

Giám định phân tử loài Săng đào (*Hopea ferrea* Pierre) dựa trên trình tự nucleotide vùng gen *ITS2* và *matK*

Keovilay Chanthalaphone, Lê Xuân Trường, Bùi Thị Mai Hương,
Nguyễn Thị Thơ, Nguyễn Thị Thu Hằng*
Trường Đại học Lâm nghiệp

Molecular identification of *Hopea ferrea* Pierre species based on nucleotide sequences of *ITS2* and *matK* genes

Keovilay Chanthalaphone, Le Xuan Truong, Bui Thi Mai Huong,
Nguyen Thi Tho, Nguyen Thi Thu Hang*
Vietnam National University of Forestry
*Corresponding author: hangntt1@vnuf.edu.vn

<https://doi.org/10.55250/jo.vnuf.15.3.2026.034-043>

Thông tin chung:

Ngày nhận bài: 11/12/2025

Ngày phản biện: 13/01/2026

Ngày quyết định đăng: 10/02/2026

Từ khóa:

Hopea ferrea, *ITS2*, *matK*,
mã vạch DNA, Ngân hàng gen.

Keywords:

DNA barcode, GenBank,
Hopea ferrea, *ITS2*, *matK*.

TÓM TẮT

Giám định chính xác loài là cơ sở thiết yếu cho nghiên cứu bảo tồn, phục hồi, nhân giống thực vật. Để nhận dạng loài thực vật, mã vạch DNA (các đoạn trình tự DNA ngắn, chuẩn hóa) thường được ứng dụng như một công cụ hữu hiệu. Trong nghiên cứu, nhóm tác giả sử dụng 2 đoạn mã vạch DNA là *ITS2* và *matK* để nhận dạng loài Săng đào (*Hopea ferrea* Pierre) ở Vườn Quốc gia Phou Khao Khouay, nước Cộng hòa Dân chủ nhân dân Lào. Mẫu phân tích có trình tự nucleotide gen *ITS2* (394 bp) tương đồng 99,24% và gen *matK* (821 bp) tương đồng 99,76% so với trình tự gen tương ứng của loài *Hopea ferrea* trên Ngân hàng gen. Cây phát sinh loài xây dựng dựa trên dữ liệu gen *ITS2* và *matK* bằng phương pháp Maximum Likelihood cũng chỉ ra mẫu nghiên cứu thuộc loài *Hopea ferrea*. So sánh hiệu quả nhận dạng loài Săng đào bằng 2 đoạn DNA mã vạch cho thấy dữ liệu di truyền vùng gen *ITS2* có hiệu quả phân biệt loài Săng đào với các loài cùng chi *Hopea* tốt hơn vùng *matK* với khoảng cách di truyền tính toán dựa trên gen *ITS2* từ 0,2% đến 7,3%, trung bình 4,2%. Bên cạnh đó, việc ứng dụng hai chỉ thị *ITS2* và *matK* làm gia tăng độ tin cậy của phương pháp nhận dạng loài Săng đào bằng mã vạch DNA.

ABSTRACT

Accurate species identification is essential for plant conservation, restoration, and propagation research. DNA barcodes (short, standardized DNA sequences) are often used as an effective tool to identify plant species. In this study, we used two DNA barcodes, *ITS2* and *matK*, to identify *Hopea ferrea* Pierre in Phou Khao Khouay National Park, Lao People's Democratic Republic. The analyzed sample had nucleotide sequences of the *ITS2* gene (394 bp) with 99.24% similarity and the *matK* gene (821 bp) with 99.76% similarity to the corresponding gene sequence of *Hopea ferrea* in GenBank. The phylogenetic tree constructed based on *ITS2* and *matK* gene data using the Maximum Likelihood method also showed that the studied sample belonged to the *Hopea ferrea* species. Comparing the effectiveness of *Hopea ferrea* species identification using two DNA barcode segments showed that genetic data from the *ITS2* gene region was more effective in differentiating *Hopea ferrea* from other species in the genus *Hopea* than the *matK* region, with calculated genetic distances based on the *ITS2* gene ranging from 0.2% to 7.3%, averaging 4.2%. In addition, the application of two markers, *ITS2* and *matK*, will increase the reliability of the method for identifying species of the genus *Hopea* using DNA barcodes.

1. ĐẶT VẤN ĐỀ

Săng đào (*Hopea ferrea* Pierre) còn được gọi bằng các tên khác như: Săng đá, sao tía, sao đá, tấu đá, chò kiền kiền. Cây Săng đào thuộc họ Dầu (Dipterocarpaceae), chi Sao (*Hopea*), phân bố tự nhiên ở các nước khu vực Đông Nam Á: Lào, Campuchia, Malaysia, Myanmar, Thái Lan và Việt Nam [1]. Cây cung cấp nguồn gỗ và nhựa quý. Gỗ từ cây Săng đào cứng, độ bền cao, khả năng chịu lực tốt nên được sử dụng nhiều trong ngành xây dựng (làm cột, kèo nhà), đóng tàu thuyền. Gỗ Săng đào không chỉ chắc mà còn có vân đẹp nên cũng thường được sử dụng để sản xuất đồ nội thất, làm đồ thủ công mỹ nghệ. Thân cây cao nên thường chiếm tầng trên của rừng và có tác dụng như cây cảnh, cây bóng mát khi trồng ở đường phố. Do có nhiều công dụng nên cây Săng đào luôn có nguy cơ bị khai thác quá mức và được xếp vào nhóm nguy cấp, mức Bị đe dọa (EN) trong Danh lục Đỏ IUCN.

