IDENTIFICAÇÃO MOLECULAR DE PLANTAS EM REGIÕES TROPICAIS – DIFICULDADES E DESAFIOS

Vânia C. R. Azevedo
Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF

Peter W. Inglis
Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF
peterwinglis@gmail.com

Fernanda A Gaiotto
Universidade Estadual de Santa Cruz, Ilhéus, BA, gaiotto@uesc.br

RESUMO - O sequenciamento de DNA surge na atualidade como ferramenta potencial na identificação correta de espécies, estudos de evolução e filogenia, quantificação e monitoramento da biodiversidade, análise forense, entre outros. As regiões *rbcL* + *matK* foram definidas pelo Consórcio Internacional de *Barcode* (CBOL) como regiões padrões para *barcode* de plantas. Em 2011 foi criado o BR-BOL com o objetivo de promover de maneira sistematizada a identificação molecular da biodiversidade brasileira. Para identificação de plantas estão sendo analisadas as regiões propostas pelo CBOL e outros autores. Para algumas espécies da família Sapotaceae, entretanto, fica evidenciado que tais regiões não são eficientes na identificação molecular dessas espécies. Outras regiões conservadas, diferentes das propostas vêm sendo testadas e apresentam resultado satisfatório na discriminação molecular de espécies de regiões tropicais, como apresentado aqui para o gênero *Manilkara*.

Palavras-chave: DNA *barcoding*, sequenciamento, identificação de espécies.

INTRODUÇÃO

O rápido avanço nas técnicas de sequenciamento de DNA e o baixo custo na obtenção de sequências têm proporcionado à ciência molecular novas oportunidades e desafios. Isso tem permitido que a cada dia novas perguntas possam ser respondidas e avanços significativos sejam alcançados. O sequenciamento de DNA surge na atualidade como ferramenta potencial na identificação correta de espécies, estudos de evolução e filogenia, quantificação e monitoramento da biodiversidade, análise forense, entre outros. Pode-se dizer, portanto, que as metodologias tradicionaispassaram a contar com ferramenta altamente informativa e moderna na solução das dificuldades de identificação e classificação correta das espécies.

O uso de sequencias de DNA para identificação de espécies (DNA *barcoding*) vem ocorrendo há quase dez anos. Em 2004 foi criado o CBOL (*Consortium for theBarcodeof Life)*, que é atualmente referência mundial no assunto e em 2009 definiram-se como padrões as regiões *rbcL* + *matK*para *barcode* de plantas. Em 2011 foi criado o BR-BOL com o objetivo de realizar a identificação molecular da biodiversidade brasileira. Sendo o Brasil um país megadiverso, com uma estimativa otimista de mais de 600.000 mil espécies ainda por serem descobertas, essa iniciativa surge como uma grande oportunidade de avanço no conhecimento da diversidade de espécies brasileiras.

Entretanto, na realidade, devido auma série de fatores, a identificação molecular de plantas ainda é uma ciência em amplo desenvolvimento e não deve ser considerada como definitiva. Enquanto a identificação molecular de animais é uma realidade, sendo possível para a grande maioria das espécies identificá-las pela sequência de um fragmento do gene codificador da proteína Citocromo oxidase I (COI), para plantas, apesar das recomendações otimistas que se baseiam em duas ou trêsregiões como supostamente universais, diversas regiões devem ser sequenciadas (ex. *RbcL*, *MatK*, *psbA-trnH*, ITS, ETS) e a eficiência, tanto do sequenciamento quanto da resolução na diferenciação molecular das espécies varia de acordo com a família.

Identificação do problema - Considerando o avanço das técnicas de biologia molecular e sequenciamento de DNA, a comunidade científica mundial vive uma era de sequenciamento em larga escala com intuito de resolver uma série de questões científicas. Entre essas questões está a identificação e quantificação da biodiversidade com o intuito de monitorar a diversidade de espécies, conhecê-la melhor e responder de maneira pró-ativa às mudanças climáticas e pressões que os ambientes naturais vêm sofrendo. Assim, foi iniciada a iniciativa BrBOL com o intuito de identificar por marcadores moleculares a biodiversidade brasileira.

Este trabalho tem como objetivo discutir a situação atual da identificação molecular de plantas, testar e validar regiões recomendadas para serem utilizadas como DNA *barcoding* de plantas, testar outras regiões potenciais e subutilizadas para esse tipo de estudo e apresentar resultados de uma pesquisa que indicam e reforçam a dificuldade na definição de DNA *barcoding* para plantas tropicais.

MATERIAL E MÉTODOS

O projeto BrBOL - Tem como objetivo agregar o maior número possível de instituições com pesquisadores interessados na aplicação da técnica de identificação molecular de espécies (DNA *barcode*) para estudo da biodiversidade brasileira. Visa integrar todas as grandes coleções de DNA dos mais diversos organismos, que são o repositório de grande parte das amostras de nossa biodiversidade. O objetivo mais audacioso está relacionado ao estudo de representantes de praticamente toda a biodiversidade brasileira, lançando as bases para sua identificação molecular. Atualmente encontra-se disponível para acesso livre o site: www.brbol.org.

