

ĐẶC ĐIỂM TRÌNH TỰ ĐOẠN GEN *ndhF* CỦA LOÀI CÁP ĐẮK NÔNG (*Capparis daknongensis* D.T. Sy, G.C. Tucker, Cornejo & Joongku Lee)

Nguyễn Thị Thu Nga^{1*}, Trần Cẩm Tú¹, Sỹ Danh Thường¹

TÓM TẮT

Bài báo cung cấp một số thông tin về trình tự đoạn gen *ndhF* của loài *Capparis daknongensis* thu thập tại xã Đắk Nia, thị xã Gia Nghĩa, tỉnh Đắk Nông. Cáp Đắk Nông là một loài thực vật đặc hữu của Việt Nam, có khu phân bố hẹp, hiện mới ghi nhận ở tỉnh Đắk Nông và Đắk Lắk. Phân lập đoạn gen *ndhF* bằng kỹ thuật PCR đã thu được trình tự có chiều dài 1105 nucleotide. Nghiên cứu về trình tự đoạn gen *ndhF* của mẫu thí nghiệm chỉ ra rằng, loài Cáp Đắk Nông có quan hệ di truyền gần nhất với 2 loài *Capparis tomentosa* và *Capparis callophylla*. Các đặc điểm nghiên cứu trên là thông tin quan trọng giúp nhận diện và xây dựng mã vạch ADN barcoding cho loài thực vật đặc hữu này.

Từ khóa: *Capparaceae*, *Capparis*, *Capparis daknongensis*, *ndhF*, phân loại.

1. BẬT VẤN ĐỀ

Do có vị trí địa lý, địa hình và khí hậu thuận lợi nên hệ thực vật Việt Nam rất phong phú và đa dạng. Số lượng các loài thực vật bậc cao có mạch ở Việt Nam khoảng 12.000 loài, chiếm khoảng 5% tổng số loài thực vật bậc cao đã biết trên toàn thế giới và khoảng trên dưới 25% số loài thực vật bậc cao đã biết ở châu Á. Bên cạnh đó, hệ thực vật Việt Nam có mức độ đặc hữu cao, chiếm khoảng 3% số chi và 27,5% số loài [12].

Công nghệ mã vạch DNA là một công nghệ nhận dạng phân tử mới, sử dụng các đoạn DNA ngắn và chuẩn để nhận dạng loài [10]. Mã vạch DNA ban đầu được sử dụng để xác định vi sinh vật [1], nhưng giờ đây nó có thể xác định nhanh chóng và chính xác các đối tượng sinh vật ở cấp độ loài với lý do không giới hạn về giai đoạn phát triển, sự đa dạng hình thái bên trong, các yếu tố môi trường... [2]. Do đó, công nghệ mã vạch DNA đã nhanh chóng được ứng dụng trong xác định loài, hệ thống học, đa dạng sinh học, tiến hóa, bảo tồn loài, xác định mẫu khảo cổ và một số lĩnh vực có liên quan khác [7]. Gen cytochrome oxidase I (COI) của ty thể được đề xuất bởi Hebert và cộng sự (năm 2003) đã được sử dụng rộng rãi trong xác định loài động vật [8]. Tuy nhiên, do tỷ lệ đột biến của DNA ty thể thấp nên không thể sử dụng COI ở thực vật [10], [11].

Trong những thập kỷ qua, nhiều nhà nghiên cứu đã có những đóng góp to lớn trong việc tìm kiếm và ứng dụng mã vạch ở thực vật. Nhiều nỗ lực đã được thực hiện để tìm ra mã vạch cho các đơn vị phân loại thực vật khác nhau. Mặc dù nhiều nghiên cứu đã tìm kiếm một mã vạch thực vật phổ biến, đến nay không có locus nào ứng dụng được trên tất cả các loài [4], [5]. Tuy nhiên, Hiệp hội mã vạch sự sống (CBOL-The Consortium for the Barcode of Life) đã chứng minh rằng, việc sử dụng nhiều locus không đưa đến kết quả rõ ràng về khả năng phân biệt ở cấp độ loài [3]. Các nhà nghiên cứu gần đây đã đề xuất việc sử dụng trình tự toàn bộ hệ gen lục lạp trong việc xác định loài thực vật [10], [11], [14], [15]. Bộ gen lục lạp chứa tất cả các trình tự DNA trong một plastid, chứa nhiều thông tin di truyền để xác định loài hơn bất kỳ dấu hiệu locus đơn lẻ nào thường được sử dụng.

Đến ngày 27 tháng 10 năm 2019, bộ gen lục lạp của 3452 loài thực vật đã được công bố trên NCBI. Đây là một hướng phát triển quan trọng của mã vạch DNA nhằm sử dụng bộ gen lục lạp để xác định và phân biệt các loài thực vật [15]. Nhiều nghiên cứu đã sử dụng hệ gen lục lạp để nhận diện loài [6], [9], [11], [14]. Zhang và cộng sự (2013) đã phát hiện ra gen *ndhF* của lục lạp có thể xác định chính xác 100% các loài thuộc chi Cà (*Solanum*) [16].

Cáp Đắk Nông là một loài thực vật đặc hữu thuộc chi *Capparis*, họ Mần mần (Capparaceae). Loài có khu phân bố hẹp, hiện mới ghi nhận ở tỉnh Đắk Nông và Đắk Lắk. Từ trước đến nay, chưa có một công trình nào nghiên cứu về loài Cáp Đắk Nông ở trên thế giới và ở Việt Nam. Trong nghiên cứu này,

¹Khoa Sinh học, Trường Đại học Sư phạm - Đại học Thái Nguyên. Email: nganttt.bio@tnue.edu.vn

bước đầu sử dụng vùng DNA barcode *ndhF* để đánh giá mối quan hệ di truyền giữa loài Cáp Đắc Nông và một số loài thuộc chi Cáp đã được định danh. Từ đó làm cơ sở cho việc phân tích di truyền và nhận diện loài.

2. NGUYÊN VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Nguyên vật liệu

Mẫu lá của loài Cáp Đắc Nông (*Capparis daknongensis* D.T. Sy, G.C. Tucker, Cornejo & Joongku Lee) được thu thập từ xã Đắc Nía, thị xã Gia Nghĩa, tỉnh Đắk Nông được sử dụng làm vật liệu nghiên cứu. Ký hiệu mẫu là VK3675. Số lượng mẫu lá được thu thập là 01 mẫu cùng với mẫu tiêu bản của loài.

Đặc điểm sinh học điển hình của loài: Cây bụi trườn; cành có gai. Lá hình bầu dục, gân bên 6-7 cặp. Cụm hoa chùy ở nách lá hoặc ở đỉnh. Đài 4, gồm 2 vòng. Tràng 4, có màu trắng. Nhị 13-18. Cuống bầu dài 3,5-4 cm. Bầu hình bầu dục. Quả hình bầu dục, chín có màu vàng. Mùa hoa và quả từ tháng 3-7 hàng năm. Loài mọc rải rác ven rừng hoặc ven đường nơi có nhiều ánh sáng.

Các hóa chất sử dụng trong nghiên cứu được mua từ các hãng uy tín như Qiagen, Fermentas, Sigma...

Cặp mỗi nhân đoạn gen *ndhF* được tổng hợp theo Kress *et al.* (2005).

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

Tên mỗi	Trình tự mỗi	Nhiệt độ gắn mỗi (Tm)	Kích thước dự kiến
<i>ndhF</i>	TTCCACTTCCAGTTCCTATGTTAATAGG	54°C	Khoảng 1,1 kb
<i>ndhR</i>	GAAATAAAGCACTTCGATAAGACCC		

2.2. Thiết bị sử dụng trong nghiên cứu

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

2.3. Phương pháp nghiên cứu

- Tách chiết DNA tổng số

Mẫu được tách chiết DNA tổng số bằng quy trình mini-CTAB và kit DNeasy® Plant Mini Kit của hãng Qiagen. Quá trình tách chiết DNA được tiến hành với riêng từng mẫu.

