

ỨNG DỤNG MÃ VẠCH DNA HỖ TRỢ ĐỊNH LOẠI LOÀI MỘT SỐ MẪU SÂM THUỘC CHI NHÂN SÂM (*PANAX L.*)

Lê Thanh Hương¹, Nguyễn Nhật Linh¹, Bùi Mạnh Minh¹, Hà Hồng Hạnh¹, Huỳnh Thị Thu Huệ¹, Nông Văn Hải¹, Hà Văn Huân², Lê Thị Thu Hiền^{1,✉}

¹Viện Nghiên cứu hệ gen, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

²Trường Đại học Lâm nghiệp, Bộ Nông nghiệp và Phát triển nông thôn

✉ Người chịu trách nhiệm liên lạc. E-mail: hienlethu@igr.ac.vn

Ngày nhận bài: 13.7.2016

Ngày nhận đăng: 30.12.2016

TÓM TẮT

Xác định mã vạch DNA (DNA barcoding) là một phương pháp định loại phân tử phổ biến thường được sử dụng để hỗ trợ định loại các loài khó nhận dạng về hình thái hay các mẫu đã qua chế biến, xử lý. Phương pháp dựa trên nguyên tắc so sánh các vùng trình tự DNA ngắn có tốc độ tiến hóa đủ nhanh để định loại các loài trong cùng chi. Trên đối tượng thực vật, các vùng DNA được sử dụng làm mã vạch trong phân loại thường là các trình tự thuộc hệ gen lục lạp và hệ gen nhân với vùng mã hóa và vùng không mã hóa như *psbA-trnH*, *matK*, *rbcL*, ITS... Trong nghiên cứu này, 5 mã vạch phân tử tiềm năng là 18S, ITS, *matK*, *psbA-trnH* và *rbcL* được sử dụng để đánh giá khả năng phân biệt loài của 11 mẫu sâm nghiên cứu. Kết quả phân tích so sánh các trình tự nhận được với 41 trình tự của 9 loài thuộc chi Nhân sâm (*Panax L.*) đã được công bố trên Ngân hàng Gen quốc tế cho thấy vùng 18S có độ tương đồng giữa các cặp trình tự cao nhất, trình tự của các loài tương đồng trung bình đạt 99,87%. Tiếp đến là vùng *rbcL*, *matK* và *psbA-trnH* với tỷ lệ tương đồng trung bình lần lượt là 99,27%, 98,66% và 96,82%. Mức độ đa hình thể hiện rõ trên vùng ITS với tỷ lệ tương đồng trung bình thấp nhất, chỉ đạt 96,50%. Các biểu đồ hình cây về mối quan hệ phát sinh loài cho thấy có 4 trong 11 mẫu sâm là Sâm Ngọc Linh (*Panax vietnamensis* Ha et Grushv.) và 3 mẫu là Tam thất hoang (*Panax stipuleanatus*). Kết quả nghiên cứu cũng cho thấy 4 mẫu có hình thái của Sâm Vũ diệp (*Panax bipinnatifidus*) được nhận dạng phân tử là loài Tam thất hoang. Ngoài ra nghiên cứu cũng cho thấy, chỉ thị ITS và *psbA-trnH* là hai chỉ thị tiềm năng, hỗ trợ định danh Sâm Ngọc Linh và Tam thất hoang.

Từ khóa: Mã vạch DNA, chi Nhân sâm, Sâm Ngọc Linh, Tam thất hoang, Sâm Vũ diệp

MỞ ĐẦU

Chi Nhân sâm (*Panax L.*) là chi gồm nhiều loài cây thuốc có giá trị cao. Trong đó có các loài chứa nhiều hợp chất tự nhiên có cấu tạo phân tử khá phức tạp, độc đáo, có hoạt tính tốt và có tác dụng tăng cường thể lực. Ở Việt Nam, một số loài mọc tự nhiên rất có giá trị làm thuốc thuộc chi *Panax L.* là Sâm Ngọc Linh hay Sâm Việt Nam (*P. vietnamensis* Ha et Grushv.), Sâm Vũ diệp (*P. bipinnatifidus*) và Tam thất hoang (*P. stipuleanatus*). Trong đó, Sâm Ngọc Linh là loài đặc biệt có giá trị khoa học và kinh tế. Sâm Ngọc Linh được xác định là một cây thuốc quý của Việt Nam với nhiều thành phần saponin, hàm lượng các acid amin, các chất khoáng vi lượng trong củ, lá và rễ hơn nhiều những loài sâm khác (Lã Đình Mới *et al.*, 2013; Phan Kế Long *et al.*, 2014). Sâm

Ngọc Linh mọc tập trung chủ yếu ở chân núi Ngọc Linh, thuộc các huyện Trà My, Trà Lĩnh, Trà Giang, tỉnh Quảng Nam. Các loài sâm này mang nhiều đặc điểm về hình dáng thân và củ tương đồng với nhau gây khó khăn cho việc nhận dạng. Việc xác định được các mã vạch DNA hỗ trợ phân loại được chính xác các loài sâm này hoặc các mẫu sâm đã qua chế biến đóng vai trò quan trọng trong việc bảo tồn, sử dụng bền vững và đảm bảo chất lượng cho các sản phẩm sâm của Việt Nam.

Xác định mã vạch DNA (DNA barcoding) là một phương pháp định loại phổ biến có thể sử dụng để phân loại phân tử. Phương pháp này có thể áp dụng cho các đối tượng sinh vật khác nhau từ động vật, thực vật, vi sinh vật và dựa trên nguyên tắc so sánh các vùng trình tự DNA ngắn có tốc độ tiến hóa đủ nhanh để phân loại các chi và các loài trong cùng

chi (Lê Thị Thu Hiền *et al.*, 2012). Trên thế giới, việc sử dụng phương pháp mã vạch DNA để phân loại các loài sâm thuộc chi Nhân sâm đã rất phổ biến và thông dụng từ những năm giữa thập kỷ 90 của thế kỷ trước. Số lượng mã vạch phân tử được sử dụng tương đối nhiều, chúng có thể nằm trong genome nhân như vùng ITS, 18S rRNA; trong ty thể như *nad1* hoặc nằm trong hệ gen lục lạp như *matK*, *psbA-trnH*, *psbK-I*, *pspM-trnD*, *rps16*, *trnC-trnD*... Trong đó, vùng *ITS* và *psbA-trnH* cho thấy nhiều đa hình nucleotide đơn hơn cả và có thể sử dụng cho định loài và phân loại nhóm chi Nhân sâm (Zhu *et al.*, 2003; Lee *et al.*, 2004; Zuo *et al.*, 2011).

