

## ĐỊNH DANH LOÀI CẤP GAI NHỎ *CAPPARIS* SP. Ở KHU BẢO TỒN THIÊN NHIÊN TÀ KÓU HUYỆN HÀM THUẬN NAM, TỈNH BÌNH THUẬN BẰNG MÃ VẠCH DNA

Nguyễn Thi Thu Ngà\*, Sỹ Danh Thường

Trường Đại học Sư phạm - Đại học Thái Nguyên

### TÓM TẮT

Đa dạng sinh học trên thế giới đang liên tục suy giảm, trong đó nhiều loài có nguy cơ dần biến mất. Một dự án chiến lược trên toàn thế giới - mã vạch DNA đã được công bố gần đây nhằm xây dựng một thư viện mã vạch của các loài đang bị đe dọa (CBOL ABS Brochure, 2012). Phương pháp mã vạch DNA sử dụng một trình tự DNA ngắn trong genome của sinh vật như chuỗi ký tự duy nhất giúp phân biệt các loài. Loài Cấp gai nhỏ *Capparis* sp. phân bố rải rác ở một số tỉnh của Việt Nam. Trong nghiên cứu này, nhằm hỗ trợ cho việc định danh loài cũng như bước đầu công bố các thông tin bằng mã vạch DNA của loài Cấp gai nhỏ thu tại Khu bảo tồn thiên nhiên Tà Kóu, huyện Hàm Thuận Nam, tỉnh Bình Thuận được chính xác, phương pháp sử dụng mã vạch DNA dựa trên đoạn gen *matK* đã được thực hiện. Đoạn gen *matK* từ loài Cấp gai nhỏ đã được phân lập dựa trên cặp đặc hiệu *matK-F/matK-R* có chiều dài 800 nucleotide. Trình tự đoạn gen *matK* từ loài Cấp gai nhỏ được so sánh với 7 trình tự của các loài thuộc chi *Capparis* công bố trên GenBank thông qua các phần mềm BLAST trong NCBI, BioEdit và MEGA5. Kết quả đã chứng minh được mẫu nghiên cứu thuộc loài *Capparis micracantha*.

*Từ khóa:* *Capparis micracantha*, gen *matK*, mã vạch, mã vạch DNA, phân biệt loài.

### MỞ ĐẦU

Đến nay, các mẫu sinh vật vẫn thường được nhận diện bằng các đặc tính hình thái bên ngoài hoặc các đặc tính sinh lý sinh hóa bên trong nhờ vào bảng hướng dẫn định danh có sẵn. Tuy nhiên, trong nhiều trường hợp, mẫu vật chưa phát triển đầy đủ các đặc tính hình thái, hoặc chúng bị hư hỏng các bộ phận ngoài, hoặc mẫu vật chết đã khiến quá trình nhận diện mẫu vật trở nên khó khăn thậm chí là không thể. Trong những trường hợp này, mã vạch DNA (DNA barcode) đã giúp giải quyết vấn đề trên vì trình tự DNA dễ dàng thu nhận từ một mẫu mô rất bé. Bên cạnh việc định danh, mã vạch DNA còn giúp phân tích quá trình tiến hóa của loài sinh vật đó trong tự nhiên.

Mã vạch DNA là một khái niệm mới được đưa ra bởi Heber vào năm 2003. Mã vạch DNA là trình tự nucleotide của một chuỗi DNA ngắn, có cùng nguồn gốc tổ tiên (orthologous), trong đó có vùng ít bị thay đổi (rất ổn định – bảo thủ) và có vùng dễ thay đổi trong quá trình tiến hóa. Dựa vào mức độ thay đổi trong trình tự DNA này để đánh giá sự sai khác di truyền giữa các sinh vật. Như vậy, sử dụng mã vạch DNA để định danh là phương pháp định danh mới sử dụng một đoạn DNA chuẩn ngắn nằm trong genome của sinh vật đang nghiên cứu, nhằm xác định sinh vật đó thuộc về loài nào (<http://www.barcodeoflife.org>). Từ cơ sở dữ liệu này, hướng nghiên cứu xây dựng cơ sở dữ liệu mã vạch DNA đang được nhiều quốc gia, nhiều nhà khoa học trên thế giới rất quan tâm phát triển, đặc biệt trong những năm gần đây và sẽ là một xu thế nghiên cứu trong thời gian tới (Valentini, 2009; Xiwen, 2014). Mã vạch DNA được xem là một công cụ mới, hỗ trợ có hiệu quả trong nghiên cứu về phân loại, phát hiện loài mới, giám định loài và các mẫu có nguồn gốc từ sinh vật sống hoặc đã chết.

