



OPEN Transcriptome analysis revealed enrichment pathways and regulation of gene expression associated with somatic embryogenesis in *Camellia sinensis*

Hao-Zhen Li^{1,6}, Hui Wu^{3,6}, Kang-Kang Song¹, Hui-Hui Zhao⁴, Xiao-Yan Tang⁵, Xiao-Hua Zhang¹, Di Wang¹, Shao-Lin Dong¹, Feng Liu², Jun Wang², Zhong-Cong Li², Long Yang¹✉ & Qin-Zeng Xiang²✉

The high frequency, stable somatic embryo system of tea has still not been established due to the limitations of its own characteristics and therefore severely restricts the genetic research and breeding process of tea plants. In this study, the transcriptome was used to illustrate the mechanisms of gene expression regulation in the somatic embryogenesis of tea plants. The number of DEGs for the (IS intermediate stage)_PS (preliminary stage), ES (embryoid stage)_IS and ES_PS stages were 109, 2848 and 1697, respectively. The enrichment analysis showed that carbohydrate metabolic processes were considerably enriched at the ES_IS stage and performed a key role in somatic embryogenesis, while enhanced light capture in photosystem I could provide the material basis for carbohydrates. The pathway analysis showed that the enriched pathways in IS_PS process were far less than those in ES_IS or ES_PS, and the photosynthesis and photosynthetic antenna protein pathway of DEGs in ES_IS or ES_PS stage were notably enriched and up-regulated. The key photosynthesis and photosynthesis antenna protein pathways and the Lhcb1 gene were discovered in tea plants somatic embryogenesis. These results were of great significance to clarify the mechanism of somatic embryogenesis and the breeding research of tea plants.

Somatic embryogenesis is a mode of stimulating plant cell totipotency¹ and is considered to be the most efficient morphogenetic pathway for plant reproduction². Gymnosperm embryogenesis, a significant breakthrough in plant tissue culture during the late twentieth century^{3,4}, offers numerous advantages over organ differentiation as a method of plant regeneration. These advantages include a high number of embryos, rapid development, structural integrity, and a high rate of regeneration⁵⁻⁷. As a result, gymnosperm embryogenesis has found extensive applications in various fields, serving as a reliable and efficient plant regeneration system and an ideal recipient system for genetic transformation^{8,9}.

The molecular biology and bioinformatics aspects of plant somatic embryogenesis has been extensively explored in order to elucidate the mechanisms involved. The formation of somatic embryos in sweet pepper (*Capsicum annuum*) was stimulated with the developmental regulator *PLETHORA* (*PLT5*) by in vitro tissue culture¹⁰. In the model plant *Arabidopsis thaliana*, the somatic embryo had been studied to originate from the leaf primordium next to the apical meristem of the stem, and growth hormone treatment could efficiently induce embryogenic responses in explants cultured in vitro^{11,12}. Transcriptome analysis was used to address the low reproductive rate of peony (*Paeonia ostii*) and revealed that genes determining cell fate and cell division dominated the formation of somatic embryos in peony¹³. Transcriptomic data from maize (*Zea mays* L.) embryonic

¹College of Plant Protection and Agricultural Big-Data Research Center, Shandong Agricultural University, Tai'an 271018, China. ²College of Horticulture Science and Engineering, Shandong Agricultural University, Tai'an 271018, China. ³AgricultureLife, Gembloux Agro-Bio Tech, Liege University, 5030 Gembloux 2, Belgium. ⁴Ri Zhao Cha Cang Tea Co. Ltd, Rizhao 276800, China. ⁵State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, He'fei 230036, China. ⁶These authors contributed equally: Hao-Zhen Li and Hui Wu. ✉email: lyang@sdau.edu.cn; qzshine@sdau.edu.cn

healing tissue and somatic embryos were studied to reveal hormonal signalling pathways and transcriptional regulation associated with somatic embryogenesis¹⁴.

Camellia sinensis (L.) O. Kuntze is a perennial evergreen plant belonging to the Camelliaceae family¹⁵. It has been cultivated for centuries and its tea leaves are highly prized for their economic value. Tea plants possess characteristic secondary metabolites such as tea polyphenols¹⁶, catechins and caffeine¹⁷, which have many health benefits¹⁸, such as cancer prevention and treatment of cardiovascular diseases¹⁹. Tea plants are self-incompatible and unaffiliated, which lead to their high degree of heterozygosity^{20,21}. Traditional breeding methods for tea have the disadvantages of being labour-intensive, blind and long breeding cycles^{22,23}, while the introduction of modern molecular techniques can speed up the selection process of good tea varieties^{24–27}. The establishment of a high-frequency somatic embryogenesis system for tea is of significant practical importance. It allows for the in vitro conservation of excellent tea germplasm resources and the development of genetic transformation and transgenic systems for tea plants.

Due to the polyphenol-rich nature of tea plants and the characteristics of perennial woody plants¹⁶, the tissue culture process of tea plants suffered from severe browning and poor reproducibility, and the difficulties in obtaining regenerated intact plants^{28–30}. Therefore, the high frequency and stable somatic embryogenesis system of tea plants has not been established so far, which greatly affects and limits the development of cell engineering and genetic engineering of tea plants and other related researches. At present, more studies on somatic embryogenesis in horticultural plants have been conducted, and the technical methods are more mature^{31–33}, but the researches on somatic embryogenesis in tea were less systematic. The majority of current studies have focused on determining the optimal conditions for somatic embryogenesis, including the ratio of plant growth regulators and the combination of basal medium^{34–36}. However, only a few studies have investigated the changes in morphological structure and endogenous hormones that occur during somatic embryogenesis.

Despite its perennial nature, tea plants still confronted difficulties in obtaining intact plants during their tissue culture process, which affected tea plants propagation. In order to study the pathways or factors that played key roles in tea plants somatic embryogenesis. In this study, transcriptomics was used to study the changes in the functions of genes of metabolic pathways and their regulation during somatic embryogenesis in order to gain an understanding of the mechanism of somatic embryogenesis in tea, which will help to further elucidate the molecular mechanisms controlling somatic embryogenesis in tea and provide ideas for reproduction and genetic engineering of tea.

Results

Quality assessment of transcriptome sequencing for tea embryogenesis

61.353G of raw data were obtained from transcriptome sequencing of nine tea samples (Table S1). The raw data for all nine samples were more than 96% for Q20 and 91% for Q30. The raw reads GC content of the nine samples ranged from 45.115 to 47.855%. The clean data mapping rates to the “shuchazao” genome were 90.59–93.78%. Non-splice reads and splice ratio reads were 51.1–55.42% and 26.53–33.12%, respectively. Tea embryogenesis transcriptome sequencing data with high quality and mapping ratios were suited for RNA analysis. Analysis of the correlation of gene abundance between samples showed that the Spearman correlation coefficients between samples were all greater than 0.81 (Fig. S1).

