

## SỬ DỤNG MÃ VẠCH DNA *matK* ĐỂ ĐỊNH DANH MẪU SUM LIÊN THU TẠI TỈNH LÀO CAI, VIỆT NAM

Nguyễn Hữu Quân\*, Kiều Thị Trà Giang  
Trường Đại học Sư phạm - ĐH Thái Nguyên

### TÓM TẮT

Loài Sum liên có tên khoa học là *Adinandra lienii* phân bố ở khu vực miền núi phía Bắc, Việt Nam. Mẫu Sum liên thu tại tỉnh Lào Cai, Việt Nam đã được nhận diện bằng phân tích các đặc điểm hình thái dựa trên Khóa phân loại. Để khẳng định mẫu Sum liên thu được thuộc loài *Adinandra lienii*, nghiên cứu này trình bày kết quả sử dụng mã vạch DNA dựa trên trình tự đoạn gen *matK* phân lập từ mẫu Sum liên thu tại Lào Cai, Việt Nam. Đoạn gen *matK* của mẫu Sum liên có kích thước 867 nucleotide. Kết quả so sánh tương đồng bằng chương trình BLAST trong NCBI cho thấy trình tự đoạn gen *matK* của mẫu Sum liên thu tại Lào Cai thuộc loài *Adinandra lienii*, chi Dương đồng (*Adinandra*). Kết quả nghiên cứu này đã mở ra triển vọng sử dụng mã vạch DNA với trình tự gen *matK* trong định danh các loài thuộc chi *Adinandra*.

**Từ khóa:** *Adinandra lienii*, DNA barcoding, gen *matK*, giám định loài, Sum liên

Ngày nhận bài: 16/4/2019; Ngày hoàn thiện: 26/4/2019; Ngày duyệt đăng: 29/4/2019

## USE OF *MATK* DNA BARCODE TO IDENTIFY *Adinandra* SAMPLES COLLECTED AT LAO CAI, VIETNAM

Nguyen Huu Quan\*, Kieu Thi Tra Giang  
University of Education - TNU

### ABSTRACT

Sum lien has scientific name as *Adinandra lienii*, distributed in the North of Vietnam. The Sum lien sample collected in Lao Cai, Vietnam has been identified by analyzing morphological characteristics based on Lock classification. In order to confirm the Sum lien sample obtained from *Adinandra lienii*, this study presents the results using DNA barcodes based on the *matK* gene sequence isolated from the Sum lien samples collected in Lao Cai, Vietnam. The *matK* gene of Sum sample has the size of 867 nucleotide. The parallel comparison by BLAST in NCBI showed that the sequence of the *matK* gene of Sum lien sample collected in Lao Cai is among *Adinandra lienii* specie, *Adinandra* genus. The results of this study open up the prospect of using DNA barcodes with the *matK* gene sequence in identifying species belonging to the *Adinandra* genus.

**Key words:** *Adinandra lienii*, DNA barcoding, *matK* gene, species identification.

Received: 16/4/2019; Revised: 26/4/2019; Approved: 29/4/2019

\* Corresponding author: Tel: 0369 238303; Email: quannah@dhsptn.edu.vn

## MỞ ĐẦU

Chi *Adinandra* thuộc họ Chè Theaceae, thường là cây gỗ ít khi là cây bụi, nhánh non, có lông nhung. Hoa mọc đơn độc ở nách lá, lưỡng tính, mẫu 5. Lá đài có lông mềm hay lông ráp. Cánh hoa không lông hay chỉ có lông ở mặt ngoài. Nhị nhiều lên tới 25 nhị. Bao phấn có lông ngắn hay dài và có mũi nhọn. Bầu trên, không lông hay có lông mềm; noãn nhiều. Quả khô không tự mở; hạt nhiều, nhỏ. Chi *Adinandra* có khoảng 85 loài phân bố ở các nước châu Phi, Trung Quốc, Nhật Bản, Ấn độ, Srilanka, Bangladesh, và một số nước Đông Nam Á [1]. Ở Việt Nam, chi *Adinandra* đã được tìm thấy có 13 loài, phân bố ở các tỉnh Lào Cai, Cao Bằng, Quảng Ninh, Vĩnh Phúc, Quảng Trị, Kon Tum, Lâm Đồng, Gia Lai, Hà Giang [2].

