



ANA BEATRIZ MONTEIRO

**CARACTERIZAÇÃO DO GENE *bZIP* EM TOMATEIRO:
DELEÇÃO DE UMA *uORF* CONSERVADA E ANÁLISE *in silico*
DE CORRELAÇÕES DO GENE *SlbZIP1***

LAVRAS-MG

2023

ANA BEATRIZ MONTEIRO

**CARACTERIZAÇÃO DO GENE bZIP EM TOMATEIRO: DELEÇÃO DE UMA *uORF*
CONSERVADA E ANÁLISE *in silico* DE CORRELAÇÕES DO GENE *SibZIP1***

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Biotecnologia Vegetal, área de concentração em Biotecnologia Vegetal, para a obtenção do título de Mestre.

Prof.^a Dr.^a. Kalynka Gabriella do Livramento
Orientadora

Prof. Dr. Luciano Vilela Paiva
Coorientador

Dr. Renan Terassi Pinto
Coorientador

LAVRAS-MG

2023

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Monteiro, Ana Beatriz.

Caracterização do gene bZIP em tomateiro : Deleção de uma
uORF conservada e análise *in silico* de correlações do gene *SlbZIP1*
/ Ana Beatriz Monteiro. - 2023.

94 p.

Orientador(a): Kalyнка Gabriella do Livramento.

Coorientador(a): Luciano Vilela Paiva, Renan Terassi Pinto.

Dissertação (mestrado acadêmico) - Universidade Federal de
Lavras, 2023.

Bibliografia.

1. CRISPR/Cas9. 2. Fator de transcrição. 3. *Solanum*
lycopersicum. I. Livramento, Kalyнка Gabriella do. II. Paiva,
Luciano Vilela. III. Pinto, Renan Terassi. IV. Título.

ANA BEATRIZ MONTEIRO

CARACTERIZAÇÃO DO GENE *bZIP* EM TOMATEIRO: DELEÇÃO DE UMA *uORF* CONSERVADA E ANÁLISE *in silico* DE CORRELAÇÕES DO GENE *SibZIP1*

CHARACTERIZATION OF THE *bZIP* GENE IN TOMATO: DELETION OF A CONSERVED *uORF* AND IN SILICO ANALYSIS OF CORRELATIONS OF THE *SibZIP1* GENE.

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Biotecnologia Vegetal, área de concentração em Biotecnologia Vegetal, para obtenção do título de mestre.

APROVADA em 18 de janeiro de 2023.

Dr. Luciano Vilela Paiva, UFLA

Dr. Filippe Elias de Freitas Soares, UFLA

Dr^a. Natália Chagas Freitas, Exitus Acadêmico

Prof.^a Dr^a. Kalynka Gabriella do Livramento
Orientadora

Prof. Dr. Luciano Vilela Paiva
Coorientador

Dr. Renan Terassi Pinto
Coorientador

**LAVRAS-MG
2023**

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG), a qual agradeço pela concessão da bolsa de estudos e às demais agências de fomento pelo financiamento concedido que contribuíram para realização deste projeto.

Agradeço a Universidade Federal de Lavras e ao Programa de Pós-Graduação em Biotecnologia Vegetal pela oportunidade de complementar minha formação acadêmica através do mestrado.;

Agradeço a orientadora Kalynka e ao coorientador Luciano pela orientação, pelo aprendizado, apoio e paciência que foram essenciais para a condução do projeto e ao coorientador Renan por sempre ter sido acolhedor e ter contribuído tão ativamente para o desenvolvimento deste trabalho;

A minha experiência no laboratório sempre foi agradável e divertida, mesmo com tantos desafios que surgiram ao longo da pesquisa. Por isso, agradeço aos colegas do LCBM pela colaboração na execução das atividades, pelos ensinamentos e conselhos ou mesmo pela amizade que foi construída com tantas pessoas. Em especial gostaria de agradecer a Maria Clara e ao Guilherme com quem dividi a bancada e os experimentos com os tomatinhos;

Agradeço as pessoas que surgiram na minha vida pelo meio acadêmico, à Giovanna com quem partilho ideias desde a iniciação científica. A Lorrana pela amizade e contribuição com o trabalho, sou muito grata pela disposição que sempre teve para me ensinar;

Agradeço a minha família a vovó Lurdes e ao vovô Hélio, aos meus tios Jacque, Thiago e Patrícia e a minha mamãe Eliane, que sempre me apoiaram e me incentivaram ao longo da jornada e, não pouparam esforços e orações para que eu pudesse chegar até aqui. Não posso deixar de mencionar os pequenos João Lucas e Luísa, por me lembrarem que eu posso encarar a vida com mais leveza e menos seriedade, espero ser sempre um exemplo para vocês;

Agradeço aos meus amigos pelo apoio e torcida, em especial a Laiza e ao Augusto, que mesmo de longe se fizeram presente durante as nossas preciosas conversas, muito obrigada pelo apoio moral;

Agradeço ao Miguel, meu amor, que esteve presente em todos os momentos, obrigada por sempre me encorajar a conquistar os meus sonhos e por ser tão dedicado e paciente;

Por fim, agradeço a todos aqueles que direta ou indiretamente contribuíram para a construção deste trabalho. Obrigada!

RESUMO

O tomate cultivado, *Solanum lycopersicum*, pertence à família Solanaceae que possui muitas espécies de importância econômica. Devido a caracterização da espécie o tomateiro além de produto comercial, também tem sido utilizado como planta modelo para o desenvolvimento de pesquisas de investigação dos fenômenos vegetais, fisiológicos e metabólicos. Em virtude do consumo elevado e para atender a demanda de produtividade, de resistência e de outras características, análises de condições que possam ser utilizadas para o melhoramento do tomateiro têm sido requisitadas. Alguns açúcares, como a sacarose, que conferem sabor adocicado aos tomates também são moléculas sinalizadoras que podem regular a expressão gênica, pelo processo de repressão da tradução induzida por sacarose em fatores de transcrição do tipo bZIP (*basic leucine zipper*). Os bZIPs estão associados a funções fundamentais na sinalização ambiental e de desenvolvimento, atuando na regulação metabólica central de carbono. Sabendo que modificações em regiões específicas dos bZIPs de classe S podem alterar a regulação pós-transcricional, o uso de ferramentas de engenharia genética para produzir uma mutagênese direcionada tem sido utilizada como estratégia. O objetivo deste trabalho foi caracterizar o gene *SlbZIP1*, utilizando a técnica CRISPR/Cas9 para a deleção de uma região *uORF* (*upstream open reading frame*) envolvida na regulação deste gene e, a identificação de uma rede de correlação de expressão de genes relacionados ao *SlbZIP1*, obtidos por meio de bibliotecas de RNA-seq para compreender suas interações gênicas. A transformação das plantas foi realizada via *Agrobacterium tumefaciens* e os eventos obtidos foram selecionados através da resistência ao agente seletivo canamicina, sendo validadas por PCR. Para a caracterização molecular da edição gênica, os fragmentos foram amplificados e sequenciados. Foi observado uma deleção de 42 pb da *uORF*, compatível com a região alvo e com a proposta apresentada. Este resultado demonstra que a estratégia de edição gênica apresentada neste estudo é efetiva e passível de ser utilizada. Quanto às análises *in silico*, foi realizada a prospecção em estudos já publicados de dados de RNA-seq em tomateiro. Foram selecionadas 193 bibliotecas classificadas em diferentes tecidos e submetidas a condições diversas. Os dados foram tratados, alinhados e a correlação de expressão entre os todos os transcritos obtidos e o gene *SlbZIP1* foi estabelecida pelo teste de *Spearman*, com um coeficiente de correlação $\rho \geq 0,8$. Foram consideradas 85 sequências de diversas funções, onde se destacaram outros fatores de transcrição e proteínas quinases. Estes resultados corroboram com demais estudos, que indicam que as proteínas quinases atuam sinergicamente com fatores específicos de transcrição e atuam em atividades importantes para a tolerância e resposta à diversas condições ambientais. Demais estudos poderão ser conduzidos com intuito de ampliar o conhecimento a respeito do gene *SlbZIP1*.

Palavras-chave: CRISPR/Cas9. Fator de transcrição. *Solanum lycopersicum*.

ABSTRACT

The cultivated tomato, *Solanum lycopersicum*, belongs to the family Solanaceae that has many species of economic relevance. Due to the characterization the tomato, besides being a commercial product, has also been used as a model plant for the development of research to investigate plant physiological and metabolic processes. Because of the high consumption and to meet the demand for productivity, resistance and other characteristics, analyses of conditions that can be used for tomato breeding have been requested. Some sugars, like sucrose, that confer sweet taste to tomatoes, are also signaling molecules that can regulate gene expression, by the process of sucrose-induced repression of translation, in bZIP (basic leucine zipper) transcription factors. The bZIPs are associated with fundamental functions in environmental and developmental signaling, acting in central carbon metabolic regulation. In the knowledge that modifications in specific regions of S-class bZIPs can alter post-transcriptional regulation, the use of genetic engineering tools to produce targeted mutagenesis has been used as a strategy. The aim of this work was to characterize the *SlbZIP1* gene, using the CRISPR/Cas9 technique for the deletion of a *uORF* (upstream open reading frame) region involved in the regulation of this gene and the identification of an expression correlation network of *SlbZIP1*-related genes obtained through RNA-seq libraries to understand their interactions. Plant transformation was performed via *Agrobacterium tumefaciens*, and the obtained events were selected through resistance to the selective agent kanamycin and validated by PCR. For the molecular characterization of the gene editing, the fragments were amplified and sequenced. A 42 bp deletion of uORF was observed, compatible with the target region and the suggested proposal. This result demonstrates that the gene editing strategy presented in this study is effective and usable. As for the in-silico analyses, a search in already published studies of RNA-seq data in tomato plants was performed. A total of 193 libraries sorted into different tissues and subjected to various conditions were selected. The data were processed, aligned, and the expression correlation between all obtained transcripts and the *SlbZIP1* gene was established by Spearman's test, with a correlation coefficient $\rho \geq 0.8$. A total of 85 sequences of various functions were considered, where other transcription factors and protein kinases stood out. These results agree with other studies, which indicate that protein kinases act synergistically with specific transcription factors and perform important activities for tolerance and response to various environmental conditions. Further studies can be conducted to increase the knowledge about the *SlbZIP1* gene.

Keywords: CRISPR/Cas9. Transcription factor. *Solanum lycopersicum*.

SUMÁRIO

	CAPÍTULO 1 - INTRODUÇÃO GERAL	10
1	INTRODUÇÃO	11
2	REFERENCIAL TEÓRICO	12
2.1	A espécie <i>Solanum lycopersicum</i> e sua importância econômica	12
2.2	A utilização do tomateiro como planta modelo	14
2.3	Influência do melhoramento genético no sabor do tomate	16
2.4	Composição do sabor do tomate	17
2.5	Os açúcares na planta: metabolismo e sinalização	18
2.6	Controle da expressão gênica por fatores de transcrição da família bZIP	21
2.7	S1-bZIPs	23
2.8	Estrutura <i>uORFs</i> únicas e mecanismo de regulação dos S1-bZIPs	24
2.9	Edição gênica pelo sistema CRISPR/Cas9	27
	REFERÊNCIAS	31
	CAPÍTULO 2 - DELEÇÃO DE UMA <i>uORF</i> CONSERVADA DO GENE bZIP EM TOMATEIRO VIA CRISPR/Cas9	37
1	INTRODUÇÃO	39
2	MATERIAIS E MÉTODOS	40
2.1	Material Vegetal	40
2.2	Construção gênica	40
2.3	Preparo da <i>Agrobacterium tumefaciens</i>	41
2.4	Transformação e regeneração	42
2.5	Extração de DNA-genômico	44
2.6	Validação molecular, detecção da mutação e sequenciamento	45
2.7	Quantificação de açúcares dos frutos	46
3	RESULTADOS E DISCUSSÃO	47
3.1	Transformação e regeneração	47
3.2	Planta T1	48
3.2.1	Geração F1	49
3.3	Plantas T2 e T3	51
3.4	Detecção e sequenciamento das mutações	54
3.5	Geração F1 proveniente da planta editada	57

4	CONSIDERAÇÕES FINAIS.....	58
5	CONCLUSÃO.....	58
	REFERÊNCIAS.....	60
	APÊNDICES	63
	APÊNDICE A - Teste de resistência a canamicina.....	63
	APÊNDICE B - Abortamento das flores e presença de tricomas	64
	APÊNDICE C - Resultado de Sequenciamento	65
	CAPÍTULO 3 - ANÁLISE DE GENES CORRELACIONADOS AO <i>SibZIP166</i>	
	RESUMO.....	67
1	INTRODUÇÃO	68
2	MATERIAIS E MÉTODOS.....	69
2.1	Seleção das bibliotecas e tratamento dos dados.....	69
2.2	Rede de correlação e caracterização <i>in silico</i>	70
3	RESULTADOS E DISCUSSÃO	71
3.1	Seleção das bibliotecas	71
3.2	Rede de correlação.....	71
3.3	Perspectivas quanto a utilização do banco de dados gerado	76
4	CONSIDERAÇÕES FINAIS.....	77
	REFERÊNCIAS.....	78
	APÊNDICES	81
	APÊNDICE A - Dados obtidos de bibliotecas selecionadas pelo banco de dados NCBI	81
	APÊNDICE B - Tabela descritiva dos 85 transcritos obtidos pela rede de correlação	87
	APÊNDICE C - Tabela de fatores de transcrição	92
	APÊNDICE D - Correlação de expressão obtida.....	93

CAPÍTULO 1 – INTRODUÇÃO GERAL

1 INTRODUÇÃO

A espécie *Solanum lycopersicum*, conhecida popularmente como tomateiro, é diploide ($2n = 24$) e faz parte da família das solanáceas, importante grupo que contém diversas plantas cultivadas de interesse comercial. Em relatório emitido pela Organização das Nações Unidas para a Alimentação e Agricultura (FAO), observou-se que a produção de tomates foi de 189,13 milhões de toneladas no mundo no ano de 2020. A alta demanda é devido a sua versatilidade, visto que os tomates são amplamente incluídos na dieta e estão presentes em diversas receitas seja como produto processado ou *in natura* (WANG; SEYMOUR, 2017a; LI et al., 2018a; QUINET et al., 2019).

Durante o processo de melhoramento e seleção dos tomateiros, características de produtividade, resistência a doenças e modificações do tamanho e forma dos frutos foram priorizadas, resultando em cultivares com diferentes morfologias (BERGOUGNOUX, 2014; LIU et al., 2020). Entretanto, o processo contribuiu para a redução da diversidade genética de características que estão associadas ao sabor dos frutos (GAO, 2019). O sabor do tomate pode ser influenciado pelo teor de açúcares, de ácidos, de sais minerais e de compostos voláteis (WANG; SEYMOUR, 2017a). Substâncias como a sacarose, a frutose e a glicose são responsáveis por conferir sabor mais adocicado e são os açúcares utilizados pela planta na síntese de compostos essenciais (MA et al., 2011a; YU; LIN; LI, 2016).

Os açúcares desempenham importantes funções no fornecimento de carbono às plantas, além de contribuírem para a percepção do sabor que proporcionam, são também moléculas sinalizadoras envolvidas na regulação de expressão de fatores de transcrição (FTs) e de outros genes de sinalização hormonal, que estão relacionados ao desenvolvimento e ao metabolismo, respondendo de acordo com a capacidade de ajuste dos níveis de expressão gênica (FICHTNER et al., 2021).

Quanto aos FTs, é importante ressaltar que estes desempenham atividades cruciais em quase todos os processos biológicos. A família gênica *basic leucine zipper*, bZIPs, são FTs associados a diversas funções centrais na sinalização ambiental e de desenvolvimento (JAKOBY et al., 2002; DRÖGE-LASER; WEISTE, 2018a). De maneira geral, os bZIPs modulam o desenvolvimento vegetal, no reajuste metabólico durante a sinalização de baixa energia ou de deficiência de nutrientes, como pode acontecer em períodos de escuro prolongado, estes FTs podem regular múltiplos genes, ter efeito ativador, repressor ou ambos (ALI et al., 2016; DRÖGE-LASER; WEISTE, 2018b).

Membros da família bZIP, do grupo S, possuem um mecanismo que atua em nível pós transcricional, através da repressão da tradução induzida por sacarose (SIRT). Dentre as particularidades deste grupo, pode-se destacar o sistema de estruturas abertas de leitura a montante, *uORF* (*upstream open-reading frames*), que são estruturas conservadas que permitem o controle da expressão gênica (ROOK et al., 1998; JAKOBY et al., 2002; MA et al., 2011a).

A crescente demanda por melhoramento de diversas espécies de plantas exige técnicas que sejam precisas e possam ser realizadas em tempo reduzido para que se possa obter novas plantas com características desejadas. A edição genômica, quando comparada as técnicas convencionais, oferece essas vantagens e pode ser aplicada para caracterização de funções gênicas e melhoramento de características agrícolas (ZHANG et al., 2018). A mutagênese direcionada pelo sistema CRISPR/Cas9 tornou-se amplamente utilizada nas plantas com diversas aplicações, incluindo o aumento do rendimento e da qualidade dos frutos, aumento da resistência a estresses, otimização do processo de domesticação de plantas selvagens e, recentemente, para personalizar cultivares para a agricultura urbana (XIA et al., 2021b).

Diante disso, o gene *SibZIP1* de *S. lycopersicum*, um bZIP de classe S foi o alvo de estudo deste trabalho. Sabendo que variações em regiões *uORFs* do gene *SibZIP1* podem alterar o teor de açúcares (XING et al., 2020a), um dos objetivos desse trabalho foi realizar uma mutação direcionada, como forma de impedir o mecanismo de repressão SIRT, utilizando a técnica CRISPR/Cas9, conforme apresentado no capítulo 02. E, visando melhor compreensão das interações de genes associados ou relacionados aos padrões de expressão do gene *SibZIP1*, foram prospectados e selecionadas bibliotecas de dados públicos de RNA-seq, com intuito de gerar uma rede de correlação de expressão de genes em diversos tecidos e condições, para comparações e estudos, como descrito no capítulo 03.

2 REFERENCIAL TEÓRICO

2.1 A espécie *Solanum lycopersicum* e sua importância econômica

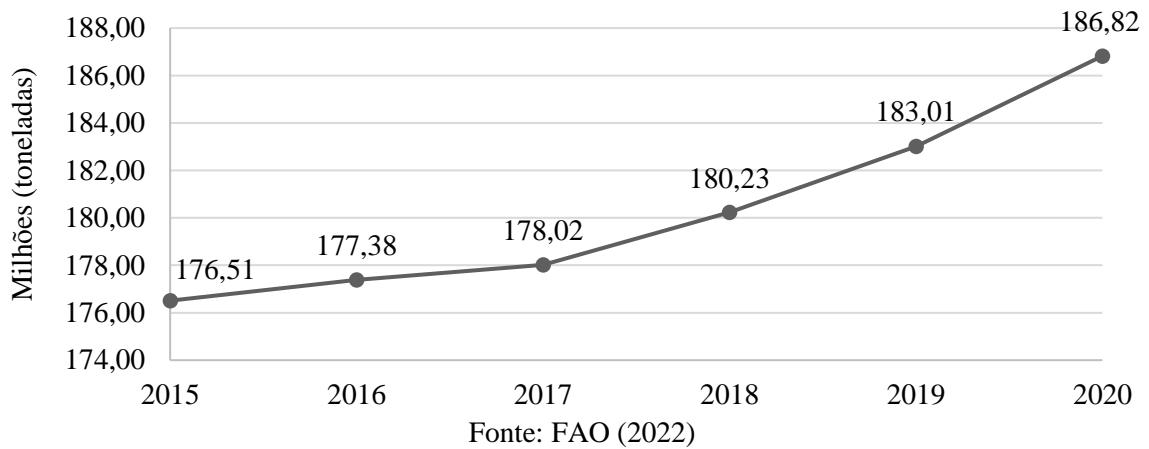
A espécie *S. lycopersicum*, o tomateiro, é derivada de uma espécie ancestral de origem tropical americana região onde estão localizados os países Peru, Equador e Chile (JENKINS, 1948; KALLOO, 1993; BERGOUGNOUX, 2014). Os primeiros relatos do tomateiro

domesticado datam da era pré-colombiana, no qual é descrito que a planta foi propagada para a América Central e cultivada em campos de milho no México, cujo frutos eram conhecidos pelos povos antigos de “*tomatl*”, uma palavra que se refere a uma planta que produz frutos suculentos e redondos (BOSWELL, 1937). O tomateiro chegou ao continente europeu no século XVI, quando passou a ser amplamente distribuído pelo mundo (MATA-NICOLÁS et al., 2020).

O tomate cultivado é diploide ($2n = 24$) e pertence à família Solanaceae, que contém mais de três mil espécies, incluindo plantas de importância econômica como batatas, berinjelas, petúnias, tabaco, pimenta e *Physalis* (KIMURA; SINHA, 2008a). O gênero *Solanum* representa o maior grupo da família das solanáceas e estão distribuídos mundialmente, destacando-se por sua diversidade morfológica e ecológica (KIMURA; SINHA, 2008a; BERGOUGNOUX, 2014)

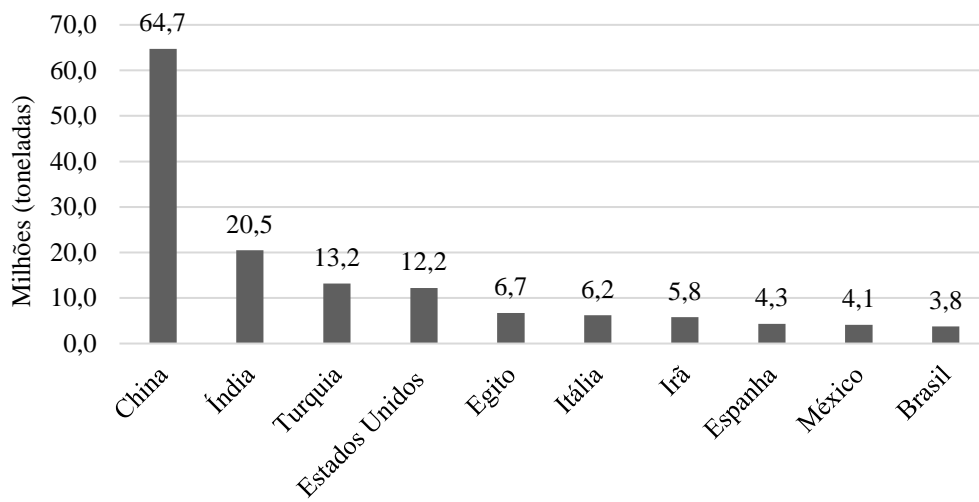
Economicamente, o tomate é um importante produto agrícola consumido mundialmente cuja produção tem aumentado anualmente (FIGURA 1.1). Em 2021 de acordo com o relatório emitido pela FAO, foram produzidos em torno de 189,13 milhões de toneladas de tomates no mundo. Devido a sua versatilidade, os tomates são amplamente consumidos, sejam os frutos frescos em saladas ou de maneira processada para fabricação de molhos, sopas, sucos, massas e temperos que compõem diversas receitas (QUINET et al., LI et al., 2018).

Figura 1.1 – Produção mundial de tomates entre os anos de 2015 à 2020



No Brasil, além de ser uma das hortaliças mais consumidas o tomate também representa um importante produto econômico (CONAB, 2019). Em 2021, a produção brasileira foi de 3,76 milhões de toneladas, gerando um valor de produtividade de 6,47 milhões de reais (IBGE, 2022). O Brasil está entre os dez maiores produtores, entretanto a China destaca-se no mercado representando 34,66% da produção mundial (FIGURA 1.2) (FAO, 2022).

Figura 1.2 – Principais países produtores de tomate em 2020



Fonte: FAO (2022)

2.2 A utilização do tomateiro como planta modelo

Por ser uma planta dicotiledônea bem caracterizada, o tomateiro tem sido bastante estudado e é de interesse como espécie vegetal modelo, uma vez que seu estudo permite a investigação de vários fenômenos fisiológicos (CARVALHO et al., 2011). As principais vantagens para o uso do tomateiro em pesquisas referem-se ao genoma relativamente compacto (900 Mb cv “Heinz 1706”), pela presença de marcadores genéticos caracterizados, pela existência de coleções bem estabelecidas de germoplasma, pela disponibilidade de protocolos de transformação de alta eficiência e, ao contrário de outros organismos modelos, o tomateiro possui características interessantes do ponto de vista agrônomo como a produção folhas compostas e frutos carnosos, que são consumidos na dieta humana (KIMURA; SINHA, 2008a; GERSZBERG et al., 2015; LIU et al., 2022).

A espécie *S. lycopersicum* contém características como a adaptabilidade a diferentes condições de fotoperíodo e estresses abióticos, tem o ciclo de vida relativamente curto, permite a polinização controlada e a hibridação, apresenta capacidade para propagação assexuada por enxertia ou regeneração a partir do cultivo *in vitro*, ou seja, possui vários fatores que contribuem para a sua utilização como planta modelo (BERGOUGNOUX, 2014; GERSZBERG et al., 2015; MATA-NICOLÁS et al., 2020). Ademais, os resultados obtidos em pesquisas com o tomateiro podem ser reproduzidos e transferidos para outras espécies vegetais (KIMURA; SINHA, 2008a; LIU et al., 2022).

O tomateiro possui o sequenciamento completo do genoma que foi disponibilizado em 2012 referente a cultivar "Heinz 1706" (TOMATO GENOME CONSORTIUM, 2012). Esse genoma de referência contribuiu para o desenvolvimento de novos dados e informações que podem ser consultados em plataformas como *Sol Genomics Network* (<https://solgenomics.net/>), *The Tomato Genomics Resources Database* (<https://59.163.192.91/tomato2/>) e *Tomato Genetic Resource Center* (<https://tgrc.ucdavis.edu/>).

Atualmente, existem algumas variedades de tomateiros mutantes que foram produzidos de forma espontânea ou por indução, disponibilizados como fonte de recursos para os pesquisadores isolarem e compreenderem funções gênicas que regulam o desenvolvimento e o crescimento (BERGOUGNOUX, 2014). Dentre as cultivares a cv Micro-tom é bastante utilizada para delineamentos experimentais, principalmente por apresentar um ciclo de crescimento e de tamanho reduzido, pela facilidade de propagação por autofecundação e pela suscetibilidade aos protocolos de transformação genética, que fornecem informações que podem ser validadas em variedades comerciais (CAMPOS et al., 2010).

