

Developmental progression toward embryogenic competence in *Theobroma cacao* L.: an integrated cytohistological and molecular analysis

Liliana Victoria Monsalve Zapata

Universidad de Antioquia

Tatiana Osorio-Montoya

Universidad de Antioquia

Liliana Alexandra Pila Quinga

Universidad Regional Amazónica IKIAM

Aura Ines Urrea Trujillo

Universidad de Antioquia

Laura Michell Carmona Rojas

`laura.carmona@ufrgs.br`

Universidade Federal do Rio Grande do Sul

Research Article

Keywords: Cacao, cytohistologic, embryogenic competence, gene expression, indirect somatic embryogenesis, recalcitrance

Posted Date: October 19th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11210265/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Somatic embryogenesis is a platform for clonal propagation and genetic improvement of cacao, but its application is limited by genotype-dependent recalcitrance and restricted acquisition of embryogenic competence. This study integrated morphological, cytohistological, and gene expression analyses to identify features associated with embryogenic competence of *Theobroma cacao* L. clone CNCH13. Embryogenic callus frequency in petal-derived calli increased from 5–8% at 30 to 28–30% at 60 days, with no new embryogenic callus formed at 90 days. Embryogenic callus was identified by its light brown-yellowish color, granular appearance, and soft consistency. Histological analysis showed a continuum from parenchyma cells to competent cells, embryogenic cell clusters, proembryos, and somatic embryos. Phenolic accumulation was positively associated with embryogenic competence, whereas starch reserves occurred broadly in both embryogenic and non-embryogenic tissues. RT-qPCR across six tissues showed that *BBM*, *ABI3*, *AGL15*, and *IAA20* were expressed in the induction-phase callus (hormone-containing medium) and in the expression-phase callus during the first 30 days of hormone-free culture; their expression declined after 60 days in expression medium, once embryogenic and non-embryogenic calli had become morphologically distinguishable and reached maximum levels in globular embryos. *LEC2* and *WUS* followed a similar pattern at lower expression levels, whereas *PIN1*, *SERK1*, and *SERK2* showed lower or broadly distributed expression across tissues, suggesting roles in additional developmental processes rather than in embryogenic competence. Phenolic accumulation and this early transcriptional signal emerge as the most reliable correlates of embryogenic competence, providing criteria for early selection of embryogenic tissue and improving somatic embryogenesis in recalcitrant cacao genotypes.

Key message

Integrated morphological, cytohistological, and molecular analyses revealed stage-dependent cellular and gene-expression changes that distinguish embryogenic competence from early somatic embryo development in cacao.

Introduction

Somatic embryogenesis (SE) is a unique developmental process in which somatic cells undergo reprogramming and acquire embryogenic competence, allowing them to develop into whole plants through an embryogenic pathway (Fehér 2019). SE has become a powerful tool for the large-scale propagation of valuable clones, germplasm conservation, plant sanitation, and the advancement of breeding programs, as well as for the production of genetically modified plants (Loyola-Vargas and Ochoa-Alejo, 2016). Appropriate stimuli, including specific culture-medium compositions, plant growth regulators (PGRs), stress-related signals, and the activity of key transcription factors (TFs), can trigger SE in a wide range of explant tissues, promoting the formation of multiple somatic embryos. These embryos may arise directly from somatic cells, through direct somatic embryogenesis (DSE), or indirectly through an intervening callus phase, through indirect somatic embryogenesis (ISE) (Yang and Zhang 2010; Horstman et al. 2017).

ISE is generally considered more productive than DSE because the extensive cellular proliferation associated with callus formation can generate large populations of potentially embryogenic cells from a single explant (Ikeuchi et al. 2013). During the early stages of culture, an unorganized mass of vacuolated cells exhibiting varying degrees of compactness may form; however, only a subset of these cells undergoes cell-fate reprogramming and acquires embryogenic competence (Ikeuchi et al. 2013). This transition is controlled by a complex regulatory network. Consequently, numerous studies have focused to integrate cellular and molecular evidence to elucidate the

mechanisms underlying the acquisition of embryogenic competence and the activation of the SE program (Mahdavi-Darvari et al. 2015).

Different molecular studies have identified a broad set of genes involved in the activation of the embryogenic program in somatic cells, including genes encoding transcription factors (TFs) such as BABY BOOM (*BBM*), WUSCHEL (*WUS*), LEAFY COTYLEDON2 (*LEC2*), ABSCISIC ACID INSENSITIVE3 (*ABI3*), and AGAMOUS-LIKE15 (*AGL15*) (Boutillier et al. 2002; Zuo et al. 2002; Wójcikowska et al. 2013; Tang et al. 2017; Joshi et al. 2022). Among these regulators, *BBM* controls cell identity, totipotency, and growth and is expressed during SE. Notably, ectopic expression of *BBM* can induce SE in *A. thaliana* and *B. napus* (Boutillier et al. 2002; Yavuz et al. 2020), while its overexpression increases SE frequency in *T. cacao* (Florez et al. 2015). Interestingly, *BBM* also interacts with the broader embryogenic regulatory network by regulating other TFs, including *ABI3* and *LEC2*. *ABI3*, a regulator expressed during seed development, has been identified as a downstream target of *LEC2* in *A. thaliana* (Mönke et al. 2012; Tang et al. 2017). *LEC2*, in turn, is a master regulator that performs critical functions in both zygotic embryogenesis (ZE) and SE (Wójcikowska et al. 2013). Overexpression of *AtLEC2* in immature Arabidopsis zygotic embryos can induce direct SE in the absence of exogenous auxin, whereas *AtLEC2* expression is significantly higher in embryogenic callus (EC) than in non-embryogenic callus (NEC) (Wójcikowska et al. 2013). Consistent with this findings, Zhang et al. (2014) demonstrated an essential role for *TcLEC2* during both ZE and SE in *T. cacao*. Moreover, *TcLEC2* overexpression significantly increased SE efficiency, and its expression was higher in EC than in NEC during secondary SE. In addition to the *BBM*-*LEC2*-*ABI3* regulatory network, *WUS* is another key regulator of the SE program, with important functions in embryogenesis and in shoot and floral meristem formation (Xiao et al. 2018). Overexpression of *WUS* has been shown to promote the transition from a vegetative to an embryogenic state in *A. thaliana* and *M. acuminata* (Zuo et al. 2002; Arroyo-Herrera et al. 2024).

In addition to these TFs, genes involved in auxin signaling and transport play crucial roles in the transition of somatic cells to an embryogenic state and in subsequent embryo development (Wójcik et al. 2020; Long et al. 2022). Aux/IAA proteins are key regulators of auxin signaling that act through their interactions with auxin response factors (ARFs) (Liu et al. 2015). Among this family, *IAA20* and *IAA30* have been shown to play physiological roles in maintaining the root stem cell niche; in addition, *IAA30* has been identified as a target of the *AGL15* transcription factor, supporting its potential involvement in SE (Sato and Yamamoto 2008). In parallel with auxin signaling, polar auxin transport establishes local auxin gradients that are essential for plant growth and development. Among the proteins involved in this process, PIN-FORMED1 (*PIN1*) is an auxin efflux carrier that contributes to embryonic, shoot, and root development. Notably, during SE induction from Arabidopsis leaf explants, *AtPIN1* is differentially expressed across somatic embryo developmental stages (Pierroz 2023).

Beyond transcriptional and hormone-mediated regulation, SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASES (SERKs) are membrane-localized proteins involved in signal transduction during SE (Mahdavi-Darvari et al. 2015). In particular, *SERK1* has been recognized as a molecular marker associated with the acquisition of embryogenic competence in *Daucus carota* (Schmidt et al. 1997). An association between *SERK1* expression and SE has subsequently been reported in several other plant species, including *Momordica charantia*, *Medicago truncatula*, *Theobroma nigrescens*, *Passiflora edulis*, and *T. cacao* (De Oliveira Santos et al. 2005; Nolan et al. 2009; Talapatra et al. 2014; Pilarska et al. 2016; Rocha et al. 2016).

Across plant species, somatic embryogenesis and its underlying gene regulatory networks share common cellular and genetic features. However, the regulation of this process can vary substantially among species and genotypes and remains poorly understood in several agriculturally important crops, including *Theobroma cacao* L. Cacao

provides the raw material for the multibillion-dollar chocolate industry and is a major source of income for approximately six million smallholder farmers worldwide (Wickramasuriya and Dunwell 2018). Beyond its use in food production, the high polyphenol content of cocoa products has attracted growing interest from the pharmaceutical and cosmetic industries, while global demand continues to increase (Wickramasuriya and Dunwell 2018; Voora et al. 2019). However, the perennial growth habit and long reproductive cycle of cacao delay conventional breeding. Moreover, conventional propagation by cuttings, grafting, or seeds is constrained by limited multiplication capacity or genetic variability, thereby restricting the availability of high-quality planting material and, ultimately, farm productivity. In this context, SE offers a promising platform for the rapid propagation and genetic improvement of selected cacao genotypes. Nevertheless, its broader application remains constrained by the highly variable and often recalcitrant embryogenic response characteristic of cacao germplasm, with many elite cultivars exhibiting particularly low embryogenic competence (Maximova et al. 2002; Garcia et al. 2018; Henao Ramírez et al. 2018). Despite the commercial importance of cacao, the cellular and molecular mechanisms underlying this genotype-dependent variability remain poorly understood, limiting the rational improvement of SE protocols, particularly for highly recalcitrant genotypes.

In cacao, callus induction is highly efficient, with nearly all cultured explants producing callus regardless of genotype (Maximova et al. 2002; Garcia et al. 2018; Henao Ramírez et al. 2018). However, only a small fraction of the resulting callus enters the embryogenic pathway, and the biological mechanisms underlying this restricted transition remain poorly understood. Thus, the major bottleneck in cacao SE appears to be not callus formation itself, but the acquisition of embryogenic competence. Previous histochemical and morphoanatomical studies have distinguished cacao embryogenic callus (EC) from non-embryogenic callus (NEC) based on its nodular morphology, differential accumulation of phenolic compounds, and distinctive cytological features (Maximova et al. 2002).

Embryogenic cells are generally characterized by a high nucleus-to-cytoplasm ratio, prominent nucleoli, fragmented vacuoles, dense cytoplasm, and callose deposition in their cell walls (Verdeil et al. 2007). These features reflect the metabolic and structural reorganization associated with embryogenic reprogramming. However, they primarily represent established embryogenic identity, whereas the sequence of cellular changes accompanying the transition from a proliferative, non-competent state to an embryogenically competent state remains largely unresolved.

Several genes associated with cacao SE have been identified, particularly through comparisons with zygotic embryogenesis (ZE) (Maximova et al. 2002; Alemanno et al. 2008; Zhang et al. 2014). Nevertheless, many key regulators of SE identified in other species remain largely unexplored in cacao. Moreover, previous studies have focused predominantly on a limited number of widely studied commercial genotypes, and comparative analyses across genotypes with contrasting embryogenic responses remain scarce. Consequently, it is unclear which regulatory patterns are broadly conserved across cacao germplasm, and which are associated with genotype-specific recalcitrance. Determining what distinguishes embryogenically competent from non-competent cells within the callus, and how their cytological and molecular states relate to the onset of embryogenic identity, is therefore essential for understanding cacao recalcitrance. Such knowledge may contribute to understand why calli derived from the same flower, explant, or donor plant can exhibit markedly different embryogenic responses, with some producing numerous somatic embryos while others remain non-embryogenic (Quiroz-Figueroa et al. 2006).

In this context, this study aimed to identify the morphological, cellular, and molecular characteristics associated with the acquisition of embryogenic competence during SE induction in *T. cacao* clone CNCH13. Callus types and

somatic embryos at early developmental stages were first characterized morphologically and then examined using cytohistological approaches to identify cellular features associated with embryogenic competence. In parallel, RT-qPCR was used to compare the expression profiles of selected embryogenesis-related genes across these tissues, including the transcription factor genes *BBM*, *ABI3*, *AGL15*, *WUS*, and *LEC2*; the auxin-related genes *IAA20* and *PIN1*; and the receptor-like kinase genes *SERK1* and *SERK2*. Collectively, our findings provide an integrated cellular and molecular view of the characteristics associated with the acquisition of embryogenic competence in *T. cacao*. To our knowledge, this is the first study to characterize the endogenous expression dynamics of multiple embryogenesis-related transcription factors and auxin-pathway genes across sequential developmental stages of conventionally induced primary somatic embryogenesis in *T. cacao*. By identifying features associated with contrasting embryogenic responses, this study provides a foundation for improving SE in recalcitrant cacao genotypes and, ultimately, expanding the application of clonal propagation and genome-engineering technologies in this economically important crop.

Materials and methods

Plant material

For this study, we selected the cacao clone CNCH13, a Colombian commercial clone originating from an open-pollinated tree selected in San Vicente de Chucurí, Santander, Colombia. CNCH13 was developed by the Compañía Nacional de Chocolates (CNCh) and registered by the Colombian Agricultural Institute (ICA). This clone is agronomically valuable because of its high productivity, self-compatibility, and favorable bean-quality traits (Oliveira-De-Andrade et al. 2025). CNCH13 is conventionally propagated by grafting and used as scion material. Its agronomic relevance to Colombian cacao production highlights the importance of developing an efficient somatic embryogenesis system for its large-scale clonal propagation.

Primary somatic embryogenesis

To produce somatic embryos of CNCH13, immature flower buds measuring 4–5 mm were collected before anthesis from trees cultivated at the Yariguíes Farm, located in the municipality of Barrancabermeja, Santander, Colombia, and owned by the Colombian National Chocolate Company (CNCh). Immediately after collection, the flower buds were immersed in a cold, sterile solution containing only half-strength DKW basal salts, placed in Falcon tubes, and transported at 4°C to the University of Antioquia in Medellín, Colombia. Somatic embryos were induced according to Fontanel et al. (2002), with some modifications. Briefly, the flower buds were disinfected for 5 min with commercial bleach diluted to a final concentration of 0.5% (v/v). The flower buds were then rinsed three times with sterile distilled water and dissected, and the petals were used as explants. Twenty-five petals were placed in each Petri dish containing “INDI” medium supplemented with 2,4-D and kinetin as plant growth regulators (PGRs) to induce callus formation (Table 1). Approximately 60 Petri dishes were established in each experimental repetition. The entire procedure, from flower-bud collection to culture establishment, was performed independently three times at different times. The explants were incubated in darkness at $27 \pm 2^\circ\text{C}$ and 60% relative humidity for 20 d.