- Kiểm tra sản phẩm DNA sau tách chiết

DNA tổng số của mẫu VK3675 sau khi tách chiết được điện di trên gel agarose 1%, trong đệm TBE 1X, ở hiệu điện thế 90V với marker 1kb.

- Phương pháp nhân gen bằng kỹ thuật PCR

Thành phần phản ứng PCR bao gồm: Master mix (2X) - 7,5 µl, mỗi xuôi (10pmol/µl) - 0,5 µl, mỗi ngược (10 pmol/µl) - 0,5µl, DNA khuôn (10ng/µl)- 0,5 µl, H₂O - 6 µl. Tổng thể tích phản ứng là 15µl.

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

Trình tự nucleotide của đoạn gen *ndhF* được gửi đi xác định trình tự với cặp mỗi đặc hiệu. Trình tự gen được phân tích, so sánh và xây dựng cây phát

sinh chủng loại bằng các chương trình Bioedit và MegaX.

3. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

3.1. Kết quả tách chiết và tinh sạch DNA tổng số

Việc lựa chọn một phương pháp tách chiết DNA 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 polysaccharides hay các sản phẩm có nguồn gốc từ thực vật.

Nhóm nghiên cứu đã tách chiết thành công DNA của mẫu lá loài *Capparis daknongensis*. Kết quả thí nghiệm được thể hiện trên hình 1.



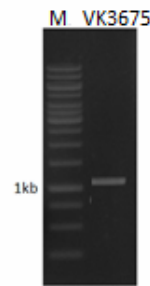
Hình 1. Kết quả điện di DNA tổng số của mẫu VK3675

Chất lượng DNA thu được sau tách chiết và tinh sạch ảnh hưởng rất lớn đến kết quả quá trình nghiên cứu. Do vậy sau khi tách chiết DNA, đã kiểm tra chất lượng DNA thu được bằng phương pháp điện di. Kết quả sau khi điện di trên gel agarose 0,8% (hình 1) cho thấy, DNA thu được sau khi tách có chất lượng tốt, DNA không bị đứt gãy, băng vạch sáng rõ. Như vậy các mẫu DNA thu được đảm bảo độ tinh sạch cho các nghiên cứu tiếp theo.

3.2. Sử dụng kỹ thuật PCR khuếch đại gen *ndhF*

Đã khuếch đại thành công vùng gen nghiên cứu với cặp mồi đặc hiệu *ndhF*; *ndhF*R, nhiệt độ bắt cặp mồi đặc hiệu ở 54°C. Hình ảnh điện di sản phẩm PCR được trình bày trên hình 2.

Kết quả khuếch đại đoạn gen *ndhF* cho thấy, trên hình ảnh điện di chỉ xuất hiện một băng sản phẩm với kích thước tính toán như dự đoán (khoảng 1,1kp).



Hình 2. Ảnh điện di sản phẩm PCR nhân đoạn gen *ndhF* của mẫu VK3675 trên gel agarose 1%, M: marker 1kp

3.3. Giải trình tự đoạn gen *ndhF* thu được từ phản ứng khuếch đại

Kết quả giải trình tự trực tiếp từ sản phẩm PCR và so sánh với một số trình tự công bố trên ngân hàng gen được trình bày trong hình 3 và hình 4.

Sequences producing significant alignments									
Download Select columns Show 100									
GenBank Graphics Distance tree of results New MSA Viewer									
☒ select all 100 sequences selected	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☒	Capparis versicolor chloroplast complete genome	Capparis versicolor	1991	1991	99%	0.0	99.19%	155051	MH142726.1
☒	Capparis spinosa var. ovata chloroplast complete genome	Capparis spinosa	1864	1864	99%	0.0	97.10%	157783	MK637690.1
☒	Capparis spinosa var. ovata chloroplast complete genome	Capparis spinosa	1864	1864	99%	0.0	97.10%	157783	NC_049616.1
☒	Capparis ovata voucher K Millenium Seed Bank 243816 NADH dehydrogenase subunit F (ndhF) gene partial	Capparis spinosa	1849	1849	98%	0.0	97.17%	1128	KU739555.1
☒	Capparis decidua chloroplast complete genome	Capparis decidua	1847	1847	99%	0.0	96.83%	157573	NC_053386.1
☒	Capparis spinosa complete genome	Capparis spinosa	1842	1842	99%	0.0	96.74%	157728	MT041701.1
☒	Capparis spinosa var. spinosa chloroplast complete genome	Capparis spinosa	1836	1836	99%	0.0	96.65%	157796	NC_047193.1
☒	Capparis spinosa var. spinosa chloroplast complete genome	Capparis spinosa	1836	1836	99%	0.0	96.65%	157796	MK639365.1
☒	Capparidastrium frondosum voucher WIS X Cornejo 7597 NADH dehydrogenase subunit F (ndhF) gene partial	Capparidastrium	1805	1805	98%	0.0	96.44%	1118	KU739554.1
☒	Crateva tapia chloroplast complete genome	Crateva tapia	1803	1803	99%	0.0	96.11%	156962	MK637696.1
☒	Crateva unilocularis chloroplast complete genome	Crateva unilocularis	1803	1803	99%	0.0	96.11%	156417	NC_050353.1
☒	Crateva unilocularis chloroplast complete genome	Crateva unilocularis	1803	1803	99%	0.0	96.11%	156417	MT679554.1
☒	Crateva tapia chloroplast complete genome	Crateva tapia	1803	1803	99%	0.0	96.11%	156962	NC_049621.1
☒	Crateva religiosa chloroplast ndhF gene for NADH dehydrogenase subunit F partial cds isolate YB035	Crateva religiosa	1770	1770	97%	0.0	96.12%	2049	LC050147.2

Hình 3. Kết quả so sánh trình tự mẫu VK3675 với một số trình tự được công bố trên Genbank

