

O gene *rbcL* como barcode para identificação forense de *Cannabis sativa*

The *rbcL* gene as a barcode tool for forensic identification of *Cannabis sativa*

Ribeiro, ASD^{1,2}; Dias, VHG¹; Mello, ICT¹; Silva, R³; Sabino, BD⁴; Garrido, RG⁵; Seldin, L⁶; Moura-Neto, RS^{1,2}

Ribeiro, ASD; Dias, VHG; Mello, ICT; Silva, R; Sabino, BD; Garrido RG; Seldin, L; Moura-Neto RS. O gene *rbcL* como barcode para identificação forense de *Cannabis sativa*. Saúde, Ética & Justiça. 2013;18(Ed. Especial):67-71.

RESUMO: *Cannabis sativa* é uma das espécies mais antigas de plantas domesticadas e permanece como uma das culturas mais amplamente difundida e a droga ilícita mais consumida no mundo. Há uma grande dificuldade em identificar e individualizar as amostras de *Cannabis sp*, dificultando a correlação a prováveis locais de plantações ilegais, o que permitiria revelar rotas de tráfico, relacionar grupos criminosos e distinguir amostras legais daquelas comercializadas como droga, onde o cultivo é permitido. A principal finalidade do DNA Barcode é o de proporcionar uma rápida e precisa identificação de organismos a partir de uma pequena região padronizada do genoma que ajuda a caracterizar e distinguir espécies e indivíduos não identificados para atribuir à espécie. Um dos genes candidatos é o ribulose-1,5-bisfosfato carboxilase/oxigenase (*rbcL*), utilizado como um sistema de DNA Barcode e presente no DNA dos cloroplastos das plantas, sendo o responsável pela produção subunidade maior da que converte dióxido de carbono e água em carboidratos. Nosso objetivo é o desenvolvimento de um protocolo eficiente de extração de DNA e de sequenciamento do gene *rbcL* para análise forense de amostras apreendidas da Polícia Judiciária. DNA de amostras de *Cannabis sativa*, caracterizadas no ICCE/DGPTC/PCERJ, foi extraído utilizando o Mini DNeasy Plant (Qiagen) no IPPGF/DGPTC/PCERJ. O material foi transferido para o LabFor/UFRJ e um fragmento de 735 pb do gene *rbcL* foi amplificado através de um par de iniciadores descritos na literatura como universais para plantas. O fragmento foi sequenciado, com o protocolo do BigDye v3.1, usando ABI 3500 (Life Technologies). As comparações das seqüências foram realizadas no software Geneious (Biomatters). A nossa seqüência consenso de 559 pb foi comparada às seqüências que correspondem às regiões *rbcL* em *Cannabis sativa* depositadas no GenBank, e observamos a ocorrência de polimorfismos (SNPs) entre as amostras dos Estados Unidos, Reino Unido e China, sugerindo que se trata de uma assinatura genética para análise forense.

PALAVRAS-CHAVE: *Cannabis sativa*; Assinatura genética; Identificação forense.

Apoio Financeiro: FAPERJ, CNPq, CAPES, PPGVB/UFRJ.

¹ Laboratório de Biologia Molecular Forense, Instituto de Biologia/UFRJ.

² Instituto Nacional de Metrologia, Qualidade e Tecnologia, InMETRO.

³ Laboratório de Metabolismo Macromolecular Firmino Torres de Castro, Instituto de Biofísica Carlos Chagas Filho/UFRJ.

⁴ Instituto de Criminalística Carlos Éboli /DGPTC/PCERJ.

⁵ Instituto de Pesquisas e Perícias em Genética Forense, /DGPTC/PCERJ.

⁶ Instituto de Microbiologia Prof. Paulo de Góes/UFRJ.

