

UNIVERSIDADE SANTO AMARO

CIÊNCIAS BIOLÓGICAS

ARTHUR VANNI LOPES

**CARACTERIZAÇÃO TRANSCRICIONAL DA VIA DA
TREALOSE-6-FOSFATO AO LONGO DO DESENVOLVIMENTO DA
CANA-DE-AÇÚCAR**

São Paulo- SP

2023

ARTHUR VANNI LOPES

**CARACTERIZAÇÃO TRANSCRICIONAL DA VIA DA
TREALOSE-6-FOSFATO AO LONGO DO DESENVOLVIMENTO DA
CANA-DE-AÇÚCAR**

Trabalho de Conclusão de Curso
apresentado ao Curso de Ciências
Biológicas da Universidade Santo Amaro –
UNISA, como requisito parcial para
obtenção do título Bacharel em ciências
Biológicas.

Orientador: Dr. Bruno Viana
Navarro

São Paulo- SP

2023

L85c

Lopes, Arthur Vanni

Caracterização transcricional da via da trealose-6-fosfato ao longo do desenvolvimento da cana-de-açúcar / Arthur Vanni Lopes. - 2023.

44 f. : il., color.

Orientador: Prof. Dr. Bruno Viana Navarro.

TCC Graduação. (Curso Superior em Ciências Biológicas) - Universidade Santo Amaro, 2023.

Bibliografia incluída.

1. Biocombustível. 2. Sinalização de açúcar. 3. Sacarose. I. Navarro, Bruno Viana. II. Universidade Santo Amaro. III. Título.

CDD 665.2

Arthur Vanni Lopes
CARACTERIZAÇÃO TRANSCRICIONAL DA VIA DA
TREALOSE-6-FOSFATO AO LONGO DO DESENVOLVIMENTO DA
CANA-DE-AÇÚCAR

Trabalho de conclusão de curso apresentado ao curso de ciências biológicas da
Universidade Santo Amaro – UNISA, como requisito parcial para obtenção do título
Bacharel em Ciências Biológicas
Orientador: Dr. Bruno Viana Navarro

São Paulo _____ de _____ de 2023

Banca Examinadora

Dr. Bruno Viana Navarro

Dr. Hellen Oliveira de Oliveira

Prof. Dr. Luiz Henrique da Silva Nali

Conceito Final

Agradecimentos

Gostaria de agradecer as pessoas que contribuíram independente da forma para realização deste trabalho.

Primeiramente gostaria de agradecer a Deus pela oportunidade da vida que me concedeu, e também a minha família principalmente meus pais, Silvia Maria Vanni Lopes e Olair Domingues Lopes, onde sempre deram apoio em minhas escolhas e meus sonhos e sempre acreditaram em tudo que eu me propusesse a fazer.

Ao Prof. Dr Marcos Silveira Buckeridge por ter permitido que eu faça parte do incrível grupo do Laboratório de Fisiologia Ecológica de Plantas (LAFIECO) e assim conseguir desenvolver meu projeto de pesquisa.

Ao Dr Bruno Viana Navarro que além de ter sido meu orientador neste projeto também foi um grande professor para mim, sou muito grato pelos ensinamentos, conversas, risadas e sermões que tivemos ao longo deste trabalho, nunca irei esquecer.

A Dr Lauana Pereira de Oliveira que também foi muito importante nesta trajetória do meu projeto, sempre me ensinando, dando conselhos e o mais importante, me incentivando para nunca desistir, mostrando sim que sou capaz de fazer o que é preciso.

A Dr Hellen Oliveira de Oliveira por ter aceitado participar da minha banca e me ajudado a analisar meus dados sempre me ensinando e proporcionando momentos alegres

Ao Dr Mauro de Medeiros me ajudando e demonstrando seu excelente trabalho nas etapas de bioinformática.

Ao Darion Samir Longhi Cervantes primeiramente por ser essa pessoa incrível e simpática e também ter proporcionado esta oportunidade de entrar nesta equipe incrível. Sou muito grato a você, obrigada por todos os conselhos.

A equipe LAFIECO, por me acolherem neste laboratório incrível e me incentivarem a sempre continuar muito obrigado mesmo: Naiara, Grayce, Eglee, João, Gabriel, Fernanda, Will. Muito obrigado por deixarem esta experiência mil vezes melhor.

A banca examinadora deste trabalho pelas contribuições e críticas necessárias.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e ao Instituto de Biociências da Universidade de São Paulo pela bolsa de iniciação científica concedida durante a minha graduação.

RESUMO

A cana-de-açúcar (*Saccharum* spp.) é uma cultivar agrícola que possui um elevado potencial bioenergético, sobretudo devido as altas concentrações de sacarose em seus colmos. Ao longo do crescimento e desenvolvimento das plantas, diferentes sinais de açúcar são gerados nos órgãos fonte e dreno em resposta aos estímulos fotossintéticos e do metabolismo do carbono. Dentre estes sinalizadores, a trealose-6-fosfato (Tre6P) é um conhecido regulador homeostático de sacarose em plantas. Este açúcar é um intermediário da via de biossíntese da trealose, a qual é sintetizada a partir da atividade da trealose-6-fosfato sintase (TPS) e hidrolisada pela trealose-6-fosfato fosfatase (TPP), formando uma molécula de trealose que posteriormente é hidrolisada pela trealase (TRE). Assim, o objetivo deste trabalho foi determinar as diferenças de expressão entre sequências de *TPS*, *TPP* e *TRE* previamente identificadas em cana-de-açúcar. Através de um banco de dados do tipo *Target Amplicon RNA-seq* (TAS) recuperamos sequências nucleotídicas e proteicas putativas de *TPS*, *TPP* e *TRE* de folha e colmo de cana-de-açúcar crescida à campo ao longo de 1, 3, 6 e 12 meses de cultivo. Concomitantemente, foram obtidos valores de expressão destes alvos em outros 17 transcriptomas da variedade SP80-3280 de cana-de-açúcar. Dentre os níveis de expressão dos diferentes alvos, um transcrito de *TPS classe I* se destacou, possuindo o maior valor de expressão entre as sequências analisadas. A partir das sequências destes alvos foi avaliada a estrutura proteica e seus respectivos domínios e análises filogenéticas foram inferidas. Valores de expressão em folhas e colmos de cana foram obtidos para oito alvos distintos da via de biossíntese de Tre6P. Observou-se padrões de expressão inverso entre folha e colmo para os alvos de *TPS classe I* analisados, em que *TPSI_A* diminuiu sua expressão ao longo do desenvolvimento e *TPSI_B* aumentou. A identificação do perfil de expressão destes alvos possibilita há identificar alvos importantes que possam auxiliar no incremento de sacarose nesta cultivar agrícola, e conseqüentemente, impactando sua produtividade.

Palavras chaves: Biocombustíveis, sinalização de açúcar, Sacarose.

ABSTRACT

Sugarcane (*Saccharum spp.*) is an agricultural crop with high bioenergy potential, primarily due to the high sucrose concentrations in its culms. During plant growth and development, different sugar signals are generated in source and sink organs in response to photosynthetic stimuli and carbon metabolism. Among these signaling molecules, trehalose-6-phosphate (Tre6P) is a well-known homeostatic regulator of sucrose in plants. This sugar is an intermediate in the trehalose biosynthesis pathway, which is synthesized through the activity of trehalose-6-phosphate synthase (TPS) and hydrolyzed by trehalose-6-phosphate phosphatase (TPP), forming a trehalose molecule that is subsequently hydrolyzed by trehalase (TRE). Thus, the aim of this study was to determine the expression differences among *TPS*, *TPP*, and *TRE* sequences previously identified in sugarcane. Using a Target Amplicon RNA-seq (TAS) database, we retrieved putative nucleotide and protein sequences of TPS, TPP, and TRE from the sugarcane leaves and culms of field-grown at 1, 3, 6, and 12 months. Simultaneously, expression values for these targets were obtained from 17 other transcriptomes of the SP80-3280 sugarcane variety. Among the expression levels of the different targets, a transcript of *TPS class I* stood out, having the highest expression value among the sequences analyzed. Protein structure and their respective domains were evaluated from the sequences of these targets, and phylogenetic analyses were inferred. Expression values in sugarcane leaves and culms were obtained for eight distinct targets in the Tre6P biosynthesis pathway. For *TPS class I* targets, an inverse expression pattern was observed between leaves and culms, with *TPSI_A* decreasing its expression over development and *TPSI_B* increasing. Identifying the expression profile of these targets enables the identification of important targets that may aid in increasing sucrose content in this agricultural cultivar, thereby impacting its productivity.

Keywords: Biofuel, sugar sensing, Sucrose.

Lista de abreviação

1G: Etanol de primeira geração

2G: Etanol de segunda geração

ACT11: Actina

BECCS: Bioenergia com captura e armazenamento de carbono

cDNA: Ácido dexossiribonucléico complementar

CDS: Coding sequence

CO₂: Dióxido de carbono

DNA: Ácido dexossiribonucléico

GAPDH: Gliceraldeído-3-fosfato desidrogenase

IPCC: Painel Intergovernamental para a Mudança de Clima

PCR: Reação de cadeia da polimerase

RNA: Ácido ribonucléico

RT-qPCR: Reação de cadeia da polimerase em tempo real

TAS: Target Amplicon RNA-seq

TPM: Valores de contagens de reads

TPP: Trealose-6-fosfato fosfatase

TPS: Trealose-6-Fosfato sintase

TRE: Trealase

Tre6P: Trealose-6-Fosfato

UBQ: Ubiquitina

UDP: Difosfato de uridina

UTR: Untranslated gene region

SUMÁRIO

1. INTRODUÇÃO.....	11
1.1 Mudanças climáticas globais	11
1.2 Cana-de-açúcar e bioenergia	12
1.3 Alternativas para maximizar a eficiência da cana-de-açúcar	13
1.4 Açúcares e sua importância nas plantas	14
1.5 Via biossíntese da trealose	15
2. OBJETIVOS.....	18
2.1 Objetivo Geral.....	18
2.2 Objetivos específicos	18
3. MATERIAL E MÉTODOS	18
3.1 Fluxograma de trabalho	18
3.2 Material vegetal	19
3.3 Isolamento de RNA, controle de qualidade e sequenciamento de vias de detecção de açúcar via TAS.....	20
3.4 Alinhamento par-a-par e análise de domínios proteicos	20
3.5 Análise de contagem de <i>Reads</i> em transcriptomas.....	21
3.6 Análise filogenética	21
3.7 Análise de expressão gênica via RT-qPCR	21
3.8 Análise estatística	22
4. RESULTADOS	22
4.1 Sequências recuperadas via TAS.....	22
4.2 Valores de identidade	24
4.3 Estrutura proteica.....	24
4.4 Valores de contagem de <i>reads</i> em bancos de transcriptoma	25
4.5 Análise filogenética	27
4.6. Extração de RNA e síntese de cDNA em amostras de folha e colmo de cana-de-açúcar	30

4.7 RT-qPCR	31
6. DISCUSSÃO.....	36
7. CONCLUSÃO	38
8. REFERÊNCIAS	38

1. INTRODUÇÃO

1.1 Mudanças climáticas globais

O desenvolvimento da vida na Terra ocorreu devido ao aumento da temperatura média global da superfície terrestre, ocasionado principalmente por gases radiativamente ativos, como dióxido de carbono (CO₂), ozônio e pelo vapor d'água (1). No entanto, desde o século XIX observa-se um aumento de até 25% nos níveis de emissão de CO₂ na atmosfera, causado principalmente pela queima de combustíveis fósseis e desmatamento provindos da atividade humana. Por consequência desses fatores, houve um aumento nos níveis de outros gases (metano e outros) que também contribuem para o efeito estufa (1). O *Intergovernmental Panel on Climate Change* (IPCC) estima que estas atividades humanas geraram resultados de aproximadamente 1,0°C de aquecimento global acima dos níveis pré-industriais (1). Caso isso não seja evitado, os efeitos serão catastróficos, podendo resultar no aumento de até 1,5°C na temperatura global entre 2030 e 2052, influenciando diretamente processos biológicos que podem ocasionar perda de culturas agrícolas, aumento da incidência de pragas e agentes patogênicos, estresses abióticos e redução da biodiversidade (2).

O setor de transporte impacta significativamente o consumo de energias não renováveis, sendo responsável por cerca de um terço das emissões totais de CO₂ (3, 4). Estas emissões de CO₂ levam ao desequilíbrio do ciclo do carbono, consequentemente aumentando sua concentração na atmosfera. Neste contexto, a produção de biocombustíveis surgiu como uma fonte viável de energia sustentável, em contraponto a natureza finita dos combustíveis fósseis (5). Estratégias como BECCS (Bioenergia com captura e armazenamento de carbono) visam alcançar as emissões negativas em relação a captura de carbono atmosférico, sendo que esta estratégia é vista como a mais viável e rentável, pois é responsável por gerar simultaneamente energia (através do uso de biomassa) à medida que o CO₂ é capturado da atmosfera por culturas agrícolas (6–9). Assim, a utilização de biomassas como fontes energéticas, não só diminuirá nossa dependência de combustíveis fósseis, como também terá impacto positivo em questões ambientais, além de ajudar a minimizar a produção líquida de gases de efeito estufa (10).

1.2 Cana-de-açúcar e bioenergia

Os biocombustíveis incluem bioprodutos ricos em energias geradas por meio de processos biológicos ou derivado da modificação química da biomassa de organismos vivos, como bactérias fotossintéticas, micro e macroalgas e plantas (5). Atualmente, um dos combustíveis mais sustentáveis é o etanol, o qual é comumente obtido do grão de milho (amido) ou da cana-de-açúcar (sacarose) (11).

A cana-de-açúcar como matéria-prima para a produção de açúcar e álcool é conhecida a quase 500 anos. Em 1930 a cana-de-açúcar se tornou uma matéria prima de produção de bioetanol em nível industrial. A cana-de-açúcar (*Saccharum* spp, Poaceae) é uma gramínea C4 perene, que possui uma alta capacidade de acúmulo de sacarose em seus colmos, no qual 64% do peso seco é armazenado na forma de sacarose e fibras (celulose, hemicelulose e lignina), enquanto o restante (~36%) são de matérias secas como folhas (12). A cana desempenha um papel central na produção de etanol, considerado sustentável, já que seus resíduos lignocelulósicos podem aumentar a produtividade das colheitas sem alterar as áreas plantadas, conseqüentemente fornecendo uma parte valiosa da compensação pelas emissões de CO₂ ocasionados pela queima de combustíveis fósseis (13). O consumo de etanol atualmente é competitivo ao uso de combustíveis derivados de fontes de energia fóssil. Uma das explicações para este é fato se deve ao melhoramento genético convencional da cana-de-açúcar feito ao longo dos anos, responsável pela geração de variedades resistentes as mais diversas condições bióticas e abióticas, aumentando assim a produtividade agrícola desta cultivar (13). Contudo, muitos desafios tecnológicos ainda são vigentes, sobretudo pelo fato de o genoma da cana-de-açúcar ainda não ser totalmente sequenciado, extremamente complexo, poliploide, aneuplóide e interespecífico (14). No Brasil a produção de bioetanol derivada da cana-de-açúcar surgiu na década de 1920 por meio do Instituto de açúcar e do Álcool no qual foi criado com objetivo de ajudar a reger a regulamentação da produção de cana-de-açúcar (15). O Brasil é o maior produtor mundial de cana-de-açúcar, sendo que nas últimas décadas o país atingiu o marco de 100% de aumento desta cultura agrícola (16). Este aporte contribui para as demandas globais de bioenergia, auxiliando na mitigação das mudanças climáticas e reduzindo o consumo de petróleo bruto (17).

