MEDDOCS Open Access Publisher

ISSN: 2637-7721

Journal of Plant Biology and Crop Research

Open Access | Research Article

Species-Associated Transcriptional Divergence in *Phoenix theophrasti* and *Phoenix pusilla*

Vaishali M Gopala[#]; Madhugiri Nageswara-Rao[#]*

Subtropical Horticultural Research Station, USDA-ARS, USA.

[#]These authors contributed equally to this article.

***Corresponding Author(s):** Madhugiri Nageswara-Rao

Subtropical Horticultural Research Station, USDA-ARS, 13601 Old Cutler Road, Miami, FL 33158, USA.

Tel: 786-573-7050;

Email: Madhugiri.Nageswara-Rao@usda.gov; mnrbhav@gmail.com

Received: Feb 05, 2026

Accepted: Mar 02, 2026

Published Online: Mar 09, 2026

Journal: Journal of Plant Biology and Crop Research

Publisher: MedDocs Publishers LLC

Online edition: http://meddocsonline.org/

Copyright: © Nageswara-Rao M (2026). *This Article is distributed under the terms of Creative Commons Attribution 4.0 International License*

Keywords: Differential genes; Molecular characterization; *Phoenix* species; Reference genome; Transcriptome.

Abstract

The genus *Phoenix* comprises ecologically important and diverse palm species, yet the molecular basis underlying functional divergence among wild *Phoenix* species remains poorly understood. In this study, we conducted a comparative transcriptome analysis of *Phoenix theophrasti* and *P. pusilla* to investigate species-associated transcriptional differences. RNA-seq approach was followed in this experiment, and cleaned reads were aligned to the *P. dactylifera* reference genome. Overall alignment rates were high and comparable between species, ranging from 84% to 85%. The number of expressed genes detected was also similar between species, with 26,000–27,000 genes expressed per library. Differential gene expression analysis identified extensive transcriptional divergence between the two species, with over 11,000 genes exhibiting significant differential expression. Hierarchical clustering of highly divergent genes revealed clear separation of samples by species, indicating consistent species-level expression patterns. Functional enrichment analyses highlighted significant differences in biological processes and pathways related to translation, stress signaling, redox regulation, defense responses, photosynthesis, lipid and cuticle metabolism, and secondary metabolism. KEGG pathway enrichment further supported divergence in phenylpropanoid biosynthesis, plant–pathogen interaction, MAPK signaling, and ribosome-associated pathways. Trait-focused curation of differentially expressed genes identified candidate genes associated with abiotic stress responses, oxidative stress regulation, immune signaling, photosynthetic machinery, and metabolic specialization. Collectively, these results demonstrate coordinated transcriptional differentiation between *P. theophrasti* and *P. pusilla* and provide novel insights into molecular processes underlying functional divergence within the genus *Phoenix*. This study establishes a foundation for future investigations into palm adaptation, physiology, and evolution.

Cite this article: Gopala VM, Nageswara-Rao M. Species-Associated Transcriptional Divergence in *Phoenix theophrasti* and *Phoenix pusilla*. J Plant Biol Crop Res. 2026; 10(1): 1114.

Introduction

The genus *Phoenix* (Arecaceae) comprises economically, ecologically, and evolutionarily significant palm species that exhibit remarkable diversity in growth form, ecological tolerance, and geographic distribution. While *Phoenix dactylifera* (date palm) has been extensively studied at the genomic [2] and transcriptomic levels [42], comparatively little is known about the molecular basis underlying functional divergence among wild and semi-wild *Phoenix* species. Understanding transcriptional variation among closely related species provides valuable insights into the genetic mechanisms associated with ecological adaptation, physiological specialization, and trait evolution [40].

Comparative transcriptome analysis using RNA sequencing (RNA-seq) has emerged as a powerful approach to investigate gene expression divergence across species, particularly in non-model plants where complete genome sequences may be unavailable [8, 30]. Reference-guided alignment strategies using high-quality genomes from related species have been successfully employed to study transcriptional variation in wild relatives of crop plants, enabling robust differential expression and functional annotation analyses [4,16]. Such approaches have been widely applied to study leaf transcriptomes, given the central role of leaves in photosynthesis, stress perception, and metabolic regulation [11,39].

Leaves are particularly informative tissues for comparative studies because they integrate developmental, physiological, and environmental signals. Comparative leaf transcriptomics has been used to uncover adaptive gene expression patterns related to photosynthesis efficiency, secondary metabolism, water-use efficiency, and abiotic stress tolerance across diverse plant lineages [9,26]. Differences in leaf gene expression between species often reflect ecological specialization and evolutionary divergence, even when overall genomic similarity remains high [34].

P. theophrasti and *P. pusilla* represent two distinct evolutionary lineages within the genus, differing in geographic distribution, ecological preferences, and morphological traits. While *P. theophrasti* is restricted to the eastern Mediterranean region, *P. pusilla* is adapted to semi-arid and coastal environments of South Asia, where it persists under contrasting climatic and edaphic conditions. Despite their close phylogenetic relationship, the molecular mechanisms underlying their ecological differentiation remain largely unexplored [6].

