Co-expression Network Analysis in Carya illinoinensis cvs. 'Tiny Tim' and 'Mahan' Identifies Cell Wall Remodeling and Inositol Metabolism Modules Associated with Nut Size

June Labbancz
Texas A&M University
Warren Chatwin
Southern Plains Agricultural Research Center
Amit Dhingra
adhingra@tamu.edu

Texas A&M University

Article

Keywords: Pecan, Nut Development, Fruit Size, Transcriptomics, WGCNA, Co-expression network

Posted Date: October 30th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7843615/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on February 10th, 2026. See the published version at https://doi.org/10.1038/s41598-026-38292-2.

Abstract

Pecan is a tree nut crop native to the United States and Mexico, with a global market of over 2 billion USD. Nut size has been the most important target trait for crop improvement during the very limited breeding cycles. However, relatively little is known about the molecular basis of pecan nut ontogeny and the mechanisms underlying pecan nut sizing. Besides nut size, pecan fruit faces myriad physiological disorders throughout the growing season, making knowledge of essential genes at each growth stage a necessary first step in developing new cultivars and management practices to overcome these issues. To develop a deeper understanding of pecan fruit development and identify candidate genes underlying the large fruit phenotype, a time-course transcriptomic study of pecan fruit in two genotypes, 'Mahan' and 'Tiny Tim', was conducted. Weighted Gene-Coexpression Network Analysis (WGCNA) was employed to group transcripts into functional clusters, and hub transcripts were identified through module correlation analysis to select those that are potential drivers of these functional clusters. Modules related to cell wall biosynthesis, cell wall organization, and inositol metabolism in 'Mahan', and proteolysis and abscisic acid response in 'Tiny Tim' were found to be associated with nut size.

Introduction

Pecan (*Carya illinoinensis*) is a large tree native to the Central United States and the hilly regions of Mexico (T. E. Thompson & Grauke, 1991) and is known for its large drupe-like fruits, commonly referred to as nuts [1]. The nuts are enclosed in husks derived from the floral involucre, commonly referred to as the "shuck" or "husk", which are inedible and are generally discarded [1, 2]. Pecan fruit development can be broadly split into two phases: fruit enlargement and kernel filling. Pecan fruit enlargement is generally linear for the first half of fruit development, slowing considerably during the latter half of development [3]. Kernel filling during the latter half of fruit development is typically subdivided into the water, gel, and dough stages, based on the consistency of the developing endosperm [4].

Pecan fruits are susceptible to various physiological disorders and diseases with respect to these phases, making understanding of pecan fruit development critical for identifying best practices for orchard management. For example, pecan scab is far more devastating to yield and quality when it initiates earlier in the season [5], while shuck split is most likely to occur during the water stage [6, 7], and vivipary is a late growing season issue, especially under conditions of irrigation and fertilization [8]. Susceptibility to these various physiological maladies varies by cultivar, yet the underlying molecular basis remains unknown.

While pecan fruits have been collected for sustenance for thousands of years by Native Americans, the history of their cultivation is less than 300 years old [9]. Pecan trees have experienced few generations of selection under cultivation due to a relatively long immature stage, with a minimum time of about 7 years to fruit bearing [10]. Despite this, nut size is one of the most striking traits that have been selected for in modern pecan cultivars, with the mature fruit of some cultivated pecans exceeding 10 grams [11, 12], as opposed to the 2–6 grams of their uncultivated counterparts [13].

Genes associated with the regulation of fruit size in higher plants are incredibly diverse in function, including hormone metabolism, transcription factors, kinases, central carbon metabolism, and polysaccharide anabolism, among others [14]. In tomato, the best-studied model for fruit development, auxin is a driver of cell division, while gibberellins are drivers of cell expansion, both influencing fruit size [15].

In the related Persian walnut, cell wall-related metabolic processes appear to be loci of interest in explaining the genetic determination of nut size [16, 17]. Genome size also appears to be correlated with fruit size in Persian walnut, but the exact mechanism driving this relationship remains unexplored [18]. The wide range of possible mechanisms influencing fruit size makes the identification of genes responsible for increased fruit size difficult and makes a transcriptomic approach appealing. Previous transcriptomic studies of pecan fruit have primarily focused on the later stages of development [19, 20], with a focus on lipid and protein content, though one study tracked fruit development biweekly through the whole growing season in the cultivar 'Annong3' [21]. Genotype-specific differences at the transcriptomic level, however, remain entirely unexplored, and genes responsible for fruit size have not been identified.

To characterize the genetic basis of pecan fruit development, we conducted a developmental time course RNA Sequencing (RNASeq) experiment on the fruits of the pecan cultivars 'Mahan' and 'Tiny Tim', followed by Weighted Gene Co-expression Network Analysis (WGCNA). 'Mahan' was selected as a cultivar with particularly large fruit, and it is the progenitor of many commercial pecan cultivars, including 'Pawnee' and 'Wichita'. 'Tiny Tim' was selected as a representative of small nut native pecans. Fruit was collected from two individuals of each genotype biweekly from the start of fruit development in May to fruit maturity in October.

We hypothesize that the previously defined stages of pecan development will correlate with modules of transcripts, and hub and other genes with high module membership can be identified as potential drivers of development stage-specific physiology, worthy of further investigation. Differences between genotypes may additionally be useful for hypothesis development in understanding the basis of the recent domestication of pecan and directions for future pecan improvement.

Methods

Sample Collection

Pecan tissue was collected from two trees per genotype, 'Mahan' and 'Tiny Tim', from the CSV block at the National Clonal Germplasm Repository pecan collection at Somerville, TX under the National Plant Germplasm System - Brownwood repository for pecans and hickories in College Station, TX. Tree location identifiers were CSV 6–6 and 18 – 12 for 'Mahan' and CSV 9–10 and 20 – 8 for 'Tiny Tim'. A minimum of three nuts were harvested at approximately two-week intervals from May 9, 2022, to October 31, 2022 (Table S1), immediately frozen in liquid nitrogen and subsequently stored at -80°C. A total of 16 'Mahan' and 20 'Tiny Tim' sample collection dates occurred.

RNA Isolation

Pecan tissue was cryomilled using a SPEX Freezer Mill 6870 (Metuchen, NJ) for 4 minutes at 15cpm. A minimum of three nuts were pooled into each sample. Milled tissue was immediately stored at -80°C until further use. A CTAB and lithium chloride-based RNA isolation method was used to prepare DNA-free samples for RNA sequencing. Approximately 100 mg of milled pooled tissue was added to 1 mL of CTAB Isolation Buffer and incubated at 60°C for 30 minutes, with 5 seconds of vortexing every 5 minutes. An equal volume of 24:1 chloroform:isoamylalcohol was added to each sample, and vortexed for 10 seconds, followed by 5 minutes of rotational mixing. Samples were centrifuged at 8000g for 6 minutes at 4°C, and then the supernatant was moved to a fresh tube. Chloroform extraction was repeated once. Once the aqueous phase was transferred to a fresh tube, 0.5 volumes of 7M DEPC-treated LiCl was added to each sample. Samples were gently mixed and then incubated at -20°C for 3 days. The samples were then centrifuged at 15,000 g for 20 minutes at 4°C.

All subsequent isolation steps were performed in a laminar flow hood to ensure sterility. Supernatant was decanted, taking care not to disturb the pellet. Samples were washed twice with 70% DEPC-treated ethanol, with intervening centrifugation at 15,000 g for 10 minutes at 4°C. The samples were then resuspended in 80 µL of DEPC-treated ultrapure water. Samples were tested for RNA integrity and RNA quantity using the RNA IQ and RNA BR assays with a QuBit4 fluorometer. RNA purity was assessed using a Nanodrop 8000 UV spectrophotometer. Any samples with RIN values < 7, 260nm/280nm absorbance ratios < 1.8, or RNA concentrations < 20ng/µL were re-isolated until a sample with sufficient quality and quantity was generated.

RNA Sequencing

RNA samples were processed at the Texas A&M AgriLife Genomics and Bioinformatics Service for RNASeq library preparation and sequencing. Quality control was performed on 11 randomly selected samples before library preparation using a Fragment Analyzer system (Agilent, USA). Library preparation was carried out on all samples using the PerkinElmer Next Flex RNA kit. The prepared library was run on a single lane of a NovaSeq S4 XP 2x150 flow cell in a NovaSeq S4 6000 instrument.

Data Analysis

The collected sequence data were assessed using FastQC to determine the optimal trimming parameters. CLC Genomics Workbench version 24 was used for read trimming and mapping of reads to the Pawnee v1.1 reference transcriptome [22], accessed at Phytozome on December 15, 2023. Reads were trimmed to remove the first 15 nucleotides, remove the adapters, and remove ambiguous or low-quality sequences (quality limit = 0.05). TPM counts of transcript abundance were log2 normalized in R. Normalized counts underwent principal component analysis, and two outlier samples were excluded from further analysis.

