

## PHÂN TÍCH VÙNG GEN *trnL-trnF* TRÊN CÂY CÀ GAI LEO (*SOLANUM PROCUMBENS* LOUR.) CỦA VIỆT NAM

Huỳnh Thị Thu Huệ<sup>1,2</sup>, Nguyễn Thị Thanh Hoa<sup>1</sup>, Lê Thị Thu Hiền<sup>1</sup>, Nguyễn Đăng Tôn<sup>1,2,✉</sup>

<sup>1</sup>Viện Nghiên cứu hệ gen, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

<sup>2</sup>Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

✉Người chịu trách nhiệm liên lạc. E-mail: dtnguyen@igr.ac.vn

Ngày nhận bài: 12.8.2020

Ngày nhận đăng: 04.12.2020

### TÓM TẮT

Chi Cà (*Solanum*) nằm trong họ Cà Solanaceae bao gồm nhiều cây trồng là lương thực, thực phẩm và cây thuốc như cà tím, cà chua, khoai tây, cà gai leo, cà hai lá v.v. Số lượng loài được sử dụng làm cây thuốc trong họ Cà ở Việt Nam khá nhiều lên tới 41 loài một trong đó là Cà gai leo (*Solanum procumbens* Lour.) thường được sử dụng để chữa trị viêm gan, xơ gan. Hiện nay, nhu cầu trồng cây Cà gai leo ngày càng tăng cao nhưng các nghiên cứu dựa trên các chỉ thị DNA của Cà gai leo nhằm phát triển chỉ thị để tránh sự nhầm lẫn về nguyên liệu trong quá trình thu mua nguyên liệu thô ở dạng phơi khô chưa có nhiều. Nghiên cứu này đã tiến hành giải trình tự, phân tích chỉ thị DNA vùng *trnL-trnF* nhằm phục vụ cho các nghiên cứu về định danh, phân loại, tiến hóa của Cà gai leo. Đoạn gen vùng *trnL-trnF* đã được nhân bản, giải trình tự và phân tích từ 10 mẫu lá khô Cà gai leo. Kết quả cho thấy độ tương đồng đoạn gen *trnL-trnF* của các mẫu và so sánh với các công bố thuộc loài *Solanum procumbens* Lour. là 100%. Ngoài ra, khoảng cách di truyền dựa trên đoạn gen *trnL-trnF* của các mẫu so với một số loài trong Chi *Solanum* dao động từ 0,0082 tới 0,0399. Các số liệu thu được từ nghiên cứu đã cung cấp thêm thông tin cho những nghiên cứu đa dạng di truyền về chi Cà (*Solanum*) cũng như cây thuốc của Việt Nam.

**Từ khóa:** Cà gai, khoảng cách di truyền, phân tích gen, *trnL*, *trnF*.

### MỞ ĐẦU

Chi Cà (*Solanum*) với khoảng 1500 loài phân bố trên toàn thế giới, là chi lớn nhất trong họ Cà Solanaceae bao gồm nhiều cây trồng là lương thực, thực phẩm và cây thuốc như cà tím, cà chua, khoai tây, cà gai leo, cà hai lá.v.v (Bennett, 2008) ngoài ra còn có các loài cà dại như *S. carolinense*, *S. elaeagnifolium*, *S. rostratum*, *S. torvum* gây ra mối quan ngại lớn đối với kiểm dịch dịch hại. Ở Việt Nam, họ Cà có tới 15 chi với 63 loài trong đó chi Cà *Solanum* có 31 loài. Số lượng loài được sử dụng làm cây thuốc trong họ Cà ở Việt Nam khá nhiều lên tới 41 loài (Pham *et al.*, 2016), một trong đó là Cà gai leo (*Solanum procumbens* Lour.). Cây thuốc này được biết đến là loại thảo dược dùng để chữa trị

viêm gan, xơ gan nhiều năm nay. Cà gai leo ban đầu là thực vật mọc hoang ở nhiều nơi trong tự nhiên, loại cây này phân bố tập trung nhất ở các tỉnh Trung du miền núi Bắc và Trung Bộ như Nghệ An, Thanh Hóa, Hà Nam, Thái Bình và một số tỉnh phía Nam. Đến nay, Cà gai leo được trồng ở nhiều vườn dược liệu Bắc Bộ cung cấp nguyên liệu cho các nhà máy sản xuất thuốc điều trị các bệnh về gan.