Differential gene expression in different tea embryogenesis stages

27,765, 25,984, 26,877 genes were expressed in the stages of PS, IS and ES, respectively. 22,045 genes were co-expressed and 2502, 1329 and 2577 genes were uniquely expressed in PS, IS and ES, respectively (Fig. 1A). The numbers of DEGs for IS_PS, ES_IS, and ES_PS were 109, 2848, and 1697, respectively (Fig. 1B). IS_PS, ES_IS and ES_PS up-regulated DEGs were 32, 1721 and 960 and down-regulated DEGs were 77, 1127 and 737, respectively. The results of the cluster analysis showed that the differences in gene expression in ES_IS were significantly greater than those in IS_PS (Fig. 1C).

Functional studies on DEGs during tea embryogenesis stages

GO annotation was used to study the DEGs function during tea embryogenesis (Fig. 2, Table S2). The main functions of the up-regulated DEGs in ES_PS were mainly enriched in photosynthesis, oxidoreductase activity, protein kinase activity, photosystem, integral component of membrane, thylakoid and chloroplast, and the number of related up-regulated genes were all greater than 39. The main functions of the down-regulated DEGs in ES_PS were mainly enriched in defense response, lipid metabolic process, ADP binding, active transmembrane transporter activity and integral component of membrane, and the number of related down-regulated genes were all greater than 17.

ES_IS was the main process of ES_PS differential, and ES_IS significantly up-regulated DEGs functions were similar to ES_PS but with increased function of carbohydrate metabolic process (Table S2). However, down-regulated DEGs functions had significant differences, and these functions were significantly up-regulated in IS_PS. The main functions of the down-regulated DEGs in ES_IS were mainly enriched in monooxygenase activity, heme binding, transferase activity, transferring hexosyl groups, iron ion binding, oxidoreductase activity, DNA binding, nucleus, intracellular membrane-bounded organelle, membrane-bounded organelle, oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, and the number of related down-regulated genes were all greater than 28. IS_PS had a significant down-regulation of a large number of DEGs enrichment functions and a non-significant up-regulation. Molecular and cellular component functions in IS_PS were performed as down-regulation, including adenylyl ribonucleotide binding, anion binding, protein kinase activity, cell periphery, membrane, integral component of membrane, intrinsic component of membrane.

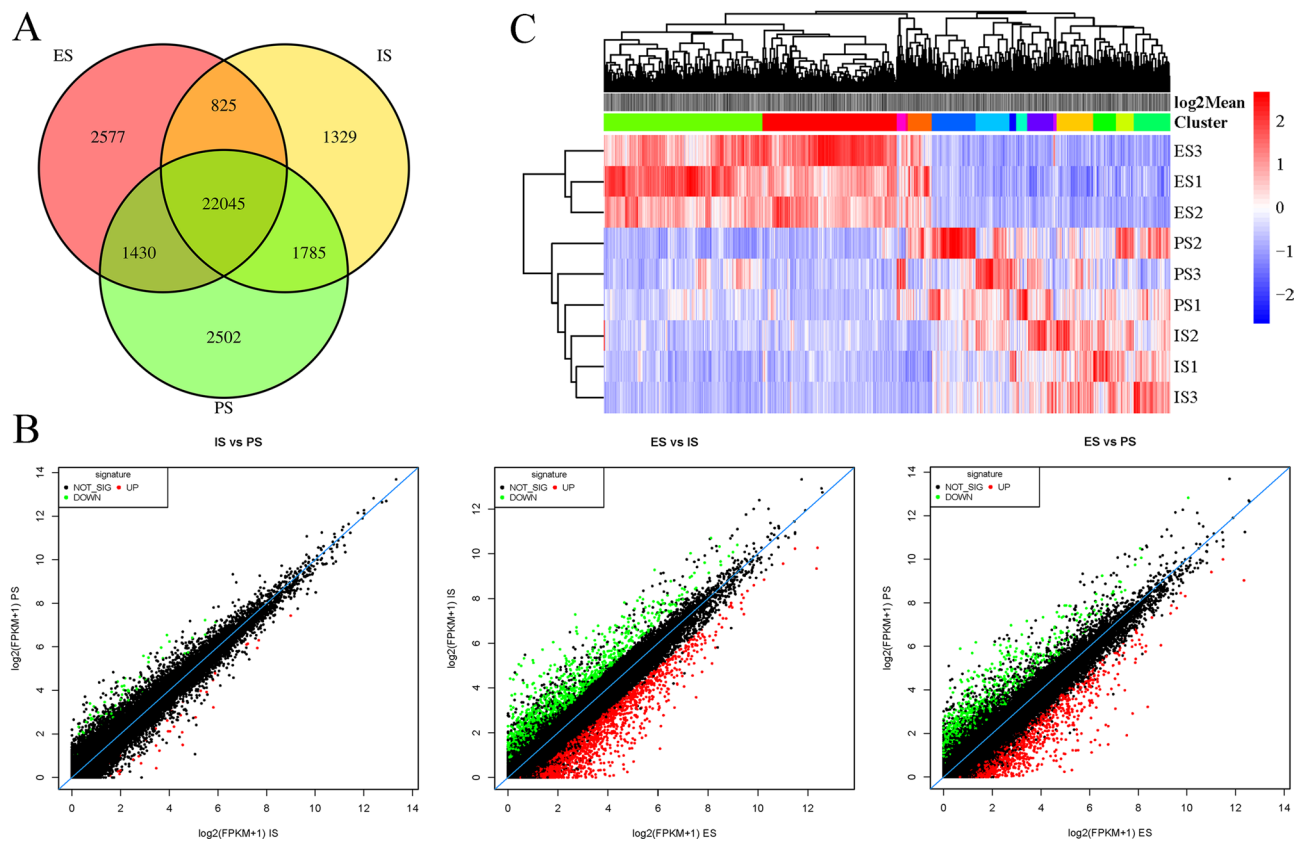


Figure 1. Correlation analyses of gene expression in the somatic embryogenesis stages of tea. (A) Venn diagram of gene expression at different stages of occurrence. (B) Scatter plot of DEGs in different stages. (C) The DEGs cluster heat map between different stages. PS preliminary stage, IS intermediate stage, ES embryoid stage.