Đoạn DNA đặc trưng được sử dụng làm mã vạch có thể là những đoạn DNA nằm ở trong nhân (nuclear DNA - nDNA), như: 18S, 5.6S, 26S, 5S spacer và vùng *ITS*; nằm ở ty thể (Mitochondrial DNA - mtDNA), như: *Cytb* và vùng kiểm soát (control region); nằm ở lục lạp (Chloroplast DNA - cpDNA), như: *matK*, *rbcl*, *atpβ*, *ndnF*, 16S (Cuenoud, 2002; Kress *et al*, 2008). Có rất nhiều gen cpDNA tham gia trong phân tích phân loại thực vật như: 16S, *rbcl*, *atpβ*, *ndhF*, intron *trnL* và *matK*,... trải rộng từ bộ cho đến mức dưới loài. Mỗi đoạn mã vạch DNA có những đặc trưng riêng và có khả năng phân biệt sinh vật ở các mức độ khác nhau. Ví dụ, các đoạn DNA như: 18S; 16S; 5.6S có khả năng phân biệt sinh vật ở mức họ và chi; các đoạn DNA, như: 26S, *rbcl*, *ndnF*, *atpβ* có khả năng phân biệt ở mức chi và loài; các đoạn DNA, như: *ITS*, *matK*, *cytb*, 5S spacer có khả năng phân biệt ở mức loài và dưới loài (subspecies, variety, strain) (Ford *et al.*, 2009; Li *et al.*, 2010; Hebert, 2003). Tuy nhiên, các nhà khoa học cũng đã khuyến cáo, chưa có đoạn DNA nào được sử dụng làm mã vạch chung cho tất cả các loài sinh vật. Vì vậy, việc lựa chọn những đoạn DNA (gen) đặc trưng để làm mã vạch và việc phối hợp giữa các đoạn mã vạch DNA là rất cần thiết và đem lại hiệu quả cao (Kress *et al.*, 2008; CBOL Plant Working Group, 2009; Vijayan, 2010).

Ở thực vật, tốc độ tiến hóa của các gen ty thể không nhanh như ở động vật, do vậy không đủ sai khác để phân biệt được các loài. Vì vậy, thường sử dụng các đoạn DNA đặc trưng ở lục lạp và nhân để làm mã vạch DNA (Kress *et al.*, 2008). Tùy vào mức độ phân tích mà mỗi vùng gen và mỗi phương pháp được áp dụng phù hợp. Tổ chức hệ thống cơ sở dữ liệu mã vạch sự sống (BOLD) khuyến cáo, ở thực vật ngoài vùng *ITS* ở trong nhân, nên dùng thêm các marker khác riêng rẽ hoặc kết hợp. Nghiên cứu của Hollingsworth và đồng tác giả (2011) đã chỉ ra sự kết hợp giữa gen *matK* và gen *rbcl* ở lục lạp để sử dụng như một mã vạch DNA có hiệu quả cao nhất trong

định danh phân tử cho nhiều loài thực vật (Gao, 2011). Sau đó, Kress và đồng tác giả (2008) bổ sung thêm vùng gen thứ ba là vùng xen *trnH-psbA*. Tính đến năm 2009, đã có khoảng 8 locus gen được sử dụng làm mã vạch DNA ở các loài thực vật, bao gồm cả hệ gen nhân và hệ gen lục lạp (vùng xen *atpF-atpH*, gen *matK*, gen *rbcl*, gen *rpoC1*, vùng xen *psbK-psbI*, vùng xen *trnH-psbA* và vùng gen nhân *ITSAB*).

Maturase K (*matK*) là một trong những trình tự được nghiên cứu và ứng dụng thành công trong việc xác định các loại thuốc thảo dược như "Dahuang" có nguồn gốc từ *Rheum palmatum* L. (Polygonaceae), *R. tanguticum* (Maxim. ex Regel) Maxim. ex Balf., *R. officinale* Baill., và loài gần gũi *Rheum* L. với mức độ biến đổi nội bộ loài và giữa các loài khác nhau là cao. Vì vậy, nó thường được sử dụng để xác định các nguyên liệu thảo dược ở những vị trí địa lý khác nhau (Yang, 2004).