The WGCNA package in R was used for network construction and visualization [23]. A soft thresholding power was selected for signed network construction. A soft thresholding power of 16 was selected, as this value maximized the scale-free tropology model fit at a value of 0.8520. WGCNA network construction was completed with the following parameters: maxBlockSize = 15000, minModuleSize = 30, power = 16, networkType = signed. Module membership was calculated for all genes, and gene correlations were calculated based on month within the growing season (May to October) and genotype (‘Mahan’ and ‘Tiny Tim’). Predicted protein sequences from the Pawnee v1.1 reference were processed by DeepGO-SE with default parameters to predict gene ontology (GO) terms for functional analysis of modules [24].

Results and Discussion

Sequencing results

A total of 534.9 GB of sequencing data was generated for 70 samples. Median sequencing depth per sample was 7.5 GB, and no sample had less than 6.3Gb of sequencing depth (Table S2). Principal components analysis of the TPM abundances of all samples shows a general progression through the growing season shared by both genotypes, which captures most of the observed variation (Figure S1).

Fruit Development

WGCNA Analysis

In this study, WGCNA was employed to identify functional gene clusters that could be organized based on their relationships to genotype and the time of year (Figure S2). A total of 105 modules were detected, with a median size of 80 and a mean size of 412 transcripts (Table S3). 50% of the present transcripts were found in the top 6 modules, and 90% of the present transcripts were found in the top 38 modules. The module grey, representing transcripts that could not be organized into modules, contained 1,192 transcripts (2.7% of the 44,626 transcripts in the dataset). A minority of transcripts in the dataset exhibited a significant genotype correlation ($p < 0.05$). A total of 5468 transcripts (12.5%) showed significant associations with ‘Mahan’, while 8386 transcripts (19.2%) showed significant associations with ‘Tiny Tim’. Correlations between module eigengenes (MEs) and traits are enumerated in Fig. 1 for modules with 100 or more members. Smaller modules tended to be composed exclusively of transcripts with very low expression levels. Most modules showed a correlation with at least one genotype or month of the growing season.

Early Season: Diverse Function for Cell Development

For both ‘Mahan’ and ‘Tiny Tim’, the early season is a time of rapid fruit development, developing from an ovary to a fruit of mostly final dimensions within about 3 months. Module turquoise, the largest module (10668 transcripts), is rich in genes of diverse biosynthetic functions, while module blue, the second largest module, is highly enriched in cell cycle-related genes (Fig. 2a and 2b), reflecting the rapid cellular division and catabolic processes occurring within pecan fruit early in the growing season. Module lightyellow is highly enriched in protein synthesis and ribosomal genes, further displaying the first months of pecan development, which is dominated by central metabolic and biosynthetic processes (Fig. 2c). As with many plant growth-related processes, auxin is a key regulatory hormone, with module turquoise’s hub gene being a highly expressed auxin response factor.

Module greenyellow’s hub gene has a predicted function as an IAA methyltransferase (Table 1), reinforcing the centrality of auxin to early fruit development in pecan independent of genotype. While this early stage of fruit development has remained unexplored in other pecan transcriptome papers, in walnut, the earliest stages of nut development are associated with diverse biosynthetic functions, cellular division, and auxin response, as well [25, 26].

Table 1
Hub transcripts of each module containing more than 100 members. Significance markers: * - p < 0.05, ** - p < 0.01, *** - p < 0.001

Transcript	Module Color	Positive Correlations	Module Correlation	Genotype	May	June	July	Aug	Sept	Oct	P
CiPaw.05G102300.2	turquoise	May, June, July	0.98	-0.04	0.29*	0.41***	0.27*	-0.14	-0.5***	-0.45***	A 4
CiPaw.13G105600.1	blue	May, June	0.99	-0.05	0.66***	0.34**	-0.02	-0.32**	-0.36**	-0.38**	C
CiPaw.09G216900.1	brown	July	0.98	-0.01	-0.12	0.16	0.39**	0.21	-0.36**	-0.41***	D L
CiPaw.16G011600.1	yellow	Tiny Tim, May, June	0.97	0.71***	0.22	0.29*	0.11	0	-0.33**	-0.39**	L C
CiPaw.15G016400.1	green	September, October	0.98	0.13	-0.21	-0.37**	-0.33**	-0.04	0.55***	0.54***	T F
CiPaw.08G177100.6	red	Mahan	0.99	-0.84***	0.08	0.07	0.17	-0.04	-0.14	-0.2	L C
CiPaw.01G154500.1	black	July	0.95	0.08	0.12	0.14	0.29*	-0.06	-0.23	-0.35**	S D
CiPaw.07G043800.2	pink	Tiny Tim, September, October	0.98	0.28*	0.15	-0.21	-0.24*	-0.21	0.23	0.38**	P F
CiPaw.04G156100.1	magenta	Tiny Tim, September	0.96	0.63***	0.1	-0.08	-0.27*	-0.16	0.29*	0.2	L R K
CiPaw.04G041200.1	purple	August, September	0.93	0.01	-0.55***	-0.37**	0.12	0.26*	0.36**	0.23	U P
CiPaw.12G133800.2	greenyellow	May	0.99	-0.11	0.84***	-0.05	-0.19	-0.25*	-0.24*	-0.17	I M
CiPaw.02G073400.1	tan	Tiny Tim, July	0.96	0.33**	0.13	0.32**	0.26*	-0.05	-0.45***	-0.29*	F C R
CiPaw.02G152500.2	salmon	Tiny Tim, July, August	0.93	0.18	-0.47***	-0.14	0.16	0.52***	0.03	-0.18	R C
CiPaw.04G111200.1	cyan	October	0.99	0.12	0.25*	-0.06	-0.06	-0.07	-0.06	-0.03	4
CiPaw.04G137800.2	midnightblue	Tiny Tim	0.92	0.43***	0.01	0.07	0.17	-0.05	-0.15	-0.06	S P
CiPaw.12G022500.1	lightcyan	Tiny Tim, May	0.94	0.34**	0.56***	0.1	-0.24*	-0.47***	0.04	0.05	E S
CiPaw.11G132500.1	grey60	September, October	0.98	-0.13	-0.24*	-0.24	-0.24*	0.14	0.37**	0.28*	C P
CiPaw.10G158400.1	lightgreen	Tiny Tim	0.97	0.94***	0.16	0.1	-0.04	-0.07	-0.11	-0.05	D F F
CiPaw.08G008700.1	lightyellow	May	0.95	-0.19	0.38**	0.16	0	-0.35**	-0.08	-0.12	4 S
CiPaw.04G066300.1	royalblue	July, August	0.95	-0.14	-0.44***	-0.09	0.33**	0.4***	-0.01	-0.31**	R
CiPaw.02G081900.1	darkred	Mahan, June, July	0.97	-0.28*	0.04	0.28*	0.34**	0.05	-0.39**	-0.44***	W P R
CiPaw.15G177100.1	darkgreen	Mahan, September, October	0.97	-0.96***	-0.11	-0.14	0.06	-0.12	0.18	0.2	D P
CiPaw.01G211800.1	darkturquoise	August, September	0.94	-0.11	-0.63***	-0.42***	0.05	0.52***	0.39***	0.09	C
CiPaw.04G162800.1	darkgrey	September	0.93	0.2	-0.03	-0.07	0.04	-0.26*	0.23	0.16	P B M P
CiPaw.07G040300.1	orange	Mahan, July	0.95	-0.13	0.06	0.26*	0.24*	0.19	-0.33**	-0.59***	L S P

