JOURNAL OF TROPICAL LIFE SCIENCE
2024, Vol. 14, No. 3, 509 – 536
http://dx.doi.org/10.11594/jtls.14.03.08

Research Article

Identification of Gene Candidates in Diterpenoid Biosynthesis of *Curcuma longa*: An mRNA Sequencing Approach

Hafizh Fadhullah [1]*, Devit Purwoko [2], Siti Zulaeha [2], Nurul Fitri Hanifah [2], Endah Dwi Hartuti [2], Gemilang Rahmadara [2], Anna Safarrida [2], Rikania Reninta [2], Evawati [3], Irwan Roza [3], Teuku Tajuddin [2]

[1] Biotechnology Graduate Program, Graduate School, Bogor Agricultural University (IPB University), Bogor 16680, Indonesia
[2] Research Centre for Applied Botany, Research Organization for Life Sciences and Environment, National Research and Innovation Agency, Science and Technology Park of Soekarno, Cibinong, Bogor West Java 16911, Indonesia
[3] Food Technology Study Program, Department of Agricultural Product Technology, Politeknik Pertanian Negeri Payakumbuh, Payakumbuh 26271, Indonesia

Article history:
Submission December 2023
Revised March 2024
Accepted March 2024

**Corresponding author:*
E-mail: hafizh.fadhullah2@gmail.com

ABSTRACT

Curcuma longa is a medicinal plant renowned for its therapeutic properties and potential treatment of cancer. This study focused on the biosynthesis of diterpenoids in the rhizome and leaves of *C. longa*. The genes responsible for producing these medicinal compounds were analyzed using BLASTx, Gene Ontology (GO) annotation, differential expression, and homology. The substantial dataset was obtained from the National Center for Biotechnology Information (NCBI), comprising 151,730,334 clean reads and 167,264 transcripts for the analysis. The results of the BLASTx analysis were as follows: NR yielded 65.93%, Swiss-Prot yielded 44.52%, and COG yielded 17.35%. Subsequently, GO annotation was performed using Blast2GO, resulting in an annotation rate of 56.79%. Differential expression analysis revealed a total of 636 genes that were significantly differentiated between the rhizome and leaves. The homology analysis resulted in 11 proteins associated with diterpenoid biosynthesis and nine proteins related to CYP450. Approximately three class I proteins were highly expressed in the rhizome. Additionally, seven CYP450 enzymes from the CYP71D and CYP726 subfamilies were identified; three were highly expressed in the rhizome. The expression patterns of these enzymes were similar to the aforementioned three class I diTPSs, indicating their potential involvement in macroditerpenoid biosynthesis in *C. longa*. These findings provide valuable genomic resources for future functional genomics research on *C. longa*, facilitating targeted efforts to enhance the production of bioactive compounds.

Keywords: Curcuma longa, Diterpenoid biosynthesis, Gene expression

Introduction

Curcuma longa, classified within the family *Zingiberaceae*, is a medicinal plant extensively employed in traditional medicine [1, 2]. This prevalence in medicinal applications [3] can be attributed to the presence of secondary metabolites resulting from metabolic processes and external factors such as abiotic and biotic stressors [4, 5]. The exploration of these secondary metabolites is of considerable interest, particularly concerning the elucidation of their biosynthetic pathways. Numerous studies have investigated the biosynthetic processes occurring in *C. longa*, shedding light on various aspects such as the biosynthesis of terpenoid [6], sesquiterpenoid [7], and curcuminoid [8]. Understanding these intricate biosynthetic pathways is crucial for unraveling the mechanisms underlying the production of bioactive compounds in *C. longa* [9, 10].

How to cite:
Fadhullah H, Purwoko D, Zulaeha S, et al. (2024) Identification of Gene Candidates in Diterpenoid Biosynthesis of Curcuma longa: an mRNA Sequencing Approach. Journal of Tropical Life Science 14 (3): 509 – 536. doi: 10.11594/jtls.14.03.08.

C. longa has been a focal point in research related to the treatment of diseases, particularly cancer, owing to the presence of bioactive compounds, including diterpenoids. This compound has been known for its potential in reducing the malignancy of cancer cells [4, 11]. Exploring the synthesis of these diterpenoids in *C. longa* has become an intriguing avenue of research. The pathway leading to diterpenoid synthesis is a specific focus, leveraging insights from previous studies that have elucidated diterpenoid pathways in other medicinal plants [12, 13].

Transcriptomics is a methodology that provides a deeper insight into the genetic processes underlying the biosynthesis of secondary metabolites through the analysis of differential gene expression [12, 14, 15]. The transcriptomic analysis yields data in the form of differentially expressed transcripts for each gene under specific treatments or in specific tissues, revealing that different genes are expressed in each tissue of *C. longa*. Previous research has demonstrated the utility of transcriptomics in gaining novel insights into abiotic stress responses [16], functional plant analysis [17], and the biosynthesis processes of secondary metabolites [12]. Such data can be instrumental in discerning variations in biosynthesis between different plant tissues based on the distinct gene expression patterns observed in the formation of biosynthetic pathways [18]. Transcriptomic analyses offer a comprehensive view of the molecular mechanisms involved in secondary metabolite biosynthesis, providing valuable information for understanding the intricacies of plant responses to various stimuli and contributing to advancements in functional genomics and plant metabolic engineering.

Through this research, the authors aimed to delve into the gene expression processes occurring in the rhizome and leaf tissues of *C. longa* using transcriptomics. The study provides a comprehensive overview of the pathways involved in the gene expression biosynthesis process expressed in both rhizome and leaf tissues of *C. longa*. The authors anticipate that this research can serve as a reference for breeders and industries, offering insights into the genes that play a crucial role in shaping diterpenoid biosynthesis. Furthermore, the functional annotation data generated from this study can aid in identifying key genes and pathways involved in diterpenoid biosynthesis, facilitating targeted efforts to enhance the production of bioactive compounds in *C. longa*.

Material and Methods

Datasets

The transcriptomic data were obtained from the Sequence Retrieval Achieve (SRA) database (https://www.ncbi.nlm.nih.gov/sra). The data were collected from six different SRA, including three for the leaves and three for the rhizome, to compare the two tissues. The rhizome data of *C. longa* were from three SRA accessions dataset [6], namely SRR495223, SRR495363, and SRR495364. Meanwhile, the leaves data were from two SRA accessions dataset, namely SRR15204660 and SRR12560783, and from the 1000 Plant (1KP) [19, 20] with one SRA accession, namely ERR2040761.

Transcriptomic analysis and functional annotation

The six SRA accessions were analyzed by uploading to a platform (www.usegalaxy.org.au). The analysis and visualization used a public server (www.usegalaxy.org.au) [21] and R software [22]. The transcriptomic data analysis used Galaxy following Bretaudeau *et al.* [23] with modifications. First, a quality control for transcriptomic data was carried out using Fastqc [24] followed by Trimmomatic [25] to obtain data cleaned from the unreadable data. Then, the assembly process was conducted using Trinity [26], resulting in a transcript. The transcript was analyzed for alignment and abundance estimation using Bowtie [27] and RSEM [28]. Subsequently, it was filtered to get the unigene sequence using "Filter Low Expression Transcript" in Trinity [26]. The unigene sequence was functionally annotated with BLASTx [29] in the protein database of non-redundant NCBI (NR), Swissprot, and Clusters of Orthologous Group (COG) (E-values < 10^{-5}) [30]. After that, the software Blast2GO [31] was used to determine GO term distribution for each unigene [32].

Differential expression analysis

The "Filter Low Expression Transcript" tool [26] was used with specific parameters to refine the transcriptome data. The input for this tool included the "Trinity assembly" file and the "Expression matrix" generated using the "Build Expression Matrix" tool [26]. The tool was configured to filter out transcripts with low expression

levels. This filtering process helps focus the analysis on transcripts with meaningful expression levels, thereby improving the accuracy and relevance of subsequent analyses. It was then analyzed to find the genes expressed with significant differences between the two tissues using DESeq2 [33]. The genes that showed a significantly different expression were visualized using volcano plots in R software to ease the interpretation [34, 35].

Gene mining diTPS, CYP450, and diterpenoid backbone

To discover possible cytochrome P450 (P450) and diterpene synthase (diTPS) candidates, diterpenoid backbone-tailored databases were built using publically accessible protein sequences, guaranteeing the representation of the least redundant sequence sets. This strategy aimed to match the diTPS and P450 gene functional ranges. After that, gene mining was done using tBLASTx to find significant candidates from the assembled transcripts. An E-value threshold of 1.0×10^{-50} and a minimum read length of 150 amino acids were used. The CLUSTAL W [36] was utilized to perform protein sequence alignments, and then MEGA 11 [37] was used to apply the maximum likelihood technique to create a phylogenetic tree.

Results and Discussion

Transcriptomic analysis

This study used two parts of *C. longa*, including the rhizomes and leaves. The data for rhizomes were collected from three accessions, namely SRR495223, SRR495363, and SRR495364. Meanwhile, the leaf data were obtained from two accessions, SRR15204660 and SRR12560783, as well as from the 1000 Plant (1KP), specifically ERR2040761. Table 1 displays the results of the assembly using Trinity, yielding a total of 277,150,123 reads. About 151,730,334 clean reads and 43.63% of GC content were attained through trimmomatic and transcript filtering with a "Filter Low Expression Transcript" using Trinity, employing a filter quality of Q < 30 and the absence of ambiguous "N". The ambiguous "N" refers to the unidentified nitrogen base in sequencing. The total number of transcripts obtained was 167,264 sequences, with a minimum transcript length of 180 bp and a maximum of 15,700 bp (Table 1). The assembled results were utilized for functional annotation, followed by an investigation into significantly differentiated transcripts.

Table 1. Transcriptome assembly statistics of *Curcuma longa*

Scaffold GC content overall	43.63%
Transcript L50	29,263
Transcript N50	1,556
Transcript L90	108,993
Transcript N90	347
Maximum transcript length	15,700
Minimum transcript length	180
Mean transcript length	907
Median transcript length	510
Raw Reads	277,150,123
Clean Reads	151,730,334
Number of transcript generated	167,264

Functional annotation

To derive the potential function of the assembled transcript using Trinity, Diamond searches were conducted using BLASTx against the NR, Swiss-Prot, and COG databases. The results of the BLASTx analysis were as follows: NR yielded 65.93%, Swiss-Prot yielded 44.52%, and COG yielded 17.35% (Figure 1). Subsequently, GO annotation was performed using Blast2GO, resulting in an annotation rate of 56.79%. These functional annotations indicate diverse transcripts in *C. longa*, and their functions are relatively conserved. Unannotated transcript data suggests a lack of genomic information regarding *C. longa*, and some of the transcripts may correspond to untranslated regions (UTRs) or non-coding RNAs.