Các nghiên cứu chiết xuất các thành phần hoạt chất ở cà gai leo đã được tiến hành nhiều năm qua nên giá trị làm thuốc chữa bệnh của loại cà này đã được hiểu biết nhiều và phục vụ tốt cho ngành sản xuất dược liệu. Do nhu cầu trồng cây Cà gai leo ngày càng tăng cao, các nghiên cứu điều tra khảo sát về đặc điểm sinh trưởng phát

triển và biện pháp kỹ thuật trồng cũng được nghiên cứu (Hoàng Thị Sáu *et al.*, 2016; Phùng Thị Thu Hà *et al.*, 2017) hay các biện pháp nhân giống nhằm cung cấp lượng lớn các cây giống (Hoàng Kim Toàn *et al.*, 2018). Tuy nhiên, các nghiên cứu dựa trên các chỉ thị DNA của Cà gai leo chưa có nhiều nên việc phát triển chỉ thị cho Cà gai leo là cần thiết nhằm tránh sự nhầm lẫn về nguyên liệu trong quá trình thu mua nguyên liệu thô ở dạng phơi khô. Vì vậy, trong nghiên cứu này chúng tôi đã tiến hành giải trình tự, phân tích các chỉ thị DNA thông dụng trên một số mẫu lá khô Cà gai leo nhằm phục vụ cho các nghiên cứu khác sâu hơn về định danh, phân loại, tiến hóa sau này.

Mã vạch DNA (DNA-barcoding) là kỹ thuật định danh loài bằng cách sử dụng các vùng DNA chuẩn trên hệ gen sinh vật, phương pháp này phục vụ cho các nghiên cứu liên quan đến phân loại nhờ tính chính xác và nhanh chóng của nó (Kress, 2017). Mã vạch DNA sử dụng vùng gen của ty thể đã được thiết lập tốt ở động vật, việc lựa chọn các mã vạch cho thực vật thường sử dụng các vùng gen ở lục lạp, cùng với sự kết hợp với một vài vùng trên hệ gen nhân như các đoạn ITS.

Vùng *trnL-trnF* nằm trong vùng sao chép đơn lớn của bộ gen lục lạp. Gen *trnL* là một phần trong vùng *trnT-L-F* bị phân chia bởi nhóm I intron, vùng không mã hóa nằm giữa các gen chức năng *trnT-trnL* và vùng mã hóa *trnF* các vùng gen này được đồng thời phiên mã (Bakker *et al.*, 2000). Vùng *trnL* intron nằm giữa nucleotide U và A của vòng lặp anticodon UAA. Các cấu trúc thứ cấp trong vùng *trnL* intron rất quan trọng vì RNA vận chuyển có chức năng mang các thông tin để mã hóa cho gen *trnL* có liên quan đến nó và vùng không mã hóa bên trong nó (Pirie *et al.*, 2007).

Cho đến nay, nhiều nghiên cứu phát sinh chủng loại sử dụng các chỉ thị DNA hiệu quả. Nghiên cứu của Levin *et al.* (2006) đã đánh giá mối quan hệ trong nhóm *Leptostemonum* thuộc chi *Solanum*, trong đó sử dụng các chỉ thị DNA là đoạn ITS, *waxy* gene và đoạn *trnS-trnL* (Levin *et al.*, 2006). Trình tự DNA của một số gen, bao

gồm vùng ITS2 của DNA nhân, gen *Waxy* và các vùng lục lạp (*ndhF* và *trnL*, *trnF*), cũng đã được sử dụng để đánh giá sự phát sinh loài của chi *Physalis* và mối quan hệ của chúng với các chi khác trong họ Solanaceae (Olmstead *et al.*, 2008; Feng *et al.*, 2016). Quan trọng hơn việc có thêm thông tin về các loài quan tâm chẳng hạn như các loài thực vật có giá trị kinh tế vào mạng lưới dữ liệu phát sinh gen sẽ làm giảm đáng kể công việc thu mẫu. Ở một mức độ nào đó, điều này sẽ giúp khắc phục các khó khăn trong việc thu mẫu, việc gây trở ngại cho các nhà nghiên cứu phân loại, tiến hóa, đặc biệt là khi nghiên cứu nhóm lớn phân bố trên toàn thế giới hoặc các loài có nguy cơ tuyệt chủng với nguyên liệu quý hiếm, các loài cây thuốc có giá trị cao.

