



ANÁLISES DO PAN-TRANSCRIPTOMA DA CANA-DE-AÇÚCAR

Hellen Regina Silvestre Silva; Felipe Vaz Peres; Jorge Mario Muñoz Perez; Diego Mauricio Riaño-Pachón

Universidade de São Paulo hellen.regina@unesp.br
Universidade de São Paulo diego.riano@cena.usp.br
Universidade de São Paulo jmmunozp@usp.br
Universidade de São Paulo felipe.vzps@usp.br

Resumo: A cana-de-açúcar desempenha um papel essencial e crucial na economia brasileira, destacando-se como um dos principais produtores globais desse cultivo. Sua ampla adoção é atribuída ao elevado teor de sacarose, que possibilita a produção de açúcar e etanol, este último sendo um biocombustível de suma importância, mas também ao fato de que seu bagaço pode ser aproveitado para gerar eletricidade e etanol de 2ª geração, reforçando sua posição como uma fonte versátil de energia renovável. No entanto, o estudo genético da cana-de-açúcar emerge como um desafio dada a existência de lacunas em nossos conhecimentos sobre seus processos fisiológicos, bioquímicos e a complexidade de seu genoma, que inclui sua natureza poliplóide e a presença de múltiplas cópias de genes, grande frequência de sequências repetitivas que podem variar em tamanho, bem como a alta heterozigosidade. Nosso grupo identificou o pan-transcriptoma de cana-de-açúcar a partir de dados de sequenciamento do transcriptoma de 50 variedades de diversas regiões do planeta. Este pan-transcriptoma compreende um total de 5,240,794 transcritos codificadores de proteínas, fornecendo uma visão mais abrangente da expressão genética dessa cultura. A fim de aprofundar nossa compreensão, empregamos recursos computacionais nas análises atuais, com foco em análise evolutiva dN/dS e no enriquecimento de categorias funcionais da Gene Ontology (GO). Para examinar as razões dN/dS dos grupos de genes identificados no pan-transcriptoma, adotamos o pacote PALM. Esse software examina sequências de genes codificadores de proteínas (CDS), calculando a relação entre as mutações que resultam em substituições sinônimas e não sinônimas. Essa análise proporciona insights valiosos sobre a pressão seletiva exercida sobre os genes, identificando padrões adaptativos e evolutivos. Complementando essas análises evolutivas, implementaremos cálculos de métricas

funcionais com o enriquecimento de termos GO, por meio do pacote de ferramentas topGO, o qual possibilita a aplicação do teste exato de Fisher por meio da função “GOFisherTest”. Esse processo visa identificar quais genes se destacam em termos de funções biológicas específicas dentro do conjunto analisado, contribuindo para uma compreensão mais aprofundada da diversidade funcional desse pan-transcriptoma. Ao integrar essas abordagens analíticas, nosso objetivo é compreender a diversidade genética e também discernir quais genes desempenham papéis fundamentais em processos biológicos específicos, proporcionando uma visão abrangente da expressão genética e da adaptação evolutiva dessa cultura. Nesse contexto, este projeto visa contribuir para o estudo da cana-de-açúcar por meio da análise do seu pan-transcriptoma, empregando ferramentas computacionais, o que promoverá uma melhor compreensão das suas características.

Palavras-chave: Pan-transcriptoma; Cana-de-açúcar; Evolução; estrutura gênica; genômica funcional