Tissue-specific Transcriptome Analysis Reveals Lignocellulose Synthesis Regulation in Elephant Grass (Pennisetum Purpureum Schum.)

Wenqing Zhang
Qilu University of Technology
Shengkui Zhang
Qilu University of Technology
Xianqin Lu
Qilu University of Technology
Can Li
Qilu University of Technology
Xingwang Liu
Qilu University of Technology
Geyu Dong
Qilu University of Technology
Tao Xia (txia@qlu.edu.cn)
Qilu University of Technology

Research article

Keywords: Elephant Grass, RNA-Seq, KEGG, WGCNA, Cellulose, Lignin

Posted Date: July 6th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-37358/v1

Version of Record: A version of this preprint was published on November 19th, 2020. See the published version at https://doi.org/10.1186/s12870-020-02735-3.

Abstract

Background: The characteristics of elephant grass, especially its stem lignocellulose, are of great significance for its quality as feed or other industrial raw materials. Because the genome of elephant grass has not been deciphered, the study of its lignocellulose synthesis pathway and key genes is limited.

Results: In this study, RNA sequencing (RNA-seq) combining with lignocellulose content analysis and cell wall morphology observation using elephant grass stems from different development stages as materials, were applied to reveal the genes regulating cellulose and lignin synthesis. A total of 3852 differentially expressed genes (DEGs) were identified in three periods of T1, T2 and T3. Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis showed that the two most abundant metabolic pathways were phenylpropanemetabolism, starch and sucrose metabolism, which closely related to cell wall development, hemicellulose, lignin and cellulose synthesis. Through weighted gene co-expression network analysis (WGCNA) of DEGs, a 'blue' module highly correlated with cellulose synthesis and a 'turquoise' module highly correlated with lignin synthesis were exhibited. A total of 43 candidate genes were screened, of which 17 had function annotations in other species. In addition, the expression of *CesA*, *PAL*, *CAD*, *C4H*, *COMT*, *CCoAMT*, *F5H*, *CAD* and *CCR* at different development stages were analyzed, and found that the content of lignocellulose was correlated with the expression levels of these structural genes.

Conclusions: This study not only provides new insights into the molecular mechanisms of cellulose and lignin synthesis pathways in elephant grass, but also offers a new and extensive list of candidate genes for more specialized functional studies in the future which may promote the development of high-quality elephant grass varieties with high cellulose and low lignin content.

Background

Fiber, which is mainly composed of three biological macromolecules of cellulose, lignin and hemicellulose, plays an important role in plant growth and stress responses [1-2].The formation offiber is a complex process that requires the coordination and balance of multiple metabolic pathways.

Cellulose is the most important component of fiber. During its synthesis, cellulose synthase(CesA) monomers form cellulosesynthase complex (CSC), which catalyzes the synthesis of dextran chain of cellulose by related substrates.To date, CesA gene have been identified and cloned in rice (*Oryza sativa* L.), *Arabidopsis thaliana*, maize (*Zea mays* L.), poplar(*Populustremuloides*)and some other plants, their function has been clarified [3-5]. Besides CesA, cellulase (Kor) and sucrose synthase (SuSy) are also found to be related to cellulose synthesis [6-7]. Although hemicellulose isa kind of heteropolysaccharide, most hemicellulose has a single skeleton. Except xylan, all the main chains of hemicellulose aresynthesized by cellulose synthase-like (CSL). Lignin, the second abundant component in plant cell walls, mainly plays roles in increasing plant cell wall strength and stem bending resistance[1]. Lignin is a kind of complex phenolic polymer, its monomer synthesis isderived from the phenylpropane pathway

[8].Lignin monomer synthase genes mainly include *PAL*, *C4H*, *C3H*, *4CL*, *COMT*, *CCoAMT*, *F5H*, *CAD* and *CCR*, etc. which have been studied in maize[9], *Arabidopsis thaliana*[10], poplar (*Populustremuloides*)[11], ryegrass[12], switchgrass [13]and other plants.The content and composition of lignin can be changed when the expression levels of these genes are up-regulated or downregulated.

Elephant grass is a perennial Poaceae C4 plant, originated in tropical Africa, and then widely distributed in tropical and subtropical climate regions of the Asia, Africa and America. It is one of the highest biomass plants in the world, with plant height up to 3 ~ 5m and annual biomass up to 4500kg / hm^{-2}[14].The main biological characteristics of elephant grass are high photosynthesis, highyield, and resistance to bioticandabiotic stress. Elephant grass is not only a high-quality forage for livestock and poultry [15], an ideal plant for soil and water conservation, a high-quality paper pulp raw material, the raw material for biofuels preparation, but also an ideal lignocellulosic energy plant [16-18].The characteristics of elephant grass, especially its stem lignocellulose, are of great significance for its quality as feed or other industrial raw materials. Improving the cellulose content and reducing the lignin content can promote the feed quality, the conversion and utilization efficiency of lignocellulose.

Elephant grass is the allotetraploid crop(A'A'BB, 2n=4x=28)with a complex genome. The species is primarily cross-pollinated due to its androgynous flowering behavior, resulting in high heterozygosity and large genetic diversity which can be utilized in breeding programs. Although the elephant grass genome has not been deciphered yet, its genetic research has been extensively focusing on evaluating genetic diversity through constructing molecular markers and fingerprints and determining genetic relationship [19-21]. Meanwhile, researchers have revealed the biosynthetic pathway of anthocyanins and the molecular mechanism of early response to cadmium in elephant grass roots and leaves through RNA_seq analysis [22-23]. Some progress has also been made in improving the agronomic traits of elephant grass by using traditional breeding methods based on phenotypic selection [24-25].

Although the regulatory network of lignocellulose synthesis has been reported in other species such as *Arabidopsis thaliana*,maize, poplarand switchgrass,research in elephant grass is still scarce. RNA-seq is the second generation of transcriptome sequencing technology, with the characteristics of high throughput and low cost, and does not need the corresponding sequencing reference genome sequence,which is very suitable for species lacking genome information. In this study, 12 samples of elephant grass at different stem development stages were used as materials for RNA-seq analysis, lignocellulose composition analysis and cell wall morphology observation. Combined with GO, KEGG, WGCNA and Q-PCR correlation analysis, thestudy identifies some candidate genes and providesvaluable genomic data for molecular mechanism of fiber formation in elephant grass, and maypromote the development and industrial application of high-quality elephant grass varietieswhich have high cellulose and low lignin content.

Results

Changes of cellulose, hemicellulose and lignin content in elephant grass stem at different development stages

The content of cellulose, hemicellulose and lignin in different stem segments ofT1, T2, T3 stage weremeasured. It was found that the content of cellulose and hemicellulose increased first and then decreased, for example,T1-S2 was higher than T1-S1, but T1-S3 was lower than T1-S2, whereas the content of lignin decreased gradually, that is, T1-S1 is the highestwhile T1-S3 is the lowest.The similar changes also appeared in different stem nodes at T2 and T3 stages. Meanwhile, by analyzing the content change of the same stem node in different development time, it was also found that with the increase of growth time of elephant grass, the content of cellulose and hemicellulose decreased, while the content of lignin increased (Fig.1b).