Trong nghiên cứu này, chúng tôi trình bày quá trình giải trình tự, phân tích chỉ thị DNA vùng *trnL-trnF* nhằm phục vụ cho các nghiên cứu về định danh, phân loại, tiến hóa của Cà gai leo được thu thập tại Việt Nam.

## VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

### Vật liệu

Lá phơi khô của 10 mẫu Cà gai leo được thu thập tại vùng miền Bắc Việt Nam, ký hiệu mẫu từ 1A đến 5A và 1B đến 5B.

### Phương pháp

#### Tách chiết DNA tổng số, khuếch đại gen và xác định trình tự

DNA tổng số được tách chiết bằng phương pháp tách chiết sử dụng CTAB (cetyl trimethylammonium bromide) theo Doyle, Doyle. (1990). Cặp mồi đặc hiệu được thiết kế dựa trên thông tin về trình tự gen trên Ngân hàng gen quốc tế (Bảng 1).

Phản ứng khuếch đại các đoạn gen được tối ưu sử dụng điều kiện như sau: 25  $\mu$ L tổng thể tích bao gồm: 50 ng DNA tổng số; 2,5  $\mu$ M mỗi loại primer; 1 unit Taq DNA polymerase; 1 mM dNTP mỗi loại và đệm tương ứng trên máy PCR với chu trình nhiệt như sau: 1 chu kỳ biến tính 95°C/ 3 phút; 35 chu kỳ (95°C/ 30 giây; 52°C/ 30

giây; 72°C/ 1 phút) và bước tổng hợp cuối cùng 72°C/ 5 phút. Sản phẩm PCR được kiểm tra trên gel agarose 0,8%, sau đó được tinh sạch sử dụng bộ hóa chất GeneJET™ PCR Purification Kit (Hãng Thermo Scientific, Mỹ). Trình tự các đoạn DNA được xác định trên hệ thống ABI 3500 Genetic Analyzer theo nguyên lý của Sanger, với bộ kit BigDye® Terminator v 3.1 Cycle Sequencing (Hãng Applied Biosystems, Mỹ) bằng cách xác định trình tự trực tiếp từ sản phẩm PCR.

**Phân tích so sánh trình tự nucleotide các đoạn *trnL-trnF***

Trình tự DNA được xử lý và phân tích bằng phần mềm BioEdit, trong đó có so sánh với các trình tự tương ứng của các loài thuộc chi Cà đã được công bố trên GenBank. Biểu đồ hình cây về mối quan hệ phát sinh loài được xây dựng theo phương pháp Maximum-Likelihood với giá trị bootstrap là 1000 lần lặp lại dựa trên cơ sở các số liệu thu được sử dụng phần mềm MEGA X (Kumar *et al.*, 2018).

**Bảng 1.** Thông tin trình tự mỗi dùng để khuếch đại gen và giải trình tự.

Tên mồi	Vùng DNA	Trình tự mồi (5'→3')
<i>trnL-trnF-F</i>	<i>trnL-trnF</i>	TGGCGAAATTGGTAGACGC
<i>trnL-trnF-R</i>		AACCATCTCGTCTCCTGA