2.3 Influência do melhoramento genético no sabor do tomate

Os tomateiros são autógamos, por isso, ao longo do histórico de cultivo o cruzamento entre dois indivíduos diferentes era raro e o desenvolvimento de plantas a partir da germinação de sementes proporcionava o fenótipo parental (BERGOUGNOUX, 2014; MATA-NICOLÁS et al., 2020). No entanto, com o avanço das indústrias de insumos agrícolas surgiram as sementes de tecnologia híbrida. Os híbridos combinam as características agronômicas dos dois progenitores e esses caracteres são segregados na progênie, dessa maneira não é possível controlar a padronização das características por meio da coleta de sementes e posterior exploração no campo (BERGOUGNOUX, 2014).

Os programas de melhoramento genético de *S. lycopersicum* avançaram em vários níveis quanto a quantidade de cultivares com diferentes características morfológicas. Do ponto de vista evolutivo, os programas de domesticação induziram mudanças fisiológicas consideráveis, entretanto a seleção artificial reduziu a diversidade genética do tomateiro cultivado (BERGOUGNOUX, 2014; GAO et al., 2019). Estima-se que a espécie cultivada de *S. lycopersicum* contenha apenas 5% da variação genética total existente em todas as espécies de tomateiro que surgiram durante sua domesticação (TIEMAN et al., 2017b; GAO et al., 2019; SALAVA et al., 2021). A diversidade genética limitada justifica o uso de engenharia genética para ampliar as possibilidades de enriquecimento da cultura, como mudanças relacionadas ao sabor dos frutos, que pode ser limitado por técnicas de melhoramento genético convencionais (BERGOUGNOUX, 2014; GAO et al., 2019; SALAVA et al., 2021).

As mudanças mais acentuadas observadas durante a domesticação e melhoramento, são as qualidades intrínsecas do fruto do tomateiro como o tamanho, a forma, a coloração, a firmeza e o tempo de prateleira (GAO et al., 2019). As espécies silvestres de tomate apresentam frutos menores quando comparados aos tomates cultivados modernos, contudo as cultivares atuais apresentam uma grande variação quanto ao tamanho, forma e coloração dos frutos (BERGOUGNOUX, 2014; LIU et al., 2020).

Em uma pesquisa apresentada sobre o pan-genoma do tomateiro construído a partir de 725 acessos, filogeneticamente e geograficamente representativos, foram apresentados 4873 genes que estavam ausentes no genoma de referência, cv. "Heinz 1706". Os autores enfatizam que parte da redução da diversidade genética pode ser atribuída a perdas substanciais de genes

durante os processos de domesticação da espécie. Como o genoma de referência trata-se de uma cultivar moderna, pelo menos alguns alelos associados ao sabor podem estar ausentes em seu genoma (GAO et al., 2019).

2.4 Composição do sabor do tomate

No processo de amadurecimento do tomate o acúmulo de açúcares e a degradação de ácidos orgânicos, bem como a proporção entre eles, afetam a qualidade do fruto (OSORIO; RUAN; FERNIE, 2014). Os fatores que mais influenciam no sabor e aroma dos tomates são os açúcares, os ácidos orgânicos, os sais minerais e os compostos voláteis (TIEMAN et al., 2017b; WANG; SEYMOUR, 2017b; GAO et al., 2019). Os frutos frescos do tomateiro são compostos de 90 a 95% de água, com concentração de aminoácidos variável entre 2,0 e 2,5%, quanto a matéria seca cerca de 50% do peso é constituído de açúcares combinados (BERTIN; GÉNARD, 2018).

Os principais açúcares solúveis encontrados nos frutos são a glicose e a frutose, formadas a partir da hidrólise de sacarose e são moléculas que desempenham papel fundamental para a percepção do sabor adocicado (TIEMAN et al., 2017b; WANG; SEYMOUR, 2017c; BERTIN; GÉNARD, 2018). Outros sólidos solúveis presente nos frutos são os ácidos orgânicos, sendo mais comuns o malato e o ácido cítrico (QUINET et al. 2019; BERGOUIGNOUX, 2014). Os aminoácidos como o glutamato, responsável pelo sabor “umami” ou salgado, a glicina, a alanina, a serina, a treonina, a prolina, a glutamina e a lisina que são associados a sabores mais adocicados, a fenilalanina e a tirosina que conferem amargor, também são compostos que interferem no sabor dos frutos (SAGOR et al., 2016).

Os frutos do tomateiro contêm diversos nutrientes que são associados a uma dieta equilibrada (TAB 1.1). Entre os principais compostos estão presentes as vitaminas, os carotenoides, o licopeno, os flavonoides e alguns aminoácidos como o glutamato, o aspartato e o ácido gama aminobutírico (GABA) (SNOWDEN et al., 2015; WANG; SEYMOUR, 2017b). Durante a última década tem-se observado que os consumidores estão mais conscientes da ingestão de alimentos como fonte de benefícios para a saúde e da possibilidade de utilizá-los na prevenção de doenças, além de demonstrarem maior interesse em escolher frutos com baixo teor de ácidos e de gosto mais doce. O aumento do brix, medida de teor de sólidos solúveis totais no fruto, tem sido alvo de pesquisas científicas para melhoria do sabor

(BERGOUGNOUX, 2014; TIEMAN et al., 2017b; WANG; SEYMOUR, 2017c; DE LAS MERCEDES CAPOBIANCO-URIARTE et al., 2021). Devido a importância da cultura do tomateiro algumas estratégias têm sido desenvolvidas para atender ao interesse comercial de maior produtividade, mais resistência a estresses bióticos ou abióticos, bem como para o aumento de compostos bioativos que alterem a cor, o sabor ou o aroma (TIEMAN et al., 2017b; LI et al., 2018b).

Tabela 1.1 – Composição dos frutos maduros de tomate

Componentes	Matéria seca (%)
Glicose	22
Frutose	25
Sacarose	1
Proteínas	8
Substâncias pécticas	7
Hemicelulose	4
Celulose	6
Ácido cítrico	9
Ácido málico	4
Principalmente K, Ca, Mg e P	8
Lipídios	2
Aminoácidos dicarboxílicos	2
Pigmentos	0,4
Ácido ascórbico	0,5
Voláteis	0,1
Outros aminoácidos, vitaminas e polifenóis	1

Fonte: Adaptado de Davies, Hobson e McGlasson (1981)

2.5 Os açúcares na planta: metabolismo e sinalização

As plantas coordenam o crescimento e o desenvolvimento de acordo com a disponibilidade de nutrientes. Durante o seu ciclo de vida, uma planta passa por diversas alterações na dinâmica de transporte e metabolismo de carbono (C), que controlam padrões cíclicos de crescimento, portanto o desenvolvimento das plantas e o metabolismo do C estão intimamente associados. Os carboidratos na forma de sacarose, gerados pela fotossíntese, representam a principal fonte de C e energia nas plantas, sendo fundamental para a produção e a manutenção da biomassa (OSORIO; RUAN; FERNIE, 2014).

O fluxo de exportação de carboidratos dos tecidos fotossintetizantes fornece o substrato para o crescimento e a manutenção de tecidos não fotossintéticos, ou seja, os tecidos “fonte” fotossinteticamente ativos, como as folhas maduras exportam o C fixado para tecidos “dreno” não fotossintéticos, como os meristemas, as raízes e demais partes da planta (OSORIO; RUAN; FERNIE, 2014). Os açúcares produzidos a partir da fotossíntese controlam os processos biológicos como a floração, o início da senescência, a embriogênese, a germinação de sementes, o desenvolvimento de mudas e a formação de tubérculos e de frutos (LUNN, 2002; YOON et al., 2021).

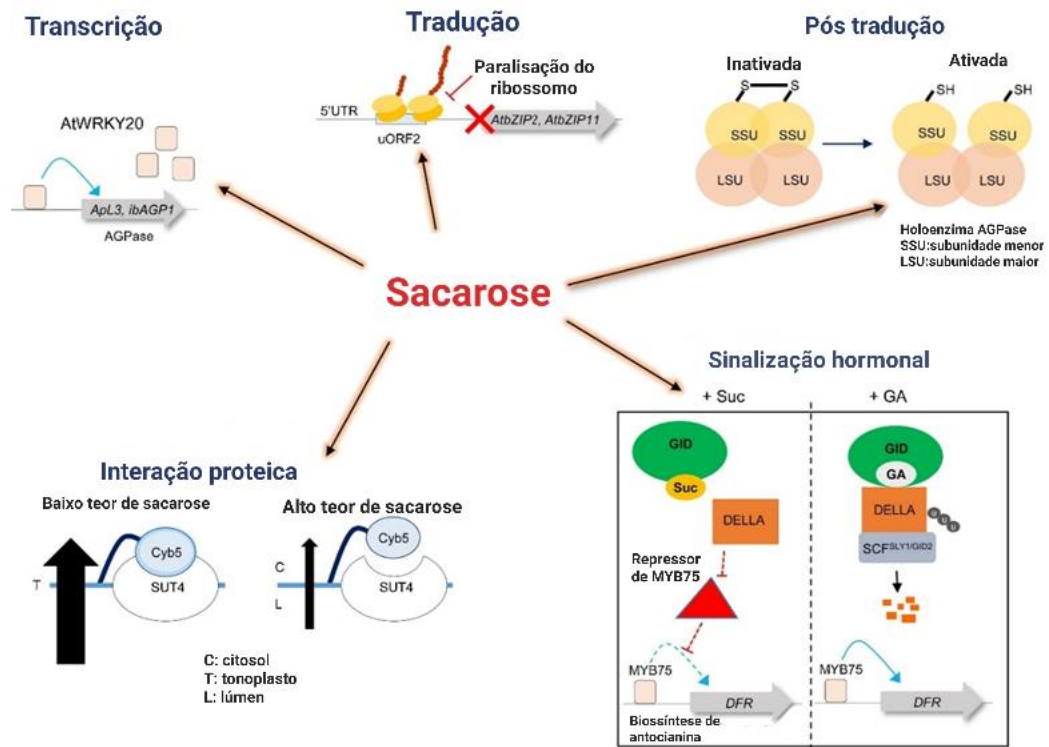
A molécula triose-P produzida pela fotossíntese é transportada para o citosol, onde ocorre a síntese da sacarose, que é translocada célula a célula, simplásticamente pela diferença de pressão hidrostática ou por via da rota apoplástica que envolve gasto de energia. Portanto, a translocação dos carboidratos como por exemplo da sacarose, está intimamente associada ao potencial osmótico das células e a manutenção do turgor celular no parênquima de armazenamento (PATRICK et al., 2013; GROF et al., 2014). Os açúcares como a sacarose contribuem não apenas para o sabor, mas também estão diretamente relacionados à qualidade dos frutos, uma vez que, o seu acúmulo nos vacúolos cria pressão osmótica que fornece turgescência para o aumento das células e de seu crescimento (STEIN; GRANOT, 2019).

O teor final de açúcares nos frutos pode ser determinado pela síntese, pelo transporte, pelo metabolismo e pelo armazenamento (OSORIO; RUAN; FERNIE, 2014). A sacarose produzida na fotossíntese é temporariamente armazenada em compartimentos vacuolares para ser utilizada em períodos de baixa ou nenhuma atividade fotossintética e é transportada para outros tecidos, como os frutos. A sacarose é a principal molécula de assimilação de C no fruto que é requisitado para a formação de esqueletos necessários para o seu desenvolvimento (TANAKA; FUJITA; KIKUCHI, 2012).

Para chegar as regiões dreno a sacarose é inicialmente hidrolisada em frutose e glicose, podendo ser sintetizada novamente antes de ser utilizada em outros processos metabólicos. A conversão de sacarose pode ser realizada por dois diferentes grupos enzimáticos como as invertases, que catalisam uma reação irreversível em que a quebra da sacarose resulta em açúcares redutores ou por meio de sintases, que catalisam a conversão reversível de sacarose em UDP glicose e frutose. As duas rotas de degradação são atribuídas a diferentes funções ao longo do desenvolvimento da planta e moduladas em respostas a fatores ambientais (SUNG; XU; BLACK, 1989; OSORIO; RUAN; FERNIE, 2014).

As plantas são capazes de perceber e responder a estímulos bióticos e abióticos por meio de mudanças fisiológicas, bioquímicas e físicas de modo que produzam um fenótipo que seja correspondente ao seu ambiente (METLEN; ASCHEHOUG; CALLAWAY, 2009). Desse modo, a sacarose também pode atuar como uma molécula sinalizadora para os processos envolvidos no crescimento vegetal, como no estímulo para a divisão celular nos meristemas apicais, na sinalização de luz vermelha distante ou na síntese de ribossomos por meio da indução de genes que codificam proteínas ribossomais (MA et al., 2011b). Como molécula sinalizadora, a sacarose atua como reguladora da expressão de mRNA, de FTs e outros genes de sinalização hormonal, oxidativos e de defesa conforme é demonstrado na Fig.1.3 (YOON et al., 2021).

Figura 1.3 – Mecanismos de sinalização da sacarose em plantas superiores



Fonte: Adaptada de Yoon, Cho, Tun, Jeon e An (2021)

Os mecanismos moleculares pelos quais a sacarose funciona como um sinal necessitam de mais estudos, entretanto até o momento, sabe-se que essa molécula influencia no aumento da expressão de alguns FTs como o *AtWRKY20* e o *MYB75*, além de controlar a

expressão do gene *AtbZIP11* no nível pós-transcricional. A regulação do gene *AtbZIP11*, de *Arabidopsis*, ocorre quando o peptídeo codificado por uma *uORF* interage com moléculas de sacarose, esse efeito regulatório também acontece em bZIPs de outras plantas. Além disso, a sacarose aumenta a atividade da AGPase por modificação redox pós-tradução, controlando assim a expressão da sacarose sintase (SUS4) (WIESE et al., HUMMEL et al., 2009; TOGNETTI; PONTIS; MARTÍNEZ-NOËL, 2013; YOON et al., 2021). Estratégias para aumentar o teor de alguns carboidratos em culturas têm sido selecionadas por meio da edição gênica de plantas, visando o aumento ou alteração da conversão de açúcares (SAGOR et al., 2016; XING et al., 2020b; VELTKAMP, 2020).

2.6 Controle da expressão gênica por fatores de transcrição da família bZIP

A regulação da expressão dos genes através da transcrição é um processo vital para os organismos. Padrões precisos da atividade dos genes resultam em tipos celulares, órgãos e estruturas diferenciadas. Há dois componentes centrais que regulam a transcrição: os FTs e os elementos cis regulatórios, que são os motivos de sequência específica regulados pelos FTs. Cerca de 8% da capacidade de codificação das proteínas nos eucariotos é atribuída a algum fator de transcrição (LI et al., 2015; SONG et al., 2016; DRÖGE-LASER; WEISTE, 2018b).

As ligações dos FTs acontecem na forma de dímeros homólogos ou heterólogos e, assim, interferem na regulação do crescimento e do desenvolvimento da planta por meio de respostas a estímulos ambientais e pela transdução de sinal (BAI et al., 2023). Por isso, os FTs são reguladores chave nas redes de sinalização em resposta ao crescimento, ao desenvolvimento e aos estresses ambientais, que ocorre através da ligação a promotores específicos, podendo atuar sob a ativação ou repressão da expressão de um gene próximo ou distal (SONG et al., 2016).

Os bZIPs são FTs compostos por 60 a 80 resíduos de aminoácidos que possuem uma região básica de ligação ao DNA e uma região adjacente chamada de zíper-leucina para homo ou heterodimerização, que permitem respostas de expressão gênica variável (EHLERT et al., 2006a; DRÖGE-LASER; WEISTE, 2018b). Foi demonstrado por estudos genéticos e moleculares, que os fatores bZIP em plantas regulam diversos processos biológicos e modulam respostas a estresses abióticos e bióticos, sendo associados a diversas funções centrais na sinalização ambiental e de desenvolvimento (NOMAN et al., 2017). Pesquisas

utilizando abordagens *knockout/knockdown* ou de super expressão em espécies de plantas modelo, demonstraram que membros da família bZIP participam da diferenciação de muitos órgãos e tecidos, na embriogênese, na maturação de sementes, no florescimento e no desenvolvimento vascular (EHLERT et al., 2006a; MA et al., 2011b; DRÖGE-LASER; WEISTE, 2018b).

As proteínas bZIP pertencem a uma grande família de FTs vegetais, em *Arabidopsis* essa família gênica compreende 78 membros que estão organizados em 13 grupos (JAKOBY, 2002). A caracterização sistemática em todo o genoma da família bZIP em tomateiro identificou 69 genes *SibZIPs*, distribuídos desigualmente entre os cromossomos e ordenados em 9 clados. Além da distribuição, foram apresentados detalhes sobre a organização do domínio proteico, da estrutura gênica, da análise filogenética e evolutiva (LI et al., 2015). O agrupamento interespecíes entre arroz, tomate e *Arabidopsis* foi realizado e, ao ser analisado é possível observar que devido a existência de genes *bZIP* homólogos entre as plantas, há indícios filogenéticos de que a estrutura e a função desses genes são conservadas entre as espécies vegetais (JAKOBY, 2002; LI et al., 2015; NOMAN et al., 2017).

Existem certas especificidades quanto aos estímulos e respostas que são gerados por membros da família bZIP, como por exemplo condições que imitam situações de baixa energia. Este estresse induz a expressão gênica de bZIP1 e bZIP63, que são genes regulados por sacarose ou glutamina, respectivamente. Como a resposta depende de moléculas que são facilmente metabolizadas, estes dados indicam que o *status* da energia celular é detectado e retransmitido para o processo de transcrição, como acontece quando há a super expressão do bZIP11 em *Arabidopsis*, em que as respostas observadas são semelhantes às situações de plantas com necessidade de carbono. Outros bZIPs regulam regiões promotoras de genes que respondem a luz, indicando que estes possuem padrões de expressão dependentes da luminosidade (WIESE et al., 2005; DRÖGE-LASER; WEISTE, 2018b).

A família bZIP fornece exemplos de como a atividade de FTs podem ser controladas em múltiplos níveis, ou seja, na transcrição, na tradução – pelo mecanismo SIRT, na pós-tradução - pela fosforilação, na translocação nuclear ou na estabilidade proteica. É evidente que o conhecimento destes mecanismos regulatórios é crucial para a compreensão da função dos bZIPs, uma vez que muitas informações ainda não foram completamente esclarecidas (WIESE et al., 2005; DRÖGE-LASER; WEISTE, 2018b).

2.7 S1-bZIPs

O grupo S é o maior subgrupo bZIP em várias espécies e contém de três a quatro subgrupos ainda menores (JAKOBY, 2002; ALI et al., 2016). Em *Arabidopsis*, os *AtbZIPs* contêm 17 membros do grupo S que são distribuídos em três subgrupos com base na homologia: S1, S2 e S3. O subgrupo S1 (S1-bZIP) em *Arabidopsis* contém cinco membros: *AtbZIP1*, -2, -11, -44 e -53 (EHLERT et al., 2006b).

Assim como nos demais bZIPs, o subgrupo S1 é caracterizado por um domínio bZIP conservado composto por dois motivos funcionalmente distintos (uma região básica e um zíper de leucina) localizados em uma α -hélice contígua. Esse grupo possui uma estrutura única de genes que codificam proteínas pequenas (20 kD), que são altamente homólogas e não possuem íntrons. Estes FTs se ligam preferencialmente aos motivos A-box, C-box e G-box dos promotores alvo que contêm sequências de DNA com um núcleo ACGT (JAKOBY, 2002; DRÖGE-LASER; WEISTE, 2018b). Através de análises de bioinformática, entende-se que essa classe atua na regulação metabólica central de carbono e de aminoácidos gerando respostas ao esgotamento gradual ou à mudança da nutrição nitrogenada (EHLERT et al., 2006b; WANG et al., 2022).

Foi verificado que membros do grupo C bZIP formam preferencialmente heterodímeros com a classe S, formando a rede C/S1 bZIP que atua ajustando o metabolismo de carbono e nitrogênio, controlando o crescimento, o desenvolvimento e as respostas a estresses. A formação de heterodímeros C/S1 tem sido observada em várias espécies, entretanto a caracterização de suas propriedades específicas de heterodimerização precisam ser mais estudadas para desvendar o controle transcricional (PEVIANI et al., 2016; DRÖGE-LASER; WEISTE, 2018b; WANG et al., 2022). Em análises filogenéticas entre bZIPs dos grupos S1 e C de angiospermas, de gimnospermas, de musgos e de algas foram observadas que os grupos S1 e C evoluíram de um proto-S/C bZIP de espécies primitivas de algas, que formavam homodímeros que se separaram e deram origem aos pares de heterodímeros existentes atualmente, este acontecimento ocorreu antes da evolução de sementes nas plantas (PEVIANI et al., 2016).

A rede bZIP C/S1 controla a expressão dos genes associados ao metabolismo de acordo com o estado nutricional das plantas, estando envolvidos no reajuste metabólico durante a sinalização de baixa energia e deficiência de nutrientes, como pode acontecer em

períodos de escuro prolongado. Existe indícios de relação entre o ritmo circadiano e a rede *bZIP C/S1* no balanço de carbono e nitrogênio, que embora estejam frequentemente ligados à falta de nutrientes ou a limitação de energia, existem aspectos funcionais adicionais que devem ser levados em consideração (DRÖGE-LASER; WEISTE, 2018b).

Os *bZIPs C/S1* são regulados por outros genes como *SnRK1*, *HXK1* ou pelo mecanismo *SIRT*. A presença de regiões *uORF* (*upstream open-reading frames*) conservadas, na UTR (*untranslated region.*), presentes em todos os genes do grupo S, permitem também a regulação pós transcricional desses genes (MA et al., 2011b; PEVIANI et al., 2016).

2.8 Estrutura *uORFs* únicas e mecanismo de regulação dos S1-*bZIPs*

Nas regiões 5' não traduzidas ou UTRs dos mRNAs, há uma variedade de elementos de sequências regulatórias que afetam o transporte, a estabilidade e a eficiência da tradução do mRNA. Um dos elementos mais caracterizados é o quadro de leitura aberto – *uORF*, que é distinto da ORF principal que é a região codificadora da proteína no mesmo mRNA policistrônico (KURIHARA, 2020). Portanto, pode-se definir que as *uORFs* são elementos cis em genes eucarióticos que normalmente suprimem a tradução de ORFs primários, as pORFs, estando associados a importantes processos biológicos como o desenvolvimento, na biossíntese e na resistência (LIN et al., 2019; ZHANG; WANG; LU, 2019).

Em S1-*bZIP* há várias *uORFs* que codificam pequenos peptídeos. Dentre elas a região correspondente a segunda *uORF* é bastante conservada e codifica um peptídeo responsável pelo controle de sacarose, essa interação que regula a tradução do pORF e conseqüentemente reduz a expressão da proteína *bZIP*, é o mecanismo de regulação *SIRT* que contribui para a homeostase da sacarose nas células (YAMASHITA et al., 2017).

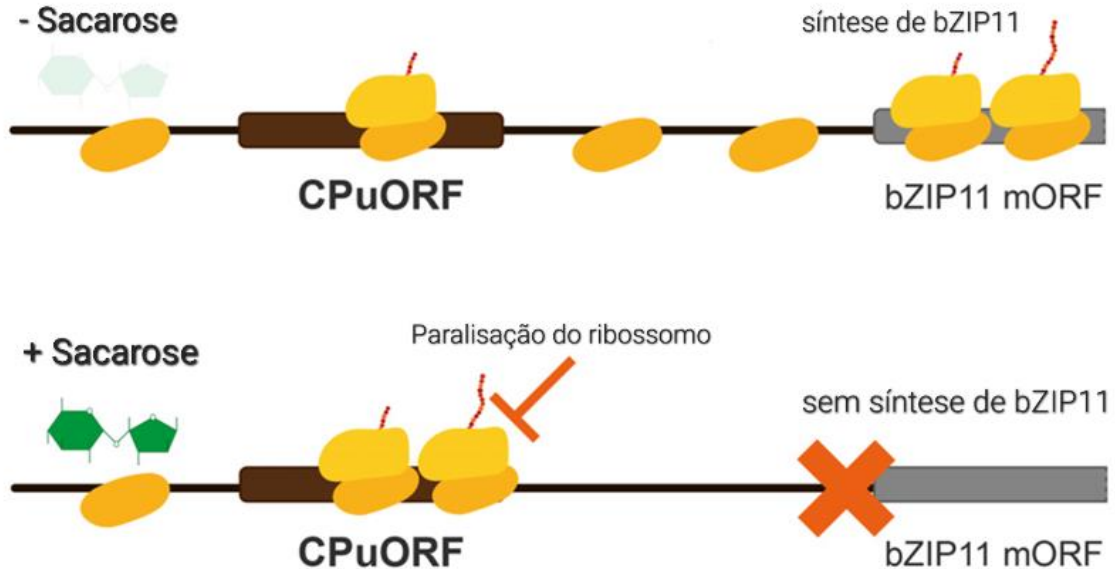
O processo de redução da expressão é controlado pela interação entre os peptídeos pequenos codificados pelo 5' *uORFs* e a sacarose, que interferem no deslocamento do ribossomo, estagnando o processo. Embora esta paralisação possa ocorrer de forma autônoma, muitas vezes é requerido a presença de um efetor, neste caso a molécula de sacarose. Dessa maneira a repressão da tradução induzida pela sacarose é responsável por reprimir a tradução ORF principal por este mesmo metabólito (YAMASHITA et al., 2017; KURIHARA, 2020).

A regulação da expressão gênica envolve diferentes camadas que incluem controles transcricionais e pós transcricionais (VAN DER HORST et al., 2020). Comparado com a

regulação transcricional, o controle na etapa de tradução permite respostas mais rápidas para ajustar a expressão de proteínas e reprogramar o metabolismo mediante sinais celulares ou estímulos ambientais (HUMMEL et al., 2009; MAZZONI-PUTMAN; STEPANOVA, 2018). Por isso, o sistema SIRT é um importante recurso em que a disponibilidade de nutrientes e a energia celular podem ser retransmitidas para a expressão gênica. Genes ortólogos aos S1-bZIPs, que são mediados pelo mecanismo SIRT, existem em todas as plantas com sementes e os aminoácidos como a serina, a leucina e a tirosina no peptídeo conservado da *uORF* são essenciais para SIRT aconteça (HUMMEL et al., 2009; PEVIANI et al., 2016).