The callus masses that developed from the petal explants were subsequently transferred to PGR-free INDl_{exp} medium to promote somatic embryo expression (Table 1). Explants that did not form callus or were lost to contamination were excluded before transfer to the expression phase. For each independent repetition, 40 Petri

dishes containing 25 explants with established callus were transferred to INDlexp medium. The cultures were maintained under the same conditions for an additional 90 d, periodically monitored, and subcultured onto fresh INDlexp medium every 30 d. The different callus types were characterized as described below.

Table 1
Composition of the culture media used for somatic embryogenesis induction (INDI) and embryo expression (INDlexp) in *cacao*.

Medium Composition (mg/L)	INDI	INDlexp
Macronutrients (Phytotech™)	DKW	DKW
Micronutrients (Phytotech™)	DKW	DKW
Myoinositol (Phytotech™)	100	100
Nicotinic acid, Vit B3 (Phytotech™)	1.0	1.0
Thiamine, Vit B1 (Phytotech™)	2.0	2.0
Glycine (Phytotech™)	2.18	-
L-lysine (Sigma-Aldrich™)	0.45	-
L-leucine (Bio Basic Inc.™)	0.32	-
L-arginine (Phytotech™)	0.43	-
L-tryptophan (Bio Basic Inc.™)	0.51	-
2,4-dichlorophenoxyacetic acid, 2,4-D (Sigma-Aldrich™)	1.0	-
α-naphthaleneacetic acid, ANA (Merck KGaA™)	-	-
Kinetin, KIN (Bio Basic Inc.™)	0.25	-
Adenine (Bio Basic Inc.™)	-	-
Glucose (Phytotech™)	30000	30000
Gellex (Caisson Labs™)	3000	3000
pH	5.8	5.8

DKW, Driver and Kuniyuki Walnut medium (1984).

Morphological analysis of callus

To characterize the morphology of the callus cultures, stereomicroscopic observations and photographic records were obtained after 60 d of culture in the expression phase on INDlexp medium. Callus samples were classified according to color, shape, and texture, and the presence or absence of somatic embryos was recorded for each morphological type. Observations and image acquisition were performed using a Nikon SMZ1270 stereomicroscope equipped with a DS-Fi3 digital camera (Nikon, Kanagawa, Japan).

Embryogenic callus frequency response

Embryogenic callus frequency (ECF) was calculated as the percentage of embryogenic calli (EC) relative to the total number of calli formed. A callus was considered embryogenic when it produced at least one somatic embryo. Each Petri dish contained 25 calli and was considered an experimental unit; 40 Petri dishes were evaluated per independent repetition, corresponding to 1,000 calli per repetition. Three independent temporal repetitions were performed, resulting in a total of 120 Petri dishes and 3,000 calli evaluated. The embryogenic response was assessed at 30, 60 and 90 days of culture during the expression phase on INDlexp medium. The EF was determined using the following equation:

$$ECF = \frac{\# \text{ embryogenic callus}}{\# \text{ total of callus}} \times 100\%$$

Histological analysis

Histological analyses were performed to characterize the cellular features of the different callus types and somatic embryos at initial developmental stages. Samples were fixed for 2 days at room temperature in FAA solution containing 3.7% formaldehyde, 50% ethanol, and 5% acetic acid in distilled water. The fixed tissues were rinsed with distilled water and dehydrated through a graded ethanol series of 70, 80, 90, 95, and 100% (v/v). Afterward, the samples were gradually infiltrated by successive transfers through a graded ethanol:HistoChoice (Sigma-Aldrich) series (% v/v), passing sequentially through the following proportions: 90:10, 70:30, 50:50, 30:70, and 10:90 (v/v). The samples were then subjected to two consecutive changes of pure HistoChoice, followed by two changes of molten paraffin at 65°C (Paraplast Plus®, P3683, Sigma-Aldrich). During dehydration, gradual infiltration, and embedding in paraffin, the samples were maintained in each solution for 2 h. After embedding, the samples were placed in metal molds and allowed to solidify into paraffin blocks for 24 hours.

The blocks were then sectioned at 10 µm thickness using a Leica RM2125 rotary microtome. Separate sections were used for two staining procedures. For general characterization of cellular and tissue organization, sections were stained with FASGA solution for 12 h according to Tolivia and Tolivia (1987), to differentiate nuclei from lignified and non-lignified tissues. For histochemical detection, additional sections were double-stained sequentially: first with periodic acid–Schiff reagent (PAS, HX 106073, Merck), according to the manufacturer's instructions, to detect starch reserves, and second with 1% (w/v) Amido black 10B prepared in 7% (v/v) acetic acid for 2 min, to detect structural and storage proteins. The stained sections were then mounted using Entellan mounting medium (EMD Millipore, Darmstadt, Germany) and examined under a Nikon Eclipse 80i light microscope. Images were acquired using a Nikon DS-Fi1 CCD camera, and image processing was conducted using NIS-Elements Basic Research software (version 4.30.02, Nikon).

Gene expression analyses by RT-qPCR

In parallel with the histological analyses, the relative gene expression of *ABI3*, *LEC2*, *WUS*, *BBM*, *AGL15*, *IAA20*, *PIN1*, *SERK1* and *SERK2* were evaluated in different callus types and somatic embryos using reverse transcription quantitative real-time PCR (RT-qPCR). Six samples were selected that which corresponded to key stages of the SE process. These samples were categorized as follows: (i) callus formed after 20 d in INDI medium (CaINDI), corresponding to the onset of petal dedifferentiation; (ii) callus cultured for 30 d on INDlexp medium (CaEXp), when somatic some embryos began to form (the calli selected for analysis at this point did not have visible embryos); (iii) embryogenic callus that produced embryos after 60 d on INDlexp medium (EC), representing the peak of somatic embryo formation; (iv) non-embryogenic brown callus (NEC), that did not form embryos after 60 d

on INDlexp medium; (v) non-embryogenic white callus (WC), that did not form embryos after 60 d in INDlexp medium; and (vi) globular embryos (GE), representing an early stage of somatic embryo development.

Total RNA was extracted using different protocols depending on the nature of the sample. For callus samples, RNA was extracted using PureLink™ Plant RNA Reagent (Invitrogen™), with some modifications. Briefly, approximately 300 mg of callus tissues was ground to a fine powder in liquid nitrogen and transferred to a 2 ml tubes containing 700 µl of cold PureLink® Plant Reagent at 4°C. The samples were gently mixed until the tissue was completely resuspended, then incubated at room temperature for 5 min, and centrifuged at 12,000 rpm for 4 min at room temperature. Next, 600 µl of the supernatant was transferred to 1.5 ml vials containing 200 µl of 5 M NaCl and mixed by inversion. Afterward, 600 µl of chloroform was added to each sample and the tubes were inverted six times, and then centrifuged at 12,000 rpm for 10 min at 4°C. The upper aqueous phase (600 µl) was transferred to tubes containing 750 µl of cold isopropanol, mixed by inversion, and incubated at room temperature for 10 min.

Following this, the mixture was centrifuged at 12000 rpm at 4°C for 10 min, the supernatant was discarded, and the pellet was washed by adding 1 ml of cold 75% ethanol and centrifuged at 12000 rpm at 4°C for 5 min. The supernatant was then discarded, and a second wash was performed with 500 µl of 75% ethanol, followed by centrifugation at 12000 rpm at 4°C for 3 min. The supernatant was removed, and the pellets were resuspended in 20 µl of DEPC-treated water.

Additionally, for embryos, RNA was obtained following the protocol proposed by Ma et al., 2015, with some modifications. For this protocol, 100 mg of tissue was macerated in 1.5 ml vials submerged in liquid nitrogen, and the powder was transferred to 2 ml vials containing 1 ml of SEBM extraction buffer (0.25 M NaCl; 0.05 M Tris-HCl (pH = 7.5); 0.02 M EDTA (pH = 8.0); 1% SDS; 1% (w/v) PVP; and 1.3 M 2-mercaptoethanol). The samples were then gently mixed until completely resuspended, incubated on ice for 10 min, and centrifuged at 12000 rpm at 4°C for 15 min. Subsequently, 500 µl of the supernatant was transferred to 1.5 ml vials containing 300 µl of 5 M NaCl and mixed 4 times by inversion. Then, 500 µl of cold chloroform was added to each sample, mixed 8 times by inversion, and centrifuged at 12000 rpm at 4°C for 7 min to separate the phases. Next, 400 µl of the aqueous phase was transferred to 1.5 ml vials containing 500 µl of a cold phenol:chloroform:isoamyl alcohol mixture (25:24:1), mixed 8 times by inversion, and centrifuged at 12000 rpm at 4°C for 7 min to separate the phases. Subsequently, approximately 400 µl of the upper aqueous phase was transferred to 1.5 ml vials containing 400 µl of cold isopropyl alcohol, mixed by inversion, and incubated at room temperature for 10 min. Afterwards, the mixture was centrifuged at 12000 rpm at 4°C for 10 min, the supernatant was discarded, and 1 ml of cold 75% ethanol (v/v) was added to each vial to wash the pellet. The samples were incubated on ice for 5 min and centrifuged at 12000 rpm at 4°C for 5 min. The supernatant was removed, and the pellets were resuspended in 20 µl of DEPC-treated water.

RNA samples were treated with DNase I enzyme (PerfeCTa® DNase I, Quantabio) to eliminate contaminating DNA, and RNA was quantified by absorbance at 260 nm in a spectrophotometer (NanoDrop 1000, Thermo Scientific). First-strand cDNA was synthesized with Reverse Transcription System A3500 (Promega, Madison, USA) using 2 µg µL⁻¹ RNA and following the manufacturer's protocol. The cDNA obtained was used for amplification by qRT-PCR using specific primers for each gene (Table 2).

All PCR reactions were performed in a JENA qTOWER³ Series thermocycler (ANALYTIK), using the intercalation dye Luminaris Color HRM as a fluorescent reporter. Each reaction was performed in a 20-µL mixture containing 1.2 µl of cDNA (final concentration of 38 ng/µl), 10 µl of Luminaris Color HRM Master Mix (2X), 0.5 µl of 40X yellow

sample buffer, and 2 µl of each forward and reverse primer (final concentration 0.3 µM). The reactions were performed in 96-well PCR plates (Sorenson, Bioscience). The amplification protocol consisted of an initial denaturation step at 95°C for 10 min, followed by 44 cycles of denaturation at 95°C for 10 s, primer-specific annealing at 55–65°C for 5 s and extension at 72°C for 20 s. Three biological replicates were analyzed for each tissue type, using three technical qPCR replicates. The mean Ct value of the technical replicates was calculated for each biological replicate. The stability of *WOX2* expression across the six tissue types was evaluated before normalization, and no significant differences in its expression were detected among the tissues. Therefore, *WOX2* was selected as the reference gene. Relative expression of *ABI3*, *LEC2*, *WUS*, *BBM*, *AGL15*, *IAA20*, *PIN1*, *SERK1*, and *SERK2* was normalized against *WOX2* and calculated using the $2^{-\Delta C_t}$ method (Schmittgen and Livak 2008).

Table 2
Primer sequences used for RT-qPCR analysis of embryogenesis-related genes in *Theobroma cacao*.

Gen	Abbreviature	Accession number	Sequence	Annealing temperature (°C)
<i>BABY BOOM</i>	<i>TcBBM</i>	Thecc1EG025114t1	5' GCAAGAACAAGCAACGGTGG 3' 3' CGGTGCTGGTTGGTTCCTTA 5'	57
<i>SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1</i>	<i>TcSERK1</i>	Thecc1EG036627t1	5' AGTCTTTGGCTTTGGTTTGGT 3' 3' GCAATGCATCACCTTCCACA 5'	55
<i>SOMATIC EMBRYOGENESIS RECEPTOR KINASE 2</i>	<i>TcSERK2</i>	XM_007043798.2	5' AAGACAACCGCAAAGAAAGC 3' 3' AATCTCCTCCCCATCCTTCA 5'	55
<i>AGAMOUS- LIKE 15</i>	<i>TcAGL15</i>	Thecc1EG005818t1	5' TCTTCATGGTGTGCTCTGC 3' 3' AGCCAATTTTGCAATATCCCTGA 5'	57
<i>ABSCISIC ACID INSENSITIVE 3</i>	<i>TcABI3</i>	Thecc1EG003056t1	5' TGGGTACATGGATCTGCTTGA 3' 3' TCGTCGGTACTTCTTGCTCC 5'	58
<i>LEAFY COTYLEDON 2</i>	<i>TcLEC2</i>	Thecc1EG029838t1	5' CTTGAGGGCAGCTTTGGATA 3' 3' ATCTGCGTCCTTGAGAGTTG 5'	58
<i>INDOLE-3-ACETIC ACID INDUCIBLE 20</i>	<i>TcIAA20</i>	Thecc1EG034196t1	5' CTATCGAGAGCAGCAACCAC 3' 3' CGATTGTTGCTCCCTTCCAT 5'	57
<i>WUSCHEL related homeobox 2</i>	<i>TcWOX2</i>	XP_007013782.2	5' CTGGCCCGAATGTTGTTTGT 3' 3' GGCCTTCCTCTCCTCTTGAC 5'	58
<i>WUSCHEL</i>	<i>TcWUS</i>	Thecc1EG042593t1	5' TCATTTGCAGAGACAACAGGA 3' 3' AGGCTGACATCCCTTAACCTG 5'	57

Statistical analyses

Embryogenic frequency was evaluated after 30, 60, and 90 d of culture on INDlexp medium. The experiment consisted of three independent temporal repetitions, each representing a separate experimental run with 40 experimental units ($n = 40$ per run). Data were analyzed using a randomized complete block design, with experimental run as the blocking factor and culture time as the treatment factor. For gene expression analysis, three biological replicates were evaluated for each of the six tissue types (CaINDI, CaEXP, EC, NEC, WC, and GE; $n = 3$), with three technical qPCR replicates per biological replicate. Relative expression values were natural-log-transformed prior to analysis to meet ANOVA assumptions of normality and homogeneity of variance, and differences among the six tissue types were assessed by one-way ANOVA using a completely randomized design. For both analyses, when significant effects were detected, means were compared using Fisher's least significant difference (LSD) post hoc test ($\alpha = 0.05$). All analyses were performed in R version 4.3.1 using the *doebio* research package version 0.1.0 (R Core Team 2021).

Results

Primary SE induction

During SE induction, petal explants underwent dedifferentiation after 20 d of culture on INDI medium, resulting in the formation of cream-colored calli (Fig. 1a-b.) Upon transfer to INDlexp medium, the calli developed distinct brown and white regions (Fig. 1b). Somatic embryos (SEs) began to form after 30 d of culture on INDlexp medium, and numerous embryos were observed by 60 d (Fig. 1c). Notably, calli derived from the basal region of the petal produced SEs, whereas calli derived from the ligule did not. The main stages from callus induction to embryo expression are illustrated in Fig. 1.