Năm 1996, Wen và Zimmer đã công bố cây phát sinh chủng loại của 12 loài sâm khác nhau thuộc chi Nhân sâm phân bố ở Bắc Mỹ và Đông Á dựa trên trình tự vùng ITS với độ dài từ 606 đến 608 bp gồm vùng ITS1, vùng xen 5,8S và ITS2. Các nghiên cứu xây dựng cây phát sinh chủng loại chi này từ đó đến nay đều sử dụng trình tự vùng *ITS* như một tiêu chuẩn để tham chiếu cũng như kết hợp với các barcode khác để có được kết quả toàn diện hơn (Chen *et al.*, 2013; Zuo *et al.*, 2011) Năm 2004, Lee *et al.*, công bố mã vạch DNA mới của chi Nhân sâm là vùng *trnC-trnD* nằm xen giữa hai gen *trnC* và *trnD* trong hệ gen lục lạp. Dựa trên trình tự vùng này kết hợp với *ITS*, nhóm nghiên cứu đã xây dựng cây phát sinh chủng loại của 18 loài trong chi Nhân sâm và 2 loài thuộc chi *Aralia* (Lee *et al.*, 2004). Các công trình trên đã chứng minh được việc sử dụng các vùng chỉ thị mã vạch DNA để phân loại chi Nhân sâm là hoàn toàn khả thi và mở ra cơ hội xây dựng một bộ mã vạch phân tử hoàn chỉnh cho chi này.

Việc nghiên cứu phân loại chi Nhân sâm ở Việt Nam đã được tiến hành nhưng với mức độ và quy mô tương đối hạn chế. Việc phân loại chủ yếu dựa trên các đặc điểm hình thái thân, lá, rễ của cây sâm kết hợp với phân tích các hợp chất saponin (Lã Đình Mới *et al.*, 2013). Năm 2007, Nguyễn Tập *et al.*, đã sử dụng kỹ thuật RAPD để xây dựng cơ sở dữ liệu DNA của một số cây thuốc quý, trong đó có Sâm Ngọc Linh. Việc sử dụng các mã vạch phân tử đã được áp dụng nhưng chưa phong phú và toàn diện. Các vùng gen mã vạch được sử dụng chủ yếu là vùng *matK* và *ITS* (Nguyễn Văn Đạt *et al.*, 2013). Vũ Huyền Trang *et al.*, (2013) đã nghiên cứu xây dựng mã vạch DNA cho Sâm Ngọc Linh trên cơ sở 5 chỉ thị mã vạch DNA là *psbA-trnH*, *matK*, *trnL*,

rbcL và *ITS*. Nhóm tác giả đã chứng minh trong 5 chỉ thị mã vạch nghiên cứu, *psbA-trnH* là chỉ thị có tiềm năng nhất, cho phép phân biệt Sâm Ngọc Linh với các loài sâm khác trên thế giới với độ chính xác cao.

Trong nghiên cứu này, 11 mẫu sâm thuộc chi Nhân sâm được định danh trên cơ sở phân tích trình tự 5 vùng mã vạch DNA là 18S, *ITS*, *matK*, *psbA-trnH* và *rbcL*, đồng thời tiềm năng của các chỉ thị mã vạch này qua đó được phân tích và đánh giá.

VẬT LIỆU VÀ PHƯƠNG PHÁP

Mẫu phân tích

Ba mẫu Sâm Ngọc Linh (PV21, PV22, PV24), 4 mẫu Sâm Vũ diệp (PB03, PB04, PB05, PB06) và 3 mẫu Tam thất hoang (PS21, PS22, PS23) do Viện Nghiên cứu và Phát triển Vùng, Viện Ứng dụng công nghệ (Bộ Khoa học và Công nghệ) cung cấp. Mẫu sâm nuôi cấy *in vitro* (PV23) do Công ty cổ phần Khoa học Công nghệ Anh Đào, TP. Hồ Chí Minh cung cấp (Hình 1).

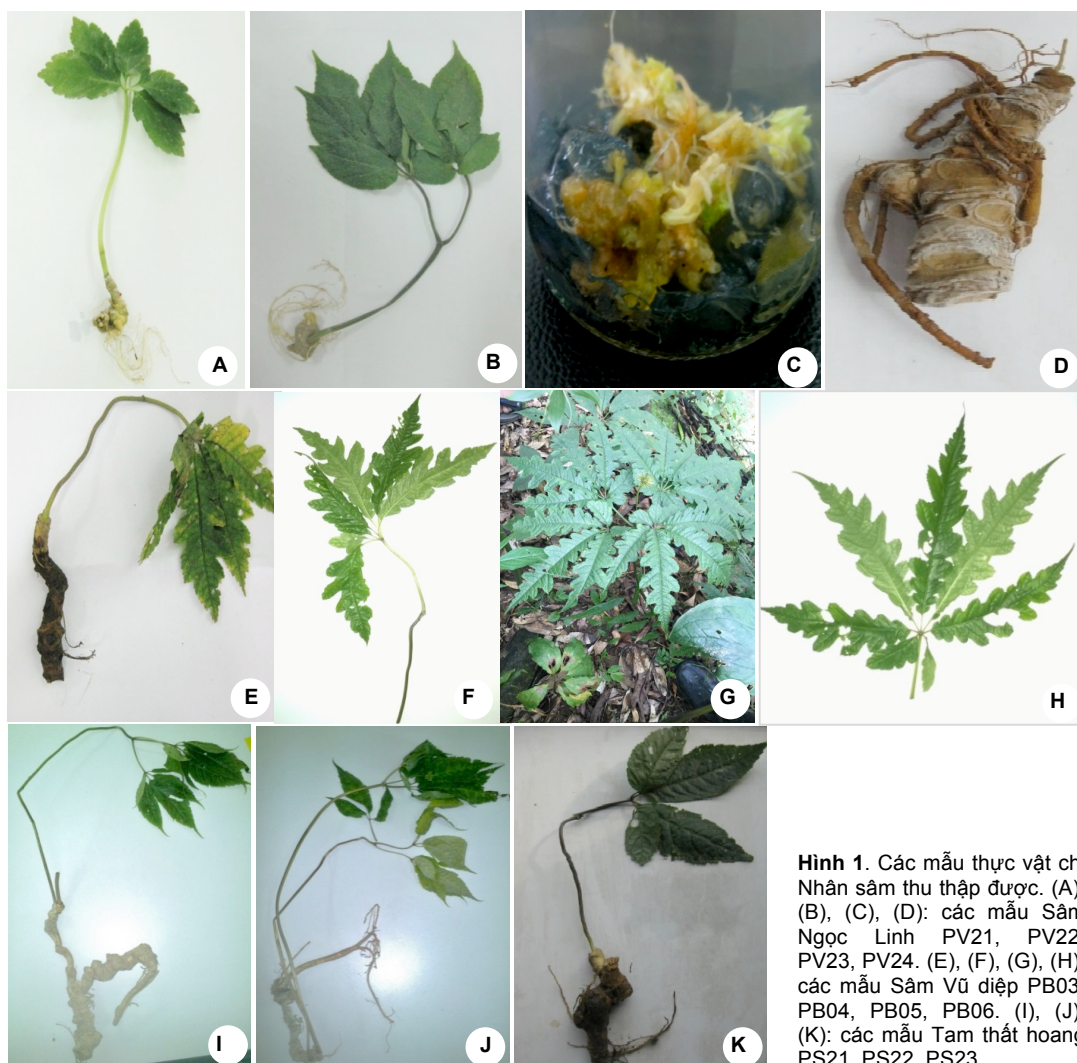
Tách chiết DNA tổng số, khuếch đại gen và xác định trình tự