A utilização de biomassa para a produção de energia não só reduz a dependência de combustíveis fósseis, mas também impacta positivamente em muitas questões ambientais, ajudando a minimizar a produção de gases do efeito estufa (18). Porém, para ser considerado uma cultura produtora ideal, é necessário que esta não utilize extensas áreas de terras, tenha crescimento rápido, alto balanço energético positivo, e baixo custo de produção de energia a partir da biomassa (19). A cana-de-açúcar possui todas estas características (20), a partir da qual pode ser obtido o etanol de primeira (1G) e segunda (2G) geração.

O bioetanol 1G é produzido a partir da fermentação alcoólica da sacarose ou açúcares simples, obtidos de diferentes fontes vegetais, como por exemplo a cana-de-açúcar (21). No entanto, a necessidade de aporte hídrico das plantações canavieiras e o ciclo de colheita dessa cultura é superior a outras culturas agrícolas, aumentando os gastos de recursos e dificultando a produção de etanol 1G (18). Já o etanol 2G provém da biomassa lignocelulósica não utilizada para a produção de 1G (12). Todavia, algumas das propriedades da biomassa utilizada (como a presença de lignina) dificulta a liberação de açúcares fermentescíveis, prejudicando o rendimento durante os processos de fermentação e conversão em etanol (18,22). Na cana-de-açúcar, algumas etapas do pré-tratamento podem levar a perda das frações mais solúveis das paredes celulares (beta-glucano e pectinas), causando a perda de uma proporção substancial dos polissacarídeos, diminuindo assim seu potencial (18).

1.3 Alternativas para maximizar a eficiência da cana-de-açúcar

Por conta dessas dificuldades, alternativas para maximizar a eficiência da conversão de biomassa foram criadas em busca não só de melhorar o seu alto rendimento e o teor de fibras, mas também o incremento de tecnologias de pré-tratamento e digestão enzimática e degradabilidade da biomassa para conversão em biocombustível (23). No Brasil, diversas estratégias relacionadas a características fisiológicas e manipulações moleculares tem sido proposta para que haja um aumento no rendimento da cana-de-açúcar em seu potencial energético (18,19). Estudos feitos por DE SOUZA e colaboradores (2013) demonstraram que ao cultivar a cana-de-açúcar sob elevada concentração atmosférica de CO₂ (o dobro da concentração normal), houve um aumento de sua fotossíntese em cerca de 30%, que conseqüentemente levou ao aumento de sua biomassa (40%) e acúmulo de sacarose

(29%). O maior desempenho da planta em CO₂ elevado foi atribuído a uma maior taxa de transporte de elétrons do sistema fotossintético de acordo com medições de níveis de expressão gênica. Este aumento de captura de CO₂ em condições em que o CO₂ estava elevado, fez com que houvesse um aumento no uso eficiente da água que, por consequência, proporciona uma melhor resposta a cana-de-açúcar em relação aos impactos negativos causados por estresse hídrico.

Estas estratégias para maximizar o rendimento da cana-de-açúcar são, em parte, associados ao entendimento das dinâmicas metabólicas dos carboidratos e de sua partição e sinalização atrelada ao longo do desenvolvimento das plantas (24,25). Como uma gramínea C₄, a cana-de-açúcar fixa o CO₂ atmosférico mais eficientemente, resultando no acúmulo de grandes quantidades de sacarose em seus colmos. Desta forma, o estudo do metabolismo de carboidratos nesta espécie torna-se imperativo, uma vez que através do conhecimento gerado acerca do acúmulo e sinalização destes açúcares, pode-se gerar meios de aumentar a produtividade da cana-de-açúcar com impacto direto na produção de etanol 1G.

1.4 Açúcares e sua importância nas plantas

Respostas adaptativas das plantas a estresses bióticos e abióticos foram adquiridas ao longo da evolução com a finalidade de garantir a correta assimilação do carbono necessário para o crescimento e desenvolvimento (26). Enquanto organismos sésseis, as plantas desenvolveram ao longo dos anos mecanismos de sensoramento e detecção, tendo assim a capacidade de detectar mudanças nos níveis de açúcares, e responder apropriadamente por meio de uma complexa rede de sinalização, afim de lidar com as variações ambientais, proporcionando assim respostas a disponibilidade de recursos para manter-se no seu equilíbrio energético (27,28). A partir destas respostas, uma rede complexa de sinalização foi estabelecida tendo por base a detecção da variação nos níveis de diferentes açúcares, direcionando modulações na expressão gênica e na atividade de proteínas que possibilitem a adaptação aos mais variados estímulos ambientais (26).

Nas plantas, os açúcares estão presentes no centro de seu metabolismo, sobretudo a sacarose. Durante seu processo fotossintético, as plantas sintetizam sacarose que são armazenadas transitoriamente em seus vacúolos, com a finalidade de ser usada em atividades celulares e sair da fonte para os tecidos dreno ajudando

na sustentação do metabolismo e no crescimento da própria planta (29). Além da sacarose, o amido também desempenha um papel crucial nas plantas, o qual pode ser armazenado no cloroplasto das folhas e degradado na fase noturna em um ciclo diuturno, atendendo a demanda contínua de carbono, e assim regulando os níveis de açúcar em escala temporais e espaciais (30).

Os açúcares além de funcionar como substrato para o crescimento e metabolismo energético, também afetam os sistemas de detecção de açúcares que regulam como, quando e onde estes são utilizados, além de desempenhar um papel fundamental como moléculas sinalizadoras (27). Assim sendo, ter um equilíbrio adequado entre a assimilação, participação e uso do carbono é essencial para o desenvolvimento, sucesso reprodutivo e sobrevivência da planta (31). Nas plantas, diferentes sinais de açúcar são gerados em resposta aos estímulos fotossintéticos e do metabolismo do carbono, tanto nos tecidos fonte quanto nos tecidos dreno, com a finalidade de modular o crescimento, desenvolvimento e as respostas ao estresse (27). Neste contexto, o açúcar trealose-6-fosfato (Tre6P), intermediário da biossíntese da trealose, tem um papel de sinalizador específico da sacarose nas plantas, influenciando o acúmulo deste açúcar e os padrões de crescimento e desenvolvimento vegetal (32).

1.5 Via biossíntese da trealose

A trealose é um dissacarídeo não redutor que pode ser utilizado como reserva, açúcar de transporte e osmoprotetor contra estresses bióticos e abióticos (33). A biossíntese de trealose acontece por duas etapas catalíticas (**Figura 1**): (1) na primeira etapa, o intermediário trealose-6-fosfato (Tre6P) é sintetizado a partir da UDP-glicose e glicose-6-fosfato pela trealose-6-fosfato sintase (TPS; EC 2.4.1.15); (2) logo em seguida, o grupo fosfato é hidrolisado pela trealose-6-fosfato fosfatase (TPP; EC 3.1.3.12), produzindo trealose livre; (3) a qual pode ser hidrolisada pela trealase (TRE; EC 3.2.1.28), formando duas glicoses (34). A grande maioria das isoformas de TPS possui função regulatória e não-catalítica, as quais respondem rapidamente aos níveis de açúcar e ao ciclo diuturno das plantas (35). Já as isoformas TPP possuem função catalítica e sua expressão responde a mudanças ambientais e estresses abióticos (36,37). Há registros claros em que houve eventos independentes de duplicação gênica, tanto para *TPS* quanto para *TPP*, em várias linhagens de plantas

(38–40). Em *Arabidopsis thaliana*, a *TPS* se agrupa em duas subfamílias distintas designadas classe I (*AtTPSI-4*) e classe II (*AtTPS5-11*) que provavelmente se originaram de um ancestral comum (36,41). Estes dois clados estão presentes em todos os principais táxons de plantas (38).

A Tre6P é um metabólito sinalizador que regula o metabolismo da sacarose em plantas (38). Além de ser um intermediário da biossíntese da trealose, também possui um papel de sinalização fundamental nas plantas, estabelecendo uma conexão entre o crescimento e o desenvolvimento ao status de carbono disponível (38,42). A Tre6P tem dupla função em relação ao sinal e ao regulador homeostático dos níveis de sacarose na planta. Nas folhas fotossinteticamente ativas, a Tre6P atua como regulador na produção de sacarose, com o objetivo de equilibrar a oferta e demanda deste açúcar nos órgãos dreno ao longo do crescimento vegetal (**Figura 1**) (42). A Tre6P é fundamental em diferentes etapas do ciclo de vida das plantas, como na embriogênese, crescimento, desenvolvimento e floração (32,42), sendo apresentada como um sensor específico de sacarose que ajuda a regular o metabolismo vegetal (43,44). Este açúcar também afeta a degradação do amido no período noturno, consequentemente fazendo parte do mecanismo de estabilidade dos níveis de sacarose em uma faixa adequada para ser utilizado como suprimento para o crescimento celular e atividades metabólicas (44,45). Além disso a Tre6P foi proposta como um sinal de disponibilidade de sacarose que também tem como função ser um regulador de *feedback* negativo sobre os níveis de sacarose, com objetivo de mantê-los dentro de uma faixa adequada de acordo com o estado metabólico celular e o estágio de desenvolvimento da planta (43,44). Porém essa função não está clara em espécies como cana-de-açúcar, pois elas acumulam grandes quantidades de sacarose em seus colmos.

Linhagens de cana-de-açúcar transgênicas que superexpressam *TPS* e *TPP* mostraram atividade aumentada de TRE e correspondente biossíntese de trealose, mas não tiveram um incremento em açúcares como sacarose e glicose (46). Por outro lado, a superexpressão independente dos genes de *Escherichia coli* *otsA* (*TPS*) e *otsB* (*TPP*) em cana-de-açúcar teve redução e aumento dos níveis de sacarose nas respectivas plantas transgênicas (47), porém os níveis de Tre6P mantiveram-se inalterados. Estes resultados demonstram que a manipulação do metabolismo de Tre6P/trealose tem o potencial de modificar o perfil de açúcares solúveis acumulados

em colmos de cana-de-açúcar. Porém, o alto grau de ploidia e aneuploidia intrínseco da cana-de-açúcar pode constituir uma problemática para a manipulação genética desta cultivar.

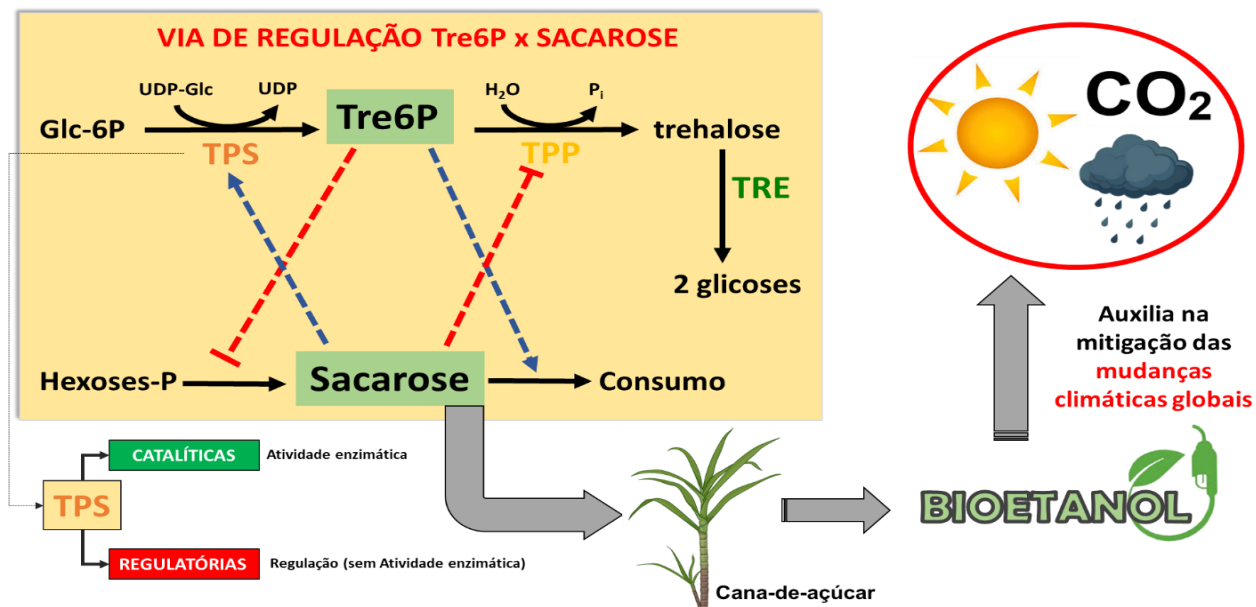


Figura 1. Regulação dos níveis de sacarose via Trealose-6-fosfato (Tre6P) em cana-de-açúcar. A Tre6P é um metabólito sintetizado na via da trealose pela enzima Trealose-6-fosfato sintase (TPS) e degradada pela enzima Trealose-6-fosfato fosfatase (TPP). Os níveis de Tre6P regulam homeostaticamente os níveis de sacarose. O aumento dos níveis de Tre6P inibe a conversão de hexoses-P em sacarose, além de induzir o seu consumo. De forma homeostática, o aumento nos níveis de sacarose induz uma maior transcrição da TPS além de reduzir a TPP. Desta forma, em cultivares que acumulam altas quantidades de sacarose, como a cana-de-açúcar, é importante saber como ocorre esse nível de regulação durante seu desenvolvimento, uma vez que esta cultivar agrícola é importante para a produção de bioetanol, a qual auxilia na mitigação dos efeitos das mudanças climáticas. Setas pontilhadas em azul indicam a possível ativação e em vermelho uma provável inibição.

Neste sentido, DE OLIVEIRA e colaboradores (2022) analisaram de forma abrangente sequências de enzimas do metabolismo da trealose em cana-de-açúcar, com a finalidade de estabelecer uma base de estudos funcionais para elucidar os papéis fisiológicos da sinalização de Tre6P na cana-de-açúcar. Foram recuperadas 71 sequências putativas de *TPS*, 93 de *TPP* e 3 *TRE* de cana-de-açúcar, mostrando todos os seus domínios conservados e suas respectivas relações filogenéticas. A partir destes dados, um painel do tipo *Targeted Amplicon RNA-Seq* (TAS) foi construído e sequenciado, o qual contém um conjunto selecionado de genes-chave relacionados ao sensoriamento de açúcar em cana-de-açúcar. Desta forma, neste

projeto pretende-se avaliar a expressão diferencial das isoformas de *TPS*, *TPP* e *TRE* identificadas por DE OLIVEIRA et al. (2022), contrastando estes níveis de expressão com os obtidos via TAS.

2. OBJETIVOS

2.1 Objetivo Geral

Determinar as diferenças de expressão entre sequências de *Trealose-6-fosfato sintase (TPS)*, *Trealose-6-fosfato fosfatase (TPP)* e *Trealase (TRE)* de cana-de-açúcar.

2.2 Objetivos específicos

- Determinar a estrutura proteica e os diferentes domínios para as sequências de *TPS*, *TPP* e *TRE* de cana-de-açúcar;
- Obter iniciadores para amplificação diferencial de cada cópia entre as sequências de *TPS*, *TPP* e *TRE* da variedade SP80-3280 de cana-de-açúcar;
- Extrair RNA e sintetizar cDNA de colmos e folhas da variedade SP80-3280 de cana-de-açúcar;
- Analisar os níveis de expressão via RT-qPCR das sequências de *TPS* catalíticas da variedade SP80-3280 de cana-de-açúcar;
- Correlacionar e validar a expressão dos mesmos alvos no banco de dados do *Targeted Amplicon RNA-Seq*.

