UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOLOGIA GERAL
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA

Área de Concentração: Biotecnologia, Genômica e Bioinformática.

TESE DE DOUTORADO

Aplicação da tecnologia de DNA Barcode em espécies vegetais aprovadas pela ANVISA e comercializadas no Brasil e elaboração de metodologia para certificação de fitoterápicos.

Orientado: Rafael Melo Palhares
Orientador: Dr. Guilherme Corrêa de Oliveira
Coorientadora: Dra. Maria das Graças Lins Brandão

Belo Horizonte
2015

Rafael Melo Palhares

Aplicação da tecnologia de DNA Barcode em espécies vegetais aprovadas pela ANVISA e comercializadas no Brasil e elaboração de metodologia para certificação de fitoterápicos.

Tese de doutorado apresentada ao Programa de Pós-Graduação em Genética do Departamento de Biologia Geral do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Doutor em Genética.

Orientador: Dr. Guilherme Corrêa de Oliveira

Coorientadora: Dra. Maria das Graças Lins Brandão

Belo Horizonte

2015

PROGRAMA DE PÓS-GRADUAÇÃO DEM GENÉTICA

Aplicação da tecnologia de DNA Barcode em espécies vegetais aprovadas pela ANVISA e comercializadas no Brasil e elaboração de metodologia para certificação de fitoterápicos.

por

Rafael Melo Palhares

Este trabalho foi realizado no grupo de Genômica e Biologia Computacional do Centro de Pesquisas René Rachou e na Faculdade de Farmácia da Universidade Federal de Minas Gerais sob orientação do Dr. Guilherme Corrêa de Oliveira e coorientação da Dra. Maria das Graças Lins Brandão.

Belo Horizonte
2015

AGRADECIMENTOS

Agradeço ao meu orientador Dr. Guilherme Corrêa de Oliveira e à minha coorientadora Dra. Maria das Graças Lins Brandão por toda a confiança e pelas oportunidades oferecidas. Esse trabalho não teria sido possível sem os conselhos e conhecimentos compartilhados por ambos.

Ao programa de Pós Graduação em Genética e todos os seus docentes que apoiaram direta ou indiretamente esse trabalho, em especial à Mary das Graças Santos.

À minha família, em especial à minha mãe, meu pai e minha irmã. As três pessoas que estão presentes em minha vida desde o início dela e que foram os responsáveis por me fazer acreditar que todo e qualquer objetivo que exista em minha vida pode ser alcançado com trabalho e dedicação.

Aos meus amigos (minha segunda família), Rafael, Marina, Lúcio, Natalie, Cleiston, Gustavo, Bruno, Cristiano e Ana Carolina. Porque desabafar e esquecer dos problemas de vez em quando também faz parte do trabalho.

À Myleus! Bruno Brasil, Marcela, Bruno Borges, Estevam, Daniel e Mariana. Esse trabalho nasceu de uma ideia conjunta. Ideia que foi levada para frente com a participação de todos e que se transformou na sociedade que é hoje. Toda a ajuda e cobrança dos amigos da Myleus me mostrou que não há limites para o quanto uma pessoa pode aprender e se aprimorar.

A todos os amigos e membros do CEBIO, Grupo de Genômica e Biologia Computacional e Laboratório de Farmacognosia. Pessoas de competência e que sempre estiveram dispostos a ajudar e contribuir.

Agradeço especialmente à Fer. Se todos acima me ajudaram de um modo ou de outro no desenvolvimento desse trabalho, você me ajudou em todos os modos, Fer. Em todos os dias, tardes e noites você fez, faz e sempre fará parte do meu desenvolvimento intelectual e pessoal. Você me faz sempre querer ser uma pessoa melhor. Muito obrigado, linda!

“Whether one is rich or poor, educated or illiterate, religious or nonbelieving, man or woman, black, white, or brown, we are all the same. Physically, emotionally, and mentally, we are all equal. We all share basic needs for food, shelter, safety, and love. We all aspire to happiness and we all shun suffering. Each of us has hopes, worries, fears, and dreams. Each of us wants the best for our family and loved ones. We all experience pain when we suffer loss and joy when we achieve what we seek. On this fundamental level, religion, ethnicity, culture, and language makes no difference.”

Dalai Lama

Sumário

RESUMO

As plantas medicinais são reconhecidas como essenciais para a complementação da saúde humana, além de potenciais fontes para novos medicamentos. O uso de plantas medicinais e medicamentos fitoterápicos tem aumentado nas últimas décadas, sendo que em alguns países até 90% da população faz uso dessa classe de medicamento. A autenticidade e qualidade desses produtos é um problema real de ordem legal, econômica, de saúde populacional e conservação da biodiversidade. Para garantir a eficácia e segurança desse segmento farmacêutico, legislações específicas, que definem os conceitos legais para estes produtos, além de determinados parâmetros de qualidade foram estabelecidas no Brasil e no mundo. Alguns desses parâmetros são a identificação da espécie vegetal e a verificação da presença, pureza e concentração de compostos químicos relacionados à eficácia do medicamento. A correta identificação da matéria prima vegetal utilizada para a produção do fitoterápico representa um desafio e o uso de espécies incorretas pode representar um risco para a saúde humana. Atualmente essa identificação é feita por métodos botânicos e, em muitos casos, ela é dificultada ou mesmo impossibilitada, dependendo do nível de processamento da matéria prima, da parte vegetal comercializada ou mesmo de variações fenotípicas presentes em alguns táxons. Em contrapartida, a identificação molecular de espécies em produtos a partir da tecnologia de DNA barcode vem crescendo no mundo. Aplicando sequências específicas do DNA, essa técnica vem sendo utilizada para catalogar as espécies de seres vivos do planeta. Seu uso de forma aplicada à identificação de produtos de consumo humano está se tornando cada vez mais comum e eficiente. A presente tese está centrada na utilização dessa técnica para identificação da matéria prima vegetal de espécies utilizadas no Brasil e no mundo para a saúde humana. Além disso, foi realizada uma correlação da identificação vegetal com análises químicas para verificação de presença e teor dos princípios ativos nas amostras analisadas. Brevemente, em dois capítulos nós demonstramos que existe no Brasil o comércio de espécies errôneas, diferentes das descritas na venda, potencialmente ameaçadas e mesmo sem registros de eficácia ou segurança. Demonstramos ainda que mesmo em espécies corretas, os padrões químicos podem estar abaixo do mínimo requerido para a obtenção dos resultados esperados do tratamento. Com esses resultados, e com a tendência observada mundialmente, concluímos que a adição de uma etapa de identificação molecular ao processo de produção de fitoterápicos agregará melhoria na qualidade desses medicamentos, além de contribuir para a conservação da biodiversidade ao inibir o uso de espécies incorretas.

ABSTRACT

Medicinal plants are recognized as essential to the completion of human health, as well as potential sources for new medicines. The use of medicinal plants and herbal medicines has increased in recent decades, and in some countries up to 90% of the population uses this class of drug. The authenticity and quality of these products is a real problem legally, economically, for populational health and biodiversity conservation. To ensure the effectiveness and safety of this pharmaceutical area, specific laws, which define the legal concepts for these products, and certain quality parameters have been established in Brazil and worldwide. Some of these parameters are the identification of plant species and checking for the presence, purity and concentration of chemical compounds related to the effectiveness of the drug. The correct identification of the vegetable raw material used for the production of herbal medicine is a challenge and the use of incorrect species may pose a risk to human health. Currently this step is done by botanical methods of identification and, in many cases it is difficult or even impossible to be made, depending on the level of processing of the raw material, the marketed plant parts, or even due to phenotypic variations present in some taxa. In contrast, the molecular identification of species in products using the DNA barcode technology is growing in the world. Applying specific DNA sequences, this technique has been used to catalog all the species of living beings on the planet. Its use to the identification of human consumer products is becoming increasingly common and efficient. This thesis focuses on the use of this technique to identify the species of vegetable raw material used in Brazil and in the world to human health. In addition, a correlation between the identification of plant species and chemical analysis for verification of presence and content of the active ingredients in the samples was performed. Briefly, in two chapters we have demonstrated that there is in Brazil trade of wrong species, different from those described in the sale, potentially threatened and even with no records effectiveness and safety. We also demonstrated that even in the correct species, chemical standards might be below the minimum required to achieve the expected results of the treatment. With these results, and with the worldwide trend, we conclude that the addition of a molecular identification step to herbal medicine production process will add improvement in the quality of these products, and contribute to biodiversity conservation by inhibiting the use of incorrect species.

I – INTRODUÇÃO

1.1 – Plantas medicinais e regulamentação para seu uso

As plantas são uma das mais importantes fontes de novas substâncias bioativas para o desenvolvimento de produtos cosméticos, farmacêuticos e nutracêuticos. Estima-se que 25% dos medicamentos modernos existentes hoje são derivados de produtos naturais e, no caso dos antitumorais, esta porcentagem chega a 60% (LI & VEDERAS; SAHOO et al., 2009). O desenvolvimento de novos produtos a partir de fonte natural vem sendo incentivado, pois calcula-se que das 300 mil espécies de plantas existentes no planeta, apenas 15% foram submetidas a algum estudo para avaliar suas potencialidades (DE LUCA et al., 2012).

Segundo a Organização Mundial da Saúde (OMS), entre 65% e 80% da população dos países em desenvolvimento depende das plantas medicinais para as suas necessidades básicas de saúde (WHO, 2005). Em países africanos, esse percentual pode atingir 90% (WHO, 2005). Além do seu uso popular e caseiro, as plantas são também utilizadas como medicamentos, chás, cosméticos e suplementos alimentares e a demanda por produtos oriundos de plantas vem crescendo significativamente (CARVALHO et al., 2011). Esse setor movimenta aproximadamente US$160 milhões por ano no Brasil, e mais de US$21,7 bilhões no mundo (CARVALHO et al., 2011) e passou por uma forte ascensão na última década (FEBRAFARMA, 2009). A expansão desse setor pode ser atribuída a diversos fatores, tais como efeitos adversos de fármacos sintéticos e a preferência dos consumidores por tratamentos "naturais". Além disso, a validação científica das propriedades farmacológicas de espécies vegetais e seus derivados, o desenvolvimento de novas formas de preparações e administrações de produtos fitoterápicos, um melhor conhecimento químico, farmacológico e clínico, além do menor custo se comparado com os fármacos sintéticos, também contribuem para essa expansão (VIEIRA, 2001; CAÑIGUERAL et al., 2003; WACHTEL-GALOR & BENZIE, 2011).

Seguindo as exigências da OMS e da Agência Nacional de Vigilância Sanitária (ANVISA), para a produção de um fitoterápico a identidade da espécie vegetal deve ser garantida, o processo de produção deve ser padronizado para a obtenção do produto final, e devem ser observadas as possibilidades de contaminação ou adulteração (WHO, 2008; CARVALHO et al., 2011). Legislações específicas para regular a produção de medicamentos fitoterápicos vêm sendo criadas no Brasil desde 1995 (BRANDÃO et al., 2008). Atualmente está em vigor, a Resolução da Diretoria Colegiada - RDC nº 26/2014 (13 de maio de 2014) (ANVISA, 2014). Esta RDC define conceitos legais para drogas vegetais (tais como: parte da planta desidratada e estabilizada que

contém o princípio ativo e que é utilizada como matéria prima na preparação de fitoterápicos) e medicamentos fitoterápicos. Além disso, determina aspectos essenciais ao seu registro e estabelece os métodos para o controle de qualidade das drogas vegetais e seus produtos (ANVISA, 2014).

Os métodos utilizados para avaliar a qualidade de drogas vegetais encontram-se especificados nas Farmacopeias, sendo no Brasil a Farmacopeia Brasileira 5ª Edição. Eles são divididos em três etapas: a primeira etapa abrange a certificação do material botânico. Para isto, avalia-se se a droga vegetal em análise corresponde à espécie descrita nos regulamentos, ou se trata de alguma falsificação. A segunda etapa consiste na avaliação dos caracteres de pureza a partir da verificação da presença de elementos estranhos (não próprios da amostra) além de verificar a ocorrência de outras adulterações. A terceira etapa visa verificar a presença e teor dos princípios ativos, necessários para as atividades farmacológicas de determinada espécie (BRANDÃO, 2007). Para o estudo de autenticidade da espécie vegetal são utilizadas caracterizações organolépticas, macroscópicas e microscópicas. Para o estudo de pureza são realizadas verificação da presença de elementos estranhos (outras partes do vegetal que não contem os marcadores químicos, ou outras sujidades) e testes de teor de umidade, já que o excesso de umidade promove o crescimento de microrganismos e a deterioração do material vegetal. São realizadas ainda análises do perfil fitoquímico para análise de presença e teor dos princípios ativos (BRANDÃO et al., 2002).

1.2 – As dificuldades na etapa de identificação

A autenticidade de espécies vegetais é feita por taxonomistas altamente treinados por meio da observação de características anatômicas e morfológicas do espécime (ANVISA, 2014). Entretanto, no caso das plantas medicinais usadas na preparação dos medicamentos, esta análise é dificultada, ou até mesmo impossibilitada uma vez que a matéria-prima vegetal encontra-se processada sob a forma de droga vegetal.

Ainda, em muitos casos, há a comercialização de partes não identificáveis botanicamente. Somada a essas restrições, alguns táxons apresentam alta variação fenotípica, dificultando sua correta identificação por observação morfológica (SUCHER & CARLES, 2008; NEWMASTER et al., 2009). As consequências de tais dificuldades para a identificação da espécie contida em determinada droga vegetal são os diversos casos de substituições ou fraudes encontradas no

mercado mundial de produtos de origem vegetal (SONG et al., 2009, STOECKLE et al., 2011; WALLACE et al., 2012, NEWMASTER et al., 2013).

Por exemplo, Coghlan et al (2012) verificaram que medicamentos tradicionais chineses apreendidos na alfândega australiana não só utilizavam espécies protegidas ambientalmente, como continham espécies tóxicas, produtoras de substâncias alergênicas e potencialmente cancerígenas. Kumar et al (2011), por sua vez, mostraram a adição, ao azeite de oliva, de óleo de canola e de girassol. Recentemente, Brandão et al (2013) demonstraram que, dentre 252 amostras vegetais obtidas em ervanarias nas cinco regiões geográficas do Brasil, apenas 126 (50%) correspondiam à espécie correta.

Esses resultados são reflexo das dificuldades em identificar as espécies contidas nas drogas vegetais, e revelam a necessidade da padronização e validação de novas metodologias que possibilitem a correta identificação.

A necessidade de identificação de espécies vegetais existe em diversas disciplinas e situações, além da certificação de medicamentos fitoterápicos. Alguns exemplos seriam: pesquisas envolvendo os hábitos alimentares de um herbívoro; pesquisas que envolvem o acesso à biodiversidade de uma determinada área; controle, por agentes alfandegários, da entrada e saída de materiais de seus países; análises forenses; identificação de espécies responsáveis por envenenamentos; identificação de ingredientes de produtos alimentícios (NEWMASTER et al., 2009).

Quando colocados em conjunto, esses cenários geram uma demanda e uma urgência muito maior que a oferta que os taxonomistas podem cobrir. Consequentemente, a falta de metodologias de controle de qualidade adequadas, em conjunto com falhas na farmacovigilância, tem possibilitado que os medicamentos fitoterápicos sejam comercializados fora dos padrões estabelecidos pela ANVISA (BRANDÃO et al., 2002).

1.3 – DNA Barcode e a identificação de espécies

O uso de ferramentas moleculares tem se tornado importante para a identificação de amostras botânicas. Métodos de identificação de espécies de plantas medicinais baseados em análise de DNA vêm sendo desenvolvidos em várias partes do mundo (SUCHER & CARLES, 2008; CBOL; NEWMASTER et al., 2009), e já estão sendo aplicados em amostras comerciais para promover

maior segurança e confiabilidade (KUMAR et al.; LI et al.; STOECKLE et al., 2011; COGHLAN et al., 2012). Tais métodos baseiam-se na amplificação do DNA presente no núcleo ou cloroplastos de células vegetais por meio da reação em cadeia da polimerase (PCR). Os produtos dessas reações são então analisados por eletroforese, sequenciamento ou hibridações com sondas espécie-específicas. As metodologias moleculares possuem, como principais atraentes, sua alta sensibilidade e especificidade, além de uma grande escalabilidade, o que acaba por tornar todo o processo mais rápido e barato.

Dentre as técnicas moleculares de identificação genética de espécies ganha destaque o DNA barcode. Esta metodologia consiste em um sistema capaz de empregar sequências de DNA específicas de forma análoga a um código de barras identificador para cada táxon (unidade de classificação biológica), funcionando como um sistema bioidentificador global prático, econômico e específico (HERBERT et al., 2003). O rápido progresso desta técnica impulsionou a criação de um consórcio mundial denominado *Consortium for barcode of Life* (CBOL – http://www.barcoding.si.edu), que objetiva a geração de *barcodes* genéticos para todas as espécies de seres vivos do planeta.

Inicialmente proposto para identificação de espécies animais, o DNA barcode para metazoários baseia-se na amplificação e sequenciamento de uma região do gene mitocondrial codificador da enzima Citocromo C Oxidase I (COI) (HERBERT et al., 2003; CARVALHO et al., 2008; AUSUBEL, 2009). Já para espécies vegetais, a definição de uma região genômica adequada para o desenvolvimento do DNA barcode revelou-se um desafio para vários grupos de pesquisa nos últimos anos (PENNISI, 2007; AUSUBEL; CHASE & FAY; NEWMASTER et al., 2009).

O Grupo de Trabalho de Plantas do CBOL propôs a utilização de um barcode padrão baseado em dois *loci* encontrados no DNA de cloroplastos. Somente com a combinação destes dois *loci* foi possível atender aos critérios de universalidade, poder de discriminação entre táxons e potencial para sequenciamento não-ambíguo e de qualidade requeridos para o barcode genético (AUSUBEL, 2009). O primeiro consiste em uma região de 599 pb localizada na extremidade 5´ do gene *rbcL*. O segundo consiste em uma região de 841 pb localizada no centro do gene *matK* e que apresenta uma das mais rápidas taxas de evolução dentre as regiões cloroplastídeas codificantes (CBOL, 2009; HOLLINGSWORTH et al., 2011).

Apesar da definição oficial desse núcleo de barcode, diversos grupos de pesquisa acreditam na possibilidade de se encontrar uma melhor combinação entre *loci*, que aprimoraria o barcode vegetal, melhorando sua eficiência (CHEN et al., 2010; GUO et al.; LI et al., 2011). Dentre outras

potenciais regiões, ganha destaque a sequência de ITS2. Inicialmente descrita por Chiou et al (2007), essa região é amplamente utilizada para reconstruções filogenéticas a nível de gênero e espécie (SCHULTZ & WOLF, 2009).

Um dos maiores ganhos da definição das regiões oficiais para o barcode vegetal foi a padronização, dentre os grupos de pesquisas mundiais, das sequências utilizadas e depositadas em bancos de dados. Isso fez com que o número de sequências geradas aumentasse rapidamente, permitindo a aplicação da tecnologia de DNA barcode de forma mais ampla (http://v3.boldsystems.org/index.php/Public_BINSearch?searchtype=records).

Atualmente, o banco de dados do CBOL, denominado BOLD (*Barcode of Life Database* – http://www.boldsystems.org/) está em sua terceira versão e conta com o depósito de sequências identificadoras para 229.379 diferentes espécies de vegetais, animais, fungos e protistas de todo o mundo, totalizando 3.721.329 *barcodes* (http://v3.boldsystems.org/index.php/TaxBrowser_Home – acessado em 03/02/2015). Por sua vez, o Brasil, por meio do Projeto BrBOL contribui com o BOLD realizando a identificação molecular da biodiversidade brasileira em onze diferentes projetos (http://brbol.org/pt-br/content/brbol-brazilian-barcode-life-0 - acessado em 08/02/2015).

Estudos recentes mostram o uso do núcleo de DNA barcode para espécies vegetais, associado a sequências complementares, na identificação de espécies medicinais (LI et al., 2011; BOONSOM et al., 2012), na identificação de fraudes e uso de espécies em extinção ou protegidas (COGHLAN, et al., 2012) e na análise de presença de adulterantes em produtos alimentícios (KUMAR et al.; STOECKLE, 2011).

Considerando a necessidade de desenvolvimento de tecnologias de identificação de espécies para o controle de autenticidade de plantas medicinais e fitoterápicos, propõe-se neste estudo a utilização do estado da arte da metodologia de DNA barcode para a padronização de testes de identificação genética práticos e precisos de espécies vegetais comumente utilizadas no Brasil. De maneira a garantir a qualidade da matéria prima utilizada para a produção de fitoterápicos, sugere-se o uso da análise molecular acompanhada das análises químicas exigidas na legislação.

II – OBJETIVOS

2.1 – Objetivo geral

Avaliar a eficácia do uso da tecnologia de DNA barcode para a identificação da matéria prima vegetal processada (droga vegetal) de espécies utilizadas na produção de fitoterápicos no Brasil.

