Analysis of protein-protein interaction and weighted co-expression networks revealed key modules and genes in multiple tissues of *Agave sisalana*

Lucas Miguel de Carvalho (lucasmigueel@gmail.com)
Center for Computing in Engineering and Sciences, State University of Campinas, Campinas, São Paulo, 13083-861, Brazil.
https://orcid.org/0000-0002-8766-0452

Nicholas Vinícius Silva
Laboratory of Genomics and bioEnergy (LGE), Department of Genetics, Evolution, Microbiology, and Immunology, Institute of Biology, State University of Campinas, Campinas, SP, 13083-862, Brazil

Luís Guilherme F. de Abreu
Laboratory of Genomics and bioEnergy (LGE), Department of Genetics, Evolution, Microbiology, and Immunology, Institute of Biology, State University of Campinas, Campinas, SP, 13083-862, Brazil

Marina Püpke Marone
Laboratory of Genomics and bioEnergy (LGE), Department of Genetics, Evolution, Microbiology, and Immunology, Institute of Biology, State University of Campinas, Campinas, SP, 13083-862, Brazil

Alexandra Russolo Cardelli
Laboratory of Genomics and bioEnergy (LGE), Department of Genetics, Evolution, Microbiology, and Immunology, Institute of Biology, State University of Campinas, Campinas, SP, 13083-862, Brazil

Fabio Trigo Raya
Laboratory of Genomics and bioEnergy (LGE), Department of Genetics, Evolution, Microbiology, and Immunology, Institute of Biology, State University of Campinas, Campinas, SP, 13083-862, Brazil

Guido Araújo
Institute of Computing, UNICAMP, Campinas, São Paulo, 13083-852, Brazil.

Marcelo Falsarella Carazzolle
Center for Computing in Engineering and Sciences, State University of Campinas, Campinas, São Paulo, 13083-861, Brazil.

Gonçalo Amarante Guimarães Pereira (goncalo@unicamp.br)
Laboratory of Genomics and bioEnergy (LGE), Department of Genetics, Evolution, Microbiology, and Immunology, Institute of Biology, State University of Campinas, Campinas, SP, 13083-862, Brazil

Research Article

Keywords: Agave, Interactome, Co-expression analysis, Protein-protein interaction network, Sisal

Posted Date: September 7th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2031343/v1

Version of Record: A version of this preprint was published at Frontiers in Chemical Engineering on May 26th, 2023. See the published version at https://doi.org/10.3389/fceng.2023.1175235.

Abstract

Agave plants are well-known for their drought resilience and commercial applications. Among them, *Agave sisalana* (sisal) is the species most used to produce hard fibers, and it is of great importance for semiarid regions. Agaves also show potential as bioenergy feedstocks, as they can accumulate large amounts of biomass and fermentable sugar. This study aimed to reconstruct the *A. sisalana* interactome, and identify key genes and modules involved in multiple plant tissues (root, stem, and leaf) through RNA-Seq analysis. We integrated *A. sisalana* transcriptome sequences and gene expression generated from stem, leaf, and root tissues to build global and conditional co-expression networks across the entire transcriptome. By combining the co-expression network, module classification, and function enrichment tools, we identified 20 functional modules related to at least one *A. sisalana* tissue, covering functions such as photosynthesis, leaf formation, auxin-activated signaling pathway, floral organ abscission, response to farnesol, brassinosteroid mediated signaling pathway, and light-harvesting. The final interactome of *A. sisalana* contains 2,582 nodes and 15,083 edges. In the reconstructed interactome, we identified submodules related to plant processes to validate the reconstruction. In addition, we identified 6 hub genes that were searched for in the co-expression modules. The intersection of hub genes identified by both the protein-protein interaction networks (PPI networks) and co-expression analyses using gene significance and module membership revealed six potential candidate genes for key genes. In conclusion, we identified six potential key genes for specific studies in *Agave* transcriptome atlas studies, biological processes related to plant survival in unfavorable environments, and provide strategies for breeding programs.

Highlights

- We built the first interactome of *Agave sisalana*, the main species used in semiarid regions to produce sisal fibers.

- PPI network data integration and co-expression analysis showed insights for viable target genes in future *Agave* studies.

- The construction of *Agave* co-expression networks to improve relationships between gene expression and tissues.

- A group of genes positively correlated with stem play an abscission role, suggesting they might be floral repression genes, which can be part of further improvement studies to accelerate or repress flowering.

Introduction

Plants of the genus *Agave* are known for their drought resistance mechanisms and commercial applications. Agaves are native to semiarid and arid regions of North and Central America and possess several morphological, anatomical, and physiological drought resistance mechanisms (Eguiarte et al., 2021). One of the most important is CAM (crassulacean acid metabolism), the most water-use efficient type of photosynthesis (Borland et al., 2014; Yin et al., 2018). Also, several *Agave* species have established commercial uses, for instance, the production of tequila (*Agave tequilana*), mezcal (*Agave angustifolia* and others), and sisal fibers (mainly *Agave sisalana*). Sisal is the main source of natural hard fibers, and, due to its drought resistance, it has a great social-economical impact on semiarid regions (Broeren et al., 2017; Monja-Mio et al., 2019). In addition, *A. sisalana* has interesting characteristics to be used as a bioenergy feedstock in semi-arid regions (Davis et al., 2011), since only 4% of *A. sisalana* leaves are used to produce commercial fibers (Suinaga et al., 2007). The remaining waste could be used to produce biogas or second-generation ethanol. Brazil is the greatest sisal fiber producer in the world and has the capacity of increasing the area in which this crop is produced to keep up with the increasing demand for biofuels (International Energy Agency (IEA), 2021).

Transcriptomic atlases (Raya et al., 2021; Sarwar et al., 2019) and phylogeny and evolution studies of *A. sisalana* (Jiménez-Barron et al., 2020) are available and provide valuable information about its genetics. However, there are no analyses or available data on protein-protein interactions (PPI) in any *Agave* species. PPI networks represent physical interactions between proteins, which can lead to useful insights into cell physiology and help identify targets for genetic engineering (Van den Broeck, L. et al., 2017; de Silva, K. K. et al., 2022; Ferreira, S.S. et al., 2016; Yu, B. et al., 2020; Khan, Z.H. et al., 2020). PPI associations, for example, may represent a functional cell action and have already been used in several areas of the biological sciences (Braun et al., 2013; Ding and Kihara, 2019; Yang et al., 2019; Yeger-Lotem and Sharan, 2015). High-throughput technologies, such as mass spectrometers and genomic sequencers, generated a large amount of data allowing for an increase in the number of known protein interactions of various organisms, especially model organisms, and are available in several databases, such as STRING (Szklarczyk et al., 2011).

However, the interactions between proteins available to non-model organisms are very limited, so the complete mapping depends on bioinformatics strategies that aim to assign relationships through orthologs (Uhrig, 2006). These strategies for assembling PPI networks assume that proteins are conserved across different species, and therefore, the interactions between two proteins may also exist in those species (Wang and Jin, 2017). Here, we predict *A. sisalana* interactome using public transcriptome sequences generated from leaf, stem, and root tissues. Moreover, we integrated network analysis with weighted gene coexpression network analysis to identify correlations between genes across tissues based on coexpression relationships focous on point out key genes and modules, which will aid current research and provide a framework for future *Agave* interactomics research.

Material And Methods

Transcriptomic analysis

The assembled transcripts and the raw RNA-Seq reads of three tissues (biological triplicates of leaf, stem, and root) of *A. sisalana* were downloaded from NCBI SRA (accession number PRJNA746623). Poor quality reads and adapters were removed with the Trimmomatic v.0.39 (Bolger et al., 2014) software. Quantification of the assembled transcripts was performed by the software kallisto v.1.0.4 (Bray et al., 2016) and normalized by TMM using the edgeR package v3.38.2 (Robinson et al., 2010). The ORFs of the assembled transcripts were identified by the Transdecoder software with the parameter -m 300 (ORFs longer than 300 nucleotides). Transcripts that did not have ORFs and whose expressions were less than 1 TPM in all conditions were removed. After this removal step, a new round of quantification was performed. *A. sisalana* proteins were obtained from the translation of ORFs.

Grouping of orthologous families

Five organisms were selected to construct the orthogroups, together with *A. sisalana* proteins, based on their wide availability of PPI data in the STRING v.11 database: *Arabidopsis thaliana, Musa acuminata*, *Oryza sativa*, *Sorghum bicolor*, and *Zea mays*. Protein sequences from each organism were retrieved from the NCBI website. OrthoFinder v.2.3.3 software was used (Emms and Kelly, 2019) to construct orthogroups (gene families) from these organisms. From the gene families obtained, we selected those that are considered one-to-one or one-to-many. One-to-one gene families are those that have one copy of a gene in *A. sisalana* and one copy in every other organism. The one-to-many gene families mean that the gene of interest has one copy in *A. sisalana* and more than one ortholog in the other species (Zahn-Zabal et al., 2020). The reconstruction of the phylogenetic tree between the species was also generated by OrthoFinder from the default parameters.

Construction of A. sisalana interactome

For each gene family containing at least one *A. sisalana* protein and proteins from other organisms, we identified the PPIs present in each orthologous organism in the STRING database, considering a cutoff score of 0.70 for each protein interaction. For example, if an *A. sisalana* gene A has as orthologs the gene B in *S. bicolor* and the gene C in *M. acuminata*, we assigned the interactions in the genes B and C to the gene A. Thus, the interactions are mapped according to their orthologs in *A. sisalana*. In the end, only interactions that have complete pairs with orthologs in *A. sisalana* were selected. To visualize the PPI network in *A. sisalana*, the predicted PPI data were loaded into Cytoscape v3.2.1 (Shannon et al., 2003).

Network analysis

After constructing the *A. sisalana* interactome and loading it into Cytoscape, the network metrics were analyzed with Network Analyzer v4.4.8. The top hub genes were selected from the intersection of top20 hub genes derived from three metrics: betweenness centrality, closeness centrality, and node connect degree, using Cytohubba v0.1 (Chin, C.H. et al., 2014). The submodules were identified using MCODE v2.0.2 in Cytoscape with a cutoff of 5 in Mcode score.

Gene ontology (GO) analysis of PPI data

A. sisalana proteins were annotated using PANNZER2 software (Törönen et al., 2018). A table < gene > < go_id > was generated from the results, considering the Gene Ontology class "Biological Process (BP)" in the output of PANNZER2. The enrichment of biological processes was performed using the topGO package v2.48.0 (Alexa et al., 2006) using R 4.1.0.

KEGG Orthology annotation and enrichment

The Kyoto Encyclopedia of Gene and Genomes (KEGG) orthology (KO) annotations were assigned using KofamKOALA 1.0.0 (Aramaki et al., 2020), which uses hmmsearch 3.1b2 against curated hidden Markov model (HMM) KO profiles. The enriched KEGG pathways and their modules have been identified through the enrichKEGG and enrichMKEGG functions of the ClusterProfiler package v4.4.4 (Wu et al., 2021) using R 4.1.0.

Tissue-specific analysis

Tau metric was used to measure expression specificity among all genes present in the *A. sisalana* interactome with the Tspex package (Camargo, A.P. et al., 2020). Tau showed a robust behavior according to data normalization (Kryuchkova-Mostacci and Robinson-Rechavi, 2017), in our case, through the TMM normalization in gene expression data. We considered genes with Tau values greater than 0.90 as tissue specific.

