DINÂMICAS DE SINALIZAÇÕES ENVOLVIDAS NA AQUISIÇÃO DA COMPETÊNCIA EMBRIOGÊNICA EM CANA-DE-AÇÚCAR: UMA ABORDAGEM FOSFOPROTEÔMICA

FELIPE ASTOLPHO DE ALMEIDA

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY RIBEIRO – UENF
CAMPOS DOS GOYTACAZES – RJ
DEZEMBRO – 2020

DINÂMICAS DE SINALIZAÇÕES ENVOLVIDAS NA AQUISIÇÃO DA COMPETÊNCIA EMBRIOGÊNICA EM CANA-DE-AÇÚCAR: UMA ABORDAGEM FOSFOPROTEÔMICA

FELIPE ASTOLPHO DE ALMEIDA

"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: PROF. DR. VANILDO SILVEIRA

CAMPOS DOS GOYTACAZES – RJ
DEZEMBRO – 2020

DINÂMICAS DE SINALIZAÇÕES ENVOLVIDAS NA AQUISIÇÃO DA COMPETÊNCIA EMBRIOGÊNICA EM CANA-DE-AÇÚCAR: UMA ABORDAGEM FOSFOPROTEÔMICA

FELIPE ASTOLPHO DE ALMEIDA

"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"

Aprovado em 16 de dezembro de 2020

Comissão Examinadora:

Prof. Miguel Pedro Guerra (D. Sc., Fisiologia Vegetal) – UFRGS

Prof. Raul Antônio Sperotto (D. Sc., Biologia Celular e Molecular) – UFRGS

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

Prof. Vanildo Silveira (D.Sc., Biotecnologia) - UENF
(Orientador)

FICHA CATALOGRÁFICA
UENF - Bibliotecas
Elaborada com os dados fornecidos pelo autor.

A447 Almeida, Felipe Astolpho de.

Dinâmicas de sinalizações envolvidas na aquisição da competência embriogênica em cana-de-açucar : uma abordagem fosfoproteômica / Felipe Astolpho de Almeida. - Campos dos Goytacazes, RJ, 2020.

88 f. : il.
Bibliografia: 59 - 77.

Tese (Doutorado em Biotecnologia Vegetal) - Universidade Estadual do Norte Fluminense Darcy Ribeiro, Centro de Biociências e Biotecnologia, 2020.
Orientador: Vanildo Silveira.

1. Fosfoproteômica. 2. ABA. 3. SnRK1. 4. HDA6. 5. HDA19. I. Universidade Estadual do Norte Fluminense Darcy Ribeiro. II. Título.

CDD - 660.6

AGRADECIMENTOS

Primeiramente gostaria de agradecer à Universidade Estadual do Norte Fluminense Darcy Ribeiro, pela excelente estrutura disponibilizada e todo aporte oferecido para o desenvolvimento de minha tese. Adicionalmente, destaco a importância do Programa de Pós-Graduação em Biotecnologia Vegetal e de todo seu corpo docente, os quais ofereceram ensino de qualidade e relevantes contribuições científicas. Também, agradeço ao grupo de Biotecnologia e Fisiologia do Desenvolvimento Vegetal em especial ao meu orientador Dr. Vanildo Silveira, por todo suporte e infraestrutura prontamente disponibilizados que permitiram o desenvolvimento de toda minha tese. Além disso, agradeço também à Universidade de Missouri – Columbia e aos professores Dr. Jay J. Thelen e Dr. Brian P. Mooney que auxiliaram em parte das análises.

Destaco ainda meu agradecimento aos órgãos de financiamento da pesquisa e bolsas concedidos pela Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Por fim, agradeço aos meus amigos e familiares, especialmente meu pai Ricardo Gomes de Almeida e mãe Maria Aparecida Astolpho de Almeida aos quais foram parte essencial nessa desafiadora jornada.

SUMÁRIO

RESUMO ..... **v**

ABSTRACT ..... **vii**

1. INTRODUÇÃO **1**

2. REVISÃO BIBLIOGRÁFICA **2**

2.1. A cultura da cana-de-açúcar e suas aplicações na biotecnologia vegetal 2

2.2 Interação entre reguladores de crescimento vegetal e a expressão de genes envolvidos na embriogênese somática 6

2.3 O controle epigenético dinâmico durante a embriogênese somática .. 12

2.4 A fosfoproteômica aplicada ao estudo e compreensão das bases moleculares durante a embriogênese somática 17

3. OBJETIVOS **22**

3.1 Objetivo geral 22

3.2 Objetivos específicos 22

4. MATERIAL E MÉTODOS **22**

4.1. Material vegetal e condições de crescimento 22

4.2 Componentes do meio de cultura e indução dos calos 23

4.3 Análises fosfoproteômicas 23

4.3.1 Extração de proteínas 23

4.3.2 Digestão de proteínas 24

4.3.3. Enriquecimento dos fosfopeptídeos 25

4.3.4. Análises por espectrometria de massas 26

4.3.5 Análise dos dados fosfoproteômicos 27

4.3.6 Análises bioinformáticas 28

5. RESULTADOS **29**

5.1. Características morfológicas dos calos embriogênicos e não embriogênicos 29

5.2. Análises por fosfoproteômica comparativa 31

5.3. Características das fosfoproteínas, seus fosfosítios e o enriquecimento dos *motifs* 36

5.4. Alinhamento das fosfoproteínas relacionadas à competência embriogênica e conservação dos fosfosítios 42

6. DISCUSSÃO 47

6.2. Caracterização das DRPs nos grupos relacionados ao EC e NEC 47

6.3. Caracterização da homologia das fosfoproteínas, conservação dos fosfosítios e enriquecimento dos *motifs* 48

6.4. A fosforilação de proteínas está envolvida nas respostas hormonais e de estresse durante a aquisição da competência embriogênica em cana-de-açúcar 50

6.5. A fosforilação controla a reprogramação da expressão genética por meio de mecanismos epigenéticos e pós-transcricionais durante a aquisição da competência embriogênica em cana-de-açúcar 53

6.6. Os processos de desenvolvimento embriogênico estão ligados com a fosforilação de proteínas e a aquisição da competência embriogênica 55

7. CONCLUSÕES 57

8. REFERÊNCIAS 59

RESUMO

ALMEIDA, Felipe Astolpho de; Dsc.; Universidade Estadual do Norte Fluminense. Darcy Ribeiro; Dezembro, 2020; Dinâmicas de sinalizações envolvidas na aquisição da competência embriogênica em cana-de-açúcar: uma abordagem fosfoproteômica; Orientador: Vanildo SIlveira

A cana-de-açúcar é uma das principais fontes de matéria-prima para a produção de bioetanol e bioeletricidade, evidenciando sua importância social-econômica e para a promoção da sustentabilidade. Seu cultivo demanda grandes áreas, com solos férteis, bem irrigados e de genótipos com alto valor genético. Diante disso, a embriogênese somática se destaca com elevado potencial de aplicação na micropropagação clonal de variedades elite, livres de doenças, como parte essencial nos protocolos de transformação genética e na produção de sementes sintéticas. No presente estudo foi realizada uma análise fosfoproteômica *label-free* para identificar os eventos de sinalização relacionados com a aquisição da competência embriogênica em cana-de-açúcar. A análise fosfoproteômica comparativa foi realizada entre calos embriogênicos e não embriogênicos na fase de multiplicação. Foram identificadas 163 fosfoproteínas exclusivas para os calos embriogênicos, nove exclusivas para os calos não embriogênicos, e 51 mais abundantes e 40 menos abundantes nos calos embriogênicos em comparação com os calos não embriogênicos. A análise de Motif-x revelou o enriquecimento dos *motifs* [xxxpSPxxx], [RxxpSxxx] e [xxxpSDxxx], que correspondem aos sítios de fosforilação preditos para quinases envolvidas com respostas à estresses. As fosfoproteínas relacionadas aos calos embriogênicos identificadas nesse estudo, correspondem aos processos biológicos de sinalização induzida pelo ácido abscísico e com respostas a estresses abióticos; das quais incluem as proteínas OSK3, ABF1, LEAs e RD29Bs. Por outro lado, as fosfoproteínas diferencialmente reguladas nos calos não embriogênicos, como as EDR1 e PP2Ac-2, apresentam função de regulação negativa da sinalização do ácido abscísico, sugerindo uma relação entre as fosfoproteínas envolvidas nas vias do ácido abscísico e nas respostas de estresse na aquisição da competência embriogênica. Além disso, foram identificadas fosfoproteínas associadas com modificações epigenéticas, tais como as HDA6, HDA19 e TPL,

bem como com o metabolismo de RNA, incluindo AGO1, SCL30 e UB2C nos calos embriogênicos. Esses resultados revelam novos sítios de fosforilação para várias proteínas e identificam potenciais candidatos a biomarcadores para a aquisição da competência embriogênica em cana-de-açúcar.
Palavras-chave: *Fosfoproteômica; ABA; SnRK1; HDA6; HDA19*

ABSTRACT

ALMEIDA, Felipe Astolpho de; Dsc.; Universidade Estadual do Norte Fluminense. Darcy Ribeiro; December, 2020; Signaling dynamics involved in the acquisition of embryogenic competence in sugarcane: a phosphoproteomic approach; Advisor: Vanildo SIlveira

Sugarcane is one of the main feedstock sources for the production of bioethanol and bioelectricity, highlighting its social-economic importance and to promote sustainability. Its cultivation demands large areas, with fertile soils, well irrigated and of genotypes with high genetic value. Therefore, somatic embryogenesis stands out with a high potential for application in the clonal micropropagation of elite varieties, free of diseases, as an essential part in the protocols of genetic transformation and in the production of synthetic seeds. In this study, a label-free phosphoproteomic analysis was performed to identify signalling events related to the acquisition of embryogenic competence in sugarcane. The comparative phosphoproteomics analysis was performed between embryogenic and nonembryogenic calli at the multiplication phase. It was identified 163 phosphoproteins unique to embryogenic calli, nine unique to nonembryogenic calli, and 51 up-regulated and 40 down-regulated in embryogenic calli compared to nonembryogenic calli. Motif-x analysis revealed the enrichment of [xxxpSPxxx], [RxxpSxxx] and [xxxpSDxxx] motifs, corresponding to the predicted phosphorylation sites for several kinases related to stress responses. The embryogenic-related phosphoproteins identified in the present study correspond to the biological process of abscisic acid -induced signalling and abiotic stress response; they include OSK3, ABF1, LEAs and RD29Bs. On the other hand, the nonembryogenic-related phosphoproteins EDR1 and PP2Ac-2 are negative regulators of abscisic acid signalling, suggesting a relationship between phosphoproteins involved in the abscisic acid pathways and stress responses in the acquisition of embryogenic competence Moreover, were identified phosphoproteins associated with epigenetic modifications, such as HDA6, HDA19, and TPL, as well as with RNA metabolism, including AGO1, SCL30 and UB2C in the embryogenic calli. These results reveal novel phosphorylation sites

for several proteins and identify potential candidate biomarkers for the acquisition of embryogenic competence in sugarcane.
Keywords: *Phosphoproteomics; ABA; SnRK1; HDA6; HDA19*

.

1. INTRODUÇÃO

A cana-de-açúcar (*Saccharum* spp.) é uma cultura economicamente importante que se destaca na produção de açúcar e como matéria prima para a geração de energia limpa, permitindo reduzir a dependência dos combustíveis fósseis de forma ambientalmente sustentável. Para aumentar sua produtividade e desempenho agronômico, abordagens biotecnológicas como a cultura de tecidos *in vitro* e a engenharia genética são ferramentas potenciais que permitem ganhos substanciais no setor agrícola.

A embriogênese somática é um processo morfogenético *in vitro* em que células isoladas, ou um pequeno grupo de células somáticas, alteram sua programação gênica e dão origem a células embriogênicas (Fehér, 2015). Embriões somáticos e zigóticos passam pelos mesmos estágios de desenvolvimento em monocotiledôneas e dicotiledôneas, tornando essa técnica um modelo ideal para o estudo dos estágios iniciais do desenvolvimento vegetal (Méndez-Hernández et al., 2019).

A mudança do estado somático para o embriogênico compreende genes que codificam proteínas envolvidas em um complexo *crosstalk* entre diferentes vias, sinalizações hormonais e respostas a estresses (Aguilar-Hernández e Loyola-Vargas, 2018). Vários genes já foram identificados durante o desenvolvimento dos embriões somáticos de maneira análoga aos embriões zigóticos (Fehér, 2015). No entanto, a expressão gênica sofre reprogramação dinâmica na via embriogênica e requer um controle complexo até que as proteínas efetoras das respostas sejam funcionais (Duarte-Aké et al., 2019).

Em várias espécies, incluindo a cana-de-açúcar, análises proteômicas já desvendaram certos aspectos da aquisição de competência embriogênica (Heringer et al., 2018). No entanto, além das mudanças na abundância de proteínas, a transição das células somáticas para embriogênicas inclui várias modificações pós-traducionais (PTM, do inglês *post-translational modifications*), que podem impactar significativamente na complexidade do proteoma, afetando a estabilidade, atividade e localização das proteínas (Aguilar-Hernández e Loyola-Vargas, 2018).

Dessa forma, a fosfoproteômica quantitativa visa identificar o perfil de proteínas fosforiladas em um dado momento, podendo melhorar a compreensão

do papel da fosforilação nas vias de sinalização envolvidas na aquisição da competência embriogênica, permitindo identificar proteínas de baixa abundância expressas de uma forma espaço-temporal dependente (Aguilar-Hernández e Loyola-Vargas, 2018), e detectar potenciais novos candidatos a biomarcadores para embriogênese somática. Em estudos anteriores utilizando a abordagem fosfoproteômica em tecidos de calo de arroz (*Oryza sativa* L.), foram encontradas duas proteínas especificamente fosforiladas, a Embryonic Abundant protein 1 (OSEM) e a Proteína viviparous homolog VP1 (OsVP1), que estão relacionadas à sinalização hormonal na embriogênese (Wang et al., 2017). Em culturas de protoplastos de *Physcomitrella patens* (Hedw.) Bruch & Schimp, foram identificadas fosfoproteínas envolvidas na transdução de sinais, como a Somatic embryogenesis receptor kinase (SERK), e em respostas hormonais, incluindo fatores de resposta à auxina, apresentando papel importante no desenvolvimento celular reprogramado (Wang, X. et al., 2014).

No presente trabalho, usamos a técnica de fosfoproteômica comparativa *label-free* de alta resolução em calos embriogênicos (EC) e não-embriogênicos (NEC) de cana-de-açúcar para identificar e quantificar fosfoproteínas diferencialmente reguladas (DRPs), bem como seus sítios de fosforilação, visando revelar novos candidatos a biomarcadores e melhor compreender a dinâmica de sinalizações na aquisição da competência embriogênica em cana-de-açúcar.

2. REVISÃO BIBLIOGRÁFICA

2.1. A cultura da cana-de-açúcar e suas aplicações na biotecnologia vegetal

A cana-de-açúcar é uma cultura com relevante destaque na economia brasileira. Para a safra 2020/21, tem-se a estimativa de que sejam colhidos cerca de 642,1 milhões de toneladas (CONAB, 2020), reafirmando o Brasil como posto de maior produtor global de cana-de-açúcar. Em relação a produção de açúcar, estima-se que 39,3 milhões de toneladas sejam produzidos no período 2020/21, representando um incremento de 32% em relação à safra anterior (CONAB, 2020).

A utilização da biomassa de cana-de-açúcar para a produção de energia limpa surge como uma importante alternativa ao estabelecimento de novas fontes de energia, possibilitando reduzir a dependência de combustíveis fósseis de maneira ambientalmente sustentável (Ayodele et al., 2020). Por ser uma rica fonte de glicose e celulose, o suco e o bagaço da cana-de-açúcar têm sido amplamente utilizados na produção do bioetanol de primeira e segunda geração, respectivamente (Ayodele et al., 2020). No Brasil, a produção estimada de bioetanol a partir de cana-de-açúcar para o período 2020/21 está estimada em torno de 30,55 bilhões de litros (CONAB, 2020). Além disso, cerca de 7% da eletricidade produzida no Brasil é gerada a partir da biomassa de cana-de-açúcar, evidenciando o importante papel do setor canavieiro na produção de energias renováveis (Watanabe et al., 2020).

Apesar de ser uma eficiente matéria prima para a produção de bioetanol e bioeletricidade, a cana-de-açúcar necessita de grandes áreas com solos férteis e bem irrigados para seu cultivo (Ayodele et al., 2020). Nesse sentido, alguns pontos ainda necessitam de desenvolvimento tecnológico para se alcançar uma maior produção de forma mais eficiente do setor sucroalcooleiro, tais como o desenvolvimento de variedades de cana-de-açúcar mais produtivas; a integração da cana-de-açúcar com o meio ambiente, através do estudo do impacto das mudanças climáticas nas lavouras; e o controle de pragas e doenças (de Matos et al., 2020).

Visando aumentar a produtividade e eficiência no uso da cana-de-açúcar como matéria-prima, diferentes abordagens de melhoramento genético clássico já foram empregadas, buscando identificar variedades mais adaptadas a diferentes condições climáticas e ao ataque de doenças e pragas (Arruda, 2012). Entretanto, os programas de melhoramento genético clássico são demorados, levando de 10 a 12 anos até se desenvolver novas variedades ou híbridos comerciais (de Souza Barbosa et al., 2020). Além disso, a maioria das variedades de cana-de-açúcar cultivadas atualmente são oriundas da hibridação interespecífica entre *Saccharum officinarum* e *S. spontaneum*, em que aproximadamente 80% dos cromossomos são oriundos de *S. officinarum*, 10% de *S. spontaneum*, e 10% recombinantes entre as duas espécies (de Souza Barbosa et al., 2020). Adicionalmente, a cana-de-açúcar possui cerca de 100 a

130 cromossomos, os quais sofrem frequentes alterações cromossômicas numéricas, denominadas aneuploidia, o que impacta diretamente no aumento da complexidade do genoma e dificulta o desenvolvimento de programas de melhoramento genético na espécie (de Souza Barbosa et al., 2020).

A implementação de ferramentas para a manipulação genética pode potencializar os ganhos já obtidos nos programas de melhoramento genético, contribuindo para o desenvolvimento de cultivares de cana-de-açúcar mais produtivas (Arruda, 2012). Essa metodologia permite inserir em cultivares elites genes que codificam proteínas de resistência a estresses abióticos, a ataques de patógenos e/ou para o aumento na performance de algum parâmetro fisiológico de interesse agronômico (Arruda, 2012). Entretanto, um dos requisitos para o sucesso da transformação genética é o estabelecimento de um eficiente sistema de regeneração das plantas transformadas (Kausch et al., 2019).

Em cana-de-açúcar, a transformação genética via embriogênese somática é considerada essencial para uma eficiente regeneração das plantas transgênicas (Snyman, 2011; Arruda, 2012; Kausch, 2019). No Brasil, duas cultivares transgênicas de cana-de-açúcar já foram liberadas para o uso comercial, a primeira em 2017, chamada CTC20BT, e em 2019 a cultivar CTC9001BT também foi aprovada (CTNBio - Comissão Técnica Nacional de Biossegurança). Ambas cultivares expressam a proteína Cry1 que confere resistência a broca da cana-de-açúcar (*Diatraea saccharalis*) e utilizaram a embriogênese somática como parte do processo de transformação (CTNBio). Nesse contexto, a embriogênese somática aplicada a cana-de-açúcar se destaca como uma técnica com alto potencial para a regeneração e micropropagação clonal das plantas transgênicas nos protocolos de transformação genética (Snyman et al., 2011).

A embriogênese somática para cana-de-açúcar é um processo bem descrito e foi primeiramente relatado há mais de 37 anos (Ahloowalia e Maretzki, 1983; Ho e Vasil, 1983). Desde então, inúmeros trabalhos vêm sendo realizados aplicando o cultivo *in vitro* como ferramenta biotecnológica para o desenvolvimento de atividades comerciais e de pesquisa (Lakshmanan, 2006; Snyman et al., 2011; Suprasanna et al., 2011; Efferth, 2019), buscando melhor caracterizar o processo da embriogênese somática em cana-de-açúcar e

otimizar os protocolos já desenvolvidos (Silveira et al., 2013; Heringer et al., 2015; Reis et al., 2016; Heringer et al., 2017; Passamani et al., 2018; Passamani et al., 2020).

Como resultado, foram obtidos avanços na compreensão dos mecanismos envolvidos nas respostas da cana-de-açúcar e associados a estresses bióticos e abióticos, na propagação clonal em massa de genótipos elite, na produção de materiais livres de doenças, na conservação de germoplasmas, na produção de sementes sintéticas, no desenvolvimento de variedades transformadas bem como na elucidação da estrutura, função e interação de diferentes genes e proteínas em cana-de-açúcar (Lakshmanan, 2006; Snyman et al., 2011; Suprasanna et al., 2011; Heringer et al., 2018; Efferth, 2019).

Durante as fases de indução e multiplicação na embriogênese somática indireta, as células especializadas revertem seu estado diferenciado original se tornando células primordiais novamente (Fehér, 2019). Sob estímulos apropriados, essas células podem se diferenciar, dando origem a massas de tecidos transientes, com alta capacidade de proliferação, conhecidos como calos (Fehér, 2015). Esses tecidos podem ser classificados e separados de acordo com sua capacidade embriogênica: as células capazes de gerarem embriões somáticos, são classificadas como embriogênicas; enquanto que os calos incapazes de gerarem embriões somáticos, são classificados como não-embriogênicos (Nieves et al., 2003; Silveira et al., 2013). Esses dois tipos de calos se diferenciam em relação aos atributos morfológicos, bioquímicos e moleculares (Silveira et al., 2013; Heringer et al., 2015).

Na embriogênese somática de cana-de-açúcar, os calos embriogênicos são formados por células com formato isodiamétrico, núcleos proeminentes, alta relação núcleo: citoplasma, pequenos vacúolos, estruturas globulares organizadas e elevados níveis de poliaminas totais livres (Silveira et al., 2013). Já os calos não-embriogênicos apresentam células dispersas, alongadas e vacuoladas, com baixa relação núcleo: citoplasma (Silveira et al., 2013). Esses calos são mais macios e friáveis do que os embriogênicos e apresentam coloração mais escurecida (Silveira et al., 2013).

Ao nível molecular, já foram identificadas proteínas específicas para os calos embriogênicos em cana-de-açúcar, como a Nitric oxide synthase, Callose synthase 1 e Desiccation-related protein (Heringer et al., 2015). Essas proteínas estão envolvidas com respostas hormonais e de estresses e foram apontadas como possíveis biomarcadoras na aquisição da competência embriogênica (Heringer et al., 2015). Para os calos não-embriogênicos, proteínas envolvidas com modificações epigenéticas tais como o fator de transcrição MYB-like DNA-binding e com degradação de proteínas, tais como a Poliubiquitin-like, indicam uma regulação diferencial nos padrões da expressão gênica e uma menor atividade metabólica, levando à baixa diferenciação (Heringer et al., 2015).

Nesse sentido, o avanço tecnológico observado nas últimas décadas, como por exemplo a conclusão dos sequenciamentos do genoma para duas variedades diferentes de cana-de-açúcar (Garsmeur et al., 2018; Zhang et al., 2018), permitirá o desenvolvimento de estudos buscando melhorar as características e desempenho das variedades de cana-de-açúcar. Em especial, o uso das “ômicas” aplicadas à cultura de tecidos e na manipulação genética, abre novas perspectivas que podem auxiliar no desenvolvendo e otimização de novas abordagens da biotecnologia vegetal para cana-de-açúcar, e consequentemente, contribuir para o aumento da produtividade das lavouras canavieiras.

2.2 Interação entre reguladores de crescimento vegetal e a expressão de genes envolvidos na embriogênese somática

Na embriogênese somática *in vitro*, para que as células somáticas se reprogramem e expressem sua competência, até adquirir o estado totipotente, é necessário um fino e complexo balanço entre diferentes reguladores de crescimento vegetal (PGRs, do inglês *Plant Growth Regulators*) e suas concentrações (Nic-Can e Loyola-Vargas, 2016). Além disso, cada espécie possui uma demanda fisiológica particular para cada PGR, nos específicos estágios de desenvolvimento, fazendo da embriogênese somática um processo altamente genótipo-dependente (Jiménez, 2005; Hazubska-Przybył et al., 2020). Neste trabalho, tanto os hormônios produzidos de forma endógena, sintetizados naturalmente pelas células vegetais, quanto os PGRs sintéticos aplicados de

forma exógena que desempenham funções análogas aos hormônios endógenos, serão frequentemente referidos como PGRs.

As auxinas (Aux) e citocininas (CKs) são os principais PGRs envolvidos na divisão e diferenciação celular, e seu balanço determina qual rota morfogenética as células irão seguir. Desde as primeiras observações por Skoog e Miller (1957), é sugerido que uma proporção equilibrada de Aux/CKs induz a formação de calos, uma relação de alta concentração de Aux em relação às CKs origina raízes, e uma alta concentração de CKs em relação às Aux induz brotações. Para cana-de-açúcar, os níveis endógenos da auxina ácido-indol-3-acético (AIA) nos calos embriogênicos são duas vezes maiores que os níveis nos calos não-embriogênicos, enquanto que no geral, os níveis das CKs são aparentemente equivalentes em ambos os tipos de calos (Guiderdoni et al., 1995).

A importância da aplicação de Aux sintéticas exógenas, principalmente o ácido 2,4-Diclorofenoxiacético (2,4-D), nas etapas de indução da embriogênese somática já é bem estabelecida (Fehér, 2015; Méndez-Hernández et al., 2019; Tang et al., 2020). Essas substâncias compartilham similaridades estruturais com o AIA, desempenhando funções análogas às Aux (Wójcik et al., 2020). Entretanto, diferentemente das Aux, que são rapidamente inativadas por conjugação ou degradação, o 2,4-D é uma molécula altamente estável, desempenhando um efeito com ação prolongada (Wójcik et al., 2020). Além disso, o transporte do 2,4-D pelas proteínas carreadoras de Aux Pin-formed (PIN) é limitado, fazendo com que o 2,4-D se acumule nas células vegetais (Wójcik et al., 2020). Essas características fazem com que essa substância seja mais efetiva na indução da embriogênese somática, quando comparada às Aux endógenas (Karami e Saidi, 2010).

O 2,4-D induz as células competentes a reentrarem no ciclo celular e começarem a se proliferar, regulando positivamente a proteína do ciclo celular Cyclin-dependent kinase (CDK) e inibindo o fator de diferenciação Proporz1 (PRZ1) (Fehér, 2015). Além disso, o 2,4-D promove a interação entre o receptor Transport inhibitor response1 (TIR1) e a proteína repressora Auxin/indole-3-acetic acid (AUX/IAA). Após a interação, TIR1 sinaliza a proteína AUX/IAA para degradação, liberando a ativação dos genes responsivos à Aux (Wójcik et al.,

2020). O 2,4-D também possui relação direta com o balanço hormonal celular, elevando os níveis endógenos de Aux, interferindo nos gradientes de Aux, e levando à perda da polaridade original das células, fato que interrompe o programa de organogênese (Nic-Can e Loyola-Vargas, 2016).

Em diversas espécies vegetais o 2,4-D também participa na regulação positiva dos genes *SERK* (Hecht et al., 2001; Nolan et al., 2003; Santa-Catarina et al., 2004; Sharma et al., 2008; Maillot et al., 2009), marcadores moleculares associados com a competência embriogênica, que codificam proteínas quinases transmembrana Leucine-rich repeat que transmitem os sinais intracelulares através de cascatas de fosforilações em resposta aos estímulos ambientais (Santos e Aragão, 2009). A super-expressão dos genes *AtSERK* aumentam a formação de embriões somáticos a partir dos meristemas apicais (Hecht et al., 2001), mediando sinalizações cruzadas com as vias das Aux, em que podem se destacar o 2,4-D, e portanto, conferir a competência embriogênica (Hecht et al., 2001; Maillot et al., 2009).