2.2 – Objetivos específicos

- Desenvolver e padronizar testes de identificação de espécies por meio da técnica de DNA barcode para espécies vegetais de importância comercial no Brasil.
- Analisar a eficiência da inclusão da região ITS2 no núcleo de DNA barcode vegetal para a identificação das espécies selecionadas.
- Analisar o perfil fitoquímico das espécies selecionadas e cruzar os resultados com aqueles obtidos nas análises de DNA barcode.
- Analisar a eficiência da inclusão da identificação molecular por DNA barcode no processo de produção de fitoterápicos.

III – CAPÍTULOS

3.1 – CAPÍTULO 1: Uma abordagem química, biológica e molecular para promover o melhor uso e conservação de espécies medicinais: o caso das quinas brasileiras

Nesse estudo, amostras de cascas de Quinas vendidas no comércio de Belo Horizonte foram analisadas por DNA barcode associado a métodos químicos e biológico. Quina é o nome popular originalmente atribuído às espécies *Cinchona pubescens* Vahl e *C. calisaya* Wedd., ambas nativas do Peru e com propriedades antimaláricas atribuídas ao alcaloide quinina. Além da introdução do plantio sem sucesso dessas espécies no Brasil pelos portugueses, espécies substitutas foram encontradas, sendo as mais conhecidas *Remijia ferruginea* (A. St.-Hil) DC. e *Strychnos pseudoquina* A. St.-Hil. A falta de informação científica que confirme a segurança e a eficácia de espécies nativas do Brasil e usadas como substitutas é um fator preocupante, pois tais plantas são amplamente comercializadas em mercados populares. Esse fato não só pode colocar em risco a saúde da população como promove a erosão genética de espécies da flora brasileira. Vinte e oito amostras vendidas em mercados populares de Belo Horizonte com o nome de quinas foram identificadas por meio de DNA barcode. Em paralelo foram determinados seus perfis fitoquímicos e eficácia antimalárica *in vitro*. Os resultados mostraram que dez amostras pertenciam ao gênero *Strychnos*, três ao gênero *Aspidosperma* e ao gênero *Coutarea*, duas ao gênero *Lamanonia* e uma ao gênero *Lacmellea*, *Ladenbergia*, *Bowdichia*, *Machaerium* e *Cissampelos*. Entre essas, apenas espécies de *Strychnos, Aspidosperma* e *Coutarea* têm representantes conhecidos como substitutas da quina verdadeira. A análise de DNA barcode evidenciou que estão presentes no comércio espécies sem qualquer relação com a medicina tradicional. A inconstância observada na composição química desses produtos, bem como na atividade antimalárica, confirma a baixa qualidade das amostras comercializadas nos mercados. A análise integrada utilizando metodologias moleculares, químicas e biológicas desses produtos se mostrou altamente eficiente na identificação das drogas vegetais e precisaria ser aplicada na identificação das plantas medicinais para garantir maior qualidade. Além disso, o melhor conhecimento das espécies que são coletadas para suprir os mercados auxilia na definição de estratégias para a sua conservação.

Journal of Ethnopharmacology 155 (2014) 815–822

Contents lists available at ScienceDirect

Journal of Ethnopharmacology

journal homepage: www.elsevier.com/locate/jep

The use of an integrated molecular-, chemical- and biological-based approach for promoting the better use and conservation of medicinal species: A case study of Brazilian quinas

Rafael M. Palhares[a], Marcela G. Drummond[b], Bruno S.A.F. Brasil[c], Antoniana U. Krettli[d], Guilherme C. Oliveira[e], Maria G.L. Brandão[f,*]

[a] Programa de Pós-graduação em Genética, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil
[b] Myleus Biotecnologia Research Team, Belo Horizonte, Brasil
[c] EMBRAPA Agroenergia, Brasília, Brazil
[d] Laboratório de Malária, Centro de Pesquisa Rene Rachou, FIOCRUZ, Belo Horizonte, Brasil
[e] Grupo de Genômica e Biologia Computacional, Centro de Pesquisa Rene Rachou, FIOCRUZ, Belo Horizonte, Brasil
[f] CEPLAMT, Museu de História Natural e Jardim Botânico & Faculdade de Farmácia, Universidade Federal de Minas Gerais, Avenida Antônio Carlos, 6627 Belo Horizonte, Brasil

ARTICLE INFO

Article history:
Received 12 April 2014
Received in revised form
15 June 2014
Accepted 15 June 2014
Available online 24 June 2014

Keywords:
Brazil
Quina
DNA barcode
Popular market
Endangered species

ABSTRACT

Ethnopharmacological relevance: Quina is a popular name originally attributed to *Cinchona pubescens* Vahl (=*Cinchona succirubra*) and *Cinchona calisaya* Wedd., species native from Peru that have the antimalarial alkaloid quinine. In Brazil, bitter barks substitutes for the Peruvian species began to be used centuries ago, and they still are sold in popular markets. To assess the authenticity and the conditions on which samples of quinas have been commercialized, using the DNA barcode, chemical and biological assays.

Materials and methods: Starting with 28 samples of barks acquired on a popular market, 23 had their DNA extracted successfully. The regions matK and rbcL were amplified and sequenced for 15 and 23 samples, respectively. Phytochemical analyses were performed by chromatographic methods, and biological essays were done by antimalarial tests in vitro.

Results: The identified species belonged to six different families, many of them endangered or with no correlation with use in traditional medicine as a Brazilian quina. The absence of typical bitter chemical substances indicated that barks have been collected from other species or from very young trees. The results of biological essays confirm the lack of standardization of the sold materials.

Conclusion: The integrated approaches proved to be efficient to evaluate medicinal plants sold in popular markets and can be useful for promoting their better use and conservation.

1. Introduction

According to the World Health Organization (WHO), between 65% and 80% of the population in developing countries currently uses medicinal plants as remedies (WHO, 2011). In Brazil, native species have been used for millennia by the Amerindians, and many important medicines, as pilocarpine, tubocurarine and emetine, were identified from their knowledge (Li and Vederas, 2009; Nogueira et al., 2010). Currently, medicinal plants are still in use by both rural and urban Brazilians, but most of the plants that are used are composed of exotic species, that were introduced to the country during the early phases of European colonization, in the 1500s. This phenomenon is also related to the intermingling of cultures that has occurred during the previous centuries, as well as the result of the continuous destruction of the rich Brazilian ecosystems. In fact, the accelerated destruction of native vegetation has contributed to a gradual loss of native medicinal species and the traditional knowledge about them (Shanley and Luz, 2003; Voeks and Leony, 2004). Today, only 7% of the Atlantic forest is preserved, and other ecosystems, such as the savannas (cerrado) and caatinga, are gradually being replaced by monocultures of eucalyptus, sugarcane, soybeans and livestock (Giulietti et al., 2005; Mittermeier et al., 2005). This situation highlights the necessity to promote the adequate use and conservation of medicinal species from Brazilian biodiversity, especially those used in traditional medicine.

Among the old-used native medicinal plants are those named quina. This name is originally attributed to *Cinchona pubescens* Vahl (=*Cinchona succirubra*) and *Cinchona calisaya* Wedd (Rubiaceae),

* Corresponding author. Tel.: +55 31 34096970; fax: +55 31 34096935.
E-mail address: mbrandao@ufmg.br (M.G.L. Brandão).

http://dx.doi.org/10.1016/j.jep.2014.06.040
0378-8741/

native from Peru and which bark furnishes the bitter alkaloid quinine. Besides it uses to treat malaria, quinine is also currently utilized as a flavor in tonics, bitters, aromatized wines and other spirits (Stewart, 2013). Knowledge of the febrifuge effect of the quina's barks was assimilated from Amerindian culture by the Spanish colonizers in the 17th century (Paz-Soldán, 1941). This remedy was introduced to Europe in 1640 by the Jesuits (named Jesuit's powder), and until the 18th century, the European countries, which at the time were expanding their population in tropical regions, sought to spread the cultivation of *Cinchona*. The Portuguese, Brazil's colonizers, also followed this practice and began growing *Cinchona* in various African colonies (Ferrão, 2005). In Brazil, however, they promoted the search for substitutes of these barks from the native vegetation (Peckolt, 1916; Souza, 1951). As a consequence, several other bitter plants were discovered, which also became referred to as quina, and were used as substitutes for *Cinchona* in Brazilian traditional medicine. In a recent study, we showed data from 29 such species that have been used historically. The study showed that *Remijia ferruginea* (A.St.-Hil.) DC. (Rubiaceae) and *Strychnos pseudoquina* A. St.-Hil. (Loganiaceae) were the most widely used as substitute of the true quina from Peru.

Since 1995, the Brazilian Ministry of Health, following the recommendations of the WHO, started to establish a set of herbal regulations to improve the quality of commercial herbal products. By these regulations, only plants with criteria of efficacy and security can be used for preparing industrialized herbal products. The lack of scientific information that confirms the efficacy and security of native Brazilian species has led, as consequence, their intensive commerce in popular markets. In a previous study, we showed the positive effect on the harvesting of native species by pharmaceutical companies, after the establishment of these regulations (Brandão et al., 2010). The impact of collecting native plants sold in popular markets, however, is more difficult to be estimated because they are obtained in the wild without control or voucher herbarium samples, for accurate identification (Melo et al., 2009).

The use of molecular techniques has been proven to be a powerful tool for the taxonomic identification of medicinal plants, including species with very similar morphological and chemical characteristics (Herbert et al., 2003; He et al., 2010; Gao et al., 2011; Zuo et al., 2011), as well as crude vegetal drugs sold in market (Kool et al., 2012; Newmaster et al., 2013). In this study, we used the molecular identification integrated to chemical and biological methods, which authenticate and verify the conditions in which species named quina have been collected and commercialized in a popular market of Brazil. The final objective is to contribute for the better use of these remedies and their conservation.

2. Materials and methods

2.1. Collection of the commercial samples

Twenty-eight samples of barks marketed with named of "quina" were purchased from the Popular Market of Belo Horizonte, a city of 2,500,000 inhabitants located in the Southwestern region of Brazil in Minas Gerais. The samples were transported to the laboratory at the vegetal drug collection (DATAPLAMT) of the Museum of Natural History and Botanic Garden of Federal University of Minas Gerais (samples Q-1 to Q-28), where they were registered and conditioned in an acclimatized room.

2.2. Reference samples

Barks of *Remijia ferruginea* (A.St.-Hil.) DC. (Rubiaceae) and *Strychnos pseudoquina* A.St.-Hil. (Loganiaceae) were collected in Diamantina (18° 5′51,1″/ 43° 27′51,3″) and Curvelo (18° 49′49,8″/ 44° 32′17,4″), respectively, and used as reference samples. These species were selected for this procedure because they were widely used in the past and considered the official Brazilian quina in Pharmacopoeia (WHO, 2005; Brandão et al., 2008). Voucher specimens were deposited in DATAPLAMT (DAT201-DAT209). Following the standards recommended by the International Consortium for the Barcode of Life (IBOL - www.ibol.org) for DNA Barcode identification, five samples of *Remijia ferruginea* and four from *Strychnos pseudoquina* were used to provide DNA barcode reference vouchers. A total of eighteen sequences were generated and deposited in GenBank (Table 1).

Table 1
Voucher identification and GenBank deposit number of the DNA Barcode sequences for reference samples.

Species	Voucher[a]	GenBank acession	Region	Sequence
Remijia ferruginea	DAT201	KF667963	*matK*	505 pb
		KF667946	*rbcL*	524 pb
Remijia ferruginea	DAT202	KF683518	*matK*	718 pb
		KF683525	*rbcL*	613 pb
Remijia ferruginea	DAT203	KF683519	*matK*	718 pb
		KF683526	*rbcL*	613 pb
Remijia ferruginea	DAT204	KF683520	*matK*	718 pb
		KF683527	*rbcL*	613 pb
Remijia ferruginea	DAT205	KF683521	*matK*	718 pb
		KF683528	*rbcL*	613 pb
Strychnos pseudoquina	DAT206	KF667964	*matK*	505 pb
		KF667947	*rbcL*	524 pb
Strychnos pseudoquina	DAT207	KF683522	*matK*	739 pb
		KF683529	*rbcL*	613 pb
Strychnos pseudoquina	DAT208	KF683523	*matK*	739 pb
		KF683530	*rbcL*	613 pb
Strychnos pseudoquina	DAT209	KF683524	*matK*	739 pb
		KF683531	*rbcL*	613 pb

[a] The Vouchers are deposited in the DATAPLAMT collection – Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais – Brazil.

2.3. DNA extraction

The DNA was extracted from the bark of plants using the DNeasy plant mini kit (QIAGEN – catalog number 69106) with modifications. Approximately 20 mg of each sample was macerated using a mortar at room temperature. The powder was mixed with 600 uL of the buffer AP1 supplied with the kit and incubated at 65 °C and 400 rpm for 1 h in Heatblock (Eppendorf – Thermomixer compact). After that, 230 uL of the buffer AP2 supplied with the kit was added, and the samples were incubated on ice for 30 min. The lysate was centrifuged for 5 min at 20.000 × g in microcentrifuge (Eppendorf 5417c Rotor FA-45-24-11). The lysate was pipetted into a QIAshredder Mini spin column with a 2 mL collection tube attached supplied with the kit and centrifuged for 2 min at 20.000 × g in microcentrifuge (Eppendorf 5417c Rotor FA-45-24-11). The flow-through was transferred into a new tube, 1.5 volumes of Buffer AP3/E supplied with the kit was added to it and mixed by pipetting. A volume of 650 uL of the mixture was transferred into a DNeasy Mini spin column attached to a 2 mL collection tube supplied with the kit and centrifuged for 1 min at 6000 × g in microcentrifuge (Eppendorf 5417c Rotor FA-45-24-11). The flow-through was discarded and this step was repeated with the reminder of the sample. The spin column containing the DNA was attached to a new 2 mL collection tube and 300 ul of buffer AW supplied with the kit was added and centrifuged for 1 min at 6000 × g in microcentrifuge (Eppendorf 5417c Rotor FA-45-24-11). This step was repeated two more times to clean the DNA. In the last repetition, the column was centrifuged at 20000 × g in microcentrifuge (Eppendorf 5417c Rotor FA-45-24-11) for two minutes to completely dry all of the ethanol. The spin column was transferred to a 1.5 mL microcentrifuge tube and 50 uL of

buffer AE supplied with the kit was added for elution. The spin column was incubated for 10 min at room temperature and centrifuged for 1 min at 6000 × g in microcentrifuge (Eppendorf 5417c Rotor FA-45-24-11). Then, 5 uL of each of the extracted DNA samples was added to 1 μL of loading buffer and visualized on a 1% agarose gel in TAE buffer with 3 uL of GelRed (Biotium). The 100 bp DNA standard from Invitrogen was used for the analysis of the genomic DNA.

2.4. PCR and sequencing

The amplification of the DNA was carried out using primers selected from the Royal Botanic Gardens Kew Phase 2 Protocols and Update on plant DNA barcoding for the *matK* (primers 3 F_KIM-r and 1 R_KIM-f) and *rbcL* (primers rbcLa_f and rbcLa_jf634R) regions (http://www.kew.org/barcoding/protocols.html). The primers sequences are F 5′ – CGTACAGTACTTTTGTGTTTACGAG – 3′ R 5′ – ACCCAGTCCATCTGGAAATCTTGGTTC – 3′ for *matK* and F 5′ – ATGTCACCACAAACAGAGACTAAAGC – 3′ R 5′ – GAAACGGTCTCTCCAACGCAT – 3′ for *rbcL*.

The DNA was amplified in a reaction with final volume of 25 uL containing 2 U of AmpliTaq Gold polymerase in GeneAmp 106 PCR Buffer II (100 mM Tris–HCl, pH 8.3; 500 mM KCl), 1,5 mM MgCl2 (Applied Biosystems, Foster City, CA), 0,2 mM dNTPs, 0,6 μM of each primer, and 100 to 200 ng of DNA of each extracted sample.

The amplification was carried out in a thermocycler GeneAmp PCR System 9700 (Applied Biosystems) with the conditions as follows: an initial denaturation step of the DNA at 95 °C for 3 min, followed by 45 cycles of 94 °C for 30 s, 50 °C for 40 s and 72 °C for 40 s, with a final extension period of five minutes at 72 °C. Then, 5 uL of each of the resulting amplifications was added to 1 uL of loading buffer and visualized on a 1% agarose gel in TAE buffer stained with GelRed (Biotium). A 100 bp DNA standard (Invitrogen) was used for the analysis of the fragments.

The sequencing reactions were performed at Myleus Biotechnology (Belo Horizonte, Brazil) on an ABI3130 automated sequencer (Applied Biosystems) using BigDye v3.1 and the same primers that were used in the amplification reactions.

2.5. Data analysis

The DNA sequences obtained were edited using the software, SeqScape v2.7 (Life Technologies) and aligned and analyzed with the program MEGA 5.2.2. The reference sequences were obtained from the Barcode of Life Data Systems (BOLD) (http://v3.boldsystems.org/index.php/databases), GenBank (http://www.ncbi.nlm.nih.gov/genbank/) and, as previously mentioned, for *Remijia ferruginea* and *Strychnos. pseudoquina*, sequences that were generated by our group.

The chosen statistical method for the phylogenetic analyses and tree assembly was neighbor-joining (K2P), which was recommended as the standard barcoding method (CBOL, 2009). The analyses were carried using the program, MEGA 5.2.2 (Tamura et al., 2011). The query sequences were identified on the basis of the reference sequences with which they formed a cluster.

2.6. Chemical analysis

The objective was to verify the presence of bitter typical chemical substances in each sample, using chromatographic methods (TLC and HPLC, respectively). The barks were pulverized and extracted to exhaustion by percolation with 80% ethanol. The solutions were dried in a rota-evaporator, wrapped and stored in a freezer −4 °C. TLC was performed using silica gel plates (Merck Darmstadt, Ref 1.05.721) and solvent systems and reagents specific for alkaloids, flavonoids and tannins (Wagner and Bladt, 1996). Quinine (Sigma Ref. 1125, St. Louis, MO, USA), quercetin (Sigma Ref. 1082, St. Louis, MO, USA) and pirogalol (Fluka, BCBG3002V) were used as standards.

For analysis in HPLC, an Agilent 1100 with a DAD UV detector was used. The column used was a Lichrospher C18 (250 um, 4.6 mm, Merck, Darmstadt, Germany) at 30 °C. Acetonitrile (A): water (B) was used as the mobile phase as follows: gradient elution at 35–100% A for 0–45 min and then at 100% A for 45–50 min. Rebalancing was carried out as follows: 100–35% A for 50–55 min, then 35% A for 55 a 60 min at a flow of 1 mL/min. The absence of quinine in all samples was confirmed by the co-analysis with a reference standard (Sigma-Aldrich®, St-Louis, MO, USA). The other substances were identified by analysis of their respective UV spectra at 205 nm. The standards were prepared in acetonitrile/water (1:1) at a concentration of 0.5 mg/mL. The injection volume was 40 uL.

2.7. Antimalarial assays in vitro

The objective was to confront the results on molecular and chemical analysis with the existence of antimalarial active materials. Chloroquine resistant *Plasmodium falciparum* strain K1 and was used in the in vitro tests. Parasites were maintained in continuous culture in A+ human erythrocytes, using RPMI medium supplemented with 10% human serum. The antiparasitic effect of the compounds was measured by growth inhibition percentage as described by Carvalho and Krettli (1991). Briefly, trophozoite-stages in sorbitol-synchronized blood were cultured at 1–2% parasitaemia and 2.5% hematocrit and then incubated with the plant extracts or isolated compounds (maximum 1 mg/ml in serial dilutions), diluted with 0.02% final concentration of DMSO in culture medium (RPMI 1640) for a total of 48 h at 37 °C. A positive control with reference antimalarial drug (chloroquine and quinine) in standard concentrations was used in each experiment. The stock solutions were further diluted in complete medium (RPMI 1640 plus 10% human serum) to each of the used concentrations (0.0001 up to 100 μg/ml in seven dilutions). The half-maximal inhibitory (IC_{50}) responses as compared to the drug-free controls were estimated by interpolation using Microcal Origin software. Each duplicate experiment was repeated three times and blood smears were read blind. The data were analyzed with the Biostat 1.0 MCT-CNPq software package using Anova and Student's *t*-test (Andrade-Neto et al., 2007).

3. Results and discussion

Plants used for the preparation of herbal medicines are commercialized as crude drugs, and the quality of these materials is currently verified by a set of botanical, physico-chemical and chemical analysis, that are established by the Pharmacopoeias and other official Compendia (Brandão et al., 2008; WHO, 2005). These analyses are very helpful for the authentication of materials from well-known vegetal species, which have had their criteria for approbation established in the bibliography. However, such methods are not efficient for identifying samples from native species of Brazilian biodiversity, for which very few or no descriptions are available. In a previous study, we have already observed a substitution of material from native species in the popular markets, and the necessity to improve methods for their identification (Brandão et al., 2013).

In the DNA barcode analysis performed in this study, the sequences obtained from *matK* amplification had 505 bp and the rbcL sequences obtained contained 524 bp. All the sequences were deposited on GenBank (Table 2). All of the amplified sequences were successfully sequenced on an ABI platform by Myleus

Table 2
Barcode identification, acession number and percentual of similiraty between the samples and the identified species on the Barcode of life Database or GenBank.