Weighted correlation network analysis (WGCNA)

Gene expression matrix (normalized gene expression of biological replicates of leaf, stem, and root tissues) was used to construct an unsigned weighted gene co-expression network using the *WGCNA* R package v1.71 (Langfelder and Horvath, 2008). Hierarchical clustering of samples was conducted to remove outliers with a cut-off value of 80 to produce two stable clusters. Then, the soft threshold power β was determined to ensure a scale-free network. The power function was used to turn the Pearson correlation matrix into an adjacency matrix, which was then transformed

into a topological overlap matrix (TOM). Using a dynamic cutting algorithm, a hierarchical clustering was performed to cluster similar genes into the same module. Subsequently, we clustered the eigengenes according to the relationship and merged them into modules with the association > 0.80. The association of each module with gene expressions in the tissues was evaluated based on Pearson correlation (|cor| > 0.5 and p-value < = 0.05). For each gene, module membership (MM) was characterized according to the association between module eigengene (ME) and its expression level. The association between gene expression and tissues was represented by gene significance (GS). Thus, for each module, GS and MM for each gene were computed and considered significant if |GS| > 0.2 and MM > 0.8.

Results And Discussion

A detailed outline of our study is summarized in Figure 1. First, we conducted an analysis of orthologous genes between several plant species and the transcriptome assembly of *A. sisalana*. From the identification of ortholog families, we mapped the interactions of these pairs of orthologs in the STRING database. In the end, we have the reconstructed *A. sisalana* interactome. From its reconstruction, we identified the interactome submodules and performed its GO enrichment to search for modules with specific functions in *A. sisalana* from known databases. In addition, we performed the analysis of hub genes, which play an important role in the interactome.

Likewise, we performed weighted gene co-expression network analysis from stem, leaf, and root tissues studied in *A. sisalana* transcriptome. From this analysis, we can identify the key moidules, which are co-expression modules correlated with each tissue, and make hypotheses about the functionalities observed in the plant.

Finally, we cross-referenced the lists of hub genes with their presence in co-expression modules, evaluating module membership (MM) and gene significance (GS). If these hub genes pass through the co-expression module evaluation cutoff, they were considered key genes.

All steps and results present in the flowchart is described in next sections.

A. sisalana orthogroups

Initially, a total of 23,794 *Agave sisalana* proteins (extracted from the transcriptomic atlas) were grouped into orthologous families using five plant species (*Arabidopsis thaliana, Musa acuminata*, *Oryza sativa*, *Sorghum bicolor*, and *Zea mays*), selected by the wide availability of interactome data. Table 1 summarizes the orthogroups containing at least one protein of plant species. From the orthogroups obtained, we selected those that are considered one-to-one or one-to-many. A total of 4,501 *A. sisalana* proteins presents in those orthogroups were eligible for the search for interactions in the STRING database.

Table 1. Overlap of *A. sisalana* orthologs between two genomes

	Agave sisalana	Arabidopsis thaliana	Musa acuminata	Oryza sativa	Sorghum bicolor	Zea mays
Agave sisalana	11,663	8,376	8,971	9,229	8,897	9,153
Arabidopsis thaliana	8,376	12,488	9,273	9,231	9,355	9,277
Musa acuminata	8,971	9,273	11,282	9,856	9,859	9,916
Oryza sativa	9,229	9,231	9,856	15,347	12,784	12,555
Sorghum bicolor	8,897	9,355	9,859	12,784	15,747	13,703
Zea mays	9,153	9,277	9,916	12,555	13,703	17,091

A. sisalana interactome

A. sisalana interactome was constructed through the mapping of the ortholog groups and protein interactions extracted from the STRING database. The number of interactions obtained from STRING dabatase was 772,055 in *A. thaliana*, 1,384,645 in *Musa acuminata*, 588, 907 in *Oryza sativa*, 618,973 in *Sorghum bicolor*, and 1,368,899 in *Zea may.* Here, it was assumed that interactions of proteins in one organism are expected to be conserved in other related organisms (Fraser, 2005; Uhrig, 2006).

After predicting PPIs in *A. sisalana*, the PPI network was visualized using Cytoscape v3.8.1. The final interactome contains 2,582 nodes and 15,083 edges (Supplementary File 1). The topological metrics of the PPI network were calculated using Network Analyzer (Table 2). Analysis of the interaction network showed the short path length distribution, the node degree distribution, the neighborhood connectivity distribution, and the clustering coefficient distribution (Figure 2). The declining neighborhood connection trend reveals a weak clustering coefficient among neighbor nodes with lesser connectedness. On the other hand, the shortest path length distribution showed that the edges have a frequency dominated by a short path length (3-5), as path length means edge. The node degree distribution shows that the *A. sisalana* interactome has many nodes with a low degree, suggesting this possible short path length. In the clustering coefficient distribution, we note the presence of nodes with a low degree (clustering coefficient equal to zero), but also a group of highly connected genes, with a clustering coefficient equal to 1.

Table 2. Analysis of the interaction network topology of *A. sisalana*

Metrics	Value
Number of nodes	2,582
Number of edges	15,083
Avg. Number of neighbors	12.422
Network diameter	16
Network radius	8
Characteristic path length	4.491
Clustering coefficient	0.303
Network density	0.005
Network heterogeneity	1.72
Network centralization	0.071
Connected components	89

At protein level, important genes are normally strongly interconnected hubs. The top 20 hubs were identified in *A. sisalana* interactome (Table "hub genes" in Supplementary File 2) and most of them are functionally related to transcription, phosphorylation, response to stimulus, and nitrate assimilation. The *D-fructose-1,6-bisphosphate 1-phosphohydrolase* gene (AS_TRINITY_DN57407_c0_g1_i1) was identified as a hub and has an important regulatory role involved in photosynthetic CO_2 assimilation (Chehebar and Wolosiuk, 1980). Another hub is an *ATPase, F1 complex, OSCP/delta subunit* (AS_TRINITY_DN7474_c0_g1_i1), which has a role in inorganic ion transmembrane transport and participates in many metabolic processes. *Protein-serine/threonine phosphatase* (AS_TRINITY_DN57726_c0_g1_i1) was identified as a hub in the *A. sisalana* interactome. For *Agave*, *Protein-serine/threonine phosphatases* are related to diel expression patterns and exhibit inverted temporal shifts in abundance when compared to the C3 plant *Arabidopsis thaliana* (Abraham et al., 2016).

GO and KEGG enrichment analysis

We performed GO enrichment analysis of all genes present in the interactome. We identified biological processes related to protein import, protein modification and methylation, such as *methylation* (GO:0032259; p-value 5.0e-07), *protein peptidyl-prolyl isomer*ization (GO:0000413; p-value 5.0e-10) and *protein import into chloroplast stroma* (GO:0045037; p-value 1.7e-04). Basal processes related to the photosystem and chloroplast were also identified, such as *photosynthesis* (GO:0015979; p-value 1.4e-21), *light harvesting* (GO:0009765; p-value 5.4e-06), *chloroplast fission* (GO:0010020; p-value 4.5e-04), *chlorophyll biosynthetic process* (GO:0015995; p-value 1.1e-07), *photosystem I assembly* (GO:0048564; p-value 9.8e-04), *photosystem II repair* (GO:0010206; p-value 9.8e-04) and *photosynthetic electron transport in photosystem I* (GO:0009773; p-value 5.3e-05).

Regarding KEGG enrichment, some basic pathways detected in plants were enriched, such as Plant hormone signal transduction (ko04075), MAPK signaling pathway - plant (ko04016), Photosynthesis (ko00195), Circadian rhythm - plant (ko04712), Ribosome (ko03010), Biosynthesis of cofactors (ko01240), Plant-pathogen interaction (ko04626), Carbon fixation in photosynthetic organisms (ko00710), Porphyrin metabolism (ko00860), Ubiquinone and another terpenoid-quinone biosynthesis (ko00130), Protein export (ko03060), Biosynthesis of amino acids (ko01230), Protein processing in endoplasmic reticulum (ko04141), Glycerolipid metabolism (ko00561), Glycolysis/Gluconeogenesis (ko00010), Photosynthesis - antenna proteins (ko00196), Carotenoid biosynthesis (ko00906) and Flavonoid biosynthesis (ko00941). In addition, an enriched metabolic pathway of Plant-pathogen interaction (ko04626; p.adjust=0.00014) was identified, suggesting the presence of a phytopathogen. Recent studies using this same transcriptomic dataset have described several viral (Quintanilha-Peixoto et al., 2021) and fungal (Marone et al., 2022) interactions in *A. sisalana*, so this metabolic pathway might be a response to these interactions. Also, Flavonoid biosynthesis (ko00941; p.adjust=0.03169) enriched may indicate the protection of *A. sisalana* against these phytopathogens (Falcone Ferreyra et al., 2012).

Tissue-specificity analysis

Tissue-specificity analysis (Tau metric of 0.90, see methods for details) identified transcripts expressed exclusively in one of the three tissues studied (leaf, stem, and root). From a total of 2,582 genes in *A. sisalana* interactome, 2, 18, and 101 were identified in leaf, stem, and root, respectively (Supplementary File 2). We found a leaf-specific LIM zinc finger protein (AS_TRINITY_DN73400_c0_g1_i1; Tau = 1), which may be involved in plant growth and development, as well as regulating resistance mechanisms to diverse biotic and abiotic stresses as observed in other plants (Gupta et al., 2012).

As a stem-specific transcript, we found an auxin transporter regulating intracellular auxin homeostasis and metabolism (AS_TRINITY_DN46645_c0_g1_i1; Tau = 0.941) which is crucial for plant development (Rosquete et al., 2012). However, auxin has differential

distribution (gradients) within plant tissues, thus this gene may be acting as a gatekeeper controlling auxin traffic in and out of plant cells. Finally, among the root-specific genes, we found *regulation of jasmonic acid mediated signaling pathway* (GO:2000022; p-value: 0.00019), *transfer of electrons from cytochrome c to oxygen* (GO:0006123; p-value: 0.02372) and *photosynthesis* (GO:0015979; p-value:0.00022).

Submodules identification

A total of 16 submodules were identified in the *A. sisalana* interactome. Each submodule has highly connected genes and can suggest specific processes and roles. The GO enrichment of each submodule (Supplementary File 2) identified two submodules related to methylation and DNA repair/recombination: submodule S9 has 5 genes related to *histone methylation* (GO:0016571; p-value 8.2e-04), whose expression values are similar in the three tissues, with maximum values ranging from 10 to 20 TPM. Submodule S5 has 21 genes related to *DNA repair* (GO:0006281; p-value 5e-04), *DNA topological change* (GO:0006265; p-value 3.7e-06), and *DNA recombination* (GO:0006310; p-value 5.3e-06), which are all related to DNA damage response. Most genes in submodule 5 have expression values above 10 TPM in all tissues, but genes related to DNA recombination and DNA repair have TPM values from 20 to 142, suggesting these processes were highly demanded by the plant at the moment of sampling. Indeed, the samples were collected at noon and UV light exposure and/or the presence of Reactive Oxygen Species (ROS) can induce the plant response mechanism to fix any damage (Nisa et al., 2019).