Pathway studies on DEGs during tea embryogenesis stages

KEGG database was used to analyze tea embryogenesis DEGs pathway annotation (Fig. 3). IS_PS DEGs were mainly enriched in Vitamin B6 metabolism, Pentose and glucuronate interconversions, and Photosynthesis—antenna proteins, but the enrichment abundance was much lower than that of ES_IS and ES_PS (Fig. 3A). ES_IS DEGs were mainly enriched in photosynthesis, photosynthesis-antenna proteins, metabolism of xenobiotics by cytochrome P450, glucosinolate biosynthesis and glutathione metabolism (Fig. 3B). The abundance of ES_PS DEGs enrichment pathways were similar to those of ES_IS, with the highest abundance being photosynthesis-antenna proteins and photosynthesis (Fig. 3C).

GSEA analysis was used to further identify the functional pathways for embryogenesis. Ribosome and SNARE interactions in vesicular transport pathways genes expression were all significantly upregulated during IS_PS. During ES_PS and ES_IS, photosynthesis and photosynthesis-antenna proteins pathways genes expression were significantly up-regulated. In addition, ribosome pathway genes were significantly up-regulated during ES_PS, steroid biosynthesis and NF-kappa B signaling pathway genes were significantly up-regulated during ES_IS.

There were 25 transcription factors in the Photosynthesis-antenna proteins pathway, 14 of which were significantly differentially expressed in the ES_PS process (Fig. 4). All 14 DEGs were significantly up-regulated and all $\log_2\text{FC}$ were greater than 2.57. Nine of the 12 light-harvesting chlorophyll protein complexes were significantly up-regulated and three were not significantly differentially expressed during ES_PS. The five *Lhcb1* genes (*LOC114297307*, *LOC114270379*, *LOC114270380*, *LOC114299811*, *LOC114299812*) were significantly up-regulated during embryogenesis and all $\log_2\text{FC}$ were greater than 4.59.

Discussion

Camellia sinensis is an essential cash crop³⁷, but researches into the breeding and propagation of superior tea varieties has been relatively slow due to the limitations of its own characteristics³⁸. Somatic embryogenesis is a key factor in plant regeneration, and tea somatic embryogenesis occurs in both direct and indirect ways. Indirect occurrence usually takes cotyledons, shoot tips or petals as explants to induce somatic embryos. The tea somatic embryos were classified into three states according to the morphology of the cotyledon explants: the preliminary stage, the intermediate transition and the somatic embryo state. Transcriptomics techniques have been widely used in the field of plants and have played a crucial role in plant gymnosperm embryogenesis^{39–42}. In this study, transcriptomics was used to investigate the mechanisms of gene expression regulation during somatic embryogenesis in tea plants.

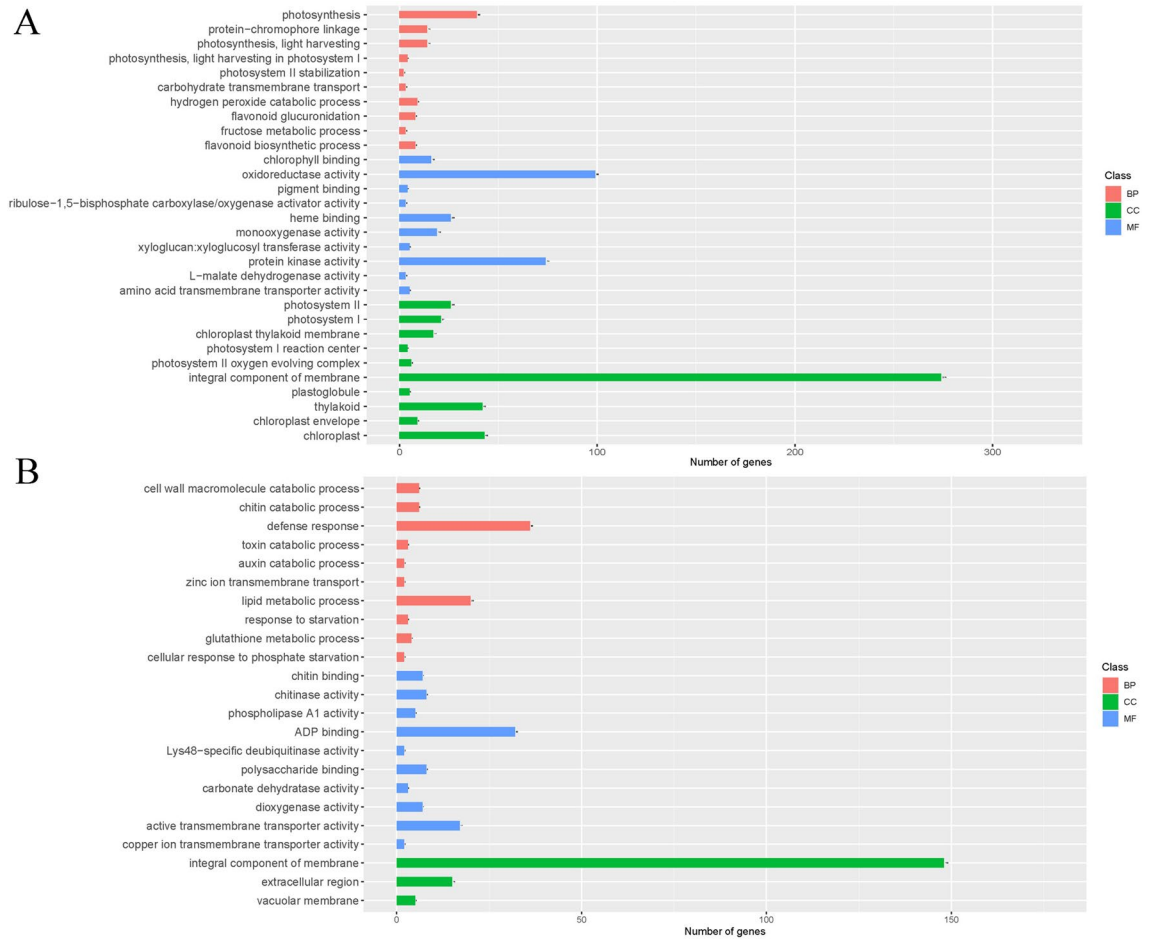


Figure 2. The function analysis of GO annotation for ES_PS stage. **(A)** Up-regulated genes enrichment analysis results. **(B)** Down-regulated genes enrichment analysis results.

