Allele-specific hormone dynamics in highly transgressive F2 biomass segregants in sugarcane (Saccharum species)

Noor-ul-Ain Ain
Fujian and Agriculture and Forestry University
Xingtan Zhang
Chinese Academy of Agricultural Sciences
Habiba Habiba
Fujian and Agriculture and Forestry University
Jishan Lin
Fujian and Agriculture and Forestry University
Ray Ming (rayming@illinois.edu)
Fujian and Agriculture and Forestry University

Research Article

Keywords: sugarcane biomass, extreme segregants, gene duplications, auxin, jasmonic acid, abscisic acid

Posted Date: December 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2329441/v1

Abstract

Background

Sugarcane is a competitive candidate to serve as feedstock for biofuel production worldwide. Plant hormones play an important role in growth and development and regulate biomass yield. Understanding hormonal dynamics and their regulation is critical for increasing sugarcane biomass. To enhance our understanding of biological pathways involved in biomass accumulation, transcriptome analysis of highly segregating F2 introgression hybrids derived from the cross of *Saccharum officinarum*'LA Purple' and wild *S. robustum* 'MOL5829' was performed in this study. Recently sequenced information rich allele-specific genome of *S. officinarum*served as a reference to identify differentially regulated genes in two groups of extreme segregants of biomass.

Results

Overall, 8059 differentially expressed genes (log2FC > 2) were grouped into distinct categories viz. Gene Models (21.5%) (GMs), alleles (68%), paralogs (10%) and tandemly duplicated genes (0.14%). KEGG analysis showed enrichment of auxin (IAA), jasmonic acid (JA) and abscisic acid (ABA) related pathways and interesting regulatory roles of three hormone repressor gene families (Aux/IAA, PP2C, and JAZ) of IAA, ABA and JA, respectively. Signaling pathways of these hormones indicated down-regulation of AUX/IAA and PP2C and up-regulation of JAZ repressor genes in high biomass group controlling the expression of downstream growth and development genes. Endogenous hormone levels show higher IAA and ABA contents in high biomass and vice versa for JA. Weighted co-expression network analysis (WGCNA) demonstrated high connectivity between identified hormones related key genes and cell wall structural genes in high biomass genotypes. FPKM and RT-PCR based expression analysis showed the up-regulation of carbohydrate structural genes and down regulation of inflorescence and senescence related genes, which indicated an extended vegetative growth phase in high biomass genotypes. Furthermore, high biomass group displayed modulated regulation of hormones which was achieved by the cumulative expression of Gene Models (GMs), dominant alleles, paralogs, and tandemly duplicated genes.

Conclusion

Our data revealed that activators and repressors of disparate hormone (IAA, JA, and ABA) signaling pathways are the points of hormone crosstalk in contrasting biomass F2 segregants and could be applied for engineering high biomass acquiring varieties.

Introduction

With the advent of the green revolution and population explosion, the energy demand has risen exponentially. Conventional methods for energy generation, for instance, burning of natural fossil fuels such as gas, coal, and petroleum, are replenishing. Moreover, they have a higher carbon footprint and add

their share as greenhouse gases (GHGs) leading to warming and climate change [1]. Thus, urges towards the environment-friendly and sustainable energy system have become inevitable and a promising area in research. Biofuels contribute (9.4%) as a 4th major source of energy followed by oil (30.9%) coal (26.8%) and natural gas (3.2%), and is gaining special attention as a critical renewable energy source in the local areas worldwide [2]. One of the important feed stocks of biofuels is the biomass generated by plants as carbohydrate "**building blocks**" at the cost of ambient air CO_2 and soil water [3]. Panicoid grasses being C4 plants are most popular as feedstocks, and their starch and lignocellulosic polymers are converted into monosaccharides prior to fermentation for bioethanol production. Sugarcane is an efficient accumulator of biomass and is considered as a benchmark in 1st and 2nd generation biofuels [4]. *S. officinarum* being the member of *Saccharum genus,* is cultivated and highly efficient in producing and storing sugars. Among the two wild species, *S. robustum* and *S. spontaneum,* the former was once thought to be the ancestor of *S. officainarum* but recent genomic analyses of these three species proved otherwise [5]. These two wild species, particularly *S. spontaneum*, were extensively involved in breeding programs for increase in sugars, biomass and stress related traits [6]. Cultivated sugarcane has many production constraints, for example biotic, abiotic stresses and long duration of the crop that decreases the sugar recovery and biomass. The genome complexity due to polyploid nature of sugarcane is a primary reason for hindrance in variety development via conventional breeding. Modern sugarcane breeding methods can help to develop a framework to dissect biological drivers for improvement of important traits. Being a principal sugar producing crop, major research challenges were stagnant yield, lower sugar recovery and disease resistance [7]. Moreover, extensive research was focused on sugar-related traits in sugarcane but there are knowledge gaps in developing high biomass sugarcanes crops called as 'energy canes' [8].

Plant growth and development are partly controlled by hormones, which act individually or in coordination, in response to both developmental and environmental cues. Auxin (IAA) is an important hormone that exerts pleiotropic effects on growth related processes in spatio-temporal pattern of regulation [9]. Several gene families regulate the biosynthesis, transport, and signaling of auxin in plants. For example, YUC, PIN, LAX, Aux/IAA, GH3, ARF, SAUR, and LBD. Aux/IAA genes, being negative regulator, contribute to the diversity of auxin-induced responses in different organs and at various developmental stages. Aux/IAA proteins repress auxin mediated gene regulation by binding Auxin Response factors (ARFs) through their conserved domains. SAUR genes are early auxin responsive and encode highly unstable mRNAs. Gretchen Hagen3 (GH3) proteins play a role for conjugation of amino acids to various acyl substrates and thus maintain the homeostasis of jasmonic acid (JA) and IAA [10]. SAURs play roles in tissues elongation and increase in biomass [11]. Mutations in different AUX-IAA genes such as IAA14/SLR [12], IAA17/ AXR3 [13], IAA19/MSG2 [14], and IAA28 [15] in *Arabidopsis* showed altered responses to auxins and modified plant morphology.

Similarly, jasmonic acid (JA) is a prototypical member of a class of related oxylipins, which are collectively termed as jasmonates and are important signaling molecules in stress and anti-oxidant response [16]. JAZ-related genes with ZIM domain regarded as TIFY genes are the repressors of JA

signaling, and are degraded by SCF complex to initiate the transcription of JA responsive genes. CORONATINE INSENSITIVE 1 (COI1–JAZ) complex acts as a substrate for the perception of incoming jasmonic acid (JA–Ile), and after ubiquitination of JAZ repressors (TIFY), MYC transcription factors initiate the transcription of JA responsive genes [17–19]. Regarding abscisic acid, the three primary proteins regulate the signaling and expression from ABA perception to transcription of ABA responsive genes. These three pillars include: pyrabactin resistance 1 (PYR)/PYR1-like (PYL)/regulatory components of ABA receptor (RCAR), protein phosphatase 2C (PP2C) and sucrose nonfermenting-1 (SNF1)-related protein kinase 2 (SnRK2) [20]. When cellular ABA level is very low under normal growth conditions, group 2 Protein Phosphatases (PP2C) genes dimerize with subclass III SnRK2s and inhibit their kinase activity. Under stress conditions, biosynthesis of ABA takes place and PYR/PYL/RCAR receptors cause the inhibition of PP2Cs enabling the activation of SnRK2s kinases and regulating the downstream genes and TFs [21, 22].

Therefore, considering the role of hormone signaling in growth processes, we performed RNA-Seq analysis of F2 introgression segregants of high and low biomass and identified different functional categories for biomass regulation. Using allele specific reference genome of *S. officinarum* we identified that hormone related Gene models (GMs), alleles, paralogs, and tandemly duplicated genes controlled by their native promoters provide regulatory role in biomass accumulation in two contrasting biomass phenotypes of F2 populations.

Results

Sequence read alignment and differential expression analysis

RNA-Seq data were generated from 22 extreme segregants of an F_2 population derived from interspecific cross between *Saccharum officinarum* 'LA Purple' and wild species *Saccharum robustum* MOL5829, including 14 high biomass and 8 low biomass segregants (Additional file 1; Table S-1). Extremely low and high biomass hybrids ranged between 18.0 to 37.0 and 78.7 to 106.5 metric tons/hectare of 12 months old plants, respectively. After mRNA extraction from the 3^{rd}, 6^{th} and 9^{th} internodes and leaf, cDNA libraries (combined of all plant sections) were constructed and used for RNA sequencing. The details of phenotypic data and sugar contents (reducing and non-reducing) was previously reported and sugarcane transcriptome was analyzed using the *Sorghum bicolor* L. Moench as a reference genome due to the unavailability of the high-quality sugarcane reference genome at that time [23]. Although the small genome of sorghum is an attractive model for tropical C4 photosynthetic grasses, but it has diploid genome and minimal duplication events in the genome. The genome of *Saccharum* species is highly polyploid and aneuploid, exhibiting variable ploidy levels among different loci, a large genome size and abundant repetitive sequences [24]. Comparing sugarcane transcript data with sorghum would compromise the contribution of allelic and duplicated copies of gene models, which are equally contributing in polyploid crops. Therefore, we re-analyzed our Illumina raw sequences of 22 RNA-seq samples (34 GB, 293.5 million reads and 151 nucleotides length) with a high-quality allele specific *Saccharum officinarum* reference genome. After quality trimming and removing the adapter

sequences, data were reduced to 19 GB and 191.3 million clean reads. Although, HISAT2 is an efficient aligner with better performance in the terms of high percentage of correctly aligned reads and high-precision rates. Nevertheless, in our data, STAR outperformed in alignment with a higher alignment rate (~85%) in all the samples followed by transcriptome assembly and differential expression analysis.

Among 22 samples, 4 samples from high biomass and 1 sample from the low biomass group were excluded, being outliers in clustering analysis of samples. Finally, differential expression of 17 samples (10 HB, 7 LB) resulted in 174593 genes and alleles (GMs (Gene Models), alleles, tandem duplicates and paralogs, (Additional file 1; Fig. S-1, Table S-1) which were sorted using the thresholds, FDR |≤0.05| and pval |<0.01| while log2FC |>2| for up-regulated and log2FC |<-2| for down-regulated. Using the above criteria of selection, 8059 genes and alleles were regarded as DEGs, 5540 were up-regulated and 2519 were down-regulated in high biomass samples (Additional file 2; Table S-3). Among these 8059 genes and alleles, 21.5% belonged to GMs, 68% was allelic contribution, and 0.14% were tandem duplicates and 10% paralogs. Further, Partial least squares discriminant analysis (PLS-DA), hierarchical clustering and volcano plot of the differentially expressed genes to visualize expression patterns in samples based on FPKM and log2FC (Fig. 1). PLS-DA plot reduced the dimensions of large data and showed distinct differences in gene expression patterns among low and high biomass segregants. PC1(26%) exhibited maximum variation in the clustering patterns and showed low biomass samples are tightly clustered on the positive x-axis (Fig.1A). Hierarchical clustering heat map showed variability between two biomass groups, by clustering all the genes in six clusters using expression values (Fig. 1B). The volcano plot visualization showed an overall trend of fold change in high biomass genotypes (Fig. 1C).

