Transcriptomes Analysis Reveals Novel Insight Into the Molecular Mechanisms of Somatic Embryogenesis in Hevea Brasiliensis

Ying Wang  
Institute of Tropical Bioscience and Biotechnology

Hui-Liang Li  
Institute of Tropical Bioscience and Biotechnology

Yong-Kai Zhou  
Hainan University

Dong Guo  
Institute of Tropical Bioscience and Biotechnology

Jia-Hong Zhu  
Institute of Tropical Bioscience and Biotechnology

Shiqing Peng (shqpeng@163.com)  
Institute of Tropical Bioscience and Biotechnology  https://orcid.org/0000-0002-3984-1902

Research article

Keywords: Heveabrasiliensis, Somatic embryogenesis, RNA-seq, Hormone signal, Transcript factor, Histone modification

DOI: https://doi.org/10.21203/rs.3.rs-44131/v1

License: ☝️ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Background

Somatic embryogenesis (SE) is a promising technology for plant vegetative propagation, which has an important role in tree breeding. Though rubber tree (Hevea brasiliensis Muell. Arg.) SE has been founded, few late SE-related genes had been identified and the molecular regulation mechanisms of late SE still not well understood.

Results

In this study, the transcriptomes of embryogenic calluses (EC), primary embryo (PE), cotyledonary embryo (CE), abnormal embryo (AE), mature cotyledonary embryo (MCE) and withered abnormal embryo (WAE) were analyzed. A total of 887,852,416 clean reads were generated, more than 85.92% of the clean reads were mapped to the rubber tree genome. The de novo assembly generated 36937 unigenes. The differentially expressed genes (DEGs) were identified in the pairwise comparisons of CE vs. AE and MCE vs. WAE, respectively. The specific common DEGs mainly involved in phytohormones signaling pathway, biosynthesis of phenylpropanoid and starch and sucrose metabolism. Among them, hormone signal transduction related genes were significantly enriched, especially the auxin signaling factors (AUX-like1, GH3.1, SAUR32-like, IAA9-like, IAA14-like, IAA27-like, IAA28-like and ARF5-like). The transcript factors including WRKY40, WRKY70, MYBS3-like, MYB1R1-like, AIL6 and bHLH93-like were characterized as molecular markers for rubber tree late SE. CML13, CML36, CAM-7, SERK1 and LEAD-29-like were related to rubber tree late SE. In addition, histone modification had crucial roles during rubber tree late SE.

Conclusions

This study provides important information to elucidate the molecular regulation during rubber tree late SE.

Background

Rubber tree (Hevea brasiliensis Muell. Arg.), a tropical rubber-producing tree within the Euphorbiaceae family, native to the great Amazonian basin of South America, is now widely cultivated to product natural rubber in Southeast Asia [1]. Rubber tree is a perennial cross-pollination tree with a long juvenile period, which takes a long period of time and makes low efficiency of hybrid breeding [2]. Rubber tree is still propagated mostly by grafting, although the interaction between scion and rootstock of grafted tree affects growth, natural rubber yield [3, 4].

Somatic embryogenesis (SE) is a promising and rapid vegetative propagation technique for plant regeneration. Plant regeneration via SE process in rubber tree had been established using different kinds
of explants including immature anthers, internal integuments of immature fruits, inflorescence, as well as root [5–8]. The regenerated plants have juvenile characters and their own roots, which are called self-rooted juvenile clones (SRJCs). Compared with donor clones, SRJCs is superior in growth, rubber yield and stress resistance [9–11], which is promising a new rubber tree planting material in the future. However, there are intrinsic limitations in SE of rubber tree, such as only a limited number of genotypes can obtain somatic embryos, different lines of friable calli from the same explant shows a different embryogenic potentials [12–14]. In addition, cotyledonary embryo can be germinate, while several abnormal embryos such as globular embryos are unable to germinate [14, 15].

To study the molecular regulation mechanism of plant SE, the transcriptomes analysis were carried out to identify SE related genes by RNA-seq in plant species, including herbaceous plants such as Arabidopsis [16], Gossypium hirsutum [17], Maize [18], Strawberry [19], Rice[20], etc., and woody plants such as norway spruce [21], coconut plam [22], brazilian pine [23], camphor tree [24], papaya [25], Dimocarpus longan [26] and so on. These study demonstrated regulation mechanisms of SE at molecular level, and several potential somatic embryogenesis key genes were identified, such as late embryogenesis abundant (LEA) protein [25], somatic embryogenesis receptor-like kinase (SERK) [27, 28], Leafy Cotyledon [28, 29], AGAMOUS-like 15 [30, 31], BABYBOOM [28, 32], WUSCHEL [33, 34], and WUSCHEL homeobox 2 [28, 35].

In rubber tree, a total of 28 differentially expressed cDNAs were identified during induction in the embryogenic regenerating line using the differential display method and five cDNAs could be enable an early diagnosis of friable rubber tree callus embryogenic potential [12]. Three MADS-box genes were differentially expressed during rubber tree SE [36]. The transcript profiling of AP2/ERF genes and ethylene biosynthesis genes were analyzed in different regeneration potential callus lines [14]. These studies about rubber tree SE-related genes mainly focused on early embryogenesis. However, the molecular regulation mechanisms of the late stage of rubber tree SE still not well understood. To clarifying the mechanism of the regenerate competence of different embryos depended on the potential genes during late SE, we investigated the expression profiling using RNA-seq technology. This study will offer valuable information for the molecular regulation mechanisms of rubber tree late SE.

**Results**

**Induction of somatic embryogenesis**

An efficient *H. brasiliensis* SE system was established (Fig. 1). The immature anthers were cultured in solid MS medium supplemented with 2, 4-D, KT and NAA for 50 days. At the end of the period, ECs were obtained. The ECs were placed in the MS medium containing IAA and GA3 for embryo induction. After 40 days, PEs were collected. The PEs were transferred to MS medium containing BA and AgNO₃ for growing. After 40 days, there were two different embryos (CEs, AEs) in the culture medium. We observed significant difference between CEs and AEs in phenotype. The CEs and AEs were placed on half-strength MS medium containing IAA and BA. The CEs turned stronger into MCE 20 days later, whereas the AEs turned
brown and hardly grown up into WAE. After 30 days, the MCEs grew into complete seedlings, whereas the WAEs turned black and died. Based on the above phenotypic observation, six different samples during SE were selected for further study.

**Transcriptome analysis of rubber tree SE**

To provide the transcriptional analysis of *H. brasiliensis* SE, the cDNA libraries of EC, PE, CE, AE, MCE and WAE (three biological replicates for each sample) were constructed and sequenced by using the Illumina HiSeq™ 2500 platform. In total, 915,535,874 sequence reads were obtained from all 18 cDNA libraries. A total of 887,852,416 clean reads were retained by filtering the reads with adaptor sequences and ambiguous “N” base more than 1% and base quality less than Q15. A quality score above 30 (Q30) percentage was over 97.92% and the GC percentage was over 43% (Table 1). On average more than 85.92% of the clean reads were mapped to *H. brasiliensis* genome.

All unigenes were annotated by the blast search against the public databases of using BLASTx (E-value—5 ≤ 10). All 36937 unigenes were annotated in 4 databases involved in COG, GO, KOG and eggNOG (Table 2). Among 36937 unigenes, 19619, 20954 and 36362 unigenes were successfully annotated in GO, KOG, eggNOG, respectively (Fig. S1, S2, S3).

According to the COG functional classification, the 13421 unigenes were categorized into 25 COG categories. The four most highly represented COG categories were “general function prediction only” (20.57%), “transcription” (11.75%), “replication, recombination and repair” (11.53%) and “signal transduction mechanisms” (10.51%)(Fig. 2).

**Global analysis of gene expression during rubber tree**

To confirm the overlap existing, statistical analysis was performed on the expressed genes in the four different developmental stages of *H. brasiliensis* SE (Fig. 3a). A total of 25841 genes overlapped in the four stages. 155 genes overlapped between EC and PE. 290 genes overlapped between PE and CE. 193 genes overlapped between CE and MCE; 388, 297, 152 and 582 genes were preferentially expressed in EC, PE, CE and MCE respectively. Statistical analysis was performed on the expressed genes in the comparisons of PE, AE and CE (Fig. 3b), 662 genes were exclusive to PE vs. AE.1369 genes were exclusive to PE vs. CE. Moreover, 365 genes were found in AE vs. CE. To evaluate the differences of molecular response among all samples, the expression level of the unigenes was calculated by FPKM. The top 20 expressed genes from EC, PE, CE and MCE libraries were shown in Table 3. Some genes including *glutathione S-transferase(GST)*, *lipid-transfer protein(LTP)*, *peroxidase*POD, *indole-3-acetic acid-amido synthetase GH3.1*, *ADP-ribosylation factor*, *catalase isozyme*, and *polyubiquitin* were highly expressed in four satges.