Capparis versicolor chloroplast, complete genome									
Sequence ID: MH142726.1 Length: 155051 Number of Matches: 1									
Range 1: 110613 to 111717 GenBank Graphics									
	Score	Expect	Identities	Gaps	Strand				
	1991 bits(1078)	0.0	1096/1105(99%)	0/1105(0%)	Plus/Minus				
VK3675	1		GGGTCTTATCGAGCCGCTTTATTTTCATTTCATCTACTCATGCTTATTCGAAAGCTTTGTG	60		VK3675	481	TTAACAATGAATAATAACAAAGAACTCTTTTTTGAAGAAAGCATATCAAGTTAGA	540
Sbjct	111717		GGGTCTTATCGAGCCGCTTTATTTTCATTTCATCTACTCATGCTTATTCGAAAGCTTTGTG	111658		Sbjct	111237	TTAACAATGAATAATAACAAAGAACTCTTTTTTGAAGAAAGCATATCAAGTTAGG	111178
VK3675	61		TTTTTAGGATCGGATCATTATTTTCATTCATGGAAGCTATAGTTGGATATCTCCCGAT	120		VK3675	541	AATAATGTAAGAAATAGAACTTTCTTACTGTGCTAATTTTCCCTTAATACAACT	600
Sbjct	111657		TTTTTAGGATCGGATCATTATTTTCATTCATGGAAGCTATAGTTGGATATCTCCCGAT	111598		Sbjct	111177	AATAATGTAAGAAATAGAACTTTCTTACTGTGCTAATTTTCCCTTAATACAACT	111118
VK3675	121		AAAAGCCAGATATGATTTTATGTTGGTTTTCGACAAACATGTCGATTACAAAAC	180		VK3675	601	TTCTCTTATCCCCAAGAAATCGGACAACTATCTTATCTATGCTGTTATGGTTTA	660
Sbjct	111597		AAAAGCCAGATATGATTTTATGTTGGTTTTCGACAAACATGTCGATTACAAAAC	111538		Sbjct	111117	TTCTCTTATCCCCAAGAAATCGGACAACTATCTTATCTATGCTGTTATGGTTTA	111058
VK3675	181		GCTTTTTTATGAGGACATTTCTCTTTGTTGTTATTTTTTCTGTTTGTGTTCT	240		VK3675	661	TTTACTTTGTTTATGGAGCTATAGGAATCTTTCAATCAAGATCAAGAAAGAAATAC	720
Sbjct	111537		GCTTTTTTATGAGGACATTTCTCTTTGTTGTTATTTTTTCTGTTTGTGTTCT	111478		Sbjct	111057	TTTACTTTGTTTATGGAGCTATAGGAATCTTTCAATCAAGATCAAGAAAGAAATAC	110998
VK3675	241		AAAGATGAAATCTTAAATGATAGTTCTTTTTCCTCAATTTTTCGAAATAGCTTGT	300		VK3675	721	TTTGATATATTATCGAGATTTTGACCCCATCTATAAATCTTTGTATAAAAATTCAAAT	780
Sbjct	111477		AAAGATGAAATCTTAAATGATAGTTCTTTTTCCTCAATTTTTCGAAATAGCTTGT	111418		Sbjct	110997	TTTGATATATTATCGAGATTTTGACCCCATCTATAAATCTTTGTATAAAAATTCAAAT	110938
VK3675	301		TCAACAGCGGATTAACCGCATTTATATGTTTTCGAAATTTTACTTACTTTTGAAGGA	360		VK3675	781	AATTTTGTGATTTGGTATGCAATTTTGAGAAATGCACTTTTTCGGTCAGTATAGCTTT	840
Sbjct	111417		TCAACAGCGGATTAACCGCATTTATATGTTTTCGAAATTTTACTTACTTTTGAAGGA	111358		Sbjct	110937	AATTTTGTGATTTGGTATGCAATTTTGAGAAATGCACTTTTTCGGTCAGTATAGCTTT	110878
VK3675	361		CATTAAACATTCATTTTCAAAATACAGTGGGAAAAAGTAGCTCGTCTATTCTCATC	420		VK3675	841	TTTGGAAATATTATAGCATACTCTTATATAAGCTTTTATCATCTTTATTAATTA	900
Sbjct	111357		CATTAAACATTCATTTTCTAAATACAGTGGGAAAAAGTAGCTCGTCTATTCTCATC	111298		Sbjct	110877	TTTGGAAATATTATAGCATACTCTTATATAAGCTTTTATCATCTTTATTAATTA	110818
VK3675	421		TCTCTATGGGTAAGAGAAAGAAAAAATAACAGAAATTTTGTAGTACCTTA	480		VK3675	901	AACCTACTTAATTCATTTCAAAATGGGGTTTGAAGAAATCTCTGGGAAAAAATAA	960
Sbjct	111297		TCTCTATGGGTAAGAGAAAGAAAAAATAACAGAAATTTTGTAGTACCTTA	111238		Sbjct	110817	AACCTACTTAATTCATTTCAAAATGGGGTTTGAAGAAATCTCTGGGAAAAAATAA	110758

VK3675 961 AATTTGATATATAATTGGTCATATAATCGTGGTTACATAGATACATTTTATAGAACATCT 1020
 Sbjct 110757 AATTTGATATATAATTGGTCATATAATCGTGGTTACATAGATACATTTTATAGAACATCT 110698
 VK3675 1021 TTAACCTGAAAGTATAAGAAGATTAGCAAAACAACTCATTTTTGATAGACAAATAATT 1080
 Sbjct 110697 TTAACCTGAAAGTATAAGAAGATTAGCAAAACAACTCATTTTTGATAGACAAATAATT 110638
 VK3675 1081 GATGGAATTACAAATGCAGTAGGTA 1105
 Sbjct 110637 GATGGAATTACAAATGCAGTAGGTA 110613

Hình 4. Kết quả giải trình tự gen *ndhF* mẫu VK3675 và so sánh với hệ gen lục lạp của trình tự mang mã số MH142726.1 (loài *Capparis versicolor*) trên Ngân hàng gen

Để xác định đoạn trình tự sản phẩm PCR có phải đoạn gen *ndhF* hay không, đã sử dụng phần mềm Blast để so sánh đoạn trình tự thu được với các trình tự đã được công bố trên Genbank. Kết quả phân tích blast cho thấy, trình tự nucleotide thu được chính là trình tự của gen *ndhF*. Từ kết quả thu được so sánh trình tự mẫu VK3675 với hệ gen lục lạp của trình tự mang mã số MH142726.1 (loài *Capparis versicolor*) trên ngân hàng gen. Hình 4 chỉ ra sự tương đồng lên

đến 99% của mẫu nghiên cứu với trình tự gen *ndhF* trong hệ gen lục lạp của loài *Capparis versicolor*. Kết quả giải trình tự cho thấy, gen *ndhF* phân lập được từ mẫu VK3675 có chiều dài 1105 nucleotide. Kết quả so sánh trên hình 4 cho thấy, sự đa hình trong trình tự nucleotid là khá cao, tại đây xảy ra các biến đổi như thay thế một vài nucleotide. Các nucleotide thay đổi nằm ở các vị trí: 47 (C-T), 336 (A-G), 380 (A-T), 800 (C-A), 873 (G-T), 933 (G-T), 945 (G-C), 1103 (A-G).

3.4. Phân tích sự đa hình di truyền trong trình tự gen *ndhF* của một số loài thuộc chi Cáp

Đã tiến hành so sánh trình tự DNA vùng gen *ndhF* của mẫu nghiên cứu VK3675 cùng với 9 trình tự DNA vùng gen *ndhF* của các loài thuộc chi Cáp đã được công bố trên Ngân hàng gen. Thông tin về các trình tự so sánh được mô tả trong bảng 2 (mã số gen, tên loài, năm công bố, kích thước gen).

Bảng 2. Thông tin các gen *ndhF* sử dụng trong so sánh với mẫu thí nghiệm (VK3675)