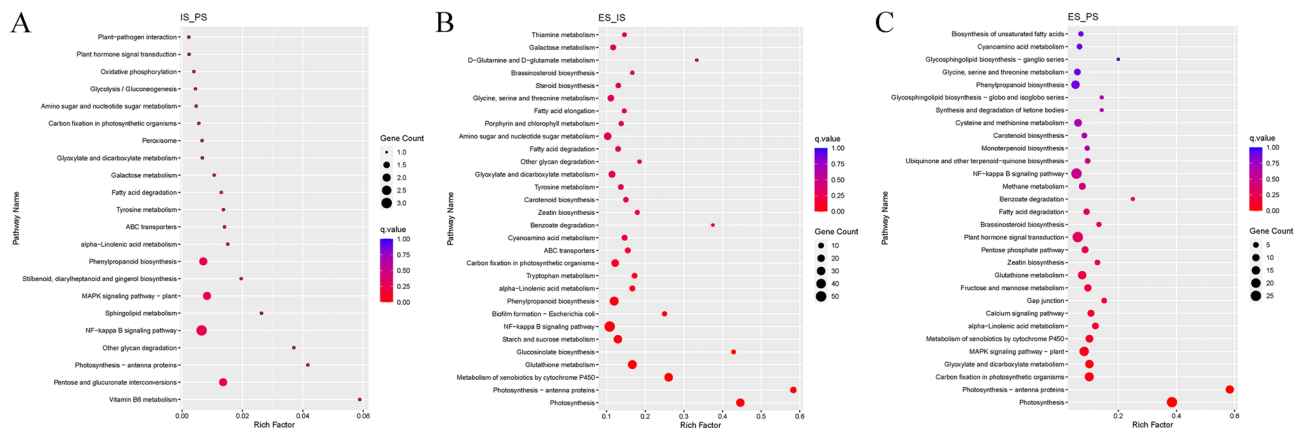


Figure 3. The enrichment analysis of KEGG pathway for the somatic embryogenesis stages of tea. **(A)** The pathway result for IP_PS stage. **(B)** The pathway result for ES_IS stage. **(C)** The pathway result for ES_PS stage.

As tea plant somatic embryogenesis was autotoxic, resulting in its inability to produce regenerating plants, key segments of the embryogenesis process were explored in stages. Transcriptome analysis showed that the numbers of DEGs in ES_IS stage were much larger than that in IS_PS and ES_PS, and the up-regulated DEGs in ES_IS stage was much larger than the down-regulated DEGs. Genes related to growth, development and metabolism were mostly found in the up-regulated DEGs. At the same time, the results of cluster thermogram also showed that ES_IS stage played a key role in the process of tea somatic embryogenesis compared with IS_PS. In addition, unlike in tea plants, the vast majority of differentially expressed genes in maize somatic embryogenesis were

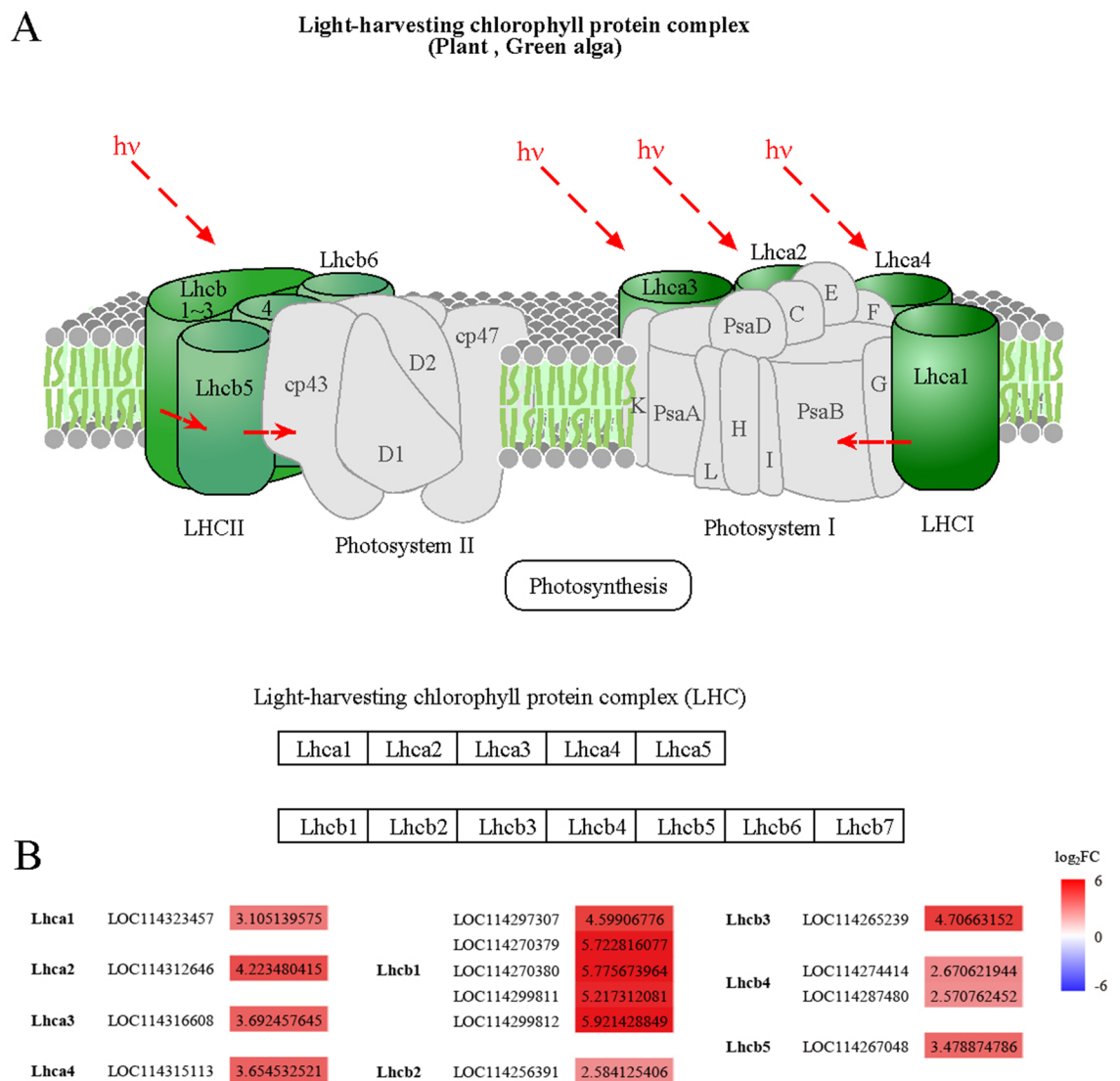


Figure 4. Gene expression regulates photosynthesis. (A) Light-harvesting chlorophyll protein complex model. (B) Expression of genes involved in light-harvesting chlorophyll protein complex.

clustered in the immature embryo to the embryogenesis healing tissue stage, with a tendency for up-regulation of expression during the dedifferentiation process¹⁴.