Exploration of biological processes involved in growth and biomass by GO

To explore the pathways underlying a large number of differentially expressed GMs, alleles and duplicated genes, GO enrichment analysis was performed (Fig. 2B, Additional file 1; Fig. S-2). Regarding GO, a significant number of genes were enriched in the biological process category in high and low biomass samples, including signaling (GO:0023052), transmembrane transport (GO:0055085), and cellular nitrogen metabolism biosynthetic (GO:0034641) processes. Significant number of genes belonging to membrane metabolism (GO:0016020), cell wall (GO:0005618) and extracellular region related processes (GO:0005576) were identified in the cellular component's category. Moreover, the genes involved in electron transport, catalysis (GO:0003824), transmembrane transporter activity (GO:0022857), transferase activity (GO:0016740) and macromolecules modification (GO:0043412) were significantly enriched in the molecular function category (Additional file 1; Fig. S-2, Additional file 2; Table S-6).

Hormone dynamics in extreme segregants:

To identify the regulatory role of hormones in growth, cell wall remodeling, and biomass accumulation, KEGG analysis was performed for the functional annotation of DEGs between high and low biomass samples (Fig. 2A, Additional file 2; Table S-4, S-7). A considerable number of genes were enriched in different metabolic pathways in KEGG analysis. For example, 96 GMs, alleles and duplicates were differentially regulated in signal transduction pathways, 346 in carbohydrate metabolism (glycolysis,

sugars, starch, pyruvate metabolism), and 42 in glycan biosynthesis. Hormone-related sequences obtained from KEGG analysis were categorized into GMs (Gene Model), allelic copies and gene duplications (tandem and paralogous duplicate) using eight-column allele table. In polyploid species, alleles are the homologous genes on the homologous chromosomes at the same loci. Tandemly duplicated genes are originated owing to unequal crossing over, categorized on the criteria if they are having a difference of configurable gene rank =1. Dispersly duplicated paralogs are neither neighbors nor on homologous chromosome, and they are re-labelled based on blastp hits [25, 26]. The details of all the discussed genes and their respective categories (GMs, alleles, tandem and paralogous duplicates) are discussed in (Additional file 2; Table S-5).

Auxin related genes:

Auxin, an essential hormone in modulating growth and development, showed the enrichment of several GMs, alleles and duplications in high biomass samples. Annotated sequences were confirmed using blastp in phytozome and NCBI (https://phytozome.jgi.doe.gov/pz/portal.html). Sorghum genome, a model for C4 plant species having high evolutionary proximity with sugarcane, has a significant number of auxin-related gene family members i.e., 26 Aux/IAA genes, 16 GH3 genes, 36 LBD genes, and 25 ARF genes [27]. Two members (Soffic.03G0032240-1A, Soffic.01G0025830-1P) of auxin influx AUX 1 (Auxin influx carrier protein) were up and down regulated in HB whereas one allele (Soffic.05G0006040-2B) related to auxin efflux PIN (PIN-FORMED) was up-regulated. There was over representation of five auxin-responsive Aux/IAA (repressor) sequences (GMs and alleles) in high biomass genotypes. Two members (Soffic.08G0009450-1P, Soffic.03G0018710-2B) of Gretchen Hagen3 (GH3) family, (indole-3-acetic acid-amido synthetase) showed a profound upregulation in high biomass genotypes. 18 members of small auxin-up RNAs (*SAUR*) family, and 6 members encoding lateral organ boundaries (*LBD*) were up-regulated in HB genotypes as compared to low biomass genotypes. Among 8 differentially regulated ARF (AUXIN RESPONSIVE FACTORS) sequences, 7 showed up-regulation in HB genotypes. Active contribution of auxin signaling and responsive genes in HB genotypes predicts auxin mediated growth changes accounting for high biomass (Fig. 3A, Additional file 2; Table S-5).

Jasmonic acid (JA) signaling pathway genes

Jasmonic acid is a lipid-derived hormone, which regulates important stress, and immunity related responses. Many jasmonic acid-related genes were highly enriched in the signal transduction category identified by KEGG analysis (Fig. 3A, Additional file 2; Table S-5). Coronatine-insensitive (COI1) protein acts essentially as a receptor of jasmonoyl–isoleucine (JA–Ile), a conjugated form of JA. Three DEGs coding COI1 were downregulated in HB genotypes indicating absence of substrate for signaling of jasmonic acid responses. There was over representation of 15 ZIM-domain (JAZ) repressor proteins, among which 13 were upregulated in HB genotypes. The action of JAZ mediated repression is accomplished by Groucho/Tup1-type co-repressor TOPLESS (TPL), one allele encoding TPL is also up regulated in HB genotypes in our data. Thanks to the upregulation of JA repressor genes, which inhibit

the activation of JA downstream stress responsive genes inhibiting the activation of stress response and allowing the normal function of growth-related genes for increased biomass.

Abscisic acid (ABA) related genes

Regarding abscisic acid signaling pathway, 23 members were differentially expressed in high and low biomass genotypes. Type 2C protein phosphatases (PP2Cs) act as a negative regulator in abscisic acid signaling pathways. Our data shows down and up-regulation of 6 and 4 PP2C coding sequences. One class of protein kinases, i.e., Snf1-related protein kinases 2 (SnRK2s), is activated by autophosphorylation after repression of PP2C. The results exhibited that 7 SnRK2 genes showed up-regulation while 3 genes were down-regulated. Moreover, two Basic leucine zipper (bZIP) and one cis-regulatory DNA element i.e Abscisic acid (ABA) response elements (ABREs) were both down-regulated in HB genotypes, respectively (Additional file 1; Fig. S-3.).

Hormones responsive genes accounting for growth

Several genes may be induced or suppressed by the biologically active hormones in their vicinity, for example, cell wall synthesis CELLULOSE SYNTHASE (CESA) and cell expansion genes xyloglucan (XTH) and Expansins (EXP) strongly respond to auxin, brassinolide and gibberellic acid. Five alleles encoding CESA proteins, were up-regulated in HB genotypes. Cell wall loosening and expansion is critical for cell growth and secondary cell wall formation. A plethora of XTH genes and EXP genes were up regulated in HB genotypes i.e., XTH =19, EXP = 4. Three genes encoding galactosyl transferase, 8 genes belonging to GDP-mannose 4, 6 dehydratase, and 14 members of glycosyltransferase family were up-regulated in HB genotypes. UDP-glycosyltransferases play their indispensable roles in glucosylation of aglycones such as hormones, secondary metabolites engaged in stress and defense responses, and xenobiotics. 14 genes encoding UGTs were downregulated, whereas 11 genes were up regulated in HB genotypes (Fig. 3B, Additional file 2; Table S-5).

Flowering and senescence genes relating vegetative growth

Flowering is an important developmental transition, and time to flowering induction is key trait for deciding assimilate portioning and biomass. Major integrators of flowering related pathways are FLOWERING PROMOTING FACTOR 1 (FPF1), EARLY FLOWERING 3 (ELF3) and AGAMOUS-LIKE MADS-BOX PROTEIN (AGL12). Two gene coding FPF1, one ELF3 and one AGL12 were down-regulated in HB genotypes. Expression of certain other genes responsible for inflorescence architecture, for example HOMEOBOX related genes, BREVIPEDICELLUS, and PROTODERMAL FACTOR 2, was also down regulated in HB genotypes. Newly identified family genes (S-40) related to senescence were down regulated in HB genotypes, which supplements the above-mentioned activity of growth-related hormones and growth responsive genes (Fig. 3B, Additional file 2; Table S-5).

Identification of WGCNA modules associated with hormone and cell wall expansion genes

Weighted gene co-expression network analysis (WGCNA) was performed using 8540 non-redundant genes to identify modules with similar expression patterns. All the genes used in WGCNA showed 21 distinct modules (Fig. 4 A, Additional file 1; Fig. S-4.). From the module-trait relationship co-relation heat map, 'blue' and 'greenyellow' modules were selected as highly correlated (r = 0.83, r= 0.91) modules in HB genotypes. Cytoscape representation of blue module genes having edge weight over 0.10 indicated high level of connections between genes. Blue module network (left) showed the connections between different hormone repressor genes (IAA, TIFY, and PP2C) and cell wall related growth genes i.e., XTH (xyloglucan endohydrolysis (XEH) and or endotransglycosylation (XET)), UDP-glycosyl-transferase, Histone-lysine N-methyltransferase and O-methyltransferase (Fig. 3.). Green yellow module network (right) comprised of co-expression of DELLA, TIFY, SAUR, PP2C(HAB2), ABCB1, and cell wall related gene Pectinesterase (PME18). Module 'blue' of 1158 genes and 'greenyellow' of 412 genes appeared to be associated with hormone regulation and biomass gain (Fig. 4 B, 4C). These modules represented highly significant gene connections having the merge cut height (~0.25).

Expression profiles of transcription factors and protein kinases (kinome)

To apprehend the regulatory roles of transcriptions factors (TFs) and protein kinases (PKs), we identified TFs from the protein sequences of differentially expressed genes. In plants, DNA transcription involves more than 1500 TFs to regulate target genes by binding with cis-regulatory elements in promoter region [28]. Mapping the important transcription factors can provide insights to understand transcriptional regulation of the plant hormone genes in achieving the enhanced biomass (Fig. 5A, Additional file 2; Table S-8). In this RNA-Seq data, a high proportion of TFs of 46 different families were differentially expressed between high and low biomass groups, for example, 45 WRKY, 16 TIFY, 38 NAC, and 18 GRAS were differentially regulated between high and low biomass groups. The over-representation of the WRKY transcription factors in high biomass genotypes showed the crucial roles in regulating genes accounting for assimilates accumulation and modulating the growth processes. Owing to their dual role, they can either mediate activation or act as repressors of target genes. For example, in rice OsWRKY72 and OsWRKY77 they activate ABA signaling and repress GA signaling. Qiao and co-workers reported the direct role of WRKY in ABA signaling and drought mediated regulation responses. 7 out of 45 WRKY TFs were down-regulated in high biomass groups whereas, 38 TFs were up-regulated in high biomass [29]. All the 16 TIFY were up regulated in HB genotypes whereas, 21 NAC TFs were down regulated among 38 differentially expressed TFs. Among 52 bHLH differentially expressed TFs, 6 were down-regulated while 46 were up-regulated in HB genotypes. bHLH TFs have a binding site for regulating MeJA inducible genes, and also reported to be transcribed in root apical meristem as a target of ARF5 downstream in auxin signaling pathway [30]. Auxin responsive TFs containing B3 domain were also differentially expressed and 12 out of 18 ARFs were up-regulated in HB genotypes. Moreover, most members of HSF (Heat Shock Factors) were down-regulated in HB genotypes, which shows the inactivated immunity and growth inhibitory processes responses.