In order to reveal the potential key factors and deep understand the regulatory network of SE, the unigenes of each libarary of *H. brasiliensis* SE were analyzed comparatively under the condition of − 1.0 ≥ Log₂ [FC] ≥ 1.0 and FDR < 0.01. A total of 9415 DEGs were obtained in EC vs. PE, PE had 5260 up-
regulated and 4155 down-regulated gene. In PE vs. CE, CE had 1483 genes up-regulated and 2366 down-regulated. In CE vs. MCE, 6449 DEGs were obtained, of which 4016 DEGs were up-regulated, whereas 2433 DEGs were down-regulated. The 2820 DEGs were found in PE vs. AE with 1300 up-regulated and 1520 down-regulated DEGs. In AE vs. WAE, 5590 DEGs were obtained, of which 3318 DEGs were up-regulated, whereas 2272 DEGs were down-regulated. In AE vs. CE, 1536 DEGs were found with 556 up-regulated and 980 down-regulated DEGs. The 3307 DEGs were found between WAE vs. MCE with 1938 up-regulated and 1369 down-regulated DEGs (Fig. 4).

**GO analysis of DEGs between cotyledonary embryo and abnormal embryo**

To further demonstrate the unigenes functions, gene ontology (GO) assignments were carried out using the Blast2GO program. In AE vs. CE, 843 DEGs were classified to three major categories: biological processes (BP), cellular components (CC) and molecular function (MF). A total of 41 GO subcategories were enriched over three major functional categories. The main subcategories were shown in Fig. 5a. The six major subcategories of biological process were metabolic process, cellular process, single-organism process, biological regulation, localization and response to stimulus. The five major subcategories of cellular component were membrane, cell, cell part, organelle and membrane part. The four major subcategories of molecular function were binding, catalytic activity, transporter activity and nucleic acid binding transcription factor activity.

In WAE vs. MCE, 1927 DEGs were classified to three major categories: biological processes (BP), cellular components (CC) and molecular function (MF). The 41 GO subcategories were enriched over three major functional categories (Fig. 5b). The major subcategories of three categories were consistent with the result in AE vs. CE.

**KEGG Pathways of DEGs between cotyledonary embryo and abnormal embryo**

There were 376 DEGs in AE vs. CE, which were assigned to 46 KEGG pathways (Fig. 6a). The most representative pathways were phenylpropanoid biosynthesis (25 unigene, Fig. S4A), plant hormone signal transduction (21 unigenes, Fig. S4B), starch and sucrose metabolism (20 unigenes, Fig. S4C), phenylalanine metabolism (19 unigenes), carbon metabolism (15 unigenes), biosynthesis of amino acid (14 unigenes) and glutathione metabolism (14 unigenes).

In WAE vs. MCE, the 771 DEGs were assigned to 57 KEGG pathways (Fig. 6b). The 7 most represented pathways were phenylpropanoid biosynthesis (63 unigenes, Fig. S5A), starch and sucrose metabolism (49 unigenes, Fig. S5B), plant hormone signal transduction (46 unigenes, Fig. S5C), carbon metabolism (31 unigenes), photosynthesis (30 unigenes), phenylalanine metabolism (29 unigenes) and cyanoamino acid metabolism (29 unigenes). The results indicated that phenylpropanoid biosynthesis, phytohormones
signaling pathway and sucrose and starch metabolism played importance roles during *H. brasiliensis* late SE.

**Differential expression of hormone signal transduction related genes between cotyledonary embryo and abnormal embryo**

Various phytohormones induced SE and regeneration in several plants have already been reported. To further understand hormone regulation, FPKMs of hormonal signal transduction related genes were analyzed (Fig. 7a and Table S1). Among all auxin signal transduction related genes, *AUX-like5, IAA9-like, IAA28-like* and *GH3.1* were up-regulated in CE. *SAUR71-like* were higher expression in AE than CE. *AUX22D-like, AUX28-like, AUX-like1, AUX-like2, SAUR32-like, IAA14-like* and *IAA27-like* were highly expressed in MCE. *ARF5-like* was low expression in CE but highly expressed in MCE. These genes participated in auxin signaling pathway, which was important for cell enlargement and plant growth (Fig. 7b).

Abscisic acid (ABA) signal transduction related genes, *PYL2-like* was down-regulated in CE. *PYL4-like* was down-regulated in AE. Jasmonic acid (JA) signal transduction related genes, *JAZ7* was higher expression in CE than AE. *JAZ5* were up-regulated in AE. Ethylene (ET) signal transduction related genes, *RAP2-3* were up-regulated in CE and AE. *RAP2-12-like* and *WR11-like* were highly expressed in CE. *ERF4-like* were up-regulated in MCE. *ERF018-like* were only up-regulated in AE. In brief, these genes involved in the hormones signaling transduction pathways, including auxin, ABA, JA, ET, implying that these hormones were indispensable role in their complicated crosstalk process during *H. brasiliensis* late SE.

**Differential expression of TFs and SE-related genes between cotyledonary embryo and abnormal embryo**

In the embryogenic callus formation, transcription factors (TFs) possessed important functions. 219 TFs were identified in CE vs. AE of *H. brasiliensis* late SE. The following TFs families were overrepresented: WRKY, MYB, MADS-box, AP2/ERF, bHLH. The expression profiles of 19 TFs in CE, AE, MCE and WAE are shown in Fig. 8a and Table S2. *WRKY40* and *WRKY70* were up-regulated in CE and down-regulated in AE. *WRKY23* were higher expression in AE than CE. *MYB26-like and MYB98-like* were up-regulated in AE. *MYBS3-like* and *MYB1R1-like* were up-regulated in MCE. *AGL11* and *AGL15* were up-regulated in AE. *BBM2* was highly expressed in AE. *AIL6* was higher expression in CE than AE. *bHLH93-like* was highly expressed in CE. The expression of *bHLH94-like* was obviously up-regulated in AE. The results implied these TFs may be play a key role in *H. brasiliensis* late SE.

Some SE-related genes, such as CAM [37], SERK [38, 39], LEA [40, 41], have been identified to play vital role during plant embryogenesis. *CML13* and *CML36* were up-regulated in CE but down-regulated in AE. *CAM-5-like* and CAM (LOC110641724) were up-regulated in AE but had not changed in CE. *CAM-7* were up-regulated in CE but down-regulated in AE. *SERK1* were up-regulated in CE. *LEAD-34-like* and *SERK2-like* showed higher expression in AE than CE. *LEAD-29-like* were up-regulated in MCE. The dynamic variation
of the FPKM of these somatic embryogenesis-related genes suggested that they were critical for *H. brasiliensis* late SE.

**Differential expression of histone modifications related genes between cotyledonary embryo and abnormal embryo**

The plant growth regulators and abiotic stress can contribute to induce SE. In the meantime these factors may contribute to induce epigenetic modifications [42]. Histone modification is one of the most important epigenetic modifications and play a key role in the regulation of gene expression [43]. Therefore, the expression levels of histone modifiers were analyzed and shown in Fig. 8b and Table S3. The most genes related to histone methylation showed higher expression in CE than AE. The histone H3 lysine 9 methyltransferase genes (*SUVH1-like*, *SUVH3-like*, *SUVH4-like* and *SUVH9*), *SUVR3-like*, *EZA1-like* and *ASHH3-like* were expressed at a higher level in CE. In addition, histone demethylation related genes, *LSD1-homolog 1-like* were highly expressed in CE. *LSD1-homolog 2* were up-regulated in MCE. The increased expression of genes in CE or MCE suggested that it is likely to have a function during late SE.

Histone acetyltransferases (HATs) may mediate hyperacetylation and activate transcription. 10 of the 11 genes related to histone acetylation showed significant differential expression in CE vs. AE. *HAG6*, *HAC12-like*, *MCC1* and *GCN5-like* were up-regulated in CE. *HAG11*, *HAG16*, *HAG18* and *HATB-like* were obvious high accumulation in AE. 7 of the 13 genes related to histone deacetylation showed obvious difference in expression in CE vs. AE. *HDAC15-like* and *HDAC19* were highly expressed in CE. *HDAC6-like*, *HDAC9* and *SAP18-like* were obvious high accumulation in AE. The histone phosphorylation related genes were only highly expressed in AE. These genes highly expressed in late SE can be used as candidate genes for in-depth study in vitro embryogenesis.

**qPCR verification of selected DEGs**

Twenty genes related to SE were selected to carry out expression level analysis using qRT-PCR across 6 different tissues of *H. brasiliensis*: EC, PE, CE, AE, MCE and WAE (Fig. 9). The results validated that the expression levels of 19 genes were highly consistent with transcriptome data.