Amostragem - Para cada espécie de planta estuda, utilizar amostras de pelo menos três indivíduos, preferencialmente de diferentes populações. Estão incluídas ações de coleta, curadoria e processamento de amostras biológicas e conservação em Bancos de Germoplasma e/ou Herbários.

Obtenção do Material Genético -O material coletado, proveniente de amostras de câmbio é acondicionado em tubo estéril contendo solução tampão. Outras fontes de DNA incluindo folhas frescas, desidratadas em sílica gel e material de herbário também são utilizadas. DNA é extraído utilizando protocolo de extração de DNA de Doyle e Doyle (1987) ou kit de extração de DNA como DNEasy (Qiagen), a depender da espécie e qualidade do material vegetal.

Banco de DNA -O DNA extraído será conservado, após o estudo, no Banco de DNA da instituição. O armazenamento de DNA total extraído e/ou purificado será realizado em ultrafreezers (-80°C).

Marcador e PCR -No caso específico de plantas, visa-se validar a proposta do CBOL de focar nas sequencias *rbcL* + *MatK*e outros autores (KRESS *et al.* 2005) que propõem também o espaçador *psbA-trn*H para identificação de plantas tropicais, e na impossibilidade dessa validação, testar outras regiões potenciais e propor alternativas para os trabalhos com plantas. Assim, para cada família alvo do estudo, essas e outras regiões específicas do DNA são testadas (**Tabela 1**).

Tabela 1. Lista de *primers* para *barcode* de plantas terrestres.

Primers	Sequências
matK 3F_Kim f	CGTACAGTACTTTTGTGTTTACGAG
matK 1R_Kim r	ACCCAGTCCATCTGGAAATCTTGGTTC
rbcLa_f	ATGTCACCACAAACAGAGACTAAAGC
rbcLa_r	GTAAAATCAAGTCCACCRCG
trnH(GUG)	CGC GCA TGG TGG ATT CAC AAT CC
PsbA	GTT ATG CAT GAA CGT AAT GCT C
ITS17SE	ACGAATTCATGGTCCGGTGAAGTGTTCG
ITS26SE	TAGAATTCCCCGGTTCGCTCGCCGTTAC
ITS92	AAGGTTTCCGTAGGTGAA
ITS75	TATGCTTAAACTCAGCGGG

Táxon estudado – Foram analisadas espécies do gênero *Manilkara* que ocorrem na região da América do Sul. Entre as amostras estão espécies da Região Amazônica (*M. huberi, M. cavalcantei, M. paraenses, M. bidentata)* e da Mata Atlântica (*M. multifida, M. triflora*) e seus grupos externos como a africana *M. butugi* e a panamenha *M. zapota.*

RESULTADOS E DISCUSSÃO

As regiões consideradas universais para identificação molecular de plantas não se mostraram eficientes nesse estudo, sendo que *rbcL* não apresentou polimorfismo e *matK* apresentou duas bases polimórficas em 857 bases alinhadas, diferenciando somente uma das espécies alvo do estudo (*M. multifida*). A região *psbA-trn*H apresentou-se mais variável, entretanto houve dificuldade de amplificação e sequenciamento devido à longa região poli-A. Além disso, grande parte da variação encontrada foi na forma de *indels*. Dentre as 645 bases alinhadas apenas 8 caracteres variáveis informativos pela análise de parcimônia foram identificados.

As regiões informativas e capazes de discriminar as espécies do gênero amostradas foram ITS e a pouco utilizada internacionalmente ETS. Ambas foram facilmente amplificadas e sequenciadas. Entre as 749 bases sequenciadas na região ITS, 54 foram variáveis e informativas (**Figura 1**).