A regulação da tradução dependente do peptídeo *uORF* em *Arabidopsis* foi bem caracterizada (FIGURA 1.4). O mRNA do gene *AtbZIP11* contém quatro regiões *uORFs* que se sobrepõem parcialmente e a deleção de partes dessa região resulta na perda do mecanismo SIRT. Uma dessas quatro *uORFs* é altamente conservada nas plantas, sendo essa uma região fundamental para a regulação da síntese de sacarose. A *uORF* reguladora contém códon inicial (AUG), onde o ribossomo inicia a tradução e porção correspondente a sequência de códons finais presentes no peptídeo dessa *uORF* interage com moléculas de sacarose que estão presentes no processo e que são responsáveis pela paralisação do ribossomo, conseqüentemente há a interrupção da tradução do gene *AtbZIP11*, este mesmo processo acontece para outros S1-bZIPs (MA et al., 2011b; YAMASHITA et al., 2017; DEVER; IVANOV; SACHS, 2020).

Figura 1.4 – Representação da repressão induzida por sacarose na tradução do mRNA de um bZIP11.



Legenda: Quando os níveis de sacarose estão baixos os ribossomos realizam a tradução até o códon de parada, permitindo que a ORF principal ou mORF seja em seguida traduzida. A tradução do mORF envolve a reinicialização da tradução e o acoplamento das subunidades do ribossomo para prosseguir. Com o aumento dos níveis de sacarose, devido a interações que acontecem na região final da molécula, ocorre a parada do ribossomo, bloqueando a tradução por impedir a varredura de outros ribossomos e, consequentemente inibindo a produção da proteína bZIP11.

Fonte: Van Der Horst et al. (2020)

Sabendo do mecanismo de regulação de alguns bZIPs através de regiões *uORFs*, algumas estratégias para aumentar o teor de açúcar foram desenvolvidas para superar SIRT. Em tabaco a super expressão do gene *tbz17*, um FT do tipo bZIP e ortólogo ao gene *AtbZIP11*, impossibilitou o processo de regulação SIRT. Foi observado também à amplificação da expressão dos genes asparagina sintetase (ASN), da prolina desidrogenase (ProDH), da sacarose fosfatase (SPP) e da sacarose fosfato sintase (SPS) que são genes regulados por *tbz17*. Os resultados apresentados demonstram o aumento do conteúdo de sacarose nas folhas, no entanto a planta transformada apresentou comprometimento de seu desenvolvimento em comparação a planta controle (THALOR et al., 2012).

Em tomateiro a expressão do gene *SlbZIP1* sem a região *uORF* usando o promotor E8, que é específico de frutos, resultou no aumento significativo da concentração de açúcares nas linhagens transgênicas. Além do aumento de 50% de teor de açúcar nos frutos, houve aumento da expressão dos genes SPP e SPS (SAGOR et al., 2016). Outros resultados

similares foram obtidos com a estratégia de edição gênica projetada para *uORF* do gene *FvebZIPs1.1* em morango (*Fragaria vesca*). Neste estudo foi utilizado o sistema de editor de base A3A-PBE, os organismos mutantes obtidos apresentaram maior teor de sacarose e não foram observados déficits de desenvolvimento ou rendimento das plantas (XING et al., 2020b). Em outra estratégia de edição gênica, acredita-se que a remoção de *uORF* no gene *bZIP2* em conjunto com a conversão da invertase vacuolar em invertase da parede celular e a exclusão do domínio N-terminal do gene *SPS*, resultaria em aumento do nível de sacarose nos frutos (MUFARRIJ et al., 2021). Estes estudos demonstram que modificações nas regiões *uORF* podem ser realizadas para se obter o aumento de açúcares.

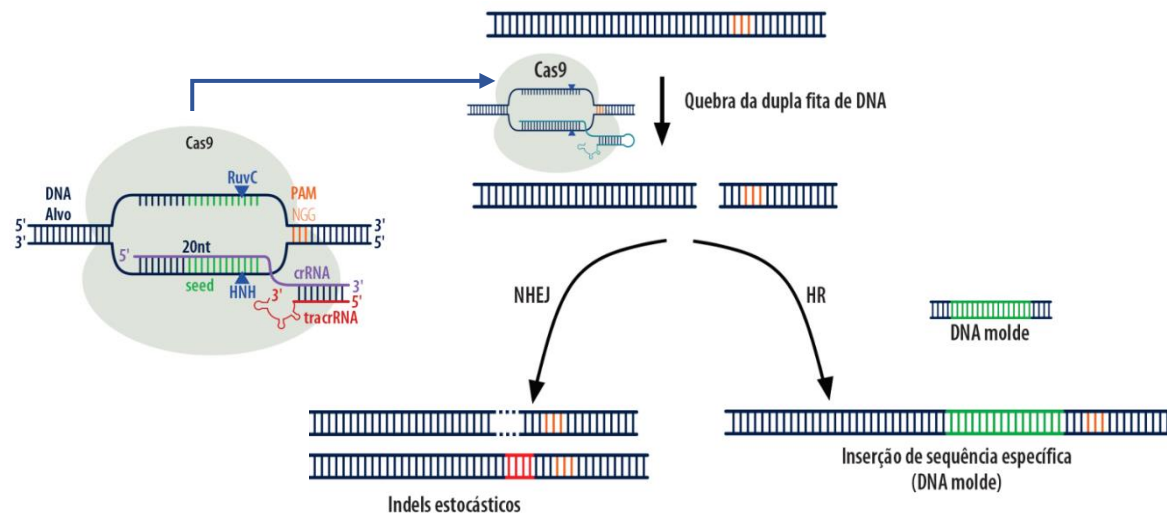
2.9 Edição gênica pelo sistema CRISPR/Cas9

Tradicionalmente, a seleção aplicada a espécies vegetais tem permitido a geração de novas variedades de culturas. As novas tecnologias permitem a introdução precisa de alelos desejáveis em muitas variedades de elite diferentes e localmente adaptadas, oferecendo uma oportunidade de gerar rapidamente variedades melhoradas com custos reduzidos (NEKRASOV et al., 2017). Existem 13 espécies reconhecidas de tomates silvestres que apresentam uma grande variedade de fenótipos e que podem ser utilizadas para o retrocruzamento com os tomateiros cultivados comercialmente (PERALTA; KNAPP; SPOONER, 2005). Esses tomates selvagens são importantes para reprodução, como fontes de características desejáveis e para estudos evolutivos, entretanto, os métodos tradicionais de melhoramento podem ser demorados (KIMURA; SINHA, 2008b).

Dessa forma, a utilização da engenharia genética tem o potencial de aumentar a produtividade através do aumento da eficiência das vias metabólicas ou fotossintéticas, aumento da resistência a fatores ambientais, como temperatura, umidade e a diferentes tipos de solo (GERSZBERG et al., 2015). A edição do genoma é uma grande inovação no melhoramento de plantas que facilita modificações eficientes, precisas e direcionadas em *loci* genômicos que permitirão a obtenção de plantas livres de transgenes. A edição do genoma permite uma mutagênese precisa de um gene de interesse, dessa forma as plantas obtidas são idênticas ou semelhantes às geradas por técnicas convencionais de melhoramento (SALAVA et al., 2021; XIA et al., 2021a).

Uma importante ferramenta aplicada a edição de genes é a técnica CRISPR/Cas9 que foi inicialmente descoberta como uma estratégia de defesa do sistema imunológico em bactérias e *Archaea* (BHAYA; DAVISON; BARRANGOU, 2011). O funcionamento da estratégia acontece da seguinte maneira: a proteína Cas9, uma endonuclease de DNA de fita dupla, é guiada por um RNA que será direcionado para uma sequência específica ou alvo (BELHAJ et al., 2013). O primeiro passo para o reconhecimento do gene alvo depende da presença de um motivo espaçador adjacente (PAM), tipicamente NGG para a nuclease do tipo Cas9 (JINEK et al., 2012). Sendo assim, a Cas9, desenrola a sequência de DNA alvo usando o emparelhamento de base sgRNA-DNA (*single-guide* RNA) e cliva ambas as fitas de DNA próxima ao local da região PAM (JINEK et al., 2012). Uma quebra da fita dupla (DSB) no DNA é reparada ou por junção final não-homóloga (NHEJ) ou por reparo homólogo direcionado (HDR). Durante o reparo NHEJ, que ocorre com maior frequência, as extremidades soltas são novamente unidas e, este mecanismo, é propenso a erros que resultam em *indels*, que são mutações com pequenas inserções ou deleções (DERIANO; ROTH, 2013) (FIGURA 1.5). Assim que foi observado que esse mecanismo possuía potencial para edição de genes, este foi rapidamente adaptado e aprimorado para as células eucarióticas (JINEK et al., 2012).

Figura 1.5: Sistema de funcionamento da técnica CRISPR/Cas9



Legenda: Representação de reparo do DNA por NHEJ ou HR. Em vermelho é indicado as deleções e as inserções, respectivamente. Em verde a sequência de interesse a ser integrada no DNA-alvo. Laranja representa a região de reconhecimento PAM. Os sítios catalíticos da proteína Cas9 estão representados por RuvC e HNH.

Fonte: Adaptado de Embrapa, 2020.

A capacidade de usar um pequeno RNA específico e não codificável, o *sgRNA* tornou a técnica de CRISPR/Cas9 ideal para a edição de genoma de maneira precisa e personalizável (BELHAJ et al., 2013). Dessa maneira, a tecnologia torna possível fazer modificações em qualquer variedade em período reduzido, além disso, podem ser criados genótipos e fenótipos completamente novos. Com boas ferramentas de predição, a técnica pode ser usada para fazer uma mutação direcionada na região de interesse e criar o fenótipo desejado (SI et al., 2020).

CRISPR/Cas9 foi validada com sucesso em tomateiro. Em um estudo foi utilizado como alvo um gene que ao ter sua função interrompida, resultaria em um fenótipo facilmente percebido e reconhecível nas primeiras fases de regeneração na cultura de tecidos. Uma construção CRISPR/Cas9 foi projetada para atingir o gene *SIAGO7*, gene recessivo, necessário para reconhecimento e resposta de genes do fator de resposta de auxina (ARF). As alterações observadas em plantas transformadas e editadas resultaram em um fenótipo característico, o surgimento de folhas estreitas com formato de agulhas (BROOKS et al., 2014). A utilização da técnica de CRISPR também foi utilizada como estratégia para edição gênica para eliminar um domínio auto inibitório que limitava a biossíntese de GABA, um

importante hipotensivo, os autores relatam que foi possível aumentar consideravelmente a concentração dos níveis de GABA nos frutos (NONAKA et al., 2017). A tecnologia de edição gênica tem potencial de aplicação para diversas finalidades, como por exemplo para a reavaliação de genes do tomateiro que estão envolvidos no processo de amadurecimento dos frutos, essa pesquisa foi realizada com o gene *RIN*, que é responsável por codificar um FT da família MADS-box. Os genes MADS-box são reguladores do amadurecimento dos frutos, por isso foram escolhidos como alvo para estratégia de edição. Os tomateiros com mutação heteróloga conseguiram expressar o gene do tipo selvagem remanescente e apresentaram a coloração vermelha característica de maturação completa, confirmando o importante papel do gene *RIN* no processo e várias mutações que foram geradas nos locais alvo foram herdadas na progênie T1, confirmando a aplicabilidade deste sistema de mutagênese em tomateiro (ITO et al., 2015). Desde então, outros trabalhos envolvendo edição gênica em tomateiro com diferentes finalidades têm sido publicados com intuito de validar características, agregar conhecimentos sobre os diferentes genótipos e para alcançar os fenótipos desejados em tempo reduzido (VELTKAMP, 2020; LI et al., 2022).

REFERÊNCIAS

- ALI, Z. *et al.* **Functions of plant's bZIP transcription factors.** Pakistan Journal of Agricultural Sciences. University of Agriculture, 1 jun. 2016. Doi: 10.21162/PAKJAS/16.2043.
- BAI, C. *et al.* **Sugar accumulation and fruit quality of tomatoes under water deficit irrigation.** Postharvest Biology and Technology, v. 195, 1 Jan. 2023. Doi: 10.1016/j.postharvbio.2022.112112.
- BELHAJ, K. *et al.* **Plant genome editing made easy: Targeted mutagenesis in model and crop plants using the CRISPR/Cas system.** Plant Methods, v. 9, n. 1, p. 1–10, 11 out. 2013. Doi: 10.1186/1746-4811-9-39/TABLES/2. Acesso em: 1 Jan. 2023.
- BERGOUGNOUX, V. **The history of tomato: From domestication to biopharming.** Biotechnology Advances, Jan. 2014. Doi: 10.1016/j.biotechadv.2013.11.003.
- BERTIN, N.; GÉNARD, M. **Tomato quality as influenced by preharvest factors.** Scientia Horticulturae, v. 233, p. 264–276, 15 mar. 2018. doi: 10.1016/J.SCIENTA.2018.01.056. Acesso em: 5 dez. 2022.
- BHAYA, D.; DAVISON, M.; BARRANGOU, R. **CRISPR-cas systems in bacteria and archaea: Versatile small RNAs for adaptive defense and regulation.** Annual Review of Genetics, v. 45, p. 273–297, 2011. Doi: 10.1146/annurev-genet-110410-132430.
- BOSWELL, V. R. **Improvement and genetics of tomatoes, peppers, and eggplant.** Principal Horticulturist, Division of Fruit and Vegetable Crops and Diseases, Bureau of Plant Industry. [s.l.: s.n.].
- BROOKS, C. *et al.* **Scientific Correspondence Efficient Gene Editing in Tomato in the First Generation Using the Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-Associated9 System 1.** [s.d.] Doi: 10.1104/pp.114.247577. Acesso em: 1 Jan. 2023.
- CAMPOS, M. L. *et al.* **The micro-tom model system as a tool to discover novel hormonal functions and interactions.** Plant Signaling and Behavior, v. 5, n. 3, p. 267–270, 2010. Doi: 10.4161/psb.5.3.10622.
- CARVALHO, R. F. *et al.* **Convergence of developmental mutants into a single tomato model system: “Micro-Tom” as an effective toolkit for plant development research.** [s.l.: s.n.].
- CONAB. **Tomate: Análise dos indicadores da produção e comercialização no mercado mundial brasileiro e catarinense.** [s.l.: s.n.].
- DAVIES, J. N.; HOBSON, G. E.; MCGLASSON, W.B. **The constituents of tomato fruit — the influence of environment, nutrition, and genotype,** C R C Critical Reviews in Food Science and Nutrition, 15:3, 205-280, (1981). Doi: 10.1080/10408398109527317
- DE LAS MERCEDES CAPOBIANCO-URIARTE, M. *et al.* **The European tomato market. An approach by export competitiveness maps.** PLoS ONE, v. 16, n. 5 May, 1 maio 2021. Doi: 10.1371/journal.pone.0250867.

- DERIANO, L.; ROTH, D. B. **Modernizing the nonhomologous end-joining repertoire: Alternative and classical NHEJ share the stage.** Annual Review of Genetics Annual Reviews Inc., 2013. Doi: 10.1146/annurev-genet-110711-155540.
- DEVER, T. E.; IVANOV, I. P.; SACHS, M. S. **Conserved Upstream Open Reading Frame Nascent Peptides That Control Translation.** 2020. Doi: 10.1146/annurev-genet-112618.
- DRÖGE-LASER, W.; WEISTE, C. **The C/S1 bZIP Network: A Regulatory Hub Orchestrating Plant Energy Homeostasis.** Trends in Plant Science, v. 23, n. 5, p. 422–433, 2018. Doi: 10.1016/j.tplants.2018.02.003.
- EHLERT, A *et al.* **Two-hybrid protein-protein interaction analysis in Arabidopsis protoplasts: Establishment of a heterodimerization map of group C and group S bZIP transcription factors.** Plant Journal, v. 46, n. 5, p. 890–900, jun. 2006a. Doi: 10.1111/j.1365-313X.2006.02731.x.
- FAO (org.). **Food and Agriculture Organization of the United Nations.** 2022. Disponível em: <http://www.fao.org/faostat/en/>. Acesso em: 1 dez. 2022
- FICHTNER, F. *et al* **Sugar and Nitrate Sensing: A Multi-Billion-Year Story.** Trends in Plant Science Elsevier Ltd, 1 abr. 2021. Doi: 10.1016/j.tplants.2020.11.006.
- GAO, L. *et al.* **The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor.** Nature Genetics, v. 51, n. 6, p. 1044–1051, 1 jun. 2019. Doi: 10.1038/s41588-019-0410-2.
- GERSZBERG, A. *et al.* **Tomato (*Solanum lycopersicum* L.) in the service of biotechnology.** Plant Cell, Tissue and Organ Culture Kluwer Academic Publishers, 1 mar. 2015. Doi: 10.1007/s11240-014-0664-4.
- HUMMEL, M. *et al.* **Sucrose-mediated translational control.** Annals of Botany, jul. 2009. Doi: 10.1093/aob/mcp086.
- IBGE (org.). **Instituto Brasileiro de Geografia e Estatística.** Produção Agropecuária de Tomate. 2022. Disponível em: <https://www.ibge.gov.br/explica/producao-agropecuaria/tomate/br>. Acesso em: 4 dez. 2022.
- ITO, Y. *et al.* **CRISPR/Cas9-mediated mutagenesis of the RIN locus that regulates tomato fruit ripening.** Biochemical and Biophysical Research Communications, v. 467, n. 1, p. 76–82, 6 nov. 2015. Doi: 10.1016/J.BBRC.2015.09.117. Acesso em: 1 Jan. 2023.
- JAKOBY, M. **bZIP transcription factors in Arabidopsis.** 2002. Acesso em: 1 Jan. 2023.
- JAKOBY, M. *et al.* **bZIP transcription factors in Arabidopsis.** Trends in Plant Science, v. 7, n. 3, p. 106–111, 2002. Doi: 10.1016/S1360-1385(01)02223-3.
- JENKINS, J. A. **The Origin of the Cultivated Tomato.** [s.l.: s.n.].
- JINEK, M. *et al.* **A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity.** [s.d.]. Acesso em: 1 Jan. 2023.
- KALLOO, G. **Tomato.** Em: Genetic Improvement of Vegetable Crops. [s.l.] Elsevier, 1993. p. 645–666. 1993.
- KIMURA, S.; SINHA, N. **Tomato (*Solanum lycopersicum*): A model fruit-bearing crop.** Cold Spring Harbor Protocols, v. 3, n. 11, nov. 2008a. Doi: 10.1101/pdb.emo105.

- KIMURA, S.; SINHA, N. **Crossing tomato plants**. Cold Spring Harbor Protocols, v. 3, n. 11, nov. 2008b. Doi: 10.1101/pdb.prot5082.
- KURIHARA, Y. ***UORF* shuffling fine-tunes gene expression at a deep level of the process**. Plants MDPI AG, 2020. Doi: 10.3390/plants9050608.
- LI, D. *et al.* **Genome-wide systematic characterization of the bZIP transcriptional factor family in tomato (*Solanum lycopersicum* L.)**. BMC Genomics, v. 16, n. 1, 12 out. 2015. Doi: 10.1186/s12864-015-1990-6.
- LI, Q. *et al.* **Increasing Fruit Weight by Editing a Cis-Regulatory Element in Tomato KLUH Promoter Using CRISPR/Cas9**. Frontiers in Plant Science, v. 13, 11 abr. 2022. Doi: 10.3389/fpls.2022.879642.
- LI, Y. *et al.* **MicroTom Metabolic Network: Rewiring Tomato Metabolic Regulatory Network throughout the Growth Cycle**. Molecular Plant, v. 13, n. 8, p. 1203–1218, 2020. Doi: 10.1016/j.molp.2020.06.005.
- LI, Y. *et al.* **Can the world's favorite fruit, tomato, provide an effective biosynthetic chassis for high-value metabolites?** Plant Cell Reports, v. 37, n. 10, p. 1443–1450, 2018. Doi: 10.1007/s00299-018-2283-8.
- LIN, Y *et al.* **Impacts of *uORF* codon identity and position on translation regulation**. Nucleic Acids Research, v. 47, n. 17, p. 9358–9367, 26 set. 2019. Doi: 10.1093/nar/gkz681.
- LIU, D. *et al.* **Domestication and breeding changed tomato fruit transcriptome**. Journal of Integrative Agriculture, v. 19, n. 1, p. 120–132, 1 Jan. 2020. Doi: 10.1016/S2095-3119(19)62824-8.
- LIU, W. *et al.* ***Solanum lycopersicum*, a Model Plant for the Studies in Developmental Biology**. Stress Biology and Food Science Foods MDPI, 1 ago. 2022. Doi: 10.3390/foods11162402.
- LUNN, J. E. **Evolution of sucrose synthesis**. Plant Physiology, v. 128, n. 4, p. 1490–1500, 2002. Doi: 10.1104/pp.010898.
- MA, J. *et al.* **The sucrose-regulated Arabidopsis transcription factor bZIP11 reprograms metabolism and regulates trehalose metabolism**. New Phytologist, v. 191, n. 3, p. 733–745, 2011a. Doi: 10.1111/j.1469-8137.2011.03735.x.
- MATA-NICOLÁS, E. *et al.* **Exploiting the diversity of tomato: the development of a phenotypically and genetically detailed germplasm collection**. Horticulture Research, v. 7, n. 1, 1 Dez. 2020. Doi: 10.1038/s41438-020-0291-7.
- MAZZONI-PUTMAN, S. M.; STEPANOVA, A. N. **A plant biologist's toolbox to study translation**. Frontiers in Plant Science, v. 9, 2 jul. 2018. Doi: 10.3389/fpls.2018.00873.
- METLEN, K. L.; ASCHEHOUG, E. T.; CALLAWAY, R. M. **Plant behavioural ecology: Dynamic plasticity in secondary metabolites**. Plant, Cell and Environment, v. 32, n. 6, p. 641–653, jun. 2009. Doi: 10.1111/j.1365-3040.2008.01910.x.
- MUFARRIJ, M., *et al.* **Modifying High Sucrose Tomatoes by Genome Editing A-Review**. 2021. doi: 10.2991/978-94-6463-062-6_31.

- NEKRASOV, V. *et al.* **Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion.** [s.d.] Doi: 10.1038/s41598-017-00578-x. Acesso em: 1 Jan. 2023.
- NOMAN, Z. *et al.* **Basic leucine zipper domain transcription factors: the vanguards in plant immunity.** *Biotechnology Letters*, v. 39, p. 1779–1791, 2017. Doi: 10.1007/s10529-017-2431-1. Acesso em: 1 Jan. 2023.
- NONAKA, S. *et al.* **Efficient increase of Γ -aminobutyric acid (GABA) content in tomato fruits by targeted mutagenesis.** *Scientific Reports*, v. 7, n. 1, p. 1–14, 2017. Doi: 10.1038/s41598-017-06400-y.
- OSORIO, S.; RUAN, Y. L.; FERNIE, A. R. **An update on source-to-sink carbon partitioning in tomato.** *Frontiers in Plant Science*. Frontiers Media S.A., 6 out. 2014. Doi: 10.3389/fpls.2014.00516.
- PERALTA, I. E.; KNAPP, S.; SPOONER, D. M. **New Species of Wild Tomatoes (*Solanum* Section *Lycopersicon*: *Solanaceae*) from Northern Peru.** *Systematic Botany*. [s.l.: s.n.].
- PEVIANI, A. *et al.* **The phylogeny of C/S1 bZIP transcription factors reveals a shared algal ancestry and the pre-angiosperm translational regulation of S1 transcripts.** *Nature Publishing Group*, 2016. Doi: 10.1038/srep30444. Acesso em: 1 Jan. 2023.
- QUINET, M. *et al.* **Tomato Fruit Development and Metabolism.** *Frontiers in Plant Science*, v. 10, n. November, p. 1–23, 2019. Doi: 10.3389/fpls.2019.01554. Acesso em: 5 Dez. 2022
- ROOK, F. *et al.* **Sucrose-specific signalling represses translation of the Arabidopsis ATB2 bZIP transcription factor gene.** *Plant Journal*, v. 15, n. 2, p. 253–263, 1998. Doi: 10.1046/j.1365-313X.1998.00205.x.
- SAGOR, G. H. M. *et al.* **A novel strategy to produce sweeter tomato fruits with high sugar contents by fruit-specific expression of a single bZIP transcription factor gene.** *Plant Biotechnology Journal*, v. 14, n. 4, p. 1116–1126, 1 abr. 2016. Doi: 10.1111/pbi.12480.
- SALAVA, H. *et al.* **Application of genome editing in tomato breeding: Mechanisms, advances, and prospects.** *International Journal of Molecular Sciences*. MDPI AG, 2 Jan. 2021. Doi: 10.3390/ijms22020682.
- SI, X. *et al.* **Manipulating gene translation in plants by CRISPR-Cas9-mediated genome editing of upstream open reading frames.** *Nature Protocols*, [s.d.] Doi: 10.1038/s41596-019-0238-3. Acesso em: 1 Jan. 2023.
- SNOWDEN, C. J. *et al.* **A tonoplast Glu/Asp/GABA exchanger that affects tomato fruit amino acid composition.** *The Plant Journal*, v. 81, n. 5, p. 651–660, 1 mar. 2015. doi: 10.1111/TPJ.12766. Acesso em: 5 dez. 2022.
- SONG, L. *et al.* **A transcription factor hierarchy defines an environmental stress response network.** *Science*, v. 354, n. 6312, 4 nov. 2016. Doi: 10.1126/science.aag1550.
- STEIN, O.; GRANOT, D. **An overview of sucrose synthases in plants.** *Frontiers in Plant Science* Frontiers Media S.A., 1 fev. 2019. Doi: 10.3389/fpls.2019.00095.
- SUNG, S.J. S.; XU, D.P.; BLACK, C. C. **Identification of Actively Filling Sucrose Sinks.** *Plant Physiol*, v. 89, p. 11–18, 1989. Acesso em: 1 Jan. 2023.

TANAKA, A.; FUJITA, K.; KIKUCHI, K. **Soil Science and Plant Nutrition Nutriophysiological studies on the tomato plant III.** Photosynthetic rate of individual leaves in relation to the dry matter production of plants. 2012. Doi: 10.1080/00380768.1974.10433240. Acesso em: 1 Jan. 2023.

THALOR, S. K. *et al.* **Deregulation of sucrose-controlled translation of a bZIP-type transcription factor results in sucrose accumulation in leaves.** PLoS ONE, v. 7, n. 3, 22 mar. 2012. Doi: 10.1371/journal.pone.0033111.

TIEMAN, D. *et al.* **The chemical interactions underlying tomato flavor preferences.** Current Biology, v. 22, n. 11, p. 1035–1039, 2012. Doi: 10.1016/j.cub.2012.04.016.

TIEMAN, D. *et al.* **A chemical genetic roadmap to improved tomato flavor.** Science (New York, N.Y.), v. 355, n. 6323, p. 391–394, 2017a. Doi: 10.1126/science.aal1556.

TOGNETTI, J. A.; PONTIS, H. G.; MARTÍNEZ-NOËL, G. M. A. **Sucrose signaling in plants: A world yet to be explored.** Plant Signaling and Behavior, mar. 2013. Doi: 10.4161/psb.23316.

