

REVIEW

Open Access



Molecular insights into genetic transformation and somatic embryogenesis in cotton (*Gossypium hirsutum* L.)

AYUBOV Mirzakamol^{1*} , SHERMATOV Shukhrat¹, MIRZAKHMEDOV Mukhammad¹, UBAYDULLAEVA Khurshida¹, OBIDOV Nurdinjon¹, BURIEV Zabardast¹ and ABDURAKHMONOV Ibrokhim^{1,2}

Abstract

Genetic transformation and somatic embryogenesis (SE) are essential biotechnological tools for cotton improvement. Recent advances in transcriptomics, proteomics, epigenetics, and RNA interference (RNAi) have enabled us to better understand the SE regulatory pathways. This review summarizes the findings of several key research studies, focusing on the roles of transcription factors (e.g., GhAGL15s, GhL1L1, GhWOX), hormonal signaling (auxin and cytokinin), DNA methylation, and small RNAs in SE initiation and embryogenic callus formation. In addition, we discuss recent breakthroughs in *Agrobacterium*-mediated and RNAi-based transformation. Challenges such as genotype recalcitrance and somaclonal variation, along with new approaches to improve SE efficiency, are also addressed.

Keywords Somatic embryogenesis, Hormonal signaling, *Agrobacterium*-mediated transformation, Transcriptomics, RNA interference

Introduction

Cotton (*Gossypium hirsutum* L.) is a leading cash crop for the global textile industry with a significant impact on Uzbekistan's economy. However, yield and fiber quality vary widely due to genotype-specific traits and complex interactions among cultivars, farming practices, and environmental conditions. In the context of a rapidly changing climate, developing stress-resistant cotton varieties is essential. Genetic transformation and somatic embryogenesis (SE) enable rapid and precise improvement of cotton traits. Recent studies on transcription factors, hormonal signaling, epigenetic modifications,

and RNA interference (RNAi) have expanded our knowledge of SE regulation. Current breeding techniques could restrict productivity improvements due to an insufficiently diverse founding gene pool (Joshi et al. 2023; Sreedasyam et al. 2024). SE is the process by which somatic cells develop into embryos, a process that is critical in plant biotechnology. Cotton SE varies by genotype, with certain types (e.g., Yuzao 1, YZ) having high embryogenic ability (>80%) and others (e.g., Lumian 1, LM) being recalcitrant (<10%) (Guo et al. 2020). This distinction raises a significant challenge for transgenic breeding and genome editing, requiring a deeper understanding of the molecular and epigenetic pathways that drive SE.

Recent innovations in omics technologies (transcriptomics, proteomics, and epigenomics) have led to the identification of key regulators of SE, including transcription factors (GhAGL15, GhWOX, and GhL1L1), hormone signaling pathways (auxin and cytokinin), and DNA methylation dynamics, particularly CHH hypomethylation in embryogenic calli. *Agrobacterium*-mediated

*Correspondence:

Ayubov Mirzakamol
mirzo.ayubov@gmail.com

¹ Center of Genomics and Bioinformatics, Academy Sciences of Uzbekistan, Kibray Region, University Street - 2, Tashkent 111215, Republic of Uzbekistan

² Ministry of Agriculture of the Republic of Uzbekistan, Tashkent 100000, Uzbekistan



transformation and RNA interference have also produced stress-tolerant cotton varieties; however, low regeneration rates persist in elite cultivars. Emerging techniques, such as single-cell RNA sequencing (scRNA-seq) and nanoparticle-enhanced regeneration [e.g., copper/silver nanoparticles (Cu/AgNPs)], provide promising ways to overcome recalcitrance (Guo et al. 2024; Malik et al. 2021). This review combines molecular and biotechnological insights from recent studies to better understand SE mechanisms and transformation approaches.

Molecular mechanisms of somatic embryogenesis

Transcriptional and epigenetic regulation

Plant gene transmission is based on molecular processes that direct SE, in which *AGAMOUS-LIKE15* (*AGL15*) plays an important regulatory role. *GhAGL15-1*, -3, and -4 in pplant cotton (*G. hirsutum*) encode conserved nuclear MADS-box proteins that are expressed at their highest levels throughout embryogenic callus formation. *GhAGL15-1* and -3 respond to both 2,4-dichlorophenoxyacetic acid (2,4-D) and kinetin, whereas *GhAGL15-4* responds exclusively to 2,4-D. According to Xu et al. (2025), the contributions of a hierarchical gene regulation network, *AGL15*, to regeneration were demonstrated.

Overexpression of *GhAGL15* genes improves callus quality and stimulates embryogenic callus development, with *GhAGL15-4* having the greatest effect. These results suggest that *GhAGL15*, namely *GhAGL15-4*, is important for cotton SE and may improve the effectiveness of genetic transformation (Fig. 1a; Yang et al. 2014). Auxin is an important regulator of SE, but its molecular mechanism is poorly understood. Xu et al. (2019) showed that the cotton *NF-YB* gene (*GhL1L1*), which is expressed in embryonic tissues, affects auxin distribution and cell fate during SE. *GhL1L1* overexpression increased indole-3-acetic acid (IAA) levels, stimulated embryogenic cell proliferation, and changed the expression of the auxin transporter GhPIN1, whereas *GhL1L1* deletion decreased IAA levels and delayed development. *GhL1L1* promotes *GhPP2AA2* by binding its promoter, thereby regulating *GhPIN1* activity. Overall, *GhL1L1* regulates auxin distribution through the *GhPP2AA2*-*GhPIN1* pathway, influencing cell fate and enhancing SE efficiency in cotton (Fig. 1b; Xu et al. 2019).