Quần thể cây Săng đào phân bố ở Vườn Quốc gia (VQG) Phou Khao Khouay, nước Cộng hòa dân chủ nhân dân Lào (CHDCND Lào) cần được bảo tồn, nhân giống. Trước khi tiến hành các nghiên cứu nhân giống Săng đào, việc thực hiện giám định hình thái và phân tử là cần thiết. Phương pháp nhận dạng hình thái thực vật có thể gặp phải những khó khăn bởi những đặc điểm hình thái tương tự của những loài trong cùng chi, họ [2]. Đặc điểm hình thái giống nhau cũng là trở ngại trong phân loại các loài cây trong chi *Hopea*, khi một số loài cùng chi rất giống nhau về hình thái [3]. Loài Săng đào (*H. ferrea*) có hình thái tương đồng (đặc biệt là đặc điểm thân, lá) với loài *H. odorata* (Sao đen) và một số loài khác trong cùng chi, như *H. chinensis* (Sao hòn gai), *H. mollissima* (Sao mặt quỷ), *H. hainanensis* (Sao hải nam). Do đó, để nhận dạng chính xác loài Săng đào, cần sử dụng công cụ mã vạch DNA để phân tích.

Mã vạch DNA là một kỹ thuật sử dụng các đoạn DNA ngắn, được chuẩn hóa trong bộ gen của sinh vật để đạt được mục tiêu nhận dạng loài [4]. Một đoạn DNA được sử dụng làm mã vạch cần có ba đặc điểm: tính linh hoạt cao, trình tự chất lượng tốt và khả năng phân biệt giữa các loài với nhau [4]. Theo Phillips và cộng sự (2022), tiêu chí quan trọng để phân định loài là dựa vào khoảng cách di truyền với khoảng cách nhỏ hơn cho thấy quan hệ gần gũi hơn và ngược lại [5].

Năm 2009, Hiệp hội Mã vạch của Sự sống (CBOL - The Consortium for the Barcode of Life) đã khuyến nghị hai locus lục lạp thích hợp làm mã vạch DNA ở thực vật: vùng gen maturase K (*matK*) và gen ribulose-bisphosphate carboxylase (*rbcl*). Bên cạnh đó, vùng gen nhân *ITS2* thường được chọn làm ứng viên mã vạch vì nhiều ưu điểm: tính phổ biến cao, kích thước nhỏ (200 - 400 bp) và khả năng nhận diện cao trong các nhóm thực vật có quan hệ họ hàng gần [6]. Đặc biệt, việc sử dụng kết hợp các đoạn DNA mã vạch khác nhau sẽ gia tăng độ tin cậy trong phân loại thực vật [7]. Mục tiêu của nghiên cứu này là phân tích trình tự hai vùng DNA mã vạch (*matK* và *ITS2*) để nhận dạng loài Săng đào phân bố ở VQG Phou Khao Khouay, nước Cộng hòa dân chủ nhân dân Lào.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đối tượng, vật liệu, hoá chất

Đối tượng nghiên cứu: Loài Săng Đào (*Hopea Ferrea* Pierre) phân bố ở VQG Phou Khao Khouay, nước Cộng hòa dân chủ nhân dân Lào, ở vị trí rừng kín mưa ẩm nhiệt đới với độ cao dưới 700 m so với mực nước biển cùng với một số nhân tố sinh thái phát sinh quần thể: nhiệt độ trung bình hằng năm 20 - 25°C, lượng mưa trung bình hằng năm 1200 - 2500 mm. Thời điểm thu mẫu là từ tháng 3 đến tháng 5 năm 2024.



Hình 1. Hình thái cây trưởng thành, cành mang hoa và lá cây Săng đào phân bố tại VQG Phou Khao Khouay (CHDCND Lào)

Vật liệu thực vật: Mẫu lá cây Săng đào (3 mẫu lá từ 3 cây phân bố ở những vị trí khác nhau trong VQG Phou Khao Khouay). Sau khi thu thập, mẫu được đựng trong túi nylon kín chứa hạt Silica gel và vận chuyển ngay theo đường hàng không từ Lào về Việt Nam. Ở Việt Nam, các mẫu lá được bảo quản ở -20°C trong

Phòng thí nghiệm của Viện Công nghệ sinh học Lâm nghiệp (Trường Đại học Lâm nghiệp) làm vật liệu tách DNA tổng số phục vụ nghiên cứu.

Cặp mồi nhân gen *ITS2* và *matK*: Trình tự nucleotide và thông tin về cặp mồi nhân gen *ITS2* và *matK* thể hiện ở Bảng 1.

Bảng 1. Thông tin về cặp mồi nhân gen *ITS2* và *matK* sử dụng trong nghiên cứu

Gen	Tên mồi	Trình tự mồi (5' – 3')	Nhiệt độ gắn mồi ($^{\circ}\text{C}$)	Kích thước gen (bp)	Nguồn tham khảo
<i>ITS2</i>	<i>IsP1_F</i>	ATGCGATACTTGGTGTGAAT	58	~ 400	White và cộng sự (1990)
	<i>IsP1_R</i>	TCCTCCGCTTATTGATATGC			
<i>matK</i>	<i>matK_F</i>	ACCCAGTCCATCTGGAAATCTTGGTTC	50	~ 800	Kress và Erickson (2007)
	<i>matK_R</i>	CGTACAGTACTTTTGTGTTTACGAG			

2.2. Phương pháp nghiên cứu

2.2.1. Tách chiết DNA tổng số

Tách chiết DNA tổng số từ lá cây Săng đào theo Kit tách DNA thực vật (innuPREP Plant DNA Kit) của hãng Analytik Jena (Đức) với các bước thực hiện như hướng dẫn của nhà sản xuất. DNA tổng số được kiểm tra nồng độ và độ tinh sạch bằng phương pháp đo mật độ quang trên máy ScanDrop² (Analytik Jena, Đức) và điện di trên gel agarose 1%. DNA tổng số sau đó được pha loãng đến nồng độ 30 ng/ μl để làm vật liệu cho phản ứng PCR khuếch đại gen *ITS2* và *matK*.