DNA tổng số được tách chiết bằng phương pháp CTAB của Doyle và Doyle (Doyle và Doyle, 1990). Các cặp mồi đặc hiệu được thiết kế dựa trên thông tin về trình tự gen trên Ngân hàng Gen quốc tế (GenBank) (Bảng 1).

Phản ứng khuếch đại các đoạn gen được tối ưu sử dụng điều kiện như sau: 20 µl tổng thể tích bao gồm: 50 ng DNA tổng số; 2,5 µM mỗi primer; 0,75 unit *Taq* DNA polymerase; 1 mM dNTPs và đệm tương ứng trên máy IBM Veriti với chu trình nhiệt như sau: 1 chu kỳ biến tính 94 °C/ 2 phút; 35 chu kỳ (94 °C/ 30 giây; 50-55 °C/ 30 giây; 72 °C/ 1 phút) và bước tổng hợp cuối cùng 72 °C/ 5 phút. Sản phẩm PCR được kiểm tra trên gel agarose 0,8 %, sau đó được tinh sạch sử dụng bộ hóa chất GeneJET™ PCR Purification Kit (Hãng Thermo Scientific, Mỹ). Trình tự các đoạn DNA được xác định trên hệ thống ABI 3500 Genetic Analyzer theo nguyên lý của Sanger, với bộ kit BigDye® Terminator v 3.1 Cycle Sequencing (Hãng Applied Biosystems, Mỹ) bằng cách xác định trình tự trực tiếp từ sản phẩm PCR.

Bảng 1. Thông tin các mẫu sử dụng trong nghiên cứu.

Vùng DNA barcodes	Trình tự môi (5'-3')	Kích thước sản phẩm PCR (bp)
18S_F	TGCAACAAACCCCGACTTCTGG	1050
18S_R	GGACCTGGTAAGTTTCCCCGTG	
ITS_F	ACGAATTCATGGTCCGGTGAAGTGTTCCG	800
ITS_R	TAGAATTCCCCGGTTCGCTCGCCGTTAC	
matK_F	ACCGTACTTTTATGTTTACGAGC	850
matK_R	TCCATCTRGAAATMTTRGTTCA	
psbA-trnH_F	ACCCGGTCTTAGTGTATACGAG	390
psbA-trnH_R	TTCACTGCCTTGATCCACTTGG	
rbcL_F	GCAAGTGTGGATTCAAAGCTGGTG	600
rbcL_R	TGGTTGTGAGTTCACGTTCT	



Hình 1. Các mẫu thực vật chi Nhân sâm thu thập được. (A), (B), (C), (D): các mẫu Sâm Ngọc Linh PV21, PV22, PV23, PV24. (E), (F), (G), (H): các mẫu Sâm Vũ diệp PB03, PB04, PB05, PB06. (I), (J), (K): các mẫu Tam thất hoang PS21, PS22, PS23.

Bảng 2. Thông tin các loài và trình tự gen đã công bố trên GenBank được phân tích trong nghiên cứu.

Tên loài	Mã số truy cập GenBank				
	18S	ITS	matK	psbA-trnH	rbcL
<i>P. vietnamensis</i>	AB033635.1	JF772113.1	AB044906.2	NC_028704.1	NC_028704.1
<i>P. stipuleanatus</i>	AB088025.1	U41695.1	AB088015.1	HQ112892.1	HQ112601.1
<i>P. notoginseng</i>	D85171.1	U41684.1	AB027527.1	HQ112884.1	JF950030.1
<i>P. ginseng</i>	D83275.1	U41680.1	AB044903.2	JX996152.1	JF950029.1
<i>P. japonicus</i>	D84100.1	U41702.1	AB044905.2	HQ112881.1	KF208384.1
<i>P. quinquefolius</i>	D85172.1	U41687.1	HM142271.1	JX996153.1	HQ112616.1
<i>P. trifolius</i>		U41697.1	KM210137.1	HQ112894.1	HQ112614.1
<i>P. pseudoginseng</i>		HQ112438.1	AB088016.1	HQ112886.1	KM210153.1
<i>P. zingiberensis</i>	AB085764.1	U41700.1	AB088005.1		

Phân tích kết quả

Trình tự DNA được xử lý và phân tích bằng phần mềm BioEdit, trong đó có so sánh với các trình tự tương ứng của các loài thuộc chi Nhân sâm đã được công bố trên GenBank (Bảng 2). Khoảng cách giữa các trình tự khi so sánh theo cặp nucleotide (Pairwise distance) được xác định sử dụng phần mềm MEGA 6.06. Biểu đồ hình cây về mối quan hệ phát sinh loài được xây dựng theo phương pháp Neighbour-Joining và Maximum-Likelihood với giá trị bootstrap là 1000 dựa trên cơ sở các số liệu thu được.

KẾT QUẢ VÀ THẢO LUẬN

Kết quả tách chiết DNA tổng số và nhân dòng các đoạn gen

Sản phẩm DNA tổng số sau khi tách chiết và tinh chế được kiểm tra bằng điện di trên gel agarose nồng độ 0,8 %. Có thể nhận thấy rằng các mẫu DNA thu được có chất lượng tốt, trên giếng điện di chỉ xuất hiện một băng DNA có trọng lượng phân tử cao, sắc nét, đảm bảo chất lượng cho các bước nghiên cứu tiếp theo (Hình 2).

