

XÁC ĐỊNH MỘT SỐ TRÌNH TỰ DNA MÃ VẠCH PHỤC VỤ CÔNG TÁC PHÂN LOẠI VÀ NHẬN DẠNG LOÀI LAN KIM TUYẾN (*Anoetochilus formosanus* Hayata) Ở THANH HÓA

Nguyễn Trọng Quyền¹, Lê Công Mạnh², Nguyễn Thị Thơ²,
Khuất Thị Hải Ninh², Hoàng Văn Sâm², Bùi Văn Thắng²

TÓM TẮT

Lan Kim tuyến (*Anoetochilus formosanus* Hayata) có giá trị dược liệu cao, loài này đã bị khai thác cạn kiệt, có thể dẫn đến nguy cơ bị tuyệt chủng ngoài tự nhiên. Loài *A. formosanus* Hayata có hình thái tương đồng với các loài thuộc chi Lan Kim tuyến (*Anoetochilus* Blume) nên việc phân loại và định danh theo phương pháp truyền thống gặp khó khăn. Vì vậy, nghiên cứu xác định một số chỉ thị DNA mã vạch loài Lan Kim tuyến thu thập tại Thanh Hóa nhằm định danh chính xác loài này là rất cần thiết cho công tác bảo tồn và phát triển nguồn gen. Kết quả nghiên cứu đã xác định được ba trình tự DNA mã vạch là *matK*, *rbcL* và *ITS2* của loài Lan Kim tuyến *A. formosanus* Hayata, các trình tự có độ tương đồng cao với các trình tự gen đã công bố trên Ngân hàng Gen quốc tế. Các trình tự DNA mã vạch thu được có từ 3 – 8 vị trí nucleotide sai khác với các trình tự tương ứng của các loài thuộc chi *Anoetochilus* có mã số công bố trong Ngân hàng Gen quốc tế. Trong ba đoạn trình tự DNA mã vạch thu được, mã vạch gen *matK* và *ITS2* là ứng viên hữu hiệu cho việc phân loại và định danh loài Lan Kim tuyến *A. formosanus* Hayata tại Thanh Hóa. Kết quả nghiên cứu này là cơ sở khoa học phục vụ cho công tác bảo tồn, khai thác và phát triển nguồn gen loài cây quý hiếm này.

Từ khóa: DNA mã vạch, định danh, Lan Kim tuyến, trình tự gen.

1. ĐẶT VẤN ĐỀ

Các loài thực vật thuộc chi Lan Kim tuyến (*Anoetochilus* Blume) thường mọc sâu trong các tầng rừng ẩm với khoảng 51 loài phân bố trên một khu vực khá rộng, từ vùng Himalaya đến Đông Nam Á, miền Nam Trung Quốc, Úc, Papua New Guinea và một số hải đảo thuộc quần đảo Thái Bình Dương; Việt Nam có 17 loài như: *A. formosanus* Hayata, *A. calcareus* Aver, *A. annamensis* Aver, *A. setaceus* Blume, v.v. (Sách Đỏ Việt Nam, 2007; Nguyễn Trọng Quyền và cộng sự, 2020). Các loài thuộc chi Lan Kim tuyến hình thái đẹp có giá trị làm cảnh, đặc biệt giá trị dược liệu quý đối với sức khỏe con người (Võ Văn Chi, 2003; Lin, 2007; Chuan Gao, 2009; Phạm Hoàng Hộ, 2010). Chúng có nhiều hoạt chất có giá trị như thân cây và lá chứa kinsenoside có khả năng bảo vệ gan, chứa rất nhiều hợp chất glycoside và flavonoids khác có khả năng hỗ trợ điều trị các bệnh như tiểu đường, cao huyết áp, viêm thận và ức chế tế bào ung thư. v.v. (Du *et al.*, 2003; Fang *et al.*, 2008). Trong chi

Lan Kim tuyến, loài *A. formosanus* Hayata có giá trị dược liệu cao nên loài này đã bị khai thác cạn kiệt ngoài tự nhiên, dẫn đến nguy cơ bị tuyệt chủng (Nghị định số 06/2019/NĐ-CP, 2019; Sách Đỏ Việt Nam, 2007). Vì vậy, việc nghiên cứu định danh rõ nguồn gốc và bảo tồn các loài Lan Kim tuyến nói chung và loài Lan Kim tuyến (*A. formosanus* Hayata) nói riêng là rất cần thiết.

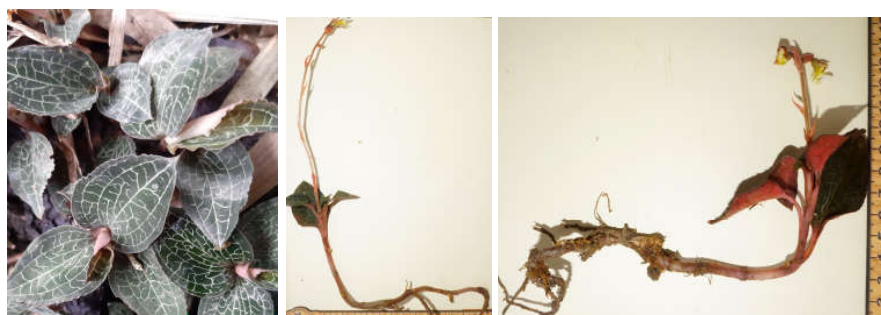
Việc phân loại và định danh loài bằng phương pháp truyền thống thường dựa trên các đặc điểm hình thái rễ, thân, lá và đặc biệt là hoa. Tuy nhiên, các loài Lan Kim tuyến rất đa dạng và phong phú đồng thời các loài gần nhau thường có hình thái tương tự nhau nên rất khó phân biệt. Hiện nay, để định danh chính xác các loài, bên cạnh việc sử dụng các đặc điểm hình thái thì việc sử dụng các chỉ thị phân tử là rất cần thiết. Trong các loại chỉ thị phân tử thì chỉ thị DNA mã vạch (DNA Barcode) đang được sử dụng phổ biến nhất để phân loại và định danh loài (Group *et al.*, 2009; Hebert, 2003) đã đề xuất DNA mã vạch như là một phương pháp để định danh loài. Định danh loài dựa vào các đoạn DNA mã vạch đặc trưng đảm bảo độ chính xác cao; đặc biệt hữu dụng và khắc phục được hạn chế của phương pháp phân