STT	Mã số gen	Tên loài	Năm công bố	Kích thước (bp)
1	Mẫu thí nghiệm (VK3675) - <i>Capparis daknongensis</i>			1105
2	KU739555.1	<i>Capparis ovata</i>	2016	1128
3	EU373685.1	<i>Capparis lucida</i>	2016	992
4	AY122364.1	<i>Capparis callophylla</i>	2016	967
5	EU373698.1	<i>Capparis tomentosa</i>	2016	994
6	EU373699.1	<i>Capparis zeylanica</i>	2016	992
7	EU373694.1	<i>Capparis spinosa</i>	2016	992
8	EU002208.1	<i>Capparis flexuosa</i> (chloroplast)	2009	2058
9	EU373690.1	<i>Capparis sandwichiana</i>	2016	992
10	EU373682.1	<i>Capparis cynophallophora</i>	2016	992

```

      5      15      25      35      45      55
C.callophy  ATTATTCATT CAATGGAAGC TATAGTTGGA TATTCTCCCG ATAAAAGCCA GAATATGATT
C.cynophal  ATTATTCATT CAATGGAAGC TATAGTTGGA TATTCTCCCG ATAAAAGCCA GAATATGATT
C.daknonge  ATTATTCATT CAATGGAAGC TATAGTTGGA TATTCTCCCG ATAAAAGCCA GAATATGATT
C.flexuosa  ATTATTCATT CAATGGAAGC TATAGTTGGA TATTCTCCCG ATAAAAGCCA GAATATGATT
C.lucida    ATTATTCATT CAATGGAAGC TATAGTTGGA TATTCTCCCG ATAAAAGCCA GAATATGATT
C.ovata     ATTATTCATT CAATGGAAGC TATAGTTGGA TATTCTCCCG ATAAAAGTCA GAATATGATT
C.sandwich  ATTATTCATT CAATGGAAGC TATAGTTGGA TATTCTCCCG ATAAAAGTCA GAATATGATT
C.spinosa   ATTATTCATT CAATGGAAGC TATAGTTGGA TATTCTCCCG ATAAAAGTCA GAATATGATT
C.tomentos  ATTATTCATT CAATGGAAGC TATAGTTGGA TATTCTCCCG ATAAAAGTCA GAATATGATT
C.zeylanic  ATTATTCATT CAATGGAAGC TATAGTTGGA TATTCTCCCG ATAAAAGTCA GAATATGATT

      65      75      85      95      105      115
C.callophy  TTTATGGGTG GTTTGACAAA ACATGTGCCG ATTACAAAAA CTGCCCTTTT TTTAGGGACA
C.cynophal  TTTATGGGTG GTTTGACAAA ACATGTGCCG ATTACAAAAA CTGCCCTTTT TTTAGGGACA
C.daknonge  TTTATGGGTG GTTTGACAAA ACATGTGCCG ATTACAAAAA CTGCCCTTTT TTTAGGGACA
C.flexuosa  TTTATGGGTG GTTTGACAAA ACATGTGCCG ATTACAAAAA CTGCCCTTTT TTTAGGGACA
C.lucida    TTTATGGGTG GTTTGACAAA ACATGTGCCG ATTACAAAAA CTGCCCTTTT TTTAGGGACA
C.ovata     TTTATGGGTG GTTTGACAAA ACATGTGCCG ATTACAAAAA CTGCCCTTTT TTTAGGGACA
C.sandwich  TTTATGGGTG GTTTGACAAA ACATGTGCCG ATTACAAAAA CTGCCCTTTT TTTAGGGACA
C.spinosa   TTTATGGGTG GTTTGACAAA ACATGTGCCG ATTACAAAAA CTGCCCTTTT TTTAGGGACA
C.tomentos  TTTATGGGTG GTTTGACAAA ACATGTGCCG ATTACAAAAA CTGCCCTTTT TTTAGGGACA
C.zeylanic  TTTATGGGTG GTTTGACAAA ACATGTGCCG ATTACAAAAA CTGCCCTTTT TTTAGGGACA

     125      135      145      155      165      175
C.callophy  CTTTCTCTTT GTGGTATTCC CCCCCTTGCT TGTTTTTGGT CTAAAGATGA AATTCTTAAT
C.cynophal  CTTTCTCTTT GTGGTATTCC CCCCCTTGCT TGTTTTTGGT CTAAAGATGA AATTCTTAAT
C.daknonge  CTTTCTCTTT GTGGTATTCC CCCCCTTGCT TGTTTTTGGT CTAAAGATGA AATTCTTAAT
C.flexuosa  CTTTCTCTTT GTGGTATTCC CCCCCTTGCT TGTTTTTGGT CTAAAGATGA AATTCTTAAT
C.lucida    CTTTCTCTTT GTGGTATTCC CCCCCTTGCT TGTTTTTGGT CTAAAGATGA AATTCTTAAT
C.ovata     CTTTCTCTTT GTGGTATTCC CCCCCTTGCT TGTTTTTGGT CTAAAGATGA AATTCTTAAT
C.sandwich  CTTTCTCTTT GTGGTATTCC CCCCCTTGCT TGTTTTTGGT CTAAAGATGA AATTCTTAAT
C.spinosa   CTTTCTCTTT GTGGTATTCC CCCCCTTGCT TGTTTTTGGT CTAAAGATGA AATTCTTAAT
C.tomentos  CTTTCTCTTT GTGGTATTCC CCCCCTTGCT TGTTTTTGGT CTAAAGATGA AATTCTTAAT
C.zeylanic  CTTTCTCTTT GTGGTATTCC CCCCCTTGCT TGTTTTTGGT CTAAAGATGA AATTCTTAAT
    
```

						
	185	195	205	215	225	235
C.callophy	GATAGTTTCT	TGTTTTCTCC	AATTTTTGCA	ATAATAGCTT	GTTCAACAGC	GGGATTAACT
C.cynophal	GATAGTTTCT	TGTTTTCTCC	AGTTTTTGCA	ATAATAGCTT	GTTCAACAGC	GGGATTAACT
C.daknonge	GATAGTTTCT	TGTTTTCTCC	AATTTTTGCA	ATAATAGCTT	GTTCAACAGC	GGGATTAACT
C.flexuosa	GATAGTTTCT	TGTTTTCTCC	AGTTTTTGCA	ATAATAGCTT	GTTCAACAGC	GGGATTAACT
C.lucida	GATAGTTTCT	TGTTTTCTCC	AATTTTTGCA	ATAATAGCTT	GTTCAACAGC	GGGATTAACT
C.ovata	GATAGTTTCT	TGTTTTCTCC	AATTTTTGCA	ATAATAGCTT	GTTCAACAGC	GGGATTAACT
C.sandwich	GATAGTTTCT	TGTTTTCTCC	AATTTTTGCA	ATAATAGCTT	GTTCAACAGC	GGGATTAACT
C.spinosa	GATAGTTTCT	TGTTTTCTCC	AATTTTTGCA	ATAATAGCTT	GTTCAACAGC	GGGATTAACT
C.tomentos	GATAGTTTCT	TGTTTTCTCC	AATTTTTGCA	ATAATAGCTT	GTTCAACAGC	GGGATTAACT
C.zeylanic	GATAGTTTCT	TGTTTTCTCC	TATTTTTGCA	ATAATAGCTT	GTTCAACAGC	GGGATTAACT
						
	245	255	265	275	285	295
C.callophy	GCATTTTATA	TGTTTCTGAT	TTATTTTACTT	ACTTTTGAAG	GACATTTTAA	CATTCAATTTT
C.cynophal	GCATTTTATA	TGTTTCTGAT	TTATTTTACTT	ACTTTTGAAG	GACATTTTAA	CATTCAATTTT
C.daknonge	GCATTTTATA	TGTTTCTGAT	TTATTTTACTT	ACTTTTGAAG	GACATTTTAA	CATTCAATTTT
C.flexuosa	GCATTTTATA	TGTTTCTGAT	TTATTTTACTT	ACTTTTGAAG	GACATTTTAA	CATTCAATTTT
C.lucida	GCATTTTATA	TGTTTCTGAT	TTATTTTACTT	ACTTTTGAAG	GACATTTTAA	CATTCAATTTT
C.ovata	GCATTTTATA	TGTTTCTGAT	TTATTTTACTT	ACTTTTGAAG	GACATTTTAA	CATTCAATTTT
C.sandwich	GCATTTTATA	TGTTTCTGAT	TTATTTTACTT	ACTTTTGAAG	GACATTTTAA	CATTCAATTTT
C.spinosa	GCATTTTATA	TGTTTCTGAT	TTATTTTACTT	ACTTTTGAAG	GACATTTTAA	CATTCAATTTT
C.tomentos	GCATTTTATA	TGTTTCTGAT	TTATTTTACTT	ACTTTTGAAG	GACATTTTAA	CATTCAATTTT
C.zeylanic	GCATTTTATA	TGTTTCTGAT	TTATTTTACTT	ACTTTTGAAG	GACATTTTAA	CATTCAATTTT
						
