

**MODULAÇÃO DO PERFIL PROTEÔMICO E FOSFOPROTEÔMICO
DURANTE A EMBRIOGÊNESE SOMÁTICA DE CANA-DE-AÇÚCAR E
MAMOEIRO**

LUCAS RODRIGUES XAVIER

**UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY
RIBEIRO – UENF**

CAMPOS DOS GOYTACAZES – RJ

MAIO – 2025

**MODULAÇÃO DO PERFIL PROTEÔMICO E FOSFOPROTEÔMICO
DURANTE A EMBRIOGÊNESE SOMÁTICA DE CANA-DE-AÇÚCAR E
MAMOEIRO**

LUCAS RODRIGUES XAVIER

“Tese apresentada ao Centro de
Biociências e Biotecnologia da
Universidade Estadual do Norte Fluminense
Darcy Ribeiro, como parte dos requisitos
para a obtenção do título de Doutor em
Biotecnologia Vegetal.”

Orientador: Dr. Vanildo Silveira

**UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY
RIBEIRO – UENF**

CAMPOS DOS GOYTACAZES – RJ

MAIO – 2025

X3

Xavier, Lucas Rodrigues.

Modulação do perfil proteômico e fosfoproteômico durante a embriogênese somática de cana-de-açúcar e mamoeiro / Lucas Rodrigues Xavier. - Campos dos Goytacazes, RJ, 2025.

264 f. : il.

Inclui bibliografia.

Tese (Doutorado em Biotecnologia Vegetal) - Universidade Estadual do Norte Fluminense Darcy Ribeiro, Centro de Biociências e Biotecnologia, 2025.

Orientador: Vanildo Silveira.

Coorientador: Vitor Batista Pinto.

1. Micropropagação. 2. Embriogênese vegetal. 3. Ômicas. 4. Biotecnologia. 5. Biologia Molecular. I. Universidade Estadual do Norte Fluminense Darcy Ribeiro. II. Título.

CDD - 660.6

MODULAÇÃO DO PERFIL PROTEÔMICO E FOSFOPROTEÔMICO DURANTE A EMBRIOGÊNESE SOMÁTICA DE CANA-DE-AÇÚCAR E MAMOEIRO

LUCAS RODRIGUES XAVIER

“Tese apresentada ao Centro de Biociências e Biotecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte dos requisitos para a obtenção do título de Doutor em Biotecnologia Vegetal.”

Aprovada em 16 de maio de 2025.

Comissão Examinadora:

Documento assinado digitalmente
 **GONCALO APOLINARIO DE SOUZA FILHO**
Data: 23/09/2025 11:22:34-0300
Verifique em <https://validar.iti.gov.br>

Gonçalo Apolinário de Souza Filho (D.Sc., Biociências e Biotecnologia) – UENF

Documento assinado digitalmente
 **FABIO LOPES OLIVARES**
Data: 23/09/2025 11:09:26-0300
Verifique em <https://validar.iti.gov.br>

Fábio Lopes Olivares (D.Sc., Ciências do Solo) – UENF

Documento assinado digitalmente
 **DANILO DE MENEZES DALOSO**
Data: 23/09/2025 10:16:15-0300
Verifique em <https://validar.iti.gov.br>

Danilo de Menezes Daloso (D.Sc., Fisiologia Vegetal) – UFC

Documento assinado digitalmente
 **VANILDO SILVEIRA**
Data: 23/09/2025 09:51:25-0300
Verifique em <https://validar.iti.gov.br>

Vanildo Silveira (D.Sc., Biotecnologia) – UENF

Quem conhece os homens é inteligente

Quem conhece a si mesmo é iluminado

Vencer os homens é ter força

Quem vence a si mesmo é forte

Quem sabe contentar-se é rico

Agir fortemente é ter vontade

Quem não perde a sua residência, perdura

Quem morre, mas não perece, eterniza-se

TAO TE CHING, capítulo 33.

Decido este trabalho à Crelia Xavier,
Antônio Xavier e Palmieres Xavier, meus
queridos familiares. Sou iluminado porque
sei o quanto amo vocês.

Agradecimentos

Ao Vanildo Silveira por me orientar, discutir e acreditar nas minhas ideias. Também agradeço por todo o cuidado e zelo comigo e todos seus alunos, sempre disponibilizando computadores novos para nosso trabalho no laboratório, aprovando financiamentos importantes para nossa rotina de uso de espectrômetros de massas e também nos direcionando para um futuro de sucesso. Te admiro muito e sempre serei grato, professor!

Agradeço ao Vitor Pinto, meu coorientador, por me aconselhar, discutir desenhos experimentais, corrigir meus manuscritos e, principalmente, ser meu amigo.

Agradeço ao professor Jay J. Thelen por me receber e aconselhar durante meu estágio de doutorado sanduíche no seu laboratório no *Bond Life Sciences* da *University of Missouri – EUA*.

Agradeço aos meus colaboradores e amigos de laboratório, Lucas Gomes, Carlos Eduardo, Karina Viana, Laura Eliza, Daniel Pacheco, Roberta da Paschoa, Caio Corrêa e Laís da Conceição. Foi muito bom beber café e discutir ciência com vocês todas as manhãs!

Agradeço à Claudete Santa-Catarina e Margareth Paes por todo suporte e empenho na coordenação e secretaria do PPGBV.

Agradeço ao professor Gonçalo Apolinário por todo o cuidado com o prédio de Biologia Integrativa da UENF, além de sempre planejar e implementar melhorias que impactam positivamente o desempenho, segurança e bem-estar dos pesquisadores que utilizam este espaço.

Agradeço aos contribuintes por darem suporte à educação pública e de qualidade no Brasil através do pagamento de seus impostos.

Agradeço à FAPERJ por financiar minha bolsa de doutorado nota 10 e também à CAPES por financiar minha bolsa de doutorado sanduíche nos Estados Unidos.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Sumário

Introdução geral	1
1. Revisão de literatura.....	1
2. Revisão em Bancos de Propriedade Intelectual	15
3. Objetivos	22
3.1 Geral	22
3.2 Específicos	22
4. Referências bibliográficas	23
Capítulo 1. Proteomics in crop improvement	30
Abstract.....	31
Proteomics toolbox for crop breeding.....	32
A workflow for plant bottom-up proteomics	33
Advances and Challenges.....	37
Proteomics Studies in Abiotic Stress	38
Proteomics and biotic stress	52
Heterosis	69
Future prospects	74
References	74
Capítulo 2. Time-dependent proteomic signatures associated with embryogenic callus induction in <i>Carica papaya</i> L.	89
Abstract.....	91
Introduction	92
Results	93
Discussion	100
Materials and Methods	104
Conclusions.....	107
References	108
Capítulo 3. Protein networks mediating embryogenic callus induction in <i>Carica papaya</i> L. respond to dose-dependent 2,4-dichlorophenoxyacetic acid treatment	111
Abstract.....	112

Introduction	113
Materials and Methods	114
Results	117
Discussion	123
Conclusions.....	128
References	129
Capítulo 4. The role of protein phosphorylation in auxin-mediated totipotency and somatic embryogenesis of <i>Carica papaya</i> L.....	134
Abstract.....	135
Introduction	136
Material and Methods	138
Results	140
Discussion	150
Conclusions.....	154
References	155
Capítulo 5. Time-based proteomic and epigenetic analysis highlights key components driving sugarcane embryogenic callus initiation	160
Abstract.....	161
Graphical abstract.....	162
Introduction	162
Material and Methods	164
Results	169
Discussion	175
Conclusions.....	181
References	182
Capítulo 6. Integration of epigenetic remodeling and phosphorylation networks drives embryogenic identity in sugarcane (<i>Saccharum</i> spp.) callus	187
Abstract.....	188
Graphical abstract.....	189
Introduction	189

Material and Methods	191
Results	194
Discussion	206
Conclusions.....	215
References	216
Capítulo 7. Disruption of auxin transport impairs somatic embryo development and proteome dynamics in sugarcane	223
Abstract.....	224
Graphical abstract.....	225
Introduction	225
Materials and Methods	227
Results	233
Discussion	240
Conclusions.....	246
References	246
Perspectivas Futuras.....	252

Resumo

Este estudo investigou os mecanismos moleculares subjacentes à embriogênese somática (ES) em cana-de-açúcar (*Saccharum* spp.) e mamoeiro (*Carica papaya* L.), com foco na regulação hormonal, epigenética e fosforilação de proteínas. Em cana-de-açúcar, o uso de 5-Azacidina (5-AzaC), um agente hipometilante, combinado com 2,4-D, resultou em maior competência embriogênica, evidenciada pelo aumento no número de embriões somáticos maduros. Análises proteômicas e fosfoproteômicas identificaram reguladores-chave envolvidos em remodelamento de cromatina, manutenção de meristemas e transporte de auxina, sugerindo um mecanismo epigenético-fosforilação dependente. Em outro experimento, o inibidor de transporte de auxina TIBA afetou negativamente a diferenciação de calos embriogênicos, alterando a abundância de proteínas associadas ao desenvolvimento embriogênico e ao metabolismo energético, além de indicar uma possível interação entre auxinas e poliaminas. Em mamoeiro, análises integradas revelaram que proteínas como PLT2, PIN1-like e GTE4 são moduladas por fosforilação e estão ligadas à aquisição da competência embriogênica, fornecendo marcadores robustos para regeneração. Estes resultados apontam caminhos moleculares críticos para a otimização da micropropagação e transformação genética em culturas de interesse agrícola. As descobertas também têm implicações na produção de sementes sintéticas e na engenharia genética de plantas tropicais, com potencial de aplicação em protocolos biotecnológicos utilizados por empresas brasileiras. Assim, o estudo contribui significativamente para o avanço científico e tecnológico da embriogênese somática, com foco na sustentabilidade e no melhoramento genético de culturas recalcitrantes.

Abstract

This study explored the molecular mechanisms underlying somatic embryogenesis (SE) in sugarcane (*Saccharum* spp.) and papaya (*Carica papaya* L.), focusing on hormonal regulation, epigenetics, and protein phosphorylation. In sugarcane, treatment with 5-Azacitidine (5-AzaC), a DNA hypomethylating agent, combined with 2,4-D, significantly enhanced embryogenic competence, increasing the number of mature somatic embryos. Proteomic and phosphoproteomic analyses identified key regulators involved in chromatin remodeling, meristem identity, and auxin transport, suggesting a phosphorylation-dependent epigenetic mechanism. In another experiment, the auxin transport inhibitor TIBA negatively affected embryogenic callus differentiation, altering the abundance of proteins associated with embryonic development and energy metabolism, and suggesting a possible link between auxins and polyamines. In papaya, integrated analyses revealed that proteins such as PLT2, PIN1-like, and GTE4 are modulated by phosphorylation and associated with the acquisition of embryogenic competence, providing robust markers for plant regeneration. These findings reveal critical molecular pathways for optimizing micropropagation and genetic transformation protocols in agriculturally important crops. The results also have implications for the production of synthetic seeds and genetic engineering of tropical plants, with potential application in biotechnological platforms used by Brazilian companies. Thus, this study significantly contributes to the scientific and technological advancement of somatic embryogenesis, with a focus on sustainability and the genetic improvement of recalcitrant crops.

Introdução geral

1. Revisão de literatura

1.1 Embriogênese vegetal

O desenvolvimento inicial do embrião em Angiospermas inicia-se com a fertilização da célula-ovo e a consequente ativação do genoma zigóticos [1]. Durante esse processo, uma célula-ovo de grande dimensão funde-se com uma célula espermática pequena, desencadeando a plasmogamia (fusão dos citoplasmas), seguida pela cariogamia (fusão dos pró-núcleos masculino e feminino) [2]. Após a fusão nuclear, ocorre intensa demetilação do DNA genômico, remodelamento da cromatina e reprogramação transcricional, culminando na ativação do zigoto [1]. Essa ativação é sucedida por uma divisão celular assimétrica, gerando uma célula basal grande e uma célula apical menor [1]. A divisão assimétrica é precedida pela polarização do zigoto, caracterizada pelo deslocamento do núcleo para a região apical e pela retenção de um grande vacúolo na região basal [2].

A embriogênese em dicotiledôneas e monocotiledôneas diverge em aspectos fundamentais. Em dicotiledôneas, como *Arabidopsis thaliana*, a célula apical origina o embrião, enquanto a célula basal dá origem ao suspensor e a parte do meristema radicular [3]. Em contrapartida, nas monocotiledôneas, ambas as células (basal e apical) contribuem para a formação do proembrião, tornando sua distinção morfológica mais complexa [2]. Além disso, embriões de monocotiledôneas possuem estruturas específicas, como o escutelo, responsável por proteger o meristema apical em desenvolvimento, e o coleóptilo, estrutura equivalente ao cotilédono das dicotiledôneas [1]. Assim, as principais diferenças entre esses grupos incluem a ausência de suspensor em monocotiledôneas, padrões de divisão celular menos previsíveis e a presença de um eixo dorso-ventral bem estabelecido [1].

Durante a embriogênese de monocotiledôneas, a divisão do zigoto origina um embrião bicelular, com uma célula basal volumosa e vacuolizada, enquanto a célula apical se divide e dá origem aos estágios subsequentes de pró-embrião, culminando no estágio de transição [3]. Nesse estágio, a padronização adaxial-abaxial torna-se evidente, e os primórdios do coleóptilo e do escutelo são observáveis. Por outro lado, em dicotiledôneas, os estágios embrionários são mais definidos e incluem as fases globular, coração (heart), torpedo e cotiledonar [3]. Apesar dos avanços, os estudos sobre embriogênese zigótica são limitados pela pequena quantidade de material disponível, o que dificulta análises moleculares e

celulares mais robustas. Nesse sentido, a indução de embriogênese somática *in vitro* representa uma alternativa viável e poderosa para investigar os processos associados à embriogênese zigótica de forma análoga [4].

Importante destacar que a competência embriogênica não é uma característica exclusiva do zigoto. Diversos estímulos físicos e químicos, como alterações hormonais e variações de temperatura, podem induzir essa competência em células somáticas [5]. Uma vez adquirida a competência, essas células são capazes de iniciar a morfogênese e originar embriões somáticos com o mesmo genótipo do tecido de origem [6]. No processo conhecido como apomixia, ocorre a formação de sementes viáveis de maneira assexuada, diretamente a partir de tecidos maternos, sem que haja meiose ou fertilização, como ocorre na embriogênese zigótica [7]. Evidências sugerem que a apomixia está relacionada à expressão de genes com função embriogênica, como BABY BOOM (BBM), cuja superexpressão leva à formação ectópica de embriões em *A. thaliana* [8]. Do ponto de vista agrônomo, a apomixia possui grande potencial para a propagação clonal de híbridos de elite, permitindo a fixação da heterose e a redução de custos no melhoramento genético [7]. Portanto, a ocorrência da apomixia oferece uma forte evidência de que a fertilização pode ser substituída por mecanismos endógenos. Isso reforça a noção de que células somáticas vegetais, quando submetidas a condições apropriadas, podem manter ou recuperar a capacidade de formar embriões [5, 9].

1.2 Papel dos hormônios na indução da embriogênese somática

A embriogênese somática constitui um notável exemplo da totipotência celular em plantas, permitindo que células somáticas vegetativas se rediferenciem em células de linhagem reprodutiva e originem embriões completos sem a necessidade de fertilização sexual [10]. Este processo envolve profundas alterações na homeostase hormonal, sendo a auxina considerada um dos principais reguladores da embriogênese somática e da morfogênese vegetal [11]. Tal importância é evidenciada pela ampla utilização de auxinas exógenas, especialmente o 2,4-diclorofenoxiacético (2,4-D), em protocolos que visam a indução de embriogênese somática ou a formação de calos com competência embriogênica [9]. Mesmo abordagens baseadas na expressão ectópica de morfogenes requerem o acúmulo de auxina endógena, como o ácido indol-3-acético (IAA), para promover a reprogramação celular [10]. Além de desencadear a diferenciação de células embriogênicas, o metabolismo e transporte da auxina são cruciais para a padronização embrionária e o estabelecimento da morfologia do embrião, regulando a diferenciação dos cotilédones, do sistema vascular, do hipocótilo e de outras estruturas a partir do estágio globular. Esse controle depende da

ativação de diferentes famílias gênicas, como os genes YUCCA (YUC, envolvidos na biossíntese de auxina), PIN-FORMED (PIN) e P-GLYCOPROTEIN (PGP), relacionados ao efluxo de auxina, além de AUXIN1/LIKE-AUX1 (AUX/LAX), envolvidos no seu influxo [12]. A partir do estabelecimento desses gradientes, a sinalização mediada por auxina ativa genes de identidade embrionária e morfogênese, conduzindo à formação do embrião somático [5].

Além da auxina, as citocininas também exercem papel fundamental na regulação da embriogênese somática e na organização dos eixos embrionários. Durante a diferenciação dos embriões somáticos, o estabelecimento de gradientes entre auxina e citocinina contribui para a definição dos meristemas apical e radicular [13]. A expressão dos genes WUSCHEL (WUS) e WUSCHEL-RELATED HOMEODOMAIN 5 (WOX5), nos domínios de concentração máxima de auxina e citocinina, respectivamente, é essencial para o padrão apical-basal do embrião [6]. Em determinadas espécies, a aplicação de citocininas exógenas, sem o uso de auxinas, é suficiente para induzir a embriogênese somática a partir de diferentes explantes [14, 15]. A interação hormonal entre citocininas e outros fitormônios também regula processos fisiológicos complexos, como a ativação de gemas laterais, divisão celular, embriogênese e resposta ao estresse [16].

O ácido abscísico (ABA), por sua vez, é amplamente reconhecido por sua atuação em processos como o acúmulo de reservas, a maturação de embriões, a indução da dormência e a tolerância à dessecação [17]. Durante a embriogênese somática de *Arabidopsis*, mutantes com insensibilidade ao ABA, como ABSCISIC ACID INSENSITIVE 3, apresentam redução na eficiência e frequência de formação de embriões somáticos [18]. Além disso, o ABA influencia a morfogênese embrionária ao modular o transporte de auxina e o estabelecimento de seus gradientes durante a maturação de calos embriogênicos [19]. Em *Daucus carota* L. (cenoura), por exemplo, a inibição da biossíntese de ABA por fluridone reduz a frequência de embriogênese somática de 17% para 5%, sendo esta restaurada ou mesmo ampliada com a adição de ABA ao meio de cultura [20].

O etileno, um hormônio gasoso produzido frequentemente em resposta à auxina, é conhecido por sua participação na senescência, amadurecimento de frutos e abscisão de órgãos. Sua biossíntese se inicia a partir da metionina e envolve a ação da enzima 1-AMINOCICLOPROPANO-1-CARBOXÍLICO SINTASE (ACCS), cuja expressão é induzida por auxinas [21]. Durante a embriogênese somática, o papel do etileno pode variar entre espécies, tendo efeitos positivos ou negativos. Para mitigar possíveis impactos deletérios, o cultivo *in vitro* com recipientes que permitam trocas gasosas tem sido empregado [22]. Em *Arabidopsis* e soja (*Glycine max*), a expressão do gene AGAMOUS-LIKE15 atua

positivamente na embriogênese ao estimular a síntese de etileno [23]. Adicionalmente, o etileno interage com o óxido nítrico (NO) e a auxina de forma dose-dependente, sendo que altas concentrações de NO induzem a produção de etileno, a qual por sua vez pode inibir a sinalização de auxina [24]. Neste contexto, o uso de inibidores da biossíntese (como o cloreto de cobalto, CoCl_2) ou da sinalização do etileno (como o nitrato de prata, AgNO_3) representa uma alternativa viável para modular suas respostas durante a micropropagação [25].

O ácido jasmônico, geralmente associado a respostas de estresse e ferimentos, também participa da regulação hormonal da embriogênese somática [26]. Estudos recentes apontam que esse fitormônio pode atuar positivamente regulando a biossíntese de auxina, sendo sua ação dependente da dose: concentrações endógenas moderadas favorecem a indução embrionária, enquanto níveis elevados reduzem a eficiência do processo [27, 28]. Considerando que a embriogênese somática resulta da integração entre sinais hormonais e de estresse, o ácido jasmônico surge como um importante modulador dessa interface [5]. Entretanto, os efeitos deste hormônio também variam conforme o estágio de aplicação. Na cana-de-açúcar, por exemplo, a aplicação de ácido jasmônico em embriões somáticos maduros, com o objetivo de induzir desidratação, resulta na redução da viabilidade e qualidade das sementes sintéticas [28]. Por fim, as giberelinas, conhecidas por seu papel na germinação de sementes, alongação de órgãos aéreos e florescimento, atuam frequentemente como antagonistas do ABA [29]. Sua função na embriogênese somática ainda é menos explorada, mas seus efeitos antagônicos sobre o ABA sugerem que também possam modular eventos-chave da maturação e desenvolvimento embrionário [29].

1.3 Aplicações da embriogênese somática de plantas

Diversas espécies arbóreas de interesse comercial e com aplicação em programas de reflorestamento enfrentam desafios como baixa produção e/ou viabilidade das sementes, além de longos períodos até atingirem a fase reprodutiva. Nesse contexto, as coníferas se destacam por sua relevância na produção de energia, madeira e alimentos, porém o desenvolvimento de abordagens de baixo custo para a propagação clonal em larga escala ainda constitui um desafio significativo [30]. A embriogênese somática tem sido apontada como uma estratégia promissora para a produção de mudas e sementes sintéticas em espécies com dificuldades reprodutivas [31]. Além disso, o desenvolvimento de protocolos de embriogênese somática tem viabilizado o melhoramento genético de espécies florestais por meio da transformação genética [32]. Em situações em que populações naturais são severamente afetadas por patógenos, a embriogênese somática constitui uma ferramenta biotecnológica essencial para a produção rápida de mudas com elevada qualidade

fitossanitária [33]. Frente às mudanças climáticas e à degradação ambiental provocadas por ações antrópicas, a propagação de espécies perenes por embriogênese somática também representa uma alternativa viável para a conservação de espécies ameaçadas [33]. Dessa forma, estudos envolvendo embriogênese somática são de grande relevância e apresentam amplo potencial de aplicação para mitigar o desmatamento e a perda de vegetação nativa.

As plantas são reconhecidas pela produção de metabólitos secundários com alto valor médico e comercial [34]. Esses compostos incluem alcaloides, glicosídeos, flavonoides, óleos essenciais, resinas, taninos, polifenóis e terpenoides, com aplicações relevantes na indústria farmacêutica e, em alguns casos, no tratamento de câncer [35]. Nesse sentido, o uso de calos e suspensões celulares *in vitro* para a produção de metabólitos secundários tem se destacado como uma alternativa de baixo custo, com extração e purificação facilitadas, independente de fatores climáticos ou ambientais, além de passível de otimização e automação por robótica [36]. Tratamentos com elicitores ou estresses abióticos também podem ser empregados para intensificar a produção de metabólitos em culturas *in vitro*, por meio da ativação de mecanismos de defesa das plantas, utilizando-se estresses oxidativos, térmicos, osmóticos e polissacarídeos como quitina, pectina e celulose [34]. A utilização de biorreatores também contribui para a escalabilidade do processo e já tem sido empregada na produção de medicamentos comerciais, como paclitaxel, berberina e podofilotoxina [35]. Além disso, o desenvolvimento de linhagens celulares transgênicas representa uma estratégia promissora para aumentar a produção de metabólitos específicos durante o cultivo *in vitro*, promovendo maior eficiência [36]. Assim, o uso de calos e suspensões celulares para a produção de metabólitos vem ganhando destaque na indústria e apresenta elevado potencial de inovação tecnológica.

O estudo do desenvolvimento embriogênico a partir de embriões zigóticos é dificultado por fatores como a escassez de material, a presença de tecidos não embrionários (ex. tegumento e endosperma) e a necessidade de dissecação para obtenção dos embriões [37]. Nesse contexto, a embriogênese somática compartilha diversas características morfológicas e fisiológicas com a embriogênese zigótica, incluindo a regulação genética, hormonal e transcricional envolvida em ambos os processos [38]. Assim, além de seu amplo potencial biotecnológico, a embriogênese somática também constitui uma ferramenta útil para o estudo do desenvolvimento embrionário em plantas. Essa abordagem tem sido empregada para investigar eventos biológicos fundamentais, como a morfogênese *in vitro*, a desdiferenciação celular, a aquisição da competência embriogênica, a diferenciação de embriões e a proliferação de linhagens celulares embriogênicas [9]. Ademais, o uso da embriogênese

somática tem possibilitado a aplicação de abordagens ômicas de alto rendimento, como a proteômica e a transcriptômica, que exigem grandes volumes de amostras [39, 40]. Dessa forma, pesquisas voltadas à regulação da embriogênese somática têm fornecido importantes insights sobre os mecanismos moleculares e celulares envolvidos no desenvolvimento embrionário, especialmente nos estágios iniciais de diferenciação, morfogênese e maturação.

1.4 Produção em larga-escala de plantas através da embriogênese somática

A embriogênese somática é o processo pelo qual células isoladas ou em grupos, após estímulo apropriado, expressam de forma ectópica o programa embriogênico e iniciam o desenvolvimento de embriões de maneira assexuada [41]. Este processo de micropropagação apresenta diversas aplicações tecnológicas e vem sendo amplamente empregado na produção de mudas em larga escala, regeneração de plantas após transformações genéticas (transformação e edição), conservação de germoplasma, melhoramento genético e também no estudo do desenvolvimento vegetal [42, 43]. Nesse contexto, a embriogênese somática constitui a principal alternativa para propagação de espécies florestais, considerando o longo ciclo de vida e a dificuldade de propagação por métodos convencionais apresentados por espécies arbóreas [32]. Além disso, o uso de biorreatores tem permitido a redução do tempo e dos custos com mão de obra, a otimização da produção, devido ao maior contato entre o meio nutritivo e o material vegetal, e a automatização de diferentes etapas da embriogênese somática [44]. Por outro lado, a aplicação comercial dessa abordagem ainda enfrenta desafios, como as baixas taxas de indução e regeneração de embriões somáticos, a escassez de explantes adequados para algumas espécies e a incidência de variação somaclonal durante a proliferação de culturas embriogênicas [45]. Assim, estudos voltados à compreensão da regulação molecular associada à progressão entre os diferentes estágios da embriogênese somática podem viabilizar sua aplicação em programas comerciais de micropropagação.

O Brasil possui a matriz energética mais limpa do mundo, sendo que parte desse sucesso se deve à utilização da cana-de-açúcar (*Saccharum spp.*) para a produção de bioenergia e etanol [46]. Um dos principais desafios na produção dessa cultura é garantir a qualidade fitossanitária das mudas, uma vez que o uso de toletes no plantio é pouco eficiente e pode favorecer a disseminação de patógenos [47]. Estudos recentes demonstram o potencial da embriogênese somática para a propagação em larga escala de clones saudáveis, contribuindo para a redução de custos na produção de mudas e na incidência de doenças [48]. Neste sentido, calos embriogênicos podem ser induzidos a partir de folhas imaturas cultivadas em meio contendo auxinas, fonte de carbono, vitaminas e sais minerais

[48]. Em seguida, o calo embriogênico é submetido a um tratamento de maturação para promover a diferenciação e o desenvolvimento dos embriões somáticos, sendo possível, com o uso de biorreatores, a produção de mais de 18 mil mudas a partir de uma única planta utilizada como fonte de explantes [49]. Dessa forma, evidencia-se o grande potencial da embriogênese somática na propagação de espécies de relevância nacional, como a cana-de-açúcar, onde novas descobertas podem gerar inovações significativas para o aumento da produtividade.

O mamoeiro (*Carica papaya* L.) é uma espécie cultivada principalmente em regiões tropicais, sendo o Brasil o segundo maior produtor mundial dessa cultura [50]. Durante seu cultivo, a segregação sexual das sementes gera prejuízos significativos, visto que as plantas podem apresentar sexo feminino, masculino ou hermafrodita, sendo que apenas estas últimas produzem frutos com valor comercial [51]. Uma alternativa promissora para reduzir os custos de produção consiste na utilização da embriogênese somática para gerar mudas com sexo definido, evitando perdas causadas pela segregação sexual [43]. Nesse contexto, calos embriogênicos podem ser induzidos a partir de embriões zigóticos cotiledonares, seguidos pela diferenciação dos embriões somáticos [52]. Durante o cultivo *in vitro*, os embriões somáticos apresentam os mesmos estágios morfológicos dos embriões zigóticos, destacando o valor da embriogênese somática como ferramenta para o estudo da morfogênese e do desenvolvimento embriogênico vegetal [53]. Adicionalmente, a combinação da embriogênese somática com técnicas de biologia molecular, como a utilização de marcadores moleculares, permite a seleção de calos com genótipo hermafrodita e a obtenção de embriões com sexo previamente definido, contribuindo para a redução de custos com produção de mudas [53]. Contudo, a aplicação comercial da embriogênese somática no mamoeiro ainda requer a otimização de protocolos de regeneração dos embriões somáticos, bem como o controle da variação somaclonal durante a cultura.

1.5 Utilização na produção de plantas transgênicas

O uso de organismos geneticamente modificados (OGMs) na agricultura tem aumentado significativamente nos últimos anos, sendo que um dos principais entraves ao desenvolvimento de novas cultivares via engenharia genética reside na regeneração eficiente das plantas transformadas [54]. Nesse contexto, a embriogênese somática desponta como uma das principais rotas de regeneração para plantas geneticamente modificadas [42]. Além disso, estudos envolvendo embriogênese somática têm levado à identificação de morfogenes cruciais, como BABY BOOM (BBM), AGAMOUS-LIKE 15 (AGL15), WUSCHEL (WUS), SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE 1 (SERK1) e LEAFY

COTYLEDON (LEC), que podem ser explorados para induzir a formação ectópica de embriões somáticos, reduzindo significativamente o tempo necessário para a regeneração de plantas após a transformação genética [54]. Novas abordagens biotecnológicas têm empregado a expressão desses morfogenes para acelerar e aumentar a eficiência da regeneração em espécies monocotiledôneas e dicotiledôneas, superando barreiras relacionadas à recalcitrância genotípica [55]. Em *Malus domestica* (macieira), por exemplo, a expressão de BBM aumentou em até dez vezes a taxa de regeneração de plantas transformadas, de 3% para 30% [42, 55]. Estratégias similares também têm sido aplicadas com sucesso em milho (*Zea mays*) e sorgo (*Sorghum bicolor*), onde a coexpressão de WUS e BBM promoveu ganhos significativos na eficiência de transformação e edição gênica [56]. Tais avanços demonstram o enorme potencial da embriogênese somática como ferramenta para a geração de plantas transgênicas e ressaltam a importância de estudos que investiguem os mecanismos moleculares envolvidos na diferenciação e desenvolvimento dos embriões somáticos.

Apesar dos progressos, um dos principais fatores limitantes à aplicação da transformação e edição genética ainda é a recalcitrância de diversos genótipos às técnicas de cultura de tecidos, o que restringe a transformação a um número reduzido de cultivares responsivos *in vitro* [57]. A otimização de protocolos para esses genótipos é frequentemente complexa, exigindo testes exaustivos com diferentes tipos de explantes, composições do meio de cultura, pH e demais fatores físico-químicos [58]. Atualmente, ao menos cinco estratégias têm sido descritas para superar essa limitação e viabilizar a regeneração de plantas de forma genótipo-independente. Todas essas estratégias envolvem a expressão de morfogenes de forma constitutiva, induzida ou transiente [57]. Essas abordagens se baseiam na indução da embriogênese somática de forma ectópica, um processo que envolve a ativação de redes complexas de transcrição relacionadas à resposta a estresses, hormônios vegetais, reprogramação epigenética e desenvolvimento embrionário [58]. Diversos estudos vêm identificando genes-chave que regulam a indução da competência embriogênica, a diferenciação e o desenvolvimento dos embriões somáticos [11]. No entanto, apesar dos avanços na elucidação desses mecanismos, ainda persistem lacunas importantes que limitam a aplicação de estratégias de transformação genética genótipo-independente [9]. Assim, pesquisas que aprofundem o conhecimento sobre o controle molecular e a progressão da embriogênese somática são fundamentais para viabilizar novas tecnologias de transformação e edição gênica em uma ampla gama de espécies vegetais.

1.6 Aplicações na conservação de germoplasma

O armazenamento de material vegetal por longos períodos é essencial para a preservação da diversidade genética, sendo uma estratégia fundamental tanto para programas de melhoramento quanto para a conservação de espécies ameaçadas de extinção [59]. Nesse contexto, a criopreservação, técnica baseada no armazenamento de tecidos em nitrogênio líquido a -196°C , tem se destacado como uma das principais abordagens para a conservação de germoplasma, sendo que o uso de embriões somáticos e linhagens de células embriogênicas, como calos embriogênicos, vem ganhando crescente relevância [60]. A adoção da criopreservação nesse contexto apresenta diversas vantagens, incluindo a redução de tempo e custos operacionais, menor risco de contaminação, diminuição da ocorrência de variação somaclonal e a preservação da competência morfogenética durante a embriogênese somática [61]. O processo de criopreservação normalmente compreende múltiplas etapas: pré-crescimento do material vegetal, tratamento com agentes crioprotetores, resfriamento lento, imersão e armazenamento em nitrogênio líquido, descongelamento rápido e posterior recuperação do material viável [62]. Apesar dos avanços já obtidos, a otimização dos protocolos de embriogênese somática específicos para diferentes espécies e cultivares permanece como um dos principais gargalos para a ampla aplicação desta tecnologia [63]. A elucidação dos mecanismos moleculares envolvidos na regulação da competência embriogênica, bem como na diferenciação e desenvolvimento dos embriões somáticos, poderá ampliar significativamente as possibilidades de aplicação da embriogênese somática na conservação de germoplasma vegetal.

Avanços nas técnicas de criopreservação representam um campo de grande interesse para a biotecnologia vegetal, especialmente à medida que cada etapa do processo vem sendo refinada. Entre os fatores físicos mais críticos para o sucesso da criopreservação estão a formação de cristais de gelo intracelulares e o processo de vitrificação, ambos fortemente influenciadores da viabilidade celular após o descongelamento [64]. Durante a fase de pré-crescimento, objetiva-se reduzir o teor de água no tecido vegetal utilizando agentes osmóticos e, simultaneamente, aumentar a tolerância ao frio por meio da adição de ácido abscísico (ABA) ao meio de cultura ou pela exposição prévia a baixas temperaturas [65]. Na etapa de crioproteção, compostos como sacarose e dimetilsulfóxido (DMSO) são adicionados para promover a desidratação controlada e uniformizar as amostras antes do resfriamento [63]. Posteriormente, o material é submetido a um resfriamento lento, geralmente a uma taxa de $0,1\text{--}1^{\circ}\text{C}$ por minuto, até atingir temperaturas entre -35°C e -40°C , antes de ser finalmente

imerso e armazenado em nitrogênio líquido por tempo indeterminado [65]. Para a recuperação do material, o descongelamento deve ocorrer de forma rápida, geralmente em banho-maria a 40 °C, a fim de evitar a recristalização da água e danos celulares subsequentes [64]. Após o descongelamento, os agentes crioprotetores são removidos e os embriões são incubados em meio de cultura apropriado por algumas semanas para permitir sua recuperação e regeneração [64].

Embora a criopreservação ofereça múltiplas vantagens em relação aos métodos convencionais de conservação *ex situ* (como os bancos de campo), seu uso ainda é limitado, em grande parte devido aos custos elevados e à necessidade de mão de obra especializada [65]. Desse modo, o desenvolvimento de novas estratégias que aumentem a taxa de recuperação e a viabilidade dos tecidos após a criopreservação poderá viabilizar a adoção mais ampla desta técnica para a conservação de diferentes espécies vegetais, especialmente aquelas de interesse agrícola e ecológico.

1.7 Abordagens proteômicas na ciência de plantas

Ferramentas ômicas foram amplamente incorporadas à pesquisa com plantas, sendo empregadas na elucidação de processos biológicos como resposta a estresses, embriogênese, fotossíntese, florescimento e reprodução [66]. No que se refere à proteômica, o desenvolvimento de plataformas livres de gel, o surgimento de novas gerações de espectrômetros de massas e sistemas de cromatografia líquida, a aplicação da separação por mobilidade iônica em abordagens *bottom-up*, os avanços em algoritmos de processamento de dados e o sequenciamento de genomas possibilitaram um aumento expressivo no número de proteínas identificadas em um único experimento [66]. Dessa forma, a proteômica consolidou-se como uma abordagem fundamental para o estudo da função gênica, contribuindo significativamente tanto para a ciência básica quanto para aplicações em biotecnologia.

Um fator crítico para o avanço da proteômica vegetal tem sido o desenvolvimento de métodos eficientes de coleta e preparo de amostras [67, 68]. Por outro lado, a identificação de modificações pós-traducionais e a quantificação de proteínas de baixa abundância ainda representam desafios técnicos importantes na proteômica vegetal [69]. A heterogeneidade celular presente em tecidos vegetais, como folhas e raízes, também gera um ruído de fundo que dificulta a interpretação precisa dos dados. Avanços recentes em técnicas de enriquecimento de peptídeos de baixa abundância, proteômica dirigida (*targeted* proteomics)

e análises em célula única representam estratégias promissoras para superar essas limitações [70, 71].

As proteínas podem ser analisadas por meio de abordagens dependentes de gel, como a eletroforese bidimensional (2DE) e a eletroforese bidimensional diferencial com marcação fluorescente (2DE-DIGE), ou por abordagens livres de gel, como a proteômica label-free, a marcação com afinidade isotópica (ICAT), as etiquetas isobáricas para quantificação relativa e absoluta (iTRAQ) e as etiquetas isobáricas para proteômica quantitativa (TMT) [39]. Dentre essas, as abordagens livres de gel (como label-free, ICAT, iTRAQ e TMT) oferecem maior reprodutibilidade e poder de resolução, resultando na identificação de um número significativamente maior de proteínas. Por esse motivo, a proteômica vegetal tem gradativamente migrado do uso das técnicas baseadas em gel para plataformas livres de gel [39]. No entanto, cada abordagem possui vantagens e limitações, sendo o uso de múltiplas estratégias de forma complementar também uma possibilidade a ser considerada [70]. Todas as etapas envolvidas no planejamento e condução de um experimento proteômico utilizando tecidos vegetais serão discutidas a seguir. Cada fase do fluxo de trabalho será abordada, com a descrição detalhada das metodologias disponíveis, de modo a auxiliar na escolha do método mais apropriado para cada projeto de pesquisa.

1.8 Abordagens Gel-dependentes e Livres de Gel na Proteômica Vegetal

A proteômica tem se consolidado como uma técnica central para a elucidação do proteoma vegetal, oferecendo subsídios essenciais para a compreensão da expressão, regulação e função de proteínas. Essa abordagem tem possibilitado a identificação de biomarcadores proteicos associados a respostas ao estresse [72], sinais de desenvolvimento [73] e condições patológicas [74], demonstrando grande potencial para a melhoria da produtividade e da resiliência de culturas agrícolas [75]. De modo geral, a análise proteômica compreende diversas etapas: dissociação dos tecidos vegetais, lise celular e desnaturação das proteínas, redução de pontes dissulfeto, alquilação de cisteínas, purificação proteica (com acetona gelada ou precipitação clorofórmio–metanol), digestão enzimática com uma ou mais proteases, dessalinização e ressuspensão dos peptídeos em solvente compatível com LC-MS, análise por LC-MS/MS, processamento dos dados brutos com softwares adequados (busca em banco de dados, controle de FDR, quantificação relativa e inferência proteica), análise estatística e interpretação biológica dos resultados [76]. Uma etapa fundamental durante esse fluxo é a escolha entre métodos gel-dependentes ou livres de gel, cada qual

com vantagens e limitações distintas [77]. Nas abordagens gel-dependentes, as proteínas são separadas em uma dimensão (1D), com base no peso molecular, ou em duas dimensões (2D), combinando ponto isoelétrico (por tiras de gradiente de pH imobilizado) e massa molecular (eletroforese em gel de poliacrilamida com SDS – SDS-PAGE). Em contrapartida, nas abordagens livres de gel, a separação de proteínas é substituída pela resolução dos peptídeos gerados na digestão enzimática via cromatografia líquida [77].

A proteômica gel-dependente foi proposta na década de 1970 e se consolidou como uma técnica de alta resolução, permitindo a identificação de até mil proteínas por amostra [78]. Nessa abordagem, a identificação proteica envolve separação por SDS-PAGE, digestão proteolítica in-gel e análise por espectrometria de massas [79]. Esse método tem sido amplamente empregado no estudo de complexos proteicos, reações enzimáticas, análises *top-down* e limpeza de amostras (remoção de detergentes, por exemplo) [80]. Além disso, técnicas gel-dependentes permitem a caracterização de complexos multiproteicos e modificações pós-traducionais, contribuindo para o entendimento de processos celulares associados ao crescimento e à adaptação de plantas [81]. Contudo, a proteômica baseada em gel apresenta diversos gargalos: baixa reprodutibilidade, dificuldades na resolução de proteínas muito ácidas ou básicas, baixa detecção de proteínas pouco abundantes (como fatores de transcrição), baixo rendimento analítico, co-migração de proteínas durante a eletroforese e dificuldade na separação de proteínas com massa molecular muito alta ou muito baixa [80]. Por esses motivos, a proteômica vegetal tem migrado para abordagens livres de gel, sendo as técnicas gel-dependentes atualmente reservadas para análises específicas, como identificação de proteínas intactas (*top-down* proteomics) e fracionamento de amostras.

Como alternativa, a proteômica livre de gel baseia-se na separação de amostras altamente complexas de peptídeos por cromatografia líquida em fase reversa acoplada à espectrometria de massas (LC-MS). Essa abordagem permite separar, sequenciar e quantificar os peptídeos gerados na etapa de digestão proteica [82]. Dentre suas principais vantagens, destacam-se o alto rendimento e reprodutibilidade, a possibilidade de detecção de proteínas de baixa abundância, o mapeamento de modificações pós-traducionais, a identificação de milhares (ou até dezenas de milhares) de proteínas, a quantificação absoluta e a redução do tempo de análise em até 20 vezes quando comparada às metodologias gel-dependentes [83]. Esses avanços só foram possíveis graças ao desenvolvimento de novas tecnologias de espectrometria de massas, sistemas de cromatografia líquida e algoritmos de bioinformática capazes de quantificar e identificar proteínas com maior sensibilidade e precisão.

1.9 Targeted vs. Untargeted Proteomics

A proteômica pode ser classificada de acordo com o delineamento experimental e os objetivos do estudo em abordagens direcionadas (*targeted*) ou não direcionadas (*untargeted*). A proteômica não direcionada, também conhecida como "discovery proteomics", consiste na aquisição de espectros dentro de uma faixa de massa específica, selecionando (DDA, data-dependent acquisition) ou não (DIA, data-independent acquisition) os íons precursores mais abundantes [84]. Por essa razão, este tipo de abordagem é considerado não orientado por hipótese, uma vez que visa identificar e quantificar o maior número possível de proteínas, sem um alvo específico previamente definido [85]. A principal vantagem da proteômica *untargeted* reside em seu caráter abrangente e de alto rendimento, permitindo uma compreensão ampla dos processos biológicos sob investigação. Contudo, uma limitação relevante refere-se à baixa capacidade de detecção e quantificação de proteínas de baixa abundância, como fatores de transcrição e proteínas reguladoras, uma vez que muitos desses alvos não produzem peptídeos em quantidade suficiente para atingir o limite mínimo de detecção dos espectrômetros de massa e acabam mascarados por proteínas altamente abundantes [86].

Como alternativa, a proteômica direcionada (*targeted*) tem sido empregada para detectar proteínas presentes em poucas cópias e realizar a quantificação absoluta de proteínas específicas nas amostras [85]. Os experimentos de proteômica direcionada geralmente envolvem a seleção de proteínas-alvo relevantes, a escolha de peptídeos únicos que funcionem como assinaturas moleculares, o desenvolvimento de métodos de LC-MS/MS com alta precisão e sensibilidade para esses peptídeos, a preparação das amostras (extração, redução, alquilação e digestão) e a aquisição dos espectros de massa [87]. Nesse contexto, a proteômica direcionada já foi aplicada com sucesso no desenvolvimento de marcadores para auxiliar programas de melhoramento genético vegetal [88].

1.10 Bottom-up, Middle-down e Top-down Proteomics

As análises proteômicas também podem ser classificadas com base no nível de processamento da proteína, sendo as principais abordagens conhecidas como *bottom-up*, *middle-down* e *top-down*. A proteômica *bottom-up*, também denominada shotgun proteomics, envolve a digestão enzimática de proteínas complexas em peptídeos, os quais são então

analisados por espectrometria de massas [82, 89]. A escolha da protease adequada é um aspecto crítico nesta abordagem, sendo a tripsina amplamente utilizada devido ao seu baixo custo, alta disponibilidade, eficiência e boa cobertura na identificação de proteínas [90]. Além disso, a utilização combinada de múltiplas proteases (como tripsina, quimotripsina e Lys-C) pode ampliar a diversidade de peptídeos gerados, aumentando a cobertura na sequência proteica [91]. Apesar das vantagens, a abordagem *bottom-up* apresenta limitações, como a redundância na inferência de proteínas a partir de peptídeos, cobertura incompleta das sequências proteicas (dificultando o estudo de modificações pós-traducionais), e a presença de peptídeos isobáricos, como os que contêm leucina e isoleucina, que compartilham o mesmo valor de m/z [92]. O desenvolvimento de técnicas de espectrometria com mobilidade iônica tem contribuído para mitigar esse problema, ao aumentar a resolução e diferenciar peptídeos isobáricos [93]. Diante disso, a proteômica *bottom-up* consolidou-se como a principal abordagem para experimentos de alta escala envolvendo identificação e quantificação de proteínas.

A proteômica *top-down*, por sua vez, caracteriza-se pela análise direta de proteínas intactas, o que confere vantagens significativas para a identificação de proteoformas e modificações pós-traducionais [94]. Apesar disso, esta abordagem enfrenta desafios, como baixa capacidade de rendimento, complexidade das amostras, insolubilidade de certas proteínas, dificuldades na análise de proteínas com massa elevada e desafios computacionais na análise dos dados gerados [95]. Mesmo assim, a proteômica *top-down* é considerada a mais adequada para o estudo de proteoformas, e avanços na preparação de amostras, instrumentação e métodos de aquisição de dados (como DIA) têm permitido a expansão de seu uso [89, 96]. Estratégias de preparo, como diluição de substâncias interferentes, ultrafiltração por corte de peso molecular (MWCO), precipitação proteica e cromatografia líquida acoplada à espectrometria de massas (LC-MS), têm sido adotadas para mitigar a supressão e dispersão de sinais [97].

Como alternativa intermediária, a proteômica *middle-down* tem ganhado destaque por analisar peptídeos maiores (entre 20 e 100 resíduos de aminoácidos), o que permite superar algumas limitações da abordagem *bottom-up*, como viés para peptídeos pequenos e perda de modificações pós-traducionais [98]. A *middle-down* é baseada na utilização de outras proteases (como Glu-C e Asp-N), obtendo peptídeos entre 2 e 20 kDa [99]. Essa abordagem é particularmente útil para estudar modificações em domínios ricos em arginina, como ocorre em histonas [100]. Portanto, enquanto a abordagem *bottom-up* tende a perder informações importantes durante a digestão, a *top-down* é ideal para o estudo de proteoformas completas.

Contudo, as limitações da *top-down* em relação à solubilidade e ao tamanho das proteínas fazem com que a *middle-down* surja como uma alternativa viável para a investigação de modificações pós-traducionais em proteínas com características estruturais complexas [101].

2. Revisão em Bancos de Propriedade Intelectual

Considerando a ampla aplicação da embriogênese somática na biotecnologia vegetal, a busca por patentes em bancos de propriedade industrial é essencial para o desenvolvimento de pesquisas originais com alto potencial de inovação. Nesse contexto, foram realizadas buscas nas bases de dados de patentes do Instituto Nacional da Propriedade Industrial (INPI) e PatentScope, com o objetivo de identificar documentos relacionados à embriogênese somática de plantas. Os resultados da pesquisa, utilizando o termo "Embriogênese somática" no INPI, revelaram 15 pedidos de patentes associados a processos biotecnológicos em plantas (Tabela 1). Os pedidos depositados no INPI estão predominantemente focados em processos e protocolos de micropropagação *in vitro*, produção de sementes sintéticas e otimização da transformação genética em diversas espécies vegetais. Dentre os pedidos encontrados, destacam-se cultivares de interesse agrônomo no Brasil, como cacau, beterraba e algodão. A presença de pedidos de patentes no INPI sublinha a relevância estratégica da inovação associada à aplicação da embriogênese somática na propagação de espécies vegetais de interesse nacional, destacando seu papel essencial no avanço da biotecnologia aplicada à agricultura e à produção de cultivos comerciais no Brasil.

Tabela 1. Busca por pedidos de patentes com os termos “embriogênese somática” no INPI. A busca foi realizada no dia 01/04/2025.

Pedido	Depósito	Título	IPC
BR 11 2024 022274 8	25/04/2023	Composições e métodos para embriogênese somática em plantas dicotiledôneas	C07K 14/415
BR 11 2023 010365 7	22/12/2021	Promoção de regeneração e transformação em beta vulgaris	C12N 15/82
BR 10 2020 001563 0	24/01/2020	Meio para cultura de tecidos vegetais	A01G 24/22
BR 11 2016 011325 0	19/11/2014	Processo para preparar embriões somáticos secundários de theobroma cacao l	A01H 4/00
PI 0906639-0	07/01/2009	Métodos para regenerar plantas férteis de jatropha curcas ou um híbrido artificial de jatropha curcas através de embriogênese somática e através de embriogênese somática e um processo de duas etapas de indução, bem como método para o estabelecimento de culturas líquidas em suspensão de plantas embriogênicas somáticas de jatropha curcas	C12N 15/87
PI 0610760-5	28/04/2006	Métodos combinatório de otimização da embriogênese somática, de preparação de embriões, de múltiplos embriões vegetais e de embriões somáticos de conífera para a produção de plantas e de obtenção de embriões germinantes, meio líquido para o cultivo de tecido embrionário e equipamentos semi-automatizado para a coleta de embriões e de preparação de múltiplos embriões vegetais para a produção de plantas, embriões e planta	A01H 4/00
PI 0406566-2	23/01/2004	Método para reproduzir coníferas por embriogênese somática empregando compostos contendo galactose como uma fonte de energia e carbono	C12N 5/00
PI 0318669-5	31/12/2003	Método para a regeneração de algodão por meio de embriogênese somática	A01H 4/00
PI 0013482-1	21/06/2000	Regeneração embriogênica somática de acacia mangium	A01H 4/00
PI 9914724-6	20/10/1999	Apomixia conferida através de expressão de proteínas de interação com serk	C12N 15/82
PI 9801485-4	10/06/1998	Clonagem de plantas adultas selecionadas de eucalyptus ssp pelo processo de regeneração in vitro por embriogênese somática	A01H 5/10
PI 9706395-9	18/12/1997	Processo aperfeiçoado para regeneração de plantas coníferas por meio de embriogênese somática utilizando polietilenoglicol.	A01H 4/00
PI 9703137-2	13/05/1997	Processo aperfeiçoado para regeneração de plantas coníferas por meio de embriogênese somática	A01H 4/00
PI 9501707-0	18/04/1995	Método para estabelecer culturas embriogênicas de coníferas em suspensão líquida e método de reprodução de plantas por embriogênese somática.	A01H 4/00
PI 9501708-9	18/04/1995	Processo para a reprodução de plantas através de embriogênese somática.	A01H 4/00

A busca realizada no PatentScope abrange a base de dados da World Intellectual Property Organization (WIPO), uma agência internacional criada pelas Nações Unidas em 1967, que atualmente conta com 193 países membros. Sendo uma base de dados mais abrangente que o INPI, a pesquisa de patentes no PatentScope foi conduzida utilizando diferentes termos e operadores booleanos. Dessa forma, a busca com o termo "*plant AND morphogene*" resultou em 259 pedidos de depósitos de patentes (Figura 1). Os cinco países com os maiores números de pedidos foram Estados Unidos, Canadá, Austrália, Índia e África do Sul (Figura 1A). Além disso, os pedidos depositados por meio do *Patent Cooperation Treaty* (PCT) geraram o maior número de patentes associadas à busca realizada no PatentScope, destacando a importância desse tratado para a proteção da propriedade industrial em escala global. A maioria desses documentos está associada aos códigos de patentes C12N (Micro-organismos ou enzimas; suas composições; propagação, preservação ou manutenção de micro-organismos; mutação ou engenharia genética; meios de cultura) e A61K (Preparações para fins médicos, dentários ou de higiene) (Figura 1B).

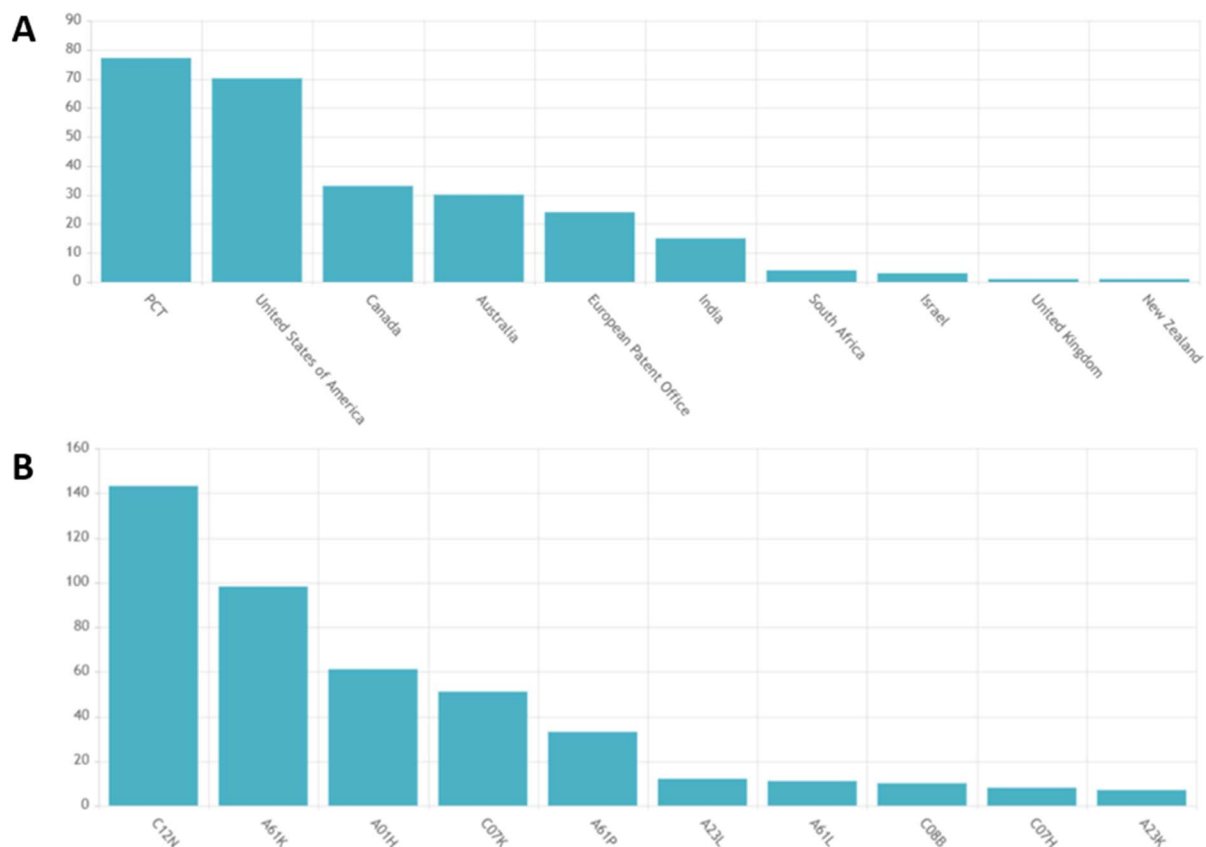


Figura 1. Gráfico de barras com os resultados da busca pelos termos "*plant AND morphogene*". (A). Países e tratados com mais pedidos de depósitos de patentes no mundo. (B). Códigos IPC mais associados aos resultados da busca.

A busca realizada com os termos “*plant AND somatic embryogenesis*” resultou em 338 pedidos de patentes (Figura 2). Esses depósitos foram predominantemente realizados por países como Estados Unidos, Canadá, Austrália, Índia e Nova Zelândia, além de órgãos internacionais como o PCT e o European Patent Office (EPO) (Figura 2A). Os códigos de patentes mais associados a esses depósitos foram CN12 e A01H. O código CN12 refere-se a “Novas plantas ou processos não transgênicos para obtê-las”, enquanto o código A01H está relacionado a “Reprodução de plantas por técnicas de cultura de tecidos” (Figura 2B).

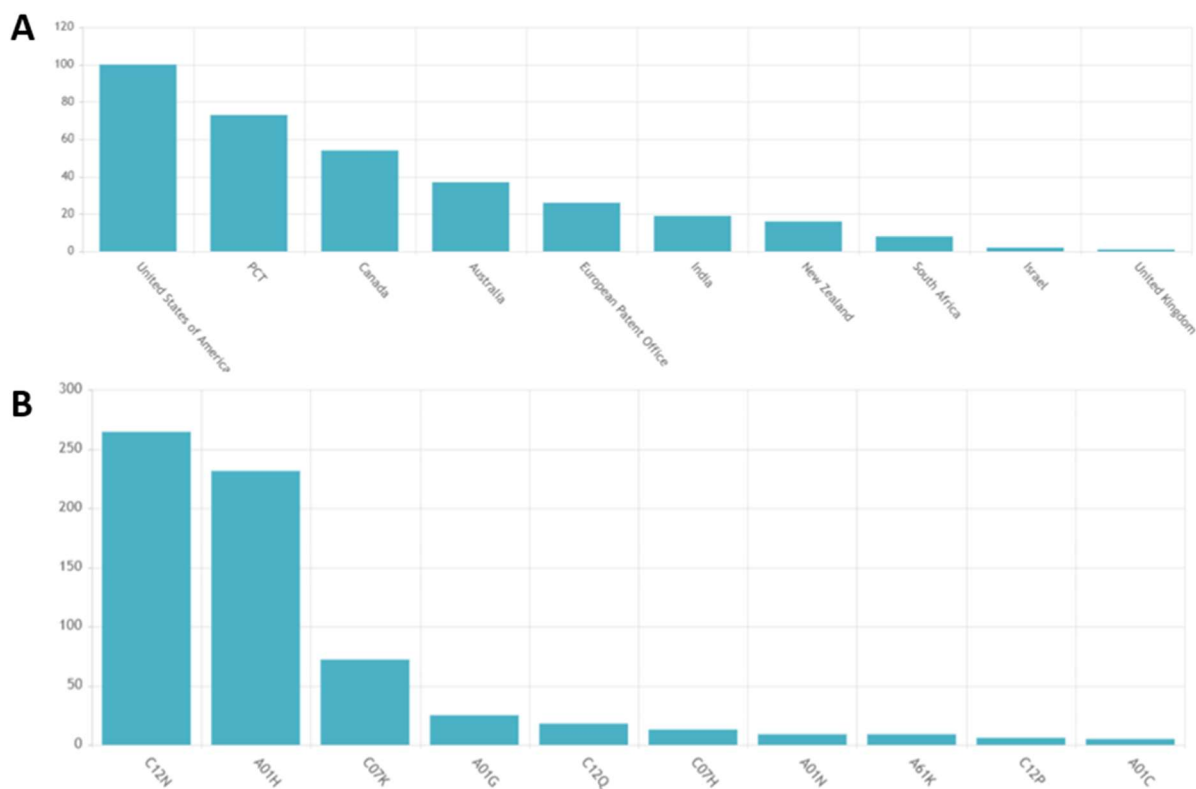


Figura 2. Gráfico de barras com os resultados da busca pelos termos “*plant AND somatic embryogenesis*”. (A). Países e tratados com mais pedidos de depósitos de patentes no mundo. (B). Códigos IPC mais associados aos resultados da busca.

A busca realizada com os termos “*plant AND babyboom*” resultou em 117 pedidos de patentes (Figura 3). Esses depósitos foram predominantemente realizados por países como Estados Unidos, Canadá, Austrália e Índia, além de órgãos internacionais como o PCT e o European Patent Office (EPO) (Figura 3A). Os códigos de patentes mais associados a esses depósitos foram CN12 e A01H (Figura 3B).

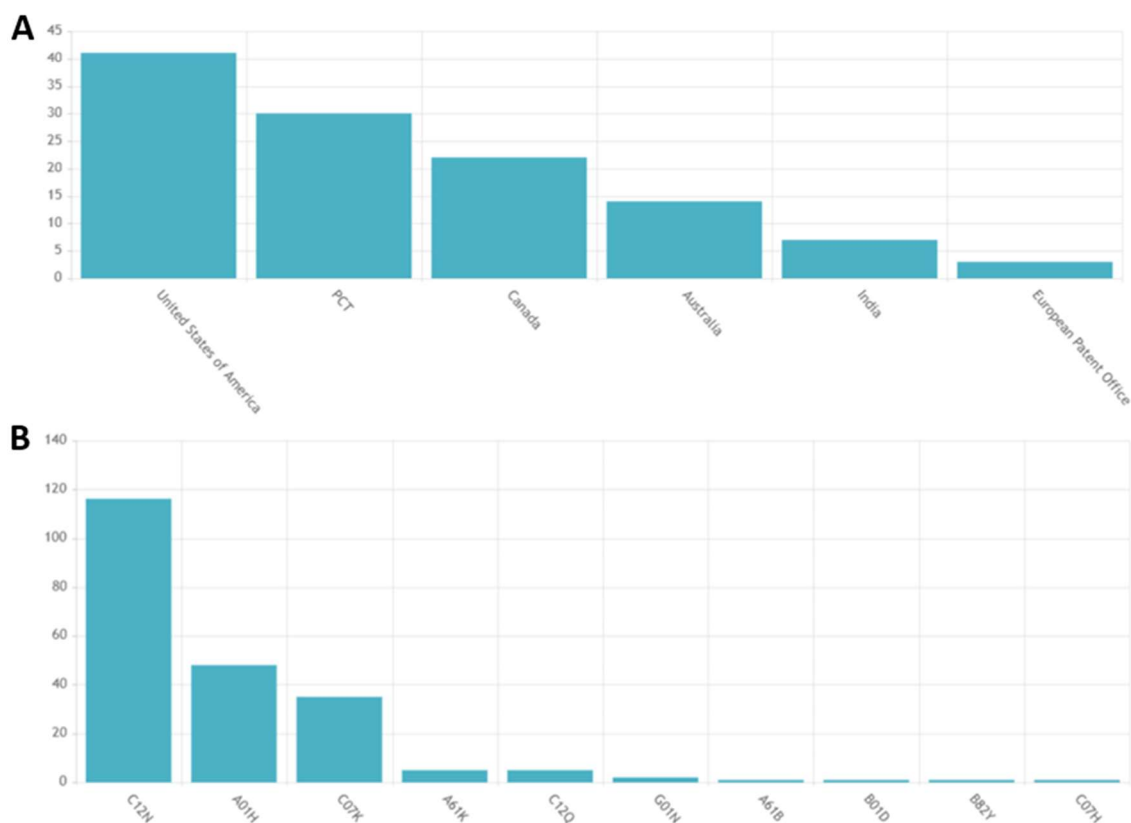


Figura 3. Gráfico de barras com os resultados da busca pelos termos “*plant AND babyboom*”. (A). Países e tratados com mais pedidos de depósitos de patentes no mundo. (B). Códigos IPC mais associados aos resultados da busca.

A busca realizada com os termos “*plant AND wuschel*” resultou em 116 pedidos de patentes (Figura 4). Esses depósitos foram predominantemente realizados por países como Estados Unidos, Canadá, Austrália, Nova Zelândia e Índia, além de órgãos internacionais como o PCT e o European Patent Office (EPO) (Figura 4A). Os códigos de patentes mais associados a esses depósitos foram CN12 e A01H (Figura 4B).

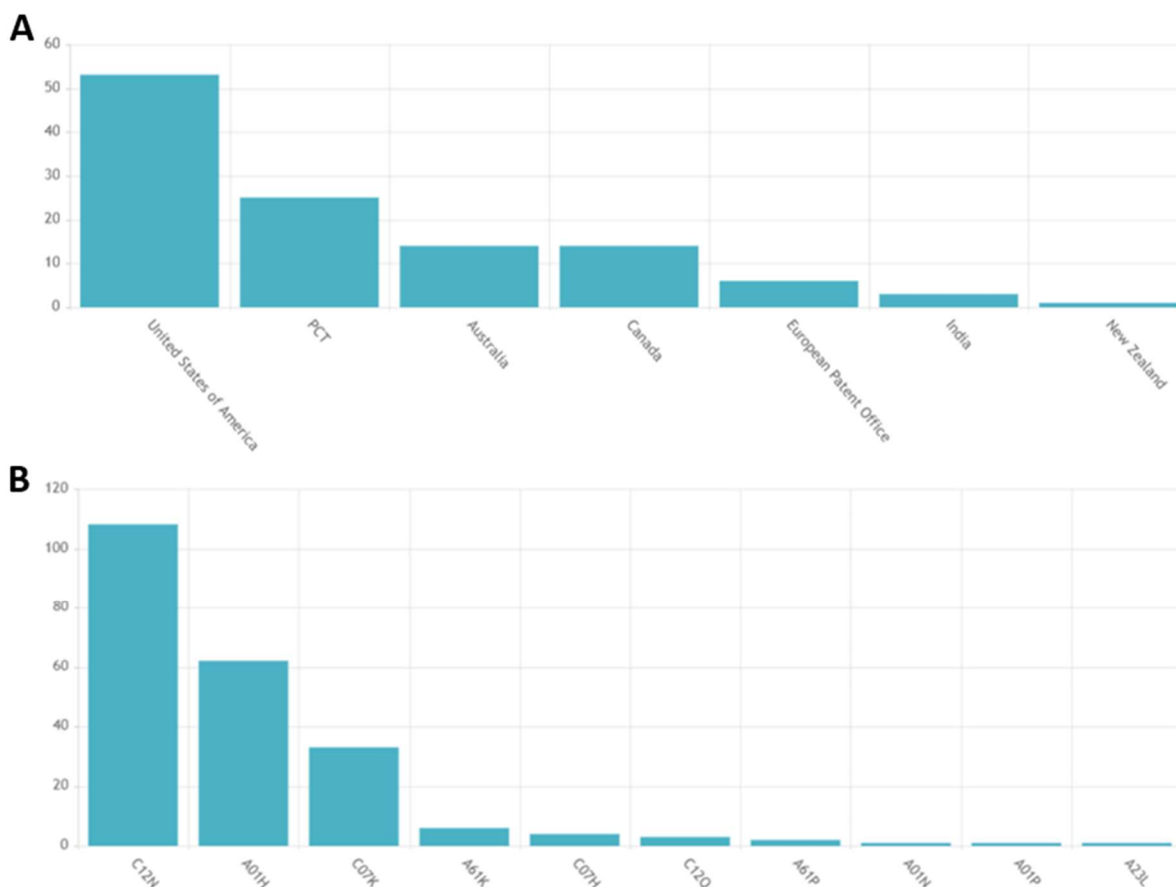


Figura 4. Gráfico de barras com os resultados da busca pelos termos “*plant AND wuschel*”. (A). Países e tratados com mais pedidos de depósitos de patentes no mundo. (B). Códigos IPC mais associados aos resultados da busca.

Como resultado das buscas de patentes na base de dados do PatentScope, observou-se que o país com maior número de depósitos de patentes associados à embriogênese somática de plantas é os Estados Unidos, seguido pela Austrália e pelo Canadá, além dos pedidos via EPO e PCT. Isso evidencia que, embora o Brasil seja um grande fornecedor de produtos agrícolas no mercado global, ainda apresenta uma baixa taxa de inovação em abordagens biotecnológicas com alto potencial para melhorar a produção nacional. Além disso, os pedidos de patentes com reivindicações relacionadas à embriogênese somática protegem principalmente a composição de meios de cultura (C12N) e a engenharia genética de plantas (A01H). Esse cenário demonstra que o potencial inovador das pesquisas em embriogênese somática impacta principalmente a produção de mudas micropropagadas e a otimização de métodos de regeneração de plantas geneticamente modificadas.

O estudo da embriogênese somática de plantas levou à identificação de morfogênicos, como BABY BOOM, AGAMOUS-LIKE 15, WUSCHEL, SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE 1 e LEAF COTYLEDON, que podem ser utilizados para induzir a

formação ectópica de embriões somáticos, reduzindo o tempo necessário para a regeneração das plantas após a transformação genética. Nesse contexto, diversas patentes protegem o uso desses genes para melhorar processos de micropropagação *in vitro* e transformação genética. Um exemplo é a patente WO2011003471A1 (Heidmann *et al.* 2011), que protege o uso do fator de transcrição BBM em plantas transgênicas. O BBM é um gene conhecido por induzir a formação espontânea de embriões somáticos quando expresso de forma ectópica (Horstman *et al.* 2017b). Esta patente descreve o uso do BBM para aumentar a capacidade regenerativa de plantas recalcitrantes aos métodos de micropropagação, como a embriogênese somática.

A patente WO2005075655A2 (Gordon-Kamm *et al.* 2006), depositada pela Pioneer Hi Bred International, descreve a utilização do gene Ap2 e da proteína codificada por esse gene em diferentes contextos. A proteína Ap2 já é conhecida na literatura por sua relação com aspectos de desenvolvimento e embriogênese somática, sendo capaz de favorecer esses processos (Tsuwamoto *et al.* 2010). A aplicação WO2005063990A2 (Bruce *et al.* 2005), depositada pela Pioneer Hi-Bred International, Inc., E. I. Dupont De Nemours & Company, protege o uso do gene WUS em plantas geneticamente modificadas para diversos fins. O gene WUS tem um papel central na regulação e promoção da embriogênese somática, além de ser essencial no crescimento celular e organogênese (Vide tópico 1.3 da introdução) (Bouchabké-Coussa *et al.* 2013). Assim, a patente descreve o uso do WUS para a transformação de plantas com o objetivo de promover a embriogênese somática, aumentar a eficiência da transformação genética, estimular o crescimento celular, organogênese e apomixia.

O registro WO2005112609A2 (Niu e Chua 2006), aplicado pela Rockefeller University, descreve o uso do gene PGA67 em plantas transgênicas com fins biotecnológicos. O gene PGA67 está relacionado ao gene WUS, o qual já foi amplamente estudado por sua capacidade de promover a embriogênese somática (Bouchabké-Coussa *et al.* 2013; Zuo *et al.* 2002). Assim, fica evidente que pesquisas em embriogênese somática de plantas podem resultar na proteção de processos e produtos industriais, ilustrando seu grande potencial inovador na biotecnologia vegetal.

3. Objetivos

3.1 Geral

Identificar e descrever redes de interação proteína-proteína envolvidas com a indução de calos embriogênicos e morfogênese in vitro de embriões de plantas.

3.2 Específicos

1. Determinar as proteínas associadas à indução de calos de *C. papaya* e cana-de-açúcar utilizando uma abordagem proteômica de séries temporais.
2. Identificar as proteínas induzidas por concentrações de 2,4-D otimizadas para indução de calos embriogênicos.
3. Estudar redes de fosforilação de proteínas com papel proativo na embriogênese somática de mamoeiro.
4. Definir as fosfoproteínas com papel regulatório envolvidas com a metilação de DNA durante a embriogênese somática de cana-de-açúcar.
5. Estabelecer alvos para o estudo da morfogênese in vitro de cana-de-açúcar.
6. Elaborar um capítulo de livro sobre as bases do uso de ferramentas proteômicas para o melhoramento de plantas.

4. Referências bibliográficas

1. Zhao, P., et al., Does early embryogenesis in eudicots and monocots involve the same mechanism and molecular players? *Plant Physiology*, 2016. 173(1): p. 130-142.
2. Zhong, S., et al., From gametes to zygote: Mechanistic advances and emerging possibilities in plant reproduction. *Plant Physiology*, 2024. 195(1): p. 4-35.
3. Armenta-Medina, A., et al., Developmental and genomic architecture of plant embryogenesis: from model plant to crops. *Plant Commun*, 2021. 2(1): p. 100136.
4. Radoeva, T., et al., Evolution, initiation, and diversity in early plant embryogenesis. *Dev Cell*, 2019. 50(5): p. 533-543.
5. Fehér, A., Somatic embryogenesis—Stress-induced remodeling of plant cell fate. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 2015. 1849(4): p. 385-402.
6. Fehér, A., Callus, Dedifferentiation, Totipotency, Somatic Embryogenesis: What these terms mean in the era of molecular plant biology? *Frontiers in Plant Science*, 2019. 10(April): p. 536.
7. Hojsgaard, D. and E. Hörandl, The rise of apomixis in natural plant populations. *Frontiers in Plant Science*, 2019. Volume 10 - 2019.
8. Boutilier, K., et al., Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *The Plant Cell*, 2002. 14(8): p. 1737-1749.
9. Fehér, A., A common molecular signature indicates the pre-meristematic state of plant calli. *International Journal of Molecular Sciences*, 2023. 24(17): p. 13122.
10. Li, M., et al., Auxin biosynthesis maintains embryo identity and growth during BABY BOOM-induced somatic embryogenesis. *Plant Physiology*, 2021. 188(2): p. 1095-1110.
11. Wójcik, A.M., B. Wójcikowska, and M.D. Gaj, Current perspectives on the auxin-mediated genetic network that controls the induction of somatic embryogenesis in plants. 2020. 21(4): p. 1333.
12. Swarup, R. and B. Péret, AUX/LAX family of auxin influx carriers—an overview. *Frontiers in Plant Science*, 2012. Volume 3 - 2012.
13. Su, Y.H., et al., Establishment of embryonic shoot–root axis is involved in auxin and cytokinin response during Arabidopsis somatic embryogenesis. *Frontiers in Plant Science*, 2015. 5.
14. Chung, H.-H., J.-T. Chen, and W.-C. Chang, Cytokinins induce direct somatic embryogenesis of *Dendrobium chingmai* pink and subsequent plant regeneration. *In Vitro Cellular & Developmental Biology - Plant*, 2005. 41(6): p. 765-769.
15. Murvanidze, N., et al., Cytokinin oxidase/dehydrogenase inhibitors stimulate 2iP to induce direct somatic embryogenesis in *Coffea arabica*. *Plant Growth Regulation*, 2021. 94(2): p. 195-200.

16. Li, S.-M., et al., Cytokinins as central regulators during plant growth and stress response. *Plant Cell Reports*, 2021. 40(2): p. 271-282.
17. Reis, R.S., et al., Pretreatment free of 2,4-dichlorophenoxyacetic acid improves the differentiation of sugarcane somatic embryos by affecting the hormonal balance and the accumulation of reserves. *Plant Cell, Tissue and Organ Culture*, 2021. 145(1): p. 101-115.
18. Gaj, M.D., et al., Hormone-response mutants of *Arabidopsis thaliana* (L.) Heynh. impaired in somatic embryogenesis. *Plant Growth Regulation*, 2006. 49(2): p. 183-197.
19. Su, Y.H., et al., Absciscic acid is required for somatic embryo initiation through mediating spatial auxin response in *Arabidopsis*. *Plant Growth Regulation*, 2013. 69(2): p. 167-176.
20. Kikuchi, A., et al., Absciscic acid and stress treatment are essential for the acquisition of embryogenic competence by carrot somatic cells. *Planta*, 2006. 223(4): p. 637-645.
21. Santner, A., L.I.A. Calderon-Villalobos, and M. Estelle, Plant hormones are versatile chemical regulators of plant growth. *Nature Chemical Biology*, 2009. 5(5): p. 301-307.
22. Neves, M., et al., Modulation of organogenesis and somatic embryogenesis by ethylene: An overview. *Plants*, 2021. 10(6): p. 1208.
23. Zheng, Q., Y. Zheng, and S.E. Perry, AGAMOUS-Like15 promotes somatic embryogenesis in *Arabidopsis* and soybean in part by the control of ethylene biosynthesis and response. *Plant Physiology*, 2013. 161(4): p. 2113-2127.
24. Mira, M.M., E.-S. Adel, and C. Stasolla, Ethylene is integrated into the nitric oxide regulation of *Arabidopsis* somatic embryogenesis. *Journal of Genetic Engineering and Biotechnology*, 2015. 13(1): p. 7-17.
25. Nowak, K., B. Wójcikowska, and M.D. Gaj, ERF022 impacts the induction of somatic embryogenesis in *Arabidopsis* through the ethylene-related pathway. *Planta*, 2015. 241(4): p. 967-985.
26. Ahmadi, B., M.E. Shariatpanahi, and J.A. Teixeira da Silva, Efficient induction of microspore embryogenesis using absciscic acid, jasmonic acid and salicylic acid in *Brassica napus* L. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 2014. 116(3): p. 343-351.
27. Mira, M.M., et al., Jasmonic acid is a downstream component in the modulation of somatic embryogenesis by *Arabidopsis* Class 2 phytochrome. *Journal of Experimental Botany*, 2016. 67(8): p. 2231-2246.
28. Zhang, G., et al., Roles of the wound hormone jasmonate in plant regeneration. *Journal of Experimental Botany*, 2021. 74(4): p. 1198-1206.
29. Gaspar, T., et al., Plant hormones and plant growth regulators in plant tissue culture. *In Vitro Cellular & Developmental Biology - Plant*, 1996. 32(4): p. 272-289.
30. Hazubska-Przybył, T., et al., Somatic embryogenesis of norway spruce and scots pine: Possibility of application in modern forestry. *Forests*, 2022. 13(2): p. 155.
31. Solís-Ramos, L.Y., et al., Somatic embryogenesis in recalcitrant plants, in *Embryogenesis*. 2012, IntechOpen.

32. Guan, Y., et al., Application of somatic embryogenesis in woody plants. *Frontiers in Plant Science*, 2016. Volume 7 - 2016.
33. Martínez, M.T., et al., Holm oak somatic embryogenesis: Current status and future perspectives. *Frontiers in Plant Science*, 2019. Volume 10 - 2019.
34. Efferth, T., Biotechnology applications of plant callus cultures. *Engineering*, 2019. 5(1): p. 50-59.
35. Ramakrishna, W., et al., Anticancer activities of plant secondary metabolites: Rice callus suspension culture as a new paradigm. *Rice Science*, 2021. 28(1): p. 13-30.
36. Babich, O., et al., Modern trends in the in vitro production and use of callus, suspension cells and root cultures of medicinal plants. *Molecules*, 2020. 25(24): p. 5805.
37. Quiroz-Figueroa, F.R., et al., Embryo production through somatic embryogenesis can be used to study cell differentiation in plants. *Plant Cell, Tissue and Organ Culture*, 2006. 86(3): p. 285-301.
38. Salaün, C., L. Lepiniec, and B. Dubreucq, Genetic and molecular control of somatic embryogenesis. *Plants*, 2021. 10(7): p. 1467.
39. Heringer, A.S., C. Santa-Catarina, and V. Silveira, Insights from proteomic studies into plant somatic embryogenesis. *Proteomics*, 2018. 18(5-6): p. e1700265.
40. Wickramasuriya, A.M. and J.M. Dunwell, Global scale transcriptome analysis of *Arabidopsis* embryogenesis in vitro. *BMC Genomics*, 2015. 16(1): p. 301.
41. Tautorus, T.E., L.C. Fowke, and D.I. Dunstan, Somatic embryogenesis in conifers. *Canadian Journal of Botany*, 1991. 69(9): p. 1873-1899.
42. Lowe, K., et al., Morphogenic Regulators Baby boom and Wuschel Improve Monocot Transformation. *The Plant Cell*, 2016. 28(9): p. 1998-2015.
43. Xavier, L., et al., Large-scale regeneration of hermaphrodite emblings of *Carica papaya* L. 'Golden' using early molecular sex determination during embryogenic callus multiplication. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 2021. 146(3): p. 643-649.
44. Valdiani, A., et al., Bioreactor-based advances in plant tissue and cell culture: challenges and prospects. *Critical Reviews in Biotechnology*, 2019. 39(1): p. 20-34.
45. Weckx, S., D. Inzé, and L. Maene, Tissue culture of oil palm: Finding the balance between mass propagation and somaclonal variation. *Frontiers in Plant Science*, 2019. Volume 10 - 2019.
46. Arruda, P., Genetically modified sugarcane for bioenergy generation. *Current Opinion in Biotechnology*, 2012. 23(3): p. 315-22.
47. Getnet, B., Review on in vitro propagation of sugarcane to advance the value of tissue culture. *Agricultural Research & Technology: Open Access Journal*, 2017. 5.
48. Heringer, A.S., et al., Label-free quantitative proteomics of embryogenic and non-embryogenic callus during sugarcane somatic embryogenesis. *PLOS ONE*, 2015. 10(6): p. e0127803.

49. Heringer, A.S., et al., Comparative proteomics analysis of the effect of combined red and blue lights on sugarcane somatic embryogenesis. *Acta Physiologiae Plantarum*, 2017. 39(2).
50. Pereira, M.G., et al., Melhoramento genético do mamoeiro: programa Uenf/Caliman. VII Simpósio do papaya brasileiro: produção e sustentabilidade hídrica. Incaper, Vitória, 2018: p. 1-59.
51. Salinas, I., et al., Assessment of two sex-determining procedures in 'BH-65' papaya from an economical and developmental point of view. *Fruits*, 2018. 73(3): p. 184-190.
52. Vale, E.M., et al., Morphological analyses and variation in carbohydrate content during the maturation of somatic embryos of *Carica papaya*. *Physiology and molecular biology of plants*, 2018. 24(2): p. 295-305.
53. Botini, N., et al., Stage-specific protein regulation during somatic embryo development of *Carica papaya* L. 'Golden'. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 2021. 1869(2): p. 140561.
54. Atkins, P.A.P. and D.F. Voytas, Overcoming bottlenecks in plant gene editing. *Current Opinion in Plant Biology*, 2020. 54: p. 79-84.
55. Chen, J., et al., Significant improvement of apple (*Malus domestica* Borkh.) transgenic plant production by pre-transformation with a Baby boom transcription factor. *Horticulture Research*, 2022. 9.
56. Wang, N., et al., Leaf transformation for efficient random integration and targeted genome modification in maize and sorghum. *Nature Plants*, 2023. 9(2): p. 255-270.
57. Lee, K. and K. Wang, Strategies for genotype-flexible plant transformation. *Current Opinion in Biotechnology*, 2023. 79: p. 102848.
58. Yan, T., et al., Promoting genotype-independent plant transformation by manipulating developmental regulatory genes and/or using nanoparticles. *Plant Cell Reports*, 2023. 42(9): p. 1395-1417.
59. Mohammadi, M.A., et al., Advances in micropropagation, somatic embryogenesis, somatic hybridizations, genetic transformation and cryopreservation for *Passiflora* improvement. *Plant Methods*, 2023. 19(1): p. 50.
60. Ballesteros, D., et al., Current status of the cryopreservation of embryogenic material of woody species. *Frontiers in Plant Science*, 2024. Volume 14 - 2023.
61. Adero, M., J.N. Tripathi, and L. Tripathi, Advances in somatic embryogenesis of banana. *International Journal of Molecular Sciences*, 2023. 24(13): p. 10999.
62. Engelmann, F., Plant cryopreservation: Progress and prospects. In *Vitro Cellular & Developmental Biology - Plant*, 2004. 40(5): p. 427-433.
63. Engelmann, F., Cryopreservation of embryos: An overview, in *Plant Embryo Culture: Methods and Protocols*, T.A. Thorpe and E.C. Yeung, Editors. 2011, Humana Press: Totowa, NJ. p. 155-184.