The efficacy of BLAST searches is contingent upon the length of the query sequence. Consequently, shorter input sequences may not align with known genes. Additionally, unhit transcripts often undergo evolutionary changes, leading to notable disparities in terms of their potential homologs. Figure 2 illustrates the distribution of E-values for the top hits in the NR database, with over half of the sequences (52.49%) displaying homology that can be confidently interpreted (>1.0 x 10-60). Meanwhile, 35.59% of the sequences exhibit E-values ranging from 1.0×10^{-60} to 1.0×10^{-15}. The distribution of similarity demonstrates that the mapped transcripts exhibit a similarity greater than 60% in 94.55% of cases, while only 5.45% exhibit a similarity below 60% (Figure 1).

In order to predict the potential gene functions

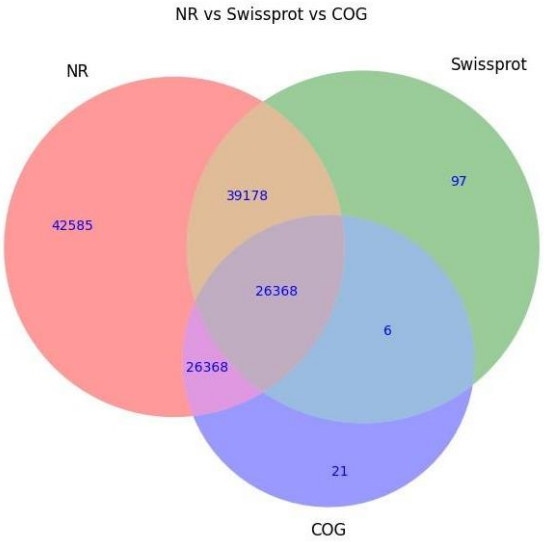


Figure 1. Functional annotation result for NR, COG and Swissprot database

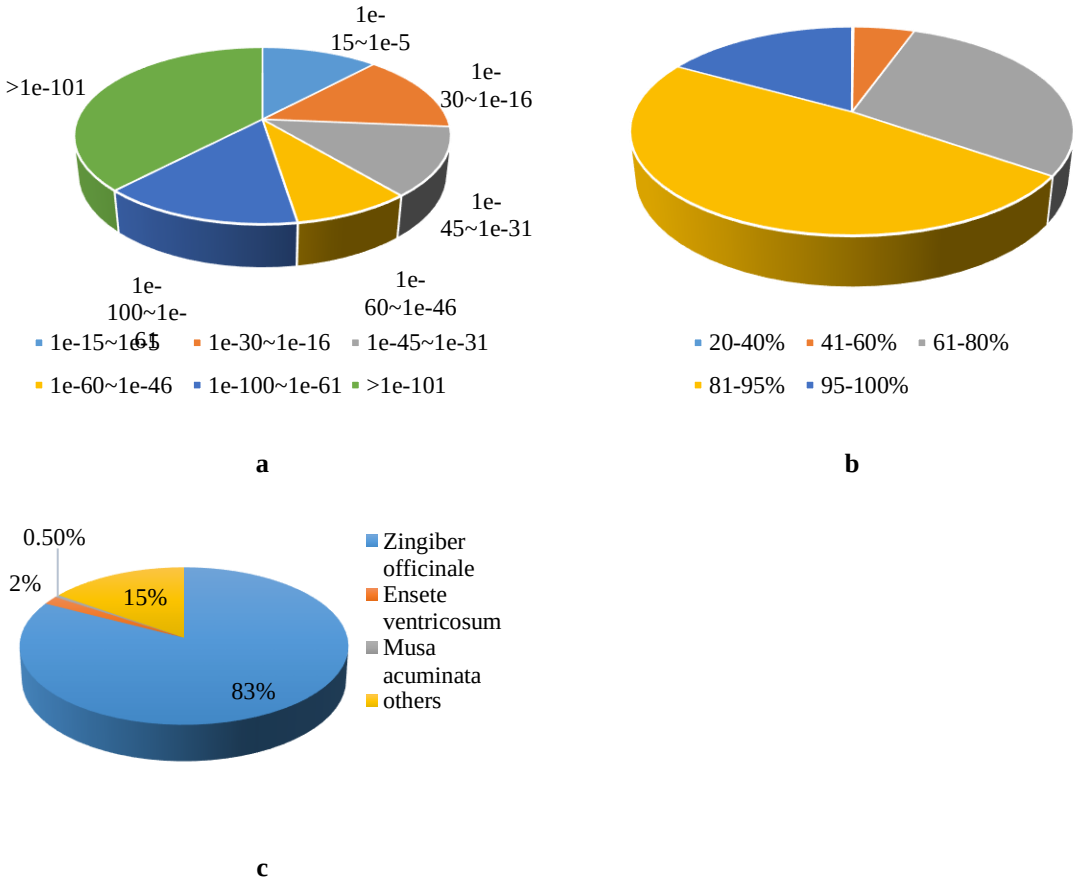


Figure 2. E-value distribution of BLAST hit for each transcript with E-value cut off at 1.0 x 10^{-5} (a), Similarity distribution of Top BLAST hit for each transcript (b), and Species distribution of top BLAST Hit (c).

and assess the transcriptome library's comprehensiveness, all assembled unigenes against clusters in the ortholog group (COG) database were searched. This analysis resulted in classifying 28,550 unique sequences into 25 functional categories (Figure 3). Among these categories, the largest is 'Signal transduction mechanisms' (27.32%), followed by 'Transcription' (21.89%), 'Posttranslational modification, protein turnover, chaperones' (21.01%), and 'RNA processing and

modification' (10.50%).

In a comprehensive manner, GO analysis provides detailed insights into the nature of genes and their associated products. By employing a rigorous and constantly updated gene grouping framework, GO analysis facilitates the division of GO annotations into 56 functional groups. Utilizing the Blast2GO software, a total of 94,997 annotated genes were obtained. The results of the GO annotation will serve as an integral component for subsequent differential analysis and gene mining endeavors, aimed at elucidating the biosynthetic processes of expressed diterpenoids in the rhizome and leaves of *C. longa*.

Differential expression of transcripts in rhizome and leaves of C. longa

The FPKM data generated using RSEM from assembly and alignment using Trinity and Bowtie provide an opportunity to observe significantly differentiated genes. DESeq2 was employed to identify genes that show significant differentiation. As shown in Figure 4, a total of 636 genes were found to be significantly differentiated. Supplementary Table 1 and Supplementary Table 2 present twenty genes each from the rhizome and leaf that are significantly differentiated. Moreover, Supplementary Table 3 and Supplementary Table 4 provide information on several genes that are differentiated considerably and involved in biosynthesis processes in the rhizome and leaf. To ensure the validity of these findings, further in-depth investigations, particularly to elucidate the pathways associated with the biosynthesis enzymes and calculate the FPKM values that drive these biosynthetic processes. This aspect will be addressed in the mining gene and homology analysis section to demonstrate the expression patterns of diterpenoid biosynthesis in the leaf and rhizome of *C. longa*.

Expression patterns of genes in the terpenoid biosynthesis pathway of C. longa

In plants, the cytosolic mevalonate pathway (MVA) and the plastid 2-C-methyl-D-erythritol-4-P pathway (MEP) are responsible for synthesizing the isoprenyl diphosphate precursor with five carbons [38]. To gain valuable insights for gene target selection in *C. longa*, we investigated the expression profiles of genes encoding enzymes involved in the biosynthesis of terpenoid framework in both types of organic tissues. All genes involved in the