Além de suas atividades relacionadas às Aux, o 2,4-D também atua como um agente estressante (Fehér, 2015). Já foi demonstrado que o 2,4-D está ligado à expressão de um grande número de fatores de transcrição, particularmente àqueles relacionados a estresses (Gliwicka et al., 2013). Os estados celulares de alta proliferação e ativo metabolismo aeróbico que ocorrem na presença do 2,4-D causam um acúmulo de espécies reativas de oxigênio (ROS, do inglês *Reactive Oxygen Species*), promovendo estresse oxidativo nas células, fato que reprograma o desenvolvimento e induz a embriogênese somática (Wójcik et al., 2020). As ROS, especialmente o H_2O_2, também atuam como mensageiros secundários durante a embriogênese induzida por Aux e estresse (Karami e Saidi, 2010).

Em nível de regulação epigenética, o 2,4-D exerce diferentes impactos na metilação do DNA e acetilação de histonas, regulando a expressão gênica durante a embriogênese somática (Fehér, 2015; Wójcikowska et al., 2020). Em culturas embriogênicas de *Acca sellowiana* (O. Berg) Burret, o tratamento com 2,4-D combinado com o agente desmetilante 5-azacitidina (5-azaC) aumentou as taxas de indução da embriogênese somática, bem como os níveis globais de metilação (Fraga et al., 2012). Em cenoura, o gene da *DNA methyltransferase*

(*Met1-5)* foi expresso transitoriamente após a indução da embriogênese somática por 2,4-D (Yamamoto et al., 2005). Recentemente, foi mostrado que o tratamento com 2,4-D está associado com a hipermetilação das regiões promotoras dos genes que codificam fatores de transcrição hormonais e relacionados a estresses, incluindo *Leafy Cotyledon (LEC1 e LEC2), Baby Boom (BBM), Agamous-like15 (AGL15)* e *Wuschel (WUS)*, que desempenham um papel central na rede regulatória que controla a indução da embriogênese somática (Grzybkowska et al., 2020). Esses dados ajudam a estabelecer uma relação direta entre o 2,4-D, as taxas de metilação do DNA e a indução da embriogênese somática.

Em paralelo com sua importância como indutor da embriogênese somática, a remoção do 2,4-D do meio de cultura desencadeia sinalizações para a expressão de genes relacionados com o desenvolvimento dos embriões somáticos (Jiménez, 2005; Fehér, 2015). Isto leva ao estabelecimento do transporte polar de Aux, formação do gradiente de Aux e indução da expressão dos genes *WUS*, que é um regulador central na determinação do meristema apical caulinar em calos embriogênicos (Su et al., 2009). Portanto, o 2,4-D, é um indutor crucial da competência embriogênica; entretanto, sua remoção ou redução no meio de cultura é necessária para o correto desenvolvimento morfogenético (Jiménez, 2005).

Diferentes estudos demonstraram que o 2,4-D e as Aux também possuem rotas interligadas com outros PGRs, tais como as CKs, ácido abscísico (ABA) e etileno, durante a embriogênese somática (Jiménez, 2005; Su et al., 2013; Fehér, 2015; Tang et al., 2020; Wójcik et al., 2020). O 2,4-D induz a expressão de genes que codificam fatores de transcrição centrais na biossíntese de ABA (Fehér, 2015). Culturas embriogênicas de cenoura tratadas com 2,4-D apresentaram um aumento nos níveis endógenos de ABA e uma diminuição nos níveis das CKs (Jiménez e Bangerth, 2001). Em sementes de soja, o tratamento com 2,4-D reprimiu a germinação, aumentando a transcrição de genes envolvidos com a biossíntese de ABA e os níveis endógenos desse PGR (Shuai et al., 2017). Além disso, o 2,4-D regulou negativamente a expressão de genes das rotas de biossíntese do ácido giberélico (GA) e diminuiu os níveis endógenos desse PGR (Shuai et al., 2017).

Já foi relatado para diversas espécies que as CKs e seus respectivos PGRs também atuam em diferentes etapas na embriogênese somática (Jiménez, 2005; Tang et al., 2020). Os sinais de resposta de CKs foram detectados em regiões específicas dos calos embriogênicos, correlacionadas com a indução da expressão dos fatores de transcrição *WUS-related homeobox5 (WOX5)*, que são reguladores chave no desenvolvimento do meristema apical radicular, e com a subsequente formação dos embriões somáticos (Su et al., 2015). A super-expressão dos genes repressores da sinalização de CKs (*Arabidopsis Response Regulator ARR7* e *ARR15*) impediu a formação do desenvolvimento do meristema apical radicular e a indução da embriogênese somática em *Arabidopsis* (Su et al., 2015).

A indução da desdiferenciação das células somáticas para a formação dos calos embriogênicos, também pode ocorrer via aplicação de diferentes condições estressantes, como o promovido por ferimento (Fehér, 2015). Esse estímulo promove tanto a biossíntese das CKs via ativação de genes que codificam suas enzimas biossintéticas (Ikeuchi et al., 2019), quanto a indução na expressão dos fatores de transcrição *Wound Induced Dedifferentiation 1 (WIND1, WIND2, WIND3* e *WIND4)*, que são reguladores chave no controle da desdiferenciação celular, e são necessários para a ativação das respostas por CKs (Iwase et al., 2011). Ambas as vias convergem na ativação da sinalização das CKs, que leva à reentrada das células no ciclo celular via regulação positiva da Cyclin D3;1 (CYCD3;1) (Ikeuchi et al., 2017).

Além das Aux e CKs, outros PGRs, como o ABA, são importantes na indução da embriogênese somática, bem como para o correto desenvolvimento dos embriões somáticos em diversas espécies (Karami et al., 2009; Karami e Saidi, 2010; Fehér, 2015). Entretanto, diferentemente das Aux, de uma forma geral os níveis endógenos de ABA parecem ser suficientes para a indução das culturas embriogênicas, especialmente em algumas monocotiledôneas, como cana-de-açúcar (Jiménez, 2005). Calos embriogênicos de cana-de-açúcar são caracterizados pelos acentuados níveis de ABA, em relação aos calos não-embriogênicos, nos quais esse PGR não foi detectado (Guiderdoni et al., 1995). Células embriogênicas de cenoura apresentam em torno de 2,5 vezes maior concentração de ABA em relação a embriões somáticos no estágio torpedo, e

concentração 67 vezes maior em relação às células não-embriogênicas (Karami et al., 2009).

Para a aquisição da competência embriogênica em cenoura, foi estabelecida uma relação direta entre os níveis endógenos de ABA e o sistema de indução por estresse (Yamamoto et al., 2005). Durante o tratamento estressante, a aplicação de fluridona, um inibidor da biossíntese de ABA, impediu a formação de embriões somáticos (Yamamoto et al., 2005). Em diversas espécies, os níveis endógenos de ABA aumentam em resposta ao estresse (Yamamoto et al., 2005; Karami et al., 2009; Karami e Saidi, 2010), porém, apenas a aplicação de ABA exógeno, sem a presença do fator estressante, não foi capaz de induzir a embriogênese somática em cenoura (Yamamoto et al., 2005).

Em resposta ao ambiente estressante da embriogênese somática, o ABA induz a expressão de quinases membros da família Sucrose non-fermenting protein (SNF1)-related kinase 2 (SnRK2), que fosforilam fatores de transcrição b-ZIP responsivos ao ABA, promovendo a sinalização que auxilia na promoção da tolerância ao estresse (Karami e Saidi, 2010). As proteínas detoxificantes Glutathione-S-Transferase (GST), que estão associadas com a tolerância às condições estressantes pela proteção celular contra os efeitos nocivos das ROS, também são induzidas por ABA durante a embriogênese somática (Karami et al., 2009). Esses dados indicam que tanto a presença do ABA, quanto as mudanças fisiológicas controladas pelo estresse, são importantes para a aquisição da competência embriogênica (Yamamoto et al., 2005).

Adicionalmente, já é bem descrito que as sinalizações por ABA em conjunto com a expressão de genes induzidos por estresse atuam de forma central nas etapas de desenvolvimento e maturação dos embriões somáticos (Yang e Zhang, 2010; Rai et al., 2011). Esses processos regulam a síntese e deposição das proteínas de reserva e das proteínas Late embryogenesis abundant (LEA), bem como das reservas de ácidos graxos e açúcares (Rai et al., 2011). Linacero et al. (2001) demonstraram que tecidos embriogênicos de cana-de-açúcar tratados com ABA tiveram um aumento considerável na expressão dos genes *LEA*.

O ABA também regula positivamente a biossíntese de Aux, o seu transporte polar e seus padrões de distribuição, controlando a resposta espacial de Aux, podendo influenciar na indução dos embriões somáticos (Su et al., 2013). Mutações no gene de biossíntese do ABA (*ABA2),* afetam a capacidade embriogênica em *Arabidopsis*, e o tratamento com fluridona interferiu na biossíntese local de Aux e no seu transporte polar, perturbando as respostas mediadas por Aux e o desenvolvimento de calos embriogênicos (Su et al., 2013).

Nesse cenário, as fases de indução, desenvolvimento, maturação e regeneração dos embriões somáticos são diretamente reguladas pela interação entre as diversas concentrações de PGRs e as condições de cultura (Rai et al., 2011). Essas intercomunicações acontecem via cascatas de sinalizações que, em nível molecular, são altamente controladas por genes, fatores de transcrição e proteínas. Além disso, as modificações epigenéticas no genoma, as pós-transcricionais no transcriptoma e as pós-traducionais no proteoma, aumentam a complexidade do controle dessas vias hormonais durante a embriogênese somática. Dessa forma, estudos moleculares e bioquímicos mais detalhados e direcionados para as espécies alvo ainda são necessários.

2.3 O controle epigenético dinâmico durante a embriogênese somática

Estresses e o balanço de PGRs são os principais indutores da embriogênese somática, uma vez que esses estímulos promovem a reprogramação da expressão gênica, levando uma célula diferenciada a se desdiferenciar (Pasternak e Dudits, 2019). Primeiramente, o ambiente estressante é percebido pelas membranas e parede celular, levando à produção de ROS pelas mitocôndrias, cloroplastos e peroxissomos (Bednarek e Orłowska, 2020). Como resultado, ocorre um desbalanço no metabolismo oxidativo que afeta a troca de sinais entre essas organelas e o núcleo, induzindo modificações epigenéticas, que controlam a reprogramação da expressão gênica e consequentemente a passagem do estado celular diferenciado para totipotente (Bednarek e Orłowska, 2020).

No processo de indução da embriogênese somática existe um complexo e dinâmico balanço entre ativação e inativação de vias gênicas essenciais para a aquisição da competência embriogênica (Ikeuchi et al., 2019; Pasternak e

Dudits, 2019). Neste contexto, já é bem estabelecido que as modificações epigenéticas são essenciais para a correta expressão de genes reguladores da embriogênese somática, como o *WUS, LEC, BBM*, *SERK* e *EMBRYONIC TEMPORAL REGULATOR (FUSCA3)* (De-la-Peña et al., 2015; Mahdavi-Darvari et al., 2015; Pasternak e Dudits, 2019).

A regulação epigenética pode ser dividida em metilações no DNA, modificações das histonas e na regulação mediada por pequenos RNAs não-codificantes (Mahdavi-Darvari et al., 2015). As modificações epigenéticas são baseadas na atividade de proteínas específicas que catalisam a adição ou remoção de grupamentos acetil ou metil nas cromatinas (histonas) e no DNA, respectivamente (Bednarek e Orłowska, 2020). Essas modificações também são acompanhadas pela expressão de pequenos RNAs não-codificantes relacionados ao sistema epigenético (Bednarek e Orłowska, 2020). A organização da cromatina permite a expressão ou repressão de genes, dependendo do grau de compactação em um *locus* específico (De-la-Peña et al., 2015). O nível em que a cromatina é remodelada antes da indução da embriogênese somática depende de uma série de fatores, como a origem do explante, a composição do meio de cultura, e principalmente, qual PGR será utilizado, e qual concentração será aplicada (Méndez-Hernández et al., 2019).

As histonas são proteínas que se associam ao DNA para compactá-lo, formando as cromatinas e auxiliando no arranjo do DNA em nucleossomos, consequentemente controlando a regulação da expressão gênica e divisão celular (Kumar e Van Staden, 2017). Essas proteínas sofrem PTMs como acetilação, metilação e fosforilação, que medeiam as interações das histonas com o DNA e com outros fatores ligantes (Ma et al., 2013). Dentre as PTMs, a acetilação de histonas tem demonstrado ser um potencial regulador da totipotência nas células vegetais, controlando a aquisição da competência embriogênica (Ma et al., 2013; De-la-Peña et al., 2015; Méndez-Hernández et al., 2019; Bednarek e Orłowska, 2020). As Histone acetyltransferases (HATs) adicionam, de forma dinâmica e reversível, grupamentos acetil em resíduos de lisina nas regiões N-terminais das histonas, sendo relacionadas com cromatinas descompactadas e com a ativação da expressão gênica (Ma et al., 2013). Por outro lado, as Histone deacetylases (HDAs) removem os grupamentos acetil das

histonas, compactando a cromatina e promovendo o silenciamento da expressão gênica (Ma et al., 2013). Essas enzimas desempenham importante papel em diferentes condições e fases do desenvolvimento vegetal, como na tolerância a condições de estresse, no controle do ciclo celular, nas respostas hormonais, no estabelecimento da polaridade em embriões zigóticos, e também na aquisição da competência embriogênica (Ma et al., 2013). A desacetilação das histonas demonstrou ser necessária para a reprogramação das células nos estágios iniciais da embriogênese somática, como a indução de calos, bem como para a manutenção da competência embriogênica (Méndez-Hernández et al., 2019; Pasternak e Dudits, 2019). Genes codificando enzimas HATs foram consideravelmente mais expressos nos estágios finais da embriogênese somática em *Arabidopsis* (Wickramasuriya e Dunwell, 2015).

Em *Arabidopsis*, linhagens duplo-mutantes para repressão das Histone Deacetylase 6 (HDA6) e 19 (HDA19) exibiram estruturas semelhantes a embriões em tecidos foliares (Tanaka et al., 2008). Além disso, esses mutantes tiveram a regulação positiva dos genes *LEC1, FUS3 e ABSICIC ACID INSENSITIVE3 (ABI3)*, demonstrando que ambas enzimas estão envolvidas no silenciamento desses genes, que são os principais reguladores na maturação de sementes e de respostas a estresses (Tanaka et al., 2008). As enzimas HDA6 e HDA19 também já foram descritas e correlacionadas com a regulação da expressão gênica durante a embriogênese zigótica de *A. thaliana* (Kim e Pandey, 2012) e nas fases iniciais da embriogênese somática de *Quercus suber* L. (Pérez et al., 2015).

Análises transcriptômicas em *Arabidopsis* identificaram que as HDAs são expressas durante toda embriogênese somática, de uma maneira relativamente estável, enquanto que as HATs demonstram maior expressão nas fases finais da embriogênese (Wickramasuriya e Dunwell, 2015). Em *Arabidopsis,* foi demonstrado que a formação de calos a partir de tecidos radiculares é dependente da acetilação de histonas (Sugimoto et al., 2010), e calos mutantes para o gene *HISTONE ACETYLTRANSFERASE (HAC1)* tiveram níveis mais reduzidos de *WUS* comparados com os selvagens (Li et al., 2011). Esses dados indicam um complexo *crosstalk* entre diferentes membros das enzimas HATs e HDAs durante a embriogênese somática.

Uma outra abordagem que tem se mostrado eficiente na investigação do papel da acetilação de histonas em diferentes fases do desenvolvimento vegetal, incluindo a embriogênese somática, é a aplicação do inibidor Tricostatina A (TSA) (Fehér, 2015; Méndez-Hernández et al., 2019; Bednarek e Orłowska, 2020). Esse composto foi originalmente identificado como um antifúngico isolado de *Streptomyces hygroscopicus* o qual inibe seletivamente e reversivelmente a atividade das HDAs pela ligação em seu sítio catalítico (Yoshida et al., 1995). Como resultado, o TSA promove uma hiperacetilação das histonas, e consequentemente um aumento da expressão gênica em diferentes níveis.

Inúmeros trabalhos vêm demonstrando que o TSA é um indutor da competência embriogênica, atuando no controle de diferentes genes relacionados com a embriogênese, como o *LEC1*, que regula a expressão das características relacionadas ao estado celular totipotente (Mahdavi-Darvari et al., 2015). Gametófitos masculinos de *Brassica napus* L. e *A. thaliana* cultivados na presença de TSA adquiriram o estado totipotente e seguiram as vias da embriogênese (Li et al., 2014). Em tecidos foliares de *Arabidopsis* tratados com o TSA, foi demonstrado que a formação de calos é dependente da desacetilação de histonas (Lee e Seo, 2018). Em *Picea abies* L. (Karst.), Uddenberg et al. (2011) relataram que embriões somáticos na fase de germinação tratados com TSA readquiriram a competência embriogênica para se diferenciarem em tecidos embriogênicos. Em *A. thaliana*, a aplicação do TSA induziu a embriogênese somática de uma maneira similar ao 2,4-D, regulando a ativação de genes responsivos às Aux e aos estresses, como os *LEC1, LEC2, BBM, MYB118, YUCCA1* e *YUCCA10* (Wójcikowska et al., 2018).

Além das modificações epigenéticas mediadas por acetilação/desacetilação das histonas, as metilações no DNA também estão intimamente ligadas com a regulação temporal de genes expressos na indução e desenvolvimento de embriões somáticos (De-la-Peña et al., 2015; Osorio-Montalvo et al., 2018; Bednarek e Orłowska, 2020). Esses processos envolvem a ação de diferentes famílias enzimáticas, que podem mediar a adição de grupamentos metil em bases de citosinas nos sítios CpG, CpHpG e CpHpH do DNA (Mahdavi-Darvari et al., 2015; Osorio-Montalvo et al., 2018). Entretanto, os eventos de metilação do DNA ocorrem de forma dinâmica, sendo específicos

para cada tecido, condição e nível hormonal, fazendo dessa modificação um processo altamente específico para cada genótipo e estágio da embriogênese somática (Nic-Can e De la Pena, 2014; De-la-Peña et al., 2015; Mahdavi-Darvari et al., 2015; Bednarek e Orłowska, 2020).

A capacidade embriogênica é associada com baixos níveis de metilação do DNA em diferentes espécies (Fehér, 2015). A desmetilação do DNA em sítios CHH foi sugerida como uma marca epigenética associada com a rediferenciação em genótipos altamente embriogênicos de algodoeiro (Guo et al., 2020). Linhagens embriogênicas de *Pinus radiata* D.don foram caracterizadas por apresentarem menores níveis de metilação global, associados com a maior expressão dos genes *SERK1, BBM* e *WOX2* (Bravo et al., 2017). Para algodão, a hipometilação aumenta as taxas da embriogênese pela ativação do gene *WUS* (Li et al., 2019). Porém, durante o desenvolvimento dos embriões somáticos de soja, foi observado um aumento na metilação do DNA ao longo do genoma, especialmente nos sítios CHH (Ji et al., 2019). Em *Cucurbita pepo* L., maiores níveis de metilação do DNA foram observados durante os estágios iniciais da embriogênese somática, e seus níveis foram reduzidos com a progressão do desenvolvimento dos embriões somáticos (Leljak-Levanić et al., 2004).

Inúmeros estudos demonstram que a aplicação de inibidores da metilação, como 5-azaC e Zebularina, pode influenciar nas diferentes fases da embriogênese somática e auxiliar na melhor compreensão dos eventos de metilação no genoma para cada espécie (Osorio-Montalvo et al., 2018; Pasternak e Dudits, 2019). O uso de 5-azaC estimulou a embriogênese somática em espécies como *Pinus pinaster* (Ait.), *B. napus, Hordeum vulgare* L. e *Cocos nucifera* L. (Osorio-Montalvo et al., 2018; Osorio-Montalvo et al., 2020). Embriões somáticos de cacau cultivados por longos períodos apresentaram declínio no potencial embriogênico, que foi revertido quando tratado com 5-azaC (Quinga et al., 2017). Além disso, calos não embriogênicos de algodão tratados com Zebularina, apresentaram formação de embriões somáticos (Li et al., 2019). O tratamento de culturas embriogênicas de cenoura com 5-azaC interrompeu a formação de embriões somáticos (Yamamoto et al., 2005). Para *Arabidopsis*, o 5-azaC reprimiu a expressão dos genes *LEC1, LEC2 e BBM*, inibindo a indução da embriogênese somática (Grzybkowska et al., 2018).

Apesar do evidente aumento no número de estudos investigando a influência da metilação do DNA durante a embriogênese somática, a maioria desses dados são baseados em abordagens em nível global (Osorio-Montalvo et al., 2018). Nesse sentido, análises do efeito da metilação em genes específicos e em quais regiões gênicas elas ocorrem, podem trazer novas informações sobre os mecanismos de regulação da embriogênese somática. Shibukawa et al. (2009) mostraram que a perda da metilação na região promotora do gene *LEC1* promoveu um aumento em sua expressão. De forma contrária, o tratamento dos explantes de *Arabidopsis* com 2,4-D promoveu a hipermetilação das regiões promotoras de genes responsivos a Aux e estresse, como *LEC1, LEC2, BBM*, *AGL15* e *WUS*, controlando a embriogênese somática (Grzybkowska et al., 2020).

Diversos autores vêm demonstrando os contrastantes resultados observados sobre os efeitos das modificações epigenéticas na embriogênese somática, como os diferentes níveis de hormônios endógenos existentes para cada genótipo, as concentrações dos PGRs utilizados e as condições de estresse *in vitro* (Fehér, 2015; Osorio-Montalvo et al., 2018). Portanto, a determinação do correto balanço entre os estados hipo/hipermetilado do DNA, bem como dos estados acetilados e desacetilados da cromatina e suas influências na regulação gênica, representam uma chave para o sucesso da embriogênese somática (Osorio-Montalvo et al., 2018). Entretanto, os mecanismos moleculares que controlam as modificações epigenéticas durante a embriogênese somática envolvem diversas rotas de sinalização em diferentes níveis de regulação que ainda não foram completamente elucidados, apesar do notório avanço nos últimos anos.

2.4 A fosfoproteômica aplicada ao estudo e compreensão das bases moleculares durante a embriogênese somática

Estudos utilizando abordagens genômicas, transcriptômicas e proteômicas vêm promovendo um ganho substancial acerca das bases moleculares que regulam a embriogênese somática. Além disso, o genoma da cana-de-açúcar foi sequenciado recentemente por dois grupos de pesquisas independentes (Garsmeur et al., 2018; Zhang et al., 2018), potencializando a

aplicação das tecnologias “ômicas” nos estudos biotecnológicos aplicados a cultura canavieira.

A reprogramação das vias celulares durante a embriogênese somática, em último estágio, requer a síntese, o enovelamento e modificações pós-traducionais nas proteínas recém-sintetizadas, bem como a remoção de proteínas que não são mais necessárias (Tchorbadjieva, 2016). Em várias espécies, incluindo a cana-de-açúcar, estudos proteômicos já abordaram os aspectos envolvidos na aquisição da competência embriogênica (Heringer et al., 2018). Dentre as proteínas diferencialmente abundantes identificadas nos calos embriogênicos destacam-se aquelas relacionadas com respostas a estresses, síntese/processamento de proteínas, modificações epigenéticas, degradação via proteassoma, metabolismo energético e de carboidratos (Heringer et al., 2018).

Sob condições de estresses abióticos e desbalanço hormonal, como o ambiente *in vitro*, o proteoma sofre mudanças consideráveis em sua abundância, e as proteínas passam por diferentes níveis de regulações até seu estado ativo (Wu et al., 2016). Essas PTMs alteram significativamente a complexidade do proteoma, afetando diretamente a estrutura, localização, estabilidade e atividade das proteínas, bem como suas interações com outros ligantes (Tchorbadjieva, 2016; Wu et al., 2016).

Dentre as PTMs, a fosforilação de proteínas desempenha um papel fundamental na sinalização hormonal, permitindo que as plantas percebam e respondam rapidamente às mudanças ambientais (Jagodzik et al., 2018). Além disso, fosforilações induzidas por estresses podem controlar a estabilidade das proteínas, atividade enzimática, e interação proteína-proteína, além de modificar o epigenoma (Kersten et al., 2009; Bigeard et al., 2014). Essa PTM é mediada por proteínas quinases e fosfatases que adicionam ou removem, reversivelmente, grupamentos fosfato nos aminoácidos serina (S), treonina (T) e tirosina (Y) (Kersten et al., 2009).

Cada proteína pode ser fosforilada por diferentes quinases, em múltiplos sítios de fosforilação, permitindo sua dinâmica regulação dependendo das condições fisiológicas em que a célula está exposta (Laugesen et al., 2006). A composição da sequência de aminoácidos que flanqueiam os sítios de

fosforilação do substrato (*motifs)* influenciam nas interações intermoleculares, e consequentemente, na ancoragem das quinases, sendo determinantes para definir a especificidade das proteínas quinases com seus substratos (Ubersax e Ferrell Jr, 2007). Os genomas de *Arabidopsis* e *O. sativa* codificam 1.019 e 1.429 proteínas quinases, respectivamente, representando cerca de 4% das proteínas anotadas nesses organismos. Esse número é aproximadamente o dobro do observado em mamíferos, indicando que nos vegetais existem complexas e específicas rotas de sinalização controladas pelas fosforilações (Vlad et al., 2008).

Receptores ligados à membrana plasmática, como as proteínas Receptor-like kinases (RLKs) e as Receptor-like proteins (RLPs), captam os estímulos exógenos e endógenos, transduzindo os sinais por uma complexa rede de fosforilações até a resposta celular (He et al., 2018). Posteriormente aos complexos de receptores de membrana, as fosfoproteínas Mitogen-activated protein kinases (MAPK) se ativam sequencialmente por meio de fosforilações, participando de forma central na sinalização de um receptor até seus efetores, amplificando os estímulos extracelulares e induzindo respostas celulares a uma variedade de estímulos, como os hormonais e os de estresses bióticos e abióticos (Kersten et al., 2009; Boudsocq et al., 2015).

Estudos aplicados a embriogênese somática identificaram fosfoproteínas membros da família RLKs que são constituintes essenciais para a formação do complexo inicial da embriogênese somática, como a proteína SERK (Méndez-Hernández et al., 2019). Em cenoura, o gene *SERK* codifica uma proteína Leucine-rich repeat (LRR), que são relacionadas com a aquisição da competência em embriões globulares (Schmidt et al., 1997). Essas proteínas se autofosforilam nos domínios quinase, controlando sua atividade e interação com duas LRRs (BRI1 e BAK1), atuando como receptoras dos brassinosteróides, que regulam diversos processos no crescimento e desenvolvimento (Karlova et al., 2009). Os genes *SERK1* apresentam altos níveis de expressão em calos de arroz durante a embriogênese somática, e também foram induzidos por moléculas de sinalização de defesa, como o ABA (Hu et al., 2005).

Quando calos embriogênicos foram comparados com calos não embriogênicos de *Vitis vinifera* L., a proteína Phosphatase 2A (PP2A) apresenta

maior acúmulo nos calos embriogênicos. Os autores associaram as PP2A ao transporte de Aux e às respostas diferenciais no alongamento celular, sendo importantes no processo embriogênico (Marsoni et al., 2008). Nos calos embriogênicos também foram identificadas proteínas Heat shock protein (HSPs), que são usualmente fosforiladas e amplamente propostas como biomarcadoras da embriogênese somática (Marsoni et al., 2008). Além disso, proteínas recombinantes da família quinase SnRK se autofosforilam e desempenham atividade quinase durante os estágios iniciais da embriogênese zigótica em arroz, sendo associadas com a regulação da divisão celular e o desenvolvimento do endosperma (Takano et al., 1998).