Similarly, protein sequences of DEGs were subjected to the identification of protein kinases (PK) from an online database iTAK. In *Sorghum bicolor* and *Saccharrum spontaneum*, kinome comprised 1210 PKs for

Sbi, and 2919 PKs besides allelic copies and segmental duplications [31]. From 8059 sequences, 132 sequences encoding for PKs were identified from an online database, iTAK. After reconfirmation by BLAST similarities, a heatmap of all the PKs having |log2FC>2|, showing 40 (30%) PKs as down-regulated and 70% were up-regulated was generated (Fig. 5B, Additional file 2; Table S-9). Among 132 PKs, 104 (78%) were receptor-like Kinase (RLKs), 6 PKs (4.8%) were calcium- and calmodulin regulated kinase (CAMK) 7 (5.3%) sequences coding cyclin dependent kinase-like kinase (CMGC) and other families belonged to serine/threonine kinase (STE), tyrosine kinase-like kinase (TKL) etc. RLKs are the kinases having P Kinase domain, either membrane localized (RLK) or membrane-associated proteins localized in cytoplasm close to plasma membrane (RLCK). Several Pelle (RLK) kinase proteins are implicated in cell wall metabolism, hormone signaling, and growth-related responses. Similarly, CAMK and CGMC protein kinases are involved in the cell cycle and certain growth responses.

Hormone quantification

To confirm the empirical evidence of FPKM profiling in signaling of IAA, ABA and JA of the hormones, we quantified three hormones from the plant leaf samples (Fig. 6A). IAA and ABA concentration was high (41.32%, 18.82) in high biomass samples. Whereas, JA contents were drastically high in low biomass samples by 155.86%. This shows highly activated defense system in low biomass genotypes, which may hinder the activity of growth responsive genes. Further, linear regression between hormone concentration and hormone signaling genes was performed, which showed auxin concentration has a positive linear relation with AUX1, ARF and XTH, whereas negative relationship was observed in JA (Fig. 6B).

RNA-seq data validation by qRT-PCR

To validate expression patterns of genes identified by RNA-seq data, 15 genes were selected and examined by qRT-PCR (Fig. 7). One gene related to ABC transporter (Soffic.10G0011210-1A), four genes from auxin signal transduction pathway (Soffic.10G0011210-1A, Soffic.03G0032240-1A, Soffic.10G0012000-1A, Soffic.01G0033270-1A), two JA (TIFY), 1 ABA (EIN) confirmed differences in transcript abundances in high and low biomass genotypes. Similarly, cell wall synthesis and remodeling genes (CSLA2, XTH) were highly expressed in high biomass genotypes. Moreover, flowering and senescence related genes i.e., ELF3, FPF1 and S-40 (Soffic.03G0016710-4D, Soffic.03G0036590-2B, Soffic.03G0024310-3H) showed consistent expression patterns in RNA-seq and qPCR quantification.

Discussion

Sugarcane holds a central position as 1st and 2nd generation feedstock crops in a bio-based energy generation system. Therefore, sugars, as well as biomass accumulating traits/factors, are of equal importance in bio-energy-based industry. Previously, sugarcane genomics research was hampered owning to polyploid and heterozygous nature, but recently published information rich allele-specific sugarcane genome opened a range of possibilities for advances in molecular and genetic studies [32]. In this study, we analyzed our transcriptome data of F2 population (extreme segregants of biomass) using

allele-specific *Saccharum officinarum* genome to dissect allele-specific genetic controls of transgressive segregation of extreme biomass segregants. Several differentially expressed GMs, alleles, and duplicated genes (tandem and paralogs) related to signal transduction, carbon metabolism, and protein synthesis were linked to modulations in biomass accumulations. Regarding hormone signaling pathways, we identified that pivotal regulatory members of signaling cascade i.e., repressors of auxin, jasmonic acid and abscisic acid perform a principal role in hormone signaling dynamics in high biomass genotypes (Fig. 4A). Moreover, results depicted the influence of auxin regulation on certain structural genes and predicted the possible role of protein kinases and transcription factors in regulating growth and biomass accumulation processes (Fig. 5).

Accumulation of plant biomass with gradual growth and maturity is a genetically regulated process involving complex crosstalk of hormones. Based upon KEGG enrichment analysis, several GMs, alleles and duplicated genes belonging to auxin families were overrepresented in high biomass group, indicating auxin mediated growth for accumulation of biomass. In the presence of auxin, the transcript levels of three gene families: AUXIN/IAA (Aux/IAA), SMALL AUXIN-UP RNAs (SAURs) and GH3 (Gretchen Hegen) were quickly and transiently induced [33]. Indeed 26-S mediated ubiquitination process led to breakdown of Aux/IAA, which is evident by down-regulation of two Aux/IAA (Soffic.10G0012000-1A (GM), Soffic.10G0013930-3C (allele)) members. Upregulation of SAURs is typically associated with accelerated cell elongation, which is evident in juvenile and in different stages of floral organ development [34–36]. On the other hand, SAURs also inhibit the activity of PP2C which ceases the phosphorylation of H^+-ATPases leading to decrease in apoplastic pH, thereby activating expansins, xyloglucans, pectinesterase and other cell wall-modifying proteins [37](Fig. 4 A, B). The activation of H^+-ATPase activity also favors the hyperpolarization of plasma membrane, thus boosting solutes and solvents uptake, turgor pressure leading to cell expansion [37]. GH3 family of auxin encodes acyl - amido synthetases which regulate the homeostasis of auxin by uniting auxin with free amino acids or mediating catalysis of conjugation [38, 39]. Down regulation of PP2C and upregulation of SAUR support the hypothesis that SAURs repress the transcription of PP2C by creating the acidic PH of apoplast, which favors the cell elongation and expansion processes (Fig. 7) [40].

Regarding JA, down regulation of JA receptor protein COI1 and up-regulation of repressor genes (JAZ) indicated the arrested transcription of jasmonic acid responsive genes. JAZ repressors containing JASMONATE ZIM domain (JAZ) binds to specific transcription factors of JA by its conserved C-terminal domain also recruiting TOPLESS corepressors or the NOVEL INTERACTOR OF JAZ (NINJA) [41]. JA is a critically important hormone in defense related responses by synthesizing low molecular weight compounds, such as polyphenols, alkaloids, quinones, terpenoids, and polypeptides [42]. Conversely, down regulation of JAZ repressors in low biomass indicates the active stress and immune response. These results are complemented in another study that JA (150–250 µM MeJ) treated as elicitor causes the complete inhibition of biomass in medicinal plant *Withania somnifera* (L.) reported by [43], in *H. hirsutum* and *H. maculatum* [44] and *Centella asiatica* [45]. JA is involved in inhibition of leaf expansion by repressing the activity of mitotic cyclin CycB1; 2 and cell division. JA

signaling cascade i.e. COI1-JAZ-MYC2 in *Arabidopsis* hampers the expansion of leaves [46]. GH3 (Gretchen Hagen3) genes act to deactivate JA by conjugation with amino acids like Asp, Met and Trp, promoting formation of adventitious roots [47]. Both jasmonate and auxin signaling processes depend on small protein ubiquitin and 26S proteasome complex (Fig. 3 A, B).

Cell wall loosening and the addition of new monomer units are essential for growth, which is predominantly activated by auxin hormone [48]. It is achieved though activation of certain genes i.e., EXP, XTH and ENDO-(1,4) - β -D- GLUCANASEs (CELLULASEs). The upregulation of sequences related to EXP (expensins) and CESA (cellulose synthase) indicates enhanced metabolism of cell wall. Moreover, XTHs control expansion of cell walls by hydrolysis and re-ligating xyloglucan polymers which are cross-linked with cellulose microfibers [49, 50]. Significantly up-regulated expression in XTH (XET, XEH) sequences might indicate increased beta-xylosidase activity, which is rate limiting in xylan hydrolysis and involved in the thickening of the secondary cell wall [51](Additional file 1; Fig. S-5).

The activity of growth responsive and cell wall related genes accompanied by delayed reproductive maturity lead to increased biomass. Down-regulation of flowering related genes i.e., FLOWERING PROMOTING FACTOR 1 (FPF1), EARLY FLOWERING 3 (ELF3), and AGAMOUS-LIKE MADS-BOX PROTEIN (AGL12, SVP) support late maturing, higher biomass and sugars in HB genotypes. Leaf senescence is an important terminal stage of plant development after reproductive phase, senescence related NAC transcription factors and a newly identified S-40 gene family also showed down-regulation [52, 53]. The expression levels of flowering and senescence related genes support the notion that HB genotypes undergo extended vegetative phase and delayed reproductive maturity. Shift of gene expression patterns i.e. from vegetative to reproductive, showed onset of early reproduction maturity in low biomass samples [54].

Hormone pathways and cell signaling processes also recruit multiple TFs, protein kinases and phosphatases, for the regulation of growth and development. Among many TFs differentially regulated in our study, ARF, bHLH, MYB, WOX, and WRKY are closely related to hormone signaling. It has been clearly demonstrated that ARF mediated transcription of target genes is important for auxin signal transduction. ARF proteins can either be transcriptional activators or repressors of downstream genes, which is determined by structure of MR domain of corresponding genes [55]. bHLH family of TFs interact with MYC TFs in jasmonic acid signaling, and up-regulation of bHLH indicate the repressed activity of JA signal transduction and responsive genes in high biomass group. A study provided with evidence that gain in function mutant ABA-INDUCIBLE BHLH-TYPE TRANSCRIPTION FACTOR/JA-ASSOCIATED MYC2-LIKE1 (JAM1) showed down-regulation of JA responses, reduced root growth and anthocyanin accumulation. MYB proteins are present in all eukaryotes and play an important role in primary and secondary metabolism, cell fate determination and developmental processes [56].

Phytohormones metabolism also involves the activation of certain protein kinases which are critical for in accumulation of sucrose and culm development processes [57]. Out of 132 protein kinase related genes, RLK was highly regulated in high biomass, whereas CAMK and CGMC related kinases were also up

regulated in high biomass group. Among CAMK, calcium-dependent protein kinases are very important in sugars metabolism and hormone signaling [58]. Sucrose synthase is an enzyme that undergoes the conversion of sucrose from glucose and fructose, whereas CDPKs phosphorylate sucrose synthase and regulate their biological activity [59]. CBL-interacting protein kinases (CIPKs) are also the calcium sensors, owing to their structural similarity these proteins are regarded as a subgroup of SNF-like kinases i.e., SnRK2 and SnRK3 [60]. Most of the members of SnRK2 and SnRK3 in our RNA-Seq data were highly expressed in high biomass samples. Although ABA is known as a growth inhibitor [61, 62], paradoxically, ABA in non-stressed conditions acts as a growth promoter [63, 64]. Calcineurin B-like (CBL) kinases also bind to calcium and mediate certain growth and stress related responses by interacting with ABA [65]. The process of ABA stimulated growth is positively influenced by kinases and negatively induced by phosphatases (PP2C) [193]. ABA signaling also recruits calmodulin sensor protein for sensing of Ca^{+} for the regulated signaling and ABA-mediated enhance growth. Up-regulation of multiple MAPK kinases belonging to CGMC category could be involved in sucrose metabolism as studies suggest that *ScMAPK-4* might be an important player in optimizing source-sink relations in sugarcane [66]. Hormone metabolism is regulated by multiple kinases, transcriptional factors and regulators (Fig. 6A, Fig. 5A, B), and crosstalk between the hormones regulates growth and development related responses. Moreover, the expression profiles in Allo or autopolyploid hybrids are the mosaic of the gene variants, and these results can be used to achieve preferential phenotype using bioengineering techniques for candidate gene variants (Fig. 7).