TOMATO GENOME CONSORTIUM. The tomato genome sequence provides insights into fleshy fruit evolution. Nature, 2012. Doi: 10.1038/nature11119. Acesso em: 5 Dez. 2022.

VAN DER HORST, S. *et al.* **Update on Translational Control Metabolite Control of Translation by Conserved Peptide *uORFs*: The Ribosome as a Metabolite Multisensor.** [s.d.]. Doi: 10.1104/pp.19.00940. Acesso em: 1 Jan. 2023.

VELTKAMP, V. **A Sweeter Tomato cracking the Cis-regulatory code of gene regulation.** 2020. Acesso em: 1 Jan. 2023.

WANG, D.; SEYMOUR, G. B. **Tomato Flavor: Lost and Found?** Molecular Plant, v. 10, n. 6, p. 782–784, 2017a. Doi: 10.1016/j.molp.2017.04.010.

WANG, H. *et al.* **S1-bZIP Transcription Factors Play Important Roles in the Regulation of Fruit Quality and Stress Response.** Frontiers in Plant Science Frontiers Media S.A., 14 Jan. 2022. Doi: 10.3389/fpls.2021.802802.

WIESE, A. *et al.* **Sucrose-induced translational repression of plant bZIP-type transcription factors.** [s.l.: s.n.].

XIA, X. *et al.* **Advances in application of genome editing in tomato and recent development of genome editing technology.** Theoretical and Applied Genetics, v. 134, n. 9, p. 2727–2747, 1 set. 2021b. Doi: 10.1007/S00122-021-03874-3. Acesso em: 22 set. 2022.

XING, S. *et al.* **Fine-tuning sugar content in strawberry.** Genome Biology, v. 21, n. 1, p. 1–14, 2020. Doi: 10.1186/s13059-020-02146-5.

YAMASHITA, Y. *et al.* **Sucrose sensing through nascent peptide-mediated ribosome stalling at the stop codon of Arabidopsis bZIP11 *uORF2*.** FEBS Letters, v. 591, n. 9, p. 1266–1277, 1 maio 2017. Doi: 10.1002/1873-3468.12634.

YOON, J. *et al.* **Sucrose signaling in higher plants.** Plant Science Elsevier Ireland Ltd, 1 Jan. 2021. Doi: 10.1016/j.plantsci.2020.110703.

YU, C. P.; LIN, J. J.; LI, W. H. **Positional distribution of transcription factor binding sites in *Arabidopsis thaliana***. Scientific Reports, v. 6, n. January, p. 1–7, 2016. Doi: 10.1038/srep25164.

ZHANG, H.; WANG, Y.; LU, J. **Function and Evolution of Upstream ORFs in Eukaryotes**. Trends in Biochemical Sciences. Elsevier Ltd, 1 set. 2019. Doi: 10.1016/j.tibs.2019.03.002.

ZHANG, Y. *et al.* **Applications and potential of genome editing in crop improvement**. Biological Sciences Genetics. BioMed Central Ltd., 30 Nov. 2018. Doi: 10.1186/s13059-018-1586-y.

**CAPÍTULO 2 – DELEÇÃO DE UMA *uORF* CONSERVADA DO GENE *bZIP* EM
TOMATEIRO VIA CRISPR/Cas9**

RESUMO

A sacarose faz parte de um grupo de moléculas que atuam na sinalização e regulação da expressão gênica de alguns fatores de transcrição. Os S-bZIPs são fatores de transcrição (FTs) associados ao metabolismo de carbono na planta, estes FTs podem sofrer regulação de expressão na etapa pós transcricional, devido a variações nos níveis de sacarose. Sabendo que modificações em regiões específicas do gene *SibZIP1* podem alterar a regulação da sua expressão gênica e influenciar na concentração de sacarose e de outros açúcares, o objetivo deste trabalho foi causar uma mutação direcionada neste gene em plantas de tomateiro cv. Micro-tom. Para a edição gênica a tecnologia CRISPR/Cas9 foi utilizada com a estratégia de mutação, contendo em sua construção gênica a presença de dois *sgRNA* complementares a região *uORF*, que está envolvida na regulação do gene *SibZIP1*. Para a montagem do vetor da construção gênica, foi utilizado a tecnologia *Golden Gate* e a transformação das plantas foi realizada via *Agrobacterium tumefaciens*. Os eventos obtidos foram selecionados através da resistência ao agente seletivo, canamicina, sendo confirmados por PCR. Para a caracterização molecular da edição gênica, os fragmentos foram amplificados, isolados, purificados e posteriormente clonados. Quatro clones por planta foram submetidos as análises de sequenciamento. Como resultado, obteve-se uma deleção de 42 pb na *uORF*, sendo compatível com a região alvo e com a proposta apresentada. Este resultado demonstra que a estratégia de edição gênica proposta neste estudo é eficiente para ser utilizada.

Palavras-chave: CRISPR/cas9. *SibZIP1*. Micro-tom.

1 INTRODUÇÃO

Os fatores de transcrição (FTs) desempenham atividades cruciais em quase todos os processos biológicos. A família bZIP é ampla e diversa, sendo dividida em 13 subfamílias. De modo geral estes FTs estão envolvidos na regulação de processos fisiológicos e de desenvolvimento nas plantas (JAKOBY, 2002; DRÖGE-LASER; WEISTE, 2018b). Dentro da subfamília S-bZIP, o grupo S1, compreende genes que não possuem íntrons e contêm *uORFs* que codificam pequenas proteínas, envolvidas com a sinalização de baixa energia. As regiões *uORFs* presentes em membros da classe S1, codificam um peptídeo que interfere no deslocamento do ribossomo durante a tradução, causando a inibição da tradução da ORF principal, este mecanismo é dependente da sacarose e é conhecido pela denominação de repressão da tradução induzida por sacarose ou SIRT. É importante ressaltar que SIRT é considerado um recurso que caracteriza esses FTs, uma vez que este mecanismo não é observado em outros grupos bZIPs (PEVIANI et al., 2016). Causar alterações neste sistema SIRT podem modificar e aumentar os teores de açúcares nas plantas (JAKOBY et al., 2002; YAMASHITA et al., 2017; DRÖGE-LASER; WEISTE, 2018a).

Devido aos avanços proporcionados pela edição gênica é possível direcionar e alterar alguns padrões de expressão gênica com a finalidade de se obter o fenótipo desejado. Dessa forma, os FTs repressivos ou ativadores, são genes alvos interessantes para a manipulação artificial da expressão gênica, sendo potencialmente aplicáveis para a otimização de características econômicas de plantas cultivadas. (LI et al., 2015; SI et al., 2020).

Atualmente, a tecnologia CRISPR/Cas9 oferece algumas vantagens sobre outros sistemas e técnicas de edição de genoma, devido ao alto potencial de aplicação para múltiplos alvos, obtenção de plantas editadas em tempo reduzido e pelo custo relativamente baixo (CORREA et al., 2020).

O tomate é um importante produto agrícola muito cultivado na horticultura, devido a sua importância econômica, a espécie *S. lycopersicum* tem sido alvo de estudos e utilizada como planta modelo para pesquisas em diversas áreas (GEBHARDT, 2016; LI et al., 2020). Os tomates são amplamente consumidos e estão incluídos na dieta de forma processada ou *in natura*, por isso a utilização de ferramentas que visem o aumento ou alteração da conversão de açúcares, tem sido alvo de estudos em biotecnologia vegetal para essa espécie (SAGOR et al., 2016; QUINET et al., 2019).

Portanto, o objetivo apresentado neste capítulo do trabalho foi realizar uma mutação direcionada à região *uORF*, regulada por sacarose no gene *bZIP1* de tomateiro. Como estratégia foi utilizado a construção CRISPR/Cas9 com a presença de dois *sgRNA*, de modo que a clivagem em dois pontos simultâneos resultasse na deleção de um fragmento.

2 MATERIAIS E MÉTODOS

Os experimentos foram conduzidos no Laboratório Central de Biologia Molecular (LCBM), da Universidade Federal de Lavras.

2.1 Material Vegetal

Plantas de *S. lycopersicum* cv Micro-tom, foram cultivadas em câmara de crescimento sob condições controladas com temperatura de $26^{\circ}\text{C} \pm 1^{\circ}\text{C}$ e fotoperíodo de 16 h. Os frutos foram coletados após estarem completamente maduros, para a remoção da polpa e extração das sementes. Para facilitar a retirada da mucilagem que recobre as sementes, estas foram colocadas em suspensão em uma solução de água e fermento biológico comercial, *Saccharomyces cerevisiae*, por um período de 2 a 4h. Posteriormente, essas sementes foram lavadas em água corrente, selecionadas e mantidas em temperatura ambiente até estarem completamente secas para serem armazenadas até a sua utilização (PINO, et al 2010).

2.2 Construção gênica

O vetor utilizado foi previamente desenvolvido pelo Dr. Renan Terassi Pinto em parceria com o professor Dr. Vagner Augusto Benedito – *West Virginia University*. A incorporação da construção gênica direta foi realizada no vetor T-DNA pDIRECT_22C, pelo método *Golden Gate* descrito por Cermák et al. (2017). Cada *sgRNA* foi separado por um sítio de ligação a *Csy4*, o promotor 35S foi usado para conduzir a expressão de *Csy4* conectado a uma proteína Cas9 (*AtCas9*). Outro promotor 35S foi incorporado para conduzir a expressão do gene de resistência ao antibiótico canamicina (FIGURA 2.1). Como estratégia para a aplicação do sistema CRISPR/Cas9, os *sgRNA* foram delineados em complementariedade a

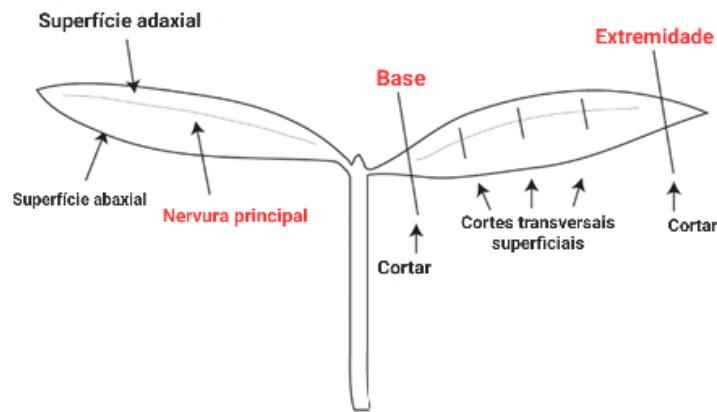
48 h a 28°C. Uma única colônia foi transferida para 3 ml de meio LB líquido suplementado com os mesmos antibióticos, onde foram mantidas em cultivo a 28°C por 48 h a 120 rpm. Dessa cultura, diferentes concentrações entre 100 e 500 µL foram adicionadas a 50 ml de meio LB líquido fresco. A cultura de *A. tumefaciens* obtida, foi incubada por 16 h a 120 rpm, sendo posteriormente centrifugada a 2000 g por 15 min. Cerca de 10 min antes da infecção, o sobrenadante foi descartado e o *pellet* foi ressuscitado em MS líquido (Murashige-Skoog) suplementado com sacarose 30 g L⁻¹, vitaminas B5 (Gamborg) e 100 µM de acetoseringona estéril (AS). A OD600 utilizada no processo de transformação foi de 0,2 a 0,4 nm.

2.4 Transformação e regeneração

O protocolo de transformação, assim como o preparo do material vegetal, seguiu as recomendações propostas por Pino et al. (2010) com algumas adaptações. As sementes de tomate foram previamente desinfestadas em solução de hipoclorito de sódio 30% por 15 min, enxaguadas em água destilada autoclavada e colocadas para germinação em meio sólido com metade das concentrações de MS, suplementado com 15 g L⁻¹ sacarose e vitaminas B5 (Gamborg) e concentração de ágar a 6 g L⁻¹. Aproximadamente 40 sementes foram inoculadas em frascos contendo 40 ml de meio de cultura, onde permaneceram por 4 dias no escuro a 27°C seguido por 4 dias sob fotoperíodo de 16 h mantendo-se a temperatura.

Após 8 dias os cotilédones foram excisados e utilizados como explantes. As pontas distal e proximal foram removidas e os cotilédones foram divididos transversalmente em duas partes (FIGURA 2.2). Os explantes foram colocados com a superfície abaxial para baixo imediatamente após o isolamento. Durante o processo de corte dos explantes, uma placa de petri estéril contendo sais de permanganato de potássio foi mantida dentro da capela de fluxo laminar, com intuito de amenizar o acúmulo de etileno, que poderia ser prejudicial para a regeneração dos explantes posteriormente. Foram distribuídos cerca de 20 explantes por placa de petri (90 × 15 mm) em meio indutor de raiz (RIM), composto por sais MS, com vitaminas B5 (Gamborg), 30 g L⁻¹ de sacarose, 6 g L⁻¹ de ágar, 0,4 µM de ANA e 100 µM de acetoseringona (KIMURA; SINHA, 2008).

Figura 2.2 - Preparo dos explantes a partir dos cotilédones



Fonte: Adaptado de Kimura; Sinha (2008)

Para o processo de transformação, três gotas de suspensão de *A. tumefaciens* em MS líquido foram aplicadas com auxílio de uma micropipeta, de modo que a solução fosse o suficiente para cobrir completamente toda a superfície dos explantes incluindo as bordas. As placas foram incubadas à temperatura ambiente por 10 min. Posteriormente, o excesso de suspensão bacteriana foi removido com uma pipeta estéril e discos de papel de filtro esterilizados foram utilizados para secar os explantes.

Para o cocultivo, as placas foram mantidas no escuro a $28^{\circ}\text{C} \pm 1^{\circ}\text{C}$ por 48 h. Os explantes foram transferidos para meio de indução de brotamentos (SIM), composto por meio MS com vitaminas B5 (Gamborg), 30 g L^{-1} de sacarose, 6 g L^{-1} de ágar, $5 \mu\text{M}$ de zeatina. Nessa etapa, para a seleção dos explantes transformados, o meio foi suplementado com 100 mg L^{-1} de canamicina e 25 mg L^{-1} de timetim. As placas foram mantidas em fotoperíodo de 16 h a $27 \pm 1^{\circ}\text{C}$, uma subcultura contendo meio SIM foi realizada quinzenalmente até o surgimento e desenvolvimento dos brotos.

Os brotos bem desenvolvidos (2-4 mm) foram separados dos explantes e transferidos para frascos contendo 40 ml de meio RIM suplementado com os antibióticos canamicina e timetim para o alongamento e enraizamento. Os brotos foram mantidos neste meio de cultura até apresentarem condições ideais de tamanho e desenvolvimento para serem aclimatizados.

As plantas obtidas pelo processo de transformação (T) enraizadas foram transferidas para o substrato e colocadas em câmara de crescimento sob condições controladas com temperatura de $26^{\circ}\text{C} \pm 1^{\circ}\text{C}$ e fotoperíodo de 16 h. As plantas transgênicas (T) obtidas foram

autopolinizadas para a produção de sementes e gerações seguintes (F1). A canamicina foi utilizada para seleção *in vitro* de linhagens transgênicas. Parte das sementes obtidas foram submetida a germinação em MS suplementado com canamicina e, outra parte das plantas validada com a aspersão de 400 mg L⁻¹ canamicina nas folhas. A determinação da eficiência de transformação (%), foi apresentada pela fórmula:

$$\frac{\text{Total do número de brotos enraizados}}{\text{Total do número de explantes submetidos a infecção com } Agrobacterium} \times 100.$$

2.5 Extração de DNA-genômico

Para a extração de DNA genômico foram coletadas folhas de plantas regeneradas e aclimatizadas. O protocolo utilizado para a extração do DNA genômico foi o CTAB, com adaptações (ROGERS; BENDICH, 1988). O tecido obtido das folhas coletadas foi macerado com *beads* utilizando o equipamento TissueLyserLT (QIAGEN) em tubos de 2 ml. Após a quebra do tecido foi adicionado 1,4 ml de tampão CTAB seguido de 28 µL do antioxidante β-mercaptoetanol. A solução obtida foi homogeneizada em vórtex e incubada a 65°C por 40 min. Em seguida, as amostras foram mantidas em gelo durante 5 min e posteriormente centrifugadas por 10 min a 11000 rpm. Após a centrifugação, a fase aquosa foi coletada utilizando novos tubos, nessa etapa 800 µL de solução de clorofórmio e álcool isoamílico (24:1) foram adicionados. Após a homogeneização as amostras foram centrifugadas por 10 min a 11000 rpm e, para melhorar a eficiência de extração, essa etapa foi repetida. Cerca de 600 µL do sobrenadante foi coletado e transferido a novos tubos de 1,5 ml previamente identificados. A mesma proporção de sobrenadante coletada foi utilizada como parâmetro para a adição de álcool isopropílico. As amostras foram homogeneizadas por inversão e incubadas a -20 °C por pelo menos 1 h. Para a precipitação do DNA, as amostras foram submetidas a centrifugação por 20 min a 11000 rpm. O sobrenadante foi descartado e foi adicionado ao tubo 500 µL de etanol 75%. O material foi mantido em temperatura ambiente para completa secagem e, posteriormente foi ressuscitado em 25 µL de água ultrapura. Para avaliar a qualidade da extração, a quantificação foi feita com auxílio do espectrofotômetro Nanodrop, onde as relações de 1,8 a 2,2 foram consideradas ideais para prosseguir com as demais etapas de análise. A integridade das amostras também foi observada via eletroforese em gel de agarose.

2.6 Validação molecular, detecção da mutação e sequenciamento

Com a finalidade de analisar a presença da deleção esperada da segunda região uORF e validar a edição gênica, foram realizadas PCRs com *primers* específicos correspondentes à região, com as seguintes sequências:

- *Forward*: 5' GCTGAAGGCACTGTTTGGATTTGG 3';
- *Reverse*: 5' GGTTGCTGAGTCGAAGCCATGTC 3'.

As PCRs foram realizadas utilizando o kit comercial GoTaq® DNA *Polymerase* (Promega). Os parâmetros utilizados para cada ciclo de amplificação foram de desnaturação inicial a 95 °C por 2 min, seguido por 35 ciclos compostos de três etapas: desnaturação a 95 °C, por 1 min; anelamento a 60°C por 1 min; extensão a 72 °C por 30 seg e uma extensão final a 72 °C durante 5 min. O produto da reação foi observado em gel de eletroforese contendo agarose a 1,5% corado com brometo de etídio para visualização.

Como parâmetro de comprovação da edição gênica foi levado em consideração o tamanho do fragmento amplificado na revelação do gel de eletroforese, as plantas possivelmente editadas apresentariam bandas com 42 pb a menos do que de plantas controle, dessa maneira, as plantas contendo a deleção esperada teriam o tamanho 353 pb observados no resultado.

As amostras selecionadas, de acordo com os resultados obtidos por PCR e eletroforese, foram preparadas para análise de sequenciamento de DNA. O DNA genômico extraído foi submetido a novas reações de amplificação utilizando a enzima Phusion™ *High-Fidelity DNA Polymerase* (Thermo Scientific). Foram realizadas 3 repetições de reações para cada amostra. A PCR seguiu os seguintes padrões para cada ciclo de amplificação, a desnaturação inicial a 98 °C por 30 seg, seguido por 35 ciclos compostos de três etapas: desnaturação a 98 °C por 10 seg; anelamento a 67,7 °C por 30 seg; extensão a 72 °C por 12 seg e uma extensão final a 72 °C durante 10 min. Para essa PCR foram utilizados os mesmos *primers* apresentados anteriormente.

Para purificação do material o kit comercial *Wizard® SV Gel and PCR Clean-Up System* (Promega) foi utilizado. Os fragmentos obtidos foram clonados em vetor do tipo pGEM®-T *Easy* (Promega) seguindo as recomendações do fabricante. A transformação foi realizada em *E. coli* e o DNA plasmidial foi extraído. As amostras foram preparadas e submetidas a análise de sequenciamento do tipo Sanger. Para cada amostra, foram selecionados 4 clones que foram confirmados por PCR de colônia.

As sequências de plantas possivelmente editadas e planta controle foram alinhadas utilizando a ferramenta *ClustalW* através do software MEGA-X (KUMAR et al., 2018). E visualizadas pelo software Gene Doc (NICHOLAS; et al 1997).

2.7 Quantificação de açúcares dos frutos

Para avaliar a edição gênica apresentava os efeitos esperados relacionados ao aumento da concentração de açúcares, os teores de sacarose, frutose, glicose foram avaliados por cromatografia líquida de alta performance (HPLC) em frutos maduros. O protocolo de coleta e seleção dos frutos maduros para as análises foi feito por meio de triagem de frutos que apresentavam alteração de coloração, do estágio de verde maduro para início do amadurecimento ou estágio *breaker*. No estágio *breaker*, os frutos foram marcados para serem coletados após 10 dias, quando apresentassem a coloração vermelha típica de frutos maduros (FIGURA 2.3). Os frutos coletados foram submetidos a secagem em estufa a $68^{\circ}\text{C} \pm 1^{\circ}\text{C}$, sendo posteriormente macerados e armazenados até serem utilizados para a extração de açúcares.

Figura 2.3 – Estágios de amadurecimento dos frutos utilizados como padrões para a coleta



Legenda: A: Fruto verde maduro (*mature green*). B: Frutos em início de amadurecimento apresentando mudança de coloração de verde para amarelo (*breaker*). C: Fruto maduro (*ripe*). Barra de escala de 1 cm.

Fonte: Do autor (2022).

O protocolo de extração dos açúcares utilizado foi realizado de acordo com a metodologia apresentada por Kawaguchi et al. (2021), com adaptações. Foram adicionados 800 μL de etanol em 100 mg de fruto macerado, mantendo as amostras incubadas a 80°C por 30 min. Depois, o material foi centrifugado a 20°C em 21000 g por 5 min. O sobrenadante foi seco a vácuo e ressuspenso em 200 μL de água miliQ, sendo homogeneizado por vórtex. A mistura obtida foi novamente centrifugada a 5 min. Cerca de do sobrenadante foi coletado e

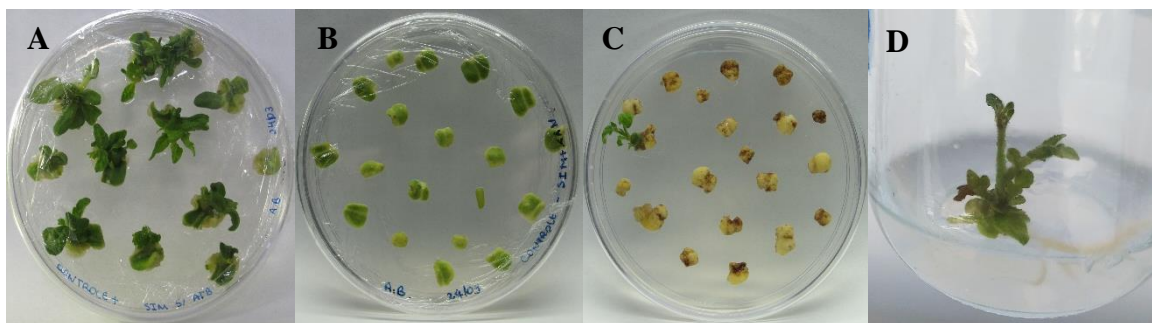
transferido para um novo tubo, onde foi adicionado 100 μ L de clorofórmio. A mistura obtida foi homogeneizada em vórtex e centrifugada novamente. Aproximadamente 20 μ L de sobrenadante foi coletado e ressuspensionado em 180 μ L de água miliQ. Para a análise de HPLC o material foi novamente diluído na proporção 1:1000.

3 RESULTADOS E DISCUSSÃO

3.1 Transformação e regeneração

Os protocolos de transformação e regeneração utilizados em tomateiro são bem estabelecidos (FIGURA 2.4). (CAMPOS et al., 2010; PINO et al., 2010; PACHECO LOMBARDI-CRESTANA et al. 2018). Contudo, é importante ressaltar que a regeneração vegetal adquirida pela organogênese é influenciada por vários fatores como o genótipo, a fonte e a idade dos explantes, a composição do meio de cultura, além das condições ambientais (KIMURA; SINHA, 2008; LIMA et al., 2009; CAMPOS et al., 2010).

Figura 2.4: Transformação e regeneração dos explantes



Legenda: Regeneração *in vitro* dos explantes. A) Explantes controle em meio sem canamicina. B) Explantes controle em meio seletivo suplementado com canamicina. C) Explantes expostos ao processo de transformação. D) Regeneração e enraizamento de um broto transformado.

Fonte: Do autor (2022).

Aproximadamente 2500 explantes foram submetidos ao protocolo de transformação, porém apenas 3 eventos sobreviveram as etapas de seleção, de regeneração, de enraizamento e de aclimatização. A pressão de seleção foi mantida sobre o tecido em regeneração para garantir que apenas brotos transformados sobrevivessem, entretanto, alguns escapes surgiram,

ou seja, foram observados a presença de brotos não transformados que não expressaram o gene marcador de seleção e que conseguiram regenerar durante o cultivo em placas contendo meio SIM. A validação dos brotos regenerados em meio com canamicina aconteceu na etapa de enraizamento, quando foram isolados dos explantes e transferidos para meio RIM, posicionados em contato direto com meio contendo o agente seletivo, onde não resistiram. Para essas situações o material biológico parcialmente regenerado, foi considerado como falso positivo.

Dessa forma, a taxa de transformação foi de 0,12%, valor bastante inferior quando comparado a outros trabalhos em que a eficiência de transformação é em média de 20 a 25% (PINO et al., 2010; GERSZBERG et al., 2015). É importante evidenciar que o aprimoramento do sistema de regeneração utilizando técnicas de cultura de tecidos em plantas, adotando estratégias para otimizar o protocolo é imprescindível para conseguir conquistar as potencialidades oferecidas pela engenharia genética (GERSZBERG et al., 2015). A otimização do protocolo e o aperfeiçoamento de técnicas de cultivo *in vitro* devem ser levados em consideração para aumentar a eficiência de obtenção de plantas transformadas, a manutenção do material vegetal com o uso de sementes que apresentem vigor e boa germinação também deve ser avaliados, uma vez que estes fatores podem ter influenciado nos resultados obtidos.

As 3 plantas transformadas (T) obtidas, foram aclimatizadas em câmara de crescimento e submetidas a autopolinização para a obtenção de sementes para as gerações seguintes (F1). Cada planta (T) recebeu a seguinte denominação T1, T2 e T3.