Forty WUSCHEL-related homeobox (*WOX*) genes were found in upland cotton and categorized into ancient, intermediate, and modern clades, with the majority belonging to the current lineage that contains the

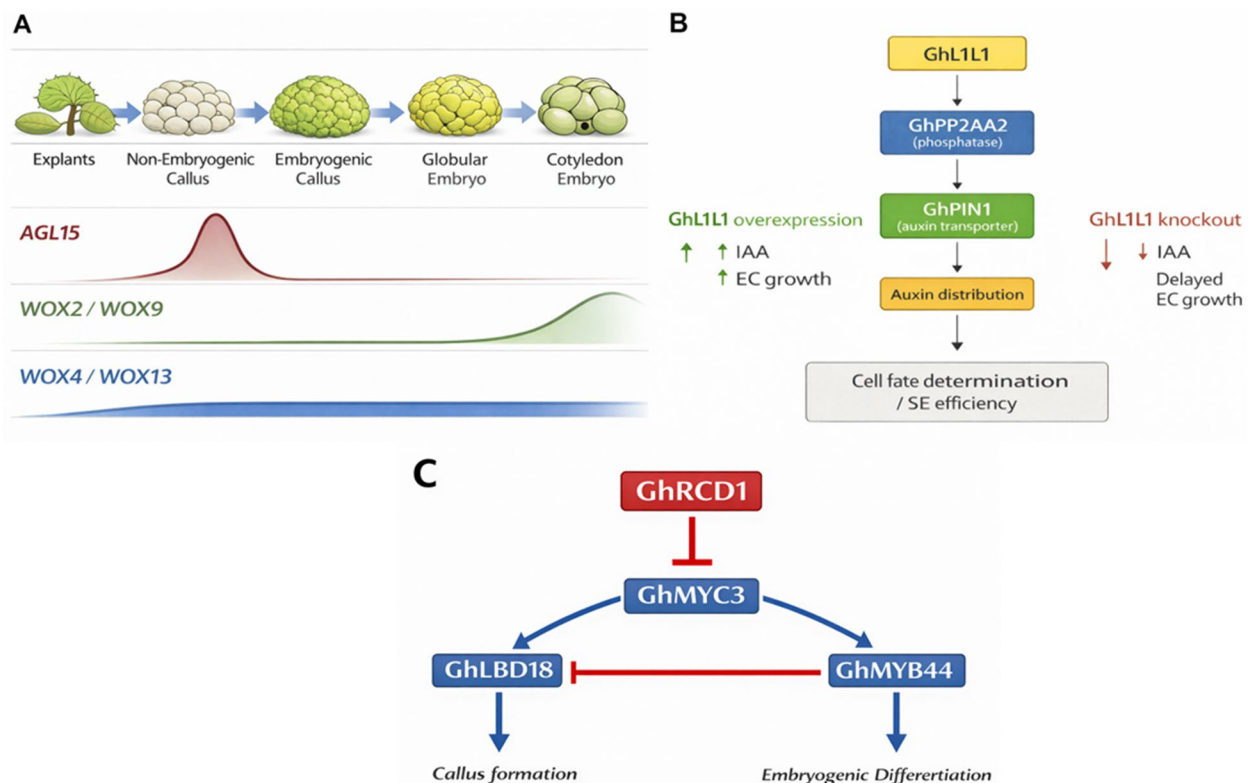


Fig. 1 Transcriptional regulatory network controlling somatic embryogenesis in cotton. **a** stage-specific transcriptional dynamics during somatic embryogenesis; **b** GhL1L1–GhPP2AA2–GhPIN1 module controls auxin distribution during somatic embryogenesis; **c** GhRCD1–GhMYC3–GhMYB44–GhLBD18 transcriptional module regulating somatic embryogenesis in cotton (*G. hirsutum*)

WUS-box domain. Gene expansion was primarily driven by duplications, and promoter analysis revealed an abundance of stress- and hormone-responsive elements. Only half of the *WOX* genes were expressed during SE in stage-specific patterns, with *GhWOX4* and *GhWOX13* being active throughout. Co-expression analysis linked *WOX* genes to hormone signaling, macromolecular processing, and metabolic pathways, highlighting their regulatory roles in cotton SE and stress tolerance (Sun et al. 2023). A tetrapartite module (*GhRCD1-GhMYC3-GhMYB44-GhLBD18*) maintains SE homeostasis by balancing callus proliferation (*GhMYC3/GhLBD18*) and embryogenic cell differentiation (*GhMYB44*). *GhRCD1* inhibits *GhMYC3*, hence inhibiting the transcription of downstream targets (*GhMYB44* and *GhLBD18*). *GhMYC3* overexpression enhances callus formation, whereas silencing has minimal morphological effects. *GhMYB44* (which is inhibited by *GhRCD1*) promotes embryogenic callus (EC) differentiation while preventing callus formation. *GhLBD18* (activated by *GhMYC3* and inhibited by *GhMYB44*) stimulates callus development. Reactive oxygen species (ROS) play a critical role in regulating cell fate. CRISPR deletion of *GhRCD1* mimics *GhMYC3* overexpression, accelerating SE via ROS dynamics. The study identifies a hierarchical regulatory cascade that fine-tunes SE through antagonistic interactions (*GhRCD1-GhMYC3*) and downstream effectors (*GhMYB44/GhLBD18*), in addition to ROS signaling. This molecular insight advances our understanding of cell fate reprogramming in cotton SE, opening the door to possibilities (e.g., *GhRCD1* editing) for optimizing in vitro regeneration for biotechnology applications (Fig. 1c; Yuan et al. 2023).