Sum liên hay còn gọi Hồng đạm liên có tên khoa học là *Adinandra lienii* được mô tả vào năm 1986 và được công nhận vào năm 2012. Loài *Adinandra lienii* phân bố ở tỉnh Hà Giang và Lào Cai, Việt Nam. Loài *Adinandra lienii* chưa có bất kì một nghiên cứu nào công bố về đặc điểm sinh học, thành phần hóa học và di truyền. Ngoài ra, dựa vào đặc điểm hình thái và cấu tạo hiển vi rễ, thân và lá chúng tôi nhận thấy mẫu Sum liên thu tại Lào Cai đều thuộc loài *Adinandra lienii*.

Gen *matK* thuộc hệ gen lục lạp, có kích thước khoảng 1550 bp và mã hóa enzyme maturase liên quan đến quá trình loại bỏ các intron loại 2 trong quá trình phiên mã RNA. Do gen *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 và phát sinh loài ở thực vật. CBOL đã thử nghiệm gen *matK* trên gần 550 loài thực vật và thấy rằng 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 gen *matK* là một trong những locus barcode chuẩn cho thực vật [3], [4], [5].

Hiện nay, chưa có công bố nào về đặc điểm phân loại học phân tử, cũng như các đặc điểm sinh học của loài Sum liên. Các mẫu Sum liên thu thập tại các địa phương khác nhau có thể có sự sai khác về đặc điểm hình thái, vì vậy, cần thiết phải căn cứ vào các dữ liệu phân loại học phân tử để nhận diện loài *Adinandra lienii*. Trong bài báo này, chúng tôi sử dụng trình tự đoạn gen *matK* để định danh loài *Adinandra lienii* thu thập tại tỉnh Lào Cai và nhằm bổ sung cơ sở dữ liệu phân loại học phân tử của loài.

## VẬT LIỆU VÀ PHƯƠNG PHÁP

**Vật liệu:** Mẫu lá non Sum liên được thu tại huyện Bát Xát, tỉnh Lào Cai sử dụng để tiến hành phân lập đoạn gen *matK*.

**Hóa chất:** Các hóa chất tách chiết ADN tổng số do hãng Wako (Nhật bản) và Merck (Đức) cung cấp; Hóa chất PCR, điện di ADN do các hãng Fermentas (Đức), Bioneer (Hàn quốc), Research Organics (Mỹ) cung cấp.

## Phương pháp

Phân lập gen *matK*: DNA tổng số được phân lập dựa trên phương pháp của Shaghai và cộng sự (1984) [6]. Khuếch đại gen *matK* bằng phản ứng PCR với cặp mồi đặc hiệu được tổng hợp theo Kress và cộng sự (2005) [7]. Kích thước mồi và trình tự mong muốn của đoạn DNA khuếch đại được mô tả theo bảng 1.

Hỗn hợp phản ứng PCR (tổng thể tích 25 µl) gồm: 12,5 µl master mix (2X); 0,5 µl mồi mỗi loại (10 pmol/µl); 1,0 µl DNA khuôn (10 ng/µl); 9,5 µl nước cất. Phản ứng PCR được thực hiện theo chương trình: 94°C/4 phút; 35 chu kỳ (94°C/30 giây; 55°C/30 giây; 72°C/45 giây); 72°C/10 phút và giữ ở 4°C. Sản phẩm PCR được điện di trên gel agarose 1,0% và được tinh sạch theo kit tinh sạch của hãng Qiagen. Trình tự DNA được xác định trên máy đọc trình tự tự động ABI PRISM 3100 Avant Genetic Analyzer. Trình tự nucleotide của gen được phân tích bằng phần mềm BLAST, BioEdit và DNASTar.