2.2.2. Nhân bản gen *ITS2* và *matK* bằng phản ứng PCR

Tổng thể tích phản ứng PCR là 50 μl với các

thành phần gồm: 25 μl Master mix 2X, 1 μl mồi xuôi (10 pmol/ μl), 1 μl mồi ngược (10 pmol/ μl), 2 μl DNA tổng số, 21 μl H₂O deion. Chu trình nhiệt của phản ứng PCR: 95 $^{\circ}\text{C}$ trong 5 phút; lặp lại 35 chu kỳ (95 $^{\circ}\text{C}$: 30 giây, 50 $^{\circ}\text{C}$ / 58 $^{\circ}\text{C}$: 30 giây, 72 $^{\circ}\text{C}$: 1 phút); 72 $^{\circ}\text{C}$ trong 10 phút; bảo quản sản phẩm PCR ở 4 $^{\circ}\text{C}$. Nhiệt độ gắn mồi để khuếch đại gen *matK* là 50 $^{\circ}\text{C}$ và gen *ITS2* là 58 $^{\circ}\text{C}$. Lặp lại 3 lần phản ứng PCR trên mỗi mẫu thí nghiệm.

Kiểm tra sản phẩm PCR bằng điện di trên gel agarose 1,5%. Phản ứng PCR dương tính với 1 băng DNA đặc hiệu và có kích thước tương ứng với kích thước dự kiến của gen *ITS2* và *matK* sẽ được tinh sạch bằng bộ Kit tinh sạch sản phẩm PCR (innuPREP PCRpure Kit) của hãng Analytik Jena (Đức). Sản phẩm PCR đã tinh sạch được

gửi đến Công ty Apical Scientific (<https://apicalscientific.com>, Malaysia) để giải trình tự theo cả hai chiều.

2.2.3. Phân tích dữ liệu mã vạch DNA để nhận dạng loài thực vật

Các trình tự nucleotide được phân tích, xử lý bằng phần mềm BioEdit v.7.2.5 [8].

So sánh với các trình tự đã có trên GenBank bằng công cụ BLAST trong NCBI (<http://www.ncbi.nlm.nih.gov>). Kết quả tham chiếu trên NCBI cho phép xác định loài dựa trên giá trị tương đồng cao nhất với loài có sẵn trên NCBI. Cây phát sinh chủng loại được xây dựng dựa trên phương pháp xác suất tối đa ML (Maximum Likelihood) với mô hình thay thế nucleotide T92, sử dụng phần mềm MEGA7.0 với giá trị ủng hộ (Bootstrap) 1.000 lần lặp lại [9].

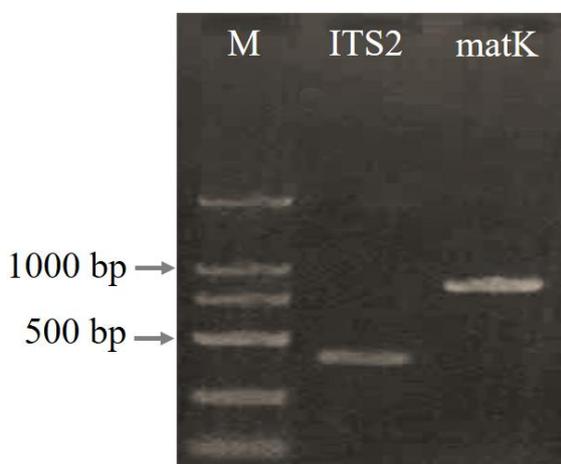
3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Tách chiết DNA tổng số và khuếch đại gen chỉ thị bằng kỹ thuật PCR

Quá trình tách chiết DNA tổng số các mẫu lá Săng đào được thực hiện theo hướng dẫn của nhà sản xuất bộ hoá chất innuPREP Plant DNA Kit. Kết quả kiểm tra chất lượng DNA tổng số

bằng phương pháp so màu quang phổ cho thấy các mẫu DNA tổng số đều đảm bảo độ tinh sạch với trị số OD_{260nm} / OD_{280nm} trong khoảng 1,8 - 2,0 và hàm lượng DNA đạt 600 - 800 ng/ μ L. Kiểm tra chất lượng DNA bằng kỹ thuật điện di trên gel agarose 1% cho thấy băng DNA tổng số của các mẫu đều rõ nét và ít bị đứt gãy, đảm bảo yêu cầu kỹ thuật làm khuôn cho thực hiện phản ứng PCR nhân bản đoạn gen nhân *ITS2* và gen lục lạp *matK*.

Kết quả phản ứng PCR nhân bản gen *ITS2* và *matK* với cặp mồi đặc hiệu tương ứng với từng gen được kiểm tra bằng điện di trên gel agarose 1,5% thể hiện ở Hình 2. Sản phẩm phản ứng PCR khuếch đại gen *ITS2* xuất hiện duy nhất một băng DNA sáng đậm với kích thước khoảng 400 bp - đúng kích thước dự kiến của gen *ITS2*. Bên cạnh đó, sản phẩm phản ứng PCR nhân gen *matK* cũng xuất hiện một băng DNA đặc hiệu sáng, rõ nét, kích thước đúng như dự tính là khoảng 800 bp. Chất lượng sản phẩm PCR là đảm bảo (đặc hiệu, không có băng DNA phụ) nên đã được tinh sạch bằng bộ hoá chất innuPREP PCRpure Kit và gửi đi giải trình tự nucleotide.



