

# PHÂN TÍCH TRÌNH TỰ DNA BARCODE CỦA MỘT SỐ MẪU ĐÌNH LĂNG ĐƯỢC THU THẬP TẠI VIỆT NAM

Trịnh Việt Nga<sup>1,2</sup>, Nguyễn Cao Kiệt<sup>1</sup>, Bùi Minh Trí<sup>1</sup>,  
Phạm Thị Minh Tâm<sup>1</sup>, Nguyễn Hữu Hồ<sup>2</sup>, Huỳnh Văn Biết<sup>1</sup>

## TÓM TẮT

Nghiên cứu này nhằm mục đích phân tích mã vạch DNA của 18 mẫu giống đình lăng được thu thập ở Việt Nam. Các mẫu đình lăng được thu thập từ nhiều vùng khác nhau ở Việt Nam. Mẫu DNA tinh sạch thu thập từ các mẫu giống sau đó được khuếch đại với các mồi *matK*, *rbcl* và *trnH-psbA*. Các sản phẩm khuếch đại sau đó được giải trình tự để phân tích những thay đổi ở mức độ các nucleotide. Các kết quả chỉ ra rằng tất cả các mẫu đình lăng thu thập có mức độ tương đồng cao và chỉ có khác biệt ở một số nucleotide. Kết quả cũng chỉ ra rằng vùng gen *matK* và *rbcl* của đình lăng có tính bảo tồn cao. Trong khi đó vùng *trnH-psbA* thường có nhiều biến đổi.

Từ khóa: Đình lăng, DNA barcode, *matK*, *rbcl*, *trnH-psbA*, khác biệt di truyền.

## 1. BẬT VẦN BÉ

Cây đình lăng có nguồn gốc từ vùng đảo Polynesie thuộc Thái Bình Dương và phân bố rải rác ở vùng nhiệt đới và cận nhiệt đới (Võ Văn Chí và Trần Hợp, 1999). Cây đình lăng thuộc họ Ngũ gia bì (Araliaceae) cùng họ với Nhân sâm, là loại cây khá quen thuộc (Võ Duy Huân, 1998).

Theo phân loại của Phạm Hoàng Hộ (2002), đình lăng bao gồm nhiều loài được trồng ở các vùng khác nhau. Mỗi loại lại có thành phần dược chất và giá trị dược liệu khác nhau. Việc phát hiện và so sánh trình tự gen, sự tương đồng và khác biệt các nucleotide của các vùng ở các loài đình lăng sẽ có giá trị hỗ trợ cho các nghiên cứu theo hướng tiêu chuẩn hóa trong kiểm định nguyên liệu dược, góp phần vào việc xác định từng loài đình lăng để tránh nhầm lẫn giống cũng như các nguồn nguyên liệu thuốc còn gặp nhiều khó khăn. Đặc biệt là góp phần xác định được các loài có đặc điểm hình thái tương tự nhau được trồng ở các vùng khác nhau. Vì vậy, các phương pháp so sánh sự trình tự các vùng gen hỗ trợ xác định các loài thực vật dựa trên hệ gen (DNA) đặc thù của chúng là nhiệm vụ cấp thiết trong công tác bảo

tồn, khai thác và phát triển nguồn gen cây đình lăng. Trong các phương pháp phổ biến hiện nay, kỹ thuật DNA barcode là một kỹ thuật hiệu quả, không chỉ cho phép kiểm tra trình tự các vùng gen của các mẫu nhanh chóng mà còn giúp mở rộng nghiên cứu. Trong nghiên cứu về đa dạng sinh học thực vật, DNA barcode rất hữu ích trong việc tìm mối quan hệ giữa các mẫu mặc dù chúng không giống nhau về hình thái. Ngoài ra, kỹ thuật này còn đóng góp vào nỗ lực xác định mã vạch cho tất cả các loài sinh vật trên trái đất nói chung. Trên cơ sở đó, *matK*, *rbcl* và *trnH-psbA* là ba vùng gen thực vật được ứng dụng nghiên cứu trong số các vùng gen thường được nghiên cứu ở kỹ thuật này (Kress và ctv, 2005). Nghiên cứu này được tiến hành nhằm kiểm tra sự tương đồng và khác biệt của ba vùng trên, góp phần vào nhận diện và phân loại các nhóm giống đình lăng.

## 2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

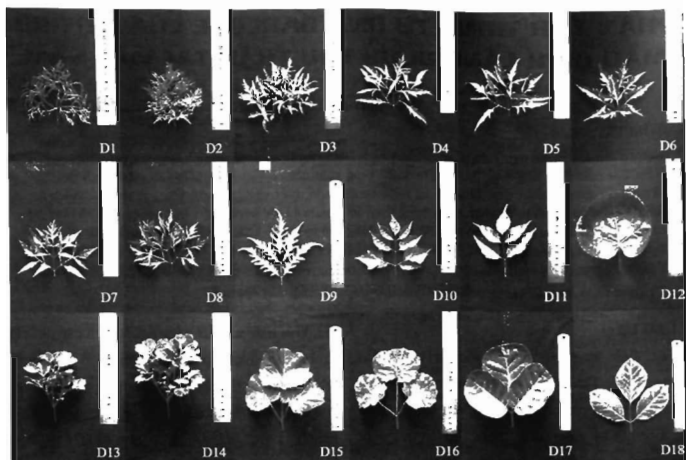
### 2.1. Vật liệu

Mười tám mẫu đình lăng đã được thu thập ở các vùng khác nhau dựa trên các đặc điểm hình thái khác nhau của các giống đình lăng như: đình lăng lá nhỏ, đình lăng lá răng, đình lăng lá trổ, đình lăng lá đĩa và đình lăng lá tròn (Hình 1, bảng 1).