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

Tên mồi	Trình tự mồi (5'→3')	Nhiệt độ gắn mồi	Kích thước đoạn gen dự kiến
<i>MatK-F</i>	5'- CGATCTATTCATTCAATATTTTC-3'	55°C	~ 800 bp
<i>MatK-R</i>	5'- TCTAGCACACGAAAGTCTGAAGT-3'		

**Bảng 2.** Các trình tự nucleotide của đoạn gen *matK* sử dụng trong phân tích

TT	Loài	Mã số trên GenBank	Tác giả	Năm	Kích thước (bp)
1	<i>Adinandra</i> sp. JH-2017	MF418698.1	Heckenhauer và cs	2017	827
2	<i>Adinandra</i> sp. JH-2017	MF418697.1	Heckenauer và cs	2017	826
3	<i>Adinandra millettii</i>	HQ427369.1	Pei và cs	2016	830
4	<i>Adinandra nitida</i>	KP093833.1	Liu và cs	2015	756
5	<i>Adinandra dumosa</i>	KU853076.1	Srivathsan và cs	2016	754

## KẾT QUẢ VÀ THẢO LUẬN

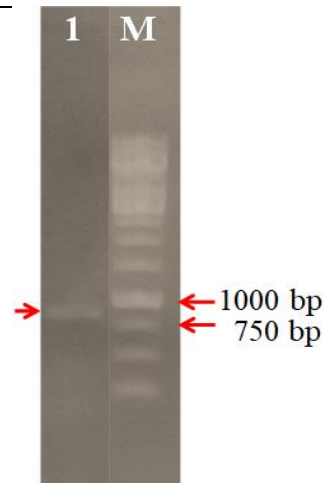
### Phân lập đoạn gen *matK* bằng phản ứng PCR

Lá non của mẫu Sum liên thu tại tỉnh Lào Cai sử dụng để tách DNA tổng số. Sản phẩm DNA tổng số được điện di trên gel agarose 0,8% và đo quang phổ để kiểm tra chất lượng tách. Kết quả kiểm tra cho thấy DNA tổng số thu được đảm bảo chất lượng cho phản ứng nhân gen.

Đoạn gen *matK* được phân lập bằng phản ứng PCR từ DNA tổng số sử dụng cặp mồi đặc hiệu *MatK-F/MatK-R*. Sản phẩm PCR được tiến hành điện di trên gel agarose 0,8%, kết quả điện di xuất hiện 1 băng có kích thước khoảng 800 bp, ứng với đoạn gen *matK* theo lý thuyết (Hình 1).

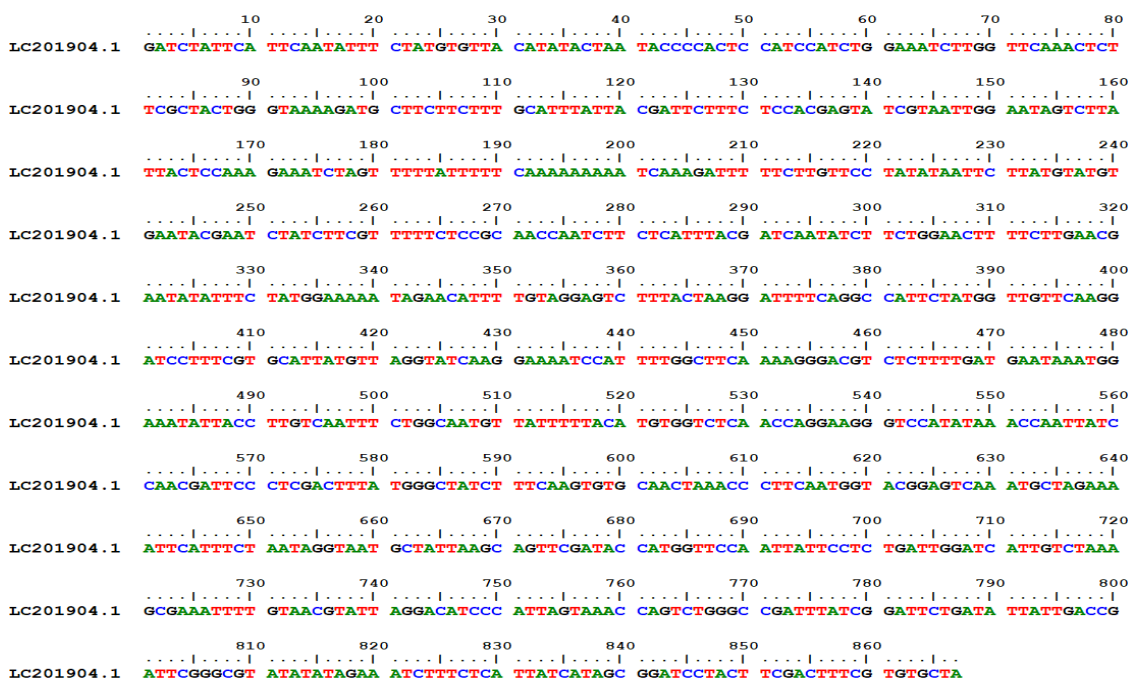
### Giải trình tự đoạn gen *matK* thu được từ phản ứng PCR