¹ Viện Nông nghiệp Thanh Hóa

² Trường Đại học Lâm nghiệp

loại chỉ dựa vào hình thái. Huỳnh Hữu Đức và cộng sự (2019), Lò Thị Mai Thu và cộng sự (2019) đã tiến hành nghiên cứu nhận dạng và định danh một số loài Lan Kim tuyến thuộc chi *Anoectochilus* Blume ở Việt Nam bằng chỉ thị DNA mã vạch. Trong nghiên cứu này, ba đoạn trình tự DNA mã vạch là *matK*, *rbcL*, *ITS2* của loài Lan Kim tuyến thu thập tại Thanh Hóa đã phân lập, xác định và phân tích trình tự DNA làm cơ sở dữ liệu phân tử phục vụ công tác phân loại, nhận dạng và bảo tồn nguồn gen.

2. VẬT LIỆU, PHƯƠNG PHÁP NGHIÊN CỨU



Hình 1. Mẫu Lan Kim tuyến thu tại Khu Bảo tồn các loài hạt trần quý hiếm Nam Động, Thanh Hóa

2.2. Phương pháp nghiên cứu

2.2.1. Phương pháp tách chiết DNA tổng số

DNA tổng số từ 3 mẫu lá Lan Kim tuyến (3 cây) được tách chiết riêng từng mẫu bằng bộ kit tách chiết DNA thực vật (Plant DNA Isolation Kit) của Hãng Norgen Biotek; các bước tách chiết theo hướng dẫn của nhà sản xuất. DNA tổng số sau tách chiết

2.1. Vật liệu nghiên cứu

Vật liệu nghiên cứu là các mẫu lá của loài Lan Kim tuyến (*A. formosanus* Hayata) thu thập tại Khu Bảo tồn các loài hạt trần quý hiếm Nam Động, Thanh Hóa. Thu 3 mẫu lá của 3 cây đại diện có điểm hình thái giống nhau (Hình 1). Đặc điểm hình thái của loài Lan Kim tuyến (*A. formosanus* Hayata) nghiên cứu được mô tả bởi Nguyễn Trọng Quyền và cộng sự (2020).

được xác định nồng độ và độ tinh sạch bằng phương pháp quang phổ kế trên máy Nanodrop 2000 và pha loãng về nồng độ 100 ng/μl để sử dụng cho phản ứng PCR.

2.2.2. Phương pháp khuếch đại PCR và giải trình tự nucleotide

Bảng 1. Trình tự các cặp mồi và kích thước vùng gen đích theo lý thuyết

Gen đích	Kí hiệu mồi	Trình tự mồi (5' – 3')	Tm (°C)	Kích thước đoạn DNA (bp) dự kiến	Tham khảo
<i>matK</i>	3F_KIM_f	CGTACAGTACTTTTGTGTTTACGAG	52	800	CBOL Plant Working Group, 2009
	1R_KIM_r	ACCCAGTCCATCTGGAAATCTTGGTTC			
<i>rbcL</i>	rbcLa-F	ATGTCACCACAAACAGAGACTAAAGC	52	600	Levin <i>et al.</i> , 2003; Kress & Erickson, 2007
	rbcLa-R	GTAAATCAAGTCCACCRGC			
ITS2	<i>ITS2PF</i>	ACGAATTCATGTCCGGTGAAGTGTTCG	55	400	Hà Văn Huân và cộng sự, 2020
	<i>ITS2PR</i>	TAGAATTCCCCGGTTCGCTCGCCGTTAC			

Nhân bản đoạn gen *matK*, *rbcL*, *ITS2* từ DNA tổng số bằng phản ứng PCR trên máy PCR 9700 Thermal Cycler Applied Biosystems (Mỹ); mỗi phản ứng PCR được thực hiện trong tổng phản ứng 25 μl,

bao gồm: H₂O deion (8,5 μl), 2x PCR Master mix Solution của Hãng Norgen Biotek (12,5 μl), 10 pmol/μl mồi xuôi (1,0 μl), 10 pmol/μl mồi ngược (1,0 μl), DNA tổng số (2 μl tương ứng 200 ng). Chu trình nhiệt phản ứng PCR: biến tính 94°C 5 phút, tiếp theo

35 chu kỳ [Biến tính 95°C - 30 giây, gắn mỗi 52°C (đối với gen *maK*, *rbcL*) và 55°C (đối với gen *ITS2*) - 50 giây, tổng hợp 72°C - 1 phút], hoàn tất kéo dài chuỗi 72°C trong 10 phút và giữ mẫu ở 4°C. Sản phẩm PCR được điện di kiểm tra trên gel agarose 1% có bổ sung thuốc Redsafe (Hãng Intron). Sau khi điện di, bản gel được soi dưới đèn UV và chụp ảnh bằng hệ thống Dolphin - Doc Image system của Hãng Wealtec. Những sản phẩm PCR sau khi được khuếch đại thành công sẽ được tinh sạch bằng bộ kit (PCR purification Kit của Hãng Norgen Biotek), các bước thực hiện theo hướng dẫn của nhà sản xuất.