<sup>1</sup> Trường Đại học Nông Lâm TP. Hồ Chí Minh

<sup>2</sup> Viện Sinh học Nhiệt đới

Email: trnhvietnga@gmail.com



Hình 1. Các mẫu lá đinh lăng D1 – D18 trong nghiên cứu

Bảng 1. Đặc điểm hình thái của 18 mẫu đinh lăng đã được sử dụng trong nghiên cứu

STT	Tên mẫu	Ký hiệu mẫu	Điểm thu mẫu	Mô tả	Phân nhóm theo kiểu hình
1	Đinh lăng lá nhỏ	D1	Nguồn giống của Viện Sinh học Nhiệt đới	Lá nhuyễn, xẻ thùy, màu xanh lá	A
2	Đinh lăng lá nhỏ	D2	Tiền Giang	Lá nhuyễn, xẻ thùy, màu xanh lá	A
3	Đinh lăng lá nhỏ	D3	Khánh Hòa	Lá nhỏ, xẻ thùy, màu xanh lá	B
4	Đinh lăng lá nhỏ	D4	Viện Dược liệu, Hà Nội	Lá nhỏ, xẻ thùy, màu xanh lá	B
5	Đinh lăng lá nhỏ	D5	Nam Định	Lá nhỏ, xẻ thùy, màu xanh lá	B
6	Đinh lăng lá nhỏ	D6	Hải Dương	Lá nhỏ, xẻ thùy, màu xanh lá	B
7	Đinh lăng lá nhỏ	D7	Thủ Đức, TP. HCM	Lá nhỏ, xẻ thùy, màu xanh lá	B
8	Đinh lăng lá nhỏ	D8	Đồng Nai	Lá nhỏ, xẻ thùy, màu xanh lá	B
9	Đinh lăng lá to	D9	Hải Dương	Lá có xẻ thùy nhiều răng cưa, màu xanh lá	C
10	Đinh lăng lá to	D10	Côn Đảo	Lá lớn, màu xanh lá	D
11	Đinh lăng lá to	D11	Nguồn giống của Viện Sinh học Nhiệt đới	Lá lớn, màu xanh lá	D
12	Đinh lăng lá đĩa	D12	Thủ Đức, TP. HCM	Lá hình đĩa, màu xanh lá	E
13	Đinh lăng lá răng	D13	Viện Dược liệu, Hà Nội	Lá hai lần kép có răng cưa, màu xanh lá	F

14	Định lăng lá rỗng	D14	Côn Đảo	Lá hai lần kép có răng cưa, màu xanh lá	F
15	Định lăng lá tròn	D15	Bình Thuận	Lá to tròn, màu xanh lá	G
16	Định lăng lá tròn	D16	Bình Thuận	Lá to tròn, màu xanh lá, có viền trắng	G
17	Định lăng lá tròn	D17	Bình Thuận	Lá to tròn, màu xanh	G
18	Định lăng lá trở	D18	Viện Dược liệu, Hà Nội	Lá hình bầu dục, có viền màu bạc răng cưa	H

2.2. Phương pháp

2.2.1. Ly trích DNA tổng số

Nghiên 50 mg mẫu lá định lăng trong 600 µl dung dịch ly trích (NaCl 0,125 M; HCl 0,01 M; EDTA 0,01 M; SDS 0,5%) sau đó vortex nhẹ và ủ mẫu ở 65°C trong 1 giờ, thêm 600 Phenol:Chloroform:Isomyly Alcohol (25:24:1), sau đó vortex và ly tâm 12.000 vòng/phút trong 10 phút, hút dịch nổi. Thực hiện thêm một lượng thể tích Chloroform:Isomyly Alcohol (24:1) bằng thể tích dịch nổi đã hút, vortex nhẹ, ly tâm 12.000 vòng/phút trong 10 phút, hút dịch nổi. Sau đó, thêm một lượng thể tích Isopropanol bằng thể tích dịch đã hút. Tiếp theo, đảo nhẹ, ủ 4°C trong 15 phút. Tiếp đến, ly tâm 12.000 vòng/phút trong 10 phút, bỏ dịch. Sau đó, rửa ethanol 70% lạnh 3 lần, ly tâm 12.000 vòng/phút trong 3 phút, bỏ dịch, phơi mẫu, thêm 50 µl TE 1X. Sau đó, bảo quản lạnh -20°C. Thực hiện kiểm tra chất lượng của các DNA thông qua chạy điện di gel agarose 1%, tỉ lệ A<sub>260</sub>/A<sub>280</sub> và nồng độ DNA trên máy BioDrop (BioDrop µLITE, Anh).

2.2.2. Phản ứng PCR và điện di

Sử dụng các primer vùng *matK* với primer *matK172F* 5'-CCCRTYCATCTGGAAATCTT

GGTTC-3' và *matK1248R* 5'-GCTRTRATAATGAGAAAGATTTCTGC-3' (Jing và ctv, 2011), vùng *rbcL* với primer *rbcL1F* 5'-ATGTCCACCACAAACAGAGACTAA-3' (Sameera và ctv, 2011) và *rbcL1460R* 5'-CTTTTAGTAAAAGATTGGGCCGAG-3' (Olmstead và ctv, 1992), vùng *trnH-psbA* với primer *psbAF* 5'-GTATGATGAACGTAATGCTC-3' và *trnHR* 5'CGCGCATGGTGGATTCAAATC-3' (Srirama R và ctv, 2010).

Thành phần phản ứng: 15 µl master mix 2X; 0,75 µl mỗi primer có nồng độ 10 µM; 12,5 µl nước nuclease free; 1 µl DNA mẫu nồng độ 50 ng/µl. Tổng thể tích 30 µl.

