CHARACTERISTICS OF *MATK* SEQUENCE ISOLATED FROM *Nervilia fordii* (Hance) Schlechter COLLECTED IN CAO BANG PROVINCE

Nguyen Thi Tam^{1*}, Bui Thi Ha²

¹TNU - University of Education,

²TNU - University of Medicine and Pharmacy

ARTICLE INFO	ABSTRACT
Received: 08/6/2021	<i>Nervilia fordii</i> (Hance) Schlechter, which belongs to the genus <i>Nervilia</i> in the family <i>Orchidaceae</i> , contains terpenoids, flavonoids, amino acids and some volatile oils. These substances have anti-inflammatory and anti-viral activities as well as reduce pain, cough, asthma and chronic bronchitis. Using <i>matK</i> sequence analyzing method, <i>matK</i> sequence dimension of LML-02-CB sample was found at 778bp, in which A type nucleotide is 236 (30.33%), C type is 120 (15.42%), G type is 115 (14.78%), T type is 307 (39.46%). As such, <i>matK</i> sequence of LML-02-CB sample is 99.08% similar to that of <i>Nervilia aragona</i> coded as JN004498 on GenBank. <i>matK</i> zone similarity index on nucleotide sequence of LML-02-CB sample compared to 5 <i>matK</i> sequences on GenBank varies between 97.4% to 99.5%. It is possible to use <i>matK</i> indicator as one of molecular indicators in assessing single leaf orchid species.
Revised: 28/7/2021	
Published: 31/7/2021	
KEYWORDS	
<i>Nervilia fordii</i> (Hance)	
Schlechter	
<i>Orchidaceae</i>	
<i>MatK</i>	
Cao Bang province	
Nucleotide	

ĐẶC ĐIỂM TRÌNH TỰ *matK* PHÂN LẬP TỪ CÂY LAN MỘT LÁ (*Nervilia fordii* (Hance) Schlechter) THU THẬP TẠI CAO BANG

Nguyễn Thị Tâm^{1*}, Bùi Thị Hà²

¹Trường Đại học Sư phạm - ĐH Thái Nguyên

²Trường Đại học Y Dược - ĐH Thái Nguyên

THÔNG TIN BÀI BÁO	TÓM TẮT
Ngày nhận bài: 08/6/2021	Lan một lá (<i>Nervilia fordii</i> (Hance) Schlechter) thuộc chi Lan một lá (<i>Nervilia</i>), họ Lan (<i>Orchidaceae</i>) chứa terpenoids, flavonoid, amino acids và một số loại dầu dễ bay hơi. Các hợp chất này có hoạt tính dược lý chống viêm, chống virus và giảm đau, giảm ho, hen suyễn và viêm phế quản mãn tính. Sử dụng phương pháp phân tích trình tự <i>matK</i> cho thấy, kích thước trình tự <i>matK</i> của mẫu LML-02-CB là 778 bp. Trong đó, nucleotide loại A là 236 (30,33%), C là 120 (15,42%), G là 115 (14,78%), T là 307 (39,46%). Trình tự <i>matK</i> của mẫu LML-02-CB có hệ số tương đồng 99,08% so với trình tự <i>matK</i> của loài <i>Nervilia fordii</i> mang mã số JX865503 và 99,19% so với trình tự <i>matK</i> của loài <i>Nervilia aragona</i> mang mã số JN004498 trên GenBank. Hệ số tương đồng về trình tự nucleotide vùng <i>matK</i> của mẫu LML-02-CB so với 5 trình tự <i>matK</i> trên GenBank dao động từ 97,4% đến 99,5%. Có thể sử dụng chỉ thị <i>matK</i> là một trong các chỉ thị phân tử trong việc giám định các mẫu Lan một lá.
Ngày hoàn thiện: 28/7/2021	
Ngày đăng: 31/7/2021	
TỪ KHÓA	
Lan một lá <i>Nervilia fordii</i> (Hance)	
Schlechter	
Họ Lan	
<i>MatK</i>	
Cao Bằng	
Nucleotide	