Việt Nam có hệ động, thực vật phong phú nhưng đến nay nhiều loài chưa có dữ liệu về mã vạch DNA. Các loài trong chi *Cap* (*Capparis*) có nhiều sự tương đồng về mặt hình thái dẫn đến có thể nhầm lẫn trong phân loại loài. Chúng tôi hy vọng việc nghiên cứu phương pháp DNA mã vạch sẽ cung cấp cái nhìn sâu hơn vào phân loại học ở cấp độ loài và đóng góp vào quá trình phân loại học, nhận diện, phân chia ranh giới giữa các loài, hỗ trợ quá trình nhận diện các mẫu vật không nhận biết được hoặc chưa xác định được thuộc loài nào.

**NGUYÊN LIỆU VÀ PHƯƠNG PHÁP**

**Nguyên liệu thực vật**

Mẫu lá cây Cáp gai nhỏ *Capparis* sp. sử dụng trong nghiên cứu do nhóm nghiên cứu thu thập trong chuyến đi thực tế tại Khu bảo tồn thiên nhiên Tà Kóu, huyện Hàm Thuận Nam, tỉnh Bình Thuận (Hình 1). Số hiệu mẫu VK 3412.



Hình 1. Cây Cáp gai nhỏ

**Hóa chất**

Các hóa chất sử dụng trong nghiên cứu được đặt từ các hãng uy tín đảm bảo về chất lượng và độ tin cậy như Qiagen, Fermentas, Sigma...

**Thiết bị sử dụng trong nghiên cứu**

Một số thiết bị chính được sử dụng trong nghiên cứu: cân điện tử Satorius (Thụy Sĩ), bể ổn nhiệt, máy đo quang phổ, máy đo pH, máy li tâm, máy PCR.

**Cặp mồi sử dụng trong nghiên cứu**

Cặp mồi được sử dụng để nhân bản đoạn gen đích là gen *matK*, trình tự và thông tin về cặp mồi được thể hiện ở Bảng 1.

Bảng 1. Thông tin về cặp mồi sử dụng trong nghiên cứu

Tên mồi	Trình tự mồi	Nhiệt độ gắn mồi	Kích thước dự kiến
matK-F	5'- CGATCTATTCATTCAATATTTTC-3'	54°C	Khoảng 800 bp
matK-R	5'- TCTAGCACACGAAAGTCGAAGT-3'		

Cặp mồi *matK-F/matK-R* sử dụng trong phản ứng PCR nhân bản gen *matK* được được tổng hợp theo Kress (2005).

**Phương pháp nghiên cứu**

Tách chiết DNA tổng số: Mẫu lá của cây *Capparis* sp. (VK 3412) được sử dụng để tách chiết DNA tổng số bằng quy trình mini-CTAB và kit DNeasy® Plant Mini Kit của hãng Qiagen.

Kiểm tra sản phẩm DNA sau tách chiết: DNA tổng số của mẫu *Capparis* sp. (VK 3412) sau khi tách chiết được điện di trên gel agarose 1%, trong đệm TBE 1X, ở hiệu điện thế 90V với marker 1 kb. Nồng độ và độ tinh sạch

của DNA tổng số được kiểm tra qua phương pháp đo quang phổ hấp phụ ở bước sóng 260 và 280 nm. Pha loãng dịch chiết 100 lần (5  $\mu\text{L}$  mẫu + 450  $\mu\text{L}$  DDW), mỗi mẫu đo ba lần và giá trị TB của ba lần đo được lấy làm kết quả cuối cùng. Độ tinh sạch của mẫu thể hiện qua thông số A260/280.

Phương pháp nhân gen bằng kỹ thuật PCR:

Phản ứng PCR được thực hiện trên máy PCR 9700 (Applied Biosystems).

Thành phần phản ứng PCR bao gồm: Master mix (2X) – 7,5  $\mu\text{L}$ , mỗi xuôi (10 pmol/ $\mu\text{L}$ ) - 0,5  $\mu\text{L}$ , mỗi ngược (10 pmol/ $\mu\text{L}$ ) - 0,5 $\mu\text{L}$ , DNA khuôn (10 ng/ $\mu\text{L}$ )- 0,5  $\mu\text{L}$ , H<sub>2</sub>O - 6  $\mu\text{L}$ . Tổng thể tích phản ứng là 15  $\mu\text{L}$ .

Chu trình nhiệt: 94°C/3 phút; lặp lại 35 chu kỳ với (94°C/30 giây, 54°C/40 giây, 72°C/30 giây); 72°C/10 phút và giữ ở 4°C.

Phương pháp giải trình tự gen.