DNA tổng số của các mẫu sau khi được xác định mức độ hấp thụ tử ngoại (OD) bằng máy đo quang phổ, được pha loãng tới nồng độ cần thiết để làm khuôn cho PCR với các cặp mồi đặc hiệu đã được thiết kế (Bảng 1). Kết quả điện di sản phẩm PCR trên gel agarose 0,8 % cho thấy các cặp mồi 18S, ITS, matK, psbA-trnH và rbcL đã được sử dụng thành công trong việc nhân bản, khuếch đại các đoạn

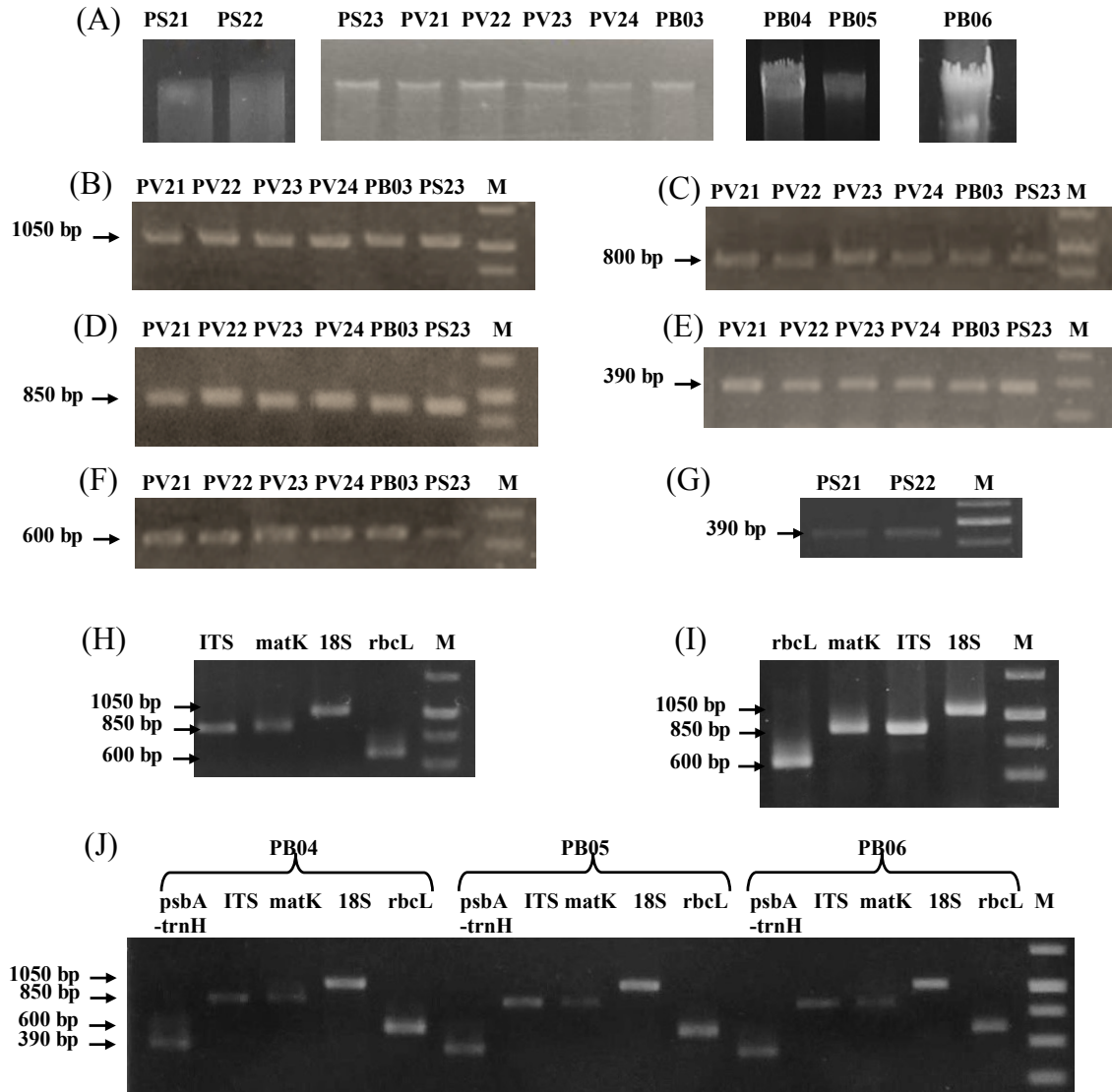
gen mong muốn từ DNA của 11 mẫu thực vật chi Nhân sâm. Các sản phẩm PCR này được tinh sạch sử dụng cho phản ứng xác định trình tự gen (Hình 2).

Kết quả phân tích trình tự

Sử dụng DNA tinh sạch và các cặp mồi đã tổng hợp, các đoạn DNA đã được khuếch đại với kích thước lần lượt là 1050 bp, 800 bp, 850 bp, 390 bp và 600 bp tương ứng với kích thước của các vùng gen 18S, ITS, matK, psbA-trnH, rbcL. Kết quả xác định trình tự cho thấy các đoạn gen đã được khuếch đại chính xác. Các trình tự gen này được so sánh với 41 trình tự tương ứng ở 9 loài thuộc chi Nhân sâm trên GenBank, trong đó có Sâm Ngọc Linh (*P. vietnamensis* Ha et Grushv.), Tam thất hoang (*P. stipuleanatus*), Tam thất (*P. notoginseng*), Nhân sâm (*P. ginseng*), Sâm Nhật Bản (*P. japonicus*), Sâm Bắc Mỹ (*P. quinquefolius*), Sâm Ba lá (*P. trifolius*), Tam thất (*P. pseudoginseng*), Tam thất gừng (*P. zingiberensis*) (Bảng 1). Kết quả so sánh cho thấy có sai khác trong trình tự các vùng gen nghiên cứu. Cụ thể, khi so sánh tỷ lệ phần trăm tương đồng giữa các cặp trình tự, vùng 18S có độ tương đồng giữa các cặp trình tự cao nhất (Trình tự của các loài tương đồng trung bình đạt 99,87 %), tiếp đến là vùng rbcL, matK và psbA-trnH với tỷ lệ tương đồng trung bình lần lượt là 99,27 %, 98,66 % và 96,82 %. Mức độ đa hình thể hiện rõ trên vùng ITS với tỷ lệ tương đồng trung bình thấp nhất, chỉ đạt 96,50 % (Bảng 3).