Transcript	Module Color	Positive Correlations	Module Correlation	Genotype	May	June	July	Aug	Sept	Oct	P
CiPaw.06G119500.1	darkorange	Tiny Tim	0.94	0.76***	-0.31*	-0.2	-0.02	0.24*	0.19	0.11	M C R
CiPaw.05G139700.1	white	August, September, October	0.97	0.07	-0.58***	-0.45***	-0.07	0.4***	0.47***	0.27*	2 d
CiPaw.14G004000.2	skyblue	July	0.94	0.03	-0.26*	-0.08	0.2	0.12	0.01	0.01	S
CiPaw.15G053600.1	saddlebrown	July, August	0.97	-0.04	-0.42***	0.07	0.29*	0.48***	-0.19	-0.35**	C
CiPaw.09G008200.1	steelblue	June, July, August	0.97	0.04	-0.17	0.28*	0.35**	0.28*	-0.32**	-0.58***	H H
CiPaw.13G089000.1	paleturquoise	August	0.98	-0.18	-0.21	-0.19	-0.18	0.37**	0.28*	-0.13	F
CiPaw.05G021100.1	violet	August, September	0.97	-0.18	-0.4***	-0.38**	-0.25*	0.44***	0.46***	0.14	A R
CiPaw.09G029300.1	darkolivegreen		0.94	0.14	-0.03	0.08	0.14	0.03	-0.04	-0.24*	P B
CiPaw.14G037300.1	darkmagenta	September, October	0.96	0.07	0.23	-0.17	-0.44***	-0.34**	0.38**	0.49***	C R
CiPaw.01G000700.1	sienna3	May	0.99	-0.13	0.25*	-0.05	-0.06	-0.07	-0.05	-0.04	P
CiPaw.03G172000.1	yellowgreen	October	0.97	-0.13	-0.06	-0.05	-0.06	-0.07	-0.05	0.39***	
CiPaw.11G154900.1	skyblue3	May, October	0.95	0.36**	0.55***	-0.07	-0.32**	-0.38**	0.07	0.22	K I
CiPaw.04G110000.1	plum1	October	0.98	0.2	0.14	-0.11	-0.13	-0.16	0.04	0.31**	1 1 3
CiPaw.02G086200.1	orangered4	May, June	0.95	0.16	0.54***	0.31*	-0.11	-0.52***	-0.19	-0.01	N D
CiPaw.02G186600.1	mediumpurple3		0.99	-0.13	-0.06	-0.05	-0.06	0.23	-0.05	-0.04	
CiPaw.15G135500.2	lightsteelblue1	Mahan, May	0.97	-0.32**	0.59***	0.04	-0.08	-0.23	-0.21	-0.15	C G
CiPaw.07G001300.1	lightcyan1	August	0.99	0.12	-0.05	-0.06	-0.06	0.23	-0.06	-0.02	I(R
CiPaw.01G001500.1	ivory	October	0.99	-0.13	-0.06	-0.05	-0.06	-0.07	-0.05	0.39***	

Mid-Season: Hardening Fruit, Developing Ovule, and Resisting Summer Heat

Growth of the ovule and its filling with free-nucleate endosperm, which will eventually become the nut within the developing pecan fruit, occurs during the middle of the growing season [4]; modules violet and royalblue, associated with this time of year, are enriched in modules that are enriched in floral organ development transcripts, and appear to be responsive to cytokinin and auxin, respectively (Fig. 3a and 3b). The largest mid-season module, brown, is enriched in transcripts related to polysaccharide biosynthesis and cell wall development (Fig. 3c).

Whereas early-season module blue represents the cellular functions necessary for the cell division stage of fruit development, module brown represents many of the functions necessary for cell expansion. Despite this clear division of stages, as in the classic double sigmoid model of fruit development, the growth of the volume of the pecan nut does not appear to follow a double sigmoid model [4, 27]. Prior work in pecan fruit has shown an elevation of polysaccharide (especially xyloglucan) metabolism-related genes during this time of year, further underlining the importance of this process in mid-season development [21].

Module violet, the module correlated most strongly with August, is enriched in transmembrane transport-related functions (Fig. 3a). Module violet's hub gene itself is an amino acid permease, a class of genes that facilitate the import of amino acids across membranes [28, 29]. This amino acid transport into the developing embryo increases seed yield and nutrient content in *Arabidopsis* and rice (*Oryza sativa*), and it is a key gene in early embryo development [30–32]. Developing embryos are regions of high solute concentration and turgor as sink tissues [33]. The increasing but incomplete rigidity of the pecan fruit, combined with continued filling mediated by solutes, can result in significant crop losses due to water splitting, particularly if rain is abundant following a dry period [6, 34].

In various fruit crops, including apple and cherry, fruit cracking has been linked to genes with functions in cell wall-related processes, polysaccharide metabolism, and ion transport [35, 36]. Amino acid permeases show an ability to modulate cytokinin levels in rice, with elevated expression being associated with lower cytokinin levels [37]. In apple, orange, persimmon, and litchi, cytokinin application can aid in the prevention of cracking disorders [38–41]. Together with module violet's cytokinin-related enriched GO terms, this hormone is an interesting candidate molecule for exploring this disorder (Fig. 3a). The interplay

between the timing of cell wall hardening, solute transport, cytokinin, and fruit splitting may be worth exploring, with the hub genes of modules violet and brown being key candidates.

A prominent aspect of July and August is the presence of modules containing highly expressed heat shock protein genes. The modules saddlebrown and steelblue are all enriched in heat stress response transcripts (Fig. 3d and 3e). The hub genes of these modules appear to be a homolog of the rice gene Os02g0158600 and a histone deacetylase, respectively (Table 1). The rice gene Os02g158600 has recently been identified as the site of a quantitative trait locus for heat stress tolerance [42], while histone deacetylases have been shown to mediate heat stress tolerance in rice and *Arabidopsis* [43, 44]. Together, this suggests that these heat stress genes and the modules they are associated with may represent widely conserved plant responses to heat stress. Pecans are native primarily to the subtropical region of North America [45], and they must endure extreme heat regularly during their normal development cycle. In the transcriptome analysis of the Chinese cultivar 'Annong3', heat shock proteins were also among the most highly expressed genes during this time of year [21]. While 'Tiny Tim' comes from a northern region of the Pecan range, there were no significant genotype differences in these modules. Click or tap here to enter text.

Late Season: Rapid Nutrient Accumulation

The late stage of pecan fruit development is characterized by the solidification of the endosperm and the rapid accumulation of lipids and proteins, resulting in a final nut that is ~ 10% protein and ~ 70% lipid by mass [46]. Previous studies into the transcriptomics of pecan development revealed an abundance of genes related to lipid and storage protein biosynthesis during this time of year [19–21]. Module green, the largest late-season module and module most associated with September, is enriched in lipid metabolism transcripts, as well as organic acid metabolism transcripts, which are upstream of lipid synthesis (Fig. 4a). Modules green and grey60 are rich in highly expressed storage proteins, together making up 8 of the 20 most highly expressed transcripts (Table 2). Other late-season associated modules, purple and white, appear to support these nutrient accumulation activities by protein localization and amino acid metabolism, respectively (Fig. 4b and 4c).

Table 2
Top 20 transcripts with the highest expression over the course of the growing season.

Name	Module color	Log_2Average Expression	Panther Annotation
CiPaw.03G264700.1	grey60	9.145	12S SEED STORAGE PROTEIN CRA1-RELATED
CiPaw.01G184200.1	grey60	9.023	12S SEED STORAGE PROTEIN CRA1-RELATED
CiPaw.16G014600.1	turquoise	9.000	
CiPaw.01G198100.1	green	8.940	Seed Storage and Functional Proteins
CiPaw.12G126700.1	grey60	8.791	2S SEED STORAGE PROTEIN 1-RELATED
CiPaw.03G275900.1	green	8.791	
CiPaw.03G264300.1	grey60	8.713	Seed Storage and Functional Proteins
CiPaw.08G107900.1	green	8.675	Seed Storage and Functional Proteins
CiPaw.01G186200.1	purple	8.646	Plant Proline-Rich Cell Wall Protein
CiPaw.04G072100.1	orange	8.571	BIFUNCTIONAL INHIBITOR/LIPID-TRANSFER PROTEIN/SEED STORAGE 2S ALBUMIN-LIKE PROTEIN
CiPaw.09G102900.1	darkturquoise	8.471	
CiPaw.05G077600.1	turquoise	8.308	PEPTIDYL-PROLYL CIS-TRANS ISOMERASE
CiPaw.13G044200.1	brown	8.269	DNAJ HOMOLOG SUBFAMILY C MEMBER
CiPaw.10G143900.1	royalblue	8.267	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37E-RELATED
CiPaw.03G262000.1	brown	8.254	METALLOTHIONEIN-LIKE PROTEIN 2A
CiPaw.07G058600.1	turquoise	8.206	NARINGENIN,2-OXOGLUTARATE 3-DIOXYGENASE
CiPaw.09G087300.1	darkgrey	8.157	UBIQUITIN
CiPaw.11G151900.1	royalblue	7.986	Plant lipid transfer/defense-related protein
CiPaw.03G264400.1	grey60	7.940	Seed Storage and Functional Proteins
CiPaw.13G096100.1	saddlebrown	7.929	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37E-RELATED

Modules correlated with October plum1 and yellowgreen show enrichment in transcripts related to 1-aminocyclopropane-1-carboxylate metabolism (Fig. 4d and 4e), a precursor to the ripening and senescence hormone ethylene, which signals the conclusion of fruit development. While ethylene is most closely associated with the ripening of climacteric fruits, it is known that ethylene signals developmental processes in diverse non-climacteric fruits as well [47–49]. In walnut, ethylene application preceding harvest stimulates the maturation of the fruit, making the hulling of nuts considerably easier while improving nut quality [50, 51]; it appears the role of ethylene in final nut maturation is shared in pecan.