3. MATERIAL E MÉTODOS

3.1 Fluxograma de trabalho

O fluxograma realizado neste trabalho está representado na **Figura 2**. Inicialmente, transcritos de cana-de-açúcar para *trealose-6-fosfato sintase (TPS class I e class II)*, *trealose-6-fosfato fosfatase (TPP)* e *trealase (TRE)* foram identificadas e recuperadas no banco *Targeted Amplicon RNA-Seq (TAS)*. A partir destas sequências, foram identificadas a presença de regiões UTRs (cDNA) e/ou codificantes (CDS). Concomitantemente, o nível de expressão de transcritos homólogos destes alvos foi avaliado em diferentes bancos de transcriptoma de cana-de-açúcar. Também foi realizada uma análise de domínios nas proteínas preditas para cada sequência recuperada, assim como um alinhamento par-a-par foi realizado com a finalidade de

se detectar o valor de identidade entre as sequências, e conseqüentemente, remover as possíveis redundâncias. Estas sequências foram então utilizadas para a construção de árvores filogenéticas para cada alvo estudado neste projeto. Após a validação por relação filogenética, iniciadores específicos foram desenhados e testados quanto à amplificação via PCR. A partir de iniciadores de qualidade, foi realizado um experimento de RT-qPCR, em que foram obtidos os valores de expressão para os alvos estudados. Por fim, estes valores foram correlacionados aos valores de expressão encontrados no banco TAS.

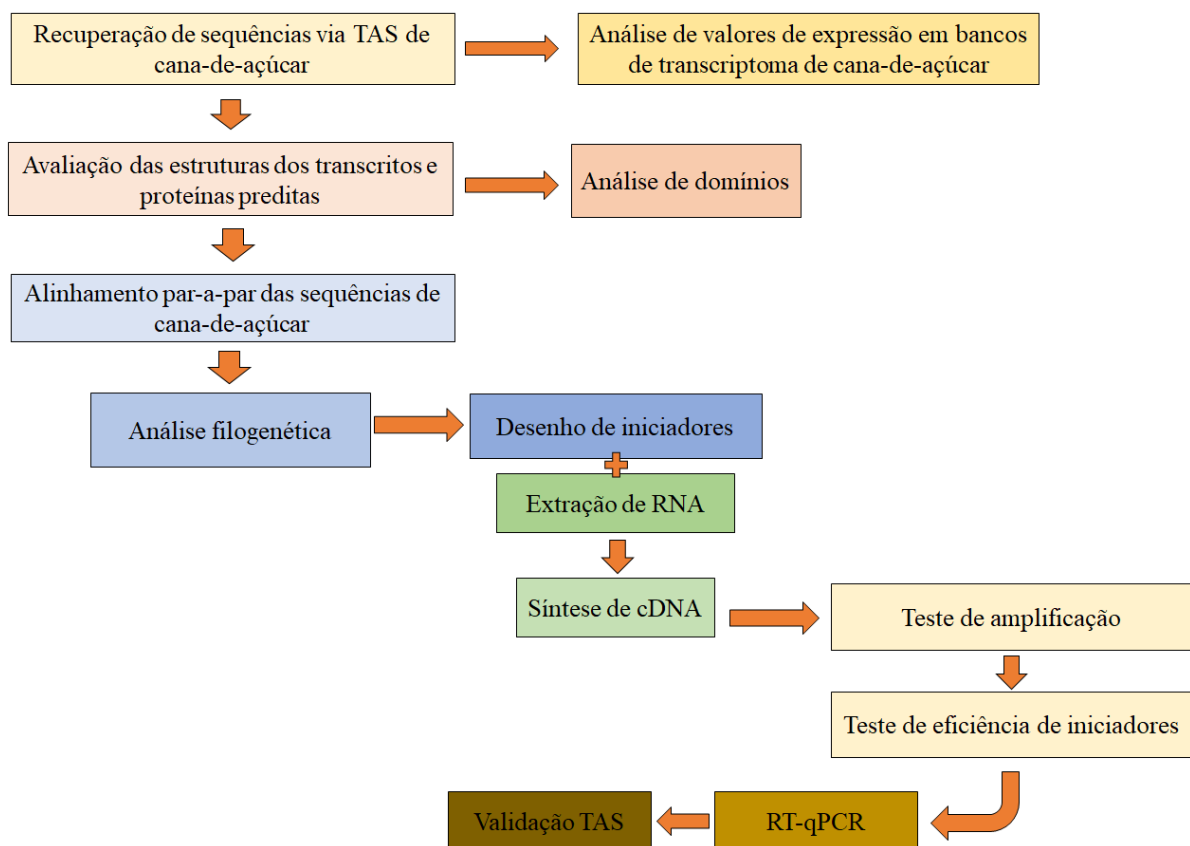


Figura 2. Fluxo de trabalho das etapas planejadas para execução e desenvolvimento do presente estudo.

3.2 Material vegetal

O material vegetal utilizado na construção do banco *Targeted Amplicon RNA-Seq* e para as análises de RT-qPCR foi obtido a partir de plantas de *Saccharum* spp. (variedade SP80–3280) cultivadas no campo em Piracicaba, São Paulo, Brasil e colhidas durante o segundo ciclo de soca (31,48). Folhas verdes totalmente expandidas (folha +1) e colmos (entrenós representando a porção madura) foram

coletados às 10h da manhã após 1, 3, 6 e 12 meses de desenvolvimento da cana-de-açúcar. Posteriormente, este material foi conservado em nitrogênio líquido, moído em pó e armazenado a -80°C até o uso.

3.3 Isolamento de RNA, controle de qualidade e sequenciamento de vias de detecção de açúcar via TAS

O RNA foi extraído utilizando-se trizol seguido de purificação utilizando-se PureLink RNA Mini Kit (ThermoFisher), seguindo o protocolo estabelecido por De Oliveira (2022). Foram utilizadas três réplicas biológicas, resultando em 72 amostras. As amostras de RNA foram quantificadas através do NanoDrop (Thermo Scientific, Wilmington, DE, EUA). Amostras com proporções de 260/280 nm e 260/230 nm entre 1,8–2,2 e 1,6–2,2, respectivamente, foram consideradas de pureza suficiente. A integridade do RNA total (1 µg) foi ainda confirmada por eletroforese em gel de agarose 1% (p/v), usando um Bioanalyzer 2100 (Agilent Technologies Inc., Santa Clara, CA, EUA). A maioria das amostras tinha número de integridade de RNA ≥ 6 e todas foram armazenadas a -80 °C até processamento posterior. Em seguida, um painel *AmpliSeq* personalizado baseado nas sequências de *Saccharum* spp. foi projetado e desenvolvido pela Illumina a nosso pedido. Este painel continha 708 conjuntos de primers para cobrir as sequências de vias de detecção de açúcar. As bibliotecas personalizadas *AmpliSeq* foram preparadas seguindo o protocolo fornecido pela empresa (*AmpliSeq for Illumina On-Demand, Custom e Community Panels*, Document # 1000000036408 v09 May 2020) com algumas modificações. Para cada biblioteca, sequências *paired-end* de 150 pb foram geradas por meio da plataforma Illumina MiSeq seguindo o manual do fabricante.

3.4 Alinhamento par-a-par e análise de domínios proteicos

Foi realizado um alinhamento *ClustalW* par-a-par através da ferramenta *needleall* separadamente para as sequências de cada alvo analisado. A partir deste alinhamento, foram recuperados os arquivos mostrando as bases alinhadas, assim como o valor de similaridade para cada combinatória. Para verificar se as sequências de TPS, TPP e TRE de cana-de-açúcar possuem os seus respectivos domínios proteicos preditos (glicosiltransferase/trealose fosfato sintase e trealose fosfatase) análises de domínios foram realizadas através da plataforma HMMER Scan v2.41.2.

3.5 Análise de contagem de Reads em transcriptomas

Os níveis de expressão de *TPS*, *TPP* e *TRE* foram recuperados de um conjunto de dados de transcriptomas publicados para a variedade SP80-3280 (49,50). Em MATIELLO e colaboradores (2015) diferentes regiões de desenvolvimento ao longo da folha +1 de plântulas de dois meses de idade foram avaliadas, no entanto este estudo teve como foco amostras da porção intermediária (4 réplicas biológicas), por ser a região com maior capacidade fotossinteticamente ativa (acesso NCBI Short Read Archive numbers: SRR1979669, SRR1979665, SRR1979662 e SRR1979660). As leituras de sequência bruta foram baixadas do SRA do NCBI e limpas com BBDuk2 (51), para remover restos de rRNA e regiões de baixa qualidade, bem como adaptadores. Salmon v1.1.0 (52) foi usado para estimar os níveis de expressão de transcrição expressos como Transcripts per Million (TPM) transformados em (LOG₁₀+1).

3.6 Análise filogenética

Um alinhamento global do tipo MUSCLE das sequências proteicas de *TPS*, *TPP* e *TRE* foram realizados a partir do programa MEGA (versão 11) (53). Para inferências filogenéticas no IQ-TREE (54), foram utilizadas a seleção automática de modelo evolutivo e a aproximação ultra-rápida por *bootstrap* (UFboot), com valores de suporte de ramo aprimorados pelo teste de razão de verossimilhança aproximada (SH-aLRT), (55).

3.7 Análise de expressão gênica via RT-qPCR

Os iniciadores foram desenhados para os alvos de acordo com as recomendações de VARKONYI-GASIC e colaboradores (2007), utilizando o programa Oligo Perfect 3.1 (idtdna.com/calc/analyzer). A transcriptase reversa foi realizada utilizando *SuperScript III First – Strand Synthesis System* (ThermoFischer Scientific) com 2 µg de RNA total previamente tratado com RNase. Os cDNAs foram diluídos e posteriormente usados para análises de RT-qPCR utilizando-se Power SYBR green. Os valores de C_q de duas réplicas técnicas e a eficiência dos iniciadores foram calculados usando o software LinRegPCR. Os valores de expressão dos genes alvos

foram normalizados contra a média geométrica dos genes de referência *Gliceraldeído-3-fosfato desidrogenase (GAPDH)*, actina (*ACT11*) e ubiquitina (*UBQ*).

3.8 Análise estatística

Ao menos três réplicas foram usadas para os experimentos. Quando necessário, os dados foram submetidos a Análise de Variância (ANOVA) seguido por teste Tukey ($p < 0,05$). Também foram realizadas comparações par-a-par pelo teste t de Student. Os valores de correlação de Pearson foram obtidos via Metaboanalyst 5.0 (PANG et al., 2021). Quando apropriado, os dados foram transformados pela função Log10. As análises foram realizadas através de script R versão 3.2.2 (R Core Team, 2019).

4. RESULTADOS

4.1 Sequências recuperadas via TAS

Com o objetivo de se avaliar as diferenças de expressão gênica de alvos relacionados aos sensores de açúcar, foi construído um banco de dados do tipo *Targeted Amplicon RNA-Seq* para cana-de-açúcar cultivada a campo. Neste banco foi analisada amostras de folha e de colmo durante o ciclo diurno e durante o ciclo de colheita da cana (1, 3, 6 e 12 meses). Entre os alvos avaliados estão sequências da via Trealose-6-fosfato. Com o objetivo de se aprofundar os conhecimentos acerca desta via, este projeto determinou o valor de expressão destes alvos via RT-qPCR para serem futuramente contrastados com o banco de dados TAS. Contudo, inicialmente se fazem necessárias análises mais profundas das sequências destes alvos em cana-de-açúcar, uma vez que cana-de-açúcar é um organismo poliploide e não tem seu genoma completamente sequenciado.

Como passo inicial, foi recuperado da literatura sequências iscas de espécies modelo com a função confirmada para os alvos da via da trealose-6-fosfato (**Tabela 1**). Os alvos analisados foram *Trealose-6-fosfato sintase (TPS)*, *Trealose-6-fosfato fosfatase (TPP)* e *Trealase (TRE)*. Para esses três alvos foram recuperadas sequências de *Arabidopsis thaliana*, *Zea mays* e *Oryza sativa*. Assim, foram recuperadas sequências homólogas a estas em cana-de-açúcar que foram utilizadas na construção do TAS. A partir deste banco, foram recuperadas 42 sequências as

quais tiveram as regiões UTR e CDS do transcrito identificadas, assim como as respectivas sequências proteica também foram obtidas.

Tabela 1. Alvos do metabolismo de Trealose-6-fosfato de *Arabidopsis thaliana*, *Oryza sativa* e *Zea mays* utilizados como iscas na identificação de genes ortólogos em Viridiplantae através do EggNOG database.

