA, Mehrtens F, Niehaus K, Weisshaar B (2007) Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. The Plant J 50:660–677

Sundari BKR, Telapolu S, Dwarakanath BS, Thyagarajan SP (2017) Cytotoxic and antioxidant effects in various tissue extracts of *Plumbago zeylanica*: implications for anticancer potential. Phcog J 9:706–712

Sunil C, Duraipandiyan V, Agastian P, Ignacimuthu S (2012) Antidiabetic effect of plumbagin isolated from *Plumbago zeylanica* L. root and its effect on GLUT4 translocation in streptozotocin-induced diabetic rats. Food Chem Toxicol 50:4356–4363

Suttipanta N, Pattanaik S, Kulshrestha M, Patra B, Singh SK, Yuan L (2011) The transcription factor CrWRKY1 positively regulates the terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. Plant Physiol 157:2081–2093

Thiel T, Michalek W, Varshney RK, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR markers in barley (*Hordeum vulgare* L.). Theor Appl Genet 106:411–422

Tholl D (2006) Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. Curr Opin Plant Biol 9:297–304

Tholl D (2015) Biosynthesis and biological functions of terpenoids in plants. Adv Biochem Eng Biotechnol 148:63–106

Upadhyay S, Phukan UJ, Mishra S, Shukla RK (2014) *De novo* leaf and root transcriptome analysis identified novel genes involved in Steroidal sapogenin biosynthesis in *Asparagus racemosus*. BMC Genomics 15:746

Vishnukanta RAC (2008) Analgesic and anti-inflammatory activity of hydroalcoholic extract of *Plumbago zeylanica* leaf extract. Pharmacogn Mag 15:S133–136

Widhalm JR, Rhodes D (2016) Biosynthesis and molecular actions of specialized 1,4-naphthoquinone natural products produced by horticultural plants. Hortic Res. https://doi.org/10.1038/hortres.2016.46

Yang H, Dou QP (2010) Targeting apoptosis pathway with natural terpenoids: implications for treatment of breast and prostate cancer. Curr Drug Targets 11:733–744

Yang L, Ding G, Lin H, Cheng H, Kong Y, Wei Y, Fang X, Liu R, Wang L, Chen X, Yang C (2013a) Transcriptome analysis of medicinal plant *salvia miltiorrhiza* and identification of genes related to tanshinone biosynthesis. PLoS ONE 8:e80464. https://doi.org/10.1371/journal.pone.0080464

Yang L, Ding G, Lin H, Cheng H, Kong Y, Wei Y, Fang X, Liu R, Wang L, Chen X, Yang C (2013b) Transcriptome analysis of medicinal plant *Salvia miltiorrhiza* and identification of genes related to tanshinone biosynthesis. PLoS ONE 8:e80464. https://doi.org/10.1371/journal.pone.0080464

Zhang H, Hedhili S, Montiel G, Zhang Y, Chatel G, Pre M, Gantet P, Memelink J (2011) The basic helix-loop-helix transcription factor CrMYC2 controls the jasmonate-responsive expression of the ORCA genes that regulate alkaloid biosynthesis in *Catharanthus roseus*. Plant J 67:61–71

Zhang K, Chen D, Ma K, Wu X, Hao H, Jiang S (2018) NAD(P)H: quinone oxidoreductase 1 (NQO1) as a therapeutic and diagnostic target in cancer. J Med Chem. https://doi.org/10.1021/acs.jmedchem.8b00124

Zhu W, Yang B, Komatsu S, Lu X, Li X, Tian J (2015) Binary stress induces an increase in indole alkaloid biosynthesis in *Catharanthus roseus*. Front. Plant Sci. https://doi.org/10.3389/fpls.2015.00582

Affiliations

Balachandran Karpaga Raja Sundari[1] · Roli Budhwar[2] · Bilikere S. Dwarakanath[1,3] · S. P. Thyagarajan[1]

Roli Budhwar
roli@bionivid.com

Bilikere S. Dwarakanath
DWARAKANATH@sphic.org.cn

S. P. Thyagarajan
deanresearch@sriramachandra.edu.in

[1] Central Research Facility, Sri Ramachandra Institute of Higher Education and Research, Chennai 600116, India

[2] Bionivid Technology [P] Limited, Kasturi Nagar, Bangalore 560043, India

[3] Shanghai Proton and Heavy Ion Center, Pudong 201321, Shanghai, China