Investigando o papel das fosforilações em culturas de *B. napus*, Cordewener et al. (2000) demonstraram que a proteína HSP70 teve um padrão mais acentuado de fosforilação em culturas embriogênicas do que nas não embriogênicas. A enzima Ascorbate peroxidase, que geralmente é associada ao estresse e à detoxificação de ROS, foi identificada fosforilada em calos na embriogênese somática de *Triticum aestivum* L. (Rampitsch et al., 2006). Douétts-Peres et al. (2019) demonstraram que a inibição da proteína quinase Monopolar Spindle 1 (Mps1) afeta o crescimento celular e a diferenciação de culturas embriogênicas em *Araucaria angustifolia* (Bert.) O. Ktze. No decorrer da embriogênese zigótica de *O. sativa*, foram identificas proteínas diferencialmente fosforiladas envolvidas nas vias hormonais de biossíntese, sinalização e transporte polar de Aux, além de fatores de transcrição bZIP e SnRKs quinases, ambos responsivos ao ABA. Também foram identificadas diversas proteínas Cyclins e Cyclin-dependent kinases (CDKs) fosforiladas, indicando um controle da proliferação celular mediada por fosforilação (Qiu et al., 2016).

Recentemente, um estudo fosfoproteômico em *Elaeis guineensis* Jacq identificou a proteína E3 ubiquitin-protein ligase fosforilada em embriões somáticos durante a fase de maturação (Aroonluk et al., 2020). Essa proteína participa da maquinaria do proteassoma mediada pela sinalização por ubiquitina, e foi proposta como um regulador central nas respostas aos PGRs na embriogênese somática de *E. guineensis* (Aroonluk et al., 2020). Em culturas de protoplastos de *P. patens* foram identificadas diversas fosfoproteínas envolvidas no controle da reprogramação da expressão gênica através de mecanismos

epigenéticos, demonstrando a importância das fosforilações regulando metilações no DNA, modificações em histonas e remodelamento da cromatina (Wang, X. et al., 2014). Em calos de *O. sativa*, a proteína OSEM da família LEA, e a proteína OsVP1 foram especificamente fosforiladas. Essas fosfoproteínas são elementos responsivos ao ABA, e os autores sugeriram o envolvimento de ambas na embriogênese do arroz (Wang et al., 2017).

Os diferentes estudos fosfoproteômicos conduzidos durante os estágios iniciais do desenvolvimento vegetal destacam a importância da fosforilação de proteínas controlando inúmeros processos biológicos, como as vias de sinalização hormonal, as respostas a estresses, as modificações epigenéticas e a divisão celular (Wang, X. et al., 2014; Qiu et al., 2016; Wang et al., 2017; Aroonluk et al., 2020). Compreender essa rede complexa de regulações mediada pela fosforilação requer conhecimento específico e técnicas analíticas com alto poder de resolução, como exemplo a espectrometria de massas aliada a metodologias de enriquecimento de fosfopeptídeos (Riley e Coon, 2015). No cenário atual, os estudos fosfoproteômicos têm se concentrado majoritariamente na investigação dos papeis dessas PTM's nas respostas aos estresses abióticos (Wu et al., 2016) e durante o desenvolvimento vegetal (Yin et al., 2018). Porém, ainda são escassos os trabalhos aplicados ao estudo da embriogênese somática, especialmente utilizando análises por espectrometria de massas livre de gel aliadas à metodologia de enriquecimento dos fosfopeptídeos.

Com os avanços nas técnicas de espectrometria de massas de alta resolução e sensibilidade, bem como o desenvolvimento de metodologias de enriquecimento dos fosfopeptídeos, é possível superar determinadas limitações inerentes às técnicas tradicionais para identificação das PTMs, permitindo a quantificação de fosfopeptídeos em baixas concentrações, com alta eficiência e precisão, assim como a predição dos sítios de modificações e suas rotas de interações (Wetie et al., 2013). Desse modo, a espectrometria de massas *label-free* representa uma promissora ferramenta para os estudos fosfoproteômicos em grande escala, auxiliando na compreensão dos eventos de fosforilação das proteínas durante a sinalização hormonal, respostas ao estresse e aquisição da competência embriogênica nas vias de regeneração na cultura de tecidos vegetais.

3. OBJETIVOS

3.1 Objetivo geral

Identificar e quantificar as fosfoproteínas diferencialmente reguladas envolvidas na aquisição da competência embriogênica em cana-de-açúcar, através da fosfoproteômica comparativa em culturas embriogênicas e não embriogênicas.

3.2 Objetivos específicos

- Definir proteínas fosforiladas que estão associadas com a aquisição da competência embriogênica com potencial de se tornarem biomarcadoras da embriogênese somática;
- Realizar a anotação funcional das fosfoproteínas diferencialmente reguladas, determinar as interações proteína-proteína e identificar os processos biológicos mais representativos dessas interações;
- Caracterizar os sítios de fosforilação das proteínas identificadas bem como revelar novos sítios de fosforilação ainda não reportados na embriogênese somática de cana-de-açúcar;
- Determinar os *motifs* mais enriquecidos e identificar potenciais novos substratos para fosforilação por quinases;
- Verificar o *status* de conservação na sequência dos sítios de fosforilação e das fosfoproteínas identificadas.

4. MATERIAL E MÉTODOS

4.1. Material vegetal e condições de crescimento

A variedade SP803280 de cana-de-açúcar foi utilizada para a indução dos calos embriogênicos (EC) e não embriogênicos (NEC), de acordo com Passamani et al. (2018). Resumidamente, segmentos nodais imaturos de cana-de-açúcar foram plantados em substrato comercial PlantMax (DDL Agroindustria, Paulínia, SP, Brasil) e foram cultivados por 60 dias em casa de vegetação até atingirem aproximadamente 45 cm.

4.2 Componentes do meio de cultura e indução dos calos

Para a indução dos calos, foram selecionadas plantas sadias da casa de vegetação. As folhas maduras foram removidas, e os cilindros foliares resultantes foram desinfetados superficialmente em etanol 70% (Sigma-Aldrich, St. Louis, EUA) por 1 min, imersos em água sanitária comercial 30% (hipoclorito de sódio de 2 a 2,5%) Qboa® (Anhembi SA, Osasco, Brasil) por 30 min e lavados 3 vezes com água deionizada autoclavada em capela de fluxo laminar. Posteriormente, os cilindros foliares foram seccionados transversalmente em fatias de 2–4 mm de espessura e cultivados em tubos de ensaio de vidro (150 × 25 mm) contendo 10 mL de meio de cultura MS (Murashige e Skoog, 1962) (Phytotechnology Lab, Overland Park, EUA), suplementado com 20 g L^{-1} de sacarose, 2 g L^{-1} de Phytagel® (Sigma-Aldrich) e 10 µM de 2,4-D (Sigma-Aldrich). O pH do meio de cultura foi ajustado para 5,8, e o meio de cultura foi esterilizado em autoclave a 121 °C por 15 min. Os tubos contendo explantes foram mantidos no escuro a 25 ± 1 ºC por 45 dias.

Os calos induzidos foram transferidos para placas de Petri (90 × 15 mm) contendo 20 mL do mesmo meio de cultura, mantidos no escuro a 25 ± 1 °C, e posteriormente subcultivados três vezes, a cada 21 dias para multiplicação. Durante esse período de multiplicação, EC e NEC foram separados de acordo com suas características morfológicas (Silveira et al., 2013). Ambos EC e NEC foram mantidos por três ciclos de subcultivo e coletados para análise fosfoproteômica (0 dias).

4.3 Análises fosfoproteômicas

4.3.1 Extração de proteínas

Cinco repetições biológicas (300 mg de matéria fresca - FM cada repetição) de EC e NEC foram liofilizadas e depois maceradas com pilão em microtubos de 1,5 mL. Em seguida, 1 mL de tampão de extração (uréia 7 M, tioureia 2 M, triton X-100 a 2%, ditiotreitol a 1% (DTT; GE Healthcare, Piscataway, EUA) e 10 µL de Coquetel Inibidor de Protease Halt e Fosfatase, EDTA-free (Thermo Fisher Scientific, Waltham, EUA), foi adicionado ao pó da

amostra e homogeneizado com auxílio do pilão. As amostras foram agitadas em vórtex por 30 min a 8° C e centrifugadas a 16.000 g por 20 min a 4 °C.

Os sobrenadantes foram coletados e a concentração das proteínas foi mensurada com o kit Bradford Protein Assay (Bio-Rad Laboratories, Hercules, EUA) utilizando a gamaglobulina bovina (BGG; Bio-Rad) como padrão.

4.3.2 Digestão de proteínas

Quatro alíquotas de 400 µg de proteína para cada repetição biológica, tanto para EC quanto para NEC, foram digeridas com tripsina. A digestão de proteínas foi realizada usando a metodologia de preparação de amostra auxiliada por filtro (FASP), conforme descrito por Wiśniewski et al. (2009) com modificações (Burrieza et al., 2019).

O volume das alíquotas de proteína foi ajustado para 400 µL com tampão de extração, adicionado às unidades de filtro Microcon-30 kDa (Merck Millipore, Darmstadt, Alemanha) e centrifugado a 12.000 g por 20 min em temperatura ambiente (salvo indicação do contrário, todas as etapas de centrifugação foram realizadas nesta condição e o filtrado foi descartado ao final). As amostras foram lavadas com 200 µL de ureia 8 M em bicarbonato de amônio 50 mM (solução B; Sigma-Aldrich) e centrifugadas. Esta etapa foi repetida uma vez. Então, as amostras foram lavadas com 200 µL de bicarbonato de amônio 50 mM (solução A) e centrifugadas. Esta etapa foi repetida mais uma vez para a remoção completa da ureia antes da redução das proteínas. Em seguida, foram adicionados 100 µL de DTT (Bio-Rad) 50 mM feito fresco na solução A, agitados suavemente com vórtex e incubados durante 20 min a 60 °C (20 min de agitação a 350 rpm). Em seguida, 200 µL da solução B fria foram adicionados e centrifugados. As amostras foram lavadas com 200 µL de ureia 8 M em bicarbonato de amônio 50 mM (solução B; Sigma-Aldrich) e centrifugadas. Esta etapa foi repetida uma vez. As amostras foram lavadas com 200 µL de bicarbonato de amônio 50 mM (solução A) e centrifugadas. Esta etapa foi repetida uma vez para a remoção completa da ureia antes da redução das proteínas. Em seguida, foram adicionados 100 µL de DTT (Bio-Rad) 50 mM feito fresco na solução A, agitados suavemente com vórtex e incubados durante 20

min a 60 °C (20 min de agitação a 350 rpm). Em seguida, 200 µL de solução B fria foram adicionados e centrifugados.

Para a alquilação das proteínas, foram adicionados 100 µL de iodoacetamida 50 mM (Sigma-Aldrich) preparados em solução B, agitados suavemente em vórtex e incubados por 20 min a 25 °C no escuro (20 min de agitação a 350 rpm). Em seguida, 200 µL de solução B foram adicionados e centrifugados. Esta etapa foi repetida uma vez. Após, 200 µL da solução A foram adicionados e centrifugados por 15 min. Este passo foi repetido duas vezes. Na última lavagem, permaneceram aproximadamente 50 µL de amostra, e as unidades do filtro foram transferidas para novos microtubos. Para a digestão proteica, 100 µL de solução de tripsina (V5111; Promega, Madison, WI, EUA) (proporção final 1:100 enzima: proteína) foram adicionados, agitados suavemente em vórtex e incubados a 350 rpm por 18 h a 37 °C. Após, as unidades dos filtros foram suavemente agitadas em vórtex e centrifugadas para eluição do peptídeo. Em seguida, 50 µL da solução A foram adicionados e centrifugados. Esta etapa foi repetida uma vez. Os filtros foram descartados e 2,5 µL de ácido fórmico (88%) foram adicionados aos peptídeos digeridos recuperados. Os peptídeos foram secos a vácuo e ressuspensos em 200 µL de uma solução contendo 5% de acetonitrila e 0,1% de ácido fórmico em água MS (Sigma-Aldrich).

Os peptídeos das quatro alíquotas de cada amostra biológica foram reunidos e a concentração de peptídeos resultante foi estimada medindo $A_{205\ nm}$ usando um espectrofotômetro NanoDrop 2000c (Thermo Fisher Scientific).

4.3.3. Enriquecimento dos fosfopeptídeos

O kit de enriquecimento de fosfopeptídeos TiO_2 High-Select (Thermo Fisher Scientific) foi usado para enriquecer os fosfopeptídeos dos peptídeos trípticos. Alíquotas de 500 µg de peptídeos digeridos de cada amostra biológica foram secas a vácuo, suspensas em 150 µL de uma solução de 5% de acetonitrila e 0,1% de ácido fórmico em água MS (Sigma-Aldrich), e colocadas nas pontas das colunas TiO_2 Spin. As etapas de enriquecimento foram realizadas de acordo com as instruções do fabricante. Após a eluição dos fosfopeptídeos, o eluído foi seco a vácuo e suspenso em 25 µL de solvente aquoso (5% de

acetonitrila, 0,1% de ácido fórmico em água MS). A concentração de fosfopeptídeos foi estimada medindo $A_{205\ nm}$ usando um espectrofotômetro NanoDrop 2000c (Thermo Fisher Scientific). Os fosfopeptídeos foram armazenados a -80 °C antes das análises.

4.3.4. Análises por espectrometria de massas

As análises por espectrometria de massas foram realizadas usando um sistema de cromatografia de nano-fluxo *nanoElute* (Bruker Daltonics, Bremen, Alemanha) acoplado *on-line* a um espectrômetro de massas híbrido *trapped ion mobility spectrometry - quadrupole time of flight* (timsTOF Pro, Bruker Daltonics) com uma fonte de ionização *nanoeletrospray CaptiveSpray* (Bruker Daltonics) (Beck et al., 2015).

Primeiro, os fosfopeptídeos foram liofilizados e ressuspensos em solvente aquoso (5% de acetonitrila, 0,1% de ácido fórmico) para atingir uma concentração final de 500 ng μL^{-1}. Uma alíquota de 2 µL equivalente a 1 µg de fosfopeptídeos digeridos foi injetada em uma coluna *trap* C18 (Thermo Fisher Scientific; PepMap 100, 5 cm x 300 µm) e separada utilizando um gradiente de 400 µL min^{-1} 70 min em uma coluna analítica C18 (BEH-C18 com partículas de 1,7 µm, 20 cm de comprimento x 75 µm de diâmetro interno; Waters). A temperatura da coluna foi mantida em 50 °C. Para eluição dos fosfopeptídeos, foi usado o gradiente binário: fase móvel A, composta por água MS e ácido fórmico 0,1%, e fase móvel B, composta por acetonitrila 100% e ácido fórmico 0,1%. O gradiente de eluição foi realizado da seguinte maneira: 2% B por 10 min seguido por um gradiente de 15 min de 2 a 17% B, aumentando de 17 a 25% B até 45 min; aumentando de 25 a 37% B até 52 min; elevando de 37 a 80% B até 58 min; mantendo constante em 80% até 70 min; diminuindo para 2% B até o equilíbrio da coluna. O tempo total de aquisição da corrida foi de 60 min (após o carregamento na coluna *trap*). Os dados de MS e MS/MS foram adquiridos usando o método PASEF (Meier et al., 2015; Meier et al., 2018). A tensão do capilar foi ajustada para 1400V, tims-on, PASEF-on Tempo de ciclo para 1 MS e 10 PASEF ~ 1,5 seg (aproximadamente 100 MS/MS adquiridos por ciclo).

4.3.5 Análise dos dados fosfoproteômicos

Os dados *raw* foram analisados com o *software* MaxQuant (Cox e Mann) V. 1.6.3.4 que foi adaptado para lidar com o formato de dados TIMS e reúne *clusters* de isótopos quadridimensionais - definidos por m/z, tempo de retenção, mobilidade iônica e intensidade - a partir dos espectros TIMS-MS e extratos dos espectros MS/MS separados por mobilidade iônica (Meier et al., 2018). A carbamidometil cisteína foi definida como modificação fixa, enquanto a oxidação de metionina, a acetilação do região N-terminal da proteína e as fosfo-STY, foram definidas como modificações variáveis. Os parâmetros de digestão foram definidos como “específicos” e “Tripsina/P” e foram permitidas até duas clivagens perdidas. O recurso de correspondência entre as corridas do MaxQuant não foi utilizado. Pelo menos um peptídeo único/*razor* foi definido para identificação das proteínas. A taxa de falsa descoberta (FDR) para a correspondência espectral dos peptídeos e identificação das proteínas foi definida para o máximo de 1%, com um comprimento mínimo de peptídeo de seis aminoácidos. Todos os outros parâmetros do MaxQuant foram mantidos com seus valores padrão para o formato de dados Bruker timsTOF-DDA. Os espectros MS/MS foram pesquisados contra o banco de dados de proteínas gerado a partir do sequenciamento do haploide *Saccharum spontaneum*, AP85-441 (versão do banco de dados: v20190103) (Zhang et al., 2018) e 245 potenciais contaminantes pelo mecanismo de pesquisa integrado Andromeda (Cox, 2011), parte do *software* MaxQuant. A quantificação das proteínas identificadas foi realizada usando a quantificação *label-free* normalizada (intensidade LFQ), usando os algoritmos MaxLFQ (Cox et al., 2014) que integram o *software* MaxQuant.

Após a análise dos dados pelo MaxQuant, os sítios de fosforilação identificados (tabela de fosforilação fosfo STY (sites).txt) foram filtrados para apenas aqueles que apresentaram probabilidade de localização significativa ≥0,75). Apenas as proteínas contendo peptídeos modificados com fosfo significativo (STY) e apresentando intensidade de LFQ em pelo menos quatro repetições biológicas ou ausentes em todas as repetições biológicas (para proteínas exclusivas), para ambos os tratamentos, foram consideradas para a análise da fosfoproteômica comparativa. O acúmulo diferencial das proteínas foi

analisado usando o teste *t* de Student (bicaudal). As proteínas (EC/NEC) com $P<0,05$ foram consideradas mais abundantes se o $\log_2$ do *fold change* (FC) >0,5, e menos abundantes se o $\log_2$ do FC <-0,5.

Os dados *raw* gerados para os espectros adquiridos pelas análises por espectrometria de massas foram depositados no *ProteomeXchange Consortium* (Deutsch et al., 2016) por meio do repositório PRIDE (Perez-Riverol et al., 2019), com o identificador PXD018054 para o conjunto de dados.

4.3.6 *Análises bioinformáticas*

As sequências das proteínas foram submetidas a pesquisa BLAST no banco de dados de proteínas vegetais não redundantes do NCBI (taxa: 33090, Viridiplantae), seguida por anotação funcional com o banco de dados eggNOG integrado ao *software* OmicsBox (BioBam Bioinformatics SL, Valencia, Espanha). Para confirmar a correta anotação dos fosfopeptídeos, as sequências de aminoácidos também foram submetidas à análise por BLAST no OmicsBox.

Para determinar as localizações subcelulares de alta significância das proteínas fosforiladas em EC e NEC foi utilizada a ferramenta SUBA4, e os dados das sequências das proteínas no formato *fasta* foram usados como entrada para a análise por BLAST contra o banco de dados de referência de *Arabidopsis thaliana*. Para determinar a similaridade estatística entre as amostras das fosfoproteínas em EC e NEC, foi realizada a análise dos componentes principais (PCA), utilizando o *software* R.

As fosfoproteínas diferencialmente reguladas foram utilizadas para a construção de uma rede de interação proteína-proteína. O primeiro grau de interação foi obtido com a ferramenta *Search Tool for the Retrieval of Interacting Genes/Proteins* (V.11.0; http://string-db.org) (STRING) (Szklarczyk et al., 2018) pesquisando no banco de dados de *Arabidopsis* e utilizando um *score* mais restritivo (0,900). As redes de interação proteína-proteína resultantes foram utilizadas como entrada para a análise *downstream* no *software* CYTOSCAPE V.3.8.2 (Shannon et al., 2003), e os plug-ins Moduland e BinGO foram utilizados para avaliar e validar as proteínas de super-representação significativas em processos biológicos.

Para identificar os *motifs* conservados nos fosfopeptídeos diferencialmente regulados, foi utilizada a plataforma MEME Suite (Bailey et al., 2009). *Motifs* enriquecidos foram estimados usando o algoritmo motif-x na ferramenta do *software* MOMO (Cheng et al., 2018). As quinases preditas para cada *motif* enriquecido foram identificadas com o NetPhos 3.1 (Ingrell et al., 2007), GPS (Xue et al., 2008) e a bibliografia disponível (van Wijk et al., 2014; Wang, K. et al., 2014; Vannini et al., 2019).

A homologia das fosfoproteínas identificadas listadas na Tabela 1, bem como a conservação dos resíduos fosforilados, entre cana-de-açúcar e os bancos de dados melhores anotados para monocotiledôneas (*Zea mays*), e dicotiledôneas (*Arabidopsis thaliana*), foram determinadas usando o alinhamento na plataforma ClustalX 2.1 (Larkin et al., 2007), e as sequências alinhadas foram editadas com o *software* GeneDoc V. 2.70.

5. RESULTADOS

5.1. Características morfológicas dos calos embriogênicos e não embriogênicos

Comparando os calos EC e NEC, os EC foram capazes de produzir grandes quantidades de embriões somáticos (Figuras 1B e 1C) regenerando plântulas somáticas normais no final do processo (Figura 1C). Em contraste, os calos NEC foram incapazes de se diferenciar em embriões somáticos e as células dos calos possivelmente sofreram oxidação, o que levou à necrose ao longo dos dias de maturação (Figuras 1D, 1E e 1F). Essas observações destacam que os EC apresentam potencial embriogênico para se diferenciar em embriões somáticos, enquanto nos NEC esse potencial está ausente.

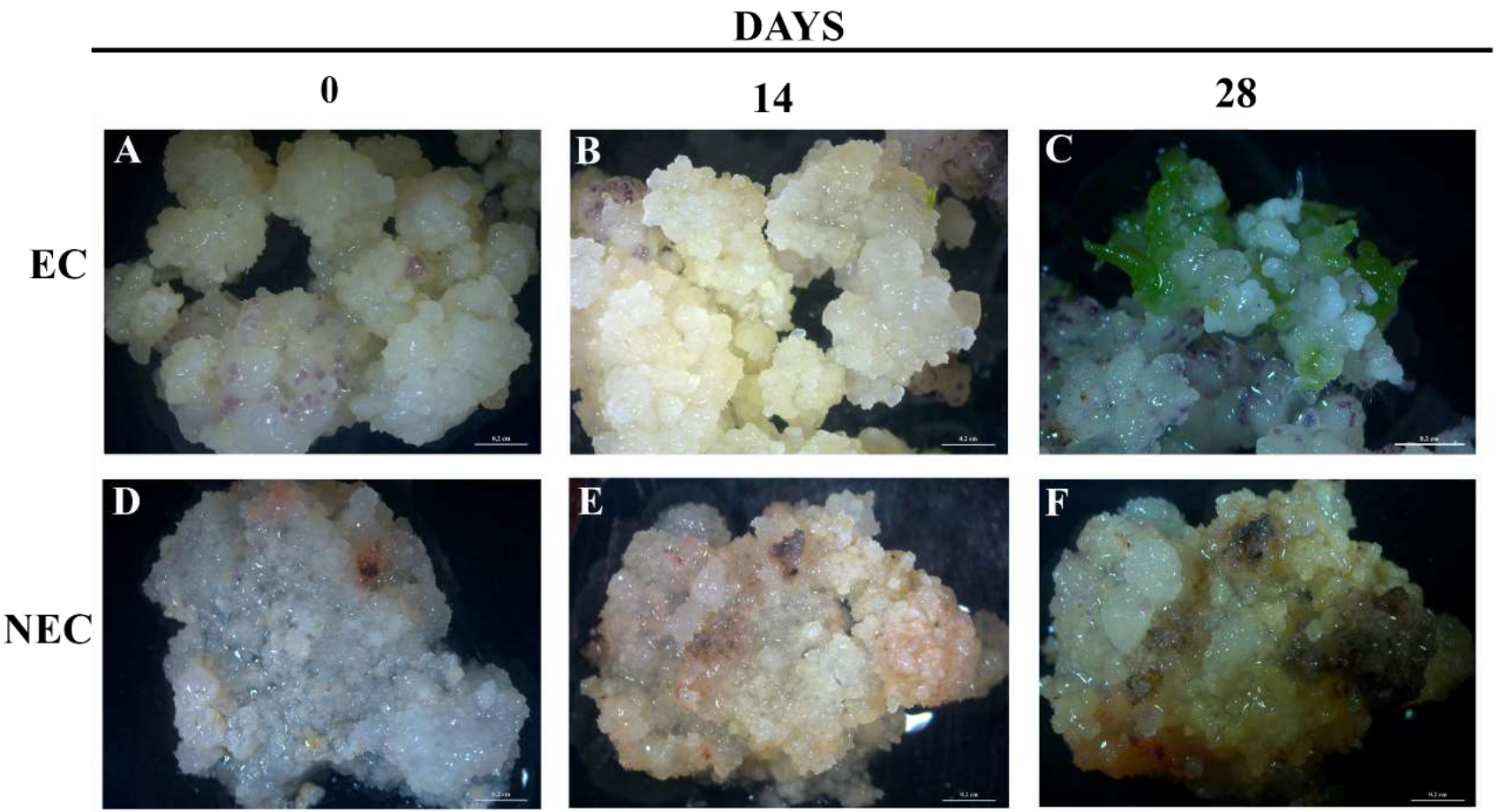

Figura 1 - Características morfológicas dos calos EC e NEC de cana-de-açúcar cv. SP803280 durante o tempo inicial 0 (A e D), 14 (B e E) e 28 (C e F) dias de maturação. Barras de escala: 1 cm.

5.2. Análises por fosfoproteômica comparativa

As medidas da concentração de proteínas totais revelaram que o EC apresentou, em média, 153,1 mg de proteínas / g de matéria seca (MS), enquanto o NEC apresentou 129,7 mg de proteínas / g de MS (Figura 2A). Após o enriquecimento com fosfopeptídeos, o EC teve uma média de 8,3 µg de fosfopeptídeos em 500 µg de peptídeos, e no NEC, esse valor foi de 7,1 µg (Figura 2A).

Para explorar a relação da fosforilação de proteínas com a aquisição da competência embriogênica, foi realizada uma análise LC-MS/MS dos fosfopeptídeos enriquecidos comparando-se EC e NEC de cana-de-açúcar na etapa de multiplicação. Os dados gerados no presente estudo permitiram a identificação de diversas fosfoproteínas diferencialmente reguladas (DRPs), bem como a determinação das posições dos resíduos fosforilados nas sequências peptídicas (Tabela 1).