Descrição do gene/proteína	Clado*	Gene	Loci	EggNOG v4.5.1	Espécie / Database / Referência	
Trehalose-6-phosphatase synthase (TPS)	Class I - B1	<i>AtTPS1</i>	At1g78580	1DVBM	<i>Arabidopsis thaliana</i> Phytozome 12.1.6 Henry et al., 2014	
	Class I - B2	<i>AtTPS2</i>	At1g16980	1EDRK		
		<i>AtTPS3</i>	At1g17000			
		<i>AtTPS4</i>	At4g27550	1EDRH		
	Class II - A1	<i>AtTPS5</i>	At4g17770	1DQAR		
	Class II - A2	<i>AtTPS6</i>	At1g68020			
	Class II - A3	<i>AtTPS7</i>	At1g06410			
	Class II - A4	<i>AtTPS8</i>	At1g70290			
		<i>AtTPS9</i>	At1g23870			
	Class II - A5	<i>AtTPS10</i>	At1g60140			
	Class II - A5	<i>AtTPS11</i>	At2g18700	1DSVD		
	Class I - B1	<i>OsTPS1</i>	LOC_Os05g44210	1DVBM	<i>Oryza sativa</i> Phytozome 12.1.6 Henry et al., 2014	
	Class II - A3	<i>OsTPS2</i>	LOC_Os01g54560	1DQAR		
		<i>OsTPS3</i>	LOC_Os01g53000			
	Class II - A2	<i>OsTPS4</i>	LOC_Os03g12360	1DZIJ		
	Class II - A4	<i>OsTPS5</i>	LOC_Os02g54820	1DSVD		
	Class II - A3	<i>OsTPS6</i>	LOC_Os05g44100	1DQAR		
	Class II - A4	<i>OsTPS7</i>	LOC_Os08g31980	1E175		
	Class II - A5	<i>OsTPS8</i>	LOC_Os08g34580	1DQDD		
		<i>OsTPS9</i>	LOC_Os09g25890			
Class II - A2	<i>OsTPS10</i>	LOC_Os09g23350	1DZIJ			
Class II - A4	<i>OsTPS11</i>	LOC_Os09g20990	1E175			
Trehalose-6-phosphatase (TPP)	Class I - B1	<i>ZmTPSI.1.1_tps1</i>	Zm00001d010755_P003	1DVBM	<i>Zea mays</i> Phytozome 12.1.6 Henry et al., 2014	
		<i>ZmTPSI.1.2</i>	Zm00001d038728_P013			
	Class II - A2	<i>ZmTPSII.2.1</i>	Zm00001d020396_P004	1DQAR		
		<i>ZmTPSII.2.2</i>	Zm00001d028267_P010	1DZIJ		
	Class II - A3	<i>ZmTPSII.3.1</i>	Zm00001d043468_P007	1DQAR		
		<i>ZmTPSII.3.2</i>	Zm00001d043469_P002			
		<i>ZmTPSII.3.3</i>	Zm00001d012748_P011			
	Class II - A4	<i>ZmTPSII.4.1</i>	Zm00001d005687_P003	1E175		
		<i>ZmTPSII.4.2</i>	Zm00001d032311_P006			
		<i>ZmTPSII.4.3</i>	Zm00001d050293_P005			
	Class II - A5	<i>ZmTPSII.5.1</i>	Zm00001d050069_P001	1DQDD		
		<i>ZmTPSII.5.2</i>	Zm00001d032118_P001			
		<i>ZmTPSII.5.3</i>	Zm00001d052060_P005			
		<i>ZmTPSII.5.4</i>	Zm00001d018342_P002			
-	<i>ZmTPS13.1</i>	Zm00001d008338_P001	1EKVF			
Trehalose-6-phosphatase phosphatase (TPP)	B1	<i>AtTPPC</i>	At1g22210	1EKDT	<i>Arabidopsis thaliana</i> Phytozome 12.1.6 Henry et al., 2014	
		<i>AtTPPD</i>	At1g35910			
		<i>AtTPPB</i>	At1g78090			
		<i>AtTPPE</i>	At2g22190			
		A1	<i>AtTPPF</i>			At4g12430
			<i>AtTPPG</i>			At4g22590
	B1	<i>AtTPPH</i>	At4g39770	1EKDT		
		<i>AtTPPI</i>	At5g10100			
	A1	<i>AtTPPA</i>	At5g51460			
	B1	<i>AtTPPJ</i>	At5g65140			
	A3	<i>OsTPP1</i>	LOC_Os02G44230	1EKDT	<i>Oryza sativa</i> Phytozome 12.1.6 Henry et al., 2014	
	A1	<i>OsTPP2</i>	LOC_Os10G40550	1EKDN		
	B1	<i>OsTPP3</i>	LOC_Os07G43160	1EKDR		
		<i>OsTPP4</i>	LOC_Os02G51680	1EKDR		
	A3	<i>OsTPP5</i>	LOC_Os04G46760	1EKDT		
		<i>OsTPP6</i>	LOC_Os08G31630			
	B1	<i>OsTPP7</i>	LOC_Os09G20390	1EKDR		
		<i>OsTPP8</i>	LOC_Os06G11840	1EKDR		
		<i>OsTPP9</i>	LOC_Os03G26910	1EKDP		
	A2	<i>OsTPP10</i>	LOC_Os07G30160	1EKDS		
A3	<i>OsTPP11</i>	LOC_Os02G44235	1EKDT			
A1	<i>OsTPP12</i>	LOC_Os10G40555				
A3	<i>OsTPP13</i>	LOC_Os12G09060				
A1	<i>ZmTPPA.1</i>	Zm00001d047110_P001	1EKDT	<i>Zea mays</i> Phytozome 12.1.6 Henry et al., 2014		
A3	<i>ZmTPPA.3</i>	Zm00001d017502_P002				
	<i>ZmTPPB.1.1</i>	Zm00001d032298_P001				
	<i>ZmTPPB.1.2</i>	Zm00001d005658_P001				
B1	<i>ZmTPPB.1.3</i>	Zm00001d020272_P001			1EKDR	
	<i>ZmTPPB.1.4</i>	Zm00001d018082_P001				
	<i>ZmTPPB.1.5</i>	Zm00001d052227_P001				
	<i>ZmTPPB.1.6</i>	Zm00001d044854_P003				
	<i>ZmTPPB.2.1_ramora3</i>	Zm00001d022193_P001				
B2	<i>ZmTPPB.2.2</i>	Zm00001d006913_P002			1EKDN	
	<i>ZmTPPB.2.3</i>	Zm00001d022192_P002				
Trehalase	-	<i>AtTRE</i>	AT4G24040	1DYMW	<i>Arabidopsis thaliana</i> Phytozome 12.1.6	
	-	<i>OsTRE</i>	LOC_Os10g37660.1	1DYMW	<i>Oryza sativa</i> MSU Rice Genome Annotation Project V.7	
	-	<i>ZmTRE</i>	Zm00001d029654_P001 Zm00001d029654_P002 Zm00001d029654_P003 Zm00001d029654_P004 Zm00001d029654_P005 Zm00001d029654_P006	1DYMW	<i>Zea mays</i> Phytozome 12.1.6 Zhou et al., 2014***	

* Os cladros são baseados em Henry et al., 2014.

** Sequências homólogas obtidas no GRAM1

***Zhou, M.; et al. Trehalose Metabolism-Related Genes in Maize. *Journal of Plant Growth Regulation*, 2014, 33, 256-271. 10.1007/s00344-013-9368-y.

4.2 Valores de identidade

Para avaliar a similaridade entre as sequências recuperadas do TAS buscou-se realizar um alinhamento do tipo par-a-par, a partir do qual foi recuperado o valor de identidade entre estas sequências. Para tanto, foi realizado um alinhamento do tipo *ClustalW* para cada sequência dentro de cada grupo descrito (TPS classe I, TPS classe II, TPP, TRE). No geral, 42 sequências foram avaliadas sendo 3 para TPS classe I, 24 para TPS classe II, 12 para TPP e 3 para TRE (**Tabela 2**). Quando avaliado o valor de identidade, observou-se que as sequências de TPS classe I foram as que tiveram os maiores valores globais de identidade com uma mediana em torno de 61.4%.

Tabela 2. Valores de identidade de sequências analisadas via alinhamento par-a-par para diferentes alvos do metabolismo de Trealose-6-fosfato.

Descrição	Sequências analisadas	Identidade (%)			Sequências remanescentes
		Maior	Menor	Mediana	
TPS Classe I	3	72.3	53.0	61.4	2
TPS Classe II	24	97.6	12.8	39.5	13
TPP	12	91.4	15.1	39.7	9
TRE	3	86.3	39.3	41.0	3
<i>Total</i>	<i>42</i>				<i>27</i>

Dentre as três sequências de TPS classe I analisadas, duas foram identificadas como redundantes já que o alinhamento par-a-par entre essas duas sequências mostrava blocos de alinhamento muito contínuos. Dentre as três sequências analisadas, duas se mantiveram nas análises posteriores. Para TPS classe II, o maior valor de identidade encontrado foi de 97.6% e o menor foi de 12.8%. Entre as 24 sequências avaliadas, 13 se mantiveram nas análises subsequentes. Para TPP, 12 sequências foram analisadas e nove continuaram nas análises remanescentes, enquanto para TRE, três sequências foram analisadas e estas foram identificadas como sequências distintas. No total, 27 sequências foram mantidas para análises posteriores.

4.3 Estrutura proteica

Com a finalidade de avaliar a estrutura das proteínas preditas mantidas para TPS, TPP e TRE de cana, foi analisado via HAMMERscan a presença dos domínios

dessas sequências e a anotação referente destes domínios (**Tabela 3**). Para TPS classe I e TPS classe II, todas as sequências mostraram domínios do tipo glycosyltransferase família 20 e Trealose fosfatase. O domínio glycosyltransferase família 20 realiza a catálise da Glicose-6-fosfato e UDP-glicose, formando a trealose-6-fosfato. Para TPP foi observado a presença do domínio Trealose fosfatase responsável pela clivagem do fosfato da Tre6P formando o produto Trealose. Para TRE foi encontrado o domínio Trealase, um domínio característico de enzimas hidrolíticas na quebra da Trealose, formando dois monômeros de glicose.

4.4 Valores de contagem de *reads* em bancos de transcriptoma

A fim de avaliar o valor de expressão dos alvos da via de Tre6P analisados nesse projeto, foram recuperados os valores de contagens de *reads* (TPM) de diferentes bancos de dados para a variedade SP80-3280 de cana-de-açúcar (**Figura 3**). Estes bancos de dados englobam dois trabalhos: (1) MATIELLO et al. (2015), que avalia a expressão de transcritos em diferentes regiões da folha +1 da cana crescida a campo; e (2) VETTORE et al. (2003), o qual foi utilizado como base para a construção do banco SUCEST.

Na **figura 3** é possível observar os níveis de expressão dos diferentes alvos de TPS classe I, TPS classe II, TPP e TRE ao longo dos 17 bancos de dados analisados, sendo que o banco de dados de DB1 se refere a MATIELLO et al. (2015) e os demais bancos de dados a VETTORE et al. (2003). É possível observar que uma sequência de TPS classe I é a que possui o maior valor de expressão entre as sequências analisadas. Por outro lado, sequências de TPS classe II possui um valor de expressão médio, enquanto TPP e TRE globalmente apresentam valores menores. A comparação entre os valores expressão foram feitas entre os valores internos de cada banco.

Tabela 3. Sequências de cana-de-açúcar (SP80-3280) utilizadas para a construção das árvores filogenéticas, seus respectivos domínios proteicos e anotação de domínios via HMMERscan.

EggNOG v.4.5.1	Identificação de seqüências	Domínios	Predição de sítio ativo (PFAM)	Família	Início do alinhamento	Final do alinhamento	Início do modelo	Final do modelo	Tamanho	Bit score	Gathering cut-off domain - PFAM	Cobertura com base no modelo (80%)	Total de aminoácidos
IDVBM - TPS class I	SCA3 SP803280_c107577_g2_i5_m.97798	Glycosyltransferase family 20 Trehalose-phosphatase		PF00982.21 PF02358.16	129 653	594 734	2 15	473 96	474 233	615.76 63.42	25.5 20.2	99.37 34.76	737
IDVBM - TPS class I	SCA1_2 Sh_224G15_p000010	Glycosyltransferase family 20 Trehalose-phosphatase		PF00982.21 PF02358.16	190 394	335 595	328 15	473 208	474 233	208.58 181.21	25.5 20.2	30.59 82.83	716
IDQAR - TPS class II	SCA3_SP803280_c95707_g1_i1_m.53094	Glycosyltransferase family 20 Trehalose-phosphatase	D587	PF00982.21 PF02358.16	52 585	535 818	2 1	473 231	474 233	570.02 251.95	25.5 20.2	99.37 98.71	853
IDQARr - TPS class II	SCA2_1_c45366f1p12001.p1	Glycosyltransferase family 20 Trehalose-phosphatase	D315	PF00982.21 PF02358.16	4 313	263 546	227 1	473 231	474 233	284.47 253.06	25.5 20.2	51.90 98.71	581
IDQAR - TPS class II	SCA2_1_c94027f1p03407.p1	Glycosyltransferase family 20 Trehalose-phosphatase	D581	PF00982.21 PF02358.16	34 579	529 812	2 1	473 231	474 233	574.50 249.63	25.5 20.2	99.37 98.71	863
IDQAR - TPS class II	SCA2_1_c95057f1p13337.p1	Glycosyltransferase family 20 Trehalose-phosphatase	D593	PF00982.21 PF02358.16	43 591	541 824	2 1	473 231	474 233	586.21 244.90	25.5 20.2	99.37 98.71	875
IDQAR - TPS class II	SCA1_2_Sh_205A13_p000010	Glycosyltransferase family 20 Trehalose-phosphatase	D611	PF00982.21 PF02358.16	67 609	559 843	2 1	473 232	474 233	593.64 247.49	25.5 20.2	99.37 99.14	878
IDQAR - TPS class II	SCA5_SCEZRZ1015H08.g.p1	Glycosyltransferase family 20 Trehalose-phosphatase	D184	PF00982.21 PF02358.16	6 182	132 271	359 1	473 91	474 233	153.91 90.18	25.5 20.2	24.05 38.63	272
IDSVD - TPS class II	SCA2_1_c87041f1p13312.p2	Glycosyltransferase family 20 Trehalose-phosphatase	D346	PF00982.21 PF02358.16	2 344	294 504	190 1	473 161	474 233	315.79 183.02	25.5 20.2	59.70 68.67	579
IDSVD - TPS class II	SCA1_2_Sh_245I08_p000020	Glycosyltransferase family 20 Trehalose-phosphatase	D584	PF00982.21 PF02358.16	50 582	532 742	2 1	473 161	474 233	524.96 182.25	25.5 20.2	99.37 68.67	777
IDZIJ - TPS class II	SCA3 SP803280_c117830_g1_i1_m.154722	Glycosyltransferase family 20 Trehalose-phosphatase	D597	PF00982.21 PF02358.16	57 595	545 829	2 1	473 232	474 233	600.55 247.53	25.5 20.2	99.37 99.14	864
IDZIJ - TPS class II	SCA3 SP803280_c112588_g1_i2_m.123748	Glycosyltransferase family 20 Trehalose-phosphatase	D337	PF00982.21 PF02358.16	2 335	285 566	207 1	473 231	474 233	328.2 221.39	25.5 20.2	56.12 98.71	602
IE175 - TPS class II	SCA3 SP803280_c113038_g2_i1_m.126277	Glycosyltransferase family 20 Trehalose-phosphatase	D163	PF00982.21 PF02358.16	1 161	111 393	377 1	473 231	474 233	110.6 251.84	25.5 20.2	20.25 98.71	429
IE175 - TPS class II	SCA2_1 c94760f1p43082.p1	Glycosyltransferase family 20 Trehalose-phosphatase	D595	PF00982.21 PF02358.16	59 593	543 829	2 1	473 233	474 233	585.67 247.27	25.5 20.2	99.37 99.57	865
IE175 - TPS class II	SCA1_2 Sh_207K10_p000010	Glycosyltransferase family 20 Trehalose-phosphatase	D595	PF00982.21 PF02358.16	58 593	543 825	2 1	473 231	474 233	572.48 250.85	25.5 20.2	99.37 98.71	861
IEKDT -TPP	SCA3_SP803280_c102057_g1_i5_m.74366	Trehalose-phosphatase	-	PF02358.16	29	131	125	232	233	112.74	20.2	45.92	154
IEKDT -TPP	SCA3_SP803280_c102057_g1_i4_m.74365	Trehalose-phosphatase	D112	PF02358.16	110	344	1	232	233	260.64	20.2	99.14	367
IEKDT -TPP	SCA3_SP803280_c110556_g1_i2_m.112799	Trehalose-phosphatase	D123	PF02358.16	121	364	1	232	233	234.63	20.2	99.14	382
IEKDT -TPP	SCA3_SP803280_c110556_g1_i1_m.112798	Trehalose-phosphatase	-	PF02358.16	1	140	88	232	233	140.91	20.2	61.80	158
IEKDT -TPP	SCA2_1_c4404f1p21394.p1	Trehalose-phosphatase	D23	PF02358.16	21	153	1	125	233	131.64	20.2	53.22	163
IEKDT -TPP	SCA2_1_c23429f1p1731.p1	Trehalose-phosphatase	-	PF02358.16	27	129	128	233	233	102.11	20.2	45.06	146
IEKDT -TPP	SCA2_1_c94242f1p13373.p1	Trehalose-phosphatase	D221	PF02358.16	219	342	1	107	233	93.3	20.2	45.49	368
IEKDT -TPP	SCA1_2_Sh_225f16_p000080	Trehalose-phosphatase	D128	PF02358.16	126	362	1	232	233	244.19	20.2	99.14	371
IEKDT -TPP	SCA1_2_Sh_210F11_p000130	Trehalose-phosphatase	D118	PF02358.16	116	349	1	233	233	244.8	20.2	99.57	366
IDYMW - TRE	SCA1_2_Sh_240H12_p000020	Trehalase	D147 and E354	PF01204.18	5	391	139	512	512	453.32	20.2	72.85	408
IDYMW - TRE	SCA5_SCEZRZ1017E02.g.p1	Trehalase	D147 and E354	PF01204.18	5	391	139	512	512	457.06	20.2	72.85	408
IDYMW - TRE	SCA3_SP803280_c107494_g1_i5_m.97331	Trehalase	D185 and E392	PF01204.18	1	429	97	512	512	527.22	20.2	81.1	446