Morphological characteristics of callus formed during SE

We identified distinct morphological changes in calli cultures over time. After 60 days of culture on INDlexp medium, four calli types were identified based on color, texture, and embryogenic response. Type (i): white callus (WC), characterized by a friable appearance and firm consistency; this callus type did not form somatic embryos (Fig. 2a). Type (ii): brown callus with a hard consistency; these calli did not produce somatic embryos and were therefore classified as non-embryogenic callus (NEC) (Fig. 2b). Type (iii): mixed callus with variable morphology, containing regions of WC and NEC together with small, light brown-yellowish regions, indicated by white circles in Fig. 2c. Type (iv): these light brown-yellowish, granular regions with a soft consistency were classified as embryogenic callus (EC), as SEs developed from them (Fig. 2c); given their embryogenic potential, EC regions were excised from the mixed calli and separated for subsequent analyses. We observed that EC regions represented only a small portion of the total callus mass. Notably, the light brown-yellowish coloration and granular appearance were consistently associated with embryogenic regions. SE developmental stages were observed simultaneously within the same cultures (Fig. 2c).

Temporal analysis of embryogenic callus frequency (ECF)

During the induction phase, a progressive increase in the embryogenic response was observed between 30 and 60 d of culture, with ECF remaining consistent across the three independent repetitions (Fig. 3). A significant difference in ECF was found between these two time points: at 30 d, ECF ranged from 5–8%, whereas by 60 d it increased to 28–30% on average. By 90 d, however, no new EC formation was observed, also after 90 d in Index medium the calli lost the capacity to generate new somatic embryos (data not shown) (Fig. 3). Additionally, embryo conversion to plants was achieved, as well *ex vitro* and field adaptation (data not shown).

Cytohystological characteristics of embryonic and non-embryonic tissues

Comparative cytohystological analyses were performed among the different calli types (WC, NEC and EC) to identify specific cellular characteristics in the cells that acquired embryogenic competence and gave rise to embryos. Additionally, histological sections were made for SE at different developmental stages to observe their characteristics and connections with EC. The WCs were observed as simple and compact tissue composed mainly of parenchyma cells (PAs) with variable shapes and sizes (Fig. 4b). Furthermore, these calli lacked accumulated storage compounds, as indicated by carbohydrate and protein staining (Fig. 4c). Similarly, NECs were composed of PAs with variable shapes, mostly larger cells with thicker walls (Fig. 4e). In contrast to WCs cells, NECs cells contained fuchsia-colored granules, which correspond to starch grains (SGs) (Fig. 4f).

In several ECs cross-sections, diverse and complex tissues comprising different cell types were observed (Fig. 5b). A relatively high prevalence of large PAs, characterized by wide intercellular spaces and abundant accumulation of SGs was observed (Fig. 5c, f). Additionally, in PA cells, the presence of two distinctive phenolic compounds (FC) types was also detected. The first ones were classified as filled complete phenolic compounds (CFCs) displaying yellow coloration into a large vacuole (Fig. 5c), and the second ones were classified as droplets of phenolic compounds (DFCs) contained into several small vacuoles characterized by a translucent yellowish appearance (Fig. 5f). Interestingly, the PA cells with FC and SG reserves had the capacity to differentiate into groups with smaller sizes, regular shapes, and visible nuclei. These groups were recognized as embryogenic competent cells (CCs) (Fig. 5d). These groups of CCs also began to separate or isolate themselves, forming smaller groups (Fig. 5g). These CCs marked the transition toward the formation of embryogenic clusters (EMCs) (Fig. 5e), eventually giving rise to independent clusters that separated from the CCs (Fig. 5h). The EMCs were formed by embryonic cells (ECEs), which are smaller than competent cells (CCs) and are isodiametric and uniformly sized, with prominent nuclei indicating high mitotic activity, resulting in rapid cell cluster proliferation. Notably, these ECEs were surrounded by larger PA cells containing SG reserves (Fig. 5i).

In the EC, before transitioning to globular embryos (GEs), proembryos (PEs) are subsequently formed from EMC (Fig. 6a). PEs structures showed that some zones remained attached to the PAs of ECs and acquired an oval-round shape like that of GEs, which lacked an epidermal layer. PEs are composed of cells that display characteristics like those of ECEs, with small sizes, regular shapes, and dense nuclei. The presence of storage compounds, such as SGs, as well as FCs, which accumulated in large vacuoles close to PE structures, was observed (Fig. 6b). Once PEs fully differentiate, they give rise to GE formation (Fig. 6c). They were characterized by an independent structure with a round shape and an epidermal layer surrounding the structure (Fig. 6d, f). At this stage, a small meristematic center (MC) formation was visible, which gave rise to the procambium (PC). This MC was characterized by small, isodiametric cells with dense nuclei and thick walls (Fig. 6e). This center was surrounded by larger embryogenic cells that accumulated SGs (Fig. 6g).

In the posterior developmental stage, early cotyledonary somatic embryos (ECTs) presented a typical embryonic structure with well-formed cotyledons (CTs) and hypocotyl (HP), and the procambium meristem was visible; additionally, a suspensor-like structure (LS) was observed at the base of the HP (Fig. 6h-i). The LS is composed of cells rich in SG reserves, which probably function as a nutritional source to mobilize carbohydrates between the EC and the SE. Notably, the LS displayed features such as the zygotic embryo (ZE) suspensor (Winkelmann 2016) (Fig. 6j). After SE at the intermediate cotyledonal (CI) stage, a bipolar configuration with two well-defined axes was observed (Fig. 6k). At the apex, the apical meristem (SAM) was observed to have a "dome" shape and was located between the cotyledons (Fig. 6l). The SAM comprises a small cluster of meristematic cells, with compact round cells and highly prominent nuclei (Fig. 6n). At the bottom, the root meristem (RAM) is fully differentiated, exhibiting a cellular morphology similar to that of the SAM (Fig. 6o). At this stage, the PC was completely formed, spreading throughout all axes of the embryo (Fig. 6k). This meristem was characterized by thin elongated cells and large nuclei (Fig. 6m-p). Additionally, SG accumulation at the lower axis of the HP was observed (Fig. 6q).

Gene expression profiling of TFs and auxin pathway genes in somatic embryogenesis

In this study, we analyzed the expression levels of genes associated with the SE process in six different types of embryogenic culture tissues: CaINDI, CaEXP, EC, NEC, WC, and GEs. We selected five transcription factors (TFs), *ABI3*, *LEC2*, *WUS*, *BBM*, and *AGL15*, which have been widely reported to be crucial during different phases of SE development (Long et al. 2022). Furthermore, we selected some genes involved in auxin biosynthesis, signaling, and transport, including *IAA20* and *PIN1*. Additionally, we included two membrane receptors, *SERK1* and *SERK2*, which are implicated in somatic embryogenesis signaling pathways.

Among the six tissues analyzed, expression of the five TFs was generally highest in GE (Fig. 7). *BBM* and *ABI3* followed a similar pattern, with the highest expression among the callus tissues in CaINDI and CaEXP; for *BBM*, CaEXP was significantly higher than EC, NEC, and WC, while CaINDI did not differ significantly from either group (Fig. 7a). *ABI3* showed a clearer separation, with CaINDI and CaEXP significantly higher than EC, NEC, and WC, and GE significantly higher than all callus tissues (Fig. 7b). *AGL15* followed a different trend, with its highest expression in GE and CaEXP, which did not differ significantly from each other; CaINDI, NEC, and WC showed the lowest expression, while EC was intermediate (Fig. 7c). *LEC2* was essentially restricted to CaEXP among the callus tissues, with CaEXP significantly higher than CaINDI and NEC, and was significantly enriched in GE relative to all callus tissues (Fig. 7d). *WUS* showed low expression overall, highest in CaINDI and CaEXP; expression in GE, although numerically the highest, was not significantly different from these two callus stages, whereas EC and WC showed the lowest expression (Fig. 7e).

IAA20 expression was significantly higher in GE than in all callus tissues, which did not differ significantly from one another (Fig. 8a). *PIN1* was detected at low levels across all six tissues, but given its high variability, no significant differences were found among any of the tissue types, including GE (Fig. 8b). *SERK1* and *SERK2* showed broadly similar patterns, with the highest expression in CaEXP and GE, which did not differ significantly from each other, and the lowest expression in CaINDI, EC, and NEC; WC was intermediate in both cases (Fig. 9). Overall, expression of the TFs and *IAA20* was consistently highest in GE, supporting their role as regulators of embryo identity and auxin-mediated signaling at the onset of embryo development, whereas *PIN1* and the *SERK* receptors were more broadly and variably expressed across embryogenic and non-embryogenic tissues alike.

Discussion

The induction of SE often represents a challenge for the propagation of *T. cacao*, especially for elite genotypes, which are recalcitrant to this process. In the present study, SE was successfully induced in the elite clone CNCH13, with embryogenic callus frequency reaching approximately 30% after 60 d of culture. The absence of a further increase at 90 days suggests that the period during which new embryogenic callus was formed was largely restricted to the first 60 days of culture. Although Henao Ramírez et al. (2018), previously reported SE induction in CNCH13 using staminodes as explants, their protocol did not result in SE under our experimental conditions. By contrast, SE was successfully induced from petal explants using the protocol described by Fontanel et al. (2002). These contrasting responses highlight the context-dependent nature of cacao SE, in which genotype may interact with explant type, culture conditions, donor-plant age and physiological status, season, and flower position to determine the embryogenic response (Boutchouang et al. 2016). Beyond establishing SE in CNCH13, this study provides, to our knowledge, the first integrated analysis of the cellular features associated with embryogenic cell cluster formation and the expression patterns of multiple embryogenesis-related transcription factor and auxin-pathway genes across successive tissue states and developmental stages of conventionally induced primary SE in *T. cacao*.

Our findings revealed morphological, cytohistological, and molecular features associated with the acquisition of embryogenic competence in *T. cacao*. During SE induction in this species, floral explants undergo cellular reprogramming and form morphologically distinct callus types (Alemanno et al. 1996). However, the morphological features of calli capable of producing somatic embryos are not consistent across cacao clones, making it difficult to distinguish embryogenic callus from non-embryogenic callus solely on the basis of external appearance (Henao Ramírez et al. 2018). In the present study, EC from clone CNCH13 was characterized by a light brown-yellowish color, granular appearance, and relatively soft consistency (Fig. 2c). Although these features have not been commonly reported during primary SE in cacao, they resemble those described for EC during secondary SE (Guillou et al. 2018). Nevertheless, during secondary embryogenesis, these features are well-recognized in ECs (Guillou et al. 2018). Similar morphological features have also been reported in embryogenic cultures of *Coffea canephora* (Ducos et al. 2007) and *Castanea mollissima* (Lu et al. 2017). Collectively, these observations suggest that a granular appearance and soft consistency may represent morphological indicators of embryogenic potential across different species. Nevertheless, because callus morphology can vary with the species, clone, explant type, and culture conditions, these traits should be interpreted together with cytohistological and molecular evidence.

Embryogenic callus comprises distinct cell types that accumulate storage reserves and phenolic compounds and exhibit different degrees of embryogenic competence (Fig. 5). We determined that EC contained large PAs with abundant starch grains and two morphologically distinct types of FCs. Although starch grains were also detected in non-embryogenic tissues, their persistence throughout embryogenic cell cluster formation and SE development is consistent with their possible function as an energy and carbon reserve during cell proliferation and tissue differentiation. By contrast, phenolic deposits were predominantly associated with EC, suggesting a closer relationship with the cellular and metabolic reorganization accompanying the acquisition of embryogenic competence (Fig. 5c,f). Cytohistological observations revealed a gradual transition from PA cells to CCs, which then form isolated EMCs and PEs with evidence of active cell division, ultimately leading to SE formation (Fig. 5). Similar anatomical features have been reported in embryogenic cultures of grapevine (He et al. 2024) and passion fruit (Rocha et al. 2016). Collectively, these findings suggest that embryogenic identity is acquired progressively through intermediate cellular states rather than through the direct conversion of differentiated cells into somatic embryos. The occurrence of organized embryogenic cell clusters also supports a multicellular origin of primary

somatic embryos. This interpretation is consistent with Maximova et al. (2002), who reported that primary somatic embryos in *T. cacao* originate from multicellular embryogenic nodules, in contrast to the unicellular origin described for secondary somatic embryogenesis in this species.

As noted above, phenolic deposits were detected primarily in the PA cells of EC and were nearly absent from those of NEC and WC (Fig. 4). Consistent with our findings, previous studies have reported an association between phenolic accumulation and the acquisition of embryogenic competence, with phenolic-containing cells occurring predominantly in calli that subsequently produced somatic embryos (Kouakou et al. 2007; Reis et al. 2008). Similarly, Gallego Rúa et al. (2016) observed that phenolic compounds accumulated mainly at the periphery of regeneration-competent somatic embryos of the cacao clone CNCH4. However, excessive phenolic accumulation can impair *in vitro* regeneration in several species, including cacao (Pearce et al. 1998; Sanbongi et al. 1998; Kroon and Williams 1999; Alemanno et al. 2003; Nascimento-Gavioli et al. 2017), suggesting that its effects may depend on concentration, localization, and developmental context. Phenolic compounds are naturally abundant in cacao seeds and participate in plant responses to biotic and abiotic stresses, including excessive ultraviolet radiation, pathogens, and herbivores (Elwers et al. 2010). Because stress-related signals contribute to the cellular reprogramming that initiates SE (Fehér 2015), the preferential localization of phenolic deposits in EC of clone CNCH13 may reflect a localized stress-related metabolic state associated with the acquisition of embryogenic competence. Nevertheless, whether phenolic accumulation promotes embryogenic reprogramming or results from the metabolic changes accompanying this developmental transition remains to be determined.

In addition to phenolic deposits, starch grains were observed in both embryogenic and non-embryogenic calli (Figs. 4 and 5). Starch constitutes an important source of metabolic energy and carbon skeletons during plant growth and may also participate in signaling processes associated with morphogenesis (Bartos et al. 2018).

In several species, including *Ormosia henryi* *Ormosia henryi* (Wu et al. 2021), *Acca sellowiana* (Cangahuala-Inocente et al. 2014), *Musa* spp. AAA (Wang et al. 2013) and *Medicago arborea* (Martin et al. 2000) starch reserves were more abundant in EC than in NEC during SE. Here, instead, starch grains occurred in both EC and NEC and therefore did not constitute a specific marker of embryogenic competence. Nevertheless, their persistence from embryogenic cell cluster formation through the globular and cotyledonary stages of SE development (Fig. 6) is consistent with their possible function as an energy and carbon reserve during cell proliferation and tissue differentiation. Similarly, He et al. (2024) reported that starch was mobilized during globular embryo formation in grapevine, contributing to energy production and cellulose biosynthesis. In cacao

Gallego Rúa et al. (2016) also described a temporal reorganization of reserve compounds during primary SE: polysaccharide content progressively decreased, whereas protein content increased between 15 and 60 days of culture in a regeneration-competent clone. By contrast, non-regenerating calli retained a disorganized mixture of polyphenols, polysaccharides, and proteins throughout culture. Collectively, these observations suggest that, although starch accumulation alone does not indicate embryogenic competence, its regulated storage and mobilization may support the energetic and structural demands of embryogenic reprogramming and somatic embryo development in cacao.