Based on GO enrichment analysis, DEGs upregulated at the ES_PS stage were significantly enriched in cellular components such as overall membrane composition, molecularly functional protein kinase activity and photosynthesis in biological processes. For example, SNF1/AMPK-associated protein kinases were linked to downstream gene expression, physiology and development through signalling⁴³; Mitogen-activated protein kinases (MAPKs) were organised into complex networks for signalling and thus regulating plant growth and development⁴⁴. In contrast, the ES_IS stage, as the main process of embryogenesis in tea somatic cells, had similar functions to other up-regulated DEGs, except for the carbohydrate metabolic process. Carbohydrates serve as the main component of cellular structure and could provide the main energy for organism development⁴⁵, as well as modifying lipids and proteins, altering their structure and functions. This might also account for the significant differences between IS_PS and ES_IS. The DEGs that were down-regulated during ES_IS were significantly different from those that occur throughout and, interestingly, these functions were significantly up-regulated in IS_PS. Light capture in cellular component-enriched functions such as photosystem I and biological processes photosystem I could promote carbohydrate biosynthesis⁴⁶ and might provide the material basis for carbohydrate metabolic processes in ES_IS. The pathway including carbohydrate metabolism was detected to maintain embryogenesis potential in *Larix kaempferi* (Lamb.) Carr⁴⁷. Carbohydrate and metabolic pathways were recognised as representative overexpression pathways in early embryogenesis of maritime pine and were valuable resources to further support the improvement of trophic reproduction with this specie⁴⁸. Similar to tea plants, crucial amino acid biosynthetic pathways were identified in conifer embryonic development⁴⁹, suggesting that metabolism might be indispensable in somatic embryo formation. Furthermore, in cotton somatic embryogenesis, genes

were significantly enriched in metabolic pathways and secondary metabolite biosynthesis⁵⁰. Genes involved in somatic embryogenesis in different species regulated embryonic development by enriching in different pathways.

Pathway enrichment analysis of DEGs showed that far fewer pathways were enriched during IS_PS than ES_IS or ES_PS, which also indicated that ES_IS was a critical stage in tea somatic embryogenesis. ES_IS and ES_PS shared similar levels of enrichment, with photosynthesis and photosynthetic antennae proteins being the most enriched. Photosynthesis and photosynthetic antenna protein pathway gene expression were also significantly up-regulated in the GSEA analysis of ES_IS and ES_PS, indicating that photosynthesis and its antenna protein pathway had an important role in tea somatic embryogenesis. Antenna proteins performed a key regulatory role in light capture in photosynthesis⁵¹ and had the role in protecting plants from high light damage and regulating the capture of solar energy by plants and its transfer to reaction centres^{52,53}. Light contributed to the early development of the somatic cell embryo, and the significant enrichment of photosynthesis in ginkgo somatic cell embryo formation promoted the induction of the cotyledonary somatic cell embryo⁵⁴. Knockdown of five genes of *Lhcb1* in *Arabidopsis* resulted in chlorophyll loss and delayed growth⁵⁵. Down-regulation of *Lhcb1*, *Lhcb3* and *Lhcb5* genes expression levels in cucumber affected its photosynthesis⁵⁶. In contrast, significant up-regulation of the five *Lhcb1* genes during embryogenesis might contribute to their light capture and photosynthesis. To date, few genes involved in embryogenesis have been characterised in tea plants, and their expression was crucial to acquisition in embryogenic competence and expression in somatic embryogenesis. In some cases, the expression of these genes might determine somatic cell fate changes. Therefore, for the discovered *Lhcb1* gene could be considered as a biomarker for somatic embryogenesis in tea plants, which could be validated for biological function and then clonally characterised to advance tea breeding and propagation strategies. Transcription factors specifically bind to cis-elements of target gene promoter to regulate gene transcription⁵⁷, and all transcription factors in the photosynthetic antennae protein pathway were significantly up-regulated during ES_PS. The mechanisms of transcriptional regulation of these transcription factors in tea somatic embryogenesis need to be further revealed. Photosynthesis acts as an irreplaceable role throughout the plant life cycle⁵⁸, providing continuous energy income for metabolism⁵⁹ and growth and development. The significant enrichment of the photosynthetic system pointed to an essential role for its regulation in somatic embryogenesis. It was worthwhile to further investigate how these significantly related differential genes and enriched pathways regulated tea somatic cell embryogenesis, and the transcriptional regulatory mechanisms of the transcription factors need to be further unravelled. Further elucidation regarding the precise roles of these pathways will facilitate the molecular understanding of tea somatic cell embryogenesis and the development of reproductive strategies. In addition, future studies should not be limited to a single omics, proteomics that identified specific proteins associated with somatic embryogenesis and development, and epigenetics techniques to serve the researches.

Materials and methods

Sample collection

The three occurrence stages used in this study to form tea somatic embryos were provided by Shandong Agricultural University. *Camellia sinensis* cv. Jinxuan tea seeds were used as the primary material for the somatic embryogenesis process, where cotyledon cuts were induced into the somatic embryo state, and all materials used in the experiments were taken from the same tea seeds. Part of the cotyledons were induced directly into the somatic embryo, a portion of the cotyledons were in the elevated state, and a percentage of the cotyledons remained in the initial state during the 6 months of induction. The outermost layer of the cotyledons was cut for the PS, the IS was cut for the uplifted portion of the leaf, and the ES was selected for the somatic embryo that had already been induced. Materials from all three states of the same tea seed were sampled with a scalpel and immediately fixed in liquid nitrogen and stored at $-80\text{ }^{\circ}\text{C}$ ⁶⁰. Transcriptome sequencing was performed on three groups of material and three replicates were used to undertake sequencing.

Transcriptome sequencing and assembly

RNA with Poly-A structure in eukaryotic total RNA was enriched using the TIANSeq mRNA capture kit (TIANGEN Biotech). Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific) was used to determine the concentration of RNA samples and assess their purity⁶¹. Agilent 2100 Bioanalyser and 2100 RNA nano 6000 Assay Kit (Agilent Technologies) were used to assess the integrity of the RNA samples⁶². Using the captured RNA as the starting sample, TIANSeq Fast RNA Library Kit (Illumina) was used to construct the transcriptome sequencing libraries. The transcriptome sequencing library was constructed through RNA randomly fragmentation, cDNA strand 1/strand 2 synthesis, end repair, A-tailing, ligation of sequencing adapters, size selection and library PCR enrichment. Library concentration was first quantified using Qubit 2.0 fluorometer (Life Technologies), and then diluted to 1 ng/μl before checking insert size on an Agilent 2100 and quantifying to greater accuracy by quantitative PCR (Q-PCR) (library activity > 2 nM). The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina sequencing platform and 150 bp paired-end reads were generated.