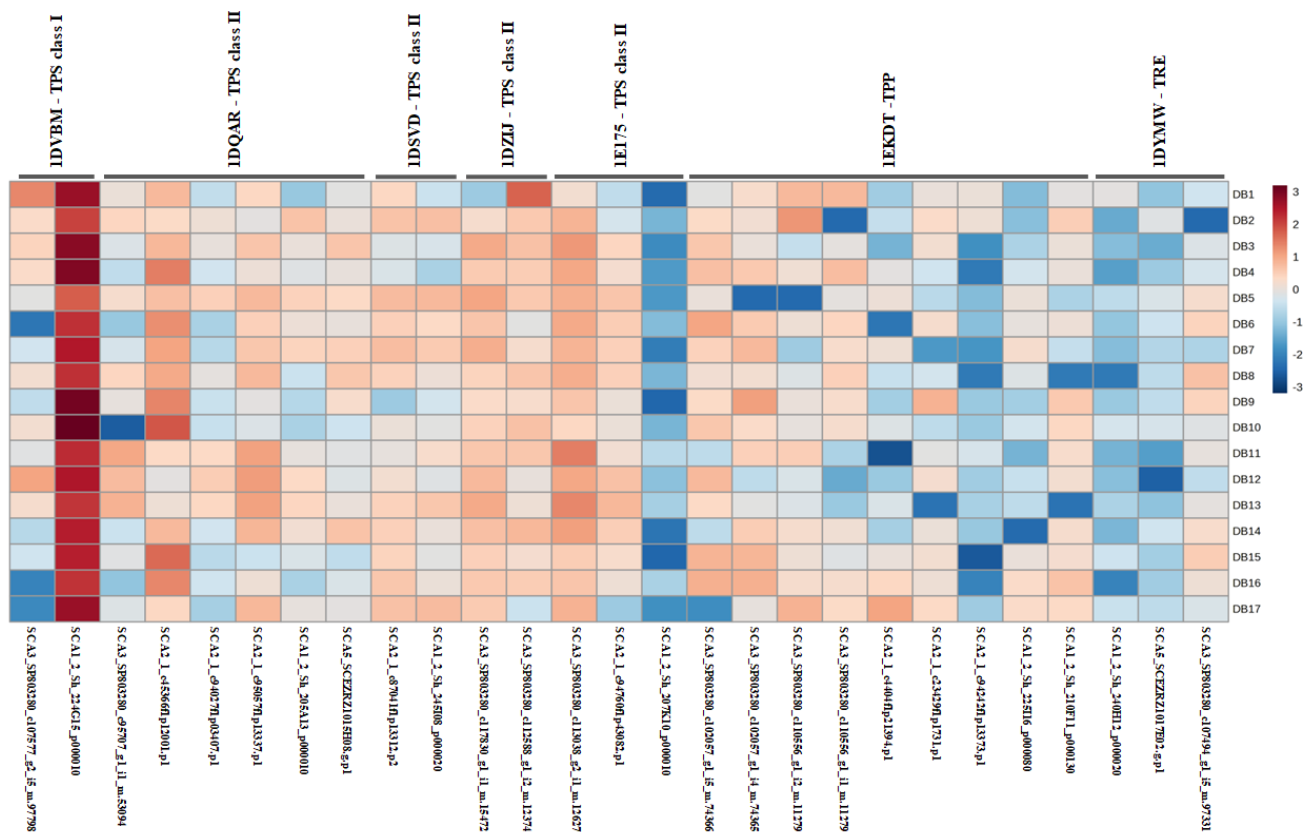


Figura 3. Valores de expressão (TPM) recuperadas de diferentes bancos de dados (DB) relacionados à variedade SP80-3280 de cana-de-açúcar. O banco de dado 1 (DB1) refere-se à MATIELLO et al. (2015), e os demais (DB2-DB17) referem-se à VETTORE et al. (2003). A intensidade das cores (vermelho mais expresso; azul menos expresso) representam valores de contagem de *reads* (TPM) normalizados por LOG10 e *auto-scaling*.

4.5 Análise filogenética

Com a finalidade de se avaliar a relação filogenética entre as diferentes sequências para TPS, TPP e TRE, foram construídas três árvores distintas para cada um dos alvos avaliados. Já é conhecido que TPS possui uma sub-funcionalização dentro de sua família multigênica (**Figura 4**). Essa sub-funcionalização é refletida em dois grandes clados (Classe I e Classe II). Para TPS classe I é reportado atividade enzimática catalítica dessas isoformas, enquanto a TPS classe II não possui atividade catalítica, desempenhando assim apenas atividades regulatórias. A TPS classe I é dividida entre os clados B1 e B2, sendo que o clado B1 possui sequências de Arabidopsis, Arroz, Milho e cana, enquanto o clado B2 apenas apresenta sequências de Arabidopsis. Para TPS classe II elas são divididas em 5 clados (A1, A2, A3, A4 e A5) e observou-se a presença de sequências de cana-de-açúcar dispersas entre os clados A2-A5. O único clado onde não foi observado uma sequência de cana-de-açúcar foi o clado A1 onde só é possível observar uma única sequência de Arabidopsis.

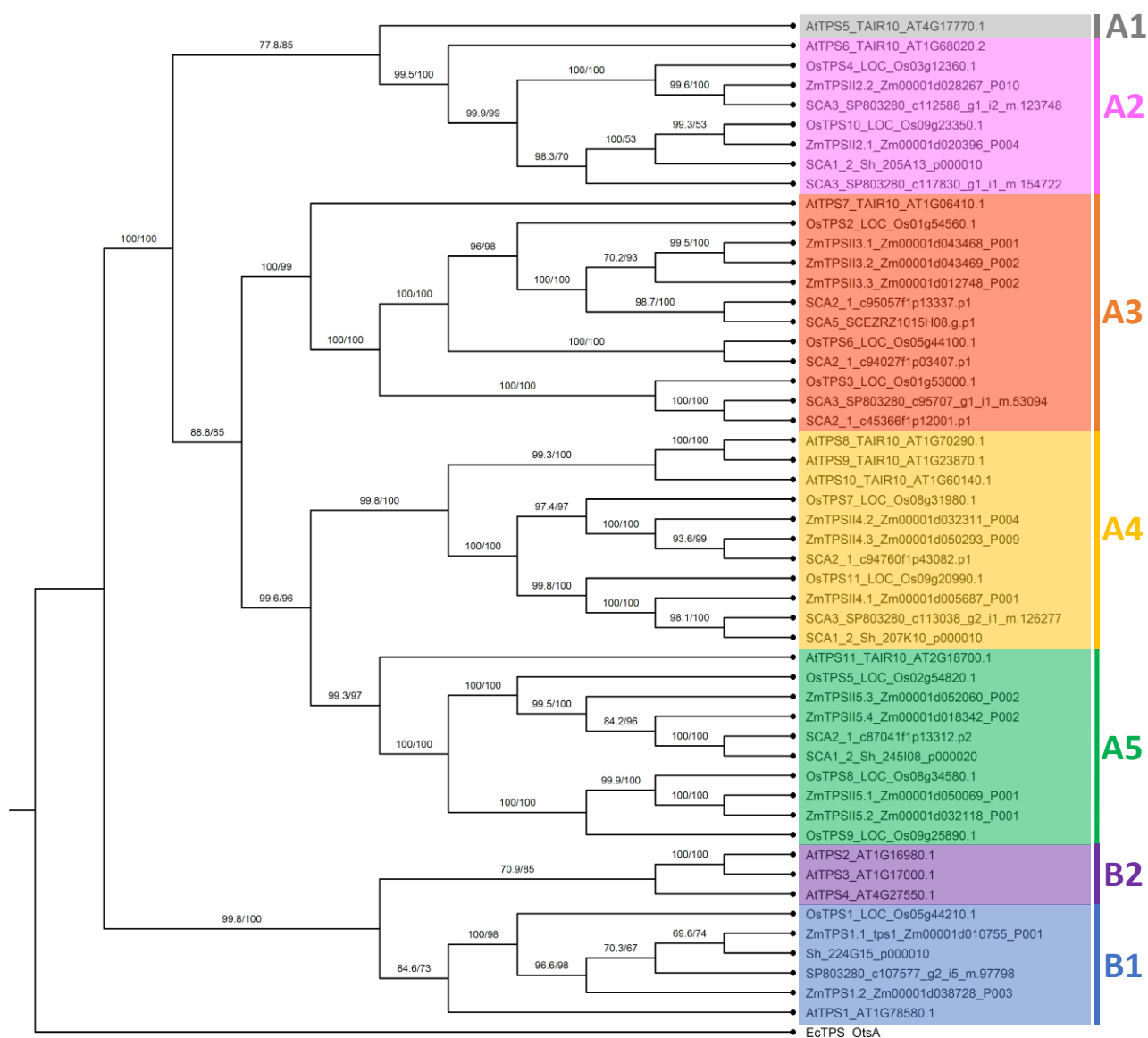


Figura 4. Árvore filogenética de trealose-6-fosfato sintase (TPS) para cana-de-açúcar (Tabela 3), *Zea mays*, *Oryza sativa*, *Arabidopsis thaliana* (Tabela 1), e *Escherichia coli*. As sequências utilizadas assim como a nomenclatura dos clados seguem trabalhos previamente estabelecidos (31,36,37). A árvore foi construída pelo IQ-TREE (54), utilizando-se seleção de modelo evolutivo automático, os valores de suporte de ramos são mostrados como SH-like aLRT (%) (55) e modelo de bootstrap ultrafast (UFboot) (%) (14).

TPP também possui um agrupamento específico dentro da família multigênica (Figura 5), porém esse agrupamento não distingue isoformas catalíticas e regulatórias, uma vez que até o presente momento só são relatadas atividades catalíticas e regulatórias conjuntas em sequências de TPP. Para TPP foram observados a presença de 5 clados sendo eles A1, A2, A3, B1 e B2. Os clados em sua maioria possuem uma definição

congruente com o observado na literatura (31), porém o clado B1 não acompanhou a topologia esperada com um deslocamento de uma sequência de *Arabidopsis* para um outro clado. Esse deslocamento pode ter ocorrido pela inclusão de sequências de cana na árvore, o que resolveria um pouco melhor a relação filogenética entre as sequências avaliadas.

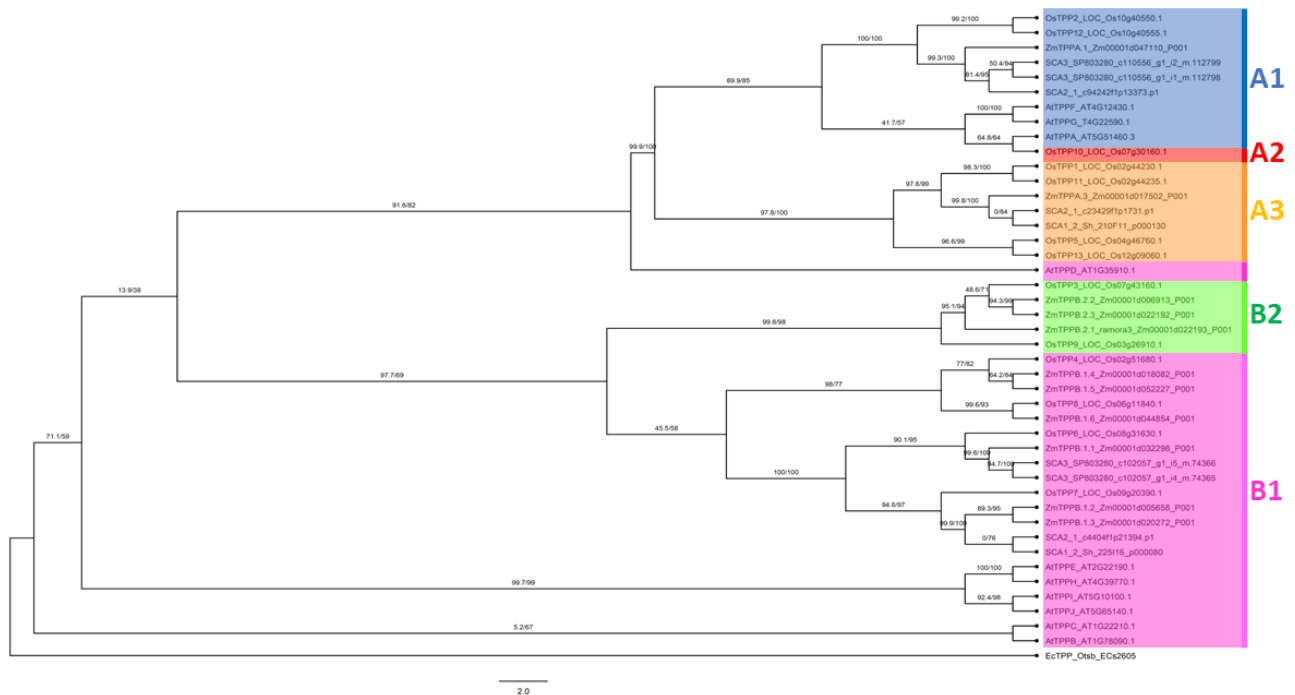


Figura 5. Árvore filogenética de trealose-6-fosfato fosfatase (TPP) para cana-de-açúcar (Tabela 3), *Zea mays*, *Oryza sativa*, *Arabidopsis thaliana* (Tabela 1), e *Escherichia coli*. As sequências utilizadas assim como a nomenclatura dos clados seguem trabalhos previamente estabelecidos (31,36,37). A árvore foi construída pelo IQ-TREE (54), utilizando-se seleção de modelo evolutivo automático, os valores de suporte de ramos são mostrados como SH-like aLRT (%) (GUINDON et al., 2010) e modelo de bootstrap ultrafast (UFboot) (%) (HOANG et al., 2018).

Para TRE foi observado uma dispersão filogenética congruente com o esperado entre a dispersão entre as espécies (**Figura 6**), em que sequências de cana e milho apresentaram uma relação de filogenia mais próxima, visto que essas espécies são monocotiledôneas gramíneas, e um deslocamento maior da sequência de *Arabidopsis*, uma vez que essa sequência se refere a uma espécie de eudicotiledônea.

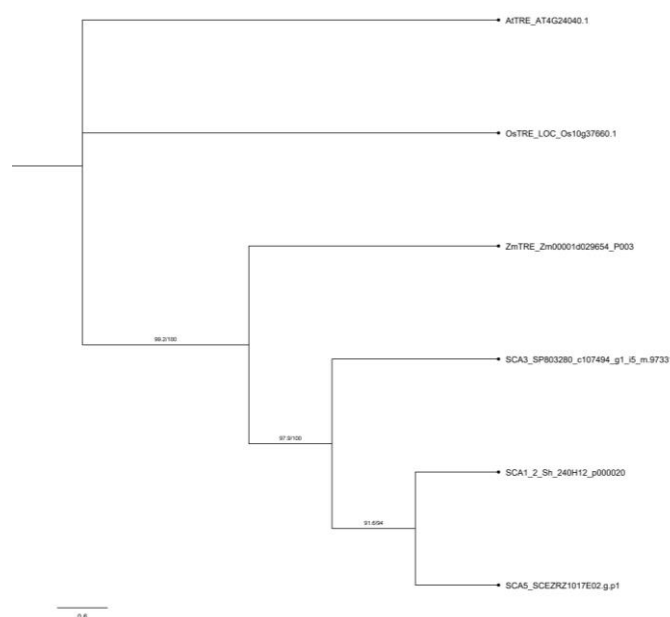


Figura 6. Árvore filogenética de trehalase (TRE) para cana-de-açúcar (Tabela 3), *Zea mays*, *Oryza sativa* e *Arabidopsis thaliana* (Tabela 1). As sequências utilizadas assim como a nomenclatura dos clados seguem trabalhos previamente estabelecidos (31,36,37). A árvore foi construída pelo IQ-TREE (54), utilizando-se seleção de modelo evolutivo automático, os valores de suporte de ramos são mostrados como SH-like aLRT (%) (GUINDON et al., 2010) e modelo de bootstrap ultrafast (UFboot) (%) (14).