Sau khi tinh sạch, sản phẩm PCR nhân gen *maK*, *rbcL*, *ITS2* được gửi cho Công ty 1st Base ở Malaysia để giải trình tự. Trình tự nucleotide của các đoạn DNA được xác định bằng máy giải trình tự tự động, sử dụng bộ kit BigDye[®] Terminator v3.1 Cycle Sequencing, theo nguyên lý Sanger.

2.2.3. Phương pháp phân tích dữ liệu DNA mã vạch

Trình tự nucleotide của các đoạn gen *maK*, *rbcL*, *ITS2* được xử lý và phân tích bằng phần mềm tin sinh học BioEdit 6.0; Mega 7 và công cụ BLAST trên website NCBI.

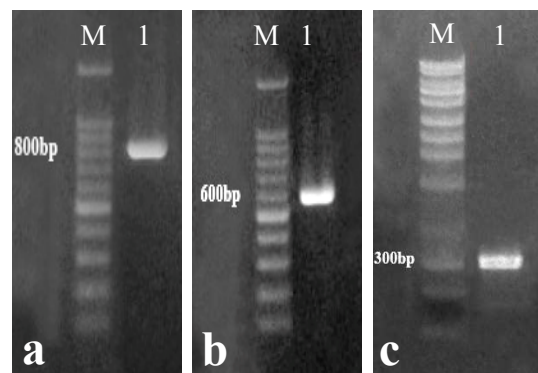
3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Nhân bản đoạn trình tự gen *maK*, *rbcL* và *ITS2*

Gen *maK* là một trong những gen mã hóa nằm trong lục lạp, có thể nhân bản một cách dễ dàng ở nhiều loài thực vật, đây là trình tự gen được sử dụng rộng rãi để định danh các loài thực vật (Vijayan và Tsou, 2010). Một số đoạn trình tự DNA mã vạch *maK* của các loài thuộc chi Lan Kim tuyến (*Anoectochilus* Blume) đã được công bố trên Ngân hàng Gen quốc tế (NCBI). Trên cơ sở các công trình đã công bố, trong nghiên cứu này đã sử dụng cặp mồi, mỗi xuôi 3F_KIM_f: 5'-CGT ACA GTA CTT TTG TGT TTA CGA G-3' và mỗi ngược: 5'-ACC CAG TCC ATC TGG AAA TCT TGG TTC-3' để nhân bản đoạn gen *maK* từ DNA tổng số của loài Lan Kim tuyến. Kết quả kiểm tra sản phẩm PCR nhân bản đoạn gen *maK* trên bản gel agarose 1% cho thấy đoạn gen *maK* đã được nhân bản đặc hiệu, băng DNA có kích thước khoảng 800 bp như tính toán lý thuyết khi so với thang DNA chuẩn 100 bp (Hình 2a); có thể khẳng định đã nhân bản thành công đoạn gen *maK* ở loài Lan Kim tuyến nghiên cứu.

Tương tự gen *maK*, gen *rbcL* cũng là một gen nằm trong lục lạp và có thể nhân bản một cách dễ dàng ở nhiều loài thực vật. Sử dụng cặp mồi, mỗi xuôi *rbcLa-F* 5'- ATG TCA CCA CAA ACA GAG ACT AAA GC-3' và mỗi ngược *rbcLa-R* GTA AAA TCA AGT CCA CCR CG-3' đã nhân bản thành công đoạn gen *rbcL* từ mẫu DNA tổng số của loài Lan Kim tuyến. Băng DNA nhân bản đặc hiệu có kích thước khoảng 600 bp khi so sánh với thang DNA chuẩn 100 bp (Hình 2b); kết quả này có thể khẳng định đã nhân bản thành công đoạn gen *rbcL* ở loài Lan Kim tuyến nghiên cứu.

Gen *ITS2* là gen nằm trong nhân tế bào và cũng có thể nhân bản một cách dễ dàng ở nhiều loài động, thực vật. Đoạn gen *ITS2* cũng được sử dụng rộng rãi và hiệu quả trong việc định danh, phân loại các loài động, thực vật (Vijayan và Tsou, 2010; Yao *et al.*, 2010; Yong *et al.*, 2010). Nhân bản đoạn gen *ITS2* từ mẫu DNA tổng số của loài Lan Kim tuyến bằng phản ứng PCR với cặp mồi, mỗi xuôi *ITS2PF*: 5'- ACG AAT TCA TGT CCG GTG AAG TGT TCG -3' và mỗi ngược *ITS2PR*: 5'- TAG AAT TCC CCG GTT CGC TCG CCG TTAC -3', kết quả điện di kiểm tra sản phẩm PCR cho thấy đoạn gen *ITS2* đã được nhân bản đặc hiệu, băng DNA có kích thước khoảng 350 bp như tính toán lý thuyết khi so với thang DNA chuẩn 100 bp (Hình 2c); có thể khẳng định đã nhân bản thành công đoạn gen *ITS2* ở loài Lan Kim tuyến nghiên cứu. Các sản phẩm nhân bản đoạn gen *maK*, *rbcL* và *ITS2* được tinh sạch và giải trình tự nucleotide.