The results indicated that with the development of elephant grass, the cell gradually matures and ages, the main macromolecular cellulose and hemicellulose content of cell wall structure gradually decreases, while the lignin content gradually increases to maintain and support the strength of elephant grass stems.

Characteristics of cell wall in different developmental stages of elephant grass stem

The changes of cell wall morphology at different development stages of elephant grass, especially the primary and secondary cell wall were observed (Fig. 2a).In T1-S1 and T1-S2, the ratio of secondary cell wall (sw) thickness to primary cell wall (pw) thickness was 1.18 and 0.85, the ratio of sw/pw thickness in T2-S1 and T2-S2 was 1.67 and 0.92, and the ratio of sw/pw thickness of T3-S1 and T3-S3 was 2.15 and 1.25, respectively (Fig.2b).The above data showed that the change trendof sw/pw thickness of T1, T2 and T3 was consistent, that is, with the increase of development time, the development speed of secondary cell wall (sw) was faster than that of primary cell wall (pw). The ratio of sw/pw in S1 node of three different development stages was alsoanalyzed. It was found that with the increase of development time,the thickness ratio of sw/pw in T1-S1T2-S1, T3-S1 increasedsubsequently.The change of sw/pw thickness ratio in S2 and S3 stages of three different development stages showed same trends.

In order to further verify the morphological changes of cell wall during the development of elephant grass, we selected S1 and S5 stem segments in T3 period with a longer development time span as samples for micro-CT observation.The stem of elephant grass consists of epidermis, parenchyma cells and vascular bundles, in which vascular bundles are composed of phloem and xylem without cambium. Vascular bundles are scattered in parenchyma cells and cannot be thickened. With the development of stem tissue, S1 vascular bundle showed regular and compact arrangement(Fig. S1), and the content of cellulose and hemicellulose in vascular bundle decreases gradually, while the content of lignin increases gradually (Fig.1b), so as to meet the mechanical support need in the process of stem maturity of elephant grass.

Differential gene expression during the development of elephant grass stems

12 deepest samples (Table S1)were selected to reassemble the elephant grass transcriptome(Fig. S2) using Trinity software.The Pearson correlation coefficient based on the expression valueofeach library indicated that there was a high correlation between sample replicates (Fig. S3).Cluster analysis among samples showed that the development time of elephant grass was the main factor affecting the clustering. The DEGs in three developmental stages of elephant grass stems were analyzed, a total of 15611, 10235and27389 DEGs were identified in T1, T2, and T3, respectively(Fig.3).In addition, we also analyzed the DEGs of different stem segments of three development stages, and found that 147 of which were co-expressed in three segments of T1, 54 in the four segments of T2,and 91 in the five segments of T3 (Fig.3). The intersection of all DEGs at three different developmental stages was compared to determine the shared core set. It was found that 3852 genes were differentially expressed in three developmental stages (Fig.3d).

GO and KEGG analysis were performed on 3852 DEGs which were co-expressed in the three developmental stages.Four of the top 10 enriched GO annotation functions are related to cell composition such as apoplast, cell wall, extracellularregion, plant-type cell wall, 4 are related to molecular functions such as peroxidase activity,xyloglucan and xyloglucosyl transferase activity,heme-binding,xyloglucan–specific endo–beta–1,4–glucanase activity, and 2 were related to biological processes such as cell wall macromolecule catabolic process, hydrogenperoxidecatabolicprocess (Fig.4a).According to KEGG analysis, all DEGs were enriched to 9 pathways, of which the two most significant were phenylpropane metabolism（23 DEGs）and starch and sucrose metabolism（23 DEGs） (Fig.4b).

Geneshighly correlated with the synthesis of cellulose, hemicellulose and lignin by WGCNA analysis

Weighted gene co-expression network analysis (WGCNA) was performed on 3852 DEGs in three stages, and the network was divided into 3 modules. The analysis of module-trait relationship showed that the ‘blue’ module was highly correlated with the synthesis of cellulose (r=0.67, $P=6.0\times10^{-6}$) and hemicellulose (r=0.51, P=0.001), whereas the ‘turquoise’ module was related to lignin synthesis (r=0.68, $P=5.0\times10^{-6}$) (Fig. S4).

The ‘blue’ module was filtered according to Module membership > 0.9, the absolute value of the correlation coefficient between the ‘turquoise’ module and lignin was greater than 0.75, and the 20 and 23 genes remained in the ‘blue’ and‘turquoise’module respectively. The WGCNA gene significance (GS) (i.e. related to traits) showed that the genes with highest GS in the ‘blue’ and ‘turquoise’ modules were Cluster-55067.0 (0.659) and Cluster-17353.3 (0.798), respectively. Six of the 20 genes in the module ‘blue’ are known to be functional, such as*GTL1* transcription factor, O-methyltransferase (*OMT*), expansin-likeA2, alpha-humulenesynthase, probable galactinolsucrose, *GhGalT1*. At the same time, 14 genes with unknown functions are also covered by the ‘blue’ module, which needs further study.The GO function annotations of these 20 genes in the module ‘blue’ included the genes which were related toextracellular region, C-4 methylsterol oxidase activity, terpene biosynthesis, etc.

Among the 23 genes in the module ‘turquoise’, there are eight coding genes such as *GTL1*, *MYB2* transcription factors, threonine-protein kinase ERECTA, probable methionine-tRNA ligase, alcohol holding hydrogenase-like2, zinc finger proteinGIS3, protein slr0074, proline-rich receptor-like protein kinase PERK8. The GO function annotations of these 23 genes in the module ‘turquoise’included genes which were related tocell growth regulation, secondary cell wall formation regulation, cell wall composition regulation, etc. 15 genes remained with unknown functions (Table S2).

Lignin and cellulose synthesis pathway during stem development of elephant grass

Cellulose synthase (CesA) is the most important enzyme in the cellulose synthesis pathway.It can directly use UDPG produced in starch and sucrose metabolism to synthesize cellulose.In elephantgrass, we identified 27 *CesA* genes. In T1 stage, the expression level of *CesA1-CesA6*was higher in the whole stem, but decreased in T2 and T3 stages, while the expression level of *CesA7-CesA27* gene in tender stems was much higher than that in mature stems in T1, T2, and T3 stages. This indicated that *CesA* gene mainly synthesizes cellulose in the tender stem tissue, when the cellulose accumulated to a certain level, its expression level gradually decreased (Fig.6a).

Lignin synthesis is one of the most important pathways of phenylpropane metabolism. In the lignin metabolism pathway, *CAD* (21 Unigenes), *4CL9* (21 Unigenes), *C4H* (8Unigenes), *PAL* (16Unigenes), *CCR* (25Unigenes), *F5H* (6Unigenes),*CCoAOMT* (4Unigenes) were identified to be related. Expression analysis found that most members of these gene families showed higher expression levels inmature stem, while a few gene members were continuously expressed in the whole tissue (Fig.6b).

The results of qRT-PCR demonstratedthat the expression trends of these genes were consistent with that of RNA-seq data. Overall, with the development of stem, the expression level of cellulose synthesis genes and the expression level of lignin synthesis related genes have the opposite trend, while the changes of stem cellulose, hemicellulose and lignin content, as well as the changes of stem primary wall and secondary cell wall thickness, are positively related to the expression of these two types of genes(Fig. S5).