Identification and analysis of differentially expressed genes

The raw data in fastq format (raw reads) were first processed by an internal perl script. In this step, clean data (clean reads) were obtained by removing reads containing articulators and trimming low quality bases with Trimmomatic⁶³. Quality control was performed using FASTQC⁶⁴. Clean data were also calculated for Q20, Q30 and GC content. All downstream analyses were based on high quality clean data. Hisat2 v2.0.5⁶⁵ was used to construct the "Shuchazao" reference genome⁶⁶ (<http://tpia.teaplant.org/download.html>, accessed 5 April 2023)

indexed on 5 April 2023. The paired-end clean reads were aligned with the reference genome and the mapping information was calculated.

Differential expression analysis of tea somatic embryo samples at each stage was performed using the DESeq2 R package (1.16.1)⁶⁷. Genes identified by DESeq2 with a P value < 0.05 were classified as differentially expressed genes. DEGs were analyzed separately by Gene Ontology (GO) (<http://www.Gene.Ontology.org>, accessed 5 April 2023) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (www.kegg.jp/kegg/pathway.html, accessed 5 April 2023) were annotated and enriched to obtain functional and pathway results for DEGs^{68,69}. ClusterProfiler⁷⁰ and ggplot2⁷¹ R packages were used for GO and KEGG enrichment analysis and visualization.

Data availability

Data from the current study are available from the corresponding author upon reasonable request.

Received: 2 June 2023; Accepted: 22 September 2023

Published online: 24 September 2023

References

- Sivanesan, I. *et al.* Genetic and epigenetic modes of the regulation of somatic embryogenesis: A review. *Biol. Future* **73**, 259–277. <https://doi.org/10.1007/s42977-022-00126-3> (2022).
- Long, Y., Yang, Y., Pan, G. & Shen, Y. New insights into tissue culture plant-regeneration mechanisms. *Front. Plant Sci.* **13**, 926752. <https://doi.org/10.3389/fpls.2022.926752> (2022).
- Feher, A. Callus, Dedifferentiation, totipotency, somatic embryogenesis: What these terms mean in the era of molecular plant biology? *Front. Plant Sci.* **10**, 536. <https://doi.org/10.3389/fpls.2019.00536> (2019).
- Guan, Y., Li, S. G., Fan, X. F. & Su, Z. H. Application of somatic embryogenesis in woody plants. *Front. Plant Sci.* **7**, 938. <https://doi.org/10.3389/fpls.2016.00938> (2016).
- Bai, Y. *et al.* Construction of ceRNA networks at different stages of somatic embryogenesis in garlic. *Int. J. Mol. Sci.* **24**, 5311. <https://doi.org/10.3390/ijms24065311> (2023).
- Egertsdotter, U., Ahmad, I. & Clapham, D. Automation and scale up of somatic embryogenesis for commercial plant production, with emphasis on conifers. *Front. Plant Sci.* **10**, 109. <https://doi.org/10.3389/fpls.2019.00109> (2019).
- Smertenko, A. & Bozhkov, P. V. Somatic embryogenesis: Life and death processes during apical-basal patterning. *J. Exp. Bot.* **65**, 1343–1360. <https://doi.org/10.1093/jxb/eru005> (2014).
- Heringer, A. S., Santa-Catarina, C. & Silveira, V. Insights from proteomic studies into plant somatic embryogenesis. *Proteomics* **18**, e1700265. <https://doi.org/10.1002/pmic.201700265> (2018).
- Raza, G., Singh, M. B. & Bhalla, P. L. Somatic embryogenesis and plant regeneration from commercial soybean cultivars. *Plants* **9**, 10038. <https://doi.org/10.3390/plants910038> (2019).
- Lian, Z. *et al.* Application of developmental regulators to improve in planta or in vitro transformation in plants. *Plant Biotechnol. J.* **20**, 1622–1635. <https://doi.org/10.1111/pbi.13837> (2022).
- Wojcikowska, B., Wojcik, A. M. & Gaj, M. D. Epigenetic regulation of auxin-induced somatic embryogenesis in plants. *Int. J. Mol. Sci.* **21**, 307. <https://doi.org/10.3390/ijms21072307> (2020).
- Kadokura, S., Sugimoto, K., Tarr, P., Suzuki, T. & Matsunaga, S. Characterization of somatic embryogenesis initiated from the Arabidopsis shoot apex. *Dev. Biol.* **442**, 13–27. <https://doi.org/10.1016/j.ydbio.2018.04.023> (2018).
- Ci, H. *et al.* A comparative transcriptome analysis reveals the molecular mechanisms that underlie somatic embryogenesis in *Peaonia ostii* “Fengdan”. *Int. J. Mol. Sci.* **23**, 595. <https://doi.org/10.3390/ijms231810595> (2022).
- Ding, M. *et al.* Transcriptomic analysis reveals somatic embryogenesis-associated signaling pathways and gene expression regulation in maize (*Zea mays* L.). *Plant Mol. Biol.* **104**, 647–663. <https://doi.org/10.1007/s11103-020-01066-z> (2020).
- Shen, J. *et al.* Dynamic changes in metabolic and lipidomic profiles of tea plants during drought stress and re-watering. *Front. Plant Sci.* **13**, 978531. <https://doi.org/10.3389/fpls.2022.978531> (2022).
- Khan, N. & Mukhtar, H. Tea polyphenols in promotion of human health. *Nutrients* **11**, 39. <https://doi.org/10.3390/nu11010039> (2018).
- Kochman, J., Jakubczyk, K., Antoniewicz, J., Mruk, H. & Janda, K. Health benefits and chemical composition of matcha green tea: A review. *Molecules* **26**, 0085. <https://doi.org/10.3390/molecules26010085> (2020).
- Zhang, L. *et al.* Chemistry and biological activities of processed *Camellia sinensis* teas: A comprehensive review. *Compr. Rev. Food Sci. Food Saf.* **18**, 1474–1495. <https://doi.org/10.1111/1541-4337.12479> (2019).
- Samanta, S. Potential bioactive components and health promotional benefits of tea (*Camellia sinensis*). *J. Am. Nutr. Assoc.* **41**, 65–93. <https://doi.org/10.1080/07315724.2020.1827082> (2022).
- Wei, C. *et al.* Draft genome sequence of *Camellia sinensis* var. *sinensis* provides insights into the evolution of the tea genome and tea quality. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E4151–E4158. <https://doi.org/10.1073/pnas.1719622115> (2018).
- Wang, X. *et al.* Population sequencing enhances understanding of tea plant evolution. *Nat. Commun.* **11**, 4447. <https://doi.org/10.1038/s41467-020-18228-8> (2020).
- Seth, R. *et al.* Global transcriptional insights of pollen-pistil interactions commencing self-incompatibility and fertilization in tea [*Camellia sinensis* (L.) O. Kuntze]. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms20030539> (2019).
- Mukhopadhyay, M., Mondal, T. K. & Chand, P. K. Biotechnological advances in tea (*Camellia sinensis* [L.] O. Kuntze): A review. *Plant Cell Rep.* **35**, 255–287. <https://doi.org/10.1007/s00299-015-1884-8> (2016).
- Ma, Q. *et al.* Transcriptomic analysis between self- and cross-pollinated pistils of tea plants (*Camellia sinensis*). *BMC Genom.* **19**, 289. <https://doi.org/10.1186/s12864-018-4674-1> (2018).
- Zhang, Z. B. *et al.* Understanding the origin and evolution of tea (*Camellia sinensis* [L.]): Genomic advances in tea. *J. Mol. Evol.* **91**, 156–168. <https://doi.org/10.1007/s00239-023-10099-z> (2023).
- Li, J. W. *et al.* Molecular markers in tea plant (*Camellia sinensis*): Applications to evolution, genetic identification, and molecular breeding. *Plant Physiol. Biochem.* **198**, 107704. <https://doi.org/10.1016/j.plaphy.2023.107704> (2023).
- Zhang, X. *et al.* Haplotype-resolved genome assembly provides insights into evolutionary history of the tea plant *Camellia sinensis*. *Nat. Genet.* **53**, 1250–1259. <https://doi.org/10.1038/s41588-021-00895-y> (2021).
- Gao, Y. *et al.* MicroRNAs and their targeted genes associated with phase changes of stem explants during tissue culture of tea plant. *Sci. Rep.* **9**, 20239. <https://doi.org/10.1038/s41598-019-56686-3> (2019).
- Mondal, T., Bhattacharya, A., Ahuja, P. & Chand, P. Transgenic tea [*Camellia sinensis* (L.) O. Kuntze cv. Kangra Jat] plants obtained by Agrobacterium-mediated transformation of somatic embryos. *Plant Cell Rep.* **20**, 712–720. <https://doi.org/10.1007/s002990100382> (2001).
- Bano, Z., Rajarathnam, S. & Mohanty, B. D. Somatic embryogenesis in cotyledon cultures of tea (*Thea sinensis* L.). *J. Hortic. Sci.* **66**, 465–470. <https://doi.org/10.1080/00221589.1991.11516175> (2015).