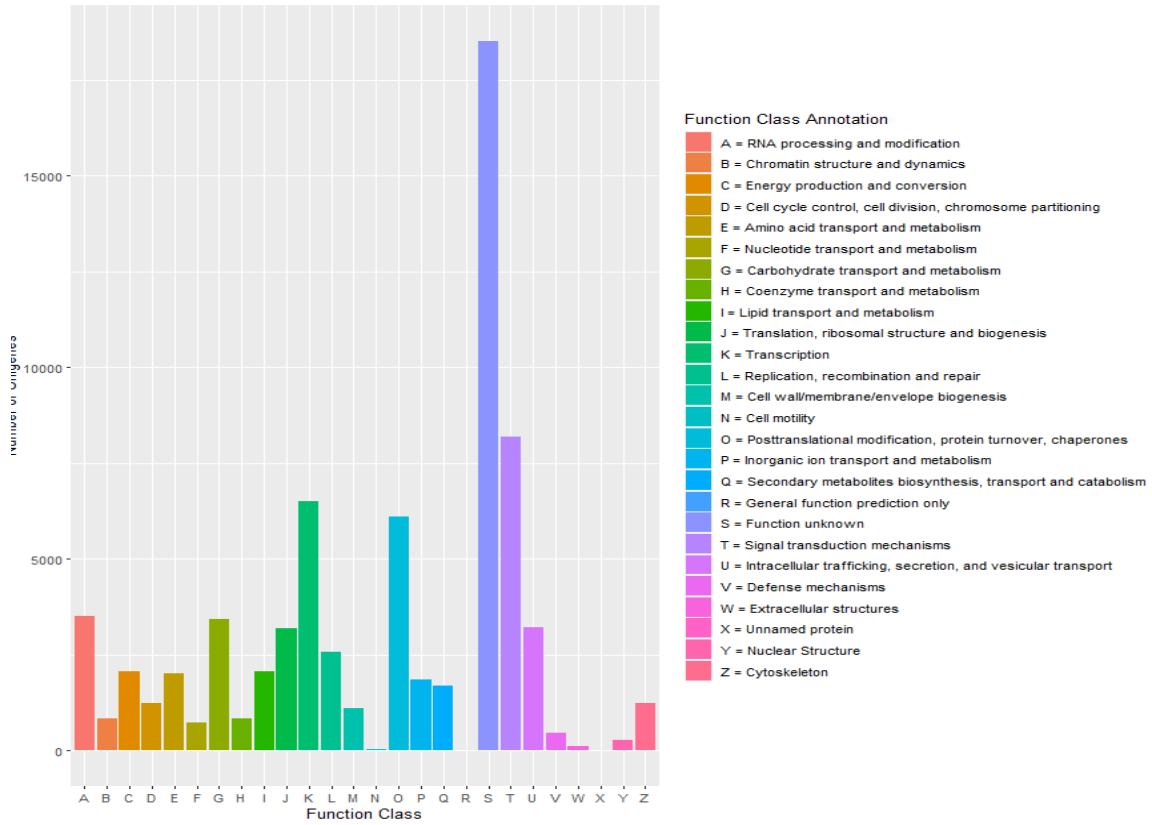


Figure 3. COG classification result of *C. longa*

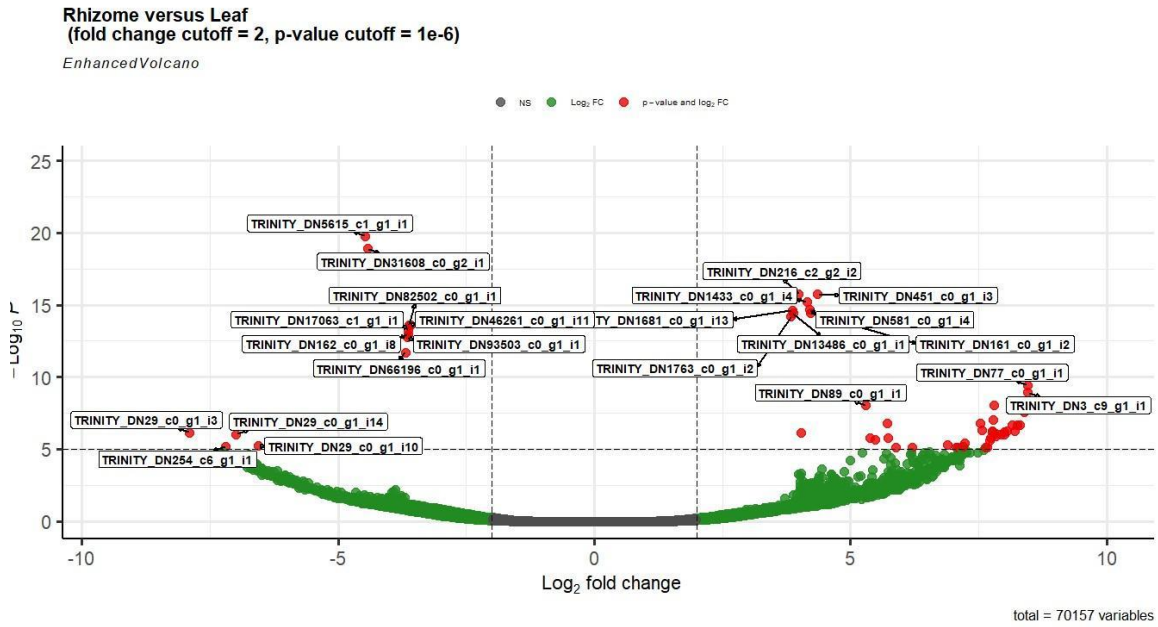


Figure 4. Differential expression analysis for rhizome against leaf in *Curcuma longa*

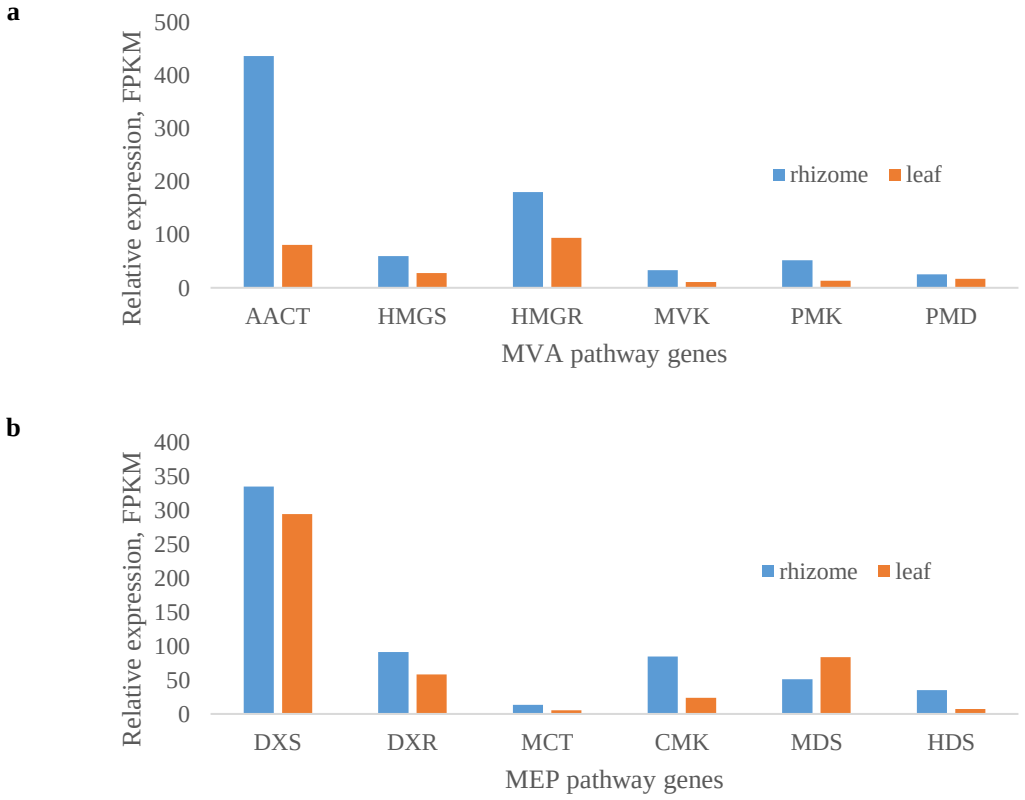


Figure 5. MVA pathway (a) and MEP pathway (b). Expression profile of diterpenoid biosynthesis genes based on FPKM in RNA-seq data. AACT: acetoacetyl- CoA thiolase, HMGS: HMG-CoA synthase, HMGR: HMG-CoA reductase, MVK: mevalonate kinase, PMK: phosphomevalonate kinase, PMD: diphospho-mevalonate decarboxylase, DXS: 1-deoxy-D-xylulose 5-phosphate synthase, DXR: 1-deoxy-D-xylulose 5- phosphate reductoisomerase, MCT: 2-C-methyl-Derythritol 4-phosphate cytidylyltransferase, CMK: 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, MDS: 2-C-methyl-derythritol 2,4- cyclodiphosphate synthase, and HDS: 1-hydroxy-2-methyl-2- (E)-butenyl 4-diphosphate synthase.

terpenoid framework biosynthesis are constitutively in the rhizome and leaves of *C. longa*, but they exhibit distinct expression patterns (Figure 5). FPKM analysis revealed that in both the MVA and MEP pathways, almost all genes exhibited similar expression patterns, and the highest expression levels were observed in the rhizome. In addition, it has the same pattern, with the total content of triterpenoids being higher in the rhizome compared to the stems and lowest in the leaves [12].

C. longa is known for its diverse terpenoid synthesis, encompassing terpenoids, sesquiterpenoids, and diterpenoids, primarily occurring in its rhizome tissue due to biological activities [6, 7]. Previous studies have investigated terpenoid, diterpenoid, and sesquiterpenoid biosynthesis pathways in various plant species, identifying several genes involved in these pathways. The biosynthesis genes expression of diterpenoid in rhizomes shows the critical role of this organ in producing compounds contained in *C. longa*. Several studies have reported the analysis of biosynthesis gene expression in various plant species. Feng *et al*. reported abundant expression of biosynthetic genes in the roots and rhizomes of *Nardostachys jatamansi* [39]. There was a positive correlation between the expression profile of genes related to biosynthesis and the level of accumulation of sesquiterpenoids in rhizome tissue. Ayer *et al*. found a higher expression of curcuminoid biosynthesis genes in rhizomes compared to leaves [40]. Furthermore, Zhu *et al*. demonstrated the differential expression of genes related to the starch synthesis pathway in lotus seeds and rhizomes, with higher expression of amylopectin synthesis genes in rhizomes [41]. Likewise, Chen *et al*. identified higher expression levels of genes related to polysaccharide biosynthesis in rhizomes compared to leaves, further supporting the specialized function of rhizomes [42]. Therefore, the higher expression of diterpenoid biosynthesis genes in *C. longa* rhizomes is in line with the findings from various studies across plant species, emphasizing the key role of rhizomes as specialized functional organs in the medicinal plant *C. longa* [1, 6, 7]. It is worth mentioning that the expression of these genes can differ among different plant species, indicating distinct mechanisms for terpenoid backbone biosynthesis [43–45].

Mining gene and homology analysis

Diterpenoid biosynthesis begins with the GGPP process, which is an isoprenoid diphosphate derived from the MEP pathway. Through the utilization of diterpene synthase (diTPS) and cytochrome P450s (CYP450), GGPP is converted into various diterpenoids [46]. Transcriptomic data analysis has revealed a high expression of diterpenoid biosynthesis in both rhizomes and leaves of *Euphorbia pekinensis* [12]. Wang et al. (2015) identified several diterpenoid compounds, such as sembrane, lathyrane, and casbane types, in the rhizome and leaves [47].

To explore the downstream processes of GGPP biosynthesis, we developed a diTPS database divided into two classes. Our findings include six proteins belonging to class I of the TPS-a subfamily (TRINITY_DN5937_c0_g1_i1, TRINITY_DN7464_c0_g1_i18, TRINITY_DN66695_c0_g1_i3, TRINITY_DN1835_c0_g1_i4, TRINITY_DN1262_c1_g1_i11, and TRINITY_DN11256_c0_g1_i2), and one protein belonging to class I of the TPS-b subfamily (TRINITY_DN72271_c0_g1_i5). For class II of the TPS-c subfamily, we identified three proteins (TRINITY_DN56327_c0_g1_i3, TRINITY_DN11310_c0_g1_i1, TRINITY_DN18553_c0_g1_i1, and TRINITY_DN37294_c0_g1_i17) (Figure 6). By calculating the FPKM values of all orthologous genes, we found that three class I proteins (TRINITY_DN5937_c0_g1_i1, TRINITY_DN66695_c0_g1_i3, TRINITY_DN11256_c0_g1_i2) are highly expressed in the rhizome (Figure 7). These results suggest that these three diTPS proteins are likely involved in diterpenoid biosynthesis in the rhizome of *C. longa*. Additionally, these findings highlight the rhizome as a specific tissue capable of revealing the biosynthetic pathway of diterpenoids [12]. We also identified six genes that share orthology with CBS. Based on the CBS of *Ricinus communis* and *Euphorbia esula*, these genes are potential candidates for identifying macrocyclases involved in *C. longa*'s macrocyclic diterpene metabolism [12, 48].