**Bảng 2.** Thông tin trình tự chuỗi gen tham chiếu trên GenBank.

TT	Tên mẫu	Mã truy cập trên GenBank
1	<i>Solanum procumbens</i>	HQ721938.1
2	<i>Solanum lycopersicum</i>	AY098703.1
3	<i>Solanum campylacanthum</i>	HQ721908.1
4	<i>Solanum burtii-davyi</i>	HQ721906.1
5	<i>Solanum caesium</i>	HM006843.1
6	<i>Solanum tuberosum</i>	HM006842.1
7	<i>Solanum dulcamara</i>	HM006840.1
8	<i>Solanum glaucophyllum</i>	HM006831.1
9	<i>Solanum betaceum</i>	HM006830.1
10	<i>Solanum_mauritianum</i>	HM006828.1
11	<i>Solanum citrullifolium</i>	HM006826.1
12	<i>Solanum granuloso-leprosum</i>	JN661825.1
13	<i>Solanum adscendens</i>	JN661822.1
14	<i>Solanum schomburgkii</i>	GU591042.1
15	<i>Solanum caricifolium</i>	GU591006.1
16	<i>Solanum villosum</i>	GU323356.1
17	<i>Solanum petraeum</i>	GQ163549.1
18	<i>Solanum virginianum</i>	EU176160.1
19	<i>Solanum violaceum</i>	EU176159.1

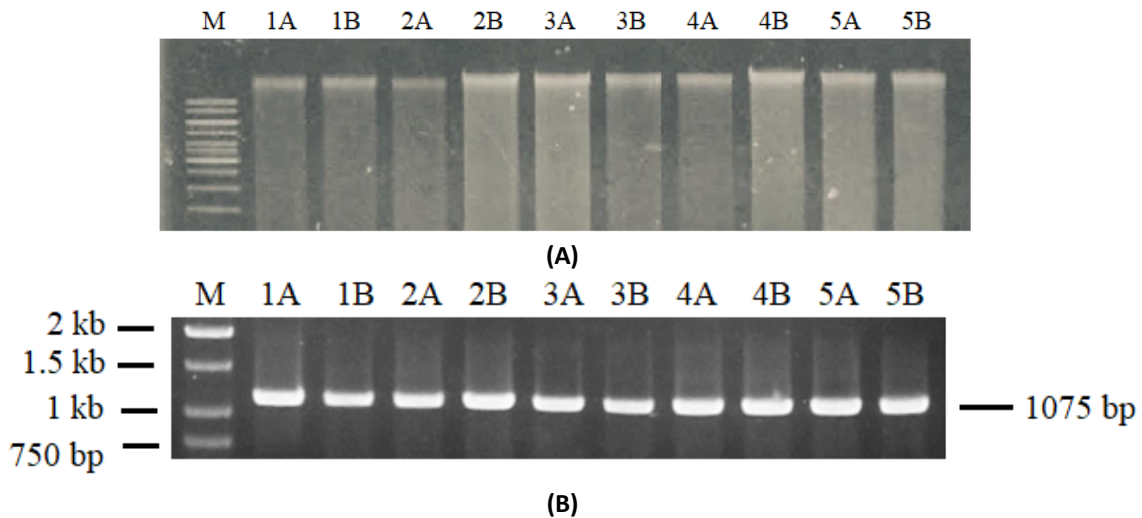
## KẾT QUẢ

### PCR nhân các đoạn gen *trnL-trnF*

Mẫu lá khô của Cà gai leo được nghiền kỹ trong N2 lỏng và đệm chiết sau đó được chiết với phenol và chloroform 2–3 lần tùy theo mẫu. Sau khi tối ưu, mỗi mẫu đều thu được DNAs với độ tinh khiết và nồng độ đáp ứng với yêu cầu nhân gen (Hình 1A). Đoạn trình tự *trnL-trnF* nằm trong lục lạp đã được nhân bản ở nhiều loài thực vật. Dựa vào thông tin về trình tự đoạn gen *trnL-trnF* ở loài Cà gai leo được công bố trên NCBI, cặp primer được thiết kế nhằm nhân một đoạn DNA trong vùng *trnL-trnF* có kích thước khoảng 1 kb. Kết quả kiểm tra sản phẩm điện di trên gel agarose 0,8% cho thấy đoạn gen *trnL-trnF* đều được khuếch đại ở các mẫu, băng DNA có độ đậm nhạt khác nhau ở một số mẫu nghiên cứu, nhưng các băng thu

được đều sáng và rõ chỉ có một băng vạch duy nhất với kích thước đúng như tính toán khi thiết kế primer (khi so sánh với thang DNA chuẩn 1 kb) vào khoảng 1 kb (Hình 1B).