Hình 2. Điện di kiểm tra sản phẩm PCR nhân bản gen *ITS2* và *matK*
(M: Thang marker 100 bp – Phusa Genomics, Việt Nam)

3.2. Phân tích trình tự nucleotide vùng gen *ITS2* của cây Săng đào phân bố ở VQG Phou Khao Khouay, Lào

Sản phẩm PCR nhân gen *ITS2* từ khuôn là DNA tổng số của các mẫu lá Săng đào sau khi được giải trình tự hai chiều đã được kiểm tra, hiệu chỉnh và loại bỏ các vùng tín hiệu nhiễu

bằng phần mềm BioEdit v.7.2.5. Sau khi xử lý, cả 3 đoạn trình tự gen *ITS2* (tương ứng với 3 mẫu lá Săng đào thu thập từ 3 cây phân bố ở vị trí khác nhau trong VQG Phou Khao Khouay) đã được so sánh mức độ tương đồng di truyền. Kết quả so sánh chỉ ra cả 3 đoạn trình tự gen *ITS2* tương ứng với 3 mẫu Săng đào đều

có sự tương đồng về trình tự nucleotide là 100% nên quá trình phân tích tiếp theo đã sử dụng 1 trình tự nucleotide đại diện để thực hiện.

Kết quả kiểm tra tính tương đồng về trình tự nucleotide của đoạn gen *ITS2* nghiên cứu (394 bp) với các trình tự sẵn có trên GenBank (Bảng 2) bằng công cụ BLAST cho thấy trình tự gen *ITS2* của loài Săng đào so với các loài trong chi *Hopea* có tỷ lệ tương đồng di truyền đạt

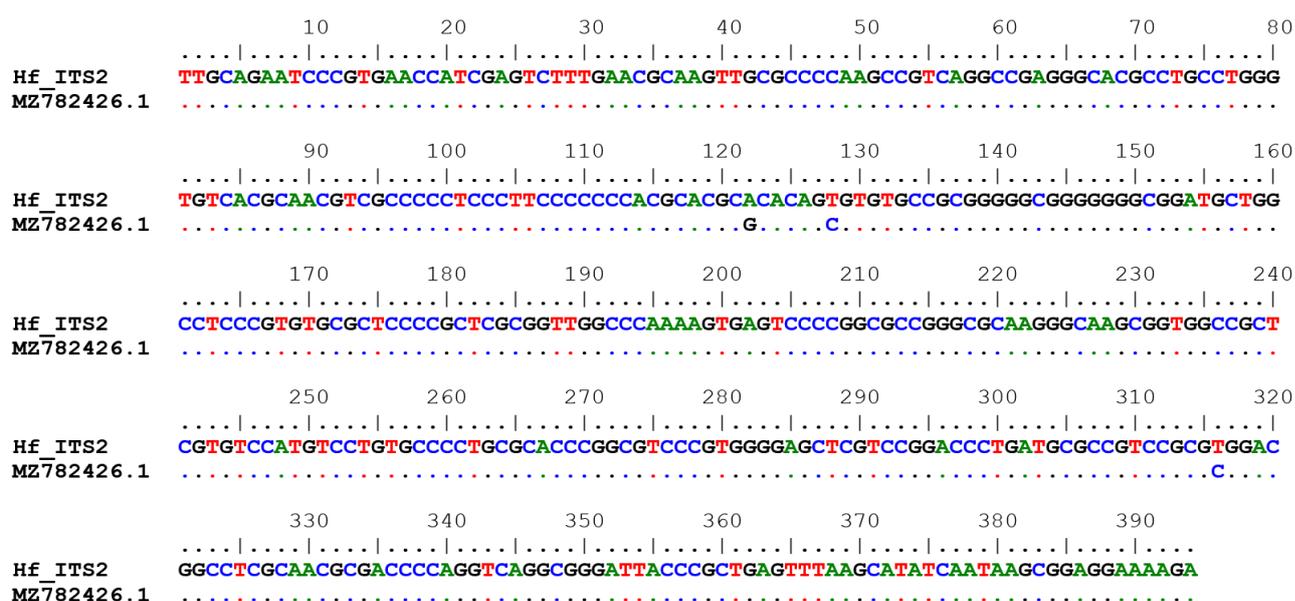
90,15% – 99,24%. Trong đó, mẫu nghiên cứu có mức độ tương đồng di truyền cao (99,24%) với loài *H. ferrea* (Săng đào) mã số MZ782426.1; mức độ tương đồng di truyền khá thấp (96,45% - 96,95%) với các loài *H. polyalthioides* MZ782441.1, *H. wightiana* MZ782446.1; mức độ tương đồng thấp (90,15% - 91,16%) với *H. kerangasensis* MZ782434.1, *H. griffithii* MZ782431.1 và *H. semicuneata* MZ782444.1.

Bảng 2. Một số loài thuộc chi sao (*Hopea*) có trình tự vùng gen *ITS2* tương đồng với mẫu nghiên cứu trên GenBank

TT	Tên loài	Mã số	Tỷ lệ tương đồng (%)
1	<i>Hopea ferrea</i>	MZ782426.1	99,24
2	<i>Hopea polyalthioides</i>	MZ782441.1	96,95
3	<i>Hopea wightiana</i>	MZ782446.1	96,71
4	<i>Hopea chinensis</i>	MK193025.1	96,45
5	<i>Hopea hainanensis</i>	MK193029.1	93,99
6	<i>Hopea kerangasensis</i>	MZ782434.1	91,16
7	<i>Hopea griffithii</i>	MZ782431.1	90,15
8	<i>Hopea semicuneata</i>	MZ782444.1	90,15

Gen *ITS2* của mẫu nghiên cứu có tỷ lệ tương đồng 99,24% với gen *ITS2* của *H. ferrea* (MZ782426.1) trên Ngân hàng gen với 3 vị trí nucleotide sai khác được thể hiện ở Hình 3: vị trí nucleotide 122 (A → G), vị trí 128 (T → C) và

vị trí 316 (T → C). Bên cạnh đó, tỷ lệ A, C, G và T trong đoạn gen *ITS2* của mẫu nghiên cứu (394 bp) lần lượt là 15,74%; 35,53%; 32,99% và 15,74%, cho thấy gen *ITS2* có hàm lượng G + C cao (68,55%).