Bảng 3. Tỷ lệ phần trăm tương đồng giữa các trình tự

	18S										
Các loài so sánh	PV21	PV22	PV23	PV24	PS21	PS22	PS23	PB03	PB04	PB05	PB06
<i>P. vietnamensis</i>	100,00	100,00	100,00	100,00	99,89	99,89	99,89	99,89	99,89	99,89	99,89
<i>P. stipuleanatus</i>	99,89	99,89	99,89	99,89	100,00	100,00	100,00	100,00	100,00	100,00	100,00
<i>P. notoginseng</i>	100,00	100,00	100,00	100,00	99,89	99,89	99,89	99,89	99,89	99,89	99,89
<i>P. ginseng</i>	99,67	99,67	99,67	99,67	99,67	99,67	99,67	99,67	99,67	99,67	99,67
<i>P. japonicus</i>	99,78	99,78	99,78	99,78	99,78	99,78	99,78	99,78	99,78	99,78	99,78
<i>P. quinquefolius</i>	100,00	100,00	100,00	100,00	99,89	99,89	99,89	99,89	99,89	99,89	99,89
<i>P. trifolius</i>											
<i>P. pseudoginseng</i>											
<i>P. zingiberensis</i>	99,89	99,89	99,89	99,89	99,89	99,89	99,89	99,89	99,89	99,89	99,89
Độ tương đồng trung bình	99,87	±	0,11								
	ITS										
Các loài so sánh	PV21	PV22	PV23	PV24	PS21	PS22	PS23	PB03	PB04	PB05	PB06
<i>P. vietnamensis</i>	99,44	99,44	99,44	99,62	95,55	95,55	95,73	95,73	95,73	95,73	95,73
<i>P. stipuleanatus</i>	96,10	96,10	96,10	95,55	99,44	99,44	99,81	99,81	99,81	99,81	99,81
<i>P. notoginseng</i>	97,40	97,40	97,40	96,85	94,99	94,99	95,18	95,18	95,18	95,18	95,18
<i>P. ginseng</i>	98,52	98,52	98,52	97,96	95,92	95,92	96,10	96,10	96,10	96,10	96,10
<i>P. japonicus</i>	97,22	97,22	97,22	97,03	94,99	94,99	95,18	95,18	95,18	95,18	95,18
<i>P. quinquefolius</i>	98,89	98,89	98,89	98,33	96,10	96,10	96,29	96,29	96,29	96,29	96,29
<i>P. trifolius</i>	94,44	94,44	94,44	93,89	94,81	94,81	94,99	94,99	94,99	94,99	94,99
<i>P. pseudoginseng</i>	94,24	94,24	94,24	93,68	96,65	96,65	96,83	96,83	96,83	96,83	96,83
<i>P. zingiberensis</i>	98,52	98,52	98,52	97,96	95,73	95,73	95,92	95,92	95,92	95,92	95,92
Độ tương đồng trung bình	96,50	±	1,62								
	matK										
Các loài so sánh	PV21	PV22	PV23	PV24	PS21	PS22	PS23	PB03	PB04	PB05	PB06
<i>P. vietnamensis</i>	100,00	100,00	100,00	99,87	98,27	98,27	98,00	98,14	98,00	98,00	98,00
<i>P. stipuleanatus</i>	98,00	98,00	98,00	98,14	99,47	99,47	100,00	99,87	100,00	100,00	100,00
<i>P. notoginseng</i>	99,60	99,60	99,60	99,73	98,40	98,40	98,16	98,27	98,16	98,16	98,16
<i>P. ginseng</i>	99,60	99,60	99,60	99,73	98,40	98,40	98,16	98,27	98,16	98,16	98,16
<i>P. japonicus</i>	99,60	99,60	99,60	99,73	98,40	98,40	98,16	98,27	98,16	98,16	98,16
<i>P. quinquefolius</i>	99,60	99,60	99,60	99,73	98,40	98,40	98,16	98,27	98,16	98,16	98,16
<i>P. trifolius</i>	98,53	98,53	98,53	98,66	97,86	97,86	97,59	97,72	97,59	97,59	97,59
<i>P. pseudoginseng</i>	98,13	98,13	98,13	98,26	98,53	98,53	98,26	98,40	98,26	98,26	98,26
<i>P. zingiberensis</i>	99,73	99,73	99,73	99,87	98,26	98,26	97,99	98,13	97,99	97,99	97,99
Độ tương đồng trung bình	98,66	±	0,76								
	psbA-trnH										
Các loài so sánh	PV21	PV22	PV23	PV24	PS21	PS22	PS23	PB03	PB04	PB05	PB06
<i>P. vietnamensis</i>	100,00	100,00	100,00	99,72	95,29	95,29	95,29	95,29	95,29	95,29	95,29
<i>P. stipuleanatus</i>	95,29	95,29	95,29	95,01	99,43	99,43	100,00	100,00	100,00	100,00	100,00
<i>P. notoginseng</i>	95,84	95,84	95,84	95,57	94,18	94,18	94,18	94,18	94,18	94,18	94,18
<i>P. ginseng</i>	96,12	96,12	96,12	95,84	98,30	98,30	98,30	98,30	98,30	98,30	98,30
<i>P. japonicus</i>	96,12	96,12	96,12	95,84	98,30	98,30	98,30	98,30	98,30	98,30	98,30
<i>P. quinquefolius</i>	96,12	96,12	96,12	95,84	98,30	98,30	98,30	98,30	98,30	98,30	98,30
<i>P. trifolius</i>	94,46	94,46	94,46	94,18	96,31	96,31	96,31	96,31	96,31	96,31	96,31
<i>P. pseudoginseng</i>	94,74	94,74	94,74	94,46	97,73	97,73	97,73	97,73	97,73	97,73	97,73
<i>P. zingiberensis</i>											
Độ tương đồng trung bình	96,82	±	1,80								
	rbcl										
Các loài so sánh	PV21	PV22	PV23	PV24	PS21	PS22	PS23	PB03	PB04	PB05	PB06
<i>P. vietnamensis</i>	100,00	100,00	100,00	99,23	99,42	99,42	99,23	99,23	99,23	99,23	99,23
<i>P. stipuleanatus</i>	99,23	99,23	99,23	98,46	99,81	99,81	100,00	100,00	100,00	100,00	100,00
<i>P. notoginseng</i>	99,81	99,81	99,81	99,04	99,23	99,23	99,04	99,04	99,04	99,04	99,04
<i>P. ginseng</i>	99,42	99,42	99,42	98,66	98,85	98,85	99,04	99,04	99,04	99,04	99,04
<i>P. japonicus</i>	99,81	99,81	99,81	99,04	99,23	99,23	99,04	99,04	99,04	99,04	99,04
<i>P. quinquefolius</i>	99,62	99,62	99,62	98,85	99,04	99,04	99,23	99,23	99,23	99,23	99,23
<i>P. trifolius</i>	99,23	99,23	99,23	98,46	99,04	99,04	98,85	98,85	98,85	98,85	98,85
<i>P. pseudoginseng</i>	99,23	99,23	99,23	98,46	99,42	99,42	99,23	99,23	99,23	99,23	99,23
<i>P. zingiberensis</i>											
Độ tương đồng trung bình	99,27	±	0,37								