Tabela 1. Fosfoproteínas diferencialmente reguladas (DRPs) relacionadas ao EC e NEC e seus respectivos resíduos fosforilados que estão relacionados à aquisição da competência embriogênica em cana-de-açúcar. EC: calos embriogênicos; NEC: calos não embriogênicos. n = 5

ID da proteína	Descrição	Abreviação	Fosfo sítios >0.75	Aminoá cido	Janela de sequência	Posição do fosfosítio	Regulação diferencial
RESPOSTAS HORMONAIS E DE ESTRESSE							
Sspon.01G0008640-2B	Late embryogenesis abundant protein 1	LEA1	1	S	ETAEATKNKLGEYKDSAVEKARETKDTVAQK	139	Exclusiva EC
Sspon.01G0003390-2B	late embryogenesis abundant protein D-34	LEA34	1	T	_MSQGQPRRPSGHEETSGEQGAVRYGDVFPA	15	Exclusiva EC
Sspon.05G0004360-3C	Late embryogenesis abundant protein 18	LEA18	1	S	AVDPAYPSAGSTYPASGKYI_____	96	Exclusiva EC
Sspon.01G0021020-1A	low-temperature-induced 65 kDa protein	RD29B	1	S	KDKAAGTVGGAAGGASYTDRLKNAAAGTTEY	248	Exclusiva EC
			2	T	AGKVQQATQSSGTATTPGAGAQQDTATPGAG	456; 467	
Sspon.01G0021020-1P	low-temperature-induced 65 kDa protein	RD29B	3	S	GDERMGDADSGSGSGSSSEEAEEDAVEREAAL	85; 87; 88	Exclusiva EC
Sspon.01G0021020-3C	low-temperature-induced 65 kDa protein	RD29B	1	T	KRDDEQRTEAVTASNTPGTEEWRDAPEATEA	377	Exclusiva EC
			1	S	KDKAADTVGGAAGGASYTDRLKNAAAGTTEY	324	
Sspon.01G0027420-2B	protein ILITYHIA	ILA	1	S	FKVAGTSGKAILEGGSDDEGASTEAQGRAII	603	DOWN
Sspon.01G0030520-2D	serine/threonine protein phosphatase PP2A-2 catalytic subunit	PP2Ac-2	1	S	_____MSSPHGGLDDQIERLMQC	3	DOWN
Sspon.01G0030850-2B	remorin 4.1	REM4.1	1	S	TPPARTRSWDTASHRSFSSEEQFMTMSREFT	472	Exclusiva EC
Sspon.01G0036970-1B	bZIP transcription factor 16	BZIP16	1	S	APTSAPSSNSRDIVLSDPAIQDERELKRQKR	315	Exclusiva EC
			2	S	SNKNKRKTLLKRSKGSLGSLDVVAVKNNKSP	173; 176	
Sspon.01G0039160-2C	serine/threonine protein kinase OSK3	OSK3 (SnRK1)	1	T	DFGLSNVMHDGHFLKTSCGSPNYAAPEVISG	173	Exclusiva EC
			1	S	TVAYYLLLDNRFRATSGYLGADYQESMDRNL	337	
Sspon.02G0000180-1A	Na+/H+ antiporter family protein 2	NHX2	2	S	RPMFGGRGFVPFSPGSPTEQSVHDGR_____	426; 431	Exclusiva EC
			2	S	RLLLPASSNTVPSEPSSPKFLHSPLLTSMQG	353; 356	
Sspon.02G0018300-1P	probable LRR receptor-like serine/threonine protein kinase	GHR1	2	S	ATGGFSPSKGSRFSWSPDSGEAYGQEGLARL	322; 324	Exclusiva EC
			1	T	GKDNKGGLVSADELVTPRKGSTSEALSQEEK	292	
Sspon.03G0009180-1A	serine/threonine protein kinase EDR1	EDR1	2	S	DSDSRNRTGSTQKAVSLPSSPHEYRGQVAPK	480; 484	DOWN
			1	S	GTSTSNARRIRRRSISITPEIGDDIVRAVRA	404	
			1	S	SQRRGVAEEPRGASSSPEHPLMRARGRSILG	351	
			1	S	SDKVEAEGIESSSNLSGRSLRNMMLRSRTFS	300	
Sspon.03G0011050-3C	calcium-dependent protein kinase 1	CDPK1	1	S	GRINYQEFVAMMRNNSPEIVPNRRRMF____	519	DOWN
			1	S	NKLKKVALKVVAENLSDEEIMGLKEMFRSLD	377	
Sspon.03G0011340-1A	UBP1-associated protein 2C	UB2C	1	S	LPGSGFRGQDPQGGMSPPGPGARAPPMYPNV	425	UP
Sspon.03G0026510-2D	villin-2	VLN2	2	S	VDRVVITPAGPSGPSSPQSEAGESNVFHQEK	719; 723	UP
Sspon.03G0028760-2D	Transcription factor bHLH128	BH128	1	S	PLAAAPAASPLFRHSSSPAGLLSRLMADPHG	118	Exclusiva EC
Sspon.03G0029320-2C	transcription initiation factor TFIID subunit 12	TAF12	1	S	AANQKNQTPKPPAPASP_____	403	Exclusiva EC
Sspon.03G0046550-1D	NEDD8-activating enzyme E1 catalytic subunit	ECR1	1	S	_____MASPDAEQPAPTEPERWR	3	DOWN
Sspon.04G0008810-4D	protein phosphatase inhibitor 2-like	IPP2	1	S	KTPYHPMIDEDEGPVSPLRLSEDSVDQSAHA	52	Exclusiva EC
Sspon.04G0030220-3D	trihelix transcription factor GTL1-like	GTL1	2	T; S	HSSSSSKRTGKEKVATPESPAPAAAANYFKK	284; 287	Exclusiva EC
			1	S	IGKGRAGDDQDDEVDSHGHDDE_____	736	
Sspon.04G0030340-1P	Heat shock 70 kDa protein 8	HSP70-8	1	S	_____MAEQFYTVASDSETTGEDKSQPSFP	10	UP
Sspon.05G0001710-1A	protein IQ-DOMAIN 32	IQD32	1	S	IIEKSIELPTQKITESPTDEPAEKINDAPTE	160	UP

			3	S	AFAAAQLKFEELSTNSIVSRSNSSSHLDGVS	553; 556; 560	
			1	S	KSHASKRSLASPGSESVGRSSTDNFSKESRH	768	
			1	S	SDSSALEAPASVPDESPKMEMRHDPESELIE	445	
			1	S	KILVCAGSGSDPAAGSDADADDHPDENKAIS	25	
Sspon.05G0006110-1A	Trihelix transcription factor GTL1	GTL1	1	S	PPKPPSRQQQPPPPPSPQATPQSKPISAAPL	430	Exclusiva EC
Sspon.05G0008220-1A	Inactive LRR receptor-like serine/threonine protein kinase BIR2	BIR2	1	S	TGKGKGGRRRHRRGASESGGGEDGSWWTERL	273	Exclusiva EC
Sspon.05G0013010-4D	SNF1-related protein kinase regulatory subunit gamma-1	KING1	2	S	_____MESPRSPEAEIGHRVEDLWEV	3; 6	DOWN
			1	S	SGAGSPVSNLVSRLGSFTFRRTSSGRVETAT	135	
			1	S	GHRVEDLWEVAEPQLSPSEKLNSCFEDIPVA	27	
Sspon.05G0018880-1P	transcription factor VIP1	VIP1	2	S	GSGSASGGPGHKRSGSMDGATSEGESALSGG	138; 144	Exclusiva EC
Sspon.08G0022270-2P	ABA responsive element binding factor 1	ABF1	1	S	AADHAAWGSSIQRQGSLTLPRTLSQKTVDEV	96	Exclusiva EC
			1	S	DEFQSSLGGAAKDFGSMNMDELLRSIWSAEE	54	
Sspon.08G0025580-2D	auxin response factor 17	ARF17	1	S	RLSSSSGSVLPAQSGSPEAVEEHKCLNSELW	17	Exclusiva EC
REGULAÇÃO DA EXPRESSÃO GÊNICA							
Sspon.01G0006700-3D	Protein TOPLESS	TPL	1	T	FGTPTNPAALLKHPRTPTTANPGMDYPSGDS	233	UP
			1	S	FVDHSCGQPNGARAPSPANNPLLGSMPKPGG	161	
Sspon.01G0039360-2P	DNA helicase INO80	INO80	1	S	SHGLHRKRKRHLDGASDDDEAEAYSNKITEE	89	Exclusiva EC
Sspon.01G0041910-2D	Chromo domain-containing protein LHP1	LHP1	1	S	DSTEGETGDKNKGEDSGNQIHMPKIIKIIKP	153	Exclusiva EC
			1	S	IRDRTSDKPGNETVDSTEGETGDKNKGEDSG	139	
Sspon.02G0026030-1A	Eukaryotic translation initiation factor 5A-2	ELF5A-2	1	S	_____MSDEEHHFESKADAGAS	2	UP
Sspon.02G0029150-2B	Serine/arginine-rich SC35-like splicing factor SCL30	SCL30	3	S	RRYSPPYRSPPRRGYGGRGRSPPPPPRRGYG	5; 10; 22	Exclusiva EC
Sspon.02G0041640-2D	chromatin remodeling protein EBS	EBS	2	S	FCQTCTAENGKMVENSHEATAQSEEKPVESK	199; 206	UP
Sspon.03G0001680-1P	PHD finger protein ALFIN-LIKE 5	ALFL5	1	S	KVTKVAAPPKDDDDESGEEYEEEEERDNTLC	223	Exclusiva EC
Sspon.03G0026060-1P	Serine/arginine-rich splicing factor SR45	SR45	2	S	PPPKRDAPQIEKGVSSAEKDAQQRPRESSPR	121; 122	Exclusiva EC
			3	S	SPRRPPDPSPRRRPDSPPIRRRPDPSPVRRG	163; 170; 180	
Sspon.04G0005920-1A	Protein photoperiod-independent early flowering 1	PIE1	1	S	VILKMYTKTKVSRESSPDSNDMLSDLGSKNL	277	Exclusiva EC
Sspon.04G0006560-1A	probable N6-adenosine-methyltransferase MT-A70-like	MTA	1	S	ARYKAAYPDVEVSPPSPPRTSTPMDVDQSSS	409	Exclusiva EC
Sspon.04G0011310-1A	TUDOR-SN protein 1	TSN1	1	S	AKKERLRIWQYGDVESDEEEQAPAARKPGGR	873	UP
Sspon.04G0015140-2P	putative histone deacetylase 19	HDA19	1	S	_____MDPSSAGSGGNSLPSVGPDGQKR	8	Exclusiva EC
			1	S	PDEDQDDPDERHDPDSDMEVDDHKAVEESAR	421	
Sspon.04G0033150-1C	probable pre-mRNA-splicing factor ATP-dependent RNA helicase DEAH5	DEAH5	1	S	EPAFLQGQSRFSIDMSPVKIFKNPEGSLSRA	457	UP
Sspon.05G0005190-1A	argonaute1b	AGO1	1	S	HLAAFRARFYMEPDTSDSGSMASGARGPPPG	1037	Exclusiva EC
Sspon.06G0006110-2B	HMG-Y-related protein A	HMGYA	1	S	PPKSRDPNAPAPPPKSPASSAGTGRGRGRPP	108	Exclusiva EC
Sspon.06G0007620-2T	histone deacetylase 6	HDA6	1	S	RPPQRSRLWSGGAYDSDTEDPDNLKSESKDV	429	UP
			1	S	___MAASGEGASLPSPAGGEDAHRRRVSYF	12	
Sspon.06G0010180-3C	Protein PAF1-like protein	PAF1	1	S	DEHPKRSRVEDIDQYSGEEYSE_____	106	UP
Sspon.06G0011170-3C	DNA topoisomerase 1 beta	TOP1	1	S	KKLKRVDTGVQKADDSDDDDKPLSLKINSSK	169	UP
			2	S	DEDSDDDKPLASRLPNNAGPKSGGDVSEDSED	210; 236	
			2	S	VDKSKLKRPLVKDERSDDSDDEVPIGLRRKA	136; 139	

Sspon.06G0013930-3D	Zinc finger CCCH domain-containing protein 33	ZFN1	1	S	TSRRMMAHVPSHPEVSPDSGSGRSRRIAHSD	379	Exclusiva EC
			1	S	QHAYQGAVTSWPLSRSASFIASPRWPGHSSY	178	
Sspon.07G0006150-1P	Putative SAC3/GANP family protein	SAC3	1	S	HDLTKYYASATALANSPEEKKRREHRSKRFE	631	Exclusiva EC
DESENVOLVIMENTO EMBRIOGÊNICO							
Sspon.01G0015180-4D	ADP-ribosylation factor GTPase-activating protein AGD10	AGD10	1	S	IEESSEARQKFSNAKSISSSQFFGDQASFEK	226	Exclusiva EC
			2	S	VAKSSTEDGNTSWPSSPVAASQAPNQDAAFP	160; 164	
Sspon.01G0041320-2C	dnaJ homolog subfamily C GRV2	GRV2	1	S	GLVAYLHTRSDEDSQSQYDEAPLSRRQRRIL	762	Exclusiva EC
Sspon.02G0011740-4D	serine/threonine protein kinase UCNL	UCNL	1	S	RSQYHHVKNIFKRSESAVTASTSGQEEEPRN	215	Exclusiva NEC
			1	S	KSARVSPMDRGKKLSSFCSAAAGERSFSFVG	265	
Sspon.03G0031400-3D	Trihelix transcription factor ASIL1	ASIL1	1	S	LEEDGDEQNNSAMDASP_____	399	Exclusiva EC
Sspon.05G0002130-2D	MAP3K epsilon protein kinase 1	M3KE1	1	S	PLLSLLEKEPPSRHVSGQLDYVRHISGLERH	859	Exclusiva EC
			1	S	SGASVTQNGSTPRRRSGQLDPSVLESCKTRL	744	
Sspon.05G0007230-2B	caleosin 1	CLO1	1	T	ARALAAPDIYHPDGTTDDEHRHHHMSVLQQH	106	Exclusiva EC
Sspon.06G0000330-3C	14-3-3-like protein GF14-A	GF14	1	S	RDNLTLWTSDMQEDGSDEMRDASKPDDE___	247	DOWN
Sspon.06G0006090-3C	14-3-3-like protein GF14-C	GF14	1	T	MQLLRDNLTLWTSDLTEDGADEGKEASKGDA	256	DOWN
Sspon.06G0014060-2C	DNA replication licensing factor MCM2	MCM2	1	S	__MDDSENNAPSTPGSPGFSTDRLPPNTTSR	14	Exclusiva EC
Sspon.07G0005070-2B	transcription elongation factor SPT6-like isoform X1	SPT6L	1	S	_____MRRTVLSDEEEDEIEADEEDPR	7	Exclusiva EC
METABOLISMO DE CARBOÍDRATOS, LIPÍDEOS E NITROGÊNIO							
Sspon.01G0050080-2C	putative clathrin assembly protein	CAP2	3	S	RDRDRWGSPDPYGRRSPSYSSPPGYGGYDDY	182; 190; 195	DOWN
			1	S	ETLEEFMRDRAKRPKSPPREPEPEPVKEEPE	358	
Sspon.02G0018020-3D	nitrate regulatory gene2 protein	NRG2	1	S	GDGGGGPAAPLHRSMSAPDIQPIRKGRSGEA	127	UP
Sspon.03G0006890-1P	ATP-dependent 6-phosphofructokinase 6	PFK6	2	S	ISRDSYPNLRALRNASSVSLADAAYVKISEG	65; 76	UP
Sspon.03G0022300-3C	probable choline kinase 2 isoform X1	CK2	1	S	PPPRGEAAPVLKRSASIDRIPEDARRILHRL	21	Exclusiva EC
Sspon.03G0028140-1B	sucrose-phosphate synthase	SPS3	1	S	LAPVETAKKKFQRNFSDLTVWSDDNKEKKLY	168	Exclusiva EC
			1	S	EQVRREATEDLAEDLSEGEKGDTLGELAPVE	142	
			2	S	GGGGGGDPRSPVAGASPTKAASPRGPHMNFN	35; 41	
Sspon.04G0003840-4P	Dynamin-2A	DRP2A	2	S; T	RASVSSYSNDTTEAESPRTPSRSGEDWRSAF	742; 745	DOWN
			1	S	YQKQSSLLSKLTRQLSIHDNRASVSSYSNDT	722	
			1	S	FKGPSTEGGSMRQSNSDGALDTMARRPADPE	636	
			1	S	KKTQEAEQSTSKRASSPQTDAEQGGGSLKSM	454	
Sspon.05G0008360-3C	glucose-6-phosphate 1-dehydrogenase, cytoplasmic isoform	G6PD6	1	S	___MSGGSTPSSRRNSFNSLSRDLDLPSEQG	13	DOWN
Sspon.06G0004940-2B	nitrate reductase [NADH]	NIA1	1	S	LETAEAAAPGLKRSTSTPFMNTTGGKQFTMS	534	Exclusiva NEC
Sspon.07G0001380-1T	sugar transporter ERD6-like 4	ERD6	2	S	GMMGSRQSSLMERLGSSAFSLRDVAISATFC	44; 55	Exclusiva EC
			1	S	_____MSFRDQESGGEDAGRTS	2	
Sspon.07G0002100-3C	Putative sucrose-phosphate synthase family protein isoform 1	SPS1	2	S	RWQKNDVATEISEADSPEDSLRDIHDISLNL	659; 663	UP
Sspon.07G0011260-1A	diacylglycerol O-acyltransferase 1-1	TAG1	1	S	SSGAAAAARASFAADSGDESGPGEPSSSRRR	53	Exclusiva EC
Sspon.07G0012190-4D	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase	F2KP	2	S	FVDRGVGSPMLPKSASACSLASGFSFGSAKT	309; 318	Exclusiva EC
			1	S	WATDSSKNSGLIESKSVGTFTPLQKLDGQKG	277	

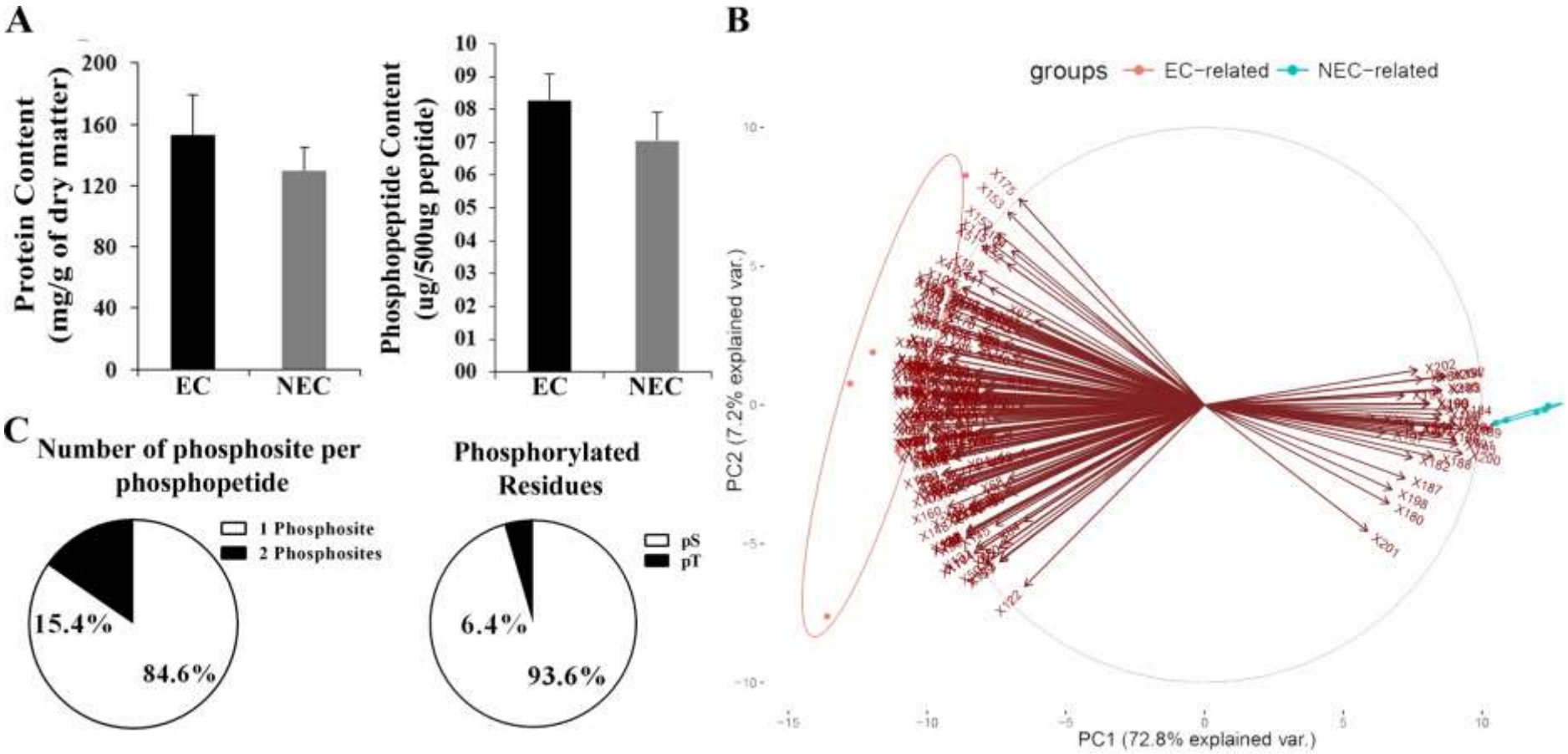

Figura 2 – Análise de fosfoproteômica comparativa dos calos embriogênicos (EC) e não embriogênicos (NEC) de cana-de-açúcar. Medidas da concentração de proteínas totais e da concentração de fosfopeptídeos enriquecidos (A). Análise dos componentes principais (PCA) entre as fosfoproteínas relacionadas com EC e NEC (B). Gráficos de pizza mostram o número de sítios de fosforilação por fosfopeptídeo, bem como os resíduos fosforilados mais representativos (C). pS: serina fosforilada; pT: treonina fosforilada.

A análise fosfoproteômica comparativa entre EC e NEC na etapa de multiplicação quantificou um total de 1.279 fosfoproteínas. A lista completa dessas fosfoproteínas identificadas e suas classificações funcionais estão disponibilizadas *online* no arquivo 2020_03_11_Supplementary_Table_S1_final.xlsx, depositado no repositório PRIDE (Perez-Riverol et al., 2019) com o identificador PXD018054 (clique aqui). Considerando os parâmetros de corte e a análise dos DRPs na comparação EC/NEC, foram separadas as fosfoproteínas em dois grupos: o primeiro grupo consiste nas fosfoproteínas exclusivas em EC ou mais abundantes no EC em comparação com o NEC; o segundo grupo incluiu as fosfoproteínas exclusivas em NEC ou aquelas menos abundantes no EC em comparação com o NEC. Daqui em diante, esses grupos serão referidos como fosfoproteínas relacionadas com EC e relacionadas com NEC, respectivamente.

No grupo relacionado ao EC, 163 fosfoproteínas foram exclusivas e 51 mais abundantes. Já no grupo relacionado ao NEC, nove fosfoproteínas foram exclusivas e 40 mais abundantes. A análise por PCA separou as fosfoproteínas relacionadas com EC e NEC em dois grupos distintos. O primeiro componente principal (PC1: 72,8%) e o segundo componente principal (PC2: 7,2%) separaram os grupos experimentais, e a maioria das fosfoproteínas foram separadas por PC1, o que distinguiu as amostras EC e NEC (Figura 2B). Além disso, a fosfoproteína com maior contribuição para a variabilidade do PCA foi a proteína UBP1-associated protein 2C (UB2C), detectada como sendo mais abundante em EC, e que foi anotada no processo biológico de regulação dos processos metabólicos do RNA.

5.3. Características das fosfoproteínas, seus fosfosítios e o enriquecimento dos *motifs*

Entre os grupos de fosfoproteínas relacionadas com EC e NEC, 84,6% dos fosfopeptídeos continham apenas um fosfosítio (Figura 2C); os resíduos fosforilados de serina (S) representaram 93,6%, os resíduos de treonina (T) representaram 6,4%, enquanto resíduos fosforilados de tirosina (Y) não foram identificados (Figura 2C). Através da análise por *motif-x,* considerando S e T como resíduos de fosfopeptídeos centrais, foram identificados três *motifs*

significativamente enriquecidos para os resíduos S, mas nenhum para os resíduos T (Figura 3). Estes *motifs* enriquecidos foram o *motif* dirigido por prolina [xxxpSPxxx] (Figura 3A), o *motif* básico [RxxpSxxx] (Figura 3B) e o *motif* ácido [xxxpSDxxx] (Figura 3C).

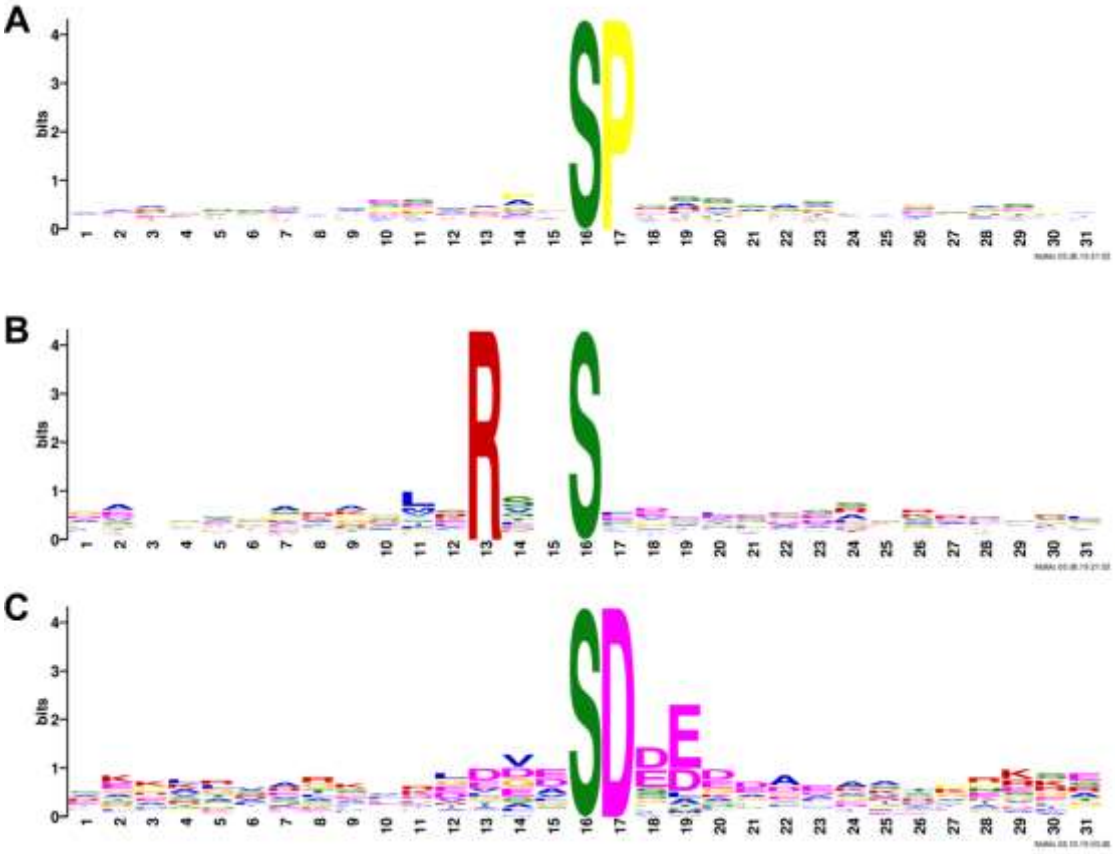

Figura 3 - Análise de enriquecimento por *Motif*-x de ambas as fosfoproteínas relacionadas ao EC e NEC com a ferramenta MOMO retornou três *motifs* mais representativos: o *motif* dirigido por prolina [xxxpSPxxx] (A), o *motif* básico [RxxpSxxx] (B), e o *motif* ácido [xxxpSDxxx] (C).