Discussion

Elephant grass is one of the highest biomassforage grasses on the earth, which is widely used in feed and bioenergyrelated industries.It is of great significance to further understand the genes and important metabolic pathways related to lignocelluloses synthesis for molecular breeding the new elephant grass varieties and make them more suitable for industrial application.

In this study, 3852 differentially expressed genes (DEGs) were identified by transcriptome analysis on 12 stem samples at different development stages of elephant grass.The GO enrichment analysis of DEGsfound that the top 10 enriched GO annotation functions included three functions that were directly related to cell wall development such as cellwall, plant-typecellwall, cellwallmacromolecule catabolicprocess. KEGG pathway analysis discovered that sucrose and starch metabolism, and

phenylpropane metabolism were the most significant metabolic pathways of differential gene enrichment.Lignin synthesis is one of the important branches in phenylpropane metabolism, the substrate for cellulose synthesis can be provided by starch and sucrose metabolic pathways [35-37], whichindicated that these DEGs may affect the synthesis of cellulose and lignin.

In the process of cellulose synthesis, sucrose is the starting substrate, and UDPG produced by its decomposition can directly synthesize dextran chain under the action of CesA gene [38]. We identified 27 CesA genes in elephant grass, which are more abundant than other species such as wheat (14 species) [4], *Arabidopsis thaliana* (10 species) [3], maize (10 species) [5], poplar (16 species) [7]. CesAswere mainly expressedin the tender stem tissue (Fig.6a). With the stem development, the cellulose content accumulated gradually (Fig.1). The high copy of CesAs in elephant grass makes it has high cellulose synthesis potential. The cellulose content in mature stem of elephant grass is higher than that of corn, wheat, reed and other plants [39].

Totally, 101 lignin-synthesizing related genes such as *PAL*, *C4H*, *4CL*, *CCoAMT*, *F5H*, *CAD*, and *CCR* were identified in elephant grass.PAL is the key enzyme in phenylpropane metabolism pathway, its inhibition will reduce lignincontent and affect plant growth and development.4CL is the key rate limiting enzyme for the production of G or S-lignin monomers. Different expression of 4CL can regulate the content of three kinds of lignin monomers, and promoting or inhibiting the expression of *4CL* gene can significantly regulate the relative proportion of lignin/cellulose [40]. In elephant grass, 16 *PAL* genes and 21 *4CL* genes were mainly expressed in mature stem tissues, thereforethe lignin content in mature stem of elephant grass was higher.It was also found that inhibition or overexpression of these genes in tobacco would affect the lignin content [41-42].

In order to further understand the relationship between these DEGs and lignocellulose components, WGCNA was conducted and found that there were 20 genes related to the synthesis of cellulose and hemicellulose, and 23 genes related to the synthesis of lignin.Among these 43 genes, 14 of which have been identified with clear functions through GO annotation, some of them such as *OMT* and *GalT*1 were directly related to the synthesis of lignocelluloses, some of them played important roles in growth and development and cell wall formation. There were 28 genes with unknown functions, which need further functional verificationin the future. In tobacco, it was reported that the sense or antisense expression of sequences encoding O-methyltransferase (*OMT*) could regulate enzyme activity of lignin synthesis[43]. The fiber length of transgenic cotton overexpressing *GhGalT1*was shorter than that of wild type, while in the *GhGalT1* silenced line, the fiber length was significantly increased than that of wild type [44].

Generally, our work indicated the dynamic changes in cell wall composition and morphology during the stemdevelopment of elephant grass are consistent with the changes and expression of cellulose and lignin related genes. These data providesthe new and extensive list of candidate genes for more specialized functional studies in the future, and provides an important theoretical basis for the genetic improvements of elephant grass lignocellulose synthesis as well.

Conclusion

RNA-seq, lignocellulose content and cell wall morphology of elephant grass stem were conducted and analyzed in this study, which provided a comprehensive insight into the underlying mechanisms of the synthesis and accumulation of lignocellulose in elephant grass. A total of 3852 common DEGs were identified and KEGG analysis showed that the two most abundant metabolic pathways were phenylpropane metabolism (23 DEGs), starch and sucrose metabolism (23 DEGs), among which phenylpropane metabolism functioned as an important pathway for the synthesis of lignin, while the latter produced UDPG for cellulose synthesis. 27*CesA* genes for cellulose synthesis and 101 related genes for lignin synthesis were identified, respectively. *CesA* genes had higher expression levels in young stems and lignin-related genes had higher expression levels in mature stems. In addition，a total of 43 candidate genes were screened by WGCNA, of which 17 had function annotations in other species. Among them, GTL1 transcription factor and O-Methyl transferase (OMT) genehave been proved to regulate the synthesis of lignocellulose in other plant species.

Methods

Experimental materials

Elephant grass was cultured in the greenhouse of Qilu University of Technology, Jinan City, Shandong Province. Stalks at seedling stage 40 days (T1 period), 80 days (T2 period), 120 days (T3 period) were sampled.The lowest node of the stem were taken and labeled as S1, samples from every other stem node were labeled as S2, S3, S4 and S5 respectively.Three samples (stem nodes)were taken in T1 period (T1-S1, T1-S2, T1-S3), four samples in T2 period(T2-S1, T2-S2, T2-S3, T2-S4) and 5 samples in T3 period (T3-S1, T3-S2, T3-S3, T3-S4, T3-S5) (Fig.1a).A total of 36 samples were collected, each including three biological replicates.All samples were immediately frozen in liquid nitrogen and stored at -80°C before total RNA extraction.

Determination of the content of cellulose, hemicellulose and lignin

The sample to be tested was naturally air-dried or placed in an oven to dry (temperature not exceeding 50°C) until the moisture is less than 10%, crushed and screened by the grinder.Took a portion between 20-80 mesh, cooled and stored in a sealed bag for further analysis.Content analysis was conducted by high-performance liquid chromatographic(HPLC) according to the NERL method [26-27]. The content of cellulose, hemicellulose and lignin are the percentage of the dry weight of the sample.

Cell wall morphology observation

The periods with the highest and lowest cellulose content in three different development stages of elephant grass stems were selected as samples,then cut thefresh stalks into 1cm×1cm pieces and placed in 2.5% glutaraldehyde phosphate buffer (0.1mol/L, pH=7.0). Samples were treated as reported [28] and observed under the EM-420transmission electron microscope (Philips Electronics, Holland).

Micro CT observation of elephant grass stem

Thestalks of elephant grass were cut into 1cm×1cm pieces and put them into 2.5% glutaraldehyde phosphate buffer (0.1mol/L, pH=7.0),fixed in a 4°Cstoragecabinet for 3h, washed the fixed tissue twice with 1mol/L phosphate buffer, and placed the sample in a 40°Coven for 24h. Cross-section of the processed sample were observed under SkyScan 2211micro-CT (Bruker, Belgium).