Hình 3. So sánh trình tự đoạn gen *ITS2* của mẫu nghiên cứu với trình tự gen *ITS2* mã số MZ782426.1 (loài *Hopea ferrea*) trên NCBI

Trình tự nucleotide gen *ITS2* của mẫu nghiên cứu (phân tích bằng phương pháp BLAST) có mức độ tương đồng cao nhất với trình tự gen *ITS2* của loài *H. ferrea* mã số MZ782426.1 trên Ngân hàng gen. Tuy nhiên, để kết quả được khẳng định về mặt phân loại, nhóm tác giả tiến hành tính toán khoảng cách di truyền và xây dựng cây phát sinh loài giữa mẫu nghiên cứu với các loài thuộc chi *Hopea* trên Ngân hàng gen bằng phần mềm MEGA7.0. Kết quả trình

bày ở Bảng 3 và Hình 4.

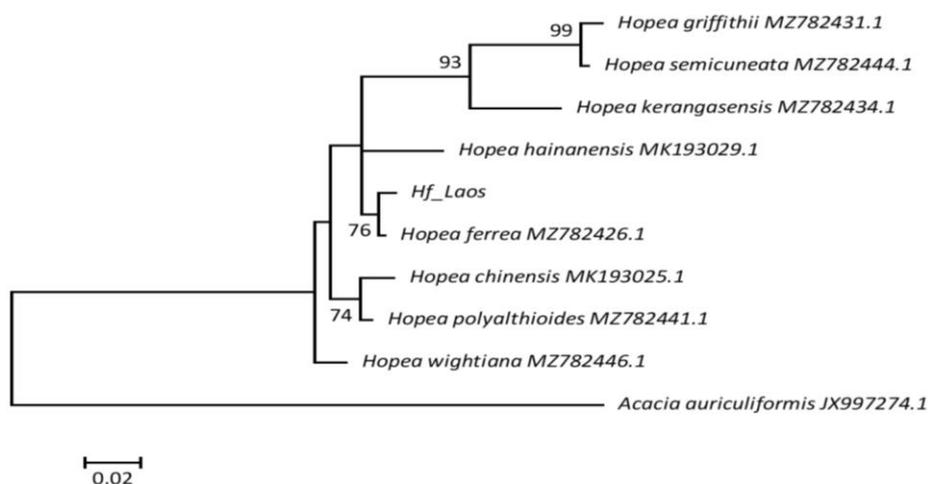
Khoảng cách di truyền giữa mẫu nghiên cứu (ký hiệu *Hf_Laos*) với các loài trong chi *Hopea* (Bảng 3) có sự dao động khá lớn (0,2% - 7,3%), trung bình 4,2%. Giá trị về khoảng cách di truyền là thấp nhất (0,2%) được ghi nhận giữa mẫu nghiên cứu với loài *H. ferrea* (MZ782426.1). Kết quả nhận được chứng minh mẫu nghiên cứu có mối quan hệ di truyền mật thiết với loài *H. ferrea*.

Bảng 3. Khoảng cách di truyền (%) của mẫu nghiên cứu với các loài trong chi *Hopea* Trên GenBank được tính toán dựa trên trình tự nucleotide vùng gen *ITS2*

Mẫu/Loài	1	2	3	4	5	6	7	8	9
<i>Hf_Laos</i>	-								
<i>H. ferrea</i> MZ782426.1	0,2	-							
<i>H. griffithii</i> MZ782431.1	6,0	5,8	-						
<i>H. hainanensis</i> MK193029.1	1,8	2,0	7,3	-					
<i>H. kerangasensis</i> MZ782434.1	5,7	5,8	5,4	6,9	-				
<i>H. polyalthioides</i> MZ782441.1	1,1	1,3	7,0	2,2	6,3	-			
<i>H. semicuneata</i> MZ782444.1	6,0	5,8	0,7	7,3	5,0	7,0	-		
<i>H. wightiana</i> MZ782446.1	1,1	1,0	6,3	2,7	6,3	1,3	6,3	-	
<i>H. chinensis</i> MK193025.1	1,8	1,7	7,2	2,4	6,9	0,7	7,2	1,7	-

Sơ đồ mối quan hệ di truyền dựa trên vùng *ITS2* giữa mẫu nghiên cứu với các loài thuộc chi *Hopea* (Hình 4) được phân thành 3 nhánh: nhánh 1 chia thành 3 nhánh phụ, trong đó nhánh phụ thứ nhất gồm *H. griffithii*, *H. semicuneata* và *H. kerangasensis*, nhánh phụ thứ hai chỉ có loài *H. hainanensis*, nhánh phụ thứ ba gồm mẫu nghiên cứu (ký hiệu *Hf_Laos*)

và loài *H. ferrea*; nhánh 2 có *H. chinensis* và *H. polyalthioides*; nhánh 3 là loài *H. wightiana*. Như vậy, kết quả xây dựng cây phát sinh chủng loại cũng chỉ ra mẫu nghiên cứu và loài *H. ferrea* (MZ782426.1) thuộc một nhóm riêng, có mức độ tương đồng di truyền cao, quan hệ mật thiết với nhau với giá trị Bootstrap 76%.