DOI: <https://doi.org/10.34238/tnu-jst.4606>

* Corresponding author. Email: tamnt@tnue.edu.vn

1. Mở đầu

Lan một lá (*Nervilia fordii* (Hance) Schlechter) thuộc chi *Nervilia*, họ Lan (*Orchidaceae*) được phân bố rải rác ở các tỉnh miền núi phía Bắc. Lan một lá chứa terpenoids, flavonoid, amino acide và một số loại dầu dễ bay hơi. Các hợp chất này có hoạt tính dược lý chống viêm, chống virus và giảm đau, giảm ho, hen suyễn và viêm phế quản mãn tính. Trong dân gian, Lan một lá được sử dụng làm thuốc giải độc, nhất là ngộ độc nấm; làm mát phổi; chữa ho lao, ho lâu năm, viêm phế quản. Nghiên cứu về thành phần hóa học và hoạt tính ức chế khối u của các chất chiết xuất từ cây Lan một lá (*Nervilia fordii*) cho thấy, dịch chiết từ cây Lan một lá có tác động ức chế đáng kể đối với sự phát triển của tế bào ung thư [1]. Các cấu trúc hóa học đã được làm sáng tỏ trên cơ sở các tính chất hóa lý và dữ liệu quang phổ và được xác định là cycloeucalenol; stigmaterol; sitosterol; axit ursolic; auranitamide; (20S, 22E, 24R)-ergosta-7,22-dien-3 β , 5 α , 6 β -triol; 6 methoxy-cerevisterol; và-daucosterol [1]. Do nạn chặt phá rừng và bị khai thác để bán làm thuốc khiến cho loài cây này trở nên quý hiếm và có nguy cơ bị đe dọa tuyệt chủng. Lan một lá được ghi vào danh sách các loài cây cần được bảo vệ trong “Sách đỏ Việt Nam” của Bộ Khoa học và Công nghệ (2007) [2].

Gần đây, việc sử dụng các DNA mã vạch (DNA barcode) để định danh loài đang được các nhà khoa học trên thế giới tập trung nghiên cứu và có những đóng góp đáng kể trong việc phân loại loài. Xác định loài bằng mã vạch DNA sẽ cho mức độ chính xác cao và đặc biệt hữu dụng với các loài gần gũi mà những quan sát hình thái, sinh trưởng và phát triển chưa đủ cơ sở để nhận dạng hoặc phân biệt loài [3]. Gen *matK* là một trong những gen tiến hoá nhanh nhất, có kích thước khoảng 1500 bp, nằm trong hệ gen lục lạp. *MaturaseK* liên quan đến quá trình loại bỏ các intron loại 2 trong quá trình hoàn thiện ARN sau phiên mã. Do *matK* tiến hoá nhanh và có mặt hầu hết trong thực vật nên đã được sử dụng như một chỉ thị trong nghiên cứu mối quan hệ giữa các loài và phát sinh loài ở thực vật. Hiệp hội mã vạch CBOL (Consortium for the Barcode of Life) đã thử nghiệm *matK* trên gần 550 loài thực vật đã cho thấy, 90% mẫu thực vật hạt kín dễ dàng khuếch đại trình tự bằng cách sử dụng một cặp mồi đơn và đề nghị sử dụng *matK* là một trong những locus barcode chuẩn cho thực vật [4], [5]. Huang và cộng sự (2013) đã nghiên cứu 4 mã vạch DNA phổ biến là *ITS2*, *rbcL*, *matK* và *LSU D1-D3* để xác định loài *Nervilia fordii* và sáu loài trong chi *Nervilia*. Kết quả nghiên cứu cho thấy, cả 4 mã vạch đều có khả năng phân biệt được ở cấp chi. Trong đó, *matK* có khả năng phân biệt hoàn toàn cho tất cả các loài được sử dụng trong nghiên cứu này. *ITS2* có khả năng phân biệt loài 66,7% đứng sau *matK* còn vùng *rbcL* và *LSU D1-D3* cho thấy khả năng phân biệt loài thấp nhất phân biệt 28,6% và 50%. Từ nghiên cứu cho thấy, gen *matK* là một mã vạch DNA tiềm năng để phân biệt các loài trong chi *Nervilia* [6]. Trong công bố trước, chúng tôi đã nghiên cứu đặc điểm hình thái, kết hợp với phân tích trình tự *ITS* của 2 mẫu Lan một lá thu thập tại Cao Bằng và Thái Nguyên đã cho thấy, các mẫu đều thuộc loài *N. fordii* (Hance) Schlechter. Hình thái của hai mẫu Lan một lá thu tại Thái Nguyên và Cao Bằng là giống nhau. Trình tự vùng *ITS* của hai mẫu Lan một lá thu thập tại Cao Bằng và Thái Nguyên có hệ số tương đồng 100% so với trình tự vùng *ITS* của *N. fordii* (Hance) Schlechter mang mã số JX011630 trên GenBank [7]. Nghiên cứu này đề cập đến đặc điểm trình tự *matK* phân lập từ cây Lan một lá thu thập tại Cao Bằng với mục tiêu góp phần tư liệu hóa nguồn gen của loài cây này trong xây dựng mã vạch DNA.