Hình 2. Hình ảnh điện di DNA tổng số và sản phẩm PCR khuếch các vùng gen nghiên cứu. (A): Kết quả điện di DNA tổng số các mẫu thực vật chi Nhân sâm. (B)-(F): Kết quả điện di sản phẩm PCR các vùng *psbA-trnH*, ITS, *matK*, 18S và *rbcL* của các mẫu Sâm Ngọc Linh (PV21-24), Sâm vũ diệp (PB03) và Tam thất hoang (PS23). (G): Kết quả điện di sản phẩm PCR vùng gen *psbA-trnH* của các mẫu Tam thất hoang (PS21, PS22). (H), (I): Kết quả điện di sản phẩm PCR các vùng gen còn lại của các mẫu Tam thất hoang (PS21, PS22). (J): Kết quả điện di sản phẩm PCR 5 vùng gen của các mẫu PB04, PB05, PB06. M: Thang chuẩn DNA 1 kb.

Trong 11 mẫu thực vật nghiên cứu, 4 mẫu PV21-24 có độ tương đồng 100% với trình tự các loài Sâm Ngọc Linh, Tam thất và Sâm bắc mỹ ở vùng gen bảo thủ như 18S. Kết quả kiểm tra vùng gen có độ biến đổi cao ITS và *psbA-trnH* cho thấy 4 mẫu đó có độ tương đồng cao nhất (99,44- 99,62 % và 99,72-100 %) với loài Sâm Ngọc Linh. Kết quả so

sánh trên chỉ ra mối quan hệ gần gũi giữa các mẫu PV21-24 với loài Sâm Ngọc Linh.

Các mẫu PS21-23, PB03-06 cho độ tương đồng 100% với loài Tam thất hoang ở vùng gen 18S. Kết quả so sánh ở vùng gen biến đổi như ITS cũng cho thấy các mẫu này có độ tương đồng cao nhất với loài Tam thất hoang (từ 99,44 - 99,81 %). Ngoài ra,

những vùng gen còn lại cũng chỉ ra quan hệ gần gũi của các mẫu trên với loài này (*matK*: 99,47 – 100 %, *psbA-trnH*: 99,43 – 100 %, *rbcL*: 99,81 - 100 %).

Những mẫu PB03-06 với những đặc điểm hình thái được nhận dạng là Sâm Vũ diệp cho kết quả phân tích mã vạch phân tử tương đồng rất cao với loài Tam thất hoang. Quy trình xác định mã vạch phân tử đã được thực hiện lại với mẫu Sâm Vũ diệp PB04 và một mẫu Tam thất hoang mới thuộc cùng khu vực phân bố (kết quả không được trình bày). Tuy nhiên, kết quả vẫn cho thấy trình tự của chúng hoàn toàn tương đồng với trình tự gen này ở các mẫu PB03-06 trước đây và giống với loài Tam thất hoang. Ngoài ra, khi nghiên cứu các vùng gen của loài Sâm Vũ diệp trên GenBank, chúng tôi nhận thấy độ tương đồng rất thấp (93,84 %) giữa 2 trình tự vùng ITS (mã số GenBank là U41679.1 và U41678.1), trong đó 1 trình tự cũng giống với loài Tam thất hoang. Kết quả nghiên cứu tương tự đã được đề cập trong một nghiên cứu trước đây của Wen và Zimmer (1996) về trình tự một số vùng gen của loài Sâm Vũ diệp. Trong nghiên cứu này, vùng gen ITS của Sâm Vũ diệp được xác định có xu hướng giống với các loài nào của chi Nhân sâm thuộc cùng khu vực phân bố trong tự nhiên, cụ thể là *P. major*, *P. omeiensis* và *P. stipuleanatus*. Khả năng cùng nguồn gốc nên Sâm Vũ diệp không được xem là một loài hay một thứ. Cho đến nay, vẫn chưa có công trình nào đưa ra được vùng gen có khả năng phân biệt được Sâm Vũ diệp với các loài khác thuộc chi Nhân sâm.

Giá trị trung bình chỉ số biến thiên nucleotide giữa các loài được thể hiện trong Bảng 4. Chỉ số biến thiên của chỉ thị ITS là cao nhất (0,032), tiếp theo lần lượt là *psbA-trnH* (0,020), *matK* (0,012), *rbcL* (0,006) và thấp nhất là 18S (0,001). Như vậy, vùng ITS là vùng mã vạch đa hình nhất trong năm vùng mã vạch được phân tích. Vùng ITS trong DNA ribosome đã được chứng minh là rất hiệu quả trong việc đánh giá mối quan hệ phát sinh giữa các loài trong các nghiên cứu về mã vạch phân tử ở thực vật. Zuo *et al.*, (2011) sau khi nghiên cứu các vùng gen *atpF-atpH*, *psbA-trnH*, *psbK-psbI*, *psbM-trnD*, *matK*, *rps16*, *rpoB*, *rpoC1*, *rbcL*, ITS và *nad1* trên quần thể sâm đã chỉ ra rằng vùng gen ITS có mức độ đa dạng di truyền cao nhất, có khả năng phân biệt loài đến 87,5 % và dưới loài đến 84,21 %. Nguyễn Thị Phương Trang *et al.*, (2011) đã phân tích mối quan hệ di truyền của Sâm Ngọc Linh với các loài trong chi Nhân sâm dựa trên kết quả phân tích thông tin di truyền đoạn gen ITS-rDNA. Kết quả cho thấy Sâm Ngọc Linh có mối quan hệ gần gũi với loài Tam thất và chúng khác nhau ở 18 nucleotide. Phan Kế Long *et al.*, (2014) đã tiến hành nghiên cứu đặc điểm di truyền của các mẫu sâm thu ở Lai Châu trên cơ sở phân tích trình tự nucleotide vùng gen *matK* và ITS-rDNA. Kết quả cho thấy trình tự vùng gen ITS của Sâm Lai Châu khác với Sâm Ngọc Linh từ 04 đến 05 nucleotide, chúng tạo nên nhánh riêng biệt và có mối quan hệ gần gũi với Tam thất gừng.