3.2 Planta T1

A planta T1 não apresentou déficits de desenvolvimento, produziu frutos e sementes que foram germinadas diretamente em substrato (FIGURA 2.5). Os frutos maduros obtidos foram coletados e submetidos a secagem para a análise dos açúcares, de acordo com a metodologia descrita anteriormente. Quanto ao resultado da quantificação dos açúcares glicose, frutose e sacarose, a planta T1 não apresentou aumento do teor de concentração destes carboidratos.

Figura 2.5 – Planta T1 e a geração F1 obtida a partir da germinação de sementes



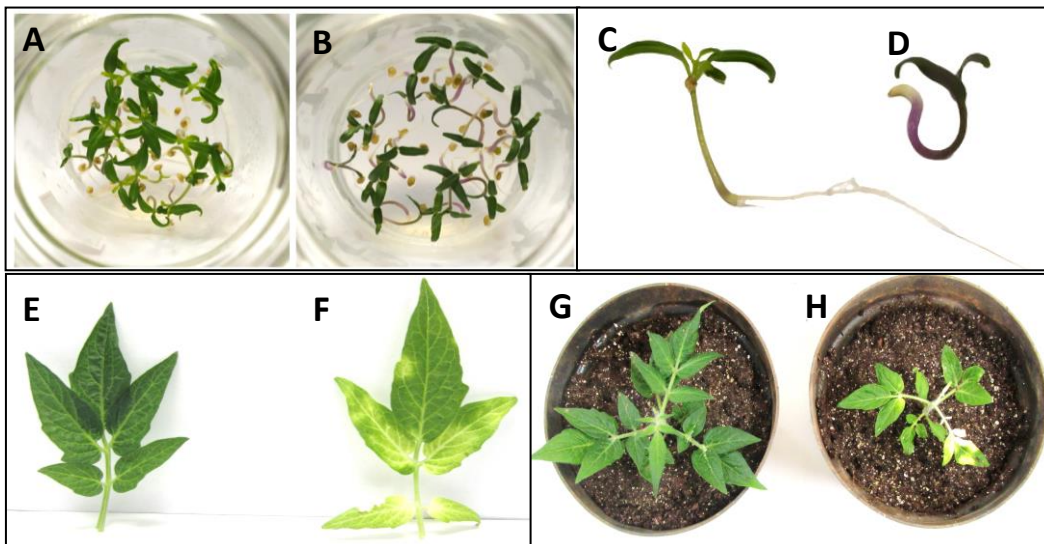
Legenda: A) Planta T1 aclimatizada. B) 18 plantas da geração F1 obtidas a partir da germinação de sementes T1.

Fonte: Do autor (2022).

3.2.1 Geração F1

Para avaliar a resistência ao agente seletivo parte das sementes obtidas da planta T1 e de plantas controle foram inoculadas *in vitro* com meio de germinação suplementado com canamicina 100 mg L^{-1} . As sementes T1 apresentaram boa taxa de germinação, desenvolvimento e enraizamento, em contrapartida, as sementes de plantas controle germinaram, porém não enraizaram e apresentaram coloração escura devido ao estresse causado pela canamicina (FIGURA 2.6 A-D). O teste de resistência a canamicina também foi aplicado às folhas de plantas controle e F1 da planta T1 durante um período de 10 dias em plantas jovens (FIGURA 2.6 E-H).

Figura 2.6 – Teste de resistência a canamicina, comparação entre sementes controle e transformadas.



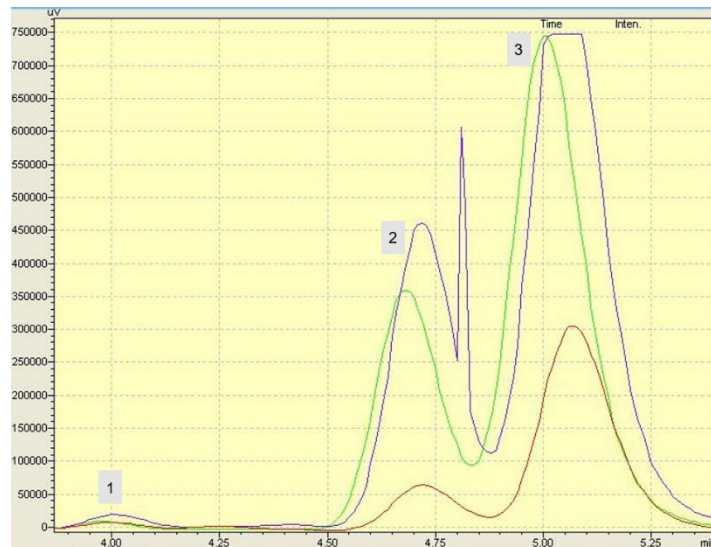
Legenda: A-D) Germinação *in vitro* de sementes em meio suplementado com canamicina. A) Plântulas provenientes da germinação de sementes T1 e em B) Plântulas provenientes da germinação de sementes controle. C) Enraizamento de F1 das sementes T1 e em D) Desenvolvimento comprometido de plântulas controle. E-H) Teste de resistência a canamicina em folhas. E e G) Plantas F1 da T1 resistentes e em G e H) Plantas controle apresentando clorose nas folhas em decorrência do contato com o antibiótico.

Fonte: Do autor (2022).

As folhas de plantas controle, sensíveis a canamicina, apresentaram branqueamento devido à inibição da síntese de proteínas associadas ao cloroplasto, efeito causado pelo contato com o antibiótico. A clorose permaneceu restrita às regiões que tiveram contato direto com o agente seletivo e não foram detectadas manchas cloróticas nas folhas emergentes após o período de aplicação. Quanto aos tomateiros obtidos de sementes T1, a maioria apresentou resistência, levando em consideração que essa era uma população segregante. A proporção de folhas resistentes e suscetíveis é concordante com a razão de segregação de 3:1, tendo em vista que a planta T1 contém apenas um *locus* de resistência à canamicina (WEIDE; KOORNNEEF; ZABEL, 1989) (APÊNDICE A).

Na Figura 2.5 são apresentadas as plantas obtidas de sementes oriundas da autofecundação da T1, os vasos foram identificados e enumerados de 01 a 18. Os tomateiros foram submetidos as análises de quantificação dos açúcares em frutos maduros. Quanto ao resultado observado, as plantas 11 e 18 apresentaram aumento significativo das concentrações de glicose e frutose (FIGURA 2.7).

Figura 2.7 - Perfil cromatográfico da quantificação de açúcares em frutos maduros



Legenda: Os picos referem-se aos açúcares analisados. 1) sacarose, 2) glicose e 3) frutose. As cores foram utilizadas para diferenciar as amostras. Vermelho – Fruto de planta controle. Verde – Fruto da planta 11. Roxo – fruto da planta 18.

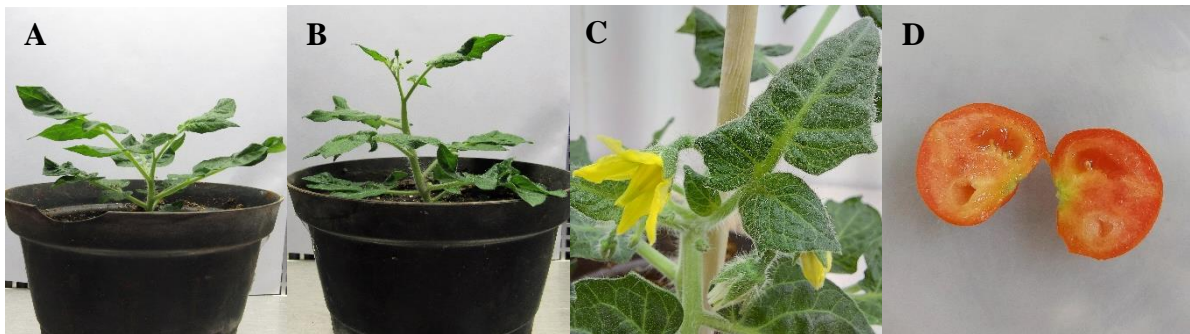
Fonte: Do autor (2021).

As plantas 11 e 18 apresentaram aumento dos açúcares frutose e glicose. O aumento da frutose apresentado pela planta 18 foi seis vezes mais alto do que a concentração da planta controle. Em outros estudos que utilizaram a estratégia de superexpressão ou de modificações da *uORF* de genes S-bZIPs, foram relatados o aumento de 50% da concentração dos açúcares totais em tomates e de 83,6% em morangos (SAGOR et al., 2016; XING et al., 2020).

3.3 Plantas T2 e T3

As plantas T2 e T3 apresentaram algumas diferenças morfológicas como o aumento da quantidade de tricomas, além da presença de baixa produtividade de frutos e de sementes (FIGURA 2.8).

Figura 2.8 – Características observadas nas plantas T2 e T3



Legenda: A) Planta T2. B) Planta T3. C) Tricomas nas folhas (T3). D) Fruto sem sementes (T2).
Fonte: Do autor (2022).

Os frutos maduros de ambas as plantas foram coletados para análises e não apresentaram aumento das concentrações de açúcares. No entanto, a quantidade de açúcares acumulados no fruto não depende apenas dos processos metabólicos endógenos, mas também do grau de deslocamento destes carboidratos pelo floema (FARRAR; WILLIAMS, 1991). Sendo assim, os açúcares podem ter sido direcionados ou armazenados em outros tecidos que não foram avaliados neste trabalho. Um desequilíbrio entre o balanço fonte-dreno, causado pela excedência da taxa de fotossíntese, pode resultar na interferência da capacidade de exportação dos produtos foto assimilados, resultando no acúmulo de carboidratos nas folhas (OSORIO; RUAN; FERNIE, 2014). O acúmulo de glicose, frutose e principalmente sacarose nas folhas pode levar a uma diminuição na fotossíntese por mecanismos de *feedback* inibitório (GAUTIER et al., 2010).

Com a possibilidade do acúmulo dos açúcares em outros tecidos e não em frutos maduros, foram visualizadas outras características nas plantas T2 e T3 que sugerem essa circunstância, como o abortamento das flores que ocorreu após o início do crescimento dos primeiros frutos. Este fato pode ter sido resultado de um desequilíbrio, uma vez que os estresses que incluem a deficiência de nutrientes podem induzir ao aborto de flores, sementes e frutos e, portanto, reduzem a produtividade (OSORIO; RUAN; FERNIE, 2014) (APÊNDICE B).

O gene *SibZIP1*, através da heterodimerização com outros bZIPs de classe C, faz parte da rede C/S1 bZIP, que atua ajustando o metabolismo de carbono e de nitrogênio, controlando o crescimento, o desenvolvimento e modulando respostas a estresses ambientais (PEVIANI et al., 2016; DRÖGE-LASER; WEISTE, 2018b). O nitrogênio está envolvido na manutenção das funções fisiológicas e no metabolismo energético de fotossíntese e respiração

(FICHTNER et al., 2021). Portanto, modificações em rotas de nitrogênio podem afetar a fixação de carbono diretamente por causa dos efeitos causados aos componentes fotossintéticos como a clorofila, o complexo coletor de luz e a rubisco. As alterações neste sistema podem ser limitantes para o crescimento e para o acúmulo de carboidratos, devido a insuficiência de nitrogênio que seria requisitado para a retroalimentação da fotossíntese (EVANS; POORTER, 2001).

Com a edição gênica direcionada para a região *uORF*, era esperado que em plantas mutantes os níveis de expressão do gene *SlbZIP1* fosse alterado e conseqüentemente a expressão de outros genes regulados por este bZIP, como a asparagina sintetase (ASN), a prolina desidrogenase (ProDH), a sacarose fosfatase (SPP) e a sacarose fosfato sintase (SPS) também seriam alterados (THALOR et al., 2012; SAGOR et al., 2016; DRÖGE-LASER; WEISTE, 2018b). O aumento da expressão do gene *tbz17* em *Nicotiana tabacum*, homólogo ao *SlbZIP1*, resultou em uma produção de plantas de tabaco com folhas mais espessas compostas por células aumentadas e com teor de sacarose 3 a 4 vezes maior em comparação com plantas de tipo selvagem, entretanto o crescimento e desenvolvimento das plantas foi comprometido, os autores observaram também um aumento significativo da expressão do gene ASN, gene que está envolvido no armazenamento e no transporte de nitrogênio na planta (THALOR et al., 2012).

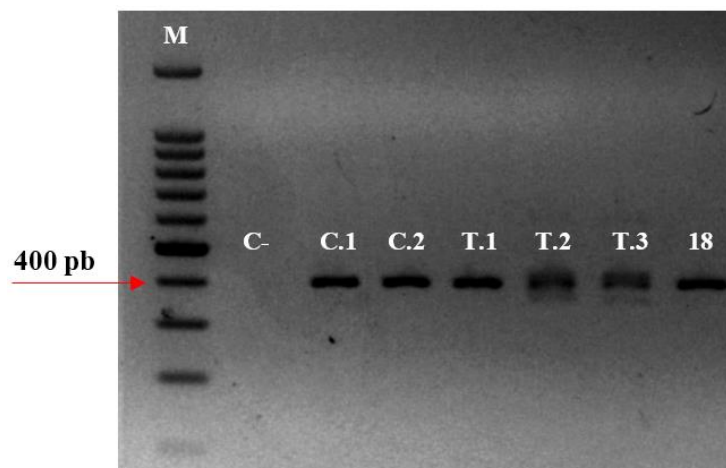
Outra característica observada quanto a estrutura das plantas T2 e T3 foi a presença em maior quantidade de tricomas (APÊNDICE B). Os tricomas são apêndices que se projetam da epiderme e podem ser classificados de acordo com a presença ou ausência de glândulas secretoras, quanto à função pode estar relacionada a proteção contra herbivoria de insetos, resistência a patógenos, controle da transpiração e tolerância a fatores abióticos como acontece no estresse hídrico (BRAND; TISSIER, 2022). Neste trabalho, não foi possível caracterizar os tricomas observados, no entanto o aumento dessas estruturas pode estar relacionado a alocação de recursos entre metabolismo primário e o especializado. A biossíntese de metabólitos especializados em plantas usa precursores que fazem parte de vias metabólicas primárias, ou seja, a produção destinada aos tricomas glandulares representam um “dreno” que desvia o fluxo do metabolismo central (BRAND; TISSIER, 2022). Em tomateiro foi realizada uma análise comparativa de dados ômicos sobre os tricomas glandulares, os resultados obtidos demonstraram que estas estruturas dependem da sacarose importada das folhas como a principal fonte de carbono para a sua formação (BALCKE et al., 2017). Outros estudos apresentam resultados que fornecem informações adicionais que

comprovam a função fundamental da sacarose como fonte de carbono nos tricomas em outros vegetais (JOHNSON et al., 2017; KUMAR NAUTIYAL et al., 2020; CONNEELY et al., 2021).

3.4 Detecção e sequenciamento das mutações

A detecção da edição gênica esperada foi inicialmente avaliada em gel de eletroforese para observar a diferença de tamanho do fragmento de amplificação da região alvo. Em organismos controle o *amplicon* correspondente seria de 395 pb enquanto em plantas editadas de 353 pb. Foram avaliadas as plantas T1, T2, T3 e 18, conforme apresentado na figura 2.9.

Figura 2.9- Reação de amplificação da região *uORF* do gene *SibZIP1*



Legenda: Visualização em gel de agarose a 1,5%. M) Marcador de peso molecular. C-) Controle negativo da reação (água). C.1) Planta controle 1. C.2) Planta controle 2. T.1) Planta transformada 1. T.2) Planta transformada 2. T.3) Planta transformada 3. 18) Planta 18 (F1-T1).

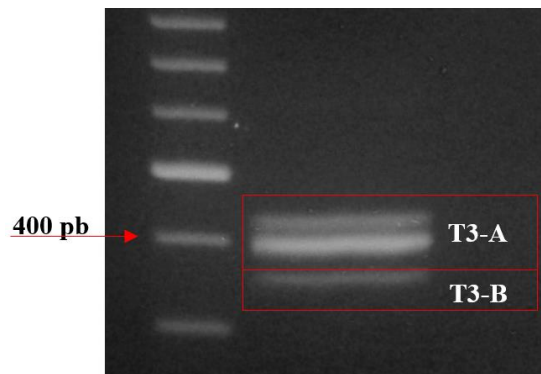
Fonte: Do autor (2022).

O resultado da estratégia utilizada, direcionando dois *sgRNAs* simultaneamente para criar uma deleção foi observado em T2 e T3. Entretanto, mutações pontuais em ambos os alvos dos *sgRNAs* também eram esperados e poderiam gerar o fenótipo desejado, visto que modificações sutis na região *uORF* são suficientes para impedir o mecanismo SIRT (XING et al., 2020).

A edição gênica pode ser classificada em bialélica quando possui dois alelos editados de forma diferente, homozigota quando há dois alelos identicamente mutados e, heterozigota quando é observado um alelo mutante junto a um alelo do tipo selvagem (BROOKS et al., 2014; D'AMBROSIO; STIGLIANI; GIORIO, 2018). De acordo com a observação da figura 2.9, pode-se inferir que as plantas T2 e T3 poderiam ter mutações bialélicas ou heterozigotas, enquanto a planta 18 por ter apresentado fenótipo com maior concentração de açúcares nos frutos, poderia apresentar mutação homozigota ou bialélica sem a presença de uma deleção que fosse possível de ser observada no gel. Portanto, a caracterização destes possíveis *indels* só poderiam ser visualizados através da análise do resultado de sequenciamento. Devido ao perfil de bandas observadas no gel e às condições das amostras de DNA, as plantas T3 e 18 foram selecionadas e conduzidas para o sequenciamento.

Com intuito de clonar individualmente cada *amplicon* observado da amostra T3, esta foi submetida a uma eletroforese em gel concentrado para separação das bandas (FIGURA 2.10).

Figura 2.10- Separação dos *amplicons* obtidos para a amostra T3



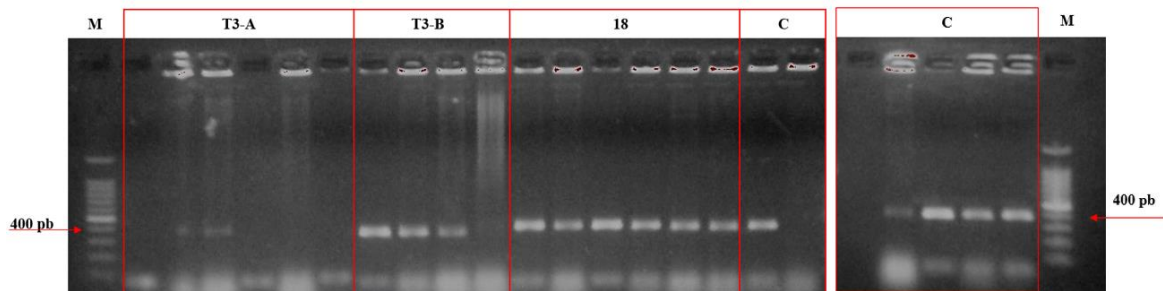
Legenda: Visualização em gel de agarose a 2,5%. T3-A) Banda dupla. T3-B) Banda menor.

Fonte: Do autor (2022).

Diferente do esperado para uma planta diploide, foram identificadas três bandas ao invés de duas, o que indica um possível quimerismo. Devido à proximidade entre as duas primeiras bandas a separação entre elas não foi possível para serem analisadas individualmente. Desta maneira para a T3, foram discriminadas duas amostras: uma contendo a banda dupla (T3-A) e outra com a banda menor (T3-B). Quatro clones para amostra T3-B, seis clones de uma planta controle (C), T3-A e planta 18, foram confirmados por PCR de colônia (FIGURA 2.11).

As amostras T3-A não apresentaram bons resultados. Por isso, foram submetidas as análises de sequenciamento Sanger, três clones da amostra T3-B e quatro das amostras 18 e C. Ao todo, 11 amostras foram submetidas ao sequenciamento.

Figura 2.11- PCR de colônias isoladas

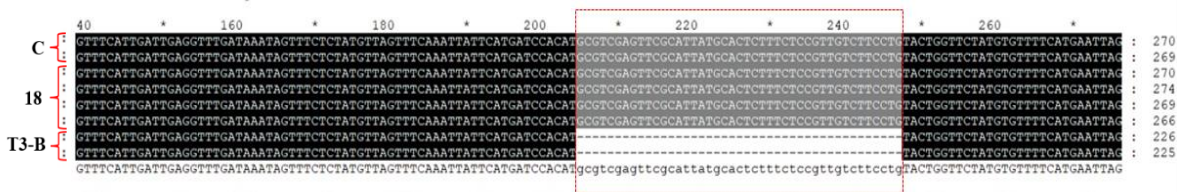


Legenda: Resultado obtido da amplificação de colônias isoladas contendo amostras de interesse para o sequenciamento. M) Marcador de peso molecular. T3-A) Amostras de banda dupla obtida para T3. T3-B) Amostras de banda isolada para T3. 18) Amostras referentes ao *amplicon* da planta 18. C) Amostras referentes ao *amplicon* de uma planta controle.

Fonte: Do autor (2022).

O resultado do alinhamento do sequenciamento demonstra que para as amostras T3, houve a deleção esperada conforme apresentado na figura 2.12. A deleção entre os dois locais de clivagem pela técnica CRISPR/Cas9 foi validada anteriormente por Cermák et al (2017), por meio de uma mutação realizada em protoplastos de tomateiro, onde a utilização de dois ou mais *sgRNA* foi eficiente para provocar uma mutagênese direcionada. A presença de *indels* curtos em um ou ambos os locais de clivagem também poderiam ser observados e eram esperados para as amostras 18, entretanto nenhuma outra mutação foi visualizada (APÊNCICE C). A ausência de edição gênica nas amostras da planta 18 podem indicar que apenas um dos alelos foi clonado e selecionado para o sequenciamento, por isso a possibilidade de outra mutação no outro alelo não deve ser descartada.

Figura 2.12- Resultado obtido pelo alinhamento entre o gene de referência *SibZIP1* (Solyc01g079480), plantas controle, plantas transformadas 18 e T3.



Legenda: Sequências que tiveram 100% de alinhamento estão indicados pela cor preta. C: indica amostras provenientes de uma planta controle. 18: planta de geração F1, que apresentou maior teor de açúcares. T3-B: parte do material genético da planta transformada 3.

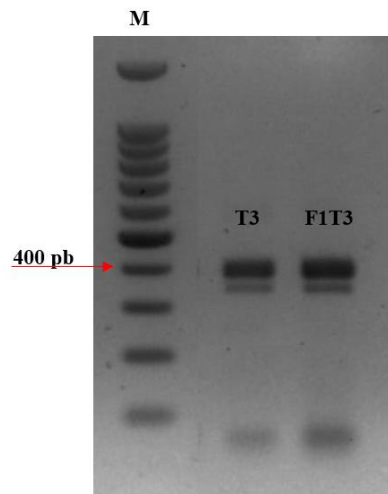
Fonte: Do autor (2022).

Foram sequenciados com sucesso duas amostras controle, quatro referentes a planta 18 e duas correspondentes ao fragmento de amplificação menor visualizado para T3. A ocorrência de setores quiméricos em tecidos vegetais transformados, como observado em T3 pode dificultar as análises e a fenotipagem (PATHAK et al., 2019). No entanto, o tomateiro é uma espécie de cultivo de reprodução sexual e os alelos editados de interesse poderiam ser transferidos por cruzamento para uma variedade de elite (BROOKS et al., 2014).

3.5 Geração F1 proveniente da planta editada

A planta T2 não produziu nenhuma semente viável. A partir da planta T3 foi possível obter uma única planta (F1), cujo perfil de bandas é similar ao da parental (FIGURA 2.13). Essa planta será utilizada para futuras análises.

Figura 2.13- Perfil de bandas visualizados para planta F1T3



Legenda: Visualização em gel de agarose a 1,5%. T3 é a planta comprovada para edição gênica, na figura é possível observar que ambas as amostras apresentam o mesmo perfil, por isso é esperado que a planta F1T3 também esteja editada geneticamente.

Fonte: Do autor (2022).

4 CONSIDERAÇÕES FINAIS

Os resultados apresentados por este trabalho permitirão a condução de novos experimentos com a intenção de se obter um número maior de mutantes, em que será possível ampliar e caracterizar as mutações mais frequentes.

Para aumentar a expressão de genes envolvidos no teor de açúcar do tomateiro as deleções observadas na região *uORF* seriam suficientes para gerar o fenótipo esperado, entretanto, apenas frutos maduros foram analisados neste trabalho e a hipótese do aumento de açúcares em outros tecidos não deve ser desconsiderada.

Existem diversos fatores associados ao aumento de açúcares, a composição genética é um dos aspectos que determinam o sabor do tomate, por isso, apesar de na planta 18 não ter sido visualizada a mutação esperada e, devido ao aumento significativo de açúcares observado em seus frutos, recomenda-se que análises adicionais sejam realizadas.

5 CONCLUSÃO

De acordo com os resultados apresentados, a estratégia de edição gênica proposta neste trabalho demonstra sua eficiência e a possibilidade de ser utilizada, uma vez que a deleção de 42 pb na *uORF* foi compatível com a região alvo. Demais estudos poderão ser realizados com intuito de identificar e caracterizar os efeitos causados pela edição gênica em plantas que possuem mutação no local genômico de interesse.