A família de quinases preditas que podem fosforilar o *motif* dirigido por prolina (Figura 3A) foram as Mitogen-activated protein kinases (MAPKs), Cyclin-dependent kinase 5 (CDK5), Glycogen synthase kinase-3 (GSK3), Extracellular signal-regulated kinase 1/2 (ERK1 e ERK2), Receptor-like protein kinases (RLKs), AGC kinase family protein kinases C/A/G (PKAs/C/G), Calcium-dependent protein kinase 5 (CDPK5), STE20-like kinase (SLKs), e SNF1-related protein kinase 2 (SnRK2). Além disso, as fosfoproteínas identificadas em EC que contêm esse *motif*, estão principalmente relacionadas ao controle da expressão

gênica por meio de mecanismos pós-transcricionais e epigenéticos, incluindo a pre-mRNA-splicing factor ATP-dependent RNA helicase DEAH5 (DEAH5), UB2C, Serine/arginine-rich SC35-like splicing factor SCL30 (SCL30), HMG-Y-related protein A (HMGYA), N6-adenosine-methyltransferase MT-A70 (MTA) e a proteína Photoperiod-independent early flowering 1 (PIE). As fosfoproteínas relacionadas ao EC e NEC que foram envolvidas com respostas hormonais e de estresses, incluindo a Auxin response factor 17 (ARF17), LRR receptor-like serine/threonine protein kinase (GHR1) e SNF1-related protein kinase regulatory subunit gamma-1 (KING1), também possuem o *motif* [SP].

Para o *motif* básico (Figura 3B), as quinases preditas foram a Calmodulin/calcium-dependent protein kinase II (CAMK-II), PKA, PKG e CDPK-SNRK. A família de quinases predita para fosforilar o *motif* ácido é a Casein kinase II (CKII). No presente estudo, o *motif* [RxxS] foi identificado como fosforilado nas fosfoproteínas relacionadas ao EC que estão envolvidas na sinalização do ABA, respostas ao estresse e desenvolvimento embriogênico, como a Serine/threonine protein kinase OSK3 (uma subunidade catalítica da SnRK1), Transcription factor VIP1 (VIP1) e MAP3K epsilon protein kinase 1 (M3KE1).

A localização subcelular das fosfoproteínas relacionadas com EC foi principalmente o núcleo, seguido pelo citosol e membrana plasmática (Figura 4). As fosfoproteínas relacionadas com NEC foram localizadas principalmente no citosol, seguidas pelo núcleo e membrana plasmática (Figura 4). De acordo com a análise *Clusters of Orthologous Groups* (COG) realizada com o eggNOG, as fosfoproteínas relacionadas ao EC estão envolvidas na transcrição, processamento e modificação de RNA, mecanismos de transdução de sinais e modificações pós-traducionais, *turnover* de proteínas e chaperonas (Figura 5A). Para as fosfoproteínas relacionadas ao NEC, as principais anotações foram mecanismos de transdução de sinais, tradução, estrutura ribossômica e biogênese, bem como modificação pós-traducional, *turnover* de proteínas e chaperonas (Figura 5B).

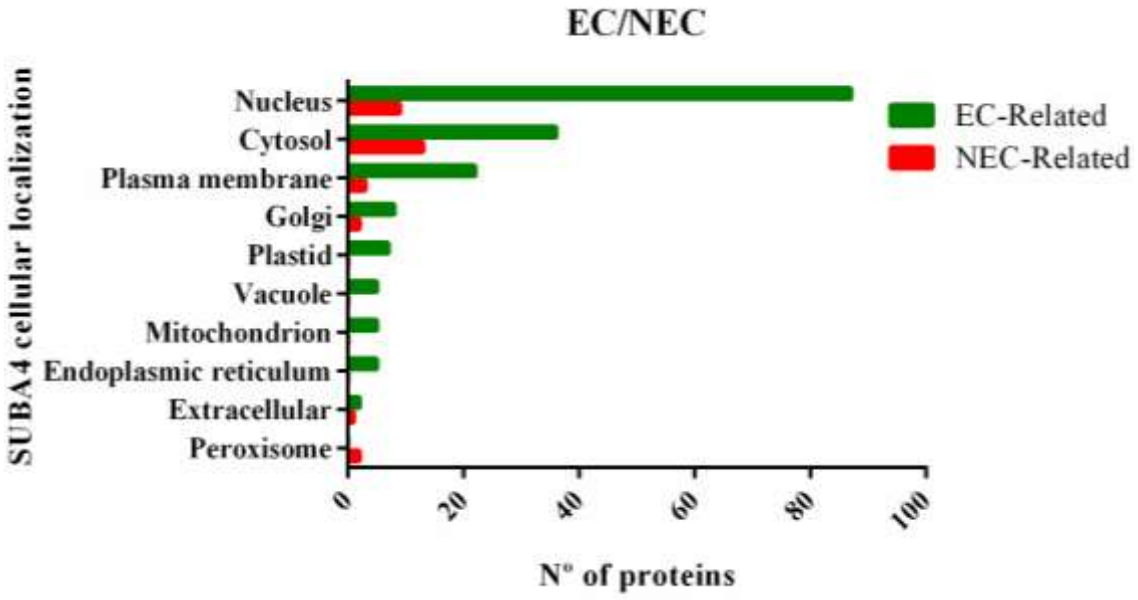

Figura 4 - Localização subcelular determinada pela ferramenta SUBA4 para as fosfoproteínas relacionadas ao EC e NEC.

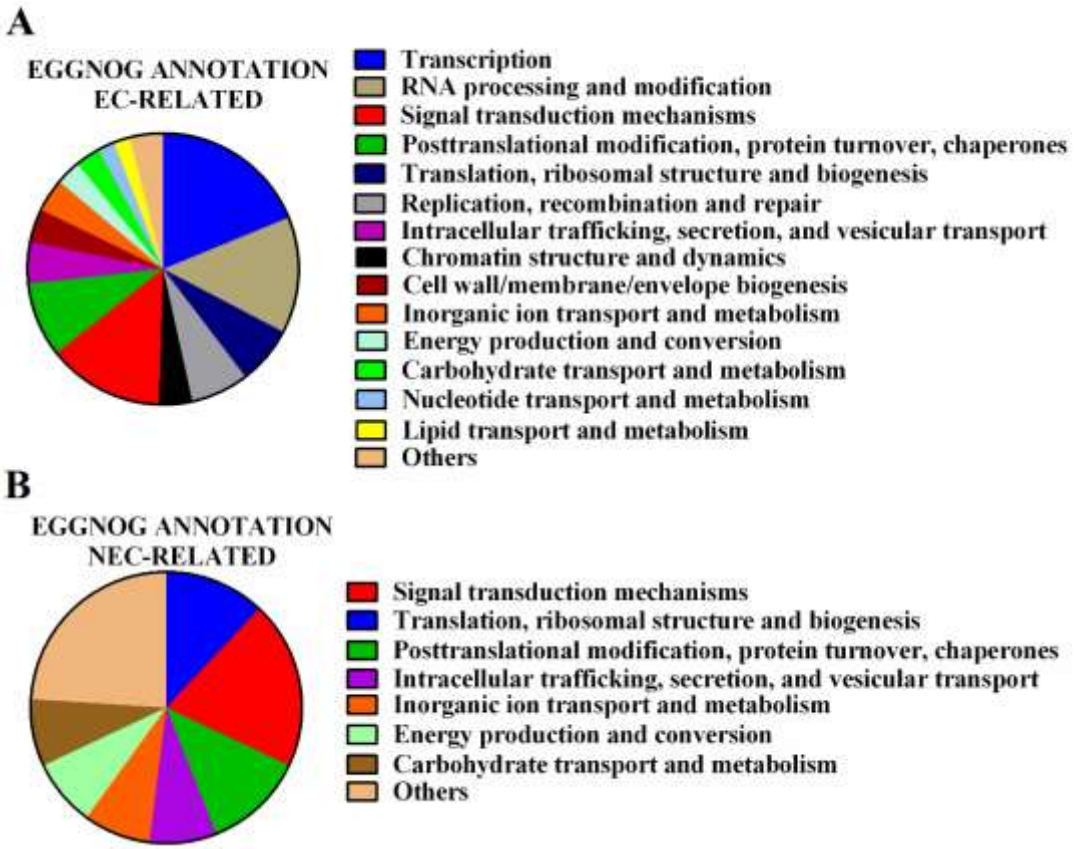

Figura 5 - Anotações funcionais do eggNOG referentes às fosfoproteínas relacionadas ao EC (A) e relacionadas ao NEC (B).

A rede de interação proteína-proteína das fosfoproteínas relacionadas ao EC, apresentou os processos biológicos enriquecidos de forma estatisticamente significativa, para os processos relacionados ao desenvolvimento pós-embrionário, respostas ao estresse e regulação da expressão gênica e epigenética (Figura 6A). No grupo relacionado ao NEC, as fosfoproteínas foram enriquecidas para os processos biossintéticos ligados a tradução (Figura 6B).

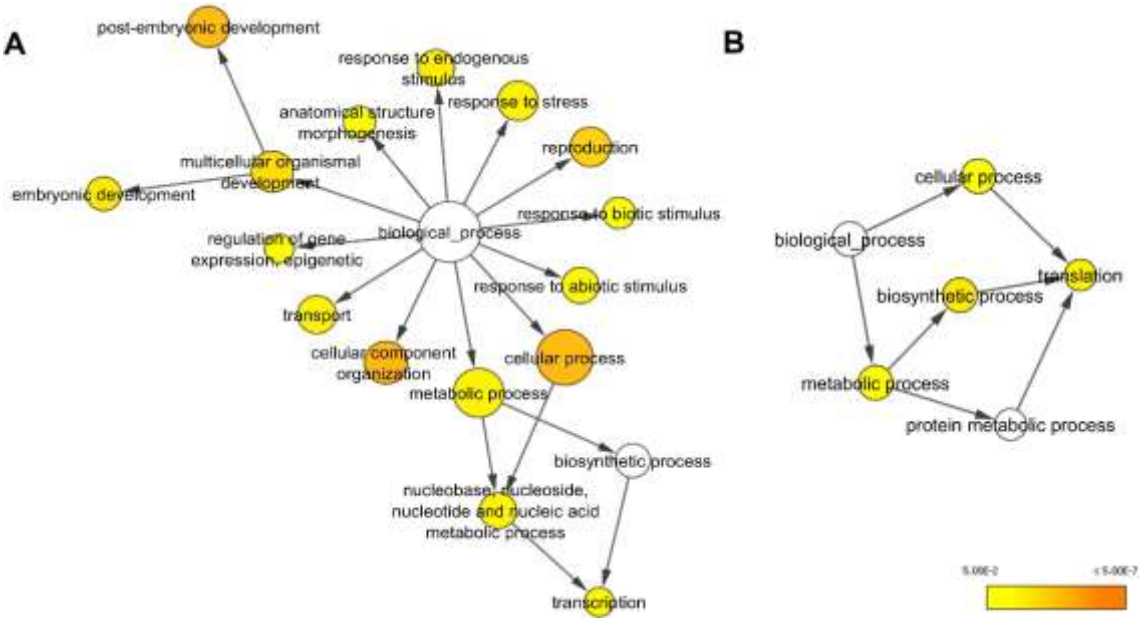

Figura 6 – As redes de interação proteína-proteína determinadas pelo STRING contrastando as sequências de fosfoproteínas diferencialmente reguladas (DRP) de cana-de-açúcar com o banco de dados de *A. thaliana*. Os resultados foram usados como entrada para a análise de enriquecimento de *Gene Ontology* com a ferramenta BiNGO. Os processos biológicos estatisticamente significativos enriquecidos para as fosfoproteínas relacionadas com EC (A) e relacionadas com NEC são mostrados (B). O tamanho do círculo para cada categoria de processo biológico aumenta com o número de proteínas anotadas para o processo dado, e a cor muda de amarelo para laranja com o nível de enriquecimento.

A Tabela 1 resume as fosfoproteínas relacionadas ao EC e NEC candidatas a desempenhar papeis essenciais nos processos biológicos envolvidos com a aquisição da competência embriogênica, como respostas a estresses, regulação da expressão gênica e desenvolvimento embriogênico. Com base nos resultados fosfoproteômicos obtidos e nas demais análises realizadas, um esquema gráfico foi construído para ilustrar as interações preditas

a nível celular e destacar as fosfoproteínas com potenciais papeis na embriogênese somática de cana-de-açúcar (Figura 7). Essas fosfoproteínas são relacionadas com respostas induzidas por ABA mediando a aquisição da tolerância ao estresse; com o controle do metabolismo de carboidratos, lipídios e nitrogênio; e com a regulação da expressão gênica por meio da remodelação da cromatina e controle da transcrição (Tabela 1, Figura 7).

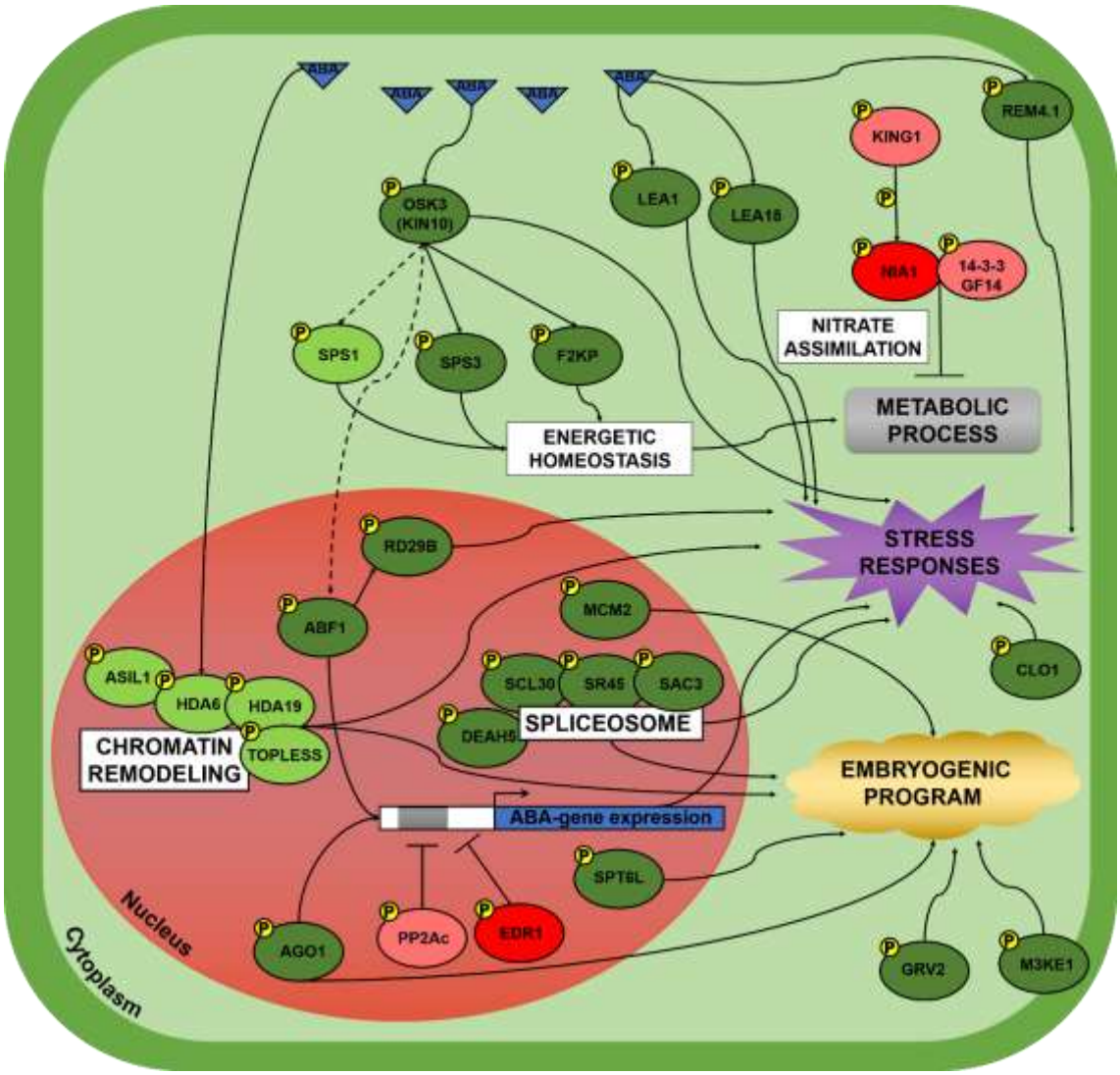

Figura 7 – Visão geral das fosfoproteínas diferencialmente reguladas comparando os EC com NEC. Essas fosfoproteínas estão principalmente relacionadas as sinalizações da via do ABA mediando as respostas ao estresse ligadas à aquisição da competência embriogênica. A cor verde escura refere-se as fosfoproteínas exclusivas em EC, e vermelho escuro as exclusivas em NEC. Verde claro representa aquelas fosfoproteínas mais abundantes e vermelho claro as menos abundantes em EC. As linhas tracejadas caracterizam as possíveis interações.

5.4. Alinhamento das fosfoproteínas relacionadas à competência embriogênica e conservação dos fosfosítios

Para verificar a homologia entre as fosfoproteínas de cana-de-açúcar e os ortólogos em monocotiledônea (*Z. mays*) e dicotiledônea (*A. thaliana*), bem como o estado de conservação dos fosfosítios, foi feito o alinhamento ClustalW das 72 fosfoproteínas listadas na Tabela 1 e seus respectivos ortólogos. Essas sequências compreendem 139 fosfosítios (Tabela 2; Figuras 8, 9 e 10). Entre as fosfoproteínas relacionadas com EC, em média, 76% compartilharam identidade com seus respectivos ortólogos em milho (*Z. mays*), enquanto 43% compartilharam identidade com sequências de *A. thaliana* (Tabela 2). Para as fosfoproteínas relacionadas ao NEC, 88% compartilharam identidade com seus respectivos ortólogos em milho, enquanto 61% compartilharam identidade com seus ortólogos em *Arabidopsis* (Tabela 2).

Tabela 2. Análise de homologia entre as fosfoproteínas em cana-de-açúcar e suas ortólogas em monocotiledôneas (*Z. mays*) e dicotiledôneas (*A. thaliana*).

Acesso	Abreviação das proteínas	Homologia *Zea mays*	Homologia *Arabidopsis*	Número de fosfosítios	Número de fosfosítios conservados para milho	Número de fosfosítios conservados para *Arabidopsis*
FOSFOPROTEÍNAS RELACIONADAS AO EC						
Sspon.08G0022270-2P	ABF1	88%	40%	1	1	1
Sspon.03G0028760-2D	BH128	61%	23%	1	1	1
Sspon.01G0039160-2C	OSK3	96%	73%	2	2	2
Sspon.01G0030850-2B	REM4.1	83%	41%	1	1	1
Sspon.05G0018880-1P	VIP1	86%	42%	2	2	2
Sspon.05G0004360-3C	LEA18	79%	43%	1	1	0
Sspon.03G0011340-1A	UB2C	92%	46%	1	1	0
Sspon.01G0021020-1A	RD29B	79%	17%	1	1	0
Sspon.01G0021020-3C	RD29B	72%	15%	2	2	0
Sspon.03G0026510-2D	VLN2	86%	52%	1	1	0
Sspon.05G0005190-1A	AGO1	94%	71%	1	1	1
Sspon.06G0007620-2T	HDA6	96%	67%	1	1	1
Sspon.04G0015140-2P	HDA19	81%	71%	1	1	0
Sspon.02G0029150-1P	SCL30	49%	44%	1	1	1
Sspon.01G0006700-3D	TPL	88%	69%	1	1	1
Sspon.04G0005920-1A	PIE1	17%	8%	1	1	1
Sspon.01G0041910-2D	LHP1	92%	18%	2	2	2
Sspon.05G0014050-1A	PUM3	59%	25%	1	1	1
Sspon.06G0002200-2B	PUM5	66%	22%	2	2	2
Sspon.03G0001680-1P	ALFL5	84%	59%	1	1	0
Sspon.04G0006560-1A	MTA	54%	40%	1	1	1
Sspon.05G0002130-2D	M3KE1	94%	60%	2	2	2
Sspon.01G0008640-2B	LEA1	89%	27%	1	1	0
Sspon.05G0007230-2B	CLO1	84%	51%	1	0	0
Sspon.06G0013930-3D	ZNF1	69%	39%	1	1	1
Sspon.02G0018020-3D	NRG2	86%	21%	1	1	0
Sspon.05G0001710-1A	IQD32	86%	25%	1	0	0
Sspon.04G0011310-1A	TSN1	77%	57%	1	1	1
Sspon.04G0033150-1C	DEAH5	83%	68%	1	1	1
Sspon.06G0011170-3C	TOP1	44%	26%	1	1	1
Sspon.01G0015180-4D	AGD10	52%	34%	1	1	1

Sspon.01G0036970-1B	BZIP16	83%	42%	1	1	1
Sspon.01G0041320-2C	GRV2	66%	49%	1	1	0
Sspon.02G0018300-1P	GHR1	57%	40%	1	1	1
Sspon.03G0022300-3C	CK2	72%	52%	1	1	1
Sspon.04G0008810-4D	IPP2	86%	29%	1	1	0
Sspon.05G0006110-1A	GTL1	84%	31%	1	1	1
Sspon.06G0014050-2B	MCM2	94%	70%	1	1	1
Sspon.08G0025580-2D	ARF17	80%	55%	1	1	0
Sspon.05G0008220-1A	BIR2	94%	49%	1	1	0
Sspon.07G0001380-1T	ERD6	62%	28%	1	1	0
Sspon.02G0026030-1A	ELF5A-2	90%	83%	1	1	1
Sspon.02G0000180-1A	NHX2	78%	59%	4	4	4
Sspon.07G0002100-3C	SPS1	81%	56%	1	1	1
Sspon.03G0028140-1B	SPS3	96%	66%	2	2	2
Sspon.07G0011260-1A	TAG1	80%	46%	1	1	0
Sspon.06G0010180-3C	PAF1	22%	8%	1	1	1
Sspon.04G0030220-3D	GTL1	86%	21%	2	0	0
Sspon.07G0005070-2B	SPT6L	93%	55%	1	1	1
Sspon.02G0041640-2D	EBS	63%	56%	1	1	0
Sspon.03G0026060-1P	SR45	47%	33%	2	2	0
Sspon.03G0029320-2C	TAF12	72%	31%	1	1	1
Sspon.06G0006110-2B	HMGYA	86%	41%	1	1	1
Sspon.01G0039360-2P	INO80	66%	40%	1	0	0
Sspon.03G0031400-3D	ASIL1	16%	18%	1	0	1
Sspon.03G0006890-1P	PFK6	91%	55%	2	2	0
Sspon.07G0012190-4D	F2KP	80%	55%	3	3	2
Sspon.04G0030340-1P	HSP70-8	94%	63%	2	2	2
Sspon.07G0006150-1P	SAC3	74%	28%	1	1	1
	TOTAL	**76%**	**43%**	**75**	**69**	**48**
					92%	**64%**
FOSFOPROTEÍNAS RELACIONADAS AO NEC						
Sspon.01G0030520-2D	PP2Ac-2	99%	92%	1	1	0
Sspon.05G0013010-4D	KING1	97%	56%	2	2	2
Sspon.03G0046550-1D	ECR1	97%	68%	1	1	0
Sspon.01G0027420-2B	ILA	90%	38%	1	1	0
Sspon.04G0007950-4D	PIP2-5	96%	79%	1	1	0
Sspon.03G0011050-3C	CDPK1	94%	71%	1	1	0
Sspon.06G0000330-3C	GF14	95%	77%	1	0	0
Sspon.05G0008360-3C	G6PD6	71%	55%	1	1	0
Sspon.06G0004940-2B	NIA1	93%	64%	1	1	1
Sspon.03G0038330-2P	IF4E1	94%	61%	1	1	0
Sspon.02G0011740-4D	UCNL	90%	44%	1	1	1
Sspon.04G0003840-4P	DRP2A	73%	62%	2	2	2
Sspon.01G0050080-2C	CAP2	95%	62%	1	1	1
Sspon.03G0009180-1A	EDR1	61%	21%	2	2	2
	TOTAL	**89%**	**61%**	**17**	**16**	**9**
					94%	**53%**

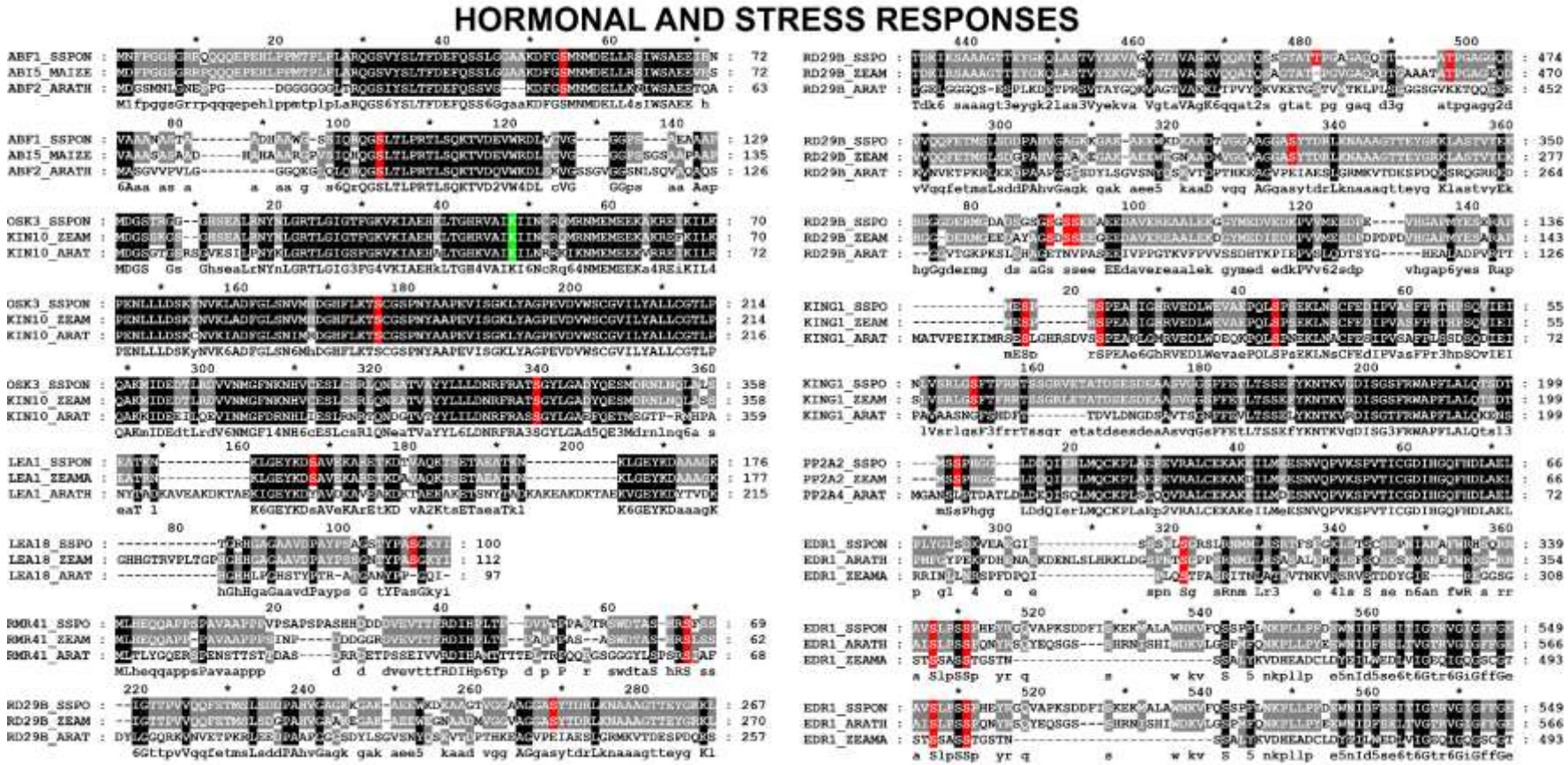

Figura 8 - Alinhamentos das fosfoproteínas relacionadas com as respostas hormonais e de estresse para EC (Calos embriogênicos) e NEC (Calos não embriogênicos) que podem estar envolvidas na aquisição da competência embriogênica. Os resíduos destacados em vermelho referem-se aos fosfosítios identificados neste estudo e seus resíduos correspondentes em *Arabidopsis* e milho. A cor verde refere-se ao resíduo conservado (K48) do sítio de ligação do ATP na SnRK1 OSK3.