Genotype: Potential Directions for Understanding Domestication and Fruit Size

Of the 43 modules with over 100 transcripts, 9 showed significant correlation with ‘Tiny Tim’ and 5 showed significant correlation with ‘Mahan’. While very few modules showed nearly exclusive expression (correlation > 0.8) in one genotype, 4 modules (red, lightgreen, darkgreen, and darkorange) did. Among these modules, 3 (red, lightgreen, and darkgreen) appear to be highly enriched in biotic response-related transcripts, reflecting individual differences in disease resistance and response. These modules are unlikely to be associated with the large fruit phenotype. Darkorange, however, appears to have a more diverse functional profile (Fig. 5a) and is associated very strongly with ‘Tiny Tim’ (correlation = 0.82).

Notably, the most overrepresented biological function in this module is potassium transport, and the hub gene is an ion transporter. Potassium and other ions are known to impact fruit quality and nutrition metrics such as soluble solid content, amino acid content, and vitamin C in different crops, including kiwifruit and pear [52, 53]. These genes may also impact fruit phenotypes beyond size.

As fruit sizing in pecan is primarily completed in the first half of fruit development [4], modules associated with both ‘Mahan’ and earlier months of the fruit development cycle (May, June, July) might be of greater interest in understanding the large fruit phenotype. Three modules fit this categorization: darkred, orange, and lightsteelblue1. Module lightsteelblue1, strongly associated with May, appears to be enriched primarily in terpenoid biosynthesis (Fig. 6a). Terpenoid biosynthesis is generally associated with plant defense against herbivory and disease, as well as tolerance to abiotic stress, and may play a role in genotype-dependent stress responses. However, evidence for its role in size determination is lacking. Module orange primarily showed enrichment of transcripts related to polysaccharide biosynthesis, cell wall biosynthesis, and cell wall organization (Fig. 6b). This module appears to be a genotype-specific counterpart to module brown, which is also most active in July and highly enriched in transcripts related to cell wall biosynthesis. The chemical and mechanical properties of the plant cell wall are factors that influence the ability of plant cells to grow, and remodeling is an ongoing process in growing plant tissues [54, 55].

In walnut, alleles in genes with functionality in cell wall modification and loosening during growth were associated with increased nut volume, and in pear, cell wall biosynthesis-related genes were identified as likely candidates for regulating fruit size [16, 56]. The enriched functions of the module darkred were more diverse (Fig. 6c), though functions related to RNA modification, inositol metabolism, and phospholipid metabolism are recurrent. Inositol is a known regulator of plant cell growth and elongation [57]. This module shows the highest expression at a time when cell division is slowing down in the developing fruit. Yet fruit expansion is ongoing, making this a promising potential process involved in the development of pecan fruit size. Hub genes and other genes with high module membership are likely to be drivers of module function. Among the transcripts with the highest module membership in modules orange and darkred, many exhibit regulatory function, which may contribute to the large fruit phenotype (Tables 1 and Supplementary Tables 4 and 5). Notably, the hub gene of the darkred module is a homolog of pleotropic regulatory locus 1 (PRL1), which has demonstrated relationships with organ size, carbon partitioning, and hormone response in *Arabidopsis* [58]. These genes are worthy of further investigation as potential drivers of fruit size.

Examination of early season ‘Tiny Tim’ associated modules may also give insight into the drivers of small fruit size. Modules yellow, tan, salmon, and lightcyan fit these criteria. Module lightcyan appears to be primarily enriched in transcripts related to biotic stress response (Fig. 5b). Module yellow appears to have many stress tolerance-related transcripts, but also has a notable elevation in photosynthesis-related transcripts as well (Figs. 5c and 5d). The role of photosynthesis in fruit development is quite variable between plant species, but is generally associated with increased oxygen supply, refixing inorganic carbon in the highly metabolically active fruit tissues, and supplying carbon intermediates to diverse biosynthetic pathways [59]. It is plausible that the small fruit phenotype of ‘Tiny Tim’ exists despite the elevated photosynthetic transcripts, rather than because of it.

Module salmon appears to be associated with proteolysis and catabolism mostly strongly; in Arabidopsis, protein degradation pathways are associated with a decrease in reproductive organ sizes (Li et al., 2008). The hub gene for module tan is FCA (Table 1), which has demonstrated interactions with abscisic acid-related transcription factors, enhancing the effect of ABA on ABA-responsive genes [61, 62]. ABA response is associated with decreased fruit size, indicating a role for this module in producing the small-fruit phenotype of ‘Tiny Tim’.

While some of these genotype-correlated genes may be involved in the development of the large fruit phenotype in pecan, the selection of only two genotypes limits the broad applicability of these results. It is also possible that biological processes external to the fruit are key drivers of this phenotype, which may not be captured in a fruit-only experiment. In tomato, for example, the final fruit weight may be influenced by the expression of genes during flowering [63].

Conclusion

To the best of our knowledge, this study is the first comprehensive evaluation of pecan fruit transcriptomics across multiple genotypes for an entire season. The early season was most strongly associated with broad anabolic processes and cell cycle functions. The mid-season was associated with cell wall development and heat stress response, and the late season was most strongly associated with lipid metabolism, transport, protein localization, and storage protein biosynthesis. While transcript expression varied enormously through the development of the pecan fruit and nut, a much smaller fraction was significantly correlated with genotype, and many of those associated with genotype appear to have roles in biotic stress response. Among these genotype-specific transcripts are potential candidates for further research into the molecular regulation of pecan fruit size, with inositol metabolism, ABA response, proteolysis, and cell wall biosynthesis and organization being particularly noteworthy biological processes.

Declarations

Competing Interests:

Authors declare no competing interests.

Funding

This research was supported by Texas A&M AgriLife Hatch Project #TEX0-9950-0 and startup funds from Texas A&M AgriLife Research and Texas A&M University to AD. JL acknowledges graduate research assistantship support from the Department of Horticultural Sciences at Texas A&M University. The work was also funded in part by the U.S. Department of Agriculture – Agriculture Research Service National Programs through CRIS project 3091-21000-046-000-D (Crop Germplasm Research Unit, TX) and a USDA ARS NACA #58-3091-3-013 award to WC and AD. The agency was not involved in the study design, collection, analysis, interpretation of data and the writing of this article. However, this manuscript was approved by the agency before submission for publication. This article reports on the results of the research only. Mention of a trademark or proprietary product is solely to provide specific information and does not constitute a guarantee or warranty of the product by the U.S. Department of Agriculture and does not imply its approval to the exclusion of other products that may also be suitable.

Author Contribution

AD and WC designed the study. WC and JL collected the samples. JL performed all the experiments and data analysis. JL and AD prepared the first draft of the manuscript. AD supervised the study. All authors reviewed, edited, and approved of the manuscript.

Data Availability

The RNA sequencing dataset generated and analyzed in this study are available at the National Center for Biotechnology Information Short Read Archive (NCBI SRA) repository under BioProject [PRJNA1312749](https:/dataview.ncbi.nlm.nih.gov/object/PRJNA1312749?reviewer=emp4pqujgjn52ip8uvqphfpsc4)
.