64. Gonzalez-Arnan, M.T., et al., Development and large scale application of cryopreservation techniques for shoot and somatic embryo cultures of tropical crops. *Plant Cell, Tissue and Organ Culture*, 2008. 92(1): p. 1-13.
65. Pence, V.C., Tissue cryopreservation for plant conservation: Potential and challenges. *International Journal of Plant Sciences*, 2014. 175(1): p. 40-45.
66. Yan, S., et al., Recent advances in proteomics and metabolomics in plants. *Molecular Horticulture*, 2022. 2(1): p. 17.
67. Isaacson, T., et al., Sample extraction techniques for enhanced proteomic analysis of plant tissues. *Nature Protocols*, 2006. 1(2): p. 769-774.
68. Wiśniewski, J.R., et al., Universal sample preparation method for proteome analysis. *Nature Methods*, 2009. 6(5): p. 359-62.
69. Picotti, P., Phosphoproteomics takes it easy. *Nature Biotechnology*, 2015. 33(9): p. 929-930.
70. Aguilar-Hernández, V. and V.M. Loyola-Vargas, Advanced proteomic approaches to elucidate somatic embryogenesis. *Frontiers in Plant Science*, 2018. 9.
71. Varnavides, G., et al., In search of a universal method: A comparative survey of bottom-up proteomics sample preparation methods. *Journal of Proteome Research*, 2022. 21(10): p. 2397-2411.
72. Rai, G.K., et al., Leaf proteome response to drought stress and antioxidant potential in tomato (*Solanum lycopersicum* L.). *Atmosphere*, 2021. 12(8): p. 1021.
73. Han, C., K. Wang, and P. Yang, Gel-based comparative phosphoproteomic analysis on rice embryo during germination. *Plant and Cell Physiology*, 2014. 55(8): p. 1376-1394.
74. YANG, F., et al., Secretomics identifies *Fusarium graminearum* proteins involved in the interaction with barley and wheat. *Molecular Plant Pathology*, 2012. 13(5): p. 445-453.
75. Kim, D.H., et al., Gel-based proteomics approach for detecting low nitrogen-responsive proteins in cultivated rice species. *Physiology and Molecular Biology of Plants*, 2009. 15(1): p. 31-41.
76. Shuken, S.R., An introduction to mass spectrometry-based proteomics. *Journal of Proteome Research*, 2023. 22(7): p. 2151-2171.
77. Abdallah, C., et al., Gel-based and gel-free quantitative proteomics approaches at a glance. *International Journal of Plant Genomics*, 2012. 2012(1): p. 494572.
78. O'Farrell, P.H., High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*, 1975. 250(10): p. 4007-21.
79. Anguraj Vadivel, A.K., Gel-based proteomics in plants: time to move on from the tradition. *Frontiers in Plant Science*, 2015. Volume 6 - 2015.
80. Kim, Y.-I. and J.-Y. Cho, Gel-based proteomics in disease research: Is it still valuable? *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 2019. 1867(1): p. 9-16.

81. Wang, Y., et al., Prestaining of glycoproteins in sodium dodecyl sulfate polyacrylamide gels by dansylhydrazine. *PROTEOMICS*, 2014. 14(11): p. 1322-1327.
82. Giansanti, P., et al., Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nature Protocols*, 2016. 11(5): p. 993-1006.
83. Ercan, H., et al., A practical and analytical comparative study of gel-based top-down and gel-free bottom-up proteomics including unbiased proteoform detection. *Cells*, 2023. 12(5): p. 747.
84. Rödiger, A. and S. Baginsky, Tailored use of targeted proteomics in plant-specific applications. *Frontiers in Plant Science*, 2018. Volume 9 - 2018.
85. Hart-Smith, G., et al., Improved quantitative plant proteomics via the combination of targeted and untargeted data acquisition. *Frontiers in Plant Science*, 2017. Volume 8 - 2017.
86. Chen, Y., et al., Exploring the diversity of plant proteome. *Journal of Integrative Plant Biology*, 2021. 63(7): p. 1197-1210.
87. Li, W. and A.A. Keller, Optimization of targeted plant proteomics using liquid chromatography with tandem mass spectrometry (LC-MS/MS). *ACS Agricultural Science & Technology*, 2023. 3(5): p. 421-431.
88. Chawade, A., et al., Targeted proteomics approach for precision plant breeding. *Journal of Proteome Research*, 2016. 15(2): p. 638-646.
89. Cupp-Sutton, K.A. and S. Wu, High-throughput quantitative top-down proteomics. *Molecular Omics*, 2020. 16(2): p. 91-99.
90. Vandermarliere, E., M. Mueller, and L. Martens, Getting intimate with trypsin, the leading protease in proteomics. *Mass Spectrometry Reviews*, 2013. 32(6): p. 453-465.
91. Tsiatsiani, L. and A.J.R. Heck, Proteomics beyond trypsin. *The FEBS Journal*, 2015. 282(14): p. 2612-2626.
92. Dupree, E.J., et al., A critical review of bottom-up proteomics: The good, the bad, and the future of this field. *Proteomes*, 2020. 8(3): p. 14.
93. Aydoğan, C., Critical review of new advances in food and plant proteomics analyses by nano-LC/MS towards advanced foodomics. *TrAC Trends in Analytical Chemistry*, 2024. 176: p. 117759.
94. Po, A. and C.E. Eyers, Top-down proteomics and the challenges of true proteoform characterization. *Journal of Proteome Research*, 2023. 22(12): p. 3663-3675.
95. Melby, J.A., et al., Novel strategies to address the challenges in top-down proteomics. *Journal of the American Society for Mass Spectrometry*, 2021. 32(6): p. 1278-1294.
96. Nickerson, J.L., et al., Recent advances in top-down proteome sample processing ahead of MS analysis. *Mass Spectrometry Reviews*, 2023. 42(2): p. 457-495.
97. Donnelly, D.P., et al., Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. *Nature Methods*, 2019. 16(7): p. 587-594.

98. Pandeswari, P.B. and V. Sabareesh, Middle-down approach: a choice to sequence and characterize proteins/proteomes by mass spectrometry. *RSC Advances*, 2019. 9(1): p. 313-344.
99. Jin, Y., Y. Yi, and B. Yeung, Mass spectrometric analysis of protein deamidation – A focus on top-down and middle-down mass spectrometry. *Methods*, 2022. 200: p. 58-66.
100. Kunder, S.R., et al., Middle-down proteomics reveals dense sites of methylation and phosphorylation in arginine-rich rna-binding proteins. *Journal of Proteome Research*, 2020. 19(4): p. 1574-1591.
101. Takemori, A., et al., GeLC-FAIMS-MS workflow for in-depth middle-down proteomics. *PROTEOMICS*, 2024. 24(3-4): p. 2200431.

Capítulo 1. Proteomics in crop improvement

Proteomics for crop improvement

Lucas Rodrigues Xavier¹, Laura Eliza de Oliveira Alves¹, Carlos Eduardo¹, Daniel Dastan Rezabala Pacheco¹, Roberta Pena da Paschoa¹, Claudete Santa-Catarina², Vanildo Silveira¹

¹Laboratório de Biotecnologia, Centro de Biociências e Biotecnologia (CBB), Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), Campos dos Goytacazes, RJ, 28013-602, Brazil

²Laboratório de Biologia Celular e Tecidual, CBB-UENF, Campos dos Goytacazes, RJ, Brazil

Abstract

Mass spectrometry has emerged as an indispensable tool in plant biology, offering exceptional versatility for investigating protein-level responses to environmental stimuli. This accessible guide provides a practical overview of quantitative proteomics for readers with limited mass spectrometry experience and a foundational understanding of protein biology. Covering the entire experimental workflow, from sample preparation to data analysis, it clarifies how protein quantities are determined, processed, and interpreted. Particular attention is given to experimental design, emphasizing the importance of biological replication (typically three or more replicates) and condition comparisons, with controls being essential for differential protein analysis. While pairwise comparisons are common, time-resolved and dose-dependent designs are increasingly favored for their robustness in revealing co-regulated proteins and reducing false positives. This chapter highlights the power of proteomics to elucidate plant responses to heterosis, abiotic stresses (e.g., drought, salinity, heat, heavy metals) and biotic stresses (e.g., pathogens), demonstrating its ability to uncover signaling pathways, regulatory proteins, and stress and trait specific markers. It also explores the central role of post-translational modifications, such as phosphorylation, glycosylation, and ubiquitination, in modulating plant immunity, cellular signaling, and stress adaptation. Advanced topics including spatial proteomics and subcellular localization are briefly introduced, alongside recommended readings for deeper exploration. The chapter concludes with a perspective on future directions, underscoring the transformative potential of proteomics in plant science and crop improvement.

Keywords: Plant breeding, sample preparation, proteomics workflow, mass spectrometry.

Corresponding author: Dr. Vanildo Silveira

E-mail: vanildo@uenf.br

Proteomics toolbox for crop breeding

It was clear from the sequencing of the first genomes that the next big challenge for biotechnology would be to discover the functions of each gene encoded in the DNA [1]. In this context, high-throughput approaches applied to investigate multiple layers of regulation of gene expression have evolved and become increasingly popular, with genomics, transcriptomics, proteomics and metabolomics receiving a lot of focus at the beginning [2]. Since then, omics tools have been rapidly incorporated into plant research and have been employed to elucidate biological processes such as stress, embryogenesis, photosynthesis, flowering and reproduction [3]. Regarding proteomics, the development of gel-free platforms, new generations of mass spectrometers and liquid chromatography systems, the application of ion mobility separation in bottom-up approaches, advances in processing algorithms and genome sequencing have led to a significant increase in the number of proteins identified in a single experiment [4]. While researchers initially identified dozens of proteins using gel-based techniques, in a few years this number has risen to tens of thousands. Thus, proteomics has established itself as a fundamental approach to the study of gene function and has provided advances in both basic science and biotechnology [3].

The application of proteomic tools in plant science has made it possible to study the regulation of metabolic pathways, transcription factors and protein kinases/phosphatases, protein-protein interaction networks, post-translational modifications and the cellular localization of proteins during different biological processes [2]. In this sense, working with plant cells, tissues and organs usually involves challenges such as the presence of cellulose, chlorophylls and very abundant proteins (such as Rubisco), where an important factor in the advancement of plant proteomics has been the development of methods for collecting and preparing samples [5, 6]. In addition, development of databases and bioinformatics tools specific to plant organisms has resulted in an increase in the depth of proteomic data analysis, leading to an expanded interpretation of changes in the proteomic profile and hypothesis testing [7-9]. On the other hand, aspects such as the identification of post-translational modifications and the quantification of low abundance proteins still represent challenges for plant proteomics [10]. The existence of cells with different functions in samples of plant tissues and organs (such as leaves and roots) also creates a background that generates doubts when interpreting data. Recently, advances in enrichment techniques for low abundance peptides, targeted proteomics and single cell analysis could help overcome these challenges [11]. This chapter aims to provide an accessible and comprehensive review of how proteomics is applied

to drive plant breeding, highlighting experimental designs, methodological advances, and applications in elucidating molecular responses to abiotic, biotic stresses and heretosis.

A workflow for plant bottom-up proteomics

Experiment design

Several factors influence the identification of proteins by mass spectrometry, including the separation technology used, the mass spectrometer, the search software and the database used [12]. On the other hand, a fundamental step during research into the regulation of the proteome consists of the most appropriate experimental design to answer the biological question. In this sense, during the planning stage of the experiments, the questions to be answered must be clear, as well as the comparisons that will be made (simple comparison, time series or dose-dependent), the sample collection (which organ/tissue, when and how to perform the sampling), the spectrometry data acquisition approach (DDA or DIA) and quantification (labeled or label-free) [2, 4, 11]. Thus, there are guidelines for the minimum information needed to publish data/articles involving proteomics, such as experimental design, the justification for the assay, sampling, parameters for liquid chromatography and mass spectrometry, processing software, electrophoresis and statistical analysis [12].

The most widely accepted experimental design in comparative proteomics involves a minimum of three biological replicates and a direct comparison between two conditions or treatments (e.g., pathogen infection or water stress) [13]. In this context, the inclusion of a control condition is essential for identifying proteins that are differentially accumulated in response to a specific stimulus. However, such pairwise comparisons are susceptible to background noise caused by stochastic fluctuations in protein abundance, which may increase the rate of false-positive identifications. As an alternative, time-resolved and dose-dependent experimental designs have become increasingly prevalent [14, 15]. These approaches enable the identification of co-accumulated proteins under defined conditions, thereby facilitating the discovery of potential regulatory proteins. Time-course and dose-dependent analyses involve the collection of samples at three or more distinct time points or treatment conditions, representing robust strategies for the more accurate identification of molecular targets.

Sample Preparation Methods

The reliability of proteomics data is directly related to the quality of the samples and each type of tissue, organ or cell studied presents unique challenges [16]. In summary, the steps that make up sample preparation for proteomics are: material collection, protein extraction and purification, and proteolytic digestion to obtain peptides [17]. During the extraction process, the method chosen must be able to capture as many substances as possible present in the sample [18]. For samples with a high concentration of contaminants, the trichloroacetic acid/acetone precipitation method followed by resuspension in urea/thiourea buffer usually generates good results [16, 18]. Alternatively, the phenol-based protein extraction method has more steps and is more laborious, but it is a process that can generate highly purified proteins [16]. On the other hand, direct protein extraction/resuspension in urea/thiourea buffer is also a widely used method for samples with few contaminants, such as embryogenic and non-embryogenic callus [19].

Once the proteins have been extracted and purified, the following step normally consists in quantifying the samples in order to measure their concentration before running on a gel or digesting with proteases [5]. The quantification stage involves some challenges such as the limited amount of material (in some cases) and the presence of interferents in the sample (such as urea, detergents and reducing agents) [20]. In this sense, there are different strategies for quantifying total proteins in samples, including methods with and without precipitation steps. Protein precipitation allows samples to be concentrated, the buffer to be altered and interferents to be removed, but these methods usually require more time and reagents than methods without precipitation (e.g. Bradford) [21].

Among the proteases used in the preparation of proteomic samples, trypsin is the most widespread and results in the cleavage of the C-terminal bond in lysine (K) and arginine (R) residues, forming peptides [17]. In addition to trypsin, other proteases are available for use in proteomics analysis, but tryptic digestion provides the highest sequence coverage rate of identified proteins [22]. In addition, trypsin is a low-cost enzyme (compared to other proteases) and easily found, factors that contribute to the high use of this protease in sample preparation for shotgun proteomics [17].

Data Acquisition and Processing

During the quantification and sequencing of peptides, two different approaches can be used for the acquisition of mass spectra, including data-dependent acquisition (DDA) or data-

independent acquisition (DIA) [23]. The DDA acquisition method is the most widely used in proteomics experiments to date, but DIA allows the identification of a greater number of proteins and is gaining ground with the development of new processing algorithms and the sensitivity of mass spectrometers [24]. While the method using DDA selects the precursors (peptides) with the highest intensity in MS1, DIA selects all precursors within a range of m/z in MS1, eliminating the bias of the most intense precursors [25]. Thus, DDA experiments are likely to have problems with missing values and perform better with a high number of biological replicates, while DIA results in the identification of a larger number of proteins, but depends on the development of spectral libraries [25].

A fundamental step during proteome studies is the selection of the protein database and the choice of the most suitable software/algorithm for mass spectra deconvolution, protein/peptide discovery and quantification [26]. During this process, the selection of the appropriate parameters is a key factor in ensuring the reliability of the data, including the tolerance for mass differences of the precursor ion and its fragments [27]. In this context, an important issue that must be considered is that each vendor/company that produces mass spectrometers stores the spectra data in different raw formats, generating incompatibility with processing software in some cases [28]. On the other hand, there are different software packages for performing database searches (DDA/DIA), de novo sequencing (DIA) and also hybrid identification approaches, where de novo sequencing and protein database searches are performed simultaneously [27].

After acquiring and processing the mass spectra of the samples, it is important to check the quality of the biological replicates in order to guarantee the reproducibility and representativeness of the proteomic analysis. For this, correlation analysis (e.g. Pearson) and principal component analysis (PCA) are usually the first choice for the data quality control and reliability stage [29].

Bioinformatics Analysis and statistics

Due to the complexity of mass spectrometry data and omics analysis, an important element when interpreting and communicating proteomics results involves plotting graphs and figures in a consistent and accessible way [30]. In this sense, graphics are present in different stages of proteomic analysis, such as quality control, identification, quantification, functional annotation, study of metabolic pathways and interaction networks, as well as data analysis in

proteome databases [31]. In addition, there are characteristic types of graphs and visualizations used to analyze data relating to different stages of the proteomics assay [30]. These plots include the total ion chromatogram (TIC) for MS1 data and fragmentation spectrum with MS2 data, lollipop plot for visualizing post-translational modifications, correlation analysis using protein intensity data, volcano plot for illustrating differential expression, principal component analysis and clustering for multidimensional experiments [31]. Thus, there are different softwares that can be used to visualize proteomics data, including DEP, Proti, Perseus, ProVision, Eatomic, amica, MSstats and FragPipe-Analyst [32].

In order to give meaning and biological interpretation to proteomics data, the use of tools for functional analysis and identification of regulatory proteins (transcription factors, kinases and phosphatases) is quite common, allowing the data to be explored in a wider, reproducible and more detailed way. For the identification of regulatory proteins, iTAK represents a specific option for plant proteomes [9]. KEGG pathways are a common option for studying metabolic pathways in different organisms, including plants [33]. For the functional annotation of proteins, tools such as g:Profiler, Metascape, OmicsBox and Mercator4 are options used in the study of gene ontology, enrichment analysis and protein domains [34-37]. During this process, it is important to define the BLAST program that will be used, the search database (taxon), the threshold (e-value) that will be used and other important parameters for the reliability of the results.

Proteins don't work alone, so they interact to form complexes with other proteins, shape cellular activity and localization through post-translational changes, control gene expression, signaling cascades and other biological processes [38]. In this context, STRING is a database that allows you to create and analyze protein-protein interaction networks from sources such as automated text mining of the scientific literature, co-expression, genomic context, interaction experiments and complexes from curated sources [8]. The use of STRING with auxiliary tools such as Cytoscape allows the customization of interaction networks and the use of plugins to identify clusters and integrate with experimental data from PTMs [39]. Databases such as Reactome, Pathway Commons and Omnipath can also be used to study interaction networks.

A fundamental step during proteomic analysis is data mining in complex tables with thousands of values, variables and information to identify differentially accumulated proteins [40]. At this point, the processing of the data prior to statistical analysis is fundamental and includes stages of normalization, cleaning, imputation and the handling of missing values [41].

However, there are bottlenecks in the analysis of proteomics data that mask potentially important discoveries due to the superficiality of functional and bioinformatics downstream to data acquisition [42]. For this reason, there are already tutorials to help novice scientists with the statistical analyses involved in comparative bottom-up proteomics [40]. In addition, different software can be used to process tables and generate visualizations for exploratory and functional analysis of proteomics data [42, 43].

Advances and Challenges

Protein post-translational modification

Since several regulatory proteins are represented by a low copy number, protein and peptide enrichment and fractionation methods can be used to reduce sample complexity and enable the identification of low abundance proteins [17]. Post-translational modifications of proteins play a key role in cellular responses and the regulation of various physiological aspects during development and environmental stresses [44]. Different proteomics methodologies can be used to identify and elucidate the role of post-translational regulation of proteins during somatic embryogenesis, including histone modifications, ubiquitination, interaction between proteins and nucleic acids and protein complex formations. Protein phosphorylation dynamics are controlled by kinases and phosphatases, rapidly coordinating changes in cell physiology according to environmental conditions [10]. These post-translational modifications of proteins can be studied using peptide enrichment techniques. In the case of phosphorylation, the main strategies available for the enrichment of phosphopeptides are Immobilized Metal Ions Affinity Chromatography (IMAC) and Metal Oxide Affinity Chromatography (MOAC) [10]. In the IMAC approach, the affinity of phosphate groups for metal ions is exploited and phosphopeptides are enriched using Fe^{3+} , Ga^{3+} and Al^{3+} ions bound to a matrix by chelating agents (such as iminodiacetic acid and nitrilotriacetic acid) [45]. The MOAC strategy uses metal oxides (mainly TiO_2 and ZrO_2) to capture and enrich phosphopeptides [45]. Thus, the study of phosphorylation using phosphopeptide enrichment and gel-free proteomics approaches makes it possible to monitor this post-translational modification on a large scale, allowing the elucidation of new players that regulate biological processes.

Single-cell proteomics analysis

Profiling cells and tissues in bulk is common in omics studies, but these analyses quantify average expression of gene products in populations of cells and tissues, undermining differences in regulation in individual cells [46]. The development of mass spectrometry platforms has made it possible to analyze proteins and their post-translational modifications in single cells, allowing the study of gene function in different cell states at high resolution [47]. In this context, even with the great advances in single-cell RNA sequencing, single-cell proteomics is now of great interest, given that proteins are the main functional players in cells [48]. On the other hand, single cell proteomics is a developing approach that is sensitive to experimental errors and requires precautions to avoid technical and data analysis artifacts that may create bias in experiment quality and reproducibility [49]. New studies have developed strategies to reduce problems with the consistency, sensitivity and coverage of protein quantification during single cell proteomic analysis [50]. In general, single cell proteomics systems involve the isolation and analysis of subpopulations of functionally similar cells or the analysis of isolated single cells when it is not possible to obtain a large sample volume [49].

Subcellular proteomics

Eukaryotic cells are characterized by the compartmentalization of their internal contents into organelles, where the transport and localization of proteins to these organelles is important for essential processes such as programmed cell death, signalling, cell growth and proliferation [51]. Thus, determining the subcellular localization of proteins is important for a comprehensive understanding of cell dynamics in plants [52]. In this context, subcellular proteomics approaches have been used to study proteins located in the nuclear membrane, plasma membrane, chloroplasts, mitochondria, nucleus, cytoplasm and peroxisomes [51]. These studies use different techniques to isolate cell compartments, including ultra-centrifugation, density gradient centrifugation and speed gradient centrifugation, each with a specific organelle purity level and limitations [53].

Proteomics Studies in Abiotic Stress

Plants are constantly exposed to abiotic stress factors such as drought, salinity, extreme temperatures, waterlogging, and nutrient imbalances, all of which can negatively impact growth, development, and productivity. These stresses influence plant physiology at

multiple levels, leading to reduced photosynthesis, altered metabolism, and impaired reproductive success [54]. The increasing frequency and intensity of abiotic stresses due to climate change have become a major concern for global agriculture. According to the Food and Agriculture Organization, climate-related factors, including extreme weather events and changing precipitation patterns, are among the primary threats to food security [55]. Estimates suggest that crop yields may decline significantly in several regions, particularly in areas where water availability and soil salinity are already limiting factors.

Understanding how plants respond and adapt to abiotic stresses is essential for developing strategies to mitigate their negative effects. Among the different approaches used to study plant stress responses, proteomics has provided valuable insights into the molecular mechanisms underlying stress adaptation. By analyzing changes in protein expression, modifications, and interactions, proteomics allows researchers to identify key proteins involved in stress responses and their potential roles in maintaining cellular homeostasis under adverse conditions. Since proteins are direct effectors of physiological processes, investigating their dynamics under stress conditions offers a more comprehensive understanding of how plants adjust their metabolism and signaling pathways to cope with environmental challenges [56].

In the following subsections of this chapter, studies dealing with plant proteome changes in response to the impact of several abiotic stress factors, water (drought, flooding/waterlogging), temperature extremes (cold and heat), salinity and enhanced concentrations of heavy metals, are summarized and, where possible, we provide an update to information about the contribution of proteomic studies for unraveling the mechanisms underlying plants' stress responses.

Water stress

Water possesses numerous chemical and physical properties that make it an essential medium for life. In plants, it is particularly vital, as it is the most abundant compound in the cytosol where critical biochemical reactions occur [57]. Furthermore, its hydraulic properties drive cell expansion and contribute to structural support. However, from an agricultural standpoint, water availability represents a major abiotic factor limiting crop productivity worldwide [58]. Water stress in plants arises under two contrasting conditions: water deficit (drought) and excess water (waterlogging). Different plant species exhibit varying levels of tolerance or sensitivity to these extremes, depending on their ecological adaptations [58].

Drought, characterized by limited water availability and cellular dehydration, is the most widespread abiotic stress globally and poses a severe threat to agricultural yields. This stress affects plants at morphological, physiological, biochemical, and molecular levels, making adaptation a complex trait [59]. Osmotic adjustment in response to drought involves metabolic reprogramming, which is often reflected in changes in the proteome [60]. Since proteins represent the active products of gene expression, proteomic approaches are particularly informative for revealing stress-related mechanisms that may not be evident through DNA or RNA analyses. Over recent decades, root-specific proteomic analyses have been instrumental in linking functional protein groups to drought response, underscoring the central role of roots in detecting and responding to soil water deficits [61-63].

In *Zea mays var. everta* (popcorn), proteomic analysis revealed significant alterations in protein profiles in response to drought stress. A total of 1838 proteins were identified, with 169 differentially accumulated proteins (DAPs) in the drought-tolerant line (L71) and 386 in the drought-sensitive line (L61) [64]. Key proteomic differences between the genotypes involved energy metabolism, photosynthesis, oxidative stress response, and protein synthesis pathways. The tolerant line demonstrated higher biomass and more efficient gas exchange, while the sensitive line exhibited early chlorophyll degradation and increased anthocyanin levels, indicating stress susceptibility [64].

Using a comparative proteomic approach in leaves and roots of *Setaria italica*, tissue-specific responses to drought stress were studied, highlighting the complexity of plant adaptation mechanisms [65]. In leaves, differentially abundant proteins (DAPs) were mainly associated with photosynthesis and carbohydrate metabolism, whereas roots showed enrichment in proteins related to cellular homeostasis and hormone biosynthesis. Notably, proteins such as MAPK6 and PAL were upregulated during drought stress, while PER12 and ADH1 were more abundant during the recovery phase. These findings suggest complementary strategies between organs: leaves prioritize the maintenance of photosynthetic capacity, while roots contribute to stress sensing and structural reinforcement [65].

Similarly, multiple studies have utilized root proteomics to investigate mechanisms of drought tolerance. In this context, Tandem Mass Tag (TMT)-based proteomic approach has been applied to analyze soybean roots subjected to PEG-simulated drought stress [66]. The analysis revealed significant alterations in proteins involved in energy metabolism, antioxidant defense, and signal transduction. Among these, components of the glutathione metabolism pathway, such as GmGSH, GmGST1, and GmCAT, were notably upregulated, suggesting a

central role in maintaining redox balance under stress. Likewise, Ren, Shi [67] employed an iTRAQ-based proteomic strategy to investigate drought stress responses in maize seedlings. This research identified key drought-responsive proteins involved in redox homeostasis, cell wall organization, photosynthesis, abscisic acid biosynthesis, and carbohydrate metabolism [67]. Molecular chaperones, heat shock proteins, and enzymes associated with primary metabolism exhibited significant differential abundance, emphasizing their roles in maintaining cellular function and enhancing drought resilience.

A broader integrative analysis combined transcriptomic and proteomic approaches to investigate drought responses in tomato (*Solanum lycopersicum*) [68]. Among the upregulated proteins, those involved in antioxidant defense mechanisms - particularly heat shock proteins (HSPs) - and osmolyte metabolism, such as enzymes associated with proline biosynthesis, played pivotal roles in enhancing drought tolerance. Additionally, proteins related to signal transduction, including abscisic acid (ABA)-responsive transcription factors, were identified as key regulators of the drought response [68]. The results suggested that drought stress led to elevated ABA levels, which in turn increased the expression of *ABA-response element binding factor*. These transcription factors regulate the expression of HSPs and other protective proteins, thereby promoting adaptive responses to water deficit conditions.

In addition, proteomic responses in shoot tissues were also explored in other species. In *Saccharum ssp*, a comprehensive proteomic analysis of sugarcane stems under drought stress was conducted using a label-free quantitative approach combined with a nuclei-enrichment strategy [69]. Proteins involved in chromatin remodeling, RNA processing, and transcriptional regulation were highlighted, such as transcription factors with domains like bZIP, NAC, and Myb-related, as well as components of the spliceosome and THO/TREX complex. These results suggest a significant role of transcriptional reprogramming and post-transcriptional regulation in sugarcane drought tolerance [69].

Futhermore, in *Nicotiana benthamiana*, chloroplast proteomics under drought stress identified 1087 DAPs, revealing extensive reprogramming of the chloroplast proteome [70]. Notably, carbonic anhydrase 1 (CA1) showed a significant decrease in chloroplasts but a notable increase in leaves under drought conditions. This contrasting distribution suggests that CA1 and similar proteins may be translocated from chloroplasts to the cytosol, acting as retrograde signal messengers to regulate nuclear gene expression. These results underscore the systemic role of chloroplast proteins in coordinating drought responses and emphasize the importance of organelle-specific proteomics in dissecting the intracellular signaling networks involved in plant stress adaptation [70].

In contrast, waterlogging results from excess water in the soil surpassing its absorption capacity, leaving the surface saturated and the root zone deprived of oxygen. The resulting hypoxic or anoxic conditions shift root metabolism from aerobic respiration to anaerobic pathways, such as glycolysis and fermentation, reducing energy efficiency [71]. This stress, often triggered by heavy or prolonged rainfall in poorly drained soils, represents one of the most severe environmental stresses, especially during early plant development [71]. It drastically impairs crop yield, particularly in flood-sensitive species such as soybean, wheat, barley, and maize. In contrast, rice demonstrates natural flood tolerance [71].

Waterlogging imposes a significant challenge to root physiology, prompting a shift in metabolic pathways to sustain energy production under low oxygen conditions. In eggplant (*Solanum melongena* L.), a proteomic investigation revealed that short-term waterlogging induced early and dynamic changes in root protein profiles [72]. The majority of the DAPs were associated with energy metabolism, amino acid biosynthesis, signal transduction, and nitrogen metabolism. Notably, enzymes involved in glycolysis and fermentation, such as fructose-bisphosphate aldolase and alcohol dehydrogenase, were upregulated in a time-dependent manner, reflecting the plant's metabolic adjustment to maintain cellular activity and flexibility under hypoxic conditions [72]. Such metabolic adjustments are not exclusive to eggplant. In soybean, mutant lines generated through gamma irradiation demonstrated enhanced tolerance to flooding [73]. Proteomic analyses comparing wild-type and mutant plants revealed critical differences in protein dynamics. Notably, the mutant lines displayed increased levels of glycoprotein-folding enzymes, such as calnexin and calreticulin, with the latter exclusively accumulating in mutant lines under flooding stress [73]. These proteins improved glycoprotein folding, contributing to endoplasmic reticulum homeostasis and enhanced stress resilience. In addition, fermentation-related proteins like alcohol dehydrogenase (ADH) were identified, although the mutant lines showed more moderate activity of ADH compared to wild-type plants, suggesting a balanced approach to anaerobic energy production. Proteins like exostosin domain-containing protein, crucial for secondary cell-wall biogenesis, also increased in abundance in mutant lines under flooding stress, unlike their reduction in wild-type plants [73]. This improved regulation of cell organization and stress-mitigation processes likely underpins the suppression of cell death observed in mutant lines, enabling a more efficient flooding tolerance mechanism .

Beyond legume and solanaceous species, grapevine has also provided valuable insights into waterlogging adaptation. In *Vitis berlandieri* × *Vitis riparia* (cv. SO4) rootstock, prolonged exposure to soil saturation led to profound changes in the proteome [74]. A total of

529 differentially expressed proteins were primarily associated with glutathione metabolism, carbon fixation, sugar metabolism, and amino acid biosynthesis, reflecting critical adjustments in energy metabolism and redox balance. The physiological impact was evident through increased oxidative damage markers and reduced chlorophyll levels. These findings illustrate how hypoxia-related protein shifts integrate with broader stress responses affecting both energy metabolism and redox balance [74].

Genotypic differences in waterlogging tolerance have also been elucidated through proteomic approaches. In wheat (*Triticum aestivum* L.), an iTRAQ-based proteomic analysis revealed contrasting responses between tolerant and sensitive genotypes under waterlogged conditions [75]. Out of 7710 proteins identified, 16 showed differential accumulation between the two genotypes. Eleven proteins were upregulated in the tolerant genotype, including Fe-S cluster assembly factor, heat shock cognate 70, GTP-binding protein SAR1A-like, and CBS domain-containing protein, molecules linked to redox regulation, protein folding, and stress signal mediation under hypoxia. In contrast, five proteins were downregulated in the sensitive genotype, such as photosystem II reaction center protein H, carotenoid cleavage dioxygenase-like, psbP-like protein 1, and mitochondrial ATPase inhibitor, suggesting impaired photosynthetic efficiency and disrupted energy metabolism in response to oxygen deprivation [75].

In peanut (*Arachis hypogaea*), waterlogging stress induces significant metabolic adjustments, particularly in energy production pathways [76]. Comparative proteomic analyses between tolerant and sensitive cultivars revealed that the tolerant genotype exhibited elevated levels of key enzymes such as alcohol dehydrogenase (ADH), malate dehydrogenase (MDH), and lactate dehydrogenase (LDH). These enzymes are essential for anaerobic respiration, malate metabolism, and the glyoxylic acid cycle, allowing the plant to efficiently generate energy from alternative carbon sources under oxygen-limited conditions [76]. Additionally, the tolerant cultivar demonstrated prolonged LDH activity and better regulation of ADH and MDH activities, contributing to reduced accumulation of toxic substances and improved root integrity .

Collectively, these findings underscore the complexity and plasticity of plant responses to water-related stress. Proteomic studies consistently reveal that successful adaptation involves coordinated changes in energy metabolism, redox balance, protein turnover, and signal transduction, pathways that are central to survival under both drought and waterlogging conditions.

Temperature Stress

Temperature is one of the most critical environmental variables influencing plant development and productivity [77]. Both high and low temperature extremes, commonly referred to as heat and cold stress, can significantly alter plant metabolism, cellular structures, and overall physiological functions [78]. Plants have evolved various mechanisms to perceive, respond to, and survive such temperature fluctuations, many of which are reflected at the proteomic level [77]. The study of temperature-related proteomic changes offers valuable insights into how plants acclimate to and tolerate thermal stress conditions.

Among temperature-related stresses, heat stress has been particularly well studied from a proteomic perspective. Defined as the exposure of plants to temperatures above their optimal physiological range, heat stress leads to a series of deleterious effects at the cellular and molecular levels [77, 78]. It disrupts protein structure, induces reactive oxygen species (ROS) accumulation, impairs photosynthesis, and affects metabolic homeostasis. As a result, plants rely on intricate stress-response mechanisms to counteract thermal damage, many of which are reflected at the proteomic level [79]. Proteomic studies have provided valuable insights into how different species, organs, and genotypes modulate their protein profiles in response to heat, shedding light on key components involved in stress perception, signaling, and tolerance [80].

For instance, a proteomic study in *Triticum aestivum* investigated the molecular basis of thermotolerance by comparing heat-tolerant and heat-sensitive genotypes [81]. Through phosphoproteomic analysis, the study identified 1594 phosphopeptides corresponding to 875 unique proteins [81]. In the heat-tolerant genotype, there was a significant enrichment of proteins involved in translational and post-translational modifications, protein turnover, and chaperone binding, particularly within the mitogen-activated protein kinase (MAPK) signaling pathway. Key phosphoproteins included ZmMKK2 and ZmMPK7, which are central to MAPK cascade-mediated stress signaling. Moreover, the accumulation of heat shock proteins, notably HSP70, along with transcription factors such as bZIP and MYB, suggested enhanced regulation of stress-responsive pathways and improved cellular adaptation [81]. In contrast, the heat-sensitive genotype displayed downregulation of enzymes linked to photosynthesis, indicating impaired energy metabolism and reduced capacity to cope with heat-induced damage.

Proteomic analysis of *Solanum tuberosum* (potato) roots under drought and high temperature stress revealed distinct molecular responses between tolerant (cv. Gwiazda) and sensitive (cv. Oberon) cultivars [82]. In the tolerant cultivar, proteins associated with energy and carbohydrate metabolism, such as enolase and malate dehydrogenase, were significantly upregulated under both stress conditions, indicating a capacity for enhanced metabolic flexibility and sustained energy production [82]. Conversely, the sensitive cultivar displayed a marked increase in proteins related to defense and detoxification, suggesting a reactive response rather than a pre-emptive or adaptive strategy. Moreover, elevated temperatures led to a decrease of proteins involved in cellular energy metabolism in the sensitive cultivar, while the tolerant one maintained higher levels of these proteins, suggesting a more robust energy production system [82].

In *Nicotiana tabacum*, integrated proteomic and physiological analyses of heat-stressed leaves treated with melatonin revealed significant modulation of photosynthesis-related proteins [83]. Among the upregulated proteins were glutamyl-tRNA reductase and monomethyl ester aerobic oxidative cyclase, both involved in chlorophyll biosynthesis, as well as PSII cytochrome b559 and fructose-1,6-bisphosphatase, associated with the photochemical activity of photosystem II and the Calvin cycle, respectively [83]. The study highlighted melatonin's role in enhancing photosynthetic efficiency and mitigating photoinhibition, emphasizing its potential in improving plant thermotolerance.

In *Gossypium hirsutum*, a proteomic analysis under heat stress identified 701 and 1270 differentially expressed proteins in tolerant and susceptible genotypes, respectively [84]. Key upregulated proteins in the tolerant genotype included those involved in photosynthesis, oxidoreductase activity, and carbohydrate metabolism, reflecting enhanced stress adaptation. In contrast, the susceptible genotype showed significant downregulation of proteins linked to these pathways, indicating impaired stress responses [84]. These findings highlight the molecular mechanisms underlying heat tolerance in cotton and provide potential targets for breeding programs.

In rice, temperature stress during reproductive development can affect grain quality. For example, elevated temperatures during the grain-filling stage in *Oryza sativa* modulate the abundance of stimulus-responsive proteins involved in storage substance accumulation and stress adaptation [85]. Proteomic analysis identified key proteins such as alpha-amylase isozyme 3E and seed allergy protein RA5, which regulate starch and protein accumulation under high temperatures. Additionally, reactive oxygen species-related proteins like cationic peroxidase SPC4 and persulfide dioxygenase ETHE1 homolog were upregulated, contributing

to oxidative stress mitigation and physiological homeostasis [85]. These findings highlight the role of signal transduction and metabolic pathways in rice quality formation under elevated temperature, offering insights into molecular targets for improving grain resilience and quality.

In contrast, cold stress, triggered by chilling ($<20\text{ }^{\circ}\text{C}$) or freezing ($<0\text{ }^{\circ}\text{C}$) temperatures, causes a different set of physiological challenges. The reduction in temperature slows down enzyme-catalyzed reactions, disrupting metabolic balance and often resulting in oxidative stress due to an accumulation of reactive oxygen species. Moreover, low temperatures impair water uptake, leading to cellular dehydration, which further exacerbates stress [86]. To counteract these adverse effects, plants activate a suite of adaptive responses, including the stimulation of primary metabolic pathways to increase energy production, upregulation of antioxidant systems and chaperone proteins, and modifications to membrane composition that help maintain osmotic balance [87]. These responses are reflected in profound changes in the plant proteome under cold stress, underscoring the importance of proteomic studies in understanding cold acclimation and tolerance.

In *O. sativa*, a crop particularly sensitive to chilling stress, proteomic analysis revealed the differential accumulation of proteins related to antioxidant defense, phenylpropanoid biosynthesis, and inositol phosphate metabolism under low-temperature conditions [88]. In seeds primed with gibberellin A₃ (GA₃) and brassinolide (BR), key proteins such as peroxidase, phenylalanine ammonia-lyase, and inositol phosphate synthase were upregulated. These findings highlight the role of enhanced antioxidant capacity and metabolic regulation in promoting germination and resilience under chilling stress. In a study with *japonica* rice, cold stress elicited significant proteomic changes, as revealed by an iTRAQ-based quantitative proteomic analysis [89]. A total of 5082 proteins were identified, with 289 being significantly regulated under cold conditions. Functional analysis highlighted the involvement of these proteins in photosynthesis, metabolic pathways, biosynthesis of secondary metabolites, and carbon metabolism [89].

Similarly, in *G. max*, low-temperature stress elicited contrasting proteomic and transcriptomic responses in cold-tolerant (HX3) and cold-sensitive (HH43) soybean varieties [90]. Proteomic analysis revealed that cold-tolerant varieties exhibited broader and sustained upregulation of proteins associated with metabolic activity and cellular homeostasis, while cold-sensitive varieties showed limited proteomic responses and degradation of early responsive proteins. Key findings included the identification of transcription factors such as ATAF1 and their potential targets, which likely contribute to the enhanced cold tolerance in HX3 [90].

In *Triticum aestivum*, label-free quantitative proteomics revealed distinct responses to spring freezing stress in freezing-tolerant (JM22) and freezing-sensitive (XM33) wheat varieties [91]. A total of 5050 proteins were identified, with 392 and 103 differentially expressed proteins (DEPs) in XM33 and JM22, respectively. Key DEPs included antioxidant enzymes, heat shock proteins, and cell wall-modifying proteins, which were enriched in processes such as redox homeostasis, protein folding, and cell wall remodeling [91].

In *Sa. spontaneum*, low-temperature stress elicited contrasting proteomic and physiological responses in cold-tolerant (1027) and cold-sensitive (3217) genotypes [92]. Proteomic analysis revealed that the cold-tolerant genotype exhibited sustained upregulation of proteins associated with antioxidant defense, signal transduction, and cytoskeletal stability, while the cold-sensitive genotype showed limited proteomic responses and higher levels of oxidative damage. Key findings included the identification of differentially abundant proteins involved in carbohydrate metabolism and amino acid transport, which likely contribute to the enhanced cold tolerance in genotype 1027 [92].

These studies reflect a highly integrated network of metabolic, homeostatic, and signaling processes, whose precise regulation underpins a plant's ability to endure thermal extremes. Beyond changes in protein abundance, thermotolerance and cold acclimation are also shaped by post-translational modifications (PTMs), organelle-specific responses, and spatially coordinated signaling events that collectively safeguard cellular function and developmental integrity under heat and cold stress.

Salt Stress

Salt stress, primarily caused by the excessive accumulation of ions such as sodium (Na^+), chloride (Cl^-), calcium (Ca^{2+}), potassium (K^+), carbonate ($(\text{CO}_3)^{2-}$), nitrate (NO_3^-), and sulfate ($(\text{SO}_4)^{2-}$) in the soil, leads to a reduction in soil water potential. This decrease hampers water uptake by plant roots, initiating osmotic stress [93, 94]. Salt stress is a major constraint on agricultural productivity, especially in hot, dry, and semi-arid regions where crop cultivation often relies on irrigation. Globally, salinity affects at least 20% of irrigated lands, significantly reducing crop yields [95]. Roots, being the first organs to sense salinity in the soil, play a central role in stress perception and response. Advances in proteomic approaches have provided valuable insights into the molecular mechanisms underlying root adaptation to salt stress.

In *T. aestivum*, salt stress disrupted chloroplast protein homeostasis, particularly affecting plastid-localized protein synthesis machinery, including 30S and 50S ribosomal proteins and elongation factors [96]. Proteomic analysis revealed significant reductions in these proteins, correlating with altered K⁺/Na⁺ ratios along the leaf gradient. Additionally, decreased abundance of plastid proteases, such as ATP-dependent caseinolytic protease, highlighted impaired protein turnover. These findings underscore the critical role of chloroplast protein homeostasis in maintaining wheat resilience under salinity stress [96].

In *Avena sativa*, label-free proteomic analysis under salt stress revealed significant alterations in root respiratory metabolism. Key findings included differential expression of proteins involved in glycolysis, the tricarboxylic acid cycle, and the pentose phosphate pathway [97]. Notably, salt-tolerant cultivars exhibited enhanced antioxidant enzyme activity and pyruvate–ethanol fermentation, while salt-sensitive cultivars relied more on pyruvate–lactic acid fermentation. These results highlight distinct metabolic strategies employed by oat genotypes to adapt to salinity stress [97].

Similarly, in *Gossypium hirsutum*, proteomic analysis during salt stress at the seedling stage identified 3107 proteins, with key pathways including carbon metabolism, glutathione metabolism, and amino acid biosynthesis [98]. Notably, proteins classified within the CAZy (Carbohydrate-Active Enzyme) family and expansin proteins were highlighted for their roles in stress adaptation. These findings provide a foundation for breeding salt-tolerant cotton varieties and improving resilience in saline environments [98].

In *Pfaffia glomerata*, proteomic analysis under salt stress revealed 282 differentially accumulated proteins [99]. Key alterations involved proteins associated with amino acid and carbohydrate metabolism, as well as stress response and detoxification mechanisms. NaCl treatment led to the upregulation of proteins involved in osmotic homeostasis, including enzymes related to sucrose and proline metabolism. Additionally, antioxidant proteins such as superoxide dismutase and ascorbate peroxidase were significantly regulated, highlighting their role in mitigating oxidative stress induced by salinity [99].

In *Saccharum spp.*, proteomic analysis under salt stress revealed significant alterations in protein abundance in micropropagated sugarcane shoots [100]. A total of 1172 non-redundant proteins were identified, with 1160 observed in both cultivars (CB38-22 and RB855536) under control and salt stress conditions [100]. The salt-tolerant cultivar RB855536 exhibited a higher abundance of proteins involved in non-enzymatic antioxidant mechanisms, ion transport, and photosynthesis, compared to CB38-22. Notably, proteins such as calcium-

dependent protein kinase, photosystem I, phospholipase D, and glyceraldehyde-3-phosphate dehydrogenase were more abundant in RB855536 under salt stress, suggesting their role in maintaining ionic and osmotic homeostasis [100].

In *S. lycopersicum*, halophyte-based crop managements under saline conditions revealed significant proteomic changes [101]. Proteomic analysis highlighted the accumulation of photosynthesis-related proteins and sugar metabolism compounds under crop rotation, enhancing photosynthetic efficiency. Additionally, acylamino acid-releasing enzymes, a class of serine-proteases, were notably increased in both intercropping and crop rotation systems, suggesting their role in modulating antioxidant capacity. These findings underscore the potential of halophyte integration in sustainable agriculture to improve crop resilience and productivity [101].

Moreover, in *Glycine max*, integrated transcriptomic, proteomic, and metabolomic analyses under salt stress revealed significant molecular adaptations [102]. Proteomic results highlighted salt-induced DNA methylation and the enrichment of phosphopyruvate hydratase, alongside interactions with enzymes from metabolic pathways such as alanine, aspartate, and glutamate metabolism. Metabolomic findings emphasized the synthesis of lipids and flavonoids, while transcriptomic data identified key transcription factors and hormone-related genes involved in stress signaling [102].

In *Medicago sativa*, integrated transcriptomic and proteomic analyses have revealed genotype-specific mechanisms of salt stress response [103]. In a comparative study between salt-tolerant (ZM-1) and salt-sensitive (AT) genotypes, proteomic profiling identified differentially expressed proteins (DEPs) involved in key metabolic pathways, including glutathione metabolism, biosynthesis of secondary metabolites, gluconeogenesis, carbon fixation, and photosynthesis [103]. In a complementary root-specific analysis, 52 DEPs were identified, with enriched pathways related to antioxidation, protein quality control, and calcium signal transduction [104]. Notably, proteins such as ascorbate peroxidase (APX), pathogenesis-related protein 10 (PR-10), and cystathionine beta-synthase (CBS) were upregulated, enhancing the plant's defense mechanisms. Additionally, the study emphasized the role of unfolded protein response (UPR) and ubiquitin-proteasome system (UPS) in maintaining protein homeostasis under stress. Annexin (MsANN2), a calcium-binding protein, was identified as a critical factor, and its overexpression in *Arabidopsis* conferred enhanced salt tolerance, underscoring its potential in improving stress resilience in alfalfa [104].

These studies reinforce the view that the proteomic response to salinity stress goes beyond basic ionic regulation. It integrates metabolic, antioxidant, and signaling processes, whose fine coordination determines a plant's ability to maintain cellular function under saline conditions. The regulation of salt stress responses involves not only changes in protein abundance but also protein turnover, redox control, and ion compartmentalization mechanisms critical for sustaining homeostasis.

Heavy metal Stress

The extensive and uncontrolled use of chemicals over recent decades has resulted in a widespread accumulation of environmental pollutants. Among these, heavy metals such as cadmium (Cd), arsenic (As), and lead (Pb) are particularly concerning due to their persistence and frequent detection in contaminated soils and groundwater. As sessile organisms, plants are continuously exposed to these toxic elements, making the investigation of their physiological and molecular responses to heavy metal stress a critical area of study [105]. Exposure to metal-contaminated soils disrupts numerous plant cellular processes. Heavy metal ions can bind to sulfhydryl groups in proteins, inactivate enzymes, displace essential cations from specific binding sites, and promote the generation of reactive oxygen species (ROS), leading to oxidative damage to lipids, proteins, and nucleic acids [106]. These harmful effects not only reduce crop productivity but also pose serious ecological risks to plants and animals alike [107].

In *Zea mays*, multiple studies have demonstrated genotype-specific and compound-specific responses to heavy metal stress [108]. In *Z. mays* var. *everta* (popcorn), label-free quantitative proteomic analysis revealed key metabolic pathways associated with aluminum (Al) tolerance in roots. A total of 489 differentially accumulated proteins (DAPs) in the Al-sensitive inbred line, 491 in the Al-tolerant inbred line, and 277 in the commercial hybrid were identified [108]. Conserved tolerance mechanisms as starch, sucrose, and carbohydrate metabolism, as well as glycolysis/gluconeogenesis, were significantly regulated in response to Al toxicity. The accumulation of sucrose synthase and increased sucrose content, along with starch degradation, suggest that these components enhance popcorn tolerance to Al stress [108]. Additionally, citrate synthase accumulation in the Al-tolerant inbred line points to its role in detoxification processes.

Another investigation demonstrated that exogenous application of cysteine (Cys) mitigated chromium (Cr) toxicity by limiting Cr translocation to aerial tissues and enhancing its sequestration in roots [109]. Sixty proteins were affected by Cr exposure, with 46 DEPs identified, primarily involved in oxidative stress response, energy and carbohydrate metabolism, protein turnover, and amino acid metabolism. Enzymes such as glutathione peroxidase, glutathione S-transferase, glyoxalase, and superoxide dismutase were regulated by Cys, contributing to Cr detoxification [109]. In addition, TMT-based comparative proteomic study explored cadmium (Cd) stress responses in two maize genotypes, B73 and ZmHMA3. B73 exhibited enriched pathways related to amino sugar and nucleotide sugar metabolism, while ZmHMA3 displayed upregulation of proteins associated with phenylpropanoid biosynthesis, nitrogen metabolism, and glyoxylate and dicarboxylate metabolism [109].

Proteomic studies in *Brassica napus* have contributed significantly to our understanding of plant tolerance to heavy metals [110]. Under Cd, Pb, and As stress, enhanced activity of antioxidant enzymes, such as peroxidase (POD), and increased abundance of transporters, particularly ATP-binding cassette (ABC) proteins, were observed, highlighting their central roles in mitigating oxidative stress and facilitating metal detoxification [110]. Absciscic acid (ABA) signaling was also implicated in modulating these responses, underscoring the importance of hormonal regulation in metal stress adaptation. Additionally, another study demonstrated that citric acid (CA) supplementation altered global protein expression under Cd stress. Key differentially expressed proteins (DEPs) were involved in energy metabolism, photosynthesis, signal transduction, protein turnover, and antioxidant defense [111]. Among them, enzymes such as superoxide dismutase, ascorbate peroxidase, glutathione S-transferase DHAR1, cysteine synthase, and S-adenosylmethionine synthase 2 played central roles in redox regulation and sulfur assimilation, contributing to improved stress resilience [111].

In *N. tabacum*, combined transcriptomic and proteomic analyses under varying zinc (Zn) availability identified molecular responses critical for Zn homeostasis and oxidative stress mitigation [112]. Notably, transport proteins such as ZIPs, NAS3s, and HMA1s were upregulated under Zn deficiency or limitation, while antioxidant enzymes like SODs, APX1s, GPX6, and GSTs were induced to counter oxidative damage [112]. Likewise, *Pogonatherum crinitum*, a plant known for its phytoremediation potential, exhibited robust proteomic responses under Pb stress [113]. A total of 118 DEPs were identified, including PcCAT3, PcCht1, PcSPS5, and PcGME-1, which contributed to energy regulation and carbohydrate metabolism. Proteins such as PcHSP26.7 and PcCAT1 were linked to protein synthesis and

oxidative stress balance. These proteins were associated with MAPK signaling, amino sugar metabolism, and starch and sucrose pathways, suggesting an integrated strategy for Pb tolerance [113].

In *Populus deltoides*, proteomic comparisons between arsenic-tolerant and -sensitive cultivars revealed the significant upregulation of PdDRT102 in the tolerant line under arsenate exposure [114]. This protein was localized to both the nucleus and plasma membrane, with increased nuclear signal under stress. Functional analysis in *Arabidopsis* confirmed its role in improving tolerance to arsenate and salt stress, indicating its relevance in broad-spectrum environmental stress responses [114]. Collectively, these studies underscore the multifaceted nature of plant responses to heavy metal stress. Proteomic investigations reveal that successful tolerance involves a coordinated network of detoxification mechanisms, antioxidant defenses, metabolic adjustments, and signal transduction pathways. As heavy metal contamination remains a major agricultural and environmental concern, proteomics provides a valuable framework for uncovering adaptive strategies and identifying candidate proteins for crop improvement programs focused on enhancing stress tolerance and phytoremediation potential.

Proteomics and biotic stress

In recent years, the application of mass spectrometry has revolutionized plant proteomics, enabling investigations at an unprecedented level of detail [4]. Once considered an isolated field, proteomics has become a central approach in biological research, offering fundamental insights into the molecular mechanisms involved in plant responses to various environmental challenges [2]. Recent studies show that this approach has been crucial for understanding plant defense mechanisms, particularly in the context of biotic stress. Among the most promising aspects, the spatial organization of the proteome and post-translational modifications (PTMs) of proteins stand out, playing key roles in plant immune responses [10]. Thus, the analysis of subcellular protein localization and their modifications has emerged as an essential element for an in-depth understanding of plant defense mechanisms. In this chapter, we explore the latest proteomic methodologies applied to the study of plant-pathogen interactions, emphasizing mass spectrometry-based approaches, spatial proteomics, and PTM identification. In addition to presenting the most relevant advances in the field, we discuss current challenges and future perspectives for plant proteomics research [10].

Proteomic responses to biotic stress

The proteomic response of plants to biotic stress is highly dynamic and multifaceted, reflecting the complex signaling networks that coordinate defense, metabolism, and cellular homeostasis [115]. Recent studies have shown that this response goes far beyond the mere induction of defense proteins, integrating metabolic pathways, PTMs, hormonal regulation, and intra- and intercellular signaling processes. Proteomic analyses in different species have revealed that, during pathogen infection, plants extensively reconfigure their proteomes, activating coordinated defense mechanisms while repressing non-essential metabolic pathways. For example, studies on *Mangifera indica* L. (manga) leaves infected by *Alternaria alternata* identified 472 differentially expressed proteins, including enzymes from the phenylpropanoid, flavonoid, and MAPK signaling pathways. Upregulated proteins were associated with tyrosine metabolism and MAPK signaling, reflecting a multifactorial response to biotic stress [116].

In *O. sativa* L. infected by *Magnaporthe oryzae*, contrasting genotypes displayed distinct temporal responses, with 18 modulated proteins related to defense (e.g., Pik-1, Argonaute proteins), jasmonate biosynthesis (AOS), and photosynthesis regulation (rubisco, cytochrome b6-f), highlighting genotype-specific expression profiles [117]. Moreover, resistance to *M. oryzae* (rice blast) and *Xanthomonas oryzae* (bacterial blight) in rice can also be modulated by post-translational mechanisms such as protein glycosylation. In the *rbb1* mutant, the loss of glucosamine-6-phosphate acetyltransferase activity reduced UDP-GlcNAc levels, compromising glycosylation of defense proteins like Prx4 and Prx13. Interestingly, deletion of these proteins increased resistance, suggesting glycosylation acts as a negative regulator of immune responses [118]. These findings reveal a novel defense mechanism in rice and indicate potential molecular targets for the development of disease-resistant varieties.

Proteomics in fruit species has also emerged as a key approach to uncover postharvest resistance mechanisms and identify biomarkers. Comparative studies in fruits like tomato, apple, banana, kiwi, peach, pitaya, citrus, and walnut show that biotic stress triggers significant alterations in antioxidant defense pathways, hormonal metabolism, ROS homeostasis, and PTMs [119-126]. Advanced techniques such as quantitative proteomics, PTM profiling, and single-cell proteomics are being applied to identify infection-sensitive proteins, with potential applications in genetic improvement and biotechnological disease control in fruit crops.

Integrated proteomic regulation of basal immunity and defense signaling

In *S. lycopersicum*, the overexpression of the BiP chaperone enhanced protein homeostasis in the endoplasmic reticulum, resulting in reduced accumulation of reactive oxygen species and mitigation of oxidative stress [127]. This regulation was accompanied by the induction of pathogenesis-related proteins (PR2, PR3, PR10), antioxidants such as SOD, CAT, and peroxidases, as well as proteins containing leucine-rich repeat (LRR) domains, indicating a state of immune readiness even in the absence of the pathogen [127].

In *Z. mays*, genotypes resistant to *Bipolaris maydis* infection exhibited a coordinated response involving photosystem II proteins (PSB27-1), antioxidants, and ribosomal proteins such as 60S P0, all associated with translational integrity and defense [13]. The enzyme DHAR3 (dehydroascorbate reductase type-3) contributed to the maintenance of redox homeostasis. In contrast, the susceptible genotype exclusively expressed GAPC2, a glycolytic enzyme, suggesting an imbalance in energy metabolism under infection conditions [13]. These findings reinforce the notion that the proteomic response to biotic stress goes beyond the mere activation of defense-related genes. It involves a complex integration of energy regulation, antioxidant mechanisms, and molecular signaling, which together define the threshold between resistance and susceptibility. In addition to protein abundance, PTMs and spatial control of cellular signaling emerge as critical elements of this refined regulation.

This complexity was further demonstrated in *Actinidia chinensis* (Kiwi), where exogenous application of salicylic acid (SA) enhanced the defense response against *Pseudomonas syringae* pv. *actinidiae* (Psa), the causal agent of bacterial canker [128]. Using 4D proteomics, it was observed that both Psa infection and SA treatment triggered profound changes in the plant's proteomic profile. SA rapidly activated key immune components, such as the MAPK signaling cascade, the biosynthesis of phenolic compounds via the phenylpropanoid pathway, and the induction of PR-proteins. Moreover, SA was able to restore pathways previously suppressed by infection, highlighting its central role in hormonal reprogramming and in promoting an immunologically primed state for future challenges [128]. Overall, the studies discussed in this section highlight that the proteomic response to biotic stress is highly integrated and multifaceted, encompassing processes from redox and energy homeostasis to the orchestrated activation of immune defense networks.

Post-Translational Modifications (PTMs) as critical regulatory layers

PTMs such as acetylation, glycosylation, phosphorylation, and ubiquitination play central roles in modulating plant immunity. In *A. thaliana*, a refined proteomic approach revealed the functional presence of extracellular proteasomes in the apoplastic fraction traditionally associated with the cytosolic matrix [129]. These complexes participate in the release of microbe-associated molecular patterns (MAMPs), such as the peptide flg22, triggering reactive oxygen species (ROS) bursts, an essential step in innate immunity. Furthermore, the bacterium *Pseudomonas syringae* was shown to inhibit this response via its effector syringolin-A, highlighting ex-proteasomes as strategic targets in plant–pathogen interactions [129].

Similarly, in *O. sativa*, resistance to the fungus *M. oryzae* was associated with reduced glycosylation of the peroxidases Prx4 and Prx13 in the *rbb1* mutant, suggesting a negative role of this modification in immunity [118]. In addition to glycosylation, other PTMs such as phosphorylation, acetylation, and ubiquitination also contribute to the fine-tuning of basal immunity against MAMPs derived from *M. oryzae*. In response to the fungal PAMP MSP1, secreted by this pathogen, a coordinated reconfiguration of the phosphoproteome, ubiquitinome, and acetylproteome profiles was identified in both wild-type and transgenic plants overexpressing MSP1 [130]. Using high-resolution mass spectrometry, the study revealed early activation of immune pathways upon PAMP recognition, underscoring the central role of PTMs in PAMP-triggered immunity (PTI) and the identification of new molecular targets for resistance against *M. oryzae* [130].

In agronomically important crops, such as *Triticum aestivum* L (wheat), infection by *wheat yellow mosaic virus* (WYMV) altered 176 glycopeptides, notably the isomerase TaPDI 1-4 [131]. In this context, the viral protein P1, encoded by WYMV, was capable of suppressing ROS production by repressing TaPOD and TaTrxR. Silencing of these genes restored ROS levels and enhanced antiviral resistance, evidencing the pathogen's manipulation of oxidative pathways [132]. This was the first study to map N-glycosylation during viral infection in wheat, revealing the critical role of PTMs in regulating antiviral responses. Complementarily, a multi-omics study mapped 44000 proteins, 20000 phosphoproteins, and 12000 acetylated proteins in wheat, showing how modifications such as the deacetylation of TaP5CS1 by TaHDA9 affect proline accumulation and stress resistance, in addition to influencing grain quality [133].

These findings are further supported by data from *A. thaliana*, where infection by *Pseudomonas syringae* led to differential acetylation and succinylation of photosynthetic

proteins [134]. In parallel, overexpression of the transcription factor NAC90 impaired NHP-induced systemic resistance [135]. These results illustrate the profound impact of PTMs on plant immune and physiological mechanisms under biotic stress. Indeed, proteomic analysis has revealed the complexity of plant responses to stress, providing a comprehensive view of the molecular mechanisms activated during interactions with pathogens.

PAMP perception remains a central trigger in the activation of basal immunity. Reinforcing this view, studies in *O. sativa* revealed that MSP1 recognition by defense receptors promotes extensive and coordinated changes in PTMs, significantly affecting immune signaling pathways [130]. The comparison of proteomic profiles between wild-type and MSP1-overexpressing plants highlights the critical regulatory function of PTMs in the context of PTI, demonstrating their potential as targets for breeding disease-resistant cultivars [130]. In summary, post-translational modifications act as a versatile and highly dynamic regulatory layer in plant defense responses. By modulating processes such as pathogen recognition, redox signaling, transcriptional regulation, and metabolic reprogramming, PTMs not only integrate but also amplify the plasticity of immune responses. The systematic identification of these modifications in multiple plant models and pathogen types strengthens their role as key components of immunity and reveals strategic opportunities for developing cultivars with greater resilience to biotic stress.

Energy metabolism, redox signaling, and metabolic reprogramming

Pathogen-induced stimuli often disrupt energy metabolism and redox homeostasis, requiring a coordinated metabolic reprogramming. In *Cajanus cajan* (pigeon pea), expression of the *Osoxo4* gene (from rice) conferred resistance to vascular wilt caused by *Fusarium udum* [136]. Proteomic and metabolomic analyses revealed the accumulation of regulatory nuclear proteins, such as helicases and nucleosome assembly factors, along with increased levels of sugars, amino acids, oxalic acids, sugar alcohols, and myo-inositol. Starch and inositol metabolic pathways were activated, reflecting a metabolic shift aimed at adaptation to oxidative stress and defense against the pathogen [136].

Extracellular matrix, apoplast, and systemic signaling

Analysis of the extracellular matrix, particularly the apoplast, has revealed critical components of plant immunity. In cereals such as rice, wheat, and barley (*Hordeum vulgare*), infection by different *M. oryzae* pathovars led to the identification of differentially abundant proteins in the apoplast, from both the host and the pathogen [137]. The response was species-specific: in wheat, there was an increase in serine proteases and peroxidases associated with defense; in rice, glycosidases and carbohydrate metabolism-related proteins stood out, suggesting their involvement in pathogen resistance; in barley, proteins involved in hydrogen peroxide catalysis were particularly enriched, indicating a central role of peroxidases in ROS detoxification during infection [137]. These variations in apoplast composition among species reflect distinct immune response strategies and provide clues about host specificity mechanisms in *M. oryzae* infections.

In *Arachis hypogaea* (peanut), elicitors such as chitosan, methyl jasmonate, and cyclodextrin triggered biphasic responses over time, with significant alterations in the root hair proteome [138]. Quantitative analyses revealed early induction of proteins associated with biotic stress responses, including heat shock proteins and wound-induced proteins, while late responses involved the regulation of antioxidant enzymes, auxin-responsive factors, and increased proteolytic activity. Notably, low-molecular-weight extracellular peptides were detected in the culture medium after 72 hours, suggesting the activation of proteolysis-based signaling cascades associated with stress perception and systemic response [138].

Specific plant–pathogen interactions: selected studies

Several specific interactions illustrate the sophistication of proteomic responses in different plant species facing biotic stress. In *S. lycopersicum*, the protein SICOP1-1, known for its role in photomorphogenesis, was shown to play a significant role in resistance to the fungus *Botrytis cinerea* [119]. Genetic manipulation of SICOP1-1 (either knockout or overexpression) significantly altered the fruit's resistance to infection without markedly affecting ripening. Proteomic analyses revealed that SICOP1-1 regulates the stability of the stress-related bZIP transcription factor SIOpaque2 via mono-ubiquitination, promoting its accumulation and contributing to pathogen resistance [119].

Infection by *Colletotrichum falcatum*, the causal agent of red rot, strongly modulated the proteome of *Saccharum spp.* (sugarcane) stalks in both resistant and susceptible cultivars. Fungal inoculation resulted in widespread repression of sucrose transporter proteins (SUT1,

5, 6, 7, and 8), affecting carbohydrate transport, while disease resistance proteins (DRP1–4) were induced particularly in the susceptible cultivar [139]. These findings suggest a coordinated reconfiguration between energy metabolism and defense, where suppression of sucrose transport may limit resource availability to the pathogen [139].

In *Carica papaya* (papaya), responses to the *papaya sticky disease complex* (PMeV-complex) were investigated throughout plant development [140]. Quantitative proteomics showed that young plants (3 to 4 months old) accumulate photosynthetic proteins and exhibit higher energy flow efficiency in photosystems I and II, while reducing activity of cell wall-degrading enzymes. In later stages (7 to 9 months), proteins related to cell wall biosynthesis were activated, although progressive loss of laticifer integrity was observed, leading to latex exudation [140]. These results suggest that infection tolerance is linked to photosynthetic and structural mechanisms present during juvenile stages.

Finally, infection by *Penicillium expansum*, the main agent of blue mold in stored apple (*Malus domestica*) fruits, triggered extensive proteomic reprogramming, with activation of MAPK signaling pathways, phenylpropanoid biosynthesis, and plant–pathogen interaction networks [120]. These responses included the accumulation of proteins related to oxidative stress, lignin synthesis, flavonoid production, and pathogenesis-related proteins (PR proteins), reinforcing structural and chemical barriers against the fungus. The study, which employed advanced techniques such as label-free quantification and parallel reaction monitoring (PRM), highlights the potential of proteomics to elucidate molecular defense mechanisms in fruits and support postharvest disease control strategies [120].

Molecular perspectives for resistance and manipulation

Molecular approaches have revealed promising targets for both engineering disease resistance in plants and controlling pathogen virulence. In *S. tuberosum*, the thylakoid protein Parakletos was recently characterized as a susceptibility factor [141]. Deletion of the gene encoding this protein resulted in enhanced resistance to oomycetes, fungi, and bacteria, as well as to abiotic stresses such as salinity and drought. Moreover, it led to increased productivity in field trials. Conversely, its overexpression suppressed ROS and calcium signaling triggered by biotic stimuli, thereby impairing the plant's defensive capacity. Homologs of this protein have been identified in various crops, reinforcing its potential as a biotechnological target in the breeding of resilient cultivars [141].

On the pathogen side, post-translational modulation has also emerged as a central strategy for successful infection. In *M. oryzae*, the mitochondrial desuccinylase enzyme Sirt5 was identified as essential for virulence [142]. Sirt5 regulates the desuccinylation of key proteins involved in oxidative phosphorylation, the Krebs cycle, and ROS detoxification, including thioredoxin (Trx2) and glutathione peroxidase (Hyr1). Its deletion led to the accumulation of PTMs, redox imbalance, and reduced invasive capacity, revealing this enzyme as a vulnerable point in fungal pathogenic physiology [142].

Beyond PTMs, endogenous plant regulatory compounds such as phytohormones also play active roles in directly suppressing pathogen virulence [143]. In *Phytophthora infestans*, salicylic acid (SA), a well-known regulator of plant immunity, was shown to directly inhibit the growth and pathogenicity of the oomycete. SA application significantly reduced sporulation, sporangial germination, and pathogen virulence. A multi-omics study integrating transcriptomics, proteomics, and phosphoproteomics revealed that SA induced autophagy while suppressing ribosome biogenesis and translation initiation in the pathogen, likely through the TOR signaling pathway. These findings open new perspectives for the use of resistance inducers that act directly on pathogen physiology [143].

In the realm of PTM regulation, mechanisms such as alternative splicing (AS) have also been exploited by pathogens as adaptive tools against host defenses[144] . In *Cladosporium fulvum*, the fungus responsible for leaf mold in tomato, an extensive AS pattern was observed throughout the infection cycle. Over 40% of protein-coding genes exhibited recurrent AS events, generating multiple protein isoforms in 59% of cases. The most affected genes were associated with nutrient transport, transcriptional regulation, and monooxygenases, conferring greater functional flexibility to the pathogen. Interestingly, these genes were located in repeat-rich genomic regions, suggesting an association between genome architecture and transcriptional plasticity [144].

Another promising avenue is the discovery of apoplastic effectors with immunogenic potential. In *Phytophthora sojae*, integrated proteomic and transcriptomic analyses identified PsCAP1, a secreted protein from the CAP superfamily, as an inducer of defense in solanaceous plants [145]. Its transient expression in *Nicotiana benthamiana* activated immune responses dependent on leucine-rich repeat (LRR)-type receptors (RCAP1) and the kinases BAK1 and SOBIR1. Pre-inoculation with PsCAP1 or its immunogenic fragment (amino acids 27–151) significantly enhanced resistance to *Phytophthora*, demonstrating the ability of such effectors to act as defense elicitors and potential tools for biocontrol strategies [145].

Lastly, studies on *Ganoderma boninense*, the causal agent of basal stem rot in *Elaeis guineensis* (oil palm), have underscored the critical role of the secretome in virulence. Using shotgun proteomics, 250 differentially expressed proteins were identified under nutrient-limited conditions, simulating the rhizospheric environment during infection [146]. In comparison with the less virulent species *G. tornatum*, *G. boninense* exhibited unique oxidoreductases and CAZymes with hydrolytic and oxidative activity involved in plant cell wall degradation. These findings suggest that species-specific differences in the secretome are directly linked to pathogenicity levels and pave the way for resistance strategies in tropical palms [146].

Recent discoveries reinforce that plant resistance and pathogen virulence are not isolated processes but rather the result of a complex network of molecular interactions. The identification of strategic targets in both hosts and pathogens including susceptibility proteins, PTM-modulating enzymes, transcriptional regulators, and immunogenic effectors broadens the repertoire of tools for integrated disease management. These molecular approaches open new possibilities for breeding resistant cultivars and suggest direct interventions on pathogen physiology as a promising frontier in plant biotechnology.

Case studies in proteomics of biotic stress

The application of proteomics to plant–pathogen interactions has enabled the identification of specific molecular targets that explain resistance mechanisms in various crop species. Cases involving fungi, bacteria, and postharvest interactions demonstrate the breadth and depth of the responses studied. Fungal infection in cultivated plants has been extensively investigated through proteomics, revealing complex mechanisms of defense, metabolic reprogramming, and post-translational regulation. These studies range from essential food crops to tropical and ornamental fruits, providing valuable insights for genetic improvement and disease management.

Fungal infections

Tropical fruit crops have significantly contributed to the understanding of proteomic responses to fungal attacks. Studies in crops such as *Mangifera indica* (mango) have revealed that infection by *Alternaria alternata* induces the expression of 472 proteins, with emphasis on

pathways such as phenylpropanoids, flavonoids, and MAPKs [116]. The accumulation of H_2O_2 , confirmed by histochemical staining, and the increase in enzymes such as SOD, peroxidases, glucanases, and chitinases highlight antioxidant responses and cell wall reinforcement [116]. Also, infection by *F. proliferatum* revealed that medium pH directly influences pathogen virulence by modulating proteins involved in cell wall degradation and redox processes [121, 147]. In papaya under infection by *F. equiseti*, proteins such as 9S-lipoxygenase and ribosomal components were associated with translational reprogramming and defense signaling [148]. In *Hylocereus polyrhizus* (dragon fruit), the presence of *Neoscytalidium dimidiatum* triggered activation of proteins related to photosynthesis, redox homeostasis, and hormonal signaling including ABP and ARP4 indicating a coordinated defense and metabolic regulation response [124].

In sugarcane, different fungal pathogens have revealed distinct infection strategies. In red rot disease caused by *Colletotrichum falcatum*, changes were observed in proteins related to sucrose transport and the activation of resistance proteins, suggesting antagonism between energy metabolism and plant defense mechanisms [139]. During the transition from sporidia to dikaryotic mycelium in *Sporisorium scitamineum*, a critical phase of infection, changes were identified in proteins associated with energy synthesis and putative effectors, with potential applications in molecular pathogen control [149].

In *Citrus limon*, infection by *Penicillium digitatum* induced a strong protein ubiquitination response, involving ATPases and membrane transporters, indicating broad metabolic reprogramming during the response to the pathogen [125]. In *Juglans regia* (walnut), infection by *C. gloeosporioides* led to the detection of 1929 differentially expressed proteins, mostly related to photosynthesis and carbon metabolism, with stronger responses in the early stages of infection [125].

In temperate fruits, postharvest resistance to fungal infection has also been extensively studied. In *Prunus persica*, the elicitor β -aminobutyric acid (BABA) activated a “primed” defense state through induction of transcription factors PpVOZ1, PpVOZ2, and interaction with PpTCP2, stimulating the salicylic acid pathway [150]. A combination of DIA-based proteomics, functional analysis, and molecular biology revealed the VOZ–TCP–ICS regulatory module as a key element of systemic acquired resistance (SAR) [150]. In *Prunus persica* (L.) Batsch (peach), infection by *Monilinia spp.* led to the accumulation of defense proteins such as HSP70, GST, and thaumatin in resistant cultivars, while susceptible genotypes showed a reduction in these markers [123].

In *Malus domestica* (apple), infection by *Penicillium expansum* resulted in the regulation of proteins involved in MAPK pathways, phenolic biosynthesis, and cell wall architecture [120]. Against *P. expansum*, 343 differentially expressed proteins were related to oxidative stress, metal transport, and oxidative phosphorylation [151]. Similarly, in *Actinidia deliciosa* 'Hayward' (kiwifruit), the response to *Botrytis cinerea* involved activation of cell wall proteins, MAPK cascade kinases, and components associated with ROS signaling [122]. In tomato, multiple studies have explored resistance mechanisms to *B. cinerea*. In ripe fruits, 186 differentially expressed proteins were detected during infection, including peroxidases, structural proteins, and defensins [152]. In this sense, response patterns varied with ripening stage, showing lower activation in green fruits [152, 153].

Another notable tomato study involved the overexpression of the BiP gene in transgenic plants infected with *Moniliophthora perniciosa*, a pathogen originally associated with witches' broom disease in *Theobroma cacao* [127]. This approach revealed the role of the endoplasmic reticulum in protein homeostasis during infection, with accumulation of PR proteins (PR2, PR3, PR10), increased antioxidant enzymes, and reduced H₂O₂ levels, indicating a protective effect against fungal stress [127]. Although most studies focus on edible fruits, ornamental species have also been targeted in proteomic investigations. In *Rosa sp.*, infection by *B. cinerea* revealed integration between phosphorylation and ubiquitination at over 10000 modified sites. PR10 proteins showed reduced ubiquitination, suggesting enhanced stability and defense reinforcement [154].

In vegetables and cereals, various defense mechanisms have been characterized with the help of proteomics. In *Cucurbita pepo* (zucchini), application of the plant elicitor *Reynoutria sachalinensis* activated induced defenses against *Podosphaera xanthii* (powdery mildew), involving pyruvate-related metabolic pathways and epigenetic repression via Polycomb complexes, suggesting an epigenetic mechanism of resistance induction [155]. In cereals, response mechanisms vary according to pathogen type and infection stage. In *Hordeum vulgare* (barley), infection by *Ramularia collo-cygni* triggered a hemibiotrophic response with distinct temporal profiles: xylanase inhibitors were induced in early phases, while senescence-related proteins appeared later, depending on plant and pathogen genotypes [156]. In wheat, the TaCRT3 protein was shown to be a positive regulator of resistance to *Blumeria graminis* (powdery mildew), acting via H₂O₂ accumulation and salicylic acid pathway activation [157].

In *Z. mays* infected with *Bipolaris maydis*, 41 proteins were differentially expressed between resistant and susceptible genotypes, with emphasis on translation networks, redox responses, and jasmonic acid signaling [13]. In ornamental grasses such as *Poa pratensis*,

resistance to powdery mildew was also associated with specific biochemical mechanisms. An integrative study combining iTRAQ-based proteomics and metabolomics revealed activation of the phenylpropanoid pathway, with emphasis on lignin biosynthesis and cell wall strengthening, forming a physical barrier against fungal penetration [158]. These findings provide valuable insights for marker-assisted breeding programs aimed at disease resistance in turfgrass species.

Among legumes, *Arachis hypogea* (peanut) has stood out in studies addressing various fungal diseases. In response to root rot caused by *F. oxysporum*, 3746 genes and 305 differentially expressed proteins were identified, with emphasis on flavonoid biosynthesis enzymes and genes associated with plant-pathogen interaction, such as PAL, CHI, and C4H [159]. In another study focused on black rot caused by *Cylindrocladium parasiticum*, contrasting cultivars revealed different response intensities: the resistant genotype showed stronger activation of hormonal pathways such as salicylic acid, jasmonic acid, and brassinosteroids as well as the MAPK cascade, suggesting that resistance success depends on the amplitude and intensity of defense signaling [160]. Similarly, in *Brassica napus*, the response to the necrotroph *Sclerotinia sclerotiorum* involved specific lipid pathways, notably the overexpression of ceramide kinase (CERK) and activation of lipases, reinforcing the role of lipids as signaling molecules in plant defense [161].

In *Cucumis sativus*, resistance to powdery mildew was investigated through targeted mutations in the MLO gene family. The Csmlo1/8/11 allele, when deleted, conferred complete resistance to infection [162]. Integrated transcriptomic and proteomic analyses revealed the involvement of a calcium-dependent signaling pathway involving CsKIC (kinesin-like calmodulin-binding protein-interacting Ca^{2+} -binding protein), CsCML28 (calmodulin-like protein 28), and CsCPK11 (Ca^{2+} -dependent protein kinase 11). Overexpression of CsCPK11 promoted resistance, while silencing increased susceptibility, confirming the positive role of these proteins in plant immunity. These findings highlight the importance of calcium signaling in MLO-mediated resistance in cucurbits and illustrate the potential of genetic editing combined with proteomics to elucidate defense mechanisms [162].

Finally, in *So. sisymbriifolium*, a wild species resistant to *Verticillium dahliae*, an iTRAQ study validated by PRM revealed differential expression of hundreds of proteins in roots at 12 and 24 hours after infection [163]. Enzymes such as PAL and SOD, PR proteins, phenylpropanoid pathways, and cell wall strengthening were activated. GO and KEGG analyses indicated early activation of metabolic and oxidative defense processes, demonstrating its potential as a genetic source for breeding cultivated solanaceous crops

[163]. These studies show that while defense strategies vary among species, tissues, and fungal types, the molecular mechanisms activated in response to fungal stress follow recurring patterns such as the activation of antioxidant, hormonal, and structural reinforcement pathways. The integration of proteomic and correlated omics approaches has broadened our understanding of these processes and opened new frontiers for the development of resistant cultivars, with practical applications in both tropical and temperate agriculture.

Oomycetes infections

Proteomic responses to oomycetes have revealed striking patterns of metabolic reprogramming and activation of antioxidant mechanisms in host plants. In *Hevea brasiliensis*, infection by *Phytophthora palmivora*, the causal agent of leaf fall disease, led to substantial changes in the foliar proteomic and metabolic profiles [164]. A total of 391 proteins were identified, with upregulation of components involved in oxidative phosphorylation, ATP synthesis, and amino acid metabolism particularly alanine, aspartate, and glutamate indicating a response focused on energy maintenance and primary metabolism adjustment as part of the plant's strategy to contain the pathogen [164].

Similarly, in *Capsicum annuum* infected by *P. capsici*, grafting with the resistant genotype 'ZCM334' proved effective in mitigating disease symptoms [165]. Proteomic analysis identified 827 differentially expressed proteins between leaves and roots, many of which were associated with antioxidant activity, cell wall stability, and responses to abiotic stress. Among them, 12 resistance-related proteins were validated by qRT-PCR, reinforcing the potential of grafting as a sustainable agronomic approach for oomycete control [165].

In *B. napus*, infection by *Plasmodiophora brassicae*, the etiological agent of clubroot disease, triggers profound and dynamic alterations in the root proteome and metabolome, especially during the early stages of infection [166]. Using nano LC-MS/MS, 2.515 differentially accumulated proteins were identified in resistant genotypes (CR) and 1.556 in susceptible ones (CS). The modulated proteins included enzymes involved in amino acid, fatty acid, lipid, and glucosinolate metabolism, as well as in cell wall reinforcement such as laccases, peroxidases, and invertase/pectin methylesterase inhibitors (PInv/PMEI) [166]. Integration with metabolomic data further revealed regulation of nucleotides, sphingolipid compounds, and intermediates of fatty acid metabolism. This study represents the first comprehensive characterization of proteomic and metabolic changes in *B. napus* roots in response to *P.*

brassicaceae infection, offering promising molecular targets for the development of cultivars with enhanced resistance [166].

Collectively, these studies demonstrate that the response to oomycetes involves profound reconfiguration of primary and secondary metabolism, with emphasis on energy, oxidative, and structural pathways. The integration of proteomic, transcriptomic, and metabolomic approaches has expanded our understanding of resistance physiology, providing valuable tools for the genetic improvement of agriculturally important species against these highly destructive pathogens.

Bacterial infections

Plant–bacteria interactions mediated by effectors, hormonal signaling, and systemic resistance have been extensively explored using high-resolution proteomic approaches. In tomato, the bacterium *Ralstonia solanacearum* secretes type III effectors such as RipU, RipD, and RipB, which are capable of interacting with proteins localized in the mitochondria, endoplasmic reticulum, and chloroplasts [167]. Using TurboID-mediated proximity labeling, researchers were able to precisely map the intracellular targets of these effectors, revealing distinct mechanisms for the suppression of both pattern-triggered immunity (PTI) and effector-triggered immunity (ETI). The study highlights the bacterium's ability to manipulate subcellular compartments to evade plant defenses [167].

On the other hand, symbiotic strategies also contribute to bacterial resistance. In *Lotus japonicus*, association with *Mesorhizobium loti* enhanced age-related resistance to *R. solanacearum* without compromising the nodulation process [168]. Proteomic analysis revealed localized activation of defense proteins in roots and nodules, illustrating the compatibility between immune responses and symbiotic maintenance a crucial balance for legume productivity [168]. In an example of a systemic disease with major agricultural impact, infection by *Candidatus Liberibacter asiaticus* in *Citrus sinensis*, the causal agent of huanglongbing (HLB), was studied in conjunction with its vector *Diaphorina citri*. Multi-omics approaches revealed time-dependent, coordinated changes in genes and proteins, identifying potential biomarkers for early diagnosis and reinforcing the role of applied proteomics in the detection and management of systemic plant diseases [125].

Viral infections

Viral infections in plants trigger highly coordinated proteomic responses, particularly involving changes in redox pathways, cellular stability, and reprogramming of vascular transport. In *Triticum aestivum*, the viral protein P1, encoded by *Wheat yellow mosaic virus* (WYMV), functioned as an efficient suppressor of plant immunity [132]. Overexpression of P1 led to a reduction in reactive oxygen species (ROS) production and repression of antioxidant genes such as TaPOD and TaTrxR, thereby facilitating infection. Silencing of these genes restored antiviral resistance, reversing the susceptible phenotype and demonstrating that manipulation of redox pathways is a key strategy in viral virulence [132]. In *N. tabacum*, infection by *Cucumber mosaic virus* (CMV) revealed contrasting proteomic profiles between resistant and susceptible cultivars. Resistant plants maintained mitochondrial and chloroplast integrity by activating mechanisms such as selective autophagy and the accumulation of antioxidant enzymes, contributing to infection containment. These findings underscore the role of cellular homeostasis and organelle recycling as important components of antiviral resistance [169].

In tomato, proteomic analysis of vascular exudates from plants infected with *Tomato spotted wilt orthotospovirus* (TSWV) identified 3305 proteins, of which 315 were differentially expressed [170]. Among these were RLK receptors, PR defense proteins, and viral proteins N, Gn, and Gc, suggesting that the xylem simultaneously functions as an active route for viral dissemination and as a molecular battleground between the plant and the pathogen [170]. In *C. papaya*, the response to the *papaya sticky disease viral complex* (PMeV) was characterized by early accumulation of proteins related to photosynthesis and cell wall synthesis in young plants, suggesting a form of tolerance linked to developmental stage. This differential response indicates that plant age may significantly influence viral infection dynamics and the effectiveness of activated proteomic responses [140].

Although not a viral infection, a relevant study on plant–bacteria interaction illustrates similar mechanisms of redox modulation. In *S. tuberosum*, the interaction with the actinobacterium *Streptomyces scabies*, the causal agent of common scab, led to a marked increase in bacterial proteins related to stress responses, secondary metabolism, and siderophore biosynthesis particularly pyochelin. This metabolic shift coincided with iron depletion in the culture medium and upregulation of pyochelin biosynthetic genes, suggesting active iron competition between pathogen and host. Additionally, increased nitrite excretion and hydrogen peroxide accumulation were observed, indicating the activation of redox defense and detoxification pathways in *S. scabies* in response to live plant tissue. These

findings highlight how host-derived signals can trigger bacterial virulence through multiple routes, including oxidative stress perception and manipulation of the ionic environment [171].