In order to identify CYP450 enzymes specifically related to diterpene biosynthesis, we conducted a tblastx search for the CYP85 and CYP71

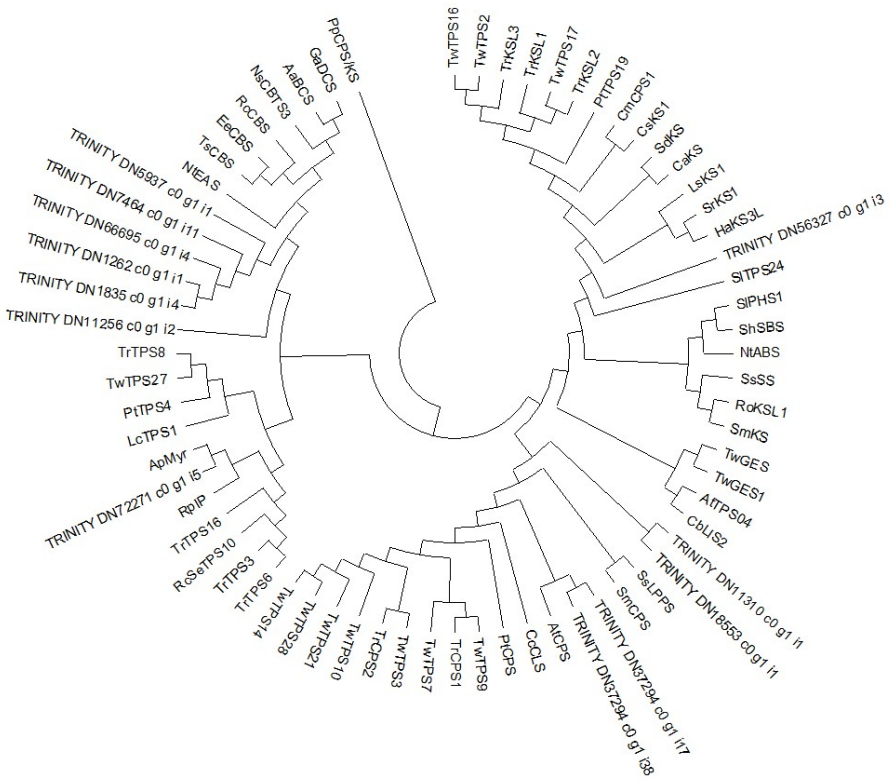


Figure 6. Phylogenetic analysis of diTPS genes

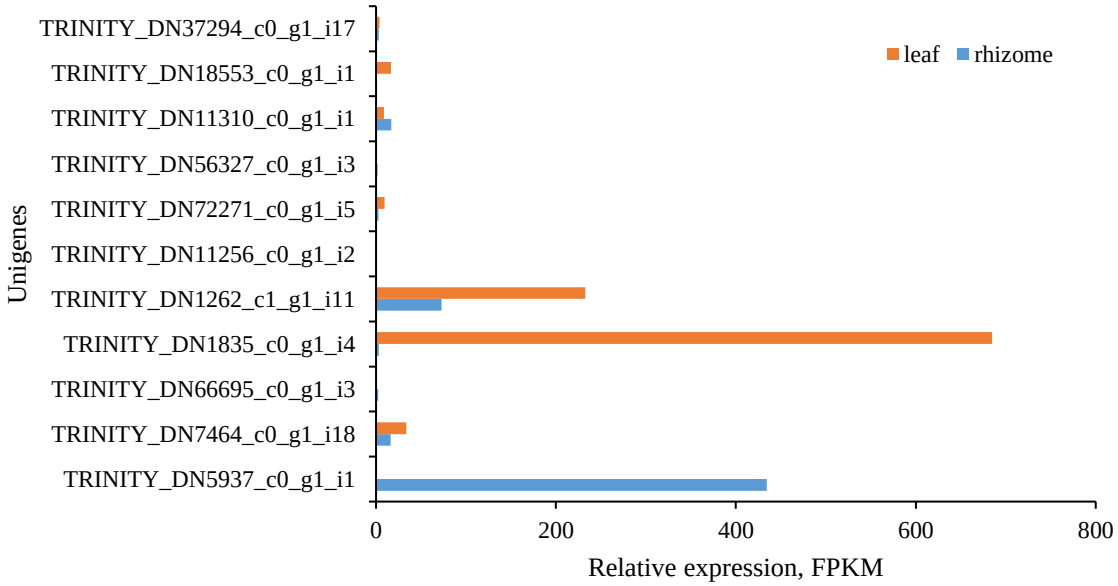


Figure 7. Expression profile of *C. longa* diterpenoid synthase based on FPKM in RNA-seq data

families. This search revealed several CYP71 and CYP85 enzymes that participate in specific diterpene biosynthesis in *C. longa* [15]. We utilized the CYP71 and CYP85 subfamily expansions reported by Trikka *et al*. as a starting point for transcriptome interrogation [15]. The results identified seven CYP450 enzymes from the CYP71 subfamily (TRINITY_DN7781_c0_g1_i2, TRINITY_DN4358_c0_g1_i2, TRINITY_DN14774_c0_g1_i12, TRINITY_DN11046_c0_g1_i5, TRINITY_DN8716_c0_g1_i3, TRINITY_DN20482_c0_g1_i1, TRINITY_DN7_c2_g1_i2,

TRINITY_DN7781_c1_g1_i1, TRINITY_DN7781_c0_g1_i7) (Figure 8). We analyzed the expression of genes encoding these CYP450 enzymes in rhizomes and leaves using FPKM. The results showed that most genes are expressed in both tissue types but exhibit distinct expression patterns (Figure 9). TRINITY_DN7781_c0_g1_i2, TRINITY_DN4358_c0_g1_i2, TRINITY_DN8716_c0_g1_i3, TRINITY_DN7781_c1_g1_i1, and TRINITY_DN7781_c0_g1_i7 showed the highest expression levels in leaves. However, TRINITY_DN14774_c0_g1_i12, TRINITY_DN11046_c0_g1_i5, TRINITY_DN20482_c0_g1_i1, and

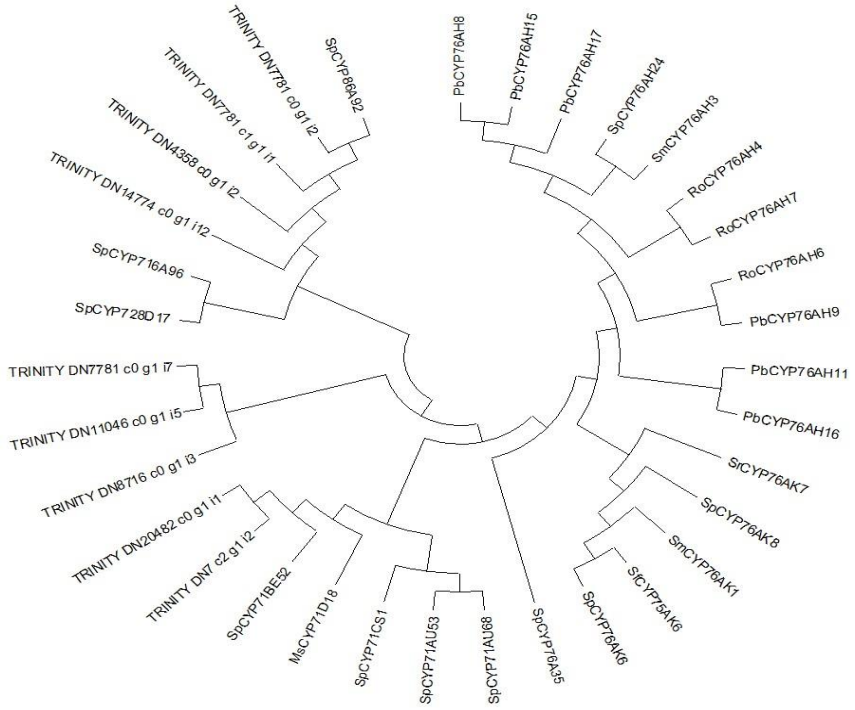


Figure 8. Phylogenetic analysis of CYP450 genes

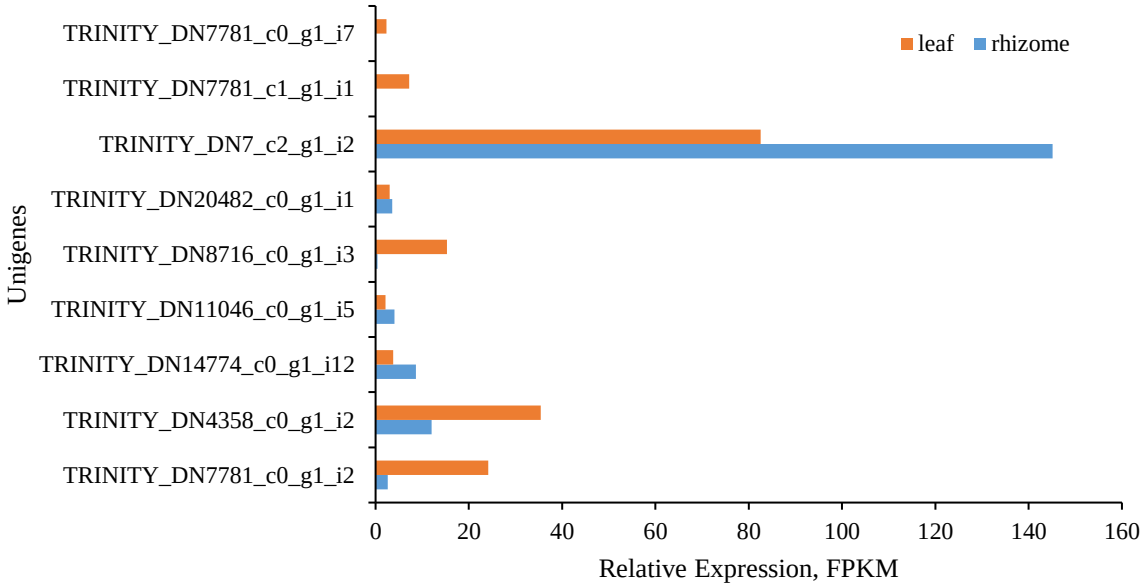


Figure 9. Expression profile of *C. longa* candidate cytochromes P450 in CYP71D and CYP726 sub-families based on FPKM in RNA-seq data

TRINITY_DN7_c2_g1_i2 were highly expressed in the rhizome of *C. longa* and exhibited expression patterns similar to diTPS proteins TRINITY_DN5937_c0_g1_i1, TRINITY_DN66695_c0_g1_i3, and TRINITY_DN11310_c0_g1_i1. Further research is required to elucidate the functions of these newly discovered genes in specific diterpene metabolism.