**Discussion**

Somatic embryogenesis is a promising and rapid vegetative propagation technique for plant regeneration. We have established an efficient rubber tree SE system using immature anthers. The transcriptome analysis of plant SE revealed a large number of potential key factors of embryogenesis [25, 26, 44–46]. In this study, we obtained the transcriptome analysis of rubber tree SE derived from EC, PE, CE, AE, MCE and WAE. The de novo assembly generated 36937 unigenes. We found the regenerate competence of cotyledonary embryo and abnormal embryo had obvious differences during late SE. Therefore, this study mainly focused on DEGs in CE vs. AE and MCE vs. WAE. In CE vs. AE, 376 DEGs were provided and assigned to 50 KEGG pathways. The 771 DEGs were also assigned to 50 KEGG pathways in MCE vs. WAE. The most representative pathways were phytohormones signaling pathway,
biosynthesis of phenylpropanoid, and sucrose and starch metabolism in CE vs. AE and MCE vs. WAE. The significant role of phenylpropanoid biosynthesis in plant SE development has been studied. The phenylpropanoid biosynthesis related genes were significantly enriched in papaya embryogenic callus [25]. The similar result was reported about enrichment of the phenylpropanoid pathway in strawberry embryogenic callus [19]. In addition, external stimuli and plant hormones related genes played a key role in the SE process [47, 48]. Sucrose were added to culture medium as exogenous carbon sources in conifers SE [49, 50]. The germination of Norway spruce (Picea abies) somatic embryos was affected by carbohydrates [51]. These suggested genes involved in the three pathways which can play important role in H. brasiliensis late SE.

Auxin are critical regulators in different developmental stages of SEs [52–54]. The addition of exogenous auxin can affect the expression level of endogenous IAA [48, 55–57]. Dynamic change of endogenous IAA has been proved to induce plant SE and improved SE competency [58]. Auxin/Indole-3-Acetic Acids (Aux/IAAs), Gretchen Hagen3s (GH3s), small auxin upregulated RNAs (SAURs) and auxin response factor (ARF) have been identified as auxin-responsive genes in auxin signaling and homeostasis [59–61], can regulate downstream genes precisely and rapidly, and further regulated plant growth and developmental processes. Aux/IAA family play a key role in inhibiting the expression levels of genes transcriptional activated by ARFs [62, 63]. In low auxin levels, Aux/IAA proteins interacted with ARFs and inhibited activation of auxin-responsive genes. In high auxin levels, these proteins can interact with TRANSPORT INHIBITOR RESPONSE 1/AUXIN SIGNALING F-BOX (TIR1/AFB) receptors to be ubiquitinated and subsequently resolved by the 26S proteasome [64–66]. The liberated ARFs regulated the expression of auxin-responsive genes (Fig. 7b). There were 29 Aux/IAA family members in Arabidopsis, but not all genes were induced by auxin [67]. Many Aux/IAA genes have also been identified in other plants, such as, Eucalyptus grandis [68], Solanum Lycopersicon [69], Cucumis sativus [70], Populus trichocarpa [71], Zea mays [72] and Oryza sativa [73–74]. SAUR genes were consisted of a large multigene family, played crucial roles in regulating plant growth and development [75, 76]. GH3 family participated in a series of hormone-dependent processes in plant, including root growth, flowering [77, 78]. In our study, Auxin signal transduction related genes, IAA9-like, IAA28-like and GH3.1 were high accumulation in CE. AUX-like1, SAUR32-like, IAA14-like, IAA27-like and ARF5-like were up-regulated in MCE. JA and ET have also been reported to play a role in SE induction [79]. JAZ7, RAP2-12-like and WRI1-like were highly expressed in CE. The phytohormones signaling pathway related genes displayed intricate regulation during H. brasiliensis late SE. The regulatory mechanisms of these genes in H. brasiliensis late SE will be confirmed in the future study.

Transcription factors are key factors in plant embryogenesis and development. Many researches on SE development showed that complicated transcription regulation networks existed in cell maintaining embryogenic competency, embryogenic callus formation [52, 80]. The WRKY family has been reported to be involved in biotic or abiotic stresses responses, and were up-regulated in plant SE [81]. The WRKY genes showed higher expression and be related to embryogenic callus formation [82]. The up-regulated WRKY genes were high ratio in embryogenic callus formation of bread wheat [44]. In addition, MYB family was also involved in plant development and growth [83–86], hormone signal transduction [87, 88].
In our study, WRKY40, WRKY70, MYBS3-like and MYB1R1-like were highly expressed in CE, they can be used as marker genes for *H. brasiliensis* late SE.

**AtEMK**, a member of the AP2/ERF family, was ectopically expressed and promote the initiation of somatic embryos in *Arabidopsis* and *H. brasiliensis* [14, 89]. BABY BOOM (BBM) had been reported as a marker in *Brassica napus* SE [90]. The over-expression of BBM can enhance SE and regeneration ability in tobacco, sweet pepper, cacao [91–93]. The bHLH family are involved in developmental, growth and abiotic stress responses [94], axillary meristem formation [52]. They also participate in abscisic acid and brassinosteroid signaling in *Arabidopsis* and rice [95]. A member of bHLH protein BIM1 regulated *Arabidopsis* SE and be involved in auxin and BR signaling pathways [96]. In our study, AIL6 and bHLH93-like were highly expressed in CE, they might play a key role in *H. brasiliensis* late SE.

SERK has been proved as a key factor in plant SE. *AtSERK1* was higher expression during *Arabidopsis* embryogenic formation [97]. SERK was abundant in embryogenic tissues in *Dactylis glomerate* [98]. However, SERKs were also tested in non-embryogenic tissues in maize, rice and wheat [38, 99, 100]. Ca$^{2+}$ has been identified to play a mediating role during plant SE [37, 101]. *LEA5*, a late embryogenesis abundant proteins gene, was highly expressed in late embryogenesis [102]. In our study, *SERK1, CML13, CML36* and *CAM-7* were up-regulated in CE. *LEAD-29-like* were up-regulated in MCE. These genes can have various regulatory functions in *H. brasiliensis* late SE.

The histone modifications played important roles in gene expression, DNA replication and transcription, chromatin compaction [103, 104]. KRYTONITE (KYP), a histone H3 lysine 9 methyltransferase, also showed higher expression level in *Arabidopsis* somatic embryos [16]. In our study, the histone methylation related genes, *SUVH1-like, SUVH3-like, SUVH4-like, SUVH9, SUVR3-like, EZA1-like* and *ASHH3-like* were expressed at a higher level in CE. In addition, histone demethylation related genes, *LSD1-homolog 1-like* were highly expressed in CE. *LSD1-homolog 2* were only up-regulated in MCE. These genes can play important role in *H. brasiliensis* late SE.

Some HATs including HAG1, HAF2, HAC1, HAC2, HAC4, HAC5 and HAC12 have been identified in *Arabidopsis* [16, 105, 106]. HAC2, HAG2 and HAG3 showed more accumulation in somatic embryos as compared to leaf tissues [16]. Similarly, in our study, histone acetylation related genes (*HAG6, HAC12-like, MCC1* and *GCN5-like*) and histone deacetylation related genes (*HDAC15-like, HDAC19*) showed obvious high accumulation in CE. The potential function histone acetylation/ deacetylation related genes in *H. brasiliensis* late SE will be further studied.

**Conclusions**

In this study, the transcriptome data for rubber tree SE was generated. A comparative analysis of gene expression profiles during rubber tree late SE provided a series of DEGs that regulated late SE in rubber tree. We revealed the expression level of some genes related to phytohormones signaling pathway such as auxin, JA and ET signaling pathway, implying their possible roles in rubber tree late SE. The transcript
factors such as WRKY, MYB, AP2 and bHLH, as well as CAM, SERK and LEA were related to rubber tree late SE, might play a key role in SE. Histone modification might have crucial roles during late SE. This study provides novel insights into the molecular regulation mechanisms during rubber tree late SE.

Materials And Methods

Plant material and induction of somatic embryogenesis

*H. brasiliensis* clones of reyan7-33-97 were planted in National Rubber Tree Varieties Resource Garden of the Chinese Academy of Tropical Agriculture Sciences, Danzhou, Hainan, China.

Immature male flowers were gathered from the rubber tree of reyan7-33-37. Immature male flowers were surface-sterilized with 75% (v/v) ethanol for 30 s, and followed to immerse in 0.2% (v/v) mercuric chloride solution for 10 min, and then washed four times with distilled water. The immature anthers were cultured in solid Murashige and Skoog (MS) medium containing 1 mg l\(^{-1}\) 2,4-dichlorophenoxyacetic acid (2,4-D), 1 mg l\(^{-1}\) kinetin (KT) and 0.5 mg l\(^{-1}\) naphthaleneacetic acid (NAA). After an additional 5–6 weeks of growth, embryogenic calluses (EC) were obtained in the darkness and 26–28°C. These samples of primary embryo (PE), cotyledonary embryo (CE), abnormal embryo (AE), mature cotyledonary embryo (MCE) and withered abnormal embryo (WAE) were collected successively. All samples were rapidly frozen in liquid nitrogen, and stored at -80°C until RNA extraction. Three biological replicates were prepared for each sample.

Construction of cDNA library and sequencing

Total RNA was extracted using RNApure Kit (Polysaccharides and Poly phenolics-rich, QIAGEN). RNA degradation and contamination was monitored on 1% agarose gels. The quality of RNA was detected by using the NanoDrop 2000 spectrophotometer (IMPLEN, CA, USA). The mRNA was enriched from total RNA by magnetic beads containing Oligo (dT) and broken into small fragments. First strand cDNA was then synthesized by these cleaved mRNA fragments. The second-strand cDNA was synthesized using DNA polymerase I and RNaseH. The purified cDNAs were carried out end repair and an addition of an “A” base. Finally, PCR was performed to generate cDNA libraries. The library quality was examined using the Qsep100 Analyzer (BIOptic Inc, Taiwan, China). The cDNA libraries were deep sequenced on the Illumina novaseq6000 cDNA sequencing platform.