Hình 2. a - Sản phẩm PCR nhân đoạn gen *maK*; b - Sản phẩm PCR nhân đoạn gen *rbcL*; c - Sản phẩm PCR nhân đoạn gen *ITS2* từ DNA tổng số của loài Lan Kim tuyến nghiên cứu trên gel agarose 1,0%. M – thang DNA chuẩn 100 bp

3.2. Phân tích trình tự các đoạn DNA mã vạch

3.2.1. Trình tự nucleotide đoạn gen *matK*

Kết quả xác định trình tự nucleotide đoạn gen *matK* của loài Lan Kim tuyến, cho thấy đoạn gen được nhân bản có kích thước 790 bp.

Sau khi giải trình tự đoạn gen *matK*, trình tự được xử lý bằng phần mềm BioEdit và công cụ

BLAST trên Ngân hàng Gen quốc tế (NCBI), thu được đoạn gen *matK* có kích thước 790 bp với chất lượng tốt. Sử dụng đoạn trình tự gen *matK* để phân tích và so sánh với một số trình tự gen khác trên Ngân hàng Gen quốc tế. Kết quả đoạn gen *matK* của loài Lan Kim tuyến thu thập tại Thanh Hóa đã được đăng trên GenBank (NCBI) với mã số MW678627, như sau:

AAGATGTTCCCTTCTTTGCATTTGTTGCGATTTATTTTCCACGAATATCATAATTTGAAGAGTATCATT
ACTTCAAATAAATCCATTACGTTTTTTCAAAAAAAGAAAAAGATTTTTTGGTTCCTACATAATTTTAA
TGTATATGAATGCGAATATCTCTTTCTTTTCTTCGTAAAAATCTTCTTATTTACGATCAACATCTTTTG
GAGTCTTTATTGAGCGAACACTTTTTATGTAAAAATGGAATCTATTCTAGTAGTATATTTTAAATCTTTT
CAGAGGATTCTCTGGTTCCTCAAAGATCCTTTCATACATTATGTTGATATCAAGGAAAAAGTAATCTGA
CTTCAAAGGTAACCTTATTCTGATGAAGAAATGGAATTTTCATGTTGTGAATTTTTGTCAATTTTATTTT
CACTTTTGGTCTCAACCTTATAGGATCCATATAAAGCAATTACCCAACTATTCCTTCTCTTTTCTGGGGT
ATTTTTTAAGTGTACAAAAAATCTTTGGTAGTAAGAAATCAAATGCTAGAGAATTCCTTTCTAATAAAT
ACTATGACTAAGAAATTAGATACCGTAGCCCCAGTTATTTCTCTTATTGGATCATTGTCGAAAGCTCAAT
TTTGTACTATATCAGGTCATCCTATTAGTAAACCCATTTGGACTGATTTTTCGGATTCTGATATTATTGAT
CGATTTTGTGCGAAATGTAGAAATCTTTGTCGTTATCACAGCGGATCCTCAAAAAAAGTTTTGTATC
GTATAAAATATATATT

Trình tự nucleotide của đoạn gen *matK* thu được này được so sánh với 7 trình tự gen *matK* của các loài khác cùng thuộc chi *Anoectochilus* và 1 loài thuộc chi *Ludisia* - đây là chi có đặc điểm hình thái khá giống với các loài chi *Anoectochilus* nhờ công cụ BLAST trên NCBI. Kết quả cho thấy, trình tự gen *matK* của loài Lan Kim tuyến nghiên cứu có mức độ tương đồng 99,87% cao nhất với trình tự gen *matK* của loài *A. formosanus* có mã số EU797513.1 trong Ngân hàng Gen quốc tế

và tương đồng với các loài thuộc chi *Anoectochilus* từ 98,72 - 99,75% (tương ứng với loài *A. lylei* mã số JN166019.1 là 98,72%; loài *A. elatus* mã số KU687098.1 là 98,95%; loài *A. albolineatus* mã số JN166018.1 là 98,98%; loài *A. koshunensis* mã số EU797512.1 là 99,62% và loài *A. roxburghii* mã số KY966708.1 là 99,75%) và loài *Ludisia discolor* mã số AJ543947.1 là 96,83%.

Bảng 2. Các vị trí nucleotide sai khác trong trình tự gen *matK* của loài Lan Kim tuyến nghiên cứu với 7 mã gen công bố trong Ngân hàng Gen quốc tế

Loài	Mã số trong NCBI	Vị trí nucleotide sai khác																	
		114	182	237	248	258	266	297	338	361	410	439	588	647	668	677	683	706	712
<i>A. formosanus</i> Hayata nghiên cứu	MW678627	A	T	T	G	C	A	T	A	T	T	C	G	A	C	T	T	G	G
<i>A. formosanus</i>	EU797513.1	A	T	T	G	C	A	T	A	T	T	T	G	A	C	T	T	G	G
<i>A. roxburghii</i>	KY966708.1	A	T	T	G	C	A	T	A	T	T	C	A	A	C	T	T	A	G
<i>A. koshunensis</i>	EU797512.1	A	T	T	G	C	A	T	-	T	T	C	A	A	C	T	T	A	G
<i>A. albolineatus</i>	JN166018.1	T	T	T	T	C	A	T	A	G	T	C	A	G	C	T	T	G	T
<i>A. lylei</i>	JN166019.1	T	T	T	G	C	A	T	A	G	T	C	A	G	C	T	T	T	G
<i>A. elatus</i>	KU687098.1	T	T	T	T	C	A	T	A	G	T	C	A	G	C	T	T	G	T
<i>L. discolor</i>	AJ543911.1	T	G	C	G	T	G	C	A	G	G	C	A	A	A	C	A	G	G