We found 3 submodules related to carbohydrate metabolism (S7, S11, and S15). Among them, S11 is the only submodule presenting only biological processes associated with *carbohydrate metabolism* (GO:0005975; p-value 6.3e-06). Interestingly, S11 submodule contains Raffinose synthase gene (AS_TRINITY_DN57582_c0_g2_i1) with TPM values 28 in the leaf tissue. Raffinose has been previously described as an important carbohydrate in Agave metabolism (Raya et al., 2021) and it is an osmolyte that can act both in drought and oxidative stress (Nishizawa et al., 2008). As for the S7 submodule, we identified GO terms *sucrose biosynthetic process* (GO:0005986; p-value 7.9e-06) and *starch biosynthetic process* (GO:0019252; p-value 1.9e-14) as well as terms related to *gene regulation, like transcription from plastid promoter* (GO:0042793; p-value 2.7e-04), *cytidine to uridine editing* (GO:0016554; p-value 7.6e-04), and *Group II intron splicing* (GO:0000373; p-value 5.1e-07). Finally, submodule S15 presented the term *regulation of starch biosynthetic process* (GO:0010581; p-value 8.8e-05) and several terms related to abiotic stress. Several heat-shock proteins (HSP) were found in this submodule in categories like *response to heat* (GO:0009408; p-value 1.4e-07), *chaperone mediated protein folding requiring cofactor* (GO:0051085; p-value 6.5e-04), and *protein folding* (GO:0006457; p-value 7.9e-06). Among this HSP, Small HSP and DnaJ molecular chaperone were more present. Curiously, the term histone H3 deacetylation was also present in this submodule. The level of histone acetylation has been shown to regulate plant response to drought stress, including ABA-responsive element binding protein (AREB) (Li et al., 2021). AREBs can regulate Dehydration responsive element binding (DREB), that are known for inducing HSP and Heat shock factors (HSF) (Agarwal et al., 2017).

Among the identified submodules, six of them were related to the photosynthesis process (S1, S2, S3, S4, S10, and S13), what was expected due to the nature of the organism. Submodule S1 showed the following unique biological processes: *nonphotochemical quenching*, *cellular response to high light intensity*, *photosystem II assembly*, *energy coupled proton transport*, *down electrochemical gradient*, *electron transport chain*, *starch biosynthetic process*, and *regulation of photosynthesis, dark reaction*. The S2 submodule presented genes related to the unique processes protoporphyrinogen IX biosynthetic process, chloroplast-nucleus signaling pathway, DNA-templated transcription, termination, and cytidine to uridine editing. The biological processes unique to the S3 submodule were *photosystem I assembly*, *protein import into chloroplast stroma*, *tetrahydrobiopterin biosynthetic process,* and *iron-sulfur cluster assembly*. Submodule 10 has an exclusive processes *chlorophyll catabolic process* and *terpenoid biosynthetic process*. The S13 submodule presented exclusive processes such as sucrose biosynthetic process, *glycolytic process, isoprenoid biosynthetic process* and *regulation of jasmonic acid mediated signaling pathway*. By comparing the biological processes in common between these submodules, except for S10, we identified *protein peptidyl-prolyl isomerization process*, with the highest abundance of *peptidylprolyl isomerase* (PPIase) expression (AS_TRINITY_DN23174_c0_g1_i1) in leaf with 12 TPM. Interestingly, PPIase is responsible for the proper folding of proteins, being the only enzyme capable of catalyzing the *cis-trans* transition without energy requirement, unlike chaperones (Fischer et al., 1984; Fischer et al., 1989; Fanghänel and Fischer, 2004). Photosynthesis is a complex process involving several proteins, so the high expression of PPIase might be related to a correct folding of proteins without spending energy (Kirschbaum, 2004; Darko et al., 2014; Martin et al., 2018).

Lastly, submodule S14 has the processes *iron-sulfur (Fe-S) regulation* and *pyruvate metabolism*. Plants have a high demand for iron in mitochondria and chloroplasts, mainly to ensure respiration and photosynthesis (Couturier et al., 2013). Iron-Sulphur is required in many metabolic pathways due to the presence of metalloproteins (Johnson et al., 2005). Moreover, Fe-S proteins play a wide range of functionality ranging from radical generation, electron transfer, participation in sulfur and nitrogen assimilation, DNA replication and repair, chlorophyll catabolism, and ribosome biogenesis (Balk and Pilon, 2011; Johnson et al., 2005). Regarding pyruvate metabolism, we found *L-cysteine catabolic process to pyruvate* was reported in this submodule, although the only gene present was *L-cysteine desulfhydrase* (LCD). LCD is responsible for cysteine degradation and is highly expressed in response to adverse environmental conditions and intracellular oxidative stress (T. Chen et al., 2020; Wang et al., 2022), which agrees with agaves having a high capacity to be cultivated in places with more adverse environmental conditions.

Co-expression network analysis

We carried out a hierarchical clustering (WGCNA) to investigate the expression profile as a function of tissues (Supplementary File 4: Figure 2S-A). During WGCNA analysis, we found β = 22 and R^2 = 0.80 as the optimal soft threshold parameters to guarantee a scale-free network (Supplementary File 4: Figure 2S-B). We set clustering height cut-off to 0.20 to merge similar modules, which resulted in 73 modules (Supplementary File 4: Figure 2S-C). Specifically, modules darkgreen, purple, darkred, darkturquoise, magenta, bisque4, mistyrose, lightyellow, pink, plum1, magenta4, darkolivegreen, blue4, turquoise, darkorange, salmon1, mediumpurple1, brown4, skyblue1, and plum4 were identified as statistically significant (|GS| > 0.2 and MM > 0.8; see methods for details) in at least one tissue (Figure 3). The significance of the genes (GS) in relation to the module membership (MM) of each module, and the highest correlation value in each tissue are summarized in Supplementary File 5.

To explore potential genes and pathways associated with each tissue, we conducted GO enrichment analysis on the modules with highest correlation with the tissues. (Supplementary File 3). A total of 8 modules showed enriched GO terms: modules darkgreen, magenta, darkred, and purple are positively correlated with root; pink and darkolivegreen are positively correlated with stem, and turquoise and blue4 are positively correlated with leaf. Leaf was the only tissue presenting negatively correlated modules: darkorange and darkred (Figure 3).

The darkgreen module genes (correlation of 0.99 with root expression profile) have enriched GO terms related to translation and cellular iron ion homeostasis (Supplementary File 4: Figure 1S-B). In addition, the magenta module showed enriched GO related to the pentose-phosphate shunt, GDP-mannose transmembrane transport, cytoplasmic translational elongation, and translation. The pentose-phosphate shunt may be acting on root ion transport systems. Studies suggest that the regulation of root nitrogen and sulfur acquisition by plant carbon status is governed by an unnamed oxidative pentose phosphate pathway-dependent sugar detection system, which coordinates the availability of these three elements for amino acid synthesis (Lejay et al., 2008).

The darkred module showed enriched GO terms related to lignin biosynthetic process, protein localization to membrane, intracellular auxin transport, S-adenosylmethionine biosynthetic process, and actin filament bundle assembly. The presence of the lignin biosynthetic process positively correlated with the root and negatively correlated with the leaf highlights the importance of lignin for different tissues in the plant. This agrees with the low contents of lignin in leaf and stem in agaves (Raya et al., 2021), although there is no available quantification of lignin content in roots. Therefore, roots might possess more lignin to ensure mechanical support (Hoson and Wakabayashi, 2015), act as a defense barrier (Underwood, 2012), and a conduit for long-distance transportation of water and essential nutrients from the roots to the shoots (Liu et al., 2018; Boerjan et al., 2003; Bonawitz and Chapple, 2010; Zhong and Ye, 2009).

In the purple module, we observed the GO process related to response to farnesol, response to chitin, potassium ion transmembrane transport, transmembrane transport, and carbohydrate catabolic process. Here we highlight that the response process to farnesol is positively correlated with root tissue. Previous studies in *Arabidopsis* have shown that folk flowers, which lack farnesol kinase activity, accumulate farnesol and develop supernumerary carpels under water stress, providing evidence of a molecular link between the farnesol metabolism, abiotic stress signaling, and flower development (Fitzpatrick et al., 2011). Therefore, as the sampled plants were in a dry region, this group of genes was probably signaling the presence of this stress to the plant and activating another warning cascade. Moreover, we observed genes related to ABA signaling in submodule S15, which also shared a role with farnesol kinase in flower development (Fitzpatrick et al., 2011).

Regarding the modules highly correlated with the stem, pink module showed enriched GO terms related to protein phosphorylation, leaf formation, floral organ abscission, mitochondrial mRNA processing, auxin-activated signaling pathway, and regulation of transcription (Supplementary File 3). The process involving auxin signaling is responsible for ramet formation in *Agave*, involved in the survival and reproduction of a whole plant (Barreto et al., 2010). Another interesting process in stem is the floral organ abscission. Abscission, in which plants shed unwanted organs, is a natural developmental program or in response to environmental stimuli (Patharkar and Walker, 2015). In *A. sisalana*, the genes related to abscission in this module, which is positively correlated with stem, could be floral repression genes, which can be part of further improvement studies to accelerate or repress flowering. Agaves are monocarpic, meaning they only flower once at the end of their lifespan, so better control of flowering time can help synchronize it in agave fields. Because agaves accumulate sugar in their stems to be used during flowering, this group of genes may be responsible for this accumulation. The relationship between each gene significance and module membership from the pink module presents a correlation of 0.9884 (Supplementary File 4: Figure 1S-C).

The GO enrichment of darkolivegreen module shows processes related to regulation of defense response to fungus, brassinosteroid biosynthetic process, and positive regulation of transcription initiation from RNA polymerase II promoter. Indeed, fungal transcripts were described in the same dataset (Marone et al., 2022). Although most fungal transcripts were root-specific, some were present in the stem.

Regarding the modules highly correlated with the leaf, GO analysis of the turquoise module (positively correlated) demonstrated that genes in this module were primarily associated with photosynthesis, photosystem I and II, light harvesting, photosynthetic electron transport, and others (Supplementary File 3). Most of the processes are linked to photosynthesis processes, which makes the turquoise module a candidate module for the reconstruction of the interactome related to leaf genes. The findings of these photosynthesis-related genes revealed the strong linkage of

these co-expressed genes to the cited GO terms in *A. sisalana* leaf. Previous findings in other plant transcriptomes also described that the primary function of light-harvesting complex (Lhc) proteins was related to light absorption by chlorophyll excitation and transfer of absorbed energy to photochemical reaction centers (Ma et al., 2018; Zhao et al., 2016). This relationship can be observed by the significance of the genes in relation to the module membership of the turquoise module, in which they present a correlation of 0.9989 (Supplementary File 4: Figure 1S-A).

The darkorange module is negatively correlated with leaf (cor=-0.96 and p-value=0.002). The GO enrichment shows processes related to membrane fusion and brassinosteroid-mediated signaling pathways. The brassinosteroid (BR) class of steroid hormones regulates plant development, growth, and physiology and is involved in mediating adaptation to abiotic stresses, such as drought, temperature changes, and salinity (Planas-Riverola et al., 2019). This result shows that these classes of genes play a key role in maintaining the delicate balance between growth and resilience to environmental threats. These mechanisms of hormonal responses and stress stimuli have already been observed in a previous study of *A. sisalana* (Sarwar et al., 2019).