2. Vật liệu và phương pháp nghiên cứu

2.1. Vật liệu

Mẫu cây Lan một lá thu thập ở xã Bình Dương, huyện Hòa An, tỉnh Cao Bằng (LML-02-CB) được dùng làm vật liệu nghiên cứu.

2.2. Phương pháp xác định trình tự *matK*

DNA tổng số được tách chiết bằng Kit GeneJET Plant Genomic DNA Purification (K0791). DNA tổng số sau đó được kiểm tra bằng điện di và bảo quản ở -20°C cho các thí nghiệm tiếp theo. Gen *matK* được khuếch đại bằng kỹ thuật PCR với cặp mồi đặc hiệu (*matK-F*, *matK-R*) có trình tự:

matK-F 5' CGATCTATTTCATTCAATATTTTC 3'
matK-R 5' TCTAGCACACGAAAGTCGAAGT 3'

Kích thước đoạn DNA dự kiến khoảng 800bp.

Tổng thể tích mỗi phản ứng PCR là 15 μl , bao gồm: 7,5 μl PCR Master Mix 2X, 0,5 μl mồi xuôi 10 pmol/ μl , 0,5 μl mồi ngược 10 pmol/ μl , 0,5 μl DNA khuôn 50-100 ng/ μl , 6 μl nước cất khử ion vô trùng. Phản ứng được thực hiện trên máy PCR Eppendorf (Đức). Chu trình nhiệt: Biến tính 95°C 5 phút, 35 chu kỳ [95°C 30 giây, 56°C 30 giây, 72°C 30 giây], 72°C 10 phút và lưu giữ ở 4°C . Sản phẩm PCR được kiểm tra bằng điện di trên gel agarose 0,8% trong đệm TAE 1X. Bản gel được nhuộm trong ethidium bromide 3-5 phút, soi dưới ánh sáng tử ngoại 254 nm. Sản phẩm PCR được tinh sạch bằng Kit GenJET PCR Purification theo hướng dẫn của hãng Thermo Scientific. Trình tự nucleotit vùng *ITS* được xác định bằng máy giải trình tự ABI PRISM® 3100 Avant Genetic Analyzer, sử dụng bộ Kit BigDye® Terminator v3.1 Cycle Sequencing với cặp mồi đặc hiệu (*matK-F*, *matK-R*). Sau đó được phân tích, so sánh bằng các chương trình Bioedit, BLAST, DNASTar.

3. Kết quả nghiên cứu

3.1. Đặc điểm trình tự gen *matK* phân lập từ mẫu Lan một lá thu thập tại Cao Bằng

Để có nguyên liệu nhân gen *matK*, chúng tôi đã tiến hành tách chiết DNA từ lá non mẫu Lan một lá, sau đó kiểm tra bằng điện di trên gel agarose 0,8%. Kết quả cho thấy, DNA tổng số thu được đủ điều kiện để tiến hành phản ứng PCR nhân gen *matK* (Hình 1).

Hình 1. Hình ảnh điện di kiểm tra sản phẩm DNA tổng số từ mẫu Lan một lá Cao Bằng

Hình 2. Kết quả điện di kiểm tra sản phẩm PCR nhân gen *matK* từ mẫu Lan một lá Cao Bằng (M: DNA marker; 1, 2: gen *matK*)

Tiến hành nhân gen *matK* bằng phản ứng PCR với cặp mồi đặc hiệu *matK-F/matK-R* từ DNA hệ gen của mẫu LML-02-CB, sau đó kiểm tra bằng điện di trên gel agarose 0,8% cho thấy, xuất hiện 1 băng DNA với kích thước ước tính khoảng 800bp (Hình 2).

Kết quả xác định trình tự *matK* từ mẫu LML-02-CB thu được 778 nucleotide. Trong đó, nucleotide loại A là 236 (30,33%), C là 120 (15,42%), G là 115 (14,78%), T là 307 (39,46%).

Kết quả so sánh độ tương đồng bằng chương trình so sánh BLAST trong NCBI thể hiện ở Hình 3.