Interactions with insects and nematodes

Plant-insect, plant-mite, and plant-nematode interactions represent complex challenges to plant defense, often characterized by temporal responses, hormonal modulation, and metabolic reprogramming. In *O. sativa*, infestation by *Nilaparvata lugens* Stål (the brown planthopper) was studied in plants containing the resistance gene Bph3. Proteomic analysis using TMT labeling indicated modulation of proteins involved in hormonal pathways, calcium signaling, and phenylpropanoid biosynthesis, highlighting the multifunctionality of Bph3 in activating defense responses against herbivory [172]. Beyond feeding, *N. lugens* oviposition also significantly influences plant immunity. Egg-associated secretions (EAS), deposited in damaged rice tissues during egg-laying, rapidly induced the production of jasmonic acid (JA), JA-Ile, and hydrogen peroxide (H₂O₂). These signals activated defensive responses not only in rice but also in other species such as tobacco, leading to reduced performance of aphids and whiteflies [173]. Proteomic analysis of the EAS identified 110 proteins, of which 53 had two or more unique peptides. Some are shared with the salivary proteome of herbivorous insects, while others are specific to oviposition and may represent novel molecular targets for inducing plant resistance [173].

Aphids also play a central role in insect–plant interactions through the secretion of salivary proteins with effector functions. In the case of the rose grain aphid, 705 potential secreted proteins were identified, with 15 validated by proteomics [174]. The high similarity with proteins from other aphid species suggests conserved functions in manipulating plant physiology, primarily through suppression of immune responses [174]. On the other hand, in *Glycine max* (soybean), infestation by the mite *Tetranychus urticae* triggered profound temporal changes in the plant proteome over 21 days [175]. More than 450 differentially regulated proteins were identified, affecting processes such as photosynthesis, amino acid metabolism, energy production, and antioxidant responses. Metabolic reprogramming was dynamic and phase-dependent, reflecting soybean's continuous adaptation to prolonged herbivore attack.

In plant–nematode interactions, oil palm (*Elaeis guineensis*) infected by *Bursaphelenchus cocophilus*, the causal agent of red ring disease, exhibited differential

expression of 1278 host proteins and 397 pathogen transcripts [176]. Multi-omics data revealed pathways involved in the production of compounds with nematicidal properties, as well as the identification of biomarkers with potential applications in early diagnosis and genetic improvement [176]. Similarly, in *Pinus koraiensis* trees, wilt caused by the nematode *Bursaphelenchus xylophilus* activated metabolic pathways associated with the biosynthesis of terpenoids, flavonoids, and toxic alkaloids, in addition to inducing proteins involved in pathogen recognition, such as receptor-like kinases, pectin modulators, and proline-rich proteins. These responses, activated even before symptom onset, illustrate an efficient model of systemic defense in conifers [177].

Mutualistic interactions and sustainable strategies

Beyond classical pathogenic interactions, recent proteomic studies have highlighted the potential of mutualistic relationships and sustainable agronomic practices as effective strategies to induce resistance to biotic stresses. In *O. sativa*, colonization by *Acrocalymma vagum*, a mutualistic endophytic fungus, promoted increased productivity, activation of antioxidant enzymes such as OsPRX70 and OsPRX95, and remodeling of the rhizospheric microbiota [178]. Remarkably, manipulation of the fungus's proteomic repertoire was sufficient to alter its lifestyle, converting it from mutualist to pathogen demonstrating molecular plasticity as a determinant of plant–microbe interaction outcomes [178]. Likewise, the symbiosis between *L. japonicus* and *Mesorhizobium loti* showed it is possible to integrate biological nitrogen fixation with age-related resistance to *R. solanacearum*. Proteomic analysis revealed localized modifications in roots and nodules, with activation of mechanisms compatible with both defense and maintenance of symbiosis, illustrating a highly advantageous functional balance for legume crops [168].

In the field of agroecological management, intercropping strategies such as the consortium between *Vicia faba* and wheat have also shown significant protective effects. This association reduced the incidence of *Fusarium* wilt and modulated the expression of 288 proteins related to the antioxidant system [179]. Additionally, classical defense genes such as VfPR1, PR2, and PR5 were activated, indicating that cropping interactions can act as natural inducers of systemic resistance [179]. Also, fungal elicitation using yeast extract in *Sorbus aucuparia* triggered an intense oxidative response, production of biphenyl-type phytoalexins, and activation of defense genes such as SaPAL, SaLOX2, SaPR1, and SaCHIB1-3 [180]. These genes are associated with central pathways such as phenylpropanoids and MAPK,

revealing the potential of exogenous stimuli in reinforcing plants' biochemical barriers against diverse stressors [180].

These examples demonstrate how mutualism-based approaches, elicitation, and sustainable agricultural practices can be strategically combined to promote pathogen resistance, reducing dependence on chemical inputs. In this context, proteomics emerges as a key tool for understanding and rationally applying these complex interactions in the agriculture of the future.

Heterosis

Heterosis, or hybrid vigor, is a genetic phenomenon in which the F_1 generation exhibits superior performance compared to its parental lines, influencing quantitative traits such as growth rate, fertility, yield, and biomass, as well as qualitative traits like disease resistance and tolerance to environmental stresses [181, 182]. However, the development of hybrids expressing heterosis requires numerous field trials, which are highly costly, time-consuming, labor-intensive, and often yield unpredictable results [183]. These challenges underscore the need to elucidate the molecular mechanisms underlying heterosis in order to enhance the efficiency and predictability of plant breeding programs.

Genetic explanations for heterosis are based on the effects of dominance, overdominance, and epistasis [184]. According to the dominance hypothesis, hybrids benefit from the complementation of deleterious recessive alleles, whose expression in homozygosity would compromise plant performance [185]. Overdominance, in turn, suggests that certain heterozygous combinations at specific loci provide a functional advantage superior to any homozygous combination [186]. Additionally, epistasis refers to the interaction between genes located at different loci, where the expression of one gene can modify or mask the expression of another. Depending on the nature of this interaction, the effects can be additive, dominant, or superdominant [187]. Over the years, considerable efforts have been made to understand the genetic and molecular basis of heterosis [188]. Although the underlying molecular mechanisms of this phenomenon remain not fully elucidated, advances in genomic and proteomic approaches have significantly contributed to deepening our understanding of this process [189].

Proteomic insights into the molecular basis of heterosis during seed germination

Germination plays a fundamental role in the plant life cycle, marking the onset of post-embryonic development through radicle protrusion. This process directly influences seedling vigor, drought tolerance, time to flowering, and ultimately crop productivity [190]. Protein expression patterns revealed the contribution of dominance, partial dominance, and overdominance effects to the differential accumulation of proteins during germination [191]. Mass spectrometry-based identification of 54 proteins followed by functional analysis showed enrichment of pathways related to primary metabolism, cellular detoxification, uncharacterized proteins, chaperones, signal transduction, developmental processes, transport, and stress responses. Notably, proteins such as 14-3-3, ABI3, ABI5, and PKABA1 were differentially accumulated in hybrids, indicating that hormonal signaling pathways mediated by abscisic acid (ABA) and gibberellins (GA) are strongly associated with the manifestation of heterosis during maize germination [191].

In a similar manner, the comparative proteomic analysis conducted on hybrid rice embryos of the LYP9 variety (Liang-You-Pei-Jiu) and its parental lines, 9311 and P64S, led to the identification of 54 differentially accumulated protein spots. The most represented functional categories included proteins related to nutrient reservoir activity and stress response, suggesting an enhanced metabolic and physiological preparedness in the hybrid [192]. Additionally, proteins involved in core metabolic pathways, such as UDP-glucose pyrophosphorylase, α -amylase, fructose-bisphosphate aldolase, and enolase, were significantly more abundant in LYP9, exhibiting an overdominant expression pattern [192]. This pattern indicates a more efficient activation of metabolic pathways associated with reserve mobilization and the resumption of growth, supporting a potential functional role of these enzymes in promoting hybrid vigor during the germination stage.

Proteomic insights into the molecular mechanisms of heterosis during root development

Enhanced protein synthesis capacity in hybrids may be a key requirement for the expression of hybrid vigor during the early stages of root development. To investigate the molecular mechanisms associated with heterosis in roots, a label-free proteomic analysis of seminal roots from maize seedlings was conducted in order to identify early manifestations of hybrid vigor during root development [193]. The study revealed the non-additive accumulation of 85 proteins in at least one of the hybrids evaluated. Among these, the most represented

functional category was “protein metabolism,” with 27 differentially accumulated proteins. Notably, of the 17 ribosomal proteins identified in the seminal roots, 16 exhibited dominant or overdominant expression patterns [193].

Similarly, a comparative proteomic analysis between the wheat hybrid 3338/2463 and its parental lines was performed to elucidate the molecular mechanisms underlying heterosis focusing on the root proteomic profile [194]. By employing two-dimensional gel electrophoresis (2-DE) coupled with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS), the authors identified 45 differentially accumulated protein spots exhibiting both qualitative and quantitative variation [194]. Among these, 25 proteins were successfully identified and categorized into functional groups related to signal transduction, primary metabolism, energy production, cell growth and division, defense response, and secondary metabolism [194]. These findings provide compelling evidence that the enhanced root vigor observed in the hybrid may be driven by the coordinated regulation of multiple biological pathways, highlighting the complex proteomic landscape associated with heterosis in wheat.

Proteomic insights into the effects of heterosis on leaf metabolism and photosynthesis

Leaves play essential roles in various physiological processes, including photosynthesis, respiration, transpiration, and guttation [195]. Final leaf size is determined by two major cellular events: cell proliferation and expansion [196]. However, a key observation in maize heterosis studies is that increased leaf area in hybrids is primarily associated with a higher number of cells, rather than increased cell size [197]. Given the pivotal role of cell proliferation in defining final leaf size and in the expression of heterosis, a comparative proteomic analysis of the nuclear fraction from the basal region of the third leaf in the maize hybrid Mo17/B73 and its parental lines, were conducted using 2DE and MALDI-TOF-MS approaches [196]. At five days after germination, the hybrid displayed significantly greater leaf length, width, and area compared to its parental lines, which was primarily attributed to an increased number of cells in the mature leaf region, with no significant differences in cell size [196]. Proteomic analysis identified 52 differentially accumulated proteins in the hybrid, 38 of which were successfully characterized. These proteins were mainly associated with RNA and protein regulation, as well as functions related to cell division, growth, and developmental mechanisms, suggesting that the differential accumulation of nuclear proteins may be directly linked to the heterotic traits observed in the hybrid's leaf architecture [196].

The investigation of protein expression patterns of the maize inbred lines Mo17 and B73, and their F1 hybrid SC704, grown under hydroponic conditions in a greenhouse, revealed a total of 114 proteins with significant changes in expression patterns [198]. Proteomic analysis was performed using two-dimensional electrophoresis (2-DE) on leaf tissue collected from the third leaf 30 days after the beginning of the experiment. The differentially expressed proteins associated with heterosis included: Gibberellin 20-oxidase-like protein, Thiamine pyrophosphokinase 1, Probable starch synthase 4, chloroplastic/amyloplastic, Basic leucine zipper 24, Cathepsin B-like protease 1, Dehydrin HIRD11, and ORGAN SIZE RELATED 1. These proteins were primarily classified into three functional categories: defense system, regulatory proteins, and energy metabolism [198]. The predominance of proteins involved in defense suggests that this functional group plays a central role in the manifestation of hybrid vigor observed in SC704. Among the components of energy metabolism, Thiamine pyrophosphokinase 1 stands out, exhibiting the highest relative expression level and a heterosis value of 5%, indicating its potential contribution to the molecular mechanisms underlying heterosis in maize [198].

A study conducted with the rice superhybrid LYP9 and its parental lines (9311 and PA64s) revealed that heterosis is associated with the differential regulation of proteins mainly involved in photosynthesis, glycolysis, and defense mechanisms. Proteomic analysis identified 384 differentially expressed proteins, of which 297 were characterized, corresponding to 222 unique proteins. The differentially expressed proteins between the hybrid and its parents (DPHP) showed greater functional enrichment in the aforementioned categories, especially during the flowering and grain-filling stages [199]. At these stages, a significant increase was observed in the abundance of most up-regulated proteins in the hybrid, including phosphoglycerate kinase (PGK), glyceraldehyde-3-phosphate dehydrogenase, triosephosphate isomerase, transketolase, sedoheptulose-1,7-bisphosphatase, phosphoribulokinase, ADP-glucose pyrophosphorylase, peroxiredoxin (2-Prx), and glutathione S-transferase. This increase was accompanied by elevated CO₂ assimilation rates and higher apparent quantum yield of photosynthesis [199]. These findings indicate that the dynamic regulation of these metabolic pathways throughout development plays a key role in the manifestation of hybrid vigor in rice.

One of the main challenges in understanding the molecular basis of heterosis is its variability among different hybrids of the same species. Through a proteomic analysis conducted on the second leaf of seedlings from two maize hybrids (ZD808 and ZD909) and their respective parental lines, it was possible to identify proteins associated with heterosis

[183]. In the hybrid ZD808, most of the differentially accumulated proteins (DAPs) showed increased levels compared to the parents, while non-additively accumulated proteins (NAPs) displayed expression patterns equal to or higher than those of the parent with greater abundance, being mainly enriched in stress response pathways [183]. In contrast, in the ZD909 hybrid, most DAPs exhibited reduced expression relative to the parental lines, and NAPs tended to show levels equal to or lower than those of the parent with the lowest expression. However, unlike ZD808, a portion of the NAPs in ZD909 were enriched in photosynthesis-related pathways, including light reactions, the Calvin cycle, and photorespiration [183]. These findings suggest that the greater vigor observed in ZD909 may be related to enhanced photosynthetic capacity, highlighting that different molecular mechanisms can underlie heterosis in distinct hybrids, while in ZD808 it may be associated with improved adaptive capacity.

Proteomic Insights into Energy Metabolism and Stress Response Mechanisms Associated with Heterosis

Total proteome and mitochondrial proteome of ear shoots between three maize hybrids that shared the female parent B73 but exhibited different levels of heterosis: B73 × N192 (low heterosis) and B73 × Mo17 and B73 × NC350 (high heterosis) [200]. In the total proteome, six proteins were differentially accumulated in the two high-heterosis hybrids compared to the low-heterosis hybrid, including translation elongation factor 2 (TEF2) subunit, t-complex protein 1 subunit c (TCP1c), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), r40c1-like protein, an unknown protein, and glutathione S-transferase III B (GSTIIIB) [200].

In the mitochondrial proteome, extracted from young inflorescences, 14 proteins showed consistent changes in the high-heterosis hybrids, including isoforms of succinate dehydrogenase flavoprotein (SDH1), aldehyde dehydrogenase 2 (ALDH2), dihydrolipoamide dehydrogenase (DLDH), formate dehydrogenase 1 (FDH1), subunits of respiratory chain complex I, a lipase domain-containing hypothetical protein, glycine-rich RNA-binding protein 7 (GRBP7), jasmonate-induced protein (JIP), and isoforms of uncharacterized proteins [200]. Additionally, the proteomic analysis between hybrids revealed several differentially expressed proteins as distinct isoforms on 2D gels, indicating post-translational modifications that alter the isoelectric point. The geLC-MS analysis showed that, in most cases, there were no changes in the total protein abundance, except for three proteins: one containing a lipase domain, one uncharacterized protein, and JIP [200]. These findings suggest that heterosis in

maize is associated with more efficient energy metabolism and enhanced stress response mechanisms, as components of these pathways showed differential accumulation patterns in high-heterosis hybrids compared to the low-heterosis hybrid.

Future prospects

The ongoing development of proteomics technologies promises to significantly advance our understanding of complex biological systems, particularly in plant research. One key area of growth is the refinement of quantitative methods, both labeled and label-free, which are crucial for accurate protein quantification. While label-free techniques offer cost-effective solutions for large sample sets, further improvements in data normalization and reducing bias will enhance their reliability. Label-based methods, although more precise in terms of absolute quantification, face challenges such as high reagent costs and limited scalability, suggesting a need for more affordable and efficient approaches in the future. Additionally, the growing interest in single-cell proteomics represents an exciting frontier. Advances in mass spectrometry have made it possible to explore protein dynamics at the single-cell level, offering unprecedented insight into cellular heterogeneity. However, challenges in data consistency, sensitivity, and reproducibility still need to be addressed for broader application.

The integration of subcellular proteomics is another promising avenue. By mapping protein localization within specific organelles, researchers can gain deeper insights into cellular function, signaling pathways, and stress responses. This subfield is poised to grow with the development of more refined isolation techniques and high-throughput methods, expanding the potential for discoveries in cell compartmentalization and protein function. Overall, as the technologies and methodologies continue to evolve, proteomics will play a pivotal role in unlocking the complexities of plant biology and beyond.

References

1. Holtorf, H., M.-C. Guitton, and R. Reski, Plant functional genomics. *Naturwissenschaften*, 2002. 89(6): p. 235-249.
2. Yan, S., *et al.*, Recent advances in proteomics and metabolomics in plants. *Molecular Horticulture*, 2022. 2(1): p. 17.

3. Mergner, J. and B. Kuster, Plant proteome dynamics. *Annual Review of Plant Biology*, 2022. 73(Volume 73, 2022): p. 67-92.
4. Aydođan, C., Critical review of new advances in food and plant proteomics analyses by nano-LC/MS towards advanced foodomics. *TrAC Trends in Analytical Chemistry*, 2024. 176: p. 117759.
5. Varnavides, G., *et al.*, In search of a universal method: A comparative survey of bottom-up proteomics sample preparation methods. *Journal of Proteome Research*, 2022. 21(10): p. 2397-2411.
6. Mikulášek, K., *et al.*, SP3 protocol for proteomic plant sample preparation prior lc-ms/ms. *Frontiers in Plant Science*, 2021. Volume 12 - 2021.
7. Tyanova, S., T. Temu, and J. Cox, The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols*, 2016. 11(12): p. 2301-2319.
8. Szklarczyk, D., *et al.*, The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*, 2022. 51(D1): p. D638-D646.
9. Zheng, Y., *et al.*, iTAK: A program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Molecular Plant*, 2016. 9(12): p. 1667-1670.
10. Picotti, P., Phosphoproteomics takes it easy. *Nature Biotechnology*, 2015. 33(9): p. 929-930.
11. Balotf, S., *et al.*, Shotgun proteomics as a powerful tool for the study of the proteomes of plants, their pathogens, and plant–pathogen interactions. *Proteomes*, 2022. 10(1): p. 5.
12. Taylor, C.F., *et al.*, The minimum information about a proteomics experiment (MIAPE). *Nature Biotechnology*, 2007. 25(8): p. 887-893.
13. Corrêa, C.C.G., *et al.*, Comparative proteomic analysis of popcorn genotypes identifies differentially accumulated proteins associated with resistance pathways to southern leaf blight disease. *Plants*, 2025. 14(3): p. 426.
14. Nikonorova, N., *et al.*, The Arabidopsis root tip (phospho)proteomes at growth-promoting versus growth-repressing conditions reveal novel root growth regulators. *Cells*, 2021. 10(7): p. 1665.
15. Xavier, L.R., *et al.*, Time-dependent proteomic signatures associated with embryogenic callus induction in *Carica papaya* L. *Plants*, 2023. 12(22): p. 3891.
16. Isaacson, T., *et al.*, Sample extraction techniques for enhanced proteomic analysis of plant tissues. *Nature Protocols*, 2006. 1(2): p. 769-774.
17. Aguilar-Hernández, V. and V.M. Loyola-Vargas, Advanced proteomic approaches to elucidate somatic embryogenesis. *Frontiers in Plant Science*, 2018. 9.
18. Heringer, A.S., C. Santa-Catarina, and V. Silveira, Insights from proteomic studies into plant somatic embryogenesis. *Proteomics*, 2018. 18(5-6): p. e1700265.

19. Almeida, F.A., *et al.*, Label-free quantitative phosphoproteomics reveals signaling dynamics involved in embryogenic competence acquisition in sugarcane. *Journal of Proteome Research*, 2020. 19(10): p. 4145-4157.
20. Ni, M.-w., *et al.*, Modified filter-aided sample preparation (FASP) method increases peptide and protein identifications for shotgun proteomics. *Rapid Communications in Mass Spectrometry*, 2017. 31(2): p. 171-178.
21. Olson, B.J.S.C., Assays for determination of protein concentration. *Current Protocols in Pharmacology*, 2016. 73(1): p. A.3A.1-A.3A.32.
22. Giansanti, P., *et al.*, Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nature Protocols*, 2016. 11(5): p. 993-1006.
23. Hu, A., W. Noble, and A. Wolf-Yadlin, Technical advances in proteomics: new developments in data-independent acquisition. *F1000Research*, 2016. 5(419).
24. Guan, S., *et al.*, Data Dependent-Independent Acquisition (DDIA) proteomics. *Journal of Proteome Research*, 2020. 19(8): p. 3230-3237.
25. Zhang, F., *et al.*, Data-Independent Acquisition mass spectrometry-based proteomics and software tools: A glimpse in 2020. *PROTEOMICS*, 2020. 20(17-18): p. 1900276.
26. Schubert, O.T., *et al.*, Quantitative proteomics: challenges and opportunities in basic and applied research. *Nature Protocols*, 2017. 12(7): p. 1289-1294.
27. Chen, C., *et al.*, Bioinformatics methods for mass spectrometry-based proteomics data analysis. *International Journal of Molecular Sciences*, 2020. 21(8): p. 2873.
28. Deutsch, E.W., *et al.*, Trans-proteomic pipeline, a standardized data processing pipeline for large-scale reproducible proteomics informatics. *PROTEOMICS – Clinical Applications*, 2015. 9(7-8): p. 745-754.
29. Lin, M.-H., *et al.*, Benchmarking differential expression, imputation and quantification methods for proteomics data. *Briefings in Bioinformatics*, 2022. 23(3).
30. Schessner, J.P., E. Voytik, and I. Bludau, A practical guide to interpreting and generating bottom-up proteomics data visualizations. *PROTEOMICS*, 2022. 22(8): p. 2100103.
31. Oveland, E., *et al.*, Viewing the proteome: How to visualize proteomics data? *PROTEOMICS*, 2015. 15(8): p. 1341-1355.
32. Hsiao, Y., *et al.*, Analysis and visualization of quantitative proteomics data using fragpipe-analyst. *Journal of Proteome Research*, 2024. 23(10): p. 4303-4315.
33. Kanehisa, M., *et al.*, KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research*, 2016. 45(D1): p. D353-D361.
34. Kolberg, L., *et al.*, g:Profiler, interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Research*, 2023. 51(W1): p. W207-W212.

35. Zhou, Y., *et al.*, Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications*, 2019. 10(1): p. 1523.
36. Schwacke, R., *et al.*, MapMan4: A refined protein classification and annotation framework applicable to multi-omics data analysis. *Molecular Plant*, 2019. 12(6): p. 879-892.
37. Conesa, A. and S. Götz, Blast2GO: A comprehensive suite for functional analysis in plant genomics. *International Journal of Plant Genomics*, 2008. 2008(1): p. 619832.
38. Van de Graaf, M.W., *et al.*, Benchmarking of protein interaction databases for integration with manually reconstructed signalling network models. *The Journal of Physiology*, 2024. 602(18): p. 4529-4542.
39. Shannon, P., *et al.*, Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 2003. 13(11): p. 2498-504.
40. Aguilan, J.T., K. Kulej, and S. Sidoli, Guide for protein fold change and p-value calculation for non-experts in proteomics. *Molecular Omics*, 2020. 16(6): p. 573-582.
41. Tyanova, S. and J. Cox, Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research, in *Cancer Systems Biology: Methods and Protocols*, L. von Stechow, Editor. 2018, Springer New York: New York, NY. p. 133-148.
42. Tyanova, S., *et al.*, The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods*, 2016. 13(9): p. 731-740.
43. Love, M.I., W. Huber, and S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 2014. 15(12): p. 550.
44. Guo, H., *et al.*, Large-scale label-free phosphoproteomics: From technology to data interpretation. *Bioanalysis*, 2014. 6(18): p. 2403-2420.
45. Low, T.Y., *et al.*, Widening the bottleneck of phosphoproteomics: Evolving strategies for phosphopeptide enrichment. *Mass Spectrometry Reviews*, 2021. 40(4): p. 309-333.
46. Clark, N.M., J.M. Elmore, and J.W. Walley, To the proteome and beyond: advances in single-cell omics profiling for plant systems. *Plant Physiology*, 2021. 188(2): p. 726-737.
47. Bennett, H.M., *et al.*, Single-cell proteomics enabled by next-generation sequencing or mass spectrometry. *Nature Methods*, 2023. 20(3): p. 363-374.
48. Ahmad, R. and B. Budnik, A review of the current state of single-cell proteomics and future perspective. *Analytical and Bioanalytical Chemistry*, 2023. 415(28): p. 6889-6899.
49. Gatto, L., *et al.*, Initial recommendations for performing, benchmarking and reporting single-cell proteomics experiments. *Nature Methods*, 2023. 20(3): p. 375-386.
50. Huffman, R.G., *et al.*, Prioritized mass spectrometry increases the depth, sensitivity and data completeness of single-cell proteomics. *Nature Methods*, 2023. 20(5): p. 714-722.
51. Christopher, J.A., *et al.*, Subcellular proteomics. *Nature Reviews Methods Primers*, 2021. 1(1): p. 32.

52. Chen, Y., *et al.*, Exploring the diversity of plant proteome. *Journal of Integrative Plant Biology*, 2021. 63(7): p. 1197-1210.
53. Breckels, L.M., *et al.*, Advances in spatial proteomics: Mapping proteome architecture from protein complexes to subcellular localizations. *Cell Chemical Biology*, 2024. 31(9): p. 1665-1687.
54. Zhang, Y., *et al.*, Plants' response to abiotic stress: Mechanisms and strategies. *International Journal of Molecular Sciences*, 2023. 24(13): p. 10915.
55. Holleman, C., *et al.*, The impact of climate variability and extremes on agriculture and food security – An analysis of the evidence and case studies. *FAO Agricultural Development Economics Technical Study*, 2020. 1(4).
56. Wang, W., *et al.*, Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends in Plant Science*, 2004. 9(5): p. 244-252.
57. Zahedi, S.M., *et al.*, Plant adaptation to drought stress: The role of anatomical and morphological characteristics in maintaining the water status. *Journal of Soil Science and Plant Nutrition*, 2025. 25(1): p. 409-427.
58. Fu, Z., *et al.*, Global critical soil moisture thresholds of plant water stress. *Nature Communications*, 2024. 15(1): p. 4826.
59. Seleiman, M.F., *et al.*, Drought stress impacts on plants and different approaches to alleviate its adverse effects. *Plants*, 2021. 10(2): p. 259.
60. Lyon, D., *et al.*, Drought and recovery: Independently regulated processes highlighting the importance of protein turnover dynamics and translational regulation in *medicago truncatula**. *Molecular & Cellular Proteomics*, 2016. 15(6): p. 1921-1937.
61. Prinsi, B., *et al.*, Root proteomic and metabolic analyses reveal specific responses to drought stress in differently tolerant grapevine rootstocks. *BMC Plant Biology*, 2018. 18(1): p. 126.
62. Schachtman, D.P. and J.Q.D. Goodger, Chemical root to shoot signaling under drought. *Trends in Plant Science*, 2008. 13(6): p. 281-287.
63. Aroca, R., R. Porcel, and J.M. Ruiz-Lozano, Regulation of root water uptake under abiotic stress conditions. *Journal of Experimental Botany*, 2011. 63(1): p. 43-57.
64. Schmitt, K.F.M., *et al.*, Decoding the effects of drought stress on popcorn (*Zea mays* var. *evarta*) flowering combining proteomics and physiological analysis. *Plant Physiology and Biochemistry*, 2024. 208: p. 108444.
65. Gao, H., *et al.*, Proteomic analysis of leaves and roots during drought stress and recovery in *Setaria italica* L. *Frontiers in Plant Science*, 2023. Volume 14 - 2023.
66. Zhou, Y., *et al.*, Proteomic investigation of molecular mechanisms in response to PEG-induced drought stress in soybean roots. *Plants*, 2022. 11(9): p. 1173.
67. Ren, W., *et al.*, iTRAQ-based quantitative proteomic analysis provides insight into the drought-stress response in maize seedlings. *Scientific Reports*, 2022. 12(1): p. 9520.

68. Liu, M., *et al.*, Candidate regulators of drought stress in tomato revealed by comparative transcriptomic and proteomic analyses. *Frontiers in Plant Science*, 2023. Volume 14 - 2023.
69. Salvato, F., *et al.*, Label-Free quantitative proteomics of enriched nuclei from sugarcane (*Saccharum* ssp) stems in response to drought stress. *PROTEOMICS*, 2019. 19(14): p. 1900004.
70. Li, P., *et al.*, Translocation of drought-responsive proteins from the chloroplasts. *Cells*, 2020. 9(1): p. 259.
71. Manghwar, H., *et al.*, Waterlogging stress in plants: Unraveling the mechanisms and impacts on growth, development, and productivity. *Environmental and Experimental Botany*, 2024. 224: p. 105824.
72. Yang, X., *et al.*, iTRAQ-based quantitative proteomics unveils protein dynamics in the root of *Solanum melongena* L. Under waterlogging stress conditions. *Life*, 2023. 13(6): p. 1399.
73. Komatsu, S., *et al.*, Proteomic and biochemical analyses of the mechanism of tolerance in mutant soybean responding to flooding stress. *International Journal of Molecular Sciences*, 2021. 22(16): p. 9046.
74. Wang, X., *et al.*, Comparative proteomic analysis of grapevine rootstock in response to waterlogging stress. *Frontiers in Plant Science*, 2021. Volume 12 - 2021.
75. Yang, R., *et al.*, iTRAQ proteomic analysis of wheat (*Triticum aestivum* L.) genotypes differing in waterlogging tolerance. *Frontiers in Plant Science*, 2022. Volume 13 - 2022.
76. Liu, D., *et al.*, Quantitative proteomics and relative enzymatic activities reveal different mechanisms in two peanut cultivars (*arachis hypogaea* L.) under waterlogging conditions. *Frontiers in Plant Science*, 2021. Volume 12 - 2021.
77. Zhao, Y., *et al.*, Fine control of growth and thermotolerance in the plant response to heat stress. *Journal of Integrative Agriculture*, 2025. 24(2): p. 409-428.
78. Distéfano, A.M., *et al.*, Heat stress in plants: sensing, signalling, and ferroptosis. *Journal of Experimental Botany*, 2024. 76(5): p. 1357-1369.
79. Zhou, R., *et al.*, Increase crop resilience to heat stress using omic strategies. *Frontiers in Plant Science*, 2022. Volume 13 - 2022.
80. Kausar, R., X. Wang, and S. Komatsu, Crop proteomics under abiotic stress: From data to insights. *Plants*, 2022. 11(21): p. 2877.
81. Ma, Z., *et al.*, Phosphoproteomic analysis of maize seedlings provides insights into the mechanisms of heat-stress tolerance. *International Journal of Molecular Sciences*, 2025. 26(6): p. 2439.
82. Boguszewska-Mańkowska, D., M. Gietler, and M. Nykiel, Comparative proteomic analysis of drought and high temperature response in roots of two potato cultivars. *Plant Growth Regulation*, 2020. 92(2): p. 345-363.

83. Song, X., *et al.*, Proteomic and physiological analyses reveal the mechanisms through which melatonin ameliorates heat stress-induced photoinhibition in *Nicotiana tabacum*. *Plant Physiology and Biochemistry*, 2025. 221: p. 109665.
84. Perveen, A., *et al.*, Integrative physiological, biochemical, and proteomic analysis of the leaves of two cotton genotypes under heat stress. *PLOS ONE*, 2025. 20(1): p. e0316630.
85. Zhao, Y., *et al.*, Stimulus-responsive proteins involved in multi-process regulation of storage substance accumulation during rice grain filling under elevated temperature. *BMC Plant Biology*, 2023. 23(1): p. 547.
86. Chinnusamy, V., J. Zhu, and J.-K. Zhu, Cold stress regulation of gene expression in plants. *Trends in Plant Science*, 2007. 12(10): p. 444-451.
87. Sharma, P., N. Sharma, and R. Deswal, The molecular biology of the low-temperature response in plants. *BioEssays*, 2005. 27(10): p. 1048-1059.
88. Zhang, H., *et al.*, Physiological and proteomic analysis of various priming on rice seed under chilling stress. *Plants*, 2024. 13(17): p. 2430.
89. Qing, D., *et al.*, ITRAQ-based quantitative proteomic analysis of japonica rice seedling during cold stress. *Breeding Science*, 2022. 72(2): p. 150-168.
90. Bukhari, B., *et al.*, Proteomic, transcriptomic, biochemical and physiological investigations shed light on responses to low temperature stress in two contrasting soybean varieties. *Plant, Cell & Environment*. n/a(n/a).
91. Wang, W., *et al.*, Label-free quantitative proteomics reveal the mechanisms of young wheat (*Triticum aestivum* L.) ears' response to spring freezing. *International Journal of Molecular Sciences*, 2023. 24(21): p. 15892.
92. Zhang, B.-Q., *et al.*, Cold-induced physiological and biochemical alternations and proteomic insight into the response of *Saccharum spontaneum* to low temperature. *International Journal of Molecular Sciences*, 2022. 23(22): p. 14244.
93. Kamran, M., *et al.*, An overview of hazardous impacts of soil salinity in crops, tolerance mechanisms, and amelioration through selenium supplementation. *International Journal of Molecular Sciences*, 2020. 21(1): p. 148.
94. Stavi, I., N. Thevs, and S. Priori, Soil salinity and sodicity in drylands: A review of causes, effects, monitoring, and restoration measures. *Frontiers in Environmental Science*, 2021. Volume 9 - 2021.
95. Tarolli, P., *et al.*, Soil salinization in agriculture: Mitigation and adaptation strategies combining nature-based solutions and bioengineering. *iScience*, 2024. 27(2): p. 108830.
96. Wijeweera, S., O. Duncan, and A.H. Millar, Spatial and development responses in the wheat leaf highlight the loss of chloroplast protein homeostasis during salt stress. *Journal of Proteomics*, 2025. 316: p. 105438.
97. Chen, X., *et al.*, Label-free proteomics reveals the response of oat (*Avena sativa* L.) seedling root respiratory metabolism to salt stress. *International Journal of Molecular Sciences*, 2025. 26(6): p. 2630.

98. Li, X., *et al.*, Evaluation of salt-tolerant germplasms and identification of salt tolerance-related proteins in upland cotton at the seedling stage. *International Journal of Molecular Sciences*, 2025. 26(5): p. 1982.
99. Fortini, E.A., *et al.*, Physiological, epigenetic, and proteomic responses in *Pfaffia glomerata* growth in vitro under salt stress and 5-azacytidine. *Protoplasma*, 2023. 260(2): p. 467-482.
100. Passamani, L.Z., *et al.*, Salt stress induces changes in the proteomic profile of micropropagated sugarcane shoots. *PLOS ONE*, 2017. 12(4): p. e0176076.
101. Barba-Espín, G., *et al.*, Halophyte-based crop managements induce biochemical, metabolomic and proteomic changes in tomato plants under saline conditions. *Physiologia Plantarum*, 2025. 177(1): p. e70060.
102. Fu, S., *et al.*, Integrated transcriptomic, proteomic, and metabolomic analyses revealed molecular mechanism for salt resistance in soybean (*Glycine max* L.) seedlings. *International Journal of Molecular Sciences*, 2024. 25(24): p. 13559.
103. Hang, J., *et al.*, Comparative transcriptomic and proteomic analyses of two salt-tolerant alfalfa (*Medicago sativa* L.) genotypes: investigation of the mechanisms underlying tolerance to salt. *Frontiers in Plant Science*, 2024. Volume 15 - 2024.
104. Gao, Y., *et al.*, Comparative proteomic discovery of salt stress response in alfalfa roots and overexpression of MsANN2 confers salt tolerance. *Plant Physiology and Biochemistry*, 2024. 215: p. 109033.
105. Peralta-Videa, J.R., *et al.*, The biochemistry of environmental heavy metal uptake by plants: Implications for the food chain. *The International Journal of Biochemistry & Cell Biology*, 2009. 41(8): p. 1665-1677.
106. Hossain, Z. and S. Komatsu, Contribution of proteomic studies towards understanding plant heavy metal stress response. *Frontiers in Plant Science*, 2013. Volume 3 - 2012.
107. Ahsan, N., J. Renaut, and S. Komatsu, Recent developments in the application of proteomics to the analysis of plant responses to heavy metals. *PROTEOMICS*, 2009. 9(10): p. 2602-2621.
108. Pinto, V.B., *et al.*, Deciphering the major metabolic pathways associated with aluminum tolerance in popcorn roots using label-free quantitative proteomics. *Planta*, 2021. 254(6): p. 132.
109. Terzi, H. and M. Yıldız, Proteomic analysis reveals the role of exogenous cysteine in alleviating chromium stress in maize seedlings. *Ecotoxicology and Environmental Safety*, 2021. 209: p. 111784.
110. Li, L., *et al.*, Conservative mechanism through various rapeseed (*Brassica napus* L.) varieties respond to heavy metal (Cadmium, Lead, Arsenic) stress. *Frontiers in Plant Science*, 2025. Volume 15 - 2024.
111. Mitra, P.K., *et al.*, Proteome insights of citric acid-mediated cadmium toxicity tolerance in *Brassica napus* L. *Environmental Science and Pollution Research*, 2023. 30(54): p. 115461-115479.

112. Lu, J., *et al.*, Combined transcriptome and proteome analysis revealed the molecular regulation mechanisms of zinc homeostasis and antioxidant machinery in tobacco in response to different zinc supplies. *Plant Physiology and Biochemistry*, 2023. 202: p. 107919.
113. Meng, W., *et al.*, Analysis of differentially expressed proteins and related metabolic pathways in response to lead stress in the leaves of *Pogonatherum crinitum*. *Ecotoxicology and Environmental Safety*, 2024. 288: p. 117438.
114. Liu, Y., R.N. Damaris, and P. Yang, Proteomics analysis identified a DRT protein involved in arsenic resistance in *Populus*. *Plant Cell Reports*, 2017. 36(12): p. 1855-1869.
115. Liu, Y., *et al.*, Proteomics: a powerful tool to study plant responses to biotic stress. *Plant Methods*, 2019. 15(1): p. 135.
116. Xie, X., *et al.*, Proteomics analysis revealed the activation and suppression of different host defense components challenged with mango leaf spot pathogen *Alternaria alternata*. *BMC Plant Biology*, 2025. 25(1): p. 227.
117. Keerthana, U., *et al.*, Comparative proteomic analysis of resistant and susceptible aromatic rice landraces in response to blast pathogen, *Magnaporthe oryzae*. *Physiological and Molecular Plant Pathology*, 2025. 137: p. 102629.
118. Zhang, B., *et al.*, RBB1 negatively regulates rice disease resistance by modulating protein glycosylation. *Journal of Integrative Plant Biology*, 2025. 67(2): p. 391-407.
119. Gao, G., *et al.*, E3 ligase SICOP1-1 stabilizes transcription factor SIOpaque2 and enhances fruit resistance to *Botrytis cinerea* in tomato. *Plant Physiology*, 2024. 196(2): p. 1196-1213.
120. Wang, K., *et al.*, The proteome of *Penicillium expansum* during infection of postharvest apple is revealed using Label-Free and Parallel Reaction Monitoring (PRM) Techniques. *Journal of Proteomics*, 2024. 298: p. 105142.
121. Xie, L., *et al.*, Proteomic and physiological analysis provides an elucidation of *Fusarium proliferatum* infection causing crown rot on banana fruit. *Microbiological Research*, 2022. 256: p. 126952.
122. Liu, J., *et al.*, Proteomic analysis of kiwifruit in response to the postharvest pathogen, *Botrytis cinerea*. *Frontiers in Plant Science*, 2018. Volume 9 - 2018.
123. Papavasileiou, A., *et al.*, Proteomic analysis upon peach fruit infection with *Monilinia fructicola* and *M. laxa* identify responses contributing to brown rot resistance. *Scientific Reports*, 2020. 10(1): p. 7807.
124. Pan, L., *et al.*, Comparative proteomic analysis of the fungal pathogen *Neoscytalidium dimidiatum* infection in the pitaya. *Horticulture, Environment, and Biotechnology*, 2021. 62(4): p. 649-659.
125. Ramsey, J.S., *et al.*, Longitudinal transcriptomic, proteomic, and metabolomic analysis of citrus limon response to graft inoculation by *Candidatus liberibacter asiaticus*. *Journal of Proteome Research*, 2020. 19(6): p. 2247-2263.

126. Fang, H., *et al.*, Transcriptome and proteome analysis of walnut (*Juglans regia* L.) fruit in response to infection by *Colletotrichum gloeosporioides*. *BMC Plant Biology*, 2021. 21(1): p. 249.
127. Alcântara, G.d.M., *et al.*, Proteomic analysis of plants with binding immunoglobulin protein overexpression reveals mechanisms related to defense against *Moniliophthora perniciosa*. *Plants*, 2025. 14(4): p. 503.
128. Qu, D., *et al.*, A 4d proteome investigation of the potential mechanisms of sa in triggering resistance in kiwifruit to *Pseudomonas syringae* pv. *Actinidiae*. *International Journal of Molecular Sciences*, 2023. 24(24): p. 17448.
129. Karimi, H.Z., *et al.*, Proteasomes accumulate in the plant apoplast where they participate in microbe-associated molecular pattern (MAMP)-triggered pathogen defense. *Nature Communications*, 2025. 16(1): p. 1634.
130. Lee, G.H., *et al.*, Dataset on post-translational modifications proteome analysis of MSP1-overexpressing rice leaf proteins. *Data in Brief*, 2023. 50: p. 109573.
131. Yang, J., *et al.*, In-depth site-specific glycoproteomic analysis reveals ER-resident protein PDI regulating wheat yellow mosaic virus infection. *International Journal of Biological Macromolecules*, 2025. 293: p. 139303.
132. Zhao, Y., *et al.*, Wheat yellow mosaic virus P1 INHIBITS ROS accumulation to facilitate viral infection. *International Journal of Molecular Sciences*, 2025. 26(4): p. 1455.
133. Zhang, N., *et al.*, Integration of multi-omics data accelerates molecular analysis of common wheat traits. *Nature Communications*, 2025. 16(1): p. 2200.
134. Ding, Y., *et al.*, Proteomic analysis of lysine acetylation and succinylation to investigate the pathogenicity of virulent *Pseudomonas syringae* pv. tomato DC3000 and avirulent line *Pseudomonas syringae* pv. tomato DC3000 avrRpm1 on *Arabidopsis thaliana*. *Genes*, 2024. 15(4): p. 499.
135. Cai, J., *et al.*, A NAC triad modulates plant immunity by negatively regulating N-hydroxy pipecolic acid biosynthesis. *Nature Communications*, 2024. 15(1): p. 7212.
136. Karmakar, S., *et al.*, A detailed proteomics and metabolomics landscape sheds light on the mechanistic insights into the resistance response of transgenic pigeon pea against wilt stress. *Plant, Cell & Environment*. n/a(n/a).
137. Wang, J., *et al.*, Proteomic identification of apoplastic proteins from rice, wheat, and barley after *Magnaporthe oryzae* infection. *Phytopathology Research*, 2024. 6(1): p. 55.
138. Pankaew, C., *et al.*, Proteomic profiling of peanut hairy root culture unveils distinctive adaptive responses induced by elicitor treatment across diverse time intervals. *Plant Stress*, 2024. 12: p. 100497.
139. Gujjar, R.S., *et al.*, *Colletotrichum falcatum* infection influences the abundance of sucrose transporters and disease resistant proteins in sugarcane stalk. *Journal of Plant Biochemistry and Biotechnology*, 2024. 33(3): p. 367-373.
140. Rodrigues, S.P., *et al.*, Juvenile-related tolerance to papaya sticky disease (PSD): proteomic, ultrastructural, and physiological events. *Plant Cell Reports*, 2024. 43(11): p. 269.

141. Zahid, M.A., *et al.*, Enhanced stress resilience in potato by deletion of Parakletos. *Nature Communications*, 2024. 15(1): p. 5224.
142. Ren, Z., *et al.*, Sirt5-mediated lysine desuccinylation regulates oxidative stress adaptation in *Magnaporthe oryzae* during host intracellular infection. *New Phytologist*, 2024. 242(3): p. 1257-1274.
143. Zhang, S., *et al.*, Anti-oomycete effect and mechanism of salicylic acid on *Phytophthora infestans*. *Journal of Agricultural and Food Chemistry*, 2023. 71(51): p. 20613-20624.
144. Zaccaron, A.Z., L.-H. Chen, and I. Stergiopoulos, Transcriptome analysis of two isolates of the tomato pathogen *Cladosporium fulvum*, uncovers genome-wide patterns of alternative splicing during a host infection cycle. *PLOS Pathogens*, 2024. 20(12): p. e1012791.
145. Jiang, H., *et al.*, The CAP superfamily protein PsCAP1 secreted by *Phytophthora* triggers immune responses in *Nicotiana benthamiana* through a leucine-rich repeat receptor-like protein. *New Phytologist*, 2023. 240(2): p. 784-801.
146. Dzulkaflı, S.B., *et al.*, Understanding the effects of carbon and nitrogen starvation on the comparative secretomes of *Ganoderma boninense* and *Ganoderma tornatum*. *Physiological and Molecular Plant Pathology*, 2023. 127: p. 102084.
147. Li, L., *et al.*, Proteomic studies on fruit ripening and senescence. *Critical Reviews in Plant Sciences*, 2017. 36(2): p. 116-127.
148. Karunamoorthy, J., *et al.*, Proteomic analysis of the Papaya-Fusarium equiseti interaction: Understanding mode of infection and plant response at the molecular level. *Physiological and Molecular Plant Pathology*, 2025. 136: p. 102583.
149. Kumaravel, N., *et al.*, Comparative proteomics of sugarcane smut fungus - *Sporisorium scitamineum* unravels dynamic proteomic alterations during the dimorphic transition. *Journal of Proteomics*, 2024. 304: p. 105230.
150. Wang, K., *et al.*, VOZ-dependent priming of salicylic acid-dependent defense against *Rhizopus stolonifer* by β -aminobutyric acid requires the TCP protein TCP2 in peach fruit. *The Plant Journal*, 2025. 121(1): p. e17176.
151. Xu, M., *et al.*, Proteomic analysis of apple response to *Penicillium expansum* infection based on label-free and parallel reaction monitoring techniques. *Journal of Fungi*, 2022. 8(12): p. 1273.
152. Shah, P., *et al.*, Proteomic analysis of ripening tomato fruit infected by *Botrytis cinerea*. *Journal of Proteome Research*, 2012. 11(4): p. 2178-2192.
153. Blanco-Ulate, B., *et al.*, Genome-wide transcriptional profiling of *Botrytis cinerea* genes targeting plant cell walls during infections of different hosts. *Frontiers in Plant Science*, 2014. Volume 5 - 2014.
154. Li, R., *et al.*, Integrated proteomic analysis reveals interactions between phosphorylation and ubiquitination in rose response to *Botrytis* infection. *Horticulture Research*, 2023. 11(1).

155. Theologidis, I., *et al.*, Proteome profiling of Cucurbita pepo phyllosphere after infection by *Podosphaera xanthii* and application of *Reynoutria sachalinensis* extract. *Applied Sciences*, 2024. 14(21): p. 10061.
156. Lemcke, R., *et al.*, Integrative transcript to proteome analysis of barley during *Ramularia collo-cygni* leaf spot development identified several proteins that are related to fungal recognition and infection responses. *Frontiers in Plant Science*, 2024. Volume 15 - 2024.
157. Ren, J., *et al.*, TaCRT3 is a positive regulator of resistance to *Blumeria graminis* f. sp. *tritici* in wheat. *Phytopathology*®, 2024. 114(3): p. 641-652.
158. Dong, W., *et al.*, Comparative proteomic and metabolomic analysis of resistant and susceptible Kentucky Bluegrass cultivars in response to infection by powdery mildew. *BMC Plant Biology*, 2024. 24(1): p. 1195.
159. Wang, M., *et al.*, Transcriptomic–proteomic analysis revealed the regulatory mechanism of peanut in response to *Fusarium oxysporum*. *International Journal of Molecular Sciences*, 2024. 25(1): p. 619.
160. Chen, X., *et al.*, Comparative proteomic analysis of the moderately resistant and susceptible peanut cultivars during infestation by *Cylindrocladium parasiticum*. *Canadian Journal of Plant Pathology*, 2024. 46(1): p. 50-64.
161. Ouyang, Z., *et al.*, Ceramide-1-phosphate enhances defense responses against *Sclerotinia sclerotiorum* in *Brassica napus*. *Plant Physiology*, 2024. 197(2).
162. Ma, M., *et al.*, Multiplex gene editing reveals cucumber MILDEW RESISTANCE LOCUS O family roles in powdery mildew resistance. *Plant Physiology*, 2024. 195(2): p. 1069-1088.
163. Wu, L., *et al.*, Comparative proteomic analysis of roots from a wild eggplant species *solanum sisymbriifolium* in defense response to *Verticillium dahliae* inoculation. *Genes*, 2023. 14(6): p. 1247.
164. Kruaweangmol, P., *et al.*, Metabolomic and proteomic changes in leaves of rubber seedlings infected by *Phytophthora palmivora*. *Tree Physiology*, 2025. 45(2).
165. Shi, F., *et al.*, Unveiling molecular mechanisms of pepper resistance to *Phytophthora capsici* through grafting using iTRAQ-based proteomic analysis. *Scientific Reports*, 2024. 14(1): p. 4789.
166. Adhikary, D., *et al.*, Proteome- and metabolome-level changes during early stages of clubroot infection in *Brassica napus* canola. *Molecular Omics*, 2024. 20(4): p. 265-282.
167. De Ryck, J., *et al.*, Exploring the tomato root protein network exploited by Core Type 3 Effectors from the *Ralstonia solanacearum* species complex. *Journal of Proteome Research*, 2025. 24(2): p. 696-709.
168. Prévitali, T., *et al.*, Lotus resistance against *Ralstonia* is enhanced by *Mesorhizobium* and does not impair mutualism. *New Phytologist*, 2025. 245(3): p. 1249-1262.

169. Shen, Z., *et al.*, Integrated transmission electron microscopy and proteomic analyses reveal the cytoarchitectural response to cucumber mosaic virus infection in tobacco. *Int J Biol Macromol*, 2024. 262(Pt 2): p. 130100.
170. Feng, H., *et al.*, Integrated biological experiments and proteomic analyses of *Nicotiana tabacum* xylem sap revealed the host response to tomato spotted wilt orthotospovirus infection. *International Journal of Molecular Sciences*, 2024. 25(20): p. 10907.
171. Giroux, L., *et al.*, Proteomics fingerprinting reveals importance of iron and oxidative stress in *Streptomyces scabies*–*Solanum tuberosum* interactions. *Frontiers in Microbiology*, 2024. Volume 15 - 2024.
172. Qing, D., *et al.*, TMT-based quantitative proteomics analysis of defense responses induced by the Bph3 gene following brown planthopper infection in rice. *BMC Plant Biology*, 2024. 24(1): p. 1092.
173. Li, J., *et al.*, Egg-associated secretions from the brown planthopper (*Nilaparvata lugens*) activate rice immune responses. *Insect Science*, 2024. 31(4): p. 1135-1149.
174. Gebrekidan, A.G., Y. Zhang, and J. Chen, A comprehensive transcriptomic and proteomics analysis of candidate secretory proteins in rose grain aphid, *Metopolophium dirhodum* (Walker). *Current Issues in Molecular Biology*, 2024. 46(12): p. 13383-13404.
175. Ruffatto, K., *et al.*, Unravelling soybean responses to early and late *Tetranychus urticae* (Acari: Tetranychidae) infestation. *Plant Biology*, 2024. 26(7): p. 1223-1239.
176. de Oliveira Braga, Í., *et al.*, A multi-omics deconstruction of red ring disease in oil palm. *Journal of Plant Pathology*, 2024.
177. Yu, L., *et al.*, Transcriptomic, metabonomic and proteomic analyses reveal that terpenoids and flavonoids are required for *Pinus koraiensis* early defence against *Bursaphelenchus xylophilus* infection. *BMC Plant Biology*, 2025. 25(1): p. 185.
178. Zeng, Y., *et al.*, Endophyte *Acrocalymma vagum* establishes the holobiont with rice to attract beneficial microorganisms and promote disease resistance. *Journal of Advanced Research*, 2025.
179. Zheng, Y., *et al.*, Intercropping and appropriate nitrogen application control faba bean *Fusarium* wilt by improving physiological and biochemical resistance and protein expression of faba bean. *Plant Pathology*, 2024. 73(3): p. 613-629.
180. Li, Y., *et al.*, Multi-omics revealed molecular mechanism of biphenyl phytoalexin formation in response to yeast extract-induced oxidative stress in *Sorbus aucuparia* suspension cells. *Plant Cell Reports*, 2024. 43(3): p. 62.
181. Lippman, Z.B. and D. Zamir, Heterosis: revisiting the magic. *Trends in Genetics*, 2007. 23(2): p. 60-66.
182. Mohayeji, M., *et al.*, Heterosis profile of sunflower leaves: A label free proteomics approach. *Journal of Proteomics*, 2014. 99: p. 101-110.
183. Wang, D., *et al.*, Comparative proteomic analysis reveals that the Heterosis of two maize hybrids is related to enhancement of stress response and photosynthesis respectively. *BMC Plant Biology*, 2021. 21(1): p. 34.

184. Zhang, H.-Y., *et al.*, A genome-wide transcription analysis reveals a close correlation of promoter indel polymorphism and heterotic gene expression in rice hybrids. *Molecular Plant*, 2008. 1(5): p. 720-731.
185. Baldauf, J.A., *et al.*, Single-parent expression complementation contributes to phenotypic heterosis in maize hybrids. *Plant Physiology*, 2022. 189(3): p. 1625-1638.
186. Shull, G.H., The composition of a field of maize. *Journal of Heredity*, 1908. 4(1): p. 296-301.
187. Powers, L., Relative yields of inbred lines and f1 hybrids of tomato. *Botanical Gazette*, 1945. 106(3): p. 247-268.
188. Liu, J., *et al.*, Exploring the molecular basis of heterosis for plant breeding. *Journal of Integrative Plant Biology*, 2020. 62(3): p. 287-298.
189. Xing, J., Q. Sun, and Z. Ni, Proteomic patterns associated with heterosis. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 2016. 1864(8): p. 908-915.
190. Harris, D., *et al.*, On-farm seed priming: using participatory methods to revive and refine a key technology. *Agricultural Systems*, 2001. 69(1): p. 151-164.
191. Fu, Z., *et al.*, Proteomic analysis of heterosis during maize seed germination. *PROTEOMICS*, 2011. 11(8): p. 1462-1472.
192. Wang, W., *et al.*, Proteomic profiling of rice embryos from a hybrid rice cultivar and its parental lines. *PROTEOMICS*, 2008. 8(22): p. 4808-4821.
193. Marcon, C., *et al.*, Heterosis-associated proteome analyses of maize (*Zea mays* L.) seminal roots by quantitative label-free LC-MS. *Journal of Proteomics*, 2013. 93: p. 295-302.
194. Song, X., *et al.*, Wheat (L.) root proteome and differentially expressed root proteins between hybrid and parents. *PROTEOMICS*, 2007. 7(19): p. 3538-3557.
195. Swanson-Wagner, R.A., *et al.*, All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents. *Proceedings of the National Academy of Sciences*, 2006. 103(18): p. 6805-6810.
196. Guo, B., *et al.*, Maize (*Zea mays* L.) seedling leaf nuclear proteome and differentially expressed proteins between a hybrid and its parental lines. *PROTEOMICS*, 2014. 14(9): p. 1071-1087.
197. UCHIMIYA, H. and N. TAKAHASHI, Kinetics of heterosis in growth of the leaf blade in *Zea mays* L. *Annals of Botany*, 1973. 37(2): p. 247-252.
198. Bandehagh, A., *et al.*, Assessment of heterosis proteins in maize (*Zea mays* L.) leaves by two-dimensional gel electrophoresis. *Plant Gene*, 2021. 28: p. 100331.
199. Zhang, C., *et al.*, Comparative proteomic study reveals dynamic proteome changes between superhybrid rice LYP9 and its parents at different developmental stages. *J Plant Physiol*, 2012. 169(4): p. 387-98.

200. Dahal, D., B.P. Mooney, and K.J. Newton, Specific changes in total and mitochondrial proteomes are associated with higher levels of heterosis in maize hybrids. *The Plant Journal*, 2012. 72(1): p. 70-83.

Acknowledgements

This research was supported by funds from the Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) (Proc E-26/211.310/2021; Proc E-26/204.185/2024) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (312258/2023-2), awarded to V.S. This study was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil (CAPES), Finance Code 001. The authors would like to thank FAPERJ, CNPq and CAPES for providing scholarships to graduate students.

Author contributions

LRX, and VS contributed to the writing, review, and editing of the manuscript. RPdP, and CS-C reviewed, edited, and translated the manuscript. CEA, LEO and DDRP contributed to writing the manuscript.

Corresponding authors

Correspondence to Vanildo Silveira

Ethics declarations

The authors declare that they have no competing interests.

Ethics approval and consent to participate

This article does not contain any studies with human or animal participants. Work with plants complies with relevant institutional, national and international guidelines and legislation.

Consent for publication

All the authors provide their consent for this publication.

**Capítulo 2. Time-dependent proteomic signatures associated with
embryogenic callus induction in *Carica papaya* L.**



plants



Article

Time-Dependent Proteomic Signatures Associated with Embryogenic Callus Induction in *Carica papaya* L.

Lucas Rodrigues Xavier, Caio Cezar Guedes Corrêa, Roberta Pena da Paschoa, Karina da Silva Vieira, Daniel Dastan Rezabala Pacheco, Lucas do Espírito Santo Gomes, Bárbara Cardoso Duncan, Laís dos Santos da Conceição, Vitor Batista Pinto, Claudete Santa-Catarina and Vanildo Silveira.

Special Issue

Advances in Somatic Embryogenesis Plant Cell Differentiation

Edited by

Dr. Víctor M. Loyola-Vargas and Prof. Dr. Neftalí Ochoa-Alejo



Time-Dependent Proteomic Signatures Associated with Embryogenic Callus Induction in *Carica papaya* L.

Lucas Rodrigues Xavier ^{1,2}, Caio Cezar Guedes Corrêa ^{1,2}, Roberta Pena da Paschoa ^{1,2}, Karina da Silva Vieira ^{1,2}, Daniel Dastan Rezabala Pacheco ^{1,2}, Lucas do Espírito Santo Gomes ^{1,2}, Bárbara Cardoso Duncan, Laís dos Santos da Conceição ^{1,2}, Vitor Batista Pinto ³, Claudete Santa-Catarina ³ and Vanildo Silveira ^{1,2,*}

¹Laboratório de Biotecnologia, Centro de Biociências e Biotecnologia (CBB), Universidade Estadual do Norte

Fluminense Darcy Ribeiro (UENF), Campos dos Goytacazes 28013-602, RJ, Brazil;

²Unidade de Biologia Integrativa, Setor de Genômica e Proteômica, UENF

³Laboratório de Biologia Celular e Tecidual, Centro de Biociências e Biotecnologia (CBB), UENF

*Correspondence: vanildo@uenf.br

Abstract

Sex segregation increases the cost of *Carica papaya* production through seed-based propagation. Therefore, in vitro techniques are an attractive option for clonal propagation, especially of hermaphroditic plants. Here, we performed a temporal analysis of the proteome of *C. papaya* calli aiming to identify the key players involved in embryogenic callus formation. Mature zygotic embryos used as explants were treated with 20 μ M 2,4-dichlorophenoxyacetic acid to induce embryogenic callus. Total proteins were extracted from explants at 0 (zygotic embryo) and after 7, 14, and 21 days of induction. A total of 1407 proteins were identified using a bottom-up proteomic approach. The clustering analysis revealed four distinct patterns of protein accumulation throughout callus induction. Proteins related to seed maturation and storage are abundant in the explant before induction, decreasing as callus formation progresses. Carbohydrate and amino acid metabolisms, aerobic respiration, and protein catabolic processes were enriched throughout days of callus induction. Protein kinases associated with auxin responses, such as SKP1-like proteins 1B, accumulated in response to callus induction. Additionally, regulatory proteins, including histone deacetylase (HD2C) and argonaute 1 (AGO1), were more abundant at 7 days, suggesting their role in the acquisition of embryogenic competence. Predicted protein–protein networks revealed the regulatory role of proteins 14-3-3 accumulated during callus induction and the association of proteins involved in oxidative phosphorylation and hormone response. Our findings emphasize the modulation of the proteome during embryogenic callus initiation and identify regulatory proteins that might be involved in the activation of this process.

Keywords: micropropagation; bottom-up proteomics; *Carica papaya*; embryogenic callus; somatic embryos; time-series analysis

Introduction

Understanding the molecular processes that control the embryogenic switch in plant somatic cells is of interest to current plant biotechnology. In this context, the costs of *Carica papaya* production through seed-based propagation are increased by sex segregation, and in vitro tissue culture are emerging as an attractive option for clonal propagation of hermaphroditic papaya plants. In this context, it is already known that using molecular markers to select sex-determined papaya plants can increase orchard yields by up to 49% when compared to the conventional method of sex determination by flowering [1]. Thus, somatic embryogenesis (SE) in papaya can be applied to produce clones on a large-scale, reducing production costs due to the sexual segregation presented by seedlings [2]. The SE in *C. papaya* already has a well-established and characterized protocol able to generate globular, heart, torpedo, and cotyledonary embryo stages, similar to those of zygotic embryogenesis [3]. Moreover, SE in papaya also serves as an efficient experimental model for genetic, physiological, and morphological studies of the formation and development of somatic embryos in plants [4,5].

SE is a morphogenetic process in which plant somatic cells undergo transcriptomic reprogramming in response to an induction signal, primarily following auxin or stress treatments, and then start to form embryos asexually. This reprogramming prompts the cells to enter the developmental embryogenic pathway, leading to the formation of structures resembling embryos, known as somatic embryos [6]. The formation of somatic embryos can occur directly in the explant or indirectly with an intermediary callus phase [7]. Only a few of the molecular events that govern embryogenic callus induction in somatic plant cells have begun to be decoded, and with the use of global transcriptome and proteomic analysis, many genes and proteins have been shown to be involved in the SE that is induced in different plant species [8,9]. On the other hand, studies on indirect SE usually use already differentiated callus, leaving gaps in the regulatory process that directs the initiation of embryogenic callus formation.

The molecular pathways involved in the embryogenic response during in vitro morphogenesis are of particular interest in the field of plant development biology. In this way, SE studies contribute to the understanding of the regulatory mechanisms and control of totipotency and pluripotency in plant somatic cells [6]. Comparative proteomics from embryogenic calli is an interesting approach allowing the identification of proteins with relevant roles during the SE process in papaya. In this context, a proteomics study using LED lamps with different wavelengths of light during the maturation of papaya somatic embryos revealed

522 proteins, where the optimization of embryo development was associated with the accumulation of energy metabolism proteins [4]. Another study identified 801 proteins at different stages of somatic embryo development, showing specific proteomic profiles for each stage [10]. Furthermore, the study of the mitochondrial proteome of callus with contrasting embryogenic competence led to the identification of 237 proteins and contributed to the discovery of proteins important for energy production and auxin homeostasis [11]. On the other hand, these studies used already differentiated callus for proteomic analyses, where the process of induction and formation of embryogenic callus is still poorly understood. In this sense, proteomics is considered a promising technique to improve our knowledge of the physiological and molecular factors that affect the plant SE process [9].

Our hypothesis is that specific protein sets function at different stages of embryogenic callus induction in order to coordinate this process. To address this question, we performed a time-course analysis of the proteomic responses during the onset of embryogenic callus formation in papaya. Our findings highlighted that induction treatment triggers hormonal responses and epigenetic modulation, leading to the accumulation of regulatory proteins for the initiation of embryogenic callus. These data can provide insights into alternatives to SE induction in papaya.

Results

Embryogenic Callus Induction

Upon removal of the seed tegument, the cotyledonary zygotic embryo used as explant exhibited a white color (T0; Figure 1A). The cotyledons bent after 7 days of inoculation (T7) in the induction treatment, while callus formation was observed 14 days after inoculation (T14; Figure 1B,C). After 21 days (T21), the embryogenic callus had already formed up to the entire explant (Figure 1D). During callus induction, the formation of a mass of yellowish and brownish cells was observed. In addition, callus formation is also characterized by an increase in cell mass and growth of the explant in the induction culture medium.

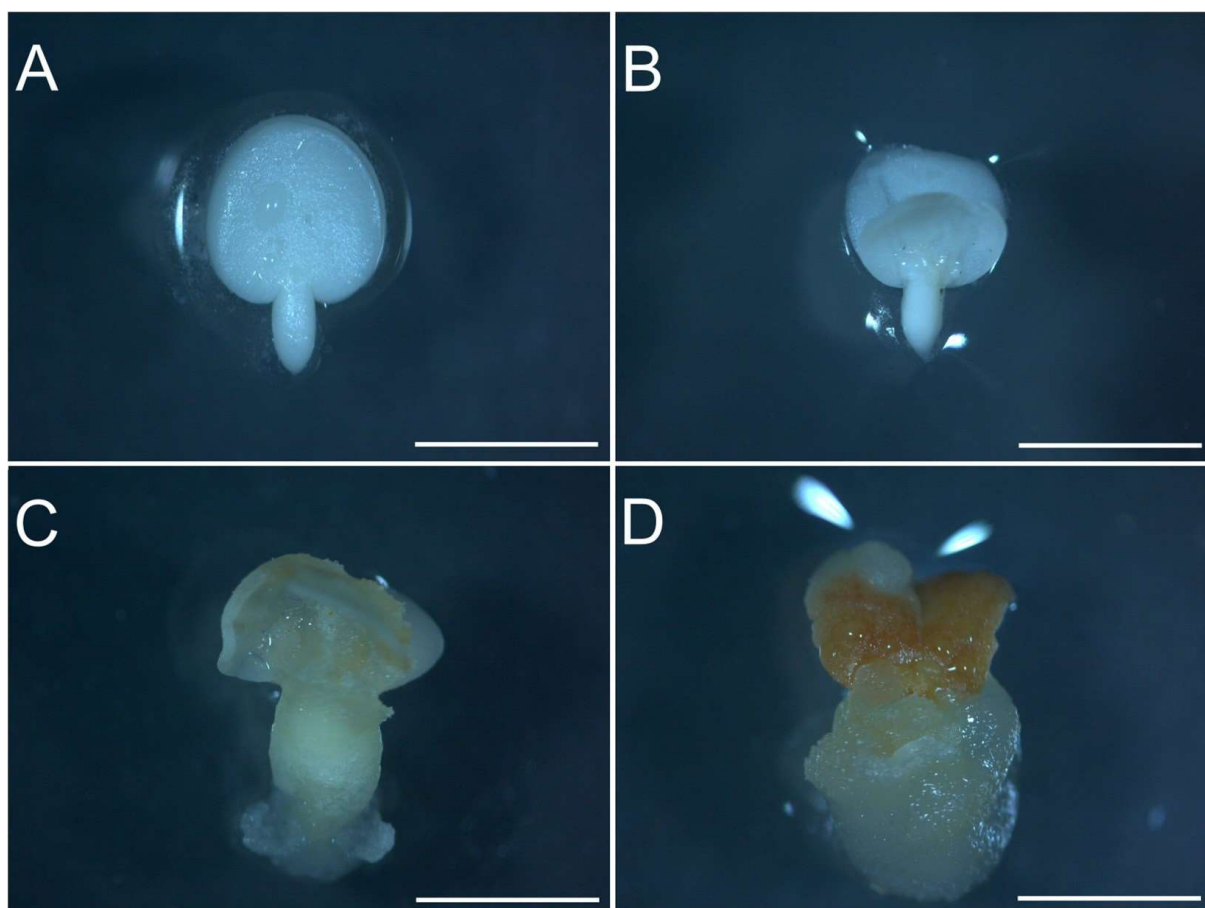


Figure 1. Morphological features of the explant during embryogenic callus induction using 2,4-D. Zygotic embryo used before incubation at time 0 (A). Initial cotyledon folding in the induction medium after 7 days of inoculation (B), followed by callus formation after 14 days of inoculation (C), and proliferation after 21 days of inoculation (D). The white bar in the images represents 0.5 cm.

Proteomics Analysis during Embryogenic Callus Induction

To understand the molecular signature of embryogenic callus formation, a time-series bottom-up proteomic approach was performed. A total of 1407 proteins were identified in bottom-up proteomic analysis of explant samples before and at 7, 14, and 21 days after induction of embryogenic callus (Supplementary Table S1). After the comparative proteomic analysis, 957 proteins presented differential accumulation in at least one comparison between the different induction times (Supplementary Table S1). The principal component analysis (PCA) showed a separation between explants before (zygotic embryo) the other sampling times in the first component (Supplementary Figure S1). In addition, callus induction at 7 days was equidistant between explants before incubation and callus after 14 and 21 days of incubation. However, explants after 7 days of induction showed a difference from the other treatments observed in the second component, probably due to a group of proteins responsible for initiating the process of dedifferentiation. The results show that callus cells at

14 and 21 days were closely grouped, possibly indicating that after 14 days of incubation, the effects of the induction treatment are smaller.

The clustering analysis revealed four clusters with distinct patterns of protein accumulation throughout embryogenic callus induction (Figure 2). Cluster 1 contained 386 proteins that increased accumulation at all analyzed times during induction. In contrast, cluster 2 contained 165 proteins showing decreased abundance during callus induction. Cluster 3 contained 251 proteins that were most abundant at the start of callus induction at 7 days of incubation and cluster 4 contained 155 proteins that accumulated after callus formation at 21 days of incubation (Figure 2; Supplementary Table S2). By differential accumulation analysis, it was also possible to identify proteins that have been associated with SE, and may be involved in embryogenic competence acquisition in the 2,4-D-mediated induction of *C. papaya* embryogenic callus (Table 1).

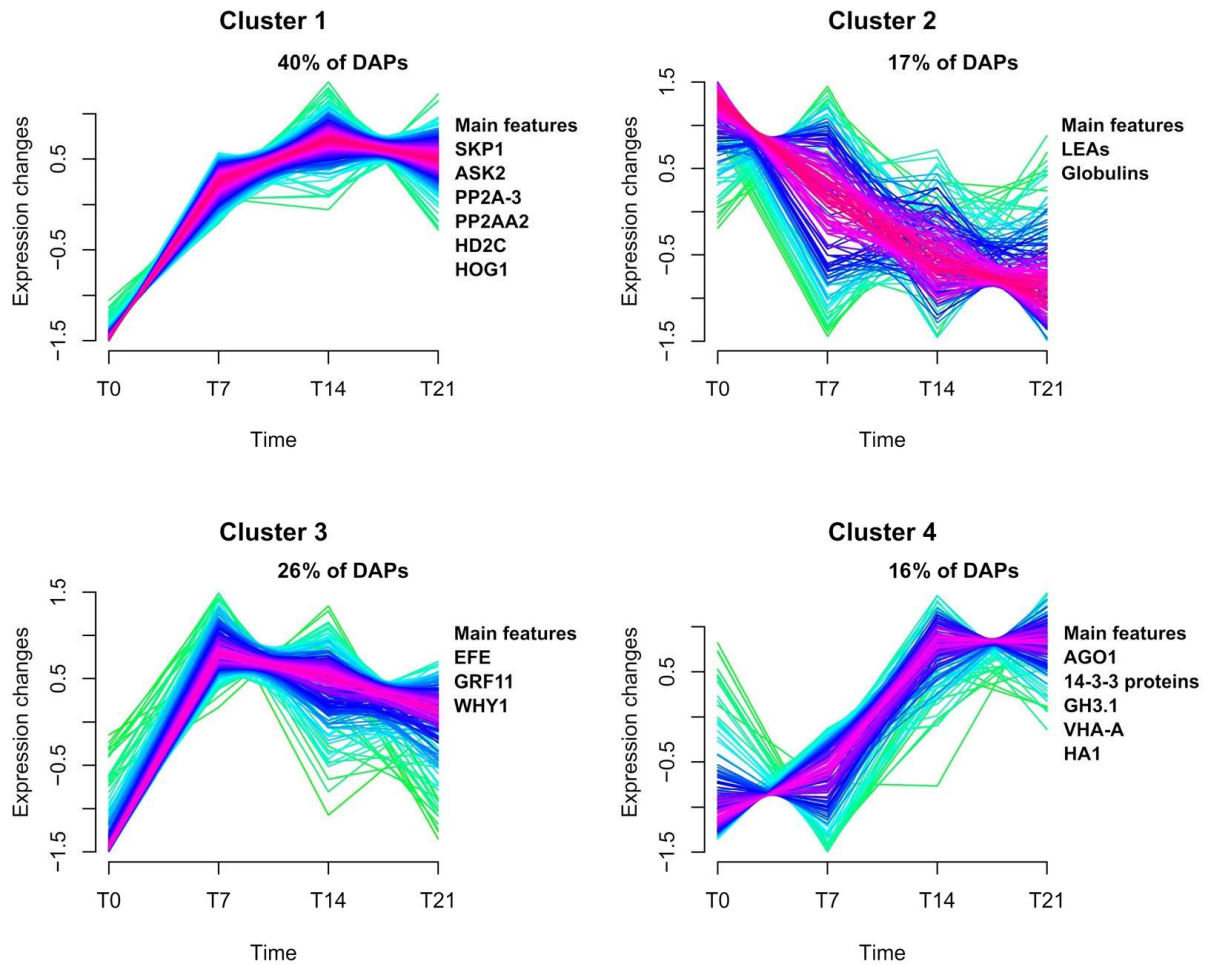


Figure 2. Results of the comparative analysis of proteomics data. Evaluation of the protein clusters revealed four patterns of accumulation throughout the time of callus induction. This analysis revealed proteins that increased in accumulation at all time points during callus induction (Cluster 1) and proteins that decreased their abundance throughout this process (Cluster 2). In addition, there are also groups of proteins that were abundant at the beginning of callus formation (7 days of induction) (Cluster 3), and proteins which accumulated after callus formation, at 21 days of incubation in the induction medium (Cluster 4).

Table 1. Differentially accumulated proteins with putative association with *C. papaya* embryogenic callus induction and their respective orthologs in *A. thaliana*. After statistical and functional analysis of the comparative proteomics data, it was possible to establish a list of proteins that may play a role in the induction of embryogenic calli. The unique 7 day+ group represents proteins that were quantified at 7, 14, and/or 21 days of incubation, but were absent in all biological replicates of explants before incubation (0 day).

Protein ID	BLAST Top Hit	<i>Arabidopsis</i> Orthologs	Cluster
evm.model.supercontig_1673.2	protein argonaute 1	AGO1	4
evm.model.supercontig_20.89	SKP1-like protein 1B	ASK2	1
evm.model.supercontig_97.102	14-3-3-like protein D isoform X2	At2g42590	1
evm.model.supercontig_309.6	14-3-3-like protein B	At5g65430	4
evm.model.supercontig_55.36	14-3-3-like protein GF14 kappa isoform X2	At5g65430	4
evm.model.supercontig_65.34	cell division cycle protein 48 homolog	CDC48C	Unique_T7+
evm.model.supercontig_1065.2	indole-3-acetic acid-amido synthetase GH3.6	DFL1	Unique_T7+
evm.TU.contig_32826.1	indole-3-acetic acid-amido synthetase GH3.6	DFL1	Unique_T7+
evm.model.supercontig_152.58	1-aminocyclopropane-1-carboxylate oxidase	EFE	3
evm.model.supercontig_36.169	14-3-3-like protein	GF14	1
evm.model.supercontig_292.1	probable indole-3-acetic acid-amido synthetase GH3.1	GH3.1	4
evm.model.supercontig_36.140	14-3-3-like protein	GRF11	3
evm.model.supercontig_7.47	putative 14-3-3 protein	GRF12	4
evm.model.supercontig_1.115	14-3-3-like protein GF14 iota	GRF12	4
evm.model.supercontig_23.10	plasma membrane ATPase 4	HA1	4
evm.model.supercontig_1377.1	plasma membrane ATPase 4	HA1	Unique_T7+
evm.model.supercontig_17.191	ATPase 11, plasma membrane-type	HA11	1
evm.model.supercontig_5.85	histone deacetylase HDT1	HD2C	1
evm.model.supercontig_157.46	adenosylhomocysteinase	HOG1	1
evm.model.supercontig_173.3	methionine S-methyltransferase	MMT	Unique_T7+
evm.model.supercontig_119.77	serine/threonine-protein phosphatase PP2A-2 catalytic subunit	PP2A-3	1
evm.model.supercontig_13.7	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A beta isoform	PP2AA2	1
evm.model.supercontig_1.195	serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A beta isoform	PP2AA2	1
evm.model.supercontig_6.142	soluble inorganic pyrophosphatase-like	PPa1	Unique_T7+
evm.model.supercontig_188.19	soluble inorganic pyrophosphatase 4	PPa4	Unique_T7+
evm.model.supercontig_179.17	14-3-3-like protein A	RC11A	1
evm.model.supercontig_67.71	S-adenosylmethionine synthase 2	SAM-2	2
evm.model.supercontig_26.289	SKP1-like protein 1B	SKP1	1
evm.model.supercontig_6.295	SKP1-like protein 1B	SKP1	Unique_T7+
evm.model.supercontig_1551.1	spermidine synthase-like	SPDS1	Unique_T7+
evm.model.supercontig_20.109	V-type proton ATPase subunit B 2	VAB2	1
evm.model.supercontig_155.13	V-type proton ATPase catalytic subunit A	VHA-A	4
evm.model.supercontig_150.48	V-type proton ATPase subunit a3	VHA-A2	1

Functional analysis of Proteomic Data

The functional analysis revealed that proteins related to seed maturation and storage reserves were more abundant in the explants before incubation, and their accumulation progressively decreased during induction treatment (Figure 3A). On the other hand, biological processes involved in carbohydrate and amino acid metabolism, aerobic respiration and protein catabolic processes were enriched at seven days of incubation (Figure 3A). Furthermore, the cellular components of secretory vesicles and lipid droplets were enriched among the most accumulated proteins in explants before incubation (0 day), while

components such as the proteasome, mitochondria, nucleolus, and cytosol were enriched among the proteins accumulated at 7, 14, and 21 days of incubation (Figure 3B).

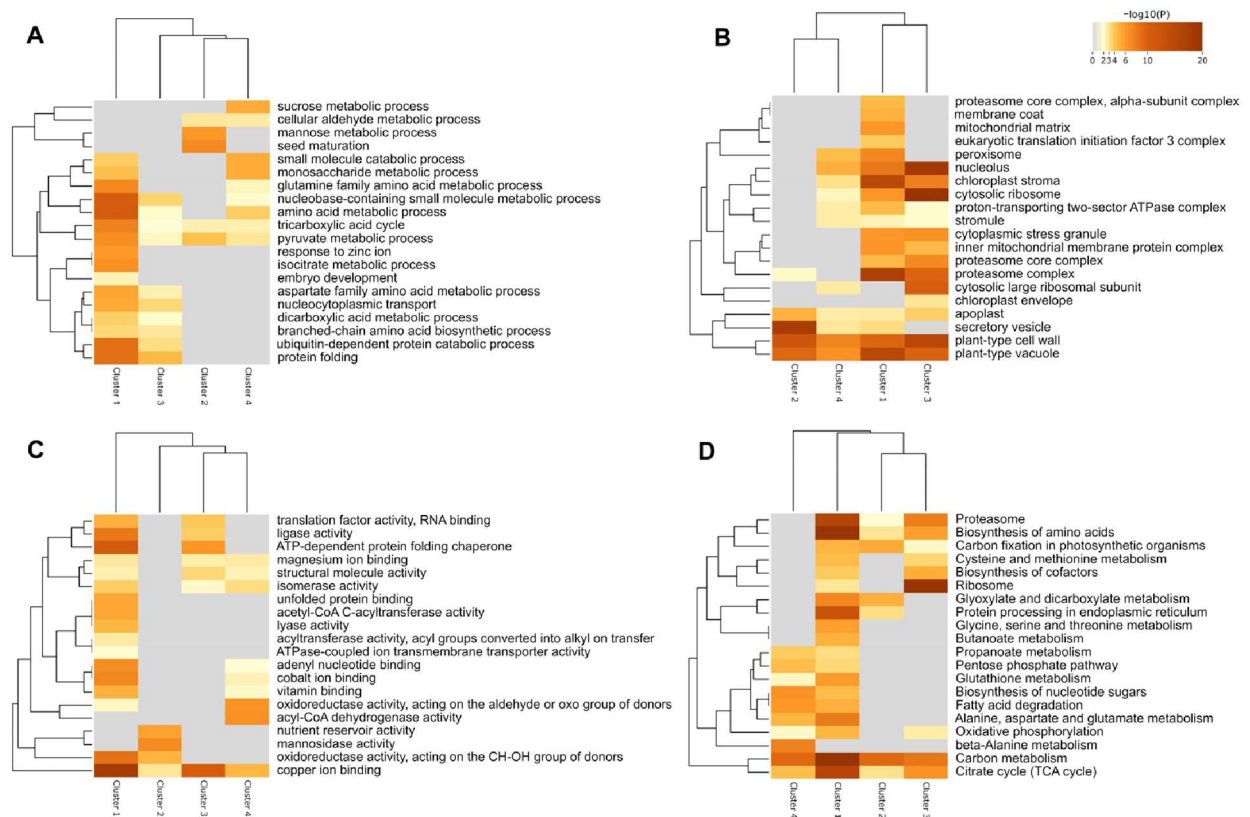


Figure 3. Gene ontology and KEGG pathways enrichment analysis among differentially accumulated proteins. The figure shows biological processes (A), cellular components (B), molecular function (C), and KEGG pathways (D) enriched among the different protein clusters differentially accumulated during embryogenic callus induction. The analysis shows that biological energy-generating processes are regulated during embryogenic callus induction, in addition to the accumulation of proteins with regulatory functions. In addition, several metabolic pathways are activated during embryogenic callus formation, including the metabolism of amino acids and reserve consumption.

The evaluation of the molecular functions of the differentially accumulated proteins showed enrichment of nutrient reservoirs in the explants before incubation (0 day), while functions involved in energy metabolism and stress response were enriched after the onset of induction (Figure 3C). In addition, enrichment analysis showed pathways associated with amino acid metabolism, reserve consumption, and ATP production during callus induction, indicating a metabolic shift associated with the process of embryogenic competence acquisition (Figure 3D).

Identification of Regulatory Proteins Differentially Accumulated during Induction

Due to the important regulatory role of protein kinases and transcription factors, a classification of the identified proteins was performed based on already known classes/types

(Supplementary Table S3). As a result, all differentially accumulated regulatory proteins were more abundant after the start of induction treatment, and no protein kinase or transcriptional regulator was identified in the explant before incubation, at 0 day. The identified regulatory proteins belong to seven different classes: CAMK_CDPK, MADS-MIKC, C2H2, CMGC_CDK-CRK7-CDK9, Whirly, CSD, and MBF1.

Prediction of Protein-Protein Interaction Networks

We used the data from functional analysis and differential accumulation to determine a set of proteins that may play a role in embryogenic callus formation (Table 1). The FASTA sequences of *C. papaya* proteins were used in STRING to predict protein–protein interactions. The search resulted in the identification of networks associated with oxidative phosphorylation, hormone response, SAM metabolism, and regulatory proteins 14-3-3 (Figure 4). In this network, the 14-3-3 proteins play an important role by connecting interactions between phosphatases, proteins associated with hormonal responses, and proteins involved in metabolic activity.

