

INTRODUÇÃO

Os genes análogos de resistência (RGAs) podem ser definidos como potenciais genes de resistência (R) que contêm domínios conservados específicos e motivos característicos de famílias de gene R conhecidas. Na ausência de sequências de genoma completo, o uso de iniciadores específicos e degenerados para identificação de motivos dentro da família NBS- LRR, como por exemplo o P-loop, kinase 2 e RNBS-D (TAKKEN; ALBRECHT; TAMELING, 2006) já foram descritos.

Até a publicação da primeira sequência do genoma de referência de dendê - *E. guineensis* (SINGH et al., 2013), apenas poucas sequências genéticas NBS-LRR em *Elaeis* spp. estavam publicamente disponíveis, derivadas de PCR (FOAN et al., 2012) ou através da anotação de genes transcritos (LOW et al., 2014).

Considerando o sequenciamento do genoma de *E. guineensis* recém concluído (SINGH et al., 2013), o draft do genoma de *E. oleifera* (Caiaué) gerado pela Embrapa Agroenergia e a escassez de estudos que caracterizaram RGAs no genoma de *Elaeis* spp., objetivou-se nesse estudo realizar a identificação e caracterização de potenciais RGAs NBS-LRR nos genomas destas duas espécies do gênero *Elaeis*.

METODOLOGIA

Para a identificação de RGAs NBS-LRR em *Elaeis* spp., foram utilizados dois conjuntos de dados: (i) sequências RGA NBS-LRR identificadas no genoma de *E. guineensis* e (ii) Sequências RGA NBS-LRR identificadas na montagem parcial do genoma de *E. oleifera*.

Para identificação de genes NBS-LRR no genoma de referência de *E. guineensis* (SINGH et al., 2013), 52 sequências proteicas de genes R NBS-LRR derivadas de espécies de plantas monocotiledôneas e dicotiledôneas foram selecionadas de acordo com Jupe et al. (2012), servindo como um conjunto de treinamento para a identificação de motivos estruturais característicos de proteínas NBS-LRR. Estas sequências foram alinhadas através do programa MAFFT (KATO et al., 2002) versão 6.717b e, posteriormente, filtradas para detecção das regiões aminoacídicas mais conservadas, ou seja, presentes em mais de 50% das sequências, por meio do programa GBLOCKS 0.91b (CASTRESANA, 2000).

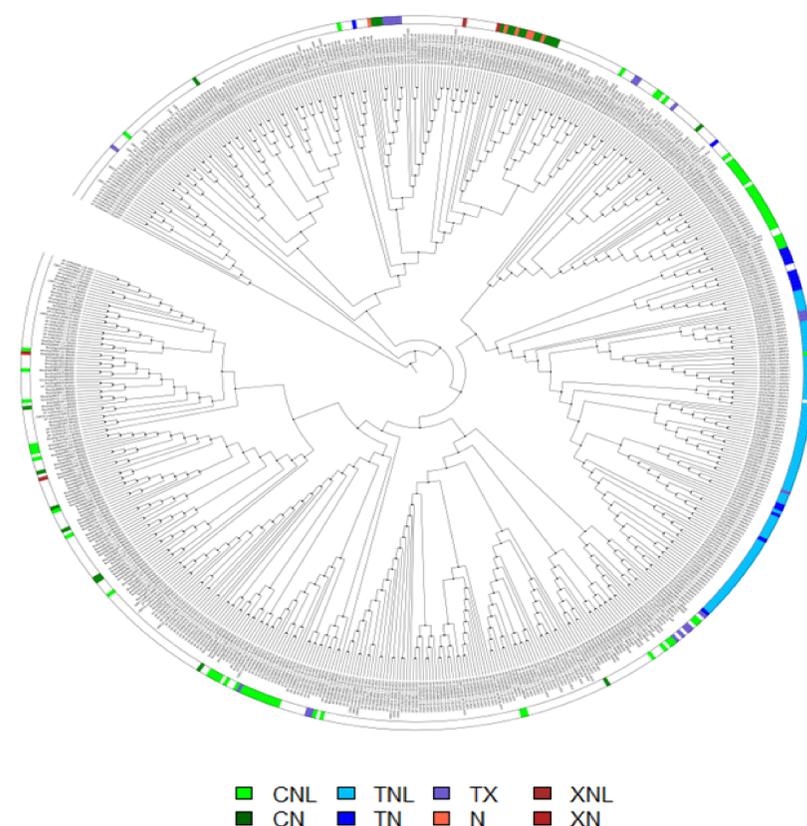
Os alinhamentos filtrados serviram como base para a criação de perfis HMM, com auxílio do programa HMM build do pacote HMMER 3.0 (EDDY, 2008). Estes perfis foram utilizados para detectar genes que codificam sequências com elementos estruturais NBS-LRR no conjunto proteico predito do genomas de *E. guineensis* e de *E. oleifera* por meio do programa HMM Search. As sequências proteicas NBS-LRR encontradas para foram utilizadas para a construção de uma árvore filogenética, onde também se incluiu *Arabidopsis thaliana* e *Brachypodium distachyon*, espécies com NBS-LRR bem caracterizados, uma representando as dicotiledônea, e a outra as monocotiledôneas.

RESULTADOS E CONCLUSÕES

Das 520 sequências encontradas pelo perfil HMM para *B. distachyon*, foi possível utilizar somente 183 sequências e das 199 encontradas em *A. thaliana* foi possível utilizar 188. O número de sequências para cada uma destas espécies supracitadas foi resultante da filtragem realizada a partir de artigos e bancos de dados de sequências NBS-LRR; busca que possibilitou a classificação destas sequências em subgrupos de CNL e TNL e, posteriormente, a ancoragem e o reconhecimento das sequências das sub-famílias NBS-LRR de *Elaeis* spp.

Um conjunto de 627 sequências aminoacídicas NBS-LRR RGAs foi alinhado para a construção de uma árvore filogenética global (Figura 1), o que inclui as sequências filtradas de *A. thaliana* e *B. distachyon*, além das 220 sequências identificadas para *E. guineensis* por meio do perfil aminoácido, e 36 sequências para *E. oleifera* identificadas por meio da análise de ORFs e domínios.

Figura 1 – Análise filogenética entre as sequências NBS-LRR de *Brachypodium distachyon*, *Arabidopsis thaliana* e RGAs NBS-LRR no gênero *Elaeis* spp.



Esta análise demonstrou divergências significativas entre as sequências regulares de CNL em monocotiledônea e dicotiledônea, ou seja, entre as sequências pertencentes a *A. thaliana* e *B. distachyon*. A topologia da árvore evidenciou também um grupo bem definido para TNL, destacado em azul claro na árvore, não sendo encontradas sequências de *Elaeis* spp. dentro deste grupo.

AGRADECIMENTOS

À CAPES-MEC, pela bolsa de mestrado de M. de L. SANTOS; ao Programa de Pós-Graduação em Biotecnologia Vegetal da UFLA; e à FINEP-MCTIC (Contrato 01.13.0315.02, Projeto Estratégias Genômicas e Agregação de Valor para a Cadeia Produtiva do Dendê).