REFERÊNCIAS

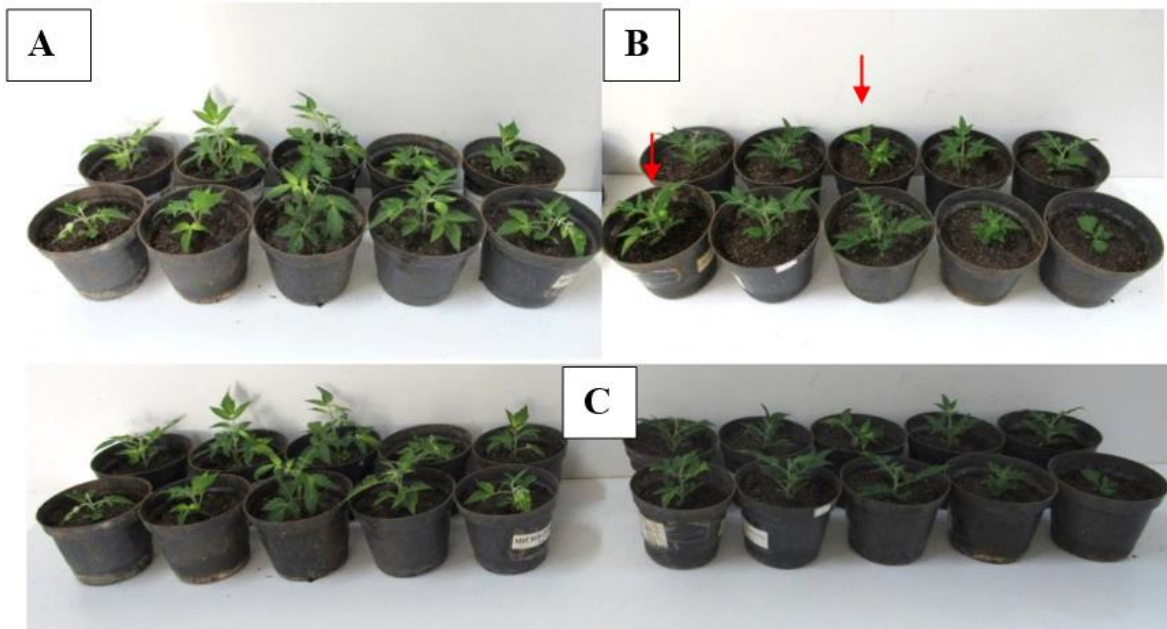
- BALCKE, G. U. *et al.* **Multi-Omics of Tomato Glandular Trichomes Reveals Distinct Features of Central Carbon Metabolism Supporting High Productivity of Specialized Metabolites OPEN.** [s.d.] Doi: 10.1105/tpc.17.00060. Acesso em: 3 jan. 2023.
- BRAND, A.; TISSIER, A. **Control of resource allocation between primary and specialized metabolism in glandular trichomes.** *Current Opinion in Plant Biology*, v. 66, p. 102172, 1 abr. 2022. Doi: 10.1016/J.PBI.2022.102172. Acesso em: 2 jan. 2023.
- BROOKS, C. *et al.* **Scientific Correspondence Efficient Gene Editing in Tomato in the First Generation Using the Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-Associated9 System 1.** [s.d.] Doi: 10.1104/pp.114.247577. Acesso em: 1 dez. 2022.
- CAMPOS, M. L. *et al.* **The micro-tom model system as a tool to discover novel hormonal functions and interactions.** *Plant Signaling and Behavior*, v. 5, n. 3, p. 267–270, 2010. Doi: 10.4161/psb.5.3.10622.
- CERMÁK, T. *et al.* **A Multipurpose Toolkit to Enable Advanced Genome Engineering in Plants.** [s.d.] Doi: 10.1105/tpc.16.00922. Acesso em: 1 jan. 2023.
- CONNELLY, L. J. *et al.* **Characterization of the Cannabis sativa glandular trichome proteome.** *PLoS ONE*, v. 16, n. 4 April, 1 abr. 2021. doi: 10.1371/journal.pone.0242633.
- CORREA, H. B. *et al.* **Tecnologia CRISPR na edição genômica de plantas Biotecnologia aplicada à agricultura.** [s.l: s.n.].
- D'AMBROSIO, C.; STIGLIANI, A. L.; GIORIO, G. **CRISPR/Cas9 editing of carotenoid genes in tomato.** *Transgenic Research*, v. 27, n. 4, p. 367–378, 1 ago. 2018. Doi: 10.1007/s11248-018-0079-9.
- DRÖGE-LASER, W.; WEISTE, C. **The C/S1 bZIP Network: A Regulatory Hub Orchestrating Plant Energy Homeostasis.** *Trends in Plant Science*, v. 23, n. 5, p. 422–433, 2018. Doi: 10.1016/j.tplants.2018.02.003.
- EVANS, J. R.; POORTER, H. **Photosynthetic acclimation of plants to growth irradiance: The relative importance of specific leaf area and nitrogen partitioning in maximizing carbon gain.** *Plant, Cell and Environment*, v. 24, n. 8, p. 755–767, 2001. Doi: 10.1046/j.1365-3040.2001.00724.x.
- FARRAR, J. F.; WILLIAMS, J. H. H. **Control of the rate of respiration in roots: compartmentation, demand and.** *Compartmentation of plant metabolism in non-photosynthetic tissues*, 1991, 42: 167.
- GAUTIER, H. *et al.* **Functional Plant Science and Biotechnology Impact of Ripening and Salinity on Tomato Fruit Ascorbate Content and Enzymatic Activities Related to Ascorbate Recycling.** [s.l: s.n.].
- GEBHARDT, C. **The historical role of species from the Solanaceae plant family in genetic research.** *Theoretical and Applied Genetics*, 2016. Doi: 10.1007/s00122-016-2804-1.
- GENEDOC: **A tool for editing and annoting multiple sequence alignments – Science Open.** Acesso em: 1 jan. 2023.

- GERSZBERG, A. *et al.* **Tomato (*Solanum lycopersicum* L.) in the service of biotechnology.** Plant Cell, Tissue and Organ Culture Kluwer Academic Publishers, 1 mar. 2015. Doi: 10.1007/s11240-014-0664-4.
- JAKOBY, M. **bZIP transcription factors in Arabidopsis. 2002.** Acesso em: 5 dez. 2022.
- JAKOBY, M. *et al.* **bZIP transcription factors in Arabidopsis.** Trends in Plant Science, v. 7, n. 3, p. 106–111, 2002. Doi: 10.1016/S1360-1385(01)02223-3.
- JINEK, M. *et al.* **A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity.** [s.d.]. Acesso em: 1 jan. 2023.
- JOHNSON, S. R.; LANGE, I.; SRIVIDYA, N.; LANGE, B. M. **Bioenergetics of Monoterpenoid Essential Oil Biosynthesis in Nonphotosynthetic Glandular Trichomes.** [s.d.]. Doi: 10.1104/pp.17.00551. Acesso em: 3 jan. 2023.
- KAWAGUCHI, K. *et al.* **Functional disruption of cell wall invertase inhibitor by genome editing increases sugar content of tomato fruit without decrease fruit weight.** Scientific Reports, v. 11, n. 1, p. 21534, 1 dez. 2021. doi: 10.1038/S41598-021-00966-4. Acesso em: 1 jan. 2023.
- KIMURA, S.; SINHA, N. **Tomato transformation.** Cold Spring Harbor Protocols, v. 3, n. 11, nov. 2008. Doi: 10.1101/pdb.prot5084.
- KUMAR N. *et al.* **Comprehensive transcriptome analysis provides insights into metabolic and gene regulatory networks in trichomes of *Nicotiana tabacum*.** Plant Molecular Biology, v. 102, p. 625–644, 2020. Doi: 10.1007/s11103-020-00968-2. Acesso em: 3 jan. 2023.
- KUMAR, S. *et al.* **MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms.** [s.d.] Doi: 10.1093/molbev/msy096. Acesso em: 1 jan. 2023.
- LI, D. *et al.* **Genome-wide systematic characterization of the bZIP transcriptional factor family in tomato (*Solanum lycopersicum* L.).** BMC Genomics, v. 16, n. 1, p. 1–18, 2015. Doi: 10.1186/s12864-015-1990-6.
- LI, Y. *et al.* **MicroTom Metabolic Network: Rewiring Tomato Metabolic Regulatory Network throughout the Growth Cycle.** Molecular Plant, v. 13, n. 8, p. 1203–1218, 2020. Doi: 10.1016/j.molp.2020.06.005.
- LIMA, J. E. *et al.* **Callus, shoot and hairy root formation in vitro as affected by the sensitivity to auxin and ethylene in tomato mutants.** Plant Cell Reports, v. 28, n. 8, p. 1169–1177, 2009. Doi: 10.1007/s00299-009-0718-y.
- OSORIO, S.; RUAN, Y. L.; FERNIE, A. R. **An update on source-to-sink carbon partitioning in tomato.** Frontiers in Plant Science. Frontiers Media S.A., 6 out. 2014. Doi: 10.3389/fpls.2014.00516.
- PACHECO LOMBARDI-CRESTANA *et al.* **Manual do modelo vegetal micro-tom capítulo 5: protocolos para cultivo in vitro e transformação de micro-tom.** [s.d.]. Acesso em: 1 jan. 2023.
- PATHAK, B. *et al.* **Dual targeting by CRISPR/Cas9 leads to efficient point mutagenesis but only rare targeted deletions in the rice genome.** 3 Biotech, v. 9, n. 4, 1 abr. 2019. Doi: 10.1007/s13205-019-1690-z.

- PEVIANI, A. *et al.* **The phylogeny of C/S1 bZIP transcription factors reveals a shared algal ancestry and the pre-angiosperm translational regulation of S1 transcripts.** Nature Publishing Group, 2016. Doi: 10.1038/srep30444. Acesso em: 1 jan. 2023.
- PINO, L. E. *et al.* **The Rg1 allele as a valuable tool for genetic transformation of the tomato “Micro-Tom” model system. 2010a** Acesso em: 22 set. 2021.
- ROGERS, S. O.; BENDICH, A. J. **Extraction of DNA from plant tissues.** Plant Molecular Biology Manual, v. 6, p. 1–10, 1988. Acesso em: 1 jan. 2023.
- SAGOR, G. H. M. *et al.* **A novel strategy to produce sweeter tomato fruits with high sugar contents by fruit-specific expression of a single bZIP transcription factor gene.** Plant Biotechnology Journal, v. 14, n. 4, p. 1116–1126, 1 abr. 2016. Doi: 10.1111/pbi.12480.
- SI, X. *et al.* **Manipulating gene translation in plants by CRISPR-Cas9-mediated genome editing of upstream open reading frames.** Nature Protocols, [s.d.] Doi: 10.1038/s41596-019-0238-3. Acesso em: 1 jan. 2023.
- THALOR, S. K. *et al.* **Deregulation of sucrose-controlled translation of a bZIP-type transcription factor results in sucrose accumulation in leaves.** PLoS ONE, v. 7, n. 3, 22 mar. 2012. Doi: 10.1371/journal.pone.0033111.
- WEIDE, R.; KOORNNEEF, M.; ZABEL, P. **A simple, nondestructive spraying assay for the detection of an active kanamycin resistance gene in transgenic tomato plants.** Theor Appl Genet. [s.l.] Springer-Verlag, 1989.
- XING, S. *et al.* **Fine-tuning sugar content in strawberry.** Genome Biology, v. 21, n. 1, 3 set. 2020. doi: 10.1186/s13059-020-02146-5.
- YAMASHITA, Y. *et al.* **Sucrose sensing through nascent peptide-mediated ribosome stalling at the stop codon of Arabidopsis bZIP11 uORF2.** FEBS Letters, v. 591, n. 9, p. 1266–1277, 2017. doi: 10.1002/1873-3468.12634.

APÊNDICES

APÊNDICE A – Teste de resistência a canamicina



Legenda: A) Plantas controle. B) F1 da planta T1. C) Visão geral entre plantas controle e F1T1. Indicado pelas setas em vermelho, são plantas suscetíveis.

Fonte: Do Autor (2022)

APÊNDICE B - Abortamento das flores e presença de tricomas

Legenda: As fotos referem-se a planta T3. A) Presença de flores e desenvolvimento de frutos. B) Abortamento das flores. C) Presença de tricomas nas folhas.

Fonte: Do Autor (2022)

CAPÍTULO 3 – ANÁLISE DE GENES CORRELACIONADOS AO *SibZIP1*

RESUMO

Uma grande quantidade de dados transcriptômicos são gerados pelo sequenciamento de RNA e a integração e interpretação desses conjuntos de dados tem sido um desafio. A busca por padrões de similaridade de expressão gênica utilizando ferramentas de coeficiente de correlação contribuem para agregar informações biológicas relevantes que podem ser aplicadas em diversos estudos. O objetivo apresentado neste capítulo do trabalho, foi realizar a prospecção em estudos já publicados de dados de RNA-seq em tomateiro, para realizar a montagem de uma rede correlação gênica em relação ao gene de interesse, o *SibZIP1*. Foram selecionadas 193 bibliotecas de diferentes tecidos e condições, em que os dados foram mapeados, alinhados ao genoma de referência e normalizados em TPMs. A estatística de correlação de *Spearman* foi utilizada para estabelecer a rede de correlação gênica, para ser visualizada por meio do software *cytoscape*. Foram identificadas e caracterizadas 85 sequências gênicas com padrão de expressão semelhante ao *SibZIP1*. Na rede de correlação obtida a presença de quinases e outros fatores transcrição se mostrou expressiva em relação aos demais transcritos. Os resultados obtidos, indicam que o gene *SibZIP1* tem forte correlação com genes que conferem tolerância a diferentes estresses abióticos, como estresse salino e hídrico ou que são responsivos a essas condições. Novas constatações e futuras interpretações podem auxiliar para a ampliação do conhecimento a respeito de genes que não são diretamente regulados por *SibZIP1*, mas que de certa forma apresentam níveis de expressão semelhantes sob as mesmas condições. A partir do banco de dados gerados neste estudo para os níveis de expressão em TPM, outras interpretações de informações biologicamente relevantes poderão ser exploradas e utilizadas para a montagem de outras redes de co-expressão ou para análises de expressão diferencial entre as bibliotecas obtidas.

Palavras-chave: Correlação gênica; RNA-seq; bZIP.

1 INTRODUÇÃO

O transcriptoma é dinâmico e determinado pontualmente de acordo com o estágio de desenvolvimento ou condição fisiológica, por isso, tem sido fundamental para a interpretação de funções e compreensão do genoma. Em virtude da grande quantidade de dados transcriptômicos gerados pelo sequenciamento de RNA a integração e a interpretação desses conjuntos de dados tem sido um desafio (TRAPNELL et al., 2013; COSTA-SILVA; DOMINGUES; LOPES, 2017).

O sequenciamento de moléculas de RNA permitiu o surgimento da técnica de análise de RNA-seq, utilizada para a caracterização e quantificação de transcritos. O RNA-seq é atualmente uma das ferramentas mais comum de ser utilizada para analisar a expressão global do gene, usando o número de leituras atribuídas a cada gene como um indicador do nível de expressão (SCHMID, 2017; ZHANG et al., 2020). O uso de estudos de análises de expressão gênica tem sido amplamente utilizado para compreender a complexidade dos organismos biológicos e gerar hipóteses testáveis para entender os mecanismos moleculares que controlam a expressão gênica. Uma estratégia para gerar hipóteses tratáveis a partir de dados transcriptômicos, tem sido a construção de gráficos de redes com base em padrões de correlação gênica. Diversos dados experimentais de RNA-seq estão disponíveis em fontes públicas, por isso os conjuntos de dados públicos de sequenciamento de RNA são um recurso valioso para análises de transcriptoma e, podem ser utilizados para a integração em um banco de dados local (SERIN et al., 2016; ZHANG et al., 2020).

A organização dos dados utilizando a construção de redes de correlação gênica, são frequentemente aplicadas para agrupar as informações de acordo com a função ou outras similaridades entre dois ou mais grupos. O resultado obtido a partir de uma rede de co-expressão gênica, exige interpretação biológica que vai além da simples visualização de genes correlacionados e o seu grau de complexidade pode ser agregado através da integração de dados, com intuito de aprimorar ainda mais a elucidação das relações entre os genes (SERIN et al., 2016).

Neste contexto, visando a elucidação de conhecimentos a respeito de genes possivelmente relacionados ao *SibZIP1*, o objetivo apresentado neste capítulo do trabalho foi a seleção de bibliotecas de dados de RNA-seq em tomateiro em condições e estágios diversos para estabelecer um banco de dados local e realizar análises de identificação de genes que

apresentassem similaridades de expressão gênica e a partir disso, estabelecer uma rede de correlação.

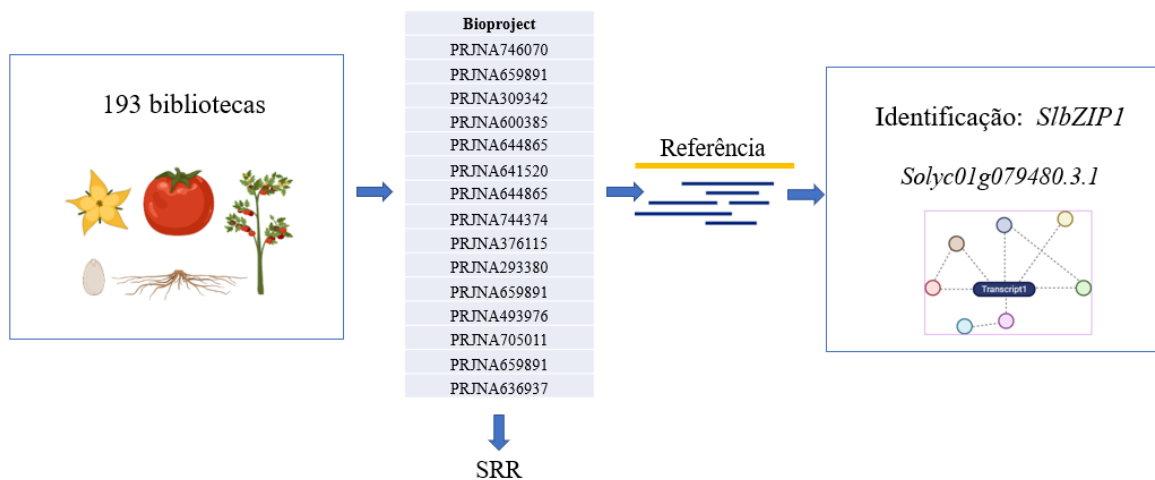
2 MATERIAIS E MÉTODOS

2.1 Seleção das bibliotecas e tratamento dos dados

Por meio do banco de dados do *National Center for Biotechnology Information* (<https://www.ncbi.nlm.nih.gov/>), foi realizada prospecção em estudos já publicados de dados de RNA-seq em tomateiro. As buscas foram realizadas na base de dados *Sequence Read Archives do National Center of Biotechnology* (SRA-NCBI), utilizando a ferramenta ‘*run selector*’ para filtrar as características de interesse como os diferentes tecidos de tomateiro e sob diferentes fases de desenvolvimento e diversas condições. Como parâmetro, foram priorizadas apenas bibliotecas obtidas de sequenciamento Illumina, por ser a plataforma de RNA-seq mais comumente utilizada, com quase 84% de todos os dados de RNA-seq gerados (CONTRERAS-LÓPEZ et al., 2018).

Os dados obtidos foram submetidos a ferramenta de controle de qualidade para dados de sequência de alto rendimento FastQC versão 0.11.9 de *reads* pareadas (WINGETT; ANDREWS, 2018). As sequências de adaptadores e de baixa qualidade foram removidos pelo *software* Trimmomatic (BOLGER; LOHSE; USADEL, 2014). Posteriormente, os dados obtidos foram mapeados e alinhados utilizando o Salmon (PATRO et al., 2017) à referência de *S. lycopersicum* versão ITAG4.0, disponível na plataforma SolGenomics (<https://solgenomics.net/>). Para estimar os níveis de expressão dos transcritos os dados foram normalizados em TPM (*transcript per million*). Em síntese, as etapas de análises utilizadas podem visualizadas na figura 3.1

Figura 3.1 - Etapas de análises dos dados de RNA-seq selecionados



Legenda: Seleção de bibliotecas em dados públicos disponibilizados no NCBI - como parâmetro de escolha foram priorizados tecidos de diferentes partes da planta. Após organizados e tratados os dados, foram mantidas apenas sequências (SRR) com boa qualidade. Posteriormente, as sequências foram identificadas por meio do alinhamento ao genoma de referência, onde o gene *SlbZIP1* foi localizado e os genes com padrões de expressão semelhantes (TPMs) foram utilizados para comparações e montagem da rede de correlação.

Fonte: Do autor (2023).

2. 2 Rede de correlação e caracterização *in silico*

A correlação de expressão entre os transcritos obtidos foi realizada utilizando o pacote do software R para a estatística de correlação de *Spearman* com um coeficiente de $\rho \geq 0,8$. Para a visualização dos dados foi utilizado o software cytoscape versão 3.9.1. (KOHL; WIESE; WARSCHEID, 2011).

A partir do banco de dados local gerado dos transcritos, foi necessário identificá-los para tal foi utilizado o software HMMER para a busca e predição por sequências homólogas, visando determinar a classificação e a função das proteínas. O software *TF prediction server* do banco de dados PlantTFDB (<http://planttfdb.cbi.pku.edu.cn/>) foi utilizado para identificar os FTs. Para predizer o número e a posição dos domínios transmembranas de cada proteína a ferramenta CCTOP (<https://cctop.ttk.hu/>) foi utilizada. Para a predição da localização subcelular de proteínas, foi utilizado o algoritmo DeepLoc2.0 (<https://services.healthtech.dtu.dk/service.php?DeepLoc-2.0>).

3 RESULTADOS E DISCUSSÃO

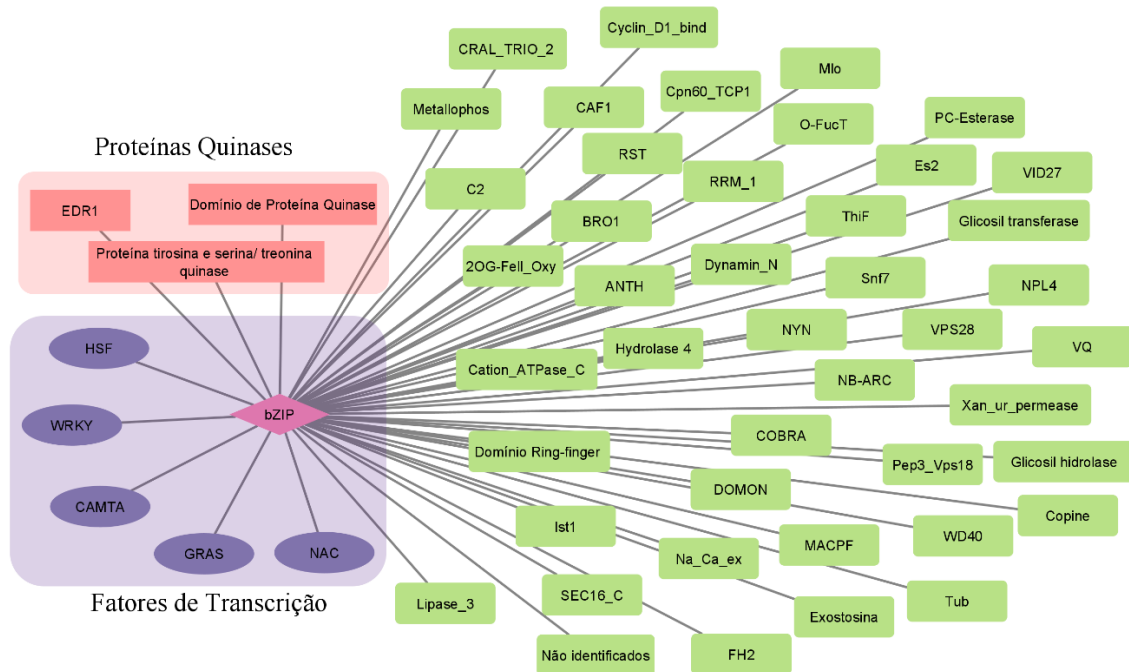
3.1 Seleção das bibliotecas

Visando diversificar a fonte de dados, foram selecionadas 193 bibliotecas referentes a diversos tecidos como folhas, frutos, meristemas, flores, raízes, tricomas e sementes que podem consultados pela tabela apresentada no apêndice A. Os conjuntos de dados públicos de RNA-seq exigem processamento e organização dos dados brutos, por isso a normalização foi realizada para gerar a matriz de dados contendo os genes, alinhados ao genoma de referência do tomateiro com os valores de transcrição em TPMs. A análise de correlação por meio do método *Spearman*, calculou o grau de similaridade entre os perfis de expressão gênica relacionados ao *SibZIP1* (Solyc01g079480.3.1) onde foram identificadas 85 sequências com similaridade $\rho \geq 0,8$. Além disso, para complementar as informações obtidas, a identificação e a classificação dos domínios transmembranas e demais informações sobre esses genes também foram apresentadas (APÊNDICES B e C).

3.2 Rede de correlação

Foram encontrados quinases e outros FTs, que em escala de hierarquia não se diferenciaram entre si (APÊNDICE D). Por isso, com intuito de facilitar a visualização, os genes foram classificados na rede e os dois grupos mais significativos, de domínios de proteínas quinases e os FTs (FIGURA 3.2).

Figura 3.2 - Representação de genes correlacionados ao *SibZIP1*



Legenda: 85 transcritos obtidos pela correlação de *Spearman* ($p \geq 0,8$). Ao centro em rosa está em evidência o gene *SibZIP1*, ligados a ele tem-se dois grupos principais - em roxo os FTs e em rosa os genes de domínios quinases. Os demais genes estão representados em verde e pertencem a diversas famílias e grupos gênicos com funções variadas.

Fonte: Do autor (2022).

A análise da rede genética de dados transcriptômicos é uma estratégia eficaz para a identificação de genes candidatos em biologia vegetal (ZAINAL-ABIDIN et al., 2022). A apresentação deste modelo de rede possui características qualitativas e, camadas de complexidade podem ser acrescidas, contudo em uma fase inicial essa estratégia pode ser aplicada para gerar hipóteses concretas testáveis para demais estudos de acompanhamento. A compreensão sobre a organização, função e evolução genética das plantas é essencial para o entendimento de processos biológicos e sua complexidade. A similaridade nos padrões de expressão gênica usando coeficiente de correlação permite a seleção de genes com altas pontuações de co-expressão e informações biológicas relevantes, que podem ser extraídas por priorização de genes. Este processo consiste em integrar e identificar grupos de genes funcionalmente relacionados, conforme foi apresentado na figura 3.2 (ZHANG et al., 2020; ZAINAL-ABIDIN et al., 2022).

Os dados obtidos a partir da visualização da rede correlação permite algumas inferências. A presença de transcritos de domínios quinases foi expressiva e sabe-se que os domínios bem conservados dessas proteínas são responsáveis pela fosforilação de outras proteínas, a partir da transferência de um grupo fosforil do trifosfato de adenosina (ATP) ou outros fosfatos de nucleosídeos, para um resíduo hidroxila. A fosforilação de proteínas é uma modificação pós-traducional bioquímica envolvida em todas as vias de sinalização e atividades celulares em organismos vivos, influenciando nas redes de sinalização celular em resposta à estimulação na própria célula ou de causa ambiental. A regulação reversível das funções das proteínas por meio de sua ativação ou inativação, atuam na formação de complexos de proteínas e no arranjo de localizações de proteínas subcelulares (APÊNDICE B). (BECRAFT, 2002; HOHMANN; LAU; HOTHORN, 2017).

Presume-se que um terço de todas as proteínas eucarióticas sejam fosforiladas. A fosforilação é resultante da ação antagônica de proteínas quinases e proteínas fosfatases, esse mecanismo é encontrado principalmente em aminoácidos hidroxilados, como serina (Ser), treonina (Thr) e tirosina (Tyr) que desempenham papel fundamental na regulação da atividade metabólica e na modulação da dinâmica do crescimento e desenvolvimento da planta. Essas quinases estão envolvidas em respostas a estresses como o déficit hídrico, a salinidade, a baixa temperatura ou o calor, ficou evidente na rede de correlação apresentada uma quantidade expressiva de domínios quinases (BECRAFT, 2002; GHELIS, 2011; HOHMANN; LAU; HOTHORN, 2017; SHAH et al., 2021).