In this study, we performed a comparative transcriptome analysis of mature leaf tissues from *P. theophrasti* and *P. pusilla* using a reference-guided RNA-seq framework based on the *P. dactylifera* genome. Differential gene expression analysis, Gene Ontology (GO) enrichment, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were conducted to identify species-specific transcriptional signatures. Furthermore, we curated differentially expressed genes associated with key functional traits to highlight biologically meaningful divergence between the two species. Our results provide novel insights into transcriptional differentiation within wild *Phoenix* species and contribute to a broader understanding of functional evolution in palms.

Materials and methods

Plant material and RNA isolation

Mature leaf tissues were collected from plants of *P. theophrasti* and *P. pusilla* growing under field conditions at the USDA-Subtropical Horticulture Research Station (SHRS), Miami, Florida (latitude N25°38'33.76" and longitude 80°17'37.86). For each species, leaves from three individuals were harvested to capture representative transcriptional profiles and pooled before RNA extraction. Total RNA was extracted from plant tissue using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) in accordance with the manufacturer's recommended protocol to ensure high-quality RNA recovery. The concentration and purity of the isolated RNA were initially evaluated using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, MA, USA), which measures absorbance ratios to assess contamination. Further assessment of RNA integrity was performed using an Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA), providing an RNA Integrity Number (RIN) as a standardized measure of RNA quality. Only samples exhibiting a RIN value greater than 7.0, indicative of intact and high-quality RNA, were selected for subsequent library preparation and sequencing steps.

RNA sequencing and quality control

RNA sequencing libraries were constructed with the Kapa Hyper Stranded mRNA library kit (Roche) following the manufacturer's guidelines to preserve strand specificity. The process involved enrichment of poly(A)-tailed mRNA, fragmentation into smaller pieces, and reverse transcription into complementary DNA (cDNA). After cDNA synthesis, sequencing adapters were ligated to the fragments, and the libraries were amplified to generate sufficient material for sequencing. High-quality libraries were then subjected to next-generation sequencing (NGS) at the University of Illinois Roy J. Carver Biotechnology Center. Sequencing was performed on the Illumina NovaSeq 6000 platform (Illumina, CA, USA) using the NovaSeq SP reagent kit, producing paired-end reads of 150 base pairs (bp), which provide comprehensive coverage and improved accuracy for downstream transcriptomic analyses. Raw sequencing reads were subjected to quality assessment using FastQC [3]. Adapter sequences and low-quality bases were removed using Trimmomatic, retaining reads with high base quality for downstream analyses [5]. All sequence reads were deposited in the public NCBI Sequence Read Archive database under BioSample accessions - PRJNA1400048.

Reference-guided read alignment

Cleaned reads from both species were aligned to the *P. dactylifera* reference genome using HISAT2, a splice-aware aligner optimized for RNA-seq data [23]. The reference-guided alignment strategy was employed due to the close phylogenetic relationship among *Phoenix* species and the availability of a high-quality *P. dactylifera* genome [2]. Default alignment parameters were used unless otherwise specified. Alignment statistics were examined to ensure mapping quality and consistency across samples.

Transcript quantification and gene expression estimation

Aligned reads were assigned to annotated gene features using featureCounts from the Subread package [27]. Raw gene-level read counts were generated for each species and used as input for differential expression analysis. Genes with extremely low read counts were filtered before downstream statistical analyses to reduce noise.

Differential gene expression analysis

Differential gene expression analysis between *P. theophrasti*

and *P. pusilla* leaf transcriptomes was performed using DESeq2 [28]. Read counts were normalized using DESeq2's median-of-ratios method. Differentially expressed genes (DEGs) were identified based on log2 fold change (Log_2FC) absolute value of ≥1 and adjusted *p*-values (0.05) calculated using the Benjamini–Hochberg False Discovery Rate (FDR) correction. Given the pooled nature of the samples, DEGs were interpreted as species-level transcriptional differences rather than population-level variation.

Functional annotation and gene ontology enrichment analysis

Functional annotation of expressed genes was derived from the *P. dactylifera* reference genome annotations. Gene Ontology (GO) enrichment analysis of DEGs was performed using the clusterProfiler R package [45]. Enrichment analyses were conducted separately for Biological Process, Molecular Function, and Cellular Component categories. GO terms with adjusted *p*-values below the selected significance threshold were considered significantly enriched.

KEGG pathway enrichment analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was conducted using clusterProfiler [45] in conjunction with KEGG annotations [21]. Differentially expressed genes were mapped to KEGG pathways to identify biological pathways showing significant enrichment between the two *Phoenix* species. Enriched pathways were interpreted in the context of leaf physiology, metabolism, and species-specific functional traits.

Trait-associated gene identification

To identify biologically relevant transcriptional differences, DEGs were manually curated into functional trait categories based on annotation descriptions, GO terms, and KEGG pathway associations. Trait-associated gene sets were selected to highlight processes relevant to leaf function, metabolism, and stress-related pathways. These curated gene lists were used for downstream visualization and interpretation.

Data visualization and statistical analysis

All statistical analyses and visualizations were performed in R. MA plots, volcano plots, heatmaps, and enrichment plots were generated using a combination of DESeq2, clusterProfiler, and ggplot2 [41]. Heatmaps were produced using Pheatmap v1.0.12 [24] with the scaled normalized expression values to visualize relative gene expression patterns between *P. theophrasti* and *P. pusilla*.