Endereço para correspondência: LabFor, Instituto de Biologia, UFRJ. Av. Carlos Chagas Filho, 373, Bl. A, sala A0-087, Ilha do Fundão. CEP. 21.941-092, Rio de Janeiro, Brasil. e-mail: rodrigomouraneto@ufrj.br

INTRODUÇÃO

C*annabis sativa* foi uma das primeiras espécies de plantas domesticadas e tem sido usada por milênios como uma fonte de fibra, óleo e por suas propriedades medicinais e psicoativas. De seu local de domesticação na Ásia Central, o cultivo de *Cannabis sp.* se espalhou nos tempos antigos por toda a Ásia e Europa e é hoje um dos cultivos mais amplamente distribuídos no mundo¹. Algumas evidências arqueológicas, encontradas em uma tumba de 2.700 anos no noroeste da China e em um túmulo judaico de 1700 anos, sugerem o uso medicinal da *Cannabis sp.*. Atualmente a *Cannabis sp.* e seus derivados como haxixe são as drogas ilícitas mais consumidas no mundo³. A propagação da *Cannabis sp.* pelo mundo ocorre através da forma vegetativa ou das sementes e, na América do Sul o que se percebe é que as sementes são colhidas anteriormente à prensagem e distribuição, sendo então exportadas para novos locais ou até mesmo usadas para dar origem a novas plantas na mesma região, sendo que na América do Sul o maior número de plantações ilegais ocorre no Brasil e Paraguai, existindo também na Bolívia e Colômbia⁴.

Polimorfismos do DNA mitocondrial (mtDNA) e cloroplástico (cpDNA), de origem geralmente uniparental (materna) em angiospermas e com baixas taxas de mutação, são capazes de individualizar haplótipos de *C. sativa*, provendo dados evolutivos e biogeográficos. Gilmore et al.⁵, a partir da análise de 12 loci do DNA mitocondrial (mtDNA) e cloroplástico (cpDNA), observaram 7 loci polimórficos, sendo 5 regiões variáveis de tamanho e 2 polimorfismos de nucleotídeo único (SNP), e que seria possível determinar sete grupos de haplótipos de *Cannabisa* partir dos tipos produtores de droga, fibra e selvagem. Desta maneira, técnicas genéticas estão cada vez mais sendo adotadas para a identificação de espécies botânicas e, dentre os métodos moleculares existentes, encontra-se o sistema chamado de DNA Barcode (o código de barras do DNA), utilizado para o estudo taxonômico da biodiversidade através da utilização de uma região curta e padronizada do DNA^{6,7}. Entretanto para ser prático e satisfatório a região do gene selecionado deve satisfazer três critérios: (i) conter suficiente

variabilidade entre as espécies; (ii) ser curto o suficiente para ser sequenciado em uma única reação; e (iii) conter regiões conservadas para o desenvolvimento de primers universais⁸.

OBJETIVOS

Nossos objetivos neste trabalho foram desenvolver e validar técnicas de identificação genética da *Cannabis sativa* através da análise do DNA Barcode, testando o desempenho do gene de *rbcL* para identificação da espécie, e encontrar haplótipos em *Cannabis sativa*, apreendidas no Estado do Rio de Janeiro.

MATERIAL E MÉTODOS

Doze amostras de *Cannabis sativa*, apreendidas e identificadas quimicamente pela presença de canabinóis, no Instituto de Criminalística Carlos Éboli, Departamento Geral de Polícia Técnico Científica (ICCE/DGPTC), foram cedidas para pesquisa. Dessas amostras, cerca de 2,0g de massa seca foram utilizadas para extração do DNA por meio do kit DNeasy Plant Mini (Qiagen). O material foi inicialmente processado nos laboratórios do Instituto de Pesquisas e Perícia em Genética Forense (IPPGF/DGPTC) e após a solubilização do material genético em solvente orgânico, inativando completamente o $\Delta 9$ THC, as amostras seguiram para o Laboratório de Biologia Molecular Forense, na UFRJ (LabFor/UFRJ).

A amplificação por PCR foi realizada utilizando os primers universais⁹ e o protocolo descrito a seguir: 25 μ l de reação continha 20ng de DNA, 0,1mM de dNTP, Tampão 1X (10 mM Tris-HCL, pH 8.5, 50 mM KCl), 2,5 mM de MgCl₂ e 2,5 U de Taq Polimerase (Ludwig Biotec). A ciclagem foi realizada no GeneAmp® PCR System 9700 (Applied Biosystems): 94°C/10min, 35 ciclos de 94°C/45s, 58°C/1min e 72°C/1min, seguido de extensão final de 72°C/10min. Uma alíquota dos produtos de PCR foram submetidos à eletroforese em gel de agarose a 1,5% e detectados por GelRed. Após confirmação da amplificação, os produtos foram purificados com a membrana AMICON® Ultra-0.5 30K (Millipore Corporation) seguindo o protocolo do fabricante.