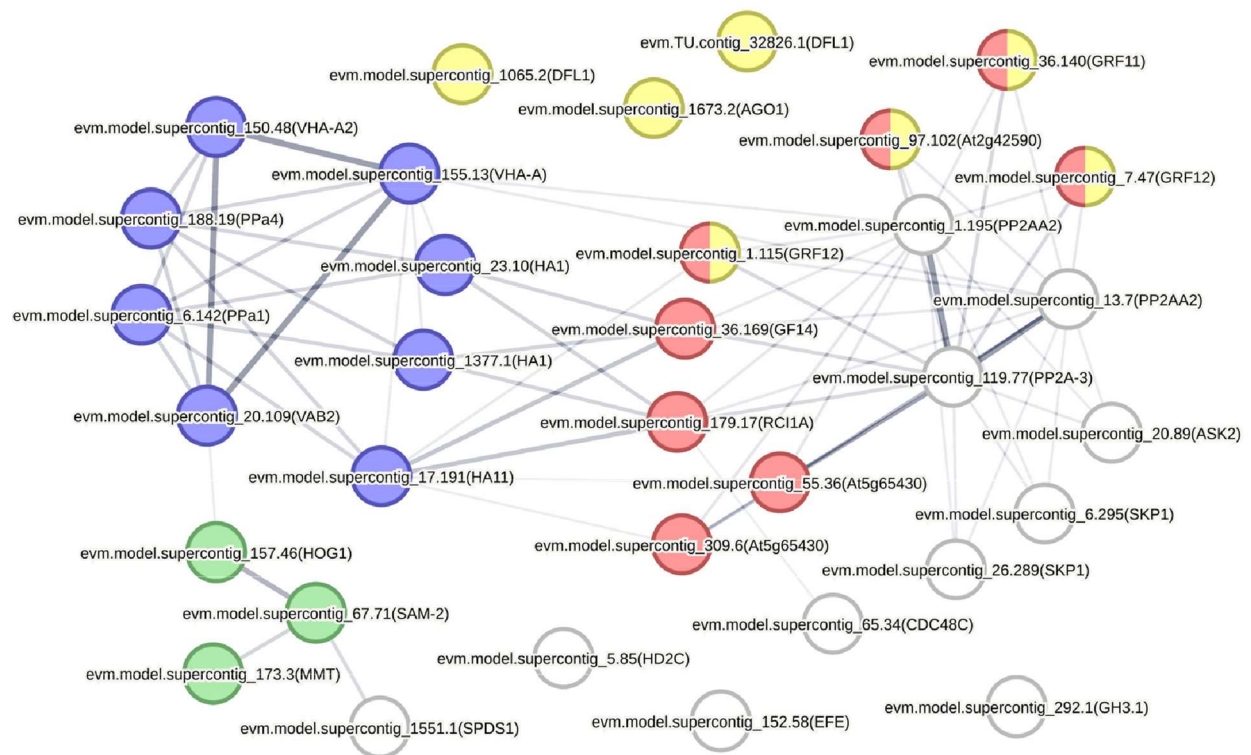


Figure 4. Protein–protein interaction network constructed with regulated DAPs during embryogenic callus induction. The network was constructed in STRING v.11.5 with *C. papaya* proteins. The names of the Arabidopsis orthologs are in parentheses. The color at the nodes indicates hormone response proteins (yellow), oxidative phosphorylation (blue), 14-3-3 (red), and SAM metabolism (green). The protein–protein interaction analysis highlights an important role for the 14-3-3 proteins in connecting hormonal signaling responses with energy metabolism during embryogenic callus induction. The network was constructed with medium confidence (combined score > 0.4).

Discussion

The Explant Contains Storage Reserves That Are Consumed during the Induction of Embryogenic Callus through Changes in Energy Metabolism

During zygotic embryogenesis, proteins and lipids accumulate throughout development [12]. Here, we identified proteins that accumulated in the explants, i.e., zygotic embryos before incubation (T0 days), that were enriched in the biological process of seed maturation (Figure 3A). Many of them have a nutrient reserve role, and their accumulation decreased over the course of embryogenic callus induction (Figure 1C). Our data reveal proteins involved in nutrient reservoir activity, such as 12S seed storage protein (evm.model.supercontig_165.8) and legumin B-like (evm.model.supercontig_165.7), which are transcriptionally active in embryogenic cultures and modulated in response to ABA [13,14]; oleosin 5-like (evm.model.supercontig_110.15) and oleosin 1 (evm.model.supercontig_13.63); and late embryogenesis abundant (LEA) proteins, such as

18 kDa seed maturation protein (evm.model.supercontig_85.72) and LEA protein 31 (evm.model.supercontig_58.99). An enhanced accumulation of storage reserves in embryogenic callus seems to be a precursor to the successful differentiation of somatic embryos [10]. Cells utilize stored carbohydrates, proteins, and other nutrients as a source of energy and building blocks for cell division and growth. During embryogenic callus induction, ABA initiates reprogramming of somatic cells, while LEA proteins protect cells from stress [15,16].

Modulation of embryogenic competence in *C. papaya* callus involves the accumulation of energy metabolism-related proteins [11]. In this regard, proteins with a role in electrochemical gradient generation and proton transport were identified, as well as pyrophosphatases regulated throughout callus induction (Table 1). Additionally, these proteins have predicted interactions that may be related to energy generation (Figure 3). This result is in agreement with the observation that biological processes such as protein catabolism and carbohydrate and amino acid metabolisms are enriched among DAPs after the start of embryogenic callus formation (Figure 3A). Furthermore, predicted interactions between oxidative phosphorylation proteins and 14-3-3 proteins were also observed, suggesting that they play an important regulatory role in integrating hormone responses with energy metabolism during embryogenic callus formation (Figure 4). These changes in energy metabolism could be related to the effects of 2,4-D and sucrose present in the induction culture medium, in addition to the catabolism of storage molecules present in the explant, such as globulins, lipids, and carbohydrates.

Acquisition of Embryogenic Competence in Somatic Cells Involves Modulation of Regulatory Proteins and Epigenetic Mechanisms

Auxin-induced SE involves the modulation of epigenetic mechanisms such as DNA methylation and miRNA-directed silencing, where histone deacetylases (HDACs) and argonaute protein 1 play important roles [6]. In this context, the phenotype of different HDAC mutants of *A. thaliana* was studied during direct SE, leading to the identification of HDACs that participate in embryo differentiation [17]. Additionally, several studies have demonstrated that histone acetylation/deacetylation is an important regulator of the in vitro induced embryogenic transition and is related to the differential expression of several elements of the HDAC gene families and histone acetyltransferase [18,19]. The increase in the accumulation of histone deacetylase HDT1 (evm.model.supercontig_5.85, orthologous to AtHD2C) protein was accumulated throughout the induction of the embryogenic callus (Supplementary Table S1). Studies in sugarcane have shown that histone deacetylase 1, belonging to the HDAC family,

is also involved in the differentiation of somatic embryos, interacting with corepressors such as TOPLESS to modulate protein post-translational regulation during somatic embryo differentiation [20]. Thus, the temporal analysis of the data in the present study suggests that histone acetylation/methylation is regulated at early stages of callus formation and acquisition of competence in *C. papaya* embryogenic callus. On the other hand, this study presents several proteins and interactions that are not observed in models of somatic embryogenesis regulation that only take into account transcriptomic and genetic studies [6,17,21]. It is clear that elucidating the regulatory mechanism involved in somatic embryogenesis depends on studying different layers of regulation of gene expression, from the epigenetic regulation of DNA, to the maturation and transport of messenger RNA, protein translation, and post-translational modifications.

Our data show that protein argonaute 1 (evm.model.supercontig_1673.2, orthologous to AtAGO1) accumulates throughout embryogenic callus induction and has predicted interactions with two SKP1-like proteins 1B (evm.model.supercontig_26.289 and evm.model.supercontig_6.295, both orthologous to AtSKP1), which are important for SCF^{TIR1/AFB} complex assembly and auxin responses, as well as SKP1-like protein 1B (evm.model.supercontig_20.89, orthologous to AtASK2) (Table 1) (Figure 4). The participation of SKP in callus induction has been suggested before, given its role in lateral root formation [7]. In this regard, the present study provides new evidence for the participation of SKP1 and ASK2 in *C. papaya* callus formation, in addition to their possible association with AGO1-mediated epigenetic regulation. In this context, DNA methylation is also a silencing mechanism known to be involved in the regulation of genes with a role in SE [22]. In the present study, S-adenosylmethionine synthase 2 (evm.model.supercontig_67.71), adenosylhomocysteinase (evm.model.supercontig_157.46), and methionine S-methyltransferase (evm.model.supercontig_173.3) proteins accumulated during embryogenic callus initiation, where they may play a role in S-adenosylmethionine (SAM) homeostasis for regulating gene transcription (Table 1) (Figure 4). These interactions may play an important role in modulating the epigenetic mechanisms associated with the cell cycle and proliferation of callus cells in *C. papaya*. In this regard, further studies should aim to investigate how 2,4-D could be associated with the accumulation of proteins involved in epigenetic regulation during the formation of embryogenic callus.

Phosphorylation is a post-translational modification involved in fine tuning the interactions, localization, activity, and shape of proteins. Several proteins with regulatory roles are phosphorylated at specific tyrosine, threonine, and serine residues in embryogenic calli

and differentiating somatic embryos [20,23]. Our data reveal the accumulation of two serine/threonine-protein phosphatase PP2A-2 catalytic subunit (evm.model.supercontig_13.7 and evm.model.supercontig_1.195, orthologous to AtPP2A-2 and AtPP2A-3, respectively) and two serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A beta isoforms (evm.model.supercontig_13.7 and evm.model.supercontig_1.195, both orthologous to AtPP2AA2) throughout embryogenic callus induction and their predicted interaction with 14-3-3 proteins (Table 1) (Figure 4). The PP2A-3 protein is known to play a role in regulating stem cell division in Arabidopsis roots [24]. In contrast, PP2AA2 regulates the phosphorylation of PIN-FORMED (PIN) proteins, which are involved in auxin flux during plant development [25]. These data provide clues to understanding how protein phosphorylation is involved in embryogenic callus formation, where PP2A and PP2AA2 phosphatases may play a role. However, the targets of these phosphatases during callus induction still need to be identified to fully understand their participation in SE. Comparative phosphoproteomics studies could help identifying phosphorylation sites and motifs involved with protein kinase and phosphatase activity associated with the induction of embryogenic callus.

Proteins Involved in Hormone Responses Are Induced Early after Callus Induction

The regulation of endogenous auxin levels is one of the main properties leading to the use of 2,4-D for the induction of embryogenic callus [21]. During the SE of *C. papaya*, it has already been demonstrated that the use of 2,4-D in induction is essential for the formation of embryogenic callus [4]. Our data reveal that several proteins involved in hormone regulation were differentially accumulated during *C. papaya* embryogenic callus initiation. These proteins are involved in indole-3-acetic acid (IAA) conjugation and can regulate free auxin levels, as well as proteins involved in ethylene responses. Auxin homeostasis is an important factor in the embryogenic competence of *C. papaya*, where the accumulation of GH3 proteins has already been associated with the differentiation of somatic embryos [11]. In the present study, a probable indole-3-acetic acid-amido synthetase GH3.1 (evm.model.supercontig_292.1) and two GH3.6 (evm.model.supercontig_1065.2 and evm.TU.contig_32826.1, orthologs of AtDFL1) accumulated during the induction of embryogenic callus (Table 1) (Figure 4). These proteins are involved in the conjugation of amino acids to free auxin, regulating the levels of responses of this hormone [26,27]. Thus, our data indicate that GH3 proteins are unique or accumulate during embryogenic callus formation and may participate in the regulation of free auxin homeostasis, as suggested by temporal analysis.

The participation of 14-3-3 proteins in the regulation of SE has been evidenced in different studies [9,28]. Here, several 14-3-3 proteins were identified throughout the process

of embryogenic callus formation (Table 1). Furthermore, these proteins have predicted interactions and are associated with hormone responses (Figure 4). The 14-3-3 proteins have a known role in modulating ethylene levels, where RCI1A and GF14 regulate 1-aminocyclopropane-1-carboxylate accumulation through interactions with 1-aminocyclopropane-1-carboxylate synthase (ACS), impacting responses involving ethylene [29,30]. Consistently, 1-aminocyclopropane-1-carboxylate oxidase (evm.model.supercontig_152.58, orthologous to AtEFE1) accumulated during embryogenic callus induction in the present study (Table 1). In this context, it is already known that 2,4-D regulates ethylene biosynthesis and that modulation of this hormone is important for auxin responses during SE [31]. Thus, 14-3-3 proteins may also play an important role in regulating ethylene responses during the formation of *C. papaya* embryogenic calli. Interestingly, our predicted network reveals an important role of 14-3-3 proteins interacting simultaneously with protein phosphatases in auxin responses and electrochemical gradient generation, indicating their participation in the integration of these processes. Given the relevance of using sex determined seedlings for papaya cultivation, these findings could contribute to the optimization of the cloning process through somatic embryogenesis and facilitate the production of cloned plantlets [1,2].

Materials and Methods

Plant Material and Callus Induction

The culture medium for callus induction was prepared using Murashige and Skoog (MS) salts and vitamins (Phytotechnology Lab, Lenexa, KS, USA) [32] supplemented with 30 g L⁻¹ sucrose (Vetec, São Paulo, Brazil) and 20 µM 2,4-dichlorophenoxyacetic acid (Sigma-Aldrich, St. Louis, MO, USA), with the pH adjusted to 5.8. Then, 2 g L⁻¹ Phytigel® (Sigma-Aldrich) was added to the medium, followed by autoclaving at 121 °C for 15 min and polymerization in Petri dishes (90 × 15 mm) on a laminar flow [10]. The mature fruits of papaya cv 'Golden' were washed in running water and transferred to a laminar flow hood for immersion in 70% ethanol (Sigma-Aldrich) for 1 min and 30 min in 50% commercial bleach Qboa® (Anhembi SA, Osasco, SP, Brazil) containing sodium hypochlorite from 1 to 1.25%. The fruit was washed three times with autoclaved distilled water and sliced in half with a sterile knife. The seeds were dissected with a sterile scalpel, and ten mature zygotic embryos were used as explants and inoculated in Petri dishes containing callus induction culture medium. The explants were maintained in darkness at 25 ± 1 °C for up to 21 days for callus induction.

Samples of explants before incubation (time 0), and after 7, 14, and 21 days of callus induction were obtained for proteomic analysis.

Protein Extraction and Digestion

Three biological replicates (each biological replicate constituted by a pool of 25 explants) collected during the time of incubation were pulverized in 1.5 mL microtubes until a powder. Next, 375 μ L extraction buffer consisting of 7 M urea (GE Healthcare, Freiburg, Germany), 2 M thiourea (GE Healthcare), 2% Triton X-100 (GE Healthcare), 1% dithiothreitol (DTT, GE Healthcare), and 1 mM phenylmethanesulfonyl fluoride (PMSF, Sigma-Aldrich) was added to the sample powder. The protein concentration of each biological replicate was estimated using the 2-D Quant Kit (Cytiva, Marlborough, MA, USA).

The protein samples were precipitated using a methanol/chloroform method to remove any interfering compounds from the samples according to Nanjo, *et al.* [33]. The samples were then resuspended in a solution consisting of 7 M urea and 2 M thiourea, after which tryptic protein digestion (1:100 enzyme:protein, V5111, Promega, Madison, WI, USA) was performed using the filter-aided sample preparation (FASP) method as described in Wisniewski, *et al.* [34] with changes described in Botini, *et al.* [10]. The peptides were vacuum-dried and solubilized in 50 μ L of a solution containing 5% (v/v) acetonitrile (Thermo Fisher Scientific, Waltham, MA, USA) and 0.1% (v/v) formic acid (Sigma-Aldrich) in mass spectrometry (MS)-grade water (Sigma-Aldrich). The peptide concentration was estimated by measuring A205 nm using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific). The peptides were stored at -80°C until mass spectrometry analyses.

Bottom-Up Proteomics Analysis

A nanoAcquity UPLC M-class (Waters, Manchester, UK) coupled to a Synapt G2-Si HDMS mass spectrometer (Waters, Manchester, UK) with an electrospray ionization (ESI)-MS/MS source was used. A total of 2 μ g of peptides was loaded onto a C18 trap column (100 $\text{\AA}$, 5 μ m, 180 μ m $\times$ 20 mm, 2D; Waters) at 5 μ L min^{-1} for 3 min and then onto an HSS T3 analytical reversed-phase column (100 $\text{\AA}$, 1.8 μ m, 75 μ m $\times$ 150 mm; Waters) at 400 nL min^{-1} heated to 45 $^{\circ}\text{C}$. For peptide elution, a binary gradient was used, with mobile phase A consisting of water and 0.1% formic acid and mobile phase B consisting of acetonitrile and 0.1% formic acid. Gradient elution was started with 5% B, which was then increased to 40% B until 92 min and from 40 to 99% B until 96 min, after which it was maintained at 99% until

100 min, posteriorly decreased to 5% B until 102 min, and maintained at 5% B until the end of the run at 118.00 min.

The mass spectrometer was set up in positive, resolution mode (V mode), and dataindependent acquisition (DIA) mode, at 35,000 FWHM, with an ion mobility (HDMSE) wave velocity program starting at 800 m s⁻¹ and ending at 500 m s⁻¹. The transfer collision energy increased from 25 to 55 V in high-energy mode, the cone voltage was 40 V, and the capillary voltage was 2800 V. The pressure of the nano-flow gas was 0.5 bar, the purge gas flow rate was 150 L h⁻¹, and the source temperature was 100 °C. For the time-of-flight (TOF) parameters, the scan time was set to 0.5 s in continuous mode, with a mass range of 50 to 2000 Da. Human [Glu1]-fibrinopeptide B at 100 fmol µL⁻¹ was used as an external calibration standard, and lock mass acquisition was performed every 30 s. Mass spectra were subsequently acquired by using MassLynx 4.1 software (Waters, Manchester, UK).

Proteomics Data Analysis

The proteomics MS/MS spectra processing and database search were performed using ProteinLynx Global SERVER (PLGS) software v.3.02 (Waters) against the *C. papaya* ASGPBv0.4 protein bank from Phytozome (<https://phytozome.jgi.doe.gov>, accessed on 1 February 2023). The raw data processing settings included a low energy threshold of 150 (counts), a high-energy threshold of 50 and an intensity threshold of 750. Additionally, the analysis was performed with the following parameters: 2 missed cleavages, at least 3 fragment ions per peptide, a minimum of 7 fragment ions per protein, a minimum of 2 peptides per protein, fixed carbamidomethyl modifications, and variable oxidation and phosphoryl modifications. The false discovery rate (FDR) for peptide and protein identification was set at a maximum of 1%, with a minimum peptide length of 6 amino acids. The label-free quantification analyses were performed using the TOP3 quantification approach, followed by the multidimensional normalization process implemented within ISOQuant software v.1.7 [35].

The protein identification criteria were defined as follows: FDR of 1%, peptide score greater than 6, minimum peptide length of 6 amino acids and at least 2 peptides per protein. All the mass spectrometry proteomics data and results tables have been deposited in the ProteomeXchange Consortium [36] via the PRIDE [37] partner repository with the dataset identifier PXD046521. To ensure the quality of the results after data processing, only proteins present in all three biological replicates or absent (for unique proteins) were considered for differential accumulation analysis. Proteins with significant Student's *t*-test (two-tailed; equal

variance; p -value < 0.05) were considered up-accumulated if the Log2 fold change (FC) was greater than 0.585 and down-accumulated if the Log2 FC was less than -0.585 .

Automated functional annotation analysis was performed using OmicsBox software V3.0 (<https://www.biobam.com>), and a manual annotation was made in UniProtKB (<https://www.uniprot.org>, accessed on 1 February 2023). *Arabidopsis* orthologs of differentially accumulated proteins (DAPs) were used for enrichment analysis (p -value < 0.05) of biological processes, cellular components, molecular functions, and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways in the Metascape tool [38]. These orthologs were also used for predicted interaction networks through a STRING v11.5 search [39] and predicted kinases, transcription factors, and transcriptional regulators using iTAK software (<http://itak.feilab.net/cgi-bin/itak/index.cgi>) [40]. The principal component analysis, Venn diagram, and heatmap were plotted in R language V.4.3.1 [41]. The Mfuzz package [42] was used for clustering analysis of the DAPs (p -value < 0.05 and Log2 FC > 0.585 or < -0.585) in at least one comparison between the induction times studied.

Conclusions

In this study, regulatory proteins such as HD2C, AGO1, SKP1, and phosphatases were accumulated during the acquisition of embryogenic competence in somatic papaya cells. In this regard, the transition from vegetative to embryogenic development could be linked to the regulation of these proteins, where modulating their activity or abundance could be a way of optimizing somatic embryogenesis in papaya and other plants. In addition, it was observed that the abundance of proteins associated with seed maturation and storage reserves decreased during callus induction, while proteins involved in energy metabolism increased to support callus multiplication. This reveals the important role of selecting optimal carbon sources to improve the in vitro conditions for growing embryogenic callus. In this context, protein–protein interactions should play a key role in integrating hormonal responses (including protein kinases and phosphatases associated with auxin responses) with energy metabolism and epigenetic regulation, where 14-3-3 proteins (accumulated during callus induction) appear to play a key role. Therefore, our results confirm that specific protein groups act at different moments to induce embryogenic callus differentiation. Further studies should aim to investigate the role of 2,4-D in the regulation of accumulation and interaction between auxin response proteins during embryogenic callus induction.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/plants12223891/s1>, Supplementary Figure S1: Principal component analysis of

time-dependent proteomics data. Supplementary Table S1: Complete list of identified proteins during embryogenic callus induction at different stages. Supplementary Table S2: Cluster information for proteins regulated in at least one time point analyzed during sugarcane callus induction. Supplementary Table S3: List of putative transcription factors, transcription regulators, and protein kinases.

Author Contributions: L.R.X., V.B.P., C.S.-C. and V.S. designed the research; K.d.S.V., L.d.E.S.G., B.C.D., L.d.S.d.C. and D.D.R.P. carried out the experiments; L.R.X., C.C.G.C., R.P.d.P. and V.S. performed the proteomics and data analysis; and all the authors read, reviewed and approved the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by funds from the Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) (Proc E-26/211.310/2021; Proc E26/200.999/2021) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (307755/2019-3), awarded to V.S. This study was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil (CAPES), Finance Code 001. The authors would like to thank FAPERJ, CNPq and CAPES for providing scholarships to L.R.X., C.C.G.C., R.P.d.P., K.S.V., D.D.R.P., L.E.S.G., B.C.D., and L.S.C.

Data Availability Statement: The data presented in this study are openly available in PRoteomics IDentifications (PRIDE) under accession number PXD046521.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Salinas, I.; Salehi, M.; Hueso, J.J.; Cuevas, J. Assessment of two sex-determining procedures in 'BH-65' papaya from an economical and developmental point of view. *Fruits* **2018**, *73*, 184–190. [\[CrossRef\]](#)
2. Xavier, L.; Almeida, F.A.; Pinto, V.B.; Silveira, V. Large-scale regeneration of hermaphrodite emblings of *Carica papaya* L. Golden using early molecular sex determination during embryogenic callus multiplication. *Plant Cell Tissue Organ Cult.* **2021**, *146*, 643–649. [\[CrossRef\]](#)
3. Gouvea, D.S.; Chagas, K.; Cipriano, J.L.D.; Lopes, J.C.; Schmildt, E.R.; Otoni, W.C.; Schmildt, O.; Araujo, C.P.D.; Alexandre, R.S. Somatic embryogenesis in the commercial papaya hybrid UENF/Caliman 01 relying on plantlet production from sexed adult hermaphrodite donor plants. *An. Acad. Bras. Cienc.* **2019**, *91*, e20180504. [\[CrossRef\]](#)
4. Almeida, F.A.; Vale, E.M.; Reis, R.S.; Santa-Catarina, C.; Silveira, V. LED lamps enhance somatic embryo maturation in association with the differential accumulation of proteins in the *Carica papaya* L. 'Golden' embryogenic callus. *Plant Physiol. Biochem.* **2019**, *143*, 109–118. [\[CrossRef\]](#)
5. Al-Shara, B.; Taha, R.M.; Mohamad, J.; Elias, H.; Khan, A. Somatic embryogenesis and plantlet regeneration in the *Carica papaya* L. Cv. Eksotika. *Plants* **2020**, *9*, 360. [\[CrossRef\]](#)
6. Wójcikowska, B.; Wójcik, A.M.; Gaj, M.D. Epigenetic regulation of auxin-induced somatic embryogenesis in plants. *Int. J. Mol. Sci.* **2020**, *21*, 2307. [\[CrossRef\]](#)
7. Fehér, A. Somatic embryogenesis, Stress-induced remodeling of plant cell fate. *Biochim. Biophys. Acta Gene Regul. Mech.* **2015**, *1849*, 385–402. [\[CrossRef\]](#)
8. Wójcikowska, B.; Jaskóła, K.; Gaj, M.; Nowak, K.; Gaj, M.D. LEAFY COTYLEDON2 (LEC2) promotes embryogenic induction in somatic tissues of Arabidopsis, via YUCCA-mediated auxin biosynthesis. *Planta* **2013**, *238*, 425–440. [\[CrossRef\]](#)
9. Aguilar-Hernández, V.; Loyola-Vargas, V.M. Advanced proteomic approaches to elucidate somatic embryogenesis. *Front. Plant Sci.* **2018**, *9*, 1658. [\[CrossRef\]](#)
10. Botini, N.; Almeida, F.A.; Cruz, K.; Reis, R.S.; Vale, E.M.; Garcia, A.B.; Santa-Catarina, C.; Silveira, V. Stage-specific protein regulation during somatic embryo development of *Carica papaya* L. 'Golden'. *Biochim. Biophys. Acta Proteins Proteom.* **2021**, *1869*, 140561. [\[CrossRef\]](#)
11. Vale, E.M.; Santana, D.B.; Reis, R.S.; Sousa, K.R.; de Souza Filho, G.A.; Oliveira, J.G.D.; Santa-Catarina, C.; Silveira, V. Mitochondrial proteomics reveals new insights into embryogenic competence acquisition in *Carica papaya* L. callus. *J. Proteom.* **2022**, *252*, 104434. [\[CrossRef\]](#)

12. Kumar, V.; Jha, P.; Van Staden, J. LEAFY COTYLEDONS (LECs): Master regulators in plant embryo development. *Plant Cell Tissue Organ Cult.* **2020**, *140*, 475–487. [[CrossRef](#)]
13. Gliwicka, M.; Nowak, K.; Ciesła, E.; Gaj, M.D. Expression of seed storage product genes (CRA1 and OLEO4) in embryogenic cultures of somatic tissues of Arabidopsis. *Plant Cell Tissue Organ Cult.* **2012**, *109*, 235–245. [[CrossRef](#)]
14. Ghelis, T.; Bolbach, G.r.; Clodic, G.; Habricot, Y.; Miginiac, E.; Sotta, B.; Jeannette, E. Protein tyrosine kinases and protein tyrosine phosphatases are involved in abscisic acid-dependent processes in arabidopsis seeds and suspension cells. *Plant Physiol.* **2008**, *148*, 1668–1680. [[CrossRef](#)]
15. Tang, L.P.; Zhang, X.S.; Su, Y.H. Regulation of cell reprogramming by auxin during somatic embryogenesis. *αBIOTECH* **2020**, *1*, 185–193. [[CrossRef](#)]
16. Amara, I.; Zaidi, I.; Masmoudi, K.; Ludevid, M.D.; Pagés, M.; Goday, A.; Brini, F. Insights into Late Embryogenesis Abundant (LEA) proteins in plants: From structure to the functions. *Am. J. Plant Sci.* **2014**, *5*, 3440–3455. [[CrossRef](#)]
17. Moronczyk, J.; Braśzewska, A.; Wójcikowska, B.; Chwiałkowska, K.; Nowak, K.; Wójcik, A.M.; Kwaśniewski, M.; Gaj, M.D. Insights into the histone acetylation-mediated regulation of the transcription factor genes that control the embryogenic transition in the somatic cells of Arabidopsis. *Cells* **2022**, *11*, 863. [[CrossRef](#)]
18. Wickramasuriya, A.M.; Dunwell, J.M. Global scale transcriptome analysis of Arabidopsis embryogenesis *in vitro*. *BMC Genom.* **2015**, *16*, 301. [[CrossRef](#)]
19. Yakovlev, I.A.; Carneros, E.; Lee, Y.; Olsen, J.E.; Fossdal, C.G. Transcriptional profiling of epigenetic regulators in somatic embryos during temperature induced formation of an epigenetic memory in Norway spruce. *Planta* **2016**, *243*, 1237–1249. [[CrossRef](#)]
20. Xavier, L.R.; Almeida, F.A.; Pinto, V.B.; Passamani, L.Z.; Santa-Catarina, C.; de Souza Filho, G.A.; Mooney, B.P.; Thelen, J.J.; Silveira, V. Integrative proteomics and phosphoproteomics reveals phosphorylation networks involved in the maintenance and expression of embryogenic competence in sugarcane callus. *J. Plant Physiol.* **2022**, *268*, 153587. [[CrossRef](#)]
21. Fehér, A. Callus, Dedifferentiation, Totipotency, Somatic Embryogenesis: What these terms mean in the era of molecular plant biology? *Front. Plant Sci.* **2019**, *10*, 536. [[CrossRef](#)]
22. Osorio-Montalvo, P.; Sáenz-Carbonell, L.; De-la-Peña, C. 5-Azacytidine: A promoter of epigenetic changes in the quest to improve plant somatic embryogenesis. *Int. J. Mol. Sci.* **2018**, *19*, 3182. [[CrossRef](#)]
23. Almeida, F.A.; Passamani, L.Z.; Santa-Catarina, C.; Mooney, B.P.; Thelen, J.J.; Silveira, V. Label-free quantitative phosphoproteomics reveals signaling dynamics involved in embryogenic competence acquisition in sugarcane. *J. Proteome Res.* **2020**, *19*, 4145–4157. [[CrossRef](#)]
24. Yue, K.; Sandal, P.; Williams, E.L.; Murphy, E.; Stes, E.; Nikonorova, N.; Ramakrishna, P.; Czyzewicz, N.; Montero-Morales, L.; Kumpf, R.; *et al.* PP2A-3 interacts with ACR4 and regulates formative cell division in the Arabidopsis root. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 1447–1452. [[CrossRef](#)]
25. Karampelias, M.; Neyt, P.; De Groeve, S.; Aesaert, S.; Coussens, G.; Rolčík, J.; Bruno, L.; De Winne, N.; Van Minnebruggen, A.; Van Montagu, M.; *et al.* ROTUNDA3 function in plant development by phosphatase 2A-mediated regulation of auxin transporter recycling. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 2768–2773. [[CrossRef](#)]
26. Chiu, L.-W.; Heckert, M.J.; You, Y.; Albanese, N.; Fenwick, T.; Siehl, D.L.; Castle, L.A.; Tao, Y. Members of the GH3 family of proteins conjugate 2,4-D and dicamba with aspartate and glutamate. *Plant Cell Physiol.* **2018**, *59*, 2366–2380. [[CrossRef](#)]
27. Park, J.-E.; Park, J.-Y.; Kim, Y.-S.; Staswick, P.E.; Jeon, J.; Yun, J.; Kim, S.-Y.; Kim, J.; Lee, Y.-H.; Park, C.-M. GH3-mediated auxin homeostasis links growth regulation with stress adaptation response in Arabidopsis. *Biol. Chem.* **2007**, *282*, 10036–10046.

[[CrossRef](#)]

28. Karlova, R.; Boeren, S.; Russinova, E.; Aker, J.; Vervoort, J.; de Vries, S. The *Arabidopsis* SOMATIC EMBRYOGENESIS RECEPTORLIKE KINASE1 protein complex includes BRASSINOSTEROID-INSENSITIVE1. *Plant Cell* **2006**, *18*, 626–638. [\[CrossRef\]](#)
29. Yoon, G.M.; Kieber, J.J. 14-3-3 regulates 1-aminocyclopropane-1-carboxylate synthase protein turnover in *Arabidopsis*. *Plant Cell* **2013**, *25*, 1016–1028. [\[CrossRef\]](#)
30. Catalá, R.; López-Cobollo, R.; Mar Castellano, M.; Angosto, T.; Alonso, J.M.; Ecker, J.R.; Salinas, J. The *Arabidopsis* 14-3-3 protein RARE COLD INDUCIBLE 1A links low-temperature response and ethylene biosynthesis to regulate freezing tolerance and cold acclimation. *Plant Cell* **2014**, *26*, 3326–3342. [\[CrossRef\]](#)
31. Bai, B.; Su, Y.H.; Yuan, J.; Zhang, X.S. Induction of somatic embryos in *Arabidopsis* requires local YUCCA expression mediated by the down-regulation of ethylene biosynthesis. *Mol. Plant* **2013**, *6*, 1247–1260. [\[CrossRef\]](#)
32. Murashige, T.; Skoog, F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* **1962**, *15*, 473–497. [\[CrossRef\]](#)
33. Nanjo, Y.; Skultety, L.; Uváčková, L.; Klubíková, K.; Hajduch, M.; Komatsu, S. Mass spectrometry-based analysis of proteomic changes in the root tips of flooded soybean seedlings. *J. Proteome Res.* **2012**, *11*, 372–385. [\[CrossRef\]](#)
34. Wisniewski, J.R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **2009**, *6*, 359–362. [\[CrossRef\]](#)
35. Distler, U.; Kuharev, J.; Navarro, P.; Levin, Y.; Schild, H.; Tenzer, S. Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat. Methods* **2014**, *11*, 167–170. [\[CrossRef\]](#)
36. Vizcaino, J.A.; Deutsch, E.W.; Wang, R.; Csordas, A.; Reisinger, F.; Ríos, D.; Dianes, J.A.; Sun, Z.; Farrah, T.; Bandeira, N.; *et al.* ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **2014**, *32*, 223–226. [\[CrossRef\]](#)
37. Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D.J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; *et al.* The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **2018**, *47*, D442–D450. [\[CrossRef\]](#)
38. Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A.H.; Tanaseichuk, O.; Benner, C.; Chanda, S.K. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **2019**, *10*, 1523. [\[CrossRef\]](#)
39. Szklarczyk, D.; Gable, A.L.; Nastou, K.C.; Lyon, D.; Kirsch, R.; Pyysalo, S.; Doncheva, N.T.; Legeay, M.; Fang, T.; Bork, P.; *et al.* The STRING database in 2021: Customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **2020**, *49*, D605–D612. [\[CrossRef\]](#)
40. Zheng, Y.; Jiao, C.; Sun, H.; Rosli, H.G.; Pombo, M.A.; Zhang, P.; Banf, M.; Dai, X.; Martin, G.B.; Giovannoni, J.J.; *et al.* iTAK: A program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol. Plant* **2016**, *9*, 1667–1670. [\[CrossRef\]](#)
41. Team, R.C. *R: A Language and Environment for Statistical Computing, R version 4.3.1 (2023-06-16 ucrt), “Beagle Scouts”*; R Foundation for Statistical Computing: Vienna, Austria, 2023.
42. Kumar, L.; Futschik, M.E. Mfuzz: A software package for soft clustering of microarray data. *Bioinformatics* **2007**, *2*, 5–7. [\[CrossRef\]](#) [\[PubMed\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Capítulo 3. Protein networks mediating embryogenic callus induction in *Carica papaya* L. respond to dose-dependent 2,4-dichlorophenoxyacetic acid treatment

Protein networks mediating embryogenic callus induction in *Carica papaya* L. respond to dose-dependent 2,4-dichlorophenoxyacetic acid treatment

Lucas Rodrigues Xavier ^{1,2}, Daniel Dastan Rezabala Pacheco ^{1,2}, Carlos Eduardo Assis da Silva ^{1,2}, Roberta Pena da Paschoa ^{1,2}, Claudete Santa-Catarina ³ and Vanildo Silveira ^{1,2}

¹Laboratório de Biotecnologia, Centro de Biociências e Biotecnologia (CBB), Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), Campos dos Goytacazes 28013-602, RJ, Brazil;

²Unidade de Biologia Integrativa, Setor de Genômica e Proteômica, UENF

³Laboratório de Biologia Celular e Tecidual, Centro de Biociências e Biotecnologia (CBB), UENF

Abstract

2,4-Dichlorophenoxyacetic acid (2,4-D), a potent herbicide, is widely used to promote both primary and secondary somatic embryogenesis in plants. However, identifying the optimal dosage of 2,4-D for effective callus or somatic embryo induction often relies on time-consuming empirical testing. This study investigates the physiological and proteomic responses of *Carica papaya* zygotic embryo explants to various concentrations of 2,4-D. Morphologically, the application of 2,4-D significantly inhibited radicle growth within seven days of incubation, in contrast to the control samples, where radicles grew unhindered. Proteomic analysis identified and quantified 1.168 proteins in *C. papaya*, of which 690 were classified as differentially accumulated proteins (DAPs) across the different treatments. Cluster analysis categorized these DAPs into six distinct patterns based on the concentration of 2,4-D: Clusters 1 and 2 comprised proteins that were upregulated even at high doses (200 μ M), while Clusters 3, 4, and 6 contained proteins that were downregulated with increasing 2,4-D concentrations. Cluster 5 included proteins that were upregulated at intermediate doses (0.2–20 μ M). Functional enrichment analysis revealed significant biological processes and metabolic pathways, such as cell division, hormonal response, and transcriptional regulation. Furthermore, nine regulatory proteins, including kinases and transcription factors, were identified as potential modulators of embryogenic callus induction. Predicted protein-protein interaction networks indicate a complex regulatory framework underlying 2,4-D-induced somatic embryogenesis in *C. papaya*. These findings enhance our understanding of the dose-dependent molecular mechanisms governing plant somatic embryogenesis, thereby facilitating the development of optimized protocols for plant propagation and genetic research.

Keywords: somatic embryogenesis, *Carica papaya*, proteomics, differential protein accumulation, regulatory networks

Corresponding author: Dr. Vanildo Silveira

E-mail: vanildo@uenf.br

Introduction

Somatic embryogenesis (SE) is a process by which haploid or diploid cells regenerate complete plants, replicating histodifferentiation patterns similar to those observed in zygotic embryos (Williams and Maheswaran, 1986). Among its variants, indirect somatic embryogenesis (ISE) stands out as a morphogenetic pathway involving callus formation followed by somatic embryo regeneration. Developed under controlled and sterile in vitro conditions (Fehér, 2015; Gladfelter *et al.*, 2021), SE holds significant potential for application in various plant species, including *Carica papaya*, where optimization and deeper understanding of its underlying mechanisms remain largely unexplored.

ISE is regulated by a complex network of interactions between genes and proteins, making it a promising approach for efficient micropropagation and biotechnological applications, such as genetic engineering (Heringer *et al.*, 2018). 2,4-Dichlorophenoxyacetic acid (2,4-D), a synthetic herbicide with auxinic properties, is widely used in plant tissue culture to induce callus formation during SE (Raghavan, 2004; Peterson *et al.*, 2016). This compound mimics the effects of natural auxins by activating the TIR1/AFB receptor-mediated signaling pathway, leading to the degradation of AUXIN/INDOLE-3-ACETIC ACID (Aux/IAA) transcriptional repressors and the subsequent activation of auxin-responsive genes (Song, 2014; Leyser, 2018).

Furthermore, 2,4-D plays a critical role in inducing cellular stress, a process closely linked to SE. Recent studies emphasize that its efficacy is directly related to its auxinic properties (Karami and Saidi, 2010; Karami *et al.*, 2023). Despite its widespread use in callus and somatic embryo induction, 2,4-D can have adverse effects that impair plant regeneration. High concentrations of 2,4-D are associated with genetic, epigenetic, and physiological disruptions, as well as morphological abnormalities in somatic embryos, often resulting in malformed structures (de Moraes Oliveira *et al.*, 2023). These findings underscore the importance of optimizing 2,4-D concentrations to enhance SE efficiency without compromising embryo quality.

Additionally, 2,4-D has been shown to alter the plant cell proteome, leading to the accumulation of proteins associated with cellular metabolism and hormonal responses. A proteomic study in *C. papaya* identified key regulatory proteins, such as histone deacetylase (HD2C) and argonaute 1 (AGO1), as crucial for the acquisition of embryogenic competence (Xavier *et al.*, 2023). However, the molecular regulation triggered by 2,4-D remains poorly

understood, necessitating significant investments of time and resources to identify optimal concentrations for various plant species and cultivars (Karami and Saidi, 2010).

Micropropagation protocols, particularly those involving SE, often rely on 2,4-D to establish embryogenic cell lines. We hypothesize that varying doses of 2,4-D regulate specific protein networks during the early stages of SE in papaya, coordinating the induction of embryogenic callus. To test this hypothesis, we performed a dose-dependent proteomic analysis over the first seven days of induction, evaluating molecular responses associated with different 2,4-D concentrations. Our findings reveal that 2,4-D modulates hormonal signaling and epigenetic mechanisms, leading to the accumulation of regulatory proteins essential for embryogenic competence. These insights provide valuable information for optimizing in vitro propagation protocols for this species.

Materials and Methods

Plant material and callus Induction

The culture medium for callus induction was prepared using Murashige and Skoog (MS) salts and vitamins (Phytotechnology Lab, Lenexa, KS, USA) as described by Murashige and Skoog, (1962). The medium was supplemented with 30 g L⁻¹ sucrose (Vetec, São Paulo, Brazil) and 2,4-dichlorophenoxyacetic acid (2,4-D) (Sigma-Aldrich, St. Louis, MO, USA) at concentrations of 0, 0.2, 2, 20, and 200 µM, with the pH adjusted to 5.8. Following this, 2 g L⁻¹ Phytigel® (Sigma-Aldrich) was added, and the medium was autoclaved at °C for 15 minutes, followed by polymerization in Petri dishes (90 × 15 mm) under a laminar flow hood (Botini *et al.*, 2021). Mature fruits of *C. papaya* cv 'Golden' were washed in running water, then transferred to a laminar flow hood for immersion in 70% ethanol (Sigma-Aldrich) for 1 minute, followed by 30 minutes in 50% commercial bleach (Qboa®; Anhembi SA, Osasco, SP, Brazil) containing sodium hypochlorite at a concentration of 1 to 1.25%. The fruits were then washed three times with autoclaved distilled water and sliced in half using a sterile knife (Vale *et al.*, 2014). The seeds were dissected with a sterile scalpel, and ten mature zygotic embryos were utilized as explants, inoculated into Petri dishes containing the culture medium supplemented with 2,4-D at the concentrations mentioned. The explants were maintained in the dark at 25 ± 1 °C for up to seven days for callus induction. Samples of the explants from each treatment were collected for proteomic analysis.

Protein extraction and digestion

Three biological replicates were prepared, each consisting of a pooled sample of fifteen independent explants, which were pulverized in 1.5 mL microtubes until a fine powder was achieved. Subsequently, 300 μ L of an extraction buffer, comprising 7 M urea (GE Healthcare, Freiburg, Germany), 2 M thiourea (GE Healthcare), 2% Triton X-100 (GE Healthcare), 1% dithiothreitol (DTT, GE Healthcare), and 1 mM phenylmethanesulfonyl fluoride (PMSF, Sigma-Aldrich), was added to the powdered samples. The protein concentration of each biological replicate was then estimated using the Bradford method.

The protein samples underwent precipitation via a methanol/chloroform method to eliminate any interfering compounds, as per the protocol outlined by Nanjo *et al.* (2012). The samples were then resuspended in a solution consisting of 7 M urea and 2 M thiourea, after which tryptic protein digestion (1:100 enzyme: protein, V5111, Promega, Madison, WI, USA) was performed using the filter-aided sample preparation (FASP) method as described by Wiśniewski *et al.* (2009) with changes described by Botini *et al.* (2021). The resulting peptides were vacuum-dried and solubilized in 50 μ L of a solution containing (v/v) acetonitrile (Thermo Fisher Scientific, Waltham, MA, USA) and 0.1% (v/v) formic acid (Sigma-Aldrich) in mass spectrometry (MS)-grade water (Sigma-Aldrich). Peptide concentrations were determined by measuring absorbance at 205 nm using a NanoDrop 2000c spectrophotometer. The peptides were stored at -80°C until mass spectrometry analyses.

Bottom-Up Proteomics analysis

A nanoAcquity UPLC M-class (Waters, Manchester, UK) coupled to a Synapt G2-Si HDMS mass spectrometer (Waters, Manchester, UK) with an electrospray ionization (ESI)-MS/MS source was used. A total of 2 μ g of peptides was loaded onto a C18 trap column (100 $\text{\AA}$, 5 μ m, 180 μ m $\times$ 20 mm, 2D; Waters) at 5 μ L min $^{-1}$ for 3 min and then onto an HSS T3 analytical reversed-phase column (100 $\text{\AA}$, 1.8 μ m, 75 μ m $\times$ 150 mm; Waters) at 400 nL min $^{-1}$ heated to 45 $^{\circ}\text{C}$. For peptide elution, a binary gradient was used, with mobile phase A consisting of water and 0.1% formic acid and mobile phase B consisting of acetonitrile and 0.1% formic acid. Gradient elution was started with 5% B, which was then increased to 40% B until 92 min and from 40 to 99% B until 96 min, after which it was maintained at 99% until 100 min, posteriorly decreased to 5% B until 102 min, and maintained at 5% B until the end of the run at 118.00 min.

The mass spectrometer was set up in positive, resolution mode (V mode), and data-independent acquisition (DIA) mode, at 35,000 FWHM, with an ion mobility (HDMSE) wave velocity program starting at 800 m s $^{-1}$ and ending at 500 m s $^{-1}$. The transfer of collision

energy increased from 25 to 55 V in high-energy mode, the cone voltage was 40 V, and the capillary voltage was 2800 V. The pressure of the nano-flow gas was 0.5 bar, the purge gas flow rate was 150 L h⁻¹, and the source temperature was 100 °C. For the time-of-flight (TOF) parameters, the scan time was set to 0.5 s in continuous mode, with a mass range of 50 to 2000 Da. Human [Glu1]-fibrinopeptide B at 100 fmol µL⁻¹ was used as an external calibration standard, and lock mass acquisition was performed every 30 s. Mass spectra were subsequently acquired by using MassLynx 4.1 software (Waters, Manchester, UK).

Proteomics data analysis

The proteomics MS/MS spectra processing and database search were performed using ProteinLynx Global SERVER (PLGS) software v.3.02 (Waters) against the *C. papaya* ASGPBv0.4 protein bank from Phytozome (<https://phytozome.jgi.doe.gov>, accessed on 30 August 2024). The raw data processing settings included a low energy threshold of 150 (counts), a high-energy threshold of 50 and an intensity threshold of 750. Additionally, the analysis was performed with the following parameters: 2 missed cleavages, at least 3 fragment ions per peptide, a minimum of 7 fragment ions per protein, a minimum of 2 peptides per protein, fixed carbamidomethyl modifications, and variable oxidation and phosphoryl modifications. The false discovery rate (FDR) for peptide and protein identification was set at a maximum of 1%, with a minimum peptide length of 6 amino acids.

Label-free quantification analyses were performed using the TOP3 quantification approach, followed by the multidimensional normalization process implemented within ISOQuant software v.1.7 (Distler *et al.*, 2014). The protein identification criteria were defined as follows: FDR of 1%, peptide score greater than 6, minimum peptide length of 6 amino acids and at least 2 peptides per protein. All the mass spectrometry proteomics data and results tables have been deposited in the ProteomeXchange Consortium (Vizcaíno *et al.*, 2014) via the PRIDE (Perez-Riverol *et al.*, 2019) partner repository with the dataset identifier PXDXXXX. To ensure the quality of the results after data processing, only proteins present in all three biological replicates or absent (for unique proteins) were considered for differential accumulation analysis. Proteins with significant Student's t-test (two-tailed; equal variance; p-value < 0.05) were considered up-accumulated if the Log2 fold change (FC) was greater than 0.585 and down-accumulated if the Log2 FC was less than -0.585.

Functional Annotation, Enrichment Analysis and TF prediction

Following protein identification, their ontology and biological functions were verified using Uniprot and TAIR (last accessed on 10 September 2024). *Arabidopsis thaliana* orthologs of differentially accumulated proteins (DAPs) were used for enrichment analysis (p-value < 0.05) of biological processes, cellular components, molecular functions, and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways via the g:Profiler tool (Kolberg *et al.*, 2023). These orthologs were also employed to predict interaction networks through STRING v12 searches (Szklarczyk *et al.*, 2021) to identify kinases, transcription factors, and transcriptional regulators using iTAK (Zheng *et al.*, 2016). Transcription factor enrichment analysis of DAPs was conducted using the PlantTFDB 4.0 tool (Jin *et al.*, 2017). The Mfuzz package (Kumar and Futschik, 2007) was utilized for clustering analysis of DAPs (p-value < 0.05 and Log2 FC > 0.585 or < -0.585) in at least one comparison across the studied induction times.

Systems biology

Protein protein interaction analysis (PPI) was performed as described by Pacheco *et al.* (2023). In brief, the Cytoscape plugin StringApp was employed to identify associations between differentially expressed genes, utilizing a confidence score cutoff of 0.700 to visualize the PPI network. The MCODE (Molecular Complex Detection) application was used to extract highly interacting proteins from the PPI network as modules, applying a degree cutoff parameter of ≥ 3 while keeping other parameters at default settings (node score cutoff ≥ 2 , K-Core ≥ 2 , and max depth = 100). Finally, the CytoHubba plugin was employed to extract hub proteins from the PPI network using the Maximum Clique Centrality (MCC) classification method, which calculates the degree of each protein node (Bader and Hogue, 2003).

Results

Application of 2,4-D induces physiological and morphological responses on the explant Carica papaya.

2,4-D is a potent herbicide frequently used to induce primary and secondary somatic embryogenesis in plants. However, its application in somatic embryogenesis often requires empirical testing to determine the optimal dose for inducing callus or somatic embryos. In this study, zygotic embryos of *C. papaya* were employed as experimental material to assess the influence of 2,4-D on radicle growth and somatic embryogenesis (Figure 1). In the control treatment, germination occurs as expected, with the resumption of growth of the embryonic axis and subsequent expansion of the radicle, following a seven-day incubation period, root hair development becomes evident (Figure 1A). While low concentrations of 2,4-D have been

shown to inhibit root growth (Figure 1B-E), the optimal dose of this growth regulator showed an expected behavior, as illustrated in Figure 1D. However, the concentration of 200 μ M of 2,4-D proved to be highly inhibitory, completely preventing germination and the formation of embryogenic callus (Figure 1E).

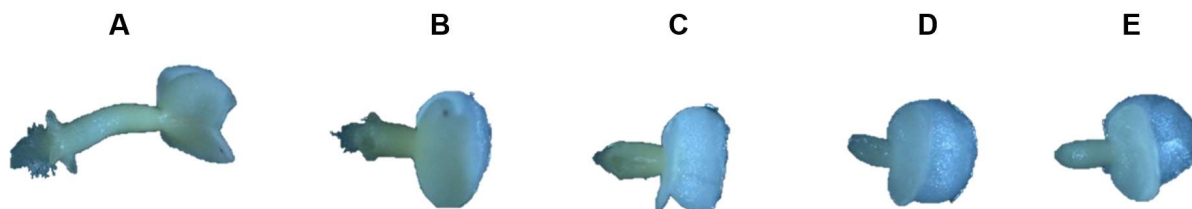


Figure 1. Morphological changes in *Carica papaya* explants during somatic embryogenesis induction with varying 2,4-D concentrations. (A) seven days of induction without 2,4-D (control). (B) seven days of induction with 0.2 μ M. (C) seven days of induction with 2 μ M. (D) seven days of induction with 20 μ M. (E) seven days of induction with 200 μ M.

The explant proteome is modulated by 2,4-D in a dose-dependent manner.

To elucidate the molecular mechanisms underlying the explant's response to 2,4-D, a bottom-up proteomic analysis was conducted. This approach enabled the identification and quantification of proteins, revealing alterations in the proteome across different herbicide concentrations. Proteomic analysis of the 2,4-D treatments resulted in the identification and quantification of 1168 *C. papaya* proteins (Supplementary Table S1). After statistical and fold change analysis, 690 were differentially accumulated proteins (DAPs) in at least one of the possible comparisons between all the 2,4-D treatments (Figure 2). The DAPs were used for cluster analysis, revealing six protein clusters with different accumulation patterns in response to 2,4-D doses. The analysis revealed an uneven distribution of proteins across groups. Group 3 had the highest number of proteins (195), followed by group 6 (190). Group 2 contained the fewest (42), while groups 1, 4, and 5 exhibited intermediate numbers (64, 105, and 94, respectively).

Clusters 1 and 2 group together proteins that increase their accumulation in response to 2,4-D even at high doses (200 μ M 2,4-D). Clusters 3, 4 and 6 group together proteins that decrease in accumulation in response to increasing doses of 2,4-D. Cluster 5 grouped together proteins that accumulate at intermediate doses of 2,4-D (0.2 - 20 μ M 2,4-D) (Figure 2 and Supplementary Table S2). In this way, it was possible to group the DAPs according to their accumulation in response to different doses of 2,4-D in the culture medium.

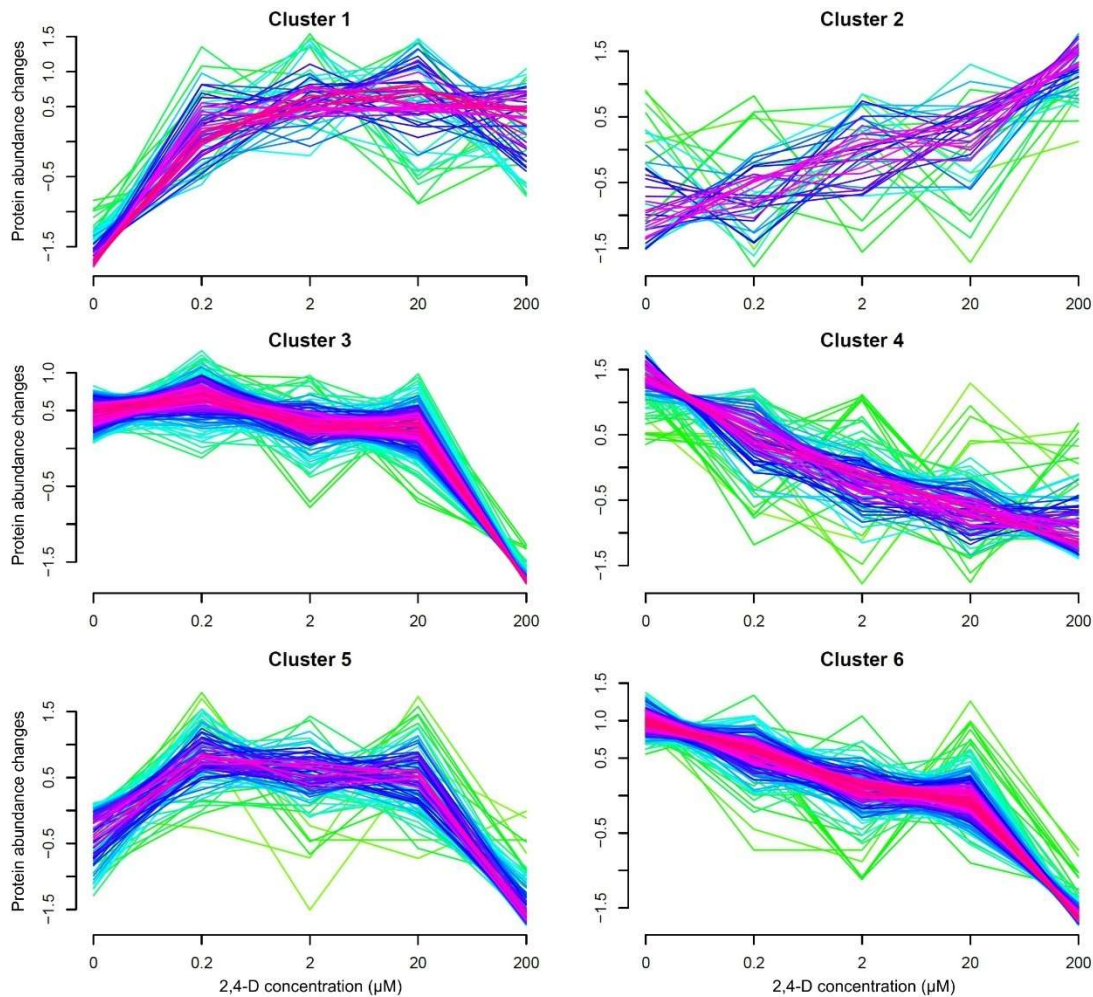


Figure 2. Dose-Dependent proteomic response to 2,4-D in *C. papaya* explants. Cluster analysis of DAPs revealed distinct response patterns to varying 2,4-D concentrations. Clusters 1 and 2 show increased protein accumulation, while Clusters 3, 4, and 6 exhibit decreased accumulation. Cluster 5 highlights a unique response at intermediate doses.

Regulatory network analysis of transcription factors and differentially accumulated proteins in response to 2,4-D

To gain deeper insight into the regulatory mechanisms governing the 2,4-D response, we employed a two-pronged approach. First, we performed a transcription factor (TF) enrichment analysis using 627 DAPs (Supplementary Table S2), as the remaining proteins lacked enrichment data. This analysis revealed 15,835 known interactions between 599 TFs and 627 input DAPs, suggesting a complex network of direct and indirect regulations (Supplementary Table S3). Among these TFs, 150 were significantly enriched ($p < 0.05$), indicating their crucial role in regulating DAPs, particularly in response to 2,4-D treatment.

These enriched TFs provide valuable insights into the underlying regulatory mechanisms and offer promising candidates for further functional studies.

Additionally, we identified nine proteins, including protein kinases, transcription factors, and other regulators, potentially involved in the 2,4-D response (Table 1). Four of these proteins exhibited differential accumulation. Notably, a transcription factor (cluster 3) and three proteins linked to transcription factors and kinases (cluster 5) emerged as potential key players in orchestrating the response to varying 2,4-D concentrations. This comprehensive analysis of regulatory proteins sheds light on the complex regulatory network underlying the 2,4-D response.

Functional analysis of Proteomic data

Gene Ontology (GO) enrichment analysis revealed that differentially abundant DAPs were significantly enriched in 177 biological process (BP) terms, 92 cellular component (CC) terms, and 76 molecular function (MF) terms. Additionally, 13 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were significantly enriched (Figure 3). Prominent biological processes included nitrogen compound metabolic and organic substance catabolism were the most significant and had the lowest gene ratio values (< 0.1). These processes are crucial for primary metabolism and provide the necessary resources for cell division, differentiation, and regeneration in response to 2,4-D treatment. The low proportion of these BPs indicates that a specific subset of genes was crucial for the observed response. This reinforces the idea that 2,4-D treatment affects specific and targeted metabolic pathways. The glyoxylate metabolic process was less significant but still relevant, despite having a lower overall impact, its high gene ratio indicates that a considerable proportion of the genes evaluated participate in this process. The glyoxylate cycle is particularly important under stress conditions, suggesting that its activation may be related to responses to stress induced by high doses of 2,4-D (Figure 3A).

Organelle, Membrane-Bounded Organelle, Intracellular Membrane-Bounded Organelle and Intracellular Anatomical Structure, these categories encompass structures such as nucleus, endoplasmic reticulum, mitochondria, and vacuoles. During somatic embryogenesis, cellular reprogramming requires intense nuclear and mitochondrial activity, as well as intracellular trafficking via membranous organelles (Figure 3B). MPs such as Small Molecule Binding and Ion include the ability of proteins to bind to small molecules, such as metabolites, cofactors and substrates. This is essential in metabolic processes and regulation. Ion binding refers to the binding to metal ions or other ions, important for enzymatic activities,

stabilization of protein structures and intracellular signaling (Figure 3C). Intermediate concentration of 2,4-D may have stimulated the expression of proteins that bind to small molecules and ions, supporting metabolic reactions and cellular reprogramming during somatic embryogenesis.

APT-Dependent Protein Folding Chaperone represents a distinct subset of molecular proteins. ATP-dependent chaperones play crucial roles in the correct folding of proteins, especially under stress conditions, such as those caused by 2,4-D concentrations. A smaller number of genes are involved, but with a strong functional impact (Figure 3C). These results indicate that DAPs are primarily involved in metabolic and biosynthetic processes (GO:1901564 and GO:1901576), intracellular localization within anatomical structures and membrane-bound organelles (including the cytosol) (GO:0005622 and GO:0043227), and small molecule binding activities (GO:0036094). Pathway enrichment analysis further highlights the significant involvement of DAPs in the biosynthesis of secondary metabolites (KEGG:01110) pathway.

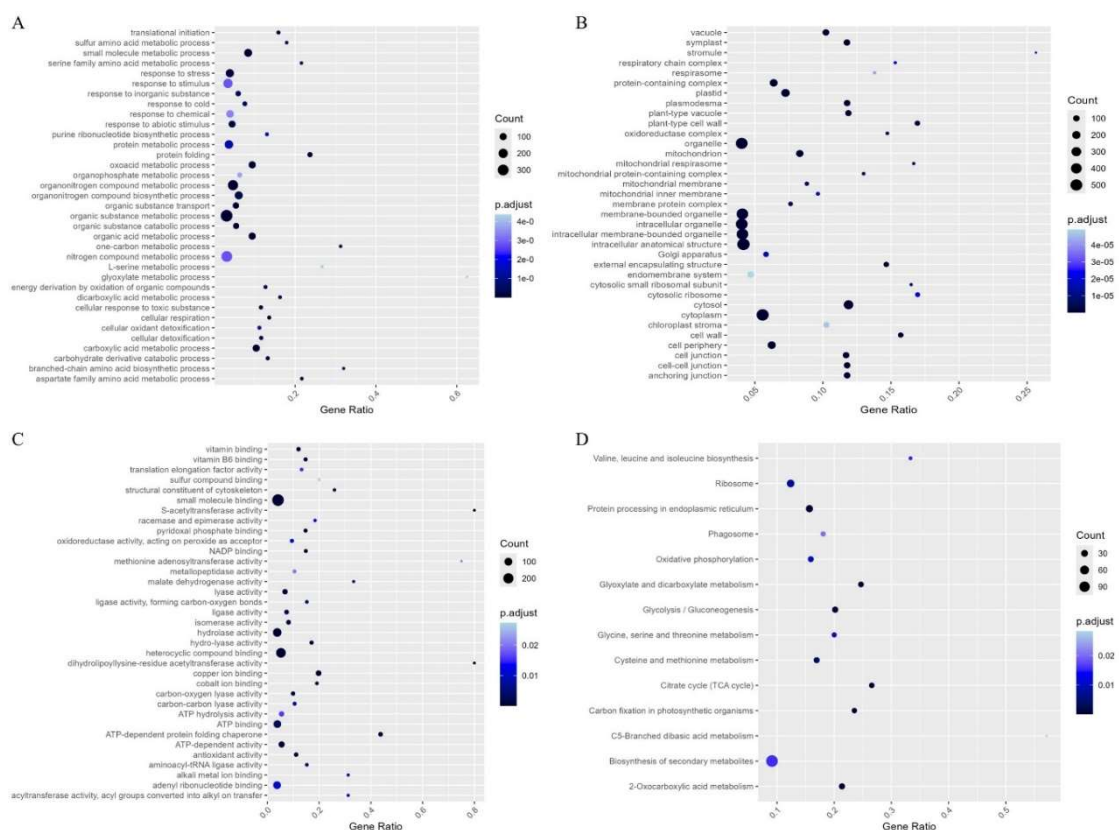


Figure 3. Functional enrichment analysis of proteomics data. The results of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses. BP: biological processes (A); CC: cellular components (B); MF: molecular function (C); and KEGG pathway

enrichment (D). The y-axis shows GO terms, while the x-axis denotes the proportion of genes. The size of each circle represents the number of genes. The shade of the circles symbolizes adjusted p-values.

Protein-protein interactions and predicted hub gene modules involved in embryogenic calli induction.

After statistical and functional analysis, it was possible to determine a list of proteins that respond to 2,4-D treatments and may be associated with the induction of embryogenic callus (Table 2). These proteins are associated with biological processes such as the control of cell division, proton transport, protein translation, hormonal response, post-translational modification of proteins and transcriptional regulation. Subsequently, MCODE analysis was used to predict the top significant modules retrieved from the protein–protein interaction (PPI) network. Module A contained 8 nodes and 28 edges, Module B had 4 nodes and 6 edges, Module C had 4 nodes and 5 edges, and Module D contained 3 nodes and 3 edges, with MCODE scores of 8.000, 4.000, 3.333, and 3.000, respectively (Figure S1).

According to the CytoHubba, Maximum Clique Centrality (MCC) method, the top 20 genes, including TIF3E1, TIF3A1, TIF3C1, TIF3F1, TIF3H1, TIF3D1, TIF3B1, TIF3K1, VDAC2, VHA-G1, VHA-E1, VDAC1, SKP1A, GRF8, MIRO1, VDAC4, VHA-C, VHA-H, EBP1, SUMO1, were classified as hub proteins (Figure S2). These hub proteins are mainly involved in translation initiation factor activity, RNA binding, translation initiation factor binding, Cullin family protein binding, auxin efflux transmembrane transporter activity, protein phosphatase regulatory activity, protein serine/threonine phosphatase activity, metal ion binding, ATP binding, transcription cis-regulatory region binding and transport of anions and cations across the outer membrane.

Furthermore, we constructed a new interactome by adjusting the maximum number of interactors displayed and setting a minimum interaction score of 0.700 (high confidence) to include proteins not previously identified as interactors. This predictive analysis revealed a regulatory network potentially involved in callus induction (Figure 5). The proteins in this network accumulate in a dose-dependent manner in response to 2,4-D, organizing themselves into clusters associated with different biological processes. The network was generated with Cytoscape's StringApp and consists of 472 nodes and 3,742 edges. The analysis highlights cluster 1, related to the brassinosteroid-mediated signaling pathway, interacting with cluster 20, associated with transition metal transport, through the proteins CDC25 (cluster 1) and GRF11 (cluster 20). Notably, CDC25 also connects cluster 1 to cluster 2, which is involved in polar auxin transport, via the hub protein GRF8.

While most clusters do not interact directly with each other, forming interactions only within their own modules, we identified several clusters associated with diverse biological processes. These clusters provide evidence for mechanisms underlying explant germination, embryogenic callus induction, and the toxic effects of high concentrations of 2,4-D. We highlight cluster 18, which is involved in protein amino acid deacetylation, a process linked to epigenetic regulation. These results provide new insights into the regulatory networks and biological processes underlying the response to 2,4-D and callus induction.

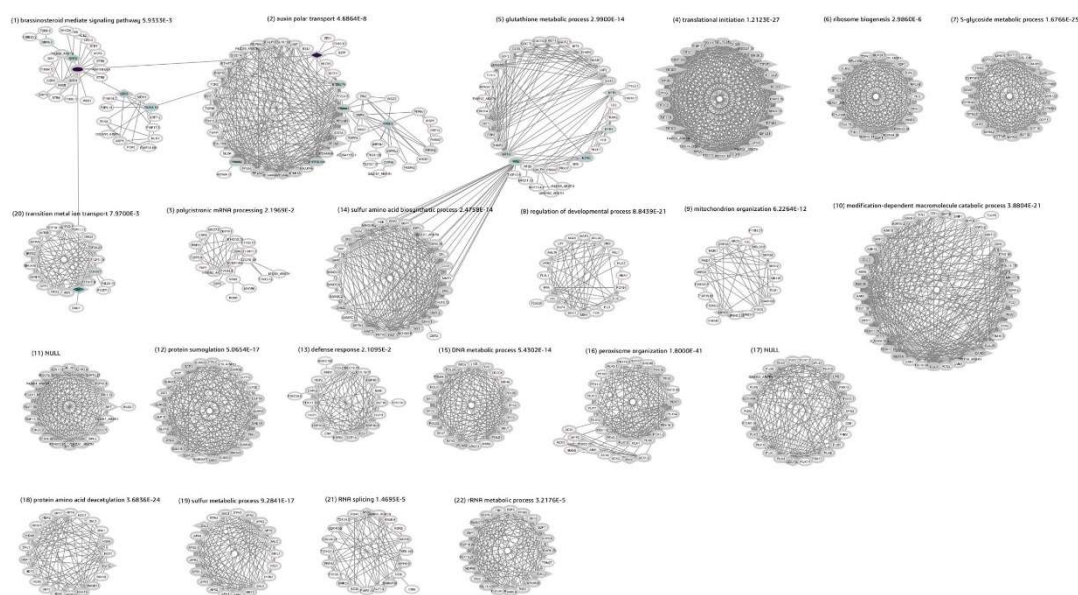


Figure 4. Predicted protein-protein interaction network of DAPs associated with embryogenic callus induction. Protein-protein interaction analysis identified 22 clusters, with each cluster being assigned a biological process based on the lowest corrected p-value, as determined using the BiNGO tool.

Discussion

Morphological and physiological effects of 2,4-D on somatic embryogenesis of C. papaya.

This study demonstrated that 2,4-D induces dose-dependent responses in explants, specifically zygotic embryos, after seven days of induction, which are critical for initiating somatic embryogenesis. In the control group, normal radicle development was observed (Figure 1B). At low concentrations, root growth was whereas the optimal dose stimulated the formation of embryogenic callus (Figure 1E), consistent with findings by Vale *et al.* (2018). However, a high concentration of 200 μ M 2,4-D was significantly inhibitory, completely suppressing germination and callus formation (Figure 1F). These results emphasize the

importance of empirical optimization to establish the ideal dose for each species and experimental objective.

Dose-dependent proteomic profile of 2,4-D in explants after seven days of induction

Proteomics has proven to be an invaluable tool for unraveling the biochemical and molecular mechanisms underlying somatic embryogenesis in plants (Heringer *et al.*, 2018). In this study, a comparative proteomic analysis was performed on *C. papaya* explants subjected to various concentrations of 2,4-D together with the control treatment, without 2,4-D (Supplementary Table S1). The results, discussed based on the differentially accumulated proteins (DAPs) and the unique proteins identified (Supplementary Table S2), highlight the relationship between the proteomic profile and the processes involved in the induction and development of somatic embryogenesis.

High doses of 2,4-D promote stress responses and reduce the accumulation of proton pumps

The use of 2,4-D as an herbicide is widespread in agriculture, where it is recognized as a defoliant agent (de Castro Marcato *et al.*, 2017). In the present study, it was observed that the use of a high dose of 2,4-D during the induction of embryogenic calluses can negatively affect the accumulation of energy metabolism and stress response proteins (Table 2 and Table S2). These proteins decrease their accumulation in high doses of 2,4-D and may be associated with the total inhibition of explant growth during the seven days of incubation (Figure 2). This effect underscores the importance of enriched biological processes, such as primary metabolism and the glyoxylate cycle, which provide essential resources for cell division, differentiation, and regeneration while also playing a key role in activating adaptation mechanisms to stress (Figure 3). Additionally, proteins such as EMB1467 (evm.model.supercontig_116.77), EMB3003 (evm.model.supercontig_66.53) and EMB1144 (evm.model.supercontig_53.71) belong to a group of genes crucial for embryogenic development in plants. Many of these genes are expressed throughout the plant's life cycle (Candela *et al.*, 2011; Meinke, 2020).

It is already known that the regulation of the accumulation and activity of proton pumps is an important factor for the expression of embryogenic competence in plants (Passamani *et al.*, 2020). In this sense, the data from the proteomic analysis of *C. papaya* explants in different treatments with 2,4-D show that proton pumps are regulated in a dose-dependent manner. These proton pumps include VDAC1, VDAC2, VDAC4, VHA-C, VHA-G1, VHA-E1 and VHA-H (evm.model.supercontig_5.272, evm.model.supercontig_6.23,

evm.model.supercontig_95.58, evm.model.supercontig_70.83,
 evm.model.supercontig_666.2, evm.model.supercontig_122.22 and
 evm.model.supercontig_43.106, respectively). Notably, these same proteins were identified
 as hub proteins with high centrality scores (Figure S1 and S2). These proton pumps are
 essential for generating electrochemical gradients, which are critical for adenosine
 triphosphate (ATP) production, molecular transport, and other vital processes that support cell
 growth and metabolism.

2,4-D modulates the post-translational modification of proteins and the accumulation of phosphatases and kinases

Proteomic studies have increasingly been employed to investigate post-translational modifications (PTMs) of proteins during somatic embryogenesis. It is well-established that regulatory proteins are phosphorylated at specific amino acid residues to modulate their activity, interactions, localization, and stability during embryogenic callus proliferation and somatic embryo differentiation (Aroonluk *et al.*, 2020; Xavier *et al.*, 2022). Thus, our data show that 2,4-D regulates the post-translational modification of proteins through changes in the accumulation of kinases, phosphatases, histone deacetylases and SUMOylation enzymes, as well as 14-3-3 phosphoproteins (Table 1). The SUMO1 and SUMO2 proteins (evm.model.supercontig_132.20 and evm.model.supercontig_8.137, respectively) are involved in different cellular processes (Vertegaal, 2022).

Phosphorylation, one of the most prevalent PTMs, can induce significant changes in a protein's three-dimensional structure, localization, and activity. The proteomics data from the present study shows that the protein kinases SKP1A (evm.model.supercontig_26.289 and evm.model.supercontig_6.295) and SKP1B (evm.model.supercontig_20.89), involved in auxin responses, are differentially accumulated in response to 2,4-D (Table 2). These proteins were grouped within the top 4 significant modules retrieved from the PPI network, and were also shown to be hub proteins, which represent proteins with many interactions (Figure S1 and Figure S2).

In addition, protein phosphatases such as PP2AA2, PP2AA2, PP2A4 and P23-1 (evm.model.supercontig_1.195, evm.model.supercontig_13.7, evm.model.supercontig_85.31 and evm.model.supercontig_9.223, respectively) are DAPs involved in the regulation of cell signaling that may be important for regulating 2,4-D-induced somatic embryogenesis. These phosphatases have predicted interactions with cell cycle control proteins and with 14-3-3 phosphoproteins associated with hormonal signaling processes (Supplementary figure 3).

Collectively, these results suggest that the induction and activation of somatic embryogenesis are mediated by the dynamic phosphorylation and dephosphorylation of proteins involved in hormonal signaling and cell cycle regulation in response to 2,4-D treatments.

2,4-D promotes protein turnover by modulating the proteasome machinery

Different hormonal pathways are often coupled to the degradation of repressor proteins in order to activate gene transcription associated with the hormonal response (Santner *et al.*, 2009). In this context, the regulation of the cellular proteome through the translation and ubiquitination of proteins is fundamental for cellular homeostasis. The data from the proteomic analysis of the explants at different doses of 2,4-D show that various members of the proteasome are differentially accumulated (Table 2 and Supplementary Table S4).

These proteins include RPN10, RPN8A, RPT1A, RPT3-2, RPN10, RPN12A, RPN1A, RPN6, RPT4B, RPT5A, RPT6A, RPN11 and RPT2A (evm.model.supercontig_37.159, evm.model.supercontig_163.7, evm.model.supercontig_25.205, evm.model.supercontig_2.273, evm.model.supercontig_37.159, evm.model.supercontig_79.5, evm.TU.contig_26338.1, evm.model.supercontig_83.90, evm.model.supercontig_106.1, evm.model.supercontig_327.2, evm.model.supercontig_334.5, evm.TU.contig_24415.6 and evm.model.supercontig_8.197, respectively). Thus, 2,4-D may regulate hormone homeostasis through the differential accumulation of proteasome members, promoting ubiquitination and protein degradation during this process.

An important point in the regulation of gene expression is the control of messenger RNA translation and protein synthesis (Kelen *et al.*, 2009). In this context, the data from the present study show that 2,4-D regulates proteins involved in translation, where intermediate doses (0.2 - 20 μ M 2,4-D) promote the accumulation of members eukaryotic translation initiation factor 3 (eIF-3) complex (evm.model.supercontig_53.167, evm.model.supercontig_48.127, evm.model.supercontig_17.146, evm.model.supercontig_5.97, evm.model.supercontig_182.25 and evm.model.supercontig_82.63) (Supplementary Table S4). In addition, our data showed predicted interactions between eIF-3 complex members and proteins involved with the cell cycle and kinases involved with auxin signaling, including EBP1 protein (Supplementary figure 3). Together, these data suggest that an important point in the induction of callus using 2,4-D is associated with the regulation of the cell proteome, where intermediate doses of 2,4-D promote translation, and high doses impair this process.

Optimal dose of 2,4-D to induce embryogenic callus promotes auxin signaling and response networks

During the induction of somatic embryogenesis using 2,4-D, one of the best-known processes is associated with the control of auxin biosynthesis, response and signaling (Yang and Zhang, 2010; Xavier *et al.*, 2023). Thus, auxin homeostasis is fundamental for the differentiation and development of the embryogenic callus and of the somatic embryo. One of the cellular responses observed in the explant in response to treatment with 2,4-D is the inhibition of germination and growth of the radicle (Figure 1). In this respect, the accumulation of GH3.1 proteins (evm.model.supercontig_292.1) and GH3.1 proteins (evm.model.supercontig_6.74) were observed in response to 2,4-D. Given that these proteins engage in the conjugation of indoleacetic acid (IAA), these data suggest that 2,4-D triggers changes in auxin homeostasis, impairing radicle growth and leading to the accumulation of GH3.1 to attenuate auxin responses. This process must also be important for the induction of embryogenic callus in response to 2,4-D.

Auxin responses have already been studied at various levels of gene expression regulation, but proteomic data is rarely incorporated into models of somatic embryogenesis regulation mechanisms. Proteomic analysis of explants subjected to different doses of 2,4-D showed that proteins important for auxin signaling are differentially accumulated in response to this growth regulator (Table supplementary S4). The proteins SKP1B (evm.model.supercontig_20.89), SKP1A (evm.model.supercontig_26.289 and evm.model.supercontig_6.295) and CAND1 (evm.model.supercontig_173.13) are regulated by auxin doses and are associated with the activation of response and hormone signaling genes. This suggests that the fine-tuning of 2,4-D concentrations for the induction of embryogenic callus may be associated with the appropriate regulation of the accumulation of auxin response proteins.

Epigenetic mechanisms associated with auxin are regulated by 2,4-D to modulate developmental responses.

Epigenetic modifications during somatic embryogenesis have already been extensively studied, where 2,4-D is known to modulate global DNA methylation levels, histone methylation and acetylation as well as miRNA-mediated regulation (Feng and Jacobsen, 2011; Smulders and de Klerk, 2011). Thus, our data show predicted interactions between proteins involved with DNA methylation that were differentially accumulated in the explant in response to 2,4-D (Figure 4). These proteins include SAHH1, SAL1, SAM1, SAM2 and SUR1

(evm.model.supercontig_157.46, evm.model.supercontig_42.165, evm.model.supercontig_146.36, evm.model.supercontig_67.71 and evm.model.supercontig_83.83, respectively). The SAM1 and SAM2 proteins participate in the metabolism of SAM, a substrate for DNA methylation. The SAHH1 protein is an adenosylhomocysteinase important for plant embryogenesis and is involved in SAM homeostasis, being important in DNA methylation-dependent gene silencing (Ouyang *et al.*, 2012; Watanabe *et al.*, 2021). Thus, the most effective doses of 2,4-D for inducing somatic embryogenesis appear to be associated with the fine-tuned regulation of gene expression through the modulation of epigenetic mechanisms, such as DNA methylation

Methylation and acetylation of histones is a process of epigenetic regulation based on the post-translational modification of proteins. In this context, histone deacetylases and methyltransferases are enzymes with a key role in controlling gene transcription. Our data show that 2,4-D regulates the accumulation of two histone deacetylases (HDACs) (evm.model.supercontig_5.85 and evm.model.supercontig_599.1) during embryogenic callus induction (Table supplementary S4). The regulation of this protein is already known in somatic embryogenesis, but little is known about the modulation of histone methylation mediated by 2,4-D. In this way, our data makes it clear that one of the first steps in the induction of embryogenic calluses is the regulation of epigenetic mechanisms to modulate transcription, but this process still needs to be further detailed in new studies using modulators of histone deacetylase activity.

Conclusions

This study highlights the profound impact of 2,4-D, particularly at higher concentrations, on various cellular processes involved in somatic embryogenesis. High doses of 2,4-D induce stress responses and reduce the accumulation of proton pumps, which are essential for maintaining the electrochemical gradients critical for cell growth and metabolism. This disruption in proton pump activity contributes to the inhibition of explant growth and embryogenic development observed during the early stages of treatment. Additionally, 2,4-D modulates protein turnover via the proteasome machinery, influencing the ubiquitination and degradation of key regulatory proteins. This regulation is dose-dependent: intermediate doses promote the accumulation of proteins involved in translation and hormonal signaling, such as members of the eukaryotic translation initiation factor 3 (eIF-3) complex, while high doses impair protein synthesis, potentially hindering proper embryogenic callus development.

Post-translational modifications, including phosphorylation and ubiquitination, are also modulated by 2,4-D. Proteins such as kinases, phosphatases, and 14-3-3 phosphoproteins were found to be differentially accumulated, underscoring their importance in regulating cell signaling and somatic embryogenesis. These modifications are critical for the proper functioning of hormonal signaling pathways and cell cycle control during callus induction. The study also demonstrated that auxin homeostasis, which is crucial for the differentiation and growth of embryogenic calluses, is finely regulated by 2,4-D. Proteins involved in auxin signaling, such as GH3.1, SKP1A, SKP1B, and CAND1, are differentially accumulated in response to 2,4-D, suggesting that the appropriate concentration of this herbicide is essential for inducing and maintaining embryogenic competence in explants. These proteins play a role in balancing auxin levels, which directly impacts cellular responses, such as the inhibition of radicle growth and the regulation of embryogenic callus formation.

Lastly, the study provides insights into the epigenetic regulation associated with 2,4-D treatment. Proteins involved in DNA methylation and histone modification were differentially accumulated, suggesting that epigenetic mechanisms play a pivotal role in the initial stages of somatic embryogenesis. The modulation of DNA methylation and histone deacetylation are essential for regulating gene expression in response to 2,4-D, although further studies are needed to fully elucidate the role of histone methylation in this process. Overall, the results demonstrate that the optimal dose of 2,4-D is critical for balancing stress responses, protein turnover, and auxin signaling, ensuring the successful induction of embryogenic calluses. This study enhances our understanding of the molecular mechanisms underlying 2,4-D-induced somatic embryogenesis and provides a foundation for future research to optimize callus induction protocols.

References

- Aroonluk, S., Roytrakul, S., and Jantasuriyarat, C. (2020). Identification and characterization of phosphoproteins in somatic embryogenesis acquisition during oil palm tissue culture. *Plants* 9. doi: 10.3390/plants9010036.
- Bader, G., and Hogue, C. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4, 1–27. doi: articles/10.1186/1471-2105-4-2.
- Botini, N., Almeida, F. A., Cruz, K. Z. C. M., Reis, R. S., Vale, E. M., Garcia, A. B., *et al.* (2021). Stage-specific protein regulation during somatic embryo development of *Carica papaya* L. 'Golden.' *Biochim. Biophys. Acta - Proteins Proteomics* 1869, 140561. doi: 10.1016/j.bbapap.2020.140561.

Candela, H., Pérez-Pérez, J. M., and Micol, J. L. (2011). Uncovering the post-embryonic functions of gametophytic- and embryonic-lethal genes. *Trends Plant Sci.* 16, 336–345. doi: 10.1016/j.tplants.2011.02.007.

de Castro Marcato, A. C., de Souza, C. P., and Fontanetti, C. S. (2017). Herbicide 2,4-D: A Review of Toxicity on Non-Target Organisms. *Water. Air. Soil Pollut.* 228. doi: 10.1007/s11270-017-3301-0.

de Morais Oliveira, J. P., Silva, A. J. da, Catrinck, M. N., and Clarindo, W. R. (2023). Embryonic abnormalities and genotoxicity induced by 2,4-dichlorophenoxyacetic acid during indirect somatic embryogenesis in *Coffea*. *Sci. Rep.* 13, 1–14. doi: 10.1038/s41598-023-36879-7.

Distler, U., Kuharev, J., Navarro, P., Levin, Y., Schild, H., and Tenzer, S. (2014). Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat. Methods* 11, 167–170. doi: 10.1038/nmeth.2767.

Fehér, A. (2015). Somatic embryogenesis - stress-induced remodeling of plant cell fate. *Biochim. Biophys. Acta (BBA)-Gene Regul. Mech.* 1849, 385–402. doi: 10.1016/j.bbagr.2014.07.005.

Feng, S., and Jacobsen, S. E. (2011). Epigenetic modifications in plants: An evolutionary perspective. *Curr. Opin. Plant Biol.* 14, 179–186. doi: 10.1016/j.pbi.2010.12.002.

Gladfelter, H. J., Johnston, J., Wilde, H. D., and Merkle, S. A. (2021). Somatic embryogenesis and cryopreservation of *Stewartia* species. *Plant Cell, Tissue Organ Cult.* 144, 211–221. doi: 10.1007/s11240-020-01834-1.

Heringer, A. S., Santa-Catarina, C., and Silveira, V. (2018). Insights from Proteomic Studies into Plant Somatic Embryogenesis. *Proteomics* 18, 1–9. doi: 10.1002/pmic.201700265.

Jin, J., Tian, F., Yang, D. C., Meng, Y. Q., Kong, L., Luo, J., *et al.* (2017). PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw982.

Karami, O., de Jong, H., Somovilla, V. J., Villanueva Acosta, B., Sugiarta, A. B., Ham, M., *et al.* (2023). Structure–activity relationship of 2,4-D correlates auxinic activity with the induction of somatic embryogenesis in *Arabidopsis thaliana*. *Plant J.* 116, 1355–1369. doi: 10.1111/tpj.16430.

Karami, O., and Saidi, A. (2010). The molecular basis for stress-induced acquisition of somatic embryogenesis. *Mol. Biol. Rep.* 37, 2493–2507. doi: 10.1007/s11033-009-9764-3.

Kelen, K. Van Der, Beyaert, R., Inzé, D., and Veylder, L. De (2009). Translational control of eukaryotic gene expression translational control of eukaryotic gene expression K. Van der Kelen *et al.* *Crit. Rev. Biochem. Mol. Biol.* 44, 143–168. doi: 10.1080/10409230902882090.

Kolberg, L., Raudvere, U., Kuzmin, I., Adler, P., Vilo, J., and Peterson, H. (2023). G:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Res.* 51, W207–W212. doi: 10.1093/nar/gkad347.