Results

Sequencing output and reference-guided alignment

RNA-seq libraries generated from mature leaf tissues of *P. theophrasti* and *P. pusilla* produced high-quality paired-end sequencing data with consistent depth across samples. Following quality filtering, reads were aligned to the *P. dactylifera* reference genome using HISAT2. Overall alignment rates were high and comparable between species, ranging from approximately 84% to 85% across all libraries (Table 1). The number of expressed genes detected was also similar between species, with approximately 26,000–27,000 genes expressed per library.

The consistently high alignment efficiency across both species indicates that the *P. dactylifera* reference genome provides adequate representation for transcriptomic analyses in closely related *Phoenix* species. These results support the suitability of a reference-guided approach for comparative gene expression analysis between *P. theophrasti* and *P. pusilla*.

Table 1: Summary of RNA-seq data.

Species	Sequencing Depth (M PE reads)	Overall alignment rate (%)	Expressed genes
Phoenix theophrasti rep1	50.4	84.45	26,128
Phoenix theophrasti rep2	50.2	84.46	26,172
Phoenix pusilla rep1	62.5	85.13	27,041
Phoenix pusilla rep2	62.9	85.16	27,083

Species-associated differential gene expression

To investigate transcriptional divergence between *P. theophrasti* and *P. pusilla*, differential gene expression analysis was performed at the gene level. After filtering low-abundance transcripts, a large number of genes exhibited significant differential expression between the two species (adjusted $p<0.05$). In total, 5,619 genes were up-regulated, and 5,901 genes were down-regulated in *P. theophrasti* relative to *P. pusilla* (Figure 1A), indicating extensive yet balanced transcriptional divergence. Visualization of differential expression using a volcano plot revealed a broad distribution of fold changes accompanied by strong statistical support (Figure 1B). More than 7,000 genes displayed large absolute log2fold changes together with extremely low adjusted *p*-values, suggesting that the observed expression differences are robust and not driven by marginal effects. Both up- and down-regulated gene sets spanned a wide range of functional annotations, indicating that multiple biological processes contribute to species-specific expression profiles.

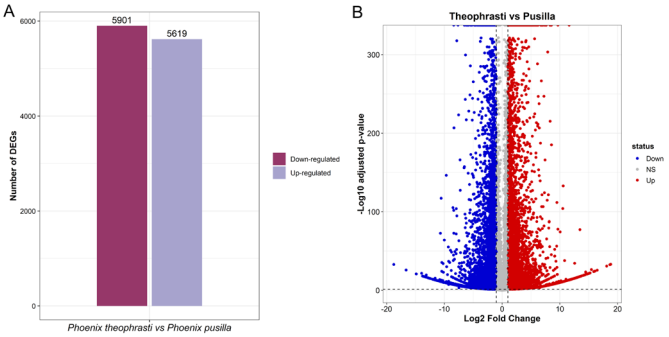


Figure 1: Differential gene expression between *P. theophrasti* and *P. pusilla*. **(A)** Bar plot showing the numbers of up- and down-regulated genes. **(B)** Volcano plot showing fold change vs significance.

Expression patterns of highly divergent genes

To visualize expression patterns among the most strongly differentiated genes, a heatmap was generated using variance-stabilized expression values for the top differentially expressed genes ranked by fold change and statistical significance. Hierarchical clustering revealed a clear grouping of samples by species, with replicates clustering closely together within each species (Figure 2).

The distinct clustering patterns observed across species, coupled with minimal variation within species, demonstrate that the transcriptional differences identified are consistent and reproducible at the species level. These results further support the presence of stable and biologically meaningful gene expression divergence between *P. theophrasti* and *P. pusilla* leaf tissues.

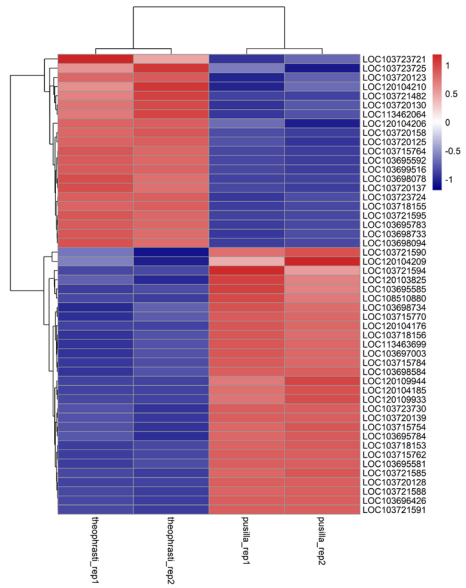


Figure 2: Heatmap of top differentially expressed genes between *P. theophrasti* and *P. pusilla.*

Functional enrichment of differentially expressed genes

To assess the biological significance of the observed transcriptional differences, functional enrichment analyses were conducted on the differentially expressed gene set. Gene Ontology (GO) enrichment analysis revealed significant over-representation of biological processes related to secondary metabolic processes, ribosome structure and biogenesis, and responses to biotic and abiotic stimuli (Figure 3A). Enriched stress-related categories included responses to jasmonic acid, wounding, hydrogen peroxide, and antimicrobial compounds, highlighting differences in stress perception and signaling between species.

Notably, multiple ribosome-associated terms, such as structural constituent of ribosome and cytosolic ribosomal subunit, were among the most significantly enriched categories. This enrichment suggests coordinated differences in translational capacity or regulation between *P. theophrasti* and *P. pusilla* leaves.