So sánh đoạn trình tự gen *matK* nghiên cứu thu được có kích thước 790 bp với các mã hiệu gen *matK* của các loài trên Ngân hàng Gen quốc tế, nhận thấy từ vị trí nucleotide thứ 8 đến vị trí 763 có mức độ tương đồng cao với cả 7 mã gen còn lại. Vì vậy, sẽ sử dụng đoạn trình tự này để so sánh và phân tích. Cụ thể trình tự gen *matK* của loài Lan Kim tuyến nghiên cứu có 18 vị trí nucleotide sai khác so với 7 mã gen so sánh (Bảng 2). Trong số 18 điểm khác biệt trên, có 8 điểm sai khác giữa mẫu Lan Kim tuyến nghiên cứu với các loài cùng chi *Anoectochilus* (vị trí: 114, 248, 338, 361, 439, 588, 647, 706 và 712). Kết quả cho thấy, đoạn gen *matK* nhân bản được có mức độ đa hình khá cao, không những có thể sử dụng làm mã vạch để phân biệt giữa các loài thuộc chi *Anoectochilus* và chi *Ludisia* mà còn có thể sử dụng để phân biệt giữa các loài thuộc chi *Anoectochilus*. Mặt khác, loài Lan Kim tuyến nghiên cứu thu thập tại Thanh Hóa có sự khác biệt nhỏ với trình tự gen *matK*

AGCTGGTGTAAAGATTACAAGTTGACTTATTATACTCCTGACTACGAAACCAAAAGTACTGATATC
TTGGCAGCATTCCGAGTAACCTCAACCGGGAGTTCCGCCTGAAGAAGCGGGCGCTGCGGTAGCAGC
CGAATCTTCTACTGGTACATGGACAGCTGTGTGGACTGATGGACTTACCAGTCTTGATCGTTACAAAGG
ACGATGCTACCACATTGAGCCCGTTGTTGGGGAGGAAATCAATATATTGCTTATGTAGCTTATCCTTTA
GACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACTTCCATTGTGGGTAATGTTTTTGGTTTCAAAG
CCCTGCGAGCTCTACGTCTGGAAGATCTGCGAATTCCTTCTTATTCCTTCAAACTTTCCAAGCCCCGCC
TCATGGCATCCAAGTTGAAAGAGATAAATTGAACAAGTATGGTCGTCCCTATTGGGATGTACTATTAA
CCAAAATTGGGATTATCCGCAAAAACTACGGTAGAGCG

Bảng 3. Các vị trí nucleotide sai khác trong trình tự gen *rbcl* của loài Lan Kim tuyến nghiên cứu với 5 mã gen công bố trong Ngân hàng Gen quốc tế

Loài	Mã số trong NCBI	Vị trí nucleotide sai khác		
		85	161	383
<i>A. formosanus</i> Hayata nghiên cứu	MW678628	A	G	A
<i>A. emeiensis</i>	LC057212.1	A	G	A
<i>A. elatus</i>	KU687104.1	A	G	C
<i>A. geniculatus</i>	JN166036.1	A	G	C
<i>A. albolineatus</i>	JN166035.1	M	R	C
<i>L. discolor</i>	MH749095.1	A	A	C

So sánh trình tự nucleotide của đoạn gen *rbcl* thu được với 4 trình tự gen *rbcl* đã được công bố trên Ngân hàng Gen quốc tế của 4 loài cùng chi *Anoectochilus* và loài *Ludisia discolor*. Kết quả cho thấy, mức độ tương đồng trình tự nucleotide của đoạn gen *rbcl* rất cao từ 99,43 - 100% (tương ứng với loài *A. emeiensis* mã số LC057212.1 là 100%; loài *A. elatus* mã số KU687104.1 là 99,81%; loài *A. geniculatus* mã số JN166036.1 là 99,81%; loài *A.*

của loài *A. formosanus* Đài Loan mã số EU797513.1, ở vị trí nucleotide 439. Điều này cho thấy, ngay cùng loài, ở các xuất xứ khác nhau cũng có thể có những khác biệt nhỏ trong cấu trúc gen *matK*. Như vậy, từ kết quả phân tích và so sánh cho thấy đoạn trình tự nucleotide gen *matK* của loài Lan Kim tuyến nghiên cứu rất đặc trưng nên có thể được sử dụng làm DNA mã vạch đối với loài Lan Kim tuyến thu thập tại Thanh Hóa.

3.2.2. Trình tự nucleotide đoạn gen *rbcl*

Sau khi giải trình tự đoạn gen *rbcl*, trình tự được phân tích và so sánh với dữ liệu trên Ngân hàng Gen quốc tế, kết quả thu được đoạn gen *rbcl* nghiên cứu có kích thước 523 bp có chất lượng tốt. Đoạn trình tự gen *rbcl* của loài Lan Kim tuyến thu thập tại Thanh Hóa đã được đăng trên GenBank (NCBI) với mã số MW678628, như sau:

albolineatus mã số JN166035.1 là 99,43% và với loài *Ludisia discolor* mã số MH749095.1 là 99,43%).