Kumar et al. (2021) investigated genotype-dependent SE in cotton by comparing transcriptomes from semi-recalcitrant Coker-312 and highly regenerative Jin668. Jin668 showed a significantly larger transcriptional response of 2 155 differentially expressed genes (DEGs) than Coker-312 (305 DEGs), indicating a greater SE capacity. Both genotypes shared 154 DEGs involved in embryo development, transportation, and transcription control. However, Jin668 revealed significantly increased expression of critical embryogenesis-promoting transcription factors, including BABY BOOM (BBM) and AGL15, whereas Coker-312 had fewer DEGs and lower transcription factor (TF) expression. These differences suggest that SE efficiency is primarily regulated at the transcriptional level, highlighting BBM and AGL15 as prospective targets for improving regeneration in resistant cotton cultivars.

A comparative RNA-Seq analysis of cotton genotypes with high- and low-SE efficiency showed key biochemical pathways that influence early SE induction. The findings indicate that SE initiation requires both protein

kinase activity (post-translational changes) and oxidoreductase activity (redox homeostasis). Stress-responsive transcription factors are differentially regulated, linking stress adaptation to embryogenic commitment. Somatic embryogenesis receptor kinases (SERKs) exhibit genotype-dependent expression, implying a role in SE competence. The auxin and ethylene signaling pathways are substantially implicated, with distinct regulatory mechanisms seen in high and low SE efficiency genotypes. SERKs and stress-responsive TFs may predict SE potential. Adjusting auxin/ethylene signaling or kinase activity may improve SE in recalcitrant crops. The identified modules (kinases, oxidoreductases, and hormone pathways) are promising targets for improving SE efficiency in genetically altered cotton. The study provides a key transcriptomic resource for understanding early SE triggers in cotton, with an emphasis on the interplay among hormone signaling, kinase cascades, and stress responses. The findings pave the way for better cotton biotechnology regeneration approaches, particularly for improving SE in resistant genotypes (Cao et al. 2017).

Further research reveals that DNA methylation dynamics, particularly CHH methylation, are important regulators of SE in cotton, distinguishing high- (YZ) and low-efficiency (LM) genotypes. SE competence varies by genotype: recalcitrant LM has increased mCG methylation, whereas YZ has hypomethylated CHH, which favors embryogenic redifferentiation. In contrast, LM hypermethylation reduces SE-related gene expression, forming epigenetic barriers to regeneration. In addition to the SE initiation-promoting methylation genes, differentially methylated genes (DMGs) were found and significantly increased in lipid biosynthesis and metabolic processes in YZ embryonic redifferentiation. Notably, unlike earlier research that focused on CHH hypermethylation, this study reveals CHH hypomethylation as a positive regulator and biomarker of SE potential. The study, which is the first genome-wide, single-base methylome analysis of cotton SE, redefines the role of DNA methylation in plant regeneration and proposes epigenetic editing—specifically targeted demethylation and lipid-pathway modulation—as promising strategies to improve SE in resistant genotypes and support precise epigenetic breeding (Guo et al. 2020).

The SE delay in cotton is genotype-dependent. Li et al. (2019) explored the epigenetic basis of accelerated regeneration in Jin668, a highly regenerative line derived from the successive regeneration acclimation (SRA) protocol. They used single-base resolution bisulfite sequencing and transcriptome analysis across SE stages to compare early-generation (R_0) and elite high-regeneration (R_4) plants. Jin668 (R_4) had considerably lower CHH DNA methylation compared with R_0 . This demonstrates that

SRA-induced hypomethylation enhances regeneration. CHH methylation increased from non-embryogenic to embryogenic callus via the RdDM and H3K9me2 pathways. Still, hypomethylation in later SRA generations activated hormone-related (auxin and cytokinin) and *WOX* genes required for embryogenesis. SE was consistently enhanced after treatment with the methylation inhibitor zebularine. This genome-wide cotton SE methylome shows that dynamic CHH methylation regulates regeneration competence and that SRA enhances SE through epigenetic reprogramming without genetic change (Li et al. 2019). Li et al. (2019) discovered that SRA-induced CHH hypomethylation epigenetically promotes the *WOX* and hormone pathways, explaining Jin668's better regeneration. Their findings provide: (i) a molecular blueprint for improving SE in crops; (ii) evidence that epigenetic modification (SRA, zebularine) can enhance regeneration efficiency; and (iii) a foundation for breeding elite cultivars without utilizing transgenic approaches.

Hormonal signaling

Auxin-response genes (*auxin response factor* (ARFs), *Auxin/Indole-3-Acetic Acid* (Aux/IAA)) and cytokinin signaling pathway components are upregulated during SE (Sun et al. 2019; Yang et al. 2012). Auxin early response genes, particularly *Aux/IAA*, play a critical role in cotton SE and may enhance cotton regeneration. Sun et al. (2019) identified 71 *ARF*, 86 *Aux/IAA*, 63 *Gretchen Hagen3* (*GH3*), and 194 small auxin-upregulated RNA (*SAUR*) genes in upland cotton. Their analysis showed that the *ARF*, *GH3*, and *SAUR* gene families diverged from *Arabidopsis*, whereas the *Aux/IAA* family remains genetically conserved. All four gene families were involved in embryogenic callus development. Notably, *Aux/IAA* genes exhibited higher expression levels in the embryogenic cultivar CCRI24 compared with the recalcitrant cultivar CCRI12, implying a role in embryogenesis.

Strong co-expression with SE-related genes indicates that auxin response genes may govern SE through interactions with these pathways (Sun et al. 2019). Yang et al. (2012) investigated the molecular events underlying this unique process in cotton, a common plant in Southeast Asia, using transcriptomic, biochemical, and histological techniques. RNA-Seq detected 5 076 DEGs during SE. Genes encoding metabolic pathways, secondary metabolite production, protein binding, and hydrolase activity were all enriched. Several transcription factor-encoding genes showed differential expression.