Conclusion

This study provided a broad view of many coding GMs, alleles and duplicated genes expressed at harvest maturity in extreme biomass segregants of *Saccharum* species. Several genes (GMs), alleles, tandemly duplicated genes, and paralogs from different families of auxin, jasmonic acid and abscisic acid were differentially expressed in contrasting biomass groups. Regulatory expression of the repressor genes of these hormones (Aux, JA, and ABA) control the signaling cascade of downstream responsive gene loci. Auxin was identified as the core hormone in modulation of growth by upregulation of genes downstream of cascade. The remarkable ontogeny of high biomass genotypes might be shaped by co-ordination of internal balance of phyto hormones i.e., activated auxin and inactivated jasmonic acid and ABA pathways. Moreover, our results highlighted a plethora of growth regulatory genes, protein kinases and transcription factors in triggering modulated growth in high biomass. Further experimentation based on the function of these genes will elucidate the basis of heterosis in contrasting genotypes in sugarcane.

Material And Methods

Background of RNA-Seq reads

Saccharum officinarum, an intensively cultivated sugarcane species, was crossed with *Saccharum robustum* species, indigenous to New Guinea. Two cultivars ‘LA Purple’ (2n = 8× = 80) and ‘MOL5829’ (2n = 8× = 80) were selected as representatives of the two species, respectively. Among 98 F1 crosses, 20

plants of F1 progeny were selected and self-fertilized after the measurements of stem weight, diameter and sugar contents. Later on, 272 F2 progenies were field-tested at different stations and 120 plants with extreme biomass were chosen for further experimentation. To confirm the performance of these 120 plants, field trials were conducted at Hawaii Agriculture Research Center Kunia station, Oahu, Hawaii. Finally, 22 plants with extreme high and low biomass were screened for digital gene expression by extracting mRNA, followed by transcriptome analysis. The first dewlap leaf, stem internodes from 3rd, 9th, and 15th node (internode at the first dew lap leaf was numbered as first internode) were collected from 8-month old F2 progenies. Plant tissues were grounded to powdered form in liquid nitrogen, and 100 mg gram sample was used for RNA extraction using Omega Bio-tek E.Z.N.A. ® Plant RNA kit (Omega Bio-tek, #R6827–02). Finally, samples from internodes and leaves were pooled from individual plants and 1ng of RNA was used for Illumina RNA sequencing after library construction. All plants materials used for RNA-seq studies were obtained from the Hawaii Agriculture Research Center Kunia substation (Hawaii, USA), and Texas A&M AgriLife Research and Extension Center at Weslaco (Texas, USA). The detailed methodology of the field experiment and physiological attributes are described in the article [23] DOI 10.1186/s12864-017-4158-8.

Sequence read alignment and analysis of differentially expressed genes

Paired-end raw RNA-Seq reads sequenced by HiSeq2500 were used for analysis, available at NCBI with BioProject ID PRJNA347369 and Short Read Archive under SRR5223340 - SRR5223361. FastQC was used for the initial visualization to access the initial quality of reads i.e., GC content, Per Base Sequence Quality, adapter Sequence (http://www.bioinformatics.babraham. ac.u k/projects/fastqc). For removing the adaptor sequences, low quality reads and to trim the long reads, we used trimmomatic [67]. Trimmed reads were aligned to *S. officinarum* genome LA Purple' (2n = 8× = 80) using STAR [68] (https://github.com/alexdobin/STAR) and HISAT2 (http://ccb.jhu.edu/software/hisat2 or http://github.com/ infphilo/hisat2) [69]. Following, HISAT2, stringtie and DeSeq2 pipeline was used for analysis, whereas additionally STAR was used as an aligner for indexing and alignment of RNA-Seq paired-end reads. STAR showed higher alignment rates and we used StringTie to assemble STAR aligned transcripts to reference genome followed by quantification of transcripts (http://ccb.jhu.edu/software/stringtie or https://github. com/gpertea/stringtie.) [70]. Whereas differential expression analysis (DEGs) was performed using DESeq2, a bio conductor package in R software. Differentially expressed genes were sorted out using the criteria of FDR value| $\leq$0.05|, p-value |<0.01| and |log2FC >2, < -2| for up and down-regulated DEGs. Partial least squares-discriminant analysis (PLS-DA), hierarchical clustering, and volcano plot were processed from FPKM and fold change and visualized using R- software.

Functional annotation and GO enrichment analysis

We used two methods to perform gene ontology (GO) enrichment analysis i.e., (1) an online database named as omicsahre (https://www.omicshare.com) and (2) windows-based software, TBtools [71] (https://github.com/CJ-Chen/TBtools). GO terms background file was generated from *Saccharum*

officinarum proteome. GO terms that showed adjusted p-value less than 5% were considered as enriched terms and visualized in REVIGO [72] (http://bioinfo.cau.edu.cn/agriGO/). The number of DEGs in different categories (cellular components, molecular functions, biological processes) was consistent in both omicshare and TBtools. The results of omicshare were used to focus the target genes involved in hormone signaling and biomass accumulation.

KEGG analysis, Transcription factor and kinome

Dynamic KEGG enrichment analysis (Kyoto Encyclopedia of Genes and Genomes) of genes was performed by an online website “omicsahre” (https://www.omicshare.com) to identify functional annotations and signaling pathways. From the results of KEGG, hormone, carbohydrate metabolism and other genes related to developmental process were selected after confirmation by blastn in Phytozome (https://phytozome-next.jgi.doe.gov/). Furthermore, transcription factors (TF) and protein kinases (PKs) within the genes were identified using ITAK, an online database (http://itak.feilab.net/cgi-bin/itak/index.cgi) and visualized by TBtools [71, 73].

Weighted gene co-expression analysis (WGCNA) analysis

To find co-expression patterns of differentially expressed genes, based on FPKM, WGCNA was performed in R. Ward's clustering method [74] was used for hierarchical clustering of contrasting biomass samples. Correlation and module detection among interacting DEGs were done via unsigned topological overlap matrix (TOM). We created adjacency matrices using adjacency function and soft power threshold as eight. To detect clusters of dendrogram, we used the Dynamic Tree Cut library in R, with a merge cut height as, 0.25 and 0.3. From the module-trait relationship heat map, highly correlated significant ($r > 0.83$, 0.91) modules "blue" and "greenyellow" were used to construct the networks in cytoscape (v. 3.9.1) [75].

Hormones quantification

In total, 5 samples from each group were selected from high and low biomass samples. The top most sugarcane leaf was selected and its middle region was used for quantification of IAA, ABA and JA previously stated by [76]. Briefly, 50 mg of the freshly powdered samples were extracted with extraction solvent (2-propanol: H_2O: HCl = 2: 1:0.002) in a 2 mL falcon tube. Simultaneously, 50 µL of the working solution of internal standards was added and tubes were shaken at 100 RPM for 30 min at 4 °C. After this, addition of 1 mL dichloromethane and another shaking was done for 30 min at 4 °C. The solutions were kept in micro-centrifuge at 4 °C and centrifugation was performed at 14000 RPM for 5 min. After discarding the supernatant, 1.2 mL of solvent was concentrated using a nitrogen evaporator with nitrogen flow. The samples were dissolved in 0.1 mL methanol (70% v/v) and centrifuged at 14,000 RPM for 5 mins. Finally, 50 µL of sample solution was injected into the reverse-phase for the quantification of acidic hormones by 6410 Triple Quad LCMS.

Quantitative RT-PCR based quantification for gene expression

First of all, RNA was extracted from the leaves of high and low biomass genotypes using Trizol reagent [77] and cDNA (first strand) was synthesized from 1 µg of RNA by using PrimeScript™ RT Reagent Kit with gDNA Eraser. The reaction volume of cDNA was diluted to 80 µl and quantitative RT-PCR was carried out using the TB Green™ Premix Ex Taq™ II kit (TaKaRa), in a CFX- 96-well Real-Time System (BioRad, USA). The reaction volume was 20 µl in total; 1 µl of cDNA, 1 µM of F and R primers and10µl of TB Green™ PCR master mix. The temperature and duration of thermal cycles during PCR were as follows; Amplification process included thermal cycles i.e., 95 °C for 3 mins, 40 cycles at 95 °C for 10 s and 50 °C for 30 s followed by disassociation stage according to instructions by the user's manual. There were three biological replicates and normalization was done by reference gene i.e., GAPDH [78]. The normalized expression level was calculated as L = 2-ΔCt and ΔCT = CT, target - CT, reference. The primers used in experiment were listed in (Additional file 1; Table S-2).

Declarations

Availability of data and materials

The transcriptome reads sequenced in this study are available at NCBI with BioProject ID PRJNA347369 and Short Read Archive under SRR5223340 - SRR5223361.

Ethics approval and consent to participate

The sugarcane species used in this study were obtained from Hawaii Agriculture Research Center Kunia substation (Hawaii, USA), and Texas A&M AgriLife Research and Extension Center at Weslaco (Texas, USA). The research on plants, including the collection of plant material complies with relevant institutional, national, and international guidelines and legislation. We confirm that all experimental protocols were approved by FAFU and UIUC-SIB Joint Center for Genomics and Biotechnology, Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, China.

Consent for publication

Not applicable.

Competing interests

The authors declare no conflicts of interest.

Funding

This work was supported by grants from National Natural Science Foundation of China to Jishan Lin (No. 2020J01593)

Author's Contributions

RM conceived the project; Ain N-u performed the transcriptome sequencing data analysis. Ain N-u and H (Fig. 6, Fig. 7) performed qPCR and hormone quantification analysis. Ain N-u wrote the manuscript, XZ and JL interpreted results, reviewed and revised the manuscript. RM supervised this project (material identification and data collection). All the authors read and approved the final manuscript.

Acknowledgements

This work was supported by startup fund from Fujian Agriculture and Forestry University.

References

1. Bessou C, Ferchaud F, Gabrielle B, Mary B. Biofuels, greenhouse gases and climate change. In: Sustainable Agriculture. Springer Netherlands; 2009. p. 365–468.
2. IEA. Key World Energy Statistics, International Energy Agency. Paris, France; 2021.
3. Work VH, D'Adamo S, Radakovits R, Jinkerson RE, Posewitz MC. Improving photosynthesis and metabolic networks for the competitive production of phototroph-derived biofuels. Curr Opin Biotechnol. 2012;23:290–7.
4. Hofsetz K, Silva MA. Brazilian sugarcane bagasse: Energy and non-energy consumption. Biomass and Bioenergy. 2012;46:564–73.
5. Zhang J, Zhang Q, Li L, Tang H, Zhang Q, Chen Y, et al. Recent polyploidization events in three *Saccharum* founding species. Plant Biotechnol J. 2019;17:264–74.
6. Correr FH, Hosaka GK, Barreto FZ, Valadaõ IB, Balsalobre TWA, Furtado A, et al. Differential expression in leaves of *Saccharum* genotypes contrasting in biomass production provides evidence of genes involved in carbon partitioning. BMC Genomics. 2020;21:1–16.
7. Qi Y, Gao X, Zeng Q, Zheng Z, Wu C, Yang R, et al. Sugarcane Breeding, Germplasm Development and Related Molecular Research in China. Sugar Tech. 2022;24:73–85.
8. Carvalho-Netto O V., Bressiani JA, Soriano HL, Fiori CS, Santos JM, Barbosa GV, et al. The potential of the energy cane as the main biomass crop for the cellulosic industry. Chemical and Biological Technologies in Agriculture. 2014;1:1–8.
9. Depuydt S, Hardtke CS. Hormone Signalling Crosstalk in Plant Growth Regulation. Curr Biol. 2011;21:R365–73.
10. Jez JM. Connecting primary and specialized metabolism: Amino acid conjugation of phytohormones by GRETCHEN HAGEN 3 (GH3) acyl acid amido synthetases. Curr Opin Plant Biol. 2022;66:102194.
11. Zhu F, Wai CM, Zhang J, Jones TC, Nagai C, Ming R. Differential expression of hormone related genes between extreme segregants of a *Saccharum* interspecific F2 population. Euphytica. 2018;214:1–12.
12. Fukaki H, Nakao Y, Okushima Y, Theologis A, Tasaka M. Tissue-specific expression of stabilized SOLITARY-ROOT/IAA14 alters lateral root development in Arabidopsis. Plant J. 2005;44:382–95.

