

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE LORENA

THAÍS CASTILHO DE ARRUDA FALCÃO

Prospecção *in silico* e análise do perfil de expressão gênica da família SQUAMOSA
Promoter-Binding Protein-Like (SPL) do gênero *Saccharum*

Lorena

2023

THAÍS CASTILHO DE ARRUDA FALCÃO

Prospecção *in silico* e análise do perfil de expressão gênica da família SQUAMOSA
Promoter-Binding Protein-Like (SPL) do gênero *Saccharum*

Dissertação apresentada à Escola de Engenharia de Lorena da Universidade de São Paulo para obtenção do título de Mestre em Ciências do Programa de Pós-Graduação em Biotecnologia Industrial na área de concentração de Biotecnologia Industrial.

Orientador: Prof. Dr. Elisson Romanel

Versão Corrigida

Lorena

2023

NÃO AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, SERÁ DISPONIBILIZADO AUTOMATICAMENTE APÓS 2 ANOS DA PUBLICAÇÃO

Ficha catalográfica elaborada pelo Sistema Automatizado
da Escola de Engenharia de Lorena,
com os dados fornecidos pelo(a) autor(a)

Falcão, Thaís Castilho de Arruda
Prospecção in silico e análise do perfil de
expressão gênica da família SQUAMOSA Promoter-Binding
Protein-Like (SPL) do gênero Saccharum / Thaís
Castilho de Arruda Falcão; orientador Elisson
Antonio da Costa Romanel - Versão Corrigida. -
Lorena, 2023.
131 p.

Dissertação (Mestrado em Ciências - Programa de Pós
Graduação em Biotecnologia Industrial na Área de
Biotecnologia Industrial) - Escola de Engenharia de
Lorena da Universidade de São Paulo. 2023

1. Cana-de-açúcar. 2. Biomassa. 3. Genômica. I.
Título. II. Romanel, Elisson Antonio da Costa, orient.

AGRADECIMENTOS

À minha família por todo apoio durante meus anos de formação, em especial aos meus pais que sempre foram meus maiores incentivadores em busca dos meus sonhos. Ao meu marido Felipe que está comigo desde o primeiro ano da faculdade e me incentiva, me dá forças e me acompanhou em cada passo dessa trajetória, me apoiando durante o desafio que foi realizar o mestrado durante a pandemia.

Ao meu orientador prof. Elisson Romanel que compartilha seus valiosos conhecimentos comigo desde a iniciação científica, auxiliando a cada passo e comemorando junto a cada resultado. Que juntamente com a prof(a) Tatiane Franca, transmitem para todos os alunos do Laboratório de Genômica de Plantas e Bioenergia (PGEMBL) suas dedicações à ciência.

À Universidade de São Paulo (USP), em especial a minha querida Escola de Engenharia de Lorena (EEL), onde eu tenho orgulho de fazer parte desde 2014. Ao Departamento de Biotecnologia (DEBIQ) e seus professores pela disponibilidade dos laboratórios e auxílios necessários. À toda equipe que esteve por trás para fazer o programa de pós-graduação em Biotecnologia Industrial não parar durante a pandemia, aos professores que se adaptaram em milhões para a realização das disciplinas online.

Agradeço também as agências de fomento que tornaram possível a realização desse projeto: FAPESP (2016/24391-7), CAPES (88887.479021/2020-00) e CNPq (444912/2014-2).

RESUMO

FALCÃO, T. **Prospecção *in silico* e análise do perfil de expressão gênica da família SQUAMOSA Promoter-Binding Protein-Like (SPL) do gênero *Saccharum***. 2023. 131 p. Dissertação (Mestrado em Ciências) – Escola de Engenharia de Lorena, Universidade de São Paulo, Lorena, 2023.