	305	315	325	335	345	355
C.callophy	CAAAATTACA	GTGGAAAAAA	AAGTAATCG	TTCTATTCCA	TCTCTCTATG	GGGTAAGAGAA
C.cynophal	CAAAATTACA	GTGGAAAAAA	AAGTAATCG	TTCTATTCCA	TCTCTCTATG	GGGTAAGAGAA
C.daknonge	CAAAATTACA	GTGGAAAAAA	AAGTAATCG	TTCTATTCCA	TCTCTCTATG	GGGTAAGAGAA
C.flexuosa	CAAAATTACA	GTGGAAAAAA	AAGTAATCG	TTCTATTCCA	TCTCTCTATG	GGGTAAGAGAA
C.lucida	CAAAATTACA	GTGGAAAAAA	AAGTAATCG	TTCTATTCCA	TCTCTCTATG	GGGTAAGAGAA
C.ovata	CAAAATTACA	GTGGAAAAAA	AAGTAATCG	TTCTATTCCA	TCTCTCTATG	GGGTAAGAGAA
C.sandwich	CAAAATTACA	GTGGAAAAAA	AAGTAATCG	TTCTATTCCA	TCTCTCTATG	GGGTAAGAGAA
C.spinosa	CAAAATTACA	GTGGAAAAAA	AAGTAATCG	TTCTATTCCA	TCTCTCTATG	GGGTAAGAGAA
C.tomentos	CAAAATTACA	GTGGAAAAAA	AAGTAATCG	TTCTATTCCA	TCTCTCTATG	GGGTAAGAGAA
C.zeylanic	CAAAATTACA	GTGGAAAAAA	AAGTAATCG	TTCTATTCCA	TCTCTCTATG	GGGTAAGAGAA
						
	365	375	385	395	405	415
C.callophy	GAAGAAAAAA	AAATTTAACAG	AAATTTTTTGT	TTAGTACCCT	TATTAACAAT	GAATAATAAC
C.cynophal	GAAGAAAAAA	AAATTTAACAG	AAATTTTTTGT	TTAGTACCCT	TATTAACAAT	GAATAATAAC
C.daknonge	GAAGAAAAAA	AAATTTAACAG	AAATTTTTTGT	TTAGTACCCT	TATTAACAAT	GAATAATAAC
C.flexuosa	GAAGAAAAAA	AAATTTAACAG	AAATTTTTTGT	TTAGTACCCT	TATTAACAAT	GAATAATAAC
C.lucida	GAAGAAAAAA	AAATTTAACAG	AAATTTTTTGT	TTAGTACCCT	TATTAACAAT	GAATAATAAC
C.ovata	GAAGAAAAAA	AAATTTAACAG	AAATTTTTTGT	TTAGTACCCT	TATTAACAAT	GAATAATAAC
C.sandwich	GAAGAAAAAA	AAATTTAACAG	AAATTTTTTGT	TTAGTACCCT	TATTAACAAT	GAATAATAAC
C.spinosa	GAAGAAAAAA	AAATTTAACAG	AAATTTTTTGT	TTAGTACCCT	TATTAACAAT	GAATAATAAC
C.tomentos	GAAGAAAAAA	AAATTTAACAG	AAATTTTTTGT	TTAGTACCCT	TATTAACAAT	GAATAATAAC
C.zeylanic	GAAGAAAAAA	AAATTTAACAG	AAATTTTTTGT	TTAGTACCCT	TATTAACAAT	GAATAATAAC
						
	425	435	445	455	465	475
C.callophy	AAAAGAACTT	CTTTTTTTTG	CAAGAAAGCA	TATCAAGTTA	CTATAATGT	AAGAAATAGA
C.cynophal	AAAAGAACTT	CTTTTTTTTG	CAAGAAAGCA	TATCAAGTTA	CTATAATGT	AAGAAATAGA
C.daknonge	AAAAGAACTT	CTTTTTTTTG	CAAGAAAGCA	TATCAAGTTA	CTATAATGT	AAGAAATAGA
C.flexuosa	AAAAGAACTT	CTTTTTTTTG	CAAGAAAGCA	TATCAAGTTA	CTATAATGT	AAGAAATAGA
C.lucida	AAAAGAACTT	CTTTTTTTTG	CAAGAAAGCA	TATCAAGTTA	CTATAATGT	AAGAAATAGA
C.ovata	AAAAGAACTT	CTTTTTTTTG	CAAGAAAGCA	TATCAAGTTA	CTATAATGT	AAGAAATAGA
C.sandwich	AAAAGAACTT	CTTTTTTTTG	CAAGAAAGCA	TATCAAGTTA	CTATAATGT	AAGAAATAGA
C.spinosa	AAAAGAACTT	CTTTTTTTTG	CAAGAAAGCA	TATCAAGTTA	CTATAATGT	AAGAAATAGA
C.tomentos	AAAAGAACTT	CTTTTTTTTG	CAAGAAAGCA	TATCAAGTTA	CTATAATGT	AAGAAATAGA
C.zeylanic	AAAAGAACTT	CTTTTTTTTG	CAAGAAAGCA	TATCAAGTTA	CTATAATGT	AAGAAATAGA
						
	485	495	505	515	525	535
C.callophy	ACTTTTCTTA	CTGTTGCTAA	TTTTGACCTT	AAT-----A	CAACAACCTT	CTCTATCCCC
C.cynophal	ACTTTTCTTA	CTGTTGCTAA	TTTTGACCTT	AAT-----A	CAACAACCTT	CTCTATCCCC
C.daknonge	ACTTTTCTTA	CTGTTGCTAA	TTTTGACCTT	AAT-----A	CAACAACCTT	CTCTATCCCC
C.flexuosa	ACTTTTCTTA	CTGTTGCTAA	TTTTGACCTT	AAT-----A	CAACAACCTT	CTCTATCCCC
C.lucida	ACTTTTCTTA	CTGTTGCTAA	TTTTGACCTT	AAT-----A	CAACAACCTT	CTCTATCCCC
C.ovata	ACTTTTCTTA	CTGTTGCTAA	TTTTGACCTT	AAT-----A	CAACAACCTT	CTCTATCCCC
C.sandwich	ACTTTTCTTA	CTGTTGCTAA	TTTTGACCTT	AAT-----A	CAACAACCTT	CTCTATCCCC
C.spinosa	ACTTTTCTTA	CTGTTGCTAA	TTTTGACCTT	AAT-----A	CAACAACCTT	CTCTATCCCC
C.tomentos	ACTTTTCTTA	CTGTTGCTAA	TTTTGACCTT	AAT-----A	CAACAACCTT	CTCTATCCCC
C.zeylanic	ACTTTTCTTA	CTGTTGCTAA	TTTTGACCTT	AAT-----A	CAACAACCTT	CTCTATCCCC
						
	545	555	565	575	585	595
C.callophy	CAAGAAATCGG	ACAATACTAT	TCTATTTTCT	ATGCTTGAT	TGGTTTTATT	TACTTTGTTT
C.cynophal	CATGAATCGG	ACAATACTAT	TCTATTTTCT	ATGCTTGAT	TGGTTTTATT	TACTTTGTTT
C.daknonge	CAAGAAATCGG	ACAATACTAT	TCTATTTTCT	ATGCTTGAT	TGGTTTTATT	TACTTTGTTT
C.flexuosa	CATGAATCGG	ACAATACTAT	TCTATTTTCT	ATGCTTGAT	TGGTTTTATT	TACTTTGTTT
C.lucida	CATGAATCGG	ACAATACTAT	TCTATTTTCT	ATGCTTGAT	TGGTTTTATT	TACTTTGTTT
C.ovata	CATGAATCGG	ACAATACTAT	TCTATTTTCT	ATGCTTGAT	TGGTTTTATT	TACTTTGTTT
C.sandwich	CATGAATCGG	ACAATACTAT	TCTATTTTCT	ATGCTTGAT	TGGTTTTATT	TACTTTGTTT
C.spinosa	CATGAATCGG	ACAATACTAT	TCTATTTTCT	ATGCTTGAT	TGGTTTTATT	TACTTTGTTT
C.tomentos	CATGAATCGG	ACAATACTAT	TCTATTTTCT	ATGCTTGAT	TGGTTTTATT	TACTTTGTTT
C.zeylanic	CATGAATCGG	ACAATACTAT	TCTATTTTCT	ATGCTTGAT	TGGTTTTATT	TACTTTGTTT
						