KEGG pathway enrichment analysis further corroborated these findings (Figure 3B). Significantly enriched pathways included phenylpropanoid biosynthesis, flavonoid and diterpenoid biosynthesis, plant–pathogen interaction, MAPK signaling, and ribosome-related pathways. Together, these results indicate that species-specific transcriptional divergence involves both core cellular processes and adaptive metabolic and signaling pathways relevant to leaf physiology.

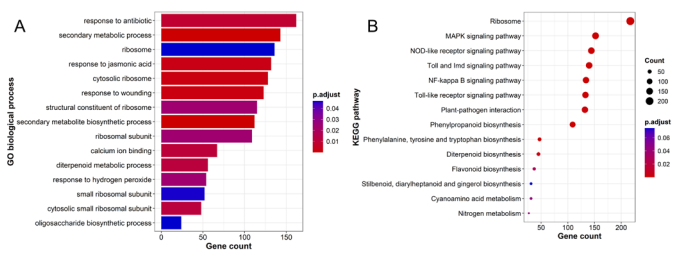


Figure 3: Functional enrichment of differentially expressed genes between *P. theophrasti* and *P. pusilla.* **(A)** Gene Ontology (GO) biological process enrichment of differentially expressed genes, showing significantly over-represented functional categories based on gene count and adjusted *p*-value. **(B)** Plant-relevant KEGG pathway enrichment of differentially expressed genes, highlighting pathways associated with metabolism, stress signaling, and defense responses.

Identification of trait-relevant candidate genes

To identify specific genes with potentially underlying functional divergence between species, differentially expressed genes were further examined at the individual gene level. High-confidence candidate genes were selected based on annotation quality, expression magnitude, and statistical significance. These genes were grouped into trait-relevant functional categories, including abiotic stress signaling, oxidative stress and Reactive Oxygen Species (ROS) regulation, defense and immunity, photosynthesis and light responses, cuticle and lipid metabolism, and secondary metabolism (Table 2 & Figure 4).

Genes associated with stress signaling and kinase-mediated pathways were prominently represented, including receptor-like kinases, calcium-associated signaling components, and heat shock proteins. Numerous genes involved in oxidative stress regulation, such as peroxidases, glutathione S-transferases, superoxide dismutases, and oxidoreductases, exhibited strong differential expression, suggesting species-specific modulation of redox homeostasis.

Defense- and immunity-related genes, including disease resistance proteins, pathogenesis-related proteins, and NB-ARC domain-containing genes, also showed pronounced expression differences, consistent with enrichment of plant–pathogen interaction pathways. In addition, multiple photosynthesis-related genes, including components of photosystem I and II, chlorophyll-binding proteins, and ATP synthase subunits, were differentially expressed, indicating divergence in photosynthetic regulation at the transcript level.

Genes associated with cuticle formation, wax and lipid metabolism, and secondary metabolic pathways, such as chalcone and terpene biosynthesis, were likewise differentially regulated between species. Collectively, these candidate genes provide mechanistic support for the enriched functional categories and suggest that transcriptional divergence between *P. theophrasti* and *P. pusilla* reflects coordinated differences in leaf metabolism, stress responses, and physiological regulation.

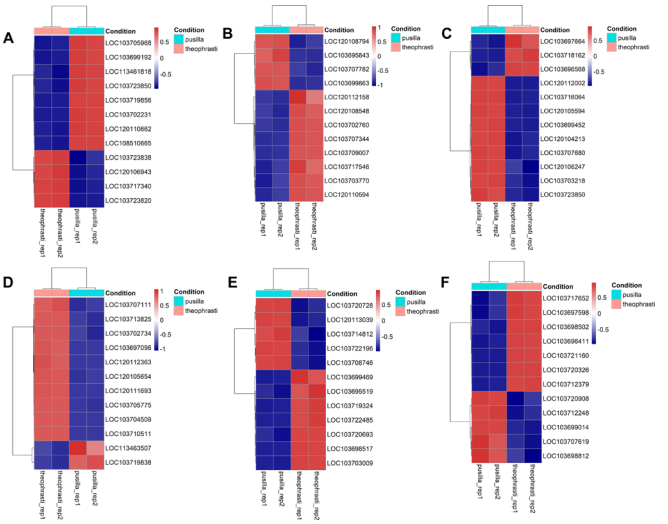


Figure 4: Differentially Expressed Genes (DEGs) between *P. theophrasti* and *P. pusilla.* **(A)** Abiotic stress signaling. **(B)** Oxidative stress / ROS. **(C)** Defense and Immunity. **(D)** Photosynthesis / Light response. **(E)** Cuticle, wax, and lipid metabolism. **(F)** Secondary metabolism.

Table 2: Candidate genes associated with trait-relevant functional categories showing differential expression between *P. theophrasti* and *P. pusilla*.