SE regulation relied heavily on complex auxin-related processes. These include IAA synthesis and metabolism, indole-3-butyric acid (IBA) metabolism and conjugate formation, auxin transport mechanisms, key regulators

(*ARF*, *Aux/IAA*, and *SAUR* genes), and Aux/IAA degradation pathways. Gene expression patterns have been connected to both morphological/histological changes and fluctuations in endogenous IAA levels. The paper reveals a key gene expression profile for cotton SE, identifying auxin signaling as the primary regulatory network orchestrating this distinct embryonic reprogramming process. The findings lay a solid foundation for future functional studies on plant embryogenesis.

The auxin-to-cytokinin ratio in culture media and endogenous hormone levels varied across sister lines and SE stages, indicating an important role in SE induction. Auxin and cytokinin production were largely found in DEGs. Gene expression patterns differ remarkably between sibling lines, indicating genotype-dependent SE responses. Genes that showed significant up- or down-regulation during SE were cloned, silenced, or overexpressed to investigate their roles. Auxin and cytokinin are the key regulators, although how they interact with other variables (such as stress responses and epigenetic modifiers) is unknown. Xu et al. (2013) provide the first global transcriptome insights into SE in cotton sister lines. It emphasizes the auxin-cytokinin relationship as a fundamental driver of SE while also encouraging additional research into cross-talk mechanisms, e.g., abscisic acid (ABA) and ethylene. Functional genomics approaches, such as clustered regularly interspaced short palindromic repeats (CRISPR) and virus-induced gene silencing (VIGS), are necessary to confirm the role of genes in SE. The study provides a critical transcriptome resource for cotton SE by focusing on hormone modulation and opening the path for mechanistic studies to enhance regeneration in recalcitrant genotypes (Xu et al. 2013).

GhmiR157a suppresses *GhSPL10* (a transcription factor) to maintain SE equilibrium. Overexpression of *GhSPL10* disrupts the balance by raising free auxin and ethylene levels and activating their signaling pathways. Flavonoid synthesis (via the *F3H* gene) stimulates callus initiation and proliferation. *GhmiR157a* reduces ethylene production, aminoethoxyvinylglycine (AVG) treatment, and callus development. Enhancing ethylene signaling ACC (1-aminocyclopropane-1-carboxylic acid) treatment, *EIN2* overexpression, or *CTR1* silencing) promotes flavonoid production and callus formation. Silencing *F3H* in *GhSPL10*-overexpressing lines prevents callus development, but applying exogenous flavonoids promotes callus growth, indicating that flavonoids stimulate the cell cycle during SE.

GhSPL10 overexpression affects SE through hormonal regulation (auxin-ethylene crosstalk). The *GhmiR157a-GhSPL10* interaction promotes proper SE progression. Wang et al. (2018) provide the first evidence linking miRNA-mediated *SQUAMOSA PROMOTER BINDING*

PROTEIN-LIKE (SPL) gene regulation to flavonoid-hormone interactions during cotton SE. These findings provide molecular targets (e.g., *GhSPL10*, *F3H*, and ethylene genes) for crop regeneration. CRISPR-edited *GhSPL10/miR157a* lines are being tested for their potential to improve SE efficiency. Further research is needed to investigate the relationships between flavonoids, auxin, and ethylene in other plant species. The *GhmiR157a-GhSPL10* module appears to act as a master regulatory switch, combining hormone signaling and flavonoid metabolism to regulate SE in cotton (Wang et al. 2018).

Proteome-level regulation of embryonic potential and regeneration

Quantitative proteomics approaches, e.g., isobaric tags for relative and absolute quantification (iTRAQ) and tandem mass tags (TMT), were employed to identify SE-associated proteins (Zhu et al. 2018; Guo et al. 2019). An iTRAQ-based proteomic analysis revealed significant protein dynamics across different SE stages in cotton, including cell dedifferentiation and redifferentiation. Zhu et al. (2018) identified 5 892 proteins in non-embryogenic callus (NEC), EC, and somatic embryos. Ninety-three and four percent of proteins were linked to catalytic activity, binding, transport, or structural roles. A total of 1 024 differentially expressed proteins (DEPs) were identified between NEC and EC, with 572 upregulated and 452 downregulated. Between EC and SE, 858 DEPs were found, including 647 downregulated and 221 upregulated proteins.

Plant growth and development are regulated by phytohormones, which have complex structures and contain small molecules. Although plants produce them in small quantities, they have significant physiological activity. Phytohormones are classified into two types based on their action: activators of plant growth and development or inhibitors, with auxins and cytokinins falling into the former category (Sosnowski et al. 2023). By promoting cell cycle reactivation and dedifferentiation, auxin and cytokinin contribute to the early cellular reprogramming events required for the development of embryogenic competence. Proteomic studies have identified GH3, an auxin-responsive protein, as being differentially expressed across developmental stages. GH3 abundance decreased in embryogenic callus and later increased in somatic embryos, a pattern consistent with the beneficial effect of low auxin concentrations on embryogenic transition (Jones et al. 2011).

Cytokinin has been shown to play a crucial role in initiating somatic embryogenesis and subsequent plant regeneration across a variety of plant species. Cotton proteins linked with cytokinin signaling, including the receptor CRE1, the phosphotransferase protein AHP, and the

response regulator B-ARR, showed variable expression patterns across embryonic stages. Both CRE1 and B-ARR were found to be increased in EC compared with NEC. At the same time, their expression was lower in somatic embryos than in EC (Zhu et al. 2018). The expression of Arabidopsis histidine phosphotransfer protein (AHP) was constant between NEC and EC.