	605	615	625	635	645	655
C.callophy	ATTGGAGCTA	TAGGAATTCC	TTTCAATCAA	GATCAAGAAG	GAATATACTT	TGATATATTA
C.cynophal	ATTGGAGCTA	TAGGAATTCC	TTTCAATCAA	GATCAAGAAG	GAATATACTT	TGATATATTA
C.daknonge	ATTGGAGCTA	TAGGAATTCC	TTTCAATCAA	GATCAAGAAG	GAATATACTT	TGATATATTA
C.flexuosa	ATTGGAGCTA	TAGGAATTCC	TTTCAATCAA	GATCAAGAAG	GAATATACTT	TGATATATTA
C.lucida	ATTGGAGCTA	TAGGAATTCC	TTTCAATCAA	GATCAAGAAG	GAATATACTT	TGATATATTA
C.ovata	ATTGGAGCTA	TAGGAATTCC	TTTCAATCAA	GATCAAGAAG	GAATATACTT	TGATATATTA
C.sandwich	ATTGGAGCTA	TAGGAATTCC	TTTCAATCAA	GATCAAGAAG	GAATATACTT	TGATATATTA
C.spinosa	ATTGGAGCTA	TAGGAATTCC	TTTCAATCAA	GATCAAGAAG	GAATATACTT	TGATATATTA
C.tomentos	ATTGGAGCTA	TAGGAATTCC	TTTCAATCAA	GATCAAGAAG	GAATATACTT	TGATATATTA
C.zeylanic	ATTGGAGCTA	TAGGAATTCC	TTTCAATCAA	GATCAAGAAG	GAATATACTT	TGATATATTA
						
	665	675	685	695	705	715
C.callophy	TCGAGATTTT	TGACCCCATC	TATAAATCTT	TTGTATAAAA	ATTCAAATAA	TTTTGTGGAT
C.cynophal	TCGAGATTTT	TGACCCCATC	TATAAATCTT	TTGTATAAAA	ATTCAAATAA	TTTTGTGGAT
C.daknonge	TCGAGATTTT	TGACCCCATC	TATAAATCTT	TTGTATAAAA	ATTCAAATAA	TTTTGTGGAT
C.flexuosa	TCGAGATTTT	TGACCCCATC	TATAAATCTT	TTGTATAAAA	ATTCAAATAA	TTTTGTGGAT
C.lucida	TCGAGATTTT	TGACCCCATC	TATAAATCTT	TTGTATAAAA	ATTCAAATAA	TTTTGTGGAT
C.ovata	TCGAGATTTT	TGACCCCATC	TATAAATCTT	TTGTATAAAA	ATTCAAATAA	TTTTGTGGAT
C.sandwich	TCGAGATTTT	TGACCCCATC	TATAAATCTT	TTGTATAAAA	ATTCAAATAA	TTTTGTGGAT
C.spinosa	TCGAGATTTT	TGACCCCATC	TATAAATCTT	TTGTATAAAA	ATTCAAATAA	TTTTGTGGAT
C.tomentos	TCGAGATTTT	TGACCCCATC	TATAAATCTT	TTGTATAAAA	ATTCAAATAA	TTTTGTGGAT
C.zeylanic	TCGAGATTTT	TGACCCCATC	TATAAATCTT	TTGTATAAAA	ATTCAAATAA	TTTTGTGGAT

	725	735	745	755	765	775
C.callophy	TGGTATCAAT	TTTTGAGAAA	TGCAACTTTT	T-----CGG	TCAGTATAGC	TTTTTTTGGG
C.cynophal	TGGTATCAAT	TTTTGAGAAA	TGCAACTTTT	T-----CAG	TCAGTATAGC	TTTTTTTGGG
C.daknonge	TGGTATCAAT	TTTTGAGAAA	TGCAACTTTT	T-----CGG	TCAGTATAGC	TTTTTTTGGG
C.flexuosa	TGGTATCAAT	TTTTGAGAAA	TGCAACTTTT	TATTTTTCAG	TCAGTATAGC	TTTTTTTGGG
C.lucida	TGGTATCAAT	TTTTGAGAAA	TGCAACTTTT	T-----CGG	TCAGTATAGC	TTTTTTTGGG
C.ovata	TGGTATCAAT	TTTTGAGAAA	TGCAACTTTT	T-----CAG	TCAGTATAGC	TTTTTTTGGG
C.sandwich	TGGTATCAAT	TTTTGAGAAA	TGCAACTTTT	T-----CAG	TCAGTATAGC	TTTTTTTGGG
C.spinosa	TGGTATCAAT	TTTTGAGAAA	TGCAACTTTT	T-----CAG	TCAGTATAGC	TTTTTTTGGG
C.tomentos	TGGTATCAAT	TTTTGAGAAA	TGCAACTTTT	T-----CGG	TCAGTATAGC	TTTTTTTGGG
C.zeilanic	TGGTATCAAT	TTTTGAGAAA	TGCAACTTTT	T-----CAG	TCAGTATAGC	TTTTTCTGGG
	785	795	805	815	825	835
C.callophy	ATATTTATAG	CATACTCTTT	ATATAAGCCT	TTTTATTTCAT	CTTTATTAAA	TTTAAACTTA
C.cynophal	ATATTTATAG	CATACTCTTT	ATATAAGCCT	TTTTATTTCAT	CTTTATTAAA	TTTAAACTTA
C.daknonge	ATATTTATAG	CATACTCTTT	ATATAAGCCT	TTTTATTTCAT	CTTTATTAAA	TTTAAACTTA
C.flexuosa	ATATTTATAG	CATACTCTTT	ATATAAGCCT	TTTTATTTCAT	CTTTATTAAA	TTTAAACTTA
C.lucida	ATATTTATAG	CATACTCTTT	ATATAAGCCT	TTTTATTTCAT	CTTTATTAAA	TTTAAACTTA
C.ovata	ATATTTATAG	CATACTCTTT	ATATAAGCCT	TTTTATTTCAT	CTTTATTAAA	TTTAAACTTA
C.sandwich	ATATTTATAG	CATACTCTTT	ATATAAGCCT	TTTTATTTCAT	CTTTATTAAA	TTTAAACTTA
C.spinosa	ATATTTATAG	CATACTCTTT	ATATAAGCCT	TTTTATTTCAT	CTTTATTAAA	TTTAAACTTA
C.tomentos	ATATTTATAG	CATACTCTTT	ATATAAGCCT	TTTTATTTCAT	CTTTATTAAA	TTTAAACTTA
C.zeilanic	ATATTTATAG	CATACTCTTT	ATATAAGCCT	TTTTATTTCAT	CTTTATTAAA	TTTAAACTTA
	845	855	865	875	885	895
C.callophy	CTTAATT---	-----CATT	TCAAAAATGG	GGTTT---AAA	GAATCCTGTG	GGAAAAACTA
C.cynophal	CTTAATT---	-----CATT	TCAAAAATGG	GGGTT---AAA	GAATTCCTGTG	GGAAAAACTA
C.daknonge	CTTAATT---	-----CATT	TCAAAAATGG	GGTTT---AAA	GAATTCCTGTG	GGAAAAACTA
C.flexuosa	CTTAATT---	-----CATT	TCAAAAATGG	GGTTT---AAA	GAATTCCTGTG	GGAAAAACTA
C.lucida	CTTAATT---	-----CATT	TCAAAAATGG	GGTTT---AAA	GAATTCCTGTG	GGAAAAACTA
C.ovata	CTTAATT---	-----CATT	TCAAAAATGG	GGTTT---AAA	GAATTCCTGTG	GGAAAAACTA
C.sandwich	CTTAATT---	-----CATT	TCAAAAATGG	GGTTT---AAA	GAATTCCTGTG	GGAAAAACTA
C.spinosa	CTTAATT---	-----CATT	TCAAAAATGG	GGTTT---AAA	GAATTCCTGTG	GGAAAAACTA
C.tomentos	CTTAATT---	-----CATT	TCAAAAATGG	GGTTT---AAA	GAATTCCTGTG	GGAAAAACTA
C.zeilanic	CTTAATT---	-----CATT	TCAAAAATGG	GGTTT---AAA	GAATTCCTGTG	GGAAAAACTA
	905	915	925	935	945	955
C.callophy	ATAAATTTGA	TATATAATTG	GTCATATAAT	CGTGGTTACA	TAGA---GTTT	TTATAGAACA
C.cynophal	ATAAATTTGA	TATATAATTG	GTCATATAAT	CGTGGTTACA	TAGA---GTTT	TTATAGAACA
C.daknonge	ATAAATTTGA	TATATAATTG	GTCATATAAT	CGTGGTTACA	TAGA---GTTT	TTATAGAACA
C.flexuosa	ATAAATTTGA	TATATAATTG	GTCATATAAT	CGTGGTTACA	TAGA---GTTT	TTATAGAACA
C.lucida	ATAAATTTGA	TATATAATTG	GTCATATAAT	CGTGGTTACA	TAGA---GTTT	TTATAGAACA
C.ovata	ATAAATTTGA	TATATAATTG	GTCATATAAT	CGTGGTTACA	TAGA---GTTT	TTATAGAACA
C.sandwich	ATAAATTTGA	TATATAATTG	GTCATATAAT	CGTGGTTACA	TAGA---GTTT	TTATAGAACA
C.spinosa	ATAAATTTGA	TATATAATTG	GTCATATAAT	CGTGGTTACA	TAGA---GTTT	TTATAGAACA
C.tomentos	ATAAATTTGA	TATATAATTG	GTCATATAAT	CGTGGTTACA	TAGA---GTTT	TTATAGAACA
C.zeilanic	ATAAATTTGA	TATATAATTG	GTCATATAAT	CGTGGTTACA	TAGA---GTTT	TTATAGAACA
	965	975	985	995	1005	1015
C.callophy	TCCTTAACTG	AAAGATAAG	AAGATTAG--			
C.cynophal	TCCTTAACTG	AAAGATAAG	AAGATTAGCA	AAACAAACTC	ATTTTTTTGA	TAG-----
C.daknonge	TCCTTAACTG	AAAGATAAG	AAGATTAGCA	AAACAAACTC	ATTTTTTTGA	TAGACAAATA
C.flexuosa	TCCTTAACTG	AAAGATAAG	AAGATTAGCA	AAACAAACTT	CTTTTTTTGA	T-----
C.lucida	TCCTTAACTG	AAAGATAAG	AAGATTAGCA	AAACAAACTC	ATTTTTTTGA	TAG-----
C.ovata	TCCTTAACTG	AAAGATAAG	AAGATTAGCA	AAACAAACTC	ATTTTTTTGA	TAGACGAATA
C.sandwich	TCCTTAACTG	AAAGATAAG	AAGATTAGCA	AAACAAACTC	ATTTTTTTGA	TAG-----
C.spinosa	TCCTTAACTG	AAAGATAAG	AAGATTAGCA	AAACAAACTC	ATTTTTTTGA	TAA-----
C.tomentos	TCCTTAACTG	AAAGATAAG	AAGATTAGCA	AAACAAACTC	ATTTTTTTGA	TAG-----
C.zeilanic	TCCTTAACTG	AAAGATAAG	AAGATTAGCA	AAACAAACTC	ATTTTTTTGA	TAG-----