The histochemical changes described above reflect the metabolic reorganization accompanying embryogenic reprogramming and embryo development. To complement this cellular evidence and investigate the associated regulatory changes, we also compared the expression patterns of embryogenesis-related genes across tissues representing distinct developmental stages and embryogenic responses. The acquisition of embryogenic competence during SE is a complex process in which external stimuli activate genetic programs that promote the

reprogramming of somatic cells toward a totipotent state (Su et al. 2021). Several TFs participate in this process and interact with plant growth regulator biosynthesis and signaling pathways, particularly those related to auxin (Méndez-Hernández et al. 2019).

Among the TF genes evaluated, *BBM* showed the highest expression in globular embryos, followed by *ABI3* and *AGL15* (Fig. 7). Within the callus tissues, *BBM* and *ABI3* were more highly expressed during the early culture phases, represented by induction-phase callus (CaINDI) and expression-phase callus (CaEXP), than in EC, NEC, and WC collected after 60 days. This pattern suggests that *BBM* and *ABI3* are activated early during SE induction and embryo expression, before EC and NEC become morphologically distinguishable, and are subsequently reactivated or maintained at high levels during globular embryo development. Consistent with this interpretation, *BBM* is expressed in embryos and root meristems, where it regulates cell identity and growth (Aida et al. 2004; Galinha et al. 2007), and is widely recognized as a regulator of SE (Rupps et al. 2016; Gulzar et al. 2020; Wójcik et al. 2020; Kumar et al. 2021; Long et al. 2022). Moreover, ectopic expression of *BBM* is sufficient to induce spontaneous SE in the absence of exogenous PGRs in both *T. cacao* (Flórez et al. 2015) and *A. thaliana* (Horstman et al. 2017b). Similarly, *ABI3* expression has been associated with embryogenic potential in several species (Freitas et al. 2019), and high transcript levels have been reported in embryogenic cell lines in cotton (Kumar et al. 2021) and coffee (Freitas et al. 2019).

In contrast, *AGL15* exhibited a distinct expression pattern. Among the callus tissues, its highest expression occurred in CaEXP, where transcript levels were statistically comparable to those detected in globular embryos, before declining in EC, NEC, and WC at 60 days. *AGL15* is predominantly expressed in embryos during early seed development (Joshi et al. 2022), and ectopic constitutive expression of *AGL15* promotes SE in *A. thaliana* and *Glycine max* (hakare et al. 2008; Zheng and Perry 2014). Thus, the elevated expression of *AGL15* during the expression phase in CNCH13 may therefore be associated with the early transition from induced callus to embryo development. Collectively, our findings support the involvement of *BBM*, *ABI3*, and *AGL15* at distinct stages of primary SE in cacao. In particular, the expression pattern of *BBM* is consistent with its established role in embryogenic reprogramming and embryo identity, whereas *ABI3* and *AGL15* may contribute to the regulatory network connecting early induction events with subsequent embryo development.

LEC2 and *WUS* are key transcription factor genes involved in SE (Long et al. 2022). Although *LEC2* transcript abundance was relatively low, its expression was highest in CaEXP among callus tissues and reached its overall maximum in globular embryos (Fig. 7). This pattern differs from that reported by Zhang et al. (2014), who detected approximately 20-fold higher *LEC2* expression in embryogenic than in non-embryogenic cacao callus. However, their study examined secondary SE from cotyledon explants, whereas ours investigated primary SE from floral explants. Moreover, their calli had been maintained for approximately one month on PGR-free medium, resembling the conditions of CaEXP, in which *LEC2* expression was also relatively high. Previous studies have described increased *LEC2* expression during early SE induction and embryo maturation (Zhang et al. 2014; Brand et al. 2019). Thus, its relatively high expression in CaEXP and globular embryos may reflect stage-dependent regulation during primary SE in CNCH13. Analysis of later embryo stages will be necessary to determine whether *LEC2* expression increases further during maturation.

WUS also showed low expression across the tissues analyzed, with its highest levels among callus tissues detected in CaINDI and CaEXP; these levels were statistically comparable to those observed in globular embryos (Fig. 7). In *Boesenbergia rotunda*, *WUS* is expressed in embryogenic callus, particularly in meristematic cell domains, supporting its involvement in shoot apical meristem establishment and maintenance (Karim et al. 2018).

Low transcript abundance, however, does not necessarily indicate limited biological relevance because *WUS* activity may depend on spatially restricted expression and dosage thresholds. For example, low *WUS* levels activate *CLV3* expression in the shoot apical meristem, whereas higher levels can repress it (Perales et al. 2016). Therefore, the low *WUS* expression detected through bulk RT-qPCR may reflect localized activity within small cell populations. Overall, the expression patterns of *LEC2* and *WUS* suggest stage- and tissue-dependent regulation rather than their use as specific markers of embryogenic competence. Spatial expression and functional analyses are needed to clarify their roles during primary SE in cacao.

In addition to regulating embryogenic identity, several embryogenesis-related transcription factors modulate genes involved in auxin biosynthesis and signaling (Wójcik et al. 2020). For example, *AGL15* binds to a cis-regulatory sequence of *IAA30* and promotes its expression (Zheng et al. 2009), while *IAA30* has been associated with competence for SE (Braybrook et al. 2006). In our analysis, *IAA20* expression was highest in globular embryos, which also exhibited high *AGL15* expression (Fig. 8). However, the high variability in *IAA20* expression among callus tissues precluded the identification of a consistent pattern associated with the acquisition of embryogenic competence. Because *IAA20* and *IAA30* are closely related members of the Arabidopsis Aux/IAA family (Sato and Yamamoto 2008), the coincident expression of *IAA20* and *AGL15* in globular embryos raises the possibility that they participate in a common regulatory pathway. Nevertheless, direct regulation of *IAA20* by *AGL15* cannot be inferred from co-expression and requires functional validation. Overall, the preferential expression of *IAA20* in globular embryos suggests a role in auxin signaling during early somatic embryo development rather than as a specific marker of competence acquisition.

By contrast, *PIN1* was expressed at low levels and showed no significant differences among the six tissues because of the high variability within samples (Fig. 8). In Arabidopsis, AtPIN1 is not required for the acquisition of embryogenic cell identity but contributes to proembryo formation and subsequent embryo development (Pierroz 2023). The absence of a tissue-specific expression pattern in our study does not support *PIN1* as a marker of embryogenic competence in cacao. However, its expression across callus and embryo tissues is consistent with a possible role in auxin distribution during embryo development. Spatial expression analyses will be necessary to determine whether PIN1 is locally expressed in developing embryogenic domains, as such patterns may not be resolved through bulk RT-qPCR.

Finally, *SERK1* and *SERK2* generally displayed low transcript abundance, with significantly higher expression in CaEXP and globular embryos than in CaINDI, EC, and NEC (Fig. 9). *SERK* family genes have been associated with the transition from somatic to embryogenically competent cells and are widely used as molecular markers of SE in several plant species (Talapatra et al. 2014). In *T. cacao*, *SERK1* and *SERK2* have been implicated in embryo development and recurrent embryogenesis but are also expressed in vegetative tissues, including leaves (De Oliveira Santos et al. 2005). In our study, both genes were also expressed in NEC and WC, which did not produce somatic embryos. Therefore, their expression was not specific to the acquisition of embryogenic competence. Instead, the higher transcript levels observed in CaEXP and globular embryos suggest that *SERK1* and *SERK2* may participate in broader signaling and developmental processes associated with SE (Hu et al. 2005; Huang et al. 2010).

Collectively, the expression patterns observed in this study suggest that the molecular changes associated with SE begin before embryogenic and non-embryogenic calli become morphologically distinguishable. Several transcription factor genes showed relatively high expression in CaINDI and CaEXP, corresponding to the induction phase and the first 30 days on PGR-free expression medium, respectively. By 60 days, when EC, NEC, and WC

could be distinguished morphologically, the expression of several genes had declined or no longer differed among callus types. In contrast, globular embryos showed high expression of *BBM*, *ABI3*, *AGL15*, and *IAA20*, indicating that these genes may also contribute to early embryo development. Together with the cytohistological differences observed among callus types, these results support a temporal progression in which early molecular responses precede visible callus differentiation, whereas distinct cellular organization and gene-expression patterns emerge during embryo formation.

The interpretation of these expression patterns must nevertheless consider the cellular heterogeneity of embryogenic cultures. As shown by our cytohistological analyses, EC contains a complex mixture of parenchyma cells, competent cells, embryogenic cell clusters, and developing embryos. Consequently, bulk RT-qPCR may dilute gene expression restricted to small cell populations and cannot assign transcripts to specific cellular. Moreover, embryogenic reprogramming is controlled by interacting gene networks that operate sequentially or in parallel. Future single-cell and spatial-expression approaches, including *in situ* hybridization, should therefore examine the genes evaluated here across closely spaced time points from initial callus formation to embryo development. Integrating these approaches with the morphological and cytohistological markers identified in this study will help resolve the regulatory programs associated with competent cells before somatic embryos become visible.

Conclusions

This study provides an integrated view of a central bottleneck in cacao somatic embryogenesis: the transition from proliferating callus to embryogenically competent tissue. Our findings indicate that this transition cannot be reliably defined by callus morphology or a single histological or molecular marker. Instead, embryogenic competence emerges through interconnected, stage-dependent changes that begin before embryogenic and non-embryogenic calli become visually distinguishable and later become apparent through cellular organization, phenolic accumulation, and embryo-associated gene-expression patterns.

The identification of intermediate cellular states linking parenchyma cells to organized embryogenic clusters provides new insight into the developmental origin of primary somatic embryos and supports their multicellular origin. At the molecular level, *BBM*, *ABI3*, and *AGL15* showed the clearest associations with early embryogenic regulation, while their high expression, together with that of *IAA20*, in globular embryos supports additional roles during early embryo development. These stage-dependent patterns demonstrate that genes associated with early embryogenic responses are not necessarily specific markers of competence acquisition and may also function during subsequent embryo development. This distinction is important for interpreting candidate molecular markers and understanding the regulatory progression from callus induction to embryo formation.

By integrating morphological, cytohistological, and molecular evidence, this study establishes a basis for identifying responsive tissues before somatic embryos become visible and for directing spatial analyses toward the cells that initiate embryogenesis. This knowledge may guide the development of more targeted SE protocols for recalcitrant cacao clones and support the broader use of embryogenic cultures in clonal propagation, genetic transformation, and genome engineering.

Declarations

Author Contribution

L.M.Z.: Conceptualization, Investigation, Data curation, Formal analysis, Writing-original draft, review and editing; T.O.M.: Investigation, Data curation and review and editing; L.P.Q.: Conceptualization, review and editing; A.U.R.: Conceptualization, Funding acquisition, Supervision, review and editing, L.C.R.: Conceptualization, Investigation, Data curation, Formal analysis, Supervision, Writing-original draft, review and editing. All authors read and approved the final version.

Acknowledgement

We thank the members of the Plant Physiology and Plant Tissue Culture Laboratory at Universidad de Antioquia and the financial entities, the General Royalties System—Science, Technology (SGR), and Innovation Fund, through the Agrobiotechnology Development Center for Innovation and Territorial Integration project (CEDAIT-BPIN 2016000100060), the National Planning Department, the Office of the Governor of Antioquia, Universidad de Antioquia, Universidad Católica de Oriente, and Compañía Nacional de Chocolates (CNCh).

Funding Declaration

This work was funded by the General Royalties System, Science, Technology and Innovation Fund of the Government of Antioquia, in collaboration with the University of Antioquia, the Catholic University of the East, and the National Chocolate Company, Colombia, under project BPIN 2016000100060, awarded to Aura Inés Urrea Trujillo.

References

1. Aida M, Beis D, Heidstra R (2004) The *PLETHORA* Genes mediate patterning of the Arabidopsis root stem cell niche poles to define the sites for postembryonic development. Both shoot and root stem cells in *Arabidopsis thaliana* are maintained by the activity of nearby organizing. *Cell* 119:109–120. <https://doi.org/10.1016/j.cell.2004.09.018>
2. Alemanno L, Berthouly M, Michaux-Ferrière N (1996) Histology of somatic embryogenesis from floral tissues cocoa. *Plant Cell Tissue Organ Cult* 46:187–194. <https://doi.org/10.1007/BF02307094>
3. Alemanno L, Ramos T, Gargadenec A (2003) Localization and identification of phenolic compounds in *Theobroma cacao* L. somatic embryogenesis. *Ann Bot* 92:613–623. <https://doi.org/10.1093/aob/mcg177>
4. Arroyo-Herrera AL, Kú-González A, Escobedo-Gracia-Medrano R (2024) Overexpression of *WUSCHEL* improves the induction of embryogenic callus in scalps of *Musa acuminata* L. AAA, CV. GRAND NAIN †. *Trop Subtropical Agroecosystems* 27:10. <https://doi.org/10.56369/tsaes.4751>
5. Bartos PMC, Gomes HT, Velho Amaral LI (2018) Biochemical events during somatic embryogenesis in *Coffea arabica* L. *3 Biotech* 8:209. <https://doi.org/10.1007/s13205-018-1238-7>
6. Boutchouang RP, Akitio OF, Germaine AO, Tchouatcheu N, Niemenak N (2016) Influence of the position of flowers buds on the tree on somatic embryogenesis of cocoa (*Theobroma cacao* L. *Int J Plant Physiol Biochem* 8(2):7–16. <https://doi.org/10.5897/ijgmb2016.0247>
7. Boutilier K, Offringa R, Sharma VK (2002) Ectopic expression of *BABY BOOM* triggers a conversion from vegetative to embryonic growth. *Plant Cell* 14:1737–1749. <https://doi.org/10.1105/tpc.001941>
8. Brand A, Quimbaya M, Tohme J, Chavarriaga-Aguirre P (2019) Arabidopsis *LEC1* and *LEC2* orthologous genes are key regulators of somatic embryogenesis in cassava. *Front Plant Sci* 10:673.