Đoạn gen *matK* được tiến hành tinh sạch và giải trình tự nucleotit trên máy giải trình tự tự động ABI PRISM 3100 Avant Genetic Analyzer. Kết quả giải trình tự đã xác định được đoạn gen *matK* phân lập từ loài Sum liên dài 867 nucleotit (Hình 2). Đoạn gen *matK* đã phân lập từ mẫu Sum liên thu tại Lào Cai được kí hiệu là LC201904.



**Hình 1.** Hình ảnh điện di sản phẩm PCR từ khuôn DNA tổng số của mẫu Sum liên  
Làn 1: đoạn gen *matK*, Làn M: DNA marker

Sử dụng phần mềm BLAST trong NCBI để phân tích sự tương đồng giữa trình tự của đoạn gen *matK* từ loài Sum liên với trình tự gen *matK* trên GenBank. Kết quả phân tích cho thấy hệ số tương đồng giữa trình tự nucleotide của đoạn gen *matK* phân lập từ mẫu Sum liên thu tại Lào Cai so với đoạn gen *matK* của các loài thuộc chi *Adinandra* trên GenBank dao động từ 99,39-100% (Hình 3).



**Hình 2.** Kết quả giải trình tự đoạn gen *matK* của loài *Sum liên* thu tại tỉnh Lào Cai

Alignments Download GenBank Graphics Distance tree of results

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/> <i>Adinandra millettii</i> voucher 14CS8453 plastid, complete genome	1526	1526	95%	0.0	100.00%	MF179492.1
<input type="checkbox"/> <i>Cleyera japonica</i> voucher Chase 1690 (K) trnK gene, partial sequence, and maturase K (matK) gene, complete cds: plastid	1526	1526	95%	0.0	100.00%	DQ924094.1
<input type="checkbox"/> <i>Cleyera japonica</i> maturase K (matK) gene, complete cds: chloroplast gene for chloroplast product	1526	1526	95%	0.0	100.00%	AF380078.1
<input checked="" type="checkbox"/> <i>Adinandra angustifolia</i> voucher 14CS9678 plastid, complete genome	1520	1520	95%	0.0	99.88%	MF179491.1
<input type="checkbox"/> <i>Cleyera japonica</i> maturase K (matK) gene, partial cds: chloroplast	1517	1517	94%	0.0	100.00%	HQ427371.1
<input checked="" type="checkbox"/> <i>Adinandra millettii</i> maturase K (matK) gene, partial cds: chloroplast	1517	1517	94%	0.0	100.00%	HQ427369.1
<input checked="" type="checkbox"/> <i>Adinandra dumosa</i> voucher Chase 1379 (K) trnK gene, partial sequence, and maturase K (matK) gene, complete cds: plastid	1504	1504	95%	0.0	99.52%	DQ924093.1
<input checked="" type="checkbox"/> <i>Adinandra millettii</i> maturase K (matK) gene, partial cds: chloroplast gene for chloroplast product	1498	1498	95%	0.0	99.39%	AF380069.1
<input type="checkbox"/> <i>Euryodendron excelsum</i> chloroplast, complete genome	1487	1487	95%	0.0	99.15%	NC_039178.1
<input type="checkbox"/> <i>Euryodendron excelsum</i> chloroplast, complete genome	1487	1487	95%	0.0	99.15%	MG770883.1
<input type="checkbox"/> <i>Eurya alata</i> chloroplast, complete genome	1482	1482	95%	0.0	99.03%	MH782188.1
<input type="checkbox"/> <i>Eurya loquaiana</i> maturase K (matK) gene, partial cds: chloroplast	1482	1482	95%	0.0	99.03%	HQ427372.1
<input type="checkbox"/> <i>Eurya rubiginosa</i> var. <i>attenuata</i> maturase K (matK) gene, partial cds: chloroplast	1482	1482	95%	0.0	99.03%	HQ427368.1
<input type="checkbox"/> <i>Eurya japonica</i> maturase K (matK) gene, complete cds: chloroplast gene for chloroplast product	1471	1471	95%	0.0	98.79%	AF380081.1
<input type="checkbox"/> <i>Eurya macarneyi</i> maturase K (matK) gene, partial cds: chloroplast	1467	1467	94%	0.0	99.02%	HQ415299.1
<input type="checkbox"/> <i>Eurya chinensis</i> maturase K (matK) gene, partial cds: chloroplast	1465	1465	94%	0.0	99.02%	HQ415300.1
<input checked="" type="checkbox"/> <i>Adinandra</i> sp. JH-2017 isolate 13-1859 maturase K (matK) gene, partial cds: chloroplast	1461	1461	91%	0.0	99.75%	MF418697.1