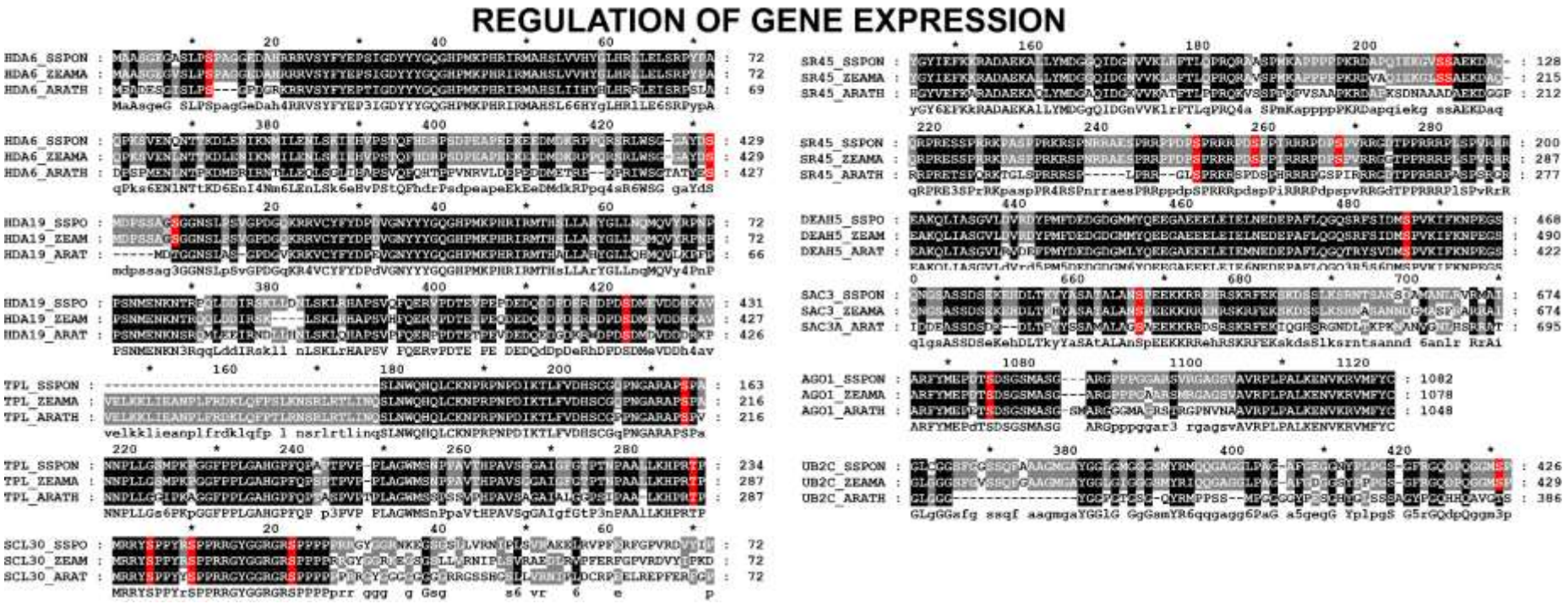

Figura 9 - Alinhamentos das fosfoproteínas relacionadas a regulação da expressão gênica no EC que foram principalmente envolvidas com modificações epigenéticas por meio da remodelação de cromatina e do metabolismo de RNA por silenciamento de RNA e controle de *splicing*. Os resíduos destacados em vermelho referem-se aos fosfosítios identificados neste estudo e seus resíduos correspondentes em *Arabidopsis* e milho.

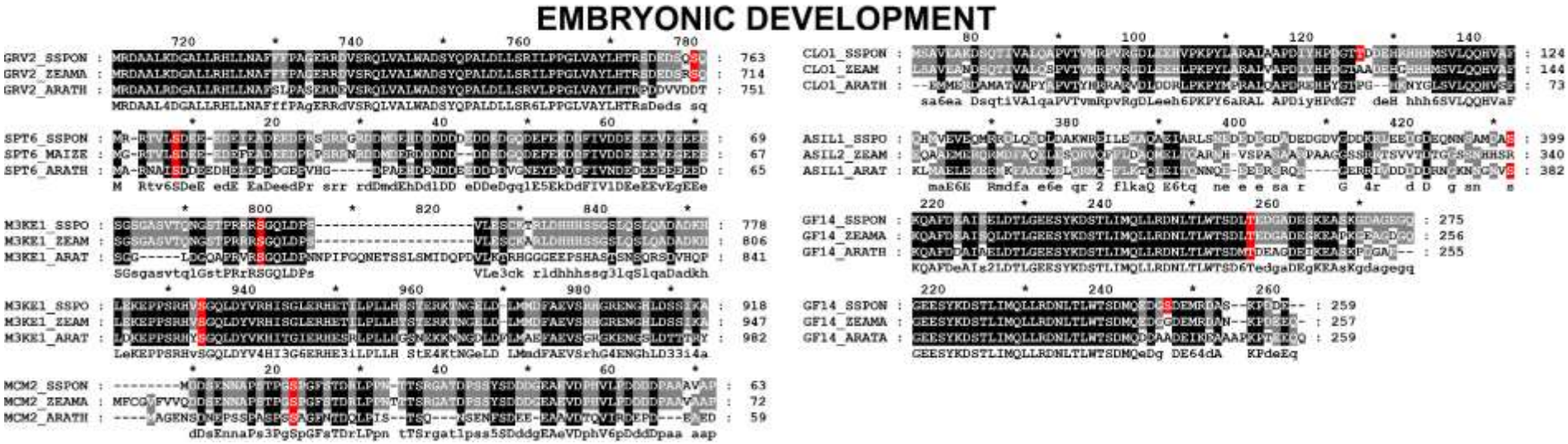

Figura 10 - Alinhamentos das fosfoproteínas no EC que estão envolvidas com a aquisição da competência embriogênica e desenvolvimento embriogênico normal. Os resíduos destacados em vermelho referem-se aos fosfosítios identificados neste estudo e seus resíduos correspondentes em *Arabidopsis* e milho.

Em relação a conservação dos fosfosítios, 95% das fosfoproteínas relacionadas ao EC demostraram ter um resíduo fosforilado na mesma posição da sequência de milho, e 52% dessas fosfoproteínas compartilharam posições com os resíduos das sequências de *Arabidopsis* (Tabela 2). Em média, 93% dos fosfosítios das fosfoproteínas relacionadas ao NEC apresentam as mesmas posições com aquelas em milho e 66% compartilham posições com aquelas em *Arabidopsis* (Tabela 2).

6. DISCUSSÃO

6.1. Análise fosfoproteômica comparativa dos calos embriogênicos (EC) e não embriogênicos em cana-de-açúcar

O processo de embriogênese somática em cana-de-açúcar é bem descrito, e as características morfológicas dos calos EC e NEC observadas no presente estudo (Figura 1) correspondem com aquelas já descritas anteriormente (Silveira et al., 2013; Heringer et al., 2015), demonstrando o claro potencial embriogênico do EC em comparação com o NEC. Para adquirirem a competência embriogênica, as células somáticas requerem certas condições hormonais e de estresse, que desencadeiam uma complexa rede de sinalização percebida e amplificada pela fosforilação de proteínas (Aguilar-Hernández e Loyola-Vargas, 2018).

Neste estudo, a análise fosfoproteômica comparativa da aquisição da competência embriogênica em cana-de-açúcar, identificou um total de 1.279 fosfoproteínas. Do mesmo modo, um estudo fosfoproteômico em calos de arroz identificou 1.310 fosfoproteínas (Wang et al., 2017), indicando que uma complexa via de transdução de sinais envolve múltiplas fosfoproteínas ligadas a diferentes processos biológicos na aquisição da competência embriogênica.

6.2. Caracterização das DRPs nos grupos relacionados ao EC e NEC

A análise SUBA4 das DRPs indicou que as fosfoproteínas relacionadas com EC estão localizadas principalmente no núcleo, citosol e membrana plasmática (Figura 4). Este padrão de distribuição observado nas fosfoproteínas

relacionadas ao EC também foi relatado em tecidos de calo (Wang et al., 2017) e durante os estágios iniciais do desenvolvimento de sementes em arroz (Qiu et al., 2016), sugerindo que a fosforilação de proteínas no núcleo e no citosol é um importante mecanismo de transdução de sinais nos tecidos em proliferação.

A análise de ontologia gênica revelou que as fosfoproteínas relacionadas ao EC foram significativamente enriquecidas nas categorias relacionadas ao “desenvolvimento pós-embrionário”, “respostas ao estresse” e “regulação da expressão gênica e epigenética” (Figura 6A). Durante a reprogramação do desenvolvimento de protoplastos do musgo (*P. patens*), fosfoproteínas relacionadas às respostas hormonais ligadas ao estresse, mecanismos epigenéticos e pós-transcricionais também foram identificadas (Wang, X. et al., 2014), como nesse trabalho, indicando um importante papel da fosforilação de proteínas na sinalização hormonal, epigenética e nas vias pós-transcricionais durante a aquisição da totipotência, corroborando com os nossas dados.

6.3. Caracterização da homologia das fosfoproteínas, conservação dos fosfosítios e enriquecimento dos *motifs*

Estudos anteriores mapearam os perfis fosfoproteômicos de diferentes espécies e como nos resultados apresentados, a maioria dos fosfopeptídeos identificados são peptídeos fosforilados em apena um resíduo, com o resíduo S representando o local principal de fosforilação, seguido pelo resíduo T em menor escala (Nakagami et al., 2010; van Wijk et al.; Wang et al., 2017). Curiosamente, nenhum resíduo Y fosforilado foi identificado entre os fosfopeptídeos relacionados com EC ou NEC; conforme relatado anteriormente em calos de arroz (Wang et al., 2017), indicando que esses resíduos podem ser menos fosforilados nos calos. Os peptídeos fosforilados em Y ocorrem em baixa abundância em comparação com os resíduos S ou T fosforilados (Nakagami et al., 2010), portanto, técnicas de enriquecimento específicas podem ser necessárias para detectar os peptídeos fosforilados em Y nas amostras de calos de cana-de-açúcar.

Em termos de conservação das fosfoproteínas e dos fosfosítios, os alinhamentos das fosfoproteínas ortólogas em cana-de-açúcar, milho e *Arabidopsis* mostraram que as fosfoproteínas relacionadas ao EC e NEC são

relativamente bem conservadas e compartilham os mesmos padrões de fosfosítios (Tabela 2, Figuras 8, 9 e 10). Uma comparação fosfoproteômica anterior entre os ortólogos de arroz e *Arabidopsis* revelou que um quarto das fosfoproteínas analisadas foram fosforiladas em locais equivalentes ou locais vizinhos (Nakagami et al., 2010). Esses resultados podem ajudar a desvendar mecanismos conservados que são regulados pela fosforilação com processos biológicos relacionados com a aquisição da competência embriogênica em cana-de-açúcar e em outras espécies.

O *motif* [SP] exibiu o maior *score* (Figura 3A) e foi comumente mais enriquecido em diferentes estudos, e as fosfoproteínas contendo esse *motif* estão principalmente envolvidas com respostas ao estresse e localizadas no núcleo e no citosol (van Wijk et al., 2014; Wang, K. et al., 2014; Qiu et al., 2017). Além disso, as fosfoproteínas contendo o *motif* [SP] identificado nas amostras EC foram principalmente relacionadas com as respostas hormonais e de estresse, bem como com o controle da expressão gênica por meio de mecanismos pós-transcricionais e epigenéticos. Esses resultados sugerem um possível relevante papel do *motif* [SP] na aquisição da competência embriogênica durante a adaptação *in vitro* sob as condições de estresse, como já observado em estudos com vegetais sob condições de estresse abiótico (Liu et al., 2019) e no desenvolvimento embriogênico (Qiu et al., 2017; Ishikawa et al., 2019).

O *motif* com o segundo maior *score* foi o [RxxS] (Figura 3B). Os *motifs* [SP] e [RxxS] aparecem com frequência durante o desenvolvimento inicial da semente de arroz, e os autores de um estudo anterior correlacionaram a fosforilação desses *motifs* com quinases *SnRK2* e com fatores de resposta ao ABA, como *ABF1* e *bZIP72* (Qiu et al., 2017). No presente estudo, o *motif* [RxxS] foi identificado como fosforilado nas fosfoproteínas relacionadas ao EC envolvidas na sinalização do ABA, nas respostas ao estresse e no desenvolvimento embrionário. A mesma tendência foi relatada durante a germinação de pólen de kiwis (*Actinidia deliciosa* (A. Chev.) CF Liang & AR Ferguson), (Vannini et al., 2019), na etapa inicial da germinação de sementes em arroz (Qiu et al., 2017), e em embriões de cevada (*Hordeum vulgare* L.)

(Ishikawa et al., 2019), indicando um possível papel da fosforilação desses *motifs* nos estágios iniciais do desenvolvimento vegetal.

6.4. A fosforilação de proteínas está envolvida nas respostas hormonais e de estresse durante a aquisição da competência embriogênica em cana-de-açúcar

Durante a aquisição da competência embriogênica, o ABA atua como um mensageiro químico crítico coordenando as respostas epigenéticas e abióticas ao estresse por meio de vias de sinalização controladas por mecanismos dependentes de fosforilação (Karami e Saidi, 2010). Nossos dados revelaram diversas fosfoproteínas relacionadas às respostas hormonais e de estresse em EC, como o ABA responsive element binding factor 1 (ABF1), três isoformas da Low-temperature-induced 65 kDa protein (RD29B), Transcription factor bHLH128 (BH128), LEA1, LEA18 e LEA34, Remorin 4.1 (REM4.1) e GHR1 (Tabela 1).

A proteína Serine/threonine protein kinase OSK3 foi identificada como exclusiva para EC (Tabela 1). Essa proteína é uma subunidade da SnRK1 e é fortemente expressa em sementes imaturas de arroz (Takano et al., 1998). A quinase SnRK1 é um regulador mestre da homeostase energética das plantas, controlando a tolerância ao estresse e as respostas ao ABA pela fosforilação de enzimas metabólicas essenciais e pela fosforilação de proteínas e transcritos regulatórios (Crepin e Rolland, 2019). A quinase OSK3 identificada no presente estudo, compartilha 73% de identidade com a subunidade α-catalítica da SnRK1 em *Arabidopsis* (AKIN10/SnRK1.1/SnRK1α1) (Figura 8; Tabela 2). É importante ressaltar que a quinase OSK3 compartilha o resíduo conservado de lisina K46 (K48 em *Arabidopsis*) e é fosforilada nos resíduos conservados T173 (T175 em *Arabidopsis*) e S337 (S339 em *Arabidopsis*) (Tabela 1; Figura 8). A fosforilação do resíduo T conservado no domínio catalítico "T-*looping*" promove uma atividade mais forte, e a lisina K48 conservada é necessária para a ligação do ATP; ambos os aminoácidos são pré-requisitos para a atividade da quinase (Baena-González et al., 2007). Em embriões de *Phaseolus vulgaris* L., a subunidade catalítica da SnRK1 é fortemente fosforilada e ativa nos estágios iniciais do desenvolvimento, e essa tendência está ligada ao acúmulo de amido

e proteínas, suportando a hipótese de que a subunidade catalítica da SnRK1 é necessária no início do desenvolvimento (Zúñiga-Sánchez et al., 2019).

A proteína KING1 foi identificada menos abundante no EC comparado ao NEC, e fosforilada em S3 (S14 em *Arabidopsis*), S6 (S23 em *Arabidopsis*), S27 (S44 em *Arabidopsis*) e S135 (nenhum fosfosítio correspondente em *Arabidopsis*) (Tabela 1; Figura 8). Essa proteína tem 56% de semelhança com a KING1 de *Arabidopsis*, e as posições dos resíduos de serina correspondem com os fosfosítios identificados nesse estudo (Tabela 2; Figura 8), sugerindo que esses resíduos também são possíveis fosfosítios em *Arabidopsis*. Recentemente foi observado que a KING1 pode regular diretamente a SnRK2, e sua ligação ao domínio N-terminal inibe funcionalmente a atividade quinase *in vitro*, influenciando na sinalização do ABA (Punkkinen et al., 2019).

Também foram identificadas duas fosfoproteínas reguladoras negativas da sinalização por ABA: a Serine/threonine protein phosphatase catalytic subunit (PP2Ac-2), que é fosforilada na S3, e a proteína Serine/threonine protein kinase EDR1 (EDR1), que é fosforilada nos resíduos S300, S351, S404, S480 e S484 (Tabela 1). A subunidade catalítica PP2Ac-2 demonstrou um papel altamente específico na repressão da expressão gênica dependente de ABA (Pernas et al., 2007). A quinase EDR1 regula negativamente a rede de sinalização do ABA, em mutantes *edr1* o tratamento com ABA promoveu a up-regulação de diferentes genes responsivos ao ABA, (Wawrzynska et al., 2008) indicando que em NEC, a sinalização por ABA pode estar menos ativa, afetando a competência embriogênica.

Estudos anteriores relataram que os substratos clássicos de fosforilação da SnRK1 são as proteínas Sucrose-phosphate synthase (SPS), Fructose-2,6-bisphosphatase (F2KP) e Nitrate reductase (NIA) (Crepin e Rolland, 2019). Corroborando com esses dados, foram identificadas as fosfoproteínas com funções metabólicas primárias da família da Sucrose-phosphate synthase isoformas 1 e 3 (SPS1 e SPS3) e a 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (F2KP) relacionadas ao EC e a Nitrate reductase (NIA1) relacionada ao NEC (Tabela 1). Essas fosfoproteínas possuem um *motif* que já foi descrito como alvo das quinases SnRK1 (Nukarinen et al., 2016). Em condições de abundância de sacarose, ausência de luz e estresse abiótico

(Nukarinen et al., 2016; Crepin e Rolland, 2019), assim como o mesmo ambiente usado para a indução da embriogênese somática, a fosforilação por SnRK1 inativa as SPS e F2KP. Esse mecanismo pode ser uma forma eficiente de reprimir processos anabólicos desnecessários que consomem energia e estão envolvidos na tolerância ao estresse durante a embriogênese somática de cana-de-açúcar (Figura 5).

Além disso, a NIA1 foi fosforilada na S534, e as duas fosfoproteínas 14-3-3-like, GF14-A e GF14-C, foram fosforiladas nas S247 e S256, respectivamente (Tabela 1). A fosforilação *downstream* da SnRK1 na proteína NIA é uma etapa regulatória na inativação da enzima. Seu estado fosforilado leva à ligação das 14-3-3 e inibição da NIA por mudança conformacional (Lambeck et al., 2012). A inativação da NIA pode impactar negativamente uma série de processos, incluindo a produção de íons de amônio a partir de nitrato e na sinalização do óxido nítrico (NO) em resposta às condições de ausência de luz e estresses (Tomanov et al., 2018). Dessa forma, considerando que o meio de cultura usado para os EC e NEC é rico em nitrogênio e os calos foram mantidos no escuro, a identificação das duas proteínas 14-3-3 GF14 menos abundantes no EC comparado ao NEC e da NIA1 presente apenas nas amostras do NEC, sugerem um impacto negativo da fosforilação dessas proteínas na via de sinalização do nitrato, podendo estar ligados com a aquisição da competência embriogênica.

A quinase OSK3, identificada como exclusiva para EC, é homóloga à KIN10 de *A. thaliana* (Figura 8), e demonstrou ser uma integradora nas vias de sinalizações por ABA, controlando a homeostase metabólica, as respostas ao estresse e o desenvolvimento embriogênico (Chan et al., 2017). Esses resultados nos levam a acreditar que a enzima *OSK3* fosforilada pode servir como um novo biomarcador da aquisição da competência embriogênica em cana-de-açúcar e um potencial alvo para os programas de transformação genética visando o desenvolvimento de plantas tolerantes a estresses.

6.5. A fosforilação controla a reprogramação da expressão genética por meio de mecanismos epigenéticos e pós-transcricionais durante a aquisição da competência embriogênica em cana-de-açúcar

Os dados fosfoproteômicos identificaram vários peptídeos fosforilados atribuídos às proteínas relacionadas com EC, que são envolvidas nas modificações epigenéticas (Tabela 1). A mudança do estado celular de somático para totipotente requer a reprogramação dos padrões da expressão genética, o que é comumente alcançado pela passagem por um estágio de calo desdiferenciado (Fehér, 2015; Méndez-Hernández et al., 2019). Essa reprogramação celular é principalmente regulada epigeneticamente por mudanças no estado da cromatina (Fehér, 2015; Bednarek e Orłowska, 2020). A desacetilação das histonas mediada pelas HDAs tem se mostrado necessária para a reprogramação das células nos estágios iniciais da embriogênese somática, como a indução de calos, bem como para a manutenção da competência embriogênica (Fehér, 2015; Méndez-Hernández et al., 2019; Bednarek e Orłowska, 2020).

As proteínas Histone deacetylase 6 (HDA6), HDA19 e TOPLESS (TPL) foram identificadas como mais abundantes no EC ou exclusivas no EC (Tabela 1). A HDA6 foi fosforilada nas S12 e S429, a HDA19 nas S8 e S421 e a TPL nas S161 e T233 (Tabela 1). A fosforilação da HDA6 na S427 em *Arabidopsis* (que corresponde a S429 em cana-de-açúcar) (Figura 9) é importante para sua atividade e função (Yu et al., 2017) e a descoberta de que a HDA6 é fosforilada no mesmo resíduo S em cana e *Arabidopsis* sugere um mecanismo conservado de controle da atividade da HDA6 entre essas espécies. As HDAs catalisam a remoção de grupamentos acetil das histonas, mediando a repressão da expressão gênica e a interação das histonas com seus ligantes (Luo et al., 2017). Já as proteínas TPL funcionam como arcabouços repressores transcricionais, interagindo com complexos modificadores da cromatina, regulando as respostas durante o desenvolvimento (Causier et al., 2012). As HDAs e TPL modulam uma gama de respostas hormonais, como a expressão de genes responsivos ao ABA e Aux, envolvidos nas vias de sinalizações da embriogênese somática (Fehér, 2015).

O fator de transcrição Trihelix (ASIL1), que é fosforilado na S399, também foi identificado como único no EC (Tabela 1). O ASIL1 atua em conjunto com a HDA6 para evitar que o programa da maturação ocorra no início da embriogênese (Willmann et al., 2011). Tanto a HDA6 quanto a HDA19, são expressas nas fases de proliferação da embriogênese somática em *Q. suber* (Pérez et al., 2015). A HDA19, juntamente com a proteína TPL, são recrutadas por *WOX* para formar um complexo que inibe a expressão dos programas de diferenciação, resultando na manutenção do estado celular totipotente (Pi et al., 2015). Além disso, a TPL trabalha em conjunto com a HDA19 para regular a diferenciação dos meristemas caulinar/radicular, sendo necessárias para a transição dos estágios na embriogênese (Long et al., 2006). Juntos, esses resultados levam a hipótese de que as HDA6, HDA19, ASIL1 e TPL fosforiladas, podem ser reguladores do estado totipotente nos calos e, portanto, podem ser potenciais novos biomarcadores para a aquisição da competência embriogênica em cana-de-açúcar (Figura 7).

Além das modificações epigenéticas, a regulação no metabolismo dos RNAs é uma etapa regulatória na expressão gênica, controlando as funções celulares durante a desdiferenciação *in vitro* e as respostas ao estresse (Ohtani, 2015; Kalinina et al., 2018). Nesse contexto, os dados gerados no presente estudo permitiram a identificação de diferentes fosfoproteínas relacionadas ao EC envolvidas com os processos metabólicos de RNAs, como a Argonaute1b (AGO1), DEAH5, SCL30, a Serine/arginine-rich splicing factor SR45 (SR45) e a UB2C (Tabela 1).

A proteína AGO1 foi identificada como exclusiva no EC e fosforilada na S1037, que corresponde a S1001 em *Arabidopsis* (Figura 9). Esse fosfosítio já foi relatado em estudos fosfoproteômicos anteriores com *Arabidopsis* (Sugiyama et al., 2008; Jones et al., 2009) sugerindo um mecanismo conservado de PTM para a AGO1 entre essas espécies. As proteínas AGO são fatores chave no silenciamento por RNA de interferência, regulando a expressão gênica ao nível transcricional, pós-transcricional e translacional (Bajczyk et al., 2019). O envolvimento do miR168 na homeostase da AGO1, controla as respostas ao estresse induzidas por ABA e a transdução de sinais (Li et al., 2012). Membros da família de proteínas AGO foram mais abundantes quando EC foi comparado

com NEC em algodão, e os autores sugeriram o envolvimento da AGO na totipotência embriogênica (Guo et al., 2019).

Além disso, as fosfoproteínas DEAH5 e UB2C foram mais abundantes no EC e fosforiladas nas S457 e S425, respectivamente. As proteínas SCL30, SR45 e SAC3/GANP family protein (SAC3) foram exclusivas no EC. A SCL30 foi fosforilada nas S5, S10 e S22; SR45 nas S121, S122, S163, S170 e S180; e S631 para SAC3 (Tabela 1). Essas fosfoproteínas estão envolvidas nos eventos de *splicing* dos mRNAs. Esse mecanismo tem se mostrado crítico para a desdiferenciação das células vegetais, nas respostas ao estresse abiótico e na formação de embriões (Ohtani, 2015; Szakonyi e Duque, 2018). Os eventos de fosforilação são sugeridos como mediadores das proteínas do spliceossomo sob estresses (Barua et al., 2019; Zhao et al., 2019) e durante o desenvolvimento embriogênico (Ishikawa et al., 2019). A proteína UB2C foi identificada fosforilada na S428 em milho sob estresse abiótico (Zhao et al., 2019), e esse resíduo corresponde a S425 identificada nesse estudo, para as células de cana-de-açúcar em condições *in vitro* (Figura 9). As proteínas do spliceossomo SCL30 e SR45 também foram fosforiladas em resposta ao estresse abiótico em *Cicer arietinum* L. (Barua et al., 2019).

Portanto, os mecanismos de fosforilação podem participar na reprogramação da expressão gênica envolvida na rediferenciação das células vegetais durante o cultivo *in vitro*, dependendo de um ajuste fino dos estímulos ambientais com o equilíbrio hormonal. Nesse sentido, os presentes resultados destacam sobre os mecanismos integrados de PTM envolvendo proteínas de remodelamento da cromatina, como as HDA6, HDA19 e TPL, bem como proteínas de controle metabólico do RNA, incluindo as DEAH5, SCL30, AGO1, UB2C e SR45, que podem estar relacionadas com a aquisição da competência embriogênica em cana-de-açúcar (Figura 7).

6.6. Os processos de desenvolvimento embriogênico estão ligados com a fosforilação de proteínas e a aquisição da competência embriogênica

A análise fosfoproteômica comparativa entre EC e NEC revelou diferentes fosfoproteínas relacionadas com EC envolvidas no desenvolvimento

embriogênico e classificadas como mais abundantes no EC, como a dnaJ homologue subfamily C protein GRV2 (GRV2), ou aquelas exclusivas no EC, como as ASIL1, transcription elongation factor SPT6-like isoform X1 (SPT6L), M3KE1, a DNA replication licensing factor MCM2 (MCM2) e a Caleosin 1 (CLO1) (Tabela 1). Os fosfosítios dessas proteínas e seus resíduos correspondentes em *Arabidopsis* são os seguintes: GRV2 é fosforilada na S762, que não possui homólogo conhecido em *Arabidopsis*; ASIL1 na S399 (S382 em *Arabidopsis*); SPT6L na S7 (S7 em *Arabidopsis*); MCM2 na S14 (S18 em *Arabidopsis*); e M3KE1 na S744 e S859, que correspondem a S788 e S922 em *Arabidopsis* (Tabela 2, Figura 10). A fosfoproteína CLO1 apresentou uma exclusiva fosforilação na T106 e não possui fosforesíduos correspondentes em *Arabidopsis* (Figura 10).