Sample n°	GenBank acession		Vernacular name[b]	Putative scientific name[c]	matK BOLD			rbcL BOLD			ID Confirmed	FINAL ID
	matK	rbcL			BOLD Records	Species	Similarity	BOLD records	Species	Similarity		
Sample 01[a]		KF667923.1	*Quina*	*Remijia ferruginea Strychnos pseudoquina*				GBVJ3569-11	*Coutarea hexandra*	99.81	Different genus	*Coutarea hexandra*
Sample 03	KF667948.1	KF667924.1	*Quina*	*Remijia ferruginea Strychnos pseudoquina*	GBVG1679-11	*Strychnos erichsonii*	100	GBVX1684-13	*Strychnos potatorum*	99.81	Species confirmed	*Strychnos pseudoquina*
					TSA163-10	*Strychnos henningsii*	100	KF667947	*Strychnos pseudoquina*	100		
					KF667964	*Strychnos pseudoquina*	100					
Sample 04[a]		KF667925.1	*Quina-cruzeiro*	*Remijia ferruginea Strychnos pseudoquina*				*GBVJ3569-11*	*Coutarea hexandra*	99.81	Different genus	*Coutarea hexandra*
Sample 05	KF667949.1	KF667926.1	*Quina rosa*	*Remijia ferruginea Strychnos pseudoquina*	GBVG1679-11	*Strychnos erichsonii*	99.8	GBVX1684-13	*Strychnos potatorum*	99.81	Species confirmed	*Strychnos pseudoquina*
					TSA163-10	*Strychnos henningsii*	99.8	KF667947	*Strychnos pseudoquina*	100		
					KF667964	*Strychnos pseudoquina*	100					
Sample 06	KF667950.1	KF667927.1	*Quina rosa*	*Remijia ferruginea Strychnos pseudoquina*	GBVJ4437-11	*Remijia pedunculata*	99.41	GBVJ4434-11	*Remijia macronemia*	99.24	Different genus	*Ladenbergia spp.*
					GBVJ4006-11	*Ladenbergia macrocarpa*	99.6	GBVJ4435-11	*Remijia pacimonica*	99.24		
					GBVJ3490-11	*Cinchona pubescens*	99.6	GBVJ4439-11	*Remijia ulei*	99.24		
								GBVJ4008-11	*Ladenbergia oblongifolia*	99.62		
								GBVJ4005-11	*Ladenbergia macrocarpa*	99.62		
								GBVJ4004-11	*Ladenbergia carua*	99.62		
Sample 07[a]		KF667928.1	*Quina rosa*	*Remijia ferruginea Strychnos pseudoquina*				GBVW1316-13	*Lamanonia ternata*	100	Different family	*Lamanonia ternata*
Sample 10	KF667951.1	KF667929.1	*Quina*	*Remijia ferruginea Strychnos pseudoquina*	GBVG1679-11	*Strychnos erichsonii*	99.41	MHPAD1545-09	*Strychnos guianensis*	100	Genus confirmed	*Strychnos spp.*
					TSA163-10	*Strychnos henningsii*	99.41					
Sample 11[a]		KF667930.1	*Quina*	*Remijia ferruginea Strychnos pseudoquina*				GBVJ3569-11	*Coutarea hexandra*	99.81	Different genus	*Coutarea hexandra*
Sample 12	KF667952.1	KF667931.1	*Quina*	*Remijia ferruginea Strychnos pseudoquina*	GBVG1679-11	*Strychnos erichsonii*	100	GBVX1684-13	*Strychnos potatorum*	99.81	Species confirmed	*Strychnos pseudoquina*
					TSA163-10	*Strychnos henningsii*	100	KF667947	*Strychnos pseudoquina*	100		
					KF667964	*Strychnos pseudoquina*	100					
Sample 13	KF667953.1	KF667932.1	*Quina*	*Remijia ferruginea Strychnos pseudoquina*	GBVG586-11	*Aspidosperma marcgravianum*	99.21	GBVG587-11	*Aspidosperma marcgravianum*	99.62	Different genus	*Aspidosperma spp.*
Sample 14	KF667954.1	KF667933.1	*Quina*	*Remijia ferruginea Strychnos pseudoquina*	GBVG1679-11	*Strychnos erichsonii*	99.21	MHPAD1545-09	*Strychnos guianensis*	99.81	Genus confirmed	*Strychnos spp.*
					TSA163-10	*Strychnos henningsii*	99.21					
Sample 15	KF667955.1	KF667934.1	*Quina-cruzeiro*	*Remijia ferruginea Strychnos pseudoquina*	GBVG1679-11	*Strychnos erichsonii*	100	GBVX1684-13	*Strychnos potatorum*	99.81	Species confirmed	*Strychnos pseudoquina*
					TSA163-10	*Strychnos henningsii*	100	KF667947	*Strychnos pseudoquina*	100		
					KF667964	*Strychnos pseudoquina*	100					
Sample 16	KF667956.1	KF667935.1	*Quina-cruzeiro*	*Remijia ferruginea Strychnos pseudoquina*	GBVG1679-11	*Strychnos erichsonii*	99.41	MHPAD1545-09	*Strychnos guianensis*	100	Genus confirmed	*Strychnos spp.*
					TSA163-10	*Strychnos henningsii*	99.41					
Sample 17	KF667957.1	KF667936.1	*Quina-cruzeiro*	*Remijia ferruginea Strychnos pseudoquina*	GBVG1679-11	*Strychnos erichsonii*	99.21	MHPAD1545-09	*Strychnos guianensis*	100	Genus confirmed	*Strychnos spp.*
					TSA163-10	*Strychnos henningsii*	99.21					

Sample 18	KF667958.1	KF667937.1	*Quina*	*Remijia ferruginea Strychnos pseudoquina*	GBVT961-13	*Cissampelos andromorpha*	99.41	GBVJ079-11	*Cissampelos andromorpha*	99.81	Different genus	*Cissampelos spp.*
					GBVJ087-11	*Cissampelos pareira*	99.41	MHPAD435-08	*Cissampelos pareira*	99.81		
								GBVJ093-11	*Cissampelos tropaeolifolia*	99.81		
								GBVJ084-11	*Cissampelos owariensis*	99.81		
Sample 19	KF667959.1	KF667938.1	*Quina do campo*	*Remijia ferruginea Strychnos pseudoquina*	GBVG1679-11	*Strychnos erichsonii*	100	GBVX1684-13	*Strychnos potatorum*	99.81	Species confirmed	*Strychnos pseudoquina*
					TSA163-10	*Strychnos henningsii*	100	KF667947	*Strychnos pseudoquina*	100		
					KF667964	*Strychnos pseudoquina*	100					
Sample 20	KF667960.1	KF667939.1	*Quina*	*Remijia ferruginea Strychnos pseudoquina*	GBVN1945-11	*Bowdichia virgilioides*	99.6	CAATB399-11	*Bowdichia virgilioides*	99.62	Different genus	*Bowdichia spp.*
					GBVW644-13	*Bowdichia nitida*	99.6	GBVO2840-11	*Pericopsis mooniana*	99.62		
								GBVO2407-11	*Ormosia macrocalyx*	99.43		
								GBVO2395-11	*Ormosia amazonica*	99.43		
								MHPAC850-08	*Acosmium panamense*	99.43		
Sample 21[a]		KF667940.1	*Quina*	*Remijia ferruginea Strychnos pseudoquina*				MHPAD210-08	*Machaerium seemannii*	100	Different family	*Machaerium spp*
								MHPAD722-09	*Machaerium pittieri*	100		
Sample 22	KF667961.1	KF667941.1	*Quina*	*Remijia ferruginea Strychnos pseudoquina*	GBVG586-11	*Aspidosperma marcgravianum*	99.6	GBVG587-11	*Aspidosperma marcgravianum*	99.81	Different genus	*Aspidosperma spp.*
					GBVG594-11	*Aspidosperma triternatum*	98.81	GBVG593-11	*Aspidosperma triternatum*	99.62		
					GBVG581-11	*Aspidosperma cruentum*	98.81	GBVG582-11	*Aspidosperma cruentum*	99.62		
					GBVG591-11	*Aspidosperma spruceanum*	99.01	GBVG590-11	*Aspidosperma spruceanum*	99.62		
Sample 23[a]		KF667942.1	*Quina rosa*	*Remijia ferruginea Strychnos pseudoquina*				GBVG870-11	*Lacmellea panamensis*	99.24	Different family	Lacmellea spp.
								GBVG865-11	*Lacmellea aculeata*	99.24		
Sample 24	KF667962.1	KF667943.1	*Quina*	*Remijia ferruginea Strychnos pseudoquina*	GBVG1679-11	*Strychnos erichsonii*	100	GBVX1684-13	*Strychnos potatorum*	99.81	Species confirmed	*Strychnos pseudoquina*
					TSA163-10	*Strychnos henningsii*	100	KF667947	*Strychnos pseudoquina*	100		
					KF667964	*Strychnos pseudoquina*	100					
Sample 26[a]		KF667944.1	*Quina-cruzeiro*	*Remijia ferruginea Strychnos pseudoquina*				GBVG590-11	*Aspidosperma spruceanum*	99.62	Different family	*Aspidosperma spp.*
								GBVG593-11	*Aspidosperma triternatum*	99.62		
								GBVG583-11	*Aspidosperma cruentum*	99.62		
Sample 28[a]		KF667945.1	*Quina rosa*	*Remijia ferruginea Strychnos pseudoquina*				GBVW1314-13	*Geissois hirsuta*	99.62	Different family	*Lamanonia spp.*
								GBVW1316-13	*Lamanonia ternata*	99.62		

[a] Samples that were only amplified for one of the two markers.
[b] Vernacular name in latinized Arabic.
[c] Putative scientific name based on vernacular names.

Biotechnology. The DNA extraction protocol was successful for 82.14% of the samples (23 of 28). The five samples for which the extraction was not effective are described as "not studied" (Table 3). The amplification of *matK* yielded PCR products for 65.22% of the samples, whereas the amplification of *rbcL* was successful for 100% of the samples. All of the PCR products that were generated were sequenced. The barcode regions of the *rbcL* and *matk* genes that were obtained for each sample were compared with the sequences available in the Barcode of Life Database – BOLD, GenBank, as well as with the sequences obtained in the barcode development stage of this work that were deposited on GenBank. For the majority of the samples, both *rbcL* and *matK* regions could be used in the phylogenetic analysis (Figs. 1S and 2S). For these samples, the information given by the analysis of the two regions were concatenated in only one phylogenetic tree (Fig. 3S), which was used to correctly identify the samples. For the samples 01, 04, 07, 11 21, 23, 26 and 28, however, the *matK* amplification was not successful, and only the *rbcL* region could be used (Fig. 2S).

The molecular analysis showed that the most commercialized samples are species of *Strychnos* (Loganiaceae, 43.5%), mainly *Strychnos. pseudoquina* A. St.-Hil (Table 3). The barks of *S. pseudoquina* were intensively used in the past as febrifuge and antimalarial in the past, and due this wide use in both traditional and conventional medicine, it was included as official medicine in the first edition of the Brazilian Pharmacopoeia, published in 1926 (Brandão et al., 2008). Barks of *Strychnos* spp. are rich in alkaloids, such as nordihydrotoxiferine, which has been associated with the bitter taste and the febrifuge activity (Bonamin et al., 2011; Frederich et al., 1999; Quetin-Leclercq et al., 1990). All the reference samples of *S. pseudoquina* showed the presence of alkaloids and tannins, while flavonoids were absent. However, three commercialized samples (15, 19 and 24), despite be identified as *Strychnos pseudoquina*, have not shown alkaloids in their composition. This result can be a consequence of the chemical variability of each individual, from which the bark was taken, or the degradation of the substances due bad conservation on the market. The result can also signalize a worse situation in which the barks of the plants have been extracted from young trees, when the substances are absents or in very low concentrations. Other species of *Strychnos* have been also sold as showed in Table 3, and one of them (14) has alkaloids and was active as antiplasmodial. Many *Strychnos* spp. are native from Cerrado, an ecosystem found in the surroundings of Belo Horizonte (Calábria et al., 2008; Rodrigues and Carvalho, 2007) and this proximity certainly contributes for the availability of different species on the market.

Species of Rubiaceae had sparked intense interest as substitute of the Peruvian quina, due its similarities with the *Cinchona* and among the commercialized samples, four were identified as Rubiaceae. Three samples are from *Coutarea hexandra* (Jacq.) K. Schum (Table 3), a species described in traditional medicine as a substitute for the true quina (Cosenza et al., 2013). Many studies have shown the potential of *Coutarea latiflora* (Sessé & Moc.) ex DC. (=*Hintonia latiflora* (Sessé &Moc.) Bullok) as an antimalarial, and this activity was correlated to the presence of alkaloids, with

Table 3
Species of quina identified by molecular analysis, results of phytochemical and Antiplasmodial studies.

Genetic identification	Code	Popular names	Phytochemical analysis			Antiplasmodial activity in vitro
Family and botanical names	Samples		Alkaloids	Flavonoids	Tannins	IC ≤ 50 μg/mL[a]
LOGANIACEAE						
Strychnos pseudoquina A. St. Hil.	03	Quina	+	–	+	Inactive
Strychnospseudoquina A. St. Hil.	05	Quina rosa	+	–	+	Inactive
Strychnos pseudoquina A. St. Hil.	12	Quina	+	–	+	Inactive
Strychnos pseudoquina A. St. Hil.	15	Quina-cruzeiro	–	–	+	Inactive
Strychnos pseudoquina A. St. Hil.	19	Quina do campo	–	–	+	Inactive
Strychnos pseudoquina A. St. Hil.	24	Quina	–	–	+	Inactive
Strychnos spp.	10	Quina	–	–	+	Inactive
Strychnos spp.	14	Quina	+	–	+	17 mg/mL
Strychnos spp.	16	Quina-cruzeiro	–	+	+	Inactive
Strychnos spp.	17	Quina-cruzeiro	–	–	+	Inactive
APOCYNACEAE						
Aspidosperma spp.	13	Quina	–	+	+	Inactive
Aspidosperma spp.	22	Quina	–	–	+	Inactive
Aspidosperma spp.	26	Quina-cruzeiro	+	–	+	20.7 mg/mL
Lacmellea spp.	23	Quina rosa	–	–	+	Inactive
RUBIACEAE					+	
Coutarea hexandra (Jacq.) K. Schum.	04	Quina-cruzeiro	+	–	+	Inactive
Coutarea. hexandra (Jacq.) K. Schum.	11	Quina	+	–	+	9.3 mg/mL
Coutarea. hexandra (Jacq.) K. Schum.	01	Quina	+	–	+	Inactive
Ladenbergia spp.	06	Quina rosa	–	–	+	Inactive
CUNONIACEAE					+	
Lamanonia ternata Vell.	07	Quina rosa	–	–	+	Inactive
Lamanonia spp.	28	Quina rosa	–	–	+	Inactive
FABACEAE					+	
Bowdichia spp.	20	Quina	–	–	+	Inactive
Machaerium spp.	21	Quina	–	–	+	Inactive
MENISPERMACEAE						
Cissampelos spp.	18	Quina	+	–	+	Inactive
Not studied	02	Quina	–	+	+	Inactive
Not studied	08	Quina rosa	+	–	+	Inactive
Not studied	09	Quina preta	–	+	+	Inactive
Not studied	25	Quina	+	+	+	Inactive
Not studied	27	Quina	–	–	+	Inactive

[a] Chloroquine 0.1 mg/mL.

structures very similar to quinine (Argotte-Ramos et al., 2006). In our study, alkaloids were detected in extracts from all *Coutarea hexandra* samples, but only one was active as antimalarial. Some *Ladenbergia* (Rubiaceae) have been described in the historical bibliography as a Brazilian quina, and one sample was identified as belonging to this genus. Barks of *Ladenbergia hexandra* Klotzch. and *Ladenbergia oblongifolia* (Humb. ex Mutis) L. Andersson have alkaloids with structures very similar to quinine (Okunade et al., 2001), but the identified sample did not show the presence of these substances. Species from *Aspidosperma* (Apocynaceae) have indole alkaloids with promising antiplasmodial activity (Dolabela et al., 2012). In our analysis, three samples were identified as *Aspidosperma* spp. by the molecular method, but only one of them showed the presence of alkaloids and was active as an antimalarial (Table 3). In the historical bibliography, only *Aspidosperma utillis* is described as a Brazilian quina (Vellozo, 1799) and efforts have been done in order to identify this active species. The results of chemical and biological essays show the complete lack of standardization in samples sold in the market and the consequences in public health are unknown.

Barks of *Remijia ferruginea* (A. St.-Hil.)(Rubiaceae) were the most frequently used Brazilian quina in traditional medicine (Cosenza et al., 2013) and as observed for *Strychnos pseudoquina*, they were also included in the 1st edition of the Brazilian Pharmacopoeia (Brandão et al., 2008). Despite this importance, curiously, none of the studied samples was identified as *Remijia ferruginea* or any other *Remijia* species. Wild populations of *Remijia ferruginea* can be also easily found in the surroundings of Belo Horizonte, and its absence in the market shown that changes in the knowledge and use of native medicinal plants are in course. The loss of traditional knowledge, specifically about plants named quina, can be also observed in recent ethnobotanical studies, including with descendants of Amerindians. Among several plants recorded in these studies, none Brazilian quina was mentioned (Borges and Peixoto, 2009; Britto and Senna-Valle, 2011, 2012; Fonseca-Kruel and Peixoto, 2004; Miranda and Hanazaki, 2008). It is currently established knowledge and use of native species that can subside their conservation (Crepaldi and Peixoto, 2010; Ibrahim et al., 2013; Joppa et al., 2013; Laersen and Olsen, 2007). It is estimated that 80% of the Brazilian population lives in urban centers, and in these areas, traditional medicine is practiced only in the popular markets. Strategies for stimulating the valorization of American medicinal plants in those places can contribute for their better use and conservation. On the other hand, there is a preoccupying risk that the herbal market poses a threat to biodiversity through overharvesting of raw materials (Moyo et al., 2013; Taylor, 2008; WHO, 2007). In the case of quina, for example, the extraction of barks leads to irreversible damage to the plant, and this predation certainly contributes for their extinction. It is necessary to establish strategies for improving the quality and the conservation of the native medicinal species sold in the popular markets and the integrate approach can be a useful tool for detecting frauds and excessive and inappropriate harvesting.

The worst situation observed in this study was the high proportion of samples (26%) identified as species that have none relation with plants named as quina from Brazilian traditional medicine (Table 3). One of them is the sample 23, identified as *Lacmellea* spp. (Apocynaceae). Accordingly the website listing the flora of Brazil (⟨www.floradobrasil.jbrj.br⟩), *Lacmellea* species are only found in north of South America (Milliken, 1997). This situation highlights a question: why species from the local ecosystems, such as *Remijia. ferruginea*, is not found in the market, while a plant that grows five thousand kilometers away is? Worse still is the observation that the barks of species from other families as Cunoniaceae (*Lamanonia* spp., samples 07 and 28), Fabaceae (*Bowdichia* sp., 20, and *Machaerium* sp., 21) and Menispermaceae (*Cissampelos* sp., 18) have been also sold as a Brazilian quina. None of these species is described as having utility as antimalarial and to treat intermittent fever in traditional medicine. Indeed, except from Menispermaceae, species from these botanical families have no alkaloids, flavonoids and other substances that could confer bitter taste. The situation is worsened when we consider that among them can be endangered species as *Bowdichia nitida* and *Machaerium ternate* (Martinelli and Moraes, 2013). The reasons why the collectors extracted these plants from wild, and sell them as a Brazilian quina, are obscure. These results show a worse situation where wild populations of other native species, that have any relation to the traditional use, are being pressured by the commerce in popular market.

Despite being considered a local level study, our results provide data that can complement international studies. We show that an integrated approach is important to authenticate the species of medicinal plants sold in the markets, as well as the conditions on which they were collected and are commercialized. The application of the methods can contribute to improve the quality of plants used in traditional medicine and promote the conservation of species from biodiversity, including those endangered.

Acknowledgment

To CNPq (563563/2010-9 and 563311/2010-0), FAPEMIG (PPM-2010) and PRPq-UFMG (PRPq-2014) for financial support, and C. Fagg and J. Paula-Souza for taxonomic identification of the reference samples.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jep.2014.06.040.

References

Andrade-Neto, V.F., Pohlit, A.M., Pinto, A.C.S., Silva, E.C.C., Nogueira, K.L., Melo, M.R. S., Henrique, M.C., Amorim, R.C.N., Silva, L.R.F., Costa, M.R.F., Nunomura, R.C.S., Nunomura, S.N., Alecrim, W.D., Alecrim, M.G.C., Chaves, F.C.M., Vieira, P.P.R., 2007. in vitro inhibition of Plasmodium falciparum by substances isolated from Amazonian antimalarial plants. Memorias do Instituto Oswaldo Cruz 102 (3), 359–366.

Argotte-Ramos, R., Ramírez-Avila, G., Rodríguez-Gutiérrez, M.C., Ovilla-Muñoz, M., Lanz-Mandonza, H., Rodríguez, M.H., González-Cortazar, M., Alvarez, L., 2006. Antimalarial 4-Phenylcoumarins from the Stem Bark of Hintonia latiflora. Journal of Natural Products 69, 1442–1444.