**Hình 3.** Kết quả phân tích sự tương đồng giữa trình tự đoạn gen *matK* của loài *Sum liên* so với trình tự đoạn gen *matK* trên GenBank bằng BLAST trong NCBI

**Phân tích mối quan hệ giữa loài *Sum liên* với các loài thuộc chi *Adinandra* dựa trên trình tự đoạn gen *matK***

Tiến hành phân tích mối quan hệ di truyền giữa loài *Sum liên* *Adinandra lienii* với một số loài *Adinandra millettii*, *Adinandra* sp. JH-2017, *Adinandra nitida* và *Adinandra dumosa* thuộc chi *Adinandra* dựa trên trình tự đoạn

gen *matK*. Kết quả so sánh đã thống kê được trình tự đoạn gen *matK* thuộc loài *Sum liên* *Adinandra lienii* và các trình tự từ các loài *Adinandra millettii* (mã số HQ427369.1), loài *Adinandra* sp. JH-2017 (mã số MF418697.1, MF418698.1), loài *Adinandra dumosa* (mã số KU853076.1) và loài *Adinandra nitida* (mã số KP093833.1) thuộc chi *Adinandra* có số

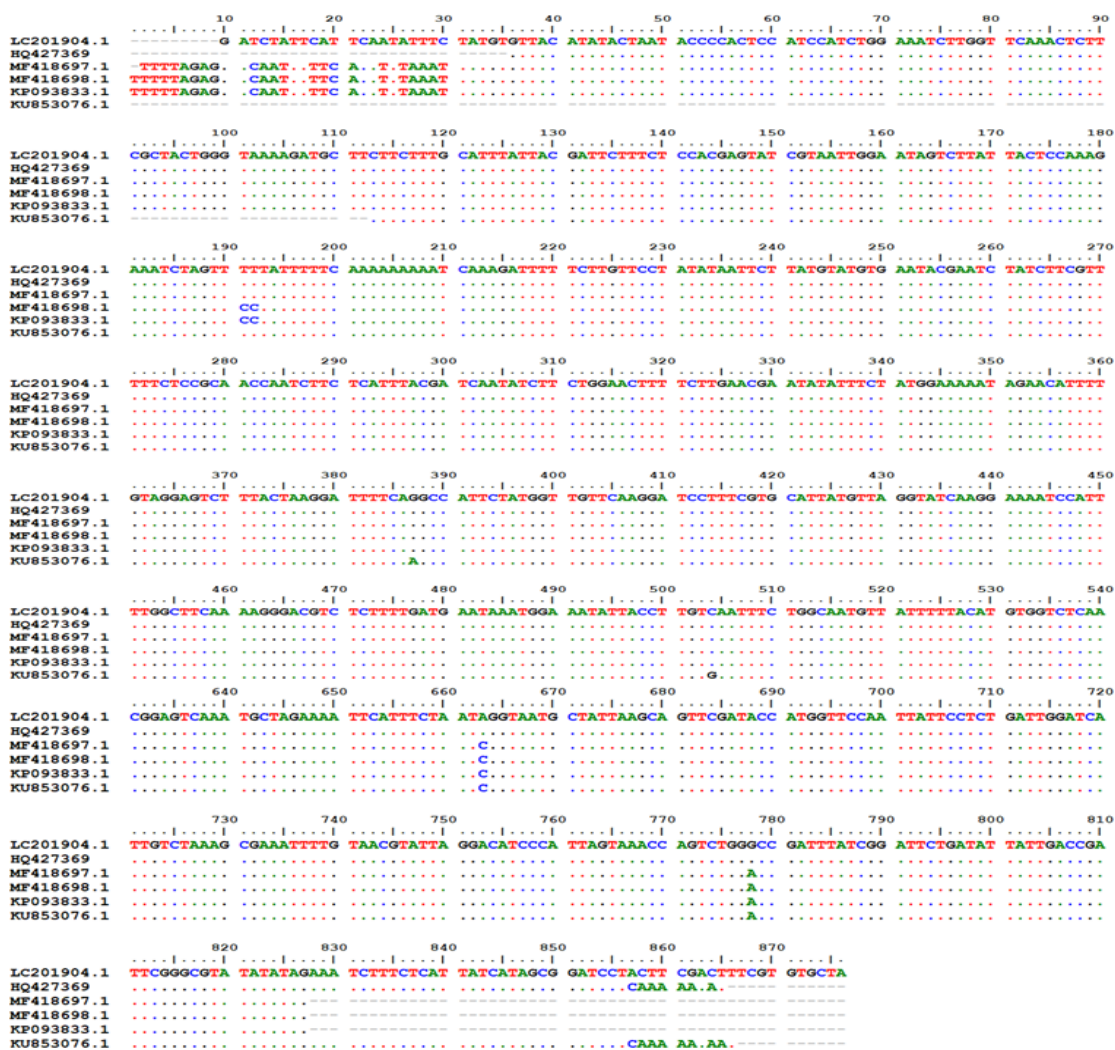
lượng nucleotide khác nhau (Bảng 2). Cụ thể, số nucleotide của các loài *Adinandra lienii* (trong nghiên cứu này) là 867, loài *Adinandra millettii* là 830, loài *Adinandra* sp. JH-2017 là 826-827, loài *Adinandra nitida* là 756 và loài *Adinandra dumosa* là 754 nucleotide. Kết quả này cho thấy, đoạn gen *matK* từ mỗi loài thuộc chi *Adinandra* có kích thước khác nhau và đặc trưng cho từng loài.