References

1. Manning, W. E. *The Morphology of the Flowers of the Juglandaceae. II. The Pistillate Flowers and Fruit*. at (1940). https://about.jstor.org/terms
2. Manning, W. E. & Missouri Botanical Garden Press. *The Classification Within the Juglandaceae*. (, at (1978). http://www.biodiversitylibrary.org/bibliography/702v.651978:http://www.biodiversitylibrary.org/item/89023 http://www.biodiversitylibrary.org/pdf3/008639400089023
3. Herrera, E. A. Fruit growth and development of 'Ideal'and 'Western'pecans. *J. Am. Soc. Hortic. Sci.* **115**, 915–923 (1990).
4. Herrera, E. A. Fruit growth and development of 'Ideal'and 'Western'pecans. *J. Am. Soc. Hortic. Sci.* **115**, 915–923 (1990).
5. Gottwald, T. R. & Bertrand, P. F. on Pecan Scab Development and Nut Quality. *Phytopathology* **73**, 714–718 (1983).
6. Wood, B. W. & Reilly, C. C. Factors influencing water split of pecan fruit. (1999).
7. Taylor, N. J. & QUANTIFYING WATER USE OF MATURE PECAN NUT TREES AND ORCHARDS IN SELECTED IRRIGATION AREAS OF THE NORTHERN CAPE. (2024).
8. Wood, B. W. REGULATION OF VIVIPARY IN PECAN. in Acta Hortic 33–42 (International Society for Horticultural Science (ISHS), Leuven, Belgium, doi:10.17660/ActaHortic.2015.1070.3 (2015).
9. True, R. H. *Notes on the early history of the pecan in America* (US Government Printing Office, 1919).
10. Thompson, T. & Grauke, L. *Pecan Tree Growth and Precocity*. *J AMER SOC. HORT SCI* **128**, (2003).
11. University of Georgia. Alphabetical Cultivar List. UGA Pecan Breeding. (2024).
12. United States Department of Agriculture. NCGR Pecan/Hickory Database. (2024).
13. Sparks, D. *Adaptability of Pecan as a Species*. *HORTSCIENCE* 40, (2005).
14. Hussain, Q. et al. Genetic and signalling pathways of dry fruit size: targets for genome editing-based crop improvement. *Plant. Biotechnol. J.* **18**, 1124–1140 (2020).
15. Mauxion, J. P., Chevalier, C. & Gonzalez, N. Complex cellular and molecular events determining fruit size. *Trends Plant. Sci.* **26**, 1023–1038 (2021).
16. Bernard, A. et al. Genome-Wide Association Study Reveals Candidate Genes Involved in Fruit Trait Variation in Persian Walnut (Juglans regia L). *Front Plant. Sci* **11**, (2021).
17. Wang, J. et al. Genome-wide association analysis of 101 accessions dissects the genetic basis of shell thickness for genetic improvement in Persian walnut (Juglans regia L). *BMC Plant. Biol.* **22**, 436 (2022).
18. Khorami, S. S., Arzani, K., Karimzadeh, G., Shojaeiyan, A. & Ligterink, W. Genome size: A novel predictor of nut weight and nut size of walnut trees. *HortScience* **53**, 275–282 (2018).
19. Huang, R., Huang, Y., Sun, Z., Huang, J. & Wang, Z. Transcriptome Analysis of Genes Involved in Lipid Biosynthesis in the Developing Embryo of Pecan (Carya illinoinensis). *J. Agric. Food Chem.* **65**, 4223–4236 (2017).
20. Mattison, C. P. et al. RNA-Seq Analysis of Developing Pecan (Carya illinoinensis) Embryos Reveals Parallel Expression Patterns among Allergen and Lipid Metabolism Genes. *J. Agric. Food Chem.* **65**, 1443–1455 (2017).
21. Xu, Z. et al. Transcriptome analysis of pecan seeds at different developing stages and identification of key genes involved in lipid metabolism. *PLoS One*. **13**, e0195913 (2018).

22. Lovell, J. T. et al. Four chromosome scale genomes and a pan-genome annotation to accelerate pecan tree breeding. *Nat Commun* **12**, (2021).
23. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, (2008).
24. Kulmanov, M. et al. Protein function prediction as approximate semantic entailment. *Nat. Mach. Intell.* **6**, 220–228 (2024).
25. Huang, R. et al. Transcriptome Analysis of Walnut (Juglans regia L.) Embryos Reveals Key Developmental Stages and Genes Involved in Lipid Biosynthesis and Polyunsaturated Fatty Acid Metabolism. *J. Agric. Food Chem.* **69**, 377–396 (2021).
26. Yu, A. et al. Global Transcriptomic Analyses Provide New Insight into the Molecular Mechanisms of Endocarp Formation and Development in Iron Walnut (Juglans sigillata Dode). *Int J. Mol. Sci* **24**, (2023).
27. Miura, H., Imada, S. & Yabuuchi, S. Double Sigmoid Growth Curve of Strawberry Fruit. *J. Japanese Soc. Hortic. Sci.* **59**, 527–531 (1990).
28. Fischer, W. N., Kwart, M., Hummel, S. & Frommer, W. B. Substrate specificity and expression profile of amino acid transporters (AAPs) in Arabidopsis. *J. Biol. Chem.* **270**, 16315–16320 (1995).
29. Su, Y. H., Frommer, W. B. & Ludewig, U. Molecular and Functional Characterization of a Family of Amino Acid Transporters from Arabidopsis. *Plant. Physiol.* **136**, 3104–3113 (2004).
30. Zhang, L. et al. Altered xylem-phloem transfer of amino acids affects metabolism and leads to increased seed yield and oil content in Arabidopsis. *Plant. Cell.* **22**, 3603–3620 (2010).
31. Sanders, A. et al. AAP1 regulates import of amino acids into developing Arabidopsis embryos. *Plant J.* **59**, 540–552 (2009).
32. Ji, Y., Huang, W., Wu, B., Fang, Z. & Wang, X. The amino acid transporter AAP1 mediates growth and grain yield by regulating neutral amino acid uptake and reallocation in Oryza sativa. *J. Exp. Bot.* **71**, 4763–4777 (2020).
33. Wolswinkel, P. Transport of nutrients into developing seeds: a review of physiological mechanisms. *Seed Sci. Res.* **2**, 59–73 (1992).
34. Allison, J. M., Prussia, S. E., Daniell, J. W. & Tollner, E. W. Measuring the internal pressure of maturing pecans. *Trans. ASAE*. **30**, 1869–1873 (1987).
35. Wang, Y. et al. Advances in mechanisms and omics pertaining to fruit cracking in horticultural plants. *Agronomy* 11, Preprint at (2021). https://doi.org/10.3390/agronomy11061045
36. Khadivi-Khub, A. Physiological and genetic factors influencing fruit cracking. *Acta Physiol. Plant.* **37**, 1718 (2014).
37. Wang, J. et al. The amino acid permease 5 (Osaap5) regulates tiller number and grain yield in rice. *Plant. Physiol.* **180**, 1031–1045 (2019).
38. Fogelman, E., Stern, R. A. & Ginzberg, I. Benzyladenine and gibberellin treatment of developing Pink Lady apples results in mature fruits with a thicker cuticle comprising clusters of epidermal cells. *Protoplasma* **252**, 1009–1017 (2015).
39. Sahu, P., Sharma, N. & Prativa Sahu, C. Fruit cracking and quality of pomegranate (Punica granatum L.) cv. Kandhari as influenced by CPPU and boron. *~ 2644 ~ J. Pharmacognosy Phytochemistry* **8**, (2019).
40. Jiang, Y. L. & Chang, P. T. Reducing Splitting of 'Murcott' Tangor Fruit (Citrus reticulate × Citrus sinensis) through Foliar Application of Forchlorfenuron (CPPU). *Horticulturae* 10, (2024).
41. Fahima, A. et al. Cytokinin treatment modifies litchi fruit pericarp anatomy leading to reduced susceptibility to post-harvest pericarp browning. *Plant Sci.* **283**, 41–50 (2019).
42. Chimthai, S. et al. Breeding for Heat Tolerant Aromatic Rice Varieties and Identification of Novel QTL Regions Associated with Heat Tolerance During Reproductive Phase by QTL-Seq. *Rice Sci.* **32**, 67–80 (2025).
43. Yang, F. et al. Plant-specific histone deacetylases associate with ARGONAUTE4 to promote heterochromatin stabilization and plant heat tolerance. *New Phytol.* **238**, 252–269 (2023).
44. Chen, Z. et al. A histone deacetylase confers plant tolerance to heat stress by controlling protein lysine deacetylation and stress granule formation in rice. *Cell Rep* **43**, (2024).
45. Thompson, T. E., Grauke, L. J., PECANS & AND OTHER HICKORIES (CARYA).. in *Acta Hortic* 839–906 (International Society for Horticultural Science (ISHS), Leuven, Belgium, (1991). 10.17660/ActaHortic.1991.290.19
46. Venkatachalam, M., Roux, K. H. & Sathe, S. K. Biochemical Characterization of Soluble Proteins in Pecan [Carya illinoinensis (Wangenh.) K. Koch]. *J. Agric. Food Chem.* **56**, 8103–8110 (2008).
47. Li, L. et al. Effects of the ethylene-action inhibitor 1-methylcyclopropene on postharvest quality of non-climacteric fruit crops. *Postharvest Biol. Technol.* **111**, 322–329 (2016).
48. Paul, V., Pandey, R. & Srivastava, G. C. The fading distinctions between classical patterns of ripening in climacteric and non-climacteric fruit and the ubiquity of ethylene—An overview. *J. Food Sci. Technol.* **49**, 1–21 (2012).
49. Chervin, C. et al. Ethylene seems required for the berry development and ripening in grape, a non-climacteric fruit. *Plant Sci.* **167**, 1301–1305 (2004).
50. Sorber, D. G. & Kimball, M. H. Use of ethylene in harvesting the Persian walnut (Juglans regia) in California. (1950).
51. Farooq, A., Hussain, S. Z., Bhat, T. A., Naseer, B. & Shafi, F. Walnut fruit: Impact of ethylene assisted hulling on in vitro antioxidant activity, polyphenols, PUFAs, amino acids and sensory attributes. *Food Chem.* **404**, 134763 (2023).
52. Shen, C. et al. Effects of foliar potassium fertilization on fruit growth rate, potassium accumulation, yield, and quality of 'Kousui'japanese pear. *Horttechnology* **26**, 270–277 (2016).
53. Ma, L., Chi, C., Lv, S., Tong, Y. & Yang, L. Effects of different sources of potassium fertiliser on yield, fruit quality and nutrient absorption in 'harward' kiwifruit (Actinidia deliciosa). *Open Life Sci* **20**, (2025).
54. Chebli, Y. & Geitmann, A. Cellular growth in plants requires regulation of cell wall biochemistry. *Curr. Opin. Cell. Biol.* **44**, 28–35 (2017).