Transcriptome de novo assembly and annotation

High quality, clean reads were acquired by filtering the reads with adaptor sequences and ambiguous “N” base more than 1% and base quality less than Q15. All unigenes were obtained by comparing with the *H. brasiliensis* genome from NCBI. All unigenes were annotated by the blast search against a number of public databases of using BLASTx (E-value–5 ≤ 10). The public databases included the Clusters of Orthologous Groups of proteins database (COG) (http://www.ncbi.nlm.nih.gov/COG), Evolutionary Genealogy of Genes: Non-supervised Orthologous Groups (eggNOG) (http://eggnog5.embl.de/), the Gene
Ontology database (GO) (http://www.geneontology.org/), the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database (http://www.genome.jp/kegg), the Clusters of Protein homology database (KOG) (http://www.ncbi.nlm.nih.gov/KOG/).

**Analysis of differentially expressed genes (DEGs)**

The expression level of the unigenes was calculated by FPKM (expected number of Fragments Per Kilobase of transcript sequence per Millions base pairs sequenced). The genes with $-1.0 \geq \log_2 [FC] \geq 1.0$ and the threshold of FDR $< 0.01$ were regarded as DEGs.

**Expression profiles of genes in** H. brasiliensis SE

FPKM was applied to analyze the gene expression level. The heat map was created using $\log_2 [\text{FPKM}]$ with the pheatmap package [107].

**Quantitative PCR (qPCR)**

Twenty genes were chosen for validation by qPCR. The samples of EC, PE, CE, AE, MCE and WAE were used for RNA extraction, and then reverse transcribed into cDNA as template. Each sample included three biological replicates. QPCR specific primers for the twenty genes were designed by using Primer Premier software 6.0 (Table S4). HbACT7 was amplified as a standard control. qPCR was performed on a Mx3005P Real-Time PCR system using a SYBR Premix EX Taq II™ Kit (TaKaRa, China). All reactions were performed at 95 °C for 30 s, 40 cycles at 95 °C for 10 s, 58 °C for 20 s, and 72 °C for 25 s. The $2^{-\Delta\Delta Ct}$ method was used to calculate the relative expression levels of genes [108]. The statistical differences were analyzed by ANOVA (One-way analysis of variance) based on Fisher's LSD test ($P < 0.05$ and $P < 0.01$) [109].

**Abbreviations**

2,4-D 2,4-dichlorophenoxyacetic acid; ABA:Abscisic acid; AE:Abnormal embryo; ARF:Auxin response factor; BP:Biological process; CC:Cellular component; cDNA:Complementary DNA; CE:Cotyledonary embryo; COG:Clusters of orthologous groups of proteins database; DEGs:Differentially expressed genes; EC:Embryogenic calluses; eggNOG:Evolutionary genealogy of genes:non-supervised orthologous groups; ET:Ethylene; FPKM:Fragments per kilo base of transcript sequence per millions base pairs sequenced; GEO:Gene expression omnibus; GO:Gene ontology; JA:Jasmonic acid; KEGG:Kyoto encyclopedia of genes and genomes; KOG:Clusters of Protein homology database; KT:Kinetin; MF:Molecular function; LEA:Late embryogenesis abundant protein; NAA:Naphthaleneacetic acid; PE:Primary embryo; qPCR:Quantitative polymerase chain reaction; SE:SERK:Somatic embryogenesis receptor-like kinase; Somatic embryogenesis; SRJC:s:Self-rooted juvenile clones; TFs:transcription factors; MCE:Mature cotyledonary embryo; Non-supervised Orthologous Groups (eggNOG)WAE:Withered abnormal embryo.

**Declarations**
Acknowledgements

We thank the GENE DENOVO Technology Corporation (Guangzhou, China) for RNA-seq sequencing.

Authors’ contributions

The study was conceived and directed by SQP. YW and SQP conducted the analysis and wrote the manuscript. YW,YKZ, HLL, DG and HJZ performed experiments and carried out the data analysis. All authors read and approved the final manuscript.

Funding

This study was supported by the National Key R&D Program of China (Grant No. 2018YFD1000502) and the National Natural Science Foundation of China (Grant No. 31670611).

Availability of data and materials

The generated RNA-seq data have been deposited in NCBI-SRA database under the accession of PRJNA646309.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Priyadarshan PM, Goncalves PS. Hevea gene pool for breeding. Genet Resour Crop Evol. 2003;50:101–14.
2. Clément-Demange A, Priyadarshan P, Hoa TTT, Venkatachalam P. Hevea rubber breeding and genetics. Plant Breeding Rev. 2007;29:177.
3. Ahmad B. Effect of rootstock on growth and water use efficiency of Hevea during water stress. J Rubber Res. 1999;2:99–119.
4. Sobhana P, Gopalakrishnan J, Jacob J, Sethuraj MR. Physiological and biochemical aspects of stock-scion interaction in Hevea brasiliensis. Indian J Nat Rubber Res. 2001;14:131–6.
5. Wang ZY, Zeng XS, Chen CQ, Wu HY, Li QY, Fan GJ, Lu WJ. Induction of rubber plantlets from anther of Hevea brasiliensis Muell. Arg. in vitro. Chinese J Trop Crops. 1980;1:25–6.
6. Carron MP, Enjalric F. Studies on vegetative micropropagation of *Hevea brasiliensis* by somatic embryogenesis and in vitro microcutting. In: Fujiwara A, editor. Plant Tissue Culture. 1982; 751–752.
7. Sushamakumari S, Sobha S, Rekha K, Jayasree R, Asokan MP. Influence of growth regulators and sucrose on somatic embryogenesis from immature inflorescence of *Hevea brasiliensis* (Muell. Arg.). Indian J Nat Rubber Res. 2000;13:19–29.
8. Zhou QN, Jiang ZH, Huang TD, Li WG, Sun AH, Dai XM, Li Z. Plant regeneration via somatic embryogenesis from root explants of *Hevea brasiliensis*. Afr J Biotechnol. 2010;9:8168–73.
9. Yang SQ, Mo YY. Some physiological properties of latex from somatic plants of *Hevea brasiliensis*. Chinese J Trop Crops. 1994;15:13–20.
10. Yuan XH, Yang SQ, Xu LY, Wu JL, Hao BZ. Characteristics related to higher rubber yield of *Hevea brasiliensis* juvenile-type clone GL1. J Rubber Res. 1998;1:125–32.
11. Chen XT, Wang ZY, Wu HD, Zhang XJ. A new planting material of *Hevea brasiliensis* self-rooting juvenile-type clone. Chinese J Trop Crops. 2002;23:192–3.
12. Charbit E, Legavre T, Lardet L, Bourgeois E, Ferrière N, Carron MP. Identification of differentially expressed cDNA sequences and histological characteristics of *Hevea brasiliensis* calli in relation to their embryogenic and regenerative capacities. Plant Cell Rep. 2004;22:539–48.
13. Lardet L, Dessailly F, Carron MP, Montoro P, Monteuuis O. Influences of aging and cloning methods on the capacity for somatic embryogenesis of a mature *Hevea brasiliensis* genotype. Tree Physiol. 2009;29:291–8.
14. Piyatrakul P, Putranto RA, Florence Martin F, Rio M, Dessailly F, Leclercq J, Dufayard JF, Lardet L, Montoro P. Some ethylene biosynthesis and AP2/ERF genes reveal a specific pattern of expression during somatic embryogenesis in *Hevea brasiliensis*. BMC Plant Biol. 2012;12:244.
15. Lardet L, Martin F, Dessailly F, Carron MP, Montoro P. Effect of exogenous calcium on post-thaw growth recovery and subsequent plant regeneration of cryopreserved embryogenic calli of *Hevea brasiliensis* (Muell. Arg.). Plant Cell Rep. 2007;26:559–69.
16. Wickramasuriya AM, Dunwell JM. Global scale transcriptome analysis of *Arabidopsis* embryogenesis in vitro. BMC Genom. 2015;16:301.
17. Jin FY, Hu LS, Yuan DJ, Xu J, Gao WH, He LR, Yang XY, Zhang XL. Comparative transcriptome analysis between somatic embryos (SEs) and zygotic embryos in cotton: evidence for stress response functions in SE development. Plant Biotechnol J. 2014;12:161–73.
18. Salvo SA, Hirsch CN, Buell CR, Kaeppler SM, Kaeppler HF. Whole transcriptome profiling of maize during early somatic embryogenesis reveals altered expression of stress factors and embryogenesis-related genes. PLoS One. 2014;9:e111407.
19. Gao LM, Zhang J, Hou Y, Yao YC, Ji QL. RNA-seq screening of differentially-expressed genes during somatic embryogenesis in *Fragaria x ananassa* Duch. ‘Benihopp’ J Hortic Sci Biotechnol. 2015;90:671–81.
20. Indoliya Y, Tiwari P, Chauhan AS, Goel R, Shri M, Bag SK, Chakrabarty D. Decoding regulatory landscape of somatic embryogenesis reveals differential regulatory networks between japonica and
indica rice subspecies. Sci Rep. 2016;6:23050.