Diterpenoid macrocycles are crucial bioactive compounds in *C. longa*'s isoprenoid production. Previous studies [49, 50] have emphasized the investigation of casbane synthase, the enzyme responsible for cyclizing the common C20 precursor, geranylgeranyl diphosphate (GGDP), into bicyclic casbane. Due to their structural complexity and the wide range of diterpenoid macrocycles, CYP450 enzymes are involved in their oxygenation. By utilizing a custom database based on relevant journals [14, 15], we identified several predicted genes associated with the diTPS (diterpene synthase) and CYP450 pathways involved in macrocyclic diterpenoid synthesis. However, despite the high similarity observed through phylogenetic analysis of closely related species, it is important to note that these genes may possess different functions [51].

Further biochemical approaches, such as metabolomics, are still required to gain a comprehensive understanding. Our study identified 11 diTPS and 9 CYP450 proteins in *C. longa*. Notably, the expression patterns of diTPS genes, specifically TRINITY_DN5937_c0_g1_i1, TRINITY_DN66695_c0_g1_i3, and TRINITY_DN11310_c0_g1_i1, exhibited high expression levels in the rhizome. This indicates their close similarity to casbane synthase, their accumulation in the rhizome, and their crucial role in the oxygenation of casbane during the diverse diterpenoid biosynthesis process.

Conclusion

Transcriptomics provides valuable insights into the biosynthesis process by sequencing transcriptomic data. These data allow us to hypothesize the active genes involved in diterpenoid biosynthesis. We can identify several genes associated with diterpenoid biosynthesis by referring to existing literature. Through a homology search, we identified a total of 11 diTPS proteins and 9 CYP450 proteins. Determining the expression levels of these genes is crucial in calculating their FPKM values, enabling a comparison of diterpenoid biosynthesis between the two tissues. This information is essential for breeders and industries aiming to enhance diterpenoid biosynthesis through genetic engineering of these specific genes, harnessing the potential of synthetic biology.

References

1. Chakraborty A, Mahajan S, Jaiswal SK, Sharma VK (2021) Genome sequencing of turmeric provides evolutionary insights into its medicinal properties. Communications Biology 4 (1): 1193. doi: 10.1038/s42003-021-02720-y.
2. Zahariah S, Winarsih S, Baktiyani SCW et al. (2017) The Effect of Turmeric Decoctum to the Angiogenic Molecules Expression on Chicken Embryo. Journal of Tropical Life Science 7 (1): 61–65. doi: 10.11594/jtls.07.01.10.
3. Suprihatin T, Widyarti S, Rifa'i M, Rahayu S (2019) Malondialdehyde (MDA) Ovary and Estradiol Blood Serum Levels of Premenopause White Rat (*Rattus norvegicus*) after Turmeric Powder (*Curcuma longa* L.) Treatment. Journal of Tropical Life Science 9 (3): 237–242.
4. Changxing L, Galani S, Hassan F, et al. (2020) Biotechnological approaches to the production of plant-derived promising anticancer agents: An update and overview. Biomedicine & Pharmacotherapy 132 110918. doi: 10.1016/j.biopha.2020.110918.
5. Imaningsih W, Ekowati N, Salamiah et al. (2023) An Overview of Stress-Tolerant Promoting Endophytic Fungal Isolates from Hiyung Chilli Grown in South Kalimantan: Endophytic Fungi of Hiyung Chilli Peppers. Journal of Tropical Life Science 13 (1): 183–192. doi: 10.11594/jtls.13.01.18.
6. Annadurai RS, Neethiraj R, Jayakumar V et al. (2013) De Novo Transcriptome Assembly (NGS) of *Curcuma longa* L. Rhizome Reveals Novel Transcripts Related to Anticancer and Antimalarial Terpenoids. PLOS ONE 8 (2): e56217. doi: 10.1371/journal.pone.0056217.
7. Chen R, Wei Q, Liu Y et al. (2021) Transcriptome sequencing and functional characterization of new sesquiterpene synthases from *Curcuma wenyujin*. Archives of Biochemistry and Biophysics 709 108986. doi: 10.1016/j.abb.2021.108986.
8. Chen R, Hu T, Wang M et al. (2022) Functional characterization of key polyketide synthases by integrated metabolome and transcriptome analysis on curcuminoid biosynthesis in *Curcuma wenyujin*. Synthetic and Systems Biotechnology 7 (3): 849–861. doi: 10.1016/j.synbio.2022.04.006.
9. Yin Y, Xie X, Zhou L, et al. (2022) A chromosome-scale genome assembly of turmeric provides insights into curcumin biosynthesis and tuber formation mechanism. Frontiers in Plant Science 13.
10. Dong Q, Zou Q-C, Mao L-H, et al. (2022) The Chromosome-Scale Assembly of the *Curcuma alismatifolia* Genome Provides Insight Into Anthocyanin and Terpenoid Biosynthesis. Frontiers in Plant Science 13.
11. Gururani S, Gairola K, Kumar R, et al. (2022) Altitudinal and geographical variations in phytochemical composition and biological activities of *Curcuma longa* accession from Uttarakhand, the Himalayan region. Journal of Food

Processing and Preservation 46 (3): e16384. doi: 10.1111/jfpp.16384.
12. Cao X, Zhang F, Yuan B, et al. (2017) De novo transcriptome sequencing and analysis of *Euphorbia pekinensis* Rupr. and identification of genes involved in diterpenoid biosynthesis. Plant Gene 12 33–42. doi: 10.1016/j.plgene.2017.07.001.
13. Zerbe P, Chiang A, Dullat H, et al. (2014) Diterpene synthases of the biosynthetic system of medicinally active diterpenoids in *Marrubium vulgare*. The Plant Journal 79 (6): 914–927. doi: 10.1111/tpj.12589.
14. Inabuy FS, Fischedick JT, Lange I, et al. (2017) Biosynthesis of Diterpenoids in *Tripterygium* Adventitious Root Cultures. Plant Physiology 175 (1): 92–103. doi: 10.1104/pp.17.00659.
15. Trikka FA, Nikolaidis A, Ignea C, et al. (2015) Combined metabolome and transcriptome profiling provides new insights into diterpene biosynthesis in *S. pomifera* glandular trichomes. BMC Genomics 16 (1): 935. doi: 10.1186/s12864-015-2147-3.
16. Rangan P, Furtado A, Henry R (2020) Transcriptome profiling of wheat genotypes under heat stress during grain-filling. Journal of Cereal Science 91:102895.
17. Jin J, Sun Y, Qu J, et al. (2017) Transcriptome and functional analysis reveals hybrid vigor for oil biosynthesis in oil palm. Scientific Reports 7 (1): 439. doi: 10.1038/s41598-017-00438-8.
18. Chu HD, Tran YTH, Pham CT, et al. (2024) Identification of Two Enzymes for Trehalose Synthesis and Their Potential Function in Growth and Development in Peanut (*Arachis hypogaea*): Genome Analysis of the Encoding Trehalose Synthesis Enzymes in Peanut. Journal of Tropical Life Science 14 (1): 83–94. doi: 10.11594/jtls.14.01.10.
19. Matasci N, Hung L-H, Yan Z, et al. (2014) Data access for the 1,000 Plants (1KP) project. Gigascience 3 (1): 2047–217X.
20. Wu X, Wafula E, Yan Z, et al. (2014) Data access for the 1,000 Plants (1KP) project.
21. Afgan E, Baker D, Batut B, et al. (2018) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Research 46 (W1): W537–W544. doi: 10.1093/nar/gky379.
22. Wickham H, Averick M, Bryan J, et al. (2019) Welcome to the Tidyverse. Journal of Open Source Software 4 (43): 1686. doi: 10.21105/joss.01686.
23. Bretaudeau A, Corguillé GL, Corre E, Liu X (2023) De novo transcriptome assembly, annotation, and differential expression analysis. https://training.galaxyproject.org/training-material/topics/transcriptomics/tutorials/full-de-novo/tutorial.html. Accessed: December 2023.
24. Andrews S (2010) FastQC: a quality control tool for high throughput sequence data.
25. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30 (15): 2114–2120.
26. Grabherr MG, Haas BJ, Yassour M, et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature biotechnology 29 (7): 644–652.
27. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nature methods 9 (4): 357–359.
28. Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12 (1): 323. doi: 10.1186/1471-2105-12-323.
29. Cock PJA, Chilton JM, Grüning B, et al. (2015) NCBI BLAST+ integrated into Galaxy. Gigascience 4 (1): s13742-015.
30. Apweiler R, Bairoch A, Wu CH (2004) Protein sequence databases. Current opinion in chemical biology 8 (1): 76–80.
31. Conesa A, Götz S, García-Gómez JM, et al. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21 (18): 3674–3676.
32. Ashburner M, Ball CA, Blake JA et al. (2000) Gene ontology: tool for the unification of biology. Nature genetics 25 (1): 25–29.
33. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15 (12): 550. doi: 10.1186/s13059-014-0550-8.
34. Kolde R, Kolde MR (2018) Package 'pheatmap.' R Package 1
35. Blighe K, Rana S, Lewis M (2019) EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. R Package Version 1
36. Sievers F, Higgins DG (2014) Clustal omega, accurate alignment of very large numbers of sequences. Methods in Molecular Biology. doi: 10.1007/978-1-62703-646-7_6
37. Tamura K, Stecher G, Kumar S (2021) MEGA11: molecular evolutionary genetics analysis version 11. Molecular biology and evolution 38 (7): 3022–3027.
38. Davis EM, Croteau R (2000) Cyclization Enzymes in the Biosynthesis of Monoterpenes, Sesquiterpenes, and Diterpenes. In: Leeper FJ, Vederas JC eds Biosynthesis: Aromatic Polyketides, Isoprenoids, Alkaloids. Berlin, Heidelberg, Springer. pp 53–95.
39. Feng M, Chen C, Qu-Bie J, et al. (2022) Metabolome and transcriptome associated analysis of sesquiterpenoid metabolism in *Nardostachys jatamansi*. Frontiers in Plant Science 13. doi: 10.3389/fpls.2022.1041321.
40. Ayer DK, Modha K, Parekh V, et al. (2020) Associating gene expressions with curcuminoid biosynthesis in turmeric. Journal of Genetic Engineering and Biotechnology 18 (1): 83. doi: 10.1186/s43141-020-00101-2.
41. Zhu F, Sun H, Wang J et al. (2022) Differential expression involved in starch synthesis pathway genes reveal various starch characteristics of seed and rhizome in lotus (*Nelumbo Nucifera*). Journal of Food Science 87 (9): 4250–4263. doi: 10.1111/1750-3841.16283.
42. Chen L, Xu S, Liu Y et al. (2022) Identification of key gene networks controlling polysaccharide accumulation in different tissues of *Polygonatum cyrtonema* Hua by integrating metabolic phenotypes and gene expression profiles. Frontiers in Plant Science 13. doi: 10.3389/fpls.2022.1012231.
43. Barrero RA, Chapman B, Yang Y, et al. (2011) De novo assembly of *Euphorbia fischeriana* root transcriptome identifies prostratin pathway related genes. BMC Genomics 12 (1): 600. doi: 10.1186/1471-2164-12-600.
44. Pandit S, Shitiz K, Sood H, et al. (2013) Expression pattern of fifteen genes of non-mevalonate (MEP) and mevalonate (MVA) pathways in different tissues of