31. Wang, G., Xu, C., Yan, S. & Xu, B. An efficient somatic embryo liquid culture system for potential use in large-scale and synchronic production of *Anthurium andraeanum* seedlings. *Front. Plant Sci.* **10**, 29. <https://doi.org/10.3389/fpls.2019.00029> (2019).
32. Zhang, J., Gai, M., Li, X., Li, T. & Sun, H. Somatic embryogenesis and direct as well as indirect organogenesis in *Lilium pumilum* DC. Fisch., an endangered ornamental and medicinal plant. *Biosci. Biotechnol. Biochem.* **80**, 1898–1906. <https://doi.org/10.1080/09168451.2016.1194178> (2016).
33. Dehestani-Ardakani, M., Hejazi, M. & Aliabad, K. K. Indirect somatic embryogenesis of purple coneflower (*Echinacea purpurea* (L.) Moench): A medicinal-ornamental plant: Evaluation of antioxidant enzymes activity and histological study. *Mol. Biol. Rep.* **47**, 6621–6633. <https://doi.org/10.1007/s11033-020-05713-y> (2020).
34. Kumari, A. *et al.* Plant growth regulator interactions in physiological processes for controlling plant regeneration and in vitro development of *Tulbaghia simmleri*. *J. Plant Physiol.* **223**, 65–71. <https://doi.org/10.1016/j.jplph.2018.01.005> (2018).
35. Valimaki, S. *et al.* Somatic embryo yield and quality from norway spruce embryogenic tissue proliferated in suspension culture. *Front. Plant. Sci.* **12**, 791549. <https://doi.org/10.3389/fpls.2021.791549> (2021).
36. Li, F., Yao, J., Hu, L., Chen, J. & Shi, J. Multiple methods synergistically promote the synchronization of somatic embryogenesis through suspension culture in the new hybrid between *Pinus elliottii* and *Pinus caribaea*. *Front. Plant Sci.* **13**, 857972. <https://doi.org/10.3389/fpls.2022.857972> (2022).
37. Xia, E. H. *et al.* The tea tree genome provides insights into tea flavor and independent evolution of caffeine biosynthesis. *Mol. Plant* **10**, 866–877. <https://doi.org/10.1016/j.molp.2017.04.002> (2017).
38. Joshi, S., Paul, P., Hartman, J. M. & Perry, S. E. AGL15 promotion of somatic embryogenesis: Role and molecular mechanism. *Front. Plant Sci.* **13**, 861556. <https://doi.org/10.3389/fpls.2022.861556> (2022).
39. Zhao, R. *et al.* Transcriptomic and physiological analysis identifies a gene network module highly associated with brassinosteroid regulation in hybrid sweetgum tissues differing in the capability of somatic embryogenesis. *Hortic. Res.* **9**, 47. <https://doi.org/10.1093/hr/uhab047> (2022).
40. Cui, Y. *et al.* Transcriptomic, metabolomic, and physiological analyses reveal that the culture temperatures modulate the cryotolerance and embryogenicity of developing somatic embryos in *Picea glauca*. *Front. Plant Sci.* **12**, 694229. <https://doi.org/10.3389/fpls.2021.694229> (2021).
41. Awada, R. *et al.* Global transcriptome profiling reveals differential regulatory, metabolic and hormonal networks during somatic embryogenesis in *Coffea arabica*. *BMC Genom.* **24**, 41. <https://doi.org/10.1186/s12864-022-09098-z> (2023).
42. Guo, H. *et al.* 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.* **20**, 2070. <https://doi.org/10.3390/ijms20092070> (2019).
43. Zhu, J. K. Abiotic stress signaling and responses in plants. *Cell* **167**, 313–324. <https://doi.org/10.1016/j.cell.2016.08.029> (2016).
44. Jiang, M. *et al.* Mitogen-activated protein kinase and substrate identification in plant growth and development. *Int. J. Mol. Sci.* **23**, 2744. <https://doi.org/10.3390/ijms23052744> (2022).
45. Naithani, S., Komath, S. S., Nonomura, A. & Govindjee, G. Plant lectins and their many roles: Carbohydrate-binding and beyond. *J. Plant Physiol.* **266**, 153531. <https://doi.org/10.1016/j.jplph.2021.153531> (2021).
46. Lima-Melo, Y., Kilic, M., Aro, E. M. & Gollan, P. J. Photosystem I inhibition, protection and signalling: Knowns and unknowns. *Front. Plant Sci.* **12**, 791124. <https://doi.org/10.3389/fpls.2021.791124> (2021).
47. Wang, J., Zhang, L., Qi, L. & Zhang, S. Integrated transcriptomic and metabolic analyses provide insights into the maintenance of embryogenic potential and the biosynthesis of phenolic acids and flavonoids involving transcription factors in *Larix kaempferi* (Lamb.) Carr. *Front. Plant Sci.* **13**, 6930. <https://doi.org/10.3389/fpls.2022.1056930> (2022).
48. Rodrigues, A. S., De Vega, J. J. & Miguel, C. M. Comprehensive assembly and analysis of the transcriptome of maritime pine developing embryos. *BMC Plant Biol.* **18**, 2. <https://doi.org/10.1186/s12870-018-1564-2> (2018).
49. Ávila, C. *et al.* Identification of metabolic pathways differentially regulated in somatic and zygotic embryos of maritime pine. *Front. Plant Sci.* **13**, 960. <https://doi.org/10.3389/fpls.2022.877960> (2022).
50. Yang, X. *et al.* Transcript profiling reveals complex auxin signalling pathway and transcription regulation involved in dedifferentiation and redifferentiation during somatic embryogenesis in cotton. *BMC Plant Biol.* **12**, 110 (2012).