Phương pháp giải trình tự gen thực hiện dựa trên kỹ thuật “chain termination- kết thúc chuỗi” bằng việc sử dụng các deoxynucleotide đã bị chỉnh sửa làm mất nhóm hydroxyl ở đầu 3' của phân tử đường. Khi một ddNTP được thêm vào chuỗi, vì không có nhóm 3'-OH nên một nucleotide không được thêm vào, phản ứng tổng hợp sẽ dừng lại. Enzyme polymerase xúc tác phản ứng gắn các dNTP vào mạch đơn của DNA để kéo dài mạch ở vị trí 3'-OH và dừng lại nếu gắn các ddNTP vào chuỗi. Trong một phản ứng chứa cả dNTP và ddNTP nên sau phản ứng tổng hợp, hình thành các đoạn DNA có độ dài khác nhau do ddNTP gắn ngẫu nhiên làm dừng phản ứng. Dựa vào sự sai khác về độ dài các đoạn DNA hiển thị trên bản gel sau điện di để xác định trình tự trong gene.

Trình tự nucleotide của đoạn gen *matK* 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 mỗi đặc hiệu. Giải trình tự gen được tiến hành tại phòng thí nghiệm trọng điểm, Viện Công nghệ Sinh học trực thuộc Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

Trình tự gen được phân tích, so sánh bằng chương trình BioEdit.

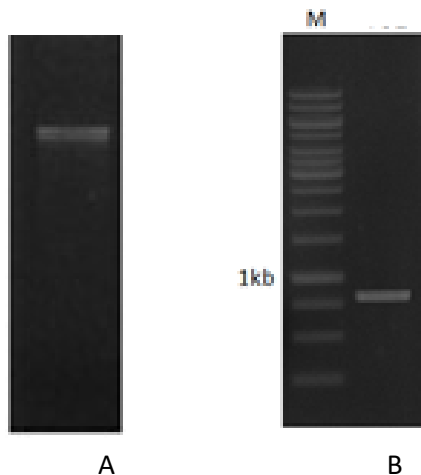
Phương pháp xây dựng cây phát sinh chủng loại: Cây phát sinh chủng loại được xây dựng bằng phần mềm MEGA5.

## KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

### Kết quả tách chiết và tinh sạch DNA tổng số

Để tách chiết DNA có độ tinh sạch đảm bảo cho thí nghiệm thì việc lựa chọn một phương pháp tách chiết phù hợp với đối tượng nghiên cứu là rất quan trọng. Hiện nay, phương pháp tách chiết DNA sử dụng CTAB là phương pháp khá phổ biến để tách chiết DNA từ các mẫu có nguồn gốc thực vật. Phương pháp sử dụng lần đầu tiên được Murray và Thompson mô tả vào năm 1980, tới nay đã được sử dụng rất phổ biến có hiệu quả với các mẫu khó tách chiết như thực vật giàu polysaccharide hay các sản phẩm có nguồn gốc từ thực vật.

Chúng tôi đã tách chiết thành công DNA của mẫu lá từ cây *Capparis* sp. (VK 3412) thu tại Khu bảo tồn thiên nhiên Tà Kóu, huyện Hàm Thuận Nam, tỉnh Bình Thuận. Kết quả thí nghiệm được thể hiện trên Hình 2 (A).



Hình 2. Kết quả điện di DNA tổng số của mẫu lá cây *Capparis* sp. (A) và sản phẩm PCR nhân đoạn gen *matK* của mẫu nghiên cứu (B) trên gel agarose 1%, M: marker 1Kb

### Khuyếch đại vùng gen nghiên cứu bằng kỹ thuật PCR

Chúng tôi đã khuyếch đại thành công vùng gen nghiên cứu với cặp mồi *matK*-F; *matK*-R, nhiệt độ bắt cặp mồi đặc hiệu ở 54°C. Hình ảnh điện di sản phẩm PCR được trình bày trên Hình 2 (B).

Kết quả khuyếch đại đoạn gen *matK* cho thấy, trên hình ảnh điện di chỉ xuất hiện một băng sản phẩm với kích thước tính toán như dự đoán (khoảng 800 bp).

### Giải trình tự đoạn gen *matK* và phân tích trình tự thu được từ phản ứng khuyếch đại

Chúng tôi đọc trình tự đoạn gen *matK* trực tiếp từ sản phẩm PCR. Kết quả giải trình tự và so sánh với một số trình tự gen *matK* được công bố trên ngân hàng gen được trình bày trong Hình 3 và Hình 4. Kết quả giải trình tự cho thấy, đoạn gen *matK* phân lập được từ mẫu nghiên cứu có chiều dài 800 bp.