GeneID	Annotation	log2FoldChange	padj	Trait
LOC103702231	LRR receptor-like serine threonine-protein kinase	-10.69	2.37e-73	Abiotic stress and signaling
LOC113461818	LRR receptor-like serine threonine-protein kinase	-8.96	1.11e-51	
LOC120110662	Serine threonine-protein kinase	-8.36	1.88e-73	
LOC108510665	Protein kinase domain	-7.26	5.16e-128	
LOC103699192	Lectin-domain containing receptor kinase	-6.58	1.33e-156	
LOC103723838	Protein tyrosine kinase	3.51	2.58e-60	
LOC103717340	Belongs to the small heat shock protein (HSP20) family	3.0	0	
LOC120106943	Calmodulin-Ca(2) complex with protein kinases and phosphatases	2.87	2.49e-78	
LOC103723820	Belongs to the protein kinase superfamily	2.34	0	
LOC103705968	inactive receptor kinase At5g58300	-1.76	4.63e-113	
LOC103719856	Possible plasma membrane-binding motif in junctophilins, PIP-5-kinases, and protein kinases	-1.19	0	
LOC103702760	Belongs to the iron ascorbate-dependent oxidoreductase family	10.06	4.85e-75	Oxidative stress / ROS
LOC120112158	NADP-dependent oxidoreductase	6.46	3.45e-56	
LOC103699863	Belongs to the GST superfamily	-5.4	0	
LOC103703770	Belongs to the peroxidase family. Classical plant (class III) peroxidase sub-family	4.76	7.18e-101	
LOC120108794	Belongs to the GST superfamily	-4.61	1.99e-87	
LOC103707344	Belongs to the peroxidase family. Classical plant (class III) peroxidase sub-family	3.81	0	
LOC103695843	Glutathione S-transferase	-3.79	1.22e-97	
LOC103707782	Belongs to the glutathione peroxidase family	-3.71	0	
LOC120108548	Superoxide dismutase	2.8	5.72e-171	
LOC103709007	Belongs to the peroxidase family	2.52	0	
LOC103717546	Belongs to the glutathione peroxidase family	2.13	2.35e-80	
LOC120110594	Belongs to the glutathione peroxidase family	1.21	1.21e-112	
LOC103716064	Involved in the modulation of pathogen defense and leaf cell death	-9.67	4.09e-147	Defense and Immunity
LOC120104213	disease resistance	-6.38	8.70e-244	
LOC120105594	Pathogenesis-related protein	-5.91	0	
LOC103703218	Receptor homology region, transmembrane domain, and RING domain-containing protein	-5.73	2.46e-116	
LOC103699452	Pathogenesis-related protein	-5.2	0	
LOC103718162	disease resistance	4.05	3.87e-259	
LOC103707680	disease resistance protein	-5.17	1.94e-217	
LOC120106247	disease resistance protein	-4.95	2.61e-51	
LOC120112002	disease resistance protein	-4.74	2.31e-117	
LOC103697664	NB-ARC domain	4.13	9.14e-81	
LOC103723850	F-box LRR-repeat protein	-1.31	1.99e-131	
LOC103696568	A Receptor for Ubiquitination Targets	3.64	9.81e-47	
LOC120105654	Photosystem I reaction center subunit V	7.09	0	Photosynthesis /Light response
LOC103702734	BEST *Arabidopsis thaliana* protein match is hydroxyproline-rich glycoprotein family protein (TAIR AT1G21695.1)	6.67	4.09e-59	
LOC103704509	Chlorophyll A-B binding protein	5.21	0	
LOC103697096	Chlorophyll A-B binding protein	3.49	0	
LOC103713825	Photosystem II 10 kDa polypeptide	2.7	0	
LOC113463507	Peptidase, trypsin-like serine and cysteine proteases	-3.77	3.22e-15	
LOC120111693	Photosystem II reaction centre X protein (PsbX)	2.16	0	

LOC120112363	Photosystem I assembly	2.06	3.48e-265	
LOC103707111	ATP synthase subunit epsilon	2.02	0	
LOC103705775	Photosystem II reaction centre X protein (PsbX)	1.86	0	
LOC103710511	Photosystem I reaction center subunit psaK, chloroplastic	1.55	0	
LOC103719838	DNA polymerase epsilon catalytic subunit A	-1.91	1.63e-135	
LOC103720728	Enoyl-(Acyl carrier protein) reductase	-7.96	1.71e-72	Cuticle, wax and lipid metabolism
LOC120113039	Enoyl-(Acyl carrier protein) reductase	-7.85	1.79e-95	
LOC103722196	Enoyl-(Acyl carrier protein) reductase	-7.48	8.21e-64	
LOC103703009	GDSL-like Lipase/Acylhydrolase family	6.81	0	
LOC103720693	Indole-3-glycerol phosphate synthase	6.41	2.65e-238	
LOC103722485	Long chain acyl-CoA synthetase	6.14	1.23e-54	
LOC103719324	glycerol-3-phosphate acyltransferase	6.06	1.95e-93	
LOC103698517	Plant non-specific lipid-transfer proteins transfer phospholipids and galactolipids	4.64	0	
LOC103708746	Hydrolyzes glycerol-phospholipids at the terminal phosphodiesteric bond	-2.5	1.08e-145	
LOC103699469	Plant non-specific lipid-transfer proteins transfer phospholipids and galactolipids	1.64	1.84e-131	
LOC103714812	3-ketoacyl-CoA synthase	-1.63	2.29e-87	
LOC103695519	Glycolipid transfer protein (GLTP)	1.32	1.07e-05	
LOC103698502	Belongs to the chalcone stilbene synthases family	5.21	0	Secondary metabolism
LOC103720908	Belongs to the terpene cyclase mutase family	-4.07	7.33e-92	
LOC103712248	Belongs to the terpene cyclase mutase family	-2.84	0	
LOC103698812	Belongs to the terpene synthase family	-2.18	1.24e-64	
LOC103720326	Belongs to the chalcone stilbene synthases family	2.18	0	
LOC103721160	Belongs to the chalcone isomerase family	1.92	0	
LOC103699014	Omega-hydroxypalmitate O-feruloyl	-1.74	0	
LOC103707619	Belongs to the DHHC palmitoyltransferase family	-1.64	4.31e-57	
LOC103696411	Belongs to the chalcone isomerase family	1.63	0	
LOC103717652	Chalcone isomerase like	1.48	6.09e-95	
LOC103712379	Belongs to the chalcone stilbene synthases family	1.4	0	
LOC103697598	palmitoyl-(protein) hydrolase activity	1.26	1.54e-172	