21. Yakovlev IA, Carneros E, Lee Y, Olsen JE, Fossdal CG. Transcriptional profiling of epigenetic regulators in somatic embryos during temperature induced formation of an epigenetic memory in Norway spruce. Planta. 2016;243:1237–49.

22. Rajesh MK, Fayas TP, Naganeswaran S, Rachana KE, Bhavyashree U, Sajini KK, Karun A. De novo assembly and characterization of global transcriptome of coconut palm (Cocos nucifera L.) embryogenic calli using Illumina paired end sequencing. Protoplasma. 2016;253:913–28.

23. Elbl P, Lira BS, Andrade SCS, Jo L, Dos Santos ALW, Coutinho LL, Floh EIS, Rossi M. Comparative transcriptome analysis of early somatic embryo formation and seed development in Brazilian pine, Araucaria angustifolia (Bertol.) Kuntze. Plant Cell Tiss Org Cult. 2015;120:903–15.

24. Shi XP, Zhang CJ, Liu QH, Zhang Z, Zheng B, Bao MZ. De novo comparative transcriptome analysis provides new insights into sucrose induced somatic embryogenesis in camphor tree (Cinnamomum camphora L.). BMC Genom. 2016;17:26.

25. Jamaluddin ND, Mohd Noor N, Goh H. Genome-wide transcriptome profiling of Carica papaya L. embryogenic callus. Physiol Mol Biol Plants. 2017;23:357–68.

26. Chen Y, Xu X, Liu Z, Zhang Z, Han X, Lin Y, Lai Z. Global scale transcriptome analysis reveals differentially expressed genes involve in early somatic embryogenesis in Dimocarpus longan Lour. BMC Genom. 2020;21:4.

27. Savona M, Mattioli R, Nigro S, Falasca G, Della Rovere F, Costantino P, De Vries S, Ruffoni B, Trovato M, Altamura MM. Two SERK genes are markers of pluripotency in Cyclamen persicum mill. J Exp Bot. 2012;63:471–88.

28. Rupps A, Raschke J, Rümmler M, Linke B, Zoglauer K. Identification of putative homologs of Larix decidua to BABYBOOM (BBM), LEAFY COTYLEDON1 (LEC1), WUSCHEL-related HOMEBOX2 (WOX2) and SOMATIC EMBRYOGENESIS RECEPTOR-like KINASE (SERK) during somatic embryogenesis. Planta. 2016;243:473–88.

29. Gaj MD, Zhang S, Harada JJ, Lemaux PG. Leafy cotyledon genes are essential for induction of somatic embryogenesis of Arabidopsis. Planta. 2005;222:977–88.

30. Thakare D, Tang W, Hill K, Perry SE. The MADS-domain transcriptional regulator AGAMOUS-LIKE15 promotes somatic embryo development in Arabidopsis and soybean. Plant Physiol. 2008;146:1663–72.

31. Zheng Q, Zheng Y, Perry SE. AGAMOUS-Like15 promotes somatic embryogenesis in Arabidopsis and soybean in part by the control of ethylene biosynthesis and response. Plant Physiol. 2013;161:2113–27.

32. Horstman A, Li M, Heidmann I, Weemen M, Chen B, Muiño JM, Angenent GC, Boutilier K. The BABYBOOM transcription factor activates the LEC1-ABI3-FUS3-LEC2 network to induce somatic embryogenesis. Plant Physiol. 2017;2175:848–57.

33. Bouchabké-Coussa O, Obellianne M, Linderme D, Montes E, Maia-Grondard A, Vilaine F, Pannetier C. Wuschel overexpression promotes somatic embryogenesis and induces organogenesis in cotton
34. Zheng Q, Perry SE. Alterations in the transcriptome of soybean in response to enhanced somatic embryogenesis promoted by orthologs of Agamouslike15 and Agamous-like18. Plant Physiol. 2014;164:1365–77.

35. Tvorogova VE, Lebedeva MA, Lutova LA. Expression of WOX and PIN genes during somatic and zygotic embryogenesis in Medicago truncatula. Russ J Genet. 2015;51:1189–98.

36. Li HL, Wang Y, Guo D, Tian WM, Peng SQ. Three MADS-box genes of Hevea brasiliensis expressed during somatic embryogenesis and in the laticifer cells. Mol Biol Rep. 2011;38:4045–52.

37. Kiselev KV, Shumakova OA, Manyakhin AY, Mazeika AN. Influence of calcium influx induced by the calcium ionophore, A23187, on resveratrol content and the expression of CDPK and STS genes in the cell cultures of Vitis Amurensis. Plant Growth Regul. 2012;68:371–81.

38. Ito Y, Takaya K, Kurata N. Expression of SERK family receptor-like protein kinase genes in rice. BBA-Gene Struct Expre. 2005;1730:253–358.

39. Singla B, Khurana JP, Khurana P. Structural characterization and expression analysis of the SERK/SERL gene family in rice (Oryza sativa). Int J Plant Genomics. 2009;10:1–8.

40. Hundertmark M, Hincha DK. LEA (late embryogenesis abundant) proteins and their encoding genes in Arabidopsis Thaliana. BMC Genom. 2008;9:118.

41. Kikawada T, Nakahara Y, Kanamori Y, Iwata K, Watanabe M, McGee B, Tunnacliffe A, Okuda T. Dehydration-induced expression of LEA proteins in an anhydrobiotic chironomid. Biochem Biophys Res Comm. 2006;348:56–61.

42. Miguel C, Marum L. An epigenetic view of plant cells cultured in vitro: Somaclonal variation and beyond. J Exp Bot. 2011;62:3713–25.

43. Berr A, Shafiq S, Shen WH. Histone modifications in transcriptional activation during plant development. Biochim Biophys Acta. 2011;1809:567–76.

44. Chu ZL, Chen JY, Sun JY, Dong ZD, Yang X, Wang Y, Xu HX, Zhang XK, Chen F, Cui DQ. De novo assembly and comparative analysis of the transcriptome of embryogenic callus formation in bread wheat (Triticum aestivum L.). BMC Plant Biol. 2017;17:244.

45. Guo HH, Guo HX, Zhang L, Tang ZM, Yu XM, Wu JF, Zeng FC. Metabolome and transcriptome association analysis reveals dynamic regulation of purine metabolism and flavonoid synthesis in transdifferentiation during somatic embryogenesis in Cotton. Int J Mol Sci. 2019; 20:2070.

46. Zhang XL, Wang YL, Yan YY, Peng H, Long Y, Zhang YC, Jiang Z, Liu P, Zou CY, Peng HW, Pan GT, Shen YO. Transcriptome sequencing analysis of maize embryonic callus during early redifferentiation. BMC Genom. 2019;20:159.

47. Zhai LL, Xu L, Wang Y, Zhu XW, Feng HY, Li C, Luo XB, Everlyne M, Liu L. Transcriptional identification and characterization of differentially expressed genes associated with embryogenesis in radish (Raphanus sativus L.). Sci Rep. 2016;6:21652–64.
48. Xu Z, Zhang C, Zhang X, Liu C, Wu Z, Yang Z, Zhou K, Yang X, Li F. Transcriptome profiling reveals auxin and cytokinin regulating somatic embryogenesis in different sister lines of cotton cultivar CCRI24. J Integr Plant Biol. 2013;55:631–42.

49. Schuller A, Reuther G. Response of Abies alba embryonal-suspensor mass to various carbohydrate treatments. Plant Cell Rep. 1993;12:199–202.

50. Lipavská H, Konrádová H. Somatic embryogenesis in conifers: the role of carbohydrate metabolism. In Vitro Cell Dev. 2004;40:23–30.

51. Businge E, Bygdell J, Wingsle G, Moritz T, Egertsdotter U. The effect of carbohydrates and osmoticum on storage reserve accumulation and germination of Norway spruce somatic embryos. Physio Plant. 2013;149:273.

52. Yang F, Wang Q, Schmitz G, Muller D, Theres K. The bHLH protein ROX acts in concert with RAX1 and LAS to modulate axillary meristem formation in Arabidopsis. Plant J. 2012;71:61–70.

53. Zhao Z, Andersen SU, Ljung K, Dolezal K, Miotk A, Schultheiss SJ, Lohmann JU. Hormonal control of the shoot stem-cell niche. Nature. 2010;465:1089–92.

54. Su YH, Zhang XS. Auxin gradients trigger de novo formation of stem cells during somatic embryogenesis. Plant Signal Behav. 2014;4:574–6.

55. Vondráková Z, Eliášová K, Fischerová L, Vágner M. The role of auxins in somatic embryogenesis of Abies alba. Cent Eur J Biol. 2011;6:587–96.

56. Michalczuk L, Ribnický DM, Cooke TJ, Cohen JD. Regulation of Indole-3-acetic acid biosynthetic pathways in carrot cell cultures. Plant Physiol. 1992;100:1346–53.

57. Ribnický DM, Illic N, Cohen JD, Cooke TJ. The effects of exogenous auxins on endogenous indole-3-acetic acid metabolism (the implications for carrot somatic embryogenesis). Plant Physiol. 1996;112:549–58.