Bonamin, F., Moraes, T.M., Kushima, H., Silva, M.A., Rozza, A.L., Pellizzon, C.H., Bauab, T.M., Rocha, L.R., Vilegas, W., Hiruma-Lima, C.A., 2011. Can a Strychnos species be used as antiulcer agent? Ulcer healing action from alkaloid fraction of *Strychnos pseudoquina* St. Hil. (Loganiaceae). Journal of Ethnopharmacology 138 (1), 47–52.

Borges, R., Peixoto, A.L., 2009. Conhecimento e uso de plantas em uma comunidade caiçara do litoral sul do Estado do Rio de Janeiro, Brasil. [Knowledge and plant use in a caiçara community in the south coast of Rio de Janeiro state]. Acta Botanica Brasilica 23, 769–779.

Brandão, M.G.L., Cosenza, G.P., Pereira, F.L., Vasconcelos, A.S., Fagg, C.W., 2013. Changes in the trade in native medicinal plants in Brazilian public markets. Environmetal Monitoring and Assessment 185, 7013–7023.

Brandão, M.G.L., Cosenza, G.P., Stanislau, A.M., Fernandes, G.W., 2010. Influence of Brazilian herbal regulations on the use and conservation of native medicinal plants. Environmental Monitoring and Assessment 164, 369–377.

Brandão, M.G.L., Zanetti, N.N.S., Oliveira, P., Grael, C.F., Santos, A.C., Montemór, R.L. M., 2008. Brazilian medicinal plants described by 19th century European naturalists and in Official Pharmacopeia. Journal of Ethnopharmacology 120 (2), 141–148.

Britto, M.R., Senna-Valle, L., 2011. Plantas medicinais utilizadas na comunidade caiçara da Praia do Sono, Paraty, Rio de Janeiro, Brasil. [Medicinal plants used in a caiçara community of the Praia do Sono, Paraty, Rio de Janeiro, Brazil]. Acta Botanica Brasilica 25, 363–372.

Britto, M.R., Senna-Valle, L., 2012. Diversity of plant knowledge in a "Caiçara" community from Brazilian Atlantic Forest Coast. Acta Botanica Brasilica 26, 735–747.

Calábria, L., Cuba, G.T., Hwang, S.M., Marra, J.C.F., Mendonça, M.F., Nascimento, R.C., Oliveira, M.R., Porto, J.P.M., Santos, D.F., Silva, B.L., Soares, T.F., Xavier, E.M., Damasceno, A.A., Milani, J.F., Rezende, C.H.A., Barbosa, A.A.A., Canabrava, H.A.N., 2008. Levantamento etnobotânico e etnofarmacológico de plantas medicinais em Indianópolis, Minas Gerais, Brasil. [Ethnobotanical and ethnopharmacological survey of medicinal plants in Indianópolis, Minas Gerais, Brazil]. Revista Brasileira de Plantas Medicinais 10 (1), 49–63.

Carvalho, L.H., Krettli, A.U., 1991. Antimalarial chemotherapy with natural products and chemically defined molecules. Memorias do Instituto Oswaldo Cruz 86 (Suppl. II), 181–184.

CBOL, 2009. Plant Working Group, 2009. A DNA barcode for land plants. Proceedings of the National Academy of Sciences USA 106 (31), 12794–12797.

Cosenza, G.P., Somavilla, N.S., Fagg, C.W., Brandão, M.G.L., 2013. Bitter plants used as substitute of Cinchona spp. (quina) in Brazilian traditional medicine. Journal of Ethnopharmacology 149 (3), 790–796.

Crepaldi, M.O., Peixoto, A.L., 2010. Use and knowledge of plants by Quilombolas as subsidies for conservation efforts in an area of Atlantic Forest in Espírito Santo State. Brazil Biodiversity and Conservation 19, 37–60.

Dolabela, M.F., Oliveira, S.G., Peres, J.M., Nascimento, J.M., Póvoa, M.M., Oliveira, A. B., 2012. in vitro antimalarial activity of six Aspidosperma species from the state of Minas Gerais (Brazil). Anais da Academia Brasileira de Ciencias 84 (4), 899–910.

Ferrão, J.E.M., 2005. A Aventura das plantas e os descobrimentos portugueses. [Plant adventures and portuguese discoveries], 3rd edition Fundação para a Ciência e a Tecnologia, Lisboa.

Fonseca-Kruel, V.S., Peixoto, A.L., 2004. Etnobotânica na reserva extrativista Marinha de Arraial do Cabo, RJ, Brasil. [Ethnobotany of the extrative reserve Marinha de Arraial do Cabo, RJ, Brasil.]. Acta Botanica Brasilica 18, 177–190.

Frederich, M., Hayette, M.P., Tits, M., De Mol, P., Angenot, L., 1999. in vitro activities of Strychnos alkaloids and extracts against *Plasmodium falciparum*. Antimicrobial Agents and Chemotherapy 43, 2328–2331.

Gao, T., Sun, Z., Yao, H., Song, J., Zhu, Y., Ma, X., Chen, S., 2011. Identification of Fabaceae plants using the DNA barcode matK. Planta Medica 77 (1), 92–94.

Giulietti, A.M., Harley, R.M., Queiroz, L.P., Wanderley, M.G.L., Berg, C.V.D., 2005. Biodiversidade e conservação das plantas no Brasil [Biodiversity and conservation of plants in Brazil]. Megadiversidade 1, 52–61.

He, J., Wong, K.L., Shaw, P.C., Wang, H., Li, D.Z., 2010. Identification of the medicinal plants in *Aconitum L.* by DNA barcoding technique. Planta Medica 76 (14), 1622–1628.

Herbert, P.D., Cywinska, A., Ball, S.L., de Waard, J.R., 2003. Biological Identification through DNA Barcodes. Proceedings of Real Society of London 270 (1512), 313–321.

Ibrahim, M.A., Na, M., Oh, J., Schinazi, R.F., McBrayer, T.R., Whitaker, T., Doerksen, R. J., Newman, D.J., Zachos, L.G., Hamann, M.T., 2013. Significance of endangered and threatened plant natural products in the control of human disease. Proceedings of National Academy of Sciences USA 110 (42), 16832–16837.

Joppa, L.N., Visconti, P., Jenkins, C.N., Pimm, S.L., 2013. Achieving the convention on biological diversity's goals for plant conservation. Science 341 (6150), 1100–1103.

Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16, 111–120.

Kool, A., de Boer, H.J., Krüger, A., Rydberg, A., Abbad, A., Björk, L., Martin, G., 2012. Molecular identification of commercialized medicinal plants in southern Morocco. PLoS One 7 (6), e39459.

Laersen, H.O., Olsen, C.S., 2007. Unsustainable collection and unfair trade?uncoveringand assessing assumptions regarding central himalayan medicinal plant conservation. Biodiversity and Conservation 16, 1679–1697.

Li, J.W.H., Vederas, J.C., 2009. Drug discovery and natural products: end of an era or endless frontier? Science 325 (5937), 161–165.

Martinelli, G., Moraes, M.S. 2013. Livro Vermelho da Flora do Brasil [Red book of Brazilian Flora], Rio de janeiro, Instituto de Pesquisas Jardim Botânico.

Melo, J.G., Amorim, E.L.C., Albuquerque, U.P., 2009. Native medicinal plants commercialized in Brazil, priorities for conservation. Environmental Monitoring and Assessment 156 (1–4), 567–580.

Milliken, W., 1997. Plants for Malaria, Plants for fever. The Royal Botanic Gardens, Kew, London p. 1997.

Miranda, T.M., Hanazaki, N., 2008. Conhecimento e uso de recursos vegetais de restinga por comunidades das Ilhas do Cardoso (SP) e de Santa Catarina (SC), Brasil. [Knowledge and use of plant resources of the restinga by communities in the Cardoso islands (SP) and Santa Catarina (SC), Brazil.]. Acta Botanica Brasilica 22, 203–215.

Mittermeier, R.A., Fonseca, G.A.B., Rylands, A.B., Brandon, K., 2005. A brief history of biodiversity conservation in Brazil. Conservation Biology 19 (3), 601–607.

Moyo, M., Aremu, A.O., Gruz, J., Subrtová, M., Szüčová, L., Doležal, K., Van Staden, J., 2013. Conservation strategy for Pelargonium sidoides DC: phenolic profile and pharmacological activity of acclimatized plants derived from tissue culture. Journal of Ethnopharmacology 2013149 (2), 557–561.

Newmaster, S.G., Grguric, M., Shanmughanandhan, D., Ramalingam, S., Ragupathy, S., 2013. DNA barcoding detects contamination and substitution in North American herbal products. BMC Medicine 11, 222–225.

Nogueira, R.C., Cerqueira, H.F., Soares, M.B.P., 2010. Patenting bioactive molecules from biodiversity: the Brazilian experience. Expert Opinion on Therapeutic Patents 20 (2), 1–13.

Okunade, A.L., Lewis, W.H., Elvin-Lewis, M.P., Casper, S.J., Goldberg, D.E., 2001. Cinchonicine-derived alkaloids from the bark of the peruvian *Ladenbergia oblongifolia*. Fitoterapia 72 (6), 717–719.

Paz-Soldán, C.E. 1941, La introducción de La quina em terapêutica [The introduction of quina therapeutics]. General Editora.

Peckolt, W.L. 1916. Monografia das falsas quinas brasileiras [Monography of the Brazilian false-quina]. Rio de Janeiro 1916.

Quetin-Leclercq, J., Angenot, L., Bisset, N.G., 1990. South American *Strychnos* species. Ethnobotany and alkaloid screening. Journal of Ethnopharmacology 28, 1–52.

Rodrigues, V.E.G., Carvalho, D.A., 2007. levantamento etnobotânico de plantas medicinais no domínio dos cerrados na região do Alto Rio Grande – Minas Gerais [Ethnobotanical survey of medicinal plants in the area of cerrados in the Rio Grande region – Minas Gerais]. Revista Brasileira de Plantas Medicinais 9 (2), 17–35.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4, 406–425.

Shanley, P., Luz, L., 2003. Eastern amazonian medicinal: marketing, use and implications of forest loss. BioScience 53 (6), 573–584.

Souza, A.H., 1951. Quinas e falsas Quinas [Quina and false-quina]. Tribuna Farmaceutica 8, 127.

Stewart, A., 2013. The Drunken Botanist, 1st edition Timberpress, London p. 2013.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28, 2731–2739.

Taylor, D.A., 2008. New yardstick for medicinal plant harvests. Environmental Health Perspectives 116 (1), 21–28.

Vellozo, F.M.C., 1799. Quinographia Portugueza. Ou Coleção de várias memórias sobre 22 espécies de quinas tendentes ao seu descobrimento nos vastos domínios do Brazil [Collection of several memories about 22 species of quinas found in the vast areas of Brazil]. Impressora da Santa Igreja Patriarcal, Lisboa.

Voeks, R.A., Leony, A., 2004. Forgetting the forest: assessing medicinal plant erosion in eastern Brazil. Economic Botany 58, 294–306.

Wagner, H., Bladt, S., 1996. Plant Drug Analysis. A Thin Layer Chromatography Atlas. Springer, Berlin.

World Health Organization, 2011. The world traditional medicines situation. Issues and Challenges, 3. WHO, Geneva, pp. 1–14.

World Health Organization, 2007. WHO guidelines on good agricultural and collection practices (GACP) for medicinal plants. WHO, Geneva p. 2007.

World Health Organization, 2005. Quality control methods for medicinal plant materials. WHO, Geneva.

Zuo, Y., Chen, Z., Kondo, K., Funamoto, T., Wen, J., Zhou, S., 2011. DNA barcoding of Panax species. Planta Medica 77 (2), 182–187.

3.2 – CAPÍTULO 2: Aplicação da técnica de DNA barcode associada ao perfil fitoquímico para a análise de qualidade de espécies vegetais aprovadas pela ANVISA e OMS para a produção de fitoterápicos

Plantas medicinais são utilizadas em todo o mundo, sendo muitas espécies validadas pela OMS e outros órgãos regulatórios para o uso como medicamentos, devido à sua validação científica ou vasto histórico de uso tradicional. No Brasil, a ANVISA é o órgão responsável pela fiscalização dos medicamentos fitoterápicos e uma lista das espécies vegetais aprovadas para a produção dessa classe de medicamentos se encontra disponível na RDC nº 26/2014 (13 de maio de 2014) (ANVISA, 2014). Grande parte das plantas medicinais aprovadas e presentes nessa lista tratam-se de espécies utilizadas em todo o mundo, sendo consideradas no Brasil como exóticas e/ou importadas. Por outro lado, raras são as espécies aprovadas pelos órgãos reguladores que compõem a biodiversidade brasileira. Nesse estudo, foram analisadas 257 amostras de drogas vegetais e produtos fitoterápicos distribuídas entre oito espécies constantes da lista da ANVISA. Cada amostra foi identificada por meio de DNA barcode e seus marcadores químicos analisados quali- e quantitativamente. Os resultados obtidos sugerem que a análise do DNA barcode é uma metodologia viável para a identificação de materiais botânicos usados na preparação de fitoterápicos. Porém sua eficiência varia de espécie para espécie, podendo não ser útil em alguns casos. Para algumas espécies, marcadores diferentes do Barcode oficial devem ser utilizados. A associação do DNA barcode com as análises químicas revelou ser uma ferramenta poderosa para a detecção de falhas na preparação desses produtos. Os resultados mostram que a inclusão de um passo de identificação molecular na identificação das espécies vegetais a serem utilizadas na produção de fitoterápicos pode levar à melhoria da qualidade desses medicamentos.

O artigo referente a este capítulo foi submetido para a revista PlosOne no dia 05 de janeiro de 2015 e encontra-se em análise pelos revisores.

Medicinal plants recommended by the World Health Organization: DNA barcode identification associated with chemical analyses guarantees their quality

Rafael Melo Palhares[1,2,3*], Marcela Gonçalves Drummond[2], Bruno dos Santos A.F. Brasil[4], Gustavo Pereira Cosenza[5], Maria das Graças Lins Brandão[5,6], Guilherme Oliveira[3]

[1]Programa de Pós-graduação em Genética, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

[2]Myleus Biotechnology Research Team, Belo Horizonte, Brazil

[3]Grupo de Genômica e Biologia Computacional, Centro de Pesquisa Rene Rachou, FIOCRUZ, Belo Horizonte, Brasil

[4]EMBRAPA Agroenergia, Brasilia, Brazil

[5]Laboratório de Farmacognosia, Faculdade de Farmácia, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil

[6]CEPLAMT, Museu de História Natural e Jardim Botânico & Faculdade de Farmácia, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil

***Corresponding author**

E-mail: rafael.melopalhares@gmail.com (RMP)

Abstract

Medicinal plants are used throughout the world, and the regulations defining their proper use, such as identification of the correct species and verification of the presence, purity and concentration of the required chemical compounds, are widely recognized. Herbal medicines are made from vegetal drugs, the processed products of medicinal species. These processed materials present a number of challenges in terms of botanical identification, and according to the World Health Organization (WHO), the use of incorrect species is a threat to consumer safety. The samples used in this study consisted of the dried leaves, flowers and roots of 257 samples from 8 distinct species approved by the WHO for the production of medicinal herbs and sold in Brazilian markets. Identification of the samples in this study using DNA barcoding (*matK*, *rbcL* and ITS2 regions) revealed that the level of substitutions may be as high as 71%. Using qualitative and quantitative chemical analyses, this study identified situations in which the correct species was being sold, but the chemical compounds were not present. Even more troubling, some samples identified as substitutions using DNA barcoding contained the chemical compounds from the correct species at the minimum required concentration. This last situation may lead to the use of unknown species or species whose safety for human consumption remains unknown. This study concludes that DNA barcoding should be used in a complementary manner for species identification with chemical analyses to detect and quantify the required chemical compounds, thus improving the quality of this class of medicines.

Introduction

The global market of products derived from plants is estimated at $83 billion US and continues to grow [1]. Furthermore, it is estimated that approximately 25% of modern drugs and as many as 60% of antitumor drugs [2] are derived from natural products [3]. According to the WHO, between 65% and 80% of the populations of developing countries currently use medicinal plants as remedies [1]. The development of new products from natural sources is also encouraged because it is estimated that of the 300,000 plant species that exist in the world, only 15% have been evaluated to determine their pharmacological potential [4]. Studies demonstrating the efficacy and importance of medicinal plants are being carried out worldwide in countries that span a wide range of developmental stages [5-9]. Due to the widespread use of medicinal plants, the WHO published the Monographs on Selected Medicinal Plants volumes 1 through 5 from 1999 to 2010; these volumes contain a list of species with recognized medicinal benefits and the accepted means to correctly use them [10-14]. In addition to following WHO recommendations, Brazil has its own agency that regulates the use of medicinal plants, the National Health Surveillance Agency (from the Portuguese ANVISA – Agência Nacional de Vigilância Sanitária). ANVISA also has its own list of approved species for manufacturing herbal medicines [15]. Although Brazil is rich in biodiversity and medicinal plants, most of these plants on this list are exotic species that were introduced to the country during the early phases of European colonization in the 1500s [16-18].

To guarantee the quality of herbal medicines, certain steps established in the Pharmacopoeias must be followed, including correct identification of the plant species, analysis of the purity and confirmation of the presence and minimum concentration of the active ingredients (chemical marker(s)) [19]. In this regard, one of the main challenges encountered in the herbal medicine industry is ensuring unequivocal species identification of the raw material that will be used to manufacture the herbal medicine. There are several plant identification techniques, but in many cases, the identification is based mainly on botanical analysis, that can be problematic due to the high phenotypic variation

among taxa, the commercialization of processed raw plant material and/or unidentifiable plant parts and the lack of highly trained professionals in plant taxonomy [20-22]. Furthermore, quality control for herbal drugs is currently performed according to a set of pharmaceutical analyses, beginning by direct observation of the morphological, sensory and microscopic characteristics of each type of plant material. If the identity of the plant part is verified, the sample is submitted for chemical characterization using chromatographic methods to verify the presence of specific substances in comparison with a chemical profile in the literature or found in standard samples [23-27]. Misidentification and substitutions are a reality, with confirmed reports from several countries [7, 28-30], including a recent study from our group in which we demonstrated substitutions of species of “quina” (*Cinchona* spp.) in Brazilian markets [31]. These issues are of high concern because they may cause fatalities among users [32, 33]. Under these conditions, DNA barcoding may be a powerful tool. The DNA barcode consists of one or more short, standardized DNA region(s) that can be used to identify a species [34]. and is a powerful tool that can be applied to address the problems in botanically identifying highly processed plant materials [35-38] in addition to other uses, such as the identification of endangered species and the use in forensic DNA researches [39, 40]. Since 2009 the Plant Working Group from the Barcode of Life project established that the official regions for DNA Barcodes of plants are *rbcL* and *matK* [41]. Despite this, those regions are not 100% efficient in discriminating plant species and other regions are used by different researchers to improve the efficiency of the official DNA Barcode [42-44].

Here we propose the use of DNA barcoding technology to identify the raw material used to manufacture herbal medicines. Along with the CBOL recommendations and based on previous studies, we evaluated the addition of the nuclear ITS2 region to the barcode core of *matK* and *rbcL* [20]. After the initial identification step, our group carried out chemical analyses to demonstrate the presence and concentration of the essential chemical compound of the herbal medicine. Our results indicate that DNA barcoding should be used as a screening step during the herbal medicine

manufacturing process, and only samples that are correctly identified should proceed to chemical validation. This proposed workflow would improve the safety, speed and reliability of this process.

Materials and Methods

Sample collection

A total of 257 samples from eight species that are recognized by the WHO and ANVISA as medicinal species and approved for use in the preparation of remedies were purchased in the Central Market (19° 55' 22.465" S 43° 56' 35.058" W) in the city of Belo Horizonte in the state of Minas Gerais, Brazil. Belo Horizonte's metropolitan region holds approximately 6 million inhabitants and possesses a large traditional popular market as well as several drugstores and pharmacies specialized in phytoterapy and medicinal plants. The collectors purchased the samples, as regular customers, from 20 stores, during the years of 2012 to 2014. Products sold as dried plant parts or as powdered tissues, either simply packed or encapsulated, were sampled. Samples were stored in an acclimatized and humidity free room before DNA extraction. The studied species, popular names and used parts were *Hamamelis virginiana* L. (Hamamelis - leaves), *Matricaria recutita* L. (Chamomile, flowers), *Maytenus ilicifolia* Mart. Ex Reiss (Espinheira Santa, leaves), *Mikania glomerata* Spreng. (Guaco, leaves), *Panax ginseng* C. A. Mey (Asian Ginseng, roots), *Passiflora incarnata* L. (passion flower, leaves), *Peumus boldus* Molina (Boldo-do-Chile, leaves) and *Valeriana officinalis* L. (Valerian, roots) (Table 1).

Characteristics of the samples

The acquired samples included, flowers, leaves and roots. The samples were collected in two forms, as the dried parts described above and as powdered tissues. No mixtures were analyzed due to limitations inherent to the Sanger sequencing method. In the laboratory, each sample was recorded and kept under uniform conditions in a climate-controlled room at DATAPLAMT (Aromatic, medicinal and poisonous center for data and sample storage at the Universidade Federal de Minas Gerais).