Biological processes enriched in blue4 (positively correlated with leaf) are related to the cellular response to hypoxia, anthocyanin-containing compound biosynthetic, RNA biosynthetic process, and auxin-activated signaling pathway (Supplementary File 3). Anthocyanin and hypoxia-related genes may play a relevant role in the tolerance and adaptation of agave plants in stress conditions, such as high temperatures, high light, drought, and low O_2 availability. Genes linked to the anthocyanin-containing compound biosynthetic process may have a role in response to high light incidence (Petrella et al., 2016; Steyn et al., 2002; Tattini et al., 2014; Trojak and Skowron, n.d.; Zheng et al., 2020) since anthocyanins are essential against UV radiation present in light stress conditions (Gould, 2004; Guo et al., 2010; Landi et al., 2015; Takahashi et al., 1991) and mitigate DNA damage in UV-B-irradiated (Kootstra, 1994; Stapleton and Walbot, 1994; Takahashi et al., 1991).

Also, plant responses to oxygen limitation (anoxia and hypoxia) are modulated by common signaling pathways, which target metabolic adaptations (Chang et al., 2011). Under oxygen limitations, tolerant plants increase glycolit flux or metabolic depression, reducing ATP consumption (Sasidharan et al., 2017). Hypoxia induces a decline in stomatal conductance with a rapid decrease in the rate of photosynthesis (Araki et al., 2012; Malik et al., 2001), however, it was shown that leaves with high CAM activity (crassulacean acid metabolism) have higher liquid photosynthesis under hypoxia (Pedersen et al., 2011). Moreover, CAM plants have their stomata closed during the day and open only at night, when they fix CO_2 (Males and Griffiths, 2017). Hypoxic pre-treated tissues have been shown to have the ability to maintain a high glycolytic rate during prolonged periods of anoxia, as well as higher levels of ATP and energy load (Mugnai et al., 2011). There is a close correlation between anoxia tolerance and carbohydrate reserves (Zahra et al., 2021). Carbohydrates, specifically sugars, increase hypoxia tolerance in many species due to their close association as an energy-providing metabolite (Ram et al., 2002; Setter et al., 1997). In addition, drought stress quickly causes an osmotic imbalance (Gurrieri et al., 2020). As the severity of stress increases, plants face drought by accumulating high intracellular levels of osmoprotective compounds (i.e. sugars) to protect cellular components and restore osmotic balance (Sharma et al., 2019; Singh et al., 2015).

Selection of Key Genes

For the selection of key genes, we combined the results of interactome with the co-expression network analysis through the selection of hub genes with high significance and module membership in the module-tissue relationship. Thus, a total of six hub genes were selected as key genes for participating in four WGCNA modules (turquoise, blue and blue4) that are positively correlated to leaf, in addition to having a negative correlation with root and stem, revealing that they might also play roles in these two tissues. The scoring of each hub gene in PPI and WGCNA is summarized in Table 3. For further validation of these potential key genes, we compared their expression values between tissue samples in the dataset. Expression levels of these four key genes in the turquoise module were markedly elevated in leaf samples compared with stem and root samples (Figure 4).

One of these transcripts is the P*rotein-serine/threonine phosphatase* gene (AS_TRINITY_DN57726_c0_g1_i1). This transcript, which was also identified as a highly connected hub, is a PP2C protein phosphatase that is a traditional regulator of the ABA receptor (*regulatory components of ABA receptor 3* - RCAR3). In *A. thaliana*, PP2C has been shown as a major negative regulator of ABA signaling (Umezawa et al., 2009). For *Agave*, the diel expression pattern of PP2C is inverted when compared to *A. thaliana,* and, for that, it has been proposed as one of the main regulators of stomata activity in CAM photosynthesis (Abraham et al., 2016). Moreover, *D-fructose-1,6-bisphosphate1-phosphohydrolase* (FBPase) (AS_TRINITY_DN57407_c0_g1_i1) is an important enzyme (EC 3.1.3.11) that acts mainly in the regulation of the Calvin cycle and glycogenesis. These reactions are involved in carbon fixation and sugar metabolism present in the chloroplast stroma and cytosol (Flechner et al., 1999). *Peptidylprolyl isomerase* (PPIase) (AS_TRINITY_DN57016_c0_g1_i1) (EC 5.2.1.8), which are the only enzymes known that can catalyze *cis-trans* transition of peptide bonds, essential for the proper folding of proteins (Fischer et al., 1984) and belong to three major classes of proteins: cyclophilins, FK506-binding proteins or FKBPs, and parvulins (Singh et al., 2020). Interestingly, in our analyses, PPIase was reported to be present in all photosynthesis submodules. *Light-harvesting chlorophyll a/b-binding protein* (Lhc) (AS_TRINITY_DN57140_c0_g1_i1) are class of antennae

proteins that play indispensable roles in capturing solar energy as well as photoprotection under stress conditions (Zhao et al., 2020). Lastly, *ATPase, F1 complex, OSCP/delta subunit* (AS_TRINITY_DN7474_c0_g1_i1), which was also selected as a key gene, is present in the blue4 module of the WGCNA analysis. This specific module has a high correlation with leaf (r=0.84 and P=0.04). cpATPase are essential for photosynthesis and their absence causes loss of photoautotrophy and increased photosensitivity. F-type ATP synthase uses the electrochemical proton gradient generated by photosynthesis to produce ATP from ADP and inorganic phosphate (Hahn et al., 2018; Maiwald et al., 2003).

Table 3. Scores of six intersecting hub genes using different ranking methods in PPI and WGCNA.

Gene ID	Gene name	Betweenness	Closeness	Degree	Module	GS-Leaf	GS-Stem	GS-Root	MM
AS_TRINITY_DN57726_c0_g1_i1	protein-serine/threonine phosphatase (PSP)	1,091,821.96	987.25	180	blue	0.54	0.09	-0.57	0.89
AS_TRINITY_DN57407_c0_g1_i1	D-fructose-1,6-bisphosphate1-phosphohydrolase (FBPase)	129,856.01	941.23	147	turquoise	0.98	-0.31	-0.70	1.00
AS_TRINITY_DN51514_c0_g1_i1	50S ribosomal protein L13, chloroplastic (RPL13)	144,691.69	940.66	163	turquoise	0.97	-0.29	-0.70	0.99
AS_TRINITY_DN7474_c0_g1_i1	ATPase, F1 complex, OSCP/delta subunit	90,648.14	917.96	164	blue4	0.93	-0.31	-0.64	0.98
AS_TRINITY_DN57016_c0_g1_i1	Peptidylprolyl isomerase (PPIase)	90,479.59	930.54	174	turquoise	0.94	-0.29	-0.67	0.97
AS_TRINITY_DN57140_c0_g1_i1	Light-harvesting chlorophyll a/b-binding protein (Lhc)	111,906.19	898.76	125	turquoise	0.95	-0.30	-0.68	0.98

PPI: protein-protein interaction network; Betweness: betweenness centrality; Closeness: closeness centrality; Degree: node connected degree; Module: module color in WGCNA; GS-Leaf: gene significance with leaf; GS-Stem: gene significance with stem; GS-Root: gene significance with root; MM: module membership;

Conclusions

In this study, we predicted 2,582 interactome components in *A. sisalana* using model organism PPI databases, with the numbers of *A. sisalana* interactome components determined from the numbers of model organism PPIs and *A. sisalana* orthologues. Also, we used RNA-Seq data from *A. sisalana* tissues to identify co-expressed genes and, consequently, co-expression modules that were positively correlated with leaf, stem, and root tissues in an analysis of the module-tissue relationship. In the co-expression analysis with WGCNA, we found 72 co-expression modules, of which turquoise, blue and blue4 were closely associated with leaf tissue.

From co-expression analysis of the module-tissue relationship, we were able to identify key modules and their associations with three different tissues of *A. sisalana* (leaf, stem, and root). These associations identified genes that play the role of abscission which can be part of further improvement studies to accelerate or repress flowering. Also, we found anthocyanin and hypoxia-related genes which may play a relevant role in the tolerance and adaptation of *Agave* plants, and genes with a role in adaptation to stress through brassinosteroid-mediated signaling pathways.

We identified key genes from the interactome network analysis through hub genes identification metrics and performed a correlation with their association in the module-tissue relationship. Further analysis suggested that six candidate genes were positively significantly related to leaf tissue and negatively related to stem and root, revealing that the key genes in the leaf have a role in the root and stem.

In conclusion, the predicted PPI network of *A. sisalana* expands the possibility of comparative analyses with other *Agave* species, thus providing additional insight into network evolution among species. Furthermore, the identification of co-expression modules and their module-tissue relationship, together with the identification of key genes, suggest target genes for specific studies in *Agave* transcriptome studies, biological processes related to plant survival in unfavorable environments, and provide strategies for breeding programs.

Declarations

Funding

This work was financed by the Center for Computational Engineering and Sciences - FAPESP/Cepid (2013/08293-7) and the São Paulo Research Foundation (FAPESP) through grants 2019/12914-3 and 2020/02524-0.

Ethics declarations

Not applicable
Competing interests: The authors declare no competing interests.

References

Abraham, P.E., Yin, H., Borland, A.M., Weighill, D., Lim, S.D., De Paoli, H.C., Engle, N., Jones, P.C., Agh, R., Weston, D.J., Wullschleger, S.D., Tschaplinski, T., Jacobson, D., Cushman, J.C., Hettich, R.L., Tuskan, G.A., Yang, X., 2016. Transcript, protein and metabolite temporal dynamics in the CAM plant Agave. Nat. Plants 2, 16178. https://doi.org/10.1038/nplants.2016.178

Agarwal, P.K., Gupta, K., Lopato, S., Agarwal, P., 2017. Dehydration responsive element binding transcription factors and their applications for the engineering of stress tolerance. J. Exp. Bot. 68, 2135–2148. https://doi.org/10.1093/JXB/ERX118

Alexa, A., Rahnenführer, J., Lengauer, T., 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 22, 1600–1607. https://doi.org/10.1093/BIOINFORMATICS/BTL140

Araki, H., Hossain, M.A., Takahashi, T., 2012. Waterlogging and Hypoxia have Permanent Effects on Wheat Root Growth and Respiration. J. Agron. Crop Sci. 198, 264–275. https://doi.org/10.1111/J.1439-037X.2012.00510.X

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., Ogata, H., 2020. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. Bioinformatics 36, 2251. https://doi.org/10.1093/BIOINFORMATICS/BTZ859

Balk, J., Pilon, M., 2011. Ancient and essential: the assembly of iron-sulfur clusters in plants. Trends Plant Sci. 16, 218–226. https://doi.org/10.1016/J.TPLANTS.2010.12.006

Barreto, R., Nieto-Sotelo, J., Cassab, G.I., 2010. Influence of plant growth regulators and water stress on ramet induction, rosette engrossment, and fructan accumulation in Agave tequilana Weber var. Azul. Plant Cell. Tissue Organ Cult. 103, 93–101. https://doi.org/10.1007/S11240-010-9758-9

Boerjan, W., Ralph, J., Baucher, M., 2003. Lignin Biosynthesis. http://dx.doi.org/10.1146/annurev.arplant.54.031902.134938 54, 519–546. https://doi.org/10.1146/ANNUREV.ARPLANT.54.031902.134938