A presença de EDR1 também foi visualizada na rede de correlação e este grupo gênico codifica uma quinase envolvida na resistência a doenças, como o oídio causado pelo fungo *Erysiphe cichoracearum*. O gene EDR1 tem como função a comunicação cruzada entre a sinalização do etileno e do ácido salicílico e está envolvido na sinalização de resposta ao estresse e na regulação da morte celular (TANG; CHRISTIANSEN; INNES, 2005; LIN et al., 2008).

Na correlação foram encontrados 7 fatores de transcrição, incluindo outro gene da família bZIP (*Solyc06g009640.2.1*). Filogeneticamente os dois bZIPs (*Solyc01g079480.3.1* e *Solyc06g009640.2.1*) são próximos e apresentam muitas similaridades, pertencendo ao mesmo grupo de classificação S que é caracterizado pela ausência de íntrons (LI et al., 2015). Os genomas vegetais têm mais de 2.000 genes que codificam os FTs de muitas famílias diferentes, por isso não é incomum que estes estejam presentes em redes de correlação gênica. A presença e a identificação dos FTs são utilizadas como parâmetros para definir os

agrupamentos em redes mais complexas (HONG, 2016). A proporção de genes FTs para o número total de genes em *Arabidopsis* é de 5 a 10% e a regulação transcricional desse grupo é bastante diversificado em plantas. Algumas famílias gênicas de FTs podem ter muitos membros como as de domínios AP2-ERF e NAC, MYB, MADS-box, bHLH, bZIP e HB. Essas famílias desempenham papéis importantes no controle do crescimento e do desenvolvimento das plantas (SCHWECHHEIMER; ZOURELIDOU; BEVAN, 1998; HONG, 2016).

Os FTs visualizados na rede de correlação, estão envolvidos com a modulação de respostas a diversos estresses. Como por exemplo, o HSF que é relacionado a resistência ao estresse térmico, sua ação acontece por meio da permissão do acúmulo rápido de proteínas de choque térmico (HSPs). Em resposta ao estresse e em estudos funcionais foi demonstrado que o HsfA1a tem uma função única como regulador fundamental da termo tolerância adquirida nas plantas (HONG, 2016).

A família WRKY é uma das famílias FTs específicas de plantas mais bem estudadas e compreende 74 membros em *Arabidopsis*. Estão envolvidos em vários processos de desenvolvimento e fisiológicos como na embriogênese, no revestimento da semente e no desenvolvimento de tricomas, na biossíntese de antocianinas e na sinalização hormonal. O zíper de leucina contido nas proteínas WRKY do Grupo IIA, forma homo ou heterodímeros complexos que cumprem um papel relevante na reação dos vegetais em condições de escassez de água, dessa forma pode-se afirmar que as proteínas WRKY são reguladas positivamente no nível proteico pelo estresse hídrico. Em tomateiro, a super expressão do gene *SIWRKY8* conferiu maior resistência ao patógeno *Pseudomonas syringae*, além de conferir também maior regulação positiva frente a estresses abióticos como seca, salinidade e frio. A abertura estomática foi menor sob estresse hídrico em plantas transgênicas para este gene, mantendo assim um maior teor de água nas folhas em comparação com plantas controle (GAO et al., 2020).

As proteínas CAMTA estão envolvidas na aclimação ao frio e fornecem um possível ponto de integração da sinalização entre cálcio e calmodulina em baixas temperaturas, tendo a sua expressão gênica regulada pelo frio. São importantes na orquestração das primeiras mudanças de resposta rápida na expressão gênica, que ocorrem após a exposição a baixas temperaturas. As proteínas CAMTA e o elemento CG-1 ao qual se ligam, são altamente conservados em organismos eucarióticos multicelulares, desde insetos, mamíferos e plantas (DOHERTY et al., 2009).

As proteínas GRAS compreendem uma importante família de proteínas específicas de plantas, participam no desenvolvimento da raiz e da parte aérea, na sinalização do ácido giberélico (GA) e na transdução de sinal do fitocromo A, na associação micorrízica nodular e arbuscular. Entretanto, uma compreensão completa de muitas de suas características e as funções biológicas específicas da maioria dos membros desta família, permanecem desconhecidas. Para uma melhor compreensão do papel dos GRAS no desenvolvimento e estresse da planta, é fundamental identificar seus parceiros de interação em níveis transcricionais (WASEEM et al., 2022). Em tomateiro, plantas transgênicas super expressando o gene *SIGRAS40* apresentaram fenótipos pleiotrópicos que conferiram maior tolerância ao estresse salino e à seca (LIU et al., 2017).

Os NAC são FTs específicos de plantas que apresentam relação ao desenvolvimento da planta, bem como na modulação de respostas de estresses abióticos e bióticos. Alguns genes NAC são regulados positivamente durante infecções bacterianas, enquanto outros medeiam a resistência viral. Como muitos NAC são induzidos por vários tratamentos de estresse, os genes responsivos ao estresse são agrupados no grupo SNAC (NAC responsivo ao estresse), que possui três subgrupos (SNAC-A, B e C). Foi relatado que a superexpressão de fatores NAC ajuda a alcançar uma tolerância melhorada ao sal em muitas plantas, como *Arabidopsis*, arroz, grão de bico, tomateiro e crisântemo (SCHWECHHEIMER; ZOURELIDOU; BEVAN, 1998; HONG, 2016; SONG et al., 2016).

Aspectos de resistência e tolerância a estresse salino e hídrico foram reportados em estudos envolvendo o bZIP. O silenciamento de *SibZIP1* em tomateiro, resultou em tolerância reduzida a sal e estresse hídrico em comparação com plantas controle. Os dados sugerem que este bZIP exerce um papel essencial na tolerância através da modulação de vias mediada por ABA (ZHU et al., 2018).

Seca, salinidade e temperaturas extremas são fatores ambientais primários que limitam o crescimento das plantas e reduzem a produção agrícola (ZHU, 2016). Diversos mecanismos de níveis moleculares, celulares e fisiológicos permitiram à adaptação das plantas para suportar essas situações desfavoráveis e essas características podem ser aprimoradas por meio da engenharia genética. Os FTs são proteínas regulatórias essenciais em redes de transdução de sinal ativadas na resposta das plantas a várias condições. Diversos componentes citoplasmáticos como cálcio e quinases são fundamentais para a sinalização e respostas moleculares em conjunto com FTs (HONG, 2016). Na rede de correlação gênica obtida, é possível visualizar a presença de quinases, FTs e demais genes que estão envolvidos na

resposta de tolerância ao estresse salino. O aumento da expressão de genes responsivos ao sal, FTs e quinases dentro da célula contribuem para a compreensão da função e de possíveis aplicações desses genes para o melhoramento genético (SHAH et al., 2021).

Diferentes componentes de sinalização juntos desempenham um papel importante na regulação da resposta ao estresse abiótico e desempenham uma ação crítica em conferir resistência ao estresse e tolerância às plantas. A integração de dados de RNA-seq e ferramentas de genômica funcional contribuem para entender os diferentes componentes envolvidos na sinalização relacionada ao estresse, por isso o uso de pesquisas abrangentes são necessárias para a seleção de candidatos que possam ser explorados na prática (ZHU, 2016; ZHANG et al., 2020; SHAH et al., 2021).

Os dados apresentados neste estudo fornecem uma base ampla para uma investigação mais aprofundada das vias de sinalização induzidas por estes estresses - salino e hídrico - e novas explorações sobre a identificação de genes regulados por *SibZIP1* serão úteis para esclarecer os mecanismos deste gene frente às condições abióticas. A integração de diferentes dados ômicos, como Mapeamento QTL e/ou abordagens GWAS, também podem ser agregadas. A representação estática da correlação gênica, não captura todas as interações possíveis, pois elas não operam simultaneamente devido à variação espacial e temporal na expressão gênica. Portanto, para descrever completamente a ligação entre genótipo e fenótipo e entender a regulação gênica subjacente é necessária a coordenação de redes em diferentes níveis moleculares (genes, proteínas e metabólitos) que podem ser integrados (SERIN et al., 2016; ZAINAL-ABIDIN et al., 2022).

3.3 Perspectivas quanto a utilização do banco de dados gerado

Obteve-se um banco de dados com níveis de expressão expressos em TPMs, que efetivamente normalizam as diferenças na composição e são considerados mais comparáveis entre amostras de diferentes origens (CONESA et al., 2016). Por isso, uma interpretação de informações biologicamente relevantes, poderá ser explorada e utilizada para montagem de outras redes de co-expressão, para análises de expressão diferencial entre as bibliotecas obtidas, para estudos e anotação funcional de genes de acordo com o tecido ou condição fornecida pela matriz de expressão gerada neste trabalho.

4 CONSIDERAÇÕES FINAIS

A interpretação de dados de genes responsivos a estresses ambientais contribui para o desenvolvimento de abordagens e aplicações futuras desses genes para o melhoramento genético e de demais estudos.

Os resultados obtidos indicam que o gene *SibZIP1* tem forte correlação com genes que são responsivos e conferem tolerância a diferentes estresses abióticos, como estresse salino e hídrico. A partir dessas observações, novas constatações e futuras interpretações podem auxiliar a ampliação do conhecimento a respeito de genes que não são diretamente regulados por *SibZIP1*, mas que apresentam níveis de expressão semelhantes sob as mesmas condições. Dessa forma, este estudo fornece dados úteis para compreensão dos mecanismos deste gene frente às condições abióticas, além da contribuição para a interpretação das informações que poderão ser obtidas com plantas editadas geneticamente neste gene, conforme abordado no capítulo 02.

REFERÊNCIAS

- BECRAFT, P. W. **Receptor kinase signaling in plant development**. Annual Review of Cell and Developmental Biology, 2002. doi: 10.1146/annurev.cellbio.18.012502.083431.
- BOLGER, A. M.; LOHSE, M.; USADEL, B. **Trimmomatic: a flexible trimmer for Illumina sequence data**. Bioinformatics, v. 30, n. 15, p. 2114–2120, 1 ago. 2014. doi: 10.1093/BIOINFORMATICS/BTU170. Acesso em: 6 jan. 2023.
- CONESA, A.; MADRIGAL, P.; TARAZONA, S.; GOMEZ-CABRERO, D.; CERVERA, A.; MCPHERSON, A.; SZCZEŚNIAK, M. W.; GAFFNEY, D. J.; ELO, L. L.; ZHANG, X.; MORTAZAVI, A. **A survey of best practices for RNA-seq data analysis**. Genome BiologyBioMed Central Ltd. 26 jan. 2016. doi: 10.1186/s13059-016-0881-8.
- CONTRERAS-LÓPEZ, O.; MOYANO, T. C.; SOTO, D. C.; GUTIÉRREZ, R. A. **Step-by-step construction of gene co-expression networks from high-throughput Arabidopsis RNA sequencing data**. Em: Methods in Molecular Biology. [s.l.] Humana Press Inc., 2018. 1761 p. 275–301. 2018.
- COSTA-SILVA, J.; DOMINGUES, D.; LOPES, F. M. **RNA-Seq differential expression analysis: An extended review and a software tool**. PLoS ONE. Public Library of Science, 1 dez. 2017. doi: 10.1371/journal.pone.0190152.
- DOHERTY, C. J.; VAN BUSKIRK, H. A.; MYERS, S. J.; THOMASHOW, M. F. **Roles for Arabidopsis CAMTA Transcription Factors in Cold-Regulated Gene Expression and Freezing Tolerance**. W OA. [s.d.] doi: 10.1105/tpc.108.063958. Acesso em: 15 nov. 2022.
- GAO, Y. F.; LIU, J. K.; YANG, F. M.; ZHANG, G. Y.; WANG, D.; ZHANG, L.; OU, Y. bin; YAO, Y. A. **The WRKY transcription factor WRKY8 promotes resistance to pathogen infection and mediates drought and salt stress tolerance in Solanum lycopersicum**. Physiologia Plantarum, v. 168, n. 1, p. 98–117, 1 jan. 2020. doi: 10.1111/ppl.12978.
- GHELIS, T. **Signal processing by protein tyrosine phosphorylation in plants**. Plant Signaling and Behavior. Landes Bioscience, 2011. doi: 10.4161/psb.6.7.15261.
- HOHMANN, U.; LAU, K.; HOTHORN, M. **The Structural Basis of Ligand Perception and Signal Activation by Receptor Kinases**. Annu. Rev. Plant Biol, v. 68, p. 109–137, 2017. doi: 10.1146/annurev-arplant-042916.
- HONG, J. C. **General Aspects of Plant Transcription Factor Families**. Em: Plant Transcription Factors: Evolutionary, Structural and Functional Aspects. [s.l.] Elsevier Inc., 2016. p. 35–56. 2016.
- KOHL, M.; WIESE, S.; WARSCHEID, B. **Cytoscape: software for visualization and analysis of biological networks**. Methods in molecular biology (Clifton, N.J.), v. 696, p. 291–303, 2011. doi: 10.1007/978-1-60761-987-1_18.
- LI, D.; FU, F.; ZHANG, H.; SONG, F. **Genome-wide systematic characterization of the bZIP transcriptional factor family in tomato (Solanum lycopersicum L.)**. BMC Genomics, v. 16, n. 1, 12 out. 2015. doi: 10.1186/s12864-015-1990-6.

- LIN, Z.; ALEXANDER, L.; HACKETT, R.; GRIERSON, D. **LeCTR2, a CTR1-like protein kinase from tomato, plays a role in ethylene signalling, development and defence.** *Plant Journal*, v. 54, n. 6, p. 1083–1093, jun. 2008. doi: 10.1111/j.1365-313X.2008.03481.x.
- LIU, Y.; HUANG, W.; XIAN, Z.; HU, N.; LIN, D.; REN, H.; CHEN, J.; SU, D.; LI, Z. **Overexpression of SLGRAS40 in tomato enhances tolerance to abiotic stresses and influences auxin and gibberellin signaling.** *Frontiers in Plant Science*, v. 8, 26 set. 2017. doi: 10.3389/fpls.2017.01659.
- PATRO, R.; DUGGAL, G.; LOVE, M. I.; IRIZARRY, R. A.; KINGSFORD, C. **Salmon provides fast and bias-aware quantification of transcript expression.** *Nature Methods*, v. 14, n. 4, p. 417–419, 2017. doi: 10.1038/nmeth.4197.
- SCHMID, M. W. **RNA-seq data analysis protocol: Combining in-house and publicly available data.** *Em: Methods in Molecular Biology*. [s.l.] Humana Press Inc., 2017. 1669 p. 309–335. 2017.
- SCHWECHHEIMER, C.; ZOURELIDOU, M.; BEVAN, M. W. **Plant transcription factor studies.** *Annu. Rev. Plant Physiol. Plant Mol. Biol.* [s.l.: s.n.].
- SERIN, E. A. R.; NIJVEEN, H.; HILHORST, H. W. M.; LIGTERINK, W. **Learning from co-expression networks: Possibilities and challenges.** *Frontiers in Plant Science*. *Frontiers Media S.A.*, 8 abr. 2016. doi: 10.3389/fpls.2016.00444.
- SHAH, W. H.; RASOOL, A.; SALEEM, S.; MUSHTAQ, N. U.; TAHIR, I.; HAKEEM, K. R.; REHMAN, R. U. **Understanding the Integrated Pathways and Mechanisms of Transporters, Protein Kinases, and Transcription Factors in Plants under Salt Stress.** *International Journal of Genomics*. Hindawi Limited, 2021. doi: 10.1155/2021/5578727.
- SONG, L.; HUANG, S. S. C.; WISE, A.; CASTANOZ, R.; NERY, J. R.; CHEN, H.; WATANABE, M.; THOMAS, J.; BAR-JOSEPH, Z.; ECKER, J. R. **A transcription factor hierarchy defines an environmental stress response network.** *Science*, v. 354, n. 6312, 4 nov. 2016. doi: 10.1126/science.aag1550.
- TANG, D.; CHRISTIANSEN, K. M.; INNES, R. **Regulation of plant disease resistance, stress responses, cell death, and ethylene signaling in Arabidopsis by the EDR1 protein kinase.** *Plant Physiology*, v. 138, n. 2, p. 1018–1026, 2005. doi: 10.1104/pp.105.060400.
- TRAPNELL, C.; HENDRICKSON, D. G.; SAUVAGEAU, M.; GOFF, L.; RINN, J. L.; PACHTER, L. **Differential analysis of gene regulation at transcript resolution with RNA-seq.** *Nature Biotechnology*, v. 31, n. 1, p. 46–53, jan. 2013. doi: 10.1038/nbt.2450.
- WASEEM, M.; NKURIKIYIMFURA, O.; NIYITANGA, S.; JAKADA, B. H.; SHAHEEN, I.; ASLAM, M. M. **GRAS transcription factors emerging regulator in plants growth, development, and multiple stresses.** *Molecular Biology Reports*. Springer Science and Business Media B.V., 1 out. 2022. doi: 10.1007/s11033-022-07425-x.
- WINGETT, S. W.; ANDREWS, S. **Fastq screen: A tool for multi-genome mapping and quality control.** *F1000Research*, v. 7, 2018. doi: 10.12688/f1000research.15931.1.
- ZAINAL-ABIDIN, R. A.; HARUN, S.; VENGATHARAJULOO, V.; TAMIZI, A. A.; SAMSULRIZAL, N. H. **Gene Co-Expression Network Tools and Databases for Crop Improvement Plants.** *MDPI*, 1 jul. 2022. doi: 10.3390/plants11131625.

ZHANG, H.; ZHANG, F.; YU, Y.; FENG, L.; JIA, J.; LIU, B.; LI, B.; GUO, H.; ZHAI, J. **A Comprehensive Online Database for Exploring ~20,000 Public Arabidopsis RNA-Seq Libraries** Molecular Plant Cell Press, 7 set. 2020. doi: 10.1016/j.molp.2020.08.001.

ZHU, J. K. **Abiotic Stress Signaling and Responses in Plants**. Cell, v. 167, n. 2, p. 313–324, 6 out. 2016. doi: 10.1016/J.CELL.2016.08.029. Acesso em: 8 jan. 2023.

ZHU, M.; MENG, X.; CAI, J.; LI, G.; DONG, T.; LI, Z. **Basic leucine zipper transcription factor *SibZIP1* mediates salt and drought stress tolerance in tomato**. BMC Plant Biology, v. 18, n. 1, 8 maio 2018. doi: 10.1186/s12870-018-1299-0.

APÊNDICES

**APÊNDICE A - Dados obtidos de bibliotecas selecionadas pelo banco de dados NCBI
(Continua).**

BioProject	Run	Instrument (Illumina)	Bases	Bytes	Tissue	DOI
PRJNA376115	SRR5274866	Genome Analyzer IIx	3,73E+09	2,29E+09	Meristem	https://doi.org/10.1016/j.cell.2017.04.032
	SRR5274867	Genome Analyzer IIx	2,94E+09	1,71E+09	Meristem	
	SRR5274868	Genome Analyzer IIx	3,21E+09	1,92E+09	Meristem	
	SRR5274869	Genome Analyzer IIx	2,11E+09	1,03E+09	Meristem	
	SRR5274870	Genome Analyzer IIx	1,98E+09	9,34E+08	Meristem	
	SRR5274871	Genome Analyzer IIx	1,86E+09	9,91E+08	Meristem	
	SRR5274872	Genome Analyzer IIx	2,22E+09	1,23E+09	Meristem	
	SRR5274873	Genome Analyzer IIx	2,27E+09	1,3E+09	Meristem	
	SRR5274874	Genome Analyzer IIx	1,19E+09	6,4E+08	Meristem	
	SRR5274875	Genome Analyzer IIx	1,32E+09	7,33E+08	Meristem	
	SRR5274887	HiSeq 2500	4,74E+09	3,07E+09	Meristem	
	SRR5274888	HiSeq 2500	6,37E+09	4,13E+09	Meristem	
	SRR5274889	HiSeq 2500	5,25E+09	3,4E+09	Meristem	
	SRR5274890	HiSeq 2500	5,86E+09	3,83E+09	Meristem	
	SRR5274891	HiSeq 2500	7,04E+09	4,6E+09	Meristem	
	SRR5274892	HiSeq 2500	5,35E+09	3,49E+09	Meristem	
	SRR5274893	HiSeq 2500	6,34E+09	4,13E+09	Meristem	
SRR5274894	HiSeq 2500	3,95E+09	2,57E+09	Meristem		
PRJNA744374	SRR15091988	NovaSeq 6000	6,49E+09	2,06E+09	Fruit	https://doi.org/10.3389/fgene.2021.714942
	SRR15091996	NovaSeq 6000	6,17E+09	1,95E+09	Fruit	
	SRR15091998	NovaSeq 6000	7,44E+09	2,36E+09	Fruit	
	SRR15092001	NovaSeq 6000	6,39E+09	2,03E+09	Fruit	
	SRR15092002	NovaSeq 6000	6,62E+09	2,11E+09	Fruit	
	SRR15091997	NovaSeq 6000	6,63E+09	2,13E+09	Fruit	
	SRR15091999	NovaSeq 6000	7,41E+09	2,34E+09	Fruit	
	SRR15092000	NovaSeq 6000	6,78E+09	2,15E+09	Fruit	
	SRR15092003	NovaSeq 6000	6,83E+09	2,16E+09	Fruit	

Dados obtidos de bibliotecas selecionadas pelo banco de dados NCBI (Continua).

BioProject	Run	Instrument (Illumina)	Bases	Bytes	Tissue	DOI
PRJNA644865	SRR12172675	HiSeq 3000	8,56E+09	3,13E+09	Leaf	https://doi.org/10.3390/genes11080900
	SRR12172676	HiSeq 3000	1,30E+10	4,66E+09	Fruit	
	SRR12172677	HiSeq 3000	1,21E+10	4,43E+09	Leaf	
	SRR12172678	HiSeq 3000	1,11E+10	4,01E+09	Fruit	
	SRR12172679	HiSeq 3000	1,10E+10	3,84E+09	Leaf	
	SRR12172680	HiSeq 3000	1,34E+10	4,79E+09	Fruit	
	SRR12172681	HiSeq 3000	1,87E+10	7,03E+09	Leaf	
	SRR12172682	HiSeq 3000	1,75E+10	6,39E+09	Fruit	
	SRR12172683	HiSeq 3000	9,46E+09	3,56E+09	Leaf	
	SRR12172684	HiSeq 3000	1,89E+10	6,83E+09	Leaf	
	SRR12172685	HiSeq 3000	1,25E+10	4,55E+09	Fruit	
	SRR12172686	HiSeq 3000	1,48E+10	5,38E+09	Leaf	
	SRR12172687	HiSeq 3000	1,38E+10	5,11E+09	Fruit	
	SRR12172688	HiSeq 3000	2,07E+10	7,54E+09	Leaf	
	SRR12172689	HiSeq 3000	1,35E+10	5,45E+09	Fruit	
	SRR12172690	HiSeq 3000	1,00E+10	3,67E+09	Leaf	
	SRR12172691	HiSeq 3000	1,15E+10	4,07E+09	Fruit	
	SRR12172692	HiSeq 3000	9,12E+09	3,42E+09	Leaf	
	SRR12172693	HiSeq 3000	1,12E+10	3,9E+09	Fruit	
	SRR12172694	HiSeq 3000	1,09E+10	4,05E+09	Fruit	
	SRR12172695	HiSeq 3000	8,35E+09	2,91E+09	Leaf	
	SRR12172696	HiSeq 3000	9,78E+09	3,43E+09	Fruit	
	SRR12172697	HiSeq 3000	9,85E+09	3,45E+09	Leaf	
	SRR12172698	HiSeq 3000	1,12E+10	3,9E+09	Fruit	
	SRR12172699	HiSeq 3000	1,18E+10	4,24E+09	Leaf	
	SRR12172700	HiSeq 3000	9,32E+09	3,31E+09	Fruit	
	SRR12172701	HiSeq 3000	6,79E+09	2,5E+09	Leaf	
	SRR12172702	HiSeq 3000	1,54E+10	5,37E+09	Fruit	
	SRR12172703	HiSeq 3000	8,74E+09	3,2E+09	Leaf	
	SRR12172704	HiSeq 3000	1,09E+10	3,95E+09	Fruit	
SRR12172705	HiSeq 3000	9,25E+09	3,39E+09	Leaf		
SRR12172706	HiSeq 3000	1,09E+10	3,93E+09	Fruit		

Dados obtidos de bibliotecas selecionadas pelo banco de dados NCBI (Continua).

BioProject	Run	Instrument (Illumina)	Bases	Bytes	Tissue	DOI
PRJNA746070	SRR15111729	HiSeq 4000	1,33E+10	4,45E+09	Anther	https://doi.org/10.3390/cells10071809
	SRR15111730	HiSeq 4000	1,30E+10	4,37E+09	Anther	
	SRR15111731	HiSeq 4000	9,66E+09	3,12E+09	Anther	
	SRR15111732	HiSeq 4000	8,54E+09	2,82E+09	Anther	
	SRR15111733	HiSeq 4000	1,25E+10	4,2E+09	Anther	
	SRR15111734	HiSeq 4000	1,30E+10	4,34E+09	Anther	
	SRR15111735	HiSeq 4000	1,20E+10	3,99E+09	Anther	
	SRR15111736	HiSeq 4000	1,30E+10	4,31E+09	Anther	
	SRR15111737	HiSeq 4000	8,42E+09	2,72E+09	Anther	
	SRR15111738	HiSeq 4000	8,75E+09	2,82E+09	Anther	
	SRR15111739	HiSeq 4000	7,17E+09	2,34E+09	Anther	
	SRR15111740	HiSeq 4000	7,57E+09	2,45E+09	Anther	
	SRR15111741	HiSeq 4000	1,30E+10	4,33E+09	Anther	
	SRR15111742	HiSeq 4000	1,28E+10	4,24E+09	Anther	
	SRR15111743	HiSeq 4000	1,31E+10	4,38E+09	Anther	
	SRR15111744	HiSeq 4000	1,30E+10	4,28E+09	Anther	
SRR15111745	HiSeq 4000	1,31E+10	4,36E+09	Anther		
SRR15111746	HiSeq 4000	1,34E+10	4,47E+09	Anther		
PRJNA705011	SRR13787016	HiSeq 4000	9,36E+09	2,94E+09	Seed	https://doi.org/10.1186/s12870-021-03405-8
	SRR13787017	HiSeq 4000	8,10E+09	2,51E+09	Seed	
	SRR13787018	HiSeq 4000	7,85E+09	2,38E+09	Seed	
	SRR13787019	HiSeq 4000	7,93E+09	2,49E+09	Seed	
	SRR13787020	HiSeq 4000	7,10E+09	2,21E+09	Seed	
	SRR13787021	HiSeq 4000	7,76E+09	2,49E+09	Seed	
	SRR13787022	HiSeq 4000	9,06E+09	2,81E+09	Seed	
	SRR13787023	HiSeq 4000	8,81E+09	2,75E+09	Seed	
	SRR13787024	HiSeq 4000	7,87E+09	2,48E+09	Seed	
	SRR13787025	HiSeq 4000	8,01E+09	2,51E+09	Seed	
	SRR13787026	HiSeq 4000	9,54E+09	3,01E+09	Seed	
SRR13787027	HiSeq 4000	7,59E+09	2,38E+09	Seed		
PRJNA293380	SRR2175729	HiSeq 2500	4,74E+09	1,58E+09	Ovule	https://doi.org/10.1038/srep23173
	SRR2175730	HiSeq 2500	5,63E+09	2,04E+09	Ovule	
	SRR2175733	HiSeq 2500	5,66E+09	1,89E+09	Ovule	
	SRR2175734	HiSeq 2500	6,32E+09	2,23E+09	Ovule	
	SRR2175737	HiSeq 2500	5,20E+09	1,99E+09	Ovule	
	SRR2175738	HiSeq 2500	5,57E+09	1,99E+09	Ovule	
	SRR2175741	HiSeq 2500	4,92E+09	1,64E+09	Ovule	
	SRR2175742	HiSeq 2500	5,34E+09	1,88E+09	Ovule	

Dados obtidos de bibliotecas selecionadas pelo banco de dados NCBI (Continua).