RNA extraction and transcriptome sequencing

Total RNA was extracted from stems using HiPure Plant RNA Kit (Magen, Guangzhou, China) according to the manufacturer's instructions. A total of 3μg ofRNA per sample was used for library preparation with insert sizes of350bp and sequenced on Novaseq 6000 (Illumina, USA). RNA-seq analysis was conducted using the Illumina platform according to the standard protocols [29].

Quality control of RNA-seq and transcriptome assembly

Raw data (raw reads) of fastq format were firstly processed through Trim Galore[30].In this step,clean data (clean reads) were obtained by removing reads containing adapter,ploy-N and low quality reads from raw data. At the same time, Q20, Q30, GC-content and sequence duplication level of the clean data were calculated. All the downstream analyses were based on clean data with high quality. To get the non-redundant transcriptome, all assembled transcriptome were clustered by coreset [31]and all parameters were set as default.

Gene function annotation

Gene function was annotated based on the following databases:Nr (NCBI non-redundant protein sequences) (https://www.ncbi.nlm.nih.gov/); Swiss-Prot(A manually annotated and reviewed protein sequence database) (http://www.gpmaw.com/html/swiss-prot.html); KO (KEGG Ortholog database) (https://www.kegg.jp/)GO (Gene Ontology) (http://geneontology.org/).

Differential expression and enrichment analysis

Gene expression levels were calculated and normalized by reads per kilobase of exon model per million mapped reads (RPKM). Genes ofRPKM>1.0 were defined tobe expressed. Differential expression analysis was performed using the R package DESeq [22].The p-values were adjusted by using Benjamini-Hochberg (BH)method, genes with an adjusted p-value <0.05 were assigned as DEGs. GO enrichment and KEGG pathway enrichment analysis of the DEGs was implemented by the R Package Goseq [32].The population set is a set with all the annotated genes, andthe study set consisted of the DEGs in the population set. The p-values were adjusted as differential expression analysis does.

Weighted gene co-expression network analysis

Gene expression patterns for common differentially expressed genes in T1, T2, T3 period were used to construct a co-expression network using WGCNA. Genes without expression detected in all tissues were

removed before analyses. Soft thresholds were set based on the scale-free topology criterion[33-34].

qRT-PCR analysis

DNase-treated RNA (2 μg) was reverse transcribed using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, USA). Gene-specific primers were designed using Primer Express (v3.0, Applied Biosystems).Quantitative reverse transcription PCR (qRT-PCR) assays were performed using SYBR Green I Master Mix (Roche, Indianapolis,USA). Three biological and three technical replicates for each reaction were analyzed on a LightCycler 480 instrument (Roche USA)with a first step of95°C for 5 min followed by 40 cycles of 95°C for 10 s, 60°C for 10 s, and 72°C for 20s. Melting curves were generated using the following programme: 95°C for 15 s, 60°C for 15 s, and 95°C for 15s. 18S was used as an internal control gene. Data analysis was calculatedusing the $2^{-\Delta\Delta CT}$ method. Significant differences between different samples were tested with IBM SPSS Statistics 19.0software.

Declarations

Acknowledgments

The RNA-seq were performed with the help of the Nextomics Biosciences Institute in Wuhan, China.

Consent to publication

Not applicable.

Authors' contributions

WQZ SKZ performed the experiments, data analysis andmanuscript writing. XQL, CL, XWL, GYD participated in some experiments and data analysis. TX designed the project and contributed to writing of the manuscript and approved the final manuscript.

Funding

This study was financially supported by the Integration of Science and Education Program Foundation for the Talents by Qilu University of Technology (No. 2018-81110268), Foundation of State Key Laboratory of Biobased Material and Green Papermaking (No. 2419010205 and No. 23190444).

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

There is no ethics approval and consent to participate in this manuscript.

Competing interests

The authors declare that they have no conflict of interests.

References

1. Hu WJ, Harding SA, Lung J. Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. Nature Biotechnology. 1999; doi:1038/11758
2. Zhang CB, Chen LH, Jiang J. Why fine tree roots are stronger than thicker roots: The role of cellulose and lignin in relation to slope stability. Geomorphology. 2014; doi:1016/j.geomorph.2013.09.024
3. Samuga, Joshi CP. A new cellulose synthase gene (*PtrCesA*2) from aspen xylem is orthologous to Arabidopsis *AtCesA*7 (irx3) gene associated with secondary cell wall synthesis. 2020; doi:1016/s0378-1119(02)00864-8
4. Kaur S, Dhugga KS, Gill K. Novel Structural and Functional Motifs in cellulose synthase (*CesA*) Genes of Bread Wheat (*Triticum aestivum*, L.). Plos One. 2016; doi:1371/journal.pone.0147046
5. Appenzeller L, Doblin M, Barreiro R. Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (*CesA*) gene family. Cellulose. 11 2004; doi:1023/b:cell.0000046417.84715.27
6. Taylor NG, Scheible WR, Cutler S. The irregular xylem3 locus of Arabidopsis encodes a cellulose synthase required for secondary cell wall synthesis. The Plant Cell. 1999; doi:2307/3870813
7. Bhandari S, Fujino T, Thammanagowda S. Xylem-specific and tension stress-responsive coexpression of KORRIGAN endoglucanase and three secondary wall-associated cellulose synthase genes in aspen trees. Planta. 2006; doi:2307/23389483
8. Bonawitz ND, Chapple C. The Genetics of Lignin Biosynthesis: Connecting Genotype to Phenotype. Annual Review of Genetics. 2010; doi:1146/annurev-genet-102209-163508
9. Silvia F, Montserrat C, En. Antonio, Kan W, Sami I, et al. Altered Lignin Biosynthesis Improves Cellulosic Bioethanol Production in Transgenic Maize Plants Down-Regulated for Cinnamyl Alcohol Dehydrogenase. Molecular Plant. 2012; doi:1093/mp/ssr097
10. Acker R, Vanholme R, Véronique S. Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in *Arabidopsis thaliana*. Biotechnology for Biofuels. 2013; doi:1186/1754-6834-6-46
11. Pilate G, Guiney E, Holt K. Field and pulping performances of transgenic trees with altered lignification. Nature Biotechnology. 2020; doi: 1038/nbt0602-607
12. Tu Y, Rochfort S, Liu Z. Functional Analyses of *Caffeic Acid O-Methyltransferase* and *Cinnamoyl-CoA-Reductase* Genes from Perennial Ryegrass (*Lolium perenne*). Plant Cell. 2020; doi:1105/tpc.109.072827
13. Wang Z, Li R, Xu J. Sodium hydroxide pretreatment of genetically modified switchgrass for improved enzymatic release of sugars. Bioresource Technology. 2012; doi: 1016/j.biortech.2012.01.097