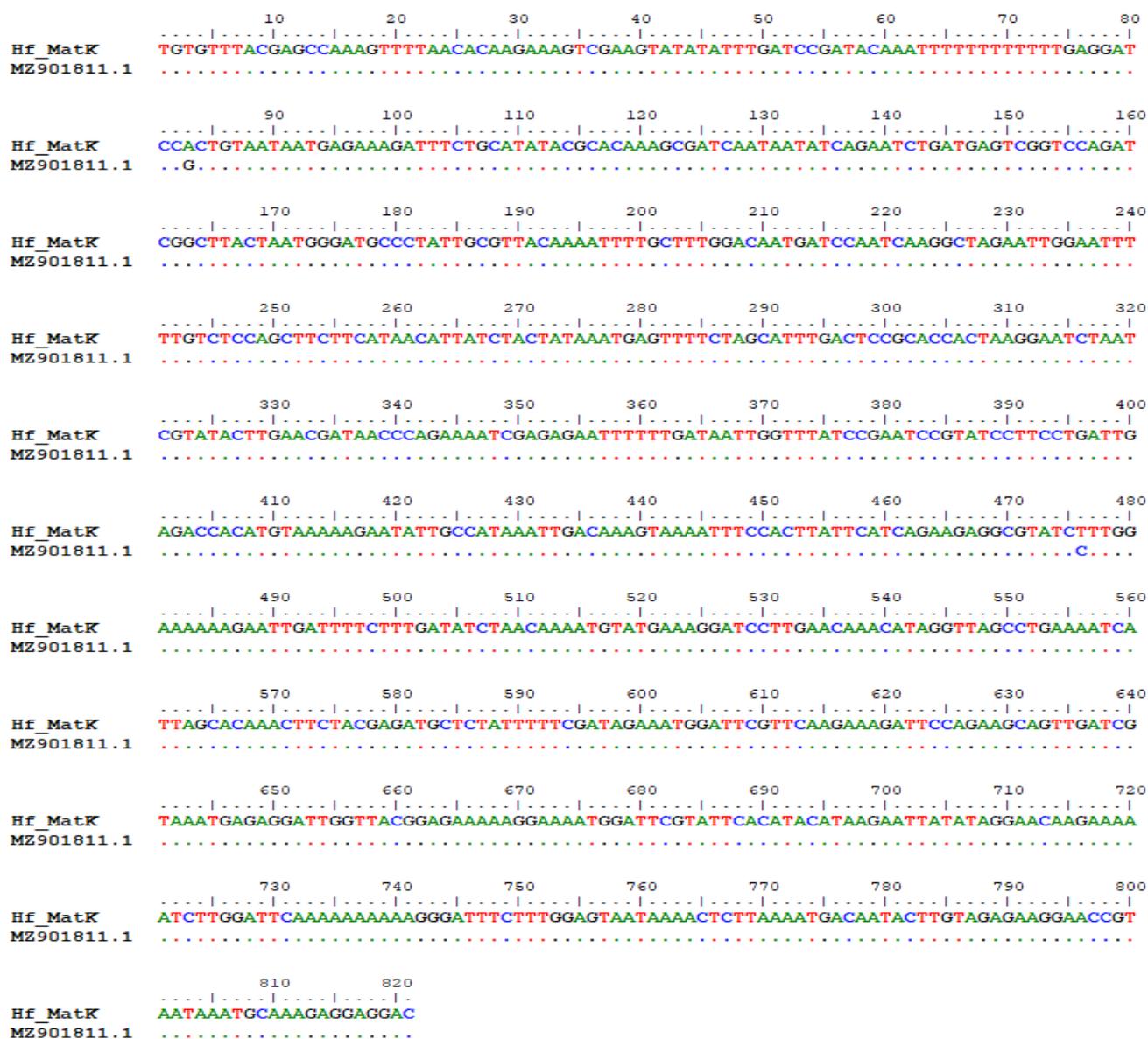
Hình 4. Cây phát sinh chủng loại xây dựng dựa trên trình tự nucleotide đoạn gen *ITS2* của mẫu nghiên cứu với các loài cùng chi *Hopea* trên GenBank (Các số trên các nhánh tượng trưng cho sự hỗ trợ Bootstrap; loài *Acacia auriculiformis* JX997274.1 là loài ngoài nhóm)

Kết quả phân tích trình tự vùng *ITS2* của mẫu nghiên cứu bằng 3 phương pháp (ứng dụng công cụ BLAST, tính toán khoảng cách di truyền theo p-Distance và xây dựng cây phát sinh loài theo Maximum Likelihood) đều cho kết quả thống nhất và cho phép khẳng định mẫu nghiên cứu thuộc loài *Hopea ferrea* (Săng đào) và chỉ thị DNA mã vạch vùng *ITS2* thích hợp cho mục tiêu nhận dạng loài Săng đào và các loài khác thuộc chi *Hopea*.

3.3. Phân tích trình tự nucleotide vùng gen *matK* của cây Săng đào phân bố ở VQG Phou Khao Khouay, Lào

Kết quả kiểm tra trình tự đoạn gen đã được nhân bản bằng cặp mồi đặc hiệu gen *matK* bằng