Trình tự nucleotide của đoạn gen *matK* của loài *Adinandra lienii* được so sánh với các loài trong bảng 2 tại vị trí nucleotide thứ 113 đến 628 cho thấy trình tự các nucleotide của 6 loài tương đối giống nhau, ngoại trừ một số điểm khác biệt tại các vị trí nucleotide số 192,

193 (T thay bằng C), 663 (A thay bằng C) và 778 (G thay bằng A). Các trình tự nucleotide còn lại liên quan tới sự sai khác số lượng nucleotide (Hình 4).

**Xây dựng cây phát sinh chủng loại**

Trình tự nucleotide của đoạn gen *matK* từ 6 loài được chúng tôi sử dụng phần mềm MegAlign để xác định hệ số tương đồng, hệ số phân li và cây phát sinh chủng loại. Kết quả bảng 3 cho thấy hệ số tương đồng và hệ số phân li trình tự của đoạn gen *matK* từ loài *Adinandra lienii* với các loài thuộc chi *Adinandra* dao động từ 95,9-100% và hệ số phân li từ 0-2,2%.



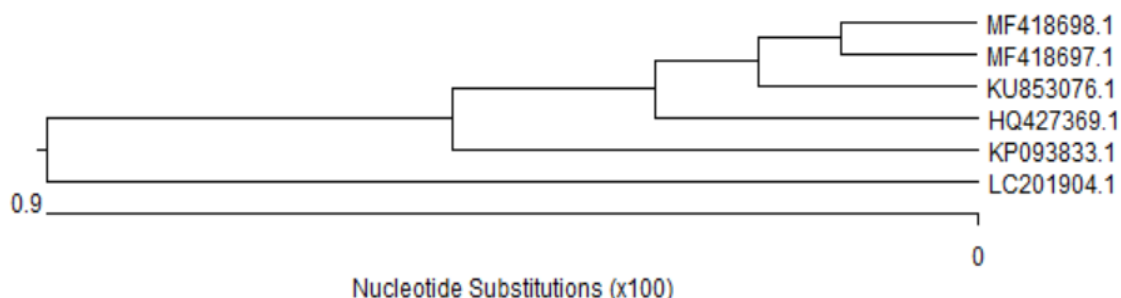
**Hình 4.** Trình tự nucleotide của đoạn gen *matK* phân lập từ loài Sum liên *Adinandra lienii* thu tại Lào Cai (Việt Nam) và trình tự đoạn gen *matK* công bố trên GenBank