A cana-de-açúcar (*Saccharum* spp.) é uma gramínea típica de climas tropicais e subtropicais e que possui variadas aplicações biotecnológicas, abrangendo as indústrias de biorrefinarias, biocombustíveis e alimentícias. Visando o aumento da produtividade da planta, é interessante que haja o acúmulo máximo de biomassa sem que haja a necessidade de aumento da área cultivada. A etapa de florescimento da cana-de-açúcar é um fenômeno biológico de alta importância, sendo indesejada em cultivos comerciais devido à perda de produtividade. O atraso do florescimento e prolongação do período vegetativo podem ocasionar maior acúmulo de biomassa por planta. Baseado nos estudos de genômica funcional de distintas espécies vegetais, sabe-se que os genes *SPL*, uma família de fatores transcricionais com o domínio SBP-box, estão envolvidos no controle de biomassa vegetal, sendo alguns membros de sua família alvos de miR156. Nas fases iniciais de desenvolvimento vegetal, têm-se uma alta expressão de miRNA156, ocasionando uma supressão dos seus genes alvo e reprimindo o florescimento. À medida que a planta cresce, têm-se um decréscimo gradual de expressão do miRNA156 e um consequente aumento da expressão dos genes *SPL*. Do ponto de vista econômico e agroindustrial, é interessante que haja esse controle de mudança de fase de desenvolvimento. Com a utilização de bancos genômicos e transcriptômicos e ferramentas de bioinformática em larga escala, como *BLAST*, *HMMER*, *PFAM*, *MUSCLE* e análises filogenéticas, foram identificados, anotados e classificados 18 genes *SPLs* presentes no gênero *Saccharum* (ex. cana-de-açúcar). Dentre esses genes, 11 genes apresentam o sítio de ligação alvo do miRNA156, responsável pela regulação entre as fases vegetativas e reprodutivas. A análise da expressão de bibliotecas transcriptômicas em distintas gramíneas, bem como em *S. spontaneum* e cana-de-açúcar híbridas nas variedades SP83-2847 e IACSP96-7569, usando as bibliotecas sobre ciclo circadiano, seções foliares e distintas fases do desenvolvimento vegetal, permitiu identificar *ScSPLs* possivelmente relacionados com a produção de biomassa vegetal, como *ScSPL3*, *ScSPL4/11*, *ScSPL6*, *ScSPL9*, *ScSPL13*, *ScSPL14*, *ScSPL15* e *ScSPL17*. Também foram identificados lncRNAs de cana-de-açúcar possivelmente envolvidos na rede regulatória de florescimento envolvendo os *ScSPLs*, ilustrando possibilidades de novos estudos na área.

Palavras-chave: Cana-de-açúcar. Biomassa. Genômica.

ABSTRACT

FALCÃO, T. ***In silico* prospecting and analysis of the gene expression profile of the SQUAMOSA Promoter-Binding Protein-Like (SPL) family in the *Saccharum* genus.** 2023. 131 p. Dissertation (Master of Science) – Escola de Engenharia de Lorena, Universidade de São Paulo, Lorena, 2023.

Sugarcane (*Saccharum* spp.) is a grass native to tropical and subtropical climates, with various biotechnological applications in biorefinery, biofuels and food industries. The goal is to increase plant productivity while maximizing biomass accumulation without the need for expanding cultivated areas. However, the flowering stage of sugarcane is a biological phenomenon of significant importance, as it negatively impacts commercial crops, leading to reduced productivity. To improve biomass accumulation per plant, delaying flowering and extending the vegetative period is essential. Functional genomic studies of different plant species have revealed the involvement of SPL genes, a family of transcription factors with SBP-box domains, in plant biomass control, with some family members being targets of miR156. During the initial stages of plant development, there is a high expression of miRNA156, which suppresses its target genes and represses flowering. As the plant matures, miRNA156 expression gradually decreases, leading to increased expression of SPL genes. This regulatory mechanism in the development phase is economically and agro-industrially advantageous. Using genomic and transcriptomic banks and large-scale bioinformatics tools, such as BLAST, HMMER, PFAM, MUSCLE and phylogenetic analyses, 18 SPL genes present in the genus *Saccharum* were identified, annotated and classified. Among these genes, 11 genes presented the miRNA156 target binding site, responsible for regulation between vegetative and reproductive phases. The expression analysis of transcriptomic libraries in different grasses, as well as in *S. spontaneum* and hybrid sugarcane in the varieties SP83-2847 and IACSP96-7569, using the libraries on circadian cycle, leaf sections and different plant development stages, allowed to identify *ScSPLs* possibly related to plant biomass production, such as *ScSPL3*, *ScSPL4/11*, *ScSPL6*, *ScSPL9*, *ScSPL13*, *ScSPL14*, *ScSPL15* and *ScSPL17*. Sugarcane lncRNAs possibly involved in the flowering regulatory network involving *ScSPLs* were also identified, illustrating possibilities for further studies in the area.

Keywords: Sugarcane. Biomass. Genomics.