

## SEQUENCIAMENTO, MONTAGEM E ANOTAÇÃO DO GENOMA NUCLEAR DE *PLINIA EDULIS* (VELL.) SOBRAL

Valdir Marcos Stefenon<sup>1</sup>; Thiago Sanches Ornellas<sup>1</sup>; Liana Bulcão Bittencourt Petrarca<sup>1</sup>; Yohan Fritsche<sup>1</sup>; Suelen Martinez Guterres<sup>1</sup>; Ingrid Lohani Degering Brand<sup>1</sup>; Ana Kelly de Sousa Silva<sup>1</sup>; Fernando Joner<sup>2</sup>; Moeses Andriago Danner<sup>3</sup>; Idemir Citadin<sup>3</sup>

<sup>1</sup>PPG Recursos Genéticos Vegetais, Universidade Federal de Santa Catarina. <sup>2</sup>PPG Agroecossistemas, Universidade Federal de Santa Catarina. <sup>3</sup>PPG Agronomia, Universidade Federal Tecnológica do Paraná. \*E-mail do autor apresentador: valdir.stefenon@ufsc.br

*Plinia edulis* (Vell.) Sobral, conhecida como cambucá, é uma mirtácea muito negligenciada, endêmica do Sul e Sudeste do Brasil, com frutos saborosos e potencial uso na indústria farmacológica. Na medicina popular, infusões de folhas de *P. edulis* são empregadas para tratar problemas estomacais, diabetes, afecções da garganta, diarreia, bronquite, além de ser utilizada como agente antipirético e diurético. É uma espécie ameaçada de extinção, classificada como vulnerável e ameaçada pelas mudanças climáticas globais. O objetivo deste trabalho foi gerar um genoma de referência para *P. edulis*, permitindo a caracterização de genes e vias metabólicas de importância agrônômica, farmacológica e biotecnológica, visando promover a valorização e conservação desta espécie. O DNA de um indivíduo adulto foi sequenciado em uma plataforma Nanopore Mk1B, montando usando Flye e Canu (mesclados com Quickmerge), filtrado com FCS-GX, BLAST e scripts de Python criados pelo grupo. A anotação foi realizada com Braker3 e Augustus. As proteínas anotadas foram identificadas utilizando os bancos de dados KEGG e InterPro. A completude foi verificada usando BUSCO e a linhagem eudicot\_odb10 como referência. O sequenciamento gerou um total de 32,1 Gbases, proporcionando uma cobertura de profundidade de 118,5x. A montagem final do genoma totalizou 278,7 Mbases. A anotação gerou 37.428 modelos gênicos. Esses modelos gênicos foram identificados em diferentes grupos de proteínas pela base de dados KEGG, incluindo 59 rotas do metabolismo de terpenoides e 40 do metabolismo de propanoides e flavonoides. Essas rotas metabólicas estão relacionadas com a produção dos compostos secundários com propriedades medicinais em *P. edulis*. O agrupamento das proteínas pelo InterPro teve maior relação com os bancos PANTHER e MobiDBLite, permitindo a verificação funcional e de desordem proteica intrínseca, respectivamente. Este genoma referência está em fase de avaliação fina das rotas metabólicas e respectivas proteínas e permitirá avanços na utilização e valorização da *P. edulis* como espécie de uso medicinal, promovendo sua conservação pelo uso. Além disso, será útil para o desenvolvimento de programas de domesticação e melhoramento da espécie.

**Palavras-chave:** NGS; Bioinformática; Myrtaceae

**Agradecimentos:** CNPq, CAPES, PPGRGV