14. José C, Prinsen P, Rencoret J. Structural characterization of the lignin in the cortex and pith of elephant grass (*Pennisetum purpureum*) stems. Journal of Agricultural and Food Chemistry. 2012; doi: 1021/jf300099g
15. Nyambati E, Nyambati M, Sollenberger L, et al. Feed intake and lactation performance of dairy cows offered na'piergrass supplemented with legume hay. Livestock Production Science. 2003; doi: 1016/s0301-6226(03)00094-0
16. Strezov V, Evans TJ, Hayman C. Thermal conversion of elephant grass (*Pennisetum Purpureum*) to bio-gas, bio-oil and charcoal. Bioresour Technol. 2008; doi:10.1016/j.biortech.2008.02.039
17. Liu X, Shen Y, Lou L. Copper tolerance of the biomass crops Elephant grass (*Pennisetum purpureum*Schumach), Vetiver grass (*Vetiveriazizanioides*) and the upland reed (*Phragmites australis*) in soil culture. Biotechnology Advances. 2009; doi: 1016/j.biotechadv.2009.04.017
18. Somerville, Youngs H, Taylor C. Feedstocks for Lignocellulosic Biofuels. Science. 2010; doi: 1126/science.1189268
19. Kawube G, Alicai T, Wanjala B. Genetic Diversity in Napier Grass (*Pennisetum purpureum*) Assessed by SSR Markers. Journal of Agricultural Science. 2015; doi: 5539/jas.v7n7p147
20. Bhandari P, Sukanya DH, Ramesh CR. Application of Isozyme Data in Fingerprinting Napier Grass (Pennisetum purpureum Schum.) for Germplasm Management. Genetic Resources and Crop Evolution. 2006; doi: 1007/s10722-004-6120-2
21. Harris K, Anderson W, Malik R. Genetic relationships among napiergrass (*Pennisetum purpureum*) nursery accessions using AFLP markers. Plant Genetic Resources. 2010; doi: 10.1017/S1479262109990165
22. Zhou S, Chen J, Lai J, Yin G, Chen P, et al. Integrative analysis of metabolome and transcriptome reveals anthocyanins biosynthesis regulation in grass species *Pennisetum purpureu* Industrial Crops & Products. 2019; doi:10.1016/j.indcrop.2019.111470
23. Zhao J, Xia B, Meng Y, Yang Z, Pan L, et al. Transcriptome Analysis to Shed Light on the Molecular Mechanisms of Early Responses to Cadmium in Roots and Leaves of King Grass (*Pennisetumamericanum × P. purpureum*). International Journal of Molecular Sciences. 2019; doi: 3390/ijms20102532
24. Bhandari P, Sukanya DH, Ramesh CR. Application of Isozyme Data in Fingerprinting Napier Grass (*Pennisetum purpureum* Schum.) for Germplasm Management. Genetic Resources and Crop Evolution. 2006; doi: 1007/s10722-004-6120-2
25. Jakob K, Zhou F, Paterson AH. Genetic improvement of C4 grasses as cellulosic biofuel feedstocks. Vitro Cellular & Developmental Biology Plant. 2009; doi: 1007/978-1-4419-7145-6_7
26. Lu X, Zheng X, Li X. Adsorption and mechanism of cellulase enzymes onto lignin isolated from corn stover pretreated with liquid hot water. Biotechnology for Biofuels. 2016; doi: 1186/s13068-016-0531-0
27. Du J, Cao Y, Liu G. Identifying and overcoming the effect of mass transfer limitation on decreased yield in enzymatic hydrolysis of lignocellulose at high solid concentrations. Bioresource Technology.

2017; doi: 1016/j.biortech.2017.01.011

28. Liu Y, Muhammad R, Yan L, Zeng Y, Jiang C. Boron and calcium deficiency disturbing the growth of trifoliate rootstock seedlings (Poncirus trifoliate L.) by changing root architecture and cell wall. Plant Physiology and Biochemistry. 2019; doi: 1016/j.plaphy.2019.10.007
29. Jiang C. Efficient extraction of RNA from various Camellia species rich in secondary metabolites for deep transcriptome sequencing and gene expression analysis. African Journal of Biotechnology. 2011; doi: 5897/AJB11.235
30. Kanno M, Kijima A. Quantitative and Qualitative Evaluation on the Color Variation of the Japanese Sea Cucumber Stichopus japonicus. Aquaculture Science. 2002; doi: 11233/aquaculturesci1953.50.63
31. Davidson NM, Oshlack A. Corset: Enabling differential gene expression analysis for de novoassembled transcriptomes. Genome Biology. 2014; doi: 1186/s13059-014-0410-6
32. Young MD, Wakefield MJ, Smyth GK. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biology. 2010; doi: 1186/gb-2010-11-2-r14
33. Zhang B, Horvath S. A general framework for weighted gene coexpression network analysis. Statal Applications in Genetics and Molecular Biology. 2005; doi: 2202/1544-6115.1128
34. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; doi: 10.1186/1471-2105-9-559
35. Douglas CJ. Phenylpropanoid metabolism and lignin biosynthesis: From weeds to trees. Trends in Plant Science. 1996; doi: 1016/1360-1385(96)10019-4
36. Kleczkowski LA. Glucose activation and metabolism through UDP-glucose pyrophosphorylase in plants. Phytochemistry. 1994; doi: 1016/s0031-9422(00)89568-0
37. Martin LK, Haigler CH. Cool temperature hinders flux from glucose to sucrose during cellulose synthesis in secondary wall stage cotton fibers. Cellulose. 2004; doi: 1023/b:cell.0000046420.10403.15
38. John W, Downton S, Hawker JS. Enzymes of starch and sucrose metabolism in *Zea mays* Phytochemistry. 1973; doi: 10.1016/0031-9422(73)80366-8
39. Lu X, Li C, Wang X, Zhang W, Xia T. Enzymatic sugar production from elephant grass and reed straw through pretreatments and hydrolysis with addition of thioredoxin-His-S. Biotechnology for Biofuels. 2019; doi:1186/s13068-019-1629-y
40. Xu B, Luis L, Escamilla T, Sathitsuksanoh N. Silencing of 4-coumarate: coenzyme A ligase in switchgrass leads to reduced lignin content and improved fermentable sugar yields for biofuel production. New Phytologist. 2011; doi: 1111/j.1469-8137.2011.03830.x
41. Dixon R, Sewalt V, Howles P. Genetic manipulation of the phenylpropanoid pathway in transgenic tobacco: new fundamental insights and prospects for crop improvement. Biotechnology & Biotechnological Equipment. 1996; doi: 1016/S0734-9750(96)00033-X

42. Paul H, Sameer A, Masoud JW. Overexpression of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase in tobacco cell suspension cultures. Plant Biotechnology and In Vitro Biology in the 21st Century. 1999; doi: 10.1007/978-94-011-4661-6_69
43. Boerjan, J. Ralph, M. Baucher, Lignin Biosynthesis. Annual Review of Plant Biology. 1999;36:519-546.
44. Qin L, Qin X, Chen Y, Zeng W. The cotton β-galactosyltransferase 1 (*GalT*1) that galactosylates arabinogalactan-proteins participates in controlling fiber development. Plant Journal. 2016; doi: 1111/tpj.13434