DNA extraction

DNA was extracted from the leaves, flowers and roots of the plants using the DNeasy plant mini kit (Qiagen, Venlo - Netherlands) with modifications. Approximately 20 mg of each sample was pulverized using a mortar at room temperature. The powder was mixed with 600 µL of buffer AP1 supplied with the kit and incubated at 65 ºC and 400 rpm for 1 hour in a heat block (Thermomixer compact; Eppendorf, Germany). After incubation, 230 µL of buffer AP2 from the DNeasy kit was added, and the samples were incubated on ice for 30 minutes. The later steps of the extraction were carried out following instructions from the manufacturer (DNeasy plant handbook, Qiagen, Venlo - Netherlands). After extraction, the DNA samples were visualized on a 1% agarose gel stained with GelRed (Biotium, California, USA). The 100-bp DNA standard from Invitrogen (California, USA) was used for the analysis of the genomic DNA. Eighteen samples did not present the total DNA band on the agarose gel and, consequently, did not yield any amplicon in the subsequent PCR reaction. These samples could not be analyzed as the correct species or a substitutions, leaving the final dataset with a total of 239 samples.

PCR and sequencing

DNA amplification was carried out using primers selected from the Royal Botanic Gardens Kew Phase 2 Protocols and Update on Plant DNA Barcoding as follows: for *matK,* forward 5’ - ACCCAGTCCATCTGGAAATCTTGGTTC - 3’ (primer 1R_KIM-f) and reverse 5’ - CGTACAGTACTTTTGTGTTTACGAG - 3’ (primer 3F_KIM-r); for *rbcL*, forward 5’ - ATGTCACCACAAACAGAGACTAAAGC - 3’ (primer rbcLa_f) and reverse 5’ - GAAACGGTCTCTCCAACGCAT - 3’ (primer rbcLa_jf634R); and for ITS2*,* forward 5' - ATGCGATACTTGGTGTGAAT - 3' (primer ITS-S2F) and reverse 5' - GACGCTTCTCCAGACTACAAT - 3' (primer ITS3R) (http://www.kew.org/barcoding/protocols.html).

The PCR reactions were performed using a final volume of 25 µL containing 2 U of AmpliTaq Gold polymerase in GeneAmp 106 PCR Buffer II (100 mM Tris-HCl, pH 8.3; 500 mM KCl), 1.5 mM

$MgCl_2$ (Applied Biosystems, Foster City, CA), 0.2 mM dNTPs, 0.6 µM of each primer, and 100 to 200 ng of DNA.

The amplification was carried out in a GeneAmp PCR System 9700 Thermocycler (Applied Biosystems, Foster City, CA) using the following conditions: *matk* - an initial denaturation step at 98 ºC for 2 minutes; followed by 40 cycles at 98 ºC for 10 seconds, 52 ºC for 30 seconds and 72 ºC for 40 seconds; with a final extension period at 72 ºC for 10 minutes; *rbcL* - an initial denaturation step at 95 ºC for 3 minutes; followed by 45 cycles at 94 ºC for 30 seconds, 50 ºC for 40 seconds and 72 ºC for 40 seconds; with a final extension period at 72 ºC for 5 minutes; ITS2 - an initial denaturation step at 95 ºC for 5 minutes; followed by 40 cycles at 94 ºC for 30 seconds, 56 ºC for 30 seconds and 72 ºC for 45 seconds; with a final extension period at 72 ºC for 10 minutes. After amplification, the DNA samples were visualized on a 1% agarose gel stained with GelRed (Biotium, California, USA). Six samples did not yield any amplicon in the subsequent PCR reaction. These samples could not be analyzed as the correct species or a substitutions, leaving the final dataset with a total of 233 samples. The sequencing reactions used 2 µL containing 10 pmol of the same amplification reaction primers. Bi-directional sequencing was performed by Myleus Biotechnology (Belo Horizonte, Brazil) using an ABI3130 automated sequencer (Applied Biosystems, Foster City, CA) with BigDye v3.1.

Data analysis

The obtained DNA sequences were edited using the SeqScape v2.7 software program (Applied Biosystems, Foster City, CA). Bases with a QV lower than 15 (i.e., a probability of error of 3.2%) were manually edited, and samples for which the entire sequence (or the majority of it) had a lower QV were discarded due to the high probability of error and/or impossibility of analyses [45]. Samples that amplified high-quality sequences from any one of the three genes (*rbcL*, *matK* or ITS2) were included in the analyses. The sequences produced in this work were submitted to GenBank (accession numbers KJ750965 through KJ751173 for *matK* sequences, KJ751175 through KJ751402 for *rbcL* sequences and KM519459 through KM519583 for ITS2 sequences). Some of the sequences for ITS2 (61

sequences/32,97% of the total ITS-2 dataset) had fewer than 200 base pairs and could therefore not be deposited in GenBank. Those sequences are available as supporting information (File S1).

The reference sequences used to identify the generated sequences were mined from the Barcode of Life Data Systems (BOLD) (http://v3.boldsystems.org/index.php/databases) for the *matK* and *rbcL* regions and from GenBank (http://www.ncbi.nlm.nih.gov/genbank/) for the ITS2 region. BOLD archives are today the more reliable databases regarding DNA barcodes for reference species, since the criteria for a researcher to deposit a sequence is carefully reviewed and the specimen must be taxonomically identified by an expert. Some of the criteria include the deposit of at least five specimens vouchers of the reference species, the personal information of the botanist that made the identification of the specimens and several metadata that brings more security as to the correct identification. Since the official DNA barcode regions chosen for plant are *matk* and *rbcL*, GenBank had to be used for the ITS2 region, but when possible the ITS2 region was mined from BOLD. The reference sequences included every species from the eight genera analyzed in this study. Every query sequence that did not group with one of these genera was submitted to a Plant Identification via BOLD and to a MEGABLAST search on GenBank; the genera returned from these identifications were added to the phylogenetic analyses. The phylogenetic analyses and tree assembly were performed using the neighbor-joining (K2P) statistical method [41] in MEGA 5.2.2 [46]. The query sequences were identified according to the reference sequences with which they formed a cluster with a 98% similarity cutoff. Samples that grouped with a genus other than the eight target genera were promptly classified as substitutions.

To better identify the samples grouped within the 8-genus set of this study, the barcode gap approach was used [34]. For the barcode gap, the pairwise distance for each of the 8 genera was calculated individually using MEGA 5.2.2. The result was then exported to PAST [47], and a frequency histogram was assembled. The barcode gap was calculated for each of the three genes individually, for *matK* and *rbcL* together, and for *matK*, *rbcL* and ITS2 together. For each calculation, a barcode gap was considered to exist if the frequency histogram showed a clear distinction between the intra-

and interspecific genetic variation. When this distinction was unclear, no barcode gap was said to exist. Samples that presented a genetic variation higher than the maximum intraspecific variation were considered to be substitutions, and samples that presented genetic variation lower than the maximum intraspecific variation could not be identified as either a substitution or the correct species.

Chemical analysis

The objective of these analyses was to verify the presence of chemical markers using thin layer chromatography (TLC) with silica gel plates (Merck Darmstadt, Ref 1.05.721). After TLC analysis, the concentration of the substances was determined using high-performance liquid chromatography (HPLC) or ultra-violet (UV) spectroscopy. The latter was performed only for a subset of the samples to demonstrate that the correctly identified samples may not have the minimum required concentration of the target chemical compounds and that samples identified as substitutions may have the chemical compounds at the minimum required concentration. Because each species has its own approved method for certification, the chemical analyses and the results interpretation followed the methods described on the American (*P. ginseng*), Brazilian (*M. ilicifolia* and *M. glomerata*) and British Pharmacopoeias (*H. virginiana, M. recutita, P. incarnata, P. boldus and V. officinalis*). Those methods are briefly detailed in Table 2.

Results

DNA barcoding efficiency

Among the 257 samples used in this study, the protocols for DNA extraction, PCR and sequencing worked for 209 (81.32%), 228 (88.72%) and 185 (71.98%) of the samples for the markers *matK*, *rbcL* and ITS2, respectively. These proportions varied greatly among the various species, with *M. ilicifolia* and *M. glomerata* yielding the best results and *V. officinalis* generating the worst ones (Figure 1). With the exception of *P. boldus*, the ITS2 marker had the fewest samples that passed

through the steps of DNA extraction, PCR and sequencing, whereas the *rbcL* region had the most samples passing through these steps (Figure 1).

In particular, the DNA barcoding protocol did not work properly for herbal medicines acquired from drugstores (*M. glomerata* sample 16, *P. ginseng* sample 07, *P. incarnata* sample 08, and *P. boldus* samples 14 and 17 through 20), with the exception of sample 08 from *P. incarnata*.

Samples that failed during the DNA barcoding protocol during the DNA extraction step, amplification step, or sequencing step were labeled as “No sequence” and were not considered in further analyses.

Molecular markers efficacy

The *matK*, *rbcL*, and ITS2 markers and their combinations achieved various levels of identification success for each of the eight medicinal species studied here (Figure 2). In many cases, identification at the species level was not possible for the species assayed in this work and with the markers used, considering the current amount of species reference sequences (DNA barcodes vouchers) deposited at BOLD and GenBank (Table S1) because the genetic diversity within the genus was not sufficient to correctly identify a given sample at the species level. Because most of the substitutions found here involved species from different genera or even families, this result did not negatively impact the substitution analyses of this study. When samples were grouped within one of the eight medicinal genera, a barcode gap analysis was applied (Table 3). In some of these cases, it was possible to reach a final conclusion regarding the species identification, e.g., samples from *Matricaria recutita*. However, in other cases, the identification remained inconclusive, again because the genetic variation within the genus was not high enough (lower than 1 %), even after applying the barcode gap.

Molecular identification and species substitution

The phylogenetic analyses applied to the sequences retrieved from the DNA barcoding methodology revealed that all eight analyzed medicinal species, with the exception of *M. glomerata*,

had samples that were substituted with other species, genera or even other families (Supporting Figures S1 through S40).

From the samples that passed through the DNA barcoding protocol, 42.06% belonged to the expected genus but could not be identified to the species level; these samples were therefore classified as “inconclusive” in terms of substitutions. The remaining samples were classified as either substitute (71.11%) or authentic (28.89%), depending on the concordance between the expected and observed species (Figure 3). The proportion of samples classified as substitutions varied greatly among the eight species. For example, 100% of the samples presented as *P. ginseng* were actually from the genus *Pfaffia*, a Brazilian ginseng, whereas only 3.45% of the samples presented as *P. boldus* were substitutions (Figure 3).

For *H. virginiana*, half of the samples (16) belonged to the genus *Hamamelis* (Hamamelidaceae), and one sample belonged to the same family but was from a genus that could not be defined. Five samples could not be identified, and the remaining ten samples were distributed among another seven different families. It is interesting to note the presence of samples identified as Brazilian native species, such as *Solanum* and *Lantana*, as well as the presence of other species that are also imported to Brazil, such as *Tilia*.

All of the samples from *M. recutita* (Asteraceae) corresponded to the correct genus, but twenty samples presented a certain level of genetic diversity for the marker *matK* (Supporting Figure S26). When the barcode gap analysis was applied, these samples were assigned to a species other than *M. recutita*. Despite these observations, those samples were not linked to any other species and their genetic diversity was found to be extremely low (lower than 0,01%).

Although some of the samples labeled as *M. ilicifolia* (Celastraceae) were found to belong to the genus *Maytenus*, the majority were identified at the family level (Fabaceae) as one of two species, *Zollernia ilicifolia* or *Lecointea peruviana,* and one sample was identified as the genus *Roupala* (Proteaceae), which includes species that are native to Brazil but morphologically distinct from *M. ilicifolia* and with no previous reports of use in folk medicine (Supporting Table S1).

Neither of the sequences for *M. glomerata* (Asteraceae) was successful as a tool able to identify substitution because it was impossible to distinguish between *M. glomerata* and *M. laevigata*.

In the case of *P. ginseng*, a species that originated in Asia and was imported to Brazil, most of the samples were identified as *Pfaffia spp.* (Amaranthaceae). This genus contains the species *Pfaffia glomerata*, a plant that is native to Brazil and popularly known as Brazilian ginseng. The only exception for this group was one sample that was identified only at the family level (Amaranthaceae) but could not be distinguished among the genera *Pfaffia*, *Hebanthe* and *Pseudoplantago*.

In the analyses of *P. incarnata* (Passifloraceae), two clear substitutions were found of the species *Senna alexandrina* (Fabaceae). All other samples belonged to the genus *Passiflora*.

Most of the samples of *Peumus* (Monimiaceae), a genus with only one species (*P. boldus,* http://www.theplantlist.org/browse/A/Monimiaceae/Peumus/), were identified as the correct species. One exception was identified as *Vernonia colorata* (Asteraceae) (Supporting Table S1).

For *V. officinalis*, the whole process of DNA extraction, amplification and sequencing did not work well and the sequences obtained were mostly low quality. From thirty-five samples, only nineteen (54,28%) could be analyzed using DNA barcoding. Of these, thirteen belonged to the genus *Valeriana* but could not be identified at the species level. Two samples that were identified only at the family level belonged to Asteraceae. One sample was identified as belonging to a different genus (*Cissampelos*). Two other samples were identified as different species: *Ageratum conyzoides* and *Stellaria vestita*. One sample could not be identified.

Chemical analysis

For most of the studied species, TLC, HPLC and UV analyses confirmed the molecular findings for samples identified as not being the true plant; many samples did not contain the expected chemical marker for the labeled medicinal species. In some cases (*H. virginiana*, *M. Recutita*, *M. ilicifolia* and *V. officinalis*), some substitutions showed a chromatography pattern resembling that of the correct species. In these cases, only molecular analysis made the correct identification possible. For *P.*

ginseng, all samples were negative for the expected chemical marker. However, all samples labeled as *M. recutita* and *M. glomerata* contained the expected chemical marker (Figure 4).

The simple presence of the chemical markers is not sufficient to validate an herbal medicine preparation, but it is mandatory that a minimal concentration of the chemical marker is present. As expected, the samples that showed negative results via TLC also showed negative results via HPLC or UV. However, for some samples that were positive via TLC, the chemical marker was not present at the minimum concentration required for validation. This finding was true for samples from *M. recutita*, *M. glomerata P. incarnata* and *V. officinalis* (Table 4).

Molecular and chemical comparison

In some cases, samples that were identified as substitutions using molecular analysis actually did contain the expected chemical marker from the labeled species. That was the case for samples from *H. virginiana*, *M. recutita* and *M. ilicifolia* (Figure 4). On the other hand, every sample that matched the labeled species according to molecular identification was also positive on the TLC analyses (Figure 4).

During the final step of concentration analyses, HPLC or UV, two interesting points arose. First, the presence of the correct chemical marker(s) in a sample does not mean that the sample contained the minimum concentration required. This result was observed for samples of *M. recutita*, *M. glomerata* and *P. incarnata* (Table 4). Second, some samples that were identified as substitutions using DNA barcoding but contained the expected chemical marker from the medicinal species also presented the minimum concentration required for validation on HPLC or UV (Table 4). This result was observed for samples from *H. virginiana*, *M. ilicifolia* and *V. officinalis*.

Overall, *V. officinalis* was the most difficult species to work with during these analyses. The medicinal part of *V. officinalis* plants is the roots. *V. officinalis* root cells contain a light brown resin [10] that was most likely responsible for the unsatisfactory results of the genetic analyses because it completely inhibited the PCR or generated problems during the amplification process. Modifications

made to the protocols to attempt to resolve this problem were not effective. Only nine samples were positive according to TLC, and of the samples submitted to HPLC, only one (sample 29) met the minimum required concentration. Curiously, this sample was identified from DNA barcoding as belonging to the *Cissampelos* genus. (Table 4).

Discussion

Plants used to prepare herbal medicines are marketed as crude drugs, and the quality of these materials is currently verified by a set of botanical, physicochemical and chemical analyses that have been established by Pharmacopoeias and other official compendia [48]. Those methods, however, are not completely reliable for species identification, and several studies have revealed species substitutions [19, 22, 31, 49].

DNA-based methods, such as the use of specific DNA sequences as markers for species identification, are used in a range of field, including agriculture [50-52] and zootechny [53-55], and comprise various methods, such as RAPF, AFLP, PCR-DGGE, real-time PCR and sequencing-based systems, such as SSR [50, 56-59]. Choosing the most appropriate method depends on several factors, including the focus of the study [60]. However, the availability of a variety of methods and approaches can also hamper research; the lack of standardization and universality decreases the reproducibility of studies. However, the proposed goal of the DNA barcode project [34] to catalogue universal markers for all life on Earth has the potential to unify DNA-based methods used for species identification.

Using common sets of primers, databases and standards to catalogue species by research groups all around the world increases the level of reliability and the number of species available for study (which has reached the greatest level ever achieved by the scientific community) while also making it possible to identify an ever-growing number of species. The definition of an official DNA barcode for plants was a crucial step, and the sequences chosen have already proven themselves to be of great value [9, 60, 61]. The discovery of universal primers for the DNA barcode would be the perfect scenario, but this goal may not be achieved. Small nuances in different families, orders and

species are responsible for different levels of amplification and in some cases the use of different primers might be the best strategy to follow to make the amplification and sequencing more efficient.

Processed samples, such as the ones analyzed on this study, are often hard-working, since the isolation of good-quality DNA may be difficult to achieve [62]. Even though we were able to analyze the majority of the samples (233 from 257) using at least one of the three markers, better ways to work with processed samples are becoming available and will be applied in future studies [43]. An example is the DNA mini-Barcode, based on the analyses of smaller regions. A DNA mini-barcode for *rbcL* is already available [63].

This study demonstrated that it is not always necessary to work with both sequences *matK* and *rbcL* when the purpose of the study is not to catalogue new species but rather to identify species from a collection of samples; this study also demonstrated that the DNA barcode approach has limitations. For all of the samples, the use of the *rbcL* and *matk* sequences together only improved species identification in two cases, one for *Hamamelis* (sample 16) and one for *Peumus* (sample 6); in both of these cases, the samples could be identified to the species level only when the two markers were used together (Supporting Figures S4 and S34).

The use of the DNA barcoding technology enabled us to detect several substitutions among the analyzed samples. Most substitutions involved species from different genera (or even a different family) than those of the expected medicinal species. When analyzing multiple species within the same genus, *matK* and *rbcL* were only rarely able to correctly identify the samples. That was the case for samples belonging to some of the analyzed species. For example, the markers could not distinguish between *M. glomerata* and *M. laevigata*. Both species are used in folk medicine in Brazil and have the same geographical distribution and several morphological and chemical similarities. For these reasons, it is believed that *M. laevigata*, which is not included in the ANVISA list of approved species for herbal medicines, is frequently used as a substitute for *M. glomerata* [64]. For *M. ilicifolia*, most of the samples belonged to *Zollernia ilicifolia* or *Lecointea peruviana*. These species share similar morphology and like *M. ilicifolia*, belong to the clade Lecointea, together with the closely related genera

Exostyles, *Harleyodendron* and *Holocalyx* [65, 66]. Most of the *P. incarnata* samples belonged to the genus *Passiflora* but could not be identified at the species level. The *matK* region showed promising results for differentiating species within the genus *Passiflora* (Supporting Figure S26), but our analysis was ultimately unsuccessful because none of the databases contained this sequence for *P. incarnata*. Brazil is one of the greatest producers of *Passiflora* species for food [67], and it is likely that some of that production ends up being marketed as herbal medicine.

The difficulty in differentiating closely related species is supported by the fact that the methodologies used to perform distance-based species discriminations based on DNA barcodes are still being worked out [68, 69]. Furthermore, the difficulty in identifying closely related species is especially pronounced in plants [70]. For this reason, the barcode sequences were only recently defined, and the search for better loci continues [43, 69]. In this study, we attempted to use the ITS2 region to improve the accuracy of species identification. However, our attempt was not successful, primarily due to difficulties encountered in working with the sequence and the fact that it did not add additional variability compared with analysis based on *matK* and *rbcL*.

Some of the substitutions that we identified, such as the genera *Solanum* and *Lantana* for *H. virginiana* or the genera *Ageratum* and *Cissampelos* for *V. officinalis*, are most likely a consequence of the ease of obtaining samples of the substitutes, which are native to Brazil. In fact, *Hamamelis* is native to North America, and *V. officinalis* is native to Europe; it is necessary to import both plants for use in Brazil. The same is also true for *P. ginseng*, but in this specific case, a mistake may have occurred because the Brazilian ginseng (*Pfaffia glomerata*) and the Asian ginseng (*P. ginseng*) are both known as ginseng. Another case of substitution due to popular name confusion may have occurred when the genus *Sorocea* (Moraceae), to which the species *Sorocea bonplandii* belongs, was used as a substituted for *M. ilicifolia*; *Sorocea bonplandii* has the same popular name as *M. ilicifolia* in Brazil (Espinheira santa) and a similar morphology [65]. Finally, a similar explanation might be responsible for the only substitution found for *P. boldus*. The genus *Vernonia* contains the species *V.*

condensata, which is known in Brazil as “Boldo baiano” and regularly used as a substitute for *P. boldus*, despite their complete lack of similarity [71].