Hình 5. Sự đa dạng trong trình tự gen *ndhF* của loài *Capparis daknongensis* với 9 trình tự trên ngân hàng gen (phân tích bằng phần mềm BioEdit)

So sánh trình tự DNA gen *ndhF* của mẫu nghiên cứu (*Capparis daknongensis*) cùng với 9 trình tự DNA vùng gen *ndhF* của các loài thuộc chi Cáp đã được công bố trên GenBank ở bảng 2 nhận thấy, chiều dài vùng gen *ndhF* giữa các loài dao động từ 992 đến 1128 bp.

Đã tiến hành phân tích sự đa dạng di truyền trong trình tự gen *ndhF* của loài *Capparis daknongensis* với 9 trình tự trên Ngân hàng gen bằng phần mềm BioEdit. Kết quả phân tích được thể hiện trên hình 5.

Tiến hành phân tích và so sánh sự đa dạng di truyền trong trình tự nucleotide của gen *ndhF* từ vị trí nu thứ nhất đến vị trí nu 990 (do kích thước của các trình tự có độ dài ngắn khác nhau). Kết quả phân tích được thể hiện trong bảng 2.

Như vậy sự đa dạng về nucleotide trong đoạn trình tự so sánh có thể quan sát thấy ở 8 vị trí. Sự thay đổi này có thể liên quan đến những biến đổi trong quá trình tiến hoá của loài dưới ảnh hưởng của môi trường sống.

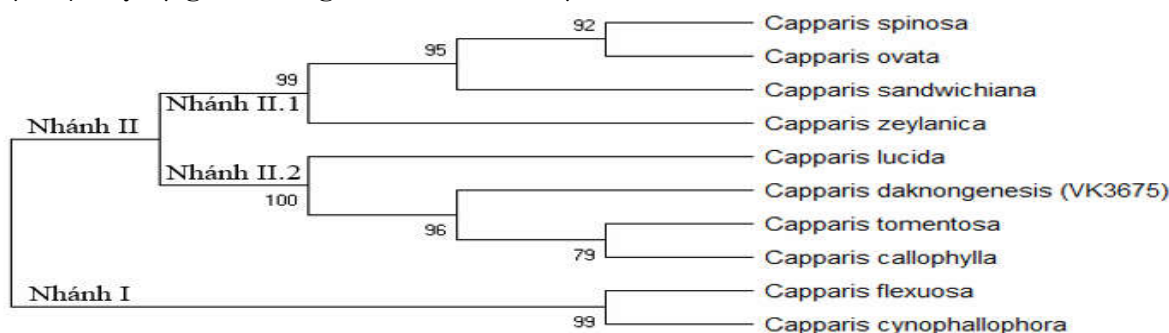
Bảng 3. Sự đa dạng di truyền trong trình tự nucleotide của gen *ndhF* khi so sánh mẫu nghiên cứu (loài *Capparis daknongensis*) với 9 trình tự trên ngân hàng gen

STT	Vị trí nucleotide	Nucleotide trên trình tự loài <i>Capparis daknongensis</i>	Nucleotide thay đổi trên một số trình tự so sánh
1	240	C	T
2	258	A	G
3	327	C	T
4	462	A	T,G
5	728	C	A
6	876	G	T
7	946	A	G
8	974	G	A

Dựa trên sự đa dạng trong trình tự nucleotide gen *ndhF* của 10 trình tự phân tích, cây phát sinh chủng loại được xây dựng với mẫu nghiên cứu và 9 trình tự

trên Genbank nhằm xác định mối quan hệ về mặt di truyền (hình 6). Hình ảnh trên sơ đồ phân loại cho thấy, cây phát sinh được chia làm 2 nhánh chính: Nhánh I gồm 2 loài *Capparis flexuosa* và *Capparis cynophallophora*, 8 trình tự còn lại là các loài thuộc nhánh II (trong đó có mẫu nghiên cứu VK3675 – loài *Capparis daknongensis*).

Cây chủng loại phát sinh cho thấy loài *Capparis daknongensis* nằm cùng nhánh phụ II.2 gần hơn với 2 loài *Capparis tomentosa* và *Capparis callophylla*, và nằm xa hơn một chút với nhánh mang loài *Capparis lucida*, điều này cũng được thể hiện rõ trong bảng hệ số sai khác di truyền giữa các loài so sánh (bảng 3). Với kết quả so sánh về đặc điểm trình tự nucleotide gen *ndhF* và dựng cây phân loại có thể khẳng định, trong phạm vi nghiên cứu, loài *Capparis daknongensis* có mối quan hệ họ hàng về mặt di truyền gần với 2 loài *Capparis tomentosa* và *Capparis callophylla*.