Alongside hormonal signals, metabolic factors also contribute significantly to somatic embryogenesis. Polyamines are known to regulate diverse physiological processes, including cell division, stress responses, and tissue differentiation. Proteomic analysis revealed that enzymes involved in polyamine biosynthesis—S-adenosylmethionine decarboxylase (SAMDC), arginine decarboxylase (ADC), and spermine synthase (SPMS)—were progressively upregulated from NEC to EC and further into somatic embryos. These findings, supported by transcriptomic evidence, suggest that polyamines positively influence embryogenic competence in cotton. Furthermore, the localization of ADC in mitotically active embryonic cells, along with enhanced growth of embryogenic cultures in response to exogenous polyamine treatment, underscores its role in promoting somatic embryo formation by modulating developmental signaling and metabolic balance (Cheng et al. 2015).

Cotton SE involves a significant shift in energy metabolism from NEC to EC and somatic embryos, which is characterized by changes in oxidative phosphorylation proteins. Fatty acid biosynthesis is also substantially active, with the major enzymes (FabH, accC, FatB, and FabD) showing increased expression in EC and somatic embryos, consistent with previous observations in other plants. The effects of ectopic *LEC2* expression lend credence to the importance of the metabolism of lipids in embryo development. When gene expression is examined at the mRNA level and compared with actual protein levels, the agreement is limited ($r=0.29$), indicating that mRNA may not always convey the entire story. That is why combining transcriptomics and proteomics is critical for understanding cotton embryogenesis. This technique identified novel proteins required for SE production, providing prospective targets for improving embryogenesis and transformation efficiency in problematic cotton cultivars. The iTRAQ-based proteomic method successfully identified key molecular actors to enhance regeneration techniques (Zhu et al. 2018).

Cotton SE, a hormone-mediated stress response, establishes totipotency through complex and dynamic protein networks. Tandem mass tag-based proteomic profiling was employed to examine protein abundance patterns at successive stages of cotton SE, including NEC, primary embryogenic callus (PEC), and globular embryos (GEs). Protein identification yielded 9 369 distinct entries, of

which 6 730 were quantitatively profiled. Comparative analysis revealed 805 differentially abundant proteins (DAPs) between PEC and NEC, 295 between GE and PEC, and 1 242 between GE and NEC. Compared with NEC, 309 proteins exhibited increased abundance in PEC, while 496 showed reduced levels. Between GE and PEC, 174 proteins were upregulated, and 121 were downregulated. In the comparison of GE and NEC, 584 proteins were more abundant, and 658 were less abundant (Guo et al. 2019). Among the identified DAPs, several functional categories were closely associated with the SE process, particularly stress adaptation pathways, including peroxidase activity, photosynthesis, and nitrogen metabolism. During the early phase of somatic embryogenesis in cotton, several peroxidase proteins were found to be downregulated in PEC compared with NEC. This reduction in peroxidase abundance is likely associated with the suppression of tissue browning, which in turn promotes embryogenic competence.

As embryogenesis progressed, a gradual upregulation of proteins involved in the photosynthetic pathway was observed. Main components such as Photosystem I P700 apoprotein, oxygen-evolving enhancer protein 2, ATP synthase gamma chain, and chlorophyll a/b-binding protein were found to be expressed at higher levels in GE compared with PEC, suggesting the progressive development of photosynthetic capacity as embryogenesis advances. Nitrogen metabolism also contributes to embryogenic development, particularly through glutamate dehydrogenase (GDH), an enzyme involved in amino acid oxidation and ammonium detoxification. Various GDH isoforms, which perform both anabolic and catabolic functions, were significantly enriched in GE relative to PEC, with a trend of downregulation. These findings suggest that GDH plays a more prominent role during the redifferentiation stage of cotton embryogenesis (Guo et al. 2019).

This study generated the first TMT-based proteome atlas for cotton SE. The identification of DAPs offers a set of potential biomarkers linked to embryogenic competence and developmental progression. By highlighting important pathways such as stress response, hormone signaling, photosynthesis, and nitrogen metabolism, the findings explain the molecular landscape underlying totipotency acquisition. This work lays the foundation for targeted strategies to improve regeneration efficiency, particularly in genotypes with low embryogenic potential, using biotechnological approaches.

***Agrobacterium*-mediated transformation**

Agrobacterium-mediated transformation (AMT) is a valuable technique in plant biotechnology for developing transgenic plants, particularly in dicots like cotton.

Zhang (2019) suggested an optimal strategy for each step of the cotton transformation process. This approach included: 1) co-cultivation of cotton explants with *Agrobacterium*, 2) selection of stable transgenic cell lines, 3) regeneration of plants through somatic embryogenesis, and 4) confirmation of transgene integration and expression. Advances in these methodologies have significantly improved transformation efficiency. Plastid transformation has emerged as a promising avenue for high-level transgene expression. The tobacco chloroplast genome has been used to express a wide range of therapeutic proteins and agronomic traits. The effective transformation of cotton plastids marks a huge step forward in chloroplast engineering, surpassing earlier limits in transgenic expression, vector design, and regeneration. Scientists showed a significant boost in transformation efficiency over conventional approaches by combining the revolutionary "Double Gene/Single Selection (DGSS)" vector with an effective somatic embryogenesis process. The resulting transgenic lines showed normal growth, flowering, and seed output while maintaining maternal inheritance of transgenes, thereby addressing concerns about pollen-mediated gene flow. This innovation has the potential to reduce environmental concerns, increase insect resistance, and improve public perception of genetically modified cotton. Future uses of this technology may include other crops, opening the door for more sustainable and secure agricultural biotechnologies (Kumar et al. 2004).