13. Rouse D, Mackay P, Stirnberg P, Estelle M, Leyser O. Changes in auxin response from mutations in an AUX/IAA gene. Science (80-). 1998;279:1371–3.
14. Tatematsu K, Kumagai S, Muto H, Sato A, Watahiki MK, Harper RM, et al. Massugu2 Encodes Aux/IAA19, an Auxin-Regulated Protein That Functions Together with the Transcriptional Activator NPH4/ARF7 to Regulate Differential Growth Responses of Hypocotyl and Formation of Lateral Roots in Arabidopsis thaliana. Plant Cell. 2004;16:379–93.
15. Rogg LE, Lasswell J, Bartel B. A Gain-of-Function Mutation in IAA28 Suppresses Lateral Root Development. 2001.
16. Mandaokar A, Thines B, Shin B, Markus Lange B, Choi G, Koo YJ, et al. Transcriptional regulators of stamen development in Arabidopsis identified by transcriptional profiling. Plant J. 2006;46:984–1008.
17. Chini A, Fonseca S, Fernández G, Adie B, Chico JM, Lorenzo O, et al. The JAZ family of repressors is the missing link in jasmonate signalling. Nature. 2007;448:666–71.
18. Sheard LB, Tan X, Mao H, Withers J, Ben-Nissan G, Hinds TR, et al. Jasmonate perception by inositol-phosphate-potentiated COI1–JAZ co-receptor. Nature. 2010;468:400–5.
19. Thines B, Katsir L, Melotto M, Niu Y, Mandaokar A, Liu G, et al. JAZ repressor proteins are targets of the SCFCOI1 complex during jasmonate signalling. Nat . 2007;448:661–5.
20. Weiner JJ, Peterson FC, Volkman BF, Cutler SR. Structural and functional insights into core ABA signaling. Curr Opin Plant Biol. 2010;13:495–502.
21. Park SY, Fung P, Nishimura N, Jensen DR, Fujii H, Zhao Y, et al. Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. Science (80-). 2009;324:1068–71.
22. Ma Y, Szostkiewicz I, Korte A, Moes D, Yang Y, Christmann A, et al. Regulators of PP2C phosphatase activity function as abscisic acid sensors. Science (80-). 2009;324:1064–8.
23. Wai CM, Zhang J, Jones TC, Nagai C, Ming R. Cell wall metabolism and hexose allocation contribute to biomass accumulation in high yielding extreme segregants of a *Saccharum* interspecific F2 population. BMC Genomics. 2017;18:773.
24. Sforça DA, Vautrin S, Cardoso-Silva CB, Mancini MC, Romero-da Cruz MV, Pereira G da S, et al. Gene Duplication in the Sugarcane Genome: A Case Study of Allele Interactions and Evolutionary Patterns in Two Genic Regions. Front Plant Sci. 2019;10:553.
25. Ganko EW, Meyers BC, Vision TJ. Divergence in Expression between Duplicated Genes in Arabidopsis. Mol Biol Evol. 2007;24:2298–309.
26. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, et al. MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40:e49.
27. Wang S, Bai Y, Shen C, Wu Y, Zhang S, Jiang D, et al. Auxin-related gene families in abiotic stress response in Sorghum bicolor. Funct Integr Genomics. 2010;10:533–46.
28. Singh KB, Foley RC, Oñate-Sánchez L. Transcription factors in plant defense and stress responses. Current Opinion in Plant Biology. 2002;5:430–6.

29. Qiao Z, Li CL, Zhang W. WRKY1 regulates stomatal movement in drought-stressed Arabidopsis thaliana. Plant Mol Biol. 2016;91:53–65.
30. Schlereth A, Möller B, Liu W, Kientz M, Flipse J, Rademacher EH, et al. MONOPTEROS controls embryonic root initiation by regulating a mobile transcription factor. Nature. 2010;464:913–6.
31. Aono AH, Pimenta RJG, Garcia ALB, Correr FH, Hosaka GK, Carrasco MM, et al. The Wild Sugarcane and Sorghum Kinomes: Insights Into Expansion, Diversification, and Expression Patterns. Front Plant Sci. 2021;12 July.
32. Zhang J, Zhang X, Tang H, Zhang Q, Hua X, Ma X, et al. Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. Nat Genet . 2018;50:1565–73.
33. Woodward AW, Bartel B. Auxin: Regulation, Action, and Interaction. Ann Bot. 2005;95:707–35.
34. Chae K, Isaacs CG, Reeves PH, Maloney GS, Muday GK, Nagpal P, et al. Arabidopsis SMALL AUXIN UP RNA63 promotes hypocotyl and stamen filament elongation. Plant J. 2012;71:684–97.
35. Spartz AK, Lee SH, Wenger JP, Gonzalez N, Itoh H, Inzé D, et al. The SAUR19 subfamily of SMALL AUXIN UP RNA genes promote cell expansion. Plant J. 2012;70:978–90.
36. van Mourik H, van Dijk ADJ, Stortenbeker N, Angenent GC, Bemer M. Divergent regulation of Arabidopsis SAUR genes: A focus on the SAUR10-clade. BMC Plant Biol. 2017;17:1–14.
37. Spartz AK, Ren H, Park MY, Grandt KN, Lee SH, Murphy AS, et al. SAUR inhibition of PP2C-D phosphatases activates plasma membrane H+-ATPases to promote cell expansion in Arabidopsis. Plant Cell. 2014;26:2129–42.
38. Staswick PE, Serban B, Rowe M, Tiryaki I, Maldonado MT, Maldonado MC, et al. Characterization of an arabidopsis enzyme family that conjugates amino acids to indole-3-acetic acid. Plant Cell. 2005;17:616–27.
39. Feng S, Yue R, Tao S, Yang Y, Zhang L, Xu M, et al. Genome-wide identification, expression analysis of auxin-responsive *GH3* family genes in maize (*Zea mays* L.) under abiotic stresses. J Integr Plant Biol. 2015;57:783–95.
40. Miedes E, Suslov D, Vandenbussche F, Kenobi K, Ivakov A, Van Der Straeten D, et al. Xyloglucan endotransglucosylase/hydrolase (XTH) overexpression affects growth and cell wall mechanics in etiolated Arabidopsis hypocotyls. J Exp Bot. 2013;64:2481–97.
41. Hickman R, Van Verk MC, Van Dijken AJH, Mendes MP, Vroegop-Vos IA, Caarls L, et al. Architecture and dynamics of the jasmonic acid gene regulatory network. Plant Cell. 2017;29:2086–105.
42. Carvalhais LC, Schenk PM, Dennis PG. Jasmonic acid signalling and the plant holobiont. Current Opinion in Microbiology. 2017;37:42–7.
43. Sivanandhan G, Arun M, Mayavan S, Rajesh M, Mariashibu TS, Manickavasagam M, et al. Chitosan enhances withanolides production in adventitious root cultures of Withania somnifera (L.) Dunal. Ind Crops Prod. 2012;37:124–9.
44. Coste A, Vlase L, Halmagyi A, Deliu C, Coldea G. Effects of plant growth regulators and elicitors on production of secondary metabolites in shoot cultures of Hypericum hirsutum and Hypericum

maculatum. Plant Cell Tissue Organ Cult. 2011;106:279–88.

45. Kim OT, Kim MY, Hong MH, Ahn JC, Hwang B. Stimulation of asiaticoside accumulation in the whole plant cultures of Centella asiatica (L.) urban by elicitors. Plant Cell Rep. 2004;23:339–44.
46. Zhang Y, Turner JG. Wound-Induced Endogenous Jasmonates Stunt Plant Growth by Inhibiting Mitosis. PLoS One. 2008;3:e3699.
47. Gutierrez L, Mongelard G, Floková K, Păcurar DI, Novák O, Staswick P, et al. Auxin controls Arabidopsis adventitious root initiation by regulating jasmonic acid homeostasis. Plant Cell. 2012;24:2515–27.
48. Cosgrove DJ. Growth of the plant cell wall. Nat Rev Mol Cell Biol 2005 611. 2005;6:850–61.
49. Wolf S, Hématy K, Höfte H. Growth control and cell wall signaling in plants. Annual Review of Plant Biology. 2012;63:381–407.
50. Vissenberg K, Oyama M, Osato Y, Yokoyama R, Verbelen J-P, Nishitani K. Differential Expression of AtXTH17, AtXTH18, AtXTH19 and AtXTH20 Genes in Arabidopsis Roots. Physiological Roles in Specification in Cell Wall Construction. Plant Cell Physiol. 2005;46:192–200.
51. Goujon T, Minic Z, El Amrani A, Lerouxel O, Aletti E, Lapierre C, et al. *AtBXL1* , a novel higher plant (*Arabidopsis thaliana*) putative beta-xylosidase gene, is involved in secondary cell wall metabolism and plant development. Plant J. 2003;33:677–90.
52. Hu Y, Xie Q, Chua NH. The Arabidopsis auxin-inducible gene ARGOS controls lateral organ size. Plant Cell. 2003;15:1951–61.
53. Habiba, Xu J, Gad AG, Luo Y, Fan C, Uddin JBG, et al. Five OsS40 Family Members Are Identified as Senescence-Related Genes in Rice by Reverse Genetics Approach. Front Plant Sci. 2021;12:1865.
54. D'Esposito D, Cappetta E, Andolfo G, Ferriello F, Borgonuovo C, Caruso G, et al. Deciphering the biological processes underlying tomato biomass production and composition. Plant Physiol Biochem. 2019;143 May:50–60.
55. Ulmasov T, Hagen G, Guilfoyle TJ. Activation and repression of transcription by auxin-response factors. Proc Natl Acad Sci U S A. 1999;96:5844–9.
56. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. MYB transcription factors in Arabidopsis. Trends in Plant Science. 2010;15:573–81.
57. De Smet I, Voß U, Jürgens G, Beeckman T. Receptor-like kinases shape the plant. Nat Cell Biol. 2009;11:1166–73.
58. Sheen J. Ca2+-dependent protein kinases and stress signal transduction in plants. Science (80-). 1996;274:1900–2.
59. Hardin SC, Winter H, Huber SC. Phosphorylation of the amino terminus of maize sucrose synthase in relation to membrane association and enzyme activity. Plant Physiol. 2004;134:1427–38.
60. Hrabak EM, Chan CWM, Gribskov M, Harper JF, Choi JH, Halford N, et al. The Arabidopsis CDPK-SnRK superfamily of protein kinases. Plant Physiol. 2003;132:666–80.