BioProject	Run	Instrument (Illumina)	Bases	Bytes	Tissue	DOI
PRJNA659891	SRR12560324	HiSeq 3000	6,20E+09	2,23E+09	Stamen	https://doi.org/10.1016/j.dib.2021.106715
	SRR12560325	HiSeq 3000	5,37E+09	2,03E+09	Pistil	
	SRR12560326	HiSeq 3000	5,41E+09	2,05E+09	Pistil	
	SRR12560327	HiSeq 3000	4,39E+09	1,66E+09	Pistil	
	SRR12560328	HiSeq 3000	5,12E+09	1,93E+09	Sepal	
	SRR12560329	HiSeq 3000	5,51E+09	2,09E+09	Sepal	
	SRR12560330	HiSeq 3000	5,89E+09	2,24E+09	Sepal	
	SRR12560331	HiSeq 3000	5,00E+09	1,9E+09	Petal	
	SRR12560332	HiSeq 3000	5,61E+09	2,03E+09	Stamen	
	SRR12560333	HiSeq 3000	5,86E+09	2,1E+09	Stamen	
	SRR12560334	HiSeq 3000	5,61E+09	2,11E+09	Petal	
	SRR12560335	HiSeq 3000	5,19E+09	1,98E+09	Petal	
PRJNA636937	SRR11912794	HiSeq 2500	1,59E+10	6,43E+09	Trichome	https://doi.org/10.1111/tpj.14902
	SRR11912795	HiSeq 2500	1,54E+10	6,24E+09	Trichome	
	SRR11912796	HiSeq 2500	1,49E+10	6,06E+09	Trichome	
	SRR11912797	HiSeq 2500	1,57E+10	6,35E+09	Trichome	
PRJNA309342	SRR3119173	HiSeq 2000	3,99E+09	2,65E+09	Pollen	https://doi.org/10.1111/tpj.14902
	SRR3119174	HiSeq 2000	4,65E+09	3,12E+09	Pollen	
	SRR3119175	HiSeq 2000	5,30E+09	3,55E+09	Style	
	SRR3119176	HiSeq 2000	4,84E+09	3,26E+09	Style	
	SRR3119179	HiSeq 2000	3,29E+09	2,21E+09	Pollen	
	SRR3119180	HiSeq 2000	5,26E+09	3,51E+09	Pollen	
	SRR3119181	HiSeq 2000	4,31E+09	2,91E+09	Style	
	SRR3119182	HiSeq 2000	3,00E+09	2,01E+09	Style	
	SRR3119185	HiSeq 2000	3,99E+09	2,68E+09	Pollen	
	SRR3119186	HiSeq 2000	5,76E+09	3,83E+09	Pollen	
	SRR3119187	HiSeq 2000	3,92E+09	2,63E+09	Style	
	SRR3119188	HiSeq 2000	3,19E+09	2,12E+09	Style	

Dados obtidos de bibliotecas selecionadas pelo banco de dados NCBI (Continua).

BioProject	Run	Instrument (Illumina)	Bases	Bytes	Tissue	DOI
PRJNA493976	SRR7959012	HiSeq 2500	7,59E+09	3,05E+09	Stem	
	SRR7959013	HiSeq 2500	7,90E+09	3,2E+09	Stem	
	SRR7959014	HiSeq 2500	7,04E+09	2,86E+09	Root	
	SRR7959015	HiSeq 2500	7,50E+09	3,04E+09	Stem	
	SRR7959016	HiSeq 2500	8,21E+09	3,35E+09	Root	
	SRR7959017	HiSeq 2500	7,20E+09	2,91E+09	Root	
	SRR7959018	HiSeq 2500	7,32E+09	2,95E+09	Stem	
	SRR7959019	HiSeq 2500	8,35E+09	3,37E+09	Stem	
	SRR7959020	HiSeq 2500	8,41E+09	3,33E+09	Root	
	SRR7959021	HiSeq 2500	7,44E+09	2,93E+09	Root	
	SRR7959022	HiSeq 2500	7,02E+09	2,78E+09	Root	
	SRR7959023	HiSeq 2500	7,73E+09	3,06E+09	Stem	
	SRR7959024	HiSeq 2500	9,28E+09	3,64E+09	Stem	
	SRR7959025	HiSeq 2500	9,21E+09	3,67E+09	Stem	
	SRR7959026	HiSeq 2500	7,52E+09	2,92E+09	Root	
	SRR7959027	HiSeq 2500	7,50E+09	2,99E+09	Root	
	SRR7959028	HiSeq 2500	9,20E+09	3,72E+09	Root	
	SRR7959029	HiSeq 2500	7,45E+09	3,02E+09	Stem	
	SRR7959030	HiSeq 2500	7,96E+09	3,2E+09	Root	
	SRR7959031	HiSeq 2500	7,27E+09	2,93E+09	Root	
	SRR7959032	HiSeq 2500	7,95E+09	3,18E+09	Stem	
	SRR7959033	HiSeq 2500	8,61E+09	3,46E+09	Stem	
	SRR7959034	HiSeq 2500	8,35E+09	3,39E+09	Root	
	SRR7959035	HiSeq 2500	7,83E+09	3,16E+09	Stem	
	SRR7959036	HiSeq 2500	8,03E+09	3,24E+09	Root	
	SRR7959037	HiSeq 2500	7,37E+09	2,99E+09	Root	
	SRR7959038	HiSeq 2500	8,34E+09	3,32E+09	Stem	
	SRR7959039	HiSeq 2500	9,31E+09	3,64E+09	Stem	
	SRR7959040	HiSeq 2500	8,19E+09	3,22E+09	Stem	
	SRR7959041	HiSeq 2500	7,14E+09	2,79E+09	Root	
	SRR7959042	HiSeq 2500	7,07E+09	2,78E+09	Root	
	SRR7959043	HiSeq 2500	7,69E+09	3,03E+09	Root	
	SRR7959044	HiSeq 2500	6,51E+09	2,59E+09	Stem	
	SRR7959045	HiSeq 2500	6,93E+09	2,73E+09	Stem	
SRR7959046	HiSeq 2500	7,45E+09	2,94E+09	Stem		
SRR7959047	HiSeq 2500	6,60E+09	2,61E+09	Root		

<https://doi.org/10.1111/tbj.14175>

Dados obtidos de bibliotecas selecionadas pelo banco de dados NCBI (Conclusão).

BioProject	Run	Instrument (Illumina)	Bases	Bytes	Tissue	DOI
PRJNA641520	SRR12095178	HiSeq 4000	6,52E+09	2,53E+09	Leaf	https://doi.org/10.1007/s11103-021-01194-0
	SRR12095179	HiSeq 4000	6,29E+09	2,41E+09	Leaf	
	SRR12095180	HiSeq 4000	5,00E+09	1,92E+09	Leaf	
	SRR12095181	HiSeq 4000	4,33E+09	1,66E+09	Leaf	
	SRR12095182	HiSeq 4000	4,90E+09	1,9E+09	Leaf	
	SRR12095183	NovaSeq 6000	1,34E+10	4,35E+09	Leaf	
	SRR12095184	NovaSeq 6000	1,27E+10	4,11E+09	Leaf	
	SRR12095185	NovaSeq 6000	1,64E+10	5,27E+09	Leaf	
	SRR12095186	HiSeq 4000	6,09E+09	2,35E+09	Leaf	
	SRR12095187	NovaSeq 6000	1,40E+10	4,51E+09	Leaf	
	SRR12095188	NovaSeq 6000	1,27E+10	4,12E+09	Leaf	
	SRR12095189	NovaSeq 6000	1,47E+10	4,75E+09	Leaf	
	SRR12095190	HiSeq 4000	1,18E+10	4,71E+09	Leaf	
	SRR12095191	HiSeq 4000	1,24E+10	4,95E+09	Leaf	
	SRR12095192	HiSeq 4000	1,16E+10	4,59E+09	Leaf	
	SRR12095193	HiSeq 4000	1,23E+10	4,96E+09	Leaf	
	SRR12095194	HiSeq 4000	6,31E+09	2,46E+09	Leaf	
	SRR12095195	HiSeq 4000	1,21E+10	4,81E+09	Leaf	
	SRR12095196	HiSeq 4000	1,13E+10	4,48E+09	Leaf	
	SRR12095197	HiSeq 4000	3,95E+09	1,52E+09	Leaf	
	SRR12095198	HiSeq 4000	5,57E+09	2,15E+09	Leaf	
	SRR12095199	HiSeq 4000	5,61E+09	2,15E+09	Leaf	
	SRR12095200	HiSeq 4000	5,13E+09	1,98E+09	Leaf	
	SRR12095201	HiSeq 4000	5,82E+09	2,23E+09	Leaf	
	SRR12095202	HiSeq 4000	6,24E+09	2,39E+09	Leaf	
	SRR12095203	HiSeq 4000	6,17E+09	2,4E+09	Leaf	
	SRR12095204	HiSeq 4000	5,01E+09	1,95E+09	Leaf	
	SRR12095205	HiSeq 4000	5,20E+09	2,01E+09	Leaf	
	SRR12095206	HiSeq 4000	4,79E+09	1,87E+09	Leaf	
	SRR12095207	HiSeq 4000	5,38E+09	2,09E+09	Leaf	
SRR12095208	HiSeq 4000	6,00E+09	2,33E+09	Leaf		
SRR12095209	HiSeq 4000	6,37E+09	2,44E+09	Leaf		
SRR12095210	HiSeq 4000	6,07E+09	2,33E+09	Leaf		

Fonte: Do autor (2022)

APÊNDICE B - Tabela descritiva dos 85 transcritos obtidos pela rede de correlação (Continua).

ID – Gene	ID- Transcrito	Classificação	Descrição	Localização subcelular	Sinal
Solyc01g01066 0.2	Solyc01g01066 0.3.1	PK_Tyr_Ser- Thr	Proteína tirosina e serina treonina quinase	Membrana Celular	Sinal de localização nuclear
Solyc01g06553 0.2	Solyc01g06553 0.3.1	COBRA	COBRA-like proteína	Membrana Celular	
Solyc01g06643 0.2	Solyc01g06643 0.3.1	zf-RING_2	Domínio Ring- finger	Membrana Celular	Peptídeo sinal
Solyc01g06810 0.2	Solyc01g06810 0.4.1			Retículo endoplasmático	Peptídeo sinal
Solyc01g07948 0.2	Solyc01g07948 0.3.1	bZIP_1	Fator de transcrição bZIP	Núcleo	Sinal de localização nuclear
Solyc01g09417 0.2	Solyc01g09417 0.3.1	Glyco_trans_1_ 4	Glicosiltransfera ses grupo1	Plastídeo Retículo endoplasmático Complexo de Golgi	
Solyc01g09441 0.2	Solyc01g09441 0.3.1	C2	Domínio C2	Retículo endoplasmático	
Solyc01g09635 0.2	Solyc01g09635 0.3.1	Pkinase	Domínio de proteína quinase	Citoplasma Membrana Celular	Sinal de localização nuclear
Solyc01g09880 0.2	Solyc01g09880 0.4.1	Na_Ca_ex	Proteína trocaadora de sódio cálcio	Retículo endoplasmático	Peptídeo sinal
Solyc01g10114 0.2	Solyc01g10114 0.4.1	NYN	Domínio NYN	Citoplasma	
Solyc01g10424 0.2	Solyc01g10424 0.3.1			Membrana Celular	Peptídeo sinal
Solyc01g10467 0.2	Solyc01g10467 0.3.1	Tub	Família Tub	Citoplasma	
Solyc01g10729 0.2	Solyc01g10729 0.3.1	zf-RING_2	Domínio Ring- finger	Núcleo	
Solyc01g10775 0.2	Solyc01g10775 0.3.1			Citoplasma Complexo de Golgi	
Solyc01g10825 0.2	Solyc01g10825 0.3.1	VID27	Domínio VID27 C-terminal tipo WD40	Citoplasma	Sinal de localização nuclear
Solyc02g02144 0.2	Solyc02g02144 0.4.1	Pkinase	Domínio de proteína quinase	Citoplasma Núcleo	Sinal de localização nuclear Sinal de exportação nuclear
Solyc02g06792 0.2	Solyc02g06792 0.4.1	Hydrolase_4	Serina amino peptidase, S33	Membrana Celular Lisossomo Vacúolo	

Tabela descritiva dos 85 transcritos obtidos pela rede de correlação (Continua).

ID – Gene	ID- Transcrito	Classificação	Descrição	Localização subcelular	Sinal
Solyc02g07136 0.2	Solyc02g07136 0.4.1	2OG-Fell_Oxy	Superfamília de oxigenase 2OG-Fe (II)	Citoplasma	
Solyc02g07178 0.2	Solyc02g07178 0.3.1	MACPF	Domínio MAC Perforin	Citoplasma	
Solyc02g07959 0.2	Solyc02g07959 0.3.1	PK_Tyr_Ser-Thr	Proteína tirosina e serina treonina quinase	Membrana Celular	Peptídeo sinal
Solyc02g08027 0.2	Solyc02g08027 0.3.1	VPS28	Proteína VPS28	Citoplasma	
Solyc02g08243 0.2	Solyc02g08243 0.4.1	Mlo	Família Mlo	Membrana Celular Lisossomo Vacúolo	
Solyc02g08826 0.2	Solyc02g08826 0.3.1			Citoplasma Núcleo	Sinal de localização nuclear
Solyc02g08929 0.2	Solyc02g08929 0.3.1	PK_Tyr_Ser-Thr	Proteína tirosina e serina treonina quinase	Membrana Celular	Peptídeo sinal
Solyc02g08968 0.2	Solyc02g08968 0.4.1	PC-Esterase	Família de Acil-esterase semelhante a GDSL SGNH encontrada em Pmr5 e Cas1p	Complexo de Golgi	Peptídeo sinal
	Solyc02g09050 0.3.1			Núcleo	
Solyc02g09352 0.2	Solyc02g09352 0.3.1	Copine	Copine	Citoplasma Lisossomo Vacúolo	Sinal de localização nuclear
Solyc03g00721 0.2	Solyc03g00721 0.3.1	Pkinase	Domínio de proteína quinase	Membrana Celular	Peptídeo sinal
Solyc03g05177 0.2	Solyc03g05177 0.2.1	Glyco_hidro_28	Família Glicosil hidrolase 28	Lisossomo Vacúolo	Peptídeo sinal
Solyc03g06375 0.2	Solyc03g06375 0.3.1	Exostosin	Família exostosina	Complexo de Golgi	Peptídeo sinal
Solyc03g08009 0.2	Solyc03g08009 0.4.1	NAM	Proteína (No Apical Meristem) sem meristema apical	Núcleo Retículo endoplasmático	Peptídeo sinal
Solyc03g12136 0.1	Solyc03g12136 0.1.1	Snf7	Snf7	Lisossomo Vacúolo	
Solyc04g00549 0.2	Solyc04g00549 0.3.1	CRAL_TRIO_2	Domínio divergente CRAL TRIO	Citoplasma	Sinal de exportação nuclear
Solyc04g00786 0.2	Solyc04g00786 0.3.1	ANTH	Domínio ANTH	Citoplasma	
Solyc04g01201 0.2	Solyc04g01201 0.3.1	NB-ARC	Domínio NB-ARC	Citoplasma	Sinal de localização nuclear

Tabela descritiva dos 85 transcritos obtidos pela rede de correlação (Continua).

ID – Gene	ID- Transcrito	Classificação	Descrição	Localização subcelular	Sinal
Solyc04g05627 0.2	Solyc04g05627 0.4.1	CG-1	Domínio CG-1	Núcleo	Sinal de exportação nuclear
Solyc04g07187 0.1	Solyc04g07187 0.1.1	PK_Tyr_Ser-Thr	Proteína tirosina e serina treonina quinase	Membrana Celular	Peptídeo sinal
Solyc04g07256 0.2	Solyc04g07256 0.3.1	Ist1	Regulador da atividade do VPS4 na via MVB	Lisossomo Vacúolo	
Solyc04g07652 0.1	Solyc04g07652 0.3.1	Cyclin_D1_bind	Domínio de ligação da ciclina D1	Citoplasma	Sinal de localização nuclear
Solyc04g07943 0.2	Solyc04g07943 0.4.1	Xan_ur_permease	Família permease	Membrana Celular	
Solyc04g08277 0.2	Solyc04g08277 0.4.1	RRM_1	Motivo de reconhecimento de RNA (a.k.a RRM, RBD, domínio RNP)	Citoplasma Núcleo	Sinal de localização nuclear
Solyc05g00656 0.1	Solyc05g00656 0.1.1	Glyco_transf_9_2	Família Glicosiltransferases 92	Retículo endoplasmático Complexo de Golgi	Peptídeo sinal
Solyc05g00820 0.2	Solyc05g00820 0.4.1	Cpn60_TCP1	Família chaperonina TCP-1 cpn60	Citoplasma	Sinal de exportação nuclear
Solyc05g00843 0.2	Solyc05g00843 0.4.1	BRO1	Domínio BRO1	Citoplasma Lisossomo Vacúolo	Sinal de exportação nuclear
Solyc05g01014 0.2	Solyc05g01014 0.3.1	Pkinase	Domínio de proteína quinase	Membrana Celular Lisossomo Vacúolo	Peptídeo sinal
Solyc05g01250 0.2	Solyc05g01250 0.3.1	WRKY	Domínio de ligação ao DNA WRKY	Núcleo	Sinal de localização nuclear
Solyc05g05417 0.2	Solyc05g05417 0.4.1	GRAS	Família de domínio GRAS	Núcleo	Sinal de localização nuclear
Solyc06g00508 0.2	Solyc06g00508 0.4.1	Pep3_Vps18	Família Pep3 Vps18 Deep laranja	Lisossomo Vacúolo	
Solyc06g00732 0.2	Solyc06g00732 0.3.1	ThiF	Família ThiF	Núcleo	Sinal de localização nuclear
Solyc06g00964 0.1	Solyc06g00964 0.2.1	bZIP_1	Fator de transcrição bZIP	Núcleo	Sinal de localização nuclear
Solyc06g05103 0.2	Solyc06g05103 0.3.1	PK_Tyr_Ser-Thr	Proteína tirosina e serina treonina quinase	Membrana Celular	Peptídeo sinal

Tabela descritiva dos 85 transcritos obtidos pela rede de correlação (Continua).

ID – Gene	ID- Transcrito	Classificação	Descrição	Localização subcelular	Sinal
Solyc06g05363 0.2	Solyc06g05363 0.3.1	Metallophos	Fosfoesterase do tipo calcineurina	Citoplasma Núcleo	Sinal de exportação nuclear
Solyc06g06119 0.1	Solyc06g06119 0.1.1	VQ	Motivo VQ	Citoplasma Núcleo	Sinal de localização nuclear
Solyc06g06572 0.1	Solyc06g06572 0.1.1	DOMON	Domínio DOMON	Membrana Celular	Peptídeo sinal
Solyc06g07305 0.2	Solyc06g07305 0.2.1	NAM	Proteína (No Apical Meristem) sem meristema apical	Retículo endoplasmático	
Solyc06g08444 0.2	Solyc06g08444 0.3.1	NPL4	Família NPL4	Citoplasma Núcleo	Sinal de localização nuclear
Solyc07g00806 0.1	Solyc07g00806 0.1.1	ANTH	Domínio ANTH	Citoplasma	
Solyc07g03223 0.2	Solyc07g03223 0.3.1	Ist1	Regulador da atividade do VPS4 na via MVB	Membrana Celular	
Solyc07g04485 0.2	Solyc07g04485 0.3.1	WD40	Domínio WD, repetição G-Beta	Núcleo	Sinal de localização nuclear
Solyc07g04798 0.1	Solyc07g04798 0.3.1	Es2	Proteína nuclear Es2	Núcleo	Sinal de localização nuclear
Solyc07g05317 0.2	Solyc07g05317 0.4.1	Pkinase	Domínio de proteína quinase	Citoplasma	Sinal de direcionamento peroxissômico
Solyc07g06195 0.2	Solyc07g06195 0.4.1	C2	Domínio C2	Membrana Celular Retículo endoplasmático	Peptídeo sinal
Solyc07g06294 0.2	Solyc07g06294 0.4.1	Pkinase	Domínio de proteína quinase	Citoplasma	Sinal de exportação nuclear
Solyc08g00736 0.2	Solyc08g00736 0.3.1	SEC16_C	Sec23-domínio de ligação de Sec16	Lisossomo Vacúolo	Sinal de exportação nuclear
Solyc08g05966 0.1	Solyc08g05966 0.3.1	ANTH	Domínio ANTH	Citoplasma	
Solyc08g06777 0.2	Solyc08g06777 0.3.1	zf-RING_2	Domínio Ring-finger	Complexo de Golgi	
Solyc08g07642 0.2	Solyc08g07642 0.3.1	RST	Domínio vegetal RCD1-SRO-TAF4 (RST)	Núcleo	Sinal de localização nuclear
Solyc09g05952 0.2	Solyc09g05952 0.3.1	HSF_DNA-bind	Ligação de DNA do tipo HSF	Núcleo	Sinal de exportação nuclear

Tabela descritiva dos 85 transcritos obtidos pela rede de correlação (Conclusão).

ID – Gene	ID- Transcrito	Classificação	Descrição	Localização subcelular	Sinal
Solyc10g00654 0.2	Solyc10g00654 0.4.1	FH2	Domínio de homologia formin 2	Membrana Celular	Peptídeo sinal
Solyc10g00654 0.2	Solyc10g01208 0.4.1			Lisossomo Vacúolo	Peptídeo sinal
Solyc10g04542 0.1	Solyc10g04542 0.3.1	O-FucT	O-FucT	Retículo endoplasmático Complexo de Golgi	Peptídeo sinal Domínio transmembrana
Solyc10g05006 0.1	Solyc10g05006 0.2.1	C2	Domínio C2	Citoplasma Núcleo	Sinal de localização nuclear
Solyc10g07913 0.1	Solyc10g07913 0.2.1	Pkinase	Domínio de proteína quinase	Membrana Celular	
Solyc10g08361 0.1	Solyc10g08361 0.2.1	EDR1	Proteína quinase Le-CTR1 responsiva ao etileno	Citoplasma	Sinal de exportação nuclear
Solyc10g08571 0.1	Solyc10g08571 0.3.1	MACPF	Domínio MAC	Citoplasma	
Solyc11g00600 0.1	Solyc11g00600 0.3.1	Pkinase	Domínio de proteína quinase	Citoplasma Núcleo	Sinal de exportação nuclear
Solyc11g01112 0.1	Solyc11g01112 0.2.1	Lipase_3	Lipase (classe 3)	Membrana Celular Lisossomo Vacúolo Complexo de Golgi	
Solyc11g03965 0.1	Solyc11g03965 0.2.1	Dynamamin_N	Família Dinamina	Citoplasma	Sinal de localização nuclear
Solyc11g04513 0.1	Solyc11g04513 0.2.1	CAF1	Família Ribonucleases CAF1	Citoplasma Núcleo	Sinal de localização nuclear
Solyc11g06489 0.1	Solyc11g06489 0.2.1	PK_Tyr_Ser-Thr	Proteína tirosina e serina treonina quinase	Membrana Celular	
Solyc11g07288 0.1	Solyc11g07288 0.2.1	Cation_ATPase_C	ATPase transportadora de cátions, C-terminal	Lisossomo Vacúolo	Peptídeo sinal
Solyc12g00642 0.1	Solyc12g00642 0.2.1			Citoplasma Núcleo	Sinal de exportação nuclear
Solyc12g04488 0.1	Solyc12g04488 0.2.1	Glyco_hidro_35	Família Glicosil hidrolase 35	Extracelular Lisossomo Vacúolo	Peptídeo sinal
Solyc12g05575 0.1	Solyc12g05575 0.3.1	Na_Ca_ex	Proteína trocadora de sódio cálcio	Membrana Celular Lisossomo Vacúolo	

APÊNDICE C - Tabela de fatores de transcrição

ID	Family	Best hit in Arabidopsis	E-value	Description
Solyc01g079480.3. 1	bZIP	AT3G62420.1	9,00E-48	basic region/leucine zipper motif 53
Solyc03g080090.4. 1	NAC	AT5G24590.2	3,00E-82	TCV-interacting protein
Solyc04g056270.4. 1	CAMT A	AT2G22300.2	0.0	signal responsive 1
Solyc05g012500.3. 1	WRKY	AT1G69310.2	2,00E-72	WRKY DNA-binding protein 57
Solyc05g054170.4. 1	GRAS	AT4G17230.1	1,00E- 144	SCARECROW-like 13
Solyc06g009640.2. 1	bZIP	AT3G62420.1	9,00E-45	basic region/leucine zipper motif 53
Solyc06g073050.2. 1	NAC	AT3G49530.1	3,00E-84	NAC domain containing protein 62
Solyc09g059520.3. 1	HSF	AT1G67970.1	8,00E-80	heat shock transcription factor A8

Fonte: Do autor (2022).