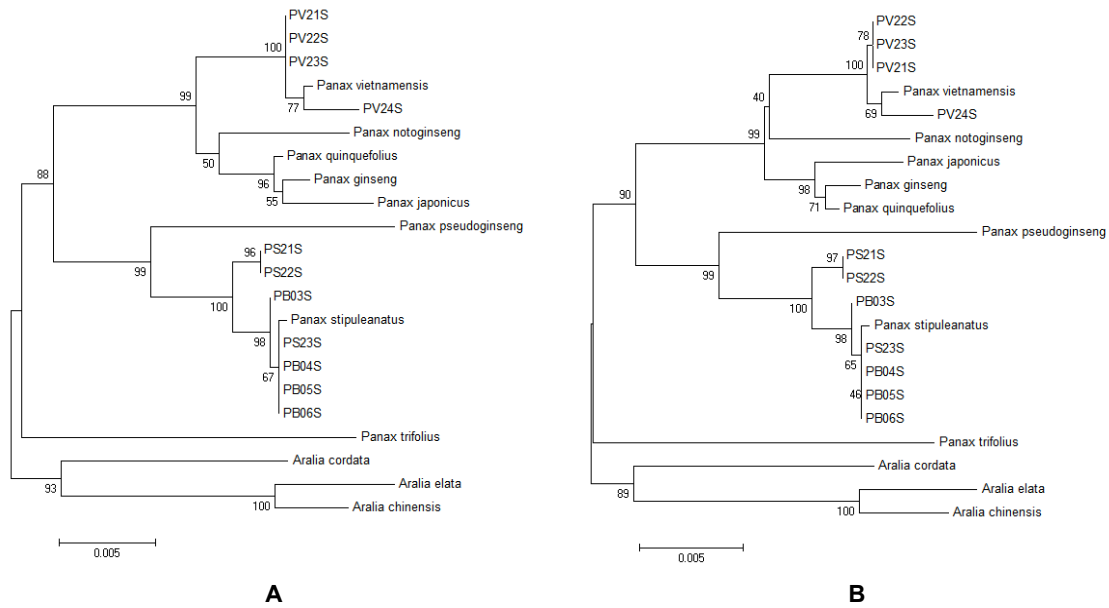
Bảng 4. Các số liệu về các chỉ thị mã vạch DNA sử dụng trong nghiên cứu.

	18S	ITS	<i>matK</i>	<i>psbA-trnH</i>	<i>rbcL</i>
Số lượng mẫu	11	11	11	11	11
Tỷ lệ nhân bản PCR thành công	100%	100%	100%	100%	100%
Tỷ lệ đọc trình tự thành công	100%	100%	100%	100%	100%
Kích thước đoạn gen so sánh (bp)	897	539	748	361	521
Khoảng cách giữa các trình tự khi so sánh theo cặp nucleotide	0,001 (0,000-0,003)	0,032 (0,000-0,072)	0,012 (0,000-0,024)	0,020 (0,000-0,041)	0,006 (0,000-0,014)

Phân tích quan hệ phát sinh loài

Dựa vào các trình tự đã xác định được và các trình tự tương đồng trên GenBank, chúng tôi đã tiến hành phân tích định danh loài dựa trên phương pháp xây dựng cây phân loại. Các trình tự Outgroup được sử dụng từ 3 loài thuộc cùng họ Araliaceae với chi Nhân sâm: *Aralia chinensis*, *Aralia elata*, *Aralia cordata*. Khi so sánh 5 cây phân loại sử dụng 5 chỉ thị DNA mã vạch

18S, ITS, *matK*, *psbA-trnH* và *rbcL* và kết hợp với kết quả so sánh trình tự nucleotide, chúng tôi nhận thấy rằng vùng trình tự ITS cho kết quả có độ tin cậy cao nhất, tiếp đến là vùng trình tự *psbA-trnH*, cho phép phân biệt Sâm Ngọc Linh với các loài sâm khác trên thế giới. Kết quả này cũng phù hợp với các nhóm nghiên cứu DNA barcode trên chi Nhân sâm (Wen và Zimmer, 1996; Zuo *et al.*, 2011; Vũ Huyền Trang *et al.*, 2013).



Hình 3. Biểu đồ hình cây về mối quan hệ phát sinh loài dựa trên trình tự nucleotide các vùng gen nghiên cứu được xây dựng bằng phương pháp Maximum-Likelihood (A) và Neighbour-Joining (B).

Quan sát biểu đồ hình cây về mối quan hệ phát sinh loài trên cơ sở so sánh trình tự các vùng gen nghiên cứu (Hình 3), chúng tôi nhận thấy các mẫu thu được thuộc 2 nhóm phân loại. Nhóm 1 gồm các mẫu PV21-24 cùng có quan hệ gần gũi nhất với loài Sâm Ngọc Linh với độ tin cậy cao (giá trị bootstrap là 100 %). Nhóm 2 gồm các mẫu PS21-23 và PB03-06 cùng có quan hệ gần gũi nhất với Tam thất hoang với giá trị bootstrap là 100%.

Với các biểu đồ hình cây về mối quan hệ phát sinh loài đã xây dựng được, chúng tôi kết luận trong 11 mẫu thực vật chi Nhân sâm có 4 mẫu (PV21-24) là Sâm Ngọc Linh, 3 mẫu (PS21-23) thuộc loài Tam thất hoang và 4 mẫu Sâm Vũ diệp (PB03-06) đã được định danh bằng hình thái nhưng có độ tương đồng cao về các mã vạch DNA nghiên cứu với loài Tam thất hoang. Để có thể tìm kiếm được các vùng mã vạch DNA đặc trưng hỗ trợ nhận dạng hai loài Tam thất hoang và Sâm Vũ diệp, toàn bộ hệ gen lục lạp của hai loài này đang được tiến hành giải mã và phân tích.

KẾT LUẬN

Bằng phương pháp phân tích trình tự các vùng mã vạch DNA 18S, ITS, *matK*, *psbA-trnH* và *rbcL*,

chúng tôi đã định tên khoa học cho 7 mẫu thực vật thuộc chi Nhân sâm, trong đó 4 mẫu là Sâm Ngọc Linh (*P. vietnamensis* Ha et Grushv.), 3 mẫu là Tam thất hoang (*P. stipuleanatus*) và 4 mẫu có hình thái của Sâm Vũ diệp nhưng có các mã vạch DNA giống với Tam thất hoang. Điều này phù hợp với một nghiên cứu trước đây về sự tương đồng trình tự vùng gen *ITS* của loài Sâm Vũ diệp với bất kỳ loài thực vật thuộc chi Nhân sâm nào cùng thuộc vùng phân bố. Đến nay, theo hiểu biết của chúng tôi, vẫn chưa có một công trình nào đưa ra được vùng mã vạch phân tử có khả năng định loại được Sâm Vũ diệp (*P. bipinnatifidus*). Việc định danh bằng phương pháp phân tử các loài thuộc chi Nhân sâm hỗ trợ các định loại hình thái và tạo cơ sở cho công tác quản lý đồng bộ và có hệ thống các loài dược liệu quý hiếm này, nâng cao giá trị của các loài sâm Việt Nam.