Kumar, L., and Futschik, M. E. (2007). Mfuzz: A software package for soft clustering of microarray data. *Bioinformatics* 2, 5–7. doi: 10.6026/97320630002005.

Leyser, O. (2018). Auxin signaling. *Plant Physiol.* 176, 465–479. doi: 10.1104/pp.17.00765.

- Meinke, D. W. (2020). Genome-wide identification of EMBRYO-DEFECTIVE (EMB) genes required for growth and development in *Arabidopsis*. *New Phytol.* 226, 306–325. doi: 10.1111/nph.16071.
- Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* 15, 1–21. doi: <https://doi.org/10.1111/j.1399-3054.1962.tb08052.x>.
- Nanjo, Y., Skultety, L., Uváčková, L., Klubicová, K., Hajduch, M., and Komatsu, S. (2012). Mass spectrometry-based analysis of proteomic changes in the root tips of flooded soybean seedlings. *J. Proteome Res.* 11, 372–385. doi: 10.1021/pr200701y.
- Ouyang, B., Fei, Z., Joung, J. G., Kolenovsky, A., Koh, C., Nowak, J., *et al.* (2012). Transcriptome profiling and methyl homeostasis of an *Arabidopsis* mutant deficient in S-adenosylhomocysteine hydrolase1 (SAHH1). *Plant Mol. Biol.* 79, 315–331. doi: 10.1007/s11103-012-9914-1.
- Pacheco, D. D. R., Santana, B. C. G., Pirovani, C. P., and de Almeida, A. A. F. (2023). Zinc/iron-regulated transporter-like protein gene family in *Theobroma cacao* L: Characteristics, evolution, function and 3D structure analysis. *Front. Plant Sci.* 14, 1–14. doi: 10.3389/fpls.2023.1098401.
- Passamani, L. Z., Reis, R. S., Vale, E. M., Sousa, K. R., Aragão, V. P. M., Santa-Catarina, C., *et al.* (2020). Long-term culture with 2,4-dichlorophenoxyacetic acid affects embryogenic competence in sugarcane callus via changes in starch, polyamine and protein profiles. *Plant Cell. Tissue Organ Cult.* 140, 415–429. doi: 10.1007/s11240-019-01737-w.
- Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., *et al.* (2019). The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* 47, D442–D450. doi: 10.1093/nar/gky1106.
- Peterson, M. A., McMaster, S. A., Riechers, D. E., Skelton, J., and Stahlman, P. W. (2016). 2, 4-D Past, Present, and Future: A Review. 303–345. doi: 10.1614/WT-D-15-00131.1.
- Raghavan, V. (2004). Role of 2,4-dichlorophenoxyacetic acid (2,4-D) in somatic embryogenesis on cultured zygotic embryos of *Arabidopsis*: Cell expansion, cell cycling, and morphogenesis during continuous exposure of embryos to 2,4-D. *Am. J. Bot.* 91, 1743–1756. doi: 10.3732/ajb.91.11.1743.
- Santner, A., Calderon-Villalobos, L. I. A., and Estelle, M. (2009). Plant hormones are versatile chemical regulators of plant growth. *Nat. Chem. Biol.* 5, 301–307. doi: 10.1038/nchembio.165.
- Smulders, M. J. M., and de Klerk, G. J. (2011). Epigenetics in plant tissue culture. *Plant Growth Regul.* 63, 137–146. doi: 10.1007/s10725-010-9531-4.
- Song, Y. (2014). Insight into the mode of action of 2,4-dichlorophenoxyacetic acid (2,4-D) as an herbicide. *J. Integr. Plant Biol.* 56, 106–113. doi: 10.1111/jipb.12131.
- Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., *et al.* (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49, D605–D612. doi: 10.1093/nar/gkaa1074.

Vale, E. D. M., Heringer, A. S., Barroso, T., Ferreira, A. T. D. S., da Costa, M. N., Perales, J. E. A., *et al.* (2014). Comparative proteomic analysis of somatic embryo maturation in *Carica papaya* L. *Proteome Sci.* 12, 1–18. doi: 10.1186/1477-5956-12-37.

Vale, E. M., Reis, R. S., Passamani, L. Z., Santa-Catarina, C., and Silveira, V. (2018). Morphological analyses and variation in carbohydrate content during the maturation of somatic embryos of *Carica papaya*. *Physiol. Mol. Biol. Plants* 24, 295–305. doi: 10.1007/s12298-017-0501-4.

Vertegaal, A. C. O. (2022). Signalling mechanisms and cellular functions of SUMO. *Nat. Rev. Mol. Cell Biol.* 23, 715–731. doi: 10.1038/s41580-022-00500-y.

Vizcaino, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Ríos, D., *et al.* (2014). ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* 32, 223–226. doi: 10.1038/nbt.2839.

Watanabe, M., Chiba, Y., and Hirai, M. Y. (2021). Metabolism and Regulatory Functions of O-Acetylserine, S-Adenosylmethionine, Homocysteine, and Serine in Plant Development and Environmental Responses. *Front. Plant Sci.* 12, 1–21. doi: 10.3389/fpls.2021.643403.

Williams, E. G., and Maheswaran, G. (1986). Somatic embryogenesis: Factors influencing coordinated behaviour of cells as an embryogenic group. *Ann. Bot.* 57, 443–462. doi: 10.1093/oxfordjournals.aob.a087127.

Wiśniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. *Nat. Methods* 6, 359–362. doi: 10.1038/nmeth.1322.

Xavier, L. R., Almeida, F. A., Pinto, V. B., Passamani, L. Z., Santa-Catarina, C., de Souza Filho, G. A., *et al.* (2022). Integrative proteomics and phosphoproteomics reveals phosphorylation networks involved in the maintenance and expression of embryogenic competence in sugarcane callus. *J. Plant Physiol.* 268, 153587. doi: 10.1016/j.jplph.2021.153587.

Xavier, L. R., Corrêa, C. C. G., da Paschoa, R. P., Vieira, K. da S., Pacheco, D. D. R., Gomes, L. do E. S., *et al.* (2023). Time-Dependent Proteomic Signatures Associated with Embryogenic Callus Induction in *Carica papaya* L. *Plants* 12, 3891. doi: 10.3390/plants12223891.

Yang, X., and Zhang, X. (2010). Regulation of somatic embryogenesis in higher plants. *CRC. Crit. Rev. Plant Sci.* 29, 36–57. doi: 10.1080/07352680903436291.

Zheng, Y., Jiao, C., Sun, H., Rosli, H. G., Pombo, M. A., Zhang, P., *et al.* (2016). iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. *Mol. Plant* 9, 1667–1670. doi: 10.1016/j.molp.2016.09.014.

Acknowledgements

This research was supported by funds from the Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) (Proc E-26/211.310/2021; Proc E-26/204.185/2024) and

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (312258/2023-2), awarded to V.S. This study was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil (CAPES), Finance Code 001. The authors would like to thank FAPERJ, CNPq and CAPES for providing scholarships to graduate students.

Author contributions

DDRP and LRX equally contributed to the development of the work. DDRP, LRX, and VS contributed to the development of the methodology, the acquisition of results, and the writing, review, and editing of the manuscript. CEAdS, RPdP, and CS-C reviewed, edited, and translated the manuscript.

Corresponding authors

Correspondence to Vanildo Silveira

Ethics declarations

The authors declare that they have no competing interests.

Ethics approval and consent to participate

This article does not contain any studies with human or animal participants. Work with plants complies with relevant institutional, national and international guidelines and legislation.

Consent for publication

All the authors provide their consent for this publication.

**Capítulo 4. The role of protein phosphorylation in auxin-mediated totipotency
and somatic embryogenesis of *Carica papaya* L.**

The role of protein phosphorylation in auxin-mediated totipotency and somatic embryogenesis of *Carica papaya* L.

Lucas Rodrigues Xavier ¹, Claudete Santa-Catarina ², Jay J. Thelen ³, Vanildo Silveira¹

¹Laboratório de Biotecnologia, Centro de Biociências e Biotecnologia (CBB), Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), Campos dos Goytacazes, RJ, 28013-602, Brazil

²Laboratório de Biologia Celular e Tecidual, CBB-UENF, Campos dos Goytacazes, RJ, Brazil

³Department of Biochemistry, Christopher S. Bond Life Sciences Center, University of Missouri, 1201 Rollins Street, 65211, Columbia, Missouri, United States of America

Abstract

Somatic embryogenesis (SE) is a cornerstone of plant biotechnology for regenerating elite genotypes and bypassing sexual variability. In *Carica papaya* L., where only hermaphrodite plants bear marketable fruit, SE offers a powerful propagation strategy. We applied integrated quantitative proteomics and phosphoproteomics to unravel the molecular basis of embryogenic competence in papaya callus. Bottom-up proteomics identified 6,182 proteins, with principal component analysis (PCA) clearly separating embryogenic and non-embryogenic samples (PC1 = 79.9% variance; PC2 = 6.5%). After normalization and statistical filtering ($p < 0.05$), 1,259 proteins were differentially accumulated and 576 were unique to one callus type, revealing extensive proteome reprogramming. Phosphoproteomic profiling detected 4,169 phosphopeptides (2,475 with probability > 0.75), of which 210 were differentially regulated and 344 were unique to embryogenic or non-embryogenic calli. Functional enrichment and motif analyses highlighted coordinated regulation of auxin transporters (ABCB19, PIN1-like), transcriptional repressors (SLK2), hormone homeostasis enzymes (GH3.1-like), and chromatin remodelers (HIRA, REF6, CTR9/VIP6), as well as epigenetic regulators (AGO1, AGO4). Phosphorylation of auxin flux facilitators and kinases (GTE4, ZAR1, CDKF-1) emerged as critical for establishing local auxin maxima and totipotency. Metabolic reprogramming was evidenced by phosphorylation of VHA-F, VDAC4, AMPD, and ABA-related factors (ABI3, SRK2E, SUA). Notably, PLT2 was detected exclusively in embryogenic calli, with predicted interactions with auxin proteins, marking it as a robust embryogenic competence marker. Collectively, this study reveals a complex phosphoregulatory network, integrating hormonal signaling, epigenetic plasticity, and metabolic shifts, that underpins somatic embryo formation, providing a molecular framework to optimize regeneration protocols in papaya and other recalcitrant tropical crops.

Keywords: Micropropagation; Bottom-up proteomics; Embryogenesis.

Corresponding author: Dr. Vanildo Silveira E-mail: vanildo@uenf.br

Introduction

The papaya tree (*Carica papaya* L.) is a species grown mainly in tropical countries, where Brazil is the world's second largest producer of this crop [1]. During papaya production, the sexual segregation of the seeds causes great losses, since the plants can be female, male or hermaphrodite, but the fruit with commercial value is only produced by the hermaphrodites. An alternative for reducing production costs in papaya is to use somatic embryogenesis to produce seedlings with a defined sex, avoiding losses caused by sexual segregation [2]. In this context, embryogenic callus can be induced using cotyledonary zygotic embryos as explants, followed by the differentiation of somatic embryos [3]. During in vitro development, somatic embryos go through the same morphological stages as zygotic embryos, demonstrating the potential of using somatic embryogenesis to study plant morphogenesis and embryogenic development [4]. Thus, the combination of somatic embryogenesis with molecular biology approaches (such as the use of molecular markers) makes it possible to select papaya callus with a hermaphrodite genotype and produce embryos with a defined sex, reducing production costs due to the sexual segregation of the seeds [5]. However, the use of somatic embryogenesis for the production of papaya seedlings still depends on the optimization of somatic embryo regeneration protocols and the reduction of risks associated with somaclonal variation.

The use of genetically modified organisms in agricultural production is increasing rapidly, and one of the biggest bottlenecks in the development of new cultivars using genetic engineering is the regeneration of transformed plants [6]. In this sense, somatic embryogenesis has been identified as one of the main regeneration routes for genetically modified plants (ref). In addition, the study of plant somatic embryogenesis has resulted in the identification of morphogens such as BABY BOOM, AGAMOUS-LIKE 15, WUSCHEL, SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE 1 and LEAF COTYLEDON, which can be applied to produce somatic embryos ectopically, reducing the time for plant regeneration after genetic transformation [6]. Thus, new approaches to the development of genetically modified plants employ morphogenesis to facilitate the regeneration stage in both monocots and dicots [7, 8].

Somatic embryogenesis is a great example of plant totipotency, where somatic (vegetative) cells differentiate into reproductive lineage cells and develop into embryos without the need for fertilization [9]. During this process, there is a major change in the hormonal balance of the cells, where auxin metabolism is identified as one of the main factors that modulates somatic embryogenesis and plant morphogenesis [10]. This is made even more

evident by the fact that the vast majority of protocols rely on the use of exogenous auxin (mainly 2,4-D) to induce somatic embryogenesis or the formation of embryogenic callus [11]. In this context, even approaches that use the ectopic expression of morphogens to initiate the differentiation of embryos into somatic cells are associated with the accumulation of endogenous auxin (indole-3-acetic acid, IAA) [9]. In addition to determining the differentiation of embryogenic cells, auxin metabolism and transport are fundamental for embryo patterning and morphogenesis, directing the differentiation of cotyledons, vascular system, hypocotyl and development from the globular embryo stage [12]. This depends on the activation of different groups of genes, including YUCCA (YUC, involved in auxin biosynthesis), PIN-FORMED (PIN) and P-GLYCOPROTEIN (PGP), responsible for auxin efflux, and AUXIN1/LIKE-AUX1 (AUX/LAX) associated with auxin influx [13]. After this process, auxin gradient-dependent expression is activated and directs the expression of embryogenic identity and morphogenesis genes, leading to the development of the somatic embryo [14].

Post-translational modifications (PTMs) of proteins, particularly phosphorylation, play a central role in modulating cellular responses and regulating diverse physiological processes during development and environmental stress adaptation [15]. Protein phosphorylation dynamics, controlled by kinases and phosphatases, enable rapid changes in cell physiology in response to environmental stimuli [16]. However, due to the often low abundance of regulatory phosphoproteins, enrichment and fractionation techniques, such as Immobilized Metal Ion Affinity Chromatography (IMAC) and Metal Oxide Affinity Chromatography (MOAC), are essential for reducing sample complexity and enhancing the detection of phosphorylated peptides [16]. Somatic embryogenesis is orchestrated by a complex regulatory network involving both transcriptional and post-translational mechanisms, offering a promising platform for micropropagation and genetic transformation in species such as *C. papaya*, where the optimization and mechanistic understanding of remain limited [17].

Given the complexity of the regulatory mechanisms involved and the pivotal role of phosphorylation in modulating embryogenic competence, the objective of this work is to elucidate protein phosphorylation networks that confer embryogenic competence to *C. papaya* callus.

Material and Methods

Plant material and callus induction

The culture medium for callus induction was prepared using Murashige and Skoog (MS) salts and vitamins (Phytotechnology Laboratories, Lenexa, KS, USA), supplemented with 30 g L⁻¹ sucrose (Vetec, São Paulo, Brazil) and different concentrations of 2,4-dichlorophenoxyacetic acid (2,4-D; Sigma-Aldrich, St. Louis, MO, USA), with the pH adjusted to 5.8 [18]. Subsequently, 2 g L⁻¹ Phytigel® (Sigma-Aldrich) was added, and the medium was autoclaved at 121 °C for 15 min and poured into 90 × 15 mm Petri dishes under laminar airflow. Mature fruits of papaya cv. 'Golden' were washed under running water, transferred to a laminar flow hood, and surface-sterilized by immersion in 70% ethanol (Sigma-Aldrich) for 1 min, followed by 30 min in 50% commercial bleach (Qboa®, Anhembí SA, Osasco, SP, Brazil) containing 1–1.25% sodium hypochlorite. The fruits were then rinsed three times with autoclaved distilled water and bisected using a sterile knife. Seeds were dissected with a sterile scalpel, and ten mature zygotic embryos were excised and inoculated onto the callus induction medium in Petri dishes. The explants were incubated in the dark at 25 ± 1 °C for up to 21 days. The embryogenic and non-embryogenic calli were submitted to three consecutive subcultures, with an interval of 21 days each, using the same culture medium and environmental conditions of the induction phase. Callus samples were collected after the third multiplication cycle.

Bottom-up proteomics

Five biological replicates, each consisting of a pool of five calli, were macerated in liquid nitrogen. From each replicate, 300 mg of powdered tissue was homogenized in an extraction buffer containing 7 M urea, 2 M thiourea, 2% Triton X-100, 1% dithiothreitol (DTT) (all from GE Healthcare), and 1 mM phenylmethanesulfonyl fluoride (PMSF; Sigma-Aldrich) [19]. The homogenate was vortexed thoroughly and then centrifuged to remove debris. The resulting supernatant was collected, and the total protein concentration was quantified using the 2-D Quant Kit (GE Healthcare). Tryptic digestion was performed using the filter-aided sample preparation (FASP) method as originally described by Wiśniewski *et al.* (2009), with adaptations from Almeida *et al.* 2020. A total of 1600 µg of protein from each biological replicate (embryogenic and non-embryogenic calli) was used. Due to the loading capacity of the filter units, each replicate was divided into four aliquots of 400 µg. Tryptic digestion was conducted separately for each aliquot, and the resulting peptides were subsequently pooled per biological replicate. The combined peptides were vacuum-dried and resuspended in

800 μ L of a solution composed of 5% (v/v) acetonitrile (Thermo Fisher Scientific) and 0.1% (v/v) formic acid in MS-grade water (Sigma-Aldrich). Peptide concentration was estimated by measuring absorbance at 205 nm using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific). Samples were stored at -80 °C until LC-MS/MS analysis. After digestion, the resulting peptides were vacuum-dried and resuspended in 50 μ L of a solution containing 95:5 (v/v) water:acetonitrile with 0.1% formic acid. Peptide concentrations were estimated by measuring absorbance at 205 nm using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific).

Phosphopeptide Enrichment

Phosphopeptides were enriched using the High-Select™ TiO₂ Phosphopeptide Enrichment Kit (Thermo Fisher Scientific) from trypsin-digested peptide samples. Aliquots of digested peptides from each biological replicate were vacuum-dried and resuspended in 150 μ L of a solution containing 5% (v/v) acetonitrile (Thermo Fisher Scientific) and 0.1% (v/v) formic acid in MS-grade water (Sigma-Aldrich). All enrichment steps were performed according to the manufacturer's instructions. After elution, phosphopeptides were vacuum-dried and resuspended in 25 μ L of a solution containing 5% (v/v) acetonitrile and 0.1% (v/v) formic acid in MS-grade water. The phosphopeptide concentration was estimated by measuring absorbance at 205 nm using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific). Samples were stored at -80 °C until analysis.

Sample analysis on the EvoSep timsTOF

From the remaining reconstituted samples were diluted to 20 μ L in 5% formic acid for loading onto EvoTip Pure tips (EvoSep, Odense, Denmark) according to the manufacturer's specifications, then stored in 0.1% formic acid until analysis. Samples were analyzed by an EvoSep One (EvoSep) LC system using four standard gradients: 100, 60, 30, and 15 SPD. Mobile phases consisted of 0.1% formic acid in water and 0.1% formic acid in ACN. Separation was performed on EvoSep EV1109 Performance columns (8 cm $\times$ 150 μ m, 1.5 μ m C18 packing) for the 100 and 60 SPD gradients and EvoSep 1137 Performance columns (15 cm $\times$ 150 μ m, 1.5 μ m C18 packing) for the 30 and 15 SPD gradients. The columns were coupled to a Bruker timsTOF Pro 2 (Bruker, Billerica, MA, USA) via a 20 μ m silica emitter (Bruker CaptiveSpray ZDV Sprayer 20) and held at 40°C in a Bruker Column Toaster.

MS/MS spectra obtained from the proteomics and phosphoproteomics data processing and database searching were performed using MaxQuant software v.1.6.3.4 (Cox and Mann, 2008), following the procedures described by Almeida *et al.* 2020. Label-free quantification was conducted using the normalized label-free quantification (LFQ intensity) approach, based on the MaxLFQ algorithm [20]. Following data processing, the output file Phospho (STY)Sites.txt was filtered to retain only phosphorylation sites with a localization probability ≥ 0.75 . For differential regulation analysis in the treatment/control comparison, only phosphoproteins containing confidently identified phosphate-modified peptides and present in at least four biological replicates, or entirely absent across all replicates in at least one condition (for unique proteins), were considered. Phosphoproteins exhibiting statistically significant differences according to a two-tailed Student's t-test ($P < 0.05$) were classified as differentially regulated phosphoproteins (DRPs).

To infer protein-protein interaction networks, Arabidopsis homologs of the sugarcane DAPs were identified via STRING, and the associated interaction data were used to construct interaction networks [21]. Prediction of kinases, transcription factors (TFs), and transcriptional regulators (TRs) among the DAPs and differentially regulated phosphoproteins (DRPs) was conducted using the iTAK software [22].

The raw and processed proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE repository under the dataset identifier PXXXXXXX [23].

Results

Embryogenic callus differentiation

Following the induction process using zygotic embryos as explants, two distinct types of callus were obtained: embryogenic and non-embryogenic (Figure 1). Embryogenic calli exhibited a characteristic dark brown coloration and friable texture, often containing somatic embryos at various developmental stages, even under continuous exposure to 2,4-D and in the absence of light. Upon transfer to a differentiation medium lacking 2,4-D and subjected to photoperiod conditions, these embryogenic calli successfully progressed toward maturation, leading to the formation of cotyledonary somatic embryos. In contrast, non-embryogenic calli

displayed a light yellow coloration and soft, watery consistency, with no evidence of somatic embryo development under the same culture conditions.

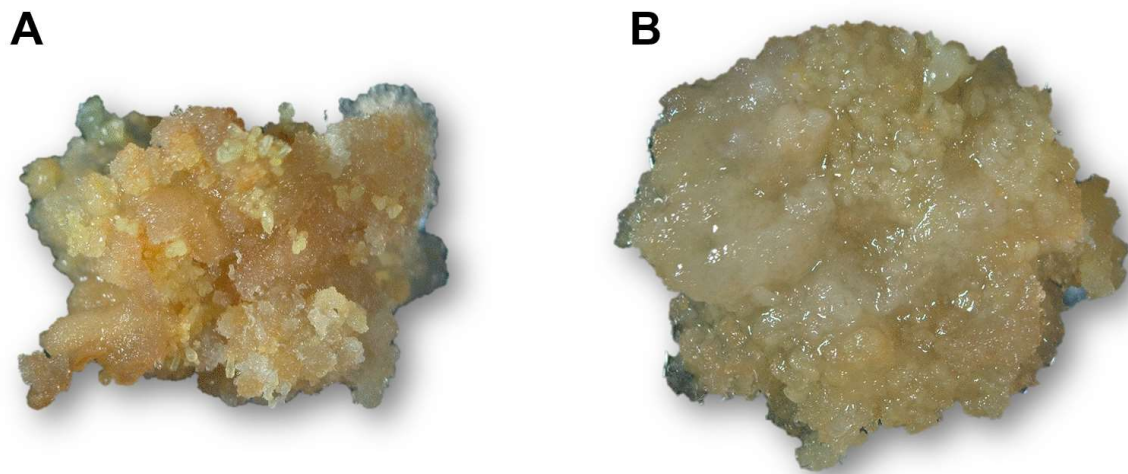


Figure 1. Morphological characterization of papaya callus. Calli were classified as embryogenic (A) or non-embryogenic (B) based on distinct morphological features. Embryogenic calli exhibit a dark brown coloration and friable texture, often associated with the presence of developing somatic embryos. In contrast, non-embryogenic calli display a light yellow color and a soft, watery consistency, with no visible signs of embryogenic structures.

To investigate the regulatory networks of proteins and phosphoproteins associated with embryogenic competence, callus samples were classified as embryogenic or non-embryogenic based on distinct morphological characteristics. These classified calli were subsequently used for total protein extraction, enabling comparative proteomic and phosphoproteomic analyses aimed at identifying key molecular signatures linked to somatic embryogenesis.

Proteomics and phosphoproteomics analysis

Proteomic analysis of papaya embryogenic and non-embryogenic calli resulted in the identification of 6,182 proteins (Supplementary Table 1). Principal component analysis (PCA) revealed clear clustering of biological replicates, with the first principal component explaining 79.9% of the total variance and effectively separating embryogenic from non-embryogenic samples. The second component accounted for an additional 6.5% of the variance, reflecting within-group biological variability (Figure 2A). Following data normalization, filtering, and statistical testing (p -value < 0.05), a total of 1,259 proteins were identified as differentially accumulated between the two callus types (Figure 2B), plus 576 uniques between embryogenic and non-embryogenic calli, suggesting widespread proteomic reprogramming associated with the acquisition of embryogenic competence.

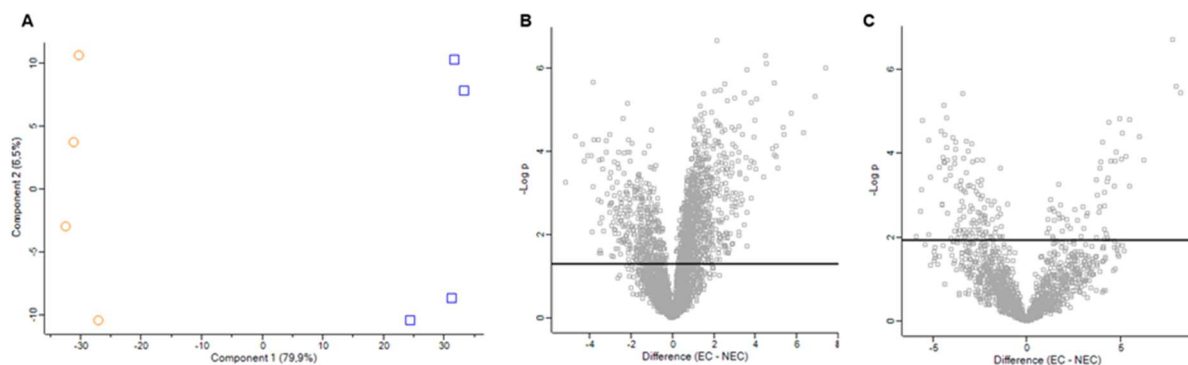


Figure 2. Overview of the proteomic and phosphoproteomic data obtained from the analysis of embryogenic and non-embryogenic papaya callus. Principal Component Analysis (PCA, A) demonstrates that biological replicates are grouped according to their respective treatments. The volcano plots display the differential accumulation and regulation of proteins (B) and phosphopeptides (C) in the comparative analysis between the two types of callus.

Phosphoproteomic profiling of embryogenic and non-embryogenic papaya callus identified a total of 4,169 phosphopeptides (2,475 with phosphorylation probability >0.75) (Supplementary Table 2). In this context, 210 phosphopeptides were differentially regulated and 344 were uniques between embryogenic and non-embryogenic callus.

Motif analysis, conducted using a 31-amino acid window centered on the phosphorylation site, revealed 48 distinct phosphorylation motifs (Table 1). Among these, the most frequently represented motifs included pSP, pSxS, pSxxS, pSxD, and pSDxE, which are commonly recognized by proline-directed kinases and other serine/threonine kinases involved in signal transduction and developmental regulation. The majority of phosphopeptides carried a single phosphorylated residue (1,229 peptides), while 908 contained two, 255 had three, 54 had four, and 15 peptides exhibited five phosphorylation sites (Figure 3A), highlighting the prevalence of multi-site phosphorylation in the regulation of protein function. Serine residues accounted for the vast majority of phosphorylation events (2,263 sites), followed by threonine (173 sites) and tyrosine (25 sites), consistent with typical plant phosphoproteome profiles (Figure 3B). Following stringent quality control and statistical filtering, 210 phosphoprotein isoforms (proteoforms) were identified as differentially regulated between embryogenic and non-embryogenic callus samples (Figure 2C), suggesting their potential involvement in the molecular reprogramming underlying somatic embryogenesis.

Table 1. Phosphorylation motifs identified in the phosphoproteomics dataset. The search for phosphorylation motifs was conducted using all the phosphopeptides identified in the proteomic analysis of embryogenic and non-embryogenic papaya callus, with a minimum occurrence threshold of five.

Phosphorylation motif	Amino acid	Score	Number of phosphopeptides	Adjusted p-value
.....SP.....	S	86.33	807	1.1e-093
.....S.S.....	S	8.32	656	4.4e-016
.....S.S.....	S	7.40	576	7.5e-006
.....S.S.....	S	7.45	458	9.8e-001
.....S.D.....	S	10.65	330	2.4e-002
.....SD.E.....	S	38.78	136	7.8e-018
.....S.S.S.....	S	19.96	120	1.1e-005
.....TP.....	T	13.23	115	3.9e-013
.....P.SP.....	S	122.71	108	7.3e-015
.....SSP.....	S	38.60	103	7.2e-010
.....S.SS.....	S	22.13	94	2.9e-003
..R.....SP.....	S	144.11	79	8.0e-012
.....S.S.E.....	S	30.38	73	1.5e-004
.....S.S..D.....	S	40.01	60	8.9e-003
.....L.R.S.....	S	26.65	59	3.0e-004
.....LSP.....	S	197.16	57	3.9e-008
.....V.....T.....	T	6.43	35	1.6e-001
.....SDEE.....	S	73.25	32	4.5e-005
.....F..S.S.....	S	51.10	31	2.2e-002
.....DSD.D.....	S	34.09	30	1.6e-005
.....K.....T.....	T	7.34	29	9.8e-001
.....I..SP.....	S	163.95	28	4.7e-004
.....F.....SP.....	S	251.48	26	2.3e-004
.....SP.H.....	S	238.05	26	2.3e-004
.....H.....SP.....	S	223.95	26	2.3e-004
.....SP.....H.....	S	177.84	26	1.7e-003
.....V...GS.....	S	15.34	25	4.9e-002
.....SD.E...E.....	S	86.89	24	1.1e-003
.....SDSD.....	S	35.52	24	6.3e-003
.....R.RS.S.....	S	34.09	23	1.7e-003
.....Y...SP.....	S	98.25	22	8.1e-002
.....F.....SP.....	S	207.28	21	6.1e-003
.....Y..SP.....	S	152.60	20	1.2e-002
.....TP...S.....	T	45.49	20	6.8e-003
M.....S.S.....	S	43.26	19	1.1e-001
.....RS.S..P.....	S	37.03	19	1.3e-001
.....SEDE.....	S	26.76	18	4.0e-002
.....T.....GS.....	S	20.14	17	1.0e+000
.....TP...S.....	T	33.30	16	8.8e-002
.....TP.....T.....	T	25.23	15	1.5e-001
.....GS.Y.....	S	17.38	12	1.0e+000
.....Y.....A.....	Y	6.25	9	4.2e-001

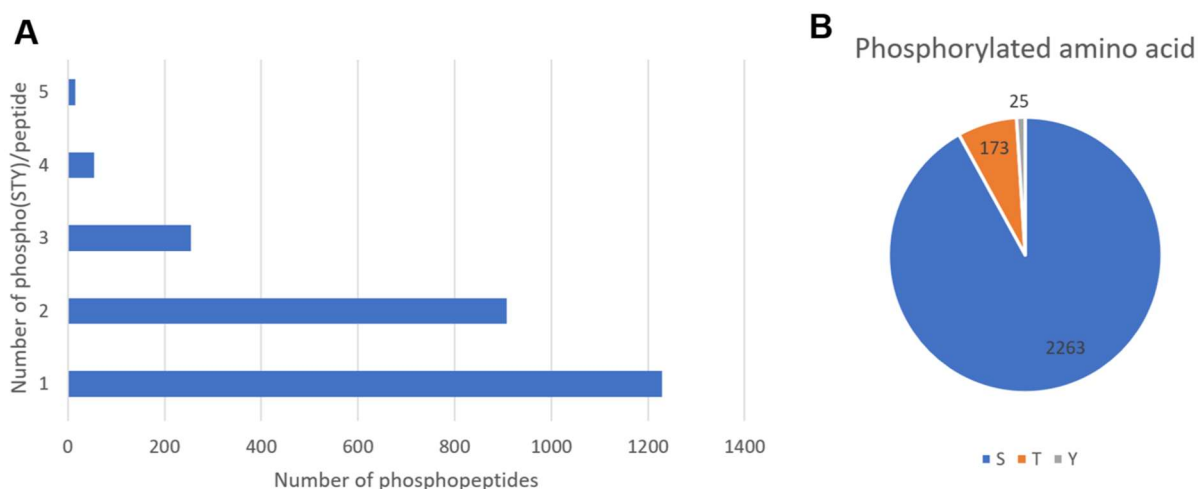


Figure 3. Bar graph showing the distribution of phosphorylation motifs identified in phosphopeptides from embryogenic and non-embryogenic callus, highlighting features of the phosphorylation sites identified in the phosphoproteomic analysis. The bar graph (A) illustrates the number of simultaneous phosphorylation sites in the phosphopeptides, while the pie chart (B) displays the proportion of phosphorylated serine, threonine, and tyrosine residues (S, T, and Y, respectively).

Functional analysis of proteomic and phosphoproteomics data

Functional enrichment analysis of differentially accumulated proteins (DAPs) revealed a strong association with metabolic biological processes, particularly those involving carboxylic acids, oxoacids, small molecules, amino acids, and starch metabolism (Figure 4A). These findings highlight the metabolic reprogramming that accompanies the acquisition of embryogenic competence in papaya callus. In terms of molecular function, the DAPs were significantly enriched for RNA binding, carbohydrate lyase activity, general catalytic activity, and NAD binding (Figure 4B). Subcellular localization analysis showed that the majority of DAPs were associated with the cytosol, followed by the vacuole, cell wall, chloroplast, and other plastid-related compartments (Figure 4C). Moreover, pathway enrichment analysis indicated that these DAPs participate in key metabolic routes, including the biosynthesis of secondary metabolites, amino acid and carbohydrate metabolism, and carbon fixation pathways (Figure 4D).

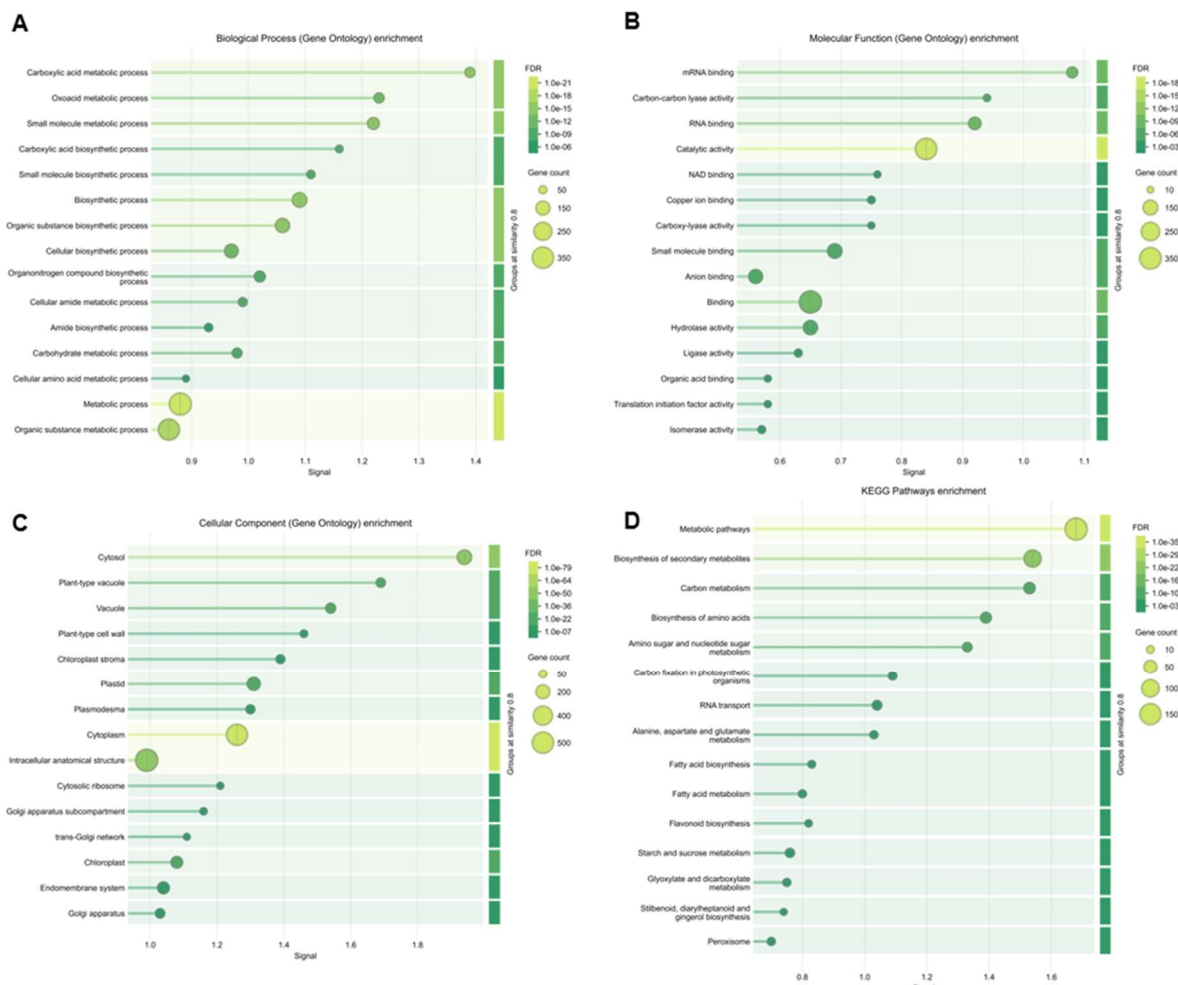


Figure 4. Functional enrichment analysis of differentially accumulated proteins between embryogenic and non-embryogenic papaya callus. Lollipop plots illustrate the enriched gene ontology (GO) categories and KEGG pathways associated with DAPs. Enriched GO terms are presented for biological processes (A), molecular functions (B), and cellular components (C), while panel (D) highlights the metabolic pathways represented.

Gene ontology enrichment analysis of differentially regulated phosphoproteins (DRPs) revealed significant enrichment in biological processes related to localization, intracellular and intercellular transport, vesicle-mediated trafficking, general cellular processes, and RNA splicing (Figure 5A). In terms of molecular function, DRPs were enriched for kinase activity, phosphotransferase activity, and nucleotide binding (Figure 5B), underscoring the relevance of phosphorylation-dependent signaling cascades regulatory mechanisms in this developmental context. Subcellular localization predictions indicated that DRPs are predominantly associated with the cytosol, nucleus, plasmodesmata, endomembrane systems, and general cytoplasmic compartments (Figure 5C), consistent with roles in signal transduction, transcriptional regulation, and cellular communication. Also, pathway enrichment

analysis identified two major pathways: the spliceosome and protein processing in the endoplasmic reticulum (Figure 5D).

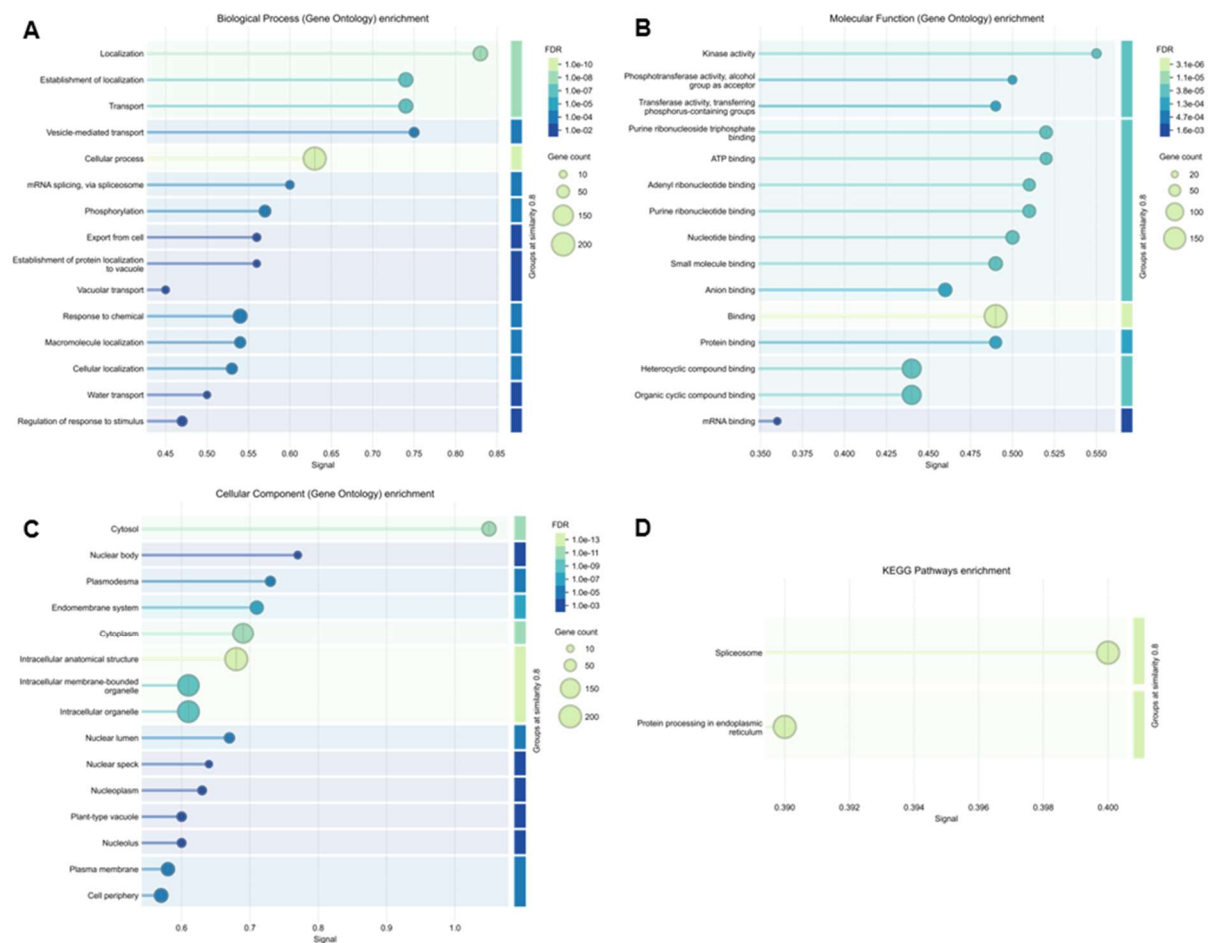


Figure 5. Enrichment analysis of gene ontology (GO) terms and KEGG pathways among differentially regulated phosphoproteins between embryogenic and non-embryogenic papaya callus. Lollipop plots summarize the GO enrichment results for biological processes (A), molecular functions (B), and cellular components (C), as well as enriched KEGG pathways (D).

Identification of regulatory proteins

To gain insight into the regulatory networks underpinning somatic embryogenesis in papaya, we conducted a comprehensive search for key transcriptional and post-translational regulators. This analysis revealed 158 transcription factors and transcriptional regulators distributed across 56 distinct families (Supplementary Table 3). Among these, the most represented classes were C3H zinc finger proteins (20 members), GNAT family histone acetyltransferases (12), basic leucine zipper (bZIP) factors (9), SET domain-containing methyltransferases (7), basic helix-loop-helix (bHLH) proteins (6), SWI/SNF-BAF60b

chromatin remodelers (6), C2H2 zinc fingers (5), and Jumonji domain-containing histone demethylases (5).

In parallel, we identified 177 protein kinases classified into 16 kinase groups (Supplementary Table 4). The receptor-like kinase (RLK)-Pelle family was the most abundant, with 80 members, followed by calcium/calmodulin-dependent protein kinases (CAMKs, 26), CMGC kinases, including CDKs and MAPKs (25), STE kinases (10), and members of the tyrosine kinase-like families TKL-PI-4 (6) and TKL (5). These kinases are central to signal transduction cascades that mediate environmental and developmental cues, potentially linking extracellular signals to the intracellular machinery that governs embryogenic competence and progression.

Predicted protein-protein interaction networks

Following the functional annotation and integration of proteomic and phosphoproteomic datasets, we performed an interaction network analysis to identify high confidence (>0.7) predicted protein-protein interactions between DAPs and DRPs in embryogenic and non-embryogenic papaya callus (Figure 6). As a result, we've identified 1741 nodes and 6575 edges in the complete network. By combining the predicted interaction data with functional annotations and evidence from the literature, we compiled a list of candidate phosphoproteins potentially involved in regulating embryogenic competence in papaya (Table 2).

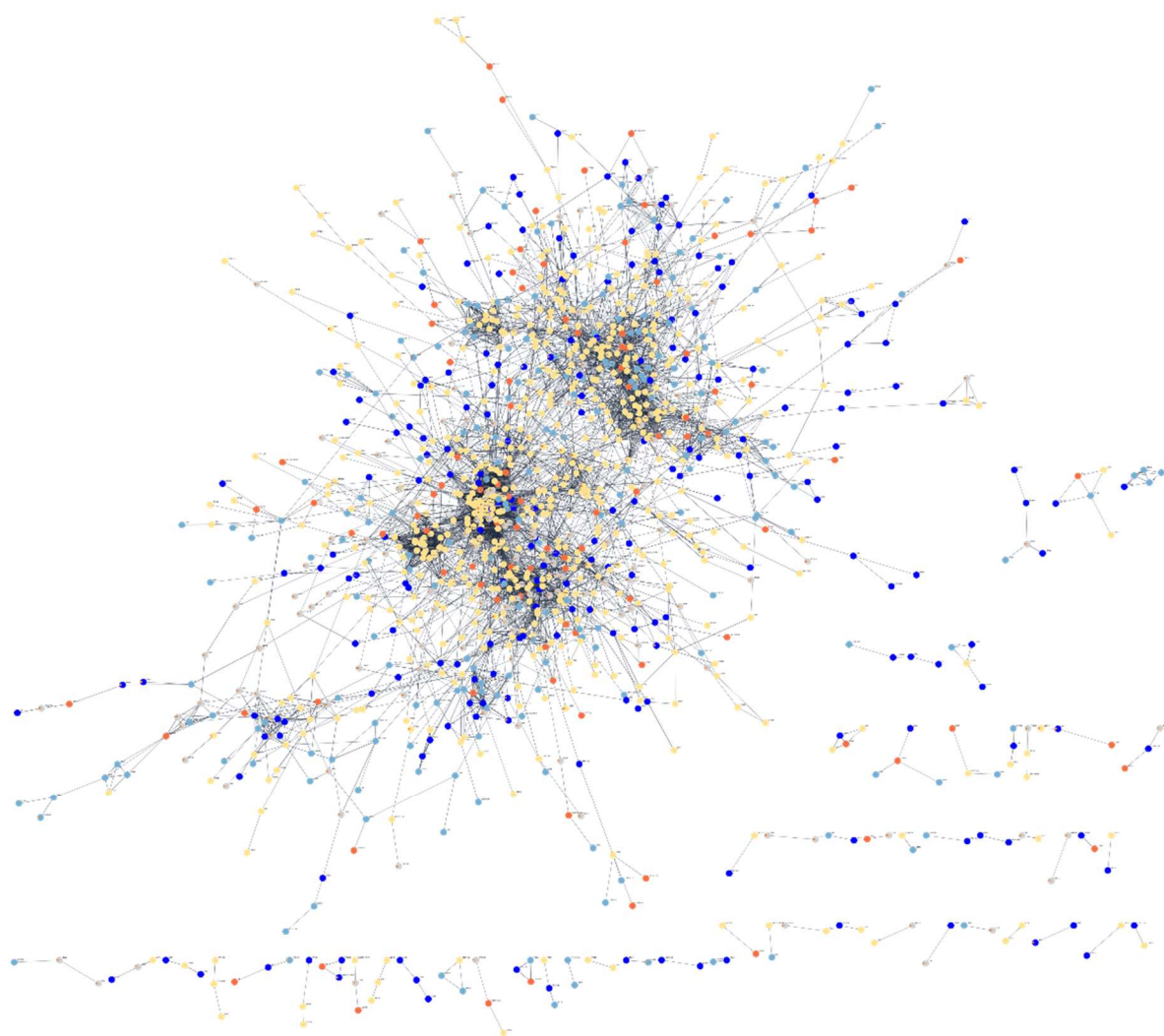


Figure 6. Prediction of interactions between regulated proteins and phosphoproteins in the comparative analysis of embryogenic and non-embryogenic callus. The orange and yellow nodes represent up-regulated and unique proteins in embryogenic calli, while the light and dark blue nodes correspond to down-regulated or unique proteins in non-embryogenic callus. Phosphoproteins with regulated phosphorylation sites are marked with a 'P' inside the node.

Table 2. Phosphopeptides associated with the embryogenic competence of papaya callus. The table shows the corresponding *Arabidopsis* orthologs, the phosphorylated amino acid (STY), the name of the BLAST top hit using the *Viridiplantae* database.

Protein ID	BLAST top hit	Arabidopsis orthologs	Phosphopeptide	DRP
evm.model.supercontig_14.14	ABC transporter B family member 19	ABCB19	NL#SYSY#STGADGR	Unique EC
evm.TU.contig_35483.1	auxin efflux carrier component 1-like	PIN1	N#SNFGANDVYGLSA#SR	Unique EC
evm.TU.contig_35483.1	auxin efflux carrier component 1-like	PIN1	N#SNFGANDVYGLSASR	Unique EC
evm.model.supercontig_56.23	cell division cycle 48 homolog	CDC48D	FAE#SG#SGGAAGSDPFAAPAGG AEEDDLYS	Unique NEC
evm.model.supercontig_13.126	SNF1-related protein kinase regulatory subunit gamma-1-like	KING1	LGMQVEDLWDTQEPQL#SPTDK	Unique NEC
evm.model.supercontig_117.89	V-type proton ATPase subunit F isoform X1	VHA-F	YLFSAE#SVASGR	Unique NEC
evm.model.supercontig_26.59	protein argonaute 4	AGO4	NNGLG#SPGGDA#SPNETDR	Yes
evm.model.supercontig_5.284	probable AMP deaminase isoform X1	AMPD	QLLDGGEDE#SEYEER	Yes
evm.TU.contig_31534.1	serine/threonine-protein kinase BSK1	BSK1	DDEGTNEL#SFQEWTTQMMR	Yes
evm.TU.contig_31534.1	serine/threonine-protein kinase BSK1	BSK1	SY#STNLAYTPPEYLK	Yes
evm.model.supercontig_103.18	BRI1 suppressor 1 (BSU1)-like 2 isoform 1	BSL3	QL#SIDQFENEGR	Yes
evm.model.supercontig_56.23	cell division cycle 48 homolog	CDC48D	FAESGSGGAAG#SDPFAAPAGGA EEDDLYS	Yes
evm.model.supercontig_292.1	probable indole-3-acetic acid-amido synthetase GH3.1	GH3.1	AVDSAIS#SPLGPSPACEK	Yes
evm.model.supercontig_6.74	Auxin-responsive GH3 family protein	GH3.1	AVDSALS#SPLGPACER	Yes
evm.model.supercontig_70.16	S-adenosylmethionine decarboxylase proenzyme	SAMDC2	MDLPVSAIGFEG#YEKR	Yes
evm.model.supercontig_107.10	probable transcriptional regulator SLK2	SLK2	SA#SINTESYMR	Yes
evm.model.supercontig_95.58	mitochondrial outer membrane protein porin 4	VDAC4	GS#SPAPFSDIGK	Yes
evm.model.supercontig_28.63	protein CTR9 homolog	VIP6	DN#STELQDNDAGVR	Yes
evm.model.supercontig_1673.2	argonaute 1	AGO1	FYMEPE#T#SD#SGSMTSGTIAAR	Unique EC
evm.model.supercontig_18.242	serine/threonine-protein kinase BSK3	BSK3	AAVLEAPDVENDEN#SEV#SDVP AFR	Unique EC
evm.model.supercontig_1624.1	protein HIRA	HIRA	VLG#SP#SRDGIVEQPGSTK	Unique EC
evm.model.supercontig_1624.1	protein HIRA	HIRA	VLG#SP#SRDGIVEQPGSTK	Unique EC
evm.model.supercontig_14.14	ABC transporter B family member 19	ABCB19	NL#SYSY#STGADGR	Yes
evm.model.supercontig_26.59	protein argonaute 4	AGO4	NNGLG#SPGGDA#SPNETDR	Yes
evm.model.supercontig_11.84	AMP deaminase isoform X1	AMPD	EREEN#SDGD#SPQHLKK	Yes
evm.model.supercontig_11.84	AMP deaminase isoform X1	AMPD	EREEN#SDGD#SPQHLKK	Yes
evm.model.supercontig_4.49	Lysine-specific demethylase REF6	REF6	YVEKEDAV#SCD#SLEQNTR	Yes
evm.model.supercontig_65.68	Receptor protein kinase-like protein	ZAR1	LANIAGG#SPTLQ#SNR	Yes
evm.model.supercontig_65.68	Receptor protein kinase-like protein	ZAR1	LANIAGG#SPTLQ#SNR	Yes
evm.model.supercontig_92.84	probable auxin efflux carrier component 1b	PIN1	PSNLTNAEIY#SLQ#S#SRNP#TP R	Unique EC
evm.model.supercontig_92.84	probable auxin efflux carrier component 1b	PIN1	PSNLTNAEIY#SLQ#S#SRNP#TP R	Unique EC
evm.model.supercontig_26.59	protein argonaute 4	AGO4	FEDA#SETSSSHGGTTTAGAVAV PQLPR	Yes

The identification of proteins and phosphoproteins potentially involved in the embryogenic competence of papaya callus enabled us to predict interactions that may be crucial to this process (Figure 7). This interaction network highlights proteins associated with DNA methylation (such as SAM1 and SAM2), auxin transport and cell differentiation (including ABCB19, PIN1, and PLT2), proton transport (VHA-1, VHA-E1, and VHA-F), and epigenetic regulation (AGO1, AGO4, and HST1).

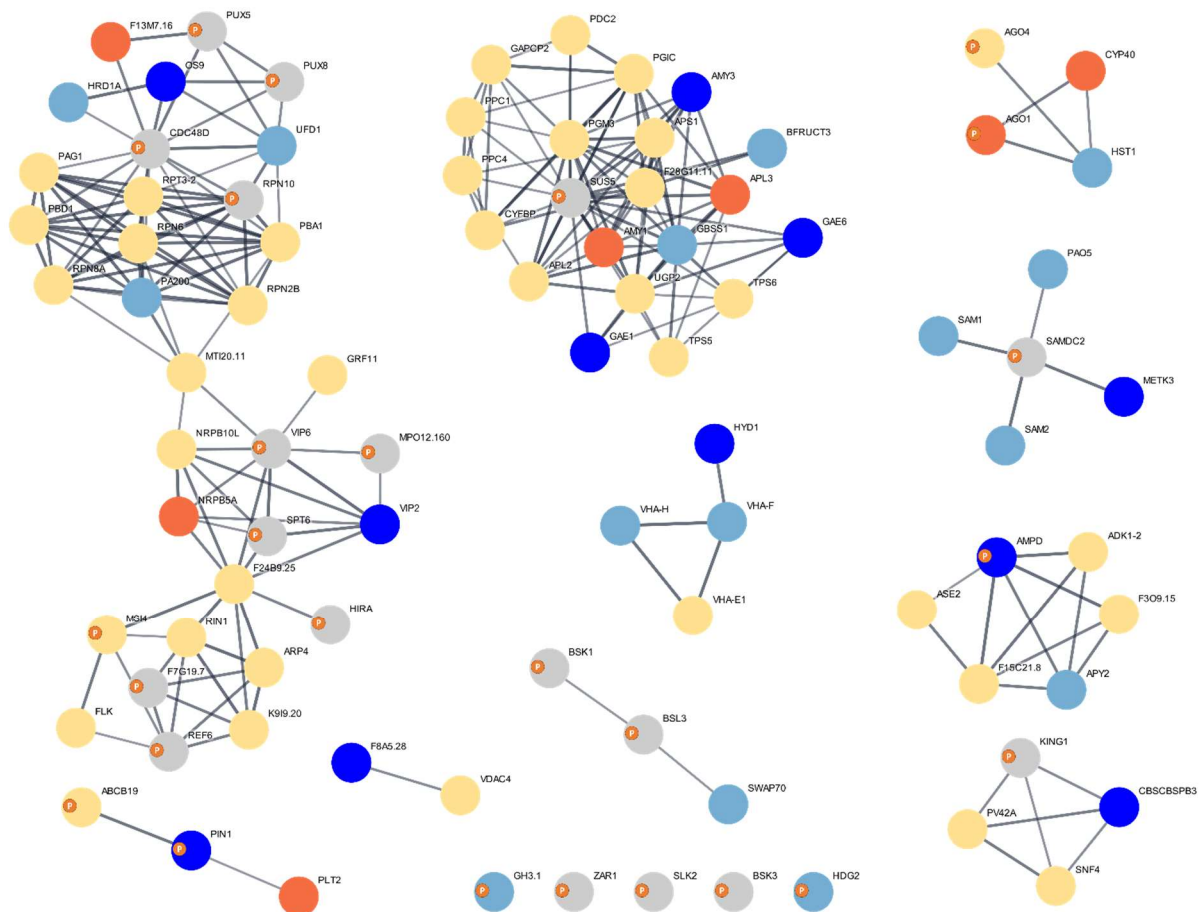


Figure 7. Subnetwork of protein-protein interactions representing targets associated with the embryogenic competence of calluses identified in the present study, along with their first neighbors in the complete network of DAPs and DRPs. The orange and yellow nodes represent up-regulated and unique proteins in embryogenic calli, while the light and dark blue nodes correspond to down-regulated or unique proteins in non-embryogenic callus. Phosphoproteins with regulated phosphorylation sites are marked with a 'P' inside the node.

Discussion

Phosphorylation-Dependent Regulation of Auxin Transport and Signaling in Somatic Embryogenesis

Auxin is a central regulator of plant development, governing various physiological and morphogenetic processes through its biosynthesis, signaling, and directional transport. The establishment of auxin gradients, essential for cell fate specification and organogenesis, is mediated by ATP-binding cassette (ABC) transporters and PIN-FORMED (PIN) efflux carriers [24]. In our phosphoproteomic analysis, papaya calli with enhanced embryogenic competence showed differential phosphorylation of key auxin transporters. The ABC transporter B family member 19 protein (evm.model.supercontig_14.14), an ortholog of *Arabidopsis thaliana* ABCB19, was significantly phosphorylated at S640 and S644. Similarly, auxin efflux carrier components (PIN1-like proteins), such as evm.TU.contig_35483.1 and evm.model.supercontig_92.84, were modulated at S282, S269, S244, and S252, respectively. ABCB19 is known to promote hypocotyl elongation through interaction with PIN1 and to negatively regulate light signaling, while PIN1 polar localization is necessary for auxin efflux and is tightly regulated by phosphorylation of its hydrophilic loop domains [25, 26]. Importantly, the spatial coordination between ABCB19 and PIN1 ensures auxin flow and stability within membrane microdomains, and phosphorylation of ABCB19 has been associated with reduced transporter activity, potentially leading to local auxin accumulation [27, 28]. This suggests a role for phospho-regulation of auxin transporters in the induction of somatic embryogenesis, possibly by establishing local auxin maxima that initiate cell reprogramming.

Auxin signaling is also modulated at the transcriptional level. The transcriptional regulator SLK2 (evm.model.supercontig_107.10), orthologous to *AtSLK2*, was phosphorylated at S151. SLK2, along with SLK1, integrates into the SEUSS–LEUNIG (SEU–LUG) corepressor complex, which modulates developmental gene expression, particularly during reproductive organ and meristem development [29]. This complex influences histone acetylation at H3K9 and H3K14, thereby altering chromatin accessibility and transcriptional activity [30]. SLK proteins also contribute to auxin gradient responses and shoot apical meristem maintenance [31]. In addition, two GH3 proteins involved in auxin homeostasis, GH3.1-like synthetases (evm.model.supercontig_292.1 and evm.model.supercontig_6.74, both orthologous to *AtGH3.1*), were differentially phosphorylated at S9. These enzymes conjugate IAA with amino acids, reducing free auxin levels and modulating hormonal responses. GH3 expression in papaya has been associated with embryogenic potential, particularly in response to synthetic auxins like 2,4-D [3]. Collectively, these data suggest that transcriptional regulators and hormone homeostasis enzymes, both modulated via phosphorylation, coordinate auxin signaling during the establishment of embryogenic competence.

Phosphorylation of Epigenetic Regulators in the Maintenance of Embryogenic Competence

Epigenetic modifications, including histone methylation/acetylation and DNA methylation, are essential for somatic embryo initiation and maintenance [10, 11]. Our study identified multiple epigenetic regulators associated with embryogenic competence. ARGONAUTE proteins, which function in RNA-directed gene silencing, were differentially phosphorylated. AGO1 (evm.model.supercontig_1673.2, orthologous to AtAGO1) was phosphorylated at S597, while AGO4 (evm.model.supercontig_26.59, orthologous to AtAGO4) was phosphorylated at S168, S174, and S889. AGO1 is involved in microRNA-mediated post-transcriptional gene silencing, while AGO4 mediates RNA-directed DNA methylation, both critical for meristem and embryo development [32, 33].

Additional chromatin remodelers were identified, including HIRA (evm.model.supercontig_1624.1, ortholog of AtHIRA), phosphorylated at S532 and S534. HIRA functions as a histone chaperone, regulating chromatin accessibility and developmental transitions [34, 35]. The lysine-specific histone demethylase REF6 (evm.model.supercontig_4.49, ortholog of AtREF6) was phosphorylated at S1038. REF6 modulates expression of auxin efflux-related genes such as PIN1, PIN3, and PIN7 [36]. Additionally, the CTR9/VIP6 homolog (evm.model.supercontig_28.63) was upregulated at S759 and is known to mediate transcriptional control of developmental phase transitions [37]. These findings indicate that epigenetic regulators, through phosphorylation and protein accumulation, modulate chromatin dynamics and gene expression programs that sustain embryogenic potential.

Regulation of Metabolic and Developmental Pathways by Protein Kinases

Protein kinases are integral to signal transduction during somatic embryogenesis, with SERK1 being a well-known marker of embryogenic competence [11, 14]. In papaya, the SNF1-related protein kinase regulatory subunit gamma-1-like (evm.model.supercontig_13.126, orthologous to AtKING1) was phosphorylated at S47. KING1 regulates energy homeostasis and interacts with SNF complex members such as PV42A and SNF4, both of which accumulated in embryogenic calli [38]. KING1 has previously been associated with embryogenic potential in sugarcane [19]. Furthermore, the receptor-like kinase ZAR1 (evm.model.supercontig_65.68, ortholog of AtZAR1) was phosphorylated at S139 and S144. ZAR1 is essential for apical-basal polarity establishment during early embryogenesis, where loss-of-function leads to developmental arrest [39, 40]. These kinases

are likely pivotal for both metabolic regulation and developmental patterning in embryogenic calli.

Metabolic Reprogramming of Energy Homeostasis During Somatic Embryogenesis

Energy metabolism is tightly regulated during callus proliferation and embryogenesis. The V-type proton ATPase subunit F isoform X1 (evm.model.supercontig_117.89, ortholog of AtVHA-F) was phosphorylated at S124 and is predicted to interact with other V-ATPase subunits, including VHA-H and VHA-E1. Similarly, the mitochondrial outer membrane protein porin 4 (evm.model.supercontig_95.58, ortholog of AtVDAC4), phosphorylated at S4, is involved in membrane potential regulation and is associated with embryogenic potential in plants [18, 41, 42]. Enzymes involved in purine metabolism, such as AMP deaminase isoforms (evm.model.supercontig_11.84 and evm.model.supercontig_5.284), were phosphorylated at S51, S55, and S84. These isoforms, orthologous to AtAMPD, are linked to ATP production and are essential for embryogenesis, interacting with proteins involved in purine biosynthesis [43]. These results highlight a metabolic shift in embryogenic calli, driven by coordinated regulation of mitochondrial and cytosolic energy-related proteins.

Lastly, we identified regulation of polyamine biosynthesis enzymes, notably the S-adenosylmethionine decarboxylase proenzyme (evm.model.supercontig_70.16, ortholog of AtSAMDC2), phosphorylated at Y13. SAMDC catalyzes a rate-limiting step in polyamine synthesis and is implicated in several developmental processes, including embryogenesis and cell division [44]. The involvement of SAMDC suggests a role for polyamines in developmental reprogramming and callus competence, aligning with the broader regulatory network controlling somatic embryo formation.

CpPLT2 as a marker of embryogenic competence in papaya callus

Together with BABY BOOM (BBM, also known as PLETHORA1, PLT1), the transcription factors PLT1 and PLT2 are known to be induced by auxin accumulation and play pivotal roles in specifying the identity of totipotent plant cells during embryogenesis [45]. In the present study, PLT2 was uniquely identified in embryogenic calluses of *Carica papaya*, suggesting a specific role in the acquisition or maintenance of embryogenic competence. Protein–protein interaction predictions revealed putative interactions between PLT2 and key auxin efflux carriers, including ABCB19 and PIN1, highlighting its possible involvement in auxin transport regulation. Phylogenetic analyses support the close evolutionary relationship between BBM and PLT2, implying potential functional redundancy in controlling auxin-

responsive developmental processes such as somatic embryogenesis and root formation [46]. Supporting this notion, the incorporation of a single BBM-derived domain (bbm-1: SLGLSMIKTWLRNQP) into the N-terminal region of PLT2 has been shown to enhance somatic embryo formation upon ectopic PLT2 expression [47]. Importantly, functional genomics studies indicate that PLT2 controls the LAFL gene network (LEC1, ABI3, FUS3 and LEC2) of embryogenic identity and has redundant functions with BBM related to somatic embryogenesis [48]. Given the well-established biotechnological applications of BBM in improving plant regeneration and large-scale micropropagation, it is plausible that PLT2 may also hold similar potential. Accordingly, we propose that PLT2 functions as a molecular marker of embryogenic competence in *C. papaya* calluses. Future studies should explore its utility in the development of genetic transformation protocols and in vitro propagation systems for this and other recalcitrant species.

Conclusions

This study provides novel insights into the phosphorylation-dependent regulation of auxin transport, signaling, and epigenetic remodeling during the acquisition and maintenance of embryogenic competence in *Carica papaya* callus. Through a comprehensive phosphoproteomic analysis, we identified multiple phosphorylated proteins involved in auxin efflux (ABCB19, PIN1-like), transcriptional repression (SLK2), hormone homeostasis (GH3.1-like), chromatin remodeling (HIRA, REF6, CTR9/VIP6), RNA-directed DNA methylation (AGO1, AGO4), and energy metabolism (VHA-F, VDAC4, AMPD), all of which are functionally associated with somatic embryo formation.

The differential phosphorylation of auxin transporters such as ABCB19 and PIN1 supports the hypothesis that phospho-modulation fine-tunes auxin fluxes and the formation of local auxin maxima, which are crucial for cell fate transitions and totipotency. Similarly, the phosphorylation of SLK2 and GH3 family members suggests that transcriptional repression and hormonal homeostasis are also dynamically regulated to coordinate auxin responses during the induction phase. Moreover, the identification of phosphorylated epigenetic regulators underscores the importance of chromatin accessibility and transcriptional plasticity in sustaining embryogenic potential, while the involvement of energy metabolism enzymes and kinases such as KING1 and ZAR1 suggests a metabolic reprogramming tightly linked to developmental reprogramming.

Importantly, the detection of the PLT2 transcription factor exclusively in embryogenic calli, alongside its predicted interactions with ABCB19 and PIN1, positions it as a promising

molecular marker of embryogenic competence in papaya. Given the well-established roles of PLT family members in auxin-mediated patterning and the induction of totipotent cell identities, PLT2 may serve not only as a diagnostic marker but also as a potential target for genetic improvement of in vitro regeneration protocols in *C. papaya* and other recalcitrant species. Together, our findings establish a functional link between phosphorylation events and the complex regulatory networks controlling somatic embryogenesis. This work contributes to a deeper understanding of the molecular determinants of embryogenic competence and lays the groundwork for developing more efficient biotechnological strategies in tropical fruit crops.

References

1. Pereira, M.G., *et al.*, Melhoramento genético do mamoeiro: programa Uenf/Caliman. VII Simpósio do papaya brasileiro: produção e sustentabilidade hídrica. Incaper, Vitória, 2018: p. 1-59.
2. Salinas, I., *et al.*, Assessment of two sex-determining procedures in 'BH-65' papaya from an economical and developmental point of view. *Fruits*, 2018. 73(3): p. 184-190.
3. Vale, E.M., *et al.*, Morphological analyses and variation in carbohydrate content during the maturation of somatic embryos of *Carica papaya*. *Physiology and molecular biology of plants*, 2018. 24(2): p. 295-305.
4. Botini, N., *et al.*, Stage-specific protein regulation during somatic embryo development of *Carica papaya* L. 'Golden'. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 2021. 1869(2): p. 140561.
5. Xavier, L., *et al.*, Large-scale regeneration of hermaphrodite emblings of *Carica papaya* L. 'Golden' using early molecular sex determination during embryogenic callus multiplication. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 2021. 146(3): p. 643-649.
6. Atkins, P.A.P. and D.F. Voytas, Overcoming bottlenecks in plant gene editing. *Current Opinion in Plant Biology*, 2020. 54: p. 79-84.
7. Lowe, K., *et al.*, Morphogenic Regulators Baby boom and Wuschel Improve Monocot Transformation. *The Plant Cell*, 2016. 28(9): p. 1998-2015.
8. Chen, J., *et al.*, Significant improvement of apple (*Malus domestica* Borkh.) transgenic plant production by pre-transformation with a BABY BOOM transcription factor. *Horticulture Research*, 2022. 9.
9. Li, M., *et al.*, Auxin biosynthesis maintains embryo identity and growth during BABY BOOM-induced somatic embryogenesis. *Plant Physiology*, 2021. 188(2): p. 1095-1110.
10. Wójcik, A.M., B. Wójcikowska, and M.D. Gaj, Current perspectives on the auxin-mediated genetic network that controls the induction of somatic embryogenesis in plants. 2020. 21(4): p. 1333.
11. Fehér, A., A common molecular signature indicates the pre-meristematic state of plant calli. *International Journal of Molecular Sciences*, 2023. 24(17): p. 13122.

12. Su, Y.H., *et al.*, Establishment of embryonic shoot–root axis is involved in auxin and cytokinin response during Arabidopsis somatic embryogenesis. *Frontiers in Plant Science*, 2015. 5.
13. Swarup, R. and B. Péret, AUX/LAX family of auxin influx carriers, an overview. *Frontiers in Plant Science*, 2012. Volume 3 - 2012.
14. Fehér, A., Somatic embryogenesis, Stress-induced remodeling of plant cell fate. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 2015. 1849(4): p. 385-402.
15. Guo, H., *et al.*, Large-scale label-free phosphoproteomics: From technology to data interpretation. *Bioanalysis*, 2014. 6(18): p. 2403-2420.
16. Picotti, P., Phosphoproteomics takes it easy. *Nature Biotechnology*, 2015. 33(9): p. 929-930.
17. Heringer, A.S., C. Santa-Catarina, and V. Silveira, Insights from proteomic studies into plant somatic embryogenesis. *Proteomics*, 2018. 18(5-6): p. e1700265.
18. Xavier, L.R., *et al.*, Time-dependent proteomic signatures associated with embryogenic callus induction in *Carica papaya* L. *Plants*, 2023. 12(22): p. 3891.
19. Almeida, F.A., *et al.*, Label-free quantitative phosphoproteomics reveals signaling dynamics involved in embryogenic competence acquisition in sugarcane. *Journal of Proteome Research*, 2020. 19(10): p. 4145-4157.
20. Tyanova, S., T. Temu, and J. Cox, The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols*, 2016. 11(12): p. 2301-2319.
21. Szklarczyk, D., *et al.*, The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*, 2022. 51(D1): p. D638-D646.
22. Zheng, Y., *et al.*, iTAK: A program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Molecular Plant*, 2016. 9(12): p. 1667-1670.
23. Perez-Riverol, Y., *et al.*, The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res*, 2019. 47(D1): p. D442-D450.
24. Cecchetti, V., *et al.*, ABCB1 and ABCB19 auxin transporters have synergistic effects on early and late Arabidopsis anther development. *Journal of Integrative Plant Biology*, 2015. 57(12): p. 1089-1098.
25. Noh, B., A.S. Murphy, and E.P. Spalding, Multidrug resistance-like genes of Arabidopsis required for auxin transport and auxin-mediated development. *Plant Cell*, 2001. 13(11): p. 2441-54.
26. Huang, F., *et al.*, Phosphorylation of conserved PIN motifs directs Arabidopsis PIN1 polarity and auxin transport. *The Plant Cell*, 2010. 22(4): p. 1129-1142.
27. Titapiwatanakun, B., *et al.*, ABCB19/PGP19 stabilises PIN1 in membrane microdomains in Arabidopsis. *The Plant Journal*, 2009. 57(1): p. 27-44.

28. Christie, J.M., *et al.*, phot1 inhibition of ABCB19 primes lateral auxin fluxes in the shoot apex required for phototropism. *PLOS Biology*, 2011. 9(6): p. e1001076.
29. Lee, J.E., *et al.*, SEUSS and SEUSS-LIKE 2 coordinate auxin distribution and KNOXI activity during embryogenesis. *The Plant Journal*, 2014. 80(1): p. 122-135.
30. Shrestha, B., B. Guragain, and V.V. Sridhar, Involvement of co-repressor LUH and the adapter proteins SLK1 and SLK2 in the regulation of abiotic stress response genes in *Arabidopsis*. *BMC Plant Biology*, 2014. 14(1): p. 54.
31. Bao, F., S. Azhakanandam, and R.G. Franks, SEUSS and SEUSS-LIKE transcriptional adaptors regulate floral and embryonic development in *Arabidopsis*. *Plant Physiology*, 2009. 152(2): p. 821-836.
32. Kidner, C.A. and R.A. Martienssen, The role of ARGONAUTE1 (AGO1) in meristem formation and identity. *Developmental Biology*, 2005. 280(2): p. 504-517.
33. Du, F., *et al.*, Dose-dependent AGO1-Mediated inhibition of the miRNA165/166 pathway modulates stem cell maintenance in *Arabidopsis* shoot apical meristem. *Plant Communications*, 2020. 1(1).
34. Phelps-Durr, T.L., *et al.*, Maize rough sheath2 and its *Arabidopsis* orthologue ASYMMETRIC LEAVES1 interact with HIRA, a predicted histone chaperone, to maintain knox gene silencing and determinacy during organogenesis. *Plant Cell*, 2005. 17(11): p. 2886-98.
35. Nie, X., *et al.*, The HIRA complex that deposits the histone H3.3 is conserved in *Arabidopsis* and facilitates transcriptional dynamics. *Biology Open*, 2014. 3(9): p. 794-802.
36. Wang, X., *et al.*, REF6 promotes lateral root formation through de-repression of PIN1/3/7 genes. *Journal of Integrative Plant Biology*, 2019. 61(4): p. 383-387.
37. Park, S., *et al.*, PLANT HOMOLOGOUS TO PARAFIBROMIN is a component of the PAF1 complex and assists in regulating expression of genes within H3K27ME3-enriched chromatin. *Plant Physiology*, 2010. 153(2): p. 821-831.
38. Ramon, M., *et al.*, The hybrid Four-CBS-Domain KIN β subunit functions as the canonical γ subunit of the plant energy sensor SnRK1. *The Plant Journal*, 2013. 75(1): p. 11-25.
39. Yu, T.-Y., *et al.*, The *Arabidopsis* receptor kinase ZAR1 is required for zygote asymmetric division and its daughter cell fate. *PLOS Genetics*, 2016. 12(3): p. e1005933.
40. Huang, Y., *et al.*, The dynamic proteome in *Arabidopsis thaliana* early embryogenesis. *Development*, 2022. 149(18).
41. Vale, E.M., *et al.*, Mitochondrial proteomics reveals new insights into embryogenic competence acquisition in *Carica papaya* L. callus. *Journal of Proteomics*, 2022. 252: p. 104434.
42. Tateda, C., *et al.*, Molecular and genetic characterization of the gene family encoding the voltage-dependent anion channel in *Arabidopsis*. *Journal of Experimental Botany*, 2011. 62(14): p. 4773-4785.

43. Xu, J., *et al.*, EMBRYONIC FACTOR 1 encodes an AMP deaminase and is essential for the zygote to embryo transition in Arabidopsis. *The Plant Journal*, 2005. 42(5): p. 743-758.
44. Hu, W.-W., H. Gong, and E.-C. Pua, Modulation of SAMDC expression in Arabidopsis thaliana alters in vitro shoot organogenesis. *Physiologia Plantarum*, 2006. 128(4): p. 740-750.
45. Aida, M., *et al.*, The PLETHORA genes mediate patterning of the Arabidopsis root stem cell niche. *Cell*, 2004. 119(1): p. 109-120.
46. Galinha, C., *et al.*, PLETHORA proteins as dose-dependent master regulators of Arabidopsis root development. *Nature*, 2007. 449(7165): p. 1053-1057.
47. El Ouakfaoui, S., *et al.*, Control of somatic embryogenesis and embryo development by AP2 transcription factors. *Plant Molecular Biology*, 2010. 74(4): p. 313-326.
48. Horstman, A., *et al.*, The BABY BOOM transcription factor activates the LEC1-ABI3-FUS3-LEC2 network to induce somatic embryogenesis. *Plant Physiology*, 2017. 175(2): p. 848-857.

Acknowledgements

This research was supported by funds from the Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) (Proc E-26/211.310/2021; Proc E-26/204.185/2024) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (312258/2023-2), awarded to V.S. This study was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil (CAPES), Finance Code 001. The authors would like to thank FAPERJ, CNPq and CAPES for providing scholarships to graduate students.

Author contributions

LRX, VS and JJT contributed to the development of the methodology, the acquisition of results, and the writing, review, and editing of the manuscript. CS-C reviewed, edited, and translated the manuscript.

Corresponding authors

Correspondence to Vanildo Silveira

Ethics declarations

The authors declare that they have no competing interests.

Ethics approval and consent to participate

This article does not contain any studies with human or animal participants. Work with plants complies with relevant institutional, national and international guidelines and legislation.

Consent for publication

All the authors provide their consent for this publication.

Capítulo 5. Time-based proteomic and epigenetic analysis highlights key components driving sugarcane embryogenic callus initiation

Time-based proteomic and epigenetic analysis highlights key components driving sugarcane embryogenic callus initiation

Lucas Rodrigues Xavier¹, Roberta Pena da Paschoa¹, Caio Cezar Guedes Corrêa¹, Karina da Silva Vieira¹, Daniel Dastan Rezabala Pacheco¹, Lucas do Espírito Santo Gomes¹, Carlos Eduardo Assis da Silva¹, Laís dos Santos da Conceição¹, Vitor Batista Pinto², Claudete Santa-Catarina², Vanildo Silveira¹

¹Laboratório de Biotecnologia, Centro de Biociências e Biotecnologia (CBB), Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), Campos dos Goytacazes, RJ, 28013-602, Brazil

²Laboratório de Biologia Celular e Tecidual, CBB-UENF, Campos dos Goytacazes, RJ, Brazil

Abstract

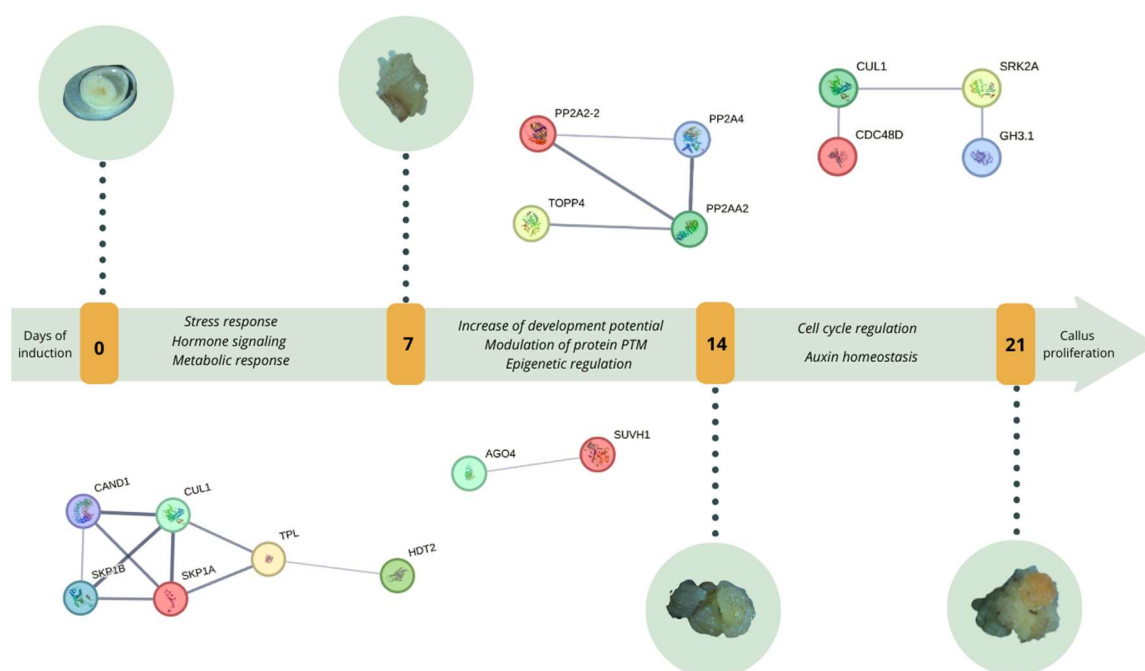
Sugarcane (*Saccharum* spp. hybrid) is an economically important crop, widely grown for its sucrose content, and somatic embryogenesis is a promising approach for efficient micropropagation and genetic transformation in this market. The induction of somatic embryogenesis is controlled by different genes and proteins involved in hormonal pathways and stress responses, acting as key regulators in the in vitro cell dedifferentiation and differentiation. In this study, we employed a temporal proteomics approach and global DNA methylation to investigate the underlying molecular mechanisms governing sugarcane embryogenic callus induction in response to a 2,4-dichlorophenoxyacetic acid throughout induction. The comprehensive proteomic profiling showed 1011 proteins differentially accumulated in at least one of the four times (0, 7, 14 and 21 days) analyzed during calli formation. These proteins were classified in different clusters based on their accumulation profile. Our study advances the understanding of the molecular mechanisms of somatic embryogenesis in sugarcane, highlighting the regulation of key proteins during this process. Proteins involved in photosynthesis and starch accumulation were not observed during the callus induction. On the other hand, proteins involved in embryogenesis, epigenetic regulation, hormone responses and protein post-translational modification were accumulated during callus induction. The predicted interactions between protein TOPLESS with auxin response proteins (SKP1, CUL1 and CAND1 proteins) are associated with the level of chromatin condensation through interaction with histone deacetylase HDT2 during embryogenic callus initiation. Additionally, global DNA methylation levels ranged from 13.56% to 16.46% during the induction period, indicating that epigenetic modulation may contribute to the acquisition of embryogenic competence. The identified proteins and methylation dynamics provide valuable targets for refining somatic embryogenesis protocols and advancing biotechnological applications in sugarcane.

Keywords: Embryo development; Bottom-up proteomics; Somatic embryogenesis; Callus induction; Cell totipotency.

Corresponding author: Dr. Vanildo Silveira

E-mail: vanildo@uenf.br

Graphical abstract



Introduction

The genetic improvement of sugarcane is always a challenge due to the complexity of the genome and poor seed viability presented by commercial hybrids (Thirugnanasambandam *et al.* 2018). Furthermore, the propagation of sugarcane for production depends on obsolete techniques, such as planting from stalks, leading to the spread of diseases and limiting the number of seedlings available (Jericó *et al.* 2018). On the other hand, in vitro micropropagation techniques, such as somatic embryogenesis, have been developed for sugarcane and are an efficient alternative to conventional vegetative propagation (Heringer *et al.* 2017). In this scenario, research on the in vitro propagation of sugarcane provides novel strategies to enhance the performance of plants in the field by genetic engineering and to generate plantlets on a large scale for sugar and bioenergy production.

Somatic embryogenesis requires the integration of hormonal, epigenetic, transcriptomic and proteomic signals that are still poorly understood (Reis *et al.* 2016). In this sense, activation of somatic embryogenesis in plants often involves the use of the synthetic auxin 2,4-dichlorophenoxyacetic acid (2,4-D), which mimics the action of the natural plant hormone auxin and interacts with auxin receptors, initiating a molecular signaling cascade (Wójcik *et al.* 2020). In *Arabidopsis*, different groups of genes are expressed at specific times during the induction of somatic embryogenesis, including transcription factors and genes

involved in embryogenic development (Wickramasuriya and Dunwell 2015). In addition, it has been shown in *Carica papaya* that specific groups of proteins play a role at specific stages of embryogenic callus induction, resulting in networks that regulate the post-translational modification of proteins, epigenetics and hormonal responses throughout callus initiation (Xavier *et al.* 2023). Together, these data show that somatic embryogenesis is a dynamic and temporally modulated process, where its elucidation depends on determining the regulatory factors present at different times.