**Bảng 3.** Hệ số tương đồng và hệ số phân ly trình tự các nucleotide của đoạn gen *matK* từ loài *Adinandra lienii* và các loài thuộc chi *Adinandra*

		Hệ số tương đồng (%)							
		1	2	3	4	5	6		
Hệ số phân ly (%)	1	█	98.9	95.9	100.0	96.2	98.1	1	LC201904.1
	2	0.8	█	95.3	100.0	95.6	99.2	2	HQ427369.1
	3	2.2	0.5	█	96.6	99.8	94.3	3	MF418698.1
	4	0.0	0.0	0.5	█	96.8	97.2	4	KP093833.1
	5	2.0	0.3	0.2	0.3	█	94.6	5	MF418697.1
	6	1.6	0.7	0.6	0.5	0.3	█	6	KU853076.1
		1	2	3	4	5	6		

Khi nghiên cứu về cây phát sinh chủng loại, trình tự nucleotide của đoạn gen *matK* từ 6 loài được xếp vào các nhánh khác nhau. Loài *Adinandra lienii* (LC201904.1) nằm một nhánh, gần với loài *Adinandra nitida* (KP093833.1). Kết quả cho thấy, loài Sum lông *Adinandra lienii* trong nghiên cứu không trùng lặp với các loài đã công bố (Hình 5). Như vậy, thông qua hệ số tương đồng, hệ số phân ly và cây phát sinh chủng loại của đoạn gen *matK* phân lập từ loài Sum lông *Adinandra lienii* thu tại Lào Cai (Việt Nam) trong nghiên cứu này đã bước đầu khẳng định được chính xác loài nghiên cứu và là công bố đầu tiên bổ sung dữ liệu về sinh học phân tử của loài *Adinandra lienii* thuộc chi *Adinandra*.



**Hình 5.** Sơ đồ cây phân loại dựa trên trình tự các nucleotide của đoạn gen *MatK*

**KẾT LUẬN**

Đã phân lập được đoạn gen *matK* từ mẫu Sum liên với kích thước là 867 nucleotide. Dựa trên trình tự đoạn gen *matK* bằng BLAST trong NCBI đã chứng minh được mẫu Sum liên thu tại tỉnh Lào Cai, Việt Nam là loài *Adinandra lienii* thuộc chi *Adinandra*. Đã xây dựng được cây phát sinh chủng loại dựa trên phần mềm MegAlige trên khung đọc gen *matK* xác định được mối quan hệ di truyền giữa các loài thuộc chi *Adinandra*.

**TÀI LIỆU THAM KHẢO**

[1]. Thực vật chí Trung quốc <http://www.efloras.org/>.  
 [2]. Phạm Hoàng Hộ, *Cây cỏ Việt Nam 2*, Nxb Trẻ, 2000.

[3]. K. W Hilu, "The *matK* gene: sequence variation and application in plant systematics", *American Journal of Botany*, 84, pp. 830-839, 1997.  
 [4]. H. L. Yong, R. Jinlan, C. Shilin (2010) "Authentication of *Taxillus chinensis* using DNA barcoding technique", *Journal of Medicinal Plants Research*, 4(24), pp. 2706 – 2709, 2010.  
 [5]. K. Vijayan and C. H. Tsou, "DNA barcoding in plants: taxonomy in a new perspective", *Current science*, 99, pp. 1530-1540, 2010.  
 [6]. M. A. Shaghai-Marroof, K. M. Soliman, R. A. Jorgensen, R. W. Allard, "Ribosomal DNA spacer-length polymorphism in barley: mendelian inheritance, chromosomal location, and population dynamics", *Proc. Natl. Acad. Sci.*, 81, pp. 8014-8019, 1984.  
 [7]. J. W. Kress, K. J. Wurdack, E. A. Zimmer, L. A. Weigh, D. H. Janzen, "Use of DNA barcodes to indentify flowering plants", *Proc. Natl. Acad. Sci. USA*, 102, pp. 8369-8374, 2005.