Phân tích trình tự nucleotid đoạn gen *rbcl* của loài Lan Kim tuyến nghiên cứu cho thấy không có nhiều điểm sai khác với các loài đã công bố. Cụ thể, đối với loài *A. emeiensis* (LC057212.1) không tìm thấy sự khác biệt nào trong 523 nucleotide được so sánh. Ba loài còn lại cùng chi *Anoectochilus* cũng có mức độ tương đồng rất cao dao động từ 99,43 -

99,81%. Trong đó, cả ba loài này khác mẫu Lan Kim tuyến nghiên cứu ở vị trí nucleotide thứ 383: ở mẫu Lan Kim tuyến nghiên cứu là A, còn ba loài còn lại là C; riêng loài *A. albolineatus* có hai vị trí thứ 85 và 161 chưa xác định được loại nucleotide chính xác trong khi mẫu Lan Kim tuyến nghiên cứu là nucleotide loại A (Bảng 3). Như vậy, với ít điểm sai khác như trên có

thể khẳng định đoạn gen *rbcL* nhân bản được ít có ý nghĩa phân định loài thuộc chi *Anoectochilus*.

3.2.3. Trình tự nucleotide đoạn gen ITS2

Đoạn gen *ITS2* được nhân bản từ mẫu Lan Kim tuyến, giải trình tự có kích thước 335 bp và đã được đăng trên GenBank (NCBI) với mã số MW663932, như sau:

AATCTTTGACGCAAGTTGCGCCTGAGGCCAATTGGCTAAGGGCACGTCCGCCTGGGCGTCAAGCAT
TACATCGCTTCATTGACACCAATTGCCAGTATTGTGCTGTGGTGTGCTGGTCTGAATGCGGAGAGTGGC
CCTTCGTGCACACTTGTGCGACGGGTTGAAGAACAATTTGCTTTCTCTGGCCATGTTTTGATAAAGGG
GTGGTGTATGCAGCCATTTGGCCACACTATCATCTCATTGCCTTGAGGAGGATAAATGTACACATTCGT
GGCTGATCACCCGATATAAATGTGCGAGGTGACGCCCTGAAATGCGACCCCAGGATGGGCG

Trình tự này được xử lý, so sánh với 4 trình tự gen tương ứng đã được công bố trên Ngân hàng Gen quốc tế NCBI của 4 loài thuộc chi *Anoectochilus*. Kết quả cho thấy, mức độ tương đồng trình tự nucleotide của đoạn gen *ITS2* rất cao từ 98,78 - 99,69% (tương ứng với loài *A. formosanus* mã số GQ396668.1 và *A.*

lylei mã số JN166060.1 là 99,69%; loài *A. koshunensis* mã số EU700340.1 là 99,39%; loài *A. albolineatus* mã số JN166058.1 là 98,78%). Kết quả so sánh trình tự nucleotide gen *ITS2* nghiên cứu với 4 mã gen công bố trong Ngân hàng Gen quốc tế bằng phần mềm BioEdit 6.0, như sau:

LKT	AATCTTTGA-ACGCAAGTTGCGCCTGAGGCCAATTGGCTAAGGGCACGTCCGCCTGGGCGT
JN166058.1	AATCTTTGAACGCAAGTTGCGCCTGAGGCCAATTGGCTAAGGGCACGTCCGCCTGGGCGT
JN166060.1	AATCTTTGAACGCAAGTTGCGCCTGAGGCCAATTGGCTAAGGGCACGTCCGCCTGGGCGT
GQ396668.1	AATCTTTGAACGCAAGTTGCGCCTGAGGCCAATTGGCTAAGGGCACGTCCGCCTGGGCGT
EU700340.1	AATCTTTGAACGCAAGTTGCGCCTGAGGCCAATTGGCTAAGGGCACGTCCGCCTGGGCGT

LKT	CAAGCATTACATCGCTTCATTGACACCAATTGCCAGTATTGTGCTGTGGTGTGCTGGTCT
JN166058.1	CAAGCATTACATCGCTTCATTGACACCAATTGCCAGTATTGTGCTGTGGTGTGCTGGTCT
JN166060.1	CAAGCATTACATCGCTTCATTGACACCAATTGCCAGTATTGTGCTGTGGTGTGCTGGTCT
GQ396668.1	CAAGCATTACATCGCTTCATTGACACCAATTGCCAGTATTGTGCTGTGGTGTGCTGGTCT
EU700340.1	CAAGCATTACATCGCTTCATTGACACCAATTGCCAGTATTGTGCTGTGGTGTGCTGGTCT

LKT	GAATGCGGAGAGTGGCCCTTCGTGCACACTTGTGCGACGGGTTGAAGAACAATTTGCTTT
JN166058.1	GAATGCGGAGAGTGGCCCTTCGTGCACACTTGTGCGACGGGTTGAAGAACAATTTGCTTT
JN166060.1	GAATGCGGAGAGTGGCCCTTCGTGCACACTTGTGCGACGGGTTGAAGAACAATTTGCTTT
GQ396668.1	GAATGCGGAGAGTGGCCCTTCGTGCACACTTGTGCGACGGGTTGAAGAACAATTTGCTTT
EU700340.1	GAATGCGGAGAGTGGCCCTTCGTGCACACTTGTGCGACGGGTTGAAGAACAATTTGCTTT