Lời cảm ơn: Công trình được thực hiện trong khuôn khổ đề tài: “Giải mã hệ gen lục lạp của Sâm Ngọc Linh (*Panax vietnamensis* Ha et Grushv.)”, mã số VAST02.01/16-17 (Viện Hàn lâm Khoa học và Công nghệ Việt Nam) và nhiệm vụ: “Tạo cơ sở dữ liệu mã vạch ADN cho các loài cây nghiên cứu” thuộc đề tài: “Xây dựng cơ sở dữ liệu mã vạch ADN (DNA barcode) cho một số loài cây lâm nghiệp gỗ lớn, lâm

sản ngoài gỗ có giá trị kinh tế” (2014-2017) (Bộ Nông nghiệp và Phát triển Nông thôn).

TÀI LIỆU THAM KHẢO

Chen X, Liao B, Song J, Pang X, Han J, Chen S (2013) A fast SNP identification and analysis of intraspecific variation in the medicinal *Panax* species based on DNA barcoding. *Gene* 530(1): 39-43.

Doyle JJ, Doyle J L (1990) Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Lã Đình Môi, Châu Văn Minh, Trần Văn Sung, Phạm Quốc Long, Phan Văn Kiệt, Trần Huy Thái, Trần Minh Hợi, Ninh Khắc Bản, Lê Mai Hương (2013) Họ Nhân sâm (Araliaceae Juss.) - Nguồn hoạt chất sinh học đa dạng và đầy triển vọng ở Việt Nam. *Hội nghị khoa học toàn quốc về sinh thái và tài nguyên sinh vật lần thứ 5*: 1152-1158.

Lee C, Wen J (2004) Phylogeny of *Panax* using chloroplast trnC–trnD intergenic region and the utility of trnC–trnD in interspecific studies of plants. *Mol Phylogenet Evol* 31: 894-903.

Lê Thị Thu Hiền, Hugo de Boer, Nông Văn Hải, Lê Thanh Hương, Nguyễn Mai Hương, Lars Bjork (2012) Mã vạch phân tử DNA và hệ thống dữ liệu mã vạch sự sống. *Tạp chí Công nghệ Sinh học* 10(3): 393-405.

Nguyễn Tập, Phạm Thanh Huyền, Lê Thanh Sơn, Ngô Đức Phương, Võ Văn Trại, Đinh Đoàn Long, Hoàng Thị Hòa (2007) Sử dụng chỉ thị ADN (RAPD-PCR) trong

nghiên cứu đa dạng di truyền và góp phần phân loại một số loài cây thuốc định hướng công tác bảo tồn và tiêu chuẩn hóa dược liệu ở Việt Nam. *Hội nghị Dược liệu toàn quốc lần thứ hai*: 288-301.

Nguyễn Văn Đạt, Trần Thị Phương Anh (2013) Bước đầu nghiên cứu xây dựng khóa định loại các chi trong họ Ngũ gia bì (Araliaceae) ở Việt Nam. *Hội nghị khoa học toàn quốc về Sinh thái và tài nguyên sinh vật lần thứ 5*: 44-51.

Phan Kế Long, Vũ Đình Duy, Phan Kế Lộc, Nguyễn Giang Sơn, Nguyễn Thị Phương Trang, Lê Thị Mai Linh, Lê Thanh Sơn (2014) Nghiên cứu đặc điểm di truyền của các mẫu sâm thu ở Lai Châu trên cơ sở phân tích trình tự nucleotide vùng gen *matK* và ITS-rRNA. *Tạp chí Công nghệ Sinh học* 12(2): 327-337.

Vũ Huyền Trang, Hoàng Đăng Hiếu, Chu Hoàng Hà (2013) Nghiên cứu xây dựng mã vạch DNA cho việc phân loại nhận dạng cây sâm Ngọc Linh. *Hội nghị khoa học Công nghệ sinh học toàn quốc*: 1100-1104.

Wen J, Zimmer EA (1996) Phylogeny and biogeography of *Panax* L. (the ginseng genus, Araliaceae): inferences from ITS sequences of nuclear ribosomal DNA. *Mol Phylogenet Evol* 6(2): 167-177.

Zhu S, Fushimi H, Cai S, Komatsu K (2003) Phylogenetic relationship in the genus *Panax*: inferred from chloroplast *trnK* gene and nuclear 18S rRNA gene sequences. *Planta Med* 69(7): 647-653.

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

APPLICATION OF DNA BARCODES IN IDENTIFICATION OF GINSENG SAMPLES IN THE GENUS *PANAX* L.

Le Thanh Huong¹, Nguyen Nhat Linh¹, Bui Manh Minh¹, Ha Hong Hanh¹, Huynh Thi Thu Hue¹, Nong Van Hai¹, Ha Van Huan², Le Thi Thu Hien¹

¹Institute of Genome Research, Vietnam Academy of Science and Technology

²Vietnam National University of Forestry

SUMMARY

Identification of species based on DNA sequences (DNA barcodes) is an aid to a taxonomic classification using morphological characteristics. DNA barcoding uses standard short genomic regions that are universally present in target lineages and has sufficient sequence variation to identify species in the genus. A variety of loci has been suggested as DNA barcodes for plants, including genes and non-coding regions in the nuclear and plastid genomes such as *psbA-trnH*, *matK*, *rbcL*, and ITS. In this study, we evaluated five potential DNA barcodes including 18S, ITS, *matK*, *psbA-trnH*, and *rbcL* for their ability to distinguish between species across our samples in the genus *Panax* L. Multiple alignments with 41 GenBank sequences selected from 9 species showed that the 18S region had the highest level of average similarity (99.87 %), followed by *rbcL*, *matK*, *psbA-trnH*, and ITS regions with lower levels of average similarities (99.27 %, 98.66 %, 96.82 % and 96.50 %, respectively). The phylogenetic trees showed that four of eleven ginseng samples were *Panax vietnamensis* Ha et Grushv., and three samples were *Panax stipuleanatus* with bootstrap values of 100 %. However, sequences

of all five screened loci in the last four samples, which were considered as *Panax bipinnatifidus* based on morphological characteristics, were highly similar to *Panax stipuleanatus* with the level of sequence similarity reached 99.81 % - 100 %. In addition, our results showed that of five investigated DNA regions, ITS, and *psbA-trnH* were the most promising barcodes that could identify *P. vietnamensis* Ha et Grushv. and *P. stipuleanatus* species within the genus *Panax* L.

Keywords: *DNA barcodes, Panax* L. genus, *Panax vietnamensis* Ha et Grushv., *Panax stipuleanatus*, *Panax bipinnatifidus*