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114. https://doi.org/10.1093/BIOINFORMATICS/BTU170

Bonawitz, N.D., Chapple, C., 2010. The Genetics of Lignin Biosynthesis: Connecting Genotype to Phenotype. http://dx.doi.org/10.1146/annurev-genet-102209-163508 44, 337–363. https://doi.org/10.1146/ANNUREV-GENET-102209-163508

Borland, A.M., Hartwell, J., Weston, D.J., Schlauch, K.A., Tschaplinski, T.J., Tuskan, G.A., Yang, X., Cushman, J.C., 2014. Engineering crassulacean acid metabolism to improve water-use efficiency. Trends Plant Sci. 19, 327–338. https://doi.org/10.1016/J.TPLANTS.2014.01.006

Boxall, S.F., Kadu, N., Dever, L. V., Knerová, J., Waller, J.L., Gould, P.J.D., Hartwell, J., 2020. Kalanchoë PPC1 Is Essential for Crassulacean Acid Metabolism and the Regulation of Core Circadian Clock and Guard Cell Signaling Genes. Plant Cell 32, 1136–1160. https://doi.org/10.1105/TPC.19.00481

Braun, P., Aubourg, S., Van Leene, J., De Jaeger, G., Lurin, C., 2013. Plant Protein Interactomes. http://dx.doi.org/10.1146/annurev-arplant-050312-120140 64, 161–187. https://doi.org/10.1146/ANNUREV-ARPLANT-050312-120140

Bray, N.L., Pimentel, H., Melsted, P., Pachter, L., 2016. Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. 2016 345 34, 525–527. https://doi.org/10.1038/nbt.3519

Broeren, M.L.M., Dellaert, S.N.C., Cok, B., Patel, M.K., Worrell, E., Shen, L., 2017. Life cycle assessment of sisal fibre – Exploring how local practices can influence environmental performance. J. Clean. Prod. 149, 818–827. https://doi.org/10.1016/J.JCLEPRO.2017.02.073

Camargo, A.P., Vasconcelos, A.A., Fiamenghi, M.B., Pereira, G.A., Carazzolle, M.F., 2020. tspex: a tissue-specificity calculator for gene expression data. Research Square.

Chang, R., Jang, C.J.H., Branco-Price, C., Nghiem, P., Bailey-Serres, J., 2011. Transient MPK6 activation in response to oxygen deprivation and reoxygenation is mediated by mitochondria and aids seedling survival in Arabidopsis. Plant Mol. Biol. 2011 781 78, 109–122. https://doi.org/10.1007/S11103-011-9850-5

Chehebar, C., Wolosiuk, R.A., 1980. Studies on the regulation of chloroplast fructose-1,6-bisphosphatase. Activation by fructose 1,6-bisphosphate. Biochim. Biophys. Acta 613, 429–438. https://doi.org/10.1016/0005-2744(80)90097-2

Chen, L.Y., Xin, Y., Wai, C.M., Liu, J., Ming, R., 2020. The role of cis-elements in the evolution of crassulacean acid metabolism photosynthesis. Hortic. Res. 7. https://doi.org/10.1038/S41438-019-0229-0/41989378/41438_2019_ARTICLE_229.PDF

Chen, T., Tian, M., Han, Y., 2020. Hydrogen sulfide: a multi-tasking signal molecule in the regulation of oxidative stress responses. J. Exp. Bot. 71, 2862–2869. https://doi.org/10.1093/JXB/ERAA093

Chin, C.H., Chen, S.H., Wu, H.H., Ho, C.W., Ko, M.T., Lin, C.Y., 2014. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC systems biology 8, 1-7.

Couturier, J., Touraine, B., Briat, J.F., Gaymard, F., Rouhier, N., 2013. The iron-sulfur cluster assembly machineries in plants: Current knowledge and open questions. Front. Plant Sci. 4, 259. https://doi.org/10.3389/FPLS.2013.00259/

Darko, E., Heydarizadeh, P., Schoefs, B., Sabzalian, M. R., 2014. Photosynthesis under artificial light: the shift in primary and secondary metabolism. Philosophical Transactions of the Royal Society B: Biological Sciences 369, 20130243. https://doi.org/10.1098/rstb.2013.0243

Davis, S.C., Dohleman, F.G., Long, S.P., 2011. The global potential for Agave as a biofuel feedstock. GCB Bioenergy 3, 68–78. https://doi.org/10.1111/J.1757-1707.2010.01077.X

de Silva, K. K., Dunwell, J. M., Wickramasuriya, A. M., 2022. Weighted Gene Correlation Network Analysis (WGCNA) of Arabidopsis somatic embryogenesis (SE) and identification of key gene modules to uncover SE-associated hub genes. International Journal of Genomics, 24. https://doi.org/10.1155/2022/7471063

Ding, Z., Kihara, D., 2019. Computational identification of protein-protein interactions in model plant proteomes. Sci. Reports 2019 91 9, 1–13. https://doi.org/10.1038/s41598-019-45072-8

Eguiarte, L.E., Jiménez Barrón, O.A., Aguirre-Planter, E., Scheinvar, E., Gámez, N., Gasca-Pineda, J., Castellanos-Morales, G., Moreno-Letelier, A., Souza, V., 2021. Evolutionary ecology of Agave: distribution patterns, phylogeny, and coevolution (an homage to Howard S. Gentry). Am. J. Bot. 108, 216–235. https://doi.org/10.1002/AJB2.1609

Emms, D.M., Kelly, S., 2019. OrthoFinder: Phylogenetic orthology inference for comparative genomics. Genome Biol. 20, 1–14. https://doi.org/10.1186/S13059-019-1832-Y/FIGURES/5

Falcone Ferreyra, M.L., Rius, S.P., Casati, P., 2012. Flavonoids: Biosynthesis, biological functions, and biotechnological applications. Front. Plant Sci. 3, 222. https://doi.org/10.3389/FPLS.2012.00222/

Fanghanel, J., Fischer, G., 2004. Insights into the catalytic mechanism of peptidyl prolyl cis/trans isomerases. Front Biosci. 9, 78. https://doi.org/10.2741/1494

Ferreira, S.S., Hotta, C.T., Poelking, V.G.D.C., Leite, D.C.C., Buckeridge, M.S., Loureiro, M.E., Barbosa, M.H.P., Carneiro, M.S., Souza, G.M., 2016. Co-expression network analysis reveals transcription factors associated to cell wall biosynthesis in sugarcane. Plant molecular biology 91, 15-35.

Fischer, G., Bang, H., Berger, E., Schellenberger, A., 1984. Conformational specificity of chymotrypsin toward proline-containing substrates. Biochim. Biophys. Acta 791, 87–97. https://doi.org/10.1016/0167-4838(84)90285-1

Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, T., Schmid, F. X., 1989. Cyclophilin and peptidyl-prolyl cis-trans isomerase are probably identical proteins. Nature 337, 476-478. https://doi.org/10.1038/337476a0

Fitzpatrick, A.H., Shrestha, N., Bhandari, J., Crowell, D.N., 2011. Roles for farnesol and ABA in Arabidopsis flower development. Plant Signal. Behav. 6, 1189. https://doi.org/10.4161/PSB.6.8.15772

Flechner, A., Gross, W., Martin, W.F., Schnarrenberger, C., 1999. Chloroplast class I and class II aldolases are bifunctional for fructose-1,6-biphosphate and sedoheptulose-1,7-biphosphate cleavage in the Calvin cycle. FEBS Lett. 447, 200–202. https://doi.org/10.1016/S0014-5793(99)00285-9

Fraser, H.B., 2005. Modularity and evolutionary constraint on proteins. Nature genetics 37, 351-352.

Gould, K.S., 2004. Nature's Swiss Army Knife: The Diverse Protective Roles of Anthocyanins in Leaves. J. Biomed. Biotechnol. 2004, 314. https://doi.org/10.1155/S1110724304406147

Guo, J., Han, W., Wang, M.-H., 2010. Ultraviolet and environmental stresses involved in the induction and regulation of anthocyanin biosynthesis: A review. African J. Biotechnol. 7, 4966–4972. https://doi.org/10.4314/ajb.v7i25.59709

Gupta, S.K., Rai, A.K., Kanwar, S.S., Sharma, T.R., 2012. Comparative analysis of zinc finger proteins involved in plant disease resistance. PLoS One 7 8, e42578. https://doi.org/10.1371/journal.pone.0042578.

Gurrieri, L., Merico, M., Trost, P., Forlani, G., Sparla, F., 2020. Impact of Drought on Soluble Sugars and Free Proline Content in Selected Arabidopsis Mutants. Biol. 2020, Vol. 9, Page 367 9, 367. https://doi.org/10.3390/BIOLOGY9110367

Hahn, A., Vonck, J., Mills, D.J., Meier, T., Kühlbrandt, W., 2018. Structure, mechanism, and regulation of the chloroplast ATP synthase. Science 360. https://doi.org/10.1126/SCIENCE.AAT4318

Hoson, T., Wakabayashi, K., 2015. Role of the plant cell wall in gravity resistance. Phytochemistry 112, 84–90. https://doi.org/10.1016/J.PHYTOCHEM.2014.08.022

International Energy Agency (IEA), 2021. Renewables 2021.

Jiménez-Barron, O., García-Sandoval, R., Magallón, S., García-Mendoza, A., Nieto-Sotelo, J., Aguirre-Planter, E., Eguiarte, L.E., 2020. Phylogeny, Diversification Rate, and Divergence Time of Agave sensu lato (Asparagaceae), a Group of Recent Origin in the Process of Diversification. Front. Plant Sci. 11, 1651. https://doi.org/10.3389/FPLS.2020.536135/

Johnson, D.C., Dean, D.R., Smith, A.D., Johnson, M.K., 2005. Structure, function, and formation of biological iron-sulfur clusters. Annu. Rev. Biochem. 74, 247–281. https://doi.org/10.1146/ANNUREV.BIOCHEM.74.082803.133518

Khan, Z.H., Agarwal, S., Rai, A., Memaya, M.B., Mehrotra, S., Mehrotra, R., 2020. Co-expression network analysis of protein phosphatase 2A (PP2A) genes with stress-responsive genes in Arabidopsis thaliana reveals 13 key regulators. Scientific reports 10, 1-16.