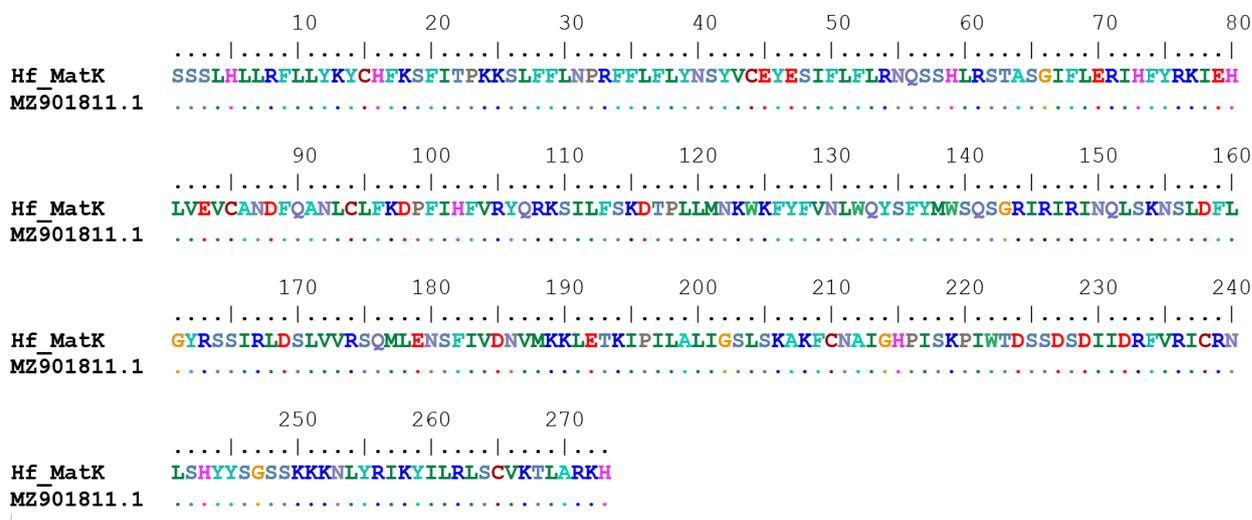
chương trình Blast/NCBI cho thấy sản phẩm PCR tương ứng với vùng gen *matK* và có độ tương đồng cao với loài *H. ferrea* và các loài khác trong chi *Hopea* trên GenBank với mức độ tương đồng đều trên 99%. Mức độ tương đồng cao giữa các loài trong chi chứng tỏ *matK* là vùng gen ít biến đổi và có sự khác biệt di truyền thấp ở các loài thuộc chi *Hopea*. Đoạn gen *matK* trong nghiên cứu có độ dài 821 bp, tương đồng 99,76% với gen *matK* của loài *Hopea ferrea* mã số MZ901811.1 trên GenBank và chỉ có 2 vị trí sai khác thể hiện rõ khi phân tích bằng phần mềm BioEdit (Hình 5): vị trí nucleotide 83 (A → G); vị trí 476 (A → T).



Hình 5. So sánh trình tự gen vùng *matK* của mẫu nghiên cứu và trình tự có mã số MZ901811.1 (*Hopea ferrea*) trên GenBank bằng phần mềm BioEdit

Trình tự gen *matK* của mẫu nghiên cứu so với loài *H. ferrea* MZ901811.1 có 2 vị trí sai khác. Tuy nhiên, trình tự amino acid suy diễn từ trình tự gen vùng *matK* của mẫu nghiên cứu so

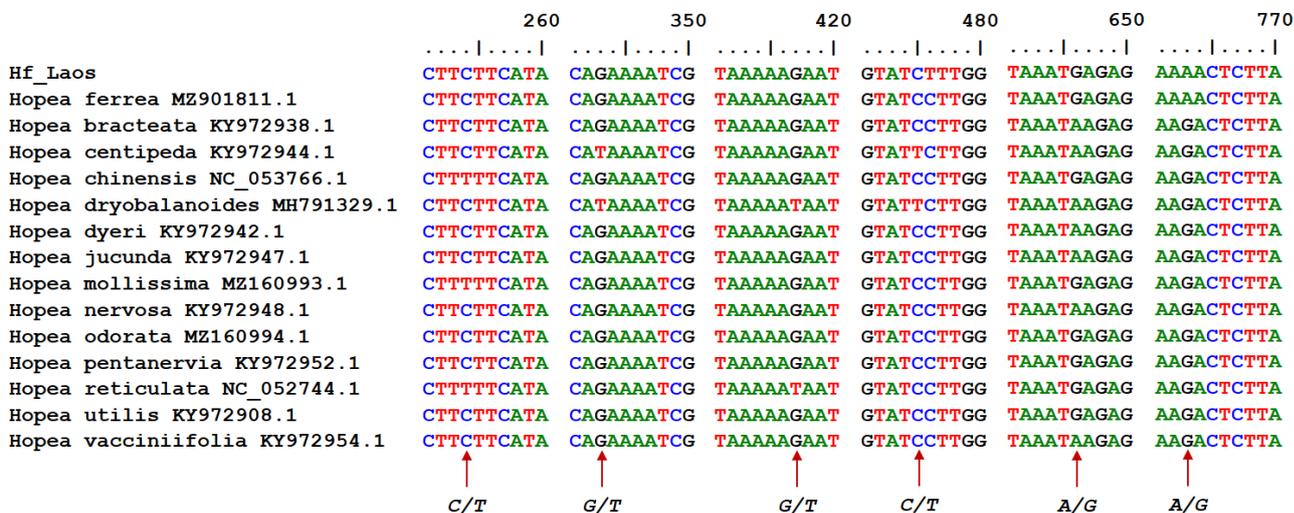
với trình tự amino acid gen *matK* của *H. ferrea* (MZ901811.1) lại có sự tương đồng 100% và không có sự sai khác về trình tự sắp xếp của các amino acid (Hình 6).



Hình 6. So sánh trình tự amino acid suy diễn của mẫu nghiên cứu với trình tự có mã số MZ901811.1 (*Hopea ferrea*) trên GenBank

Trình tự gen *matK* của mẫu nghiên cứu có sự tương đồng cao (>99%) so với loài *H. ferrea* mã số MZ901811.1 và 13 loài khác trong chi *Hopea*. Tuy nhiên, khi so sánh vùng gen *matK* của mẫu nghiên cứu với trình tự gen tương ứng của các loài có sự tương đồng cao trên GenBank, nhóm tác giả nhận thấy có 6 vị trí sai khác thường xuất hiện ở vùng gen *matK* (Hình 7): vị trí nucleotide 254 và 475 là C hoặc T; vị trí 343 hoặc 417 là nucleotide G hoặc T; vị trí nucleotide 646 và 763 là A hoặc G. Đặc biệt, có 1 vị trí nucleotide của mẫu nghiên cứu giống với loài *H. ferrea*

nhưng khác với tất cả các loài còn lại trong chi *Hopea* trên GenBank là vị trí nucleotide 763: thay thế T (chỉ có ở gen *matK* của mẫu nghiên cứu và loài *H. ferrea* MZ901811.1) bằng G ở tất cả các loài còn lại trong chi *Hopea*. Bên cạnh đó, gen *matK* của mẫu nghiên cứu cũng có 1 vị trí nucleotide khác biệt so với trình tự *H. ferrea* MZ901811.1 và tất cả các loài khác trong chi *Hopea*, bước đầu nhận định có thể là 1 vị trí sai khác nucleotide: vị trí 476 (nucleotide T) chỉ xuất hiện ở mẫu nghiên cứu và ở các loài khác vị trí này là C.