4.6. Extração de RNA e síntese de cDNA em amostras de folha e colmo de cana-de-açúcar

Com a finalidade de realizar as análises de expressão gênica via RT-qPCR, o RNA total foi extraído de amostras de folha e colmo de cana-de-açúcar, crescida à campo. Após a extração do RNA total, as amostras foram quantificadas e validadas quanto a sua pureza e integridade. Na **tabela 4** encontram-se os valores da quantificação do RNA e das razões 260/280 e 260/230, as quais indicam a pureza das amostras. No geral, foi obtido um alto rendimento de extrato, com medidas de pureza para a razão 260/280 dentro dos padrões de qualidade requeridos. A razão 260/230 apresentou um tendência de índices menores, porém a integridade do RNA mantevesse quando analisado em gel de agarose.

A partir da síntese de cDNA das respectivas amostras biológicas em que foi obtido o RNA, foi possível realizar uma PCR para avaliar a especificidade dos iniciadores desenhados a partir das sequências dos respectivos alvos (**Tabela 5**). Ao longo do processo de avaliação de possíveis iniciadores, observamos que sete sequências estavam contidas dentro de outras sequências de maior tamanho, e por isso não era possível retornar um único produto de amplificação. Desta forma, 20 pares de iniciadores foram sintetizados e posteriormente testados. Dos 20 pares de iniciadores, oito apresentaram amplificação

específica e condizente com o tamanho predito do amplicon. Assim, para a análise de RT-qPCR foram mantidos os seguintes alvos: *TPSI_A*, *TPSI_B*, *TPSII_A*, *TPSII_E*, *TPSII_G*, *TPSII_J*, *TPP_F* e *TRE*.

Tabela 4. Quantificação e pureza do RNA extraído para amostras de folha e colmo de cana-de-açúcar. A quantidade de RNA total é representada em ng/μL. A pureza é determinada pelas razões 260/280nm e 260/230nm.

Órgão	Meses	Amostras	260/280 nm	260/230 nm	Rendimento (ng/μL)
Folha	01	F011	2,01	1,27	330,61
		F015	2,04	1,53	259,71
		F017	2,04	1,84	196,29
	03	F111	2,16	2,29	702,79
		F113	2,16	2,22	595,44
		F117	2,12	1,71	590,39
	06	F211	2,08	1,59	599,55
		F213	2,06	1,99	356,01
		F217	2,08	1,21	312,66
	12	F413	2,06	1,56	374,47
F415		2,05	2,13	392,70	
F417		2,16	2,32	514,08	
Colmo	01	C011	2,10	1,81	615,67
		C015	2,09	1,57	562,02
		C017	1,69	1,50	216,28
	03	C111	2,00	0,79	103,02
		C113	1,99	1,07	74,78
		C211	1,97	0,98	89,28
	06	C215	1,92	0,76	77,00
		C217	2,03	1,19	122,43
		C413	1,86	0,88	73,32
	12	C415	1,97	0,73	150,18

4.7 RT-qPCR

Com o intuito de analisar o perfil de expressão gênica dos alvos *TPS* classe I e II, *TPP* e *TRE*, uma análise de RT-qPCR foi realizada para amostras biológicas de folhas e colmos coletados ao longo do desenvolvimento à campo da cana-de-açúcar (1, 3, 6 e 12 meses) (**Figura 7**).

Tabela 5. Lista de iniciadores utilizados neste trabalho. A tabela abaixo evidencia as sequências dos iniciadores utilizados na análise de RT-qPCR, assim como os parâmetros físico-químicos atrelados a cada sequência e o tamanho do amplicon esperado para cada alvo.

ID da sequência	Nome do alvo	5'-Sequência-3'	Pares de base	GC (%)	Temperatura de Melting (°C)	Hairpin (ΔG)	Temperatura de Harpin (°C)	Self-dimer (ΔG)	Número de ligações	Hetero-dimer (ΔG)	Número de ligações	Amplicon
SCA3 SP803280_c107577_g2_i5_m.97798	TPS/I/A-F	GAGAAGGTGATGTGAAAGG	20	45	50.9	1.98	-57.7	-1.94	2	-3.89	3	201
	TPS/I/A-R	CTGGGATTTGGAGAGGAG	18	55.6	51.9	1.69	-13.4	-1.60	2			
Sh_224G15_p000010	TPS/I/B-F	CTCCAGCCTCAACATCACC	19	57.9	55.6	-	-	-3.14	2	-3.55	3	215
	TPS/I/B-R	ACCGCACTGACATTATCG	18	50	52.3	0.66	12.7	-3.61	2			
SCA3_SP803280_c95707_g1_i1_m.53094	TPS/II/A-F	CAAATCGCTGCTCAAAGG	18	50	51.8	2.06	-27.7	-3.61	2	-3.61	2	152
	TPS/II/A-R	TAGGAGGTGTAGTCGTG	18	55.6	52.7	1.55	6.7	-3.61	2			
SCA2_1_c94027f1p03407.p1	TPS/II/B-F	GGATACATTGGATTGGAGTAC	21	42.9	50.3	0.56	17.3	-3.65	4	-3.30	3	196
	TPS/II/B-R	GGAAACGGACTGTGGAGGAAG	21	57.1	57.7	0.15	20.7	-3.61	2			
SCA2_1_c95057f1p13337.p1	TPS/II/C-F	CTCTTTGTTCTGCTCGGTG	19	52.6	53.5	0.98	6.2	-3.61	2	-3.54	3	161
	TPS/II/C-R	GAAGGAGATGGTGGGTTC	18	55.6	52.4	1.71	-1.7	-3.52	3			
SCA1_2_Sh_205A13_p000010	TPS/II/D-F	GTTTAGCAGCAGGATACAAG	21	47.6	53.4	1.28	-7.7	-3.61	2	-3.29	3	227
	TPS/II/D-R	CAGTAGACAACCACAACAG	19	47.4	50.5	0.34	19.3	-1.34	2			
SCA5_SCEZRZ1015H08.g.p1	TPS/II/E-F	CGATGTTTCTACTATCTGAGGC	22	45.5	52.6	-0.02	25.3	-3.61	2	-4.03	4	231
	TPS/II/E-R	GGACTTGITGGGGCATAG	18	55.6	53.1	0.93	9.7	-3.14	2			
SCA2_1_c87041f1p13312.p2	TPS/II/F-F	CTAATCCAATCCAATCCTCATC	22	45.5	52.5	2.04	-106.2	-1.60	2	-3.90	3	183
	TPS/II/F-R	GTAGTGTGACGGGTTCCTTTC	21	52.4	56.1	1.03	-4.5	-3.61	2			
SCA3 SP803280_c117830_g1_i1_m.154722	TPS/II/G-F	CTCTATCTCTCTGAACCTCTTCC	23	43.5	51.8	0.01	24.9	-3.52	3	-3.89	3	180
	TPS/II/G-R	CTGGCTTGTGGTCAAATG	18	50	51.6	0.05	24.4	-3.90	3			
SCA3 SP803280_c112588_g1_i2_m.123748	TPS/II/H-F	CTGAGAATGAGAAGGTTG	20	45	50.1	1.27	-9.6	-1.94	2	-3.90	3	185
	TPS/II/H-R	GAAGTGAAGGTCAAGCGAG	19	52.6	53.3	0.51	16.1	-3.61	2			
SCA3 SP803280_c113038_g2_i1_m.126277	TPS/II/I-F	GGAGTTGACAAGGGGATTTCTG	22	50	55.6	1.11	4.6	-3.90	3	-3.90	3	207
	TPS/II/I-R	GTAGTTCGCTTTGCTGGC	18	55.6	54.4°	0.27	19.5	-3.61	2			
SCA2_1_c94760f1p43082.p1	TPS/II/J-F	GTCTCTTCCTCGTGTGTAAC	21	47.6	52.9	0.41	18.5	-3.61	2	-3.30	3	233
	TPS/II/J-R	CTTCTTCCTCCTCACCTG	19	52.6	52.2	-	-	-1.95	2			
SCA1_2 Sh_207K10_p000010	TPS/II/K-F	CTCTTCCTCGTGTGATGACC	20	55	54.9	0.69	13.6	-3.61	2	-3.30	3	167
	TPS/II/K-R	CACCAACCCTCTCTCCTCTGTAC	24	54.2	58.5	1.77	-23.5	-3.65	4			
SCA3_SP803280_c102057_g1_i4_m.74365	TPP/A-F	CAGTTGGATGACGAAGC	18	55.6	53.5	1.26	-0.7	-3.61	2	-3.61	2	179
	TPP/A-R	GAAGAAGGTGCGGAGGAGAG	19	57.9	54.9	-	-	-3.61	2			
SCA3_SP803280_c110556_g1_i2_m.112799	TPP/B-F	GTTATGAGCAAAGGAGCCG	19	52.6	53.7	-0.20	28.2	-3.61	2	-3.90	3	212
	TPP/B-R	TTGAACCTGATTGAGAAACCC	20	40	50.3	-0.28	29.8	-1.94	2			
SCA2_1_c4404f1p21394.p1	TPP/C-F	CTGAGGTGATGGAGTCTTTG	20	50	52.7	0.98	11.5	-1.60	2	-3.90	3	196
	TPP/C-R	CGGGACAGTAGAGACAATAC	20	50	51.9	0.02	24.7	-3.61	2			
SCA2_1_c94242f1p13373.p1	TPP/D-F	CCCCCTTCCAATACAGAATG	20	50	52.9	0.10	22.8	-3.52	3	-1.95	2	206
	TPP/D-R	CAATCCAGCCTACACTACAG	20	50	52.5	1.16	-10.9	-3.14	2			
SCA1_2_Sh_225I16_p000080	TPP/E-F	GTCCAGCCAAGTCTCTTTC	19	57.9	55.3	0.04	24.4	-3.54	3	-3.54	3	192
	TPP/E-R	CTCGCCCCACATCTTTTC	18	55.6	53.6	-	-	-3.61	2			
SCA1_2_Sh_210F11_p000130	TPP/F-F	CTCTTCCTGTTGAGCTATTG	21	47.6	52.9	1.20	6.0	-3.55	3	-3.14	2	174
	TPP/F-R	GTGTTCCTTCTTGCCATC	19	52.6	53.7	2.19	-17.1	-3.14	2			
SCA5_SCEZRZ1017E02.g.p1	TRE-F	CAATAAACACAAACCCACGAAC	22	40.9	52.8	-	-	-3.61	2	-3.07	2	212
	TRE-R	CAGAGGACAAGGAAATAGC	19	47.4	50.2	-	-	-3.14	2			

No total foram obtidos valores de expressão para 8 alvos distintos. O alvo *TPSI_A* apresentou um decréscimo de expressão nas folhas ao longo do desenvolvimento da cana à campo, enquanto para o colmo os maiores valores foram encontrados no terceiro e décimo segundo mês (**Figura 7A**). Já o alvo *TPSI_B*, também pertencente à TPS classe I, apresentou um perfil crescente ao longo do desenvolvimento à campo, sendo oposto à *TPSI_A*, porém congruente entre folha e colmo (**Figura 7B**).

Para os alvos de TPS classe II, não foi possível observar um padrão único de expressão entre estes transcritos, porém é possível afirmar que os maiores valores de expressão para *TPSII_A*, *TPSII_G* e *TPSII_J* foram no colmo aos 12 meses de cultivo, enquanto para *TPSII_E* foi neste mesmo órgão, porém no primeiro mês de cultivo da cana à campo. Para *TPP_F* não se observou diferenças significativas entre os meses analisados para cada órgão, sendo que a única diferença encontrada foi entre folha e colmo no terceiro mês de cultivo (**Figura 7G**). Para *TRE* foram obtidos valores mais expressivos no colmo em comparação a folha em todos os meses analisados (**Figura 7H**).

A partir destes dados, uma análise multivariada foi realizada com a finalidade de identificar padrões de diferenciação com base em múltiplas variações (**Figura 8**). Tanto para folha (**Figura 8A**) quanto para colmo (**Figura 8B**), o PLS-DA explicou mais que 40% das variações observadas, sendo que o mês 01 é o único que apresentou um agrupamento diferencial, explicado por ambas as componentes 1 e 2. Dentre as variáveis com maior grau de contribuição no agrupamento, *TPSI_B* e *TPSII_A* foram as que apresentaram o maior VIP score em ambos os órgãos, tanto para o componente 1, quanto para o componente 2.

Com a finalidade de validar os dados de expressão gênica do TAS via RT-qPCR, foram obtidos valores de correlação de Pearson para cada alvo, não discriminando os meses de coleta (**Tabela 6**). Nesta análise foi possível observar que três alvos (*TPSI_A*, *TPSI_B* e *TPP_F*) apresentaram valores de correlação (r) maior que 0.50. Dentre estes alvos, *TPSI_A* /foi o que apresentou o maior valor de correlação ($r = 0.92$).

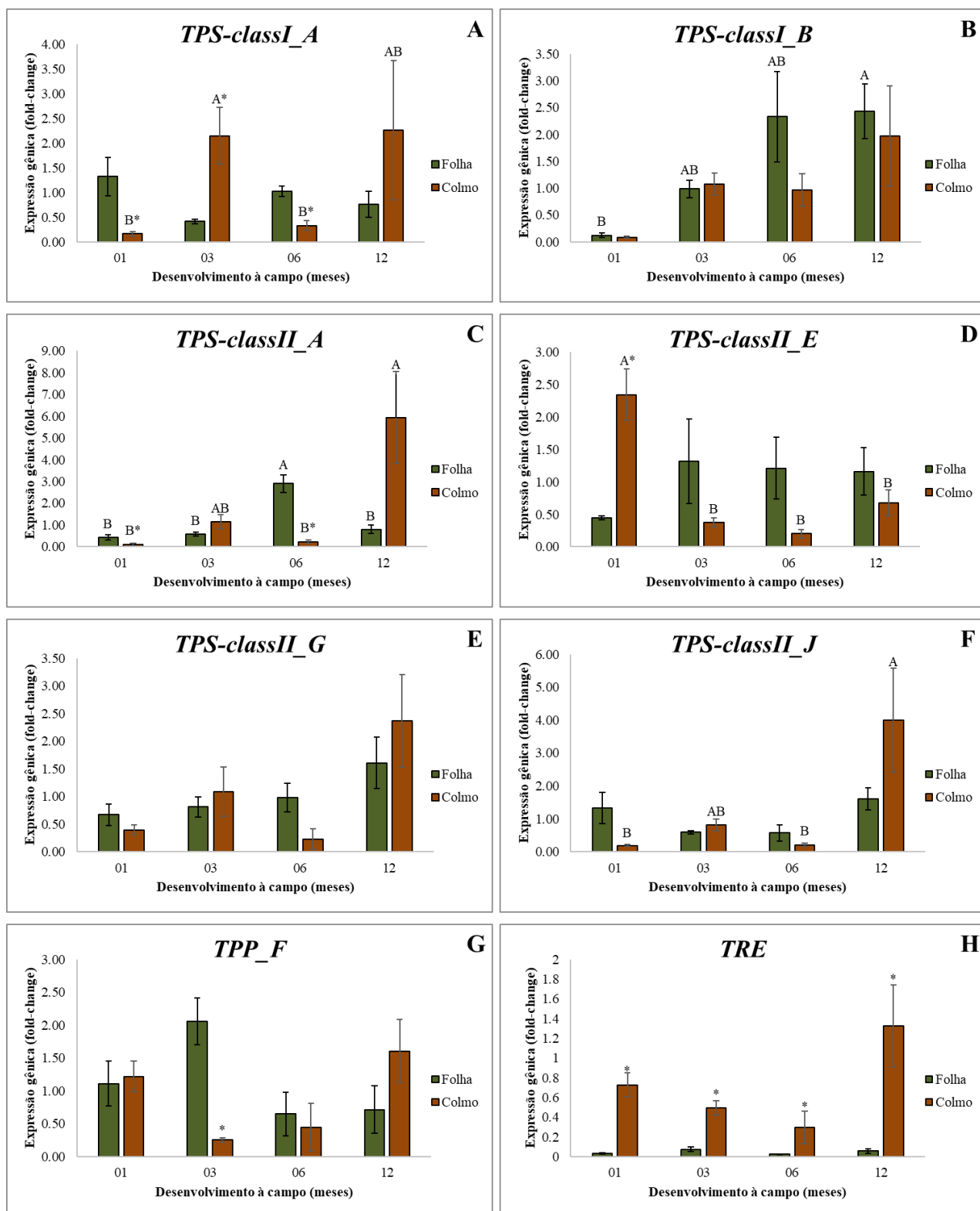


Figura 7. Valores de expressão gênica (*fold-change*) para alvos da via de biossíntese da trealose-6-fosfato (Tre6P). No total, foram avaliados oito alvos, duas *TPS* classe I – *TPSI_A* (A) e *TPSI_B* (B); quatro *TPS* classe II – *TPSII_A* (C), *TPSII_E* (D), *TPSII_G* (E) e *TPSII_J* (F); uma *TPP* – *TPP_F* (G); e uma *TRE* (H). Barras representam média \pm erro padrão. Letras maiúsculas indicam diferenças significativas ($p < 0,05$) entre os meses individualmente para cada órgão. Asteriscos (*) indicam diferenças significativas ($p < 0,05$) entre os órgãos para cada mês.