Sequences producing significant alignments		Download	Manage Columns	Show	100		
<input checked="" type="checkbox"/>	select all 100 sequences selected			GenBank	Graphics	Distance tree of results	
	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/>	<a href="#">Capparis micracantha chloroplast matK gene for maturase K, partial cds, specimen voucher: KYUM&lt;JPN&gt; 481</a>	1408	1408	98%	0.0	99.87%	<a href="#">AB925022.1</a>
<input checked="" type="checkbox"/>	<a href="#">Capparis zeylanica voucher MO.S.M. Phillips and A. Weerasooriya 82 maturase K (matK) gene, complete cds, chloroplast</a>	1380	1380	100%	0.0	98.59%	<a href="#">EU371779.1</a>
<input checked="" type="checkbox"/>	<a href="#">Capparis spinosa voucher WIS.J. Rodman 533 maturase K (matK) gene, complete cds, chloroplast</a>	1358	1358	100%	0.0	98.08%	<a href="#">EU371772.1</a>
<input checked="" type="checkbox"/>	<a href="#">Capparis ovata voucher KMillenium Seed Bank 243816 trnK-UUUU gene, partial sequence, and maturase K (matK) gene, complete cds, chloroplast</a>	1352	1352	100%	0.0	97.95%	<a href="#">KJ739546.1</a>
<input checked="" type="checkbox"/>	<a href="#">Capparis spinosa maturase (matK) gene, partial cds, chloroplast</a>	1349	1349	100%	0.0	97.82%	<a href="#">AY491650.1</a>
<input checked="" type="checkbox"/>	<a href="#">Capparis micracantha chloroplast gene for maturase K, partial cds, isolate, CHUIA-050</a>	1339	1339	94%	0.0	99.46%	<a href="#">LC438891.1</a>
<input checked="" type="checkbox"/>	<a href="#">Capparis sandwichiana voucher WIS.H.H. Iltis 30502 maturase K (matK) gene, complete cds, chloroplast</a>	1330	1330	100%	0.0	97.44%	<a href="#">EU371768.1</a>
<input checked="" type="checkbox"/>	<a href="#">Capparis elaeagnoides voucher MO.DeChamps 11698 maturase K (matK) gene, complete cds, chloroplast</a>	1330	1330	100%	0.0	97.56%	<a href="#">EU371759.1</a>
<input checked="" type="checkbox"/>	<a href="#">Capparis sepiaria voucher OM2746 maturase K (matK) gene, partial cds, chloroplast</a>	1325	1325	96%	0.0	98.53%	<a href="#">JX517328.1</a>
<input checked="" type="checkbox"/>	<a href="#">Capparis urophylla isolate L074 maturase K (matK) gene, partial cds, chloroplast</a>	1314	1314	94%	0.0	98.91%	<a href="#">KR530525.1</a>

Hình 3. Kết quả so sánh trình tự đoạn gen *matK* nghiên cứu với một số trình tự được công bố trên NCBI