58. Thomas C, Bronner R, Molinier J, Prinsen E, van Onckelen H, Hahne G. Immuno-cytochemical localization of indole-3-acetic acid during induction of somatic embryogenesis in cultured sunflower embryos. Planta. 2002;215:577–83.

59. Goldental-Cohen S, Israeli A, Ori N, Yasuor H. Auxin response dynamics during wild-type and entire flower development in tomato. Plant Cell Physiol. 2017;58:1661–72.

60. Guilfoyle TJ, Hagen G. Auxin response factors. Curr Opin Plant Biol. 2007;10:453–60.

61. Abel S, Nguyen MD, Theologis A. The PS-IAA4/5-like family of early Auxin-inducible mRNAs in Arabidopsis thaliana. J Mol Biol. 1995;251:533–49.

62. Abel S, Oeller PW, Theologis A. Early auxin-induced genes encode short-lived nuclear proteins. Proc Natl Acad Sci. 1994;91:326–30.

63. Dreher KA, Brown J, Saw RE, Callis J. The Arabidopsis Aux/IAA protein family has diversified in degradation and auxin responsiveness. Plant Cell. 2006;18:699–714.

64. Weijers D, Benkova E, Jäger KE, Schlereth A, Hamann T, Kientz M, Wilmoth JC, Reed JW, Jürgens G. Developmental specificity of auxin response by pairs of ARF and Aux/IAA transcriptional regulators.
EMBO J. 2005;24:1874–85.

65. Parry G, Calderon-Villalobos LI, Prigge M, Peret B, Dharmasiri S, Itoh H, Lechner E, Gray WM, Bennett M, Estelle M. Complex regulation of the TIR1/AFB family of auxin receptors. Proc Natl Acad Sci. 2009;106:22540–5.

66. Tan X, Calderon-Villalobos LIA, Sharon M, Zheng C, Robinson CV, Estelle M, Zheng N. Mechanism of auxin perception by the TIR1 ubiquitin ligase. Nature. 2007;446:640–5.

67. Overvoorde PJ, Okushima Y, Alonso JM, Chan A, Chang C, Ecker JR, Hughes B, Liu A, Onodera C, Quach H, Smith A, Yu G, Theologis A. Functional genomic analysis of the AUXIN/INDOLE-3-ACETIC ACID gene family members in Arabidopsis thaliana. Plant Cell. 2005;17:3282.

68. Yu H, Soler M, Clemente H, Mila I, Paiva JAP, Myburg AA, Bouzayen M, Grima-Pettenati J, Cassan-Wang H. Comprehensive genome-wide analysis of the Aux/IAA gene family in eucalyptus: Evidence for the role of EgrIAA4 in wood formation. Plant Cell Physiol. 2015;56:700–14.

69. Audran-Delalande C, Bassa C, Mila I, Regad F, Zouine M, Bouzayen M. Genome-wide identification, Functional analysis and expression profiling of the Aux/IAA gene family in tomato. Plant Cell Physiol. 2012;53:659–72.

70. Gan D, Zhuang D, Ding F, Yu Z, Zhao Y. Identification and expression analysis of primary auxin-responsive Aux/IAA gene family in cucumber (Cucumis sativus). J Genet. 2013;92:513–21.

71. Kalluri UC, DiFazio SP, Brunner AM, Tuskan GA. Genome-wide analysis of Aux/IAA and ARF gene families in Populus trichocarpa. BMC Plant Biol. 2007;7:59.

72. Ludwig Y, Zhang Y, Hochholdinger F. The maize (Zea mays L.) AUXIN/INDOLE-3-ACETIC ACID gene family: Phylogeny, synteny, and unique root-type and tissue-specific expression patterns during development. PLoS one. 2013;8:e78859.

73. Jain M, Kaur N, Garg R, Thakur JK, Tyagi AK, Khurana JP. Structure and expression analysis of early auxin-responsive Aux/IAA gene family in rice (Oryza sativa). Funct Integr Genom. 2006;6:47–59.

74. Wu W, Liu Y, Wang Y, Li H, Liu J, Tan J, He J, Bai J, Ma H. Evolution analysis of the Aux/IAA gene family in plants shows dual origins and variable nuclear localization signals. Int J Mol Sci. 2017;18:2107.

75. Ren H, Gray WM. SAUR proteins as effectors of hormonal and environmental signals in plant growth. Mol Plant. 2015;8:1153–64.

76. Kant S, Bi Y, Zhu T, Rothstein SJ. SAUR39, a small Auxin-up RNA gene, acts as a negative regulator of Auxin synthesis and transport in Rice. Plant Physiol. 2009;151:691–701.

77. Khan S, Stone JM. Arabidopsis thaliana GH3.9 influences primary root growth. Planta. 2007;226(1):21–34.

78. Pierdonati E, Unterholzner SJ, Salvi E, Svolacchia N, Bertolotti G, Ioio RD, Sabatini S, Mambro RD. Cytokinin-dependent control of GH3 group II family genes in the Arabidopsis root. Plants (Basel). 2019;8:94.
79. Linkies A, Leubner-Metzger G. Beyond gibberellins and abscisic acid: how ethylene and jasmonates control seed germination. Plant Cell Rep. 2012;31:253–70.
80. Li QF, Zhang SG, Wang JH. Transcriptome analysis of callus from *Picea Balfouriana*. BMC Genom. 2014;15:553.
81. Tripathi P, Rabara RC, Rushton PJ. A systems biology perspective on the role of WRKY transcription factors in drought responses in plants. Planta. 2014;239:255–66.
82. Lagace M, Matton DP. Characterization of a WRKY transcription factor expressed in late torpedo-stage embryos of *Solanum Chacoense*. Planta. 2004;219:185–9.
83. Cominelli E, Tonelli C. A new role for plant R2R3-MYB transcription factors in cell cycle regulation. Cell Res. 2009;19:1231–32.
84. Huang C, Hu G, Li F, Li Y, Wu J, Zhou X. NbPHAN, a MYB transcriptional factor, regulates leaf development and affects drought tolerance in *Nicotiana benthamiana*. Physiol Plant. 2013;149:297–09.
85. Cai H, Tian S, Dong H, Guo C. Pleiotropic effects of TaMYB3R1 on plant development and response to osmotic stress in transgenic *Arabidopsis*. Gene. 2015;558:227–34.
86. Oh JE, Kwon Y, Kim JH, Noh H, Hong SW, Lee H. A dual role for MYB60 in stomatal regulation and root growth of *Arabidopsis thaliana* under drought stress. Plant Mol Biol. 2011;77:91–103.
87. Shin R, Burch AY, Huppert KA, Tiwari SB, Murphy AS, Guilfoyle TJ, Schachtman, DP. The *Arabidopsis* transcription factor MYB77 modulates auxin signal transduction. Plant Cell. 2007;19:2440–53.
88. Zhao Y, Xing L, Wang XG, Hou YJ, Gao JH, Wang PC, Duan CG, Zhu XH, Zhu JK. The ABA receptor PYL8 promotes lateral root growth by enhancing MYB77-dependent transcription of auxin-responsive genes. Sci Signal. 2014;7:ra53.
89. Tsuwamoto R, Yokoi S, Takahata Y. Arabidopsis EMBRYOMAKER. Encoding an AP2 domain transcription factor plays a key role in developmental change from vegetative to embryonic phase. Plant Mol Biol. 2010;73:481–92.
90. El Ouakfaoui S, Schnell J, Abdeen A, Colville A, Labbe H, Han S, Baum B, Laberge S, Miki B. Control of somatic embryogenesis and embryo development by AP2 transcription factors. Plant Mol Biol. 2010;74:313–26.
91. Srinivasan C, Liu Z, Heidmann I, Supena ED, Fukuoka H, Joosen R, Lambalk J, Angenent G, Scorza R, Custers JB, Boutilier K. Heterologous expression of the BABY BOOM AP2/ERF transcription factor enhances the regeneration capacity of tobacco (*Nicotiana Tabacum* L). Planta. 2007;225:341–51.
92. Heidmann I, de Lange B, Lambalk J, Angenent GC, Boutilier K. Efficient sweet pepper transformation mediated by the BABY BOOM transcription factor. Plant Cell Rep. 2011;30:1107–15.
93. Florez SL, Erwin RL, Maximova SN, Guiltinan MJ, Curtis WR. Enhanced somatic embryogenesis in *Theobroma Cacao* using the homologous BABY BOOM transcription factor. BMC Plant Biol. 2015;15:121.
94. Zhou J, Li F, Wang JL, Ma Y, Chong K, Xu YY. Basic helix-loop-helix transcription factor from wild rice (OrbHLH2) improves tolerance to salt and osmotic stress in Arabidopsis. J Plant Physiol. 2009;166:1296–306.

95. Zhang LY, Bai MY, Wu J, Zhu JY, Wang H, Zhang Z, Wang W, Sun Y, Zhao J, Sun XH, Yang HJ, Xu YY, Kim SH, Fujioka S, Lin WH, Chong K, Lu TG, Wang ZY. Antagonistic HLH/bHLH transcription factors mediate brassinosteroid regulation of cell elongation and plant development in rice and Arabidopsis. Plant Cell. 2009;21:3767–80.