Chu trình nhiệt cho phản ứng PCR của cặp primer *matK*

Biến tính ban đầu	95°C	5 phút	
Biến tính	95°C	30 giây	} 35 chu kỳ
Gắn mồi	52°C	30 giây	
Kéo dài chuỗi	72°C	50 giây	
Ổn định	72°C	7 phút	
Kết thúc phản ứng	4°C	-	

Chu trình nhiệt cho phản ứng PCR của cặp primer *rbcL*

Biến tính ban đầu	95°C	5 phút	
Biến tính	95°C	1 phút	} 35 chu kỳ
Gắn mồi	57°C	1 phút	
Kéo dài chuỗi	72°C	1 phút	
Ổn định	72°C	7 phút	
Kết thúc phản ứng	4°C	-	

Chu kỳ nhiệt cơ bản cho phản ứng PCR của cặp primer *trnH-psbA*

Biến tính ban đầu	95°C	5 phút	
Biến tính	95°C	1 phút	} 35 chu kỳ
Gắn mồi	52°C	1 phút	
Kéo dài chuỗi	72°C	1 phút	
Ổn định	72°C	7 phút	
Kết thúc phản ứng	4°C	-	

Sản phẩm PCR được phát hiện bởi điện di trên gel agarose 1,5% trong thời gian 30 phút, nhuộm với gel red và chụp hình trên máy soi gel UV Transilluminator.

2.2.3. Phân tích và đánh giá mối quan hệ di truyền của 18 mẫu định lăng

Phản ứng PCR của từng mẫu phân tích được tiến hành 3 lần lặp lại, các mẫu có sản phẩm DNA khuếch đại có kích thước bằng nhau được sử dụng cho việc giải trình tự. Lựa chọn 1 sản phẩm PCR ở mỗi mẫu để tinh sạch và giải trình tự hai chiều nhằm đảm bảo tính chính xác của các nucleotide. Các trình tự được hiệu chỉnh bằng phần mềm Bioedit, so sánh với nhau bằng công cụ Clustal omega (European Bioinformatics Institute, 2017) và so sánh với các trình tự tương đồng có sẵn trên cơ sở dữ liệu GenBank

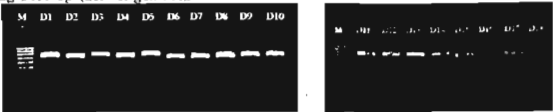
thông qua việc sử dụng công cụ BLAST. Xây dựng cây phát sinh loài bằng phần mềm MEGA 7.1 với thuật toán Construct/Test Maximun Likelihood Tree from DNA sequences với hệ số bootstrap 1000.

### 3. KẾT QUẢ VÀ THẢO LUẬN

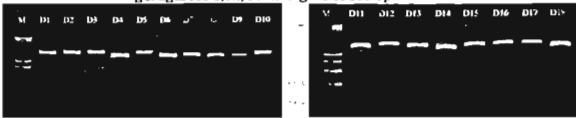
#### 3.1. Khuếch đại gen

Kết quả kiểm tra sản phẩm khuếch đại gen *matK*, *rbcL* và *trnH-psbA* trên gel agarose (hình 2, hình 3, hình 4) cho thấy gen *matK*, *rbcL* và *trnH-psbA* được khuếch đại tốt ở tất cả các mẫu đỉnh lã nghiên cứu. Các băng thu được đều sáng và rõ, tạo ra một sản phẩm duy nhất với kích thước khoảng 700 bp (đối với gen *matK* khi so sánh với thang DNA chuẩn 100 bp), kích thước khoảng 1400 bp (đối với gen *rbcL* khi so

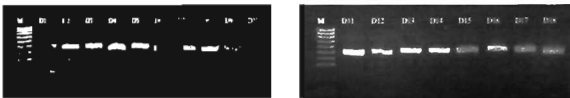
sánh với thang DNA chuẩn 1 kb), kích thước khoảng 500 bp (đối với gen *trnH-psbA* khi so sánh với thang DNA chuẩn 100 bp), phù hợp để thực hiện tiếp tục bước giải trình tự. Kết quả khuếch đại vùng gen *matK* tốt hơn nghiên cứu của Yu và ctv (2011) khi sử dụng cặp primer tương tự chỉ đạt 93,1% ở nhóm thực vật hạt kín, tương tự vùng gen *rbcL* được khuếch đại tốt hơn nghiên cứu của Sameera và ctv (2011) chỉ đạt 88% ở 26 loài thực vật khác nhau (bao gồm 14 họ) từ Ả Rập Saudi khi dùng một primer tương tự. Đồng thời, kết quả khuếch đại vùng gen *trnH-psbA* tốt tương ứng nghiên cứu của Srirama và ctv (2010) đạt 100% ở nhóm loài *Phyllanthus*.