Kết quả so sánh với một số trình tự được công bố trên ngân hàng gen cho thấy, trình tự gen *matK* nghiên cứu có sự tương đồng lớn nhất với trình tự mang mã số AB925022.1 (đến 99,87%) của loài *C. micracantha* (Hình 2), chỉ có sự thay đổi một nucleotide (T thành G) (vị trí đánh dấu trên Hình 3).

```

Query 5 GCCTCTCTTTGGCATTATTGGCATTCTGCTATATGAGTATTGGAATTGGAAGAATTTT 64
      |||
Sbjct 1 GCCTCTCTTTGGCATTATTGGCATTCTGCTATATGAGTATTGGAATTGGAAGAATTTT 60
Query 65 TATACTCAAAAAAAAAATCAATTTTGAATCAAGATTTTCTGTTCTTATATAATTCTCAT 124
      |||
Sbjct 61 TATACTCAAAAAAAAAATCAATTTTGAATCAAGATTTTCTGTTCTTATATAATTCTCAT 120
Query 125 GTATGTGAATACGAATCCATCTTTTTCTACGCAAGCGGCTTCTCATTACGATCC 184
      |||
Sbjct 121 GTATGTGAATACGAATCCATCTTTTTCTACGCAAGCGGCTTCTCATTACGATCC 180
Query 185 ACATCTTCTGCAGTCTTTTTGAGCGAAATTTTTCTATGGAAAAATAGAACATCTTGTA 244
      |||
Sbjct 181 ACATCTTCTGCAGTCTTTTTGAGCGAAATTTTTCTATGGAAAAATAGAACATCTTGTA 240
Query 245 AAAGTCTTTGTTAATGATTTTCAGGACATCCTAGGATGCTCAAGGATCCTTTCATACAT 304
      |||
Sbjct 241 AAAGTCTTTGTTAATGATTTTCAGGACATCCTAGGATGCTCAAGGATCCTTTCATACAT 300
Query 305 TATGTTAGATATCACGCAAAATGATTTTGGTAGCAAGGATACGCCGCTTCAATGAAT 364
      |||
Sbjct 301 TATGTTAGATATCACGCAAAATGATTTTGGTAGCAAGGATACGCCGCTTCAATGAAT 360
Query 365 AAATGGAAATATTATTTGTTAATTTATGGCAATGGCATTCTTCCGTATGGTTCAACCG 424
      |||
Sbjct 361 AAATGGAAATATTATTTGTTAATTTATGGCAATGGCATTCTTCCGTATGGTTCAACCG 420
Query 425 CAAAAGGTTTCATATAAATGAATATCTAAAGATAATTTATACCTTTCTGGGCTATCTCTCA 484
      |||
Sbjct 421 CAAAAGGTTTCATATAAATGAATATCTAAAGATAATTTAGACTTTCTGGGCTATCTCTCA 480
Query 485 AGTTTGGCATTAAATCCTTTACTGGTAGTACTCAAATGCTAGAAATCTCATTCTAATA 544
      |||
Sbjct 481 AGTTTGGCATTAAATCCTTTACTGGTAGTACTCAAATGCTAGAAATCTCATTCTAATA 540
Query 545 GATAATGTTAGAAAGAACTCGATACAAAAATCCAAATTTGTTCTAATTATGGGTCATTG 604
      |||
Sbjct 541 GATAATGTTAGAAAGAACTCGATACAAAAATCCAAATTTGTTCTAATTATGGGTCATTG 600
Query 605 GCTAAAGAAAGATTTGTAATGTATTAGGACATCCCATAGTAAATCTACCTGGATGGAT 664
      |||
Sbjct 601 GCTAAAGAAAGATTTGTAATGTATTAGGACATCCCATAGTAAATCTACCTGGATGGAT 660
Query 665 TCATCAGATCTGATATCTCGACCGATTTGTGCGTATATCCAGAAATCTTCTCATTAT 724
      |||
Sbjct 661 TCATCAGATCTGATATCTCGACCGATTTGTGCGTATATCCAGAAATCTTCTCATTAT 720
Query 725 CACAGTGGATCTTCAAAAAAAAAAATTTGATTCGAATAAATAT 769
      |||
Sbjct 721 CACAGTGGATCTTCAAAAAAAAAAATTTGATTCGAATAAATAT 765 |
    
```

Hình 4. Kết quả so sánh trình tự đoạn gen *matK* nghiên cứu với trình tự gen *matK* mã số AB925022.1 của loài *C. micracantha* được công bố trên ngân hàng gen

Chúng tôi cũng đã tiến hành so sánh trình tự amino acid suy diễn từ mẫu nghiên cứu và trình tự mang mã số AB925022.1 của loài *Capparis micracantha* được công bố trên ngân hàng gen (Hình 5). Kết quả so sánh cho thấy, từ sự sai khác trong trình tự nucleotide đã dẫn đến sự sai khác trong trình tự amino acid (D - Asparagine thành Y - Tyrosine).

Capparis m	5	15	25	35	45	55
VK3412	ASLLHLLRFC	LVEYVNHKNF	YTOQKKSILNT	RFLLFLYNISH	VCEYESIFFF	LRKRSSHLSRS
	ASLLHLLRFC	LVEYVNHKNF	YTOQKKSILNT	RFLLFLYNISH	VCEYESIFFF	LRKRSSHLSRS
Capparis m	65	75	85	95	105	115
VK3412	TSSAVLFERI	FFYKGIEHLV	KVFVNDFOI	LGLLKDFFIH	YVRYHAKCIL	VAKDTPLLMN
	TSSAVLFERI	FFYKGIEHLV	KVFVNDFOI	LGLLKDFFIH	YVRYHAKCIL	VAKDTPLLMN
Capparis m	125	135	145	155	165	175
VK3412	KMKYYFVNLW	QVHFSVWFQP	QKVHINELSK	DNLDFLGYLA	SLRLNPLLVR	TQMLNENFLI
	KMKYYFVNLW	QVHFSVWFQP	QKVHINELSK	DNLDFLGYLS	SLRLNPLLVR	TQMLNENFLI
Capparis m	185	195	205	215	225	235
VK3412	DINVRKLDTK	IPICSIIGSL	AKERFCHVLG	HPISKSTWMD	SSDSIDLDRF	VRISRNLSHY
	DINVRKLDTK	IPICSIIGSL	AKERFCHVLG	HPISKSTWMD	SSDSIDLDRF	VRISRNLSHY
Capparis m	245	255				
VK3412	HS6SSKKNL	YRIKY				
	HS6SSKKNL	YRIKY				