Análises genéticas indicaram que a quinase *M3KE1* é expressa em embriões e está envolvida no processo de expansão e alongamento celular; a letalidade dos embriões ocorre nos duplo mutantes *map3kε1*$^{-/-}$*;map3kε2*$^{-/-}$ (Chaiwongsar et al., 2012). Como já descrito, o complexo HDA6/ASIL1 atua *downstream* aos miRNAs, para reprimir o programa de maturação durante os estágios iniciais da embriogênese (Willmann et al., 2011). A SPT6L desempenha um papel crítico no desenvolvimento do embrião, especificando o destino apical-basal (Gu et al., 2012). Embriões de *Arabidopsis* mutantes para *spt6l* mostraram proeminente expressão dos genes *PHABULOSA (PHB)* e *PHAVOLUTA (PHV)*, demonstrando um importante papel da SPT6L na especificação dos eixos apical-basal dos embriões (Gu et al., 2012). As proteínas MCM funcionam como DNA helicases durante a fase S do ciclo celular (Ni et al., 2009). Também foi demonstrado que a perturbação da MCMC2 em *Arabidopsis* é letal em um estágio muito inicial da embriogênese e que a super-expressão da *AtMCM2* induz a produção de células indiferenciadas na columela da coifa da raiz (Ni et al., 2009). A GRV2 é necessária para a formação dos endossomos envolvidos no transporte de proteínas para vacúolos de armazenamento de proteínas, como no tráfego vesicular dos transportadores de Aux, que são essenciais para a correta distribuição de Aux e, consequentemente, determinação do eixo de crescimento do embrião. Dessa forma, a identificação das proteínas fosforiladas M3KE1, HDA6/ASIL, SPT6L, MCM2 e GRV2, como up-reguladas ou exclusivas

no EC, sugere a atividade dessas fosfoproteínas na aquisição da competência embriogênica (Figura 7).

A proteína CLO1 é reconhecida como uma proteína embrião-específica em *Arabidopsis* (Nuccio e Thomas, 1999) A CLO1 é importante na degradação e tráfego de lipídios, acúmulo ou *turnover* de triacilglicerol, na atividade de peroxigenase mediada por cálcio nas vias de sinalizações das oxilipinas e nas respostas ao estresse (Shen et al., 2016). Estudos anteriores identificaram a CLO1 durante a embriogênese somática em milho e na embriogênese inicial de micrósporos em *B. napus* (Che et al., 2006; Malik et al., 2007). Esses dados refletem um possível papel da CLO1 nas vias dos lipídios que sustentam a competência embriogênica em cana-de-açúcar. No entanto, o impacto da fosforilação na atividade da CLO1 ainda não foi caracterizado e necessita de maiores investigações.

7. CONCLUSÕES

A análise fosfoproteômica comparativa entre EC e NEC na fase de multiplicação da embriogênese somática de cana-de-açúcar, permitiu a identificação de proteínas fosforiladas envolvidas nas respostas ao ABA ligadas com a tolerância ao estresse e na regulação da expressão gênica através da remodelação da cromatina e controle transcricional. A quinase OSK3, da família SnRK1, é um regulador central da homeostase metabólica e da sinalização por ABA em resposta às condições de estresse. Assim, a identificação da OSK3 exclusiva no EC e com a fosforilação ocorrendo em seu resíduo conservado (T173) do domínio catalítico "*T-loop*", sugere um papel chave dessa quinase na aquisição da competência embriogênica em cana-de-açúcar. Por outro lado, as proteínas PP2Ac-2 e EDR1, que são reguladores negativos das respostas ao ABA, foram identificadas entre as fosfoproteínas relacionadas ao NEC, reforçando a hipótese de que as respostas induzidas por ABA têm funções importantes na embriogênese somática da cana-de-açúcar. As fosfoproteínas envolvidas com modificações epigenéticas, como as HDA6, HDA19 e TPL, bem como no metabolismo de RNAs, incluindo a AGO1, DEAH5, SCL30, UB2C e SR45, também foram identificadas como fosfoproteínas relacionadas ao EC. Esses resultados sugerem um papel da fosforilação no controle da

reprogramação da expressão gênica nos níveis pós-transcricionais e epigenético, para mediar as respostas ao estresse durante a aquisição da competência embriogênica. Além disso, as fosfoproteínas relacionadas ao EC GRV2, ASIL1, SPT6L, M3KE1 e MCM2, estão diretamente ligadas com a formação e padronização dos embriões e podem representar vias de sinalizações adicionais envolvidas com a aquisição da competência embriogênica em cana-de-açúcar. Este é o primeiro estudo fosfoproteômico em larga escala para cana-de-açúcar. Embora análises adicionais para validação sejam necessárias, novos sítios de fosforilação e fosfoproteínas específicas relacionadas ao EC com potencial para serem candidatas a biomarcadoras para a aquisição da competência embriogênica foram identificadas para cana-de-açúcar.

O presente trabalho está publicado no periódico científico internacional *Journal of Proteome Research*:

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. **Journal of Proteome Research,** v. 19, n. 10, p. 4145-4157, 2020/10/02 2020. DOI: 10.1021/acs.jproteome.0c00652.

8. REFERÊNCIAS

AGUILAR-HERNÁNDEZ, V.; LOYOLA-VARGAS, V. M. Advanced proteomic approaches to elucidate somatic embryogenesis. **Frontiers in Plant Science,** v. 9, p. 1658, 2018.

AHLOOWALIA, B.; MARETZKI, A. Plant regeneration via somatic embryogenesis in sugarcane. **Plant Cell Reports,** v. 2, n. 1, p. 21-25, 1983.

AROONLUK, S.; ROYTRAKUL, S.; JANTASURIYARAT, C. Identification and Characterization of Phosphoproteins in Somatic Embryogenesis Acquisition during Oil Palm Tissue Culture. **Plants,** v. 9, n. 1, p. 36, 2020.

ARRUDA, P. Genetically modified sugarcane for bioenergy generation. **Current Opinion in Biotechnology,** v. 23, n. 3, p. 315-322, 2012.

AYODELE, B. V.; ALSAFFAR, M. A.; MUSTAPA, S. I. An overview of integration opportunities for sustainable bioethanol production from first-and second-generation sugar-based feedstocks. **Journal of Cleaner Production,** v. 245, p. 118857, 2020.

BAENA-GONZÁLEZ, E.; ROLLAND, F.; THEVELEIN, J. M.; SHEEN, J. A central integrator of transcription networks in plant stress and energy signalling. **Nature,** v. 448, n. 7156, p. 938, 2007.

BAILEY, T. L.; BODEN, M.; BUSKE, F. A.; FRITH, M.; GRANT, C. E.; CLEMENTI, L.; REN, J.; LI, W. W.; NOBLE, W. S. MEME SUITE: tools for motif discovery and searching. **Nucleic Acids Research,** v. 37, n. suppl_2, p. W202-W208, 2009.

BAJCZYK, M.; BHAT, S. S.; SZEWC, L.; SZWEYKOWSKA-KULINSKA, Z.; JARMOLOWSKI, A.; DOLATA, J. Novel nuclear functions of Arabidopsis ARGONAUTE1: Beyond RNA interference. **Plant Physiology**, p. pp. 01351.2018, 2019.

BARUA, P.; LANDE, N. V.; SUBBA, P.; GAYEN, D.; PINTO, S.; KESHAVA PRASAD, T.; CHAKRABORTY, S.; CHAKRABORTY, N. Dehydration-responsive nuclear proteome landscape of chickpea (*Cicer arietinum* L.) reveals phosphorylation-mediated regulation of stress response. **Plant, Cell & Environment,** v. 42, n. 1, p. 230-244, 2019.

BECK, S.; MICHALSKI, A.; RAETHER, O.; LUBECK, M.; KASPAR, S.; GOEDECKE, N.; BAESSMANN, C.; HORNBURG, D.; MEIER, F.; PARON, I. The

Impact II, a very high-resolution quadrupole time-of-flight instrument (QTOF) for deep shotgun proteomics. **Molecular & Cellular Proteomics,** v. 14, n. 7, p. 2014-2029, 2015.

BEDNAREK, P. T.; ORŁOWSKA, R. Plant tissue culture environment as a switch-key of (epi) genetic changes. **Plant Cell, Tissue Organ Culture,** v. 140, n. 2, p. 245-257, 2020.

BIGEARD, J.; RAYAPURAM, N.; PFLIEGER, D.; HIRT, H. Phosphorylation-dependent regulation of plant chromatin and chromatin-associated proteins. **Proteomics,** v. 14, n. 19, p. 2127-2140, 2014.

BOUDSOCQ, M.; DANQUAH, A.; DE ZÉLICOURT, A.; HIRT, H.; COLCOMBET, J. Plant MAPK cascades: Just rapid signaling modules? **Plant signaling & behavior,** v. 10, n. 9, p. e1062197, 2015.

BRAVO, S.; BERTÍN, A.; TURNER, A.; SEPÚLVEDA, F.; JOPIA, P.; PARRA, M. J.; CASTILLO, R.; HASBÚN, R. Differences in DNA methylation, DNA structure and embryogenesis-related gene expression between embryogenic and non embryogenic lines of *Pinus radiata* D. don. **Plant Cell, Tissue Organ Culture,** v. 130, n. 3, p. 521-529, 2017.

BURRIEZA, H. P.; RIZZO, A. J.; VALE, E. M.; SILVEIRA, V.; MALDONADO, S. Shotgun proteomic analysis of quinoa seeds reveals novel lysine-rich seed storage globulins. **Food Chemistry,** v. 293, p. 299-306, 2019.

CAUSIER, B.; ASHWORTH, M.; GUO, W.; DAVIES, B. The TOPLESS interactome: a framework for gene repression in *Arabidopsis*. **Plant Physiology,** v. 158, n. 1, p. 423-438, 2012.

CHAIWONGSAR, S.; STROHM, A.; SU, S.-H.; KRYSAN, P. J. Genetic analysis of the *Arabidopsis* protein kinases MAP3Kε1 and MAP3Kε2 indicates roles in cell expansion and embryo development. **Frontiers in Plant Science,** v. 3, p. 228, 2012.

CHAN, A.; CARIANOPOL, C.; TSAI, A. Y.-L.; VARATHARAJAH, K.; CHIU, R. S.; GAZZARRINI, S. SnRK1 phosphorylation of FUSCA3 positively regulates embryogenesis, seed yield, and plant growth at high temperature in *Arabidopsis*. **Journal of Experimental Botany,** v. 68, n. 15, p. 4219-4231, 2017.

CHE, P.; LOVE, T. M.; FRAME, B. R.; WANG, K.; CARRIQUIRY, A. L.; HOWELL, S. H. Gene expression patterns during somatic embryo development and

germination in maize Hi II callus cultures. **Plant Molecular Biology,** v. 62, n. 1-2, p. 1-14, 2006.

CHENG, A.; GRANT, C. E.; NOBLE, W. S.; BAILEY, T. L. MoMo: Discovery of statistically significant post-translational modification motifs. **bioRxiv**, p. 410050, 2018.

CONAB, C. N. D. A. **Acompanhamento da safra brasileira Cana-de-açúcar**. Brasília: Segundo levantamento. 7**:** 1-64 p. 2020.

CORDEWENER, J.; BERGERVOET, J.; LIU, C.-M. Changes in protein synthesis and phosphorylation during microspore embryogenesis in *Brassica napus*. **Journal of plant physiology,** v. 156, n. 2, p. 156-163, 2000.

COX, J.; HEIN, M. Y.; LUBER, C. A.; PARON, I.; NAGARAJ, N.; MANN, M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. **Molecular & Cellular Proteomics,** v. 13, n. 9, p. 2513-2526, 2014.

COX, J.; MANN, M. MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. **Nature Biotechnology,** v. 26, n. 12, p. 1367, 2008.

CREPIN, N.; ROLLAND, F. SnRK1 activation, signaling, and networking for energy homeostasis. **Current Opinion in Plant Biology,** v. 51, p. 29-36, 2019.

DE-LA-PEÑA, C.; NIC-CAN, G. I.; GALAZ-ÁVALOS, R. M.; AVILEZ-MONTALVO, R.; LOYOLA-VARGAS, V. M. The role of chromatin modifications in somatic embryogenesis in plants. **Frontiers in plant science,** v. 6, n. 635, 2015.

DE MATOS, M.; SANTOS, F.; EICHLER, P. Sugarcane world scenario. In: (Ed.). **Sugarcane Biorefinery, Technology and Perspectives**: Elsevier, 2020. p.1-19.

DE SOUZA BARBOSA, G. V.; DOS SANTOS, J. M.; DINIZ, C. A.; CURSI, D. E.; HOFFMANN, H. P. Energy cane breeding. In: SANTOS, F.;RABELO, S. C.*, et al* (Ed.). **Sugarcane Biorefinery, Technology and Perspectives**: Academic Press, 2020. p.103-116. ISBN 978-0-12-814236-3.

DEUTSCH, E. W.; CSORDAS, A.; SUN, Z.; JARNUCZAK, A.; PEREZ-RIVEROL, Y.; TERNENT, T.; CAMPBELL, D. S.; BERNAL-LLINARES, M.; OKUDA, S.; KAWANO, S.; MORITZ, R. L.; CARVER, J. J.; WANG, M.; ISHIHAMA, Y.;

BANDEIRA, N.; HERMJAKOB, H.; VIZCAÍNO, J. A. The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. **Nucleic Acids Research,** v. 45, n. D1, p. D1100-D1106, 2016.

DOUÉTTS-PERES, J. C.; ARAGÃO, V. P. M.; CRUZ, M. A. L.; REIS, R. S.; ELBL, P.; DOS SANTOS, A. L. W.; FLOH, E. I. S.; SILVEIRA, V.; SANTA-CATARINA, C. AaMps1 protein inhibition regulates the protein profile, nitric oxide, carbohydrate and polyamine contents in embryogenic suspension cultures of *Araucaria angustifolia* (Bertol.) Kuntze (Araucariaceae). **Plant Cell, Tissue and Organ Culture** p. 1-14, 2019.

DUARTE-AKÉ, F.; NIC-CAN, G.; DE-LA-PEÑA, C. Somatic Embryogenesis: Polycomb Complexes Control Cell-to-Embryo Transition. In: (Ed.). **Epigenetics in Plants of Agronomic Importance: Fundamentals and Applications**: Springer, 2019. p.339-354.

EFFERTH, T. **Biotechnology applications of plant callus cultures**. Engineering 5**:** 50–59 p. 2019.

FEHÉR, A. Somatic embryogenesis—Stress-induced remodeling of plant cell fate. **Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms,** v. 1849, n. 4, p. 385-402, 2015.

FEHÉR, A. Callus, dedifferentiation, totipotency, somatic embryogenesis: what these terms mean in the era of molecular plant biology? **Frontiers in plant science,** v. 10, p. 536, 2019.

FRAGA, H. P.; VIEIRA, L. N.; CAPRESTANO, C. A.; STEINMACHER, D. A.; MICKE, G. A.; SPUDEIT, D. A.; PESCADOR, R.; GUERRA, M. P. 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,** v. 31, n. 12, p. 2165-2176, 2012.

GARSMEUR, O.; DROC, G.; ANTONISE, R.; GRIMWOOD, J.; POTIER, B.; AITKEN, K.; JENKINS, J.; MARTIN, G.; CHARRON, C.; HERVOUET, C. A mosaic monoploid reference sequence for the highly complex genome of sugarcane. **Nature communications,** v. 9, n. 1, p. 1-10, 2018.

GLIWICKA, M.; NOWAK, K.; BALAZADEH, S.; MUELLER-ROEBER, B.; GAJ, M. D. Extensive modulation of the transcription factor transcriptome during somatic embryogenesis in *Arabidopsis thaliana*. **Plos one,** v. 8, n. 7, p. e69261, 2013.

GRZYBKOWSKA, D.; MOROŃCZYK, J.; WÓJCIKOWSKA, B.; GAJ, M. D. Azacitidine (5-AzaC)-treatment and mutations in DNA methylase genes affect embryogenic response and expression of the genes that are involved in somatic embryogenesis in *Arabidopsis*. **Plant Growth Regulation,** v. 85, n. 2, p. 243-256, 2018.

GRZYBKOWSKA, D.; NOWAK, K.; GAJ, M. D. Hypermethylation of Auxin-Responsive Motifs in the Promoters of the Transcription Factor Genes Accompanies the Somatic Embryogenesis Induction in *Arabidopsis*. **International journal of molecular sciences,** v. 21, n. 18, p. 6849, 2020.

GU, X.-L.; WANG, H.; HUANG, H.; CUI, X.-F. SPT6L encoding a putative WG/GW-repeat protein regulates apical–basal polarity of embryo in *Arabidopsis*. **Molecular Plant,** v. 5, n. 1, p. 249-259, 2012.

GUIDERDONI, E.; MEROT, B.; EKSOMTRAMAGE, T.; PAULET, F.; FELDMANN, P.; GLASZMANN, J.-C. Somatic embryogenesis in sugarcane (*Saccharum* species). In: (Ed.). **Somatic Embryogenesis and Synthetic Seed II**: Springer, 1995. p.92-113.

GUO, H.; FAN, Y.; GUO, H.; WU, J.; YU, X.; WEI, J.; LIAN, X.; ZHANG, L.; GOU, Z.; FAN, Y. Somatic embryogenesis critical initiation stage-specific mCHH hypomethylation reveals epigenetic basis underlying embryogenic redifferentiation in cotton. **Plant Biotechnology Journal**, 2020.

GUO, H.; GUO, H.; ZHANG, L.; FAN, Y.; FAN, Y.; TANG, Z.; ZENG, F. Dynamic TMT-Based Quantitative Proteomics Analysis of Critical Initiation Process of Totipotency during Cotton Somatic Embryogenesis Transdifferentiation. **International Journal of Molecular Sciences,** v. 20, n. 7, p. 1691, 2019.

HAZUBSKA-PRZYBYŁ, T.; RATAJCZAK, E.; OBARSKA, A.; PERS-KAMCZYC, E. Different Roles of Auxins in Somatic Embryogenesis Efficiency in Two Picea Species. **International Journal of Molecular Sciences,** v. 21, n. 9, p. 3394, 2020.

HE, Y.; ZHOU, J.; SHAN, L.; MENG, X. Plant cell surface receptor-mediated signaling–a common theme amid diversity. **Journal of Cell Science,** v. 131, n. 2, p. jcs209353, 2018.

HECHT, V.; VIELLE-CALZADA, J.-P.; HARTOG, M. V.; SCHMIDT, E. D.; BOUTILIER, K.; GROSSNIKLAUS, U.; DE VRIES, S. C. The *Arabidopsis*

SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture. **Plant Physiology,** v. 127, n. 3, p. 803-816, 2001.

HERINGER, A. S.; BARROSO, T.; MACEDO, A. F.; SANTA-CATARINA, C.; SOUZA, G. H. M. F.; FLOH, E. I. S.; DE SOUZA-FILHO, G. A.; SILVEIRA, V. Label-free quantitative proteomics of embryogenic and non-embryogenic callus during sugarcane somatic embryogenesis. **PloS One,** v. 10, n. 6, p. e0127803, 2015.

HERINGER, A. S.; REIS, R. S.; PASSAMANI, L. Z.; DE SOUZA-FILHO, G. A.; SANTA-CATARINA, C.; SILVEIRA, V. Comparative proteomics analysis of the effect of combined red and blue lights on sugarcane somatic embryogenesis. **Acta Physiologiae Plantarum,** v. 39, n. 2, p. 52, 2017.

HERINGER, A. S.; SANTA-CATARINA, C.; SILVEIRA, V. Insights from proteomic studies into plant somatic embryogenesis. **Proteomics,** v. 18, n. 5-6, p. 1700265, 2018.

HO, W.-J.; VASIL, I. K. Somatic embryogenesis in sugarcane (*Saccharum officinarum* L.) I. The morphology and physiology of callus formation and the ontogeny of somatic embryos. **Protoplasma,** v. 118, n. 3, p. 169-180, 1983.

HU, H.; XIONG, L.; YANG, Y. Rice SERK1 gene positively regulates somatic embryogenesis of cultured cell and host defense response against fungal infection. **Planta,** v. 222, n. 1, p. 107-117, 2005.

IKEUCHI, M.; FAVERO, D. S.; SAKAMOTO, Y.; IWASE, A.; COLEMAN, D.; RYMEN, B.; SUGIMOTO, K. Molecular mechanisms of plant regeneration. **Annual review of plant biology,** v. 70, p. 377-406, 2019.

IKEUCHI, M.; IWASE, A.; RYMEN, B.; LAMBOLEZ, A.; KOJIMA, M.; TAKEBAYASHI, Y.; HEYMAN, J.; WATANABE, S.; SEO, M.; DE VEYLDER, L. Wounding triggers callus formation via dynamic hormonal and transcriptional changes. **Plant physiology,** v. 175, n. 3, p. 1158-1174, 2017.

INGRELL, C. R.; MILLER, M. L.; JENSEN, O. N.; BLOM, N. NetPhosYeast: prediction of protein phosphorylation sites in yeast. **Bioinformatics,** v. 23, n. 7, p. 895-897, 2007.

ISHIKAWA, S.; BARRERO, J.; TAKAHASHI, F.; PECK, S.; GUBLER, F.; SHINOZAKI, K.; UMEZAWA, T. Comparative phosphoproteomic analysis of

barley embryos with different dormancy during imbibition. **International Journal of Molecular Sciences,** v. 20, n. 2, p. 451, 2019.

IWASE, A.; OHME-TAKAGI, M.; SUGIMOTO, K. WIND1: a key molecular switch for plant cell dedifferentiation. **Plant Signaling Behavior,** v. 6, n. 12, p. 1943-1945, 2011.

JAGODZIK, P.; TAJDEL-ZIELIŃSKA, M.; CIEŚLA, A.; MARCZAK, M.; LUDWIKOW, A. Mitogen-activated protein kinase cascades in plant hormone signaling. **Frontiers in Plant Science,** v. 9, p. 1387, 2018.

JI, L.; MATHIONI, S. M.; JOHNSON, S.; TUCKER, D.; BEWICK, A. J.; DO KIM, K.; DARON, J.; SLOTKIN, R. K.; JACKSON, S. A.; PARROTT, W. A. Genome-wide reinforcement of DNA methylation occurs during somatic embryogenesis in soybean. **The Plant Cell,** v. 31, n. 10, p. 2315-2331, 2019.

JIMÉNEZ, V. M. Involvement of plant hormones and plant growth regulators on in vitro somatic embryogenesis. **Plant Growth Regulation,** v. 47, n. 2-3, p. 91-110, 2005.

JIMÉNEZ, V. M.; BANGERTH, F. Endogenous hormone levels in explants and in embryogenic and non-embryogenic cultures of carrot. **Physiologia Plantarum,** v. 111, n. 3, p. 389-395, 2001.

JONES, A. M.; MACLEAN, D.; STUDHOLME, D. J.; SERNA-SANZ, A.; ANDREASSON, E.; RATHJEN, J. P.; PECK, S. C. Phosphoproteomic analysis of nuclei-enriched fractions from *Arabidopsis thaliana*. **Journal of Proteomics,** v. 72, n. 3, p. 439-451, 2009.

KALININA, N. O.; MAKAROVA, S.; MAKHOTENKO, A.; LOVE, A. J.; TALIANSKY, M. The multiple functions of the nucleolus in plant development, disease and stress responses. **Frontiers in Plant Science,** v. 9, p. 132, 2018.

KARAMI, O.; AGHAVAISI, B.; POUR, A. M. Molecular aspects of somatic-to-embryogenic transition in plants. **Journal of chemical biology,** v. 2, n. 4, p. 177-190, 2009.

KARAMI, O.; SAIDI, A. The molecular basis for stress-induced acquisition of somatic embryogenesis. **Molecular Biology Reports,** v. 37, n. 5, p. 2493-2507, 2010.

KARLOVA, R.; BOEREN, S.; VAN DONGEN, W.; KWAAITAAL, M.; AKER, J.; VERVOORT, J.; DE VRIES, S. Identification of *in vitro* phosphorylation sites in

the *Arabidopsis thaliana* somatic embryogenesis receptor-like kinases. **Proteomics,** v. 9, n. 2, p. 368-379, 2009.

KAUSCH, A. P.; NELSON-VASILCHIK, K.; HAGUE, J.; MOOKKAN, M.; QUEMADA, H.; DELLAPORTA, S.; FRAGOSO, C.; ZHANG, Z. J. Edit at will: Genotype independent plant transformation in the era of advanced genomics and genome editing. **Plant science**, 2019.

KERSTEN, B.; AGRAWAL, G. K.; DUREK, P.; NEIGENFIND, J.; SCHULZE, W.; WALTHER, D.; RAKWAL, R. Plant phosphoproteomics: an update. **Proteomics,** v. 9, n. 4, p. 964-988, 2009.

KIM, M. S.; PANDEY, A. Electron transfer dissociation mass spectrometry in proteomics. **Proteomics,** v. 12, n. 4-5, p. 530-542, 2012.

KUMAR, V.; VAN STADEN, J. New insights into plant somatic embryogenesis: an epigenetic view. **Acta Physiologiae Plantarum,** v. 39, n. 9, p. 194, 2017.

LAKSHMANAN, P. Somatic embryogenesis in sugarcane—An addendum to the invited review ‘sugarcane biotechnology: The challenges and opportunities,’in vitro cell. Dev. Biol. Plant 41 (4): 345–363; 2005. **In Vitro Cellular & Developmental Biology,** v. 42, n. 3, p. 201-205, 2006.

LAMBECK, I. C.; FISCHER-SCHRADER, K.; NIKS, D.; ROEPER, J.; CHI, J.-C.; HILLE, R.; SCHWARZ, G. Molecular mechanism of 14-3-3 protein-mediated inhibition of plant nitrate reductase. **Journal of Biological Chemistry,** v. 287, n. 7, p. 4562-4571, 2012.

LARKIN, M. A.; BLACKSHIELDS, G.; BROWN, N.; CHENNA, R.; MCGETTIGAN, P. A.; MCWILLIAM, H.; VALENTIN, F.; WALLACE, I. M.; WILM, A.; LOPEZ, R. Clustal W and Clustal X version 2.0. **Bioinformatics,** v. 23, n. 21, p. 2947-2948, 2007.

LAUGESEN, S.; MESSINESE, E.; HEM, S.; PICHEREAUX, C.; GRAT, S.; RANJEVA, R.; ROSSIGNOL, M.; BONO, J.-J. Phosphoproteins analysis in plants: a proteomic approach. **Phytochemistry,** v. 67, n. 20, p. 2208-2214, 2006.

LEE, K.; SEO, P. J. Dynamic epigenetic changes during plant regeneration. **Trends in plant science**, 2018.

LELJAK-LEVANIĆ, D.; BAUER, N.; MIHALJEVIĆ, S.; JELASKA, S. Changes in DNA methylation during somatic embryogenesis in *Cucurbita pepo* L. **Plant Cell Reports,** v. 23, n. 3, p. 120-127, 2004.

LI, H.; SORIANO, M.; CORDEWENER, J.; MUIÑO, J. M.; RIKSEN, T.; FUKUOKA, H.; ANGENENT, G. C.; BOUTILIER, K. The histone deacetylase inhibitor trichostatin a promotes totipotency in the male gametophyte. **The Plant Cell,** v. 26, n. 1, p. 195-209, 2014.