<https://doi.org/10.3389/fpls.2019.00673>

9. Braybrook SA, Stone SL, Park S, Bui AQ, Le BH, Fischer RL, Goldberg RB, Harada JJ (2006) Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. *Proc Natl Acad Sci USA* 103:3468–3473. <https://doi.org/10.1073/pnas.0511331103>
10. Cangahuala-Inocente GC, Silveira V, Caprestano CA (2014) Dynamics of physiological and biochemical changes during somatic embryogenesis of *Acca sellowiana*. *Vitro Cell Dev Biology - Plant* 50:166–175. <https://doi.org/10.1007/s11627-013-9563-3>
11. De Oliveira Santos M, Romano E, Yotoko KSC (2005) Characterization of the cacao *SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE (SERK)* gene expressed during somatic embryogenesis. *Plant Sci* 168:723–729. <https://doi.org/10.1016/j.plantsci.2004.10.004>
12. Driver J, Kuniyuki A (1984) *In vitro* propagation of *Paradox walnut* rootstock. *HortScience* 19:507–509
13. Ducos JP, Labbe G, Lambot C, Pétiard V (2007) Pilot scale process for the production of pre-germinated somatic embryos of selected robusta (*Coffea canephora*) clones. *Vitro Cell Dev Biology - Plant* 43:652–659. <https://doi.org/10.1007/s11627-007-9075-0>
14. Durgaprasad K, Roy MV, Venugopal MA, Kareem A, Raj K, Willemsen V, Mähönen AP, Scheres B, Prasad K (2019) Gradient expression of transcription factor imposes a boundary on organ regeneration potential in plants. *Cell Rep* 29:453–463. <https://doi.org/10.1016/j.celrep.2019.08.099>
15. Elwers S, Zambrano A, Rohsius C, Lieberei R (2010) Histological features of phenolic compounds in fine and bulk cocoa seed (*Theobroma cacao* L). *J Appl Bot food Qual* 83:182–188
16. Fehér A (2015) Somatic embryogenesis - stress-induced remodeling of plant cell fate. *Biochim Biophys Acta Gene Regul Mech* 1849:385–402. <https://doi.org/10.1016/j.bbagrm.2014.07.005>
17. Fehér A (2019) 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>
18. Florez SL, Erwin RL, Maximova SN (2015) Enhanced somatic embryogenesis in *Theobroma cacao* using the homologous *BABY BOOM* transcription factor. *BMC Plant Biol* 15:121. <https://doi.org/10.1186/s12870-015-0479-4>
19. Fontanel A, Gire-Bobin S, Labbé G, Favereau P, Álvarez M, Rutte S, Pétiard V (2002) *vitro* multiplication and plant regeneration of *Theobroma cacao* L. via stable embryogenic calli. 10 IAPTC Congress. *Plant Biotechnol* 1:23–28
20. Freitas NC, Barreto HG, Torres LF (2019) In silico and in vivo analysis of *ABI3* and *VAL2* genes during somatic embryogenesis of *Coffea arabica*: Competence acquisition and developmental marker genes. *Plant Cell Tissue Organ Cult* 37:599–561. <https://doi.org/10.1007/s11240-019-01594-7>
21. Galinha C, Hofhuis H, Luijten M (2007) PLETHORA proteins as dose-dependent master regulators of Arabidopsis root development. *Nature* 449:1053–1057. <https://doi.org/10.1038/nature06206>
22. Gallego Rúa AM, Henao Ramírez AM, Urrea Trujillo AI, Atehortúa Garcés L (2016) Análisis de distribución de polifenoles y sustancias de reserva en embriogénesis somática de cacao. *Acta Biolo Colomb* 21:335–346. <https://doi.org/10.15446/abc.v21n2.50196>
23. Garcia C, Marelli JP, Motamayor JC, Villela C (2018) Somatic embryogenesis in *Theobroma cacao* L. *Methods Mol Biol* 1815:227–245. https://doi.org/10.1007/978-1-4939-8594-4_15
24. Guillou C, Fillodeau A, Brulard E (2018) Indirect somatic embryogenesis of *Theobroma cacao* L. in liquid medium and improvement of embryo-to-plantlet conversion rate. *Vitro Cell Dev Biology - Plant* 54:377–391.

<https://doi.org/10.1007/s11627-018-9909-y>

25. Gulzar B, Mujib A, Malik MQ (2020) Genes, proteins and other networks regulating somatic embryogenesis in plants. *J Genetic Eng Biotechnol* 18(1):31. <https://doi.org/10.1186/s43141-020-00047-5>
26. He X, He J, Luo F (2024) Impact of activated charcoal on the progression of somatic embryogenesis in grape cells: Histological and transcriptomic perspectives. *Sci Hortic* 336:113367. <https://doi.org/10.1016/j.scienta.2024.113367>
27. Henao Ramírez AM, De la Hoz Vasquez T, Ospina Osorio TM (2018) Evaluation of the potential of regeneration of different Colombian and commercial genotypes of cocoa (*Theobroma cacao* L.) via somatic embryogenesis. *Sci Hortic* 229:148–156. <https://doi.org/10.1016/j.scienta.2017.10.040>
28. Horstman A, Berner M, Boutilier K (2017a) A transcriptional view on somatic embryogenesis. *Regeneration* 4:201–216. <https://doi.org/10.1002/reg2.91>
29. Horstman A, Li M, Heidmann I (2017b) The *BABY BOOM* transcription factor activates the *LEC1-ABI3-FUS3-LEC2* network to induce somatic embryogenesis. *Plant Physiol* 175:848–857. <https://doi.org/10.1104/pp.17.00232>
30. Hu H, Xiong L, Yang Y (2005) Rice *SERK1* gene positively regulates somatic embryogenesis of cultured cell and host defense response against fungal infection. *Planta* 222:107–117. <https://doi.org/10.1007/s00425-005-1534-4>
31. Huang X, Lu XY, Zhao JT (2010) *MaSERK1* gene expression associated with somatic embryogenic competence and disease resistance response in banana (*Musa* spp). *Plant Mol Biol Rep* 28:309–316. <https://doi.org/10.1007/s11105-009-0150-z>
32. Ikeuchi M, Sugimoto K, Iwase A (2013) Plant callus: mechanisms of induction and repression. *Plant Cell* 25:3159–3173. <https://doi.org/10.1105/tpc.113.116053>
33. Ikeuchi M, Shibata M, Rymen B (2018) A gene regulatory network for cellular reprogramming in plant regeneration. *Plant Cell Physiol* 59:765–777. <https://doi.org/10.1093/pcp/pcy013>
34. Joshi S, Paul P, Hartman JM, Perry SE (2022) *AGL15* promotion of somatic embryogenesis: Role and molecular mechanism. *Front Plant Sci* 13:861556. <https://doi.org/10.3389/fpls.2022.861556>
35. Karim R, Tan YS, Singh P (2018) Expression and DNA methylation of *SERK*, *BBM*, *LEC2* and *WUS* genes in *in vitro* cultures of *Boesenbergia rotunda* (L.) Mansf. *Physiology and Molecular Biology of Plants* 24:741–751. <https://doi.org/10.1007/s12298-018-0566-8>
36. Kouakou TH, Waffo-Tégou P, Kouadio YJ (2007) Phenolic compounds and somatic embryogenesis in cotton (*Gossypium hirsutum* L). *Plant Cell Tissue Organ Cult* 90:25–29. <https://doi.org/10.1007/s11240-007-9243-2>
37. Kroon PA, Williams G (1999) Hydroxycinnamates in plants and food: Current and future perspectives†. *J Sci Food Agric* 79:355–361. [https://doi.org/10.1002/\(SICI\)1097-0010\(19990301\)79:3%3C355::AID-JSFA255>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1097-0010(19990301)79:3%3C355::AID-JSFA255>3.0.CO;2-G)
38. Kumar S, Ruggles A, Logan S (2021) Comparative transcriptomics of non-embryogenic and embryogenic callus in semi-recalcitrant and non-recalcitrant upland cotton lines. *Plants* 10:1775. <https://doi.org/10.3390/plants10091775>
39. Liu S, Hu Q, Luo S, Li Q, Yang X, Wang X, Wang S (2015) Expression of wild-type *PtrIAA14.1*, a poplar *Aux/IAA* gene causes morphological changes in *Arabidopsis*. *Front Plant Sci* 6:388. <https://doi.org/10.3389/fpls.2015.00388>

40. Long Y, Yang Y, Pan G, Shen Y (2022) New insights into tissue culture plant-regeneration mechanisms. *Front Plant Sci* 13:926752. <https://doi.org/10.3389/fpls.2022.926752>
41. Lu D, Wei W, Zhou W (2017) Establishment of a somatic embryo regeneration system and expression analysis of somatic embryogenesis-related genes in Chinese chestnut (*Castanea mollissima* Blume). *Plant Cell Tissue Organ Cult* 130:601–616. <https://doi.org/10.1007/s11240-017-1250-3>
42. Mahdavi-Darvari F, Noor NM, Ismanizan I (2015) Epigenetic regulation and gene markers as signals of early somatic embryogenesis. *Plant Cell Tissue Organ Cult* 120:407–422. <https://doi.org/10.1007/s11240-014-0615-0>
43. Martin AB, Cuadrado Y, Guerra H, Gallego P, Hita O, Martin L, Dorado A, Villalobos N (2000) Differences in the contents of total sugars, reducing sugars, starch and sucrose in embryogenic and non-embryogenic calli from *Medicago arborea* L. *Plant science: Int J experimental plant biology* 154(2):143–151. [https://doi.org/10.1016/s0168-9452\(99\)00251-4](https://doi.org/10.1016/s0168-9452(99)00251-4)
44. Maximova SN, Alemanno L, Young A (2002) Efficiency, genotypic variability, and cellular origin of primary and secondary somatic embryogenesis of *Theobroma Cacao* L. *Vitro Cell Dev Biology - Plant* 38:252–259. <https://doi.org/10.1079/IVP2001257>
45. Méndez-Hernández HA, Ledezma-Rodríguez M, Avilez-Montalvo RN (2019) Signaling overview of plant somatic embryogenesis. *Front Plant Sci* 10:77. <https://doi.org/10.3389/fpls.2019.00077>
46. Mönke G, Seifert M, Keilwagen J (2012) Toward the identification and regulation of the *Arabidopsis thaliana* *ABI3* regulon. *Nucleic Acids Res* 40:8240–8254. <https://doi.org/10.1093/nar/gks594>
47. Nascimento-Gavioli MCA, Cangahuala-Inocente GC, Steinmacher D (2017) Physiological and biochemical features of embryogenic and non-embryogenic peach palm (*Bactris gasipaes Kunth*) cultures. *Vitro Cell Dev Biology - Plant* 53:33–40. <https://doi.org/10.1007/s11627-017-9805-x>
48. Nolan KE, Kurdyukov S, Rose RJ (2009) Expression of the *SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE1 (SERK1)* gene is associated with developmental change in the life cycle of the model legume *Medicago truncatula*. *J Exp Bot* 60:1759–1771. <https://doi.org/10.1093/jxb/erp046>
49. Oliveira-de-Andrade D, Torres-Bazurto J, Cayón-Salinas DG, Piamba O, Alvarado LM, Hincapié OD (2025) Componentes del rendimiento en tres clones de cacao (*Theobroma cacao* L.) bajo fertilización química fraccionada. *Rev UDCA Act Div Cient* 28(1):e2672. <https://doi.org/10.31910/rudca.v28.n1.2025.2672>
50. Pearce G, Marchand PA, Griswold J (1998) Accumulation of feruloyltyramine and p-coumaroyltyramine in tomato leaves in response to wounding. *Phytochemistry* 47:659–664. [https://doi.org/10.1016/S0031-9422\(97\)00620-1](https://doi.org/10.1016/S0031-9422(97)00620-1)
51. Perales M, Rodriguez K, Snipes S, Yadav RK, Diaz-Mendoza M, Reddy GV (2016) Threshold-dependent transcriptional discrimination underlies stem cell homeostasis. *Proc Natl Acad Sci USA* 113:E6298–E6306. <https://doi.org/10.1073/pnas.1607669113>
52. Pierroz G (2023) Making babies: how auxin regulates somatic embryogenesis in *Arabidopsis* tissue culture. *Plant journal: cell Mol biology* 113(1):5–6. <https://doi.org/10.1111/tpj.16062>
53. Pilarska M, Malec P, Salaj J (2016) High expression of *SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE* coincides with initiation of various developmental pathways in *in vitro* culture of *Trifolium nigrescens*. *Protoplasma* 253:345–355. <https://doi.org/10.1007/s00709-015-0814-5>
54. Quiroz-Figueroa FR, Rojas-Herrera R, Galaz-Avalos RM, Loyola-Vargas VM (2006) Embryo production through somatic embryogenesis can be used to study cell differentiation in plants. *Plant Cell Tissue Organ Cult* 86:285–301. <https://doi.org/10.1007/s11240-006-9139-6>

55. Reis E, Batista MT, Canhoto JM (2008) Effect and analysis of phenolic compounds during somatic embryogenesis induction in *Feijoa sellowiana* Berg. *Protoplasma* 232:193–202. <https://doi.org/10.1007/s00709-008-0290-2>
56. Rocha DI, Pinto DLP, Vieira LM (2016) Cellular and molecular changes associated with competence acquisition during passion fruit somatic embryogenesis: ultrastructural characterization and analysis of *SERK* gene expression. *Protoplasma* 253:595–609. <https://doi.org/10.1007/s00709-015-0837-y>
57. Rupps A, Raschke J, Rümmler M (2016) 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* 243:473–488. <https://doi.org/10.1007/s00425-015-2409-y>
58. Sanbongi C, Osakabe N, Natsume M (1998) Antioxidative polyphenols isolated from *Theobroma cacao*. *J Agric Food Chem* 46:454–457. <https://doi.org/10.1021/jf970575o>
59. Sato A, Yamamoto KT (2008) Overexpression of the non-canonical *Aux/IAA* genes causes auxin-related aberrant phenotypes in *Arabidopsis*. *Physiol Plant* 133(2):397–405. <https://doi.org/10.1111/j.1399-3054.2008.01055.x>
60. Schmidt ED, Guzzo F, Toonen MA, de Vries SC (1997) A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. *Development* 124(10):2049–2062. <https://doi.org/10.1242/dev.124.10.2049>
61. Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative CT method. *Nat Protoc* 3:1101–1108. <https://doi.org/10.1038/nprot.2008.73>
62. Su YH, Tang LP, Zhao XY, Zhang XS (2021) Plant cell totipotency: Insights into cellular reprogramming. *J Integr Plant Biol* 63:228–243. <https://doi.org/10.1111/jipb.12972>
63. Talapatra S, Ghoshal N, Raychaudhuri S (2014) Molecular characterization, modeling and expression analysis of a *SOMATIC EMBRYOGENESIS RECEPTOR KINASE* (*SERK*) gene in *Momordica charantia* L. during somatic embryogenesis. *Plant Cell Tissue Organ Cult* 116:271–283. <https://doi.org/10.1007/s11240-013-0401-4>
64. Tang LP, Zhou C, Wang S, Yuan J, Zhang XS, Su YH (2017) *FUSCA3* interacting with *LEAFY COTYLEDON2* controls lateral root formation through regulating *YUCCA4* gene expression in *Arabidopsis thaliana*. *New Phytol* 213(4):1740–1754. <https://doi.org/10.1111/nph.14313>
65. Thakare D, Tang W, Hill K, Perry SE (2008) The MADS-domain transcriptional regulator *AGAMOUS-LIKE15* promotes somatic embryo development in *Arabidopsis* and soybean. *Plant Physiol* 146:1663–1672. <https://doi.org/10.1104/pp.108.115832>
66. Tolivia D, Tolivia J (1987) Fasga: A new polychromatic method for simultaneous and differential staining of plant tissues. *J Microsc* 148:113–117. <https://doi.org/10.1111/j.1365-2818.1987.tb02859.x>
67. Verdeil JL, Alemanno L, Niemenak N, Tranbarger TJ (2007) Pluripotent versus totipotent plant stem cells: dependence versus autonomy? *Trends Plant Sci* 12:245–252. <https://doi.org/10.1016/j.tplants.2007.04.002>
68. Voora V, Bermúdez S, Larrea C (2019) Global market report: cocoa. International Institute for Sustainable Development, Winnipeg
69. Wang X, Shi L, Lin G (2013) A systematic comparison of embryogenic and non-embryogenic cells of banana (*Musa* spp. AAA): Ultrastructural, biochemical and cell wall component analyses. *Sci Hortic* 159:178–185. <https://doi.org/10.1016/j.scienta.2013.05.008>