61. Trewavas A, Jones H, Davies WJ. Abscisic acid: physiology and biochemistry. An Assess role ABA plant Dev Oxford BIOS. 1991;:169–88.
62. Munns R, Cramer GR. Is coordination of leaf and root growth mediated by abscisic acid? Opinion. Plant Soil. 1996;185:33–49.
63. Cheng W-H, Endo A, Zhou L, Penney J, Chen H-C, Arroyo A, et al. A Unique Short-Chain Dehydrogenase/Reductase in Arabidopsis Glucose Signaling and Abscisic Acid Biosynthesis and Functions. Plant Cell. 2002;14:2723 LP – 2743.
64. Sharp RE, LeNoble ME. ABA, ethylene and the control of shoot and root growth under water stress. J Exp Bot. 2002;53:33–7.
65. Boudsocq M, Laurière C. Osmotic signaling in plants. Multiple pathways mediated by emerging kinase families. Plant Physiology. 2005;138:1185–94.
66. McCormick AJ, Cramer MD, Watt DA. Changes in photosynthetic rates and gene expression of leaves during a source-sink perturbation in sugarcane. Ann Bot. 2008;101:89–102.
67. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.
68. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.
69. Kim D, Langmead B, Salzberg SL. HISAT: A fast spliced aligner with low memory requirements. Nat Methods. 2015;12:357–60.
70. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33:290–5.
71. Chen C, Chen H, He Y, Xia R. TBtools, a Toolkit for Biologists integrating various biological data handling tools with a user-friendly interface. bioRxiv. 2018;:289660.
72. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. PLoS One. 2011;6:e21800.
73. Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, Zhang P, et al. iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. Mol Plant. 2016;12:1667–70.
74. Milligan GW, Sokol LM. A Two-Stage Clustering Algorithm with Robust Recovery Characteristics. Educ Psychol Meas. 1980;40:755–9.
75. Saito R, Smoot ME, Ono K, Ruscheinski J, Wang PL, Lotia S, et al. A travel guide to Cytoscape plugins. Nat Methods. 2012;9:1069–76.
76. Pan X, Welti R, Wang X. Quantitative analysis of major plant hormones in crude plant extracts by high-performance liquid chromatography–mass spectrometry. Nat Protoc 2010 56. 2010;5:986–92.
77. Ma J, He Y, Wu C, Liu H, Hu Z, Sun G. Cloning and Molecular Characterization of a SERK Gene Transcriptionally Induced During Somatic Embryogenesis in Ananas comosus cv. Shenwan. Plant Mol Biol Report. 2012;30:195–203.

78. Ling H, Wu Q, Guo J, Xu L, Que Y. Comprehensive selection of reference genes for gene expression normalization in sugarcane by real time quantitative RT-PCR. PLoS One. 2014;9.

Figures

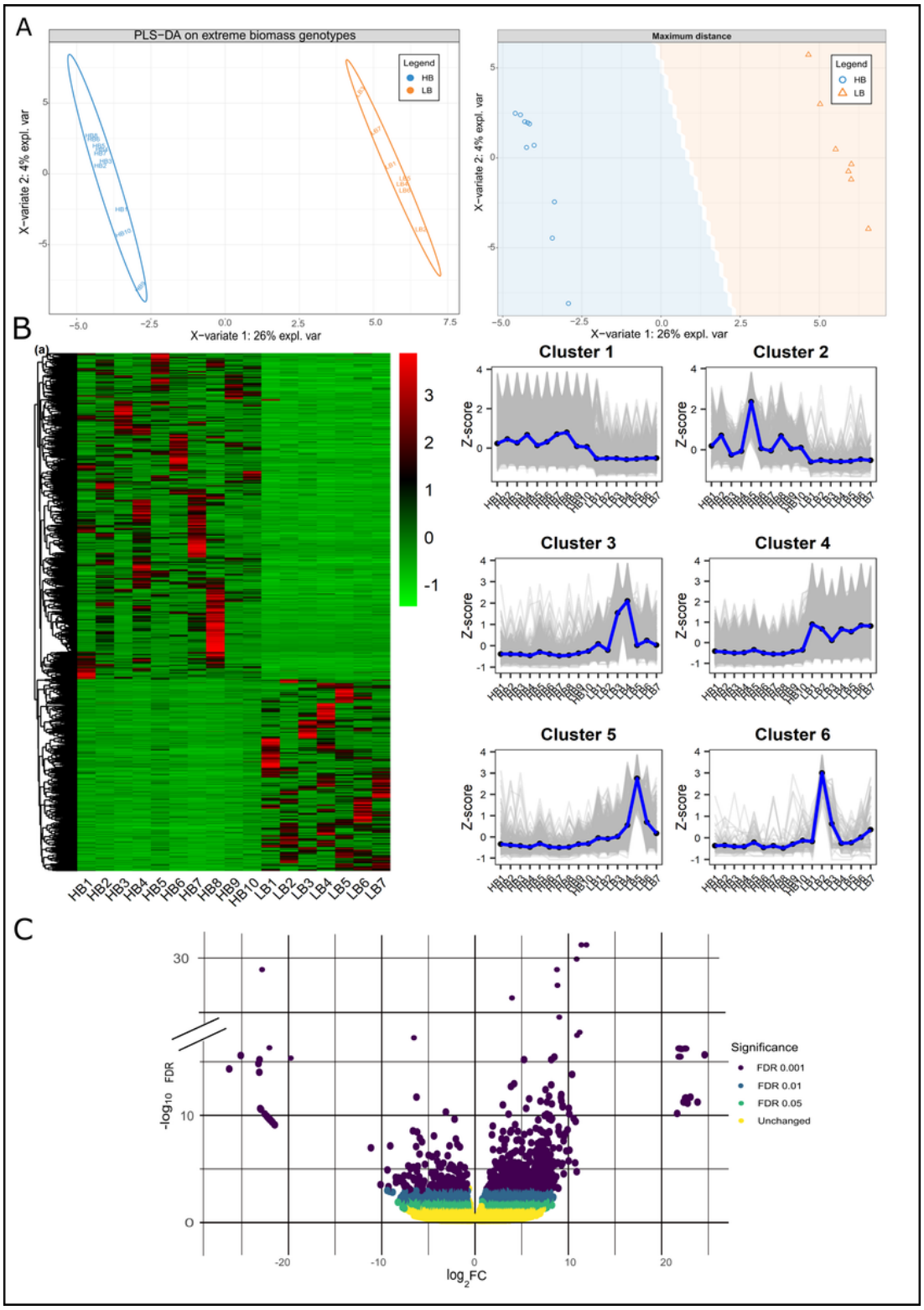

Figure 1

Partial least squares-discriminant analysis (PLS-DA), Hierarchical clustering and volcano plot of DEGs (A) PLS-DA score plot of FPKM data of DEGs generated by using preprocessed original data, shows the clustering of extreme biomass segregants. Component 1 (26%) is clearly distinguishing the two biomass groups. (B) Hierarchical clustering heatmap of the DEGs based on FPKM expression values. (C) Volcano plot overall shows the range of log2FC of DEGs i.e., 24 to -26 from right to left

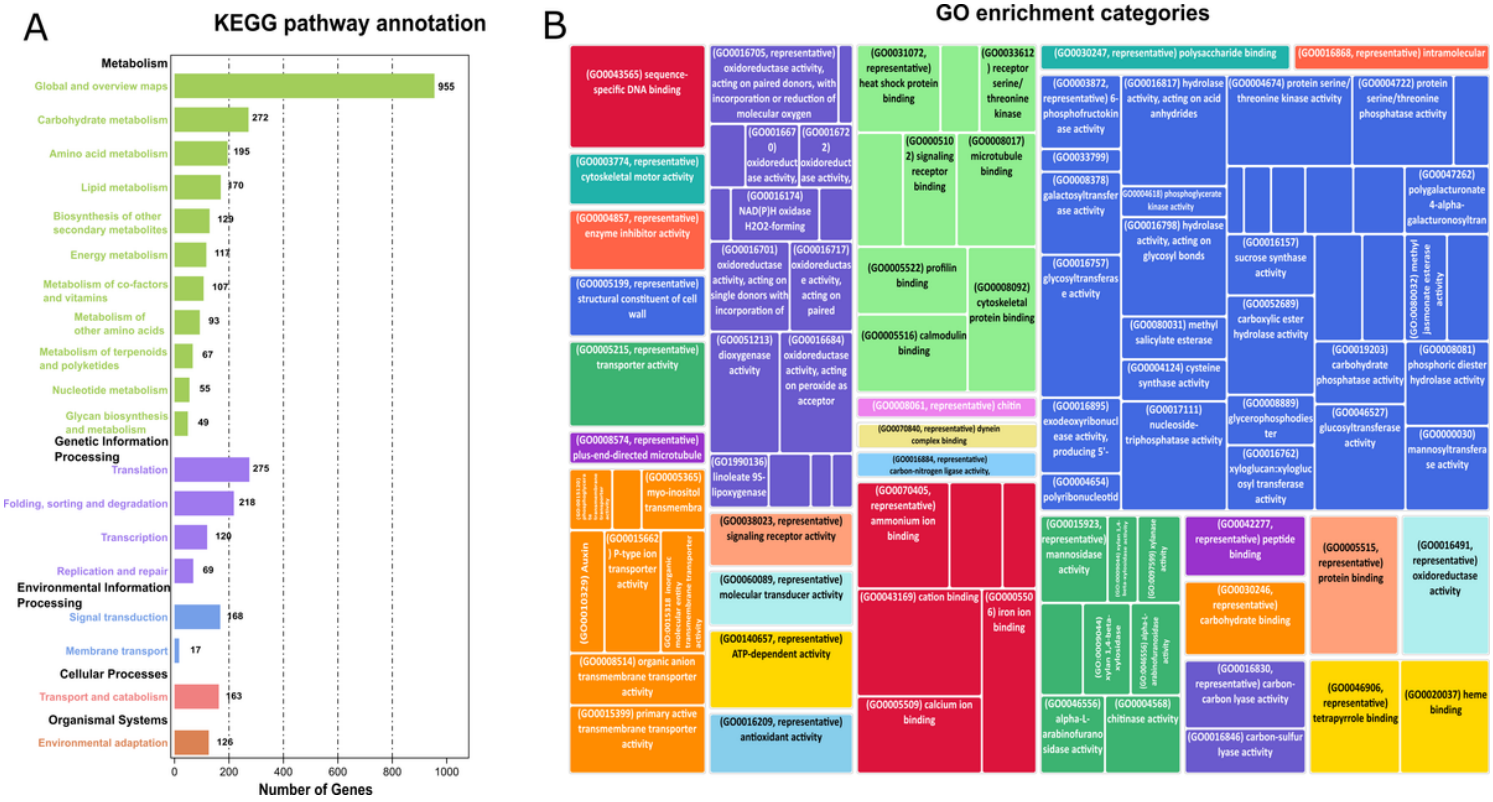

Figure 2

KEGG and GO analysis of DEGs. (A) Fig. 3: Bar plot for the KEGG categories and enriched DEGs in KEGG analysis. The abscissa depicts enriched pathways while gene number is plotted on ordinate axis (B) GO shows the enrichment of potential DEGs in different categories i.e. Biological processes, cellular component and molecular function