55. Cosgrove, D. J. Structure and growth of plant cell walls. *Nat Rev Mol Cell Biol* 25, 340–358 Preprint at (2024). https://doi.org/10.1038/s41580-023-00691-y
56. Wang, G. et al. Transcriptome analysis reveals gene associated with fruit size during fruit development in pear. *Sci. Hortic.* **305**, 111367 (2022).
57. Stevenson, J. M., Perera, I. Y., Heilmann, I., Persson, S. & Boss, W. F. Inositol signaling and plant growth. *Trends Plant. Sci.* **5**, 252–258 (2000).
58. Németh, K. et al. Pleiotropic control of glucose and hormone responses by PRL1, a nuclear WD protein, in Arabidopsis. *Genes Dev.* **12**, 3059–3073 (1998).
59. Garrido, A., Conde, A., Serôdio, J., De Vos, R. C. H. & Cunha, A. Fruit Photosynthesis: More to Know about Where, How and Why. *Plants* 12, Preprint at (2023). https://doi.org/10.3390/plants12132393
60. Li, Y., Zheng, L., Corke, F., Smith, C. & Bevan, M. W. Control of final seed and organ size by the DA1 gene family in Arabidopsis thaliana. *Genes Dev.* **22**, 1331–1336 (2008).
61. Lee, S., Lee, H. J., Jung, J. H. & Park, C. M. The rabidopsis thaliana RNA-binding protein FCA regulates thermotolerance by modulating the detoxification of reactive oxygen species. *New Phytol.* **205**, 555–569 (2015).
62. Kumar, S., Jiang, S., Jami, S. K. & Hill, R. D. Cloning and characterization of barley caryopsis FCA. *Physiol. Plant.* **143**, 93–106 (2011).
63. van der Knaap, E. et al. What lies beyond the eye: The molecular mechanisms regulating tomato fruit weight and shape. *Front Plant. Sci* **5**, (2014).

Figures

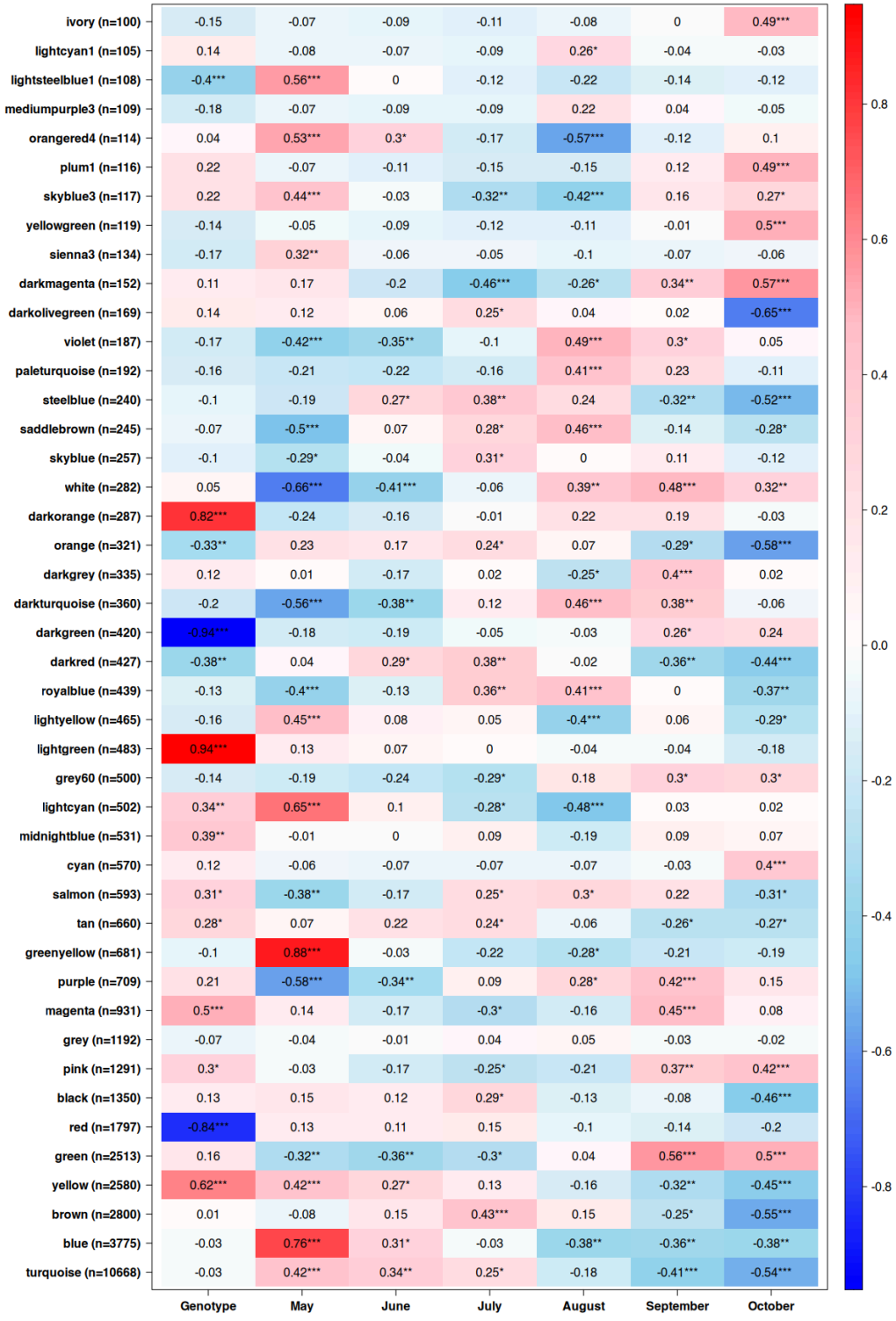

Figure 1

Module-trait correlations for all modules containing more than 100 transcripts. In the “Genotype” column, negative values correspond to ‘Mahan’, while positive values correspond to ‘Tiny Tim’. Number of transcripts per module indicated to the right of the module name. Significance markers: * - p<0.05, ** - p<0.01, *** - p <0.001.