Discussion

Comparative transcriptome analysis of mature leaf tissues from *P. theophrasti* and *P. pusilla* revealed extensive and coordinated transcriptional divergence between these two closely related palm species. Despite the use of a single reference genome (*P. dactylifera*), high alignment rates and consistent expression profiles across libraries indicate that the observed differences reflect genuine species-associated transcriptional variation rather than technical artifacts. Together, differential expression, functional enrichment, and trait-level gene analyses point to divergence in multiple core biological processes, including stress signaling, metabolic regulation, photosynthesis, and defense responses. The suitability of *P. dactylifera* as a reference genome for comparative analyses within the genus is supported by the availability of a high-quality genome assembly and extensive gene model annotation [1,2].

Species-level transcriptional divergence

The identification of more than 11,000 differentially expressed genes between *P. theophrasti* and *P. pusilla* highlights substantial transcriptional differentiation at the leaf level. The relatively balanced distribution of up- and down-regulated genes suggests widespread regulatory divergence rather than directional shifts confined to a single pathway. Similar magnitudes of interspecific transcriptional divergence have been reported in comparative leaf transcriptome studies of closely related plant species, even when overall genome structure remains highly conserved [7,34].

Clear species-specific clustering in the expression heatmap further supports the reproducibility and coherence of these transcriptional differences. Comparable patterns of clustering by species have been observed in RNA-seq studies of wild and cultivated monocots, including rice, maize, and other grasses, where leaf transcriptomes reflect both developmental programming and environmental adaptation [19,37]. In palms, transcriptome-level differentiation has been less extensively explored, making the present study an important contribution to understanding molecular divergence within the genus *Phoenix* [46]. Similar leaf-level transcriptome coherence has also been reported in date palm and royal palm under salinity and drought stress conditions [14,18,35,43].

Divergence in translational capacity and cellular processes

One of the findings of the functional enrichment analysis

was the significant over-representation of ribosome-related GO terms and KEGG pathways. Enrichment of genes encoding ribosomal proteins and translational machinery suggests differences in protein synthesis capacity or regulation between the two species. Differential regulation of ribosomal genes has been reported in multiple plant comparative transcriptomics studies and is often associated with variation in growth rates, metabolic activity, or stress responsiveness [13,14,22,44].

In leaves, translational regulation plays a central role in coordinating photosynthesis, stress responses, and metabolic fluxes. The observed divergence in ribosome-associated genes therefore likely reflects broader differences in cellular regulation and resource allocation between *P. theophrasti* and *P. pusilla*, rather than isolated effects on individual pathways. Enrichment of ribosome-related transcripts has similarly been observed in reference-guided transcriptome analyses of date palm tissues [31].

Divergence in stress signaling and redox regulation

Genes associated with abiotic stress signaling, oxidative stress, and Reactive Oxygen Species (ROS) regulation constituted a major component of the differentially expressed gene set. A number of receptor-like kinases, calcium-associated signaling proteins, and heat shock proteins exhibited strong species-specific expression patterns. Such signaling components are central to plant perception and integration of environmental cues, particularly in leaf tissues exposed to fluctuating light, temperature, and water availability [29,47].

Similarly, the prominent representation of peroxidases, glutathione S-transferases, superoxide dismutases, and oxidoreductases indicates divergence in redox homeostasis between species. Comparative studies in monocots have shown that variation in ROS-related gene expression often underlies differences in stress tolerance and acclimation capacity [15]. Date palm transcriptome studies under heat, drought, salinity, and cadmium stress have consistently reported differential regulation of antioxidant enzymes, stress-responsive transcription factors, and signaling components in leaf tissues [18,33].

Defense- and immunity-related transcriptional differences

Defense-related genes, including disease resistance proteins, pathogenesis-related proteins, and NB-ARC domain-containing genes, were among the most strongly differentiated between *P. theophrasti* and *P. pusilla*. Enrichment of plant–pathogen interaction pathways further support divergence in immune-related transcriptional programs. Similar patterns have been observed in comparative transcriptome analyses across plant species, where immune genes are often among the most rapidly evolving and differentially regulated components of the genome [17,20].

Leaf tissues represent a primary interface between plants and their biotic environment, and variation in immune gene expression may reflect differences in pathogen pressure or ecological context experienced by the two species. While functional assays would be required to confirm adaptive significance, the transcriptional patterns observed here suggest that defense signaling constitutes a key axis of divergence within *Phoenix*. Genome-wide analyses of date palm have also reported expansion and diversification of resistance gene families, underscoring the importance of immune regulation in palms [2].

Photosynthesis and leaf metabolic regulation

Differential expression of genes encoding photosystem components, chlorophyll-binding proteins, ATP synthase subunits, and photosystem assembly factors indicate divergence in photosynthetic regulation between species. Comparative transcriptomics in leaves has repeatedly demonstrated that even subtle differences in photosynthesis-related gene expression can influence photosynthetic efficiency, light utilization, and overall carbon metabolism [11,25].