Kirschbaum, M. U. F., 2004. Direct and indirect climate change effects on photosynthesis and transpiration. Plant Biology 6, 242-253. https://doi.org/10.1055/s-2004-820883

Kootstra, A., 1994. Protection from UV-B-induced DNA damage by flavonoids. Plant Mol. Biol. 26, 771–774. https://doi.org/10.1007/BF00013762

Kryuchkova-Mostacci, N., Robinson-Rechavi, M., 2017. A benchmark of gene expression tissue-specificity metrics. Brief. Bioinform. 18, 205–214. https://doi.org/10.1093/BIB/BBW008

Landi, M., Tattini, M., Gould, K.S., 2015. Multiple functional roles of anthocyanins in plant-environment interactions. Environ. Exp. Bot. 119, 4–17. https://doi.org/10.1016/J.ENVEXPBOT.2015.05.012

Langfelder, P., Horvath, S., 2008. WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics 9. https://doi.org/10.1186/1471-2105-9-559

Lejay, L., Wirth, J., Pervent, M., Cross, J.M.F., Tillard, P., Gojon, A., 2008. Oxidative Pentose Phosphate Pathway-Dependent Sugar Sensing as a Mechanism for Regulation of Root Ion Transporters by Photosynthesis. Plant Physiol. 146, 2036. https://doi.org/10.1104/PP.107.114710

Li, S., He, X., Gao, Y., Zhou, C., Chiang, V.L., Li, W., 2021. Histone Acetylation Changes in Plant Response to Drought Stress. Genes (Basel). 12. https://doi.org/10.3390/GENES12091409

Liu, Q., Luo, L., Zheng, L., 2018. Lignins: Biosynthesis and Biological Functions in Plants. Int. J. Mol. Sci. 19. https://doi.org/10.3390/IJMS19020335

Ma, X., Zhao, H., Xu, W., You, Q., Yan, H., Gao, Z., Su, Z., 2018. Co-expression Gene Network Analysis and Functional Module Identification in Bamboo Growth and Development. Front. Genet. 9, 574. https://doi.org/10.3389/FGENE.2018.00574

Maiwald, D., Dietzmann, A., Jahns, P., Pesaresi, P., Joliot, P., Joliot, A., Levin, J.Z., Salamini, F., Leister, D., 2003. Knock-Out of the Genes Coding for the Rieske Protein and the ATP-Synthase δ-Subunit of Arabidopsis. Effects on Photosynthesis, Thylakoid Protein Composition, and Nuclear Chloroplast Gene Expression. Plant Physiol. 133, 191. https://doi.org/10.1104/PP.103.024190

Males, J., Griffiths, H., 2017. Stomatal Biology of CAM Plants. Plant Physiol. 174, 550–560. https://doi.org/10.1104/PP.17.00114

Malik, A.I., Colmer, T.D., Lambers, H., Schortemeyer, M., 2001. Changes in physiological and morphological traits of roots and shoots of wheat in response to different depths of waterlogging. Funct. Plant Biol. 28, 1121–1131. https://doi.org/10.1071/PP01089

Mallona, I., Egea-Cortines, M., Weiss, J., 2011. Conserved and Divergent Rhythms of Crassulacean Acid Metabolism-Related and Core Clock Gene Expression in the Cactus Opuntia ficus-indica. Plant Physiol. 156, 1978–1989. https://doi.org/10.1104/PP.111.179275

Marone, M. P., Campanari, M. F. Z., Raya, F. T., Pereira, G. A. G., Carazzolle, M. F., 2022. Fungal communities represent the majority of root-specific transcripts in the transcriptomes of Agave plants grown in semiarid regions. *PeerJ 10*, e13252. https://doi.org/10.7717/peerj.13252

Martin, W. F., Bryant, D. A., Beatty, J. T., 2018. A physiological perspective on the origin and evolution of photosynthesis. FEMS microbiology reviews 42, 205-231. https://doi.org/10.1093/femsre/fux056

Mielenz, J.R., Rodriguez, M., Thompson, O.A., Yang, X., Yin, H., 2015. Development of Agave as a dedicated biomass source: Production of biofuels from whole plants. Biotechnol. Biofuels 8, 1–13. https://doi.org/10.1186/S13068-015-0261-8/FIGURES/5

Monja-Mio, K.M., Herrera-Alamillo, M.A., Sánchez-Teyer, L.F., Robert, M.L., 2019. Breeding Strategies to Improve Production of Agave (spp.). Adv. Plant Breed. Strateg. Ind. Food Crop. 6, 319–362. https://doi.org/10.1007/978-3-030-23265-8_10

Moseley, R.C., Mewalal, R., Motta, F., Tuskan, G.A., Haase, S., Yang, X., 2018. Conservation and diversification of circadian rhythmicity between a model crassulacean acid metabolism plant kalanchoë fedtschenkoi and a model C3 photosynthesis plant arabidopsis thaliana. Front. Plant Sci. 871, 1757. https://doi.org/10.3389/FPLS.2018.01757/

Mugnai, S., Marras, A.M., Mancuso, S., 2011. Effect of Hypoxic Acclimation on Anoxia Tolerance in Vitis Roots: Response of Metabolic Activity and K+ Fluxes. Plant Cell Physiol. 52, 1107–1116. https://doi.org/10.1093/PCP/PCR061

Nisa, M.U., Huang, Y., Benhamed, M., Raynaud, C., 2019. The plant DNA damage response: Signaling pathways leading to growth inhibition and putative role in response to stress conditions. Front. Plant Sci. 10, 653. https://doi.org/10.3389/FPLS.2019.00653/

Nishizawa, A., Yabuta, Y., Shigeoka, S., 2008. Galactinol and raffinose constitute a novel function to protect plants from oxidative damage. Plant Physiol. 147, 1251–1263. https://doi.org/10.1104/PP.108.122465

Owen, N.A., Fahy, K.F., Griffiths, H., 2016. Crassulacean acid metabolism (CAM) offers sustainable bioenergy production and resilience to climate change. GCB Bioenergy 8, 737–749. https://doi.org/10.1111/GCBB.12272

Parascanu, M.M., Sanchez, N., Sandoval-Salas, F., Carreto, C.M., Soreanu, G., Sanchez-Silva, L., 2021. Environmental and economic analysis of bioethanol production from sugarcane molasses and agave juice. Environ. Sci. Pollut. Res. 28, 64374–64393. https://doi.org/10.1007/S11356-021-15471-4/TABLES/12

Patharkar, O.R., Walker, J.C., 2015. Floral organ abscission is regulated by a positive feedback loop. Proc. Natl. Acad. Sci. U. S. A. 112, 2906–2911. https://doi.org/10.1073/PNAS.1423595112/SUPPL_FILE/PNAS.1423595112.SD02.XLSX

Pedersen, O., Rich, S.M., Pulido, C., Cawthray, G.R., Colmer, T.D., 2011. Crassulacean acid metabolism enhances underwater photosynthesis and diminishes photorespiration in the aquatic plant Isoetes australis. New Phytol. 190, 332–339. https://doi.org/10.1111/J.1469-8137.2010.03522.X

Pérez-Pimienta, J.A., López-Ortega, M.G., Sanchez, A., 2017. Recent developments in Agave performance as a drought-tolerant biofuel feedstock: agronomics, characterization, and biorefining. Biofuels, Bioprod. Biorefining 11, 732–748. https://doi.org/10.1002/BBB.1776

Petrella, D.P., Metzger, J.D., Blakeslee, J.J., Nangle, E.J., Gardner, D.S., 2016. Anthocyanin production using rough bluegrass treated with high-intensity light. HortScience 51, 1111–1120. https://doi.org/10.21273/HORTSCI10878-16

Planas-Riverola, A., Gupta, A., Betegoń-Putze, I., Bosch, N., Ibaneş, M., Cano-Delgado, A.I., 2019. Brassinosteroid signaling in plant development and adaptation to stress. Development 146. https://doi.org/10.1242/DEV.151894

Quintanilha-Peixoto, G., Fonseca, P.L.C., Raya, F.T., Marone, M.P., Bortolini, D.E., Mieczkowski, P., Olmo, R.P., Carazzolle, M.F., Voigt, C.A., Soares, A.C.F., Pereira, G.A.G., Góes-Neto, A., Aguiar, E.R.G.R., 2021. The sisal virome: Uncovering the viral diversity of agave varieties reveals new and organ-specific viruses. Microorganisms 9. https://doi.org/10.3390/MICROORGANISMS9081704/S1

Ram, P.C., Singh, B.B., Singh, A.K., Ram, P., Singh, P.N., Singh, H.P., Boamfa, I., Harren, F., Santosa, E., Jackson, M.B., Setter, T.L., Reuss, J., Wade, L.J., Pal Singh, V., Singh, R.K., 2002. Submergence tolerance in rainfed lowland rice: physiological basis and prospects for cultivar improvement through marker-aided breeding. F. Crop. Res. 76, 131–152. https://doi.org/10.1016/S0378-4290(02)00035-7

Raya, F.T., Marone, M.P., Carvalho, L.M., Rabelo, S.C., de Paula, M.S., Campanari, M.F.Z., Freschi, L., Mayer, J.L.S., Silva, O.R.R.F., Mieczkowski, P., Carazzolle, M.F., Pereira, G.A.G., 2021. Extreme physiology: Biomass and transcriptional profiling of three abandoned Agave cultivars. Ind. Crops Prod. 172, 114043. https://doi.org/10.1016/J.INDCROP.2021.114043

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139. https://doi.org/10.1093/BIOINFORMATICS/BTP616

Rosquete, M.R., Barbez, E., Kleine-Vehn, J., 2012. Cellular auxin homeostasis: gatekeeping is housekeeping. Molecular plant 5 4,772-786. https://doi.org/10.1093/mp/ssr109

Sarwar, M.B., Ahmad, Z., Rashid, B., Hassan, S., Gregersen, P.L., Leyva, M.D. la O., Nagy, I., Asp, T., Husnain, T., 2019. De novo assembly of Agave sisalana transcriptome in response to drought stress provides insight into the tolerance mechanisms. Sci. Reports 2019 91 9, 1–14. https://doi.org/10.1038/s41598-018-35891-6

Sasidharan, R., Bailey-Serres, J., Ashikari, M., Atwell, B.J., Colmer, T.D., Fagerstedt, K., Fukao, T., Geigenberger, P., Hebelstrup, K.H., Hill, R.D., Holdsworth, M.J., Ismail, A.M., Licausi, F., Mustroph, A., Nakazono, M., Pedersen, O., Perata, P., Sauter, M., Shih, M.C., Sorrell, B.K., Striker, G.G., van Dongen, J.T., Whelan, J., Xiao, S., Visser, E.J.W., Voesenek, L.A.C.J., 2017. Community recommendations on terminology and procedures used in flooding and low oxygen stress research. New Phytol. 214, 1403–1407. https://doi.org/10.1111/NPH.14519

Setter, T.L., Ellis, M., Laureles, E. V., Ella, E.S., Senadhira, D., Mishra, S.B., Sarkarung, S., Datta, S., 1997. Physiology and Genetics of Submergence Tolerance in Rice. Ann. Bot. 79, 67–77. https://doi.org/10.1093/OXFORDJOURNALS.AOB.A010308

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res. 13, 2498. https://doi.org/10.1101/GR.1239303

Sharma, A., Shahzad, B., Kumar, V., Kohli, S.K., Sidhu, G.P.S., Bali, A.S., Handa, N., Kapoor, D., Bhardwaj, R., Zheng, B., 2019. Phytohormones Regulate Accumulation of Osmolytes Under Abiotic Stress. Biomol. 2019, Vol. 9, Page 285 9, 285. https://doi.org/10.3390/BIOM9070285

Sharma, A., Wai, C.M., Ming, R., Yu, Q., 2017. Diurnal Cycling Transcription Factors of Pineapple Revealed by Genome-Wide Annotation and Global Transcriptomic Analysis. Genome Biol. Evol. 9, 2170–2190. https://doi.org/10.1093/GBE/EVX161

Singh, H., Kaur, K., Singh, M., Kaur, G., Singh, P., 2020. Plant Cyclophilins: Multifaceted Proteins With Versatile Roles. Front. Plant Sci. 11, 1558. https://doi.org/10.3389/FPLS.2020.585212/