RNA interference

RNAi was used in cotton to investigate gene function, enhance insect and disease resistance, and improve its fiber quality and early maturity (Salvador et al. 2022; Abdurakhmonov et al. 2014). *Agrobacterium*-mediated transformation and VIGS are both effective methods for delivering double-stranded RNA into cotton plants, each with specific advantages. While stable transgenic integration provides long-term RNAi effects for pest resistance, VIGS enables rapid evaluation of target gene silencing efficacy and phenotypic effects in herbivorous insects. The combination of these techniques, as demonstrated by Salvador et al. (2022), accelerates the development of RNAi-based cotton plant protection, bridging the gap between initial validation and long-term pest control solutions. Significant cotton research activities were examined, with a focus on the development of RNAi technology, accomplishments, limits, and future needs in the functional characterization of cotton genes. The need for initiatives to develop RNAi-based cotton cultivars, including risk and safety assessment, local and large-scale

field testing, and commercialization, was also highlighted (Abdurakhmonov et al. 2016).

Challenges and solutions

Genotype recalcitrance

The transcriptomic analysis of Jin668, a highly regenerable *G. hirsutum* cultivar, reveals essential molecular pathways involved in SE. DEGs between NEC and EC reveal key regulatory pathways, such as transcription factor upregulation (LEC1, BBM, FUS3, and AGAMOUS-LIKE15), phytohormone responses, and oxidative reduction processes, as well as downregulation of methylation-related genes. These results reveal vital insights into the genetic control of cotton SE and suggest potential targets to improve the effectiveness of plant regeneration (Wen et al. 2020). Great progress has been made in plant transformation techniques, but the successful regeneration of some genotypes remains a major challenge in attempts to transform recalcitrant species.

Regeneration ability has been shown in embryogenic genotypes. The results demonstrated that *GhSERK1* can be used to assess cotton plants' SE potential (de Cunha Soares et al. 2018). The *GhSERK1* gene is critical for embryo development, and overexpression increases embryogenic competence. Researchers employed reverse transcription quantitative polymerase chain reaction (RT-qPCR) and blotting assays to verify a short *GhSERK1* probe for identifying embryogenic cotton genotypes. Cotton genotypes with variable SE potential were studied using in vitro methods. Transcripts were found to be strongly expressed in embryogenic genotypes using RT-PCR and blotting with a nonradioactive probe.

Malik et al. (2021) studied recalcitrance to avoid callus-based genetic transformation. Copper nanoparticles (CuNPs) and AgNPs enhance embryogenic callus growth and regeneration in wheat mature embryo cultures, exceeding traditional treatments (CuSO₄ and AgNO₃). Optimized concentrations (e.g., 0.015 mg·L⁻¹ CuNPs + 4.0 mg·L⁻¹ AgNPs) increased regeneration efficiency by 21% compared with controls, demonstrating nanotechnology's capacity to overcome wheat's genetic transformation resistance and promote agricultural sustainability.

Overcoming genotype restrictions in cotton genetic transformation entails moving away from time-consuming SE, which is only effective on a few genotypes, such as 'Coker', and toward genotype-independent approaches that use in planta transformation or shoot-based regeneration. Key ways for accomplishing this include the use of shoot apical meristem (SAM) cells, sonication-assisted transformation, and optimizing tissue culture conditions for elite, recalcitrant cultivars. Shoot apical meristem cell-mediated transformation (SAMT) technology

successfully changed the use of previously resistant cotton genotypes (TM-1, XinHai56, and CCRI49). Nonetheless, efficient SAMT-based T-DNA delivery is still dependent on critical factors such as age or type of explant, infection time, and culture media (Ge et al. 2022).

Regeneration is commonly achieved using experimentally developed tissue culture methods based on totipotency induction, in which cells are first guided toward a root stem cell-like identity before being reprogrammed to take on a shoot fate. Alternatively, pluripotent or totipotent cells can undergo somatic embryogenesis to become full plantlets. Knowledge of root stem cell niche biology was used to directly utilize key stem cell factors, including *RETINOBLASTOMA (RBR)*, *SCARECROW (SCR)*, *SHORT ROOT (SHR)*, and members of the *AINTEGUMENTA-LIKE/PLETHORA (AIL/PLT)* and *WUSCHEL-related homeobox (WOX)* gene families, to induce regeneration in a manner similar to induced pluripotent stem cells. It was discovered that these variables work together to encourage SE-mediated regeneration and overcome recalcitrance in *Arabidopsis thaliana* and pepper (*Capsicum annuum*) (Wittmer et al. 2025). This novel technique has the potential to help cotton regenerate and overcome recalcitrance.

A telomere-to-telomere and single-cell RNA-sequencing

A telomere-to-telomere assembly of Jin668 and a high-quality assembly of YZ1, two highly regenerative allotetraploid cotton germplasms, were demonstrated. The Jin668 genome is complete, allowing for analysis of ~30.1 Mb of centromeric regions invaded by maize and Tekay retrotransposons, a ~8.1 Mb 5S rDNA array with 25 190 copies, and a ~75.1 Mb main 45S rDNA array with 8 131 copies. A comparison of regenerative and recalcitrant genotypes reveals changing transcriptional patterns and chromatin accessibility during the first regeneration process. Furthermore, genetic diversity has an effect on single-guide RNA (sgRNA) target sites, and Jin668 genome assembly showed a lower probability of off-target effects in CRISPR-based genome editing. The full Jin668 genome highlights the complexity of genomic areas and cotton regeneration while also improving genome editing precision (Xu et al. 2025). Plants exhibit strong totipotency, as demonstrated by single-cell differentiation during SE. It has been difficult to investigate cell identity and gene regulation in this complex process. Single-cell RNA-seq was employed in cotton (*G. hirsutum*) to characterize cellular topography and gene regulation during embryonic development. NEC and primary embryogenic calli (PEC) were heterogeneous and were split into four populations and six clusters. Totipotency was associated with specific cell states and

enriched pathways, including energy metabolism, signaling, and transcription factors. The SE-associated lipid transfer protein (*SELTP*) gene characterized embryogenic stages in a dose-dependent, parabolic way, suggesting its potential as a quantitative marker. RNA velocity and Scissor verified the pseudotemporal trajectory and the correctness of the data. The result offers marker genes and a cellular atlas for SE in plant regeneration (Guo et al. 2024). Since the lack of some SE-related TF motifs, including WOX, BBM, LEC1, ERF59, ERF107, and PLT5 in either the JASPAR2022 (Castro-Mondragon et al. 2022) or HOMER databases, Xu et al. (2025) gathered each of the SE-related motifs from the Plant Transcription Factor Database and converted them to Homer format using TFBStools.