Alignments Download GenBank Graphics Distance tree of results						
Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
Nervilia aragoana voucher SBB-0946 maturase K (matK) gene, partial cds; chloroplast	1330	1330	94%	0.0	99.19%	JN004498.1
Nervilia fordii voucher NFJX01 maturase K (matK) gene, partial cds; chloroplast	1360	1360	97%	0.0	99.08%	JX865503.1
Nervilia gammieana voucher SBB-0882 maturase K (matK) gene, partial cds; chloroplast	1323	1323	94%	0.0	99.05%	JN004506.1
Nervilia gammieana voucher SBB-0881 maturase K (matK) gene, partial cds; chloroplast	1323	1323	94%	0.0	99.05%	JN004505.1
Nervilia gammieana voucher SBB-0884 maturase K (matK) gene, partial cds; chloroplast	1319	1319	94%	0.0	99.05%	JN004508.1
Nervilia scottii maturase K (matK) gene, partial cds; chloroplast	1303	1303	93%	0.0	99.04%	KF421853.1
Nervilia mekongensis isolate Sando02 maturase K (matK) gene, partial cds; plastid	1380	1380	99%	0.0	98.97%	MG452081.1
Nervilia mekongensis isolate CC01 maturase K (matK) gene, partial cds; plastid	1380	1380	99%	0.0	98.97%	MG452080.1
Nervilia aragoana voucher NAGX01 maturase K (matK) gene, partial cds; chloroplast	1349	1349	97%	0.0	98.82%	JX865513.1

Hình 3. Kết quả phân tích sự tương đồng giữa trình tự *matK* của mẫu Lan một lá LML-02-CB với một số trình tự *matK* trên GenBank bằng BLAST trong NCBI

Kết quả cho thấy, đoạn gen *matK* phân lập từ mẫu Lan một lá LML-02-CB có độ tương đồng trên 98% so với 9 trình tự *matK* trên GenBank. Đặc biệt, trình tự *matK* của mẫu LML-02-CB có hệ số tương đồng 99,08% so với trình tự *matK* của loài *Nervilia fordii* mang mã số JX865503 và 99,19% so với trình tự *matK* của loài *Nervilia aragoana* mang mã số JN004498 trên GenBank. Như vậy, kết quả so sánh bằng BLAST trong NCBI đã khẳng định trình tự phân lập từ mẫu Lan một lá LML-02-CB là đoạn gen *matK*. Kết quả so sánh trình tự *matK* cho biết, mẫu Lan một lá LML-02-CB thuộc chi *Nervilia*.

Kết quả so sánh trình tự nucleotide của đoạn gen *matK* của mẫu Lan một lá LML-02-CB với hai trình tự gen *matK* mang mã số JX865503 và JN004498 trên GenBank bằng phần mềm BioEdit được thể hiện ở Hình 4.

Kết quả so sánh trình tự nucleotide trên Hình 4 cho thấy, đoạn gen *matK* của mẫu Lan một lá LML-02-CB và hai trình tự gen mang mã số JX865503 và JN004498 có độ tương đồng khá cao. Các vị trí nucleotide sai khác được thể hiện ở Bảng 1.

Bảng 1. Các vị trí sai khác giữa trình tự nucleotide của đoạn gen *matK* từ mẫu *matK*-02-CB với JX865503 và JN004498

Vị trí	JX865503	JN004498	<i>matK</i> -02-CB
372	-	-	T
378	G	G	T
451	-	-	G
486	C	C	T
487	T	T	C
493	G	G	T
568	-	-	A
775	T	T	G
776	G	G	A
778	A	A	T
779	T	T	G

Hình 4. So sánh trình tự matK của mẫu LML02-CB và hai trình tự mang mã số JX865503 và JN004498 trên GenBank

3.2. Sự đa dạng về trình tự nucleotide của đoạn gen *matK* của cây Lan một lá

Bảng BLAST trong NCBI xác định được trình tự đoạn gen *matK* phân lập từ mẫu Lan một lá LML02-CB có hệ số tương đồng cao hơn 97% so với 5 trình tự đoạn gen *matK* trên GenBank. Các trình tự gen này được sử dụng trong phân tích đa dạng về trình tự nucleotide của đoạn gen *matK* từ mẫu Lan một lá LML02-CB (Bảng 2).

Bảng 2. Mã số, năm công bố, quốc gia và tác giả của 5 trình tự gen *matK* trên GenBank

STT	Mã số	Loài	Năm công bố	Quốc gia	Tác giả
1	JN004498	<i>Nervilia aragona</i>	2016	Ấn Độ	Parveen I.
2	JX865503	<i>Nervilia fordii</i>	2012	Trung Quốc	Huang Q.
3	JN004505	<i>Nervilia gammieana</i>	2016	Ấn Độ	Parveen I.
4	MG452080	<i>Nervilia mekongensis</i>	2018	Trung Quốc	Gale S.W.
5	JN004513	<i>Nervilia aragona</i>	2016	Ấn Độ	Parveen I.