Các sản phẩm PCR nhân các đoạn gen trên sau khi tinh sạch được xác định trình tự trực tiếp theo phương pháp Sanger. Kết quả Blast trên NCBI cho thấy toàn bộ các sản phẩm PCR nhân lên đều chính xác là vùng gen đích *trnL-trnF*, tuy nhiên 1 mẫu 4A tín hiệu nhiễu nên được loại bỏ khi phân tích. Đoạn gen sau khi phân tích loại bỏ các đoạn đọc nhiễu đầu và cuối còn lại gồm toàn bộ phần intron và exon 2 của gen *trnL* cùng với đoạn *trnL-trnF* intergenic spacer, kích thước đoạn DNA này dài 1075 bp. Qua so sánh bằng phần mềm Bioedit các trình tự *trnL-trnF* trên các mẫu đã thu thập cho thấy độ tương đồng 100% giữa các mẫu nghiên cứu và với trình tự HQ721938.1 của loài *Solanum procumbens*.



**Hình 1.** Kết quả tách DNA tổng số (A) và sản phẩm PCR nhân vùng *trnL-trnF* (B). Chú thích: M: thang DNA chuẩn (marker 1kb), 1A-5B: sản phẩm từ mẫu 1A đến 5B.

### Phân tích và so sánh trình tự đoạn *trnL-trnF*

Vùng gen *trnL-trnF* lục lạp này có tính đa hình cao, để so sánh đoạn gen *trnL-trnF* của các mẫu nghiên cứu với đoạn gen này của các loài trong cùng chi *Solanum* (các trình tự tham chiếu thể hiện ở Bảng 2) để thấy sự đa hình của chúng với nhau. Qua phân tích cho thấy, trong

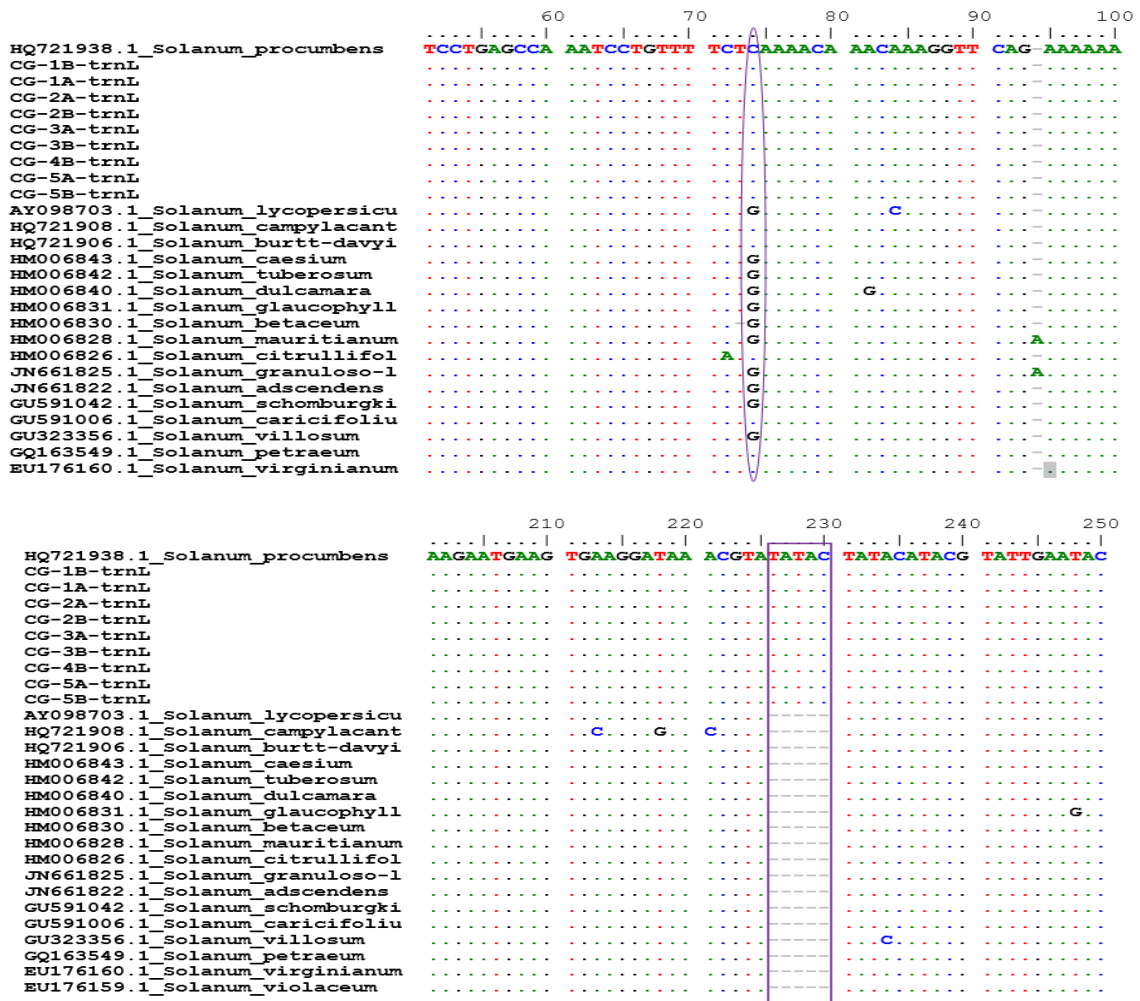
vùng exon 2 của gen *trnL* không có điểm sai khác về nucleotide giữa các trình tự được so sánh, toàn bộ những đa hình nucleotide đều nằm trong vùng intron của gen *trnL* và vùng *trnL-trnF* intergenic spacer. Mặt khác, các đa hình, insert hay delete nucleotide xảy ra nhiều ở vùng đệm giữa *trnL* và *trnF* hơn là vùng intron. Trong đó, có những vị trí thể hiện sự