Curiously, we also detected *Tilia* among the samples of *Hammamelis*. This plant does not occur in Brazil, and its presence in the market here indicates that substitutions are sometimes occurring outside Brazil, which may also be the case for the sample of *S. alexandrina* that was found among the samples sold as *P. incarnata*. This species is popularly used in certain countries (including Brazil) for constipation, but recent studies have revealed toxic effects in mouse models [72, 73].

The parallels between the genetic and chemical analyses proved that it is possible for a sample to pass quality control tests even if it does not belong to the correct species. This result was observed for samples that were identified as substitutions using DNA barcoding but exhibited similarity with the correct species according to TLC and contained concentrations of chemical markers that were above the required minimums (*H. virginiana* samples 08 and 17, *M. ilicifolia* sample 06, and *V. officinalis* sample 29). These results may be attributed to the specificity of the chemical markers; even though some of these chemicals substances used as markers, such as valerenic acid, are very specific, others (such as tannins) are common to a large variety of plants. However, these analyses also demonstrated that correct species identification is not sufficient because the active compound may not be present in the samples or may be below the minimum required concentration. Thus, when taking into account the results of DNA barcoding, TLC and HPLC or UV, the complementarity of the tests becomes clear.

In addition to the health implications of the correct use of the approved medicinal species, another factor that should be considered is the possible environmental impacts to these plants. It is estimated that one in every five plant species in the world is threatened. It has been suggested that the herbal market poses a threat to biodiversity through the over-harvesting of raw materials [31, 74-76]. In a previous study, our group demonstrated positive results regarding the inhibition of native species collection in the wild by pharmaceutical companies following the establishment of rules from the Brazilian Health Ministry [77]. The impact of the use of native materials sold in popular markets, however, is difficult to estimate because these materials are obtained from various suppliers and from

unmanaged forests. If the findings of this study are cross-referenced with the Official List of Endangered Species of the Brazilian Flora [78], the genera *Solanum* (one species), *Maytenus* (four species), *Mikania* (six species), *Pfaffia* (three species), *Passiflora* (five species), and *Vernonia* (fifteen species) are all represented, demonstrating that correct species identification is required to prevent the use of threatened species.

Conclusions

The present study showed a great number of species substitutions and mislabeling, demonstrating that the current surveillance methods are not being efficient to control he herbal medicine market. Also, we showed that the traditional methodologies of species identification using chemical analysis are, in the majority of cases, not adequate to correctly identify a plant species. Thus, we propose the use of DNA barcode as a powerful first screening step. Applying the DNA barcode technique to the quality control of herbal medicine production will make the process safer, more reliable, and cheaper because substitutions will be promptly discarded without requiring more expensive chemical analyses that are otherwise necessary.

Acknowledgements

This project was funded by CNPq (Conselho Nacional de Pesquisa – Brasil) (INCT 573899/2008-8), CAPES and FAPEMIG (Fundação de apoio à pesquisa do Estado de Minas Gerais Brasil) (INCT APQ-0084/08). MGD is grateful to CNPQ for the research fellowship provided (process number 300702/2012-4). The funding agencies played no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

References

1. WHO (World Health Organization) (2011) The World Traditional Medicines Situation, in Traditional medicines: Global Situation, Issues and Challenges. Geneva 3:1-14.

2. Brower V (2008) Back to nature: extinction of medicinal plants threatens drug discovery. J Natl Cancer Inst 100: 838-9.
3. Newman DJ, Cragg GM (2012) Natural products as sources of new drugs over the 30 years from 1981 to 2010. J Nat Prod 75: 311-35.
4. De Luca V, Salim V, Atsumi SM, Yu F (2012) Mining the biodiversity of plants: a revolution in the making. Science 336: 1658-61.
5. Aslanargun P, Cuvas O, Dikmen B, Aslan E ,Yuksel MU (2012) Passiflora incarnata Linneaus as an anxiolytic before spinal anesthesia. J Anesth 26: 39-44.
6. Butnariu M, Coradini CZ (2012) Evaluation of Biologically Active Compounds from Calendula officinalis Flowers using Spectrophotometry. Chem Cent J 6: 35.
7. Coghlan ML, Haile J, Houston J, Murray DC, White NE*, et al.* (2012) Deep sequencing of plant and animal DNA contained within traditional Chinese medicines reveals legality issues and health safety concerns. PLoS Genet 8: e1002657.
8. Gurib-Fakim A (2006) Medicinal plants: traditions of yesterday and drugs of tomorrow. Mol Aspects Med 27: 1-93.
9. Mankga LT, Yessoufou K, Moteetee AM, Daru BH ,van der Bank M (2013) Efficacy of the core DNA barcodes in identifying processed and poorly conserved plant materials commonly used in South African traditional medicine. Zookeys: 215-33.
10. WHO (World Health Organization) (1999) Monographs on selected medicinal plants Volume 1. Geneva, 295p.
11. WHO (World Health Organization) (2003) Monographs on selected medicinal plants Volume 2. Geneva, 375p.
12. WHO (World Health Organization) (2007) Monographs on selected medicinal plants Volume 3. Geneva, 390p.
13. WHO (World Health Organization) (2009) Monographs on selected medicinal plants Volume 4. Geneva, 456p.
14. WHO (World Health Organization) (2010) Monographs on selected medicinal plants Volume 5. Geneva, 452p.
15. ANVISA. Agência Nacional de Vigilância Sanitária (Brasil) (2008) Instrução Normativa nº 5 de 11 de Dezembro de 2008: Lista de medicamentos fitoterápicos de registro simplificado.
16. Begossi A, Hanazaki N, Tamashiro J (2002) Medicinal plants in the Atlantic forest (Brazil): knowledge, use and conservation. Hum Ecol 30: 19.

17. Shanley P, Luz L (2003) Eastern Amazonian Medicinals: Marketing, Use and Implications of Forest Loss. BioScience 53: 12.
18. Voeks R, Leony A (2004) Forgetting the Forest: Assessing Medicinal Plant Erosion in Eastern Brazil. Econ Bot 58: 13.
19. Brandao M, Cosenza GP, Pereira FL, Vasconcelos AS ,Fagg CW (2013) Changes in the trade in native medicinal plants in Brazilian public markets. Environ Monit Assess 185: 7013-23.
20. Chen S, Yao H, Han J, Liu C, Song J*, et al.* (2010) Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One 5: e8613.
21. Newmaster S, Ragupathy S, Janovec J (2009) A botanical renaissance: state-of-the-art DNA barcoding facilitates an Automated Identification Technology system for plants. Int J Comput Appl T 35: 1.
22. Sucher NJ, Carles MC (2008) Genome-based approaches to the authentication of medicinal plants. Planta Med 74: 603-23.
23. Eschricher W (1988) Pulver Atlas der Drogen des Deutschen Arzneibuches. Stuttgart: Gustav Fischer, 335p.
24. Langhammer L (1989) Bildatlas zur mikroscopischen Analytik pflanzlicher Arzneidrogen. Berlin: Walter de Gruyter, 241p.
25. Oliveira F, Akisue G, Akisue M (1991) Farmacognosia. São Paulo: Livraria Atheneu Editora, 421p.
26. Wagner H, Bladt S (1996) Plant drug analyses. A thin layer chromatography atlas: Springer, 368p.
27. WHO (World Health Organization) (1998) Quality control methods for medicinal plants materials. Geneva, 114p.
28. Guo X, Wang X, Su W, Zhang G ,Zhou R (2011) DNA barcodes for discriminating the medicinal plant Scutellaria baicalensis (Lamiaceae) and its adulterants. Biol Pharm Bull 34: 1198-203.
29. Heubl G (2010) New aspects of DNA-based authentication of Chinese medicinal plants by molecular biological techniques. Planta Med 76: 1963-74.
30. Newmaster SG, Grguric M, Shanmughanandhan D, Ramalingam S ,Ragupathy S (2013) DNA barcoding detects contamination and substitution in North American herbal products. BMC Med 11: 222.
31. Palhares RM, Drummond MG, Brasil BS, Krettli AU, Oliveira GC*, et al.* (2014) The use of an integrated molecular-, chemical- and biological-based approach for promoting the

better use and conservation of medicinal species: a case study of Brazilian quinas. J Ethnopharmacol 155: 815-22.

32. Bruni I, De Mattia F, Galimberti A, Galasso G, Banfi E*, et al.* (2010) Identification of poisonous plants by DNA barcoding approach. Int J Legal Med 124: 595-603.
33. WHO (World Health Organization) (2004) Medicinal plants – guidelines to promote patient safety and plant conservation for a U$ 60 billion industry. http://www.who.int/mediacentre/news/notes/2004/np3/en/.
34. Herbert P, Cywinska A, Ball S ,deWaard J (2003) Biological Identification through DNA Barcodes. Proc Biol Sci 270: 9.
35. Mahadani P, Ghosh S (2013) DNA Barcoding: A tool for species identification from herbal juices. DNA Barcodes: 4.
36. Seethapathy GS, Ganesh D, Santhosh Kumar JU, Senthilkumar U, Newmaster SG*, et al.* (2014) Assessing product adulteration in natural health products for laxative yielding plants, Cassia, Senna, and Chamaecrista, in Southern India using DNA barcoding. Int J Legal Med.
37. Stoeckle MY, Gamble CC, Kirpekar R, Young G, Ahmed S*, et al.* (2011) Commercial teas highlight plant DNA barcode identification successes and obstacles. Sci Rep 1: 42.
38. Wallace L, Boilard S, Eagle S, Spall J, Shokralla S*, et al.* (2012) DNA barcodes for everyday life: Routine authentication os natural health products. Food Res Int 49: 8.
39. Nithaniyal S, Newmaster SG, Ragupathy S, Krishnamoorthy D, Vassou SL*, et al.* (2014) DNA barcode authentication of wood samples of threatened and commercial timber trees within the tropical dry evergreen forest of India. PLoS One 9: e107669.
40. Verma SK, Goswami GK (2014) DNA evidence: current perspective and future challenges in India. Forensic Sci Int 241: 183-9.
41. Group CPW (2009) A DNA barcode for land plants. Proc Natl Acad Sci U S A 106: 12794-7.
42. Ferri G, Corradini B, Ferrari F, Santunione AL, Palazzoli F*, et al.* (2015) Forensic botany II, DNA barcode for land plants: Which markers after the international agreement? Forensic Sci Int Genet 15: 131-6.
43. Hollingsworth PM, Graham SW, Little DP (2011) Choosing and using a plant DNA barcode. PLoS One 6: e19254.
44. Purushothaman N, Newmaster S, Ragupathy S, Stalin N, Suresh D*, et al.* (2014) A tiered barcode authentication tool to differentiate medicinal *Cassia* species in India. Gen Mol Res 13: 10.

45. Prosdocimi F, Peixoto FC, Ortega JM (2004) Evaluation of window cohabitation of DNA sequencing errors and lowest PHRED quality values. Genet Mol Res 3: 483-92.
46. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, *et al.* (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731-9.
47. Hammer Ø, Harper D, Ryan P (2001) PAST: Paleontological Statistics software package for education and data analysis. Paleont Electron 4: 1.
48. WHO (World Health Organization) (2005) Quality control methods for medicinal plant materials, Geneva.
49. Zuccolotto T, Apel M, Rates S (1999) Avaliação da qualidade de produtos fitoterápicos comercializados em Porto Alegre- RS. Rev Inst Adolfo Lutz 58: 1.
50. Chuang H, Lur H, Hwu K, Chang M (2011) Authentication of domestic Taiwan rice varieties based on fingerprinting analysis of microsatellite DNA markers. Bot Stud 52: 13.
51. Doulaty B, Grassi F, Mohammadi A, Nazemieh A, De Mattia F, *et al.* (2007) The use of AFLP and morphological markers to study Iranian grapevine germplasm to avoid genetic erosion. J Hortic Sci 82: 8.
52. Grassi F, Labra M, Minuto L (2006) Molecular diversity in Ligurian local races of common beans (Phaseolus vulgaris L.). Plant Biosyst 140: 4.
53. Carvalho D, Seerig A, Melo D, de Souza A, Pimenta D (2008) Identificação molecular de peixes: o caso do Surubim (Pseudoplatystoma spp.). Rev Bras Reprod Anim 32: 5.
54. Drummond M, Brasil B, Dalsecco L, Brasil R, Teixeira L, *et al.* (2013) A versatile real-time PCR method to quantify bovine contamination in buffalo products. Food Control 29: 7.
55. Mane B, SK M, AKT (2009) Polymerase chain reaction assay for identification of chicken in meat products. Food Chem 116: 5.
56. Dalmacio LM, Angeles AK, Larcia LL, Balolong MP, Estacio RC (2011) Assessment of bacterial diversity in selected Philippine fermented food products through PCR-DGGE. Benef Microbes 2: 273-81.
57. Fajardo V, Gonzáles I, Rojas M, García T, Martín R (2010) A review of current PCR-based methodologies for the authentication of meats from game animal species. Trends Food Sci Tech 21: 14.
58. Kumar P, Gupta V, Misra A, Modi D ,Pndey B (2009) Potential of molecular markers in plant biotechnology. Plant Omics 2: 23.

59. Zheng XW, Yan Z, Han BZ, Zwietering MH, Samson RA*, et al.* (2012) Complex microbiota of a Chinese "Fen" liquor fermentation starter (Fen-Daqu), revealed by culture-dependent and culture-independent methods. Food Microbiol 31: 293-300.
60. Galimberti A, De Mattia F, Losa A, Bruni I, Federici S*, et al.* (2013) DNA barcoding as a new tool for food traceability. Food Res Int 50: 9.
61. Little DP, Knopf P, Schulz C (2013) DNA barcode identification of Podocarpaceae--the second largest conifer family. PLoS One 8: e81008.
62. Srivastava S, Mishra N (2009) Genetic Markers - A cutting-edge technology in herbal drug research. J Chem Pharm Res 1: 18.
63. Little DP (2014) A DNA mini-barcode for land plants. Mol Ecol Resour 14: 437-46.
64. Lima N, Biasi L (2002) Estaquia semilhenhosa e comparação de metabólitos secundários em Mikania glomerata Sprengel e Mikania laevigata. Sci Agric 1: 20.
65. Alberton M, de Souza E, Falkenberg D, Falkenberg M (2002) Identificação de marcadores cromatográficos de Zollernia ilicifolia e Sorocea bonplandii para o controle de qualidade de espinheira-santa. Rev Bras Farmacogno 12: 3.
66. Herendeen PS (1995) Phylogenetic relationships of the tribe Swartzieae. In M. Crisp and J. J. Doyle [eds.], Advances in legume systematics, Part 7, Phylogeny, Royal Botanic Gardens, Kew, Richmond, Surrey, UK. pp. 123-132.
67. Zeraik M, Pereira C, Zuin V, Yariwake J (2010) Maracujá: um alimento funcional? Rev Bras Farmacogno 20: 13.
68. DeSalle R, Egan MG, Siddall M (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding. Philos Trans R Soc Lond B Biol Sci 360: 1905-16.
69. Rubinoff D, Cameron S, Will K (2006) Are plant DNA barcodes a search for the Holy Grail? Trends Ecol Evol 21: 1-2.
70. Laiou A, Mandolini LA, Piredda R, Bellarosa R ,Simeone MC (2013) DNA barcoding as a complementary tool for conservation and valorisation of forest resources. Zookeys: 197-213.
71. da Silva JB, Temponi Vdos S, Fernandes FV, de Assis Dias Alves G, de Matos DM*, et al.* (2011) New approaches to clarify antinociceptive and anti-inflammatory effects of the ethanol extract from Vernonia condensata leaves. Int J Mol Sci 12: 8993-9008.
72. Hanlidou E, Karousou R, Kleftoyanni V, Kokkini S (2004) The herbal market of Thessaloniki (N Greece) and its relation to the ethnobotanical tradition. J Ethnopharmacol 91: 281-99.

73. Surh I, Brix A, French J, Collins B, Sanders J*, et al.* (2013) Toxicology and carcinogenesis study of senna in C3B6.129F1-Trp53 tm1Brd N12 haploinsufficient mice. Toxicol Pathol 41: 9.
74. Ibrahim MA, Na M, Oh J, Schinazi RF, McBrayer TR*, et al.* (2013) Significance of endangered and threatened plant natural products in the control of human disease. Proc Natl Acad Sci U S A 110: 16832-7.
75. Taylor DA (2008) New yardstick for medicinal plant harvests. Environ Health Perspect 116: 21-28.
76. WHO (World Health Organization) (2007) WHO guidelines on good agricultural and collection practices (GACP) for medicinal plants. Geneva.
77. Brandao MG, Cosenza GP, Stanislau AM ,Fernandes GW (2010) Influence of Brazilian herbal regulations on the use and conservation of native medicinal plants. Environ Monit Assess 164: 369-77.
78. Ministério do Meio Ambiente (2008) Instrução normativa de setembro de 2008 (Normative instruction of september of 2008).

Supporting information captions

Figure S1: Phylogenetic tree *Hammamelis virginiana matK*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 1.47656508 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 442 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 205 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S2: Phylogenetic tree *Hammamelis virginiana rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.67568534 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer

the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 370 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 215 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S3: Phylogenetic tree *Hammamelis virginiana* ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.71702123 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 207 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 13 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S4: Phylogenetic tree *Hammamelis virginiana matK* + *rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.86650001 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 189 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 620 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S5: Phylogenetic tree *Hammamelis virginiana matK* + *rbcL* + ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.51134035 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site.

The analysis involved 54 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 726 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S6: Phylogenetic tree *Matricaria recutita matK*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.00478244 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 36 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 631 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S7: Phylogenetic tree *Matricaria recutita rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.00211645 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 36 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 473 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S8: Phylogenetic tree *Matricaria recutita* ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.12287265 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 27 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All

positions containing gaps and missing data were eliminated. There were a total of 191 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S9: Phylogenetic tree *Matricaria recutita matK* + *rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.00363776 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 36 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 1104 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S10: Phylogenetic tree *Matricaria recutita matK* + *rbcL* + ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.00534820 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 24 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 1319 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S11: Phylogenetic tree *Maytenus ilicifolia matK*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.39435507 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 112 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 98 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S12: Phylogenetic tree *Maytenus ilicifolia rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.16105704 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 84 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 370 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S13: Phylogenetic tree *Maytenus ilicifolia* ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 1.85107254 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 77 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 100 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S14: Phylogenetic tree *Maytenus ilicifolia matK + rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.22908763 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 75 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 518 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S15: Phylogenetic tree *Maytenus ilicifolia matK + rbcL* + ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length

= 0.31908268 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 39 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 858 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S16: Phylogenetic tree *Mikania glomerata matK*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.00738890 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 46 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 679 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S17: Phylogenetic tree *Mikania glomerata rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.00755703 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 47 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 531 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S18: Phylogenetic tree *Mikania glomerata* ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.21453668 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with

branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 49 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 220 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S19: Phylogenetic tree *Mikania glomerata matK + rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.00165371 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 45 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 1210 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S20: Phylogenetic tree *Mikania glomerata matK + rbcL* + ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.02267872 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 41 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 1435 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S21: Phylogenetic tree *Panax ginseng matK*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.24952737 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite

Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 180 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 490 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S22: Phylogenetic tree *Panax ginseng rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.10182289 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 72 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 360 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S23: Phylogenetic tree *Panax ginseng* ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 1,22004815 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 234 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 78 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S24: Phylogenetic tree *Panax ginseng matK + rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.27861795 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 72 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All

positions containing gaps and missing data were eliminated. There were a total of 850 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S25: Phylogenetic tree *Panax ginseng matK* + *rbcL* + ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.29889235 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 65 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 935 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S26: Phylogenetic tree *Passiflora incarnata matK*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.83753081 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 191 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 530 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S27: Phylogenetic tree *Passiflora incarnata rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.64321259 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 234 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 512 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S28: Phylogenetic tree *Passiflora incarnata* ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 2.20614488 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 259 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 77 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S29: Phylogenetic tree *Passiflora incarnata matK + rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.46000144 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 64 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 1045 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S30: Phylogenetic tree *Passiflora incarnata matK* + *rbcL* + ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.49341320 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 46 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 1191 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S31: Phylogenetic tree *Peumus boldus matK*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.46390557

is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 53 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 421 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S32: Phylogenetic tree *Peumus boldus rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.13298911 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 53 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 502 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S33: Phylogenetic tree *Peumus boldus* ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.40028624 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 92 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 65 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S34: Phylogenetic tree *Peumus boldus matK + rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.22642831 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn

to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 48 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 928 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S35: Phylogenetic tree *Peumus boldus matK* + *rbcL* + ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.17636441 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 31 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 998 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S36: Phylogenetic tree *Valeriana officinalis matK*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 1.06223373 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 125 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 525 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S37: Phylogenetic tree *Valeriana officinalis rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.45263449 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter

method and are in the units of the number of base substitutions per site. The analysis involved 155 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 474 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S38: Phylogenetic tree *Valeriana officinalis* ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 1.70998580 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 31 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 50 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S39: Phylogenetic tree *Valeriana officinalis matK + rbcL*. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.74255654 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 113 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 1001 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Figure S40: Phylogenetic tree *Valeriana officinalis matK + rbcL* + ITS2. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.48881377 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 18 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All

positions containing gaps and missing data were eliminated. There were a total of 1121 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Tables

Table 1: Species analyzed in this study and their therapeutical recommendations.