Hình 5. Kết quả so sánh trình tự amino acid suy diễn từ mẫu nghiên cứu với trình tự mang mã số AB925022.1 của loài *C. micracantha* được công bố trên ngân hàng gen bằng phần mềm Bioedit

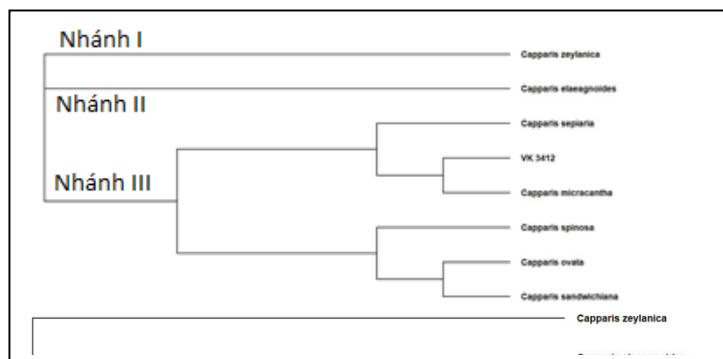
Như vậy, từ sự sai khác trình tự nucleotide và trình tự amino acid suy diễn giữa mẫu nghiên cứu và trình tự tương đồng nhiều nhất trên ngân hàng gen (AB925022.1) có thể nhận thấy, sự sai khác về mặt di truyền là rất ít.

### Xây dựng cây phát sinh chủng loại

Nhằm khẳng định một cách chính xác về mặt phân loại, chúng tôi đã tiến hành dựng cây phát sinh chủng loại giữa mẫu nghiên cứu với 7 trình tự trên ngân hàng gen (các trình tự với mã số được thể hiện trên Hình 3) bằng phần mềm MEGA5 dựa trên trình tự nucleotide thu được của đoạn gen *matK*. Kết quả được thể hiện trên Hình 6.

Dựa trên sự sai khác trong trình tự gen *matK*, cây phát sinh chủng loại được xây dựng với mẫu nghiên cứu và 7 trình tự trên GenBank. Hình ảnh trên sơ đồ phân loại cho thấy, cây phát sinh được chia làm 3 nhánh chính: Nhánh I là loài *Capparis zeylanica*, nhánh II là loài *Capparis elaeagnoides*. Sáu trình tự còn lại thuộc nhánh thứ III.

Nhánh thứ III của cây phát sinh chủng loại lại được chia thành 2 nhánh phụ: tại nhánh phụ thứ nhất cho thấy, mẫu nghiên cứu (VK 3412) thuộc cùng loài với *C. micracantha* (điều này phù hợp với kết quả phân tích tính đa dạng trong trình tự vùng gen *matK*) và nằm cùng nhánh gần hơn với *Capparis sepiaria*. Ba trình tự còn lại thuộc nhánh phụ thứ 2. Với kết quả so sánh trình tự và dựng cây chủng loại phát sinh ở trên, có thể tạm xác định và kết luận mẫu nghiên cứu của chúng tôi (VK 3412) thuộc loài *C. micracantha*.



Hình 6. Cây phát sinh chủng loại được xây dựng dựa trên trình tự đoạn gen *matK* của mẫu nghiên cứu bằng phần mềm MEGA5

### KẾT LUẬN

Từ các kết quả thu được, chúng tôi đưa ra một số kết luận chính như sau: Đã khuếch đại thành công đoạn gen *matK* của cây Cáp gai nhỏ *Capparis* sp. thu thập tại Khu bảo tồn thiên nhiên Tà Kóu, huyện Hàm Thuận Nam, tỉnh Bình Thuận. Kết quả giải trình tự cho thấy vùng gen *matK* có kích thước 800 bp. Từ kết quả so sánh và phân tích trình tự đoạn gen *matK* có thể xác định mẫu nghiên cứu (VK 3412) thuộc loài *Capparis micracantha*.