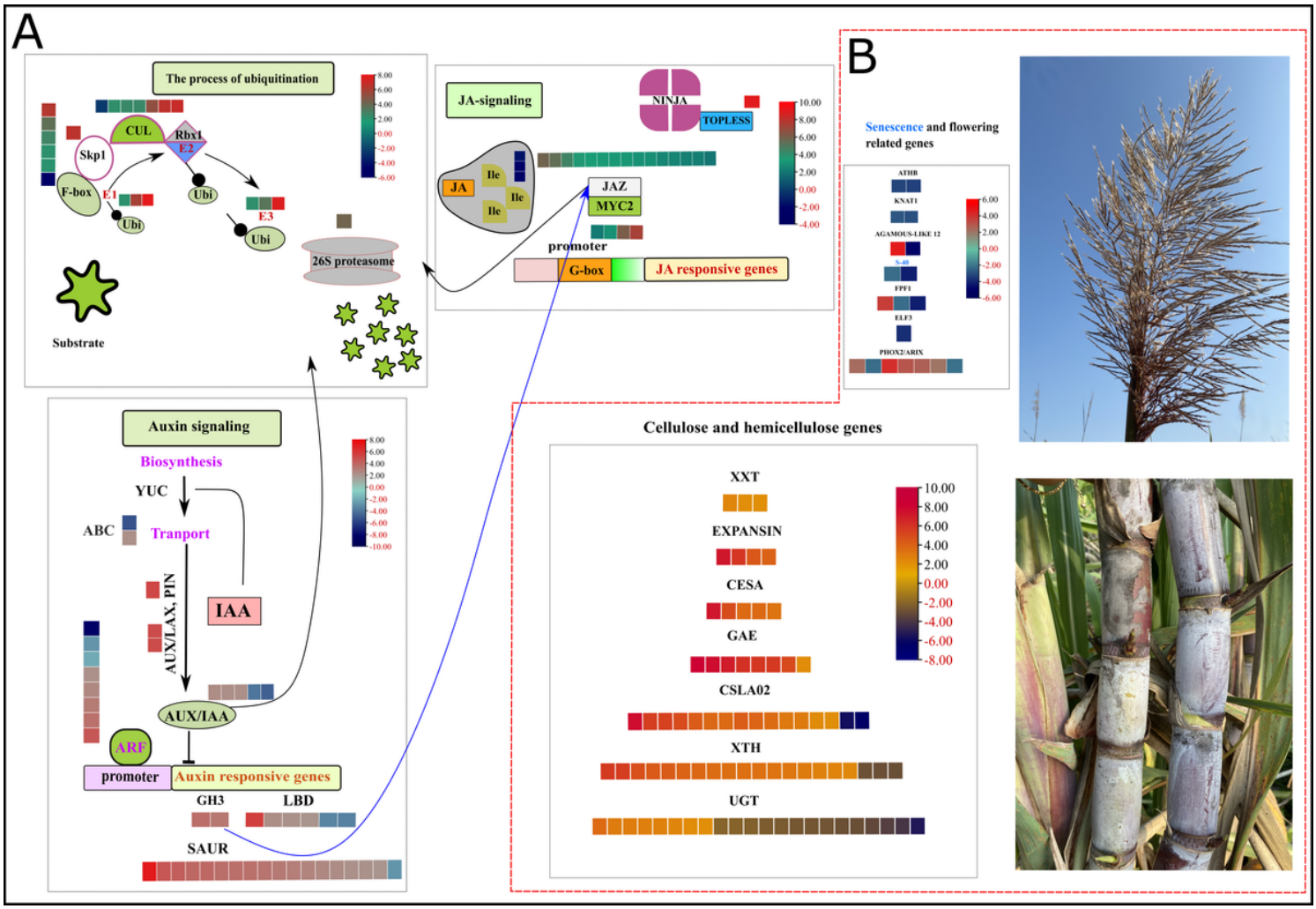

Figure 3

Heatmap of log2FC values in HB segregants, Red and black values in legend show down and up-regulated genes. 3A shows ubiquitin mediated signaling of auxin and jasmonic acid signaling pathway. 3B. depicts the expression patterns of hormone responsive growth-related genes in cell wall and terminal developmental phases i.e., inflorescence and senescence. S-40: Senescence regulator, ATHB: ARABIDOPSIS THALIANA HOMEOBOX 7, KNAT1: BREVIPEDICELLUS, PHOX2/ARIX: Transcription factor PHOX2/ARIX, ELF3: EARLY FLOWERING 3 (ELF3), FPF1: FLOWERING PROMOTING FACTOR 1, AGAMOUS-LIKE 12: AGAMOUS-LIKE 12, XXT: galactosyl transferase GMA12/MNN10 family, EXPANSIN: EXPANSIN, CESA : cellulose synthase subfamily, GAE: GDP-mannose 4,6 dehydratase, CSLA02: Belongs to the glycosyltransferase 2 family, XTH: Xyloglucan endohydrolysis (XEH) and or endotransglycosylation (XET), UGT: UDP-glycosyltransferase family.

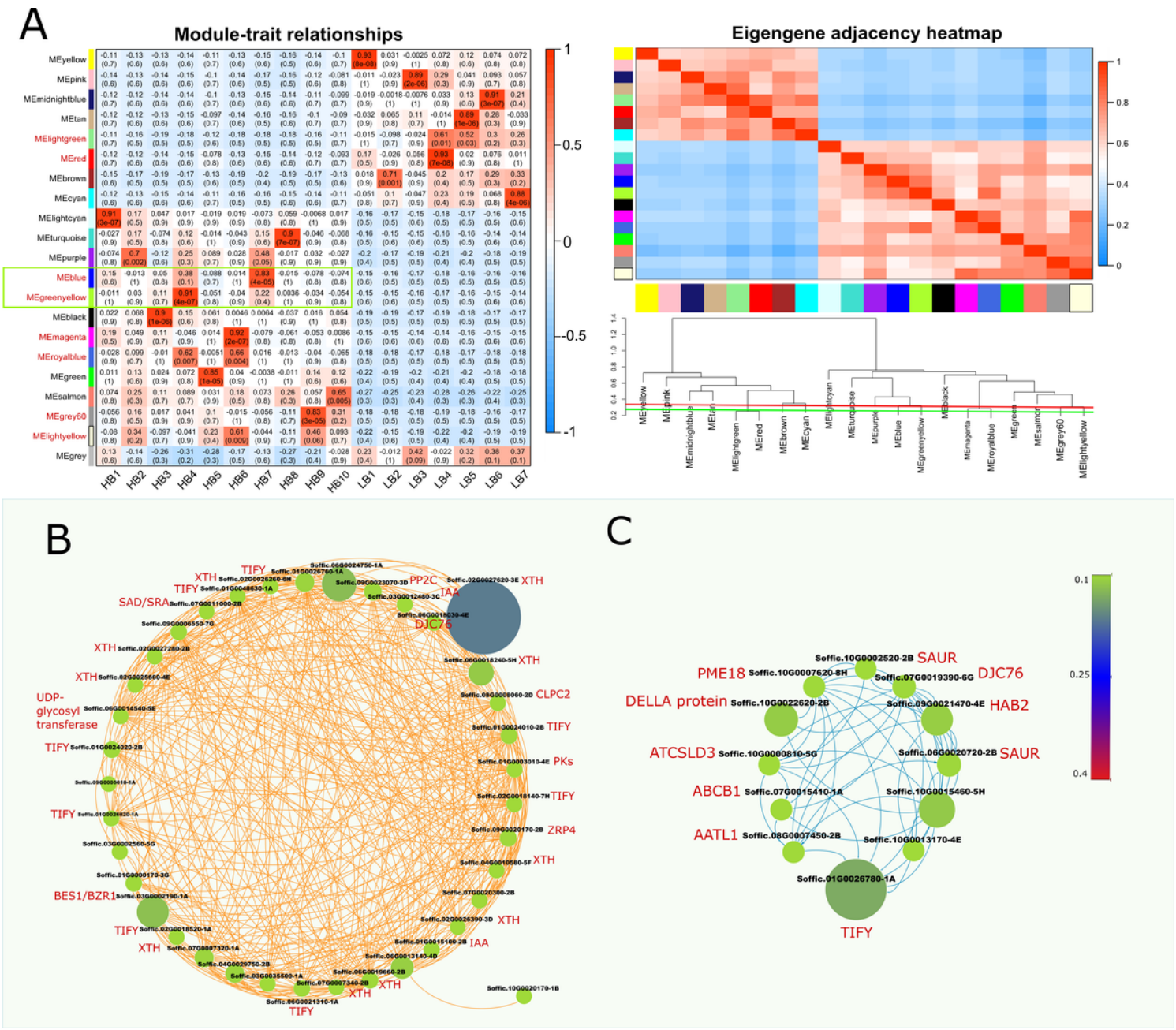

Figure 4

Weighted gene co-expression network analysis (WGCNA) of DEGs between HB and LB segregants at harvest maturity. (A) eigengene modules, Module-trait relationship and module-module relationship. Module-trait heatmap shows the pearson's correlation of all the modules with samples whereas in module-module relationship progressive saturation in blue and red color points high co-expression interconnectedness. Additionally, it shows dendrogram of module clustering, green and red horizontal line represent threshold (0.25, 0.3). (B) Cytoscape network shows co-expression network of blue module which depicts highly upregulated genes in HB segregants (C) Cytoscape network of overrepresented DEGs in module 'greenyellow'. Size and color of nodes is proportional to weights whereas edges color correspond to module names