Species	Recommended uses			Number of samples
H. virginiana L.	Topically for minor skin lesions, bruises and sprains, local inflammation of the skin and mucous membranes, hemorrhoids and varicose veins [11]			32
	Internal uses	**External uses**	**Inhalation**	
M. recutita L.	Symptomatic treatment of digestive ailments, treatment of restlessness and insomnia due to nervous disorders [10]	Inflammation and irritations of the skin and mucosa, including irritations and infections of the mouth and gums, and hemorrhoids [10]	Symptomatic relief on irritations of the respiratory tract due to common cold [10]	31
M. ilicifolia Mart. Ex Reiss	Treatment of dyspepsia, gastritis and gastroduodenal ulcer [15]			33
M. glomerata Sprengl.	Bronchodilatador and expectorant [15]			31
P. ginseng C. A. Mey	Prophylactic and restorative agent for enhancement of mental and physical capacities, in cases of weakness, exhaustion, tiredness, and loss of concentration, and during convalescence [10]			31
P. incarnata L.	Mild sedative for nervous restlessness, insomnia and anxiety. Treatment of gastrointestinal disorders of nervous origins [12]			30
P. boldus Molina	Treatment of functional dyspepsia and gastrointestinal disorders, cholagogue and choleretic [15]			34
V. officinalis L.	Mild sedative and sleep promoting agent. Often used as a milder alternative or a possible substitute for stronger synthetic sedatives in treatment of nervous excitation and anxiety-induced sleep disturbances [10]			35

Table 2: Conditions used for the chemical analyses

Species	TLC		HPLC	UV
	Mobile phase	Developer solution	Mobile phase	Diluent and absorbancy
H. virginiana	Formic acid anidrous : water : ethyl acetate (10:10:80)	Ferric chloride	-	Phosphomolybdotungistc 760 nm
M. recutita	Ethyl acetate : toluene (5:95)	Anisaldehyde	Phase A (Phosphoric acid : water) 0,5:99,5 Phase B (Phosphoric acid : acetonitrile) 0,5:99,5	-
M. ilicifolia	Ethyl acetate : Formic acid : water (90:5:5)	Vanillin sulfuric	Phase A (water : trifluoroacetic acid 0,05%) Phase B (Acetonitrile : trifluoroacetic acid 0,05%)	-
M. glomerata	Toluene : dicholomethane : acetone (45:25:30)	Ethanolic KOH (50%)	Methanol : water (47:53)	-
P. ginseng	Butyl alcohol : ethyl acetate : water (10:2,5:5)	Anisaldehyde in glacial acetic acid plus methanol	Phase A (water) Phase B (Acetonitrile : water) 4:1	-
P. incarnata	Anhydrous formic acid : water : methyl ethyl ketone : ethyl acetate (10:10:30:50)	Diphenyl boric acid amino ethyl ester in methanol plus macrogol in methanol	-	Methanol (10 volumes) : Glacial acetic acid (100 volumes) : Boric acid 25 g/L (10 mL) : oxalic acid (20 g/L) in formic acid anidrous 401nm
P. boldus	Diethylamine : metanol : toluene (10:10:80)	Potassium iodobismutate	Phase A (0,2 mL Diethylamine : 99,8 mL acetonitrile) Phase B (0,2 mL Diethylamine : 99,8 mL water)	-
V. officinalis	Glacial acetic acid : ethyl acetate : cyclohexane (2:38:60)	Anisaldehyde	Phase A (Acetonitrile R1 + Phosphoric acid solution 5 g/L) 20:80 Phase B (Phosphoric acid solution 5 g/L + Acetonitrile R1) 20:80	-

Table 3: Barcode gap analyses

SPECIES	MAXIMUM INTRASEPECIFIC DIVERGENCE				
	matK	*rbcL*	ITS2	*matK* + *rbcL*	*matK* + *rbcL* + ITS2
Hamamelis virginiana	0,026	X	0,019	0,026	0,033
Matricaria recutita	0,00042	0,00068	0,0058	0,00024	0,0014
Maytenus ilicifolia	X	0,016	X	0,028	X
Mikania glomerata	0,0017	0,0005	0,02	0,0005	0,0058
Panax ginseng	0,024	0,021	X	0,028	0,046
Passiflora incarnata	X	X	X	0,035	0,038
Peumus boldus			Not applicable*		
Valeriana officinalis	X	X	0,08	X	0,046

The numbers represent the maximum intra-specific divergence. Values above this number were considered as a different species.

X – The Barcode Gap was not calculated because there was no clear division between intra- and interspecific genetic divergence.

*The genus *Peumus* possess only one specie, which makes the Barcode Gap not applicable.

Table 4: Molecular identification versus TLC, HPLC and UV analyses

Species, minimal [ ] of chemical markers and method of dosage	Samples	Molecular identification	Chemical Markers TLC	Content
H. virginiana (3,0% of tannins/ UV)	01	*Hamamelis spp.*	Present	3,59%
	02	*Betula spp.*	Not Present	0,57%
	07	*Tilia spp.*	Not Present	0,23%
	08	*Solanum spp.*	Present	4,22%
	09	Unidentified	Not Present	0,61%
	12	*Hamamelis spp.*	Present	4,04%
	17	*Lantana spp.*	Present	4,43%
	22	*Persicaria spp.*	Not Present	0,84%
	23	*Sterculia urens*	Not Present	0,20%
	28	*Hamamelis spp.*	Present	4,13%
M. recutita (0,25% Apigenin 7-glucoside/ HPLC)	05	*Matricaria spp.*	Present	0,005%
	20	*Matricaria spp.*	Present	0,001%
	24	*Matricaria spp.*	Present	0,001%
	27	*Matricaria recutita*	Present	0,002%
M. ilicifolia (2,8 mg/g of Epicatechin, HPLC)	06	*Roupala spp.*	Present	26,32 mg/g
	10	*Maytenus spp.*	Present	79,80 mg/g
	22	*Sorocea affinis*	Not Present	-
	24	Fabaceae	Not Present	-
	28	*Maytenus spp.*	Present	107,44 mg/g
M. glomerata (0,1% of Cumarin, HPLC)	04	*Mikania spp.*	Present	0,038%
	15	*Mikania spp.*	Present	0,020%
	26	*Mikania spp.*	Present	0,011%
P. ginseng (0,2% ginsenoside Rg_1 and 0,1% ginsenoside Rb_1, HPLC)	05	*Pfaffia dunaliana*	Not Present	-
	08	*Pfaffia dunaliana*	Not Present	-
	10	Amaranthaceae	Not Present	-
	13	*Pfaffia dunaliana*	Not Present	-
	25	*Pfaffia dunaliana*	Not Present	-
P. incarnata (1,5% of total flavonoids, UV)	01	*Passiflora spp.*	Not Present	0,126%
	03	*Passiflora spp.*	Present	1,614%
	08	*Passiflora spp.*	Present	1,58%
	13	*Passiflora spp.*	Present	0,833%
	18	*Senna alexandrina*	Not Present	0,154%
	21	*Passiflora spp.*	Present	0,97%
	29	*Senna alexandrina*	Not Present	0,229%
P. boldus (0,1% of total alkaloids, HPLC)	06	*Vernonia colorata*	Not Present	-
	08	*Peumus boldus*	Present	0,33%
	26	*Peumus boldus*	Present	0,72%
	31	*Peumus boldus*	Present	0,29%
V. officinalis (0,05% of valerenic acid, HPLC)	02	*Ageratum conyzoides*	Not present	-
	04	No sequence	Not present	-
	06	*Valeriana spp.*	Not present	-
	09	Unidentified	Not present	-
	22	*Valeriana spp.*	Not present	-
	24	*Valeriana spp.*	Present	0,037%
	25	*Asteraceae*	Present	0,023%
	29	*Cissampelos spp.*	Present	0,050%
	32	*Valeriana spp.*	Not present	-

Figures

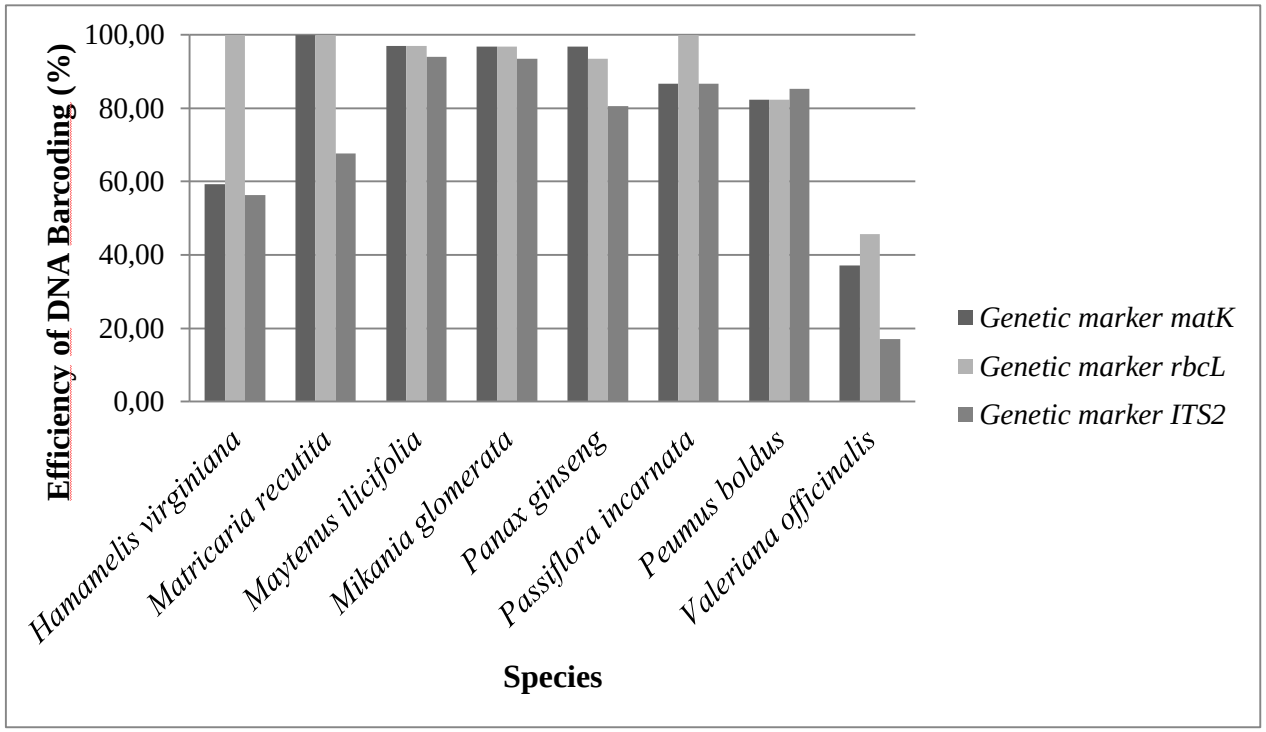

Figure 1: Percentage of samples analyzed according to species and genetic marker.

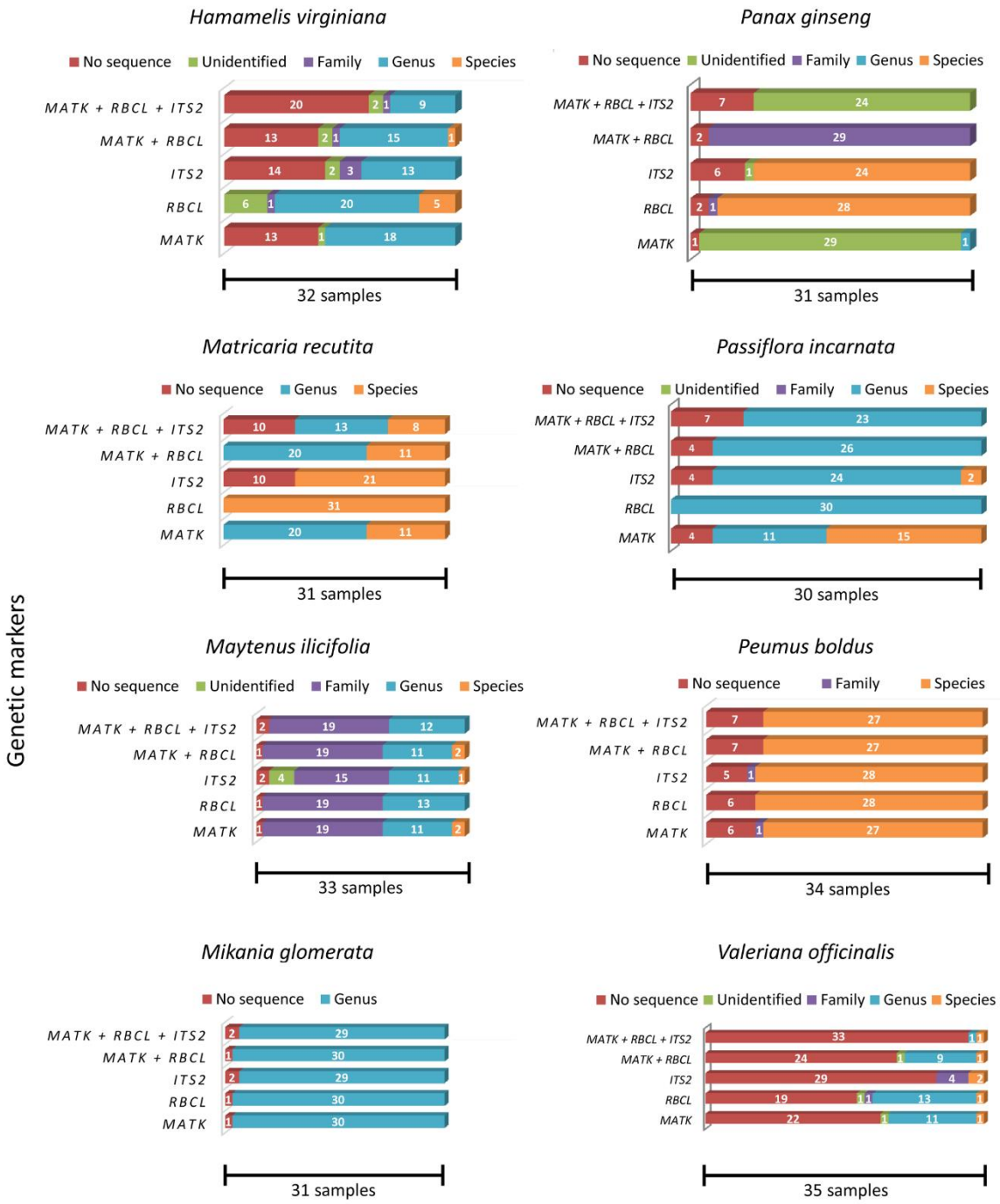

Figure 2: Identification levels for the analyzed samples when using each or a combination of the chosen markers. **No sequence:** samples for which the DNA barcoding protocol did not work. **Unidentified:** samples that could not be identified. The sequences from these samples did not show similarity levels above 98% to any of the sequences within the databases. **Family:** samples that could be identified at the family level. The sequences from these samples showed equal similarity levels to database sequences from multiple species belonging to the same family.

Genus: samples that could be identified to the genus level. The sequences from these samples showed equal similarity levels to database sequences from multiple species belonging to the same genus. **Species:** samples that could be identified to the species level. The sequences from these samples showed similarity levels above 98% to database sequences from a unique species.

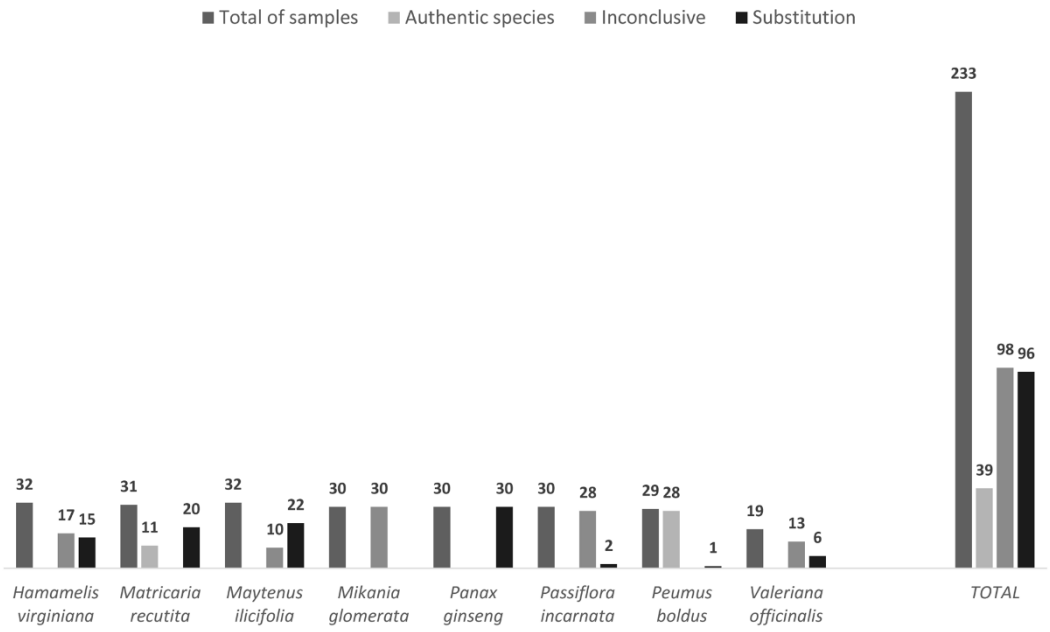

Figure 3: DNA final barcode identification of the analyzed samples.

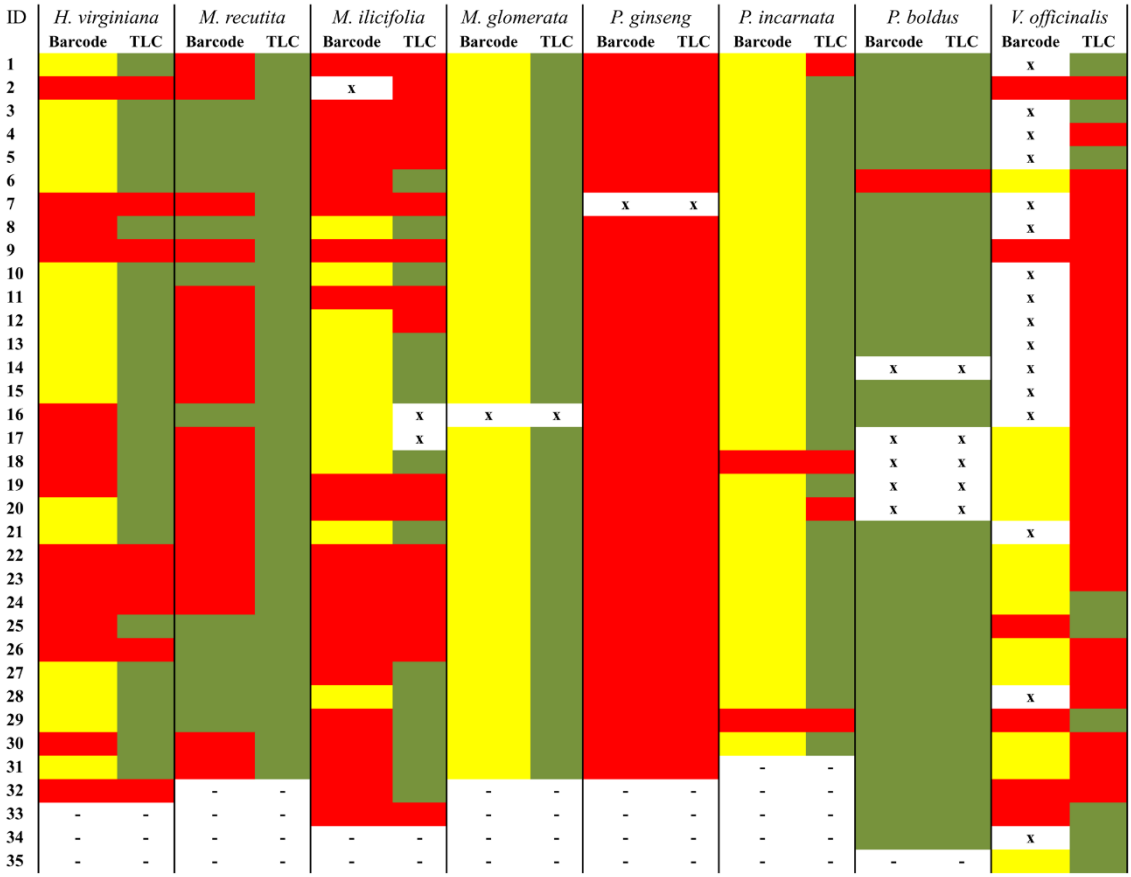

Figure 4: Comparison between the DNA barcode and TLC findings. **ID**: sample number. **Green**: samples that were identified as the expected medicinal species using DNA barcoding and that contained the expected chemical marker from the medicinal species according to TLC. **Yellow**: samples that were not identified within the genus of the medicinal species using DNA barcoding. **Red**: samples that were identified using DNA barcoding as a genus or family that varied from the expected one and that did not contain the chemical marker according to TLC. **X**: samples that did not generate any sequence using DNA barcoding or that could not be tested using TLC. -: absent samples.