Conclusion

SE and genetic transformation remain critical for cotton improvement, especially in light of climate change, genotype recalcitrance, and the requirement for quick trait deployment. Cotton SE is regulated by a complex network that includes key transcription factors (e.g., GhAGL15, GhWOX, GhL1L1, and BBM), phytohormone signaling, epigenetic reprogramming (particularly dynamic CHH DNA methylation), small RNAs, and metabolic changes that all contribute to embryogenic competence. Multi-omics research demonstrates that efficient SE necessitates coordinated transcriptional, post-transcriptional, and protein-level regulation, emphasizing the necessity of integrative rather than single-layer analysis. Recent epigenetic research has identified CHH hypomethylation as a good sign of SE potential and proven that epigenetic reprogramming can increase regeneration capacity without changing DNA sequence, providing a non-transgenic option for improvement. Although improvements in *Agrobacterium*-mediated transformation, plastid engineering, and RNAi have increased cotton biotechnology, their usefulness is restricted by the ability to regenerate elite cultivars. Emerging methods, like single-cell RNA sequencing, telomere-to-telomere genome assembly, nanotechnology-assisted regeneration, and focused manipulation of stem cell regulators, offer fresh insights into cellular heterogeneity and developmental paths during SE. Integrating high-resolution genomes, epigenomics, single-cell studies, and novel regeneration techniques is predicted to result in more efficient and genotype-independent cotton transformation systems.

Acknowledgements

We thank the research team at the Center for Genomic and Bioinformatics for their support in evaluating the study data presented in these papers.

Authors' contributions

Ayubov MS wrote the manuscript; Shermatov SE, Mirzakhmedov MK, and Ubaydullaeva KA acquired and analyzed word literature; Buriev ZT and Abdurakhmonov IY read and edited the paper critically, creating a subsection; and Abdurakhmonov IY thoroughly revised and approved the article. All authors read and approved the final manuscript.

Funding

This work was supported by the World Bank and the Government of the Republic of Uzbekistan under Grant «2/4» (REP-25112021/116) as part of the Modernizing Uzbekistan's National Innovation System (MUNIS) Project. The statements do not always represent the official position of the World Bank or the Government of the Republic of Uzbekistan.

Data availability

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 14 August 2025 Accepted: 19 February 2026

Published online: 03 April 2026

References

- Abdurakhmonov I, Buriev Z, Saha S, et al. Phytochrome RNAi enhances major fibre quality and agronomic traits of the cotton *Gossypium hirsutum* L. *Nat Commun*. 2014;5:3062. <https://doi.org/10.1038/ncomms4062>.
- Abdurakhmonov IY, Ayubov MS, Ubaydullaeva KA, et al. RNA interference for functional genomics and improvement of cotton (*Gossypium* sp.). *Front Plant Sci*. 2016;7:202. <https://doi.org/10.3389/fpls.2016.00202>.
- Cao A, Zheng Y, Yu Y, et al. Comparative transcriptome analysis of SE initial dedifferentiation in cotton of different SE capability. *Sci Rep*. 2017;7:8583. <https://doi.org/10.1038/s41598-017-08763-8>.
- Castro-Mondragon JA, Riudavets-Puig R, Raulusevičiute I, et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res*. 2022;50(D1):D165–73. <https://doi.org/10.1093/nar/gkab1113>.
- Cheng WH, Wang FL, Cheng XQ, et al. Polyamine and its metabolite H₂O₂ play a key role in the conversion of embryogenic callus into somatic embryos in upland cotton (*Gossypium hirsutum* L.). *Front Plant Sci*. 2015;6:1063. <https://doi.org/10.3389/fpls.2015.01063>.
- da Cunha Soares T, da Silva CRC, Chagas Carvalho JMF, et al. Validating a probe from *GhSERK1* gene for selection of cotton genotypes with somatic embryogenic capacity. *J Biotechnol*. 2018;270:44–50. <https://doi.org/10.1016/j.jbiotec.2018.02.002>.
- Ge X, Xu J, Yang Z, et al. Efficient genotype-independent cotton genetic transformation and genome editing. *J Integr Plant Biol*. 2022;65(4):907–17. <https://doi.org/10.1111/jipb.13427>.
- Guo H, Guo H, Zhang L, et al. Dynamic TMT-based quantitative proteomics analysis of critical initiation process of totipotency during cotton somatic embryogenesis transdifferentiation. *Int J Mol Sci*. 2019;20(7):1691. <https://doi.org/10.3390/ijms20071691>.
- Guo H, Fan Y, Guo H, et al. Somatic embryogenesis critical initiation stage-specific mCHH hypomethylation reveals epigenetic basis underlying embryogenic redifferentiation in cotton. *Plant Biotechnol J*. 2020;18(8):1648–50. <https://doi.org/10.1111/pbi.13336>.
- Guo H, Zhang L, Guo H, et al. Single-cell transcriptome atlas reveals somatic cell embryogenic differentiation features during regeneration. *Plant Physiol*. 2024;195(2):1414–31. <https://doi.org/10.1093/plphys/kiad107>.