70. Wickramasuriya AM, Dunwell JM (2018) Cacao biotechnology: Current status and future prospects. *Plant Biotechnol J* 16:4–17. <https://doi.org/10.1111/pbi.12848>
71. Wójcik AM, Wójcikowska B, Gaj MD (2020) Current perspectives on the auxin-mediated genetic network that controls the induction of somatic embryogenesis in plants. *Int J Mol Sci* 21:1333. <https://doi.org/doi:10.3390/ijms21041333>
72. Wójcikowska B, Jaskóła K, Gąsiorek P (2013) *LEAFY COTYLEDON2 (LEC2)* promotes embryogenic induction in somatic tissues of *Arabidopsis*, via *YUCCA*-mediated auxin biosynthesis. *Planta* 238:425–440. <https://doi.org/10.1007/s00425-013-1892-2>
73. Wu G, Wei X, Wang X, Wei Y (2021) Changes in biochemistry and histochemical characteristics during somatic embryogenesis in *Ormosia henryi* Prain. *Plant Cell Tissue Organ Cult* 144:505–517. <https://doi.org/10.1007/s11240-020-01973-5>
74. Xiao Y, Chen Y, Ding Y (2018) Effects of *GhWUS* from upland cotton (*Gossypium hirsutum* L.) on somatic embryogenesis and shoot regeneration. *Plant Sci* 270:157–165. <https://doi.org/10.1016/j.plantsci.2018.02.018>
75. Yang X, Zhang X (2010) Regulation of somatic embryogenesis in higher plants. *Crit Rev Plant Sci* 29:36–57. <https://doi.org/10.1080/07352680903436291>
76. Yavuz C, Tillaboeva S, Bakhsh A (2020) Apprehending the potential of *BABY BOOM* transcription factors to mitigate cotton regeneration and transformation. *J Cotton Res* 3:29. <https://doi.org/10.1186/s42397-020-00071-3>
77. Zhang Y, Clemens A, Maximova SN, Guiltinan MJ (2014) The *Theobroma cacao* B3 domain transcription factor *TcLEC2* plays a dual role in control of embryo development and maturation. *BMC Plant Biol* 14:106. <https://doi.org/10.1186/1471-2229-14-106>
78. Zheng Q, Perry SE (2014) Alterations in the transcriptome of soybean in response to enhanced somatic embryogenesis promoted by orthologs of *AGAMOUS-Like15* and *AGAMOUS-Like18*. *Plant Physiol* 164:1365–1377. <https://doi.org/10.1104/pp.113.234062>
79. Zheng Y, Ren N, Wang H (2009) Global identification of targets of the arabidopsis MADS domain protein *AGAMOUS-LIKE15*. *Plant Cell* 21:2563–2577. <https://doi.org/10.1105/tpc.109.068890>
80. Zuo J, Niu QW, Frugis G, Chua NH (2002) The *WUSCHEL* gene promotes vegetative-to-embryonic transition in *Arabidopsis*. *Plant J* 30:349–359. <https://doi.org/10.1046/j.1365-3113X.2002.01289.x>

Figures

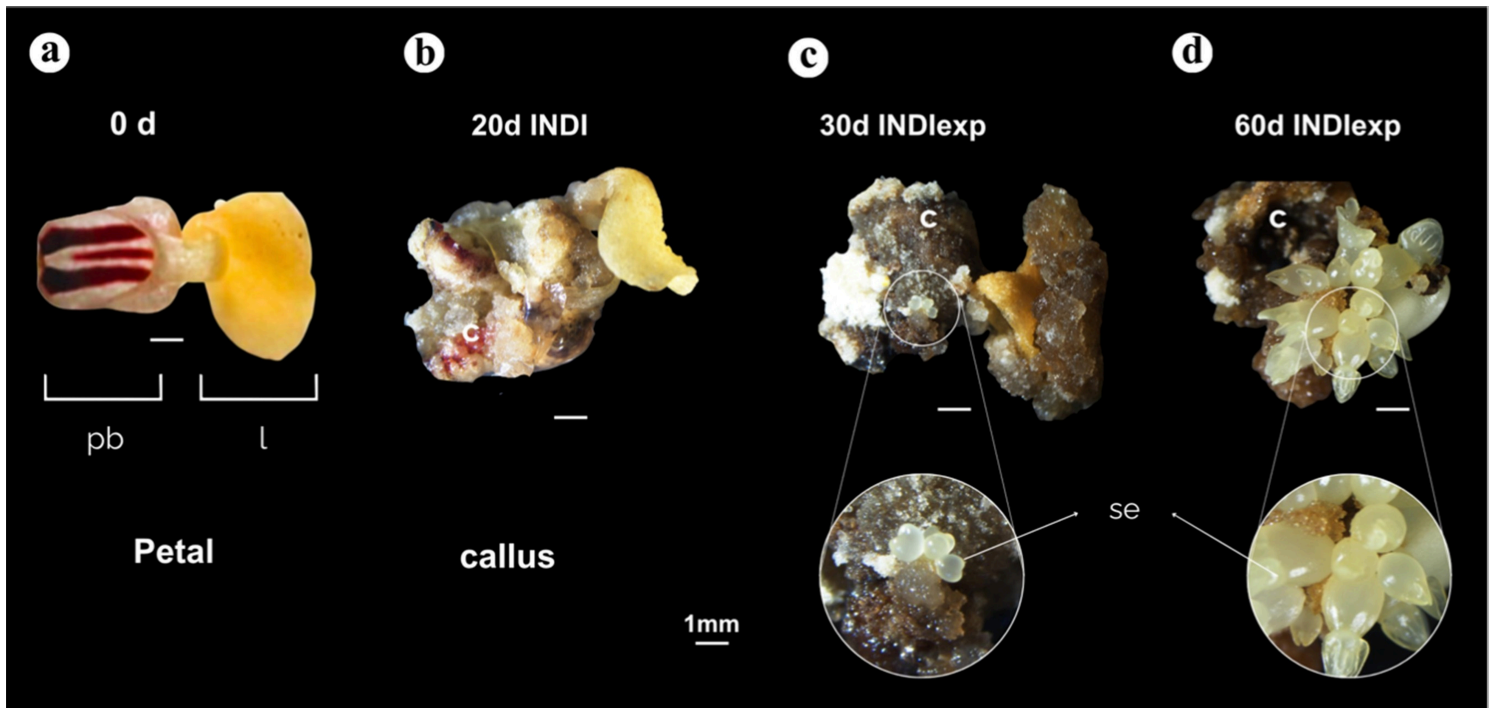


Figure 1

Morphological progression of somatic embryogenesis (SE) from petal explants of cacao. (a) Petal explant at day 0, showing the petal base (pb) and ligule (l) regions. (b) Cream-colored callus (c) formed after 20 days of culture on INDI medium. (c) Initial somatic embryos (se) developing from callus after 30 days on PGR-free INDlexp medium. (d) Numerous somatic embryos formed after 60 days on INDlexp medium. Circular insets show enlarged views of the developing somatic embryos. Scale bars = 1 mm.

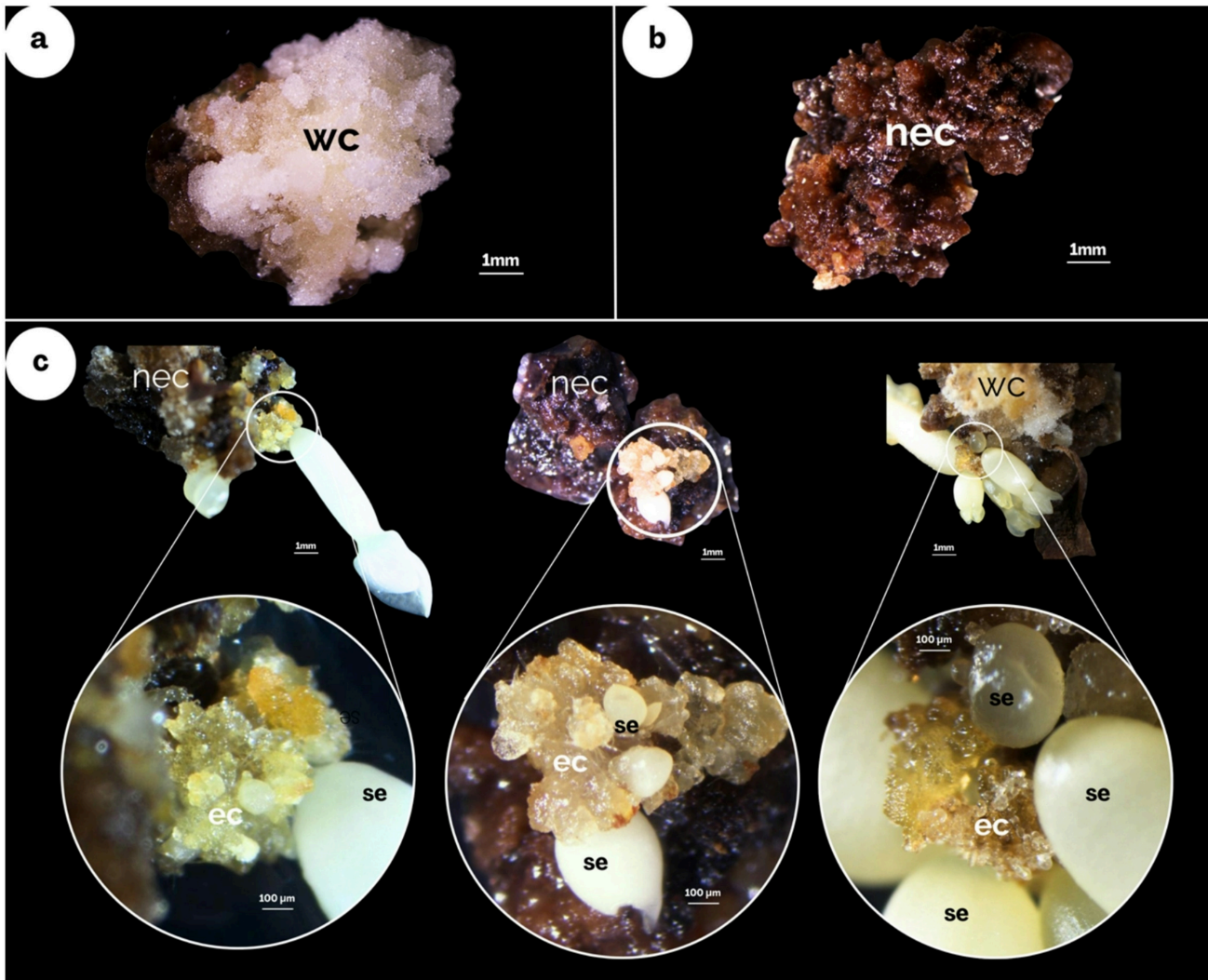


Figure 2

Morphological callus types identified in cacao after 60 days of culture on INDlexp medium. **(a)** White callus (WC). **(b)** Non-embryogenic callus (NEC). **(c)** Representative mixed calli containing WC and/or NEC regions together with embryogenic callus (EC), from which somatic embryos (se) developed. Outlined regions are shown at higher magnification in the corresponding circular insets. Scale bars = 1 mm (a, b), 2 mm (whole calli in c), and 100 μm (magnified insets in c).