In addition to photosynthesis, genes associated with cuticle formation, wax biosynthesis, lipid metabolism, and secondary metabolism were differentially expressed. Cuticle and wax components play critical roles in water retention, protection from environmental stress, and leaf surface interactions [36]. Divergence in these pathways has been reported in monocots occupying contrasting environments and is often linked to ecological specialization [32]. Leaf transcriptome analyses in date palm have similarly shown coordinated regulation of photosynthesis- and cuticle-related genes under salinity stress [18].

Secondary metabolism and species-specific chemical profiles

The enrichment of phenylpropanoid, flavonoid, and terpene biosynthesis pathways, together with differential expression of chalcone synthases, terpene synthases, and related enzymes, suggests divergence in secondary metabolic capacity between *P. theophrasti* and *P. pusilla*. Secondary metabolites contribute to defense, stress tolerance, and interactions with herbivores and microbes, and their regulation is frequently species-specific [10,38]. Comparative transcriptomic studies in monocots and woody plants have shown that variation in secondary metabolism often reflects both evolutionary history and environmental adaptation [12,26]. Enrichment of secondary metabolic pathways has also been reported in date palm transcriptomes under abiotic stress, suggesting conserved regulation of these pathways across palms [12].

Methodological considerations

Because RNA samples were pooled for each species, the present study captures species-level transcriptional differences rather than within-species variation. Similar pooled sample strategies have been successfully employed in comparative transcriptomics when the primary objective is interspecific comparison [8,30]. The use of *P. dactylifera* as a reference genome represents a pragmatic and widely accepted strategy in non-model plant transcriptomics. The high alignment rates observed here support the validity of this approach, although species-specific transcripts absent from the reference genome may not be fully captured. Comparable reference-guided approaches have been successfully applied in multiple date palm transcriptome studies across tissues and stress conditions [18,31].

Conclusions

This comparative transcriptome analysis reveals extensive and coordinated transcriptional divergence between *P. theophrasti* and *P. pusilla* palms. Differences in translational machinery, stress signaling, redox regulation, defense responses, photosynthesis, and secondary metabolism collectively suggest that species-specific transcriptional programs underlie functional differentiation within the genus *Phoenix*. By integrating differential expression, functional enrichment, and trait-focused gene analysis, this study provides a comprehensive framework

for understanding transcriptome divergence in palms and establishes a foundation for future ecological, physiological, populational and evolutionary investigations.

Author declarations

Author contributions

VMG and MNR analyzed, reviewed, and drafted the manuscript. MNR conceptualized the experiments, edited the manuscript, and contributed to the logistical support for the execution of the experiment.

Conflict of interest

Authors in this research work declare no conflict of interest.

Acknowledgements

This research was supported through appropriated funds from the United States Department of Agriculture - Agricultural Research Service (USDA-ARS) under Project Number 603813210-004-000-D. Various biological technicians who maintain and help with sampling palms are highly acknowledged.