Significant changes in gene expression profiles during somatic embryogenesis induction are attributed to hormonal and epigenetic factors, which regulates gene expression, essential for plant embryonic development (Wójcikowska *et al.* 2020). Also, studies in sugarcane are showing an integrated network of proteins and post-translational modifications associated with the embryogenic competence and differentiation of somatic embryos, where it's now clear that several layers of regulation of gene expression are associated (Almeida *et al.* 2020; Heringer *et al.* 2015; Reis *et al.* 2016; Xavier *et al.* 2022). In this context, proteomic approaches have been fundamental in shedding light on the mechanisms that control the differentiation and development of sugarcane somatic embryos in previous works (Heringer *et al.* 2017). This research into the modulation of the sugarcane proteome during somatic embryogenesis has revealed interaction networks that integrate different hormonal, epigenetic, metabolic and cell signaling responses (Almeida *et al.* 2020; Xavier *et al.* 2022). However, regulation of the initial stages of embryogenic callus induction in sugarcane is still poorly understood.

Proteomic studies comparing embryogenic and non-embryogenic callus revealed that unique proteins accumulated only in embryogenic callus, such as the dehydrogenase-like protein, the desiccation-related protein 62 pcc13, the catalytic subunit of callose synthase 1, and the NOS-interacting protein, may play a role in the maturation process of sugarcane somatic embryos (Heringer *et al.* 2015). In addition, regulatory phosphoproteins, such as the serine/threonine protein kinase OSK3, cyclin-dependent kinase 5, calcium-dependent protein kinase 5, SNF1-related protein kinase 2 and SNF1-related protein kinase regulatory subunit gamma-1, as well as abscisic acid responses were associated with embryogenic callus competence (Almeida *et al.* 2020). On the other hand, the phosphorylation of proteins such as ABA-responsive element binding factor 1, histone deacetylases 6 and 9, TOPLESS and argonaute 1b is regulated during sugarcane somatic embryo differentiation and development (Xavier *et al.* 2022). These proteomics studies on somatic embryogenesis generally makes comparisons of tissues/treatments, such as embryogenic callus versus non-embryogenic

callus or the presence and absence of some treatment (Heringer *et al.* 2015; Polesi *et al.* 2023). In this context, a time series proteomics analysis could be applied to clarify the mechanism of in vitro cell dedifferentiation and induction of embryogenic cells.

This study presents the time-resolved dynamics of the sugarcane proteome and global DNA methylation during the first 21 days of embryogenic callus induction in response to 2,4-D. Through this integrative approach, novel proteomic signatures and a progressive methylation pattern were associated with explant cell dedifferentiation and callus formation, highlighting the roles of phosphatases, auxin response proteins, and epigenetic regulators in the early stages of somatic embryogenesis.

Material and Methods

Plant material and callus induction

For callus induction, two-month-old sugarcane plants (SP80–3280) were used as the source of explants. The outer mature leaves were removed, and the 20 cm resulting leaf rolls were surface sterilized in 70 % ethanol for 1 min and posteriorly in 0.75 % sodium hypochlorite for 15 min, and then water-washed three times. The explants were obtained from 3 mm thick slices of leaf rolls from cross-sections. The explants were transferred in test tubes (150 × 25 mm) containing 10 mL of Murashige and Skoog (Murashige and Skoog 1962) (MS medium - Phytotechnology Lab, KS, USA), supplemented with 30 g.L⁻¹ sucrose (Vetec, SP, Brazil), 10 µM 2,4-D (Sigma-Aldrich, MO, USA). The pH of culture medium was adjusted to 5.8 before adding 2 g.L⁻¹ Phytigel®, followed autoclaving at 121 °C for 15 min. The test tubes with explants were maintained in darkness at 25 °C for callus induction. The time-lapse treatments were the explants in the 0 day (immediately before cutting the leaf rolls) and after 7, 14 and 21 days of induction.

Protein extraction and digestion

Protein extraction and digestion for proteomics analyses were performed with three biological replicates (each replicate is a pool of two immature leaf rolls for day 0 or five explants for days 7, 14 and 21) (Figure 1) and the samples were pulverized in liquid nitrogen and 150 mg of fresh mass tissue were collected for each treatment. For samples before the induction treatment, two whole immature leaf rolls were used for protein extraction in each biological repetition (Figure 1A). Next, 500 µL of extraction buffer consisting of 7 M urea (GE Healthcare, Germany), 2 M thiourea (GE Healthcare, Germany), 2 % Triton X-100 (GE Healthcare,

Germany), 1 % dithiothreitol (DTT, GE Healthcare, Piscataway, USA), 1 mM phenylmethanesulfonyl fluoride (PMSF, Sigma-Aldrich) was added, and the mixture was vortexed and centrifuged. The supernatants were collected, and the protein concentration was measured using Bradford reagent (Bio-Rad, Hercules, USA) (Bradford 1976).

Following, 100 µg of protein from each biological repetitions was digested using Microcon-30 kDa filter (Millipore) with the filter-aided sample preparation (FASP) methodology described by Wiśniewski *et al.* (2009), with the following modifications: Protein aliquots were washed with 200 µL of 50 mM ammonium bicarbonate (solution A), and centrifuged at 12,000 x g for 15 min at 25 °C (unless stated otherwise, all centrifugation steps were performed under this condition). This step was repeated once for complete removal of urea before reduction of proteins. Next, 100 µL of 50 mM DTT, freshly made in solution A, was added, gently vortexed, and incubated for 20 min at 60 °C (with repetitive cycles of 1 min agitation and 4 min resting, until the end of the 20 minutes). Then, 200 µL of 8 M Urea in 50 mM ammonium bicarbonate (solution B) was added, and centrifuged for 15 min. For protein alkylation, 100 µL of 50 mM iodoacetamide, freshly prepared in solution B, was added, gently vortexed, and incubated for 20 min at 25 °C in the dark (1 min agitation and 19 min resting, at 47 x g). Next, 200 µL of solution B was added and centrifuged for 15 min. This step was repeated once. Then, 200 µL of solution A was added and centrifuged for 15 min. This step was repeated twice. Approximately 50 µL of sample should remain in the last washing. For protein digestion, the filter units were transferred to new microtubes and 25 µL of trypsin (V5111; Promega, Madison, WI, USA) solution (1:100 enzyme:protein) were added, gently vortexed, and incubated for 16 h at 37 °C (with repetitive cycles of 1 min agitation and 4 min resting, at 47 x g). For peptide elution, 50 µL of solution A was added and centrifuged for 15 min. This step was repeated once. For trypsin inhibition, 5 µL of 15% of trifluoroacetic acid was added, gently vortexed, and incubated for 30 min at 37 °C. Then, samples were centrifuged for 15 min, and the supernatants were collected and vacuum dried. Peptides were resuspended in 50 µL solution of 95 % water and 5 % acetonitrile plus 0.1 % formic acid in mass spectrometry (MS)-grade (Sigma-Aldrich). The peptide concentration was estimated by measuring A205 nm using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific).

Proteomics analysis

A nanoAcquity UPLC M-class (Waters, Manchester, UK) coupled to a Synapt G2-Si HDMS mass spectrometer (Waters) with electrospray ionization (ESI)-MS/MS source was used. An amount of 2 µg of peptides was loaded onto a c18 trap column (100Å, 5 µm, 180 µm x 20 mm, 2D; Waters) at 5 µL min⁻¹ for 3 min and then onto a HSS T3 analytical reversed-

phase column (100Å, 1.8 µm, 75 µm x 150 mm; Waters) at 400 nL min⁻¹ heated to 45 °C. For peptide elution, a binary gradient was used, with mobile phase A consisting of water and 0.1% formic acid and mobile phase B consisting of acetonitrile and 0.1% formic acid. Gradient elution was started with 5% B, which was then increased to 40% B until 92 min and from 40 to 99% B until 96 min, after which it was maintained at 99% until 100 min, posteriorly decreased to 5% B until 102 min, and maintained at 5% B until the end of the run at 118.00 min.

The mass spectrometer has been set up in positive, resolution mode (V mode), and data-independent acquisition (DIA) mode, at 35,000 FWHM, with ion mobility (HDMSE) wave velocity program starting at 800 m s⁻¹ and ending at 500 m s⁻¹. The transfer collision energy increased from 25 to 55 V in high-energy mode, the cone voltage was 40 V, and capillary voltage was 2800 V. The pressure of the nano-flow gas was 0.5 bar, the purge gas flow rate was 150 L.h⁻¹, and the source temperature was 100 °C. For the time-of-flight (TOF) parameters, the scan time was set to 0.5 s in continuous mode, with a mass range of 50 to 2000 Da. Human [Glu¹]-fibrinopeptide B at 100 fmol.µL⁻¹ was used as an external calibrant, and lock mass acquisition was performed every 30 s. Mass spectra were subsequently acquired by using MassLynx 4.1 software (Waters, Manchester, UK).

Proteomics data analysis

The proteomics MS/MS spectra processing and database search were performed using ProteinLynx Global SERVER v.3.02 (Waters, Manchester, UK) against the protein bank generated from the sequencing of the haploid *Saccharum spontaneum* AP85-441 (Zhang *et al.* 2018). The quantification was performed using ISOQuant software v.1.7 (Distler *et al.* 2014). The threshold parameters for intensity, energy, peptides, fragments, and modifications are described in our previous work (Reis *et al.* 2021). All the proteomics data, and results tables have been deposited in the *ProteomeXchange Consortium* (Vizcaíno *et al.* 2014) via the PRIDE (Perez-Riverol *et al.* 2019) repository with identifier PXXXXXXX.

To guarantee the quality of the data processing, only proteins present in the three biological repetitions were accepted for the differential accumulation analysis. Unique proteins were defined as those absent in all biological repetitions of one treatment and present in another. Differentially accumulated proteins (DAPs) were those that showed significantly different means from the triplicates of another treatment by the Student's t-test (two-tailed, equal variance; *p-value* < 0.05), and with a Log2 fold change greater than 1.5 times (Log2FC > 0.585) for up-accumulated proteins and less than 0.667 times (Log2FC < -0.585) for down-

accumulated proteins. To understand the proteome modulation during sugarcane callus induction, a soft clustering protein analysis was performed. Soft clustering can assign a protein to many clusters using the fuzzy c-means algorithm with time-course data on the protein accumulation (Kumar and M 2007). The R package Mfuzz (version 2.34.0, <http://bioconductor.org/packages/release/bioc/html/Mfuzz.html>) was used for cluster analysis of DAPs.

Automated functional annotation analysis was performed using OmicsBox software (<https://www.biobam.com/omicsbox>) and a manual annotation was made in UniProtKB (<https://www.uniprot.org>). Sugarcane orthologous DAPs in Arabidopsis were obtained using Metascape (Zhou *et al.* 2019). These orthologs were used for enrichment analysis of biological processes, cellular components, molecular functions, and KEGG pathways. The identification of kinases, transcription factors and transcriptional regulators among DAPs using iTAK tool (Zheng *et al.* 2016). The DAPs were also used for predicted interaction networks through STRING search (Szklarczyk *et al.* 2020).

DNA extraction

DNA extraction was performed according to (Doyle & Doyle, 1987) with modifications. A total of 200 mg of fresh matter was ground in liquid nitrogen and transferred to a microtube (2 ml). Then, 700 µl of extraction buffer (2 % CTAB, 1.4 M NaCl, 20 mM EDTA, 100 mM Tris–HCl, pH 8.0, 1 % polyvinylpyrrolidone and 0,2 % 2-mercaptoethanol) was added, vortex and incubated at 65°C for 30 min. Subsequently, 650 µl of chloroform–isoamyl alcohol (CIA) (24:1) was added and subjected to stirring for 10 min by inversion and centrifuged at 10,000 rpm for 9 min at 4°C. The aqueous phase was transferred to a new tube to which was added 200 µl of extraction buffer and 650 µl of CIA and subjected to stirring for 5 min by inversion. The tubes were centrifuged again at 10,000 rpm for 9 min at 4°C. The supernatant was transferred to a new microtube, and DNA was precipitated with 500 µL of cold isopropanol. For DNA precipitation, the tubes were incubated at -20°C for 1 h to form the pellet. Next, the tubes were centrifuged at 10,000 rpm for 9 min at 4°C, and the supernatant was discarded. The pellet was washed three times with 250 µL of 70% ethanol. After washing, the ethanol was removed, and the pellet was air-dried at 37°C for 1 h. The pellet was then resuspended in TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0). and quantified in a spectrophotometer Nanodrop One C (Thermo Fisher Scientific®), providing nucleic acid concentration of samples.

Nucleic acid digestion

Digestion procedures were based on the protocol described by (Fraga *et al.*, 2012). Digestions with RNase were performed with 50.000 ng of nucleic acids in TE buffer. The samples were heated to 65 °C for 5 min to denaturation and incubated at 4 °C for 10 min. Subsequently, 1 µl of RNase A (Sigma-Aldrich®) (1 µg µl⁻¹) and 5 µl of RNase T1 (Sigma-Aldrich®) (20 µg µl⁻¹) were added, vortexed and incubated at 37 °C for 17 h. After digestion, 10 µl of sodium acetate (3 M, pH 5.4) was added and the microtubes were vortexed. DNA and RNA resistant to RNase were precipitated with 68 µl of cold isopropanol for 30 min at -20 °C. The samples were centrifuged at 5,000×g for 10 min to form a pellet, washed with ethanol 70 % (v/v) and re-centrifuged at 5,000×g for 5 min. The pellets were dried at room temperature and resuspended in 50 µl of water for LC/MS.

Digestions with nuclease P1 and alkaline phosphatase (Sigma-Aldrich®) were performed using the 50 µl of nucleic acid resuspended previously. Nucleic acids were denatured by heating at 100 °C for 2 min and incubated at 4 °C for 5 min. 2,5 µl of 10 mM ZnSO₄ and 5 µl of nuclease P1 (1.0 U ml⁻¹ in 30 mM NaOAc, pH 5.4) were added, followed by sample vortexing and incubating for digestion reaction at 37 °C for 17 h. After this period, 5 µl of 0.5 M Tris-HCl (pH 8.3) and 5 µl of alkaline phosphatase (10 U ml⁻¹ (NH₄)₂SO₄ 2.5 M) were added and incubated for 2 h at 37 °C. The samples were centrifuged at 10,000×g for 5 min and the supernatant was stored at -80 °C until Mass spectrometer analysis.

Mass spectrometer analysis

The runs consisted of three biological replicates. The dNTPs were separated by liquid chromatography by loading 10 µL of the solution recovered in the extraction onto a Acquity UPLC BEH C18 1.7 µm column (2.1 µm × 50 mm) (Waters) heated to 40 °C, with a flow rate of 300 µL.min⁻¹. The binary gradient consisted of LC-MS grade water plus 0.01% formic acid, and LC-MS grade methanol plus 0.01% formic acid as the eluent. The elution gradient was initiated at 100% water and decrease to 95% over five minutes, and further reduced to 5% over the next one minute, maintained at 5% for one minute, returning to 100% in 1 minute and maintained at 100% for 2 minutes to balance the column, totaling 10 minutes of running.

Using a high-performance dual orthogonal Z-spray electrospray ionization (ESI) source (Waters), the effluent from the Acquity UPLC I-Class FTN (Waters) was introduced into a Xevo TQ-XS triple quadrupole mass spectrometer (Waters). A pressurized nitrogen gas nebulization of 2.8 bar and a desolvation flow of 550 L h⁻¹ were used, and a cone flow rate of

150 L h⁻¹ was used. The desolvation gas temperature was set to 500 °C, the source temperature was 150 °C, and the capillary voltage was 2.8 kV. The argon gas was used as the collision gas with a pressure of 0.8 kgf cm², and 0.15 ml/min in the collision cell. The spectrometer was operated in multiple reaction monitoring (MRM) mode, with 20 V for cone voltage, positive in MassLynx software (v.4.2, Waters) for MS data acquisition. The spectra were integrated, and the lines were smoothed twice with averages between three scans per points.

The area under the peak signal was compared through a linear equation estimated from a standard curve (concentrations: 0.6, 1.2, 2.3, 4.7, 7.0, 9.4 and 11.7 ng for Desoxycitidina and 0.7, 1.5, 3, 5.9, 8.9, 11.9 and 14.8 ng for 5-Methyl-2'-deoxycytidine). The 5-Methylcytosine (5mdC) and cytosine (dC) (Zimo Research® - D5405) used as standards (dA, dT, dC and dG, 5mdC) were digested for 2 h with alkaline phosphatase (10 U ml⁻¹) and Tris-HCl (0.5 M, pH 8.3) at 37 °C, to obtain the nucleosides. The standard nucleosides (2000 ng) were prepared in water for LC/MS and stored at -20 °C. The identification of each nucleoside was made according to the known mass of each (5mdC: 242/126; dC: 228/112; dG: 268/152; dA: 252/136; dT: 243/127 m/z).

Statistics of Global DNA Methylation Analysis

Statistical analysis was carried out using R software (version 4.3.2) (Team, 2022). For the statistical analysis, the homogeneity of variances was assessed using Bartlett's test ($p < 0.05$). Multiple comparisons between time points were performed using the Student-Newman-Keuls (SNK) test, with a significance level of 5% ($p < 0.05$). The linear model was fitted using the `lm` function in R, and an analysis of variance (ANOVA) was conducted to assess the significance of differences between treatments. Additionally, adjusted means and confidence intervals for each induction time were calculated using the `emmeans` function, which estimates the least-squares means. Data visualization was performed using the `ggplot2` package.

Results

Morphological features of the explant during embryogenic callus induction

The somatic embryogenesis protocol using immature leaf roll resulted in a shift from vegetative development to embryogenic development, yielding embryogenic callus (Figure 1). Thus, it was possible to observe the beginning of callus formation after 7 days of induction

treatment (Figure 1B). After 14 days, the callus becomes prominent on the explant and proliferates until day 21 (last time analyzed) (Figures 1C and D).

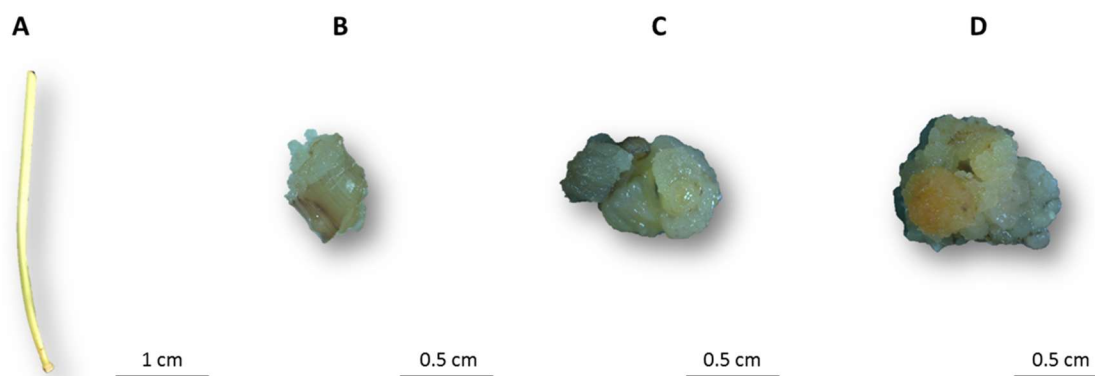


Figure 1. Morphological aspects of sugarcane callus during induction. (A) The sugarcane explant used to start the experiment showing immature leaf foliage from the central part of the plants with the mature green leaves removed before slicing from transverse sections; (B) Explants at 7; (C) 14; and (D) 21 days after placing in the culture medium with 2,4-dichlorophenoxyacetic acid (2,4-D) for callus induction.

Proteomics analysis during embryogenic callus initiation

A time-lapse proteome analysis was performed during sugarcane callus induction on the explant before induction (0 day), and 7, 14 and 21 days after the start of the induction treatment. A total of 1787 proteins were identified, of which 1011 showed differential accumulation in at least one of the treatment comparisons (Supplementary Table 1).

Clustering analysis revealed 4 clusters with a distinct pattern of protein accumulation during embryogenic callus induction (Figure 2, Supplementary Table 2). These protein groups contain 244, 256, 204 and 307 proteins (clusters 1, 2, 3 and 4, respectively). Cluster 1 shows the proteins that decline in abundance throughout the induction process, while the cluster 3 profile includes proteins that decrease in abundance during the multiplication of the embryogenic callus after 14 days of treatment. On the other hand, the cluster 2 shows proteins that were increasing in accumulation throughout the induction treatment. Cluster 4 shows proteins that increase in abundance in the first 7 days of induction, and then decrease as the callus grows.

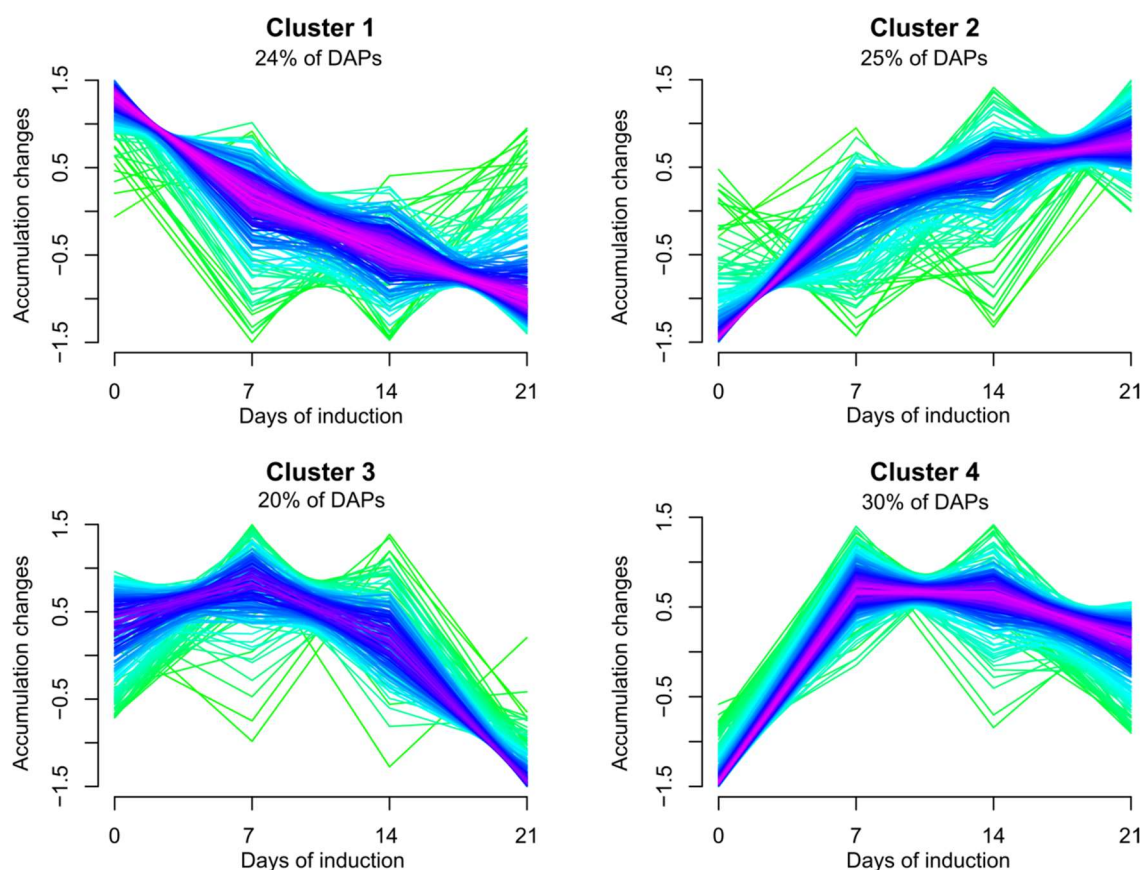


Figure 2 - Cluster analysis of differentially accumulated proteins (DAPs) during sugarcane embryogenic callus induction with 2,4-D. The DAPs were grouped into four similar patterns (Cluster 1, Cluster 2, Cluster 3, and Cluster 4). The accumulation changes axis represents how many standard deviations the log2 of the ion count deviates from the average of the DAPs. Yellow or green colored lines correspond to genes with low membership value; red and purple colored lines correspond to genes with high membership value. Supplementary Table 2 provides detailed information on the proteins included in each cluster.

Functional analysis of proteomic data

The investigation of proteins regulated during the induction of sugarcane embryogenic calli yielded insights into enriched gene ontology terms and metabolic pathways associated with this process (Supplementary Figure 1). This analysis showed that only clusters 2 and 4 have proteins that enrich the biological processes of embryogenesis and detoxification, in addition to increasing the process of generating metabolites and energy (Supplementary Figure 1A). The proteins in clusters 1 and 3 enriched the acetyl-CoA biosynthetic process and the nucleotide-sugar metabolic process (Supplementary Figure 1A). In addition, the cellular components in the clusters 1, 3 and 4 enriched the chloroplast stroma, proteasome complex and apoplast proteins, while the cluster 2 showed enrichment in nucleolar proteins (Supplementary Figure 1B).

The molecular functions of translation initiation factor, and ligase activity were enriched only among the proteins in clusters 2 and 4, while structural constituent of cytoskeleton was enriched in cluster 3 (Supplementary Figure 1C). The molecular function of salicylic acid binding was enriched only in proteins at cluster 2. Cluster 1 did not display any enrichment greater than the other clusters in any molecular function (Supplementary Figure 1C). The pathway analysis revealed that the metabolism of starch and sucrose was enriched only among the proteins in cluster 1, while the metabolism of amino acids, including glycine, serine, threonine, cysteine, methionine, valine, leucine and isoleucine were most enriched in cluster 2 (Supplementary Figure 1D). Cluster 3 was most enriched in ribosomal and the citric acid (TCA) cycle proteins, whereas nucleocytoplasmic transport was enriched in clusters 2 and 4 (Supplementary Figure 1D). After the functional and statistical analysis, it was possible to determine a list of proteins that are important for the dedifferentiation process and increase in the developmental potential presented by the callus (Table 1).

Table 1. Differentially accumulated proteins (DAPs) associated with sugarcane embryogenic callus induction during the first 21 days of treatment with 2,4-D.

accession	BLAST Top Hit	Arabidopsis ortholog	Cluster
Sspon.04G0017510-3C	Cullin-associated NEDD8-dissociated protein 1	CAND1	4
Sspon.01G0018700-1A	Cell division control protein 48 homolog D	CDC48D	4
Sspon.03G0044930-1D	cell division cycle 5-like protein	CDC5	4
Sspon.07G0013010-4D	cullin-1	CUL1	2
Sspon.01G0012730-1A	probable indole-3-acetic acid-amido synthetase GH3.8	GH3.1	-
Sspon.07G0003330-2B	Histone deacetylase HDT2	HDT2	2
Sspon.06G0013840-1A	Adenosylhomocysteinase	HOG1	4
Sspon.08G0006940-2C	serine/threonine-protein phosphatase PP2A-1 catalytic subunit	PP2A-2	2
Sspon.01G0017760-1A	putative serine/threonine-protein phosphatase PP2A-4 catalytic subunit	PP2A-4	-
Sspon.01G0030520-1P	Serine/threonine-protein phosphatase PP2A-2 catalytic subunit	PP2A-4	2
Sspon.01G0003870-1A	Serine/threonine-protein phosphatase PP2A-4 catalytic subunit	PP2A-4	4
Sspon.08G0016530-2B	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A beta	PP2AA2	4
Sspon.03G0015850-1A	S-adenosylmethionine synthase 3	SAM-2	2
Sspon.03G0042160-1C	SKP1-like protein 1	SKP1a	1
Sspon.04G0021150-1A	SKP1-like protein 1	SKP1a	2
Sspon.06G0013850-2B	SKP1-like protein 1	SKP1b	4
Sspon.05G0011250-2P	putative snRK/SAPK family protein kinase	SNRK2.4	4
Sspon.02G0022290-3C	histone-lysine N-methyltransferase, H3 lysine-9 specific SUVH1-like	SUVH1	2
Sspon.01G0007220-1A	Serine/threonine-protein phosphatase PP1	TOPP4	4
Sspon.01G0006700-1A	Protein TOPLESS	TPL	2

Identification of regulatory proteins

In this study, several classes of protein kinases and transcription factors were identified throughout the induction of sugarcane embryogenic callus. The major group of kinases comprises the membrane receptor protein kinases (10 members), followed by the CMGC family (4 members), calcium/calmodulin-dependent protein kinases (3 members) and one Aurora kinase (Supplementary Table 1). In addition, 10 different classes of transcriptional

regulators were found, including the SWI/SNF-BAF60b, SET, NF-YC, MYB-related, Jumonji, HMG, CSD, C2H2, bHLH and AUX/IAA families (Supplementary Table 1).

Predicted of protein-protein interaction networks

The search for predicted interactions between proteins (from Table 1) revealed a series of interaction networks that may be important for the transition from vegetative growth to embryogenic development and callus formation (Figure 3). We were able to find interactions between master regulators of processes such as embryo development, protein phosphorylation, hormonal responses and methionine metabolism.

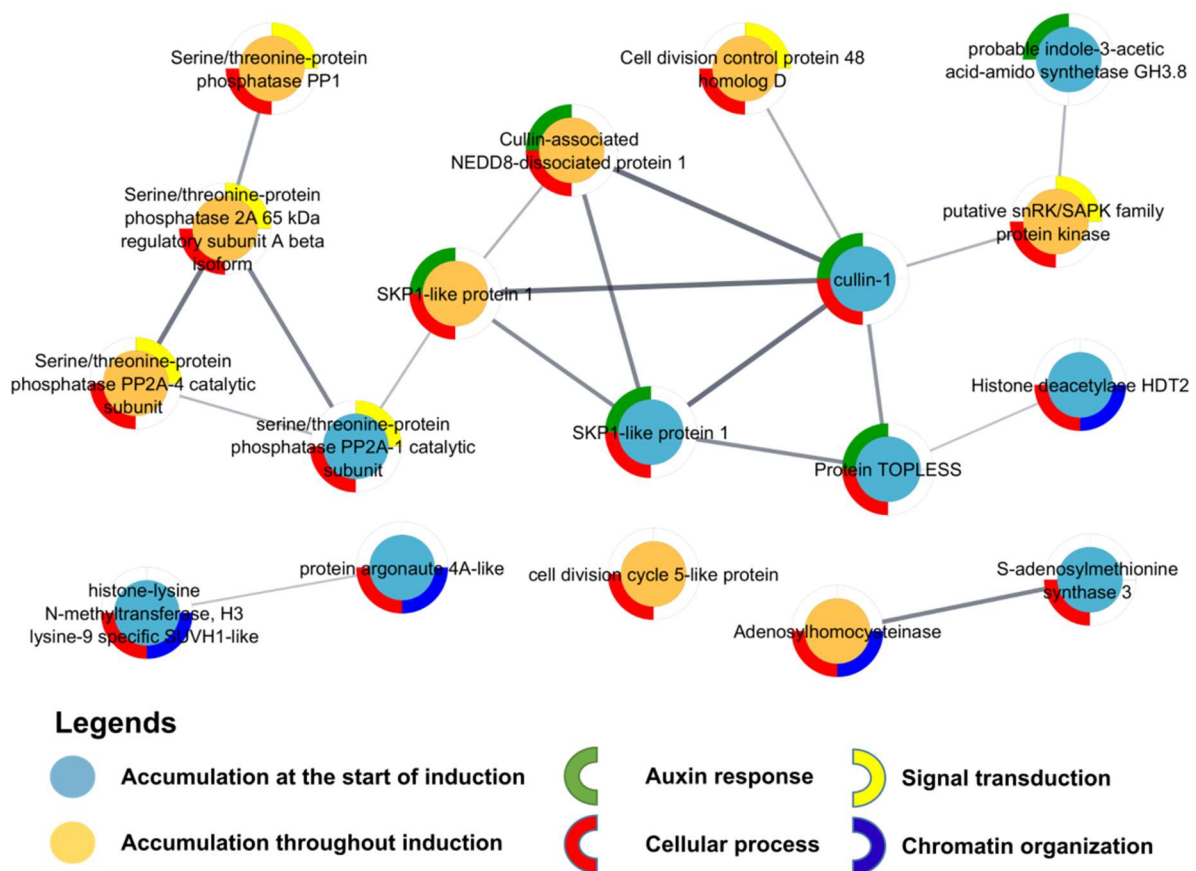


Figure 3. Predicted protein-protein interaction networks involved in induction of sugarcane embryogenic callus. The search for interactions in STRING (medium confidence) using proteins from Table 1 has revealed a network with a series of proteins that may play a role in the transition from vegetative growth to embryogenic development and callus induction. The blue and orange nodes represent proteins from clusters 2 and 4, respectively. The circle around nodes represents features of proteins based on STRING annotation.

In the protein-protein interaction analysis, TOPLESS connects hormonal responses (SKP1, CAND1 and CUL1 proteins) with the level of chromatin condensation through the histone deacetylase HDT2 (Figure 3). Interactions involved in the phosphorylation of proteins

were also found, including protein kinases, phosphatases and 14-3-3 phosphoproteins, revealing a network of post-translational modifications during embryogenic callus induction (Figure 3). In this context, the interaction between SNRK2A and GH3.1 may play a role in regulating auxin homeostasis and the response levels of this hormone (Figure 3). The interaction between methionine metabolism proteins should regulate the level of DNA methylation during cell division in embryogenic callus (Figure 3).

Moreover, the time-lapse analysis of the sugarcane proteome revealed the depletion of proteins involved in photosynthesis, sucrose and starch metabolism in the explant during induction treatment (Supplementary Table 1). Proteins related to photosynthetic machinery, as the proteins Cytochrome b-c1 complex subunit 7 (Sspon.01G0030710-2P), cytochrome b-c1 complex subunit 9 (Sspon.05G0020120-1A), mitochondrial-like cytochrome b-c1 complex subunit Rieske (Sspon.04G0011390-2B), chloroplastic chlorophyll a-b binding protein 1 (Sspon.02G0036200-1B) and chlorophyll a-b binding protein M9 (Sspon.07G0002040-2B) decreasing their abundance during induction of embryogenic calli (Figure 2) (Supplementary table 2). In addition, accumulation of proteins involved in sucrose and starch biosynthesis decreases throughout the induction treatment (Figure 2) (Supplementary table 2) such as sucrose synthase 2 (Sspon.01G0011850-1T), sucrose synthase 3 (Sspon.01G0009300-1A) and starch phosphorylase1 (Sspon.01G0049600-1P) (Supplementary table 1). On the other hand, two sucrose synthase 1 proteins (Sspon.08G0022540-2C and Sspon.08G0022540-3D).

Moreover, several phosphatases proteins, the serine/threonine-protein phosphatase PP2A-4 catalytic subunit (Sspon.01G0003870-1A), serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A beta isoform (Sspon.08G0016530-2B), serine/threonine-protein phosphatase PP1 (Sspon.01G0007220-1A), serine/threonine-protein phosphatase PP2A-1 catalytic subunit (Sspon.08G0006940-2C) and serine/threonine-protein phosphatase PP2A-2 catalytic subunit (Sspon.01G0030520-1P) were accumulated during embryogenic callus induction.

Our data show also several DAPs associated with auxin during callus induction, among them the probable indole-3-acetic acid-amido synthetase GH3.8 (Sspon.01G0040550-2C). Moreover the predicted interactions among three SKP1-like protein 1 (Sspon.06G0013850-2B, orthologous to AtSKP1b, Sspon.04G0021150-1A and Sspon.03G0042160-1C, both orthologous to AtSKP1a), protein TOPLESS (Sspon.01G0006700-1A), cullin-1 (Sspon.07G0013010-4D) and Histone deacetylase HDT2 are probably associated with the integration of hormonal signaling and epigenetic regulation during somatic embryogenesis. The proteins involved in epigenetic control during the induction of sugarcane

embryogenic callus, such as protein argonaute 4A-like (Sspon.02G0016200-2D) was accumulated throughout callus induction and multiplication (Table 1) while the protein argonaute4a (Sspon.03G0016720-2B) decreased its accumulation during the induction treatment (Supplementary Table 2).

Global DNA methylation

Global DNA methylation levels progressively increased during somatic embryogenesis induction in *Saccharum spp. L.*, as shown in Supplementary Figure 2. On day 0, the average methylation was the lowest (13.56%), reflecting the baseline methylation status of the initial callus tissue. By day 7, an increase in methylation was detected (14.26). This trend continued from day 14 through day 21, with the highest DNA methylation levels recorded at the latter time point (15.76% and 16.46%, respectively). Although no statistically significant differences were observed between the time points, the average methylation levels showed a clear increasing trend over time.

Discussion

In the present investigation, we advanced the understanding of the molecular mechanisms associated with the initiation of sugarcane embryogenic calli in response to induction with 2,4-D. This study offers important insights into the dynamic proteomic changes and the role of epigenetic regulation that occur during the induction of embryogenic callus. Proteins associated with photosynthesis and starch accumulation consistently demonstrated a depletion pattern throughout the induction process. Conversely, proteins involved in embryogenesis, epigenetic regulation, hormonal responses, and post-translational modifications exhibited an accumulation profile, underscoring their pivotal roles in the pathways of callus induction. Additionally, global DNA methylation levels were observed to range from 13.56% to 16.46% during the callus induction, with a progressive increase over time. These findings suggest a link between the modulation of DNA methylation and the regulation of proteins involved in cellular reprogramming and embryogenic competence.

Proteins from energy metabolism and photosynthesis are left behind during dedifferentiation of the explant cells and callus induction

The induction of embryogenic calli depends on increasing the development potential of the competent cells in the explant (Fehér 2015). During this process, the explant cells must be able to respond to the induction treatment and change their functions, differentiating into

embryogenic or totipotent cell lines (Suprasanna and Bapat 2006). In this context, time-lapse analysis of the sugarcane proteome revealed the depletion of proteins involved in photosynthesis, sucrose and starch metabolism during induction (Supplementary Table 1). Moreover, the callus induction process is conducted under dark conditions, which modulate energy metabolism and deactivate the photosynthetic machinery, leading to decrease in abundance of several proteins not relevant for the induction step of embryogenic calli (Figure 2; Supplementary Table 2). Although the importance of photosynthetic potential has been reported in late stages of somatic embryogenesis (Afreen *et al.* 2002), photochemical activity remains relatively low in the early stages of callus proliferation (Rival *et al.* 1997). This suggests that sugarcane embryogenic calli do not have photosynthetic activity.

One of the main metabolic changes that occur during the initiation process of embryogenic calli is the disabling of the photosynthetic machinery present in the explant, along with the activation of heterotrophic metabolism (Xu *et al.* 2021). This metabolic adaptation is essential for utilization of the carbon sources present in the induction culture medium. Our findings indicate a reduction in the accumulation of proteins associated with sucrose and starch biosynthesis throughout the induction of embryogenic calli (Figure 2). This decline in protein levels suggests a reconfiguration of metabolic pathways, potentially prioritizing the utilization of exogenous sucrose within the culture medium. These observed changes likely reflect a strategic adjustment in the metabolism of embryogenic calli to efficiently utilize the provided sucrose and support their developmental processes.

The early post-translational modification of proteins drives the shift from vegetative to embryonic development

The post-translational modification of proteins is one of the first steps in cellular signaling perception and responses. The extensive regulation of receptor protein kinases found in this study clearly demonstrates the role of protein phosphorylation in the regulation of somatic embryogenesis (Table 1) (Supplementary Table 1). On the other hand, the signaling cascades involved in the response to 2,4-D are still unknown during the induction of embryogenic callus. Furthermore, a large number of protein phosphatases are accumulated during embryogenic callus induction and could be associated with modulation of the phosphorylation level of regulatory proteins (Table 1). Among the phosphatases, the serine/threonine-protein phosphatase PP2A-4 catalytic subunit (Sspon.01G0003870-1A) is involved in the regulation of auxin distribution through the dephosphorylation of PIN1 proteins (Ballesteros *et al.* 2013). This protein was grouped in cluster 4 and accumulated over the course of the induction treatment (Table 1). Moreover, others Serine/threonine-protein

phosphatases accumulated during embryogenic callus induction have a known role in regulating responses to abscisic acid and auxin transport (Ballesteros *et al.* 2013; Guo *et al.* 2015; Pernas *et al.* 2007). Together, these findings show the key roles of protein phosphatases in modulating and integrating hormonal responses during the initiation of embryogenic callus and in the transition from vegetative growth to embryogenic development. We were able to predict interactions between these proteins and further studies could demonstrate the targets and function of phosphatases during somatic embryogenesis (Figure 3).

In terms of post-translational protein modifications, histone methylation and acetylation play a central role in regulating the level of chromatin condensation and also in activating somatic embryogenesis. In this regard, the simple addition of trichostatin A (a potent inhibitor of histone deacetylases) is capable of inducing direct somatic embryogenesis in *Arabidopsis* (Wójcikowska *et al.* 2018). Our data show that the protein Histone deacetylase HDT2 (Sspon.07G0003330-2B) is accumulated during the whole process of sugarcane embryogenic callus induction (Table 1). The silencing of this protein causes *Arabidopsis* seeds to abort, demonstrating the role of this histone deacetylase during embryogenic development (Wu *et al.* 2000). New research using inhibitors and promoters of histone acetylation and methylation during sugarcane somatic embryogenesis could result in the discovery of novel targets to optimize this process.

Cell division and auxin responses work together throughout the initiation of sugarcane embryogenic calli

The use of 2,4-D to induce embryogenic calli is associated with the modulation of endogenous auxin levels and also of the cell cycle, causing an increase in the development potential and callus proliferation (Fehér 2019). Once embryogenic callus has been induced with 2,4-D, the cells must be able to respond to auxin signals and remain in homeostasis. Our data show that the probable indole-3-acetic acid-amido synthetase GH3.8, accumulated during callus induction, may be associated with the conjugation of IAA with amino acids to regulate the level of auxin response. In this context, the predicted interaction between probable indole-3-acetic acid-amido synthetase GH3.8 (Sspon.01G0040550-2C) and putative snRK/SAPK family protein kinase (Sspon.05G0011250-2P) may be associated with the activation of some homeostasis mechanism due to the presence of 2,4-D in the induction culture medium (Figure 3).

Our data showed a wide regulation of auxin responses and also of cell division throughout the induction treatment (Table 1). In addition, the proteins involved in these processes have predicted interactions that may be important for the activation and growth of embryogenic callus. The predicted interactions among three SKP1-like protein 1 (Sspon.06G0013850-2B, orthologous to AtSKP1b, Sspon.04G0021150-1A and Sspon.03G0042160-1C, both orthologous to AtSKP1a), protein TOPLESS (Sspon.01G0006700-1A), cullin-1 (Sspon.07G0013010-4D) and Histone deacetylase HDT2 is probably associated with the integration of hormonal signaling and epigenetic regulation during somatic embryogenesis (Figure 3). In this sense, SKP1 proteins are important kinases in auxin responses and cell division (Rao *et al.* 2018). In parallel, the TOPLESS protein modulates auxin responses during Arabidopsis embryogenesis by mediating transcriptional repression of target genes (Szemenyei *et al.* 2008). On the other hand, the mechanism of action and the targets of these proteins may be different during somatic embryogenesis and have not yet been studied.

Epigenetic changes are the keys to the doors of the development of embryogenic cell lines

The state of differentiation of plant tissues is strongly controlled by DNA methylation and the embryogenic competence is determined epigenetically (Wójcikowska *et al.* 2020). Here, we tracked the accumulation of several proteins involved in epigenetic control during the induction of sugarcane embryogenic callus (Table 1). These proteins are involved in DNA methylation, RNA messenger processing, histone acetylation and cell cycle regulation and accumulate at the onset of induction treatment (Supplementary Table 2). However, little is known about the epigenetic mechanisms and their targets during sugarcane somatic embryogenesis. Specifically, one argonaute protein exhibited an increase in accumulation, while another showed a decrease (Table 1; Supplementary Table 2). This dynamic regulation suggests differential roles for these proteins at different stages of callus development, potentially indicating distinct functions in the underlying molecular processes. The argonaute protein is involved in gene silencing mediated by small RNAs and is important for modulating various cellular responses (Hutvagner and Simard 2008; Peters and Meister 2007). In addition, the argonaute1 protein has already been found phosphorylated in sugarcane embryogenic callus, where it may be associated with somatic embryo differentiation (Xavier *et al.* 2022). In this context, it is possible that argonaute paralogs in sugarcane have developed different functions in sugarcane, where there must be genes with a role in vegetative growth and also during embryogenesis and other processes. Further studies should focus on

identifying the targets of argonaute-induced gene silencing and characterizing the microRNAs expressed during somatic embryogenesis.

The adenosylhomocysteinase protein accumulated during embryogenic callus induction (Table 1) is known for its role in DNA methylation-dependent gene silencing, as well as being associated with embryogenic development in *Arabidopsis* (Tzafrir *et al.* 2004). Previous studies have shown that Adenosylhomocysteinase is phosphorylated during the differentiation of sugarcane somatic embryos (Xavier *et al.* 2022). In this context, the interaction predicted between Adenosylhomocysteinase and S-adenosylmethionine synthase 3 in this study (Figure 3), both accumulated from the beginning of induction, may be related to DNA phosphorylation events and also histone proteins and may be associated with the induction of sugarcane embryogenic callus.

In addition to DNA methylation, histone methylation also takes place during callus initiation with the protein histone-lysine N-methyltransferase, H3 lysine-9 specific SUVH1-like (Sspon.02G0022290-3C) being accumulated throughout the induction process (Table 1) (Supplementary Table 2). In *Arabidopsis thaliana*, SUVH1 is responsible for gene silencing and is involved in the control of cell division, expansion and differentiation (Naumann *et al.* 2005). In this context, the cell division cycle 5-like protein and Cell division control protein 48 homolog D were accumulated during induction and might be involved in integrating cell cycle responses and promoting callus growth (Table 1). These findings imply that fine-tuning of gene expression is necessary for the formation and proliferation of sugarcane embryogenic callus. This process is involved in the transition from vegetative growth to embryogenic development and in vitro cell proliferation.

Global DNA Methylation and Its Interaction with Regulatory Proteins During Somatic Embryogenesis in Saccharum spp. L.

Our results showed that global DNA methylation levels ranged from 13.56% to 16.46% during the induction of somatic embryogenesis in *Saccharum* spp (Supplementary figure 2). *L.* This range is approximately three times higher than that observed in *Arabidopsis thaliana* (5.3%) and is comparable to other species such as *Oryza sativa* (16.3%), *Brassica oleracea* (16%), *Lepidium sativum* (14.43%), and *Primula vulgaris* (14.32%) (Alonso *et al.*, 2015). These findings support the existence of interspecies variation in global DNA methylation levels, which may be associated with genome size in angiosperms (Alonso *et al.*, 2015).

Although a progressive increase in global DNA methylation was observed during the induction of somatic embryogenesis in *Saccharum* spp. L., the differences were not statistically significant. Two hypotheses may explain this lack of significance. First, global DNA methylation represents an average of epigenetic modifications across the entire genome, which may mask relevant local variations in specific genes or regulatory regions. In plants, cytosine methylation occurs not only at symmetrical CG sites but also at symmetrical CHG and asymmetrical CHH contexts, where H represents any nucleotide except G (Agius *et al.*, 2023). As such, localized changes in DNA methylation may remain undetected in global analyses. Further studies focusing on locus-specific methylation and its correlation with the expression of key regulatory genes may provide deeper insights into the role of these subtle changes in embryogenic competence.

Alternatively, the variation in global DNA methylation levels may be intrinsically linked to embryogenic callus induction. Although the observed differences were not statistically significant, it is possible that even small changes in global methylation levels play a critical role in this process. This suggests that subtle shifts in %mC may be sufficient to modulate embryogenic competence, even if they remain undetectable by conventional statistical analyses.

This increase is correlated with the regulation of proteins involved in epigenetic metabolism and cell cycle control. For example, Adenosylhomocysteinase (Sspon.06G0013840-1A) appears to play a role in modulating DNA methylation by influencing the metabolism of S-adenosylmethionine, a key methyl-group donor in methylation reactions. In addition, the histone deacetylase HDT2 (Sspon.07G0003330-2B) and the histone-lysine N-methyltransferase, H3 lysine-9 specific SUVH1-like (Sspon.02G0022290-3C), suggest critical epigenetic regulation by modulating chromatin compaction and gene expression during the early phase of somatic embryogenesis.

Proteins associated with hormonal regulation and cell cycle control, such as SKP1-like protein 1 (Sspon.04G0021150-1A, Sspon.06G0013850-2B) and Protein TOPLESS (Sspon.01G0006700-1A), also exhibited significant accumulation, suggesting their involvement in the integration of hormonal and epigenetic signals required for the activation and growth of embryogenic calli. In parallel, phosphorylation and dephosphorylation events mediated by phosphatases, including serine/threonine-protein phosphatases PP2A-4, PP2A-2, and PP1 (Sspon.08G0006940-2C, Sspon.01G0017760-1A, Sspon.01G0030520-1P, Sspon.01G0003870-1A), appear to contribute to the transition from vegetative metabolism to

embryogenic development, possibly by fine-tuning hormonal responses during somatic embryo induction.

Overall, the findings of this study indicate that DNA methylation and the regulation of proteins involved in cell cycle progression and hormonal signaling play essential roles in controlling somatic embryogenesis in sugarcane. The interaction between these proteins and methylation pathways suggests a complex molecular regulatory network that facilitates cellular reprogramming and the acquisition of embryogenic competence, providing a robust foundation for optimizing somatic embryogenesis protocols and strategies for the propagation and improvement of *Saccharum spp.* L.

Conclusions

Our study sheds light on key aspects of sugarcane somatic embryogenesis, providing insights into the molecular mechanisms underlying this process. During the induction of embryogenic calli, proteins associated with energy metabolism and photosynthesis are downregulated, reflecting the deactivation of the photosynthetic apparatus and adaptation to a heterotrophic metabolism. The accumulation of phosphatases like PP2A-4, PP1, and PP2A-2, suggests their involvement in hormonal responses and the transition from vegetative growth to embryonic development.

Proteins involved in auxin and cell cycle regulation, such as like SKP1-like protein 1, TOPLESS, and Histone deacetylase HDT2, which integrate hormonal signaling and epigenetic regulation, exhibit interactions that likely contribute to the activation and growth of embryogenic callus. Proteins involved in DNA methylation, RNA messenger processing, histone acetylation, and cell cycle regulation accumulate at the onset of induction treatment. The Argonaute 4A-like protein and Adenosylhomocysteinase protein are implicated in gene silencing and embryogenic development, while histone methylation mediated by SUVH1-like protein may play a role in cell division, expansion, and differentiation.

Our study reveals that global DNA methylation levels remain relatively stable but increasing during the induction phase of somatic embryogenesis in *Saccharum spp.* L., aligning with methylation levels observed in other plant species with large genomes. Although no statistically significant differences were detected over time, the correlation between methylation dynamics and the accumulation of proteins involved in epigenetic regulation, cell cycle control, and hormonal signaling suggests that subtle epigenetic modulations may contribute to the acquisition of embryogenic competence. The involvement of key regulators,

such as Adenosylhomocysteinase, HDT2, SUVH1-like, SKP1-like proteins, and serine/threonine phosphatases, points to a coordinated network integrating chromatin remodeling, signal transduction, and metabolic reprogramming.

These findings provide a foundation for further research into the specific targets and functions of these molecular components, potentially leading to the optimization of somatic embryogenesis protocols and the development of novel strategies for sugarcane propagation and breeding.

References

Agius, D.R. *et al.* (2023) 'Exploring the crop epigenome: a comparison of DNA methylation profiling techniques', *Frontiers in Plant Science*, 14, p. 1181039. doi:<https://doi.org/10.3389/fpls.2023.1181039>

Afreen F, Zobayed SMA & Kozai T (2002) Photoautotrophic culture of *Coffea arabusta* somatic embryos: Photosynthetic ability and growth of different stage embryos. *Annals of Botany* 90(1):11-19 doi:<https://doi.org/10.1093/aob/mcf150>

Almeida FA, Passamani LZ, Santa-Catarina C, Mooney BP, Thelen JJ & Silveira V (2020) Label-free quantitative phosphoproteomics reveals signaling dynamics involved in embryogenic competence acquisition in sugarcane. *J. Proteome Res.* 19(10):4145-4157 doi:<https://doi.org/10.1021/acs.jproteome.0c00652>

Alonso, C. *et al.* (2015) 'Global DNA cytosine methylation as an evolving trait: phylogenetic signal and correlated evolution with genome size in angiosperms', *Frontiers in Genetics*, 6, p. 4. doi:<https://doi.org/10.3389/fgene.2015.00004>

Ballesteros I, Domínguez T, Sauer M, Paredes P, Duprat A, Rojo E, Sanmartín M & Sánchez-Serrano JJ (2013) Specialized functions of the PP2A subfamily II catalytic subunits PP2A-C3 and PP2A-C4 in the distribution of auxin fluxes and development in *Arabidopsis*. 73(5):862-872 doi:<https://doi.org/10.1111/tpj.12078>

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 72(1):248-254 doi:[https://doi.org/10.1016/0003-2697\(76\)90527-3](https://doi.org/10.1016/0003-2697(76)90527-3)

Distler U, Kuharev J, Navarro P, Levin Y, Schild H & Tenzer S (2014) Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat. Methods* 11(2):167-70 doi:<https://doi.org/10.1038/nmeth.2767>

Doyle, J.J. and Doyle, J.L. (1987) 'A rapid DNA isolation procedure for small quantities of fresh leaf tissue', *Phytochemical bulletin* [Preprint].

Fehér A (2015) Somatic embryogenesis, Stress-induced remodeling of plant cell fate. *Biochim. Biophys. Acta - Gene Regul. Mech.* 1849(4):385-402 doi:<https://doi.org/10.1016/j.bbagrm.2014.07.005>

Fehér A (2019) Callus, Dedifferentiation, Totipotency, Somatic Embryogenesis: What these terms mean in the era of molecular plant biology? *Front. Plant Sci.* 10(April):536 doi:<https://doi.org/10.3389/fpls.2019.00536>

Fraga, H.P.F. *et al.* (2012) '5-Azacytidine combined with 2, 4-D improves somatic embryogenesis of *Acca sellowiana* (O. Berg) Burret by means of changes in global DNA methylation levels', *Plant Cell Reports*, 31, pp. 2165–2176. doi:<https://doi.org/10.1007/s00299-012-1327-8>

Guo X, Qin Q, Yan J, Niu Y, Huang B, Guan L, Li Y, Ren D, Li J & Hou S (2015) TYPE-ONE PROTEIN PHOSPHATASE4 regulates pavement cell interdigitation by modulating PIN-FORMED1 polarity and trafficking in *Arabidopsis*. *Plant Physiol* 167(3):1058-75 doi:<https://doi.org/10.1104/pp.114.249904>

Heringer AS, Barroso T, Macedo AF, Santa-Catarina C, Souza GHMF, Floh EIS, de Souza-Filho GA & Silveira V (2015) Label-free quantitative proteomics of embryogenic and non-embryogenic callus during sugarcane somatic embryogenesis. *PloS one* 10(6):e0127803 doi:<https://doi.org/10.1371/journal.pone.0127803>

Heringer AS, Reis RS, Passamani LZ, de Souza-Filho GA, Santa-Catarina C & Silveira V (2017) Comparative proteomics analysis of the effect of combined red and blue lights on sugarcane somatic embryogenesis. *Acta Physiol. Plant* 39(2) doi:<https://doi.org/10.1007/s11738-017-2349-1>

Hutvagner G & Simard MJ (2008) Argonaute proteins: key players in RNA silencing. *Nature Reviews Molecular Cell Biology* 9(1):22-32 doi:<https://doi.org/10.1038/nrm2321>

Jericó JB-B, Maurilio M-M & Juan AP-S (2018) In vitro propagation of sugarcane for certified seed production. *Sugarcane:Ch. 6* doi:<http://dx.doi.org/10.5772/intechopen.74037>

Kumar L & M EF (2007) Mfuzz: a software package for soft clustering of microarray data. *Bioinformatics* 2(1):5-7 doi:<https://doi.org/10.6026/97320630002005>

Murashige T & Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* 15(3):473-497 doi:<https://doi.org/10.1111/j.1399-3054.1962.tb08052.x>

Naumann K, Fischer A, Hofmann I, Krauss V, Phalke S, Irmeler K, Hause G, Aurich AC, Dorn R, Jenuwein T & Reuter G (2005) Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in *Arabidopsis*. *The EMBO journal* 24(7):1418-29 doi:<https://doi.org/10.1038/sj.emboj.7600604>

Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G & Eisenacher M (2019) The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* 47(D1):D442-D450 doi:<https://doi.org/10.1093/nar/gky1106>

Pernas M, García-Casado G, Rojo E, Solano R & Sánchez-Serrano JJ (2007) A protein phosphatase 2A catalytic subunit is a negative regulator of abscisic acid signalling. *The Plant journal : for cell and molecular biology* 51(5):763-78 doi:<https://doi.org/10.1111/j.1365-313x.2007.03179.x>

Peters L & Meister G (2007) Argonaute proteins: mediators of RNA silencing. *Molecular cell* 26(5):611-23 doi:<https://doi.org/10.1016/j.molcel.2007.05.001>

Polesi LG, Fraga HPdF, Almeida FA, Silveira V & Guerra MP (2023) Comparative proteomic analysis and antioxidant enzyme activity provide new insights into the embryogenic competence of *Guadua chacoensis* (Bambusoideae, Poaceae). *Journal of Proteomics* 273:104790 doi:<https://doi.org/10.1016/j.jprot.2022.104790>

Rao V, Petla BP, Verma P, Salvi P, Kamble NU, Ghosh S, Kaur H, Saxena SC & Majee M (2018) *Arabidopsis* SKP1-like protein13 (ASK13) positively regulates seed germination and seedling growth under abiotic stress. *Journal of Experimental Botany* 69(16):3899-3915 doi:<https://doi.org/10.1093/jxb/ery191>

Reis RS, Vale Ede M, Heringer AS, Santa-Catarina C & Silveira V (2016) Putrescine induces somatic embryo development and proteomic changes in embryogenic callus of sugarcane. *J. Proteomics* 130:170-9 doi:<https://doi.org/10.1016/j.jprot.2015.09.029>

Reis RS, Vale EM, Sousa KR, Santa-Catarina C & Silveira V (2021) Pretreatment free of 2,4-dichlorophenoxyacetic acid improves the differentiation of sugarcane somatic embryos by affecting the hormonal balance and the accumulation of reserves. *Plant Cell Tissue Organ Cult.* 145(1):101-115 doi:<https://doi.org/10.1007/s11240-020-01995-z>

Rival A, Beulé T, Lavergne D, Nato A, Havaux M & Puard M (1997) Development of photosynthetic characteristics in oil palm during in vitro micropropagation. *Journal of Plant Physiology* 150(5):520-527 doi:[https://doi.org/10.1016/S0176-1617\(97\)80314-8](https://doi.org/10.1016/S0176-1617(97)80314-8)

Suprasanna P & Bapat VA (2006) Differential gene expression during somatic embryogenesis. *Somatic Embryogenesis*:305-320 doi:https://doi.org/10.1007/7089_038

Szemenyei H, Hannon M & Long JA (2008) TOPLESS mediates auxin-dependent transcriptional repression during *Arabidopsis* embryogenesis. *Science* 319(5868):1384-1386 doi:<https://doi.org/10.1126/science.1151461>

Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, Jensen LJ & von Mering C (2020) The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49(D1):D605-D612 doi:<https://doi.org/10.1093/nar/gkaa1074>

Team, R.C. (2022) 'A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2022'. URL:<https://www.R-project.org/>

Thirugnanasambandam PP, Hoang NV & Henry RJ (2018) The challenge of analyzing the sugarcane genome. *PLoS* 9 doi:<https://doi.org/10.3389/fpls.2018.00616>

Tzafrir I, Pena-Muralla R, Dickerman A, Berg M, Rogers R, Hutchens S, Sweeney TC, McElver J, Aux G, Patton D & Meinke D (2004) Identification of genes required for embryo development in *Arabidopsis* *Plant Physiology* 135(3):1206-1220 doi:<https://doi.org/10.1104/pp.104.045179>

Vizcaíno JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Ríos D, Dienes JA, Sun Z, Farrah T, Bandeira N, Binz P-A, Xenarios I, Eisenacher M, Mayer G, Gatto L, Campos A, Chalkley RJ, Kraus H-J, Albar JP, Martinez-Bartolomé S, Apweiler R, Omenn GS, Martens L, Jones AR & Hermjakob H (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* 32(3):223-226 doi:<https://doi.org/10.1038/nbt.2839>

Wickramasuriya AM & Dunwell JM (2015) Global scale transcriptome analysis of Arabidopsis embryogenesis in vitro. BMC Genom. 16(1):301 doi:<https://doi.org/10.1186/s12864-015-1504-6>

Wiśniewski JR, Zougman A, Nagaraj N & Mann M (2009) Universal sample preparation method for proteome analysis. Nat. Methods 6(5):359-62 doi:<https://doi.org/10.1038/nmeth.1322>

Wójcik AM, Wójcikowska B & Gaj MD (2020) Current perspectives on the auxin-mediated genetic network that controls the induction of somatic embryogenesis in plants. 21(4):1333 doi:<https://doi.org/10.3390/ijms21041333>

Wójcikowska B, Botor M, Morończyk J, Wójcik AM, Nodzyński T, Karcz J & Gaj MD (2018) Trichostatin A triggers an embryogenic transition in Arabidopsis explants via an auxin-related pathway. 9 doi:<https://doi.org/10.3389/fpls.2018.01353>

Wójcikowska B, Wójcik AM & Gaj MD (2020) Epigenetic regulation of auxin-induced somatic embryogenesis in plants. Int. J. Mol. Sci. 21(7):2307 doi:<https://doi.org/10.3390/ijms21072307>

Wu K, Tian L, Malik K, Brown D & Miki B (2000) Functional analysis of HD2 histone deacetylase homologues in Arabidopsis thaliana. 22(1):19-27 doi:<https://doi.org/10.1046/j.1365-313x.2000.00711.x>

Xavier LR, Almeida FA, Pinto VB, Passamani LZ, Santa-Catarina C, de Souza Filho GA, Mooney BP, Thelen JJ & Silveira V (2022) Integrative proteomics and phosphoproteomics reveals phosphorylation networks involved in the maintenance and expression of embryogenic competence in sugarcane callus. J. Plant Physiol. 268:153587 doi:<https://doi.org/10.1016/j.jplph.2021.153587>

Xavier LR, Corrêa CCG, da Paschoa RP, Vieira KdS, Pacheco DDR, Gomes LdES, Duncan BC, da Conceição LdS, Pinto VB, Santa-Catarina C & Silveira V (2023) Time-dependent proteomic signatures associated with embryogenic callus induction in *Carica papaya* L. Plants 12(22):3891 doi:<https://doi.org/10.3390/plants12223891>

Xu X, Legay S, Berni R, Hausman J-F & Guerriero G (2021) Transcriptomic changes in internode explants of stinging nettle during callogenesis. 22(22):12319 doi:<https://doi.org/10.3390/ijms222212319>

Zhang JZhang XTang HZhang QHua XMa XZhu FJones TZhu XBowers JWai CMZheng CShi YChen SXu XYue JNelson DRHuang LLi ZXu HZhou DWang YHu WLin JDeng YPandey NMancini MZerpa DNguyen JKWang LYu LXin YGe LArro JHan JOChakrabarty SPushko MZhang WMa YMa PLv MChen FZheng GXu JYang ZDeng FChen XLiao ZZhang XLin ZLin HYan HKuang ZZhong WLiaw PWang GYuan YShi JHou JLin JJin JCao PShen QJiang QZhou PMa YZhang XXu RLiu JZhou YJia HMa QQi RZhang ZFang JFang HSong JWang MDong GWang GChen ZMa TLiu HDhungana SRHuss SEYang XSharma ATrujillo JHMartinez MCHudson MRiascos JJSchuler MChen L-QBraun DMLi LYu QWang JWang KSchatz MCheckerman DVan Sluys M-ASouza GMMoore PHSankoff DVanBuren RPaterson AHNagai C & Ming R (2018) Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. Nature Genetics 50(11):1565-1573 doi:<https://doi.org/10.1038/s41588-018-0237-2>

Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, Zhang P, Banf M, Dai X, Martin GB, Giovannoni JJ, Zhao PX, Rhee SY & Fei Z (2016) iTAK: A program for genome-wide prediction

and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol. Plant* 9(12):1667-1670 doi:<https://doi.org/10.1016/j.molp.2016.09.014>

Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C & Chanda SK (2019) Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10(1):1523 doi:<https://doi.org/10.1038/s41467-019-09234-6>

**Capítulo 6. Integration of epigenetic remodeling and phosphorylation networks
drives embryogenic identity in sugarcane (*Saccharum* spp.) callus**

Integration of epigenetic remodeling and phosphorylation networks drives embryogenic identity in sugarcane (*Saccharum* spp.) callus

Lucas Rodrigues Xavier ¹, Lucas Gomes do Espirito Santo ¹, Vitor B. Pinto ²,
Claudete Santa-Catarina ², Jay J. Thelen ³, Vanildo Silveira¹

¹Laboratório de Biotecnologia, Centro de Biociências e Biotecnologia (CBB), Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), Campos dos Goytacazes, RJ, 28013-602, Brazil

²Laboratório de Biologia Celular e Tecidual, CBB-UENF, Campos dos Goytacazes, RJ, Brazil

³Department of Biochemistry, Christopher S. Bond Life Sciences Center, University of Missouri, 1201 Rollins Street, 65211, Columbia, Missouri, United States of America

Abstract

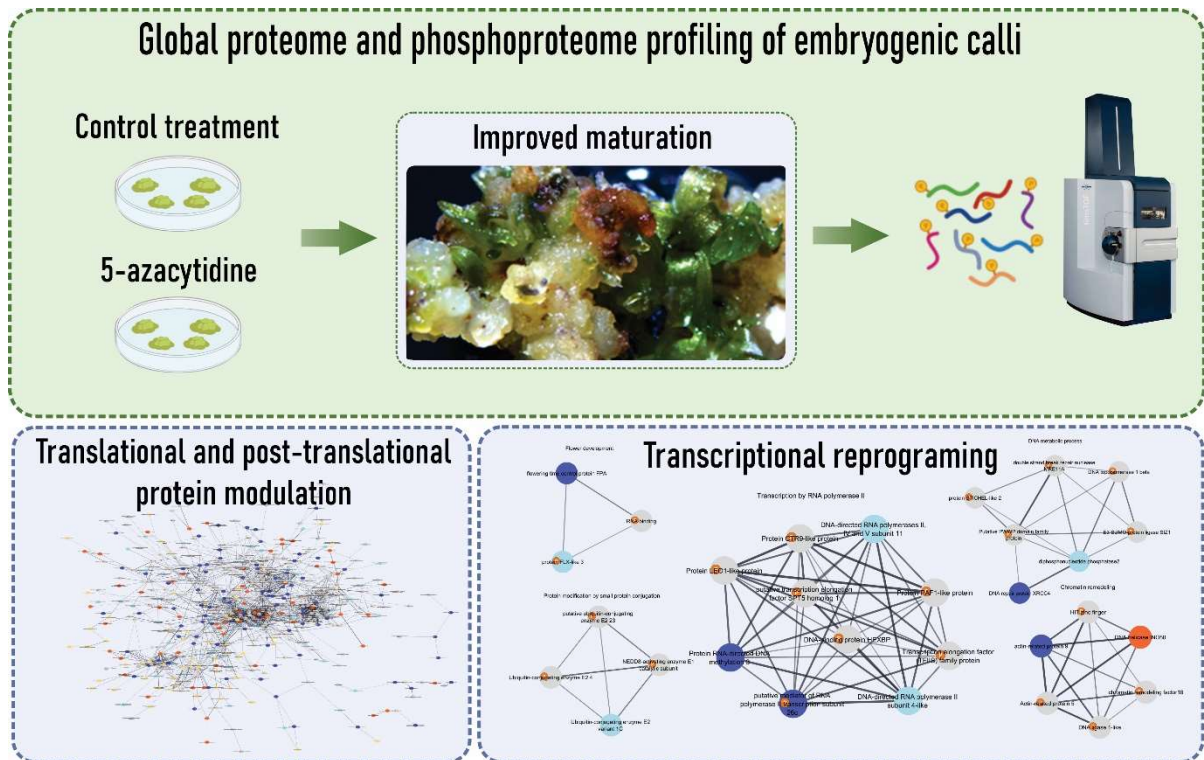
Micropropagation enables the large-scale production of genetically uniform and pathogen-free plants, offering significant advantages for agronomic improvement and biotechnological applications. In *Saccharum* spp., somatic embryogenesis (SE) represents a key regeneration pathway, yet its efficiency remains limited by the embryogenic potential of cultured tissues. Here, we investigated the effects of DNA hypomethylation induced by 5-Azacytidine (5-AzaC) in combination with 2,4-D on embryogenic competence in sugarcane calli, using comparative proteomic and phosphoproteomic approaches. Treatment with 50 μ M 5-AzaC significantly increased the number of mature somatic embryos, indicating enhanced cellular reprogramming. A total of 7,106 proteins were identified, with 427 differentially accumulated proteins (DAPs) between control and treated calli. Phosphoproteomic analysis identified 5,324 phosphopeptides, of which 3,028 had a phosphorylation probability >0.75 , with 650 differentially regulated. Key regulators, such as GTE4, ZAR1, and CDKF-1, were upregulated, implicated in chromatin remodeling, meristem identity, and cell cycle progression. Phosphorylation of auxin transporters (ABCB1, ABCB19, PIN1, PIN4) highlighted their role in auxin flux modulation and embryogenic cell fate specification. Additionally, phosphorylation of metabolic enzymes (VHA-F, VDAC4, AMPD) and ABA-related factors (ABI3, SRK2E, SUA) suggested metabolic and hormonal reprogramming during embryo induction. Transcription factors linked to meristem maintenance (HDG1, HDG5, LUG, NERD) were differentially phosphorylated, supporting a root apical meristem (RAM)-like program. Despite no significant changes in global DNA methylation, post-translational modifications indicated a compensatory mechanism fine-tuning epigenetic states. These findings demonstrate that 5-AzaC enhances embryogenic competence via phosphorylation-dependent regulation of key developmental networks, offering insights for optimizing regeneration protocols in sugarcane and other recalcitrant crops.

Keywords: Micropropagation; Bottom-up proteomics; Embryogenesis.

Corresponding author: Dr. Vanildo Silveira

E-mail: vanildo@uenf.br

Graphical abstract



Introduction

As a renewable and sustainable energy crop, sugarcane is pivotal not only for sugar and ethanol production but also as a feedstock for electricity generation and second-generation biofuels derived from lignocellulosic biomass [1]. These features place sugarcane at the forefront of strategies to reduce fossil fuel consumption and greenhouse gas emissions, with projections indicating it could replace up to 13.7% of global crude oil consumption and cut CO₂ emissions by 5.6% by 2045 [2]. To meet increasing demands for sustainable biomass and improved crop varieties, in vitro propagation has become a key tool in sugarcane biotechnology. These techniques facilitate the mass propagation of elite genotypes while capturing and fixing genetic gains [3]. The success of such biotechnological approaches hinges on precise control of in vitro morphogenesis, especially through somatic embryogenesis (SE), which is considered a cornerstone for effective regeneration systems and genetic transformation protocols in sugarcane [4]. Cultivars such as SP80-3280 have demonstrated consistent performance in in vitro regeneration and transformation studies, further supporting their utility in genomic studies [5].

Somatic embryogenesis is an intricate developmental process influenced by a network of signaling pathways and cellular reprogramming events. Central to SE induction are phytohormones such as auxins and cytokinins, which regulate cell division and tissue differentiation [6]. Typically, high concentrations of auxins (e.g., 2,4-D) are used to initiate embryogenic calli, followed by transfer to low-auxin or auxin-free media to promote embryo maturation [7]. Throughout this process, cellular and molecular reprogramming is tightly regulated, with a growing body of evidence pointing to epigenetic regulation as a key mechanism.

Epigenetic modifications, such as DNA methylation and histone modifications, have emerged as crucial regulators of gene expression during somatic embryogenesis. DNA methylation, in particular, influences the acquisition of embryogenic competence and morphogenic capacity in a dose- and context-dependent manner [8, 9]. The application of DNA methylation inhibitors like 5-azacytidine (5-AzaC) has shown promising results in enhancing SE in several plant species [10]. In sugarcane, although global methylation dynamics have been studied the specific role of DNA methylation in SE remains to be fully elucidated. Molecular components involved in epigenetic silencing include DNA methyltransferases (e.g., MET1, DRM1/2), histone deacetylases (HDACs), and polycomb group proteins (PRC1 and PRC2), which together orchestrate the remodeling of chromatin and regulation of gene expression [8, 9, 11]. The opposing Trithorax group (TrxG), histone acetyltransferases (HATs), and DNA demethylases (e.g., ROS1, DML2) act as transcriptional activators, ensuring a dynamic balance in chromatin state [12, 13]. Finding the correct balance between hypo- and hypermethylated DNA states and between acetylated and deacetylated chromatin is therefore crucial for successful SE [10].

Advances in proteomics have enabled the identification of key proteins involved in the acquisition of embryogenic competence. Differentially accumulated proteins during sugarcane SE include those related to auxin signaling (IAA-amido synthetases), stress response (peroxidases), lipid metabolism (oleosins), storage (vicilin-like), and chromatin remodeling (LEA proteins, HDA6, HDA19, TOPLESS) [14-17]. In this context, among post-translational modifications (PTMs), phosphorylation plays a pivotal role in modulating protein activity, stability, and interaction networks during developmental and stress responses [18]. The dynamic nature of phosphorylation, regulated by kinases and phosphatases, provides a rapid means of signal transduction in response to environmental cues [19]. Comparative phosphoproteomics allows for the large-scale identification of phosphorylation events, offering

insights into molecular mechanisms that are otherwise masked in traditional proteomic analyses.

Enrichment of phosphopeptides is critical for the detection of low-abundance phosphorylated proteins. Strategies such as Immobilized Metal Ion Affinity Chromatography (IMAC) and Metal Oxide Affinity Chromatography (MOAC) are routinely used to isolate phosphopeptides based on their affinity to metal ions (Fe^{3+} , Ga^{3+} , Al^{3+}) or metal oxides (TiO_2 , ZrO_2) [20]. These approaches, when integrated with gel-free mass spectrometry-based proteomics, provide robust datasets that reveal regulatory nodes and pathways relevant to SE and plant biotechnology [19].

In this context, our study applies label-free comparative proteomics to analyze sugarcane embryogenic calli treated with 2,4-D and the hypomethylating agent 5-AzaC. By integrating phosphoproteomic and proteomic data, we aim to identify novel biomarker candidates and gain a deeper understanding of how methylation states affect the protein signaling networks underpinning somatic embryogenesis in sugarcane. Such insights are essential for improving plant regeneration protocols, optimizing genetic transformation systems, and advancing sugarcane biotechnology as a whole.

Material and Methods

Obtention of explants

Callus induction was performed using the sugarcane cultivar SP-8032-80, following the protocol described by Passamani *et al.* 2018. Nodal segments were cultivated for 60 days in a greenhouse, after which internal immature leaf rolls were excised and used as explants. These explants were sealed with paraffin and subjected to surface sterilization by immersion in 70% ethanol (Sigma-Aldrich, St. Louis, USA) for 1 minute, followed by 30 minutes in 30% commercial bleach (sodium hypochlorite, 2-2.5%), and subsequently rinsed three times with autoclaved water under a laminar flow hood. The sterilized explants were transversely sectioned into 3 mm thick slices and cultured in glass test tubes (150 × 25 mm).

Induction and multiplication of embryogenic callus

Each test tube contained 10 mL of Murashige & Skoog (MS) basal medium (Phytotechnology Lab, Overland Park, USA) supplemented with 20 g L⁻¹ sucrose, 2 g L⁻¹ Phytigel® (Sigma-Aldrich), and 10 µM 2,4-dichlorophenoxyacetic acid (2,4-D; Sigma-Aldrich). The pH was adjusted to 5.8 prior to autoclaving (121°C for 15 min). Explants were incubated

in the dark at 25 °C for 45 days to induce callogenesis, leading to the formation of both embryogenic and non-embryogenic calli, as previously reported [14]. Calli were selected and transferred to Petri dishes (90 × 15 mm) containing 20 mL of the same induction medium. These calli were subcultured every 21 days for three successive cycles to promote multiplication. At the beginning of the third multiplication cycle, embryogenic and non-embryogenic calli were treated with 0 (control), 1, 10, 50, and 100 µM of 5-AzaC to investigate the effects of this hypomethylation agent in competence embryogenic induced by 2,4-D.

Maturation of embryogenic callus

Following the multiplication phase, four embryogenic calli (250 mg each) were transferred to fresh MS medium lacking 2,4-D and maintained in darkness for 21 days, as described by Passamani *et al.* 2018. Subsequently, the calli were transferred to a hormone-free maturation medium containing 30 g L⁻¹ sucrose and 2 g L⁻¹ Phytigel, and incubated in darkness for 7 days. Afterward, the cultures were exposed to a 16-hour photoperiod for 42 days using GreenPower TLED 20-W WmB lamps (Koninklijke Philips Electronics NV, Amsterdam, Netherlands) at an intensity of 55 µmol m⁻² s⁻¹. Two sets of experiments were conducted to evaluate the effects of DNA methylation inhibitors. In the first, 5-AzaC was applied at concentrations of 0, 1, 10, 50, and 100 µM. In the second, a narrower range was tested: 0, 0.1, 1, and 10 µM. In both experiments, 0 µM served as the control, and each Petri dish was considered an experimental unit (n = 10). At the end of the maturation period, the number and mass of somatic embryos were recorded.

Bottom-up proteomics

Five biological replicates, each consisting of a pool of five calli, were macerated in liquid nitrogen. From each replicate, 300 mg of powdered tissue was homogenized in an extraction buffer containing 7 M urea, 2 M thiourea, 2% Triton X-100, 1% dithiothreitol (DTT) (all from GE Healthcare), and 1 mM phenylmethanesulfonyl fluoride (PMSF; Sigma-Aldrich). The homogenate was vortexed thoroughly and then centrifuged to remove debris. The resulting supernatant was collected, and the total protein concentration was quantified using the 2-D Quant Kit (GE Healthcare). Tryptic digestion was performed using the filter-aided sample preparation (FASP) method as originally described by Wiśniewski *et al.* (2009), with adaptations from Almeida *et al.* 2020. A total of 1600 µg of protein from each biological replicate (Embryogenic and non-embryogenic calli) was used. Due to the loading capacity of the filter units, each replicate was divided into four aliquots of 400 µg. Tryptic digestion was conducted separately for each aliquot, and the resulting peptides were subsequently pooled

per biological replicate. The combined peptides were vacuum-dried and resuspended in 800 μ L of a solution composed of 5% (v/v) acetonitrile (Thermo Fisher Scientific) and 0.1% (v/v) formic acid in MS-grade water (Sigma-Aldrich). Peptide concentration was estimated by measuring absorbance at 205 nm using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific). Samples were stored at -80 °C until LC-MS/MS analysis. After digestion, the resulting peptides were vacuum-dried and resuspended in 50 μ L of a solution containing 95:5 (v/v) water:acetonitrile with 0.1% formic acid. Peptide concentrations were estimated by measuring absorbance at 205 nm using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific). *Phosphopeptide Enrichment*

Phosphopeptides were enriched using the High-Select™ TiO₂ Phosphopeptide Enrichment Kit (Thermo Fisher Scientific) from trypsin-digested peptide samples. Aliquots of digested peptides from each biological replicate were vacuum-dried and resuspended in 150 μ L of a solution containing 5% (v/v) acetonitrile (Thermo Fisher Scientific) and 0.1% (v/v) formic acid in MS-grade water (Sigma-Aldrich). All enrichment steps were performed according to the manufacturer's instructions. After elution, phosphopeptides were vacuum-dried and resuspended in 25 μ L of a solution containing 5% (v/v) acetonitrile and 0.1% (v/v) formic acid in MS-grade water. The phosphopeptide concentration was estimated by measuring absorbance at 205 nm using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific). Samples were stored at -80 °C until analysis.

Sample analysis on the EvoSep timsTOF

Each aliquot was diluted to a final volume of 20 μ L in 5% formic acid and subsequently loaded onto EvoTip Pure tips (EvoSep, Odense, Denmark) following the manufacturer's protocol. The prepared tips were then stored in 0.1% formic acid until mass spectrometry analysis. Sample analysis was conducted using the EvoSep One liquid chromatography system (EvoSep) operated under the standard "30 samples per day" gradient program. Chromatographic separation was achieved using EvoSep 1137 Performance columns (15 cm $\times$ 150 μ m, packed with 1.5 μ m C18 particles), with mobile phases composed of 0.1% formic acid in water (aqueous phase) and 0.1% formic acid in acetonitrile (organic phase). The column was maintained at 40 °C in a Bruker Column Toaster and interfaced to a Bruker timsTOF Pro 2 mass spectrometer (Bruker, Billerica, MA, USA) via a 20 μ m silica emitter (Bruker CaptiveSpray ZDV Sprayer 20).

MS/MS spectra obtained from the proteomics and phosphoproteomics data processing and database searching were performed using MaxQuant software v.1.6.3.4 (Cox and Mann, 2008), following the procedures described by Almeida, Passamani [15]. Label-free quantification was conducted using the normalized label-free quantification (LFQ intensity) approach, based on the MaxLFQ algorithm [21]. Following data processing, the output file Phospho (STY)Sites.txt was filtered to retain only phosphorylation sites with a localization probability ≥ 0.75 . For differential regulation analysis in the treatment/control comparison, only phosphoproteins containing confidently identified phosphate-modified peptides and present in at least four biological replicates, or entirely absent across all replicates in at least one condition (for unique proteins), were considered. Phosphoproteins exhibiting statistically significant differences according to a two-tailed Student's t-test ($P < 0.05$) were classified as differentially regulated phosphoproteins (DRPs).

To infer protein-protein interaction networks, Zea mays homologs of the sugarcane DAPs were identified via STRING, and the associated interaction data were used to construct interaction networks [22]. Prediction of kinases, transcription factors (TFs), and transcriptional regulators (TRs) among the DAPs and differentially regulated phosphoproteins (DRPs) was conducted using the iTAK software [23].

The raw and processed proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE repository under the dataset identifier PXXXXXXX [24].

Results

Embryogenic callus differentiation

All multiplication treatments resulted in calli with a yellowish-white appearance and cell mass growth, with no evident morphological differences. After the 49-day maturation period, distinct chlorophyll spots were observed throughout the callus, and the number of somatic embryos was quantified (Figure 1). The 50 μM concentration of 5-AzaC yielded statistically superior results, with an average of 54 somatic embryos per callus, compared to the other treatments. In contrast, the treatment with 100 μM of 5-AzaC resulted in the lowest number of somatic embryos, with only 29 embryos per callus (Figure 2).

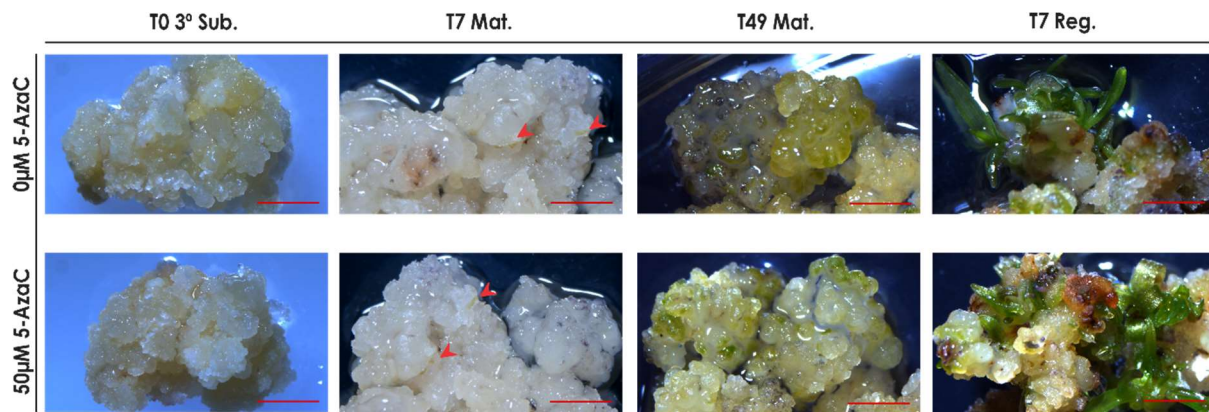


Figure 1. Morphology of sugarcane embryogenic calli and somatic embryos following treatment with 5-AzaC. Sugarcane embryogenic calli were cultured for 21 days in the third multiplication cycle in the presence of 0 (control), 1, 10, 50, and 100 μ M 5-AzaC. After maturation, the callus treated with 50 μ M of 5-AzaC displayed a higher number of somatic embryos compared to the control culture medium. Furthermore, treatment with 50 μ M 5-AzaC did not impair somatic embryo regeneration. The bar represents 1cm.