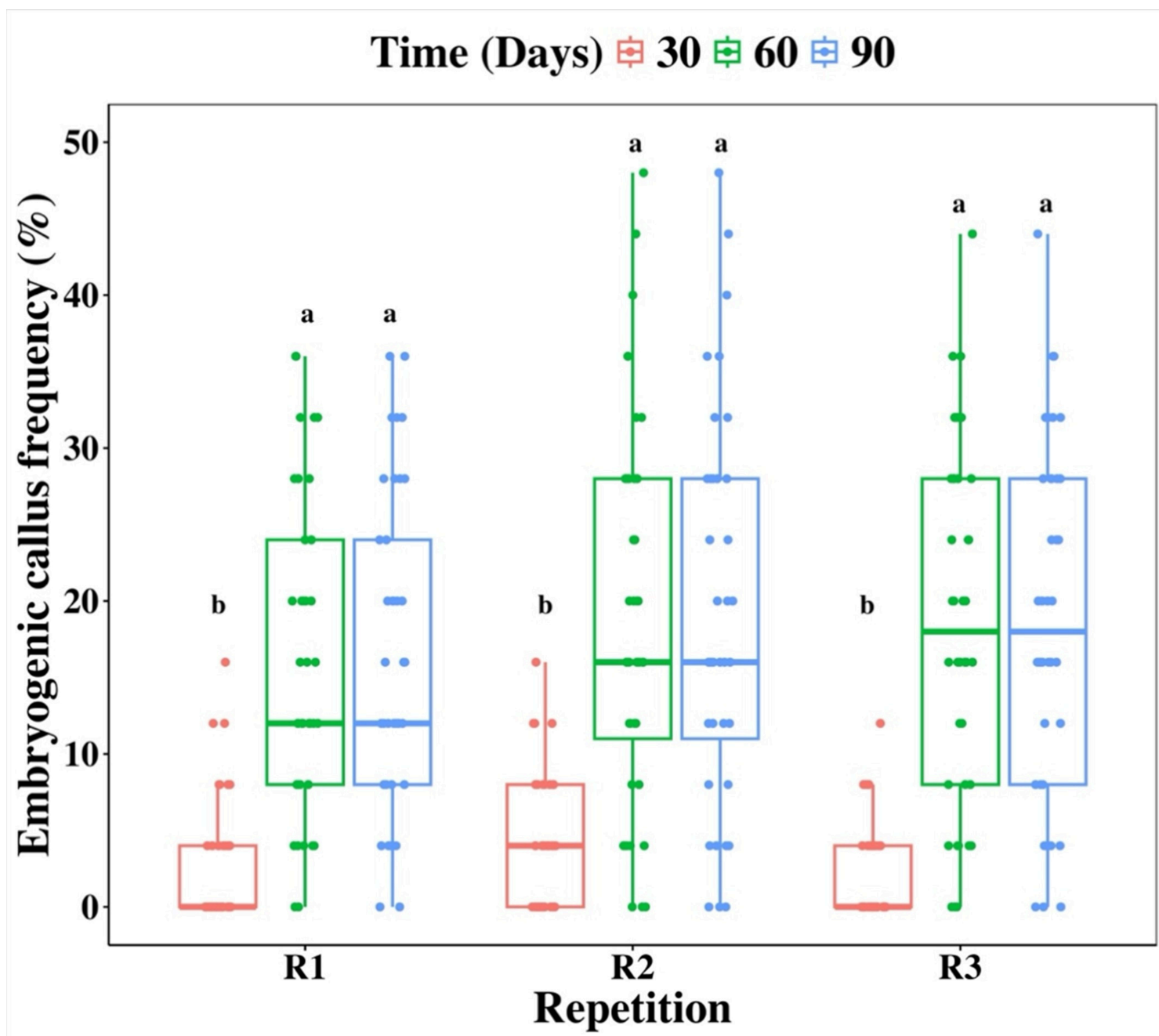


Figure 3

Embryogenic callus frequency (ECF) of calli from cacao after 30, 60, and 90 days of culture on INDlexp medium in three independent experimental repetition (R1–R3). Different lowercase letters indicate significant differences among culture times within each experimental repetition according to Fisher's least significant difference (LSD) test ($p < 0.05$).

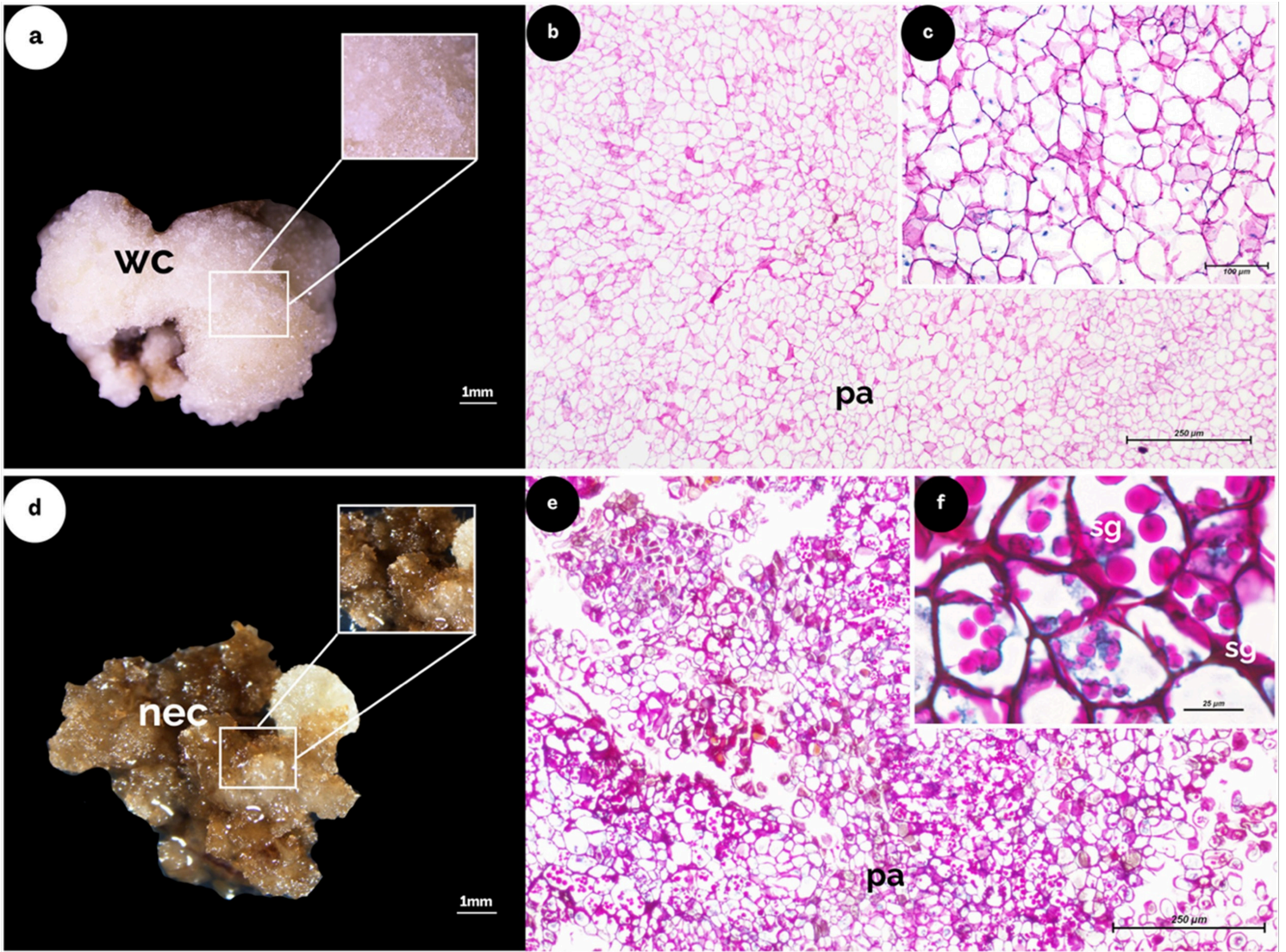


Figure 4

Cytohistological characterization of white callus (WC) and non-embryogenic callus (NEC) from cacao. **(a)** Macroscopic view of WC. **(b)** Histological section of WC showing parenchymatous cells (pa). **(c)** Higher-magnification view of the WC cells shown in b. **(d)** Macroscopic view of NEC. **(e)** Histological section of NEC showing parenchymatous cells (pa). **(f)** Higher-magnification view of the NEC cells shown in e, showing fuchsia-stained starch grains (sg). Insets in a and d show enlarged views of the callus surface. Scale bars = 1 mm (a, d), 250 μm (b, e), 100 μm (c), and 25 μm (f).

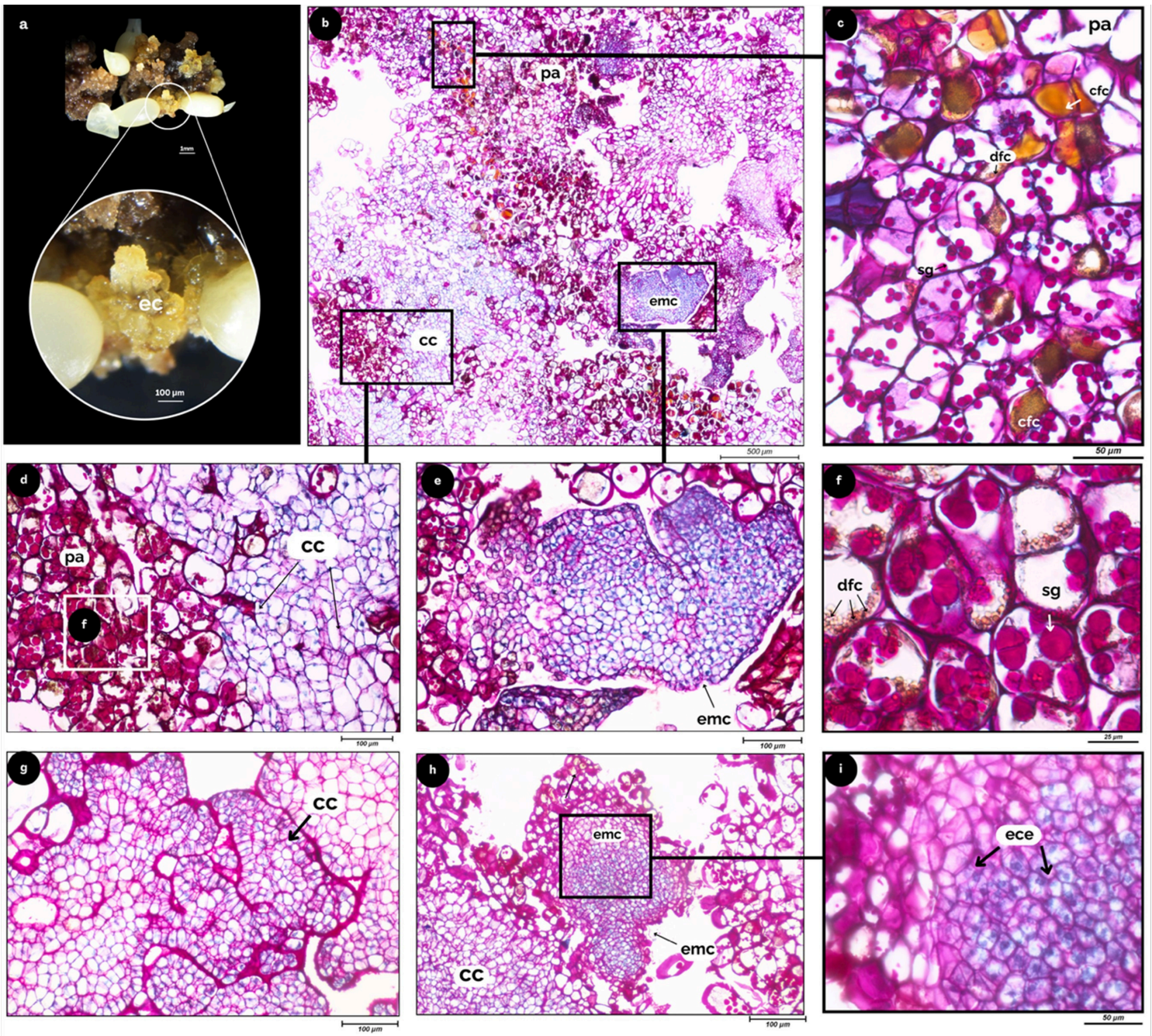


Figure 5

Cytohistological characterization of embryogenic callus (EC) from cacao, showing cellular features associated with the transition from parenchymatous tissue to embryogenic cell clusters. (a) Macroscopic view of EC within a mixed callus; the circular inset shows an enlarged view of the EC region. (b) Histological overview of EC showing parenchyma cells (pa), competent cells (cc), and embryogenic cell clusters (emc). (c) parenchyma cells containing abundant starch grains (sg) and two morphologically distinct forms of phenolic deposits: densely filled phenolic inclusions (cfc) and phenolic droplets (dfc). (d) Transitional region containing parenchyma cells and competent cells with conspicuous nuclei. (e) Embryogenic cell cluster associated with competent cells. (f) Higher-magnification view of the boxed region in d, showing phenolic droplets (dfc) and starch grains (sg). (g) Groups of competent cells becoming progressively separated from the surrounding tissue. (h) Well-defined embryogenic cell cluster separated from the surrounding tissue. (i) Higher-magnification view of the boxed region in h, showing

embryogenic cells (ece) within the cluster. Scale bars = 1 mm (a, main image), 100 μ m (a, inset), 500 μ m (b), 50 μ m (c, i), 100 μ m (d, e, g, h), and 25 μ m (f).

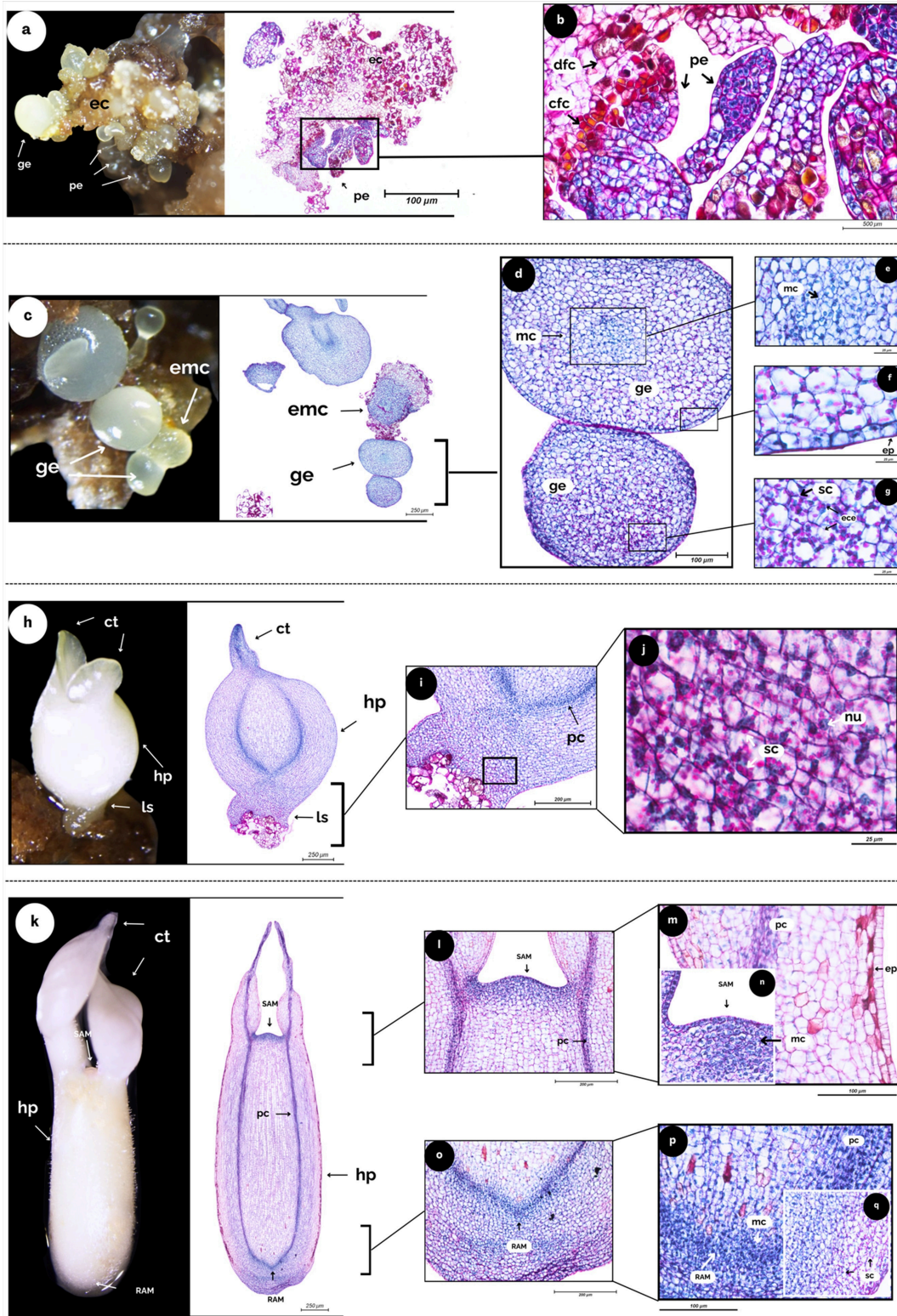


Figure 6

Cytohistological characterization of somatic embryo development in cacao, from the proembryo to the cotyledonary stage. *Proembryo stage*: (a) Macroscopic and histological views of proembryos (pe) associated with embryogenic callus (EC). (b) Higher-magnification view of a proembryo showing small cells with conspicuous