Hình 7. Một số vị trí nucleotide biến đổi trong gen *matK* của các loài trong chi *Hopea*

mẫu Săng đào (*Hopea ferrea*) phân bố ở VQG Phou Khao Khouay, nước Cộng hòa dân chủ nhân dân Lào. Kết quả xác định trình tự nucleotide gen *ITS2* và *matK* ở cây Săng đào (Lào) cho thấy gen *ITS2* có kích thước 394 bp và gen *matK* kích thước 821 bp. Phân tích gen bằng công cụ BLAST/NCBI cho kết quả gen *ITS2* và *matK* của mẫu nghiên cứu có tỷ lệ tương đồng cao với trình tự gen tương ứng của loài Săng đào trên GenBank: gen *ITS2* tương đồng 99,24% và gen *matK* tương đồng 99,76%. Cây phân loại di truyền được xây dựng dựa trên trình tự gen *ITS2* và *matK* đều thể hiện mẫu nghiên cứu cùng nhánh và có khoảng cách di truyền thấp nhất với loài Săng đào. Trong 2 đoạn trình tự DNA mã vạch, vùng gen *ITS2* cho hiệu quả phân biệt các loài thuộc chi *Hopea* tốt hơn vùng gen *matK*, với khoảng cách di truyền được tính toán dựa trên gen *ITS2* của các loài cùng chi dao động từ 0,2 - 7,3%, trung bình 4,2%. Để gia tăng độ chính xác của phương pháp nhận dạng loài Săng đào bằng DNA mã vạch, nên sử dụng kết hợp cả 2 chỉ thị *ITS2* và *matK*.

TÀI LIỆU THAM KHẢO

- [1]. N. V. Quy, Y. X. Kang, I. Ashrafu, M. Li, N. T. Tuan, V. Q. Nguyen & N. V. Hop (2022). Spatial distribution and association patterns of *Hopea pierrei* hance and other tree species in the Phu Quoc island evergreen broadleaved forest of Vietnam. *Applied Ecology and Environmental Research*. 20(2): 1911–1933. DOI: 10.15666/aeer/2002_19111933.
- [2]. P. B. Breslin, M. F. Wojciechowski & L. C. Majure (2021). Molecular phylogeny of the Mammilloid clade (Cactaceae) resolves the monophyly of Mammillaria. *Taxon*. 70(2): 308-323. DOI: 10.1002/tax.12451.
- [3]. N. T. P. Trang, N. Duc, N. Sinh & L. Triest (2015). Application of DNA barcoding markers to the identification of *Hopea* species. *Genetics and Molecular Research*. 14(3): 9181-9190. DOI: 10.4238/2015.August.7.28.
- [4]. P. D. Hebert, A. Cywinska, S. L. Ball & J. R. deWaard (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*. 270(1512): 313-321. DOI: 10.1098/rspb.2002.2218.
- [5]. J. D. Phillips, D. J. Gillis & R. H. Hanner (2022). Lack of statistical rigor in DNA barcoding likely invalidates the presence of a true Species' Barcode Gap. *Frontiers in Ecology and Evolution*. 10: 859099. DOI: 10.3389/fevo.2022.859099.
- [6]. S. Chen, H. Yao, J. Han, C. Liu, J. Song, L. Shi, Y. Zhu, X. Ma, T. Gao & X. J. P. o. Pang (2010). Validation of the *ITS2* region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE*. 5(1): e8613. DOI: 10.1371/journal.pone.0008613.
- [7]. R. A. Mir, K. A. Bhat, G. Rashid, L. B. Ebinezer, A. Masi, R. Rakwal, A. A. Shah & S. M. Zargar (2021). DNA barcoding: a way forward to obtain deep insights about the realistic diversity of living organisms. *The Nucleus*. 64(2): 157-165. DOI: 10.1007/s13237-020-00330-3.
- [8]. T. A. Hall (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series*. 41: 95-98.
- [9]. S. Kumar, G. Stecher & K. Tamura (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*. 33(7): 1870-1874. DOI: 10.1093/molbev/msw054.
- [10]. L. Wei, F. C. Pacheco-Reyes, J. A. Villarreal-Quintanilla, V. Robledo-Torres, J. A. Encina-Dominguez, E. E. Lara-Ramirez & M. A. Perez-Rodriguez (2024). Effectiveness of DNA barcodes (rbcL, matK, ITS2) in identifying genera and species in Cactaceae. *Pakistan Journal of Botany*. 56(5): 1911-1928. DOI: 10.30848/PJB2024-5(11).
- [11]. E. Prasetya, L. Lazuardi, F. Harahap, Y. Rachmawati, Y. Yusuf, S. I. Al Idrus & P. Prastowo (2023). Region of Nuclear Ribosomal DNA (*ITS2*) and Chloroplast DNA (rbc L and trn LF) as A Suitable DNA Barcode for Identification of *Zingiber loerzingii* Valetton From North Sumatera, Indonesia. *Journal of Tropical Biodiversity and Biotechnology*. 8(3): 76956. DOI: 10.22146/jtbb.76956.