**Lời cảm ơn:** Nghiên cứu này được tài trợ bởi Quỹ Phát triển Khoa học và Công nghệ Quốc gia (NAFOSTED) trong đề tài mã số 106.03-2019.10.

## TÀI LIỆU THAM KHẢO

- Cuenoud PS, Chatrou LW (2002). Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcl*, *atpB*, and *matK* DNA sequences. *Amer J of Bot* 89: 132-144.
- Ford CS, Karen LA (2009). Selection of candidate DNA barcoding regions for use on land plants. *Bot J Linn Soc* 159(1): 11.
- Gao T, Sun Z, Yao H, Song J, Zhu Y, Ma X, Chen S (2011). Identification of fabaceae plants using the DNA barcode *matK*. *Plant Med* 77: 92-94.
- Hebert PDN, Cywinska A, Ball SL, Waard JR (2003). Biological identifications through DNA barcodes. *Proc R Soc Lond B Biol Sci* 270: 313-321.
- Hollingsworth PM, Little DP (2011). Choosing and using a Plant DNA Barcode. *PLoS* 6(5): e19254
- Kress WJ, Erickson DL (2008). DNA barcodes: Genes, genomics, and bioinformatics. *PNAS* 105(8): 2761-2762.
- Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005). Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci USA* 102: 8369-8374.
- Li M, Cao H, But PP, Shaw PC (2010). Identification of herbal medicinal materials using DNA barcodes. *J Syst Evol* 49(3): 271-283.
- Murray MG, Thompson WF (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acid Res* 8(19): 4321-5.
- Valentini A, Pompanon F, Taberlet P (2009). DNA barcoding for ecologists. *Trend Ecol Evol* 24(2): 110-7.
- Vijayan K, Tsou C (2010). DNA barcoding in plants: taxonomy in a new perspective. *Curr Sci (Bangalore)* 99: 1530-1541.
- Xiwen L, Yang Y, Robert JH, Maurizio R, Yitao W, Shilin C (2014). Plant DNA barcoding: from gene to genome. *Biol Rev* 923-931.
- Yang DY, FH, Cai SQ, Komatsu K (2004). Molecular analysis of Rheum species used as Rhei Rhizoma based on the chloroplast *matK* gene sequence and ITS application for identification. *Biol Pharma Bull* 27: 375-383.
- Web: [www.barcodeoflife.org](http://www.barcodeoflife.org); <http://blast.ncbi.nlm.nih.gov/Blast.cgi>.
- CBOL ABS Brochure, 2012, [http://www.imb.dvo.ru/misc/barcoding/files/CBOL&Fish-BOL\\_Projects/CBOL-ABS\\_DNA\\_Barcoding.pdf](http://www.imb.dvo.ru/misc/barcoding/files/CBOL&Fish-BOL_Projects/CBOL-ABS_DNA_Barcoding.pdf)
- CBOL Plant Working Group, 2009, A DNA Barcode for Land Plants. *Proc Natl Acad Sci USA* 106: 12794-12797. <http://dx.doi.org/10.1073/pnas.0905845106>

## USING DNA BARCODE TO IDENTIFY *CAPPARIS* SP. IN TA-KOU NATURE RESERVE, HAM THUAN NAM DISTRICT, BINH THUAN PROVINCE, VIETNAM

Nguyen Thi Thu Nga\*, Sy Danh Thuong

*Thai Nguyen university of education*

### SUMMARY

Biodiversity around the world is constantly declining, including many endangered species. A strategic project worldwide - DNA barcoding was announced recently to build a barcode library of endangered species (CBOL ABS Brochure, 2012). DNA barcoding using a short DNA sequence in the genome of the organism as the only string that distinguish species. *Capparis* sp. distributes in some provinces in Vietnam. In this study, in order to assist in the identification of species and initially published information using DNA barcode of *Capparis* sp. was collected at Ta-Kou Nature Reserve, Ham Thuan Nam district, Binh Thuan province is accurate, the method of using DNA barcode is based on the *matK* gene was performed. The *matK* gene from *Capparis* sp. was isolated based on *matK-F/matK-R* primers with a length of 800 bp. The *matK* gene from *Capparis* sp. was compared with 7 sequences of genus *Capparis* published on GenBank via BLAST software in NCBI, BioEdit, and MEGA5. The results have proved the research sample belongs to species *Capparis micracantha*.

**Keywords:** Barcode, *Capparis micracantha*, DNA barcode, *matK* gene, species identification.

\* Author for correspondence: Tel: +84-976714982; Email: [ngantt.bio@tnue.edu.vn](mailto:ngantt.bio@tnue.edu.vn)