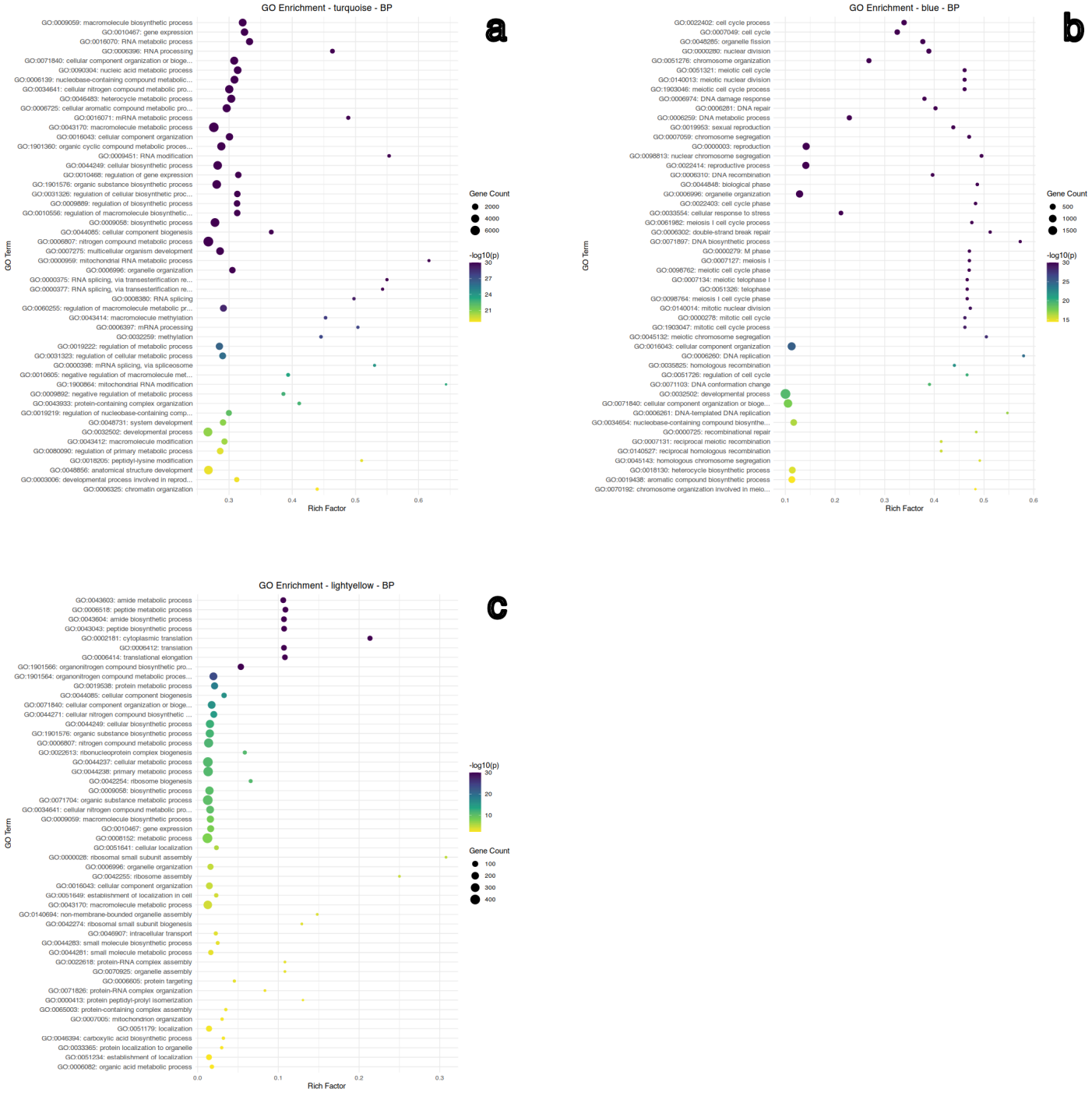

Figure 2

Gene ontology term enrichment analysis for biological processes in early season specific modules a) turquoise, b) blue, and c) lightyellow.

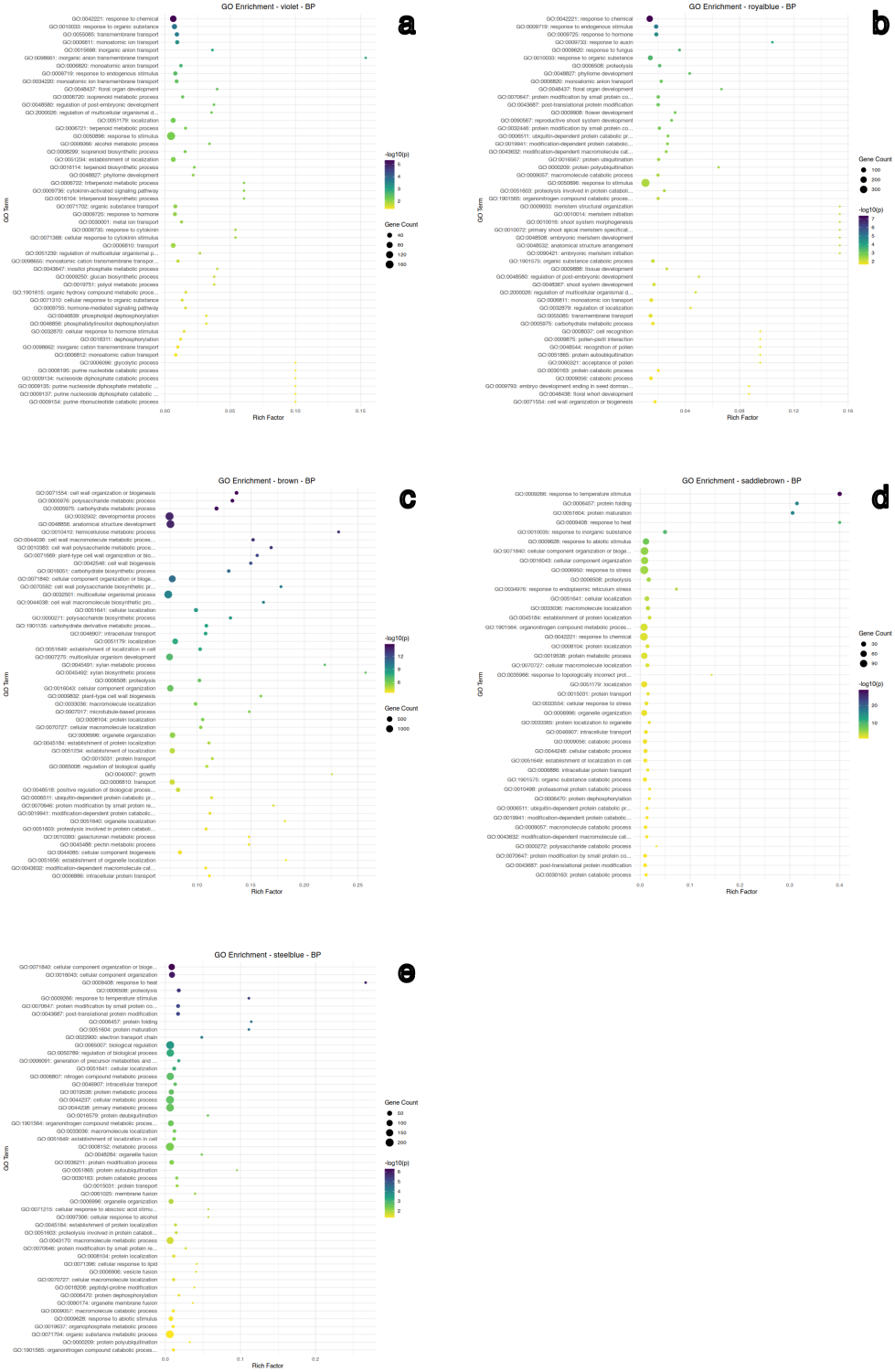

Figure 3

Gene ontology term enrichment analysis for biological processes in mid-season season specific modules a) violet, b) royalblue, c) brown, d) saddlebrown, and e) steelblue.

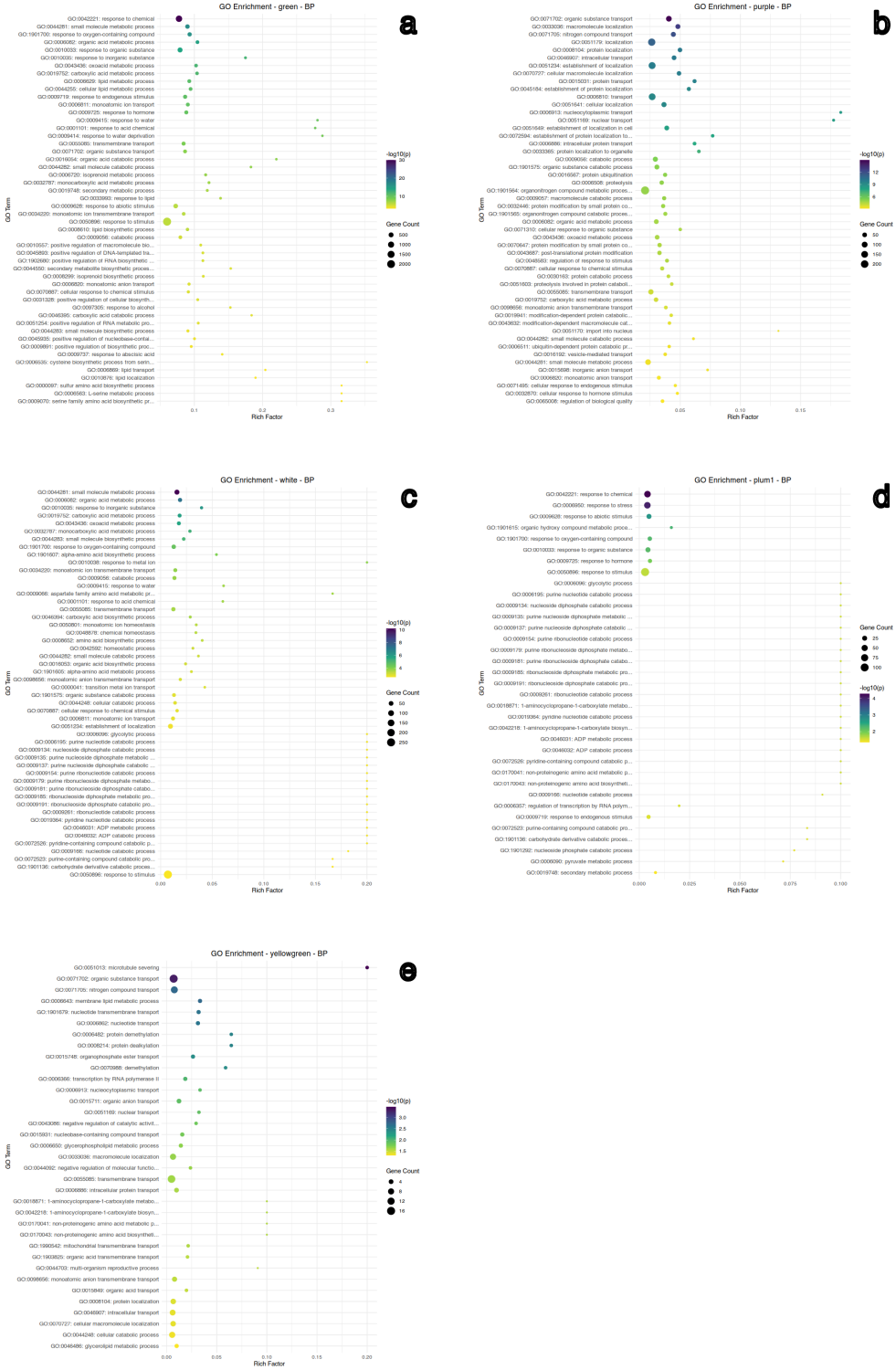

Figure 4

Gene ontology term enrichment analysis for biological processes in late season specific modules a) green, b) purple, c) white, d) plum1, and e) yellowgreen.