O sequenciamento foi realizado utilizando o BigDye Terminator v 3.1 (Life Technologies) com aproximadamente 10ng do produto de PCR.

As sequências foram detectadas no ABI 3500, usando Data Collection e Sequencing Analysis (Applied Biosystems). As análises das sequências e haplótipos obtidos foram realizadas no Software Geneious (BioMatters) e as árvores filogenéticas geradas pelo Mega 5.1.

RESULTADOS

Os sequenciamentos obtidos de cada uma das 12 amostras foram alinhados a uma sequência consenso foi extraída com um tamanho de 559pb (*Cannabis sativa* Rio de Janeiro).

Após a obtenção da sequência consenso foi realizado um BLAST e no resultado observaram-se cinco sequências de *Cannabis sativa*: duas pertencentes ao Reino Unido (Access#AJ390068 e AJ402933), duas da China (Access#JN040400 e JN040401) e uma dos EUA (Access#AF500344). O alinhamento realizado entre as sequências revelou duas diferenças (Tabela 1) no gene do *rbcL* da *C. sativa*. A notação dada foi relativa ao sequenciamento completo deste gene em *Morus indica* (Access#DQ226511.1), pertencente à família Moraceae e filogeneticamente mais próxima a Cannabaceae¹⁰. Os SNP's encontrados foram: 58233(G/A) e 58493 (T/C). As amostras do Rio de Janeiro possuíam o haplótipo GC, enquanto que as amostras do Reino Unido e Estados Unidos o haplótipo GT e, a China com o haplótipo AC.

TABELA 1. Diferenças encontradas nas sequências do gene de *rbcL* de *C. sativa*, das amostras do Rio de Janeiro, Reino Unido, Estados Unidos e China

| Sequência | Posição no Gene | |
|----------------|-----------------|------------|
| | 58233(G/A) | 58493(T/C) |
| Rio de Janeiro | G | C |
| Reino Unido | | G T |
| Estados Unidos | G | T |
| China | A | C |

Em relação à cadeia de aminoácidos encontramos apenas uma única diferença. Na posição 19411 existe a variação de aminoácidos Valina (V) e Isoleucina (I). Para as amostras do Rio de Janeiro, Reino Unido e Estados Unidos da

América encontramos a Valina, e para a China, a Isoleucina (Tabela 2).

TABELA 2. Diferenças encontradas na sequência de aminoácidos entre as amostras do Rio de Janeiro, Reino Unido, Estados Unidos e da China

| Sequência | Posição na Cadeia Polipeptídica 19411(V/I) |
|----------------|---|
| Rio de Janeiro | V |
| Reino Unido | V |
| Estados Unidos | V |
| China | I |

Finalizando as nossas análises comparamos as sequências obtidas com todos os membros da família Cannabaceae, disponíveis no GenBank. A Figura 1 apresenta a árvore UPGMA, com bootstrap de 1000 réplicas, em que observamos o gênero *Cannabis* como monofilético próximo ao gênero *Humulus* corroborando as análises com dados morfológicos¹⁰.

DISCUSSÃO E CONCLUSÃO

As doze amostras de *C. sativa*, apreendidas no Rio de Janeiro, mostraram-se idênticas após a análise da sequência do gene de *rbcL*, contido no fragmento amplificado de 735pb. Possivelmente ao fato de todas serem originárias de cultivares geneticamente próximos ou idênticos.

Em comparação com outras amostras do gene de *rbcL* de *C. sativa*, depositadas no GenBank, observamos três haplótipos distintos. Um diferenciando as amostras do Rio de Janeiro, outro da China e, um terceiro comum entre os Estados Unidos e Reino Unido.

Ainda de acordo com nossos resultados, essas diferenças nos haplótipos encontrados levam a uma única mudança de aminoácido. A Valina foi encontrada na posição 19411 nas amostras do Rio de Janeiro, Estados Unidos e Reino Unido, enquanto que a Isoleucina foi encontrada na amostra da China. Dados preliminares sugerem que essa mutação tenha ocorrido após a expansão da *Cannabis* a partir da Ásia (dado não mostrado).

Finalmente a ocorrência de polimorfismos (SNP) encontrados no gene de *rbcL*, entre as amostras do Rio de Janeiro, Estados Unidos, Reino Unido e China, pode ser usado como marcador biogeográfico, sugerindo que se trata de uma assinatura genética para análise forense.