endangered medicinal herb *Picrorhiza kurroa* with respect to picrosides content. Molecular Biology Reports 40 (2): 1053–1063. doi: 10.1007/s11033-012-2147-1.
45. He M, Wang Y, Hua W, et al. (2012) De Novo Sequencing of *Hypericum perforatum* Transcriptome to Identify Potential Genes Involved in the Biosynthesis of Active Metabolites. PLoS ONE 7 (7): e42081. doi: 10.1371/journal.pone.0042081.
46. Davis EM, Croteau R (2000) Cyclization Enzymes in the Biosynthesis of Monoterpenes, Sesquiterpenes, and Diterpenes. In: Leeper FJ, Vederas JC eds Biosynthesis. Berlin, Heidelberg, Springer Berlin Heidelberg. pp 53–95.
47. Wang K, Yu H, Wu H, et al. (2015) A new casbane diterpene from *Euphorbia pekinensis*. Natural Product Research 29 (15): 1456–1460. doi: 10.1080/14786419.2015.1027704.
48. Mau CJ, West CA (1994) Cloning of casbene synthase cDNA: evidence for conserved structural features among terpenoid cyclases in plants. Proceedings of the National Academy of Sciences 91 (18): 8497–8501. doi: 10.1073/pnas.91.18.8497.
49. Kirby J, Nishimoto M, Park JG, et al. (2010) Cloning of casbene and neocembrene synthases from Euphorbiaceae plants and expression in *Saccharomyces cerevisiae*. Phytochemistry 71 (13): 1466–1473. doi: 10.1016/j.phytochem.2010.06.001.
50. Luo D, Callari R, Hamberger B, et al. (2016) Oxidation and cyclization of casbene in the biosynthesis of *Euphorbia* factors from mature seeds of *Euphorbia lathyris* L. Proc Natl Acad Sci USA. doi: 10.1073/pnas.1607504113
51. Zerbe P, Bohlmann J (2015) Plant diterpene synthases: exploring modularity and metabolic diversity for bioengineering. Trends in Biotechnology 33 (7): 419–428. doi: 10.1016/j.tibtech.2015.04.006.

Table 1. Top twenty differential expressed genes in leaf against rhizome of *Curcuma longa*

TRINITY_DN1_c4_g1_i6	-21,4981	1,65E-11	AMM06269.1'photosystem I P700 apoprotein A1	P:GO:0015979; P:GO:0022900; F:GO:0000287; F:GO:0009055; F:GO:0016168; F:GO:0051539; C:GO:0009522; C:GO:0009535	P:photosynthesis; P:electron transport chain; F:magnesium ion binding; F:electron transfer activity; F:chlorophyll binding; F:4 iron, 4 sulfur cluster binding; C:photosystem I; C:chloroplast thylakoid membrane
TRINITY_DN7_c0_g2_i3	-21,0857	9,1E-10	QDA80787.1photosystem II CP43 chlorophyll apoprotein	P:GO:0009772; F:GO:0016168; F:GO:0045156; F:GO:0046872; C:GO:0009523; C:GO:0009535	P:photosynthetic electron transport in photosystem II; F:chlorophyll binding; F:electron transporter, transferring electrons within the cyclic electron transport pathway of photosynthesis activity; F:metal ion binding; C:photosystem II; C:chloroplast thylakoid membrane
TRINITY_DN1_c52_g1_i2	-19,263	8,4E-09	QIQ34539.1ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit	P:GO:0009853; P:GO:0019253; F:GO:0000287; F:GO:0004497; F:GO:0016984; C:GO:0009507	P:photorespiration; P:reductive pentose-phosphate cycle; F:magnesium ion binding; F:monooxygenase activity; F:ribulose-bisphosphate carboxylase activity; C:chloroplast
TRINITY_DN241080_c58_g1_i3	-19,0845	1,99E-07	AES88250.150S ribosomal protein L33	P:GO:0006412; F:GO:0003735; F:GO:0019843; C:GO:0005840; C:GO:0009507; C:GO:1990904	P:translation; F:structural constituent of ribosome; F:rRNA binding; C:ribosome; C:chloroplast; C:ribonucleoprotein complex

TRINITY_DN1_c4_g1_i4	-18,8543	1E-11	AMM06269.1'photosystem I P700 apoprotein A1	P:GO:0015979; P:GO:0022900; F:GO:0000287; F:GO:0009055; F:GO:0016168; F:GO:0051539; C:GO:0009522; C:GO:0009535	P:photosynthesis; P:electron transport chain; F:magnesium ion binding; F:electron transfer activity; F:chlorophyll binding; F:4 iron, 4 sulfur cluster binding; C:photosystem I; C:chloroplast thylakoid membrane
TRINITY_DN587_c0_g1_i1	-18,0141	7,37E-14	XP_009390221.1 ribulose bisphosphate carboxylase small chain, chloroplastic	P:GO:0009853; P:GO:0019253; F:GO:0016984; C:GO:0009507	P:photorespiration; P:reductive pentose-phosphate cycle; F:ribulose-bisphosphate carboxylase activity; C:chloroplast
TRINITY_DN587_c0_g1_i5	-18,0063	2,58E-08	XP_009390221.1 ribulose bisphosphate carboxylase small chain, chloroplastic	P:GO:0009853; P:GO:0019253; F:GO:0016984; C:GO:0009507	P:photorespiration; P:reductive pentose-phosphate cycle; F:ribulose-bisphosphate carboxylase activity; C:chloroplast
TRINITY_DN1_c14_g1_i1	-17,678	1,03E-11	ANP25791.1photosystem II protein D1	P:GO:0009635; P:GO:0009772; F:GO:0005506; F:GO:0010242; F:GO:0016168; F:GO:0016682; F:GO:0045156; C:GO:0009523; C:GO:0009535	P:response to herbicide; P:photosynthetic electron transport in photosystem II; F:iron ion binding; F:oxygen evolving activity; F:chlorophyll binding; F:oxidoreductase activity, acting on diphenols and related substances as donors, oxygen as acceptor; F:electron transporter, transferring electrons within the cyclic electron transport pathway of photosynthesis activity; C:photosystem II; C:chloroplast thylakoid membrane

TRINITY_DN9_c1_g1_i22	14,71221	9,61E-08	A5HMM7.2RecName: Full=Mannose-specific lectin 2; AltName: Full=Agglutinin; Contains: RecName: Full=Mannose-specific lectin 2 chain 1; Contains: RecName: Full=Mannose-specific lectin 2 chain 2; Flags: Precursor	F:GO:0005537; F:GO:0030246; F:GO:0046872; C:GO:0005576; C:GO:0016020; C:GO:0016020	F:mannose binding; F:carbohydrate binding; F:metal ion binding; C:extracellular region; C:membrane; C:membrane
TRINITY_DN9_c0_g1_i28	13,90045	1,24E-06	KAF5202293.1Mannose-binding lectin	F:GO:0005537; F:GO:0030246; C:GO:0005576; C:GO:0016020; C:GO:0016020	F:mannose binding; F:carbohydrate binding; C:extracellular region; C:membrane; C:membrane
TRINITY_DN9_c0_g1_i27	13,55091	4,58E-06	BAD67183.1mannose specific lectin	F:GO:0030246; C:GO:0016020	F:carbohydrate binding; C:membrane
TRINITY_DN9_c1_g1_i2	13,31432	3,85E-06	XP_020571214.1mannose-specific lectin-like	F:GO:0030246; C:GO:0016020	F:carbohydrate binding; C:membrane
TRINITY_DN75_c0_g1_i19	13,14408	1,38E-07	XP_009398364.1 YTH domain-containing family protein 1-like isoform X1	F:GO:0003723	F:RNA binding
TRINITY_DN29_c0_g1_i6	12,65661	7,53E-08	XP_010937631.1metallothionein-like protein type 2	F:GO:0046872	F:metal ion binding
TRINITY_DN1132_c0_g1_i3	12,63024	0,000163	XP_009390266.1 patatin-like protein 2 isoform X2	P:GO:0016042; F:GO:0016298	P:lipid catabolic process; F:lipase activity
TRINITY_DN29_c0_g1_i2	12,36979	0,001284	XP_010937631.1metallothionein-like protein type 2	P:GO:0006412; F:GO:0046872; C:GO:0005840; C:GO:1990904	P:translation; F:metal ion binding; C:ribosome; C:ribonucleoprotein complex