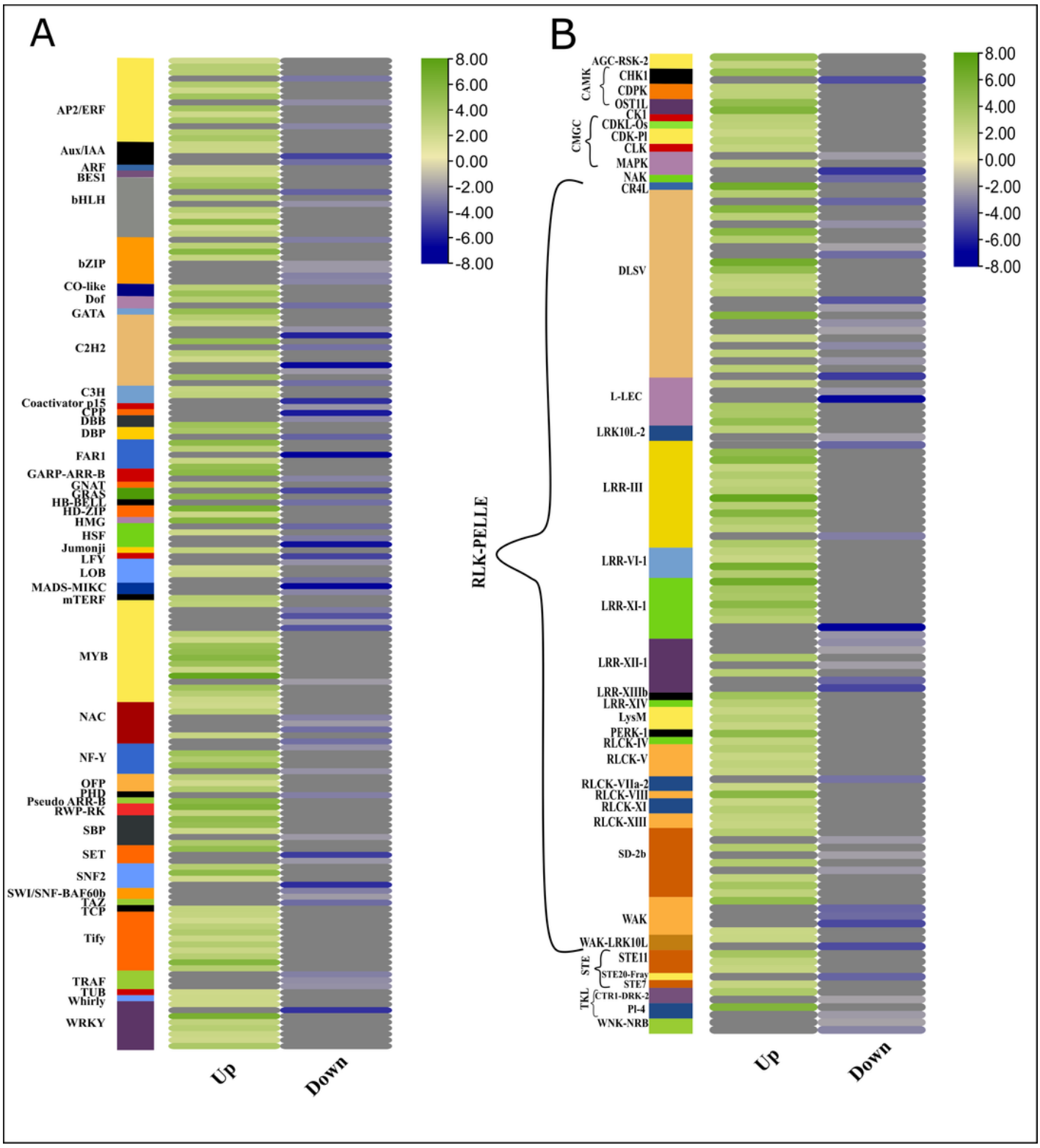

Figure 5

Heatmap generated using log2FC. Expression dynamics of TFs and PKs involved in high biomass samples. Green and blue scales represent up and down regulated TFs in HB samples whereas grey color indicates blanks

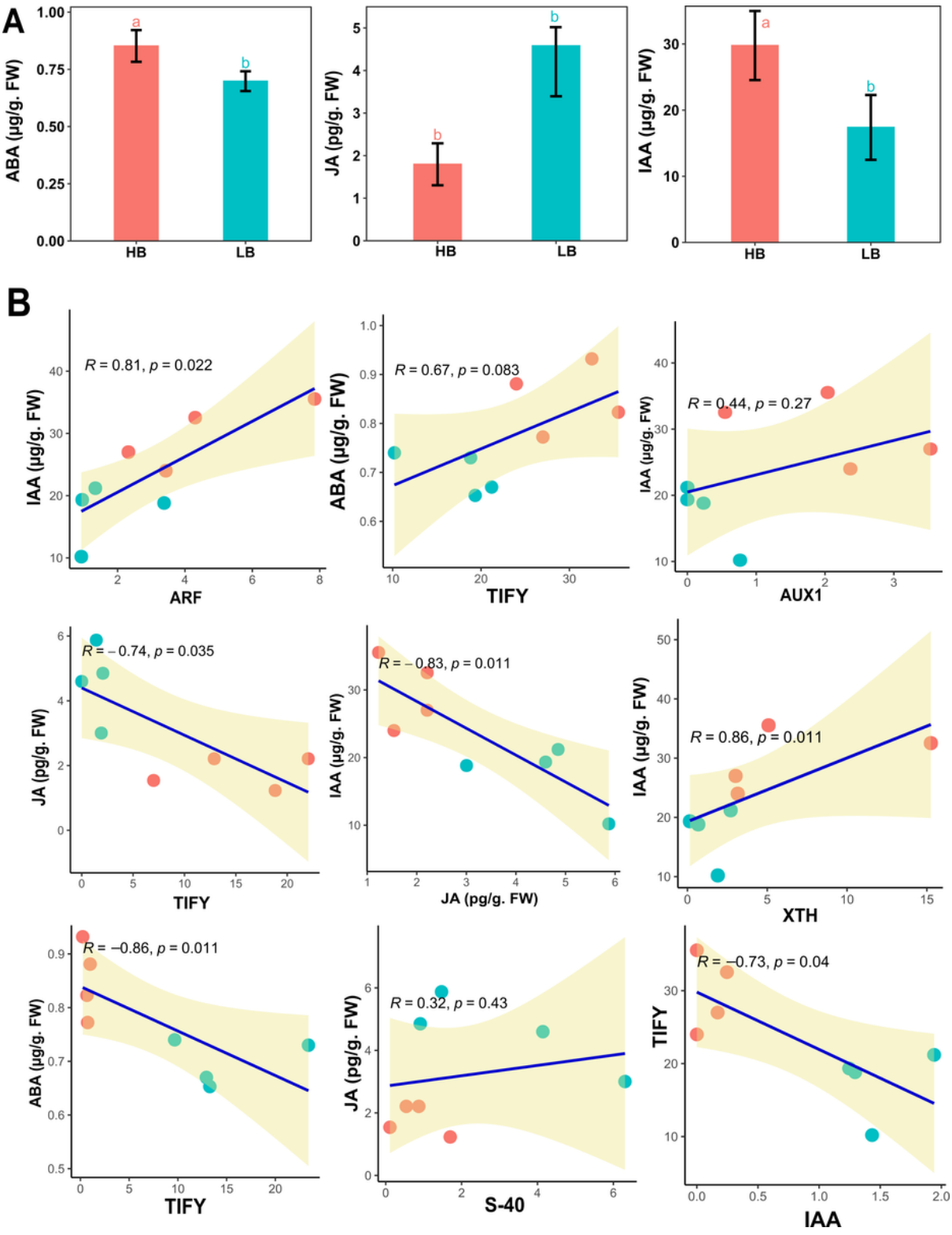

Figure 6

(A) Concentrations of the endogenous hormones i.e., ABA, JA and IAA in the leaves of high and low biomass genotypes. (B) Linear regression model of hormone content and qPCR values of the genes identified in RNA-Seq. analysis in the respective signaling pathway.

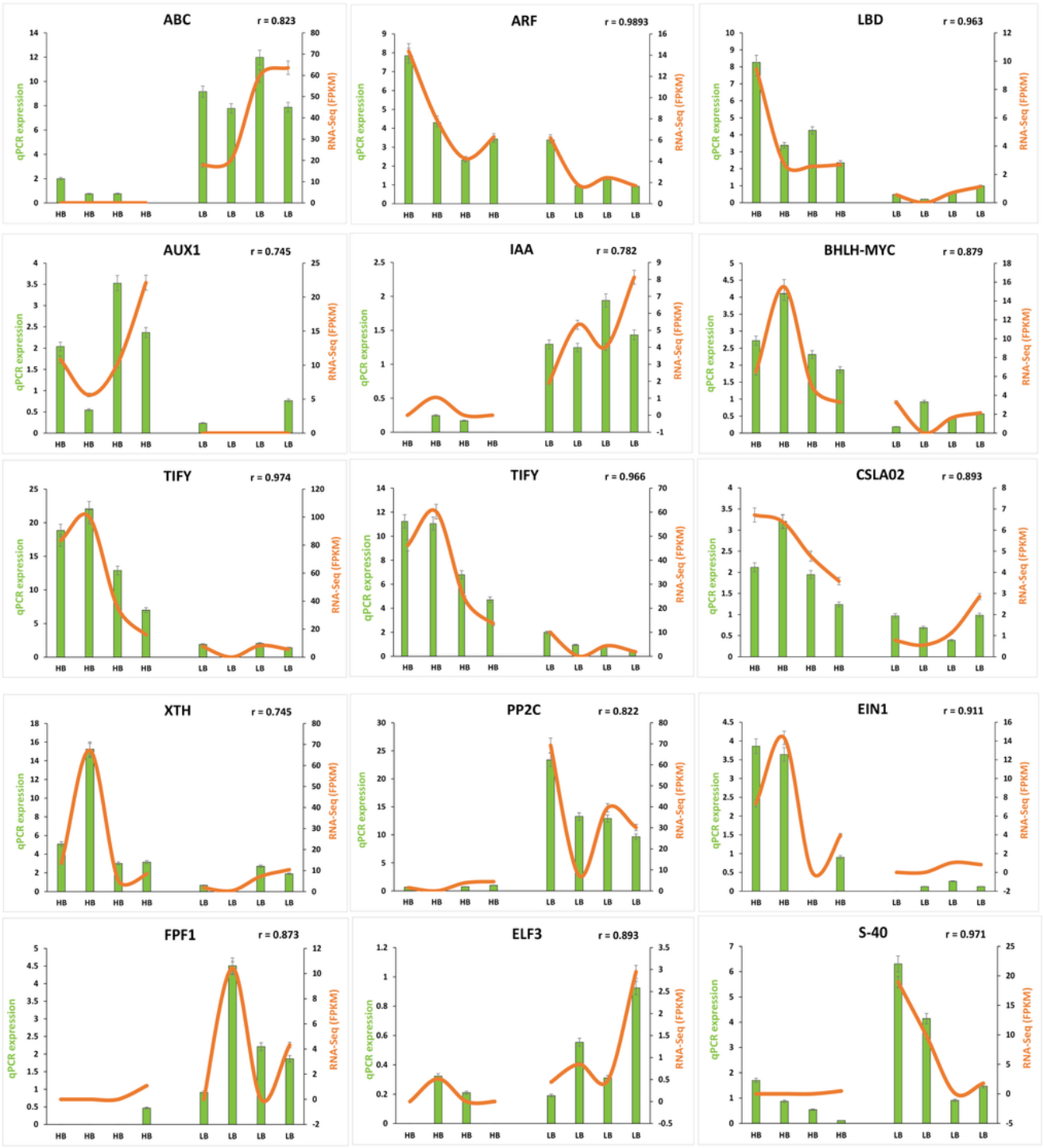

Figure 7

Confirmation of FPKM by qPCR expression, green bars represent the relative expression in qPCR, and orange lines represent FPKM values in transcriptome for corresponding genes. Values on the y-axis indicate relative expression levels of qPCR (left) and RNA-Seq (right). Error bars show standard error of the means at ($P < 0.05$) and “r” is indicative of correlation between qPCR and FPKM expression values.

Supplementary Files

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

- Additionalfile1.docx
- Additionalfile2.xlsx