Figures

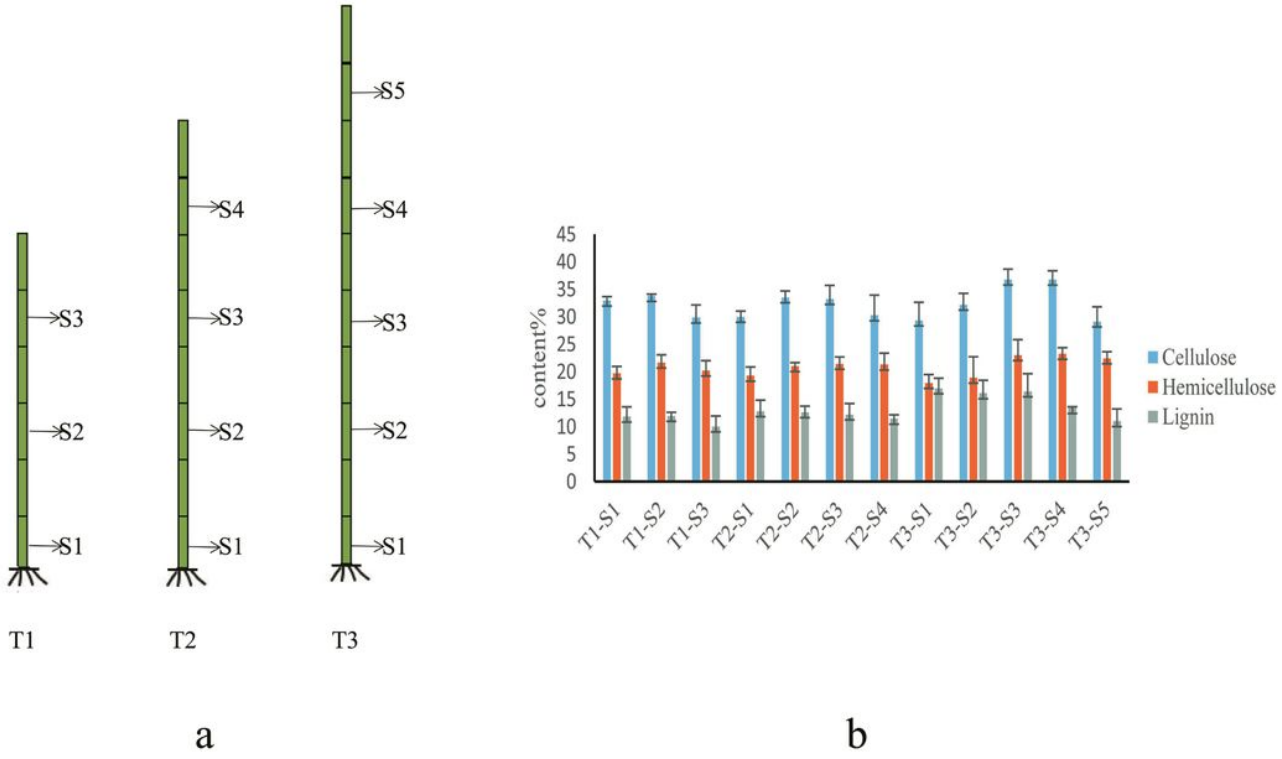
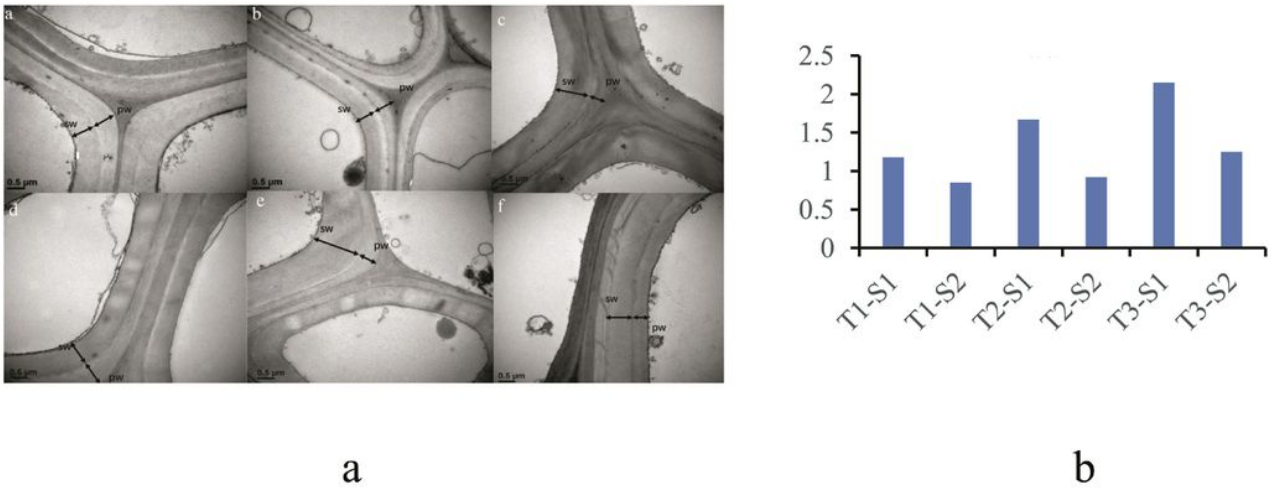


Figure 1

Contents of cellulose, hemicellulose and lignin in different development stages of elephant grass stems. (a) Sampling period and position. (b) Cellulose, hemicellulose and lignin contents



a b

Figure 2

The cell wall morphology and thickness changes in different development stages of elephant grass stems. (a) a.T1-S1, b.T1-S2, c.T2-S1, d.T2-S2, e. T3-S1, f. T3-S3. SW:Secondary cell wall,PW:Primary cell wall. (b) thickness ratio of secondary cell wall to primary cell wall.