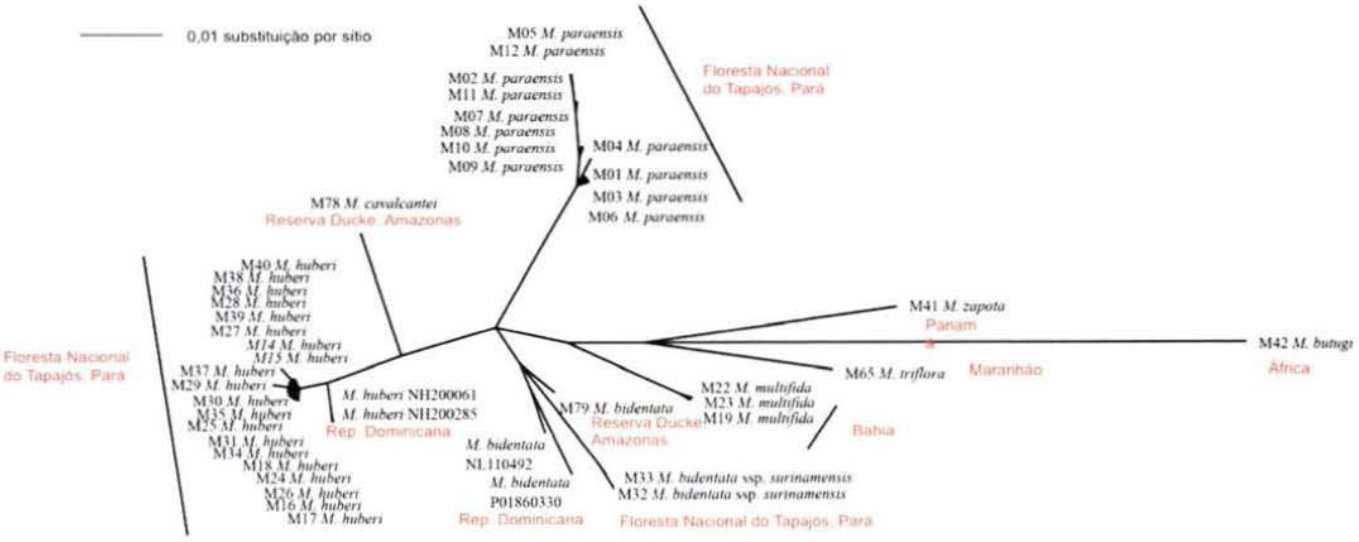

Figura 1. Árvore gerada pela análise Bayesiana com base no marcador ITS.

ETS gerou sequências de 543 bases, das quais 38 foram variáveis e informativas. A vantagem da região ETS está no fato de possuir baixa variação intraespecífica se comparada com ITS.

Em 2007,Kress e Erichson recomendaram as regiões *rbcL* e *trnH-psbA* como regiões universais para DNA *barcoding*de plantas terrestres. Já em 2009, o CBOL afirmou que as regiões *rbcL* + *MatK*são na verdade os marcadores universais. Fica evidenciado, entretanto, neste e em outros trabalhos (Jeanson *et al* 2011; Melo, 2012) que para espécies tropicais, esses marcadores não são eficientes na identificação ao nivel de espécie, sendo fundamental o uso de outros marcadores, conforme já previsto na criação do BR-BOL. Resultados semelhantes vêm sendo encontrados por diferentes grupos de pesquisa no Brasil (comunicação pessoal).

Evidenciou-se com este estudo que de fato novas regiões devem ser testadas e que ETS, apesar de pouco utilizado para este fim, apresenta grande potencial para ser utilizado como DNA *barcoding* de plantas tropicais, conforme também evidenciado por Melo (2012).

É fato que a correta identificação de espécies ainda depende de ferramentas adequadas e mais informativas e o sequenciamento de DNA tem grande potencial de auxiliar nesse processo. Fica evidenciado que muito trabalho ainda deve ser feito no desenvolvimento de metodologia adequada e eficiente. O sequenciamento completo do cloroplasto tem sido considerado importante no desenvolvimento de metodologia eficiente de identificação de espécies de plantas (Dong *et al*, 2012), assim como novas tecnologias de genotipagem em larga escala como GBS (*GenotypebySequencing*) e DArT(*DiversityArrays Technology*) podem vir a ser consideradas, devido à enorme quantidade de informação genética gerada por indivíduo analisado.

REFERÊNCIAS BIBLIOGRÁFICAS

DONG W, LIU J, YU J, WANG L, ZHOU S. **Highly Variable Chloroplast Markers for Evaluating Plant Phylogeny at Low Taxonomic Levels and for DNA Barcoding**. PLoS ONE 7(4): e35071. doi:10.1371/journal.pone.0035071, 2012.

DOYLE, J.; J. DOYLE. **A rapid DNA isolation procedure for small amounts of leaf tissue**. Phytochemical Bulletin,19:810–815, 1987.

JEANSON, M. L.; LABAT, JEAN-NOEL; LITTLE, D. P. **DNA barcoding: a new tool for palm taxonomists?**Annals of Botany, doi:10.1093/aob/mcr158, acessoem 2012).

KRESS, W. JOHN; WURDACK, K. J.; ZIMMER, E. A.; WEIGT, L. A.; JANZEN, D. H. **Use of DNA barcodes to identify flowering plants**. *Proceedings of the National Academy of Sciences*, 102(23):8369–8374, 2005.

KRESS, W. JOHN; ERICKSON, D. L. **A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region**. PLoS ONE, 2(6):e508. doi:10.1371/journal.pone.0000508, 2007.

MELO, C. V. V. **DNA Barcode em espécies arbóreas de Sapotaceae da Mata Atlântica do Sul da Bahia**. (dissertação de mestrado), Universidade Estadual de Santa Cruz, Ilhéus, BA, 2012.