Hình 2. Các băng sản phẩm PCR cặp primer *matK* của các mẫu lá đỉnh lã D1 – D18 được điện di trên gel agarose 1,5%, M: thang DNA 100 bp



Hình 3. Các băng sản phẩm PCR cặp primer *rbcL* của các mẫu lá đỉnh lã D1 – D18 được điện di trên gel agarose 1,5%, M: thang DNA 1 kb



Hình 4. Các băng sản phẩm PCR cặp primer *trnH-psbA* của các mẫu lá đỉnh lã D1 – D18 được điện di trên gel agarose 1,5%, M: thang DNA 100 bp

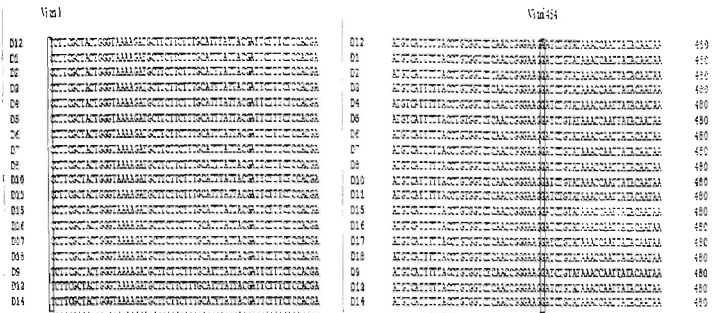
#### 3.2. Phân tích và đánh giá mối quan hệ di truyền của 18 mẫu đỉnh lã

Trên cơ sở kết quả khuếch đại các vùng DNA barcode *matK*, *rbcL*, *trnH-psbA*, trình tự 18 mẫu đỉnh lã đã được phân tích sự khác biệt di truyền.

Gen *matK* có hầu hết trong thực vật, có tính phổ quát và có tính đa dạng hơn những gen khác có trong lục lạp, do vậy gen *matK* trở thành gen chỉ thị quan trọng để giúp phân loại thực vật, đặc biệt là cho sự nghiên cứu giữa các loài (Asahina và ctv, 2010; Sharma và ctv, 2012). Trình tự vùng *matK* sau khi xử lý có độ dài 648 bp. Kết quả so sánh bất cấp cho thấy,

trình tự của 18 mẫu đỉnh lã có độ tương đồng cao, chỉ khác biệt ở hai vị trí.

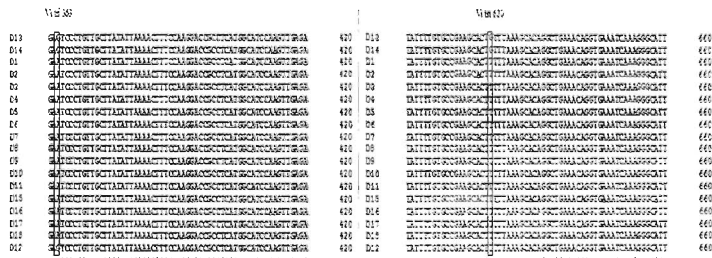
Tại vị trí khác biệt 1 có thể phân chia các mẫu đỉnh lã thành hai nhóm bao gồm nhóm thứ nhất (D9, D12, D13, D14) có chứa nucleotide T và nhóm thứ hai với các mẫu còn lại chứa nucleotide C. Tại vị trí khác biệt 454 có thể phân thành hai nhóm bao gồm nhóm thứ nhất (D12) chứa nucleotide A và nhóm thứ hai với các mẫu còn lại chứa nucleotide G. Dựa vào hai vị trí trên, có thể thấy rằng D12 chứa 2 nucleotide T, A khác biệt với các nhóm còn lại chứa 2 nucleotide T, G (D9, D13, D14) và nhóm chứa 2 nucleotide C, G (các mẫu còn lại).



Hình 5. Vị trí các điểm nucleotide khác biệt trên vùng gen *matK*

Hollingsworth (2009) cho rằng *rbcL* là một trong những trình tự gen bền vững nhất cho các nghiên cứu DNA barcode ở thực vật. Trình tự vùng *rbcL* sau

khí xử lý có độ dài trong khoảng 1.375 bp. Kết quả so sánh khả năng cho thấy, trình tự của 18 đoạn trình tự có độ tương đồng cao, chi khác biệt ở bốn vị trí.

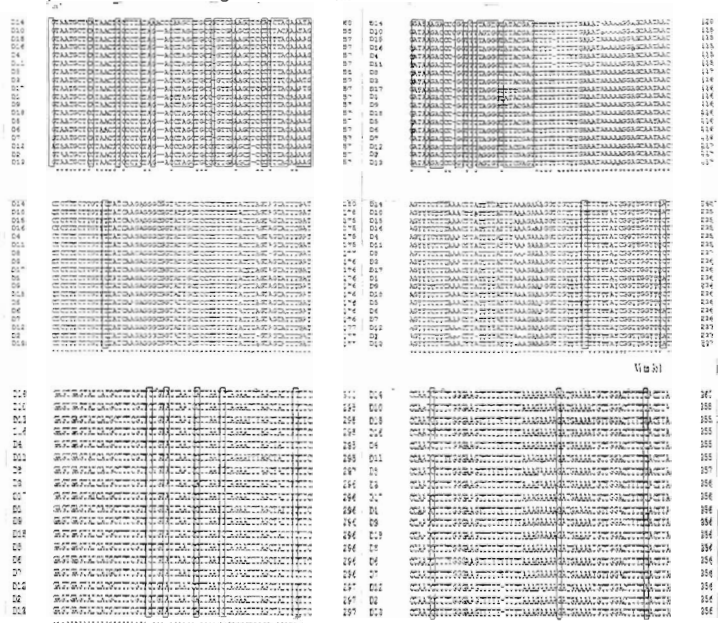


Hình 6. Vị trí các điểm nucleotide khác biệt trên vùng gen *rbcl*

Tại vị trí khác biệt 363 có thể phân chia các mẫu định lượng thành hai nhóm bao gồm nhóm thứ nhất (D12, D13, D14) có chứa nucleotide G và nhóm thứ hai với các mẫu còn lại chứa nucleotide A. Tại vị trí khác biệt 620 có thể phân thành hai nhóm bao gồm nhóm thứ nhất (D13, D14) chứa nucleotide G và nhóm thứ hai với các mẫu còn lại chứa nucleotide T. Tại vị trí khác biệt 665 có thể phân chia các mẫu định lượng thành hai nhóm bao gồm nhóm thứ nhất (D13, D14) có chứa nucleotide C và nhóm thứ hai với các mẫu còn lại chứa nucleotide T. Tại vị trí khác biệt 895 có thể phân thành hai nhóm bao gồm nhóm thứ nhất

(D12, D13, D14) chứa nucleotide C và nhóm thứ hai với các mẫu còn lại chứa nucleotide G. Như vậy dựa vào 4 vị trí khác biệt, có thể phân chia 3 nhóm gồm (D12), (D13, D14) và nhóm các mẫu còn lại.