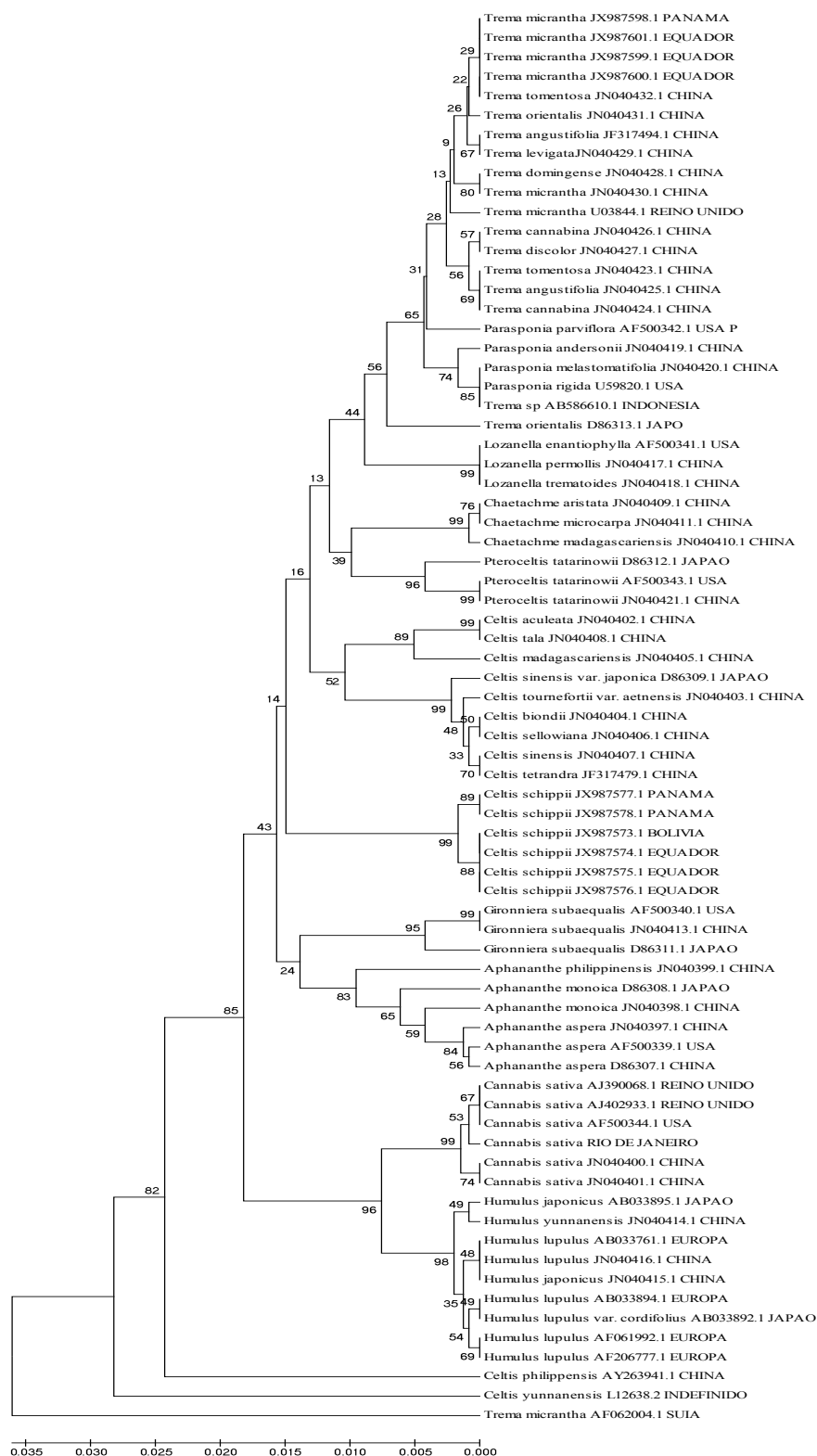


Figura 1. Árvore filogenética da família Cannabaceae, usando as diferenças par-a-par, com 1000 réplicas de bootstrap (números nos cladogramas). As sequências foram alinhadas e trimadas em Mega v5.1. A amostra do Rio de Janeiro é mostrada em destaque

Agradecimentos: Ao apoio técnico excepcional de Priscila Menezes, do Laboratório de Biologia Molecular Forense do IB/UFRJ, e de Cesar Schmidt, do Laboratório de Metabolismo Macromolecular Firmino Torres de Castro do IBCCF/UFRJ.