IV – DISCUSSÃO

As plantas medicinais são utilizadas em diversas culturas em todo o mundo, seja na forma de medicamento fitoterápico ou na forma de produtos fitoterápicos relacionados a conhecimento popular. Desde a década de 70, a Organização Mundial da Saúde (OMS) reconhece o valor das plantas medicinais como recurso terapêutico importante, e vem estimulando os países ricos em biodiversidade a elaborar políticas e desenvolver produtos de qualidade para uso em saúde pública. Desde então as plantas medicinais estão sendo reconhecidas como essenciais para a complementação da saúde humana, além de potenciais fontes para novos medicamentos, o que tem incentivado estudos de identificação e validação de mais e mais espécies medicinais (WHO, 2013).

O Brasil tem uma das floras mais diversas do planeta, constituída de cerca de 40.000 espécies, o que corresponde a 20% da flora mundial. A Floresta Amazônica é conhecida no mundo por sua vastidão ainda intocada e a Floresta Atlântica e o Cerrado brasileiro são considerados hotspots de biodiversidade. Esses biomas estão incluídos entre os mais ricos de vida animal e vegetal no planeta (CONSERVATION INTERNATIONAL, 2010). A Caatinga e o Pantanal abrangem quase 15% do território brasileiro e também contém vasta diversidade biológica. Diferentes ecossistemas produzem uma variedade enorme de substâncias, com estruturas químicas diferentes, que podem ser úteis para os mais variados fins, o que faz do Brasil um dos países de maior potencial biotecnológico no mundo. A despeito de toda esta riqueza, a maioria dos medicamentos fitoterápicos comercializados aqui provém de espécies exóticas e, ou importadas (ANVISA, 2014). Esse fato evidencia a falta de investimentos em pesquisas científicas que poderiam contribuir na descoberta de espécies medicinais e padronização de medicamentos fitoterápicos a partir delas.

Somado à falta de pesquisas nessa área, os resultados reportados no Capítulo 1 (Tabela 2), que trata mais especificamente de espécies nativas da nossa flora, apontam que o conhecimento popular, adquirido e transmitido desde a época da colonização brasileira, pode estar se perdendo. Essa hipótese é corroborada pelos resultados de outro trabalho (BRANDÃO et al., 2013) no qual verificou-se que 50% de um total de 252 amostras de 40 espécies nativas da flora brasileira e comercializadas em mercados de todo o país tratavam-se de espécies identificadas incorretamente. Outra hipótese que deve ser considerada é a substituição intencional de espécies.

Desde a primeira edição da Farmacopeia Brasileira, quando foram estabelecidos os primeiros critérios para a garantia da qualidade no uso de espécies medicinais, diversos avanços foram feitos. Apesar disso, a etapa de identificação das espécies componentes de determinado produto continua dependendo de métodos químicos, como o perfil cromatográfico e métodos botânicos clássicos. Na última década, a ascensão das metodologias de identificação genética como uma alternativa tornou as fraudes mais evidentes e comuns (KUMAR et al.; STOECKLE et al., 2011; COGHLAN et al.; WALLACE et al., 2012; SEETHAPATHY et al., 2014). Associado a esses e outros trabalhos publicados em todo o mundo, nossos resultados sugerem que o DNA barcode é uma técnica que, se incluída no processo de produção de fitoterápicos, levará a uma melhoria na qualidade e segurança desses medicamentos.

Os resultados obtidos demonstram a eficácia dos processos de extração do DNA, amplificação e sequenciamento para as mais diversas espécies trabalhadas (Tabela 3 – Capítulo 1 e Figura 1 – Capítulo 2), principalmente em se tratando do marcador *rbcL*. Apesar de o DNA barcode empregar oficialmente duas regiões distintas (*matK* e *rbcL*), ficou demonstrado que nem sempre as duas regiões são necessárias (Figuras suplementares – Capítulos 1 e 2). Sendo assim, é possível padronizar qual sequência é mais eficiente para determinada espécie, tornando o processo ainda mais rápido e barato.

Apesar de a metodologia de DNA barcode ter sido eficaz na detecção de um alto percentual de substituições de espécies nas amostras analisadas (Tabela 2 – Capítulo 1 e Figura 3 – Capítulo 2), os resultados levantam um problema: a identificação de espécies dentro de um mesmo gênero. A maioria das substituições confirmadas ocorreu com espécies de gêneros ou famílias diferentes da espécie esperada, o que tornou mais fácil a detecção destas substituições (Figuras suplementares – Capítulo 2). Já a diferenciação entre espécies muito próximas mostrou-se mais difícil (Tabela 2 – Capítulo 1 e Figuras suplementares – Capítulo 2). Mesmo utilizando análises de *Barcode gap* para espécies dentro de um mesmo gênero, as variabilidades intra e interespecífica mostraram, em alguns casos, uma sobreposição que impossibilitou a diferenciação (Tabela 3 – Capítulo 2). Sendo assim, melhorias no processo de diferenciação entre espécies próximas são uma necessidade (RUBINOFF et al., 2006, BHARGAVA & SHARMA, 2013). Dentre as espécies estudadas, o guaco (*M. glomerata*) chama a atenção, pois é o melhor exemplo de que a busca por marcadores para identificação de espécies deve sempre continuar. O guaco é uma espécie nativa do Brasil utilizado para o tratamento de dispepsias, gastrites e úlcera gastroduodenal (ANVISA, 2008). Nenhum dos marcadores utilizados, individualmente ou em associações, foi capaz de diferenciar entre as espécies *M. glomerata* e *M, laevigata*, sendo

que ambas obtiveram sempre 100% de identidade entre elas. Nesse ponto, análises de outras sequências, incluindo possivelmente o sequenciamento cloroplastídeo dessas espécies para identificar uma região mais eficiente na diferenciação entre essas duas espécies.

De fato, o DNA barcode tem sido foco de estudo e aprimoramento constante dos pesquisadores da rede mundial, pois existe uma necessidade de maior eficiência no depósito de sequências provenientes de novas espécies (HAJIBABAEI et al., 2011; KVIST, 2013). Isso abriu caminho para o emprego de metodologias de sequenciamento de larga escala na tentativa de melhorar o processo como um todo. Dentre as diversas vantagens do uso de tais metodologias, está a possibilidade de se analisar diversos marcadores de uma só vez, o que possibilita o aumento da resolução da técnica com relação às espécies muito próximas filogeneticamente. Outra questão relevante, principalmente para análises em casos de misturas de duas ou mais espécies diferentes, que impulsionou a padronização do sequenciamento de nova geração para o DNA barcode é a limitação da técnica de Sanger para a identificação de espécies em produtos industrializados, pois esse método de sequenciamento impossibilita a identificação em amostras compostas por mais de uma espécie (TILLMAR et al., 2013).

A correlação entre as análises de identificação molecular e as análises químicas reportada nesse estudo apresentou resultados alarmantes. Diversas amostras que apresentaram, na análise de identificação molecular, espécies diferentes das esperadas, apresentaram, nas análises químicas, os marcadores esperados para a espécie declarada (Tabela 2 – Capítulo 1 e Figura 4 – Capítulo 2). Em alguns casos, como para a espécie *Roupala spp.*, a espécie identificada pela análise molecular não possuía nenhum registro de utilização prévia como planta medicinal ou registros de testes de segurança para o consumidor. Isso significa que tais amostras podem ter passado por análises químicas, que resultaram positivas para a espécie esperada, apesar de conterem outras espécies. Ainda na análise comparativa, foram observadas espécies corretas que não apresentaram os efeitos biológicos esperados (Tabela 3 – Capítulo 1) ou que apresentaram os marcadores químicos corretos, porém em concentrações abaixo do mínimo exigido (tabela 4 – Capítulo 2). Esses resultados sugerem que o preparo, ou mesmo as condições de armazenamento das drogas vegetais estão fora dos padrões. Isso comprova a necessidade da integração entre as técnicas de identificação molecular e as técnicas químicas de comprovação de pureza, presença e concentração dos princípios ativos.

A aplicação de metodologias moleculares para identificação de espécies em produtos de consumo humano não só é uma realidade em outros países do mundo, com casos de sucesso

como melhorias no processo já estão sendo implementadas (GALIMBERTI et al.; NEWMASTER et al.; TILLMAR et al., 2013). O DNA barcode já não é apenas presente em publicações científicas e vem sendo empregado por governos em alguns países para a identificação de espécies em produtos processados de origem animal ou vegetal utilizados tanto para consumo humano quanto animal. Nos Estados Unidos, o órgão de fiscalização Food and Drugs Administration (FDA) adotou oficialmente o DNA barcode para a identificação de pescados (YANCY et al., 2008). Recentemente, a procuradoria geral do Estado de Nova York (EUA) utilizou a identificação por DNA barcode para impedir a comercialização de suplementos à base de vegetais naquele Estado. Não só os suplementos continham espécies diferentes das anunciadas como também continham espécies potencialmente tóxicas (http://www.iflscience.com/health-and-medicine/herbal-supplements-dont-contain-what-they-claim - acessado em 08/02/2015). Na China, metodologias moleculares para a certificação de vegetais utilizados na produção de medicamentos são recomendadas desde 2010 na Farmacopeia oficial do país (LI et al., 2011). No Brasil o DNA barcode foi utilizado para identificação de pescados comercializados na cidade de Florianópolis em uma ação que resultou na multa, pelo PROCON, de estabelecimentos que comercializavam espécies diferentes das rotuladas (CARVALHO et al., 2014).

Por fim, os resultados aqui apresentados indicam não só que a identificação molecular deve ser implementada no processo de produção de fitoterápicos, como também que mais estudos voltados para a melhoria desse processo e, principalmente para catalogar a biodiversidade nativa do Brasil, devem ser abordados de uma forma mais rápida e eficiente.

V – REFERÊNCIAS

ANVISA. AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA (Brasil) – Instrução Normativa nº 5 de 11 de Dezembro de 2008: Lista de medicamentos fitoterápicos de registro simplificado. 2008.

ANVISA. AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA (Brasil). Resolução RDC nº 26 de 13 de maio de 2014. Dispõe sobre o registro de medicamentos fitoterápicos. Brasília, 14 de maio de 2014.

AUSUBEL, J. H. A botanical macroscope. Proc Natl Acad Sci, v. 106, n. 31, p. 12569-12570. 2009.

BHARGAVA, M.; SHARMA, A. DNA barcoding in plants: Evolution and applications of *in silico* approaches and resources. Mol Phylogenet Evol, v. 67, p. 631-641. 2013.

BOONSOM, T.; WARANUCH, N.; INGKANINAN, K.; DENDUANGBORIPANT, J.; SUKRONG, S. Molecular analysis of the genus Asparagus based on matK sequences and its application to identify *A. racemosus*, a medicinally phytoestrogenic species. Fitoterapia, v. 83, n. 5, p. 947-953. 2012.

BRANDÃO, M. G. L. ALVES, R. M. S.; MOREIRA, R. A.; OLIVEIRA, P.; MANUELA, T.; et al. Qualidade de amostras comerciais de chás de plantas medicinais. Rev Bras Plantas Med, v. 5, n. 1, p. 56-59. 2002.

BRANDÃO, M. G. L. Dossiê Técnico: Produção de chás e extratos de plantas medicinais. Fundação Centro Tecnológico de Minas Gerais, 2007. 24 p.

BRANDÃO, M. G. L.; ZANETTI, N. N.; OLIVEIRA, P.; GRAEL, C.F.; SANTOS, A. C.; et al. Brazilian medicinal plants described by the 19th century European naturalists and in the Official Pharmacopoeia. J Ethnopharmacol, v. 120, p. 141-148. 2008.

BRANDÃO, M. G. L.; COSENZA, G. P.; PEREIRA, F. L.; VASCONCELOS, A. S.; FAGG, C. W. Changes in the trade in native medicinal plants in Brazilian public markets. Environ Monit Assess, v. 185, p. 7013-7023. 2013.

CAÑIGUERAL, S.; DELLACASSA, E.; BANDONI, A.L. Plantas Medicinales y Fitoterapia: ¿indicadores de dependencia o factores de desarrollo? Acta Farm Bonaer, v. 22, n. 3, p. 265-278. 2003.

CARVALHO, D. C.; SEERIG, A. S.; MELO, D. C.; DE SOUZA, A. B.; PIMENTA, D.; et al. Identificação molecular de peixes: o caso do Surubim (*Pseudoplatystoma* spp.). Rev Bras Reprod Anim, v. 32, p. 215-219. 2008.

CARVALHO, A. C. B.; PERFEITO, J. P. S.; COSTA E SILVA, L. V.; RAMALHO, L. S.; MARQUES, R. F. O.; et al. Regulation of herbal medicines in Brazil: advances and perspectives. Braz J Pharm Sci, v. 47, n. 3, p. 467-473. 2011.

CARVALHO, D. C.; PALHARES, R. M.; DRUMMOND, M. G.; FRIGO, T. B. DNA Barcoding identification of commercialized seafood in South Brazil: A governamental regulatory forensic program. Food control, v. 50, p. 784-788. 2014.

CBOL Plant Working Group. A DNA barcode for land plants. P Natl Acad Sci USA, v. 106, p. 12794–12797. 2009.

CHASE, M. W.; FAY, M. F. Ecology. Barcoding of plants and fungi. Science, v. 325, n. 5941, p. 682-683. 2009.

CHEN, S.; YAO, H.; HAN, J.; LIU, C.; SONG, J.; et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One, v. 5, n. 1, p. e8613. 2010.

CHIOU, S. J.; YEN, J. H.; FANG, C. L.; CHEN, H. L.; LIN, T. Y. Authentication of medicinal herbs using PCR-amplified ITS2 with specific primers. Planta Med, v. 73, n. 13, p. 1421-1426. 2007.

COGHLAN, M. L.; HAILE, J.; HOUSTON, J.; MURRAY, D. C.; WHITE, N. E.; et al. Deep sequencing of plant and animal DNA contained within traditional Chinese medicines reveals legality issues and health safety concerns. PLoS Genet, v. 8, n. 4, p. e1002657. 2012.

CONSERVATION INTERNATIONAL. Biodiversity hotspots. Disponível em: <http://www.biodiversityhotspots.org/xp/hotspots/Pages/default.aspx>. Acessado em 08/02/2013. 2010.

DE LUCA, V.; SALIM, V.; ATSUMI, S. M.; YU, F. Mining the biodiversity of plants: a revolution in the making. Science, v. 336, p. 1658-1661. 2012.

FEBRAFARMA. FEDERAÇÃO BRASILEIRA DA INDÚSTRIA FARMACÊUTICA. Disponível em: <http://www.febrafarma.org.br>. Acesso em: 20 abr. 2009.

GALIMBERTI, A.; DE MATTIA, F.; LOSA, A.; BRUNI, I.; FEDERICI, S.; et al. DNA barcoding as a new tool for food traceability. Food Res Int, v. 50, p. 55-63. 2013.

GUO, X.; WANG, X.; SU, W.; ZHANG, G.; ZHOU, R. DNA barcodes for discriminating the medicinal plant *Scutellaria baicalensis* (Lamiaceae) and its adulterants. Biol Pharm Bull, v. 34, n. 8, p. 1198-1203. 2011.

HAJIBABAEI, M.; SHOKRALLA, S.; ZHOU, X.; SINGER, G. A. C.; BAIRD, D.J. Environmental Barcoding: A Next-Generation Sequencing approach for biomonitoring applications using river benthos. Plos One, v. 6, n. 4, e17497. 2011.

HERBERT, P. D.; CYWINSKA, A.; BALL, S. L.; DEWAARD, J. R. Biological Identification through DNA Barcodes. Proc. R. Soc. Lond, v. 270, n. 1512, p. 313-321. 2003.

HOLLINGSWORTH, P. M.; GRAHAM, S. W.; LITTLE, D. P. Choosing and using a plant DNA barcode. PLoS One, v. 6, n. 5, p. e19254. 2011.

KUMAR, S.; KAHLON, T.; CHAUDHARY, S. A rapid screening for adulterants in olive oil using DNA barcodes. Food Chem, v. 127, p. 1335-1341. 2011.

KVIST, S. Barcoding in the dark? A critical view of the sufficiency of zoological DNA barcoding databases and a plea for broader integration of taxonomic knowledge. Mol Phylogenet Evol, v. 69, n. 1, p. 39-45. 2013.

LI, J. W. H.; VEDERAS, J. C. Drug discovery and natural products: End of an Era or Endless Frontier? Science, v. 325, n. 5937, p. 161-165. 2009.

LI, M.; CAO, H.; BUT, P. P. H. Identification of herbal medicinal materials using DNA barcodes. J Syst Evol, v. 49, n. 3, p. 271-283. 2011.

NEWMASTER, S. G.; RAGUPATHY, S.; JANOVEC, J. A botanical renaissance: state-of-the-art DNA barcoding facilitates an Automated Identification Technology system for plants. Int. J. Computer Applications in Technology, v. 35, n. 1, p. 50-60. 2009.

NEWMASTER, S. G.; GRGURIC, M.; SHANMUGHANANDHAN, D.; RAMALINGAM, S.; RAGUPATHY, S. DNA barcoding detects contamination and substitution in North American herbal products. BMC Med, v. 11 n. 222. p. 1-13. 2013.

PENNISI, E. Taxonomy. Wanted: a barcode for plants. Science, v. 318, n. 5848, p. 190-191. 2007.

RUBINOFF, D.; CAMERON, S.; WILL, K. Are plant DNA barcodes a search for the Holy Grail? Trends Ecol Evol, v. 21, p. 1-2. 2006.

SAHOO, N.; CHOUDHURY, K.; MANCHIKANTI, P. Manufacturing of biodrugs: Need for harmonization in regulatory standards. Bio Drugs, v. 23, n. 4 p. 217-229. 2009.

SCHULTZ, J.; WOLF, M. ITS2 sequence-structure analysis in phylogenetics: a how-to manual for molecular systematics. Mol Phylogenet Evol, v. 52, n. 2, p. 520-523. 2009.

SEETHAPATHY, G. S.; GANESH, D.; KUMAR, J. U. S.; SENTHILKUMAR, U.; NEWMASTER, S. G.; et al. Assessing product adulteration in natural health products for laxative yielding plants, *Cassia*, *Senna*, and *Chamaecrista*, in Southern India using DNA barcoding. Int J Legal Med, 2014 Nov 26 [epub ahead of print].

SONG, J.; YAO, H.; LI, Y.; LI, X.; LIN, Y.; et al. Authentication of the family Polygonaceae in Chinese pharmacopoeia by DNA barcoding technique. J Ethnopharmacol v. 124, p. 434-439. 2009.

STOECKLE, M. Y.; GAMBLE, C. C.; KIRPEKAR, R.; YOUNG, G.; AHMED, S.; et al. Commercial teas highlight plant DNA barcode identification successes and obstacles. Sci Rep, v. 1, n. 42, p. 1-7. 2011.

SUCHER, N. J.; CARLES, M. C. Genome-based approaches to the authentication of medicinal plants. Planta Med, v. 74, n. 6, p. 603-623. 2008.

TILLMAR, A. O.; DELL´AMICO, B.; WELANDER, J.; HOLMLUND, G. A universal method for species identification of mammals utilizing next generation sequencing for the analysis of DNA mixtures. Plos One, v. 9, n. 12, e83761. 2013.

VIEIRA, R. A. Validação científica de plantas medicinais como fator catalisador no desenvolvimento da indústria farmacêutica nacional. Rev Saúde Amb, v. 2, n. 1/2, p.57-64. 2001.

WACHTEL-GALOR, S.; BENZIE, I. F. F. Herbal Medicine: An Introduction to Its History, Usage, Regulation, Current Trends, and Research Needs. In: Benzie IFF, Wachtel-Galor S, editors. Herbal Medicine: Biomolecular and Clinical Aspects. 2nd edition. Boca Raton (FL): CRC Press; 2011. Chapter 1. Available from: http://www.ncbi.nlm.nih.gov/books/NBK92773/

WALLACE, L. J.; BOILARD, S. M. A. L.; EAGLE, S. H. C.; SPALL, J. L.; SHOKRALLA, S.; et al. DNA barcodes for everyday life: Routine authentication of Natural Health Products. Food Res Int, v. 49, p. 446-452. 2012.

WORLD HEALTH ORGANIZATION (WHO). Quality control methods for medicinal plant materials. Geneva, 2005.

WORLD HEALTH ORGANIZATION (WHO). Instruções operacionais: informações necessárias para a condução de ensaios clínicos com fitoterápicos. Brasília: Ministério da Saúde. 2008.

WORLD HEALTH ORGANIZATION (WHO). WHO Traditional medicine strategy 2014 - 2023. Geneva, 2013.

YANCY, H. F.; ZEMLAK, T. S.; MASON, J. A.; WASHINGTON, J. D.; TENGE, B. J.; et al. Potential Use of DNA Barcodes in Regulatory Science: Applications of the Regulatory Fish Encyclopedia. J Food Prot, v. 71, n. 1, p. 210-217. 2008.