Gen *trnH-psbA* là vùng gen có kích thước trung bình khoảng 450 bp, nhưng thay đổi từ 296 đến 1120 bp, *trnH-psbA* được chứng minh là có khả năng xác định loài cao (Vijayan K, 2010). Trình tự gen *trnH-psbA* sau khi xử lý có độ dài 504 bp. Kết quả so sánh bất cặp cho thấy, trình tự của 18 mẫu khác nhau 68 vị trí nucleotide.



Hình 7. Vị trí các điểm nucleotide khác biệt trên vùng gen *trnH-psbA* từ vị trí 1 - 356

Tại vị trí khác biệt 1 có thể chia thành hai nhóm gồm nhóm thứ nhất (D12) chứa nucleotide A và nhóm thứ hai với các mẫu còn lại chứa nucleotide G. Tại vị trí khác biệt 351, có thể chia thành thành hai

nhóm gồm nhóm thứ nhất (D12) chứa nucleotide A và nhóm thứ hai với các mẫu còn lại chứa nucleotide T.

014	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	014	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	473
015	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	015	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	473
016	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	016	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
017	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	017	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
018	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	018	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
019	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	019	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
020	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	020	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
021	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	021	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
022	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	022	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
023	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	023	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
024	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	024	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
025	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	025	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
026	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	026	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
027	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	027	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
028	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	028	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
029	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	029	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
030	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	030	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
031	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	031	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
032	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	032	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474
033	GTAAATTAACTAACTAACTTTACTGTGAAA--MAGAAAGCTAAGAAAGAAAT	459	033	GATGACGCTGAAAATTTACTTTTGGATGAAAGAAAAGGAGGAGGAGGAGG	474

459

014	TTTCCCAAGTGGATCAGGCGCGAATTT	504
015	TTTCCCAAGTGGATCAGGCGCGAATTT	504
016	TTTCCCAAGTGGATCAGGCGCGAATTT	504
017	TTTCCCAAGTGGATCAGGCGCGAATTT	504
018	TTTCCCAAGTGGATCAGGCGCGAATTT	504
019	TTTCCCAAGTGGATCAGGCGCGAATTT	504
020	TTTCCCAAGTGGATCAGGCGCGAATTT	504
021	TTTCCCAAGTGGATCAGGCGCGAATTT	504
022	TTTCCCAAGTGGATCAGGCGCGAATTT	504
023	TTTCCCAAGTGGATCAGGCGCGAATTT	504
024	TTTCCCAAGTGGATCAGGCGCGAATTT	504
025	TTTCCCAAGTGGATCAGGCGCGAATTT	504
026	TTTCCCAAGTGGATCAGGCGCGAATTT	504
027	TTTCCCAAGTGGATCAGGCGCGAATTT	504
028	TTTCCCAAGTGGATCAGGCGCGAATTT	504
029	TTTCCCAAGTGGATCAGGCGCGAATTT	504
030	TTTCCCAAGTGGATCAGGCGCGAATTT	504
031	TTTCCCAAGTGGATCAGGCGCGAATTT	504
032	TTTCCCAAGTGGATCAGGCGCGAATTT	504
033	TTTCCCAAGTGGATCAGGCGCGAATTT	504

Hình 8. Vị trí các điểm nucleotide khác biệt trên vùng gen *trnH-psbA* từ vị trí 357 – 504

Tại vị trí khác biệt 462, có thể chia thành thành hai nhóm gồm nhóm thứ nhất (D12) chứa nucleotide A và nhóm thứ hai với các mẫu còn lại chứa nucleotide T. Tương tự, tại vị trí khác biệt 473, có thể chia thành thành hai nhóm thứ nhất (D12) chứa nucleotide G và nhóm thứ hai với các mẫu còn lại chứa nucleotide C. Tại vị trí khác biệt 499, có thể chia thành hai nhóm gồm nhóm thứ nhất (D1, D2) chứa nucleotide G và nhóm thứ hai với các mẫu còn lại chứa nucleotide T. Các vị trí khác có sự khác biệt nucleotide, tuy nhiên vẫn chưa có sự đặc trưng cho từng mẫu, nhóm mẫu. Nhìn chung gen *trnH-psbA* có sự khác biệt nucleotide nhiều hơn so với hai vùng gen *matK* và *rbcL*.

Kiểm tra bằng công cụ BLAST trên ngân hàng GenBank với trình tự vùng *matK* cho thấy cả 18 mẫu nói trên đều thuộc các loài chi đình lẵng *Polyscias* có trong GenBank với mức độ bao phủ và độ tương đồng từ 99 – 100%. Mười bốn mẫu đình lẵng bao gồm D1, D2, D3, D4, D5, D6, D7, D8, D18, D10, D11, D15, D16, D17 có độ tương đồng và độ bao phủ 100% với giống *Polyscias fruticosa* voucher VTN993 và giống *Polyscias balfouriana* voucher Hosam00272. Ba mẫu D13, D14, D9 có độ tương đồng 100% và độ bao phủ 99% với giống *Polyscias fruticosa* voucher VTN993 và giống *Polyscias balfouriana* voucher Hosam00272. Riêng mẫu D12 có độ tương đồng 99,85% và độ bao phủ 99% với giống *Polyscias fruticosa* voucher VTN993 và giống *Polyscias balfouriana* voucher