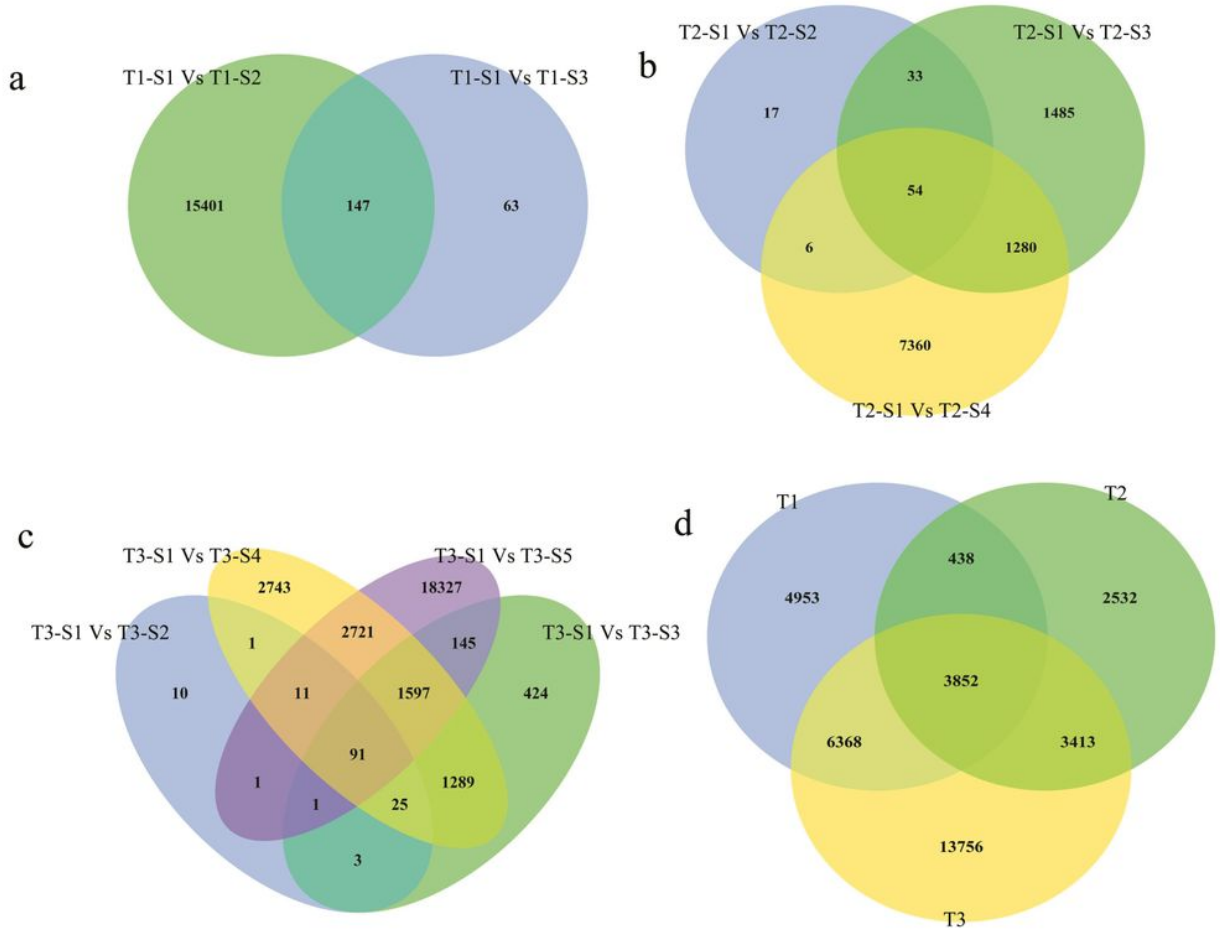


Figure 3

The differentially expressed genes (DEGs) identified by RNA sequence analysis in stem tissues of development stages of elephant grass. (a) The number of DEG between three adjacent stages (T1-S1, T1-S2, T1-S3).green:T1-S1 Vs T1-S2 , blue:T1-S1 Vs T1-S3. (b) The number of DEG between the four adjacent stages (T2-S1, T2-S2, T2-S3, T2-S4).green:T2-S1 Vs T2-S2,blue:T2-S1 Vs T2-S3,yellow:T2-S1 Vs T2-S4. (c) The number of DEG between the five adjacent stages (T3-S1, T3-S2, T3-S3, T3-S4). Purple:T3-S1 Vs T3-S5,green:T3-S1 Vs T3-S2, blue:T3-S1 Vs T3-S3, yellow:T3-S1 Vs T3-S4. (d) DEGs between T1, T2 and T3 stages.blue:T1 (DEGs),green:T2 (DEGs),yellow: T3 (DEGs).

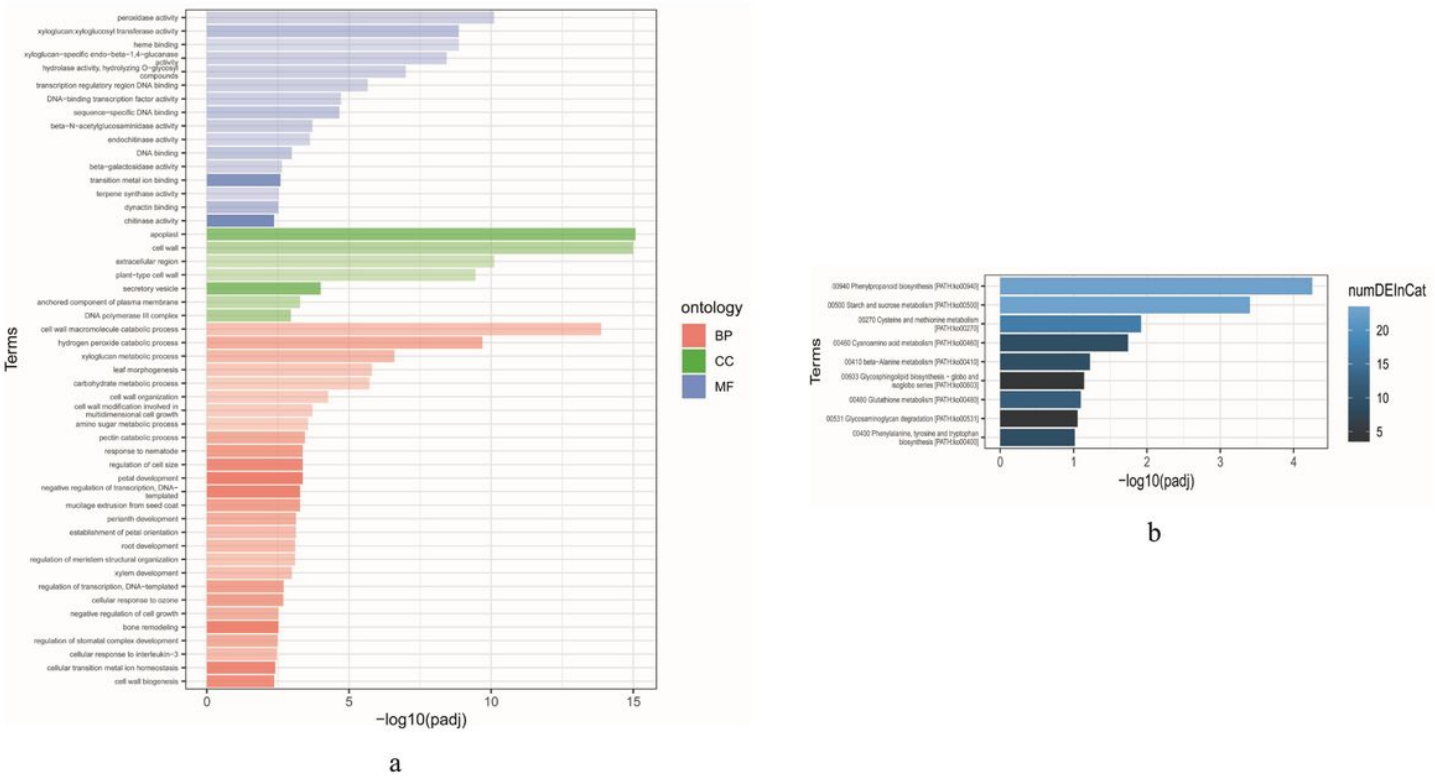


Figure 4

The 3852 DEGs co-expressed in three developmental stages of T1, T2, and T3 were enriched by GO and KEGG. (a) GO enrichment analysis.(b) KEGGenrichment analysis.