LKT	CCTCTGGCCATGTTTTGATAAAGGGGTGGTGTATGCAGCCATTGGCCACACTATCATC
JN166058.1	CCTCTGGCCATGTTTTGATAAAGGGGTGGTGTATGCAGCCATTGGCCACACTATCATC
JN166060.1	CCTCTGGCCATGTTTTGATAAAGGGGTGGTGTATGCAGCCATTGGCCACACTATCATC
GQ396668.1	CCTCTGGCCATGTTTTGATAAAGGGGTGGTGTATGCAGCCATTGGCCACACTATCATC
EU700340.1	CCTCTGGCCATGTTTTGATAAAGGGGTGGTGTATGCAGCCATTGGCCACACTATCATC

LKT	TCATTGCCTTGAGGAGGATAAATGTACACATTCTGGGCTGATCACCCGATATAAATGTCG
JN166058.1	TCATTGCCTTGAGGAGGATAAATGTACACATTCTGGGCTGATCACCCGATATAAATGTCG
JN166060.1	TCATTGCCTTGAGGAGGATAAATGTACACATTCTGGGCTGATCACCCGATATAAATGTCG
GQ396668.1	TCATTGCCTTGAGGAGGATAAATGTACACATTCTGGGCTGATCACCCGATATAAATGTCG
EU700340.1	TCATTGCCTTGAGGAGGATAAATGTACACATTCTGGGCTGATCACCCGATATAAATGTCG

LKT	CAGGTGACGCCCTGAAATGCGACCCC
JN166058.1	CAGGTGACGCCCTGAAATGCGACCCC-
JN166060.1	CAGGTGACGCCCTGAAATGCGACCCC-
GQ396668.1	CAGGTGACGCCCTGAAATGCGACCCC-
EU700340.1	CAGGTGACGCCCTGAAATGCGACCCC-

Bảng 4. Các vị trí nucleotide sai khác trong trình tự gen *ITS2* của loài Lan Kim tuyến nghiên cứu với 4 mã gen công bố trong Ngân hàng Gen quốc tế

STT	Tên loài	Mã hiệu	Vị trí nucleotide sai khác				
			10	112	221	231	273
1	<i>A. formosanus</i> Hayata nghiên cứu	MW663932	-	G	C	T	G
2	<i>A. formosanus</i>	GQ396668.1	A	T	C	T	G
3	<i>A. albolineatus</i>	JN166058.1	A	T	T	A	G
4	<i>A. lylei</i>	JN166060.1	A	T	C	T	G
5	<i>A. koshunensis</i>	EU700340.1	A	T	C	T	A

Trình tự nucleotide gen *ITS2* của loài Lan Kim tuyến nghiên cứu có 5 vị trí nucleotide sai khác đối với 4 mã hiệu gen so sánh (Bảng 4). Trong đó, so sánh với trình tự nucleotide gen *ITS2* của loài *A. formosanus* Đài Loan công bố thì mẫu Lan Kim tuyến nghiên cứu có 2 vị trí nucleotide sai khác là vị trí 10 và 112. Như vậy, gen *ITS2* có thể dùng để phân biệt các xuất xứ loài *A. formosanus* với nhau. Đối với mức độ loài, Lan Kim tuyến nghiên cứu có sự khác biệt với loài *A. albolineatus* ở 4 vị trí (10, 112, 221 và 231), với loài *A. koshunensis* có sự khác biệt ở 3 vị trí (10, 112 và 273) và với loài *A. lylei* có sai khác ở 2 vị trí (số 10 và 112). Với những phân tích trên, cho thấy có thể dùng gen *ITS2* để phân biệt ở mức độ loài hoặc dưới loài trong chi *Anoectochilus*.

4. KẾT LUẬN

Đã xác định được 3 đoạn trình tự DNA mã vạch cho loài Lan Kim tuyến *A. formosanus* Hayata, đoạn gen *matK* kích thước 790 bp (mã số trong Genbank: MW678627), gen *rbcL* kích thước 523 bp (mã số trong Genbank: MW678628) và gen *ITS2* kích thước 335 bp (mã số trong Genbank: MW663932). Trình tự nucleotide của các đoạn gen này đều có sự sai khác với các trình tự gen tương đồng công bố trên Ngân hàng Gen quốc tế. Sử dụng chỉ thị gen *matK* và *rbcL* nghiên cứu có thể phân loại và định danh chính xác loài Lan Kim tuyến *A. formosanus* Hayata tại Thanh Hóa, góp phần quan trọng cho công tác bảo tồn, khai thác và phát triển nguồn gen loài cây dược liệu quý này.

TÀI LIỆU THAM KHẢO

1. Bộ Khoa học và Công nghệ (2007). Sách Đỏ Việt Nam. Phần II: Thực vật. NXB Khoa học và Công nghệ, Hà Nội, 401-402.
2. CBOL Plant Working Group (2009). A DNA barcode for land plants. Proc Natl Acad Sci USA 106: 12794-12797.
3. Chính phủ nước CHXHCN Việt Nam (2019). Nghị định số 06/2019/NĐ - CP ngày 22 tháng 01

năm 2019 về quản lý thực vật rừng, động vật rừng nguy cấp, quý, hiếm và thực thi Công ước về buôn bán quốc tế các loài động vật, thực vật hoang dã nguy cấp.