- Jones B, Ljung K. Auxin and cytokinin regulate each other's levels via a metabolic feedback loop. *Plant Signal Behav.* 2011;6(6):901–4. <https://doi.org/10.4161/psb.6.6.15323>.
- Joshi B, Singh S, Tiwari GJ, et al. Genome-wide association study of fiber yield-related traits uncovers the novel genomic regions and candidate genes in Indian upland cotton (*Gossypium hirsutum* L.). *Front Plant Sci.* 2023;14:1252746. <https://doi.org/10.3389/fpls.2023.1252746>.
- Kumar S, Dhingra A, Daniell H. Stable transformation of the cotton plastid genome and maternal inheritance of transgenes. *Plant Mol Biol.* 2004;56(2):203–16. <https://doi.org/10.1007/s11103-004-2907-y>.
- Kumar S, Ruggles A, Logan S, et al. Comparative transcriptomics of non-embryogenic and embryogenic callus in semi-recalcitrant and non-recalcitrant upland cotton lines. *Plants.* 2021;10(9):1775. <https://doi.org/10.3390/plants10091775>.
- Li J, Wang M, Li Y, et al. Multi-omics analyses reveal epigenomics basis for cotton somatic embryogenesis through successive regeneration acclimation process. *Plant Biotechnol J.* 2019;17(2):435–50. <https://doi.org/10.1111/pbi.12988>.
- Malik WA, Mahmood I, Razzaq A, et al. Exploring potential of copper and silver nano particles to establish efficient callogenesis and regeneration system for wheat (*Triticum aestivum* L.). *GM Crops Food.* 2021;12(1):564–85. <https://doi.org/10.1080/21645698.2021.1917975>.
- Salvador R, Maskin L, Niz J, et al. RNAi expression in cotton for control of herbivorous insects. In: Vaschetto LM, editor. *RNAi strategies for pest management*. New York, USA: Humana; 2022. p. 217–33. https://doi.org/10.1007/978-1-0716-1633-8_17.
- Sosnowski J, Truba M, Vasileva V. The impact of auxin and cytokinin on the growth and development of selected crops. *Agriculture.* 2023;13(3):724. <https://doi.org/10.3390/agriculture13030724>.
- Sreedasyam A, Lovell JT, Mamidi S, et al. Genome resources for three modern cotton lines guide future breeding efforts. *Nat Plants.* 2024;10(6):1039–51. <https://doi.org/10.1038/s41477-024-01713-z>.
- Sun R, Wang S, Ma D, et al. Genome-wide analysis of cotton auxin early response gene families and their roles in somatic embryogenesis. *Genes.* 2019;10(10):730. <https://doi.org/10.3390/genes10100730>.
- Sun R, Zhang X, Ma D, Liu C. Identification and evolutionary analysis of cotton (*Gossypium hirsutum*) *WOX* family genes and their potential function in somatic embryogenesis. *Int J Mol Sci.* 2023;24(13):11077. <https://doi.org/10.3390/ijms241311077>.
- Wang L, Liu N, Wang T, et al. The *GhmiR157a-GhSPL10* regulatory module controls initial cellular dedifferentiation and callus proliferation in cotton by modulating ethylene-mediated flavonoid biosynthesis. *J Exp Bot.* 2018;69(5):1081–93. <https://doi.org/10.1093/jxb/erx475>.
- Wen L, Li W, Parris S, et al. Transcriptomic profiles of non-embryogenic and embryogenic callus cells in a highly regenerative upland cotton line (*Gossypium hirsutum* L.). *BMC Dev Biol.* 2020;20:25. <https://doi.org/10.1186/s12861-020-00230-4>.
- Xu Z, Zhang C, Zhang X, et al. Transcriptome profiling reveals auxin and cytokinin regulating somatic embryogenesis in different sister lines of cotton cultivar CCR124. *J Integr Plant Biol.* 2013;55(7):631–42. <https://doi.org/10.1111/jipb.12073>.
- Xu J, Yang X, Li B, et al. *GhL1L1* affects cell fate specification by regulating GhPIN1-mediated auxin distribution. *Plant Biotechnol J.* 2019;17(1):63–74. <https://doi.org/10.1111/pbi.12947>.
- Xu Z, Wang G, Zhu X, et al. Genome assembly of two allotetraploid cotton germplasms reveals mechanisms of somatic embryogenesis and enables precise genome editing. *Nat Genet.* 2025;57:2028–39. <https://doi.org/10.1038/s41588-025-02258-3>.
- Yang X, Zhang X, Yuan D, 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.* 2012;12:110. <https://doi.org/10.1186/1471-2229-12-110>.
- Yang Z, Li C, Wang Y, et al. *GhAGL15s*, preferentially expressed during somatic embryogenesis, promote embryogenic callus formation in cotton (*Gossypium hirsutum* L.). *Mol Genet Genomics.* 2014;289(5):873–83. <https://doi.org/10.1007/s00438-014-0856-y>.
- Yuan J, Liu X, Zhao H, et al. GhRCD1 regulates cotton somatic embryogenesis by modulating the GhMYC3-GhMYB44-GhLBD18 transcriptional cascade. *New Phytol.* 2023;240(1):207–23. <https://doi.org/10.1111/nph.19120>.
- Zhang B. *Agrobacterium*-mediated genetic transformation of cotton. In: Zhang B, editor. *Transgenic cotton*. New York, USA: Humana; 2019. p. 19–33. https://doi.org/10.1007/978-1-4939-8952-2_2.
- Zhu HG, Cheng WH, Tian WG, et al. iTRAQ-based comparative proteomic analysis provides insights into somatic embryogenesis in *Gossypium hirsutum* L. *Plant Mol Biol.* 2018;96:89–102. <https://doi.org/10.1007/s11103-017-0681-x>.