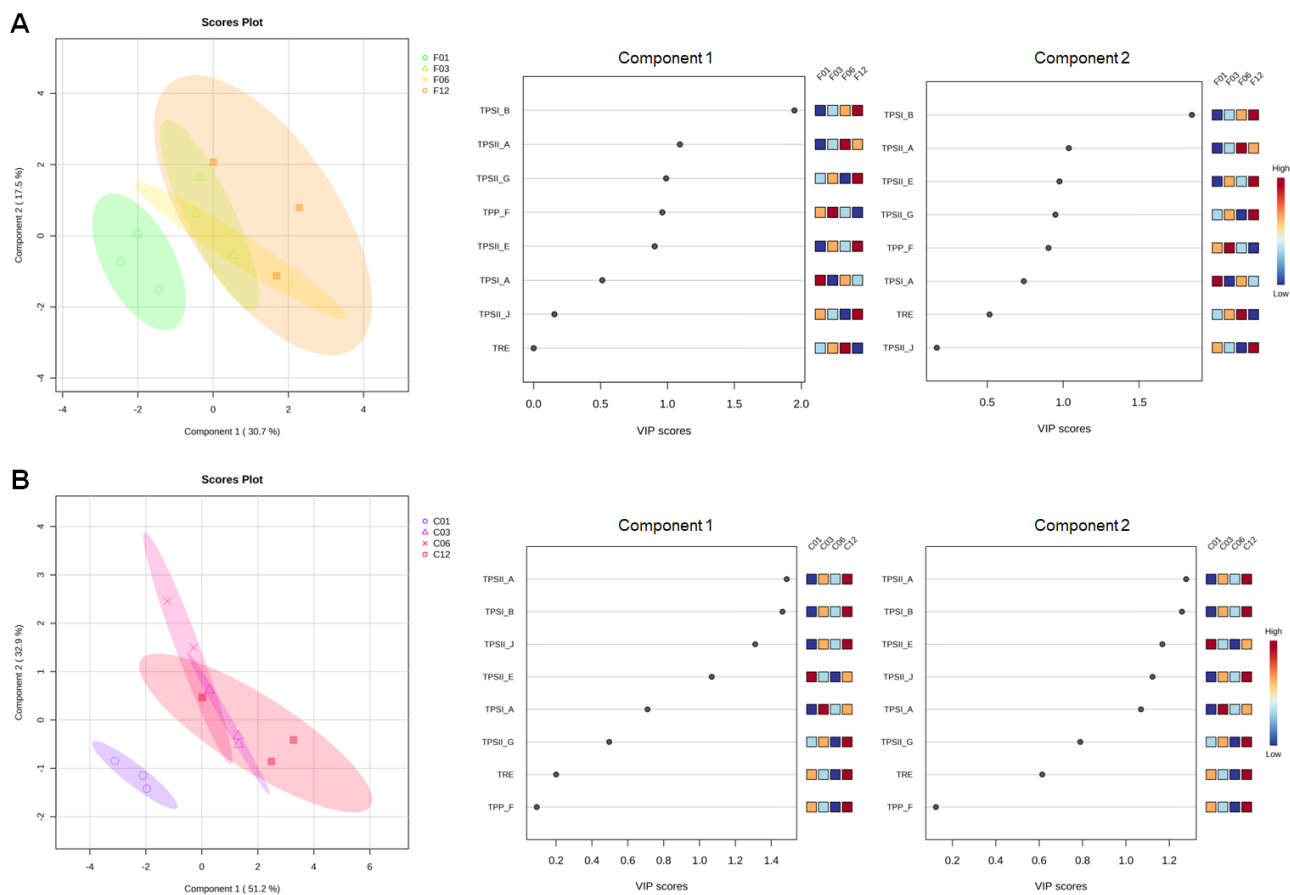


Figura 8. PLS-DA e VIP score para folha (A) e colmo (B) de cana-de-açúcar (variedade SP80-3280) crescida à campo. Os valores de expressão gênica de alvos da via de biossíntese da trealose-6-fosfato (Tre6P) foram utilizados na análise.

Tabela 6. Correlação entre os valores de expressão obtidos via *Target Amplicon RNA-Seq* (TAS) e RT-qPCR (qPCR) para alvos da via de biossíntese da trealose-6-fosfato (Tre6P) em cana-de-açúcar crescida à campo (01, 03, 06 e 12 meses). Valores indicam média \pm erro padrão. Os valores de (r) foram obtidos através de correlação de Pearson independentemente para cada alvo.

Mês	Órgão	Dados	<i>TPSI_A</i>	<i>TPSI_B</i>	<i>TPSII_A</i>	<i>TPSII_E</i>	<i>TPSII_G</i>	<i>TPSII_J</i>	<i>TPP_F</i>	<i>TRE</i>
01	Folha	TAS	0.00 \pm 0.00	1.28 \pm 1.01	375.56 \pm 367.75	13983.52 \pm 7895.39	0.00 \pm 0.00	883.58 \pm 444.82	4.18 \pm 2.15	28.90 \pm 23.27
		qPCR	1.32 \pm 0.39	0.12 \pm 0.04	0.43 \pm 0.11	0.45 \pm 0.03	0.67 \pm 0.20	1.32 \pm 0.47	1.11 \pm 0.34	0.75 \pm 0.11
	Colmo	TAS	215.61 \pm 215.61	0.30 \pm 0.30	0.25 \pm 0.25	7905.17 \pm 3643.00	0.00 \pm 0.00	1287.52 \pm 440.07	1.14 \pm 1.14	11.35 \pm 2.86
		qPCR	0.16 \pm 0.05	0.09 \pm 0.02	0.11 \pm 0.06	2.34 \pm 0.40	0.40 \pm 0.09	0.19 \pm 0.02	1.22 \pm 0.23	0.73 \pm 0.12
03	Folha	TAS	0.00 \pm 0.00	1.76 \pm 1.76	149.55 \pm 112.46	39949.88 \pm 19693.08	0.00 \pm 0.00	276.47 \pm 276.47	10.20 \pm 6.05	5.90 \pm 2.95
		qPCR	0.41 \pm 0.045	0.99 \pm 0.16	0.59 \pm 0.09	1.32 \pm 0.65	0.81 \pm 0.19	0.59 \pm 0.05	2.06 \pm 0.36	0.81 \pm 0.14
	Colmo	TAS	163.55 \pm 161.98	0.60 \pm 0.43	4.94 \pm 4.43	14933.51 \pm 3003.65	0.00 \pm 0.00	519.67 \pm 87.71	4.07 \pm 1.22	31.84 \pm 15.12
		qPCR	2.15 \pm 0.57	1.08 \pm 0.20	1.16 \pm 0.33	0.37 \pm 0.07	1.08 \pm 0.45	0.82 \pm 0.17	0.25 \pm 0.04	0.49 \pm 0.07
06	Folha	TAS	1886.33 \pm 1885.92	1.08 \pm 0.64	3.87 \pm 3.87	5212.09 \pm 4006.92	0.00 \pm 0.00	160.02 \pm 160.02	2.20 \pm 2.20	15.68 \pm 15.68
		qPCR	1.02 \pm 0.11	2.33 \pm 0.84	2.90 \pm 0.42	1.21 \pm 0.48	0.98 \pm 0.26	0.57 \pm 0.24	0.65 \pm 0.33	1.45 \pm 0.42
	Colmo	TAS	0.00 \pm 0.00	0.53 \pm 0.27	477.38 \pm 200.45	9460.65 \pm 3117.37	13.67 \pm 7.20	390.38 \pm 71.08	3.81 \pm 0.44	62.06 \pm 13.15
		qPCR	0.33 \pm 0.10	0.97 \pm 0.31	0.23 \pm 0.07	0.20 \pm 0.06	0.23 \pm 0.19	0.21 \pm 0.05	0.45 \pm 0.36	0.30 \pm 0.16
12	Folha	TAS	0.00 \pm 0.00	0.00 \pm 0.00	0.00 \pm 0.00	8803.69 \pm 775.59	0.00 \pm 0.00	711.24 \pm 370.66	9.72 \pm 5.37	2.20 \pm 2.20
		qPCR	0.76 \pm 0.27	2.43 \pm 0.51	0.79 \pm 0.19	1.16 \pm 0.36	1.61 \pm 0.46	1.61 \pm 0.34	0.71 \pm 0.36	0.77 \pm 0.33
	Colmo	TAS	0.00 \pm 0.00	0.14 \pm 0.14	497.56 \pm 495.98	3474.29 \pm 1543.68	0.00 \pm 0.00	865.87 \pm 103.77	5.46 \pm 2.80	11.06 \pm 2.68
		qPCR	2.27 \pm 1.41	1.97 \pm 0.94	5.94 \pm 2.10	0.67 \pm 0.20	2.37 \pm 0.84	4.00 \pm 1.58	1.61 \pm 0.48	1.32 \pm 0.41
Correlação (r)			0.92	0.84	0.12	0.32	0.25	0.22	0.67	0.26

6. DISCUSSÃO

A cana-de-açúcar é de cultura agrícola de extrema importância no setor sucroalcooleiro do Brasil, impactando significativamente em aspectos econômicos e sociais. Além de ser uma alternativa eficaz para mitigar as emissões de CO₂, contribuindo para o combate às mudanças climáticas, outros bioprodutos provindos da cana-de-açúcar também são importantes e auxiliam na rentabilidade associada à esta cultivar (18). Entretanto, estudos acerca de sua fisiologia e bioquímica associadas ao crescimento e desenvolvimento ainda são necessários, a fim de se alcançar uma melhoria no seu potencial produtivo. Um dos maiores desafios atuais é o mapeamento do genoma desta espécie, uma vez que há altos níveis de ploidia, presença de aneuploidia, alta taxa de polimorfismo e conteúdo de repetição em suas sequências gênicas, dificultando assim estratégias de sequenciamento do genoma, e conseqüentemente, fazendo com que haja lacunas para a compreensão da função de sua estrutura molecular (56). O aumento do mapeamento do genoma da cana-de-açúcar possui um potencial biotecnológico inigualável, uma vez que, a partir do conhecimento atrelado ao genoma, será possível desenvolver estratégias de engenharia genética para esta cultura (14). Desta forma, será possível aumentar o ganho de características agrícolas de interesse, como por exemplo, o aumento dos teores de sacarose no colmo.

Um dos processos biológicos mais importantes que contribuem para o acúmulo de sacarose nas plantas é o mecanismo de sensoriamento e percepção de açúcares. Entre os sinalizadores de açúcar mais estudados em plantas, encontra-se a Tre6P, um intermediário da via da biossíntese da trealose, que possui uma relação de controle de níveis de sacarose em diferentes órgãos da planta (43,44). Assim, estudos que visam a identificação dos genes da via de Tre6P (*TPS*, *TPP* e *TRE*) podem auxiliar no entendimento da regulação dos níveis de sacarose em cana-de-açúcar. DE OLIVEIRA e colaboradores (2022), realizaram análises abrangentes de sequências de enzimas do metabolismo da trealose em cana-de-açúcar, assim possibilitando análises e estudos funcionais para elucidar os papéis fisiológicos da sinalização de Tre6P em cana. Neste trabalho foram recuperadas sequências putativas de *TPS*, *TPP* e *TRE* de cana-de-açúcar, evidenciando seus respectivos domínios conservados e suas relações filogenéticas. A partir destas sequências, um painel do tipo *Target Amplicon RNA-Seq* (TAS) foi obtido e utilizado neste projeto como base inicial para recuperação das sequências dos alvos relacionados ao sensoriamento de açúcar em cana-de-açúcar.

Neste sentido, este projeto teve como objetivo caracterizar a expressão gênica de diferentes transcritos de (*TPS*, *TPP* e *TRE*) ao longo do desenvolvimento da cana a campo.

Primeiramente foi recuperado sequências iscas de espécies modelos com funções confirmadas para alvos da via da trealose, assim possibilitando a recuperação de sequências homólogas a estas em cana-de-açúcar (**Tabela 1**). Inicialmente foram recuperadas 42 sequências, mas a partir de um alinhamento par-a-par foi possível observar a presença de redundâncias entre estas sequências, das quais 27 foram mantidas (**Tabela 2**). Também foram analisados os respectivos domínios proteicos destes alvos, sendo que para TPS apresentou os domínios *glycosyltransferase família 20* e *Trealose fosfatase*, enquanto para TPP foi possível identificar o domínio *Trealose fosfatase* e para TRE o domínio *trealase* (**Tabela 3**).

Em uma inferência filogenética, foi possível observar que as sequências de TRE recuperadas para cana-de-açúcar seguiram a congruência da dispersão das espécies em Viridiplantae descrito na literatura (**Figura 6**). Da mesma forma, árvores filogenéticas também foram construídas para TPS (**Figura 4**) e TPP (**Figura 5**). Estes dois alvos apresentam sequências ortólogas e parálogas em diversas espécies de plantas (31,36,37), sendo sua dispersão agrupada em clados bem definidos. A TPP apresenta 5 clados sem diferenciação de função ainda relatada na literatura (36). Já TPS é subdivida em dois grandes clados (A e B) que representam sequências catalíticas (clado A) e regulatórias (clado B), definidas respectivamente como TPS classe I e classe II. As TPS I foram divididas entre B1 e B2, sendo que o clado B1 possui sequências de Arabidopsis, Arroz, Milho e cana, enquanto o clado B2 apenas apresenta sequências de Arabidopsis. Já a classe II é dividida em 5 clados (A1, A2, A3, A4 e A5), em que se observou que o clado A1 foi o único não possuía sequências de cana-de-açúcar.