Ribeiro, ASD; Dias, VHG; Mello, ICT; Silva, R; Sabino, BD; Garrido, RG; Seldin, L; Moura-Neto, RS. The *rbcl* gene as a barcode tool for forensic identification of *Cannabis sativa*. *Saúde, Ética & Justiça*. 2013;18(Ed. Especial):67-71.

ABSTRACT: *Cannabis sativa* is one of the oldest species of domesticated plants and crops remains one of the most widely deployed and most used illicit drug in the world. There is a great difficulty in identifying and differentiating the samples of *Cannabis* sp, complicating the correlation of the likely places illegal plantations, which would reveal trafficking routes, criminal groups relate and distinguish those samples legal marketed as a drug, where cultivation is permitted. The main purpose of the DNA barcode is to provide a rapid and accurate identification of organisms from a standard small region of the genome that help characterize and distinguish between species and individuals not to assign the identified species. A candidate gene is the ribulose-1,5-bisphosphate carboxylase/oxygenase (*rbcl*), used as a DNA barcode system and present in chloroplast DNA of plants, and are responsible for producing large subunit of converting carbon dioxide and water into carbohydrates. Our goal is to develop an efficient protocol for DNA extraction and sequencing of *rbcl* gene for forensic analysis of seized samples of the Judicial Police. DNA samples from *Cannabis sativa* were characterized at ICCE/DGPTC/PCERJ and extracted using the DNeasy Plant Mini kit (Qiagen) at IPPGF/DGPTC/PCERJ. The material was transferred to LabFor/UFRJ and a fragment of 735 bp *rbcl* gene was amplified using a primer pair described in the literature as universal plants. The fragment was sequenced with the BigDye v3.1 protocol using ABI 3500 (Life Technologies). Comparisons of the sequences were performed in software Geneious (Biomatters). Our consensus sequence of 559 bp was compared to sequences that correspond to regions *rbcl* in *Cannabis sativa* deposited in GenBank, and observe the occurrence of polymorphisms (SNPs) between samples of the United States, Britain and China, suggesting that it is a genetic signature for forensic analysis.

KEYWORDS: *Cannabis sativa*; Genetic signature; Forensic identification.

REFERÊNCIAS

1. Van Bakel H, Stout JM, Cote AG, Tallon CM, Sharpe AG, Hughes TR, Page JE. The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol*. 2011;12(10):R102.
2. Russo EB, Jiang H-E, Li X, Sutton A, Carboni A, Bianco F del, Mandolino G, Potter DJ, Zhao Y-X, Bera S, Zhang Y-B, Lü E-G, Ferguson DK, Hueber F, Zhao L-C, Liu C-J, Wang Y-F, Li C-S. Phytochemical and genetic analyses of ancient cannabis from Central Asia. *J Exp Bot*. 2008;59:4171-82.
3. World Drug Report 2011. e-book. Available from: <http://www.unodc.org/>
4. Castro JLO. Desempenho forense de microssatélites para a investigação da origem de *Cannabis sativa* no Brasil e no Paraguai [tese]. Brasília: Universidade Católica de Brasília; 2006.
5. Gilmore S, Peakall R, Robertson J. Organelle DNA haplotypes reflect crop-use characteristics and geographic origins of *Cannabis sativa*. *Forensic Sci Int*. 2007;172:179-90.
6. Hebert PD, Cywinska A, Ball SL, de Waard JR. Biological identifications through DNA barcodes. *Proc Biol Sci*. 2003; 270:313-21.
7. Ratnasingham S, Hebert PD. BOLD: the barcode of life data system. *Mol Ecol Notes*. 2007; 7:355-64. Available from: <http://www.barcodinglife.org>
10. Clement LW, Donoghue MJ. Barcoding success as a function of phylogenetic relatedness in *Viburnum*, a clade of woody angiosperms. *BMC Evol Biol*. 2012;12:73.
9. Schuettpelz E, Pryer KM. Fern phylogeny inferred from 400 leptosporangiate species and three plastid genes. *Taxon*. 2007;56:1037-50.
10. Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ. *Sistemática vegetal: um enfoque filogenético*. 3a ed. Porto Alegre: Artemed; 2009. p.39.