LI, J.; WANG, M.; LI, Y.; ZHANG, Q.; LINDSEY, K.; DANIELL, H.; JIN, S.; ZHANG, X. Multi-omics analyses reveal epigenomics basis for cotton somatic embryogenesis through successive regeneration acclimation process. **Plant biotechnology journal,** v. 17, n. 2, p. 435-450, 2019.

LI, W.; CUI, X.; MENG, Z.; HUANG, X.; XIE, Q.; WU, H.; JIN, H.; ZHANG, D.; LIANG, W. Transcriptional regulation of *Arabidopsis* MIR168a and argonaute1 homeostasis in abscisic acid and abiotic stress responses. **Plant Physiology,** v. 158, n. 3, p. 1279-1292, 2012.

LI, W.; LIU, H.; CHENG, Z. J.; SU, Y. H.; HAN, H. N.; ZHANG, Y.; ZHANG, X. S. DNA methylation and histone modifications regulate de novo shoot regeneration in *Arabidopsis* by modulating WUSCHEL expression and auxin signaling. **PLoS Genet,** v. 7, n. 8, p. e1002243, 2011.

LINACERO, R.; LOPEZ-BILBAO, M.; VÁZQUEZ, A. Expression of different abscisic acid-responsive genes during somatic embryogenesis in sugarcane (*Saccharum officinarum*). **Protoplasma,** v. 217, n. 4, p. 199-204, 2001.

LIU, W.; WANG, C.; SHEN, X.; LIANG, H.; WANG, Y.; HE, Z.; ZHANG, D.; CHEN, F. Comparative transcriptome analysis highlights the hormone effects on somatic embryogenesis in *Catalpa bungei*. **Plant Reproduction,** v. 32, n. 2, p. 141-151, 2019.

LONG, J. A.; OHNO, C.; SMITH, Z. R.; MEYEROWITZ, E. M. TOPLESS regulates apical embryonic fate in *Arabidopsis*. **Science,** v. 312, n. 5779, p. 1520-1523, 2006.

LUO, M.; CHENG, K.; XU, Y.; YANG, S.; WU, K. Plant responses to abiotic stress regulated by histone deacetylases. **Frontiers in Plant Science,** v. 8, p. 2147, 2017.

MA, X.; LV, S.; ZHANG, C.; YANG, C. Histone deacetylases and their functions in plants. **Plant cell reports,** v. 32, n. 4, p. 465-478, 2013.

MAHDAVI-DARVARI, F.; NOOR, N. M.; ISMANIZAN, I. Epigenetic regulation and gene markers as signals of early somatic embryogenesis. **Plant Cell, Tissue and Organ Culture,** v. 120, n. 2, p. 407-422, 2015.

MAILLOT, P.; LEBEL, S.; SCHELLENBAUM, P.; JACQUES, A.; WALTER, B. Differential regulation of SERK, LEC1-Like and Pathogenesis-Related genes during indirect secondary somatic embryogenesis in grapevine. **Plant Physiology and Biochemistry,** v. 47, n. 8, p. 743-752, 2009.

MALIK, M. R.; WANG, F.; DIRPAUL, J. M.; ZHOU, N.; POLOWICK, P. L.; FERRIE, A. M.; KROCHKO, J. E. Transcript profiling and identification of molecular markers for early microspore embryogenesis in *Brassica napus*. **Plant Physiology,** v. 144, n. 1, p. 134-154, 2007.

MARSONI, M.; BRACALE, M.; ESPEN, L.; PRINSI, B.; NEGRI, A. S.; VANNINI, C. Proteomic analysis of somatic embryogenesis in *Vitis vinifera*. **Plant cell reports,** v. 27, n. 2, p. 347-356, 2008.

MEIER, F.; BECK, S.; GRASSL, N.; LUBECK, M.; PARK, M. A.; RAETHER, O.; MANN, M. Parallel accumulation–serial fragmentation (PASEF): multiplying sequencing speed and sensitivity by synchronized scans in a trapped ion mobility device. **Journal of Proteome Research,** v. 14, n. 12, p. 5378-5387, 2015.

MEIER, F.; BRUNNER, A.-D.; KOCH, S.; KOCH, H.; LUBECK, M.; KRAUSE, M.; GOEDECKE, N.; DECKER, J.; KOSINSKI, T.; PARK, M. A. Online parallel accumulation–serial fragmentation (PASEF) with a novel trapped ion mobility mass spectrometer. **Molecular & Cellular Proteomics,** v. 17, n. 12, p. 2534-2545, 2018.

MÉNDEZ-HERNÁNDEZ, H. A.; LEDEZMA-RODRÍGUEZ, M.; AVILEZ-MONTALVO, R. N.; JUÁREZ-GÓMEZ, Y. L.; SKEETE, A.; AVILEZ-MONTALVO, J.; DE-LA-PEÑA, C.; LOYOLA-VARGAS, V. M. Signaling Overview of Plant Somatic Embryogenesis. **Frontiers in plant science,** v. 10, 2019.

MURASHIGE, T.; SKOOG, F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. **Physiologia Plantarum,** v. 15, n. 3, p. 473-497, 1962.

NAKAGAMI, H.; SUGIYAMA, N.; MOCHIDA, K.; DAUDI, A.; YOSHIDA, Y.; TOYODA, T.; TOMITA, M.; ISHIHAMA, Y.; SHIRASU, K. Large-scale

comparative phosphoproteomics identifies conserved phosphorylation sites in plants. **Plant Physiology,** v. 153, n. 3, p. 1161-1174, 2010.

NI, D. A.; SOZZANI, R.; BLANCHET, S.; DOMENICHINI, S.; REUZEAU, C.; CELLA, R.; BERGOUNIOUX, C.; RAYNAUD, C. The *Arabidopsis* MCM2 gene is essential to embryo development and its over-expression alters root meristem function. **New Phytologist,** v. 184, n. 2, p. 311-322, 2009.

NIC-CAN, G. I.; DE LA PENA, C. Epigenetic advances on somatic embryogenesis of agronomical and important crops. In: (Ed.). **Epigenetics in Plants of Agronomic Importance: Fundamentals and Applications**: Springer, 2014. p.91-109.

NIC-CAN, G. I.; LOYOLA-VARGAS, V. M. The Role of the Auxins During Somatic Embryogenesis. In: (Ed.). **Somatic Embryogenesis: Fundamental Aspects and Applications**: Springer, 2016. p.171-182.

NIEVES, N.; SEGURA-NIETO, M.; BLANCO, M. A.; SÁNCHEZ, M.; GONZÁLEZ, A.; GONZÁLEZ, J. L.; CASTILLO, R. Biochemical characterization of embryogenic and non-embryogenic calluses of sugarcane. **In Vitro Cellular Developmental Biology,** v. 39, n. 3, p. 343-345, 2003.

NOLAN, K. E.; IRWANTO, R. R.; ROSE, R. J. Auxin Up-Regulates *MtSERK1* Expression in Both *Medicago truncatula* Root-Forming and Embryogenic Cultures. **Plant Physiology,** v. 133, n. 1, p. 218-230, 2003.

NUCCIO, M. L.; THOMAS, T. L. ATS1 and ATS3: two novel embryo-specific genes in *Arabidopsis thaliana*. **Plant Molecular Biology,** v. 39, n. 6, p. 1153-1163, 1999.

NUKARINEN, E.; NÄGELE, T.; PEDROTTI, L.; WURZINGER, B.; MAIR, A.; LANDGRAF, R.; BÖRNKE, F.; HANSON, J.; TEIGE, M.; BAENA-GONZALEZ, E. Quantitative phosphoproteomics reveals the role of the AMPK plant ortholog SnRK1 as a metabolic master regulator under energy deprivation. **Scientific Reports,** v. 6, p. 31697, 2016.

OHTANI, M. Regulation of RNA metabolism is important for in vitro dedifferentiation of plant cells. **Journal of Plant Research,** v. 128, n. 3, p. 361-369, 2015.

OSORIO-MONTALVO, P.; DE-LA-PEÑA, C.; OROPEZA, C.; NIC-CAN, G.; CÓRDOVA-LARA, I.; CASTILLO-CASTRO, E.; SÁENZ-CARBONELL, L. A peak

in global DNA methylation is a key step to initiate the somatic embryogenesis of coconut palm (*Cocos nucifera* L). **Plant Cell Reports,** v. 39, n. 10, p. 1345-1357, 2020.

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. **International journal of molecular sciences,** v. 19, n. 10, p. 3182, 2018.

PASSAMANI, L. Z.; BERTOLAZI, A. A.; RAMOS, A. C.; SANTA-CATARINA, C.; THELEN, J. J.; SILVEIRA, V. Embryogenic Competence Acquisition in Sugar Cane Callus Is Associated with Differential H^+-Pump Abundance and Activity. **Journal of Proteome Research,** v. 17, n. 8, p. 2767-2779, 2018/08/03 2018.

PASSAMANI, L. Z.; REIS, R. S.; VALE, E. M.; SOUSA, K. R.; ARAGÃO, V. P. M.; SANTA-CATARINA, C.; SILVEIRA, V. 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,** v. 140, n. 2, p. 415-429, 2020/02/01 2020.

PASTERNAK, T. P.; DUDITS, D. Epigenetic Clues to Better Understanding of the Asexual Embryogenesis in planta and in vitro. **Frontiers in plant science,** v. 10, p. 778, 2019.

PEREZ-RIVEROL, Y.; CSORDAS, A.; BAI, J.; BERNAL-LLINARES, M.; HEWAPATHIRANA, S.; KUNDU, D. J.; INUGANTI, A.; GRISS, J.; MAYER, G.; EISENACHER, M. The PRIDE database and related tools and resources in 2019: improving support for quantification data. **Nucleic Acids Research,** v. 47, n. D1, p. D442-D450, 2019.

PÉREZ, M.; CAÑAL, M. J.; TOOROP, P. E. Expression analysis of epigenetic and abscisic acid-related genes during maturation of *Quercus suber* somatic embryos. **Plant Cell, Tissue and Organ Culture,** v. 121, n. 2, p. 353-366, 2015.

PERNAS, M.; GARCÍA-CASADO, G.; ROJO, E.; SOLANO, R.; SÁNCHEZ-SERRANO, J. J. A protein phosphatase 2A catalytic subunit is a negative regulator of abscisic acid signalling. **The Plant Journal,** v. 51, n. 5, p. 763-778, 2007.

PI, L.; AICHINGER, E.; VAN DER GRAAFF, E.; LLAVATA-PERIS, C. I.; WEIJERS, D.; HENNIG, L.; GROOT, E.; LAUX, T. Organizer-derived WOX5

signal maintains root columella stem cells through chromatin-mediated repression of CDF4 expression. **Developmental Cell,** v. 33, n. 5, p. 576-588, 2015.

PUNKKINEN, M.; DENESSIOUK, K.; FUJII, H. *Arabidopsis* KIN gamma subunit 1 has a potential to regulate activity of sucrose nonfermenting 1-related protein kinase 2s (SnRK2s) *in vitro*. **Biologia Plantarum,** v. 63, n. 1, p. 54-58, 2019.

QIU, J.; HOU, Y.; TONG, X.; WANG, Y.; LIN, H.; LIU, Q.; ZHANG, W.; LI, Z.; NALLAMILLI, B. R.; ZHANG, J. Quantitative phosphoproteomic analysis of early seed development in rice (*Oryza sativa* L.). **Plant Molecular Biology,** v. 90, n. 3, p. 249-265, 2016.

QIU, J.; HOU, Y.; WANG, Y.; LI, Z.; ZHAO, J.; TONG, X.; LIN, H.; WEI, X.; AO, H.; ZHANG, J. A comprehensive proteomic survey of ABA-induced protein phosphorylation in rice (*Oryza sativa* L.). **International Journal of Molecular Sciences,** v. 18, n. 1, p. 60, 2017.

QUINGA, L. A. P.; DE FREITAS FRAGA, H. P.; DO NASCIMENTO VIEIRA, L.; GUERRA, M. P. Epigenetics of long-term somatic embryogenesis in *Theobroma cacao* L.: DNA methylation and recovery of embryogenic potential. **Plant Cell, Tissue Organ Culture,** v. 131, n. 2, p. 295-305, 2017.

RAI, M. K.; SHEKHAWAT, N.; GUPTA, A. K.; PHULWARIA, M.; RAM, K.; JAISWAL, U. The role of abscisic acid in plant tissue culture: a review of recent progress. **Plant Cell, Tissue Organ Culture,** v. 106, n. 2, p. 179-190, 2011.

RAMPITSCH, C.; BYKOVA, N. V.; MAUTHE, W.; YAKANDAWALA, N.; JORDAN, M. Phosphoproteomic profiling of wheat callus labelled in vivo. **Plant science,** v. 171, n. 4, p. 488-496, 2006.

REIS, R. S.; DE MOURA, V. E.; HERINGER, A. S.; SANTA-CATARINA, C.; SILVEIRA, V. Putrescine induces somatic embryo development and proteomic changes in embryogenic callus of sugarcane. **Journal of proteomics,** v. 130, p. 170-179, 2016.

RILEY, N. M.; COON, J. J. Phosphoproteomics in the age of rapid and deep proteome profiling. **Analytical chemistry,** v. 88, n. 1, p. 74-94, 2015.

SANTA-CATARINA, C.; HANAI, L. R.; DORNELAS, M. C.; VIANA, A. M.; FLOH, E. I. SERK gene homolog expression, polyamines and amino acids associated

with somatic embryogenic competence of *Ocotea catharinensis* Mez.(Lauraceae). **Plant cell, tissue organ culture,** v. 79, n. 1, p. 53-61, 2004.

SANTOS, M. O.; ARAGÃO, F. J. Role of SERK genes in plant environmental response. **Plant signaling behavior,** v. 4, n. 12, p. 1111-1113, 2009.

SCHMIDT, E.; GUZZO, F.; TOONEN, M.; DE VRIES, S. A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. **Development,** v. 124, n. 10, p. 2049-2062, 1997.

SHANNON, P.; MARKIEL, A.; OZIER, O.; BALIGA, N. S.; WANG, J. T.; RAMAGE, D.; AMIN, N.; SCHWIKOWSKI, B.; IDEKER, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. **Genome Research,** v. 13, n. 11, p. 2498-2504, 2003.

SHARMA, S. K.; MILLAM, S.; HEIN, I.; BRYAN, G. J. Cloning and molecular characterisation of a potato SERK gene transcriptionally induced during initiation of somatic embryogenesis. **Planta,** v. 228, n. 2, p. 319, 2008.

SHEN, Y.; LIU, M.; WANG, L.; LI, Z.; TAYLOR, D. C.; LI, Z.; ZHANG, M. Identification, duplication, evolution and expression analyses of caleosins in Brassica plants and *Arabidopsis* subspecies. **Molecular Genetics and Genomics,** v. 291, n. 2, p. 971-988, 2016.

SHIBUKAWA, T.; YAZAWA, K.; KIKUCHI, A.; KAMADA, H. Possible involvement of DNA methylation on expression regulation of carrot LEC1 gene in its 5′-upstream region. **Gene,** v. 437, n. 1-2, p. 22-31, 2009.

SHUAI, H.; MENG, Y.; LUO, X.; CHEN, F.; ZHOU, W.; DAI, Y.; QI, Y.; DU, J.; YANG, F.; LIU, J. Exogenous auxin represses soybean seed germination through decreasing the gibberellin/abscisic acid (GA/ABA) ratio. **Scientific reports,** v. 7, n. 1, p. 1-11, 2017.

SILVEIRA, V.; DE VITA, A. M.; MACEDO, A. F.; DIAS, M. F. R.; FLOH, E. I. S.; SANTA-CATARINA, C. Morphological and polyamine content changes in embryogenic and non-embryogenic callus of sugarcane. **Plant Cell, Tissue and Organ Culture,** v. 114, n. 3, p. 351-364, 2013.

SKOOG, F.; MILLER, C. Chemical regulation of growth and organ formation in plant tissues cultured. Vitro Symp Soc Exp Biol, 1957.

SNYMAN, S. J.; MEYER, G. M.; KOCH, A. C.; BANASIAK, M.; WATT, M. P. Applications of in vitro culture systems for commercial sugarcane production and

improvement. **In Vitro Cellular & Developmental Biology,** v. 47, n. 2, p. 234-249, 2011/05/01 2011.

SU, Y. H.; LIU, Y. B.; BAI, B.; ZHANG, X. S. Establishment of embryonic shoot–root axis is involved in auxin and cytokinin response during *Arabidopsis* somatic embryogenesis. **Frontiers in plant science,** v. 5, p. 792, 2015.

SU, Y. H.; SU, Y. X.; LIU, Y. G.; ZHANG, X. S. Abscisic acid is required for somatic embryo initiation through mediating spatial auxin response in *Arabidopsis*. **Plant growth regulation,** v. 69, n. 2, p. 167-176, 2013.

SU, Y. H.; ZHAO, X. Y.; LIU, Y. B.; ZHANG, C. L.; O'NEILL, S. D.; ZHANG, X. S. Auxin-induced WUS expression is essential for embryonic stem cell renewal during somatic embryogenesis in *Arabidopsis*. **The Plant Journal,** v. 59, n. 3, p. 448-460, 2009.

SUGIMOTO, K.; JIAO, Y.; MEYEROWITZ, E. M. *Arabidopsis* regeneration from multiple tissues occurs via a root development pathway. **Developmental cell,** v. 18, n. 3, p. 463-471, 2010.

SUGIYAMA, N.; NAKAGAMI, H.; MOCHIDA, K.; DAUDI, A.; TOMITA, M.; SHIRASU, K.; ISHIHAMA, Y. Large-scale phosphorylation mapping reveals the extent of tyrosine phosphorylation in *Arabidopsis*. **Molecular Systems Biology,** v. 4, n. 1, p. 193, 2008.

SUPRASANNA, P.; PATADE, V.; DESAI, N.; DEVARUMATH, R.; KAWAR, P.; PAGARIYA, M.; GANAPATHI, A.; MANICKAVASAGAM, M.; BABU, K. Biotechnological developments in sugarcane improvement: an overview. **Sugar Tech,** v. 13, n. 4, p. 322-335, 2011.

SZAKONYI, D.; DUQUE, P. Alternative splicing as a regulator of early plant development. **Frontiers in Plant Science,** v. 9, 2018.

SZKLARCZYK, D.; GABLE, A. L.; LYON, D.; JUNGE, A.; WYDER, S.; HUERTA-CEPAS, J.; SIMONOVIC, M.; DONCHEVA, N. T.; MORRIS, J. H.; BORK, P. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. **Nucleic Acids Research,** v. 47, n. D1, p. D607-D613, 2018.

TAKANO, M.; KAJIYA-KANEGAE, H.; FUNATSUKI, H.; KIKUCHI, S. Rice has two distinct classes of protein kinase genes related to SNF1 of *Saccharomyces*

cerevisiae, which are differently regulated in early seed development. **Molecular and General Genetics,** v. 260, n. 4, p. 388-394, 1998.

TANAKA, M.; KIKUCHI, A.; KAMADA, H. The *Arabidopsis* histone deacetylases HDA6 and HDA19 contribute to the repression of embryonic properties after germination. **Plant physiology,** v. 146, n. 1, p. 149-161, 2008.

TANG, L. P.; ZHANG, X. S.; SU, Y. H. Regulation of cell reprogramming by auxin during somatic embryogenesis. **aBIOTECH**, p. 1-9, 2020.

TCHORBADJIEVA, M. I. Advances in proteomics of somatic embryogenesis. In: (Ed.). **Somatic Embryogenesis in Ornamentals and Its Applications**: Springer, 2016. p.67-90. ISBN 813222681X.

TOMANOV, K.; NUKARINEN, E.; VICENTE, J.; MENDIONDO, G. M.; WINTER, N.; NEHLIN, L.; WECKWERTH, W.; HOLDSWORTH, M. J.; TEIGE, M.; BACHMAIR, A. Sumoylation and phosphorylation: hidden and overt links. **Journal of Experimental Botany,** v. 69, n. 19, p. 4583-4590, 2018.

UBERSAX, J. A.; FERRELL JR, J. E. Mechanisms of specificity in protein phosphorylation. **Nature reviews Molecular cell biology,** v. 8, n. 7, p. 530-541, 2007.

UDDENBERG, D.; VALLADARES, S.; ABRAHAMSSON, M.; SUNDSTRÖM, J. F.; SUNDÅS-LARSSON, A.; VON ARNOLD, S. Embryogenic potential and expression of embryogenesis-related genes in conifers are affected by treatment with a histone deacetylase inhibitor. **Planta,** v. 234, n. 3, p. 527-539, 2011.

VAN WIJK, K. J.; FRISO, G.; WALTHER, D.; SCHULZE, W. X. Meta-analysis of *Arabidopsis thaliana* phospho-proteomics data reveals compartmentalization of phosphorylation motifs. **The Plant Cell,** v. 26, n. 6, p. 2367-2389, 2014.

VANNINI, C.; MARSONI, M.; SCOCCIANTI, V.; CECCARINI, C.; DOMINGO, G.; BRACALE, M.; CRINELLI, R. Proteasome-mediated remodeling of the proteome and phosphoproteome during kiwifruit pollen germination. **Journal of Proteomics,** v. 192, p. 334-345, 2019.

VLAD, F.; TURK, B. E.; PEYNOT, P.; LEUNG, J.; MERLOT, S. A versatile strategy to define the phosphorylation preferences of plant protein kinases and screen for putative substrates. **The Plant Journal,** v. 55, n. 1, p. 104-117, 2008.

WANG, K.; ZHAO, Y.; LI, M.; GAO, F.; YANG, M. K.; WANG, X.; LI, S.; YANG, P. Analysis of phosphoproteome in rice pistil. **Proteomics,** v. 14, n. 20, p. 2319-2334, 2014.

WANG, X.; ZHOU, S.; CHEN, L.; QUATRANO, R. S.; HE, Y. Phospho-proteomic analysis of developmental reprogramming in the moss *Physcomitrella patens*. **Journal of Proteomics,** v. 108, p. 284-294, 2014.

WANG, Y.; TONG, X.; QIU, J.; LI, Z.; ZHAO, J.; HOU, Y.; TANG, L.; ZHANG, J. A phosphoproteomic landscape of rice (*Oryza sativa*) tissues. **Physiologia Plantarum,** v. 160, n. 4, p. 458-475, 2017.

WATANABE, M. D.; MORAIS, E. R.; CARDOSO, T. F.; CHAGAS, M. F.; JUNQUEIRA, T. L.; CARVALHO, D. J.; BONOMI, A. Process simulation of renewable electricity from sugarcane straw: Techno-economic assessment of retrofit scenarios in Brazil. **Journal of Cleaner Production**, p. 120081, 2020.

WAWRZYNSKA, A.; CHRISTIANSEN, K. M.; LAN, Y.; RODIBAUGH, N. L.; INNES, R. W. Powdery mildew resistance conferred by loss of the ENHANCED DISEASE RESISTANCE1 protein kinase is suppressed by a missense mutation in KEEP ON GOING, a regulator of abscisic acid signaling. **Plant Physiology,** v. 148, n. 3, p. 1510-1522, 2008.

WETIE, A. G. N.; SOKOLOWSKA, I.; WOODS, A. G.; DARIE, C. C. Identification of post-translational modifications by mass spectrometry. **Australian Journal of Chemistry,** v. 66, n. 7, p. 734-748, 2013.

WICKRAMASURIYA, A. M.; DUNWELL, J. M. Global scale transcriptome analysis of *Arabidopsis* embryogenesis in vitro. **BMC genomics,** v. 16, n. 1, p. 301, 2015.

WILLMANN, M. R.; MEHALICK, A. J.; PACKER, R. L.; JENIK, P. D. MicroRNAs regulate the timing of embryo maturation in *Arabidopsis*. **Plant Physiology,** v. 155, n. 4, p. 1871-1884, 2011.

WIŚNIEWSKI, J. R.; ZOUGMAN, A.; NAGARAJ, N.; MANN, M. Universal sample preparation method for proteome analysis. **Nature Methods,** v. 6, n. 5, p. 359, 2009.

WÓJCIK, A. M.; WÓJCIKOWSKA, B.; GAJ, M. D. Current perspectives on the auxin-mediated genetic network that controls the induction of somatic

embryogenesis in plants. **International journal of molecular sciences,** v. 21, n. 4, p. 1333, 2020.

WÓJCIKOWSKA, B.; BOTOR, M.; MOROŃCZYK, J.; WÓJCIK, A. M.; NODZYŃSKI, T.; KARCZ, J.; GAJ, M. D. Trichostatin A triggers an embryogenic transition in *Arabidopsis* explants via an auxin-related pathway. **Frontiers in plant science,** v. 9, p. 1353, 2018.

WÓJCIKOWSKA, B.; WÓJCIK, A. M.; GAJ, M. D. Epigenetic Regulation of Auxin-Induced Somatic Embryogenesis in Plants. **International Journal of Molecular Sciences,** v. 21, n. 7, p. 2307, 2020.

WU, X.; GONG, F.; CAO, D.; HU, X.; WANG, W. Advances in crop proteomics: PTMs of proteins under abiotic stress. **Proteomics,** v. 16, n. 5, p. 847-865, 2016.

XUE, Y.; REN, J.; GAO, X.; JIN, C.; WEN, L.; YAO, X. GPS 2.0, a tool to predict kinase-specific phosphorylation sites in hierarchy. **Molecular & Cellular Proteomics,** v. 7, n. 9, p. 1598-1608, 2008.

YAMAMOTO, N.; KOBAYASHI, H.; TOGASHI, T.; MORI, Y.; KIKUCHI, K.; KURIYAMA, K.; TOKUJI, Y. Formation of embryogenic cell clumps from carrot epidermal cells is suppressed by 5-azacytidine, a DNA methylation inhibitor. **Journal of plant physiology,** v. 162, n. 1, p. 47-54, 2005.

YANG, X.; ZHANG, X. Regulation of somatic embryogenesis in higher plants. **Critical Reviews in Plant Science,** v. 29, n. 1, p. 36-57, 2010.

YIN, X.; WANG, X.; KOMATSU, S. Phosphoproteomics: Protein Phosphorylation in Regulation of Seed Germination and Plant Growth. **Current protein & peptide science,** v. 19, n. 4, p. 401, 2018.

YOSHIDA, M.; HORINOUCHI, S.; BEPPU, T. Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function. **Bioessays,** v. 17, n. 5, p. 423-430, 1995.

YU, C.-W.; TAI, R.; WANG, S.-C.; YANG, P.; LUO, M.; YANG, S.; CHENG, K.; WANG, W.-C.; CHENG, Y.-S.; WU, K. HISTONE DEACETYLASE6 acts in concert with histone methyltransferases SUVH4, SUVH5, and SUVH6 to regulate transposon silencing. **The Plant Cell,** v. 29, n. 8, p. 1970-1983, 2017.

ZHANG, J.; ZHANG, X.; TANG, H.; ZHANG, Q.; HUA, X.; MA, X.; ZHU, F.; JONES, T.; ZHU, X.; BOWERS, J. Allele-defined genome of the autopolyploid

sugarcane *Saccharum spontaneum* L. **Nature Genetics,** v. 50, n. 11, p. 1565, 2018.

ZHAO, X.; BAI, X.; JIANG, C.; LI, Z. Phosphoproteomic analysis of two contrasting maize inbred lines provides insights into the mechanism of salt-stress tolerance. **International Journal of Molecular Sciences,** v. 20, n. 8, p. 1886, 2019.

ZÚÑIGA-SÁNCHEZ, E.; RODRÍGUEZ-SOTRES, R.; COELLO, P.; MARTÍNEZ-BARAJAS, E. Effect of catalytic subunit phosphorylation on the properties of SnRK1 from *Phaseolus vulgaris* embryos. **Physiologia Plantarum,** v. 165, n. 3, p. 632-643, 2019.