Com o intuito de analisar o perfil de expressão dos alvos TPS I e II, TPP e TRE ao longo do desenvolvimento da cana-de-açúcar entre 1, 3, 6 e 12 meses foi realizado uma análise de RT-qPCR (**Figura 7**). Observou-se padrões de expressão inverso entre folha e colmo para os alvos de TPS classe I analisados, em que *TPSI_A* diminuiu sua expressão ao longo do desenvolvimento e *TPSI_B* aumentou (**Figuras 7A e 7B**). Estes dois alvos também apresentaram os maiores valores de correlação entre TAS e RT-qPCR (0.92 e 0.84, respectivamente) (**Tabela 6**). Tre6P é um conhecido regulador dos níveis de sacarose em plantas (29). Ao ponto que os níveis de sacarose se elevam, os níveis de Tre6P também aumentam, principalmente pela ativação da transcrição de TPS e inibição da transcrição de TPP (44). Em cana-de-açúcar, os níveis de sacarose aumentam progressivamente no colmo e na folha ao longo do desenvolvimento à campo, atingindo valores máximos no décimo segundo mês de cultivo (48). Neste sentido, o padrão de expressão da *TPSI_B* está mais de

acordo com a teoria do *nexus* sacarose/Tre6P. De fato, a expressão de *TPSI_B* é o que melhor explica a diferenciação entre os meses de cultivo da cana à campo (**Figura 8**). Além de ser um dos alvos que possui um dos maiores valores de correlação que validam o TAS (**Tabela 6**). Desta forma, este alvo possui um potencial para ser selecionado como promissor para futuramente ser empregado protocolos de transformação genética em cana-de-açúcar.

7. CONCLUSÃO

Nossos resultados exploram as diferenças nos níveis de expressão para alvos da via de biossíntese de Tre6P em cana-de-açúcar. O estudo desta via em cana se faz necessário, uma vez que a Tre6P é um conhecido regulador dos níveis de sacarose nas plantas. A identificação do perfil de expressão de *TPS*, *TPP* e *TRE* em cana-de-açúcar possui um potencial de identificar alvos importantes que possam auxiliar no incremento de sacarose neste cultivar agrícola, impactando significativamente sua produtividade. De forma geral, foram obtidos valores de expressão em folhas e colmos de cana crescida à campo (1, 3, 6 e 12 meses) para oito alvos distintos da via de biossíntese de Tre6P. Além da identificação do perfil de expressão destes alvos, também foi possível correlacionar estes valores com os encontrados no banco *Target Amplicon RNA-Seq*. A partir dos resultados obtidos via RT-qPCR, os alvos *TPSI_A*, *TPSI_B* e *TPSII_G* apresentaram altos valores de correlação com o TAS, podendo assim validar a acurácia deste banco de dados.

8. REFERÊNCIAS

1. Mitchell JFB, Senior CA, Ingram WJ. CO₂ and climate: a missing feedback? *Nature*. Setembro de 1989;341(6238):132–4.
2. Dufresne JL, Foujols MA, Denvil S, Caubel A, Marti O, Aumont O, et al. Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Clim Dyn*. 1º de maio de 2013;40(9):2123–65.
3. Mazzarino M. The economics of the greenhouse effect: evaluating the climate change impact due to the transport sector in Italy. *Energy Policy*. 1º de novembro de 2000;28(13):957–66.
4. Mileva A, Johnston J, Nelson JH, Kammen DM. Power system balancing for deep decarbonization of the electricity sector. *Appl Energy*. 15 de janeiro de 2016;162:1001–9.
5. Rodionova MV, Poudyal RS, Tiwari I, Voloshin RA, Zharmukhamedov SK, Nam HG, et al. Biofuel production: Challenges and opportunities. *Int J Hydrog Energy*. 23 de março de 2017;42(12):8450–61.

6. Kemper J. Biomass and carbon dioxide capture and storage: A review. *Int J Greenh Gas Control*. 1º de setembro de 2015;40:401–30.
7. Ranjan M, Herzog HJ. Feasibility of air capture. *Energy Procedia*. 1º de janeiro de 2011;4:2869–76.
8. Gough C, Upham P. Biomass energy with carbon capture and storage (BECCS or Bio-CCS). *Greenh Gases Sci Technol*. 2011;1(4):324–34.
9. Azar C, Lindgren K, Obersteiner M, Riahi K, van Vuuren DP, den Elzen KMGJ, et al. The feasibility of low CO₂ concentration targets and the role of bio-energy with carbon capture and storage (BECCS). *Clim Change*. 1º de maio de 2010;100(1):195–202.
10. McKendry P. Energy production from biomass (part 1): overview of biomass. *Bioresour Technol*. 1º de maio de 2002;83(1):37–46.
11. Gray KA, Zhao L, Emptage M. Bioethanol. *Curr Opin Chem Biol*. 1º de abril de 2006;10(2):141–6.
12. Singh A, Pant D, Korres NE, Nizami AS, Prasad S, Murphy JD. Key issues in life cycle assessment of ethanol production from lignocellulosic biomass: Challenges and perspectives. *Bioresour Technol*. 1º de julho de 2010;101(13):5003–12.
13. Grandis A, Fortirer JS, Navarro BV, de Oliveira LP, Buckeridge MS. Biotechnologies to Improve Sugarcane Productivity in a Climate Change Scenario. *BioEnergy Res* [Internet]. 30 de agosto de 2023 [citado 19 de outubro de 2023]; Disponível em: <https://doi.org/10.1007/s12155-023-10649-9>
14. Thirugnanasambandam PP, Hoang NV, Henry RJ. The Challenge of Analyzing the Sugarcane Genome. *Front Plant Sci* [Internet]. 2018 [citado 19 de outubro de 2023];9. Disponível em: <https://www.frontiersin.org/articles/10.3389/fpls.2018.00616>
15. Martines-Filho J, Burnquist HL, Vian CEF. Bioenergy and the Rise of Sugarcane-Based Ethanol in Brazil. *Choices*. 2006;21(2):91–6.
16. Cursi DE, Hoffmann HP, Barbosa GVS, Bressiani JA, Gazaffi R, Chapola RG, et al. History and Current Status of Sugarcane Breeding, Germplasm Development and Molecular Genetics in Brazil. *Sugar Tech*. 1º de fevereiro de 2022;24(1):112–33.
17. Bordonal R de O, Carvalho JLN, Lal R, de Figueiredo EB, de Oliveira BG, La Scala N. Sustainability of sugarcane production in Brazil. A review. *Agron Sustain Dev*. 27 de fevereiro de 2018;38(2):13.
18. de Souza AP, Leite DCC, Pattathil S, Hahn MG, Buckeridge MS. Composition and Structure of Sugarcane Cell Wall Polysaccharides: Implications for Second-Generation Bioethanol Production. *BioEnergy Res*. 1º de junho de 2013;6(2):564–79.
19. Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content - Waclawovsky - 2010 - *Plant Biotechnology Journal* - Wiley Online Library [Internet]. [citado 21 de maio de 2023]. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/j.1467-7652.2009.00491.x>

20. Mohan C, Easterling M, Yau YY. Gene Editing Technologies for Sugarcane Improvement: Opportunities and Limitations. *Sugar Tech.* 1º de fevereiro de 2022;24(1):369–85.
21. Demirbas A. Progress and recent trends in biofuels. *Prog Energy Combust Sci.* 1º de fevereiro de 2007;33(1):1–18.
22. Schmatz AA [UNESP. Synergy between chemical additives and pretreatments of sugarcane biomass favoring enzymatic hydrolysis and production of fuel ethanol. Sinergia entre aditivos químicos e pré-tratamentos da biomassa de cana-de-açúcar favorecendo hidrólise enzimática e produção de etanol de combustível [Internet]. 26 de novembro de 2020 [citado 17 de maio de 2023]; Disponível em: <https://repositorio.unesp.br/handle/11449/204238>
23. Yuan JS, Tiller KH, Al-Ahmad H, Stewart NR, Stewart CN. Plants to power: bioenergy to fuel the future. *Trends Plant Sci.* 1º de agosto de 2008;13(8):421–9.
24. Fischer RA (Tony), Edmeades GO. Breeding and Cereal Yield Progress. *Crop Sci.* 2010;50(S1):S-85-S-98.
25. Slewinski TL. Non-structural carbohydrate partitioning in grass stems: a target to increase yield stability, stress tolerance, and biofuel production. *J Exp Bot.* 1º de agosto de 2012;63(13):4647–70.
26. Martínez-Noël GMA, Tognetti JA. Chapter 22 - Sugar Signaling Under Abiotic Stress in Plants. Em: Ahmad P, Ahanger MA, Singh VP, Tripathi DK, Alam P, Alyemeni MN, organizadores. *Plant Metabolites and Regulation Under Environmental Stress* [Internet]. Academic Press; 2018 [citado 21 de maio de 2023]. p. 397–406. Disponível em: <https://www.sciencedirect.com/science/article/pii/B9780128126899000224>
27. Rolland F, Baena-Gonzalez E, Sheen J. SUGAR SENSING AND SIGNALING IN PLANTS: Conserved and Novel Mechanisms. *Annu Rev Plant Biol.* 2006;57(1):675–709.
28. Lastdrager J, Hanson J, Smeeckens S. Sugar signals and the control of plant growth and development. *J Exp Bot.* 1º de março de 2014;65(3):799–807.
29. Lunn JE. Sucrose Metabolism. Em: *Encyclopedia of Life Sciences* [Internet]. John Wiley & Sons, Ltd; 2016 [citado 19 de outubro de 2023]. p. 1–9. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470015902.a0021259.pub2>
30. Ruan YL. Sucrose Metabolism: Gateway to Diverse Carbon Use and Sugar Signaling. *Annu Rev Plant Biol.* 2014;65(1):33–67.
31. Bioinformatic analyses to uncover genes involved in trehalose metabolism in the polyploid sugarcane | *Scientific Reports* [Internet]. [citado 21 de maio de 2023]. Disponível em: <https://www.nature.com/articles/s41598-022-11508-x>
32. Fichtner F, Lunn JE. The Role of Trehalose 6-Phosphate (Tre6P) in Plant Metabolism and Development. *Annu Rev Plant Biol.* 2021;72(1):737–60.
33. Paul MJ, Primavesi LF, Jhurrea D, Zhang Y. Trehalose Metabolism and Signaling. *Annu Rev Plant Biol.* 2008;59(1):417–41.

34. Goddijn OJM, Dun K van. Trehalose metabolism in plants. *Trends Plant Sci.* 1º de agosto de 1999;4(8):315–9.
35. Sugars and Circadian Regulation Make Major Contributions to the Global Regulation of Diurnal Gene Expression in Arabidopsis | The Plant Cell | Oxford Academic [Internet]. [citado 19 de outubro de 2023]. Disponível em: <https://academic.oup.com/plcell/article/17/12/3257/6114688>
36. Yang HL, Liu YJ, Wang CL, Zeng QY. Molecular Evolution of *Trehalose-6-Phosphate Synthase (TPS)* Gene Family in Populus, Arabidopsis and Rice. *PLOS ONE.* 8 de agosto de 2012;7(8):e42438.
37. Henry C, Bledsoe SW, Siekman A, Kollman A, Waters BM, Feil R, et al. The trehalose pathway in maize: conservation and gene regulation in response to the diurnal cycle and extended darkness. *J Exp Bot.* 1º de novembro de 2014;65(20):5959–73.
38. Lunn JE. Gene families and evolution of trehalose metabolism in plants. *Funct Plant Biol.* 1º de junho de 2007;34(6):550–63.
39. Ramon M, De Smet I, Vandesteene L, Naudts M, Leyman B, Van Dijck P, et al. Extensive expression regulation and lack of heterologous enzymatic activity of the Class II trehalose metabolism proteins from *Arabidopsis thaliana*. *Plant Cell Environ.* 2009;32(8):1015–32.
40. Vandesteene L, López-Galvis L, Vanneste K, Feil R, Maere S, Lammens W, et al. Expansive Evolution of the *TREHALOSE-6-PHOSPHATE PHOSPHATASE* Gene Family in Arabidopsis. *Plant Physiol.* 1º de outubro de 2012;160(2):884–96.
41. Avonce N, Leyman B, Mascorro-Gallardo JO, Van Dijck P, Thevelein JM, Iturriaga G. The Arabidopsis *Trehalose-6-P Synthase AtTPS1* Gene Is a Regulator of Glucose, Abscisic Acid, and Stress Signaling. *Plant Physiol.* 1º de novembro de 2004;136(3):3649–59.
42. Tale of Two Sugars: Trehalose 6-Phosphate and Sucrose | Plant Physiology | Oxford Academic [Internet]. [citado 21 de maio de 2023]. Disponível em: <https://academic.oup.com/plphys/article/172/1/7/6115767>
43. Lunn JE, Delorge I, Figueroa CM, Van Dijck P, Stitt M. Trehalose metabolism in plants. *Plant J.* 2014;79(4):544–67.
44. Yadav UP, Ivakov A, Feil R, Duan GY, Walther D, Giavalisco P, et al. The sucrose–trehalose 6-phosphate (Tre6P) nexus: specificity and mechanisms of sucrose signalling by Tre6P. *J Exp Bot.* 1º de março de 2014;65(4):1051–68.
45. Martins MCM, Hejazi M, Fettke J, Steup M, Feil R, Krause U, et al. Feedback Inhibition of Starch Degradation in Arabidopsis Leaves Mediated by Trehalose 6-Phosphate. *Plant Physiol.* 1º de novembro de 2013;163(3):1142–63.
46. O'Neill BP, Purnell MP, Nielsen LK, Brumbley SM. RNAi-mediated abrogation of trehalase expression does not affect trehalase activity in sugarcane. *SpringerPlus.* 21 de dezembro de 2012;1(1):74.

47. Gabriel C, Fernhout JJ, Fichtner F, Feil R, Lunn JE, Kossmann J, et al. Genetic manipulation of *trehalose-6-phosphate synthase* results in changes in the soluble sugar profile in transgenic sugarcane stems. *Plant Direct*. 2021;5(11):e358.
48. Souza APD, Grandis A, Arenque-Musa BC, Buckeridge MS, Souza APD, Grandis A, et al. Diurnal variation in gas exchange and nonstructural carbohydrates throughout sugarcane development. *Funct Plant Biol*. 27 de março de 2018;45(8):865–76.
49. Mattiello L, Riaño-Pachón DM, Martins MCM, da Cruz LP, Bassi D, Marchiori PER, et al. Physiological and transcriptional analyses of developmental stages along sugarcane leaf. *BMC Plant Biol*. 29 de dezembro de 2015;15(1):300.
50. Vettore AL, da Silva FR, Kemper EL, Souza GM, da Silva AM, Ferro MIT, et al. Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. *Genome Res*. dezembro de 2003;13(12):2725–35.
51. Bushnell B, Rood J, Singer E. BBMerge – Accurate paired shotgun read merging via overlap. *PLOS ONE*. 26 de outubro de 2017;12(10):e0185056.
52. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. abril de 2017;14(4):417–9.
53. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol*. 1º de julho de 2021;38(7):3022–7.
54. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol Biol Evol*. 1º de janeiro de 2015;32(1):268–74.
55. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst Biol*. 1º de maio de 2010;59(3):307–21.
56. Souza GM, Van Sluys MA, Lembke CG, Lee H, Margarido GRA, Hotta CT, et al. Assembly of the 373k gene space of the polyploid sugarcane genome reveals reservoirs of functional diversity in the world's leading biomass crop. *GigaScience*. 1º de dezembro de 2019;8(12):giz129.