51. Crepin, A. *et al.* Size and fluorescence properties of algal photosynthetic antenna proteins estimated by microscopy. *Int. J. Mol. Sci.* **23**, 20778. <https://doi.org/10.3390/ijms23020778> (2022).
52. Su, Q. *et al.* Physiological and comparative transcriptome analysis of the response and adaptation mechanism of the photosynthetic function of mulberry (*Morus alba* L.) leaves to flooding stress. *Plant Signal Behav.* **17**, 2094619. <https://doi.org/10.1080/15592324.2022.2094619> (2022).
53. Crepin, A., Cunill-Semanat, E., Kuthanova Trskova, E., Belgio, E. & Kana, R. Antenna protein clustering in vitro unveiled by fluorescence correlation spectroscopy. *Int. J. Mol. Sci.* **22**, 969. <https://doi.org/10.3390/ijms22062969> (2021).
54. Chen, Y. *et al.* Transcriptome and physiological analysis highlights the hormone, phenylpropanoid, and photosynthesis effects on early somatic embryogenesis in *Ginkgo biloba*. *Ind. Crops Prod.* **203**, 7176. <https://doi.org/10.1016/j.indcrop.2023.117176> (2023).
55. Sattari Vayghan, H. *et al.* Photosynthetic light harvesting and thylakoid organization in a CRISPR/Cas9 *Arabidopsis thaliana* LHCB1 knockout mutant. *Front. Plant Sci.* **13**, 833032. <https://doi.org/10.3389/fpls.2022.833032> (2022).
56. Wu, P. *et al.* Unveiling molecular mechanisms of nitric oxide-induced low-temperature tolerance in cucumber by transcriptome profiling. *Int. J. Mol. Sci.* **23**, 5615. <https://doi.org/10.3390/ijms23105615> (2022).
57. Weidemuller, P., Kholmatov, M., Petsalaki, E. & Zaugg, J. B. Transcription factors: Bridge between cell signaling and gene regulation. *Proteomics* **21**, e2000034. <https://doi.org/10.1002/pmic.202000034> (2021).
58. Du, K., Wu, W., Liao, T., Yang, J. & Kang, X. Transcriptome analysis uncovering regulatory networks and hub genes of *Populus* photosynthesis and chlorophyll content. *Genomics* **114**, 110385. <https://doi.org/10.1016/j.ygeno.2022.110385> (2022).
59. Cazzaniga, S. & Ballottari, M. Photosynthesis 2.0. *Int. J. Mol. Sci.* **24**, 4355. <https://doi.org/10.3390/ijms24054355> (2023).
60. Passow, C. N. *et al.* Nonrandom RNAseq gene expression associated with RNAlater and flash freezing storage methods. *Mol. Ecol. Resour.* **19**, 456–464. <https://doi.org/10.1111/1755-0998.12965> (2019).
61. Ferreira, M. R. S., Galvao, A. P. O., Valenca, M. M. & Martins, D. B. G. Storage and purification adaptations for the isolation of total RNA from the dura mater. *Arq. Neuropsiquiatr.* **80**, 1213–1219. <https://doi.org/10.1055/s-0042-1758865> (2022).
62. Soler-Garcia, A. A., De Jesus, A. J., Taylor, K. & Brown, E. W. Differentiation of *Salmonella* strains from the SARA, SARB and SARC reference collections by using three genes PCR-RFLP and the 2100 Agilent bioanalyzer. *Front. Microbiol.* **5**, 417. <https://doi.org/10.3389/fmicb.2014.00417> (2014).
63. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170> (2014).
64. Ward, C. M., To, T. H. & Pederson, S. M. ngsReports: A bioconductor package for managing FastQC reports and other NGS related log files. *Bioinformatics* **36**, 2587–2588. <https://doi.org/10.1093/bioinformatics/btz937> (2020).
65. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915. <https://doi.org/10.1038/s41587-019-0201-4> (2019).
66. Xia, E. *et al.* The reference genome of tea plant and resequencing of 81 diverse accessions provide insights into its genome evolution and adaptation. *Mol. Plant* **13**, 1013–1026. <https://doi.org/10.1016/j.molp.2020.04.010> (2020).
67. McDermaid, A., Monier, B., Zhao, J., Liu, B. & Ma, Q. Interpretation of differential gene expression results of RNA-seq data: Review and integration. *Brief. Bioinform.* **20**, 2044–2054. <https://doi.org/10.1093/bib/bby067> (2019).

68. Gene Ontology, C. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* **32**, 258–261. <https://doi.org/10.1093/nar/gkh036> (2004).
69. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361. <https://doi.org/10.1093/nar/gkw1092> (2017).
70. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287. <https://doi.org/10.1089/omi.2011.0118> (2012).
71. Tyner, S., Briatte, F. & Hofmann, H. Network visualization with ggplot2. *R. J.* **9**, 27 (2017).

Acknowledgements

This work was supported by the Shandong Province Improved Seed Project (2017LZN026) and Special Funds for Local Scientific and Technological Development Guided by the Central Government (YDZX2022123). And we thank Supercomputing Center in Shandong Agricultural University for technical support.

Author contributions

H.L. and H.W. collected the data and performed the experiment. H.L. analyzed the results and wrote the manuscript. K.S., H.Z. and X.T. conceptualized it. X.Z., D.W. and S.D. visualized data. F.L, J.W. and Z.L. checked manuscript. L.Y. and Q.X. supervised the project and provided resources. All authors participated in the revision of the manuscript and approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-43355-9>.

Correspondence and requests for materials should be addressed to L.Y. or Q.-Z.X.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023