Singh, M., Kumar, J., Singh, S., Singh, V.P., Prasad, S.M., 2015. Roles of osmoprotectants in improving salinity and drought tolerance in plants: a review. Rev. Environ. Sci. Bio/Technology 2015 143 14, 407–426. https://doi.org/10.1007/S11157-015-9372-8

Stapleton, A.E., Walbot, V., 1994. Flavonoids can protect maize DNA from the induction of ultraviolet radiation damage. Plant Physiol. 105, 881. https://doi.org/10.1104/PP.105.3.881

Steyn, W.J., Wand, S.J.E., Holcroft, D.M., Jacobs, G., 2002. Anthocyanins in vegetative tissues: a proposed unified function in photoprotection. New Phytol. 155, 349–361. https://doi.org/10.1046/J.1469-8137.2002.00482.X

Suinaga, F.A., Silva, O.R.R.F., Coutinho, W.M., Cartaxo, W.V., Costa, L.B., 2007. Avaliação Agronômica de Oito Genótipos de Sisal (Agave spp.). Comun. Técnico.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., Jensen, L.J., Von Mering, C., 2011. The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res. 39.

https://doi.org/10.1093/nar/gkq973

Takahashi, A., Takeda, K., Ohnishi, T., 1991. Light-Induced Anthocyanin Reduces the Extent of Damage to DNA in UV-Irradiated Centaurea cyanus Cells in Culture. Plant Cell Physiol. 32, 541–547. https://doi.org/10.1093/OXFORDJOURNALS.PCP.A078113

Tattini, M., Landi, M., Brunetti, C., Giordano, C., Remorini, D., Gould, K.S., Guidi, L., 2014. Epidermal coumaroyl anthocyanins protect sweet basil against excess light stress: multiple consequences of light attenuation. Physiol. Plant. 152, 585–598. https://doi.org/10.1111/PPL.12201

Törönen, P., Medlar, A., Holm, L., 2018. PANNZER2: a rapid functional annotation web server. Nucleic Acids Res. 46, W84–W88. https://doi.org/10.1093/NAR/GKY350

Trojak, M., Skowron, E., n.d. Role of anthocyanins in high-light stress response.

Uhrig, J.F., 2006. Protein interaction networks in plants. Planta 2006 2244 224, 771–781. https://doi.org/10.1007/S00425-006-0260-X

Umezawa, T., Sugiyama, N., Mizoguchi, M., Hayashi, S., Myouga, F., Yamaguchi-Shinozaki, K., Ishihama, Y., Hirayama, T., Shinozaki, K., 2009. Type 2C protein phosphatases directly regulate abscisic acid-activated protein kinases in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. 106, 17588–17593. https://doi.org/10.1073/PNAS.0907095106/SUPPL_FILE/0907095106SI.PDF

Underwood, W., 2012. The plant cell wall: A dynamic barrier against pathogen invasion. Front. Plant Sci. 3, 85. https://doi.org/10.3389/FPLS.2012.00085

Van den Broeck, L., Dubois, M., Vermeersch, M., Storme, V., Matsui, M., Inzé, D., 2017. From network to phenotype: the dynamic wiring of an Arabidopsis transcriptional network induced by osmotic stress. Molecular systems biology 13, 961.

Wai, C.M., Weise, S.E., Ozersky, P., Mockler, T.C., Michael, T.P., Vanburen, R., 2019. Time of day and network reprogramming during drought induced CAM photosynthesis in Sedum album. PLOS Genet. 15, e1008209. https://doi.org/10.1371/JOURNAL.PGEN.1008209

Wang, L., Mu, X., Chen, X., Han, Y., 2022. Hydrogen sulfide attenuates intracellular oxidative stress via repressing glycolate oxidase activities in Arabidopsis thaliana. BMC Plant Biol. 2022 221 22, 1–12. https://doi.org/10.1186/S12870-022-03490-3

Wang, X., Jin, Y., 2017. Predicted networks of protein-protein interactions in Stegodyphus mimosarum by cross-species comparisons. BMC Genomics 18. https://doi.org/10.1186/S12864-017-4085-8

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., Yu, G., 2021. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innov. 2, 100141. https://doi.org/10.1016/J.XINN.2021.100141

Yan, X., Corbin, K.R., Burton, R.A., Tan, D.K.Y., 2020. Agave: A promising feedstock for biofuels in the water-energy-food-environment (WEFE) nexus. J. Clean. Prod. 261, 121283. https://doi.org/10.1016/J.JCLEPRO.2020.121283

Yang, S., Li, H., He, H., Zhou, Y., Zhang, Z., 2019. Critical assessment and performance improvement of plant–pathogen protein–protein interaction prediction methods. Brief. Bioinform. 20, 274–287. https://doi.org/10.1093/BIB/BBX123

Yeger-Lotem, E., Sharan, R., 2015. Human protein interaction networks across tissues and diseases. Front. Genet. 6, 257. https://doi.org/10.3389/FGENE.2015.00257/

Yin, H., Guo, H.B., Weston, D.J., Borland, A.M., Ranjan, P., Abraham, P.E., Jawdy, S.S., Wachira, J., Tuskan, G.A., Tschaplinski, T.J., Wullschleger, S.D., Guo, H., Hettich, R.L., Gross, S.M., Wang, Z., Visel, A., Yang, X., 2018. Diel rewiring and positive selection of ancient plant proteins enabled evolution of CAM photosynthesis in Agave. BMC Genomics 19, 1–16. https://doi.org/10.1186/S12864-018-4964-7/FIGURES/6

Yu, B., Liu, J., Wu, D., Liu, Y., Cen, W., Wang, S., Li, R., Luo, J., 2020. Weighted gene coexpression network analysis-based identification of key modules and hub genes associated with drought sensitivity in rice. BMC plant biology 20, 1-21.

Zahn-Zabal, M., Dessimoz, C., Glover, N.M., 2020. Identifying orthologs with OMA: A primer. F1000Research 9. https://doi.org/10.12688/F1000RESEARCH.21508.1

Zahra, N., Hafeez, M.B., Shaukat, K., Wahid, A., Hussain, S., Naseer, R., Raza, A., Iqbal, S., Farooq, M., 2021. Hypoxia and Anoxia Stress: Plant responses and tolerance mechanisms. J. Agron. Crop Sci. 207, 249–284. https://doi.org/10.1111/JAC.12471

Zhao, H., Lou, Y., Sun, H., Li, L., Wang, L., Dong, L., Gao, Z., 2016. Transcriptome and comparative gene expression analysis of Phyllostachys edulis in response to high light. BMC Plant Biol. 16, 1–17. https://doi.org/10.1186/S12870-016-0720-9/TABLES/3

Zhao, Y., Kong, H., Guo, Y., Zou, Z., 2020. Light-harvesting chlorophyll a/b-binding protein-coding genes in jatropha and the comparison with castor, cassava and arabidopsis. PeerJ 2020, e8465. https://doi.org/10.7717/PEERJ.8465/SUPP-7

Zheng, X.T., Yu, Z.C., Tang, J.W., Cai, M.L., Chen, Y.L., Yang, C.W., Chow, W.S., Peng, C.L., 2020. The major photoprotective role of anthocyanins in leaves of Arabidopsis thaliana under long-term high light treatment: antioxidant or light attenuator? Photosynth. Res. 2020 1491 149, 25–40. https://doi.org/10.1007/S11120-020-00761-8

Zhong, R., Ye, Z.-H., 2009. Secondary Cell Walls. eLS. https://doi.org/10.1002/9780470015902.A0021256

Figures

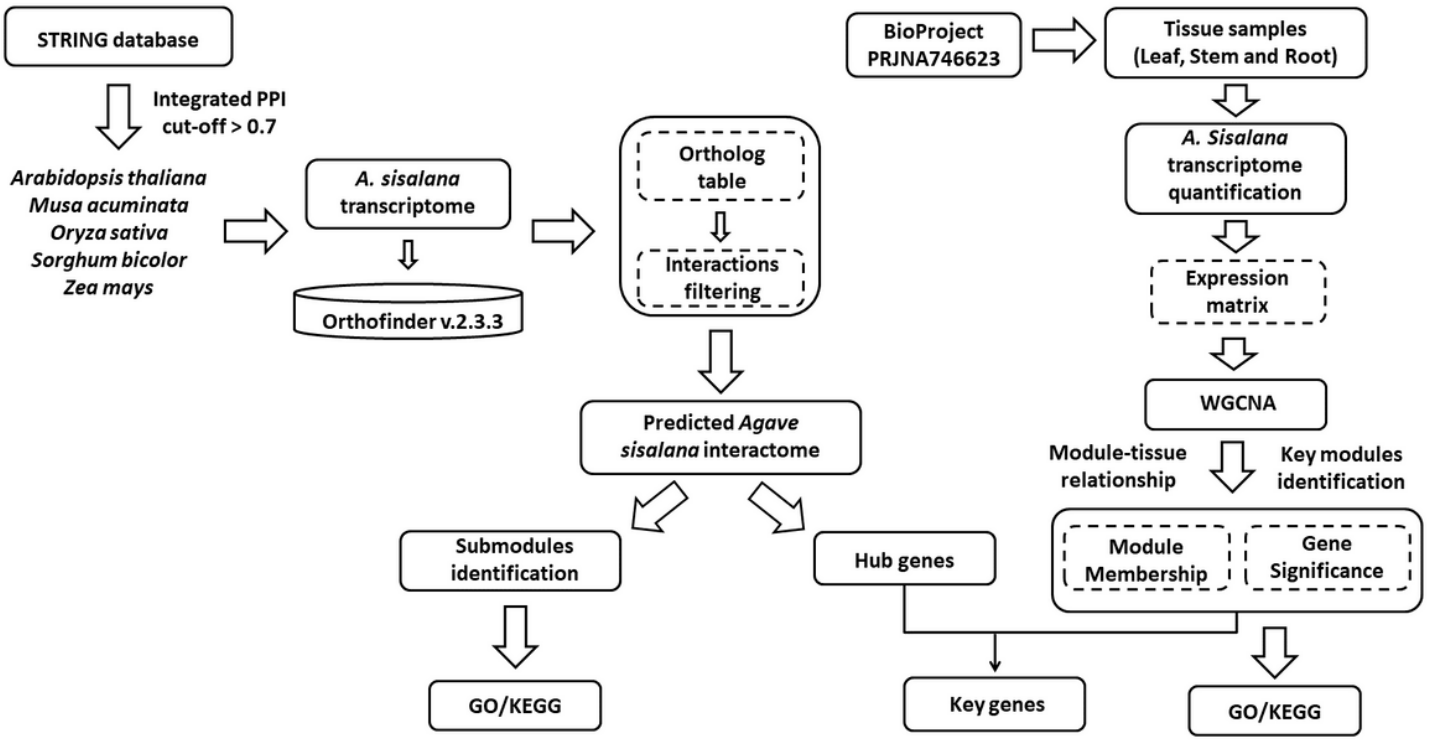

Figure 1

Flowchart illustrating the study design. *Agave sisalana* interactome was predicted from five orthologs organisms based on STRING database interactions in order to identify hub genes. Also, a WGCNA analysis was perfomed to identify key modules based on module-tissue relationship from three tissues of *A. sisalana* (leaf, stem and root). Both analyses were integrated to find key genes. The dashed lines indicate software outputs or in-house scripting processes.