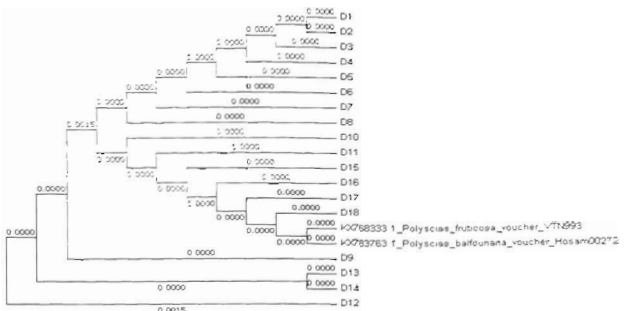
Hosam00272. Như vậy, vùng gen *matK* của các mẫu đình lẵng trong nghiên cứu có trình tự gen tương đồng cao với hai giống *Polyscias fruticosa* voucher VTN993 và *Polyscias balfouriana* voucher Hosam00272. Điều này chứng tỏ, vùng gen *matK* có tính bảo tồn, ít có sự biến đổi giữa các mẫu đình lẵng trong nghiên cứu và cần kết hợp các chỉ thị phân tử khác để hỗ trợ việc đánh giá sự khác biệt di truyền chuyên sâu hơn.

Trình tự vùng *rbcL* sau khi xử lý có kích thước là 1375 bp. Tuy nhiên, các trình tự gen *rbcL* nhóm *Polyscias* trên GenBank có kích thước khoảng 500 – 900 bp, chỉ có một loài *Polyscias guilfoylei* (U50251.1) có kích thước 1428 bp. Do đó khi thực hiện BLAST, chỉ xuất hiện sự tương đồng với các nhóm thuộc họ Araliaceae và loài *Polyscias guilfoylei*. Điều này chứng tỏ, cơ sở dữ liệu vùng gen *rbcL* thuộc chi *Polyscias* trên GenBank chưa cao, không đủ cơ sở để so sánh với các mẫu trong nghiên cứu.

Thực hiện BLAST vùng gen *trnH-psbA* của 18 mẫu trên cho thấy 18 mẫu có độ bao phủ 92% và độ tương đồng 95,72% – 99,15% với giống *Polyscias australiana* voucher Wen 1070 (JX106123.1), có độ bao phủ 92% và độ tương đồng 95,09% – 98,09% với giống *Polyscias* sp. Wen 1070 (JX106126.1), có độ bao phủ 83% – 84% và độ tương đồng 96,74% – 99,29% với loài *Polyscias fruticosa* (HQ220595.1). Trong cơ sở dữ liệu của GenBank vẫn còn các loài *Polyscias*

khác đã công bố trình tự gen *trnH-psbA*, tuy nhiên kết quả BLAST chỉ xuất hiện các giống loài trên và họ Araliaceae cho thấy 18 mẫu trên có sự tương đồng

cao với ba giống loài nói trên trên thế giới và dữ liệu các giống loài định làng tại Việt Nam vẫn còn nhiều hạn chế.

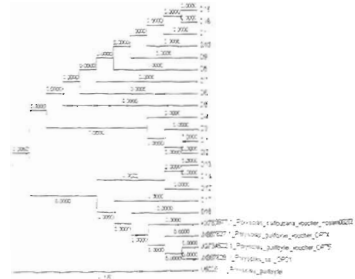


Hình 9. Cây phát sinh loài dựa vào kết quả giải trình tự gen *matK*

Dựa vào cây phân loại vùng gen *matK* (Hình 9) cho thấy các mẫu có xu hướng xếp gần nhau tương ứng khi phân loại theo hình thái (Bảng 1) là các mẫu lá nhỏ (lá nhuyễn) (D1, D2), mẫu lá nhỏ (D3, D4, D5, D6, D7, D8), mẫu lá to (xẻ thùy) (D9), mẫu lá to (không xẻ thùy) (D10, D11), mẫu lá đĩa (D12), mẫu lá răng (D13, D14), mẫu lá tròn (D15, D16, D17), mẫu lá trổ (D18). Tuy nhiên, giá trị khác biệt di truyền giữa 18 mẫu không cao. Hai giống *Polyscias fruticosa* voucher VTN993 và giống *Polyscias balfouriana* voucher Hosam00272 trên GenBank tương đồng cao, không khác biệt di truyền. Từ đó cho thấy, 18 mẫu nghiên cứu đều có sự tương đồng cao ở vùng gen *matK*, giá trị di truyền ít bị biến đổi bởi hình thái mẫu và khu vực thu mẫu, đồng thời, cả 18 mẫu đều tương đồng cao với bốn giống *Polyscias guilfoylei* voucher OP75, *Polyscias* sp. OP31, *Polyscias guilfoylei* voucher OP74 và *Polyscias balfouriana* voucher Hosam00272 trên GenBank. Riêng loài *Polyscias guilfoylei* trên GenBank có sự khác biệt nhất định với các mẫu và giống con lại.

*guilfoylei* voucher OP74 và *Polyscias balfouriana* voucher Hosam00272 trên GenBank có sự tương đồng cao. Từ đó cho thấy, 18 mẫu nghiên cứu đều có sự tương đồng cao ở vùng gen *rbcL*, giá trị di truyền ít bị biến đổi bởi hình thái mẫu và khu vực thu mẫu, đồng thời, cả 18 mẫu đều tương đồng cao với bốn giống *Polyscias guilfoylei* voucher OP75, *Polyscias* sp. OP31, *Polyscias guilfoylei* voucher OP74 và *Polyscias balfouriana* voucher Hosam00272 trên GenBank. Riêng loài *Polyscias guilfoylei* trên GenBank có sự khác biệt nhất định với các mẫu và giống con lại.

