

## Mineração de novas espécies virais em dados de sequenciamento em banco de dados públicos: Ora-pronóbis como um caso de estudo

Fernandes. JEA<sup>1</sup>, Nagata. T, Melo. FL<sup>2</sup>

<sup>1</sup>Universidade de Brasília

<sup>2</sup>OnSite Genomics

**Resumo:** Ora-pro-nóbis é a denominação dada a um grupo de espécies cactáceas do gênero *Pereskia*. Ela pode ser classificada como uma planta alimentícia não convencional (PANCs), que se refere a plantas com alto valor nutritivo que não são comumente usadas para alimentação. No Brasil, essa planta é usada tanto na alimentação quanto para uso medicinal. Devido a essas aplicações, essa planta tem sido alvo de estudos genômicos, particularmente, sequenciamento de RNA, mas a diversidade de vírus infectando essas plantas ainda é amplamente desconhecida. Esse estudo objetivou encontrar novas espécies de vírus presentes em dados de sequenciamento genômico de Ora-pro-nobis depositados em bancos de dados públicos. Um pipeline computacional previamente desenvolvido pelo nosso grupo utilizado para minerar possíveis sequências virais foi adaptado para encontrar possíveis espécies virais associadas a 3 espécies de Ora-pro-nobis: *Pereskia aculeata*, *P. grandifolia* e *P. bleu*. Ao todo foram analisados 9 diferentes acessos. Ao final das análises, 7 possíveis genomas virais foram recuperados, 4 em *P. bleu*, 1 em *P. grandifolia* e 2 em *P. aculeata*. Os prováveis genomas foram anotados manualmente na ferramenta Geneious usando busca por similaridade com o programa BLAST. Foi identificado que três deles são isolados de espécies de vírus pertencentes ao gênero Potexvirus (família *Alphaflexiviridae*), especificamente, um isolado de Schlumbergera virus X e dois isolados de Cactus virus X, encontrados em *P. aculeata* e *P. bleu*, respectivamente. Também foi encontrado um isolado do vírus Epiphyllum badnavirus 1 que pertence ao gênero *Badnavirus*, que faz parte da família *Caulimoviridae* em *P. aculeata*. Os três genomas remanescentes foram encontrados em amostras sequenciadas de *P. grandifolia* (um genoma) e *P. bleu* (dois genomas) e possivelmente são novas espécies virais pertencentes ao gênero *Badnavirus*. Nossos resultados indicam o poder que o pipeline desenvolvido tem de detectar genomas virais em dados de sequenciamento depositados em bancos de dados públicos. Os genomas virais pertencem a grupos de vírus bem distintos, os membros da família *Alphaflexiviridae* pertencem à ordem *Tymovirales* e têm o genoma de RNA fita simples no sentido positivo. Enquanto isso, os membros da família *Caulimoviridae* têm seu genoma circular formado por uma fita simples de DNA. A possibilidade de recuperar genomas que já foram descritos anteriormente também mostra a precisão da ferramenta, dando suporte e confiança nos genomas que foram detectados que possivelmente devem ser classificados como novas espécies. Os planos para a continuidade do trabalho envolvem propor a classificação dos genomas encontrados como novas espécies de vírus aplicar o pipeline em novas espécies de PANCs.

**Palavras-chave:** PANCs, diversidade viral, SRA, Ora-pro-nóbis.