TRINITY_DN2_c1_g1_i9	-17,5684	1,13E-06	QCE13735.1DNA-directed RNA polymerase subunit beta'	P:GO:0006351; P:GO:0046034; P:GO:1901566; F:GO:0000166; F:GO:0003677; F:GO:0003899; F:GO:0008270; F:GO:0097367; C:GO:0000428; C:GO:0009507; C:GO:0016469	P:DNA-templated transcription; P:ATP metabolic process; P:organonitrogen compound biosynthetic process; F:nucleotide binding; F:DNA binding; F:DNA-directed 5'-3' RNA polymerase activity; F:zinc ion binding; F:carbohydrate derivative binding; C:DNA-directed RNA polymerase complex; C:chloroplast; C:proton-transporting two-sector ATPase complex
TRINITY_DN2_c2_g2_i1	-17,2975	7,91E-10	NBI56447.1photosystem I core protein PsaA	P:GO:0015979; F:GO:0046872; C:GO:0005739; C:GO:0009507; C:GO:0009579; C:GO:0016020	P:photosynthesis; F:metal ion binding; C:mitochondrion; C:chloroplast; C:thylakoid; C:membrane
TRINITY_DN49_c0_g1_i19	-17,0002	3,63E-08	XP_009396492.1 chlorophyll a-b binding protein of LHCII type 1-like	P:GO:0009416; P:GO:0009768; F:GO:0016168; C:GO:0009522; C:GO:0009523; C:GO:0009535	P:response to light stimulus; P:photosynthesis, light harvesting in photosystem I; F:chlorophyll binding; C:photosystem I; C:photosystem II; C:chloroplast thylakoid membrane
TRINITY_DN587_c0_g1_i2	-16,9245	1,37E-09	XP_009390221.1 ribulose bisphosphate carboxylase small chain, chloroplastic	P:GO:0009853; P:GO:0019253; F:GO:0016984; C:GO:0009507	P:photorespiration; P:reductive pentose-phosphate cycle; F:ribulose-bisphosphate carboxylase activity; C:chloroplast

TRINITY_DN69_c0_g1_i3	-16,6855	3,55E-12	XP_009384868.1 ribulose bisphosphate carboxylase/oxygenase activase 2, chloroplastic	F:GO:0005524; F:GO:0016887; C:GO:0009570	F:ATP binding; F:ATP hydrolysis activity; C:chloroplast stroma
TRINITY_DN372_c0_g1_i2	-16,428	3,09E-06	XP_009382124.1 protochlorophyllide reductase-like	P:GO:0015979; P:GO:0015995; F:GO:0016630; C:GO:0009507	P:photosynthesis; P:chlorophyll biosynthetic process; F:protochlorophyllide reductase activity; C:chloroplast
TRINITY_DN395_c0_g1_i2	-16,3139	2,25E-10	KAG1369863.1fructose-bisphosphate aldolase, chloroplastic	P:GO:0006096; P:GO:0030388; F:GO:0004332; C:GO:0005829	P:glycolytic process; P:fructose 1,6-bisphosphate metabolic process; F:fructose-bisphosphate aldolase activity; C:cytosol
TRINITY_DN3023_c1_g2_i3	-16,2234	2,4E-06	XP_009417196.1 probable pyridoxal 5'-phosphate synthase subunit PDX1.1	P:GO:0042823; F:GO:0036381	P:pyridoxal phosphate biosynthetic process; F:pyridoxal 5'-phosphate synthase (glutamine hydrolysing) activity
TRINITY_DN237_c0_g1_i10	-15,9253	2,22E-08	XP_009388762.1 chlorophyll a-b binding protein 4, chloroplastic	P:GO:0009409; P:GO:0009644; P:GO:0009645; P:GO:0009768; F:GO:0003729; F:GO:0016168; F:GO:0019904; C:GO:0009522; C:GO:0009523; C:GO:0009535	P:response to cold; P:response to high light intensity; P:response to low light intensity stimulus; P:photosynthesis, light harvesting in photosystem I; F:mRNA binding; F:chlorophyll binding; F:protein domain specific binding; C:photosystem I; C:photosystem II; C:chloroplast thylakoid membrane

TRINITY_DN164_c0_g1_i8	-15,7915	9,48E-11	XP_009415903.1 chlorophyll a-b binding protein 7, chloroplastic-like isoform X2	P:GO:0009416; P:GO:0009768; F:GO:0016168; C:GO:0009522; C:GO:0009523; C:GO:0009535	P:response to light stimulus; P:photosynthesis, light harvesting in photosystem I; F:chlorophyll binding; C:photosystem I; C:photosystem II; C:chloroplast thylakoid membrane
TRINITY_DN33_c19_g1_i1	-15,779	6,78E-07	OTF84420.1putative photosystem antenna protein-like protein	P:GO:0009767; F:GO:0016168; C:GO:0009507; C:GO:0009523	P:photosynthetic electron transport chain; F:chlorophyll binding; C:chloroplast; C:photosystem II
TRINITY_DN542_c0_g1_i3	-15,677	3,31E-13	XP_009420999.1 glyceraldehyde-3-phosphate dehydrogenase A, chloroplastic	P:GO:0006006; F:GO:0004365; F:GO:0050661; F:GO:0051287	P:glucose metabolic process; F:glyceraldehyde-3-phosphate dehydrogenase (NAD+) (phosphorylating) activity; F:NADP binding; F:NAD binding

Supplementary 2

Table 2. Top twenty differential expressed genes in rhizome against leaf of *Curcuma longa*

Genes	log2FoldChange	padj	Description	GO.IDs	GO.Names
TRINITY_DN9_c1_g1_i24	28,88607	1,76E-13	P49329.1RecName: Full=Mannose-specific lectin; AltName: Full=Agglutinin; Contains: RecName: Full=Mannose-specific lectin heavy chain; Contains: RecName: Full=Mannose-specific lectin light chain; Flags: Precursor	F:GO:0005537; F:GO:0030246; C:GO:0005576	F:mannose binding; F:carbohydrate binding; C:extracellular region
TRINITY_DN9_c0_g1_i4	27,97115	8,69E-09	P49329.1RecName: Full=Mannose-specific lectin; AltName: Full=Agglutinin; Contains: RecName: Full=Mannose-specific lectin heavy chain; Contains: RecName: Full=Mannose-specific lectin light chain; Flags: Precursor	F:GO:0005537; F:GO:0030246; C:GO:0005576	F:mannose binding; F:carbohydrate binding; C:extracellular region
TRINITY_DN9_c1_g1_i7	25,00564	4,22E-08	A5HMM7.2RecName: Full=Mannose-specific lectin 2; AltName: Full=Agglutinin; Contains: RecName: Full=Mannose-specific lectin 2 chain 1; Contains: RecName: Full=Mannose-specific lectin 2 chain 2; Flags: Precursor	F:GO:0005537; F:GO:0030246; F:GO:0046872; C:GO:0005576	F:mannose binding; F:carbohydrate binding; F:metal ion binding; C:extracellular region
TRINITY_DN196_c0_g2_i2	24,81102	4,94E-08	XP_009382058.1 dehydrin COR410	P:GO:0009415	P:response to water

TRINITY_DN9_c1_g1_i22	14,71221	9,61E-08	A5HMM7.2RecName: Full=Mannose-specific lectin 2; AltName: Full=Agglutinin; Contains: RecName: Full=Mannose-specific lectin 2 chain 1; Contains: RecName: Full=Mannose-specific lectin 2 chain 2; Flags: Precursor	F:GO:0005537; F:GO:0030246; F:GO:0046872; C:GO:0005576; C:GO:0016020; C:GO:0016020	F:mannose binding; F:carbohydrate binding; F:metal ion binding; C:extracellular region; C:membrane; C:membrane
TRINITY_DN9_c0_g1_i28	13,90045	1,24E-06	KAF5202293.1Mannose-binding lectin	F:GO:0005537; F:GO:0030246; C:GO:0005576; C:GO:0016020; C:GO:0016020	F:mannose binding; F:carbohydrate binding; C:extracellular region; C:membrane; C:membrane
TRINITY_DN9_c0_g1_i27	13,55091	4,58E-06	BAD67183.1mannose specific lectin	F:GO:0030246; C:GO:0016020	F:carbohydrate binding; C:membrane
TRINITY_DN9_c1_g1_i2	13,31432	3,85E-06	XP_020571214.1mannose-specific lectin-like	F:GO:0030246; C:GO:0016020	F:carbohydrate binding; C:membrane
TRINITY_DN75_c0_g1_i19	13,14408	1,38E-07	XP_009398364.1 YTH domain-containing family protein 1-like isoform X1	F:GO:0003723	F:RNA binding
TRINITY_DN29_c0_g1_i6	12,65661	7,53E-08	XP_010937631.1metallothionein-like protein type 2	F:GO:0046872	F:metal ion binding
TRINITY_DN1132_c0_g1_i3	12,63024	0,000163	XP_009390266.1 patatin-like protein 2 isoform X2	P:GO:0016042; F:GO:0016298	P:lipid catabolic process; F:lipase activity
TRINITY_DN29_c0_g1_i2	12,36979	0,001284	XP_010937631.1metallothionein-like protein type 2	P:GO:0006412; F:GO:0046872; C:GO:0005840; C:GO:1990904	P:translation; F:metal ion binding; C:ribosome; C:ribonucleoprotein complex

TRINITY_DN34_c3_g1_i1	12,33947	3,48E-06	CEE24891.1hypothetical protein XAC908_1160001	C:GO:0016020	C:membrane
TRINITY_DN99_c0_g1_i8	12,33578	4,65E-08	RDX60239.1ADP-ribosylation factor 2	P:GO:0015031; P:GO:0016192; F:GO:0003924; F:GO:0005525; C:GO:0005794	P:protein transport; P:vesicle-mediated transport; F:GTPase activity; F:GTP binding; C:Golgi apparatus
TRINITY_DN9_c0_g1_i8	12,10239	0,000179	KAF5202293.1Mannose-binding lectin	F:GO:0005537; F:GO:0030246; C:GO:0005576; C:GO:0016020; C:GO:0016020	F:mannose binding; F:carbohydrate binding; C:extracellular region; C:membrane; C:membrane
TRINITY_DN565_c0_g1_i2	11,96799	6,91E-05	XP_009408264.1 polyphenol oxidase, chloroplastic-like	P:GO:0046148; F:GO:0004097; F:GO:0046872	P:pigment biosynthetic process; F:catechol oxidase activity; F:metal ion binding
TRINITY_DN9_c0_g1_i15	11,9383	0,001915	AKT75734.1mannose-binding lectin precursor	P:GO:0009615; F:GO:0005537	P:response to virus; F:mannose binding
TRINITY_DN2516_c0_g1_i9	11,9194	1,17E-07	XP_009388575.1 protein NETWORKED 1A	F:GO:0003779; C:GO:0110165	F:actin binding; C:cellular anatomical entity
TRINITY_DN586_c0_g1_i10	11,79934	0,000585	XP_009405950.1 sulfate transporter 3.1	P:GO:1902358; F:GO:0008271; C:GO:0009507; C:GO:0016020	P:sulfate transmembrane transport; F:secondary active sulfate transmembrane transporter activity; C:chloroplast; C:membrane