nuclei (nu), starch grains (sg), densely filled phenolic inclusions (cfc), and phenolic droplets (dfc). *Globular stage*: (c) Macroscopic and histological views of globular embryos (ge) arising from embryogenic cell masses (emc). (d) Globular embryos showing a meristematic center (mc). (e) Higher-magnification view of the meristematic center. (f) Epidermal layer (ep) surrounding the globular embryo. (g) Peripheral embryogenic cells (ece) containing starch grains and surrounding the meristematic center. *Early cotyledonary stage*: (h) Macroscopic and histological views of an early cotyledonary embryo showing cotyledons (ct), hypocotyl (hp), and a suspensor-like structure (ls). (i) Procambium (pc) differentiation. (j) Higher-magnification view of the suspensor-like structure showing conspicuous nuclei. *Intermediate cotyledonary stage*: (k) Macroscopic and histological views of an embryo exhibiting bipolar organization, with cotyledons, hypocotyl, and a root apical meristem (RAM). (l) Shoot apical meristem (SAM). (m) Epidermal layer and meristematic cells (mc) surrounding the SAM. (n) Higher-magnification view of the SAM. (o) Root apical meristem and procambium. (p) Higher-magnification view of the procambium. (q) Accumulation of starch grains at the base of the hypocotyl. Scale bars: 100 μ m (a), 500 μ m (b), 250 μ m (c), 100 μ m (d), 25 μ m (e, g), variable (f), 250 μ m (h), 200 μ m (i), 25 μ m (j), 250 μ m (k), 200 μ m (l), 100 μ m (m), 25 μ m (n), 200 μ m (o), 150 μ m (p).

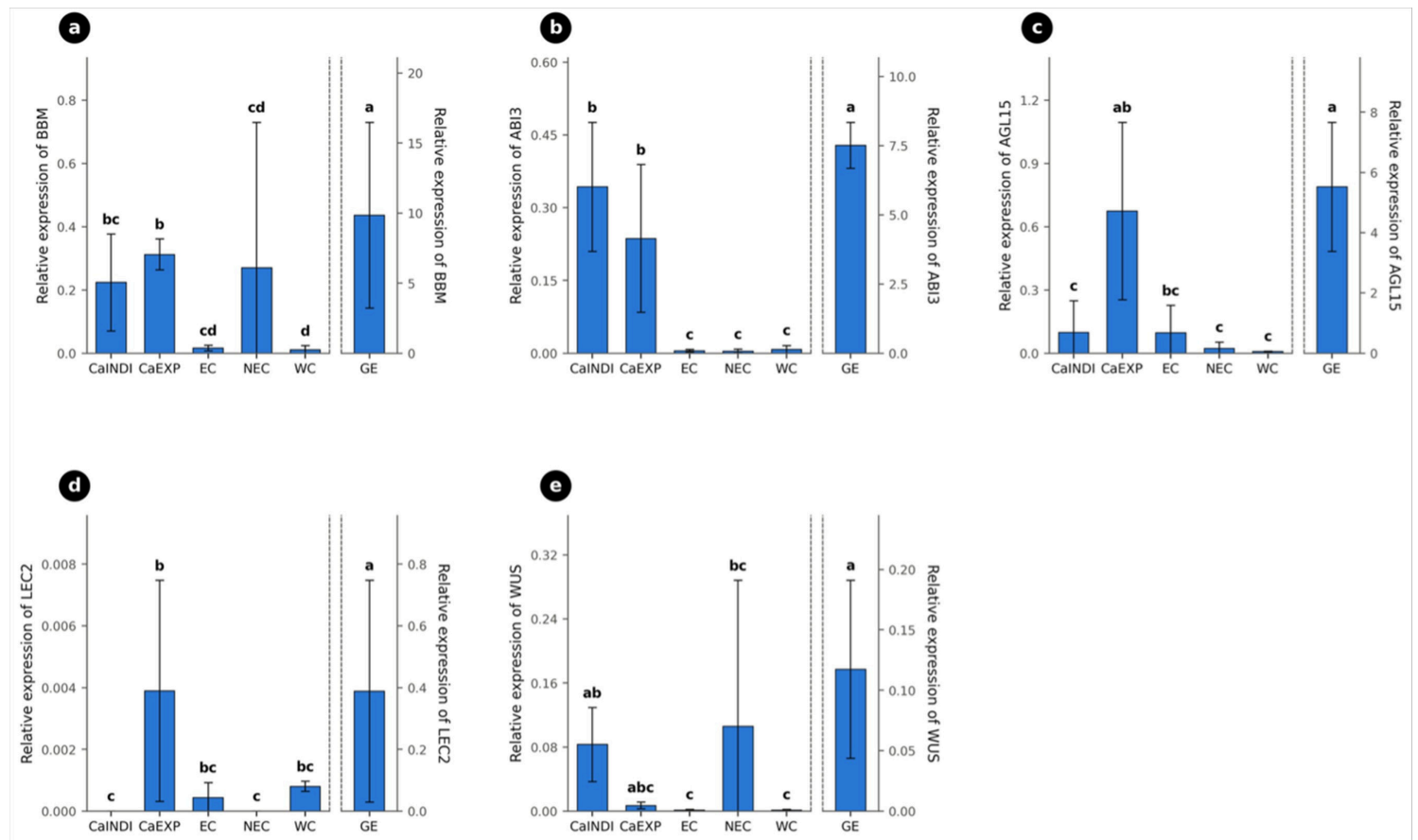


Figure 7

Relative expression of the transcription factor genes (a) *BBM*, (b) *ABI3*, (c) *AGL15*, (d) *LEC2* and (e) *WUS* in six tissue types of cacao: induction-phase callus (CaINDI), expression-phase callus (CaEXP), embryogenic callus (EC), non-embryogenic callus (NEC), white callus (WC), and globular embryos (GE). Expression levels were analyzed by qRT-PCR and the relative abundance of mRNA was normalized against the *WOX2* gene in the corresponding samples. The data represent mean of three biological replicates $\pm$ SD with three technical replicates. Different

lowercase letters indicate significant differences among tissue types according to Fisher's least significant difference (LSD) test ($p < 0.05$).

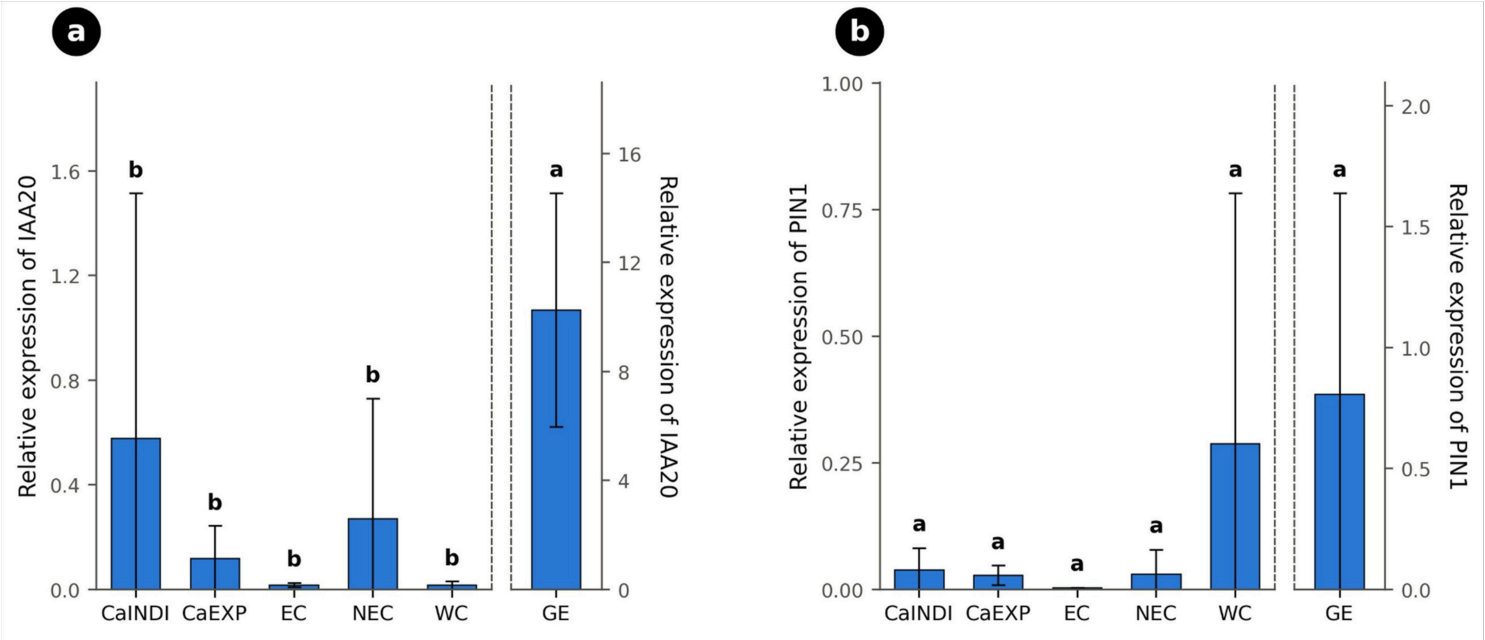


Figure 8

Relative expression of the auxin-related genes **(a) *IAA20*** and **(b) *PIN1***, in six tissue types of cacao: induction-phase callus (CaINDI), expression-phase callus (CaEXP), embryogenic callus (EC), non-embryogenic callus (NEC), white callus (WC), and globular embryos (GE). Expression levels were analyzed by qRT-PCR and the relative abundance of mRNA was normalized against the *WOX2* gene in the corresponding samples. The data represent mean of three biological replicates $\pm$ SD with three technical replicates. Different lowercase letters indicate significant differences among tissue types according to Fisher's least significant difference (LSD) test ($p < 0.05$).

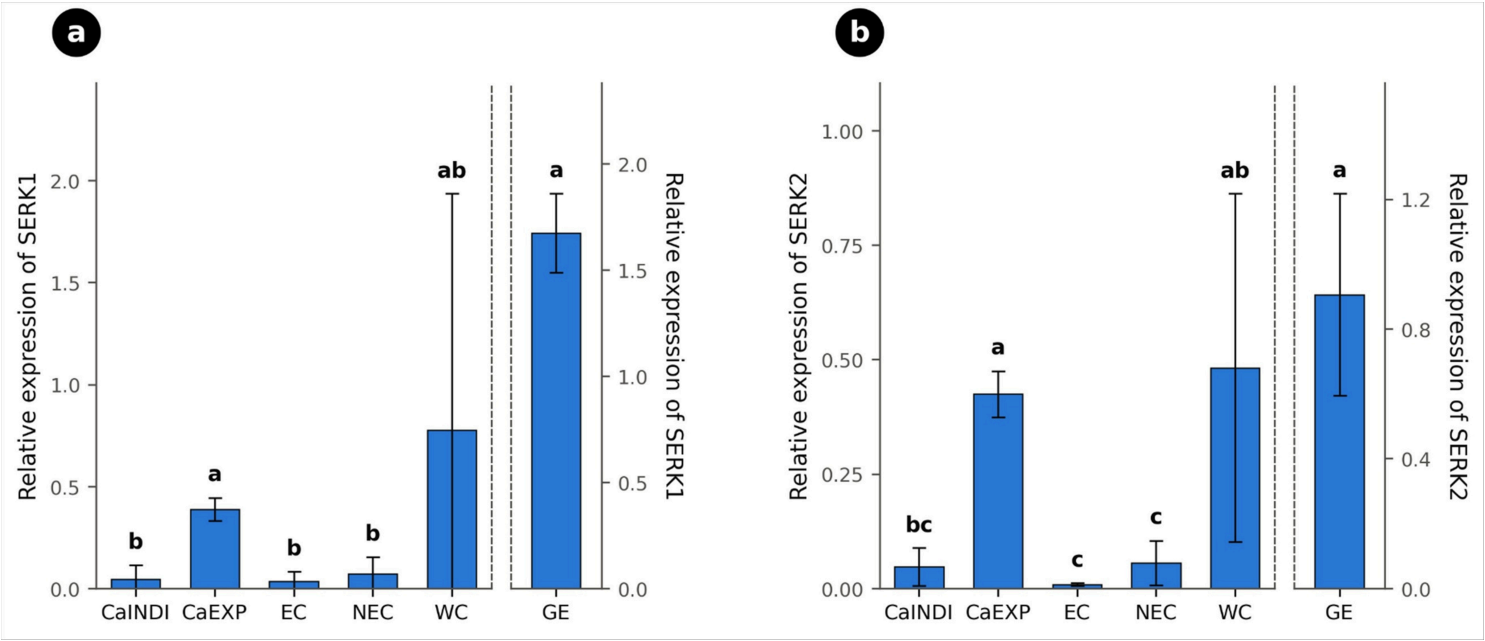


Figure 9

Relative expression of the membrane receptors genes (**a**) *SERK1* and (**b**) *SERK2*, in six tissue types of cacao: induction-phase callus (CaINDI), expression-phase callus (CaEXP), embryogenic callus (EC), non-embryogenic callus (NEC), white callus (WC), and globular embryos (GE). Expression levels were analyzed by qRT-PCR and the relative abundance of mRNA was normalized against the *WOX2* gene in the corresponding samples. The data represent mean of three biological replicates $\pm$ SD with three technical replicates. Different lowercase letters indicate significant differences among tissue types according to Fisher's least significant difference (LSD) test ($p < 0.05$).