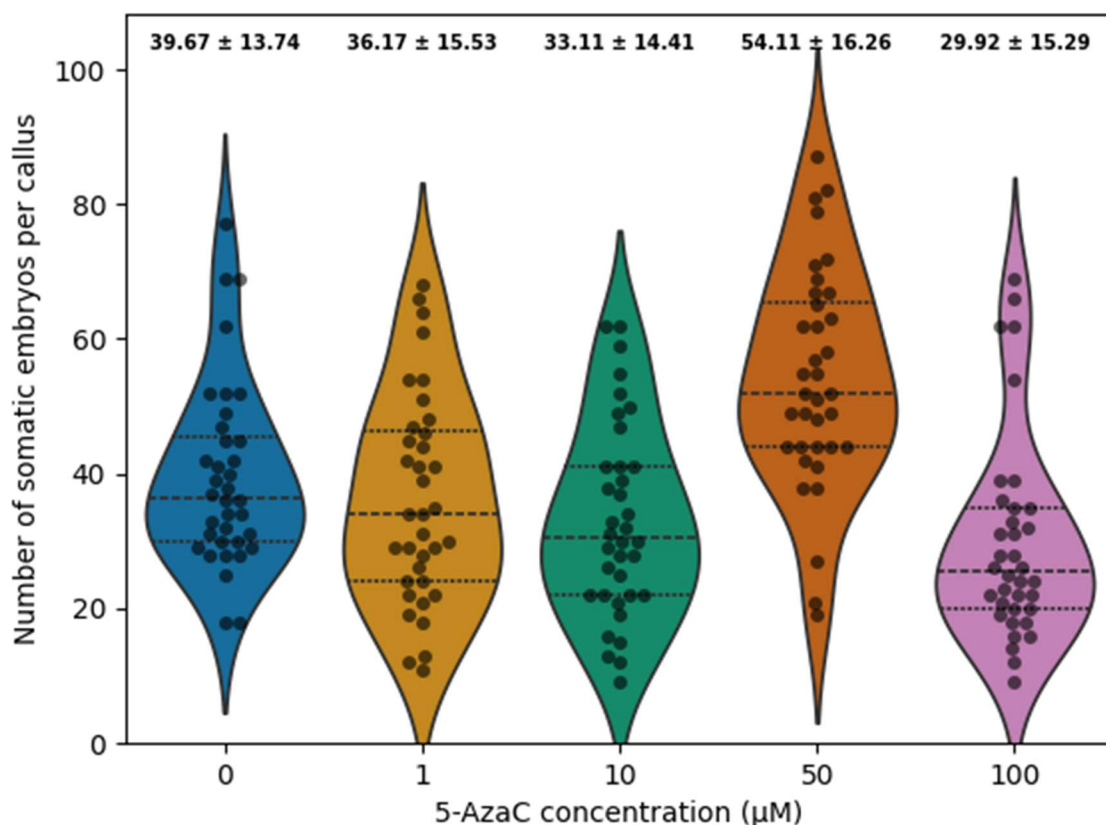


Figure 2. Number of somatic embryos after 49 days of maturation treatment in sugarcane callus. Data showing the number of somatic embryos in calli treated with Control, 1, 10, 50, and 100 µM of 5-AzaC after 49 days of maturation. A total of 180 calli were analyzed.

Since the treatment with 50 µM 5-AzaC resulted in higher averages compared to the control treatment (39 embryos per callus), we proceeded to regenerate the somatic embryos to evaluate whether this treatment could enhance the embryogenic competence of callus in somatic embryogenesis protocols. The regenerated embryos developed successfully and were similar to those from the control treatment (Figure 1). Based on these results, we selected the 50 µM 5-AzaC treatment for a comparative proteomic and phosphoproteomic analysis to identify regulatory factors associated with the increased embryogenic competence induced by this DNA hypomethylating agent.

Bottom-up proteomics and phosphoproteomics analysis

A total of 7,106 proteins were identified through bottom-up analysis (Supplementary Table 1). The comparative proteomic analysis revealed 427 differentially accumulated proteins (DAPs, including uniques) (Figure 3) (Supplementary table 1). Also, 5,324 phosphopeptides

(3,028 with phosphorylation probability >0.75) were identified, including 650 differentially regulated phosphopeptides (Supplementary table 2).

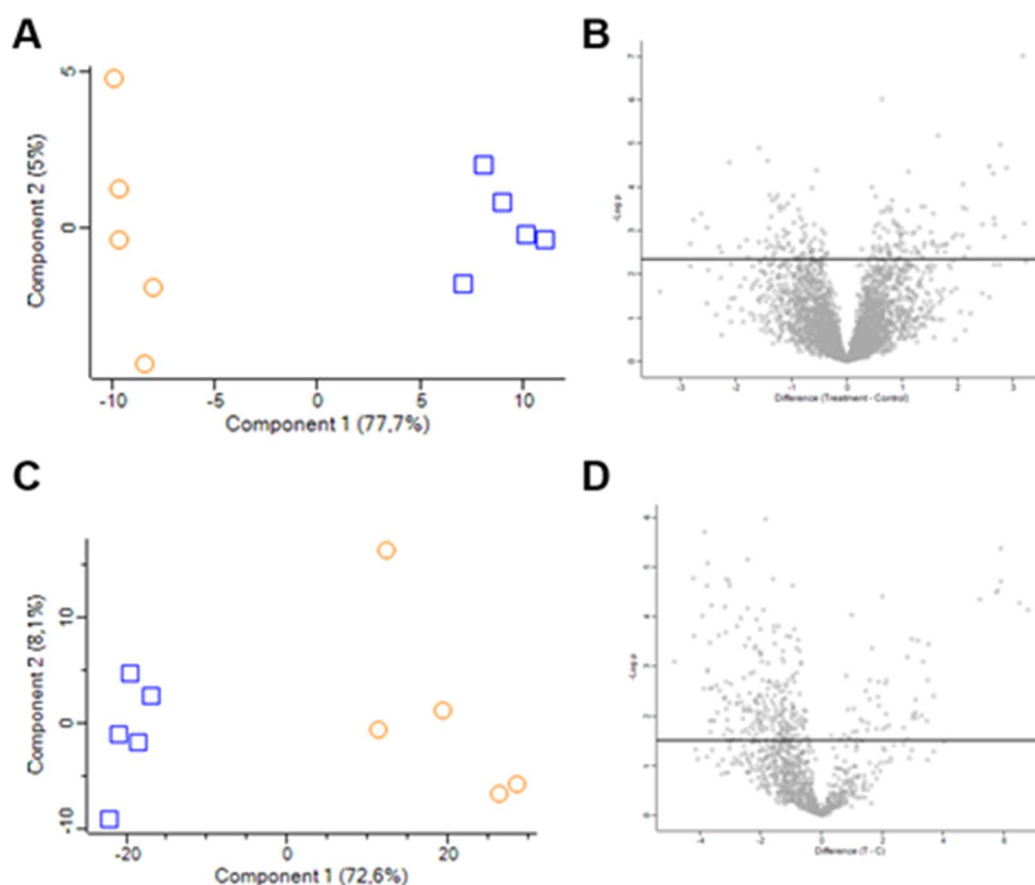


Figure 3. Overview of the proteomic and phosphoproteomic data obtained from the analysis of embryogenic and non-embryogenic papaya callus. Principal Component Analysis (PCA, A) shows that biological replicates are grouped based on the treatment. Volcano plots illustrate the differential accumulation and regulation of proteins (B) and phosphopeptides (C) in the comparative analysis between the callus types.

A total of 51 phosphorylation motifs were detected among these phosphopeptides (Table 1). The most prevalent phosphorylation motifs were pSP (1,059 occurrences), pSxS (846), pSxxS (592), pSG (268), and pSxSP (247). After filtering for phosphopeptides with at least 75% phosphorylation probability, 2,461 phosphorylation sites with high confidence were identified (Supplementary Table 2). Our analysis revealed 1,598 phosphopeptides with one phosphorylation site, 1,166 phosphopeptides with two phosphorylation sites, 223 with three phosphorylation sites, 41 with four phosphorylation sites, and three with five phosphorylation sites (Figure 4A). Furthermore, the majority of the phosphopeptides were phosphorylated on serine (S, 2,703 sites), followed by threonine (T, 278 sites) and tyrosine (Y, 50 sites) residues (Figure 4B).

Table 1. Phosphorylation motifs identified in the phosphoproteomics dataset.

Phosphorylation motif	Amino acid	Score	Number of phosphopeptides	Adjusted p-value
..F.....SP.....	S	251.48	26	2.3e-004
.....SP.H.....	S	238.05	26	2.3e-004
.....H.....SP.....	S	223.95	26	2.3e-004
.....F.....SP.....	S	207.28	21	6.1e-003
.....LSP.....	S	197.16	57	3.9e-008
.....SP.....H.....	S	177.84	26	1.7e-003
.....I..SP.....	S	163.95	28	4.7e-004
.....Y..SP.....	S	152.60	20	1.2e-002
..R.....SP.....	S	144.11	79	8.0e-012
.....P.SP.....	S	122.71	108	7.3e-015
.....Y...SP.....	S	98.25	22	8.1e-002
.....SP.....	S	86.33	807	1.1e-093
.....F..S.S.....	S	51.10	31	2.2e-002
.....SD.E...E.....	S	86.89	24	1.1e-003
M.....S.S.....	S	43.26	19	1.1e-001
.....SDEE.....	S	73.25	32	4.5e-005
.....S.S...D.....	S	40.01	60	8.9e-003
.....RS.S..P.....	S	37.03	19	1.3e-001
.....SD.E.....	S	38.78	136	7.8e-018
.....S.S..E.....	S	30.38	73	1.5e-004
.....SDSD.....	S	35.52	24	6.3e-003
.....L.R..S.....	S	26.65	59	3.0e-004
.....R.RS.S.....	S	34.09	23	1.7e-003
.....S.SS.....	S	22.13	94	2.9e-003
.....S.S.S.....	S	19.96	120	1.1e-005
...T.....GS.....	S	20.14	17	1.0e+000
.....SEDE.....	S	26.76	18	4.0e-002
.....GS.Y.....	S	17.38	12	1.0e+000
.....DSD.D.....	S	34.09	30	1.6e-005
.....V...GS.....	S	15.34	25	4.9e-002
.....S.S.....	S	7.40	576	7.5e-006
.....SSP.....	S	38.60	103	7.2e-010
.....S.D.....	S	10.65	330	2.4e-002
.....S.S.....	S	8.32	656	4.4e-016
.....S..S.....	S	7.45	458	9.8e-001
.....TP...S.....	T	45.49	20	6.8e-003
.....TP...S.....	T	33.30	16	8.8e-002
.....TP.....T.....	T	25.23	15	1.5e-001
.....TP.....	T	13.23	115	3.9e-013
...K.....T.....	T	7.34	29	9.8e-001
.....V.....T.....	T	6.43	35	1.6e-001
.....Y.....A.....	Y	6.25	9	4.2e-001

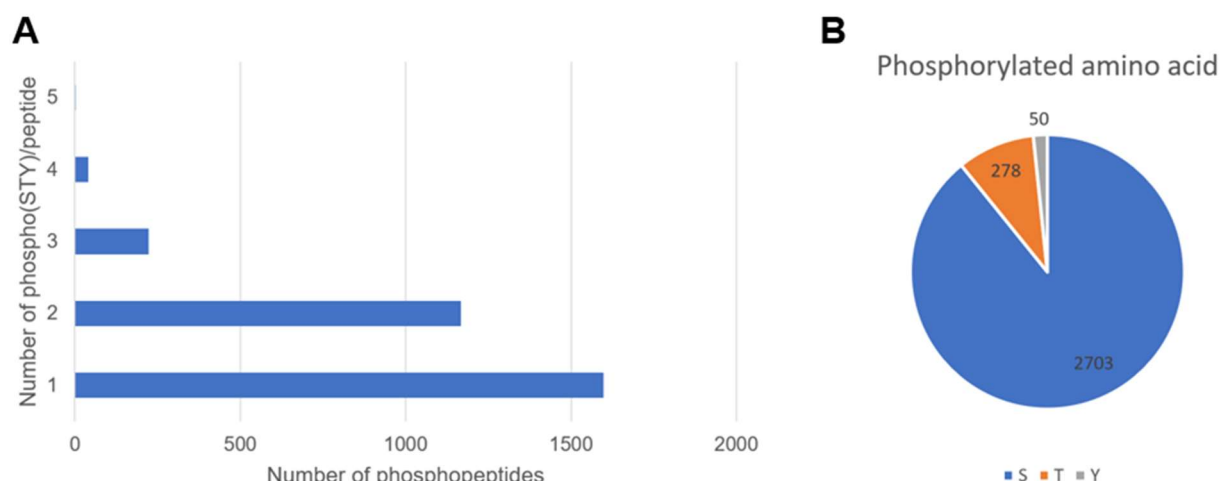


Figure 4. Bar graph showing the distribution of phosphorylation motifs identified in the control and 5-AzaC treatments, highlighting the features of the phosphorylation sites identified in the phosphoproteomic analysis. The bar graph (A) illustrates the number of simultaneous phosphorylation sites in the phosphopeptides, while the pie chart (B) displays the proportion of phosphorylated serine, threonine, and tyrosine residues (S, T, and Y, respectively).

Functional analysis of proteomic data

Gene ontology (GO) enrichment analysis of the differentially accumulated proteins (DAPs) in the control and 5-AzaC-treated calli revealed that detoxification, carbohydrate metabolism, and DNA modification and methylation were the most prominent biological processes (Figure 5A). Among the molecular functions, the most enriched terms included oxidoreductase activity, xyloglycan transferase activity, ligand binding, glycosyltransferase activity, and carbohydrate binding (Figure 5B). The most enriched cellular compartments among the DAPs were the extracellular region, apoplast, and cell lumen (Figure 5C). Furthermore, the most enriched metabolic pathways were related to the biosynthesis of secondary metabolites, starch and sucrose metabolism, and purine metabolism (Figure 5D).

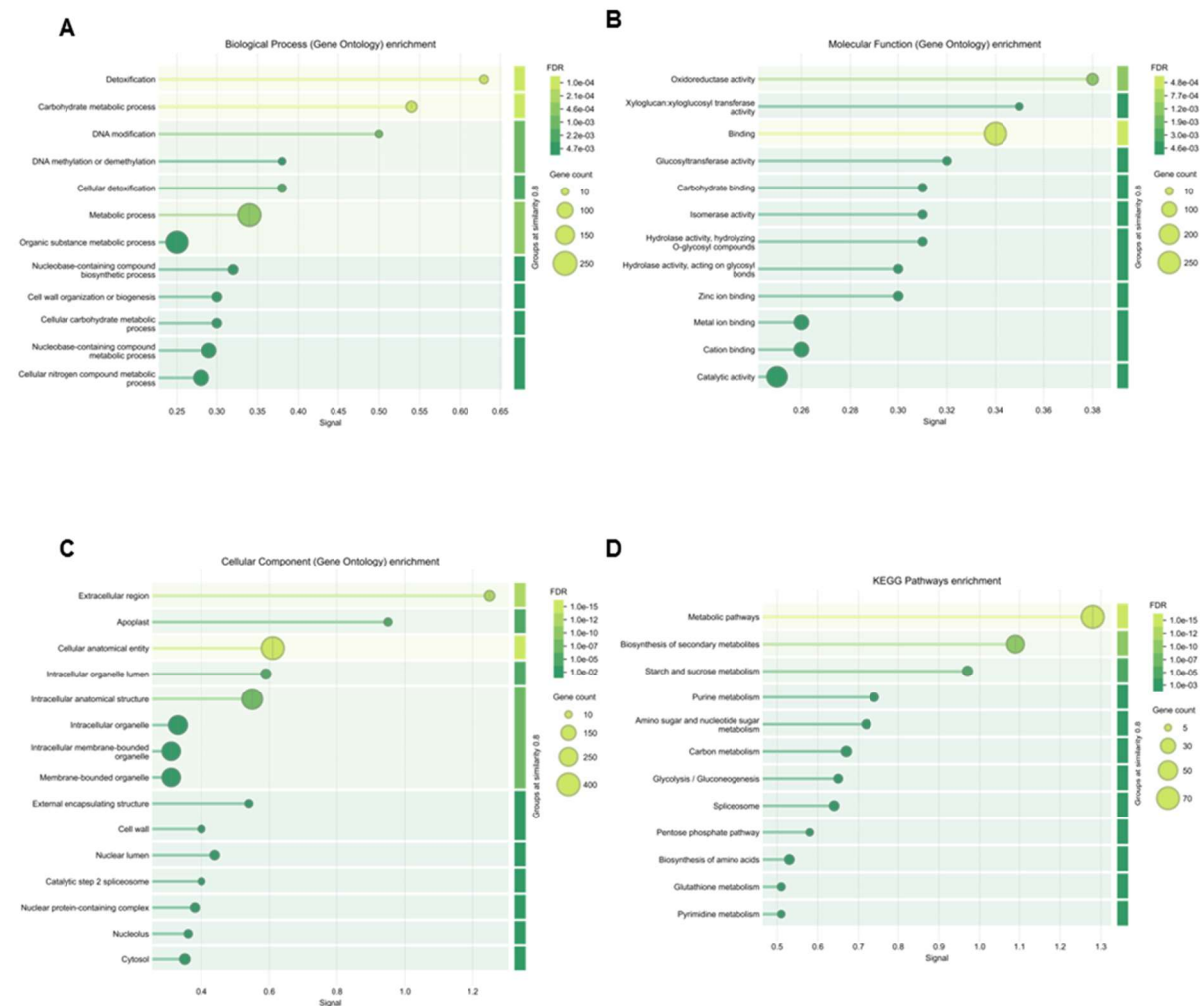


Figure 5. Enrichment analysis of gene ontology terms and metabolic pathways among the proteins differentially accumulated between embryogenic and non-embryogenic callus. The lollipop plots illustrate the enriched biological process terms (A), molecular functions (B), cellular components (C), and KEGG pathways (D) among the DAPs.

We performed an enrichment analysis of gene ontology terms among the differentially regulated phosphoproteins (DRPs) and found that RNA processing, regulation of gene expression, macromolecule metabolism, and regulation of metabolic processes were the most enriched biological processes (Figure 6A). The most enriched molecular functions among the DRPs included RNA binding, nucleic acid binding, RNA polymerase binding, and binding to heterocyclic and organic cyclic compounds (Figure 6B). The most enriched cellular compartments were the nuclear lumen, intracellular organelle lumen, nucleus, nucleoplasm, and intracellular anatomical structure (Figure 6C). The only KEGG pathway enriched among the DRPs was the spliceosome (Figure 6D).

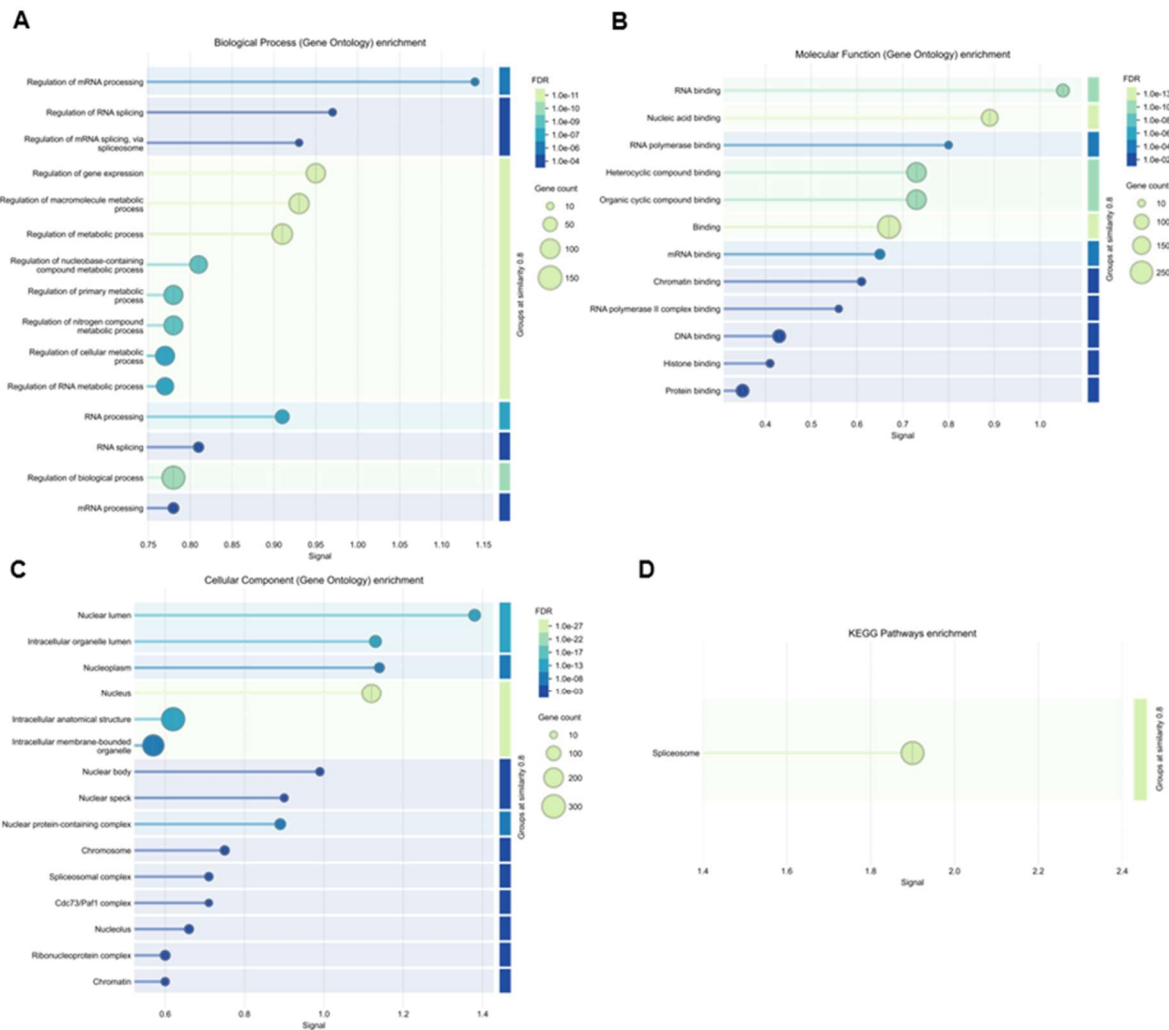


Figure 6. Enrichment analysis of gene ontology terms and metabolic pathways among the phosphopeptides regulated between embryogenic and non-embryogenic callus. The lollipop plots illustrate the enriched biological process terms (A), molecular functions (B), cellular components (C), and KEGG pathways (D) among the phosphopeptides.

Identification of regulatory proteins

Through proteomic and phosphoproteomic analysis, we identified a diverse array of protein kinases, transcription factors, and transcriptional regulators in sugarcane embryogenic callus (Supplementary Table 1). Among the protein kinase families, the most abundant were RLK-Pelle (105 members), CAMK (26), CMGC (25), STE (16), and TKL-PI-4 (11). Regarding transcriptional regulators, the most represented families included C3H zinc finger proteins (32 members), bZIP transcription factors (19), bHLH proteins (15), Trihelix factors (15), SET domain-containing methyltransferases (14), and C2H2 zinc fingers (14).

Prediction of protein-protein interaction networks associated with DNA methylation responses during somatic embryo differentiation

Following the statistical analysis of the differentially accumulated proteins (DAPs) and differentially regulated phosphoproteins (DRPs), we constructed a protein-protein interaction network that integrates data on protein abundance and phosphorylation at serine, threonine, and tyrosine (STY) sites (Figure 7). This network comprises 726 nodes and 1872 edges, offering a detailed view of the complex epigenetic mechanisms governing multiple layers of gene expression. Cluster analysis further revealed functional modules within the interaction network, resulting in the identification of 19 distinct clusters (Figure 8). These clusters are primarily involved in key biological processes such as RNA processing, transport, flower development, carbohydrate metabolism, and DNA modification.

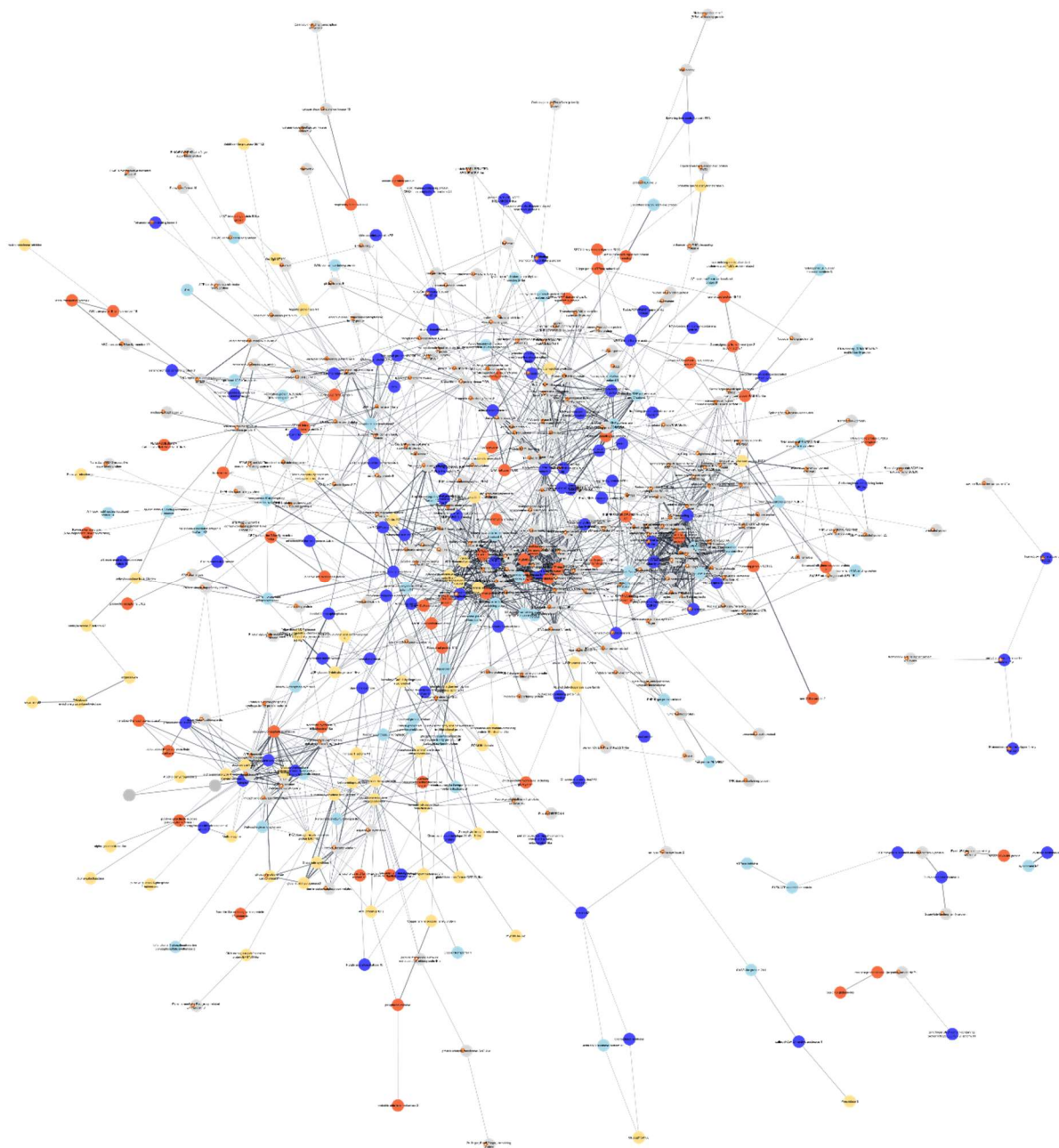


Figure 7. Prediction of interactions between regulated proteins and phosphoproteins in the comparative analysis of embryogenic and non-embryogenic callus. The orange and yellow nodes represent up-regulated and unique proteins in embryogenic calli, while the light and dark blue nodes correspond to down-regulated or unique proteins in non-embryogenic callus. Phosphoproteins with regulated phosphorylation sites are denoted with a 'P' inside the node.

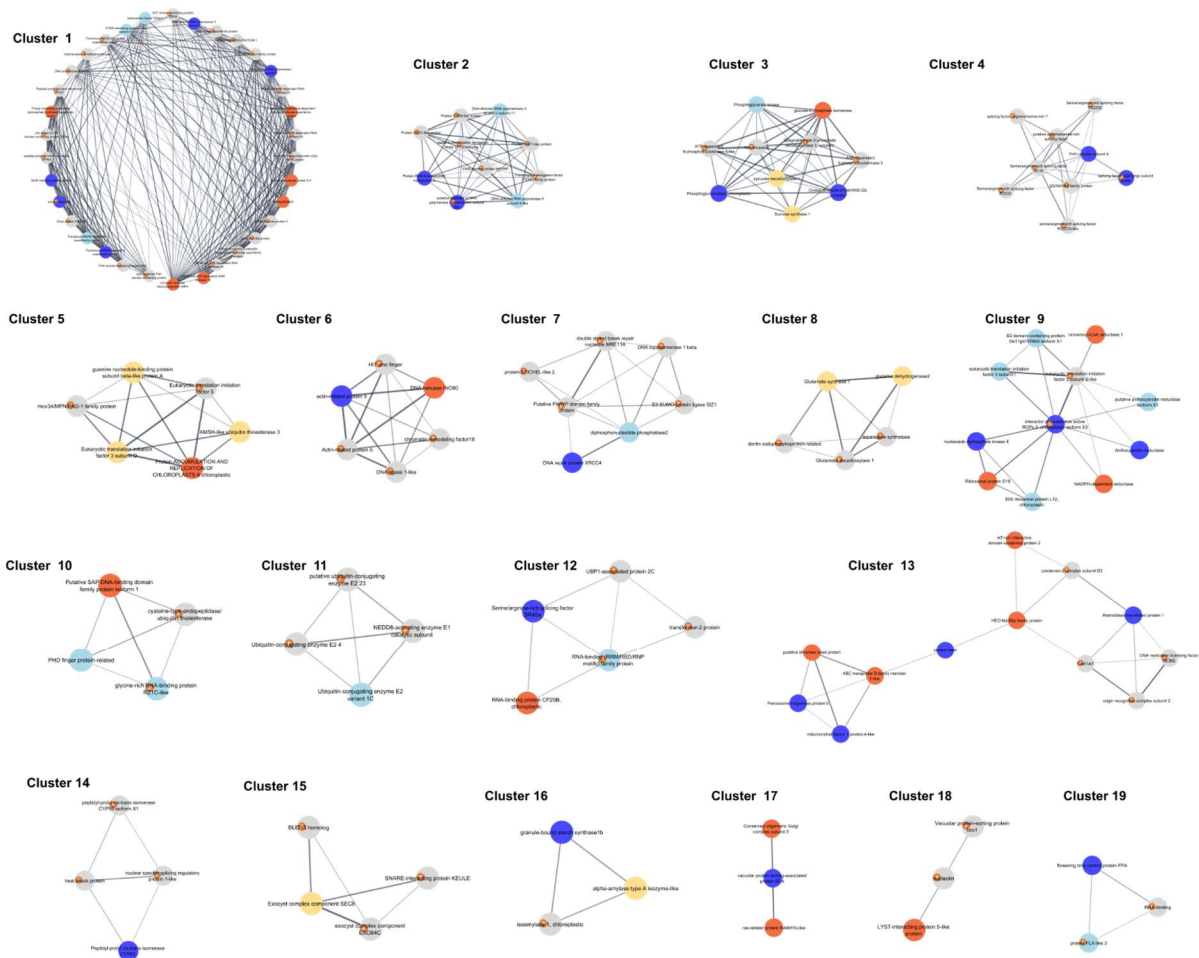


Figure 8. Analysis of molecular components and display of predicted clusters in the interaction network between regulated proteins and phosphoproteins in response to treatment with 5-AzaC. The orange and yellow nodes represent up-regulated and unique proteins in embryogenic calli, while the light and dark blue nodes represent down-regulated or unique proteins in non-embryogenic callus. Phosphoproteins with regulated phosphorylation sites are marked with a 'P' inside the node.

Beyond the functional analysis of proteins and phosphoproteins, the examination of predicted interactions between differentially accumulated proteins (DAPs) and differentially regulated phosphoproteins (DRPs) led to the identification of a list of potential targets that may play crucial roles in regulating the embryogenic competence of callus in response to 5-AzaC treatment (Table 2).

Table 2. List of DAPs and DRPs regulated in sugarcane callus after 5-AzaC treatment.

Protein ID	Arabidopsis orthologs	BLAST top hit (Viridiplantae)	DRP	Accumulation (treatment/control)
Sspon.07G0016240-2C	AGO1	AGO1a	Yes	Down
Sspon.05G0022390-2D	ABCB19	ABC transporter B family member 19	Unique Control	Down
Sspon.06G0031100-1C	ABI3	putative B3 DNA binding domain family protein	Unique Control	Down
Sspon.07G0011500-1P	CHR12	chromatin complex subunit A	Unique Control	Down
Sspon.06G0007620-4D	HDA6	Histone deacetylase	Yes	Down
Sspon.06G0011620-4D	LDL2	lysine-specific histone demethylase 1 homolog 2-like		Down
Sspon.03G0006600-1A	NERD	Zinc finger CCCH domain-containing protein 19	Unique Control	Down
Sspon.08G0012870-2B	PKL	Putative homeodomain-like transcription factor superfamily protein	Yes	Down
Sspon.02G0009460-2P	SHI	SHI	Yes	Down
Sspon.01G0014770-1A	SRK2E	putative serine/threonine-protein kinase	Unique Treatment	Down
Sspon.06G0020970-2C	SUVR4	Histone-lysine N-methyltransferase SUV4	Unique Control	Down
Sspon.06G0020980-3D	SUVR4	Histone-lysine N-methyltransferase SUV4		Down
Sspon.06G0010720-3D	TPR1	ramosa 1 enhancer locus 2	Yes	Down
Sspon.05G0023810-1B	VIP6	Protein CTR9-like protein	Yes	Down
Sspon.05G0005190-1A	AGO1	argonaute1c	Yes	Unique Control
Sspon.02G0009710-1A	HDG1	Homeodomain leucine zipper family IV protein	Unique Control	Unique Control
Sspon.03G0011690-2B	LUG	Transcriptional corepressor LEUNIG	Yes	Unique Control
Sspon.05G0020310-4D	PIN4	probable auxin efflux carrier component 1b	Unique Control	Unique Control
Sspon.07G0020350-3C	SGS3	protein SUPPRESSOR OF GENE SILENCING 3 homolog		Unique Control
Sspon.03G0030150-2D	SUVH1	Histone-lysine N-methyltransferase H3 lysine-9 specific SUVH3		Unique Control
Sspon.05G0027780-3D	SWI3D	SWI/SNF complex subunit SWI3D-like	Unique Control	Unique Control
Sspon.05G0027530-2D	ABCB19	ABC transporter B family member 19		Unique Treatment
Sspon.01G0057760-1D	MAIL3	putative polypeptide	Unique Treatment	Unique Treatment
Sspon.05G0032690-1C	MAIL3	serine/threonine-protein phosphatase 7 long form homolog		Unique Treatment
Sspon.02G0019430-4D	SUA	SUPPRESSOR OF ABI3-5 isoform X1	Yes	Unique Treatment
Sspon.03G0017500-4D	TPR2-2	topless-related1	Yes	Unique Treatment
Sspon.03G0003300-4D	VHA-a1	V-type proton ATPase subunit a1	Yes	Unique Treatment
Sspon.02G0000320-2D	AMPD	AMP deaminase	Yes	Up
Sspon.03G0002830-2C	APUM3	pumilio homolog 1-like isoform X1	Unique Control	Up
Sspon.08G0008450-2B	CDKF-1	CAK1AT	Unique Treatment	Up
Sspon.06G0010440-1A	CUL3A	Cullin-3B		Up
Sspon.02G0009460-1A	SRS1	Protein SHI RELATED SEQUENCE 1	Unique Control	Up
Sspon.06G0010180-1P	VIP2	Protein PAF1-like protein	Yes	Up
Sspon.06G0004630-3C	ABCB1	brachytic2	Unique Control	Unchanged
Sspon.04G0020470-3D	BRM	brahma1	Unique Treatment	Unchanged
Sspon.07G0010050-1P	EBP1	ERBB-3 BINDING PROTEIN 1	Unique Treatment	Unchanged
Sspon.03G0043250-1C	HDG1	homeobox-leucine zipper protein ROC5-like	Unique Control	Unchanged

Sspon.01G0015190-2B	HDG5	Homeodomain leucine zipper family IV protein	Unique Control	Unchanged
Sspon.01G0036140-2C	OBE4	Protein OBERON 4	Unique Treatment	Unchanged
Sspon.04G0003600-4D	PIN1	auxin efflux carrier component 1a	Unique Control	Unchanged
Sspon.04G0035280-1D	TSK	Protein TONSOKU	Unique Treatment	Unchanged
Sspon.02G0025830-1A	VIP4	Protein LEO1-like protein	Unique Treatment	Unchanged
Sspon.03G0013960-2B	ZAR1	receptor protein kinase-like protein ZAR1	Yes	Unchanged

Discussion

The interaction between 2,4-D and 5-AzaC alters the embryogenic competence of sugarcane callus

The combination of 2,4-D and 5-AzaC at specific concentrations has previously been shown to enhance somatic embryo formation in *Acca sellowiana* [25]. Similarly, in *Daucus carota*, transient expression of a DNA methyltransferase gene (Met1-5) was observed following the induction of somatic embryogenesis (SE) by 2,4-D, suggesting a role for DNA methylation in this process [26]. It is hypothesized that this methylation pattern is associated with a synergistic interaction between 5-AzaC and 2,4-D, a phenomenon also reported in somatic embryogenesis of *Cucurbita pepo* L. [27]. However, in sugarcane callus, no differences in DNA methylation status were detected between the control and treatment with 50 μ M 5-AzaC during the third subculture, suggesting a stage- or species-specific response.

DNA methylation has been widely recognized as an important regulatory mechanism during both zygotic and somatic embryogenesis [28]. However, the effects of 5-AzaC in combination with 2,4-D are still debated, with studies reporting both positive and neutral outcomes. For example, in *A. sellowiana*, the application of 200 μ M 2,4-D and 50 μ M 5-AzaC increased DNA methylation levels and enhanced the induction of somatic embryogenesis [25]. In contrast, in *D. carota*, treatment with 4.52 mM 2,4-D and 20.5 μ M 5-AzaC resulted in a similar number of embryos compared to the control (without 5-AzaC), and embryogenesis was actually inhibited by day 14 post-treatment in the absence of 2,4-D [26].

Treatment with 2,4-D has been linked to hypermethylation of promoter regions of key transcription factor genes involved in hormonal responses and stress signaling, such as LEC1, LEC2, BBM, AGL15, and WUS [8]. In cotton, hypomethylation was found to promote embryogenesis by activating WUS expression [29]. These genes play a central role in the regulatory network that governs somatic embryogenesis induction. Despite these findings, our

understanding of auxin's role, particularly 2,4-D, in the regulatory network of somatic embryogenesis remains incomplete. Beyond gene transcription, auxins also modulate a range of cellular processes, including chromatin state, protein stability, protein-protein interactions, and metabolic pathways [8].

Modulation of proteins involved in epigenetic responses increase somatic embryo differentiation

Epigenetic regulation plays a pivotal role in various stages of somatic embryogenesis, involving mechanisms such as histone acetylation and methylation, DNA methylation, and microRNA activity [6, 9]. In this context, ARGONAUTE (AGO) proteins are key regulators within the RNA interference (RNAi) pathway and epigenetic silencing machinery. Our results revealed phosphorylation of ARGONAUTE1c (Sspon.05G0005190-1A, ortholog of AtAGO1) at S1037, while AGO1a (Sspon.07G0016240-2C, ortholog of AtAGO4) was upregulated at T1048 and T1049. AGO1 primarily mediates microRNA-guided post-transcriptional gene silencing, a function essential for maintaining meristematic and totipotent cell identity during development [30, 31]. Conversely, AGO4 participates in RNA-directed DNA methylation (RdDM), contributing to transcriptional silencing through epigenetic modifications. Notably, AGO proteins act in a dose-dependent manner to regulate plant development and cellular totipotency, suggesting their critical involvement in embryogenic competence [31]. Post-translational modifications (PTMs), including phosphorylation, are known to modulate AGO1 function by influencing protein-protein and protein-RNA interactions [32]. Although studies in animal systems have demonstrated that AGO phosphorylation affects subcellular localization and protein stability, the functional consequences of such modifications in plants remain largely unexplored [33]. Taken together, our findings emphasize the potential regulatory roles of AGO protein phosphorylation and accumulation in somatic embryogenesis and highlight the need for further studies to elucidate their specific mechanisms of action in plant developmental reprogramming.

Our data indicate that a putative homeodomain-like transcription factor superfamily protein (Sspon.08G0012870-2B, orthologous to *AtPKL*) was regulated in sugarcane callus treated with 5-AzaC. PKL is a well-characterized chromatin remodeling factor known for its role in the embryogenesis-to-vegetative growth transition, where it actively represses

embryogenic identity genes during post-embryonic development [34]. Loss-of-function mutations in *PKL* are sufficient to trigger the formation of ectopic somatic embryos, highlighting its key role in maintaining developmental boundaries [6]. In *Arabidopsis*, *PKL*-defective mutants ectopically express embryogenesis-related genes such as *LEC1*, *LEC2*, and *FUS3* in vegetative tissues due to the failure to repress these targets after embryogenesis [35]. In our study, treatment with 5-AzaC resulted in a reduction in the abundance of this *PKL*-like protein, as well as regulation of its phosphorylation status, suggesting that both transcriptional and post-translational mechanisms are involved in its modulation. *PKL*'s repressor function is known to depend on the activity of HISTONE DEACETYLASE6 (*HDA6*), which acts in concert with *PKL* to reduce histone H3 lysine 9 acetylation (H3K9ac), thereby contributing to chromatin compaction and gene repression [36]. *HDA6* also participates in repressing embryogenic identity genes during the embryogenesis-vegetative phase transition [37]. In addition, *HDA6* is implicated in embryogenic callus induction by epigenetically regulating auxin biosynthesis genes such as *YUCCA1* (*YUC1*), *YUC7*, and *YUC9*. This occurs via the formation of a protein complex with ENHANCER OF SHOOT REGENERATION 2 (*ESR2*), indicating a feedback mechanism linking chromatin regulation to hormone signaling [38]. Our interaction analysis predicted a functional association between *PKL* and *HDA6*, both of which were regulated following 5-AzaC treatment. The observed modulation of their phosphorylated states suggests a potential mechanism by which these epigenetic regulators coordinate the acquisition of embryogenic competence in sugarcane callus.

Phosphoproteomic analysis identified several proteins associated with epigenetic regulation that were modulated in sugarcane callus treated with 5-AzaC. One such protein was a chromatin complex subunit A (Sspon.07G0011500-1P, orthologous to *AtCHR12*), which was phosphorylated at S321. *CHR12* encodes a chromatin remodeling ATPase that integrates plant growth and development in response to environmental stimuli [39]. Loss-of-function mutations in *CHR12* in *Arabidopsis* result in defects in the maintenance of totipotent cell lines and impair embryo development, suggesting a conserved role in the regulation of embryogenesis [40]. Another epigenetic regulator affected by 5-AzaC treatment was a lysine-specific histone demethylase 1 homolog 2-like protein (Sspon.06G0011620-4D, orthologous to *AtLDL2*), which showed reduced accumulation. In *Arabidopsis*, *LDL2* is implicated in germination control in response to abscisic acid (ABA) and cooperates with HISTONE DEACETYLASE6 (*HDA6*) to mediate histone modifications involved in various biological processes, including circadian regulation [41, 42]. In our analysis, *LDL2* and *HDA6* were both regulated in response to 5-AzaC, and predicted to interact with other epigenetic regulators

such as PKL and CHR12. This coordinated regulation supports a central role for chromatin remodeling in establishing embryogenic competence.

Our findings also revealed modulation of DNA methylation-associated proteins. DNA methylation (specifically 5-methylcytosine modification) is known to decrease during embryogenic callus induction, a process essential for the reprogramming of gene expression and cellular differentiation in plant development [43]. DNA methyltransferases such as *CMT3* are linked to morphogenesis and organ development, while methylation marks serve as recruitment platforms for effector proteins like *SUVH1*, which can activate gene expression even in methylated promoters [44, 45]. Proteomic data showed that 5-AzaC treatment led to increased abundance of a sugarcane ortholog of *AtSUVH1*, the Histone-lysine N-methyltransferase H3K9-specific SUVH3 (Sspon.03G0030150-2D), along with regulation of another methyltransferase, SUVH4 (Sspon.06G0020970-2C, orthologous to *AtSUVH4*), phosphorylated at S129 and T133. Both SUVH and SUVH proteins contribute to chromatin condensation via histone methyltransferase activity and promotion of DNA methylation [46]. Additionally, the BRAHMA1 protein (Sspon.04G0020470-3D, orthologous to *AtBRM*) was uniquely regulated at S1370 following 5-AzaC treatment. *BRM* is involved in auxin signaling and the regulation of PIN transporter gene expression, playing a crucial role in the establishment of directional auxin flow. It also contributes to the expression of *PLETHORA1* (*PLT1*) and *PLT2*, members of a gene cluster that includes *BABY BOOM* (*BBM*), a transcription factor that can induce ectopic somatic embryo formation and is widely used in in vitro regeneration protocols [47, 48].

Together, these findings suggest that treatment with 5-AzaC triggers a wide-ranging reprogramming of chromatin-related factors in sugarcane callus. Our analysis uncovered an integrated network of differentially phosphorylated and accumulated proteins (DRPs and DAPs) involved in epigenetic regulation, including *CMT3*, *SUVH1*, *SUVH4*, *AGO1*, *LDL2*, *HDA6*, and *PKL*. These proteins represent key regulators of chromatin state and gene expression, and offer promising targets for modulating embryogenic competence in plant tissue cultures.

Protein post-translational regulation networks are involved in the expression of embryogenic competence

The identity of embryogenic callus cells is often associated with traits characteristic of the root apical meristem (RAM) [7]. In this study, the phosphoprotein transcription factor GTE4 (Sspon.03G0023940-4D, ortholog of *AtGTE4*) was found to be upregulated in S7 following

treatment with 5-Azacytidine (5-AzaC). GTE4 is known to induce and regulate cell division in meristems and is essential for controlling cell number during organ development [49]. Its expression has also been linked to the acquisition of embryogenic identity in cells of the root quiescent center [50]. GTE4 exerts its function by recognizing acetylated histone H3 (specifically H3K4 trimethylation), thereby facilitating the transcriptional activation of target genes [51]. Through this epigenetic mechanism, GTE4 promotes the mitotic activity of meristematic cells by regulating the cell cycle [50]. Additionally, our phosphoproteomic data revealed that the receptor-like kinase ZAR1 (Sspon.03G0013960-2B, ortholog of *AtZAR1*) was also upregulated following 5-AzaC treatment. ZAR1 is a key regulator of early embryogenesis, particularly by mediating the asymmetric division of the zygote, an essential process for establishing the embryo's apical-basal polarity [52]. Localized at the plasma membrane, this receptor is critical for embryo viability, as homozygous *zar1* mutants exhibit embryonic lethality at early developmental stages [53]. Perturbations in ZAR1 function lead to abnormal embryonic patterning, including symmetric zygotic divisions and reversed polarity, highlighting its role in cell fate specification during plant embryogenesis [52].

Phosphoproteomic analysis identified phosphorylation of the CAK1AT protein (Sspon.08G0008450-2B, orthologous to *AtCDKF-1*) at serine residue S458. *CDKF-1* encodes a cyclin-dependent kinase implicated in the regulation of root cell differentiation and mitotic activity. Loss-of-function mutants exhibit premature differentiation of meristematic cells, highlighting its role in maintaining cell proliferation within the root apical meristem (RAM) [54]. Another protein modulated in response to 5-Azacytidine (5-AzaC) treatment was a putative protein phosphatase polyprotein (Sspon.01G0057760-1D, orthologous to *AtMAIL3*), which was phosphorylated at residues S1303 and S1304. *MAIL3* is one of three Arabidopsis orthologs of the *MAINTENANCE OF MERISTEMS (MAIN)* gene and functions in the regulation of meristem cell division and the coordination of cell differentiation programs [55]. Together, these findings suggest that the phosphorylation state of regulatory proteins such as ZAR1, CDKF-1, and MAIL3 plays a pivotal role in maintaining the identity and developmental potential of embryogenic callus tissue. Their established roles in root meristem activity support the hypothesis that 5-AzaC-induced modulation of DNA methylation may activate a gene regulatory network resembling that of the RAM, thereby facilitating the acquisition of embryogenic competence and promoting somatic embryogenesis.

Phosphorylation level of auxin transporters is associated with the embryogenic competence of callus

Phosphoproteomic analysis identified phosphorylation events in key auxin transport regulators, including *brachytic2* (Sspon.06G0004630-3C, orthologous to *AtABCB1*), phosphorylated at S312, and ABC transporter B family member 19 (Sspon.05G0022390-2D, orthologous to *AtABCB19*), regulated at residues S657 and S672. Additionally, auxin efflux carrier component 1a (Sspon.04G0003600-4D, orthologous to *AtPIN1*) exhibited increased phosphorylation at S242 and T250. The probable auxin efflux carrier component 1b (Sspon.05G0020310-4D, orthologous to *AtPIN4*) showed multiple phosphorylation sites, including serines 212, 213, 251, 254, 260, 263, and 394, as well as threonine 268. Auxin plays a central role in plant development by coordinating physiological and morphogenetic processes through its biosynthesis and polar transport. The establishment of auxin gradients is mediated by ATP-binding cassette (ABC) transporters and PIN-FORMED (PIN) efflux carriers, which direct auxin flow necessary for cell fate determination and organogenesis [56]. *ABCB19* acts as a pivotal component in auxin transport and functions as a negative regulator of light signaling, promoting hypocotyl elongation through interactions with *PIN1* [57]. *ABCB1* is similarly involved in developmental processes, and its loss-of-function mutants exhibit dwarf phenotypes and reduced colony growth, indicative of impaired auxin-related responses [58].

PIN1 is essential for establishing auxin maxima, a prerequisite for the activation of somatic embryogenesis genes such as *WUSCHEL*, which defines embryogenic cell domains [59]. During embryogenesis, auxin gradients guide tissue patterning, with *PIN1* and *PIN4* playing major roles in maintaining directional auxin flow [60]. The polar localization of *PIN1* at the plasma membrane, crucial for directional transport, is finely regulated by phosphorylation within its hydrophilic loop regions [61]. Moreover, *ABCB19* contributes to the spatial organization and functional stability of *PIN1* in membrane microdomains, pointing to a cooperative mechanism that reinforces auxin flux [62]. Notably, phosphorylation of *ABCB19* has been shown to reduce its transporter activity, which may result in localized auxin accumulation, an essential cue for cellular reprogramming and somatic embryo induction [63]. Collectively, these findings indicate that the regulation of auxin transporters at both the transcriptional and post-translational levels, particularly through phosphorylation of *ABCB19* and *PIN1*, plays a critical role in establishing embryogenic competence in *Carica papaya* callus. These modifications likely alter intracellular auxin dynamics, thereby promoting the initiation of somatic embryogenesis.

The application of 5-azac alters the metabolism of embryogenic callus and promotes the development of somatic embryos

Energy metabolism is a key determinant in callus proliferation and the maintenance of embryogenic competence. In *Carica papaya*, proton pumps accumulate as early as the initial stages of callus induction, highlighting their importance in cellular reprogramming [64]. In this context, our phosphoproteomic analysis revealed that the V-type proton ATPase subunit F isoform X1 (evm.model.supercontig_117.89, ortholog of AtVHA-F) was differentially phosphorylated at S124 in embryogenic versus non-embryogenic calli. VHA-F is predicted to interact with additional subunits of the proton pump complex, including VHA-H and VHA-E1, indicating a coordinated regulation of proton transporters during callus proliferation. Another regulated proton pump was the mitochondrial outer membrane protein porin 4 (evm.model.supercontig_95.58, ortholog of AtVDAC4), which was phosphorylated at S4. In Arabidopsis, VDAC4 participates in various developmental processes and contributes to the maintenance of membrane potential homeostasis [65]. The accumulation of mitochondrial proton pumps has previously been linked to embryogenic competence in other plant systems, further supporting their role in metabolic reprogramming during somatic embryogenesis [66].

Beyond membrane transport, enzymes involved in purine metabolism also appear to be modulated. Two AMP deaminase isoforms, AMP deaminase isoform X1 (evm.model.supercontig_11.84, phosphorylated at S51 and S55) and probable AMP deaminase isoform X1 (evm.model.supercontig_5.284, phosphorylated at S84), both orthologs of AtAMPD, were identified. AMPD, also known as EMBRYOGENIC FACTOR 1 (FAC1), plays a pivotal role in ATP generation during early embryogenesis. Loss-of-function AtAMPD mutants fail to develop beyond the zygote stage, underscoring its essential role in embryo viability [67]. Further supporting a role for metabolism in embryo maturation, phosphoproteomic data showed that a putative B3 DNA-binding domain family protein (Sspon.06G0031100-1C, orthologous to AtABI3) was phosphorylated at S26. ABI3 is a central component of the LAFL (LEC1, ABI3, FUS3, LEC2) transcriptional regulatory network, which governs seed development, maturation, and germination [68]. Although BBM can activate somatic embryo formation by modulating auxin and LAFL gene expression, ABI3 alone does not induce somatic embryogenesis [69]. Instead, it is primarily involved in ABA-mediated processes and the accumulation of storage reserves critical for embryo maturation [70]. Interestingly, ABI3 was up-regulated in response to 5-AzaC treatment, suggesting that its expression may be under epigenetic control.

The predicted interaction of ABI3 with PKL and HDA6 supports the hypothesis that epigenetic mechanisms, including histone remodeling, influence ABI3-mediated accumulation of reserves during somatic embryo maturation. Moreover, the regulation of SUPPRESSOR

OF ABI3-5 isoform X1 (Sspon.02G0019430-4D, ortholog of AtSUA) provides another layer of control. SUA is a splicing factor that modulates ABI3 mRNA processing, thereby fine-tuning ABA responses during seed maturation and potentially during somatic embryo development as well [71]. Also, our data show that the putative serine/threonine-protein kinase (Sspon.01G0014770-1A, ortholog of SRK2E) was regulated at S297 and have a predicted interaction with ABI3. SRK2E transduces ABA signals, triggering phosphorylation cascades that control seed development and dormancy [72]. Together, these results highlight a tightly regulated network of metabolic enzymes, transcriptional regulators, and epigenetic modulators that orchestrate the metabolic reprogramming necessary for the establishment and maintenance of embryogenic competence in sugarcane callus.

Transcription factors associated with embryo development

Phosphoproteomic analysis revealed that the homeobox-leucine zipper protein ROC5-like (Sspon.03G0043250-1C, orthologous to *AtHDG1*) was up-regulated at residues S72 and S104, whereas a homeodomain leucine zipper family IV protein (Sspon.01G0015190-2B, orthologous to *AtHDG5*) was down-regulated at S720. *HDG1* belongs to the class IV homeodomain-leucine zipper (HD-ZIP IV) gene family and is implicated in meristem-associated cell differentiation and embryogenic development [73]. Notably, *HDG1* has a reciprocal regulatory relationship with the transcription factor *BABY BOOM (BBM)*, and both share targets involved in the control of cell proliferation, differentiation, and meristematic identity [74]. Furthermore, protein interaction predictions suggest that HDG1 may interact with the auxin transporter *PIN4*, reinforcing its potential involvement in developmental reprogramming during somatic embryogenesis. In contrast, no published studies have yet clarified the role of *HDG5* in embryogenesis. Taken together, the investigation of *HDG1* and *HDG5*, and their interaction with *BBM*, may uncover novel mechanisms governing the regulation of embryogenic competence in callus tissue.

Our phosphoproteomic data showed that the transcriptional corepressor *LEUNIG (LUG)* (Sspon.03G0011690-2B, orthologous to *AtLUG*) was up-regulated at S274, and the regulatory protein Zinc finger CCCH domain-containing protein 19 (*NERD*; Sspon.03G0006600-1A, orthologous to *AtNERD*) was differentially phosphorylated at T613. Both *LUG* and *NERD* act as transcriptional repressors involved in regulating gene expression programs that direct embryogenic and floral development. *NERD* has recently been identified as a positive regulator of N⁶-methyladenosine (m⁶A) methylation in messenger RNA through

its interaction with and stabilization of mRNA methyltransferases, contributing to post-transcriptional repression [75]. Moreover, *NERD* is involved in RNA interference (RNAi)-mediated gene silencing through interactions with Argonaute proteins, functioning in both pre- and post-transcriptional gene regulation [76]. *LUG*-mediated repression relies on its interaction with histone deacetylases (HDACs) and is part of a gene family that includes *TOPLESS* and *WUSCHEL*-interacting proteins. These factors are known to participate in totipotency acquisition, embryo patterning, and the modulation of auxin responses [77]. Protein-protein interaction analysis further revealed that *LUG* interacts with Ramosa1 enhancer locus 2 (*REL2*; Sspon.06G0010720-3D, orthologous to *AtTPR1*), a *TOPLESS*-related protein that may contribute to *LUG*'s regulatory function in embryogenic callus. These interactions suggest that *LUG* may have broader roles beyond transcriptional repression, including the maintenance of meristematic identity [78].

Our results also revealed that Cullin-3B (Sspon.06G0010440-1A, orthologous to *AtCUL3A*) was accumulated in calli treated with 5-AzaC. *CUL3* proteins are essential regulators of embryo and endosperm development, with loss-of-function mutants exhibiting embryo patterning defects and lethality [79]. These proteins function as core scaffolds of E3 ubiquitin ligase complexes and are tightly regulated post-translationally, mediating the ubiquitination and degradation of target proteins. This activity promotes early morphogenesis and the transition from zygote to heart-stage [80]. Moreover, *CUL3* has been shown to regulate ethylene biosynthesis and is required to maintain cell division and root apical meristem integrity during root development [81]. Overall, the regulation of *HDG1*, *CUL3*, *LUG*, and *NERD* in response to 5-AzaC treatment suggests that epigenetic reprogramming modulates complex transcriptional and post-transcriptional networks that underlie the acquisition of embryogenic competence in *Carica papaya* callus.

Phosphoproteomic analysis revealed that the protein TONSOKU (Sspon.04G0035280-1D), an ortholog of *Arabidopsis thaliana* TSK (*AtTSK*), was differentially phosphorylated in the S628 genotype following treatment with 5-Azacytidine (5-AzaC). TSK is known to play a crucial role in meristem maintenance by providing positional information necessary for the expression of *WUSCHEL* (*WUS*) and for the specification of meristematic cell identity [82]. In addition to its role in meristem regulation, TSK has also been implicated in the control of cell cycle progression and mitosis. Functional studies indicate that loss of TSK function leads to genome instability and delays in cell cycle, suggesting its importance in safeguarding genomic integrity during cell division [83]. Another phosphoprotein potentially associated with *WUS* regulation is Protein OBERON 4 (Sspon.01G0036140-2C), which

exhibited phosphorylation exclusively in 5-AzaC-treated callus. In *Arabidopsis*, OBE4 functions in conjunction with OBE3 in a positive feedback regulatory loop with *WUS*, whereby the expression levels of these genes are co-regulated to maintain the homeostasis of the apical meristem [84]. Collectively, these findings indicate that *WUS* expression may be modulated by multiple regulators involved in meristem identity and cell cycle control. The differential phosphorylation of TSK and OBE4 under hypomethylating conditions suggests that 5-AzaC treatment promotes both translational and post-translational modulation of key *WUS* regulators, potentially contributing to enhanced meristematic competence and embryogenic potential.

Conclusions

This study provides comprehensive phosphoproteomic evidence that treatment with the DNA hypomethylating agent 5-Azacytidine (5-AzaC) enhances embryogenic competence in *Saccharum* spp. calli by modulating key regulatory networks at multiple levels, including chromatin remodeling, transcriptional control, and protein post-translational modification. The identification of numerous phosphorylated proteins implicated in epigenetic regulation, auxin transport, meristem identity, and metabolism supports the hypothesis that 5-AzaC triggers a broad reprogramming of cellular states conducive to somatic embryo formation. One of the central conclusions of this study is that protein post-translational regulation, particularly via phosphorylation, plays a pivotal role in the establishment of embryogenic competence. Key phosphoproteins such as the transcription factor GTE4, the receptor-like kinase ZAR1, and the cyclin-dependent kinase CDKF-1 (CAK1AT) were up-regulated following 5-AzaC treatment. These proteins are directly involved in cell cycle control, meristem maintenance, and asymmetric cell division, hallmarks of embryogenic identity. Notably, the identity of embryogenic callus cells was found to be closely aligned with features characteristic of the root apical meristem (RAM), indicating that embryogenic calli may activate a RAM-like transcriptional program to achieve totipotency.

Phosphorylation of auxin transporters such as ABCB1, ABCB19, and the PIN-FORMED proteins PIN1 and PIN4 suggests that post-translational modulation of auxin dynamics contributes critically to the acquisition of embryogenic identity. These proteins regulate the formation of auxin maxima, spatiotemporal hormone gradients essential for cell fate specification and somatic embryo induction. In addition to hormonal regulation, 5-AzaC-induced reprogramming of metabolic pathways, as reflected in the phosphorylation of VHA-F, VDAC4, AMPD isoforms, and ABI3, highlights the importance of energetic homeostasis and

storage reserve accumulation in supporting embryo development and maturation. ABI3 and its regulatory partners (e.g., SUA and SRK2E) integrate ABA signaling with chromatin state regulation, reinforcing the concept that epigenetic control and post-translational modifications operate in concert to fine-tune developmental gene networks.

Lastly, the differential regulation of transcription factors such as HDG1, HDG5, LUG, and NERD, many of which exhibit established or predicted interactions with auxin-related proteins or embryogenesis regulators, points to a highly coordinated transcriptional network that reinforces cell proliferation and meristematic identity in embryogenic calli. Altogether, this work underscores the interconnectedness of epigenetic modifications, phosphorylation-dependent signaling, hormonal transport and perception, and metabolic reprogramming in the orchestration of somatic embryogenesis. The activation of RAM-like identity, mediated by a phospho-regulated network of transcription factors, kinases, transporters, and metabolic enzymes, may represent a core regulatory strategy employed by plant cells to regain totipotency and initiate somatic embryo formation. These findings provide a valuable molecular framework for improving tissue culture and regeneration protocols in sugarcane and other recalcitrant crops.

References

1. You, Z., *et al.*, Developing a new ethylene glycol/H₂O pretreatment system to achieve efficient enzymatic hydrolysis of sugarcane bagasse cellulose and recover highly active lignin: Countercurrent extraction. *Separation and Purification Technology*, 2025. 354: p. 128564.
2. Jaiswal, D., *et al.*, Brazilian sugarcane ethanol as an expandable green alternative to crude oil use. *Nature Climate Change*, 2017. 7(11): p. 788-792.
3. Kumar, T., *et al.*, Genetic engineering for enhancing sugarcane tolerance to biotic and abiotic stresses. *Plants*, 2024. 13(13): p. 1739.
4. Arruda, P., Genetically modified sugarcane for bioenergy generation. *Current Opinion in Biotechnology*, 2012. 23(3): p. 315-322.
5. Cursi, D.E., *et al.*, History and current status of sugarcane breeding, germplasm development and molecular genetics in Brazil. *Sugar Tech*, 2021. 24: p. 112–133.
6. Fehér, A., Somatic embryogenesis, Stress-induced remodeling of plant cell fate. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 2015. 1849(4): p. 385-402.
7. Fehér, A., Callus, Dedifferentiation, Totipotency, Somatic Embryogenesis: What these terms mean in the era of molecular plant biology? *Frontiers in Plant Science*, 2019. 10(April): p. 536.

8. Wójcik, A.M., B. Wójcikowska, and M.D. Gaj, Current perspectives on the auxin-mediated genetic network that controls the induction of somatic embryogenesis in plants. 2020. 21(4): p. 1333.
9. Wójcikowska, B., A.M. Wójcik, and M.D. Gaj, Epigenetic regulation of auxin-induced somatic embryogenesis in plants. *International Journal of Molecular Sciences*, 2020. 21(7): p. 2307.
10. Osorio-Montalvo, P., L. Sáenz-Carbonell, and C. De-la-Peña, 5-azacytidine: A promoter of epigenetic changes in the quest to improve plant somatic embryogenesis. *International Journal of Molecular Sciences*, 2018. 19(10): p. 3182.
11. Morończyk, J., *et al.*, Insights into the histone acetylation-mediated regulation of the transcription factor genes that control the embryogenic transition in the somatic cells of *Arabidopsis*. *Cells*, 2022. 11(5): p. 863.
12. Mahdavi-Darvari, F., N.M. Noor, and I. Ismanizan, Epigenetic regulation and gene markers as signals of early somatic embryogenesis. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 2015. 120(2): p. 407-422.
13. Greer, E.L. and Y. Shi, Histone methylation: a dynamic mark in health, disease and inheritance. *Nat Rev Genet*, 2012. 13(5): p. 343-57.
14. Passamani, L.Z., *et al.*, Long-term culture with 2,4-dichlorophenoxyacetic acid affects embryogenic competence in sugarcane callus via changes in starch, polyamine and protein profiles. *Plant Cell, Tissue and Organ Culture*, 2019. 140(2): p. 415-429.
15. Almeida, F.A., *et al.*, Label-free quantitative phosphoproteomics reveals signaling dynamics involved in embryogenic competence acquisition in sugarcane. *Journal of Proteome Research*, 2020. 19(10): p. 4145-4157.
16. Reis, R.S., *et al.*, Pretreatment free of 2,4-dichlorophenoxyacetic acid improves the differentiation of sugarcane somatic embryos by affecting the hormonal balance and the accumulation of reserves. *Plant Cell, Tissue and Organ Culture*, 2021. 145(1): p. 101-115.
17. Xavier, L.R., *et al.*, Integrative proteomics and phosphoproteomics reveals phosphorylation networks involved in the maintenance and expression of embryogenic competence in sugarcane callus. *Journal of Plant Physiology*, 2022. 268: p. 153587.
18. Guo, H., *et al.*, Large-scale label-free phosphoproteomics: From technology to data interpretation. *Bioanalysis*, 2014. 6(18): p. 2403-2420.
19. Picotti, P., Phosphoproteomics takes it easy. *Nature Biotechnology*, 2015. 33(9): p. 929-930.
20. Low, T.Y., *et al.*, Widening the bottleneck of phosphoproteomics: Evolving strategies for phosphopeptide enrichment. *Mass Spectrometry Reviews*, 2021. 40(4): p. 309-333.
21. Tyanova, S., T. Temu, and J. Cox, The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols*, 2016. 11(12): p. 2301-2319.
22. Szklarczyk, D., *et al.*, The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*, 2022. 51(D1): p. D638-D646.

23. Zheng, Y., *et al.*, iTAK: A program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Molecular Plant*, 2016. 9(12): p. 1667-1670.
24. Perez-Riverol, Y., *et al.*, The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res*, 2019. 47(D1): p. D442-D450.
25. Fraga, H.P., *et al.*, 5-Azacytidine combined with 2,4-D improves somatic embryogenesis of *Acca sellowiana* (O. Berg) Burret by means of changes in global DNA methylation levels. *Plant Cell Rep*, 2012. 31(12): p. 2165-76.
26. Yamamoto, N., *et al.*, Formation of embryogenic cell clumps from carrot epidermal cells is suppressed by 5-azacytidine, a DNA methylation inhibitor. *J Plant Physiol*, 2005. 162(1): p. 47-54.
27. Leljak-Levanić, D., *et al.*, Changes in DNA methylation during somatic embryogenesis in *Cucurbita pepo* L. *Plant Cell Rep*, 2004. 23(3): p. 120-7.
28. Nic-Can, G.I., *et al.*, New insights into somatic embryogenesis: LEAFY COTYLEDON1, BABY BOOM1 and WUSCHEL-RELATED HOMEBOX4 are epigenetically regulated in *Coffea canephora*. *PLOS ONE*, 2013. 8(8): p. e72160.
29. Li, J., *et al.*, Multi-omics analyses reveal epigenomics basis for cotton somatic embryogenesis through successive regeneration acclimation process. *Plant Biotechnol J*, 2019. 17(2): p. 435-450.
30. Kidner, C.A. and R.A. Martienssen, The role of ARGONAUTE1 (AGO1) in meristem formation and identity. *Developmental Biology*, 2005. 280(2): p. 504-517.
31. Du, F., *et al.*, Dose-dependent AGO1-Mediated inhibition of the miRNA165/166 pathway modulates stem cell maintenance in *Arabidopsis* shoot apical meristem. *Plant Communications*, 2020. 1(1).
32. Martín-Merchán, A., *et al.*, *Arabidopsis* AGO1 N-terminal extension acts as an essential hub for PRMT5 interaction and post-translational modifications. *Nucleic Acids Research*, 2024. 52(14): p. 8466-8482.
33. Mallory, A. and H. Vaucheret, Form, function, and regulation of ARGONAUTE proteins. *The Plant Cell*, 2010. 22(12): p. 3879-3889.
34. Dean Rider Jr, S., *et al.*, Coordinate repression of regulators of embryonic identity by PICKLE during germination in *Arabidopsis*. *The Plant Journal*, 2003. 35(1): p. 33-43.
35. Rider, S.D., *et al.*, Metabolic profiling of the *Arabidopsis* pkl mutant reveals selective derepression of embryonic traits. *Planta*, 2004. 219(3): p. 489-499.
36. Li, W., *et al.*, PICKLE and HISTONE DEACETYLASE6 coordinately regulate genes and transposable elements in *Arabidopsis*. *Plant Physiology*, 2024. 196(2): p. 1080-1094.
37. Tanaka, M., A. Kikuchi, and H. Kamada, The *Arabidopsis* histone deacetylases HDA6 and HDA19 contribute to the repression of embryonic properties after germination. *Plant Physiology*, 2007. 146(1): p. 149-161.

38. Lee, K., *et al.*, ESR2–HDA6 complex negatively regulates auxin biosynthesis to delay callus initiation in Arabidopsis leaf explants during tissue culture. *Plant Communications*, 2024. 5(7): p. 100892.
39. Mlynárová, L., J.-P. Nap, and T. Bisseling, The SWI/SNF chromatin-remodeling gene AtCHR12 mediates temporary growth arrest in Arabidopsis thaliana upon perceiving environmental stress. *The Plant Journal*, 2007. 51(5): p. 874-885.
40. Sang, Y., *et al.*, Mutations in two non-canonical Arabidopsis SWI2/SNF2 chromatin remodeling ATPases cause embryogenesis and stem cell maintenance defects. *The Plant Journal*, 2012. 72(6): p. 1000-1014.
41. Zhao, M., *et al.*, Arabidopsis histone demethylases LDL1 and LDL2 control primary seed dormancy by regulating DELAY OF GERMINATION 1 and ABA signaling-related genes. *Frontiers in Plant Science*, 2015. Volume 6 - 2015.
42. Hung, F.-Y., *et al.*, The Arabidopsis LDL1/2-HDA6 histone modification complex is functionally associated with CCA1/LHY in regulation of circadian clock genes. *Nucleic Acids Research*, 2018. 46(20): p. 10669-10681.
43. Huang, X., *et al.*, Dynamic changes in DNA methylation play a regulatory role in gene expression during the formation of callus from immature barley embryos. *BMC Plant Biology*, 2025. 25(1): p. 515.
44. Huang, J., *et al.*, Developmental expression of Arabidopsis methyltransferase genes MET1, DRM2, and CMT3. *Molecular Biology*, 2014. 48(5): p. 681-687.
45. Li, S., *et al.*, SUVH1, a Su(var)3–9 family member, promotes the expression of genes targeted by DNA methylation. *Nucleic Acids Research*, 2015. 44(2): p. 608-620.
46. Thorstensen, T., *et al.*, The Arabidopsis SUVR4 protein is a nucleolar histone methyltransferase with preference for monomethylated H3K9. *Nucleic Acids Research*, 2006. 34(19): p. 5461-5470.
47. Yang, S., *et al.*, The Arabidopsis SWI2/SNF2 chromatin remodeling ATPase BRAHMA targets directly to PINs and is required for root stem cell niche maintenance. *The Plant Cell*, 2015. 27(6): p. 1670-1680.
48. Boutilier, K., *et al.*, Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *The Plant Cell*, 2002. 14(8): p. 1737-1749.
49. Della Rovere, F., *et al.*, The Arabidopsis BET bromodomain factor GTE4 regulates the mitotic cell cycle. *Plant Signaling & Behavior*, 2010. 5(6): p. 677-680.
50. Airoidi, C.A., *et al.*, The Arabidopsis BET Bromodomain Factor GTE4 is involved in maintenance of the mitotic cell cycle during plant development. *Plant Physiology*, 2009. 152(3): p. 1320-1334.
51. Qian, F., *et al.*, The GTE4–EML chromatin reader complex concurrently recognizes histone acetylation and H3K4 trimethylation in Arabidopsis. *The Plant Cell*, 2024. 37(1).
52. Yu, T.-Y., *et al.*, The Arabidopsis receptor kinase ZAR1 is required for zygote asymmetric division and its daughter cell fate. *PLOS Genetics*, 2016. 12(3): p. e1005933.

53. Huang, Y., *et al.*, The dynamic proteome in *Arabidopsis thaliana* early embryogenesis. *Development*, 2022. 149(18).
54. Umeda, M., C. Umeda-Hara, and H. Uchimiya, A cyclin-dependent kinase-activating kinase regulates differentiation of root initial cells in *Arabidopsis*. *Proceedings of the National Academy of Sciences*, 2000. 97(24): p. 13396-13400.
55. Ühlken, C., *et al.*, MAIN-LIKE1 is a crucial factor for correct cell division and differentiation in *Arabidopsis thaliana*. *The Plant Journal*, 2014. 78(1): p. 107-120.
56. Cecchetti, V., *et al.*, ABCB1 and ABCB19 auxin transporters have synergistic effects on early and late *Arabidopsis* anther development. *Journal of Integrative Plant Biology*, 2015. 57(12): p. 1089-1098.
57. Noh, B., A.S. Murphy, and E.P. Spalding, Multidrug resistance-like genes of *Arabidopsis* required for auxin transport and auxin-mediated development. *Plant Cell*, 2001. 13(11): p. 2441-54.
58. Ye, L., *et al.*, Characterization of a dwarf mutant allele of *Arabidopsis* MDR-like ABC transporter AtPGP1 gene. *Biochemical and Biophysical Research Communications*, 2013. 441(4): p. 782-786.
59. Su, Y.H., *et al.*, Auxin-induced WUS expression is essential for embryonic stem cell renewal during somatic embryogenesis in *Arabidopsis*. *The Plant Journal*, 2009. 59(3): p. 448-460.
60. Weijers, D., *et al.*, Maintenance of embryonic auxin distribution for apical-basal patterning by PIN-FORMED-dependent auxin transport in *Arabidopsis*. *The Plant Cell*, 2005. 17(9): p. 2517-2526.
61. Huang, F., *et al.*, Phosphorylation of conserved PIN motifs directs *Arabidopsis* PIN1 polarity and auxin transport. *Plant Cell*, 2010. 22(4): p. 1129-42.
62. Titapiwatanakun, B., *et al.*, ABCB19/PGP19 stabilises PIN1 in membrane microdomains in *Arabidopsis*. *The Plant Journal*, 2009. 57(1): p. 27-44.
63. Christie, J.M., *et al.*, phot1 inhibition of ABCB19 primes lateral auxin fluxes in the shoot apex required for phototropism. *PLOS Biology*, 2011. 9(6): p. e1001076.
64. Xavier, L.R., *et al.*, Time-dependent proteomic signatures associated with embryogenic callus induction in *Carica papaya* L. *Plants*, 2023. 12(22): p. 3891.
65. Tateda, C., *et al.*, Molecular and genetic characterization of the gene family encoding the voltage-dependent anion channel in *Arabidopsis*. *Journal of Experimental Botany*, 2011. 62(14): p. 4773-4785.
66. Vale, E.M., *et al.*, Mitochondrial proteomics reveals new insights into embryogenic competence acquisition in *Carica papaya* L. callus. *Journal of Proteomics*, 2022. 252: p. 104434.
67. Xu, J., *et al.*, EMBRYONIC FACTOR 1 encodes an AMP deaminase and is essential for the zygote to embryo transition in *Arabidopsis*. *The Plant Journal*, 2005. 42(5): p. 743-758.

68. Tian, R., *et al.*, Direct and indirect targets of the Arabidopsis seed transcription factor ABSCISIC ACID INSENSITIVE3. *The Plant Journal*, 2020. 103(5): p. 1679-1694.
69. Horstman, A., *et al.*, The BABY BOOM transcription factor activates the LEC1-ABI3-FUS3-LEC2 network to induce somatic embryogenesis. *Plant Physiol*, 2017. 175(2): p. 848-857.
70. Salaün, C., L. Lepiniec, and B. Dubreucq, Genetic and molecular control of somatic embryogenesis. *Plants*, 2021. 10(7): p. 1467.
71. Sugliani, M., *et al.*, The conserved splicing factor SUA controls alternative splicing of the developmental regulator ABI3 in Arabidopsis. *The Plant Cell*, 2010. 22(6): p. 1936-1946.
72. Nakashima, K., *et al.*, Three Arabidopsis SnRK2 protein kinases, SRK2D/SnRK2.2, SRK2E/SnRK2.6/OST1 and SRK2I/SnRK2.3, involved in ABA signaling are essential for the control of seed development and dormancy. *Plant and Cell Physiology*, 2009. 50(7): p. 1345-1363.
73. Nakamura, M., *et al.*, Characterization of the Class IV homeodomain-leucine zipper gene family in Arabidopsis. *Plant Physiology*, 2006. 141(4): p. 1363-1375.
74. Horstman, A., *et al.*, AIL and HDG proteins act antagonistically to control cell proliferation. *Development*, 2015. 142(3): p. 454-464.
75. Shao, Y., *et al.*, NERD-dependent m6A modification of the nascent FLC transcript regulates flowering time in Arabidopsis. *Nature Plants*, 2025. 11(3): p. 468-482.
76. Pontier, D., *et al.*, NERD, a plant-specific gw protein, defines an additional rna-dependent chromatin-based pathway in Arabidopsis. *Molecular Cell*, 2012. 48(1): p. 121-132.
77. Sitaraman, J., M. Bui, and Z. Liu, LEUNIG_HOMOLOG and LEUNIG perform partially redundant functions during Arabidopsis embryo and floral development. *Plant Physiology*, 2008. 147(2): p. 672-681.
78. Stähle, M.I., *et al.*, YABBYs and the transcriptional corepressors LEUNIG and LEUNIG_HOMOLOG maintain leaf polarity and meristem activity in Arabidopsis. *The Plant Cell*, 2009. 21(10): p. 3105-3118.
79. Thomann, A., *et al.*, Arabidopsis CUL3A and CUL3B genes are essential for normal embryogenesis. *The Plant Journal*, 2005. 43(3): p. 437-448.
80. Figueroa, P., *et al.*, Arabidopsis has two redundant cullin3 proteins that are essential for embryo development and that interact with RBX1 and BTB PROTEINS TO FORM MULTISUBUNIT E3 ubiquitin ligase complexes in vivo. *The Plant Cell*, 2005. 17(4): p. 1180-1195.
81. Thomann, A., *et al.*, Arabidopsis CULLIN3 genes regulate primary root growth and patterning by ethylene-dependent and -independent mechanisms. *PLOS Genetics*, 2009. 5(1): p. e1000328.
82. Suzuki, T., *et al.*, A novel Arabidopsis gene TONSOKU is required for proper cell arrangement in root and shoot apical meristems. *The Plant Journal*, 2004. 38(4): p. 673-684.

83. Suzuki, T., *et al.*, TONSOKU is expressed in S phase of the cell cycle and its defect delays cell cycle progression in Arabidopsis. *Plant and Cell Physiology*, 2005. 46(5): p. 736-742.
84. Lin, T.-F., *et al.*, OBE3 and WUS interaction in shoot meristem stem cell regulation. *PLOS ONE*, 2016. 11(5): p. e0155657.

Acknowledgements

This research was supported by funds from the Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) (Proc E-26/211.310/2021; Proc E-26/204.185/2024) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (312258/2023-2), awarded to V.S. This study was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil (CAPES), Finance Code 001. The authors would like to thank FAPERJ, CNPq and CAPES for providing scholarships to graduate students.

Author contributions

LRX, VS and JJT contributed to the development of the methodology, the acquisition of results, and the writing, review, and editing of the manuscript. VPB, LGES, and CS-C reviewed, edited, and translated the manuscript.

Corresponding authors

Correspondence to Vanildo Silveira

Ethics declarations

The authors declare that they have no competing interests.

Ethics approval and consent to participate

This article does not contain any studies with human or animal participants. Work with plants complies with relevant institutional, national and international guidelines and legislation.

Consent for publication

All the authors provide their consent for this publication.

**Capítulo 7. Disruption of auxin transport impairs somatic embryo development
and proteome dynamics in sugarcane**

Disruption of auxin transport impairs somatic embryo development and proteome dynamics in sugarcane

Lucas Rodrigues Xavier^{1,2}, Daniel Dastan Rezabala Pacheco^{1,2}, Laís dos Santos Conceição^{1,2}, Caio Cezar Guedes Corrêa^{1,2}, Roberta Pena da Paschoa^{1,2}, Lucas Gomes do Espírito Santo^{1,2}, Felipe Astolpho Almeida^{1,2}, Vitor Batista Pinto^{1,2}, Tadeu Reis³, Claudete Santa-Catarina³ and Vanildo Silveira^{1,2*}

¹Laboratório de Biotecnologia, Centro de Biociências e Biotecnologia (CBB), Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), Campos dos Goytacazes, RJ, 28013-602, Brazil

²Unidade de Biologia Integrativa, Setor de Genômica e Proteômica, UENF, Campos dos Goytacazes, RJ, 28013-602, Brazil

³Laboratório de Biologia Celular e Tecidual, CBB-UENF, Campos dos Goytacazes, RJ, Brazil

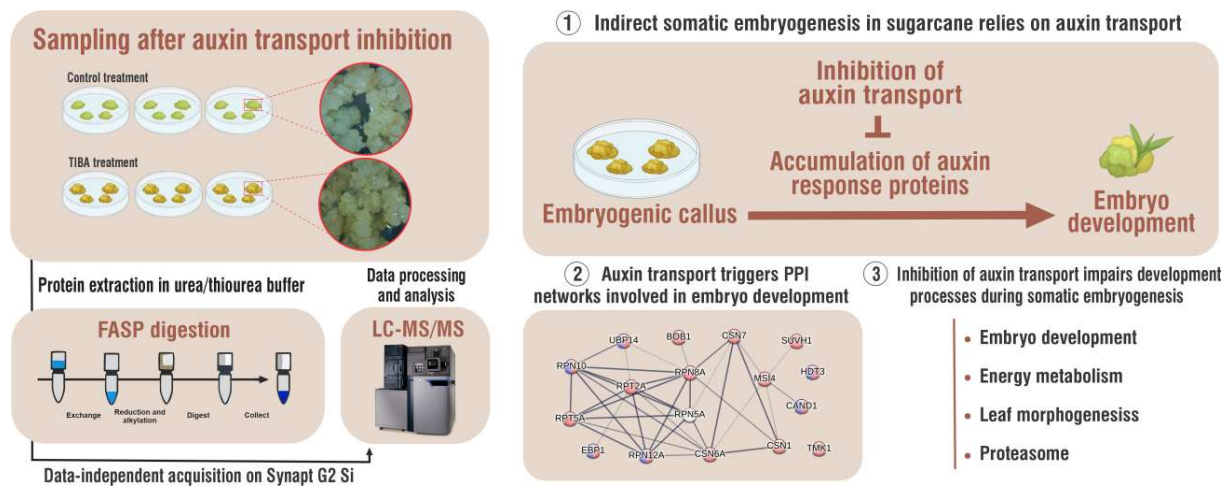
*Corresponding author: Dr. Vanildo Silveira E-mail: vanildo@uenf.br

Abstract

Auxin biosynthesis and transport are critical processes for in vitro tissue culture development, but the regulatory responses associated with these processes are still poorly understood in plants. We investigated the role of auxin transport during somatic embryo differentiation in sugarcane callus using an auxin transport inhibitor through a high-throughput proteomic approach. Four concentrations of 2,3,5-triiodobenzoic acid (TIBA 0, 1, 10, and 100 μ M) were tested on sugarcane embryogenic callus during the pre-maturation phase. Treatment with 100 μ M TIBA inhibited the differentiation of embryogenic callus into somatic embryos. Then, embryogenic callus from control and 100 μ M TIBA treatments were compared via proteomic analysis. A total of 1221 proteins were found, of which 740 were differentially accumulated (DAP) in the 100 μ M TIBA/Control comparison. Among the DAPs, 33 were unique to the control treatment, 31 were unique to the 100 μ M TIBA pre-maturation treatment, and 415 down-regulated and 261 up-regulated. Inhibition of auxin transport impaired the accumulation and predicted protein interactions involved in embryonic development, ubiquitination, and energy metabolism. The levels of spermidine synthase and spermidine were reduced, suggesting an association between auxin and polyamines during development. The down-regulation of the proteins chaperone BOBBER1, Cullin-associated NEDD8-dissociated protein 1, and Receptor-like kinase TMK2 could be associated with auxin responses and the predicted interactions between proteasome members found during the differentiation of somatic embryos. These new findings demonstrate the role of auxin transport in downstream responses during somatic embryo differentiation and open new avenues for exploring the potential of sugarcane for biotechnological purposes.