TRINITY_DN234_c0_g1_i7	11,75383	1,7E-05	XP_009387625.1 NADP-dependent malic enzyme	P:GO:0006090; P:GO:0006108; F:GO:0004471; F:GO:0004473; F:GO:0046872; F:GO:0051287; C:GO:0009507	P:pyruvate metabolic process; P:malate metabolic process; F:malate dehydrogenase (decarboxylating) (NAD+) activity; F:malate dehydrogenase (decarboxylating) (NADP+) activity; F:metal ion binding; F:NAD binding; C:chloroplast

Supplementary 3

Table 3. Top ten differential expressed biosynthetic genes in rhizome against leaf of *Curcuma longa*

Genes	log2FoldChange	padj	Description	GO.IDs	GO.Names
TRINITY_DN565_c0_g1_i2	11,96799	6,91E-05	XP_009408264.1 polyphenol oxidase, chloroplastic-like	P:GO:0046148; F:GO:0004097; F:GO:0046872	P:pigment biosynthetic process; F:catechol oxidase activity; F:metal ion binding
TRINITY_DN417_c0_g1_i20	11,40198	0,002223	XP_009407551.1 probable linoleate 9S-lipoxygenase 5	P:GO:0006633; P:GO:0031408; P:GO:0034440; F:GO:0016702; F:GO:0046872	P:fatty acid biosynthetic process; P:oxylipin biosynthetic process; P:lipid oxidation; F:oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen; F:metal ion binding
TRINITY_DN565_c0_g2_i7	11,09088	1,38E-05	KAG1328177.1polyphenol oxidase, chloroplastic	P:GO:0046148; F:GO:0004097; F:GO:0046872	P:pigment biosynthetic process; F:catechol oxidase activity; F:metal ion binding
TRINITY_DN25787_c1_g1_i2	11,06851	0,000116	XP_009405800.1 soluble starch synthase 2-2, chloroplastic/amyloplastic-like	P:GO:0019252; F:GO:0004373; C:GO:0009501; C:GO:0009507	P:starch biosynthetic process; F:glycogen (starch) synthase activity; C:amyloplast; C:chloroplast
TRINITY_DN8056_c0_g1_i2	11,03414	0,000159	XP_009405800.1 soluble starch synthase 2-2, chloroplastic/amyloplastic-like	P:GO:0019252; F:GO:0004373; C:GO:0009501; C:GO:0009507	P:starch biosynthetic process; F:glycogen (starch) synthase activity; C:amyloplast; C:chloroplast
TRINITY_DN207_c0_g1_i3	10,57649	0,002744	KAG1354106.1trans-resveratrol di-O-methyltransferase	P:GO:0019438; P:GO:0032259; F:GO:0008171; F:GO:0008757	P:aromatic compound biosynthetic process; P:methylation; F:O-methyltransferase activity; F:S-adenosylmethionine-dependent methyltransferase activity

TRINITY_DN2480_c0_g1_i5	10,41047	0,00563	XP_018686191.1 threonine synthase, chloroplastic-like	P:GO:0009088; F:GO:0004795; F:GO:0030170; C:GO:0016020	P:threonine biosynthetic process; F:threonine synthase activity; F:pyridoxal phosphate binding; C:membrane
TRINITY_DN11571_c0_g1_i8	10,35322	1,18E-05	XP_009400322.1 callose synthase 9	P:GO:0006075; P:GO:0008360; P:GO:0071555; F:GO:0003843; C:GO:0000148	P:(1->3)-beta-D-glucan biosynthetic process; P:regulation of cell shape; P:cell wall organization; F:1,3-beta-D-glucan synthase activity; C:1,3-beta-D-glucan synthase complex
TRINITY_DN2896_c0_g1_i1	10,34368	0,000401	XP_009408489.1 FRIGIDA-like protein 4a	P:GO:0009086; P:GO:0009908; P:GO:0030154; P:GO:0032259; F:GO:0003871; F:GO:0008270	P:methionine biosynthetic process; P:flower development; P:cell differentiation; P:methylation; F:5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase activity; F:zinc ion binding
TRINITY_DN119790_c0_g1_i2	10,21852	0,000209	XP_009416548.1 glucose-1-phosphate adenylyltransferase large subunit 1-like	P:GO:0005978; P:GO:0019252; F:GO:0005524; F:GO:0008878; C:GO:0009507	P:glycogen biosynthetic process; P:starch biosynthetic process; F:ATP binding; F:glucose-1-phosphate adenylyltransferase activity; C:chloroplast

Supplementary 4

Table 4. Top ten differential expressed biosynthetic genes in leaf against rhizome of *Curcuma longa*

Genes	log2Fold-Change	padj	Description	GO.IDs	GO.Names
TRINITY_DN2_c1_g1_i9	-17,5684	1,13E-06	QCE13735.1DNA-directed RNA polymerase subunit beta'	P:GO:0006351; P:GO:0046034; P:GO:1901566; F:GO:0000166; F:GO:0003677; F:GO:0003899; F:GO:0008270; F:GO:0097367; C:GO:0000428; C:GO:0009507; C:GO:0016469	P:DNA-templated transcription; P:ATP metabolic process; P:organonitrogen compound biosynthetic process; F:nucleotide binding; F:DNA binding; F:DNA-directed 5'-3' RNA polymerase activity; F:zinc ion binding; F:carbohydrate derivative binding; C:DNA-directed RNA polymerase complex; C:chloroplast; C:proton-transporting two-sector ATPase complex
TRINITY_DN372_c0_g1_i2	-16,428	3,09E-06	XP_009382124.1 protochlorophyllide reductase-like	P:GO:0015979; P:GO:0015995; F:GO:0016630; C:GO:0009507	P:photosynthesis; P:chlorophyll biosynthetic process; F:protochlorophyllide reductase activity; C:chloroplast
TRINITY_DN3023_c1_g2_i3	-16,2234	2,4E-06	XP_009417196.1 probable pyridoxal 5'-phosphate synthase subunit PDX1.1	P:GO:0042823; F:GO:0036381	P:pyridoxal phosphate biosynthetic process; F:pyridoxal 5'-phosphate synthase (glutamine hydrolysing) activity

TRINITY_DN2_c1_g1_i22	-14,9105	3,82E-05	QCE13735.1DNA-directed RNA polymerase subunit beta'	P:GO:0034654; P:GO:1901564; F:GO:0003899; F:GO:0036094; F:GO:0046872; F:GO:0097159; F:GO:0097367; F:GO:1901363; C:GO:0009536; C:GO:0016020; C:GO:0043232; C:GO:1902494	P:nucleobase-containing compound biosynthetic process; P:organonitrogen compound metabolic process; F:DNA-directed 5'-3' RNA polymerase activity; F:small molecule binding; F:metal ion binding; F:organic cyclic compound binding; F:carbohydrate derivative binding; F:heterocyclic compound binding; C:plastid; C:membrane; C:intracellular non-membrane-bounded organelle; C:catalytic complex
TRINITY_DN609_c0_g1_i5	-14,3074	3,1E-07	XP_009398831.1 sedoheptulose-1,7-bisphosphatase, chloroplastic	P:GO:0005986; P:GO:0006000; P:GO:0006002; P:GO:0006094; P:GO:0019252; P:GO:0030388; F:GO:0042132; F:GO:0046872; F:GO:0050278; C:GO:0005737	P:sucrose biosynthetic process; P:fructose metabolic process; P:fructose 6-phosphate metabolic process; P:gluconeogenesis; P:starch biosynthetic process; P:fructose 1,6-bisphosphate metabolic process; F:fructose 1,6-bisphosphate 1-phosphatase activity; F:metal ion binding; F:sedoheptulose-bisphosphatase activity; C:cytoplasm
TRINITY_DN2211_c0_g1_i5	-13,832	4,19E-09	XP_009391592.1 glutamine synthetase leaf isozyme, chloroplastic-like	P:GO:0006542; F:GO:0004356; F:GO:0005524; C:GO:0005737; C:GO:0016020	P:glutamine biosynthetic process; F:glutamate-ammonia ligase activity; F:ATP binding; C:cytoplasm; C:membrane

TRINITY_DN609_c0_g1_i4	-13,6679	1,75E-07	XP_009398831.1 sedoheptulose-1,7-bisphosphatase, chloroplastic	P:GO:0005986; P:GO:0006000; P:GO:0006002; P:GO:0006094; P:GO:0019252; P:GO:0030388; F:GO:0042132; F:GO:0046872; F:GO:0050278; C:GO:0005737	P:sucrose biosynthetic process; P:fructose metabolic process; P:fructose 6-phosphate metabolic process; P:gluconeogenesis; P:starch biosynthetic process; P:fructose 1,6-bisphosphate metabolic process; F:fructose 1,6-bisphosphate 1-phosphatase activity; F:metal ion binding; F:sedoheptulose-bisphosphatase activity; C:cytoplasm
TRINITY_DN459_c0_g1_i6	-13,5424	7,13E-06	XP_009416289.1 omega-3 fatty acid desaturase, chloroplastic	P:GO:0006096; P:GO:0006119; P:GO:0006636; P:GO:0009088; P:GO:0009229; P:GO:0009234; P:GO:0036108; F:GO:0016717; C:GO:0016020	P:glycolytic process; P:oxidative phosphorylation; P:unsaturated fatty acid biosynthetic process; P:threonine biosynthetic process; P:thiamine diphosphate biosynthetic process; P:menaquinone biosynthetic process; P:4-amino-4-deoxy-alpha-L-arabinopyranosyl undecaprenyl phosphate biosynthetic process; F:oxidoreductase activity, acting on paired donors, with oxidation of a pair of donors resulting in the reduction of molecular oxygen to two molecules of water; C:membrane
TRINITY_DN3524_c0_g1_i11	-13,2403	9,55E-07	XP_009383479.1 magnesium-chelatase subunit ChlI, chloroplastic isoform X1	P:GO:0015979; P:GO:0015995; F:GO:0005524; F:GO:0016851; C:GO:0009570	P:photosynthesis; P:chlorophyll biosynthetic process; F:ATP binding; F:magnesium chelatase activity; C:chloroplast stroma

TRINITY_DN582_c0_g1_i2	-13,1722	9,04E-08	XP_010906115.1photosystem II 5 kDa protein, chloroplastic	P:GO:0006412; P:GO:0010468; P:GO:0010556; P:GO:0031326; P:GO:0051171; P:GO:0080090; F:GO:0003676; C:GO:0005634; C:GO:0016020	P:translation; P:regulation of gene expression; P:regulation of macromolecule biosynthetic process; P:regulation of cellular biosynthetic process; P:regulation of nitrogen compound metabolic process; P:regulation of primary metabolic process; F:nucleic acid binding; C:nucleus; C:membrane