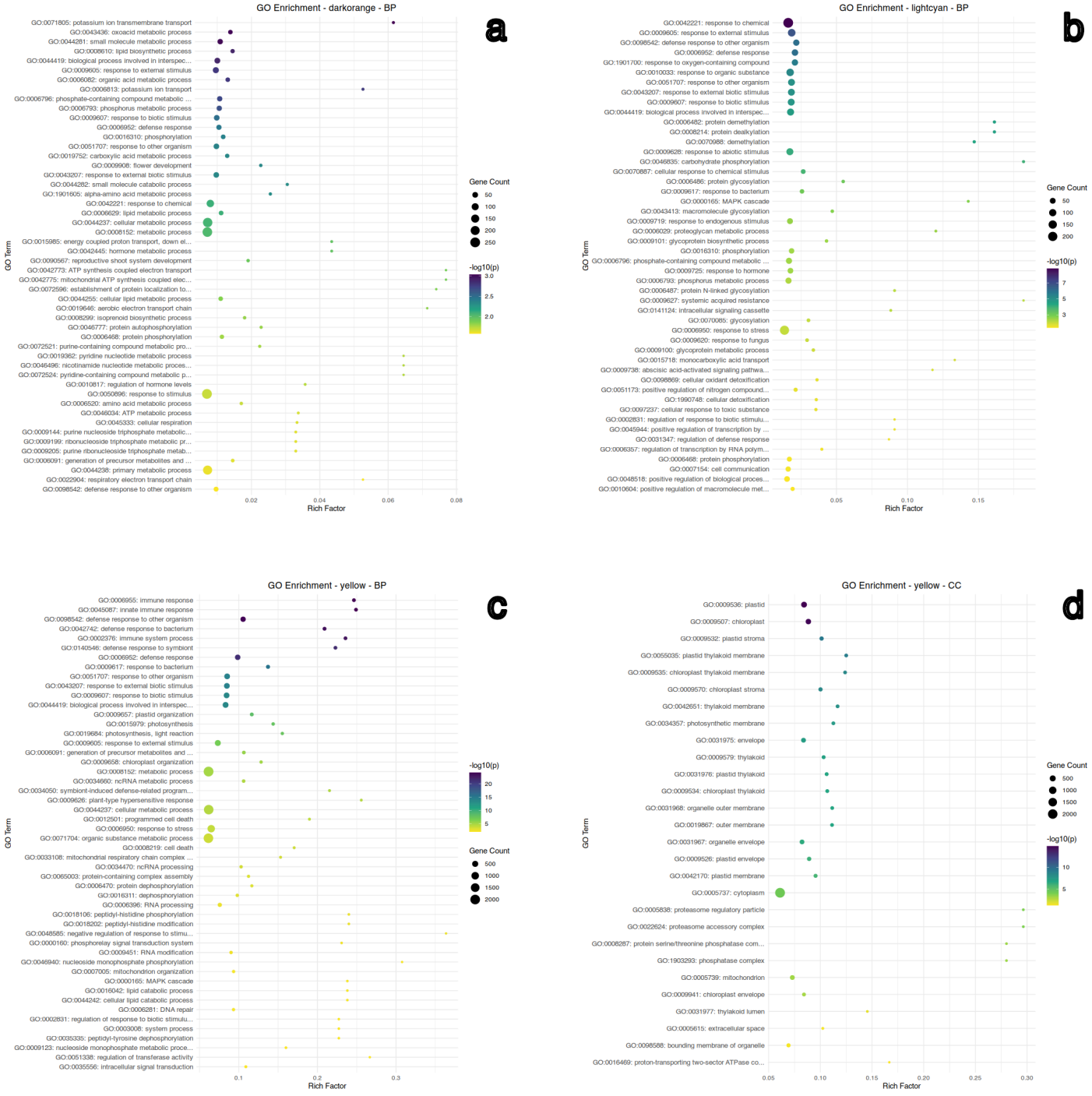

Figure 5

Gene ontology term enrichment analysis for biological processes in 'Tiny Tim' specific modules a) darkorange, b) lightcyan and c) yellow. Gene ontology term enrichment analysis for cellular components are provided for module d) yellow.

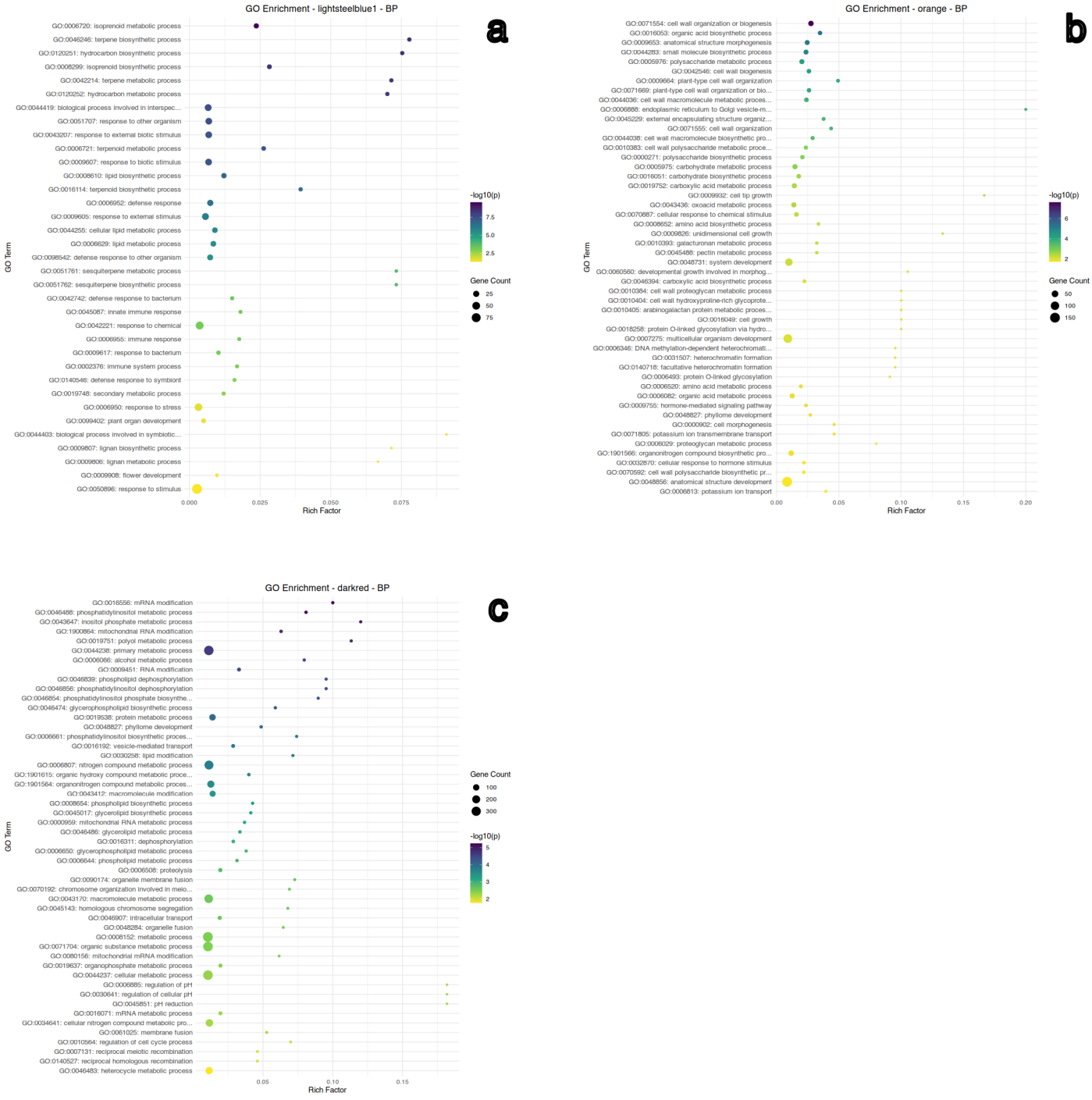

Figure 6

Gene ontology term enrichment analysis for biological processes in 'Mahan' specific modules a) lightsteelblue1, b) orange, and c) darkred.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SRSupData.pdf