Dựa trên trình tự nucleotide của đoạn gen *matK* từ mẫu Lan một lá LML-02-CB và 5 trình tự trên GenBank tiến hành thiết lập bảng ma trận về hệ số tương đồng di truyền và hệ số phân ly về trình tự nucleotide bằng phần mềm DNASTAR. Kết quả được trình bày ở Bảng 3.

Bảng 3. Hệ số tương đồng và hệ số phân ly dựa trên trình tự đoạn gen *matK* từ mẫu LML-02-CB với các trình tự đoạn gen *matK* trên GenBank

		Percent Identity							
Divergence		1	2	3	4	5	6		
	1		99.5	97.4	97.9	99.3	98.1	1	matK-02-CB
	2	0.5		97.9	99.2	99.9	99.9	2	JN004498
	3	2.6	2.1		97.8	97.9	97.9	3	JN004513
	4	1.2	0.0	2.1		99.2	99.9	4	JX865503
	5	0.7	0.1	2.1	0.1		100.0	5	JN004505
	6	1.4	0.1	2.1	0.1	0.0		6	MG452080
		1	2	3	4	5	6		

Kết quả ở Bảng 3 cho thấy, hệ số tương đồng di truyền dựa trên trình tự nucleotide của đoạn gen *matK* từ mẫu Lan một lá LML-02-CB và các trình tự đoạn gen *matK* trên GenBank dao động từ 97,4% đến 100%, hệ số phân ly dao động từ 0,0% đến 2,6%. Trình tự đoạn gen *matK* của mẫu LML-02-CB có hệ số tương đồng so với các trình tự đoạn gen *matK* trên GenBank dao động từ 97,4% đến 99,5%.

4. Kết luận

Kích thước trình tự vùng *matK* của Lan một lá mẫu LML-02-CB là 778 bp. Trong đó, nucleotide loại A là 236 (30,33%), C là 120 (15,42%), G là 115 (14,78%), T là 307 (39,46%). Trình tự *matK* của mẫu LML-02-CB có hệ số tương đồng 99,08% so với trình tự *matK* của loài *Nervilia fordii* mang mã số JX865503 và 99,19% so với trình tự *matK* của loài *Nervilia aragona* mang mã số JN004498 trên GenBank. Hệ số tương đồng về trình tự nucleotide vùng *matK* của mẫu LML-02-CB so với 5 trình tự *matK* trên GenBank dao động từ 97,4% đến 99,5%. Có thể sử dụng chỉ thị *matK* là một trong các cơ sở trong việc giám định các mẫu Lan một lá.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] C.-L. Lu, "Studies on chemical constituents of petroleum ether extract with anti-tumor activity from *Nervilia fordii*," 2009. [Online]. Available: http://en.cnki.com.cn/Article_en/CJFDTOTAL-JNDX200905026.htm. [Accessed Apr. 2021].

-
- [2] Ministry of Science and Technology, *Vietnam's Red Book*. Publishing House of Natural Science and Technology, 2005, pp. 456-457.
 - [3] M. W. Chase, R. S. Cowan, and P. M. Hollingsworth, "A proposal for a standard protocol to barcode all land plants," *Taxon*, vol. 56, pp. 295-299, 2007.
 - [4] K. Vijayan and C. H. Tsou, "DNA barcoding in plants: taxonomy in a new perspective," *Current science*, vol. 99, pp. 1530-1540, 2010.
 - [5] H. L. Yong, R. Jinlan, C. Shilin, S. Jingyuan, L. Kun, L. Dong, and Y. Hui, "Authentication of *Taxillus chinensis* using DNA barcoding technique," *Journal of Medicinal Plants Research*, vol. 4, no. 24, pp. 2706-2709, 2010.
 - [6] Q. Huang, L. Liang, R. He, X. Ma, R. Zhan, and W. Chen, "Applying DNA barcoding to identify *Nervilia fordii* and six congeneric species," *POJ*, vol. 6, no. 5, pp. 325-332, 2013.
 - [7] T. T. Nguyen and T. T. L. Tran, "Morphological characteristics and ITS sequence isolated from *Nervilia fordii* (Hance) Schlechter collected in Cao Bang and Thai Nguyen province," *Proceeding of the 4th national conference on biology research and teaching in Vietnam*, 2020, pp. 216-222.