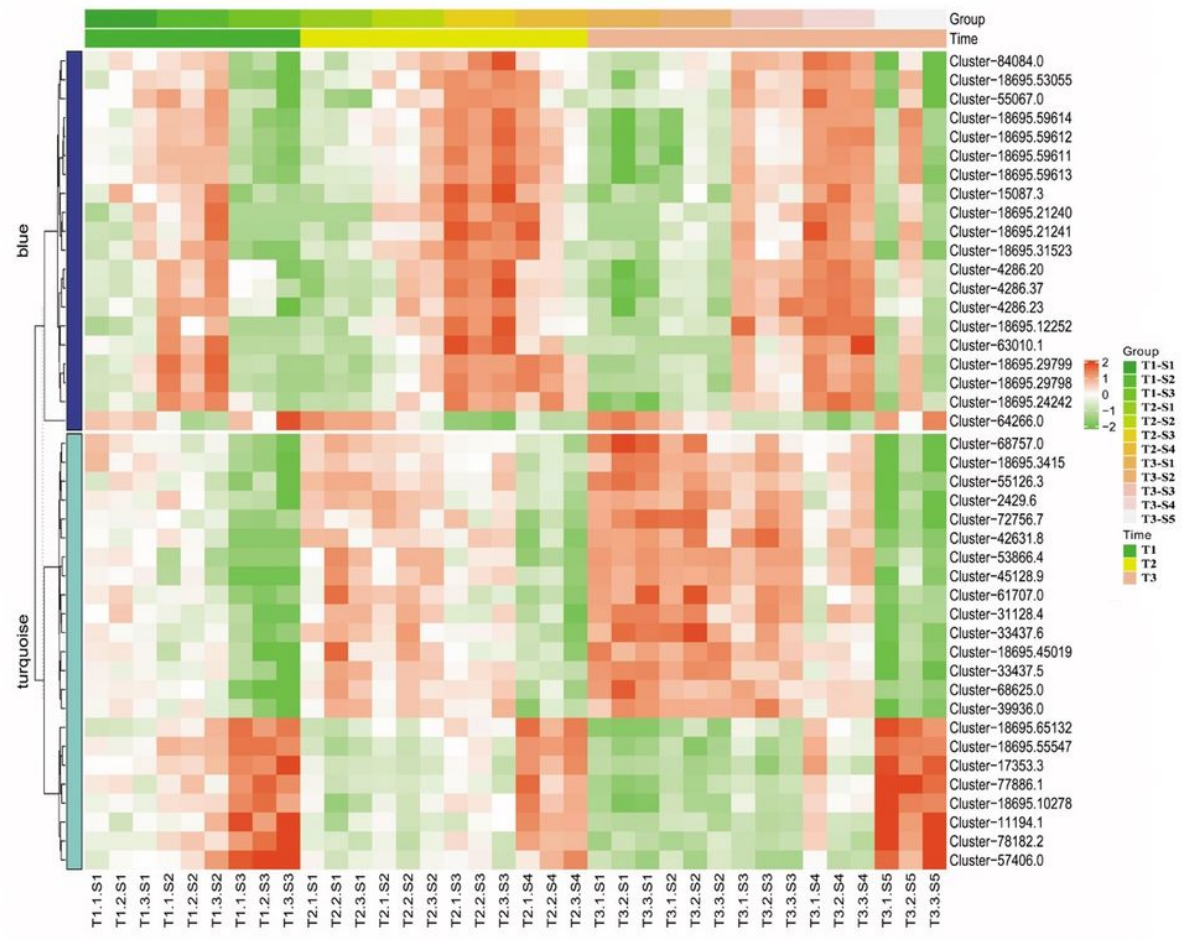


Figure 5

The heat maps of the expression pattern of 43 genes in WGCNA analysis at different development stages of elephant grass stems.

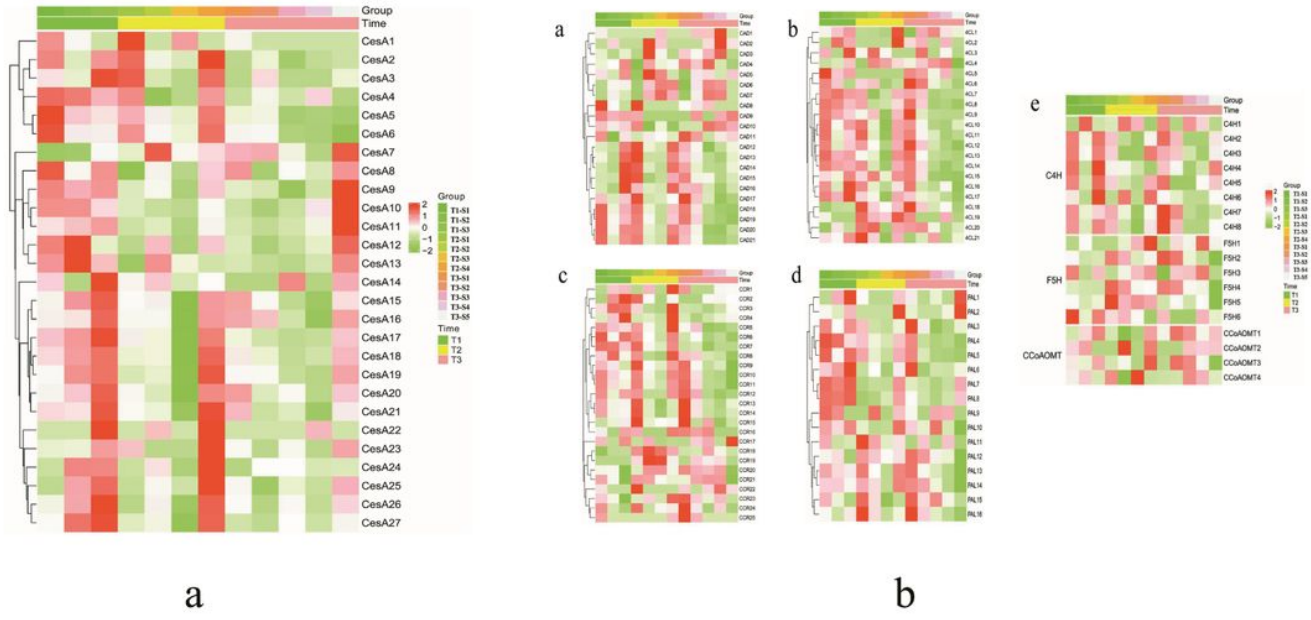


Figure 6

The heat maps of the expression pattern of cellulose synthase genes andlignin synthesis-related genes at different developmental stages of elephant grass stems. (a) Heat map of the expression pattern of 27 cellulose synthase genes. (b) Heat map of the expression pattern of some lignin synthesis-related genes. The grid with four different colors shows the absolute expression of genes, which are represented by different scale levels 2, 1, 0, -1, and -2, respectively. PAL:PhenylalanineAmmonia-Lyase;C4H: CinnamicAcid4-Hydroxylase;CCR:Cinnamoyl-CoAReductase;CAD:CinnamylAlcoholDehydrogenase;F5H:FerulicAcid5-Hydroxylase;CoAOMT:Caffeoyl-CoAO-methytransferase; COMT:5-HydroxyferulicAcidO-MethyltransferaseBispecificCaffeicAcid.

Supplementary Files

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

- Additionalfile1.tif
- Additionalfile2.tif
- Additionalfile3.tif
- Additionalfile4.tif
- Additionalfile5.tif
- Additionalfile6.rtf

- Additionalfile7.rtf
- Additionalfile8.rtf