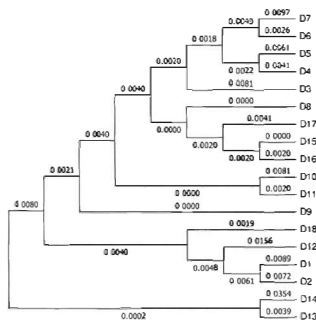
Tương tự cây phân loại vùng gen *rbcL* (Hình 10) có các mẫu lá nhỏ (lá nhuyễn) (D1, D2), mẫu lá nhỏ (D3, D3, D5, D6, D7, D8), mẫu lá to (xẻ thùy) (D9), mẫu lá to (không xẻ thùy) (D10, D11), mẫu lá đĩa (D12), mẫu lá tròn (D16, D17) và mẫu lá trổ (D18), có xu hướng xếp gần nhau tương ứng khi phân loại theo hình thái (Bảng 1). Riêng các mẫu lá răng (D13, D14) xếp gần mẫu lá tròn (D15). Nhìn chung, giá trị khác biệt di truyền không cao giữa 18 mẫu. Bốn giống *Polyscias guilfoylei* voucher OP75, *Polyscias* sp. OP31, *Polyscias*



Hình 10. Cây phát sinh loài dựa vào kết quả giải trình tự gen *rbcL*



Đồng thời, kết quả cây phân loài trình tự *trnH-psbA* (Hình 11) có sự tương đồng với vùng gen *matK* và *rbcL*. Các nhóm mẫu xếp cạnh nhau tương tự phân loại theo hình thái, nhóm mẫu lá nhỏ (lá nhuyễn) (D1, D2), mẫu lá nhỏ (D3, D4, D5, D6, D7, D8), mẫu lá to (xẻ thùy) (D9), mẫu lá to (không xẻ thùy) (D10, D11), mẫu lá răng (D13, D14), mẫu lá tròn (D15, D16, D17). Riêng mẫu lá đĩa (D12) và mẫu lá trổ (D18) xếp cạnh mẫu lá nhỏ (lá nhuyễn) (D1, D2).



Hình 11. Cây phát sinh loài dựa vào kết quả giải trình tự gen *trnH-psbA*

Vùng gen *matK* trên cây đinh lăng ở các mẫu có sự tương đồng cao khác với kết quả nghiên cứu của Nguyễn Thị Thanh Nga (2012) về một số loài cây dược liệu thuộc chi Đẳng Sâm (*Codonopsis* sp.) có vùng gen *matK* tương đối khác biệt (61 điểm đa hình trên 871 bp). Vùng gen *rbcL* trên cây đinh lăng ở các mẫu có sự tương đồng cao tương tự nghiên cứu của Hoàng Đăng Hiếu (2012) về tập đoàn cây dó bầu (*Aquilaria* sp.) có vùng gen *rbcL* với độ tương đồng cao (99,0 - 100%). Tuy nhiên, vùng gen *trnH-psbA* cho kết quả khác kết quả nghiên cứu của Hoàng Đăng Hiếu về tập đoàn cây dó bầu (*Aquilaria* sp.) - có độ tương đồng rất cao (99,1-100%).

Như vậy, cả hai vùng gen *matK* và *rbcL* đều có tính bảo tồn và ít bị biến đổi bởi hình thái và khu vực thu các mẫu đinh lăng nghiên cứu. Vùng gen *trnH-psbA* chưa thể hiện rõ sự tương quan giữa trình tự gen và hình thái. Tuy nhiên, vùng gen *trnH-psbA* có sự khác biệt nhiều hơn về trình tự gen giữa các mẫu đinh lăng phân tích so với trình tự gen *matK* và *rbcL*.

Đây có thể là vùng gen DNA barcode tiềm năng để phát triển trở thành marker phân tử trong việc phân biệt loài đinh lăng.

#### 4. KẾT LUẬN

Kết quả khuếch đại và phân tích 3 vùng DNA barcode *matK*, *rbcL*, *trnH-psbA* ở 18 mẫu đinh lăng cho thấy có sự giống nhau và khác nhau về trình tự nucleotide giữa các mẫu đã thu thập. Vùng gen *matK* có sự khác biệt nucleotide vị trí 1 và vị trí 454. Vùng gen *rbcL* có sự khác biệt nucleotide ở vị trí 363, vị trí 620, vị trí 665, vị trí 895. Vùng gen *trnH-psbA* có sự khác biệt nucleotide nhiều vị trí hơn hai vùng *matK* và *rbcL*. Trong đó, chú ý các vị trí 1, 351, 462, 473, 499 có chứa các nucleotide đặc trưng cho một nhóm mẫu. Sự khác biệt trình tự vùng gen *trnH-psbA* là cơ sở để thực hiện các nghiên cứu về marker phân tử trong việc phân biệt các loài đinh lăng.

Kết quả cây phát sinh loài của vùng gen *matK*, *rbcL* có tính bảo tồn, có tính tương đồng cao và ít sự khác biệt giữa các mẫu đinh lăng trong nghiên cứu. Vùng gen *trnH-psbA* có sự khác biệt di truyền nhiều hơn.