96. Chandler JW, Cole M, Flier A, Werr W. BIM1, a bHLH protein involved in brassinosteroid signalling, controls Arabidopsis embryonic patterning via interaction with DORNROSCHEN and DORNROSCHEN-LIKE. Plant Mol Biol. 2009;69:57–68.

97. Hecht V, Vielle-Calzada JP, Hartog MV, Schmidt ED, Boutlier K, Grossniklaus U, de Vries SC. The Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture. Plant Physiol. 2001;127:803–16.

98. Somleva MN, Schmidt EDL, de Vries SC. Embryogenic cells in Dactylis glomerata L. (Poaceae) explants identified by cell tracking and by SERK expression. Plant Cell Rep. 2000;19:718–26.

99. Zhang SZ, Liu XG, Lin YA, Xie GN, Fu FL, Liu HL, Wang J, Gao SB, Lan H, Rong TZ. Characterization of a ZmSERK gene and its relationship to somatic embryogenesis in a maize culture. Plant Cell Tiss Org Cult. 2011;105:29–37.

100. Singla B, Khurana JP, Khurana P. Characterization of three somatic embryogenesis receptor kinase genes from wheat, Triticum aestivum. Plant Cell Rep. 2008;27:833–43.

101. Yang X, Zhang X. Regulation of somatic embryogenesis in higher plants. Crit Rev Plant Sci. 2010;29:36–57.

102. Liu Y, Xie L, Liang X, Zhang S. CpLEA5, the late embryogenesis abundant protein gene from Chimonanthus praecox, possesses low temperature and osmotic resistances in prokaryote and eukaryotes. Int J Mol Sci. 2015;16:26978–90.

103. Kouzarides T. Chromatin modifications and their function. Cell. 2007;128:693–705.

104. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. Nat Rev Genet. 2016;17:487–500.

105. Earley KW, Shook MS, Brower-Toland B, Hicks L, Pikaard CS. In vitro specificities of Arabidopsis co-activator histone acetyltransferases: implications for histone hyperacetylation in gene activation. Plant J. 2007;52:615–26.

106. Perrella G, Consiglio MF, Aiese-Cigliano R, Cremona G, Sanchez-Moran E, Barra L, Errico A, Bressan RA, Franklin FCH, Conicella C. Histone hyperacetylation affects meiotic recombination and chromosome segregation in Arabidopsis. Plant J. 2010;62:796e806.

107. Kolde R, pheatmap. Pretty Heatmaps. https://cran.r-project.org/web/packages/pheatmap/index.html. 2015.
108. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods. 2001;25:402–8.

109. Quirk TJ, Quirk MH, Horton HF. One-way analysis of variance (ANOVA). In book: Excel 2019 for Biological and Life Sciences Statistics. 2020; 169–185.

### Tables

| Sample | Raw Reads | Clean Reads | Raw Bases(Gb) | Clean Bases(Gb) | Effective Rate(%) | Q30 content(%) |
|--------|-----------|-------------|---------------|-----------------|-------------------|----------------|
| EC-1   | 5.2E+07   | 50059934    | 7.86          | 7.56            | 96.21            | 94.81          |
| EC-2   | 5.1E+07   | 49524648    | 7.73          | 7.48            | 96.73            | 94.81          |
| EC-3   | 5.1E+07   | 49118950    | 7.68          | 7.42            | 96.61            | 94.78          |
| PE-1   | 5E+07     | 48319634    | 7.53          | 7.25            | 96.29            | 97.01          |
| PE-2   | 5.1E+07   | 49061282    | 7.64          | 7.36            | 96.33            | 96.86          |
| PE-3   | 5.1E+07   | 48891852    | 7.6           | 7.33            | 96.46            | 96.9           |
| CE-1   | 5.1E+07   | 49805096    | 7.73          | 7.52            | 97.32            | 94.74          |
| CE-2   | 5.2E+07   | 50906314    | 7.88          | 7.69            | 97.56            | 94.91          |
| CE-3   | 5.1E+07   | 50054842    | 7.76          | 7.56            | 97.4             | 94.84          |
| MCE-1  | 5.1E+07   | 49771578    | 7.7           | 7.47            | 96.96            | 95.89          |
| MCE-2  | 5E+07     | 48654566    | 7.54          | 7.3             | 96.85            | 94.92          |
| MCE-3  | 5E+07     | 48974062    | 7.52          | 7.35            | 97.72            | 95.62          |
| AE-1   | 5E+07     | 48881230    | 7.56          | 7.33            | 97.05            | 96.81          |
| AE-2   | 5.1E+07   | 48970492    | 7.6           | 7.35            | 96.7             | 96.75          |
| AE-3   | 5.1E+07   | 48844568    | 7.59          | 7.33            | 96.52            | 96.88          |
| WAE-1  | 5.1E+07   | 49843978    | 7.71          | 7.53            | 97.67            | 94.71          |
| WAE-2  | 5E+07     | 49076246    | 7.6           | 7.41            | 97.49            | 94.65          |
| WAE-3  | 5E+07     | 49093144    | 7.59          | 7.41            | 97.71            | 94.72          |
Table 2 The number and distribution of unigenes annotated in the databases

| Database | Annotated Number | 300<=length<1000 | length>=1000 |
|----------|------------------|-------------------|--------------|
| COG      | 13421            | 4142              | 9153         |
| GO       | 19619            | 5980              | 13639        |
| KOG      | 20954            | 7547              | 13097        |
| eggNOG   | 36362            | 14578             | 21038        |
| All      | 36937            | 14983             | 21176        |

Table 3 The top 20 expressed genes in EC, PE, CE and MCE library
| No | Gene-ID   | Database-ID          | FPKM-EC | Description                                      |
|----|-----------|----------------------|---------|--------------------------------------------------|
| 1  | gene10318 | XM_021818345.1       | 3266.37 | metallothionein-like protein type 2              |
| 2  | gene23077 | XM_021779607.1       | 2803.41 | probable indole-3-acetic acid-amido synthetase GH3.1 |
| 3  | gene24550 | XM_021781891.1       | 2587.98 | peptidyl-prolyl cis-trans isomerase-like         |
| 4  | gene37167 | XM_021801099.1       | 5591.156667 | pathogenesis-related protein PR-4-like         |
| 5  | gene41379 | XM_021807510.1       | 3354.676667 | metallothionein-like protein type 3              |
| 6  | gene41538 | XM_021807736.1       | 1336.78 | peroxidase 12-like                               |
| 7  | gene42156 | XM_021808475.1       | 3150.052519 | L-ascorbate peroxidase, cytosolic-like         |
| 8  | gene548   | XM_021811448.1       | 1719.636667 | thioredoxin H-type-like                         |
| 9  | gene11066 | XM_021819455.1       | 1216.968149 | catalase isozyme 2-like                         |
| 10 | gene1185  | XM_021821602.1       | 6138.533333 | metallothionein-like protein type 2              |
| 11 | gene15002 | XM_021825368.1       | 2578.13 | glucan endo-1,3-beta-glucosidase, basic isoform-like |
| 12 | gene18326 | XM_021830411.1       | 2459.693333 | endochitinase EP3-like                         |
| 13 | gene19193 | XM_021831939.1       | 1835.053333 | glutathione S-transferase F9-like               |
| 14 | gene33311 | XM_021795239.1       | 1328.469977 | pathogenesis-related protein PR-4-like         |
| 15 | gene3644  | XM_021801975.1       | 1588.716667 | thaumatin-like protein 1b                      |
| 16 | gene41464 | XM_021807622.1       | 2882.38 | endochitinase EP3-like                         |
| 17 | gene5134  | XM_021810359.1       | 2157.947846 | catalase isozyme 2                              |
| 18 | gene12558 | XM_021821637.1       | 1745.217667 | cysteine synthase                              |
| 19 | gene21974 | XM_021836019.1       | 1238.357898 | 40S ribosomal protein S25-3-like                |
| 20 | gene24408 | XM_021781690.1       | 1518.806667 | polyubiquitin                                  |