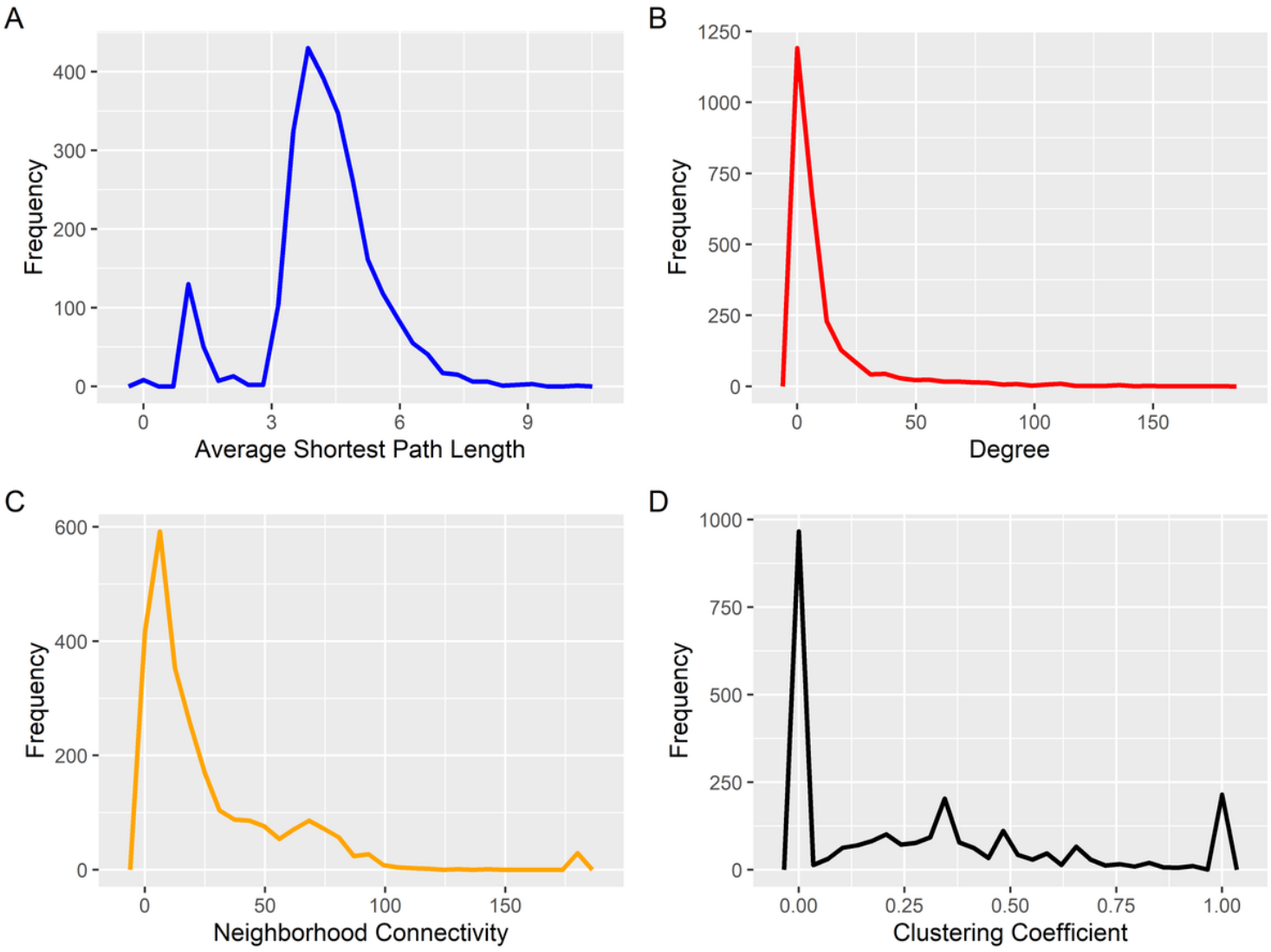

Figure 2

Agave sisalana interactome metrics. (a) short path length distribution (B) the node degree distribution, (C) the neighborhood connectivity distribution, and (D) clustering coefficient distribution.

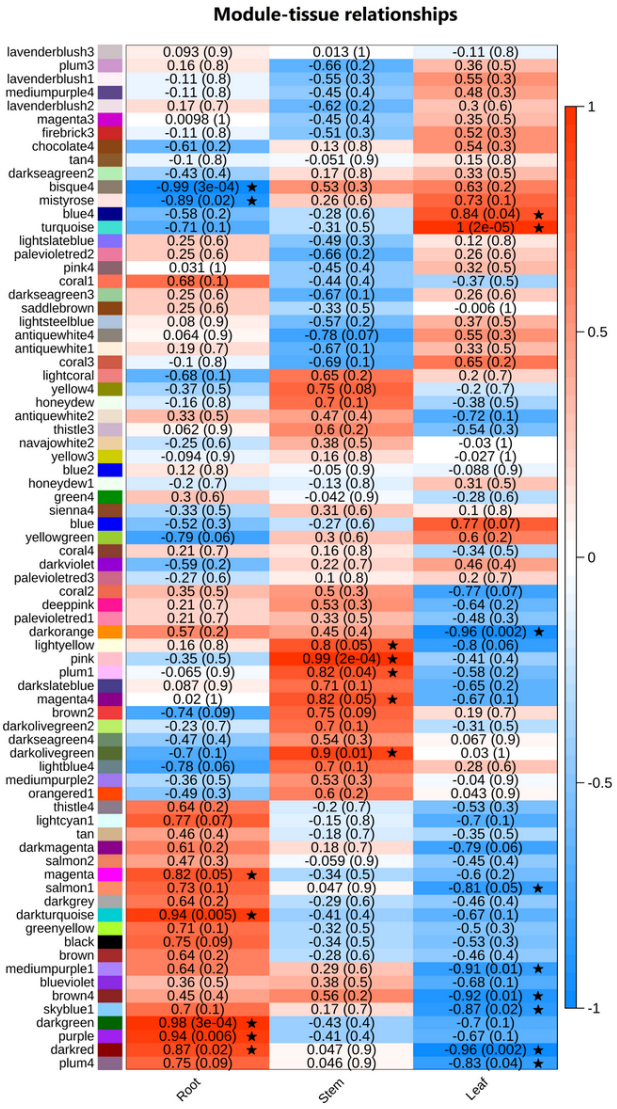

Figure 3

Heatmap of associations among module eigengenes in tissue samples (stem, leaf, and root) from weighted correlation network analysis (WGCNA). A black star marks the key modules, i.e., modules with |cor| > 0.5 and p-value <= 0.05 at least one tissue.

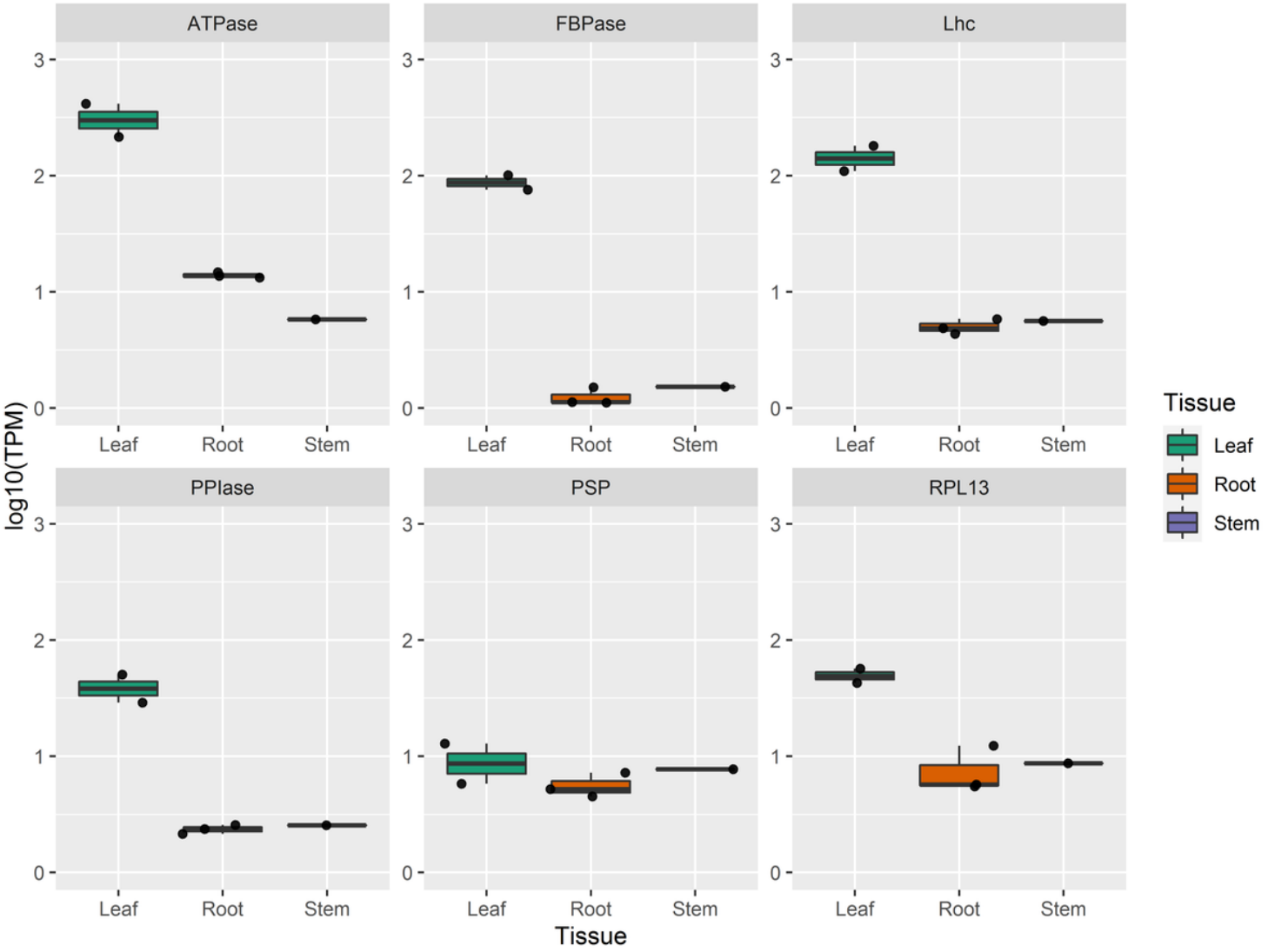

Figure 4

Per tissue normalized expression of six hub genes identified in PPI of *A. sisalana*. AS_TRINITY_DN57726_c0_g1_i1 (PSP); AS_TRINITY_DN57407_c0_g1_i1 (FBPase); AS_TRINITY_DN51514_c0_g1_i1 (RPL13); AS_TRINITY_DN7474_c0_g1_i1 (ATPase); AS_TRINITY_DN57016_c0_g1_i1 (PPIase); AS_TRINITY_DN57140_c0_g1_i1 (Lhc).

Supplementary Files

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

- SupplementaryFile1.xlsx
- SupplementaryFile2.xlsx
- SupplementaryFile3.xlsx
- SupplementaryFile4.docx
- SupplementaryFile5.xlsx