#### TÀI LIỆU THAM KHẢO

- Asahina H., Shinozaki J., Masuda K., Morimitsu Y., and Satake M., 2010. Identification of medicinal *Dendrobium* species by phylogenetic analyses using *matK* and *rbcL* sequences, J. Nat. Med, 64(2): 133-138.
- European Bioinformatics Institute, 2017. Programmatic access to bioinformatics tools from EMBL-EBI update. Truy cập từ <https://www.ebi.ac.uk/Tools/msa/clustalo/> ngày 10/4/2019.
- Hoàng Đăng Hiếu, 2012. Sử dụng các kỹ thuật sinh học phân tử trong phân tích đa dạng và định danh loài ở tập đoàn cây dó bầu (*Aquilaria* sp.) tại Hà Tĩnh, Luận văn thạc sĩ Sinh học, Trường Đại học Khoa học Tự nhiên, ĐHQG Hà Nội.
- Hollingsworth P. M., Forrest L. L., Spouge J. L., Hajibabaei M., Rasmasingham S., van der Bank M., Chase M. W., Cowan R. S., Erickson D. L., Fazekas A. J., 2009. A DNA barcode for land plants, Proceedings of the National Academy of Sciences, 106, pp. 2794-2797.
- Jing YU, Jian-Hua XUE, Shi-Liang ZHOU, 2011. New universal *matK* primers for DNA barcoding angiosperms, Journal of Systematics and Evolution 49 (3): 176 - 181.

6. Kress J. W., Wurdack K. J., Zimmer E. A., Weigt L. A. & Janzen D. H., 2005. *Use of DNA barcodes to identify flowering plants*. Proc. Natl. Acad. Sci. USA, 102, pp. 8369-8374.
7. Nguyễn Thị Thanh Nga. 2012. *Đánh giá đa dạng di truyền một số loài cây dược liệu Việt Nam thuộc chi Đương Sâm (Codonopsis sp.) bằng kỹ thuật ADN mã vạch*, Luận văn Thạc sĩ Khoa học, Trường Đại học Khoa học Tự nhiên, ĐHQG Hà Nội.
8. Olmstead R. G., H. J. Michaels, K. M. Scott, and J. D. Palmer. 1992. *Monophyly of the Asteridae and identification of their major lineages inferred from DNA sequences of rbcL*. Ann. Missouri Bot. Gard. 79: 249-265.
9. Phạm Hoàng Hộ, 2002. *Cây cỏ Việt Nam, quyển III*. Nhà xuất bản Trẻ, trang 516 – 518.
10. Sachidanandam R, Weissman D, Schmidt S. C., Kakol J. M., Stein L. D., Marth G., Sherry S., Mullikin J. C., Mortimore B. J., and Willey D. L., 2001. *A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms*. Nature 409: 928-933.
11. Sameera O. Bafeel, Ibrahim A. Arif, Mohammad A. Bakir, Haseeb A. Khan, Ahmad H. Al Farhan, Ali A. Al Homaidan, Anis Ahamed, Jacob Thoma, 2011. *Comparative evaluation of PCR success with universal primers of maturase K (matK) and ribulose-1, 5-bisphosphate carboxylase oxygenase large subunit (rbcL) for barcoding of some arid plants*, Plant Omics Journal. POJ 4(4):195-198.
12. Sharma S. K., Dkhar J., Kumaria S., Tandon P., and Rao S. R., 2012. *Assessment of phylogenetic inter-relationships in the genus Cymbidium (Orchidaceae) based on internal transcribed spacer region of rDNA*. Gene. 495(1): 10-15.
13. Srirama R., Senthilkumar U., Sreejayan N., Ravikanth G., Gurumurthy B. R., Shivanna M. B., Sanjappa M., Ganeshiah K. N., Shaanker R. U., 2010. *Assessing species admixtures in raw drug trade of Phyllanthus, a hepato-protective plant using molecular tools*, J. Ethnopharmacol., Jul 20; 130(2): 208-15. doi: 10.1016/j.jep.2010.04.042. Epub 2010 May 8.
14. Võ Duy Huân, Satoshi Yamamura, Kazuhiko Ohtani, Ryoji Kasai, Kazuo Yamasaki, Nguyen Thai Nham and Hoang Minh Chau. 1998. *Oleanane saponin from Polyscias fruticosa*, Phytochemistry, 47(3): 451-457.
15. Võ Văn Chi và Trần Hop, 1999. *Từ điển cây thuốc Việt Nam*. NXB Y học, trang 425 – 428.

## DNA BARCODE ANALYSIS OF POLYSCIAS SAMPLES COLLECTED IN VIETNAM

Trình Việt Nga<sup>1,2</sup>, Nguyễn Cao Kiệt<sup>1</sup>, Bùi Minh Trí<sup>1</sup>,  
Phạm Thị Minh Tâm<sup>1</sup>, Nguyễn Hữu Ho<sup>2</sup>, Huỳnh Văn Biet<sup>1</sup>

<sup>1</sup> Nông Lâm University

<sup>2</sup> Institute of Tropical Biology

## Summary

This study aimed to analyze the DNA barcode of eighteen *Polyscias* spp accessions collected in Vietnam. *Polyscias* samples were collected from various regions in Viet Nam. The DNA extracts of the samples then amplified with *matK*, *rbcL* and *trnH-psbA* primers. Subsequently, the amplicons were sequenced for nucleotide variation analysis. The results indicated that the all *Polyscias* accessions had high level of similarity with some nucleotide variations. It was also shown that *matK* and *rbcL* regions of *Polyscias* were highly conservative. While, *trnH-psbA* region of *Polyscias* was more variable.

**Keywords:** *Polyscias*, DNA barcode, *matK*, *rbcL*, *trnH-psbA*, genetic variation.

Người phản biện: PGS.TS. Nguyễn Văn Đồng

Ngày nhận bài: 26/4/2019

Ngày thông qua phản biện: 29/5/2019

Ngày duyệt đăng: 5/6/2019