Hình 6. Sơ đồ phân loại hình cây được xây dựng dựa trên trình tự gen *ndhF* của mẫu VK3675 - *Capparis daknongensis* bằng phần mềm Mega X

Bảng 3. Hệ số sai khác di truyền giữa các trình tự so sánh

	1	2	3	4	5	6	7	8	9	10
1. <i>Capparis zeylanica</i>		0.00568	0.00354	0.00392	0.00373	0.00537	0.00691	0.00557	0.00621	0.00601
2. <i>Capparis tomentosa</i>	0.02912		0.00550	0.00570	0.00553	0.00383	0.00711	0.00279	0.00546	0.00263
3. <i>Capparis spinosa</i>	0.01530	0.02905		0.00332	0.00152	0.00517	0.00689	0.00535	0.00625	0.00588
4. <i>Capparis sandwichiana</i>	0.01635	0.03121	0.01017		0.00290	0.00528	0.00699	0.00567	0.00613	0.00612
5. <i>Capparis ovata</i>	0.01530	0.02905	0.00202	0.00812		0.00521	0.00689	0.00542	0.00617	0.00587
6. <i>Capparis lucida</i>	0.02677	0.01440	0.02671	0.02779	0.02671		0.00677	0.00367	0.00532	0.00378
7. <i>Capparis flexuosa</i>	0.04517	0.04657	0.04622	0.04735	0.04619	0.04402		0.00740	0.00486	0.00737
8. <i>Capparis daknongensis</i> (VK3675)	0.02896	0.00818	0.02889	0.03103	0.02921	0.01432	0.04843		0.00568	0.00318
9. <i>Capparis cynophallophora</i>	0.03316	0.03129	0.03629	0.03418	0.03523	0.02891	0.02585	0.03111		0.00584
10. <i>Capparis callophylla</i>	0.03075	0.00623	0.03068	0.03288	0.03068	0.01362	0.04740	0.01042	0.03297	

4. KẾT LUẬN

Đã khuếch đại thành công vùng gen *ndhF* của mẫu VK3675 (*Capparis daknongensis*) thu thập tại xã Đắc Nĩa, thị xã Gia Nghĩa, tỉnh Đắk Nông. Kết quả giải trình tự cho thấy, vùng gen *ndhF* của mẫu nghiên cứu có kích thước 1105 nucleotide. Sự phân tích và so sánh trình tự gen đã chỉ ra sự đa dạng di truyền trong quá trình tiến hoá của vùng gen *ndhF*. Thông tin về vùng gen *ndhF* của loài *Capparis daknongensis* được sử dụng làm mã vạch DNA barcoding giúp nhận dạng loài. Bên cạnh đó còn là cơ sở khoa học cho công tác thu thập, bảo tồn nguồn gen và nhân giống loài thực vật đặc hữu này ở Việt Nam.

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

1. Abel S., Wiesch P. A., Chang H. H., Davis B. M., Lipsitch M., Waldor M. K. (2015), "Sequence tag-based analysis of microbial population dynamics", *Nature Methods*, (12), pp.223-226.
2. Bhargava M., Sharma A. (2013), "DNA barcoding in plants: evolution and applications of in silico approaches and resources", *Molecular Phylogenetics and Evolution*, (67), pp.631-641.
3. CBOL Plant Working Group1, Hollingsworth P.M., Laura L. F., John L. S., Mehrdad H., Sujevan R., Michelle B., Mark W. (2009), "A DNA barcode for land plants", *PNAS*, 106(31), pp.12794-12797.
4. Chen S.L., Yao H., Han J.P., Liu C., Song J.Y., Shi L.C et al (2010), "Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species", *PloS One*, Article e8613.
5. Cuenoud P.S.V., Chatrou L.W. (2002), "Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *ndhF* DNA sequences", *American Journal of Botany*, (89), pp.132-144.
6. David L. E., John S., Alissa R., Lee A. W., Kress W. J. (2008), "DNA barcoding in land plants: developing standards to quantify and maximize success", *Taxon*, 57(4), pp.1304-1316.
7. Hebert P. D. N., Cywinska A., Ball S.L., Waard J. R. (2003), "Biological identifications through DNA barcodes", *Proceedings of the Royal Society Biological Sciences*, (270), pp. 313-321.
8. Hilu K. W., L. H., (1997), "The *ndhF* gene: sequence variation and application in plant systematics", *American Journal of Botany*, (84), pp. 830-839.
9. Hollingsworth P.M., Little D.P. (2011), "Choosing and using a Plant DNA Barcode", *PloS One*, 6(5), e19254.
10. Li M., Cao H., But P. P. H., Shaw P. C. (2011), "Identification of herbal medicinal materials using DNA barcodes", *Journal of Systematics and Evolution*, 49(3), pp.271-283.
11. Sucher N. J., Carles M.C (2008), "Genome-based approaches to the authentication of medicinal plants", *Planta Medica*, 74(6), pp.603-623.
12. Lã Đình Mõi, Trần Minh Hợi, Dương Đức Huyền, Trần Huy Thái, Ninh Khắc Bản (2005), *Tài nguyên thực vật Việt Nam* (Những cây chứa các hợp chất có hoạt tính sinh học), tập 1, Nhà xuất bản Nông nghiệp, Hà Nội.
13. Trần Minh Hợi (chủ biên), Lã Đình Mõi, Trần Huy Thái, Ninh Khắc Bản (2013), *Tài nguyên thực vật Việt Nam* (Giáo trình giảng dạy dùng cho học viên cao học và nghiên cứu sinh), Nhà xuất bản Khoa học Tự nhiên và Công nghệ, Hà Nội.
14. Yang D. Y., Cai S. Q., Komatsu K. (2004), "Molecular analysis of Rheum species used as Rhei Rhizoma based on the chloroplast *ndhF* gene sequence and ITS application for identification", *Biological and Pharmaceutical Bulletin*, 27(3), pp.375383.
15. Yang Y., Dang Y. Y., Li Q., Lu J. J., Li X. W., Wang Y. T. (2014), "Complete chloroplast genome sequence of poisonous and medicinal plant *Datura stramonium*: organizations and implications for genetic engineering", *PLoS One*, 9(11), Article e110656.
16. Zhang W., Fan X. H., Zhu S. F. (2013), "Species-specific identification from incomplete sampling: Applying DNA barcodes to monitoring invasive *Solanum* plants", *PLoS ONE*, 8(2), e55927.
17. Web: <https://www.ncbi.nlm.nih.gov/genome/browse#!/organelles/>.

CHARACTERISTICS OF NDHF GENE SEQUENCE OF *Capparis daknongensis* D.T. Sy, G.C. Tucker, Cornejo & Joongku Lee

Nguyen Thi Thu Nga^{1*}, Tran Cam Tu¹, Sy Danh Thuong¹

¹*University of Education - Thai Nguyen University*

This paper provides some information on the *ndhF* gene sequences of *Capparis daknongensis* collected in Dak Nia commune, Gia Nghia town, Dak Nong province. *Capparis daknongensis* is a plant endemic to Vietnam, with a narrow distribution, only recently recorded in Dak Nong and Dak Lak provinces. Isolation of the *ndhF* gene fragment by PCR technique obtained a sequence with a length of 1105 nucleotides. The study on the *ndhF* gene sequence of the experimental sample showed that the *Capparis daknongensis* has the closest genetic relationship to the *Capparis tomentosa* and *Capparis callophylla*. The above research characteristics are important information to help identify and build DNA barcoding for this endemic species.

Keywords: *Capparaceae*, *Capparis*, *Capparis daknongensis*, *ndhF*, *identification*.

Người phản biện: PGS.TS. Khuất Hữu Trung

Ngày nhận bài: 13/6/2021

Ngày thông qua phản biện: 8/7/2021

Ngày duyệt đăng: 8/7/2021