| No | Gene-ID   | Database-ID          | FPKM-PE | Description                                      |
|----|-----------|----------------------|---------|--------------------------------------------------|
| 1  | gene17338 | XM_021828886.1       | 448.8675164 | ADP-ribosylation factor                           |
| 2  | gene24550 | XM_021781891.1       | 1090.893333 | peptidyl-prolyl cis-trans isomerase-like         |
| 3  | gene25944 | XM_021784022.1       | 517.586 | polyubiquitin                                   |
| 4  | gene37168 | XM_021801110.1       | 1051.049333 | pathogenesis-related protein PR-4-like           |
| 5  | gene37235 | XM_021801218.1       | 424.1643333 | probable glutathione S-transferase             |
| 6  | gene5278  | XM_021810573.1       | 700.6816667 | probable aquaporin TIP3-2                        |
| No | Gene-ID  | Database-ID          | FPKM-CE     | Description                  |
|----|----------|----------------------|-------------|------------------------------|
| 1  | gene458  | XM_021811448.1       | 651.8526667 | thioredoxin H-type-like      |
| 2  | gene1185 | XM_021821602.1       | 691.8516667 | metallothionein-like protein type 2 |
| 3  | gene17500| XM_021829184.1       | 651.8516667 | uncharacterized              |
| 4  | gene19193| XM_021831939.1       | 444.690335  | glutathione S-transferase F9-like |
| 5  | gene19425| XM_021832135.1       | 4129.713333 | non-specific lipid-transfer protein 1-like |
| 6  | gene22222| XM_021836400.1       | 475.7673333 | histone H2B                  |
| 7  | gene23940| XM_021780963.1       | 563.119     | osmotin-like protein         |
| 8  | gene37576| XM_021801775.1       | 574.5693333 | thaumatin-like protein       |
| 9  | gene12558| XM_021821637.1       | 419.428     | cysteine synthase            |
| 10 | gene35575| XM_021798790.1       | 464.6649333 | copper transport protein ATX1-like |
| 11 | gene30702| XM_021791318.1       | 1738.72     | peroxidase 42-like           |
| 12 | gene23545| XM_021780391.1       | 2407.276667 | peroxidase 42-like           |
| 13 | gene33942| XM_021796208.1       | 577.7063333 | peptidyl-prolyl cis-trans isomerase 1 |
| 14 | gene24408| XM_021781690.1       | 472.838     | polyubiquitin                |

| No | Gene-ID  | Database-ID          | FPKM-CE     | Description                  |
|----|----------|----------------------|-------------|------------------------------|
| 1  | gene17338| XM_021828886.1       | 1145.143911 | ADP-ribosylation factor      |
| 2  | gene18178| XM_021830179.1       | 943.5483996 | protein translation factor SUI1 homolog 2-like |
| 3  | gene25944| XM_021784022.1       | 1762.396667 | polyubiquitin                |
| 4  | gene37168| XM_021801110.1       | 9026.456667 | pathogenesis-related protein PR-4-like |
| 5  | gene37235| XM_021801218.1       | 2538.03333  | probable glutathione S-transferase |
| 6  | gene5278 | XM_021810573.1       | 1959.873667 | probable aquaporin TIP3-2    |
| 7  | gene5809 | XM_021811329.1       | 1199.044333 | metallothionein-like protein type 2 |
| 8  | gene7973 | XM_021814772.1       | 1108.26     | glutaredoxin                 |
| 9  | gene9140 | XM_021816591.1       | 4259.97     | metallothionein-like protein type 2 |
| 10 | gene17500| XM_021829184.1       | 995.952     | uncharacterized              |
| 11 | gene19425| XM_021832135.1       | 3340.51     | non-specific lipid-transfer protein 1-like |
| 12 | gene20309| XM_021833577.1       | 1361.93341  | ubiquitin-conjugating enzyme E2 28 |
| 13 | gene12558| XM_021821637.1       | 1265.565333 | cysteine synthase            |
| 14 | gene25797| XM_02183808.1        | 1133.149667 | L-ascorbate peroxidase, cytosolic |
| No | Gene-ID | Database-ID       | FPKM-MCE  | Description                                    |
|----|---------|-------------------|-----------|------------------------------------------------|
| 1  | gene17338 | XM_021828886.1    | 1126.976  | ADP-ribosylation factor                        |
| 2  | gene18178 | XM_021830179.1    | 1004.106  | protein translation factor SUI1 homolog 2-like |
| 3  | gene25944 | XM_021784022.1    | 2101.800  | polyubiquitin                                  |
| 4  | gene33318 | XM_021795235.1    | 1202.313  | pro-hevein                                     |
| 5  | gene37168 | XM_021801110.1    | 18664.897 | pathogenesis-related protein PR-4-like         |
| 6  | gene39161 | XM_021804156.1    | 918.860   | 2-methylbutanal oxime monooxygenase            |
| 7  | gene41379 | XM_021807510.1    | 864.198   | metallothionein-like protein type 3            |
| 8  | gene41597 | XM_021807803.1    | 791.620   | elicitor-responsive protein 3-like             |
| 9  | gene42156 | XM_021808475.1    | 2095.956  | L-ascorbate peroxidase, cytosolic-like          |
| 10 | gene548  | XM_021811448.1    | 831.757   | thioredoxin H-type-like                        |
| 11 | gene9140 | XM_021816591.1    | 5217.597  | metallothionein-like protein type 2            |
| 12 | gene11066 | XM_021819455.1    | 1380.437  | catalase isozyme 2-like                       |
| 13 | gene1185  | XM_021821602.1    | 1598.003  | metallothionein-like protein type 2            |
| 14 | gene19425 | XM_021832135.1    | 2686.840  | non-specific lipid-transfer protein 1-like     |
| 15 | gene20309 | XM_021833577.1    | 1215.717  | ubiquitin-conjugating enzyme E2 28             |
| 16 | gene23545 | XM_021780391.1    | 1473.353  | peroxidase 42-like                             |
| 17 | gene24345 | XM_021781508.1    | 1151.487  | translationally-controlled tumor protein homolog |
| 18 | gene30702 | XM_021791318.1    | 2923.16667| peroxidase 42-like                             |
| 19 | gene31451 | XM_021792523.1    | 821.413   | probable aquaporin PIP1-2                     |
| 20 | gene36607 | XM_021800241.1    | 1155.01333| aquaporin TIP1-1-like                         |
| 21 | gene41316 | XM_021807427.1    | 1017.60334| aquaporin PIP1-3-like                         |
| 22 | gene41597 | XM_021807803.1    | 864.198   | metallothionein-like protein type 3            |
| 23 | gene41597 | XM_021807803.1    | 791.620   | elicitor-responsive protein 3-like             |
| 24 | gene41597 | XM_021807803.1    | 791.620   | elicitor-responsive protein 3-like             |
Figures

Figure 1

The cultures during H. brasiliensis SE. EC: embryogenic calluses; PE: primary embryo; CE: cotyledonaryembryo; MCE: mature cotyledonaryembryo; AE: abnormal embryo; WAE: witheredabnormal embryo
Figure 2

The COG assignments of assembled unigenes. Out of 36937 de novo assembled unigenes, 13421 were assigned to 25COG categories GO annotation of assembled unigenes by Blast2GO during H. brasiliensis SE.
Figure 3

Statistical analysis of the DEGsduring SE stages. a The venn diagram of expressed genes in four developmental stages. b The venn diagram of expressed genes in PE vs. AE, PE vs. CE and CE vs. AE. EC: embryogenic calluses; PE: primary embryo; CE: cotyledonaryembryo; AE: abnormal embryo; MCE: mature cotyledonaryembryo; WAE: withered abnormal embryo
Figure 4

The number of up- or down-regulated DEGs in EC vs. PE, PE vs. CE, CE vs. MCE, PE vs. AE, AE vs. WAE, CE vs. AE, MCE vs. WAE. EC: embryogenic calluses; PE: primary embryo; CE: cotyledonary embryo; AE: abnormal embryo; MCE: mature cotyledonary embryo; WAE: withered abnormal embryo.
Figure 5

Molecular functions and biological processes of DEGs in CE vs. AE (a) and MCE vs. WAE (b) based on gene ontology categories. CE: cotyledonary embryo; AE: abnormal embryo; MCE: mature cotyledonary embryo; WAE: withered abnormal embryo
Figure 6

KEGG annotation of DEGs in CE vs. AE (a) and MCE vs. WAE(b) based on gene ontology categories. CE: cotyledonary embryo; AE: abnormal embryo; MCE: mature cotyledonary embryo; WAE: withered abnormal embryo
Figure 7

Heatmap of the differentially expressed genes in hormonal signaling transduction and putative pathway for AUX signaling. Heatmap indicate the gene expression level by Log2[FPKM] with a rainbow color scale. CE: cotyledonary embryo; AE: abnormal embryo; MCE: mature cotyledonary embryo; WAE: withered abnormal embryo
Figure 8

Analysis of the differentially expressed genes in CE, AE, MCE and WAE. (a) Heatmap of the differentially expressed TFs and SE-related genes. (b) Heatmap of the differentially expressed in histone modifications. Heatmap indicate the gene expression level by Log2[FPKM] with a rainbow color scale. CE: cotyledonary embryo; AE: abnormal embryo; MCE: mature cotyledonary embryo; WAE: withered abnormal embryo.
Figure 9

qRT-PCR verification of the selected DEGs involved in EC, PE, CE, AE, MCE, and WAE. The H. brasiiliensisDEGs selected on the basis of their annotation. The data of polyline derived from FPKM of each gene. The 2-ΔΔCt method was used to calculate the relative expression levels of genes. The statistical differences were analyzed by ANOVA (One-way analysis of variance) based on Fisher’s LSD (P < 0.05 and P < 0.01). EC: embryogenic calluses; PE: primary embryo; CE: cotyledonary embryo; AE: abnormal embryo; MCE: mature cotyledonary embryo; WAE: withered abnormal embryo.

Supplementary Files

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

- Fig.S5.pdf
- Fig.S4.pdf
• FigureS3.tif
• FigureS2.tif
• FigureS1.tif
• TableS4.docx
• TableS3.xlsx
• TableS2.xlsx
• TableS1.xlsx