References

1. Al-Dous EK, George B, Al-Mahmoud ME, Al-Jaber MY, Wang H, et al. De novo genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). Nat Biotechnol. 2011; 29: 521-527.
2. Al-Mssallem IS, Hu S, Zhang X, Lin Q, Liu W, et al. Genome sequence of the date palm *Phoenix dactylifera* L. Nat Commun. 2013; 4: 2274.
3. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010.
4. Behera S, Voshall A. Plant transcriptome assembly: review and benchmarking. Bioinformatics. 2022.
5. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30: 2114-2120.
6. Chaluvadi SR, Young P, Thompson K, Bahri BA, Gajera B, et al. Phoenix phylogeny, and analysis of genetic variation in a diverse collection of date palm (*Phoenix dactylifera*) and related species. Plant Divers. 2019; 41: 330-339.
7. Chang YM, Lin HH, Liu WY, Yu CP, Chen HJ, et al. Comparative transcriptomics method to infer gene coexpression networks and its applications to maize and rice leaf transcriptomes. Proc Natl Acad Sci USA. 2019; 116: 3091-3099.
8. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, et al. A survey of best practices for RNA-seq data analysis. Genome Biol. 2016; 17: 13.
9. Des Marais DL, McKay JK, Richards JH, Sen S, Wayne T, et al. Physiological genomics of response to soil drying in diverse *Arabidopsis* accessions. Plant Cell. 2012; 24: 893-914.
10. Dixon RA, Paiva NL. Stress-induced phenylpropanoid metabolism. Plant Cell. 1995; 7: 1085.
11. Du K, Wu W, Liao T, Yang J, Kang X. Transcriptome analysis uncovering regulatory networks and hub genes of Populus photosynthesis and chlorophyll content. Genomics. 2022; 114: 110385.
12. Fang C, Fernie AR, Luo J. Exploring the diversity of plant metabolism. Trends Plant Sci. 2019; 24: 83-98.
13. Gopala VM, Nageswara-Rao M, Govindarajulu R, Meerow AW, Singh S. Genetic insights into *Hippeastrum* species: a comparative transcriptomic analysis. J Biosci. 2025; 50: 82.
14. Gopala VM, Nageswara-Rao M, Griffith MP, Tucker DA, Sukhwinder S. Transcriptomic insights into salinity responses in *Roystonea oleracea* under freshwater and saline field conditions. J Plant Biol Crop Res. 2026; 10: 1113.
15. Gill SS, Tuteja N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiol Biochem. 2010; 48: 909-930.
16. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013; 8: 1494-1512.
17. Han X, Tsuda K. Evolutionary footprint of plant immunity. Curr Opin Plant Biol. 2022; 67: 102209.
18. Hazzouri KM, Flowers JM, Nelson D, Lemansour A, Masmoudi K, et al. Prospects for the study and improvement of abiotic stress tolerance in date palms in the post-genomics era. Front Plant Sci. 2020; 11: 293.
19. Huang CF, Chang YM, Lin JJ, Yu CP, Lin HH, et al. Insights into the regulation of C4 leaf development from comparative transcriptomic analysis. Curr Opin Plant Biol. 2016; 30: 1-10.
20. Jones JD, Dangl JL. The plant immune system. Nature. 2006; 444: 323-329.
21. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. Nucleic Acids Res. 2021; 49.
22. Kawaguchi R, Bailey-Serres J. Regulation of translational initiation in plants. Curr Opin Plant Biol. 2002; 5: 460-465.
23. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019; 37: 907-915.
24. Kolde R. pheatmap: Pretty Heatmaps. R package version 1.0.12. 2018.
25. Lawson T, Kramer DM, Raines CA. Improving yield by exploiting mechanisms underlying natural variation of photosynthesis. Curr Opin Biotechnol. 2012; 23: 215-220.
26. Li X, Cai K, Zhao Q, Li H, Wang X, et al. Morphological and comparative transcriptome analysis of three species of five-needle pines: insights into phenotypic evolution and phylogeny. Front Plant Sci. 2022; 13: 795631.
27. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014; 30: 923-930.
28. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15: 550.
29. Mittler R, Vanderauwera S, Gollery M, Van Breusegem F. Reactive oxygen gene network of plants. Trends Plant Sci. 2004; 9: 490-498.
30. Mutz KO, Heilkenbrinker A, Lönne M, Walter JG, Stahl F. Transcriptome analysis using next-generation sequencing. Curr Opin Biotechnol. 2013; 24: 22-30.
31. Naganeeswaran S, Fayas TP, Rajesh MK. Dataset of transcriptome assembly of date palm embryogenic calli and functional annotation. Data Brief. 2020; 31: 105760.

32. Qiao P, Bourgault R, Mohammadi M, Matschi S, Philippe G, et al. Transcriptomic network analyses shed light on the regulation of cuticle development in maize leaves. Proc Natl Acad Sci USA. 2020; 117: 12464-12471.

33. Rekik I, Chaâbene Z, Kriaa W, Rorat A, Franck V, et al. Transcriptome assembly and abiotic related gene expression analysis of date palm reveal candidate genes involved in response to cadmium stress. Comp Biochem Physiol C Toxicol Pharmacol. 2019; 225: 108569.

34. Romero IG, Ruvinsky I, Gilad Y. Comparative studies of gene expression and the evolution of gene regulation. Nat Rev Genet. 2012; 13: 505-516.

35. Safronov O, Kreuzwieser J, Haberer G, Alyousif MS, Schulze W, et al. Detecting early signs of heat and drought stress in *Phoenix dactylifera* (date palm). PLoS One. 2017; 12: 177883.

36. Shepherd T, Wynne Griffiths D. The effects of stress on plant cuticular waxes. New Phytol. 2006; 171: 469-499.

37. Street NR, Sjödin A, Bylesjö M, Gustafsson P, Trygg J, et al. A cross-species transcriptomics approach to identify genes involved in leaf development. Genomics. 2008; 9: 589.

38. Tohge T, Alseekh S, Fernie AR. On the regulation and function of secondary metabolism during fruit development and ripening. J Exp Bot. 2013; 65: 4599-4611.

39. Vercruysse J, Van Bel M, Osuna-Cruz CM, Kulkarni SR, et al. Comparative transcriptomics enables the identification of functional orthologous genes involved in early leaf growth. Plant Biotechnol J. 2020; 18: 553-567.

40. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009; 10: 57-63.

41. Wickham H. Data analysis. In: ggplot2: elegant graphics for data analysis. Cham: Springer International Publishing; 2016: 189-201.

42. Xu Z, Zhang N, Fu H, Wang F, Wen M, et al. Salt stress modulates the landscape of transcriptome and alternative splicing in date palm (*Phoenix dactylifera* L.). Front Plant Sci. 2022; 12: 807739.

43. Yaish MW, Patankar HV, Assaha DVM, Zheng Y, Al-Yahyai R, et al. Genome-wide expression profiling in leaves and roots of date palm (*Phoenix dactylifera* L.) exposed to salinity. BMC Genomics. 2017; 18: 246.

44. Yang X, Song B, Cui J, Wang L, Wang S, et al. Comparative ribosome profiling reveals distinct translational landscapes of salt-sensitive and tolerant rice. BMC Genomics. 2021; 22: 612.

45. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16: 284-287.

46. Zhang G, Pan L, Yin Y, Liu W, Huang D, et al. Large-scale collection and annotation of gene models for date palm (*Phoenix dactylifera* L.). Plant Mol Biol. 2012; 79: 521-536.

47. Zhu JK. Abiotic stress signaling and responses in plants. Cell. 2016; 167: 313-324.