Keywords: Auxin transport; Embryo development; Plant morphogenesis; Comparative proteomics.

Graphical abstract



Introduction

Sugarcane is a source of renewable energy due to its potential for biomass and ethanol production, but the development of new varieties is limited by the complexity of the genome and the infertility of the seeds (Cursi *et al.* 2021). Transformation and gene editing are strategies for accelerating this process but require extensive use of plant tissue culture protocols. In this context, somatic embryogenesis is a plant tissue culture approach in which the differentiation of somatic embryos occurs either directly from cultured cells or indirectly via an intermediate callus phase (Fehér 2019). SE has become indispensable for the genetic improvement and genetic transformation of sugarcane genotypes (Arruda 2012). Sugarcane somatic embryogenesis is already well established in sugarcane and can also be used for large-scale clonal propagation of elite genotypes (Heringer *et al.* 2017; Passamani *et al.* 2018; Passamani *et al.* 2019; Reis *et al.* 2021). Although somatic embryogenesis is widely studied in biotechnology, much of the information is based on trial and error, leaving us with limited understanding of the underlying mechanisms. It is still unclear how the totipotency of plant cells relates to embryogenic competence and the differentiation of somatic sugarcane embryos.

There are different ways to induce somatic embryogenesis, but exposing explant cells to auxin treatments is required in the majority of plant species (Fehér 2015). In sugarcane, the protocols widely used include incubation of explants in culture medium supplemented with 2,4-dichlorophenoxyacetic acid (2,4-D) to induce embryogenic competence and callus multiplication, followed by the differentiation of somatic embryos callus in 2,4-D-free culture medium (Almeida *et al.* 2020; Xavier *et al.* 2022). The removal of 2,4-D from the culture

medium results in increased biosynthesis of endogenous auxin (mainly indole-3-acetic acid) and the restoration of auxin concentration gradients in differentiating embryos (Fehér 2019). The role and biosynthesis of auxin is one piece of the puzzle and involves several other factors, such as polar transport, perception and response to stress, signaling, and DNA methylation, among others (Su *et al.* 2009; Wójcikowska *et al.* 2020). In this sense, studies on the metabolism of auxin during somatic embryogenesis can help to fit one more piece together and provide a broader view of why somatic embryogenesis occurs.

Auxin transport is carried out cell-to-cell by the action of PIN-FORMED (PIN) and P-GLYCOPROTEIN (PGP) influx transporters and the AUXIN1/LIKE-AUX1 (AUX/LAX) efflux transporter (De Rybel *et al.* 2009). The family of PIN proteins in Arabidopsis contains 8 members and can be divided into PIN1-like (PIN1, PIN2, PIN3, PIN4 and PIN7) and PIN5-like (PIN5, PIN6 and PIN8) proteins, which are located in the plasma membrane and in the endoplasmic reticulum, respectively (Mravec *et al.* 2009). The mechanism of action of PIN1-like proteins is known, as its phosphorylation by the protein kinase PINOID is necessary for its asymmetric distribution in the plasma membrane, which is central to the unidirectional transport and generation of auxin gradients during development (Huang *et al.* 2010). Transcriptional responses to auxin depend on ubiquitination of the auxin/indole-3-acetic acid (Aux/IAA) repressors and release of the TOPLESS (TPL) corepressor from target gene promoters, followed by expression mediated by auxin response factors (ARF) (Wójcikowska *et al.* 2020). In the specific case of Aux/IAA proteins, ubiquitination is signaled by the binding of indole-3-acetic acid (IAA) to the SKP/CULLIN/F-BOX (SCFTIR1) complex and results in degradation by the proteasome (De Rybel *et al.* 2009). ARFs then recruit SWITCH/SUCROSE NONFERMENTING (SWI/SNF) to alter the chromatin state and allow the expression of target genes (Strader and Zhao 2016). In this context, the biosynthesis and transport of auxins are critical factors for the expression of meristematic identity genes and shoot-root axis establishment, which results in the development of bipolar somatic embryos (Su *et al.* 2015).

In Arabidopsis, incubation of embryogenic callus in 2,4-D-free culture medium leads to auxin accumulation at the callus periphery, and inhibition of auxin transport suppresses the differentiation of somatic embryos (Su *et al.* 2009). In this context, the use of transport inhibitors is useful for studying hormonal responses during the embryo development process (Elhiti and Stasolla 2011). The effects of auxin transport inhibitors, such as 2,3,5-triiodobenzoic acid (TIBA) and 2-(1-pyrenoyl) benzoic acid (PBA), are related to changes in the dynamics of actin organization and consequently the mobility of organelles and vesicles (Dhonukshe *et al.* 2008). Thus, the use of auxin transport inhibitors has revealed that the

efflux/influx of this hormone is important for maintaining the embryogenic identity of cells and for the correct development of somatic embryos (Karami *et al.* 2023).

On the other hand, even though they play a critical role in the differentiation and development of somatic embryos, responses related to auxin transport regulating the proteomic profile of sugarcane during somatic embryo differentiation have not yet been studied. We hypothesized that auxin transport in the early stages of sugarcane somatic embryo development is crucial for the accumulation of proteins related to morphogenesis and hormonal responses, and knowledge of the molecular processes related to the regulation of somatic embryogenesis and embryonic development is essential. For this purpose, we performed a high-throughput proteomic analysis of embryogenic callus after TIBA pre-maturation treatment to understand the role of auxin transport in the molecular mechanisms that drive the formation of sugarcane somatic embryos. These data provide new insights into proteins regulated by auxin and the functions of ubiquitination, the proteasome machinery, energy metabolism and other important processes related to embryo development.

Materials and Methods

Embryogenic callus induction and multiplication

Sixty-day-old sugarcane plants of the SP80-7280 cultivar were used as an explant source for embryogenic callus induction. The mature leaves were removed, and the resulting immature leaf rolls were surface-disinfected in 70% ethanol (Sigma-Aldrich, Saint Louis, USA) for 1 min and then in 0.75% sodium hypochlorite Qboa (Anhembi SA, Osasco, Brazil) for 30 min, followed by three washes with autoclaved water in a laminar flow hood. Afterwards, the leaf rolls were cross-sectioned into 3 mm thick slices and grown in 150 × 25 mm glass test tubes containing 10 mL of MS (Murashige and Skoog 1962) culture medium (Phytotechnology Lab, Overland Park, USA). The culture medium was supplemented with 20 g L⁻¹ sucrose (Labsynth, Diadema, Brazil), 2 g L⁻¹ Phytagel® (Sigma-Aldrich), and 10 µM 2,4-D (Sigma-Aldrich), the pH was adjusted to 5.8, and the culture media were sterilized by autoclaving at 121 °C for 15 min.

The explants in the tubes were maintained in the dark at 25 °C for 45 days for callus induction. Afterwards, the induced callus were transferred to 90 × 15 mm Petri dishes containing 20 mL of the same culture medium, kept in the dark at 25 °C, and then subcultured three times every 21 days for multiplication. During the multiplication period, embryogenic and

non-embryogenic callus were separated according to their morphological characteristics (Silveira *et al.* 2013).

Pre-maturation treatment

For somatic embryo differentiation embryogenic callus, were placed into pre-maturation treatment with different concentrations of TIBA (Sigma-Aldrich). Each treatment consisted of 10 biological replicates, with each replicate comprising one Petri dish containing five embryogenic callus (200 mg of fresh matter - FM - per callus) and this experiment was performed twice in order to confirm the effect of TIBA on the measured number of somatic embryos. The Petri dishes contained 20 mL of MS 2,4-D-free culture medium supplemented with 20 g L⁻¹ sucrose and 2 g L⁻¹ Phytigel. The pH was adjusted to 5.8, and the culture medium was sterilized by autoclaving at 121 °C for 15 min. TIBA was filter-sterilized and added to the culture medium after autoclaving, at concentrations of 0 µM (control), 1 µM, 10 µM, and 100 µM. Petri dishes were maintained at 25 °C in the dark during 21 days. After, ten replicates were used for the maturation experiment, and three of the 13 replicates were used for proteomic analysis and polyamines determination. For each biological replicate, the five embryogenic calli were macerated under liquid nitrogen, and the powder was aliquoted into three aliquots of 300 mg FM for further analyses.

Maturation of somatic embryos

Embryogenic callus grown in pre-maturation with different TIBA treatments, including the control, were used for the maturation of somatic embryos. Ten biological replicates for each treatment were used, with each replicate consisting of four embryogenic calli (250 mg FM) per Petri dish. The Petri dishes containing 20 mL of MS 2,4-D-free culture medium, supplemented with 20 g L⁻¹ sucrose, 2 g L⁻¹ Phytigel (pH adjusted to 5.8), were sterilized by autoclaving at 121 °C for 15 min. The embryogenic callus were maintained in the dark for the first 7 days and subsequently transferred to a 16-h photoperiod established with GreenPower TLED 20-W W_mB (Koninklijke Philips Electronics NV, Amsterdam, Netherlands) at 55 µmol m⁻² s⁻¹.

After 42 days, considered the end of the maturation phase, the number of somatic embryos per callus and the number of abnormal roots per dish were measured to determine which contrasting treatments would be used in the proteomics and PAs analyses from samples collected at the end of pre-maturation. The treatments were compared by Tukey's test (*p value*

< 0.05) using the R language (Team 2023). After analyze the proteomics results samples five biological replicates were collected for IAA determination. For each biological replicate, the five calli were macerated under liquid nitrogen, and the powder was aliquoted into aliquots of 300 mg FM aliquots.

Scanning electron microscopy

Somatic embryo samples were collected from callus after 42 days of the maturation treatment and immediately fixed in 4.0% formaldehyde and 2.5% glutaraldehyde in a 50 mM phosphate buffer solution at pH 6.8 for 24h. Next, the samples were washed four times with the same buffer (one hour each). Later, samples were sequentially dehydrated in a solution with increasing levels of ethanol (30%, 50%, 70%, 90%, and 100%) and were retained twice for one hour in each solution. Dehydrated samples imbibed in pure ethanol were dried using a Critical Point Drier device (Bal-Tec E 3000 CPD). The samples were mounted in aluminum stubs and subsequently coated with platinum in a Sputter Coater Device (Bal-tec SCD 050). Specimens were viewed on a scanning electron microscope (SEM) (Zeiss EVO 040 SEM at 15 kV).

Bottom-up proteomics analysis

Protein extraction and digestion for proteomics analyses were performed on samples of 300 mg embryogenic callus FM collected at the end of pre-maturation from 0 μ M (control) and 100 μ M TIBA treatments. The powdered samples were homogenized with 1 mL of extraction buffer, consisting of 7 M urea (Cytiva, Hoegaarden, Belgium), 2 M thiourea (Cytiva), 2% Triton X-100 (Sigma–Aldrich), 1% dithiothreitol (Bio-Rad, Hercules CA, USA), and 1 mM phenylmethanesulfonyl fluoride (PMSF, Sigma-Aldrich) was added to the sample powder to extract the proteins (Reis *et al.* 2021). The protein concentrations of the biological replicates were estimated using a 2-D Quant Kit (Cytiva).

To prepare for protein digestion, 100 μ g aliquots of each sample were used. First, the protein samples were precipitated using the methanol/chloroform protocol to remove any detergent contaminants according to Nanjo *et al.* (2012). Next, tryptic digestion was performed using the filter-aided sample preparation (FASP) method, adapted from (Wiśniewski *et al.* 2009). Briefly, the proteins were double desalted with 50 mM ammonium bicarbonate (Sigma-Aldrich) at pH 8.5 on Microcon-30 kDa intact filters (Merck Millipore, Darmstadt, Germany). The proteins were then incubated at 60 °C for 20 min with 100 μ L of 50 mM DTT reduction solution prepared in 8 M urea and 50 mM ammonium bicarbonate at pH 8.5. They were then

cooled with 200 μL of 8 M urea and 50 mM ammonium bicarbonate. The proteins were then alkylated for 20 min at 25 °C in the dark with 100 μL of 50 mM iodoacetamide (Sigma-Aldrich) prepared in 50 mM ammonium bicarbonate. The samples were then washed twice with 8 M urea and 50 mM ammonium bicarbonate and three more times with 50 mM ammonium bicarbonate. Finally, the proteins were digested for 18 h at 37 °C in a mixture of 25 μL of 0.2% RapiGest solution (Waters, Manchester, UK) and 20 μL of trypsin solution (trypsin:sample ratio 1:100 μg , Promega, Madison, USA). To precipitate RapiGest and inhibit trypsin activity, 5 μL of 15% trifluoroacetic acid (TFA) was added to bring the final concentration to 0.5%, and the mixture was incubated for 30 min at 37 °C. Finally, the peptides were vacuum-dried and solubilized at 0.7-1 μg μL^{-1} in 95% water and 5% acetonitrile plus 0.1% formic acid, all mass spectrometry (MS)-grade (Sigma-Aldrich). The peptide concentration was estimated by measuring at the A205 nm using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, USA).

Mass spectrometry analysis for proteomic data acquisition

Proteomic analysis was performed using a nanoAcquity UPLC M-class (Waters) coupled to a Synapt G2-Si mass spectrometer (Waters) with an electrospray ionization (ESI)-MS/MS source, with the parameters described by Passamani *et al.* 2018. Briefly, the runs were performed by loading 2 μg of digested peptides onto a C18 trap column (100Å, 5 μm , 180 μm x 20 mm, 2D; Waters) at 5 μL min^{-1} for 3 min and then onto a HSS T3 analytical reversed-phase column (100Å, 1.8 μm , 75 μm x 150 mm; Waters) at 400 nL min^{-1} heated to 45 °C. For peptide elution, a binary gradient was used, with mobile phase A consisting of water (Tedia, Fairfield, USA) and 0.1% formic acid and mobile phase B consisting of acetonitrile and 0.1% formic acid. Gradient elution was started with 5% B, which was then increased from 5 to 40% B until reaching 92 min and from 40 to 99% B until 96 min, after which it was maintained at 99% B until 100 min. The elution was then decreased to 5% B until 102 min, after which it was maintained at 5% B until the end of the run at 118.00 min.

The HDMSE MS data was obtained in data-independent acquisition (DIA) mode using ion mobility with the following parameters: positive mode, resolution (V mode) at 35,000 FWHM, and an ion mobility wave velocity program starting at 800 m s^{-1} to 500 m s^{-1} . The transfer collision energy increased from 25 to 55 V in high-energy mode, and the cone and capillary voltages were 40 and 2800 V, respectively. The pressure of the nanoflow gas was 0.5 bar, the purge gas flow rate was 150 L h^{-1} , and the source temperature was 80 °C. For the time-of-flight (TOF) parameters, the scan time was set to 0.6 s in continuous mode, with a mass range of 50 to 2000 Da. Human [Glu1]-fibrinopeptide B (Waters) at 100 fmol μL^{-1} was

used as an external calibrant, and lock mass acquisition was performed every 30 s. Mass spectra were acquired by MassLynx v4.1 software (Waters).

MS data analysis

The proteomics spectra were searched against the protein databank generated from the sequencing of the haploid *Saccharum spontaneum* AP85-441 (databank version: v20190103) (Zhang *et al.* 2018). The proteomic MS data have been deposited in the ProteomeXchange Consortium via the PRIDE (Perez-Riverol *et al.* 2019) partner repository with the dataset identifier PXD057907.

Proteomic MS/MS spectra processing and database searches were performed using ProteinLynx Global SERVER (PLGS) software v.3.03 (Waters), and label-free quantification analyses were performed using ISOQuant software v.1.6 (Distler *et al.* 2014) with previously described parameters (Passamani *et al.* 2018). Briefly, the PLGS was processed using a low-energy threshold of 150 (counts), an elevated energy threshold of 50, and an intensity threshold of 750. In addition, the analysis was performed using the following parameters: two missed cleavages, a minimum fragment ion per peptide equal to three, a minimum fragment ion per protein equal to seven, a minimum peptide per protein equal to two, fixed modifications of carbamidomethyl (C) and variable modifications of oxidation (M) and phosphoryl (STY). The FDR for peptide and protein identification was set to a maximum of 1%, with a minimum peptide length of six amino acids. The ISOQuant analysis included retention time alignment, exact mass retention time, IMS clustering, data normalization and protein homology filtering. ISOQuant annotates the resulting feature clusters by evaluating consensus peptide identifications and identification probabilities. The protein identification parameters in ISOQuant were set to an FDR of 1%, a peptide score greater than six, a minimum peptide length of six amino acids, and at least two peptides per protein. Label-free quantifications were estimated using the TOP3 quantification approach, followed by the multidimensional normalization process implemented within ISOQuant.

To ensure the quality of the results after data processing, only proteins present in all three runs or absent in all biological replicates (for unique proteins) were considered for differential accumulation analysis in the 100 μ M TIBA/control treatments comparison. Proteins with significant Student's t test (two-tailed; $P < 0.05$) results and Log₂ fold change (FC) greater than 0.585 (for up-regulated proteins) or less than -0.585 (for down-regulated proteins) were considered differentially accumulated proteins (DAPs).

Functional annotation analysis was performed using OmicsBox software (<https://www.biobam.com/omicsbox>) and UniProtKB (<https://www.uniprot.org>). The protein sequences were subjected to a BLAST search against the NCBI non-redundant green plant protein database (taxa: 33090, Viridiplantae). An analysis of the enrichment of biological processes, cellular components, molecular functions and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways of DAPs (p-value < 0.01) was performed in Metascape (Zhou *et al.* 2019) using Arabidopsis orthologues of sugarcane proteins. The DAPs were also used for prediction of interaction networks through a STRING (version 12.0) search (Szklarczyk *et al.* 2022). The predicted kinases, transcription factors and transcriptional regulators among the DAPs were identified using the iTAK tool (Zheng *et al.* 2016). A PCA was performed considering all protein TICs (missing values = 10) using prcomp::stats v.3.5.1 in the R language (Team 2023). The log2 FC and log10 p value values from the Student's t test of proteomics data were used as variables in a cartesian plane to construct a volcano plot using Python.

Polyamine quantification

The polyamine contents were determined from three biological replicates (300 mg FM) collected at the end of pre-maturation from 0 μ M (control) and 100 μ M TIBA treatments. The analysis of free polyamines (PAs) was carried out using high-performance liquid chromatography (HPLC; Shimadzu, Japan) as previously described (Silveira *et al.* 2004). The concentrations of putrescine (PUT; Sigma-Aldrich), spermidine (SPD; Sigma-Aldrich), and spermine (SPM; Sigma-Aldrich) were determined using a standard curve. A T-test ($P < 0.05$) was performed to determine the significant differences between treatments.

IAA quantification

IAA contents were determined using three biological samples (300 mg FM each) of embryogenic callus from both control and 100 μ M TIBA pre-maturation treatments, collected after 42 days of maturation stage in somatic embryogenesis. Then, 100 mg of sample N₂-powdered callus was homogenized with 1 mL of extraction buffer containing 80% ethanol and 1% polyvinylpyrrolidone (Sigma-Aldrich). The extracts were sonicated, followed by centrifugation at 16000 x g, both steps were performed at 4 °C for 10 min. The supernatant was collected and vacuum-dried at 45 °C for 160 min until it reached approximately 200 μ L in volume. At this point, the volume of the tubes was individually adjusted with water to 500 μ L (using a scale, considering 1:1 mass and volume). Subsequently, 1 mL of ether was added, vortexed, and frozen in liquid nitrogen for 10 s. During freezing, only the water was frozen, the

ether containing the metabolites was transferred to a new tube, and the addition of ether, stirring, freezing and collection was repeated. The ether was completely vacuum air-dried at 45 °C, and the pellet was solubilized in 300 µL of 90:10 water:methanol (MS grade; Tedia, Fairfield, USA) and filtered through a 0.2 µm PVDF filter.

The metabolites were separated by liquid chromatography by loading 10 µL of the solution recovered in the extraction onto a BEH C18 1.7 µm column (2.1 µm × 50 mm) heated to 40 °C, with a flow rate of 300 µL.min⁻¹. The binary gradient consisted of LC-MS grade water plus 0.01% formic acid, and LC-MS grade methanol plus 0.01% formic acid as the eluent. The elution gradient was started at 10 % methanol and maintained for two minutes, increasing linearly to 90 % for four minutes, then increasing linearly to 100 % for one minute, and reducing to 10 % for 50 seconds, maintained at 10 % for 70 seconds, totaling nine minutes of running.

Using a high-performance dual orthogonal Z-spray electrospray ionization (ESI) source (Waters), the effluent from the Acquity UPLC I-Class FTN (Waters) was introduced into a Xevo TQ-XS triple quadrupole mass spectrometer (Waters). A pressurized nitrogen gas nebulization of 7 bar and a desolvation flow of 1000 L h⁻¹ were used, and a cone flow rate of 150 L h⁻¹ was used. The desolvation gas temperature was set to 650 °C, the source temperature was 150 °C, and the capillary voltage was 3 kV. For the double quadrupole tandem mass spectrometer (MS/MS), argon gas was used as the collision gas with a pressure of 0.8 kgf cm² in the collision cell. The spectrometer was operated in multiple reaction monitoring (MRM) mode, with 25 V for cone voltage, 15 V for collision energy, positive mode for 176 m/z precursor and 130 m/z for fragment mass of IAA quantification.

MassLynx software (v.4.2, Waters) was used for MS data acquisition. The spectra were integrated, and the lines were smoothed with averages between the points. The area under the peak of the IAA signal was compared through a linear equation estimated from a standard curve (concentrations: 0, 0.5, 1, 2, 4, 6, 8, 10, 25, 50, 75, 100, 150, and 200 pg µL⁻¹) using TargetLynx XS software (Waters). A T-test (p value 0.05) was used to compare the average of 5 biological replicates of the control and 100 µM TIBA maturation treatments. Tukey's test was performed to assess whether the differences in means were significantly different with a p-value <0.05.

Results

TIBA pre-maturation treatment in embryogenic callus impairs protein accumulation and the differentiation of somatic embryos

To study the role of auxin transport in sugarcane somatic embryogenesis, embryogenic callus were transferred to a pre-maturation medium without 2,4-D and supplemented with different concentrations of TIBA (0 μ M - control, 1, 10, and 100 μ M). After 21 days of culture, morphological differences between the callus were observed, including browning of callus treated with 100 μ M TIBA compared to the control and those treated with 1 and 10 μ M TIBA (Fig. 2 A-D). The callus were subjected to maturation treatment for 42 days. Callus from the control, 1 and 10 μ M TIBA pre-maturation treatments presented the formation of somatic embryos, with no significant differences between treatments (Fig. 2 E-G, Fig. 3). On the other hand, embryogenic callus treated with 100 μ M TIBA did not differentiate into somatic embryos (Fig. 2H, Fig. 3). The morphology of mature somatic embryos that developed from the control, 1 and 10 TIBA μ M treatments was similar among them, with prominent shoots (Fig. 2I and K). In contrast, more abnormal roots (with no scutellum/cotyledon) developed in the callus of the 100 μ M TIBA pre-maturation treatment than in those of the other treatments and, and these roots presented evident darkening after differentiation (Fig. 2 H and J, Fig. 3). Details from scanning electron microscopy show the lack of organized structures on the developing roots in the culture media containing 100 μ M TIBA (Fig. 2 L-N).

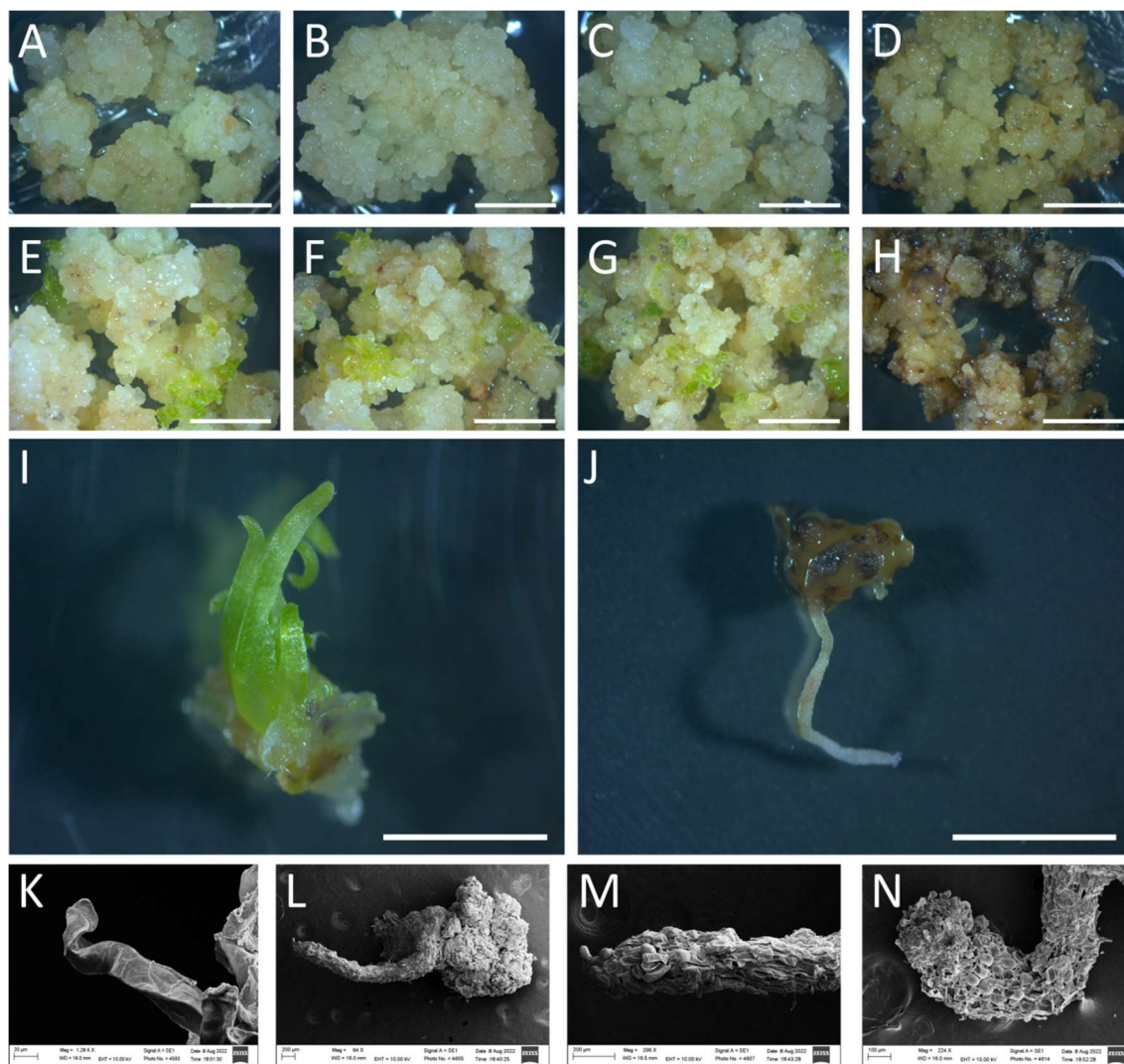


Fig. 2. Morphology of cultures during the maturation of somatic embryos from embryogenic callus subjected to pre-maturation treatment with different concentrations of TIBA. Embryogenic callus from the control (A), 1 μ M (B), 10 μ M (C) and 100 μ M (D) TIBA pre-maturation treatments after 21 days of culture in the dark. Embryogenic callus from the control (E), 1 μ M (F), 10 μ M (G) and 100 μ M (H) TIBA pre-maturation treatments after 42 days of maturation in differentiation culture medium with a 16-hour photoperiod. Mature somatic embryo developed from embryogenic callus from the control pre-maturation treatment. (I). Abnormal radicle developed from embryogenic callus from 100 μ M TIBA pre-maturation treatment (J). Pictures from scanning electron microscopy of somatic embryos with normal development (K) or with defects caused by inhibition of auxin transport (L-N). The white bars represent 0.5 cm in A-H and 0.1 cm in I and J.

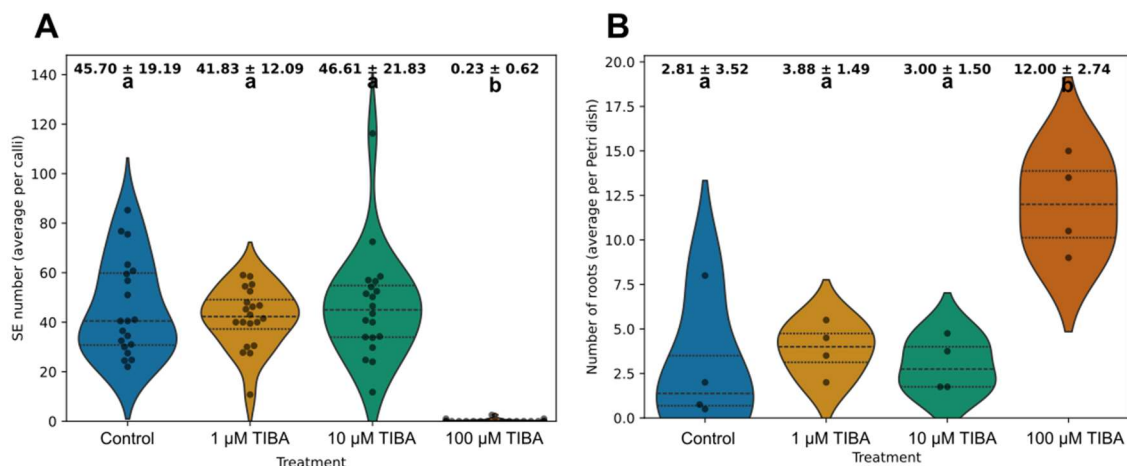


Fig. 3. Influence of auxin transport inhibition on sugarcane somatic embryogenesis. The boxplot shows the average number of somatic embryos per callus (A) and number of abnormal roots per Petri dish (B) after different TIBA treatments. Each point in the boxplot represents a biological replicate. Values above the plots indicate the median $\pm$ standard deviation for each treatment and the values in the same line followed by different letters are significantly different according to Tukey's test ($p < 0.05$).

Effects of TIBA pre-maturation treatment on differential protein accumulation

Since embryo differentiation was contrasting between callus from the control and 100 μ M TIBA pre-maturation treatments, a shotgun proteomic analysis was conducted to identify and quantify proteins related to auxin transport in sugarcane embryogenic callus. In the proteomic analysis of the control and 100 μ M TIBA samples, 1221 proteins were identified (Supplementary Table 1). Principal component analysis (PCA) separated the samples into two groups (Supplementary Fig. 1A). The proteins that most contributed to the variance observed in PC1 were elongation factor 1-alpha, basic 7S globulin, oryzain alpha chain and alpha-amylase type A isozyme. The complete list of the 15 proteins that most contributed to the observed variance in PC1 can be found in Supplementary Table 2.

The comparative proteomic analysis allowed the identification of 740 DAPs between the 100 μ M TIBA/control comparison (Supplementary Table 1, Supplementary Fig. 1B). Among these, 33 were unique to the Control, 31 were unique to the 100 μ M TIBA pre-maturation treatments, and 415 down-regulated and 261 up-regulated (Supplementary Table 1).

Proteomic functional analysis

Functional proteomic analysis of the DAPs from the 100 μ M TIBA/control comparison revealed enrichment of biological processes related to embryo development, regulation of translation, ubiquitin-dependent protein catabolic processes and leaf morphogenesis biological process in the control treatment (Fig. 4A). On the other hand, amino acid,

nucleotide-sugar and small molecule metabolic processes were enriched among the DAPs in the 100 μ M TIBA pre-maturation treatment (Fig. 4A). The analysis of metabolic pathways shows that proteasome, DNA replication and chromatin remodeling were enriched among down-regulated or unique proteins in the control, while the fatty acid degradation, biosynthesis and metabolism of amino acids pathways were enriched among up-regulated and unique proteins in 100 μ M TIBA pre-maturation treatment (Fig. 4B). The analysis of the molecular functions of the DAPs revealed that translation regulator, oxidoreductase, antioxidant and ion-binding activities were enriched among the down-regulated or unique proteins in the control treatment (Fig. 4C).

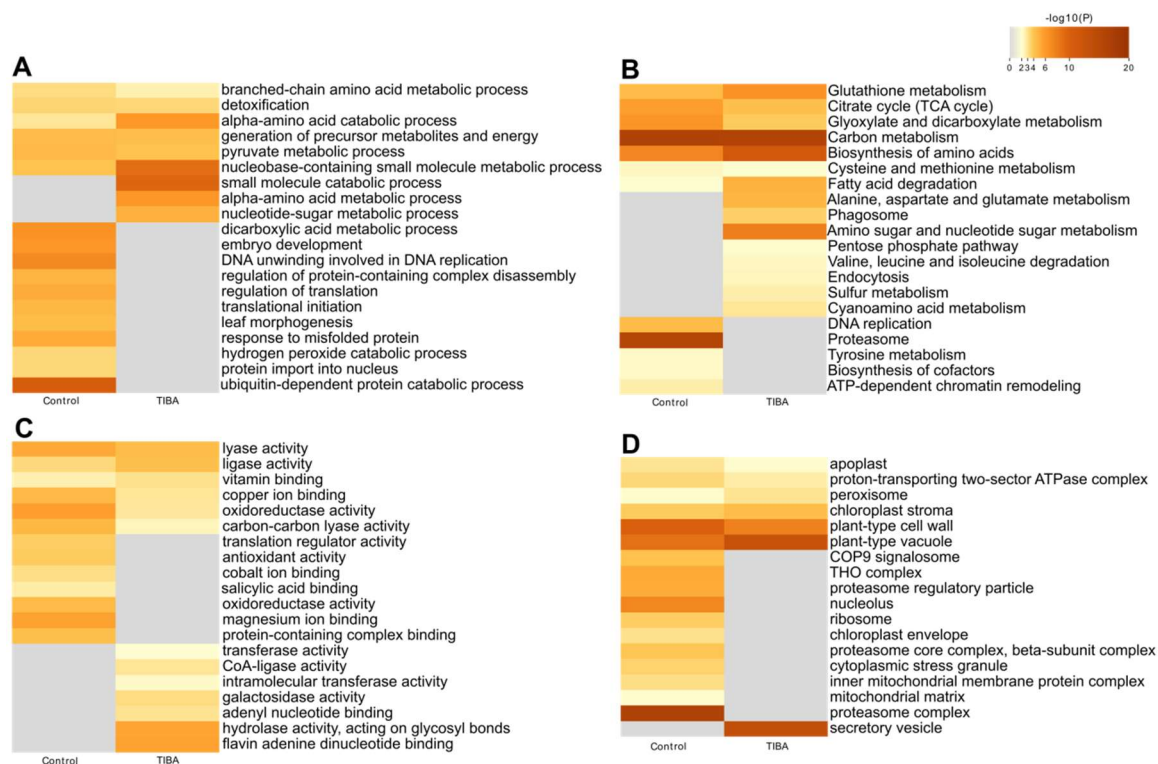


Fig. 4. Proteomics functional analysis of DAPs from the 100 μ M TIBA/control comparison. The top 20 enriched biological processes (A), KEGG pathways (B), molecular functions (C) and cellular components (D) between the DAPs. The analyses were performed in Metascape using a p value < 0.01, a minimum count of 3, and an enrichment factor > 1.5.

The molecular functions of transferase, COA-ligase, galactosidase, adenyl nucleotide binding and hydrolase were enriched among the up-regulated or unique DAPs in the 100 μ M TIBA pre-maturation treatment. Cellular components involved in energy generation and protein turnover, such as the ATPase complex, chloroplast envelope, mitochondrial inner membrane and matrix, COP9 signalosome and proteasome, were enriched among the DAPs accumulated in the control treatment (Fig. 4D). Proteasome protein analysis revealed that several regulatory and core particles were down-regulated in response to the 100 μ M TIBA

pre-maturation treatment (Supplementary Fig. 2). In addition, many TCA cycle reactions are carried out by proteins that were down-regulated after 100 μ M TIBA pre-maturation treatment (Supplementary Fig. 3). Among the DAPs that were up-accumulated and unique to the 100 μ M TIBA pre-maturation treatment, proteins from the vacuole, secretory vesicle and peroxisome were enriched (Fig. 4D).

The search for regulatory proteins led to the identification of 3 kinases, 3 transcription factors and 5 transcriptional regulators between DAPs (Supplementary Table 3). In this context, the receptor-like kinase group (RLK-Pelle), nuclear transcription factor/regulators nuclear factor Y subunit C (NF-YC), high mobility group (HMG), SET domain containing protein (SET), Plant Homeodomain (PHD) and switch/sucrose non-fermentable (SWI/NSF) families were down-regulated in the 100 μ M TIBA/control treatments comparison (Supplementary Table 3). The remaining two groups, CAMK_CDPK and MYB-related transcription factor were up-regulated in the same comparison (Supplementary Table 3).

After the statistical and functional analysis of the DAPs, it was possible to select those proteins down-regulated or unique to the control that were associated with the effects of auxin transport inhibition by 100 μ M TIBA pre-maturation treatment during sugarcane somatic embryogenesis (Table 1). These proteins that are down-regulated in response to 100 μ M TIBA include receptor kinases, epigenetic regulators, auxin signaling and protein degradation.

Table 1. List of proteins associated with auxin responses during differentiation in sugarcane. The use of the auxin efflux inhibitor TIBA compromises the differentiation of sugarcane somatic embryos and decreases the abundance of proteins involved in auxin signaling and distribution.

Protein ID	BLAST top hit	<i>Arabidopsis</i> ortholog	Accumulation 100 μ M TIBA/control
Sspon.08G0011080-2B	Protein BOBBER 1	BOB1	Down
Sspon.04G0017510-2B	Cullin-associated NEDD8-dissociated protein 1	CAND1	Down
Sspon.01G0000970-1A	COP9 signalosome complex subunit 1	CSN1	Down
Sspon.06G0001040-2B	COP9 signalosome complex subunit 6a	CSN6A	Down
Sspon.02G0007340-1A	COP9 signalosome complex subunit 7 isoform X3	CSN7	Down
Sspon.07G0010050-3C	ERBB-3 BINDING PROTEIN 1	EBP1	Down
Sspon.07G0003330-2B	Histone deacetylase HDT2	HDT3	Down
Sspon.01G0037080-1B	26S proteasome non-ATPase regulatory subunit 4	RPN10	Down
Sspon.06G0009170-2B	26S proteasome non-ATPase regulatory subunit 8 homolog A	RPN12A	Down
Sspon.01G0032970-2D	26S proteasome non-ATPase regulatory subunit 12 homolog A-like	RPN5A	Down
Sspon.05G0001800-3D	26S proteasome non-ATPase regulatory subunit 7 homolog A	RPN8A	Down
Sspon.02G0000500-2P	26S proteasome regulatory subunit 4 homolog	RPT2A	Unique control
Sspon.04G0001300-1A	26S proteasome regulatory subunit 6A homolog	RPT5A	Down
Sspon.07G0005310-1A	histone-lysine N-methyltransferase, H3 lysine-9 specific SUVH1-like	SUVH1	Down
Sspon.01G0027230-3D	Receptor-like kinase TMK2	TMK1	Down
Sspon.03G0022020-1P	ubiquitin carboxyl-terminal hydrolase 14	UBP14	Down
Sspon.03G0007870-1A	WD-40 repeat-containing protein MSI4	MSI4	Down

Predicted protein-protein interaction networks related to auxin signaling and somatic embryo development

A protein-protein interaction analysis was conducted to discern networks associated with the differentiation and patterning of somatic embryos, aiming to identify key associations in these processes (Table 1; Fig. 4). The results revealed predicted associations between proteins involved in cellular responses, embryo development, proteasome machinery and hormone responses. Proteasome-related proteins (RPN10, RPN12A, RPN5A, RPN8A, RPT2A and RPT5A) act as central hubs, interacting with proteins associated with auxin signaling, epigenetic regulation and morphogenesis.

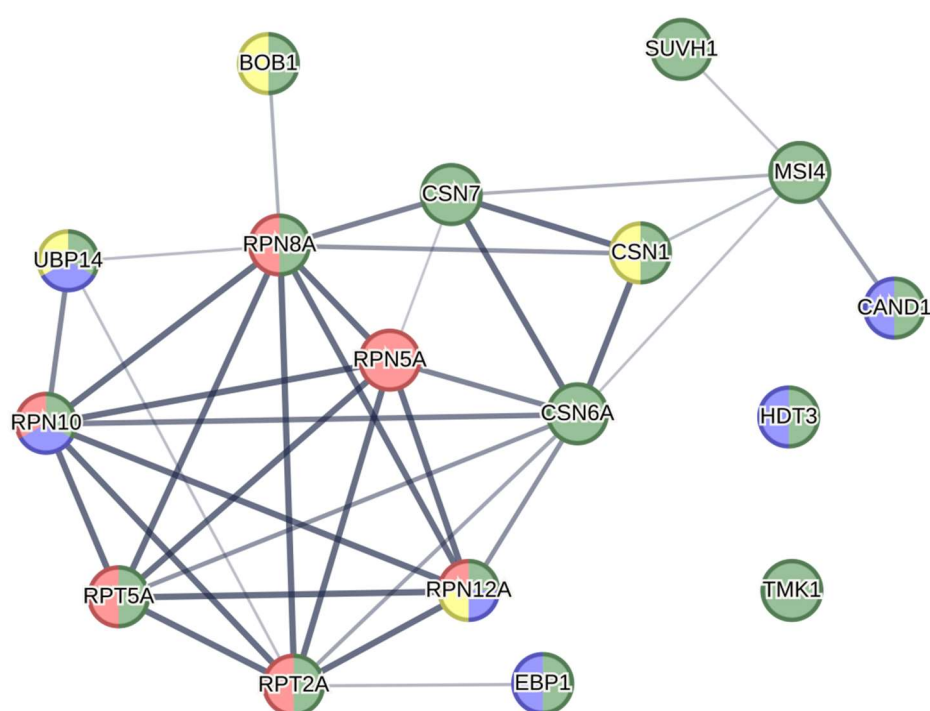


Fig. 5. Predicted protein-protein interaction networks and biological processes impaired by inhibition of auxin transport during the expression of embryogenic callus competence. Orthologues of sugarcane DAPs were identified in *Arabidopsis thaliana* using STRING and used to construct the network. Yellow: Embryo development; Blue: Response to hormone; Red: Proteasome; Green: Cellular process. The legends are in Table 1.

Inhibition of auxin transport affects the endogenous contents of polyamines and IAA in embryogenic callus

Total free PAs were quantified to observe whether auxin transport regulation is associated with polyamine metabolism during embryo maturation. All quantified PAs were more abundant in embryogenic callus from the control treatment than in those from the 100 μ M TIBA treatment (Fig. 6). Treatment with 100 μ M TIBA reduced the endogenous levels of

putrescine, spermidine, spermine and total free polyamines by 34, 66, and 37, respectively, compared to those in the control treatment. Additionally, it was observed that spermidine synthase (Sspon.02G0020240-4D) was down-regulated (Supplementary Table 1), corroborating the PAs quantification results.

Additionally, the endogenous IAA content was quantified in embryogenic callus to verify whether inhibiting auxin transport also promoted an imbalance in the accumulation of this hormone. As a result, we observed that pre-maturation treatment with 100 μ M TIBA resulted in a significant decrease in the endogenous content of IAA in embryogenic callus by approximately 61% after 42 days in maturation culture medium (Fig. 6).

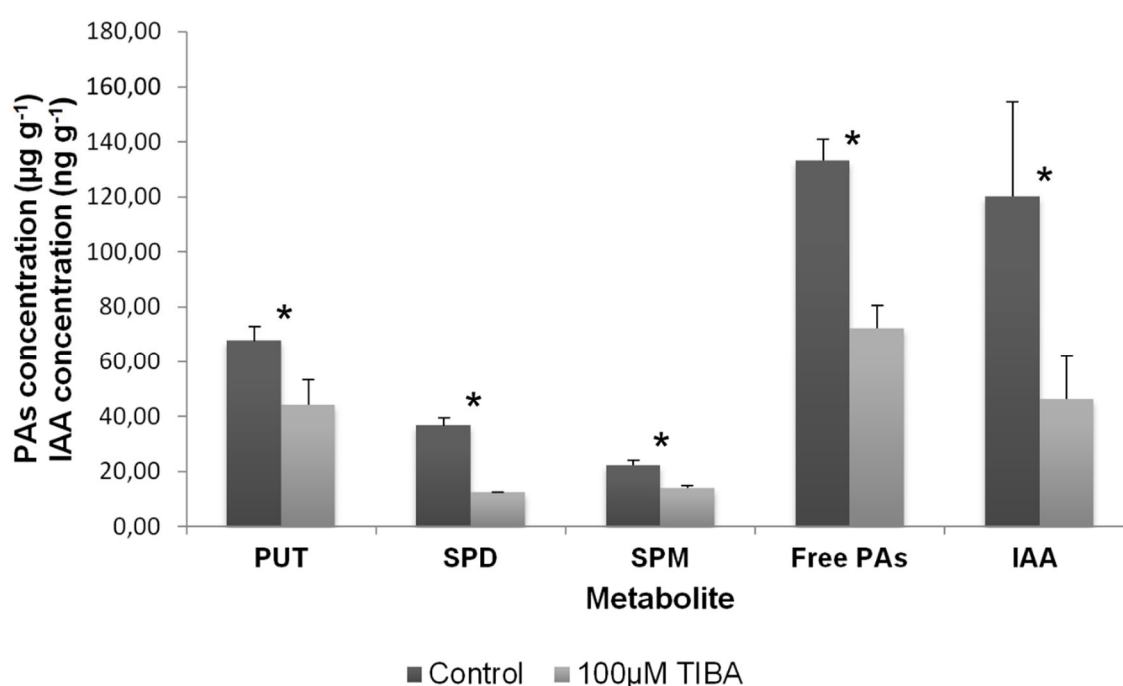


Fig. 6. Endogenous contents of putrescine (PUT), spermidine (SPD), spermine (SPM), total free polyamines (PAs; μ g g⁻¹ FM) and auxin (IAA) (ng g⁻¹ FM) in embryogenic callus from the control and 100 μ M TIBA pre-maturation treatments. The quantification of PAs was performed immediately after 21 days of pre-maturation treatment with TIBA, while the quantification of IAA was performed on the callus after 42 days of maturation. The means indicated by asterisks are significantly different according to T-test ($P < 0.05$). Mean $\pm$ standard deviation. $n = 3$. Coefficient of variation (CV)_{PUT} = 15%, CV_{SPD} = 5%, CV_{SPM} = 7%, and CV_{Free PAs} = 9%. CV_{IAA} = 31%.

Discussion

In sugarcane, the use of 2,4-D is required for embryogenic callus induction and multiplication phases during somatic embryogenesis (Silveira *et al.* 2013). However, during the multiplication phase, there is a progressive decrease in embryogenic competence after the fifth subculture (Passamani *et al.* 2019). Additionally, it has been shown that a pre-

maturation treatment of embryogenic callus without 2,4-D, prior to their transfer to a differentiation culture medium, significantly improves the production of somatic embryos in sugarcane during the maturation phase (Reis *et al.* 2021).

The transfer of embryogenic callus to a culture medium without 2,4-D promotes auxin biosynthesis (Fehér 2019). Auxin gradients are then re-established via PIN-mediated transport, activating WUSCHEL expression and the differentiation of somatic embryos at the callus periphery (Su *et al.* 2015). Our results showed that treatment of embryogenic callus with the auxin transport inhibitor TIBA compromised the formation of sugarcane somatic embryos and led to abnormal embryo development (Fig. 2, Fig. 3). Additionally, our data showed that the inhibition of auxin transport at pre-maturation step could also disrupt embryo development by reducing the accumulation of IAA in embryogenic callus under maturation conditions (Fig. 6). The inhibition of auxin transport during the induction of embryogenic callus also resulted in the disruption of PIN1 polar distribution and completely inhibited the regeneration of somatic embryos in *Arabidopsis* (Su *et al.* 2009). These findings confirm that generation of auxin gradients, mediated by polar auxin transport, are crucial during sugarcane somatic embryo induction.

Proteomic analysis showed that several kinases and transcription factors, including the Receptor-like kinase TMK2 protein (Sspon.01G0027230-3D, orthologous to AtTMK1), which is involved in the perception of auxin responses, were down-regulated in response to the 100 μ M TIBA pre-maturation treatment (Table 1). This transmembrane receptor kinase plays a role in the perception and transduction of signals during the differentiation process, as it controls root and hypocotyl growth, regulating cell expansion in response to auxin (Dai *et al.* 2013). In addition, TMK1 is involved in cell polarization and the distribution of PIN proteins to control auxin distribution during root growth in response to gravitropism (Wang *et al.* 2023). These observations imply that the Receptor-like kinase TMK2 might also be involved in the patterning of somatic embryos in sugarcane through auxin perception and the coordinated distribution of PIN proteins during development.

Proteomic analysis also revealed that different COP9 signalosome (CSN) components, including COP9 signalosome complex subunit 1 (Sspon.01G0000970-1A, orthologous to AtCSN1), COP9 signalosome complex subunit 6a (Sspon.06G0001040-2B, orthologous to AtCSN6A), and COP9 signalosome complex subunit 7 isoform X3 (Sspon.02G0007340-1A, orthologous to AtCSN7), were down-regulated after the 100 μ M TIBA pre-maturation treatment. The CSN is a complex present in eukaryotes that directs protein degradation via the ubiquitin/26S proteasome pathway, modulating auxin responses

in plants (Qin *et al.* 2020). These proteins are predicted to interact with various members of the proteasome machinery (Fig. 5). Silencing of CSN6A in *Arabidopsis* leads to various defects in auxin-associated patterning and developmental processes, demonstrating that CSN plays a role in protein degradation during hormone responses (Peng *et al.* 2001). Furthermore, auxin responses involve the interaction between CSN and the SCFTIR1 complex, leading to proteasome-mediated degradation of the AUX/IAA repressors (Qin *et al.* 2020). In this context, Cullin-associated NEDD8-dissociated protein 1 (Sspon.04G0017510-2B, orthologous to AtCAND1) is required for the activity and functions of the SCFTIR1 complex during auxin responses (Zhang *et al.* 2008). Additionally, 100 μ M TIBA pre-maturation treatment reduced the regulation of CAND1 in embryogenic callus (Table 1). In this sense, the inhibition of auxin transport in embryogenic callus undermines the differentiation and development of sugarcane somatic embryos, which may be associated with altered auxin responses due to a reduction in CAND1 and CSN proteins.

Epigenetic regulation through DNA and histone methylation/acetylation is known to modulate auxin responses associated with somatic embryogenesis (Morończyk *et al.* 2022). Accordingly, our study showed that the inhibition of auxin transport in embryogenic callus causes a reduction in the accumulation of proteins involved in the embryogenesis process (Fig. 4A) (Table 1). The histone-lysine N-methyltransferase, H3 lysine-9 specific SUVH1-like (Sspon.07G0005310-1A, orthologous to AtSUVH1) protein promotes histone H3 lysine 9 (H3K9) methylation-dependent gene silencing and is associated with plant development (Shen and Meyer 2004). In addition, SUVH1 has a predicted interaction with the WD-40 repeat-containing protein MSI4 (Sspon.03G0007870-1A, orthologous to AtMIS4), which has also been down-regulated and is involved in RNA-directed DNA methylation-mediated gene silencing (Zhou *et al.* 2021). Similarly, the Histone deacetylase HDT2 protein (Sspon.07G0003330-2B, orthologous to AtHDT3) was down-regulated and is involved in regulating the level of H3K9 methylation and acetylation, as well as being associated with abscisic acid responses (Luo *et al.* 2012). Our understanding of the epigenetic mechanisms orchestrating embryogenic development is currently evolving. Nevertheless, a wealth of indirect evidence highlights the significance of histone protein methylation/acetylation in auxin signaling pathways during somatic embryogenesis (Wójcikowska *et al.* 2020). Considering the observed reduction in the accumulation of histone deacetylase/acetyltransferase proteins in sugarcane embryogenic callus upon 100 μ M TIBA pre-maturation treatment, our findings offer valuable insights into this process. Future investigations on sugarcane somatic embryogenesis could employ modulators of histone methylation and/or acetylation to

systematically assess the direct involvement of these post-translational modifications in the differentiation and development of somatic embryos.

Hormonal responses usually involve the ubiquitination and degradation of negative regulators by the proteasome, allowing signal transduction (Silverblatt-Buser *et al.* 2018). In this context, ubiquitination is a post-translational modification of proteins that plays a role in different phases of somatic embryogenesis (Aguilar-Hernández and Loyola-Vargas 2018). Our data revealed that different proteins involved in hormone-responsive and developmental processes, including 26S proteasome regulatory subunit 4 homolog (Sspon.02G0000500-2P, orthologous to AtRPT2A), 26S proteasome non-ATPase regulatory subunit 8 homolog A (Sspon.06G0009170-2B, orthologous to AtRPN12A) and 26S proteasome non-ATPase regulatory subunit 4 (Sspon.01G0037080-1B, orthologous to AtRPN10), were down-regulated in response to the 100 μ M TIBA pre-maturation treatment. AtRPT2a is known to participate in root development (Yu *et al.* 2019) and is required for the maintenance of Arabidopsis meristems (Ueda *et al.* 2004). Additionally, AtRPN12A mutants show reduced sensitivity to auxin and cytokinin, altering their responses to these hormones during root and apical growth (Smalle *et al.* 2002), and AtRPN10 is involved in germination and abscisic acid and auxin responses in the roots (Smalle *et al.* 2003). In addition, ubiquitin carboxyl-terminal hydrolase 14 (Sspon.03G0022020-1P, orthologous to AtUBP14) was also down-accumulated in response to the 100 μ M TIBA pre-maturation treatment. Arabidopsis UBP14 mutants are embryonic lethal and do not develop beyond the globular stage, indicating the participation of this protein during morphogenesis and early development (Doelling *et al.* 2001). In this context, UBP14 also participates in auxin responses, and mutants of this gene show defects in embryo patterning due to the presence of repressors that are thought to be degraded via the proteasome (Majumdar *et al.* 2020). In this way, auxin transport should be important for regulating the proteasome and reducing the transcriptional repression of genes that promote somatic embryogenesis.

Auxin signaling leads to the ubiquitination of Aux/IAA repressor proteins, which then target the 26S proteasome for degradation and reduce the transcriptional repression of target genes (Santner and Estelle 2010). During somatic embryo development, auxin is synthesized and transported, creating auxin concentration gradients that determine the polarity of embryo (Su *et al.* 2009). In this sense, the inhibition of somatic embryo formation by TIBA may be caused by the loss of embryogenic competence or by the altered polarity of embryogenic cells due to blocked or delayed auxin transport (Choi *et al.* 2001). We observed both the suppression of somatic embryo differentiation and changes in development after the pre-

maturation treatment of embryogenic callus with 100 μ M TIBA (Fig. 2, Fig. 3). Accordingly, in the present study, the treatment of sugarcane embryogenic callus with 100 μ M TIBA resulted in the down-regulation of the Protein BOBBER 1 (Sspon.08G0011080-2B, orthologous to AtBOB1) followed by the inhibition of cotyledon differentiation (Fig. 2, Supplementary Table 1). The BOB1 protein is an important chaperone for correct auxin distribution and cotyledon cell specification during *Arabidopsis* zygotic embryogenesis (Jurkuta *et al.* 2009). Furthermore, BOB1 interacts with the proteasome subunits RPT2A, RPN1A and RPN8A during *Arabidopsis* zygotic embryo development (Silverblatt-Buser *et al.* 2018). Our data revealed a predicted interaction between the BOB1 and RPN8A proteins during the differentiation of sugarcane somatic embryos (Fig. 5). Thus, during the development of sugarcane somatic embryos, BOB1 may integrate auxin transport with proteasome-mediated protein catabolism during embryonic development. This process must be important for morphogenesis and embryo progression beyond the globular stage.

The ERBB-3 BINDING PROTEIN 1 (EBP1) is expressed mainly in the root meristem and is part of a mechanism that regulates the rate of cell differentiation (Lokdarshi *et al.* 2020). Functional analysis of the proteins down-accumulated in embryogenic callus after 100 μ M TIBA pre-maturation treatment revealed that EBP1 (Sspon.07G0010050-3C, orthologous to AtEBP1) likely interacts with members of the proteasome that play a role in plant development (Fig. 5). The effect of auxin on EBP1 is dose-dependent and is necessary for the expression of cell cycle genes, the regulation of cell growth and proliferation, and meristematic competence in *Arabidopsis* (Horváth *et al.* 2006). These functions in plant development seem to be conserved between monocots and dicots, given that the expression of *Zea mays* EBP1 (ZmEBP1) in *Arabidopsis* results in an increase in the size of leaves and siliques, as well as an increase in the number of seeds (Wang *et al.* 2016). This mechanism of growth regulation involves FERONIA, a plasma membrane receptor protein kinase with a central role in plant development that phosphorylates EBP1, promoting its localization in the cell nucleus to modulate the expression of target genes (Li *et al.* 2018). Given that EBP1 is associated with auxin responses and cell cycle regulation, the down-regulation of this protein may negatively affect the differentiation and development of sugarcane somatic embryos.

The embryogenic competence of sugarcane callus and somatic embryo differentiation are directly related to proton pump activity and accumulation (Passamani *et al.* 2018). Similarly, down-regulated proteins after 100 μ M TIBA pre-maturation treatment were enriched in cellular components and metabolic pathways involved in energy generation (Fig. 4) (Supplementary Fig. 2 and 3). This finding suggests that auxin responses play a crucial role

in regulating energy metabolism during sugarcane somatic embryogenesis. In this context, the disruption of energy-generating processes due to the inhibition of auxin transport may lead to a mismatch, potentially inducing stress and the darkening of sugarcane embryogenic callus, thereby hindering the differentiation of somatic embryos (Fig. 2). In addition, the accumulation of glutathione metabolism proteins in embryogenic callus from the 100 μ M TIBA pre-maturation treatment may indicate increased cellular stress caused by these changes in metabolism (Fig. 4B). Glutathione (γ -L-glutamyl-L-cysteinylglycine, GSH) is a multifunctional tripeptide involved in defense responses, reactive oxygen species homeostasis, plant development, and somatic embryogenesis (Pasternak *et al.* 2020). Therefore, the differentiation and development of somatic embryos are energy demanding, and auxin could be essential for signaling and coordinating metabolic pathways throughout these processes, in which inhibiting the transport of this hormone by using TIBA triggers stress-mitigating responses due to metabolic disorders.

Studies in papaya and sugarcane have shown that the concentrations of endogenous polyamines in embryogenic callus increase during development (Botini *et al.* 2021; Reis *et al.* 2016). Polyamines play an important role in the somatic embryogenesis of sugarcane, and the use of exogenous putrescine increases the efficiency of this process (Reis *et al.* 2016). The 100 μ M TIBA pre-maturation treatment resulted in the down-regulation of Spermidine synthase associated with a decrease in the endogenous PAs callus content (Fig. 6, Supplementary Table 1). Polyamines have several functions in plant development, representing nitrogen reserves, participating in antioxidant metabolism and stress signaling and contributing to growth (Rakesh *et al.* 2021). Thus, the decrease in the levels of spermine, spermidine and putrescine in embryogenic callus after treatment with TIBA may be related to the inhibition of sugarcane somatic embryo development. Polyamine biosynthesis competes with ethylene production through the precursor S-adenosyl methionine (SAM), which can be used in both pathways (Bitrián *et al.* 2012). Proteomic data revealed that 1-aminocyclopropane-1-carboxylate oxidase 1 (Sspon.03G0042170-2D), a protein that converts SAM to ethylene, was up-regulated in callus treated with TIBA. In this way, the decrease in polyamine accumulation during the inhibition of auxin transport may be related to an increase in ethylene accumulation, which disrupts hormonal responses and affects the development of sugarcane somatic embryos.

Conclusions

In the present study, the use of TIBA, an inhibitor of auxin transport, during the pre-maturation treatment, impaired the development of sugarcane somatic embryos. The comparative proteomic analysis revealed that auxin transport is involved in the regulation of proteins related to embryo development, ubiquitination, mitochondrial organization, proton transport, and pyruvate metabolism. Spermidine synthase was down-regulated in response to the 100 μ M TIBA pre-maturation treatment, and the resulting decrease in endogenous polyamine content suggests the integration between auxin transport and polyamine biosynthesis during somatic embryo differentiation. The chaperone protein BOBBER1 was down-accumulated after TIBA treatment and is associated with generation of auxin gradients, playing a role in responses related to proteasome activity. Accordingly, the Receptor-like kinase TMK2 may be important for auxin perception while Cullin-associated NEDD8-dissociated protein 1 modulates transcriptional responses. These results demonstrate the fine-tuning of auxin transport regulation and its importance during sugarcane somatic embryogenesis and somatic embryo differentiation, in addition to revealing new targets involved in the in vitro development of somatic embryos.

References

- Aguilar-Hernández V, Loyola-Vargas VM (2018) Advanced proteomic approaches to elucidate somatic embryogenesis. *Front Plant Sci* 9 doi:<https://doi.org/10.3389/fpls.2018.01658>.
- Almeida FA, Passamani LZ, Santa-Catarina C, Mooney BP, Thelen JJ, Silveira V (2020) Label-free quantitative phosphoproteomics reveals signaling dynamics involved in embryogenic competence acquisition in sugarcane. *J Proteome Res* 19: 4145-4157. doi:<https://doi.org/10.1021/acs.jproteome.0c00652>.
- Arruda P (2012) Genetically modified sugarcane for bioenergy generation. *Curr Opin Biotechnol* 23: 315-322. doi:<https://doi.org/10.1016/j.copbio.2011.10.012>.
- Bitrián M, Zarza X, Altabella T, Tiburcio AF, Alcázar R (2012) Polyamines under abiotic stress: Metabolic crossroads and hormonal crosstalks in plants. *Metabolites* 2: 516-528. doi:<https://doi.org/10.3390/metabo2030516>.
- Botini N, Almeida FA, Cruz KZCM, Reis RS, Vale EM, Garcia AB, Santa-Catarina C, Silveira V (2021) Stage-specific protein regulation during somatic embryo development of *Carica papaya* L. 'Golden'. *Biochim Biophys Acta* 1869: 140561. doi:<https://doi.org/10.1016/j.bbapap.2020.140561>.
- Choi YE, Katsumi M, Sano H (2001) Triiodobenzoic acid, an auxin polar transport inhibitor, suppresses somatic embryo formation and postembryonic shoot/root development in *Eleutherococcus senticosus*. *Plant Sci* 160: 1183-1190. doi:[https://doi.org/10.1016/S0168-9452\(01\)00357-0](https://doi.org/10.1016/S0168-9452(01)00357-0).

Cursi DE, Hoffmann HP, Barbosa GVS, Bressiani JA, Gazaffi R, Chapola RG, Fernandes Junior AR, Balsalobre TWA, Diniz CA, Santos JM, Carneiro MS (2021) History and current status of sugarcane breeding, germplasm development and molecular genetics in Brazil. *Sugar Tech* 24: 112–133. doi:<https://doi.org/10.1007/s12355-021-00951-1>.

Dai N, Wang W, Patterson SE, Bleecker AB (2013) The TMK subfamily of receptor-like kinases in *Arabidopsis* display an essential role in growth and a reduced sensitivity to auxin. *PLOS ONE* 8: e60990. doi:<https://doi.org/10.1371/journal.pone.0060990>.

De Rybel B, Audenaert D, Beeckman T, Kepinski S (2009) The past, present, and future of chemical biology in auxin research. *ACS Chem Biol* 4: 987-998. doi:<https://doi.org/10.1021/cb9001624>.

Dhonukshe P, Grigoriev I, Fischer R, Tominaga M, Robinson DG, Hašek J, Paciorek T, Petrášek J, Seifertová D, Tejos R, Meisel LA, Zažímalová E, Gadella TWJ, Stierhof Y-D, Ueda T, Oiwa K, Akhmanova A, Brock R, Spang A, Friml J (2008) Auxin transport inhibitors impair vesicle motility and actin cytoskeleton dynamics in diverse eukaryotes. *Proc Natl Acad Sci USA* 105: 4489-4494. doi:<https://doi.org/10.1073/pnas.0711414105>.

Distler U, Kuharev J, Navarro P, Levin Y, Schild H, Tenzer S (2014) Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat Methods* 11: 167-170. doi:<https://doi.org/10.1038/nmeth.2767>.

Doelling JH, Yan N, Kurepa J, Walker J, Vierstra RD (2001) The ubiquitin-specific protease UBP14 is essential for early embryo development in *Arabidopsis thaliana*. *Plant J* 27: 393-405. doi:<https://doi.org/10.1046/j.1365-313X.2001.01106.x>.

Elhiti M, Stasolla C (2011) Ectopic expression of the *Brassica* SHOOTMERISTEMLESS attenuates the deleterious effects of the auxin transport inhibitor TIBA on somatic embryo number and morphology. *Plant Sci* 180: 383-390. doi:<https://doi.org/10.1016/j.plantsci.2010.10.014>.

Fehér A (2019) Callus, Dedifferentiation, Totipotency, Somatic Embryogenesis: What these terms mean in the era of molecular plant biology? *Front Plant Sci* 10: 536. doi:<https://doi.org/10.3389/fpls.2019.00536>.

Fehér A (2015) Somatic embryogenesis, Stress-induced remodeling of plant cell fate. *Biochim Biophys Acta* 1849: 385-402. doi:<https://doi.org/10.1016/j.bbagr.2014.07.005>.

Heringer AS, Reis RS, Passamani LZ, de Souza-Filho GA, Santa-Catarina C, Silveira V (2017) Comparative proteomics analysis of the effect of combined red and blue lights on sugarcane somatic embryogenesis. *Acta Physiol Plant* 39 doi:<https://doi.org/10.1007/s11738-017-2349-1>.

Horváth BM, Magyar Z, Zhang Y, Hamburger AW, Bakó L, Visser RG, Bachem CW, Bögre L (2006) EBP1 regulates organ size through cell growth and proliferation in plants. *EMBO J* 25: 4909-4920. doi:<https://doi.org/10.1038/sj.emboj.7601362>.

Huang F, Kemel Zago M, Abas L, van Marion A, Galván-Ampudia CS, Offringa R (2010) Phosphorylation of conserved PIN motifs directs *Arabidopsis* PIN1 polarity and auxin transport. *Plant Cell* 22: 1129-1142. doi:<https://doi.org/10.1105/tpc.109.072678>.

Jurkuta RJ, Kaplinsky NJ, Spindel JE, Barton MK (2009) Partitioning the apical domain of the Arabidopsis embryo requires the BOBBER1 nudc domain protein. *Plant Cell* 21: 1957-1971. doi:<https://doi.org/10.1105/tpc.108.065284>.

Karami O, Philipsen C, Rahimi A, Nurillah AR, Boutilier K, Offringa R (2023) Endogenous auxin maintains embryonic cell identity and promotes somatic embryo development in Arabidopsis. *Plant J* 113: 7-22. doi:<https://doi.org/10.1111/tpj.16024>.

Li C, Liu X, Qiang X, Li X, Li X, Zhu S, Wang L, Wang Y, Liao H, Luan S, Yu F (2018) EBP1 nuclear accumulation negatively feeds back on FERONIA-mediated RALF1 signaling. *PLoS Biol* 16: e2006340. doi:<https://doi.org/10.1371/journal.pbio.2006340>.

Lokdarshi A, Papdi C, Pettkó-Szandtner A, Dorokhov S, Scheres B, Magyar Z, von Arnim AG, Bögre L, Horváth BM (2020) ErbB-3 BINDING PROTEIN 1 regulates translation and counteracts RETINOBLASTOMA RELATED to maintain the root meristem. *Plant Physiol* 182: 919-932. doi:<https://doi.org/10.1104/pp.19.00805>.

Luo M, Wang Y-Y, Liu X, Yang S, Lu Q, Cui Y, Wu K (2012) HD2C interacts with HDA6 and is involved in ABA and salt stress response in Arabidopsis. *J Exp Bot* 63: 3297-3306. doi:<https://doi.org/10.1093/jxb/ers059>.

Majumdar P, Karidas P, Siddiqi I, Nath U (2020) The ubiquitin-specific protease TNI/UBP14 functions in ubiquitin recycling and affects auxin response. *Plant Physiol* 184: 1499-1513. doi:<https://doi.org/10.1104/pp.20.00689>.

Morończyk J, Brąszewska A, Wójcikowska B, Chwiałkowska K, Nowak K, Wójcik AM, Kwaśniewski M, Gaj MD (2022) Insights into the histone acetylation-mediated regulation of the transcription factor genes that control the embryogenic transition in the somatic cells of Arabidopsis. *Cells* 11: 863. doi:<https://doi.org/10.3390/cells11050863>.

Mravec J, Skůpa P, Bailly A, Hoyerová K, Křeček P, Bielach A, Petrášek J, Zhang J, Gaykova V, Stierhof Y-D, Dobrev PI, Schwarzerová K, Rolčík J, Seifertová D, Luschnig C, Benková E, Zažímalová E, Geisler M, Friml J (2009) Subcellular homeostasis of phytohormone auxin is mediated by the ER-localized PIN5 transporter. *Nature* 459: 1136-1140. doi:<https://doi.org/10.1038/nature08066>.

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol Plant* 15: 473-497. doi:<https://doi.org/10.1111/j.1399-3054.1962.tb08052.x>.

Nanjo Y, Skultety L, Uváčková Lu, Klubíková K, Hajduch M, Komatsu S (2012) Mass spectrometry-based analysis of proteomic changes in the root tips of flooded soybean seedlings. *J Proteome Res* 11: 372-385. doi:<https://doi.org/10.1021/pr200701y>.

Passamani, Bertolazi AA, Ramos AC, Santa-Catarina C, Thelen JJ, Silveira V (2018) Embryogenic competence acquisition in sugarcane callus is associated with differential H(+)-pump abundance and activity. *J Proteome Res* 17: 2767-2779. doi:<https://doi.org/10.1021/acs.jproteome.8b00213>.

Passamani LZ, Reis RS, Vale EM, Sousa KR, Aragão VPM, Santa-Catarina C, Silveira V (2019) Long-term culture with 2,4-dichlorophenoxyacetic acid affects embryogenic competence in sugarcane callus via changes in starch, polyamine and protein profiles. *Plant Cell Tiss Org Cult* 140: 415-429. doi:<https://doi.org/10.1007/s11240-019-01737-w>.

Pasternak T, Palme K, Paponov IA (2020) Glutathione enhances auxin sensitivity in Arabidopsis roots. *Biomolecules* 10: 1550. doi:<https://doi.org/10.3390%2Fbiom10111550>.

Peng Z, Serino G, Deng XW (2001) Molecular characterization of subunit 6 of the COP9 signalosome and its role in multifaceted developmental processes in Arabidopsis. *Plant Cell* 13: 2393-2407. doi:<https://doi.org/10.1105/tpc.010248>.

Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M (2019) The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* 47: D442-D450. doi:<https://doi.org/10.1093/nar/gky1106>.

Qin N, Xu D, Li J, Deng XW (2020) COP9 signalosome: Discovery, conservation, activity, and function. *J Integr Plant Biol* 62: 90-103. doi:<https://doi.org/10.1111/jipb.12903>.

Rakesh B, Sudheer WN, Nagella P (2021) Role of polyamines in plant tissue culture: An overview. *Plant Cell Tiss Org Cult* 145: 487-506. doi:<https://doi.org/10.1007/s11240-021-02029-y>.

Reis RS, Vale Ede M, Heringer AS, Santa-Catarina C, Silveira V (2016) Putrescine induces somatic embryo development and proteomic changes in embryogenic callus of sugarcane. *J Proteomics* 130: 170-179. doi:<https://doi.org/10.1016/j.jprot.2015.09.029>.

Reis RS, Vale EM, Sousa KR, Santa-Catarina C, Silveira V (2021) Pretreatment free of 2,4-dichlorophenoxyacetic acid improves the differentiation of sugarcane somatic embryos by affecting the hormonal balance and the accumulation of reserves. *Plant Cell Tiss Org Cult* 145: 101-115. doi:<https://doi.org/10.1007/s11240-020-01995-z>.

Santner A, Estelle M (2010) The ubiquitin-proteasome system regulates plant hormone signaling. *Plant J* 61: 1029-1040. doi:<https://doi.org/10.1111/j.1365-313X.2010.04112.x>.

Shen W-H, Meyer D (2004) Ectopic expression of the NtSET1 histone methyltransferase inhibits cell expansion, and affects cell division and differentiation in tobacco plants. *Plant Cell Physiol* 45: 1715-1719. doi:<https://doi.org/10.1093/pcp/pch184>.

Silveira V, de Vita AM, Macedo AF, Dias MFR, Floh EIS, Santa-Catarina C (2013) Morphological and polyamine content changes in embryogenic and non-embryogenic callus of sugarcane. *Plant Cell Tiss Org Cult* 114: 351-364. doi:<https://doi.org/10.1007/s11240-013-0330-2>.

Silveira V, Iochet Segal Floh E, Handro W, Pedro Guerra M (2004) Effect of plant growth regulators on the cellular growth and levels of intracellular protein, starch and polyamines in embryogenic suspension cultures of *Pinus taeda*. *Plant Cell Tiss Org Cult* 76: 53-60. doi:<https://doi.org/10.1023/A:1025847515435>.

Silverblatt-Buser EW, Frick MA, Rabeler C, Kaplinsky NJ (2018) Genetic interactions between BOB1 and multiple 26S proteasome subunits suggest a role for proteostasis in regulating Arabidopsis development. *G3: Genes Genomes Genetics* 8: 1379-1390. doi:<https://doi.org/10.1534/g3.118.300496>.

Smalle J, Kurepa J, Yang P, Babiychuk E, Kushnir S, Durski A, Vierstra RD (2002) Cytokinin growth responses in Arabidopsis involve the 26S proteasome subunit RPN12. *Plant Cell* 14: 17-32. doi:<https://doi.org/10.1105/tpc.010381>.

Smalle J, Kurepa J, Yang P, Emborg TJ, Babiychuk E, Kushnir S, Vierstra RD (2003) The pleiotropic role of the 26S proteasome subunit RPN10 in Arabidopsis growth and development supports a substrate-specific function in abscisic acid signaling. *Plant Cell* 15: 965-980. doi:<https://doi.org/10.1105/tpc.009217>.

Strader LC, Zhao Y (2016) Auxin perception and downstream events. *Curr Opin Plant Biol* 33: 8-14. doi:<https://doi.org/10.1016/j.pbi.2016.04.004>.

Su YH, Liu YB, Bai B, Zhang XS (2015) Establishment of embryonic shoot–root axis is involved in auxin and cytokinin response during Arabidopsis somatic embryogenesis. *Front Plant Sci* 5 doi:<https://doi.org/10.3389/fpls.2014.00792>.

Su YH, Zhao XY, Liu YB, Zhang CL, O'Neill SD, Zhang XS (2009) Auxin-induced WUS expression is essential for embryonic stem cell renewal during somatic embryogenesis in Arabidopsis. *Plant J* 59: 448-460. doi:<https://doi.org/10.1111/j.1365-3113X.2009.03880.x>.

Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva Nadezhda T, Pyysalo S, Bork P, Jensen Lars J, von Mering C (2022) The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* 51: D638-D646. doi:<https://doi.org/10.1093/nar/gkac1000>.

Team RC R: A language and environment for statistical computing, R Foundation for Statistical Computing: 2023.

Ueda M, Matsui K, Ishiguro S, Sano R, Wada T, Paponov I, Palme K, Okada K (2004) The HALTED ROOT gene encoding the 26S proteasome subunit RPT2a is essential for the maintenance of Arabidopsis meristems. *Development* 131: 2101-2111. doi:<https://doi.org/10.1242/dev.01096>.

Wang T, Sui Z, Liu X, Li Y, Li H, Xing J, Song F, Zhang Y, Sun Q, Ni Z (2016) Ectopic expression of a maize hybrid up-regulated gene, ErbB-3 binding Protein 1 (ZmEBP1), increases organ size by promoting cell proliferation in Arabidopsis. *Plant Sci* 243: 23-34. doi:<https://doi.org/10.1016/j.plantsci.2015.11.002>.

Wang Y, Yan X, Xu M, Qi W, Shi C, Li X, Ma J, Tian D, Shou J, Wu H, Pan J, Li B, Wang C (2023) Transmembrane kinase 1-mediated auxin signal regulates membrane-associated clathrin in Arabidopsis roots. *J Integr Plant Biol* 65: 82-99. doi:<https://doi.org/10.1111/jipb.13366>.

Wiśniewski JR, Zougman A, Nagaraj N, Mann M (2009) Universal sample preparation method for proteome analysis. *Nat Methods* 6: 359-362. doi:<https://doi.org/10.1038/nmeth.1322>.

Wójcikowska B, Wójcik AM, Gaj MD (2020) Epigenetic regulation of auxin-induced somatic embryogenesis in plants. *Int J Mol Sci* 21: 2307. doi:<https://doi.org/10.3390/ijms21072307>.

Xavier LR, Almeida FA, Pinto VB, Passamani LZ, Santa-Catarina C, de Souza Filho GA, Mooney BP, Thelen JJ, Silveira V (2022) Integrative proteomics and phosphoproteomics reveals phosphorylation networks involved in the maintenance and expression of embryogenic competence in sugarcane callus. *J Plant Physiol* 268: 153587. doi:<https://doi.org/10.1016/j.jplph.2021.153587>.

Yu M, Meng B, Wang F, He Z, Hu R, Du J, Lai J, Yang C (2019) A SUMO ligase AtMMS21 regulates activity of the 26S proteasome in root development. *Plant Sci* 280: 314-320. doi:<https://doi.org/10.1016/j.plantsci.2018.12.014>.

Zhang J, Zhang X, Tang H, Zhang Q, Hua X, Ma X, Zhu F, Jones T, Zhu X, Bowers J, Wai CM, Zheng C, Shi Y, Chen S, Xu X, Yue J, Nelson DR, Huang L, Li Z, Xu H, Zhou D, Wang Y, Hu W, Lin J, Deng Y, Pandey N, Mancini M, Zerpa D, Nguyen JK, Wang L, Yu L, Xin Y, Ge L, Arro J, Han JO, Chakrabarty S, Pushko M, Zhang W, Ma Y, Ma P, Lv M, Chen F, Zheng G, Xu J, Yang Z, Deng F, Chen X, Liao Z, Zhang X, Lin Z, Lin H, Yan H, Kuang Z, Zhong W, Liang P, Wang G, Yuan Y, Shi J, Hou J, Lin J, Jin J, Cao P, Shen Q, Jiang Q, Zhou P, Ma Y, Zhang X, Xu R, Liu J, Zhou Y, Jia H, Ma Q, Qi R, Zhang Z, Fang J, Fang H, Song J, Wang M, Dong G, Wang G, Chen Z, Ma T, Liu H, Dhungana SR, Huss SE, Yang X, Sharma A, Trujillo JH, Martinez MC, Hudson M, Riascos JJ, Schuler M, Chen LQ, Braun DM, Li L, Yu Q, Wang J, Wang K, Schatz MC, Heckerman D, Van Sluys MA, Souza GM, Moore PH, Sankoff D, VanBuren R, Paterson AH, Nagai C, Ming R (2018) Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. *Nat Genet* 50: 1565-1573. doi:<https://doi.org/10.1038/s41588-018-0237-2>.

Zhang W, Ito H, Quint M, Huang H, Noël LD, Gray WM (2008) Genetic analysis of CAND1-CUL1 interactions in Arabidopsis supports a role for CAND1-mediated cycling of the SCFTIR1 complex. *Proc Natl Acad Sci USA* 105: 8470-8475. doi:<https://doi.org/10.1073/pnas.0804144105>.

Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, Zhang P, Banf M, Dai X, Martin GB, Giovannoni JJ, Zhao PX, Rhee SY, Fei Z (2016) iTAK: A program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol Plant* 9: 1667-1670. doi:<https://doi.org/10.1016/j.molp.2016.09.014>.

Zhou J-X, Du P, Liu Z-W, Feng C, Cai X-W, He X-J (2021) FVE promotes RNA-directed DNA methylation by facilitating the association of RNA polymerase V with chromatin. *Plant J* 107: 467-479. doi:<https://doi.org/10.1111/tpj.15302>.

Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK (2019) Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 10: 1523. doi:<https://doi.org/10.1038/s41467-019-09234-6>.

Perspectivas Futuras

As descobertas deste estudo reforçam o papel central de proteínas, modificações pós-traducionais e hormônios no controle da embriogênese somática em cana-de-açúcar, oferecendo uma base molecular sólida para o aprimoramento de protocolos de regeneração vegetal *in vitro*. A identificação de reguladores epigenéticos, fatores de transcrição, quinases e transportadores hormonais fornece alvos promissores para a engenharia metabólica e genética, ampliando as fronteiras da biotecnologia vegetal.

Dentre os achados mais relevantes, destaca-se o uso de 5-AzaC na reprogramação celular de calos embriogênicos de cana-de-açúcar. O tratamento com 50 μ M de 5-AzaC, em combinação com 2,4-D, promoveu um aumento significativo na formação de embriões somáticos maduros, demonstrando que a indução de um estado epigeneticamente permissivo favorece a aquisição de competência embriogênica. A análise proteômica e fosfoproteômica revelou a ativação de vias associadas à manutenção meristemática, remodelamento da cromatina, modulação hormonal e metabolismo energético. A fosforilação de transportadores de auxina (ABCB1, ABCB19, PIN1, PIN4), enzimas metabólicas (VHA-F, VDAC4, AMPD) e fatores relacionados à sinalização de ABA (ABI3, SRK2E, SUA) destacou o papel integrativo das modificações pós-traducionais na reprogramação celular. Esses resultados oferecem subsídios para o refinamento de protocolos de micropropagação e transformação genética em culturas recalcitrantes como *Saccharum* spp..

Além disso, nossos dados demonstram a importância do transporte de auxina durante a diferenciação embrionária somática. O uso de TIBA, em diferentes concentrações, evidenciou que o bloqueio do fluxo de auxina (especialmente com 100 μ M) compromete a formação de embriões somáticos, alterando o acúmulo proteico e a previsão de interações moleculares envolvidas com desenvolvimento embriogênico, ubiquitinação e metabolismo energético. A redução de espermidina e de proteínas como BOBBER1, TMK2 e CAND1 sugere uma conexão funcional entre auxina, poliaminas e o proteossoma, destacando uma complexa rede de regulação dependente do transporte hormonal. Esses achados ampliam a compreensão dos mecanismos *downstream* da sinalização de auxina durante a embriogênese, abrindo novas possibilidades para manipulação genética de vias hormonais.

Do ponto de vista agrônomo e industrial, a embriogênese somática é essencial para a regeneração de plantas geneticamente uniformes e livres de patógenos, com aplicação direta na micropropagação, produção de sementes sintéticas e transformação genética. Empresas de biotecnologia como SBW, Syngenta e CTC têm interesse direto em inovações

nos protocolos de embriogênese somática que possam ser patenteadas e aplicadas na produção agrícola em larga escala. As informações aqui geradas representam avanços tanto no nível fundamental quanto aplicado, e são altamente relevantes para o desenvolvimento de cultivares elite com maior valor agregado e adaptabilidade.

No contexto de outras culturas tropicais, como o mamoeiro, os achados também possuem alto potencial de translação. A integração de proteômica e fosfoproteômica permitiu elucidar redes regulatórias relacionadas à competência embriogênica, destacando o papel de proteínas associadas à homeostase hormonal, remodelamento epigenético e manutenção da totipotência, como os reguladores PLT2, GTE4 e ZAR1. A detecção de PLT2 exclusivamente em calos embriogênicos e sua interação predita com proteínas da via da auxina sugerem seu uso como marcador molecular para seleção de tecidos competentes, o que é particularmente promissor em culturas como o mamoeiro, onde a propagação de genótipos hermafroditas é crítica para a produção comercial.

Adicionalmente, a resposta dose-dependente ao 2,4-D em explantes embrionários de *C. papaya* revelou que diferentes concentrações desse regulador impactam diretamente a expressão de proteínas-chave envolvidas na divisão celular, sinalização hormonal e transcrição. Esses dados, aliados aos conhecimentos derivados de cana-de-açúcar, indicam que abordagens baseadas em modulação epigenética e hormonal podem ser amplamente adaptáveis a outras espécies, potencializando estratégias para otimização de protocolos de cultura de tecidos e engenharia genética.

Em suma, este trabalho contribui de forma significativa para o avanço da biotecnologia vegetal, fornecendo um repertório de alvos moleculares e mecanismos regulatórios que podem ser explorados para o desenvolvimento de sistemas de regeneração mais eficientes. A integração de dados ômicos, incluindo proteômica, fosfoproteômica e epigenômica, permitirá o desenvolvimento de estratégias racionais para indução de embriogênese somática, produção de sementes sintéticas e transformação genética em culturas economicamente relevantes, promovendo a sustentabilidade e inovação na agricultura tropical.