4. Chuan Gao (2009). Identification of *Anoectochilus* based on rDNA ITS sequences alignment and SELDI-TOF-MS. Int J Biol Sci. 5(7): 727 – 735.

5. Du, X. M., Sun, N. Y., Hayashi, J., Chen, Y., Sugiura, M., and Shoyama, Y. (2003). Hepatoprotective and antihyperliposis activities of *in vitro* cultured *Anoectochilus formosanus*. Phytotherapy Research: An International Journal Devoted to Pharmacological and Toxicological Evaluation of Natural Product Derivatives. 17(1): 30 - 33.

6. Fang H. L., Wu J. B., Lin W. L., Ho H. Y. and Lin W. C. (2008). Further studies on the hepatoprotective effects of *Anoectochilus formosanus*. Phytotherapy Research: An International Journal Devoted to Pharmacological and Toxicological Evaluation of Natural Product Derivatives. 22(3): pp. 291 - 296.

7. Group C. P. W., Hollingsworth P. M., Forrest L. L., et al. (2009). A DNA barcode for land plants. Proceedings of the National Academy of Sciences. 106 (31): 12794 - 12797.

8. Hà Văn Huân, Hoàng Minh Trang, Bùi Thị Mai Hương (2020). Nghiên cứu xác định đoạn DNA barcode cho loài Hoàng đàn (*Cupressus tonkinensis*) phục vụ giám định loài. Tạp chí Khoa học và Công nghệ Lâm nghiệp số 1: 3 - 10.

9. Hebert P. D. N., Alina C., Shelley L. B., Jeremy R. (2003). Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B, 270: 313 - 321.

10. Kress W. J., Erickson D. L. (2007). A two - locus global DNA barcode for land plants: the coding

rbcl gene complements the non - coding trnH - psbA spacer region. *PLoS One* 2: e508.

11. Levin R. A., Wagner W. L., Hoch P. C., *et al.* (2003). Family - level relationships of Onagraceae based on chloroplast rbcL and ndhF data. *American Journal of Botany*, 90: 107 - 115.

12. Lin Shun - Fu, Tsay Hsin - Sheng, Chou Tsui - wei, Yang Ming - Jing, Cheng Kur - ta (2007). Genetic variation of *Anoetochilus formosanus* revealed by ISSR and AFLP analysis. *Journal of Food and Drug Analysis* 15 (2): 156 - 162.

13. Nguyễn Trọng Quyền, Hoàng Văn Sâm, Bùi Văn Thắng (2020). Đa dạng và phân bố các loài thuộc chi Lan Kim tuyến (*Anoetochilus* Blume) tại Thanh Hóa, Tạp chí Nông nghiệp và PTNT, 12: 80 - 86.

14. Phạm Hoàng Hộ (2010). Cây cỏ Việt Nam III. Nhà xuất bản Trẻ, thành phố Hồ Chí Minh.

15. Vijayan K. and Tson C. H. (2010). DNA barcoding in plants: taxonomy in a new perspective. *Current science*, 99: 1530 - 1540.

16. Võ Văn Chi (2003). Từ điển thực vật thông dụng, tập I. Nxb Khoa học và Kỹ thuật.

17. Yao H., Song J., Liu C., Luo K., Han J. (2010). Use of ITS2 region as the universal DNA barcode for plants and animals, *PLoS ONE* (5): 13.

18. Yong H. L., Jinlan R., Shilin C., Jingyuan S., Kun L., Dong L. and Hui Y. (2010). Authentication of *Taxillus chinensis* using DNA barcoding technique, *Journal of Medicinal Plants Research*, 4(24): 2706 - 2709.

DETERMINATING DNA BARCODE FOR CLASSIFICATION AND IDENTIFICATION OF *Anoetochilus formosanus* Hayata IN THANH HOA PROVINCE

Nguyen Trong Quyen¹, Le Cong Manh², Nguyen Thi Tho²,
Khuat Thi Hai Ninh², Hoang Van Sam², Bui Van Thang²

¹Thanh hoa Agriculture Institute

² Vietnam National University of Forestry

Summary

Anoetochilus formosanus Hayata, a highly valued medicinal plant, has been over-exploited, which can lead to the extinction of this species in the wild. However, it is difficult to distinguish this plant from other species of genus *Anoetochilus* because of the similar morphological characteristics of these species. This study aimed to develop some DNA barcode markers of *A. formosanus* Hayata collected in Thanh Hoa province to accurately identify this species for the conservation and development of the genetic resources. In the outcome of this study, we have identified three DNA barcode sequences from *A. formosanus* Hayata genome, which were *matK*, *rbcL* and *ITS2*. The three sequences had high similarities with the respective sequences published on the International GeneBank (NCBI) with the differences were only from 3 to 8 nucleotides. Out of the three DNA barcode sequences, *matK* and *ITS2* were effective candidates for the classification and identification of *A. formosanus* Hayata in Thanh Hoa. These findings are the scientific basis for the conservation, exploitation and development of the genetic resources of of this plant species.

Keywords: *Anoetochilus formosanus* Hayata, DNA barcode, identification, gene sequences.

Người phản biện: PGS.TS. Khuất Hữu Trung

Ngày nhận bài: 4/02/2021

Ngày thông qua phản biện: 4/3/2021

Ngày duyệt đăng: 11/3/2021