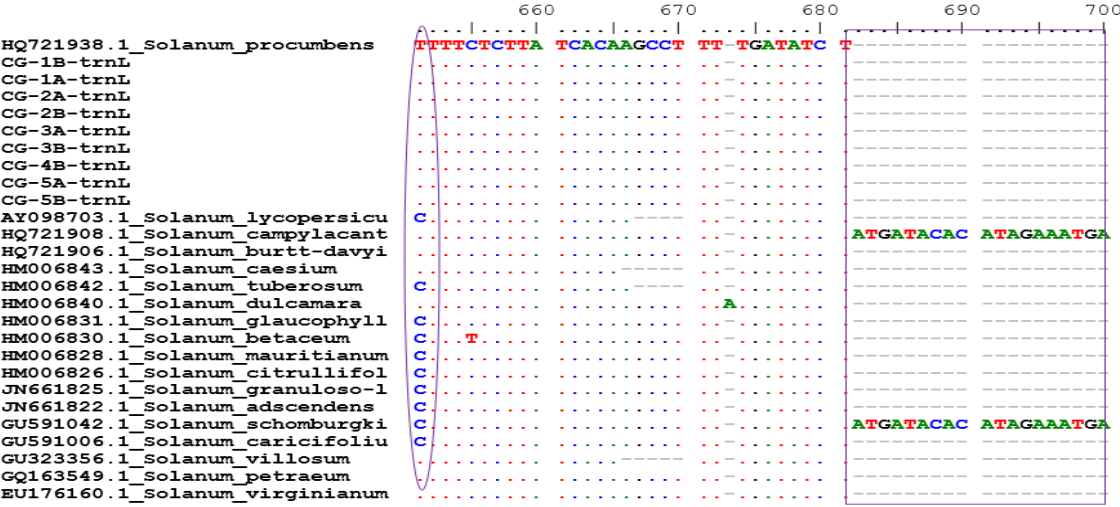
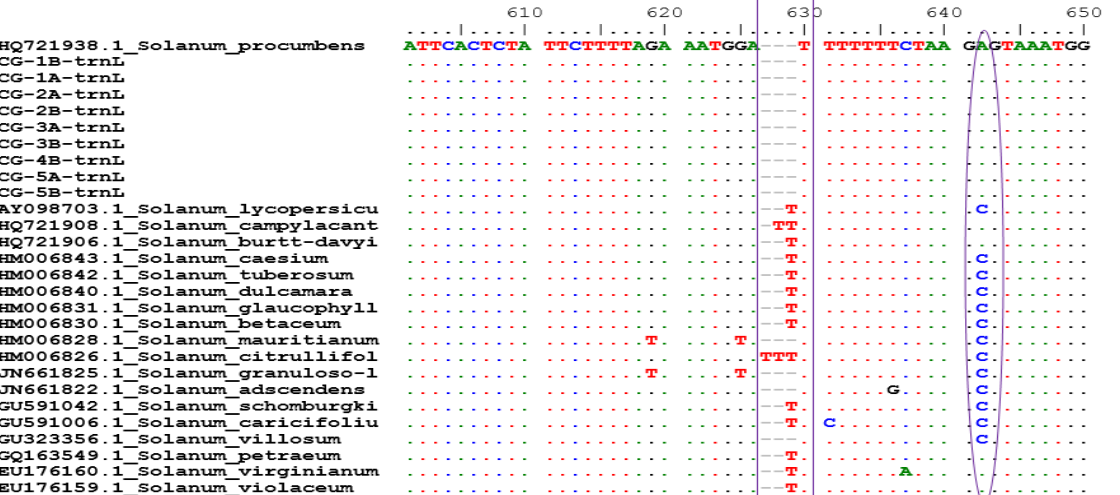
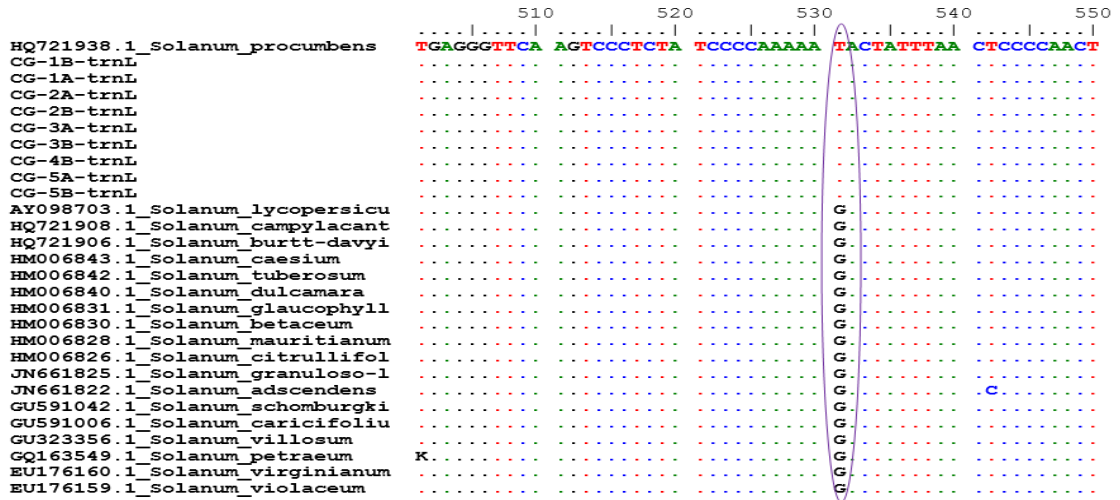
khác biệt khá rõ nét phân biệt giữa loài *S. procumbens* với một số loài khác cùng chi, như vị trí đa hình 74 C>G; insert 226 TATAC; đa hình 531 T>G; del 627 TTT; đa hình 642 A>C; đa hình 651 T>C; del 869 TTTTGTCTT; đa hình 883 C>A; đa hình 905 G>T; đa hình 925 A>G; đa hình 936 C>G; đa hình 946 C>G; đa hình 968 A>G; đa hình 979 T>C; đa hình 983 A>G; del 984TTGGGAATAGCCGGGAT; đa hình 1002 T>G; đa hình 1004-1007 CAGA>TCAG; đa hình 1015-1017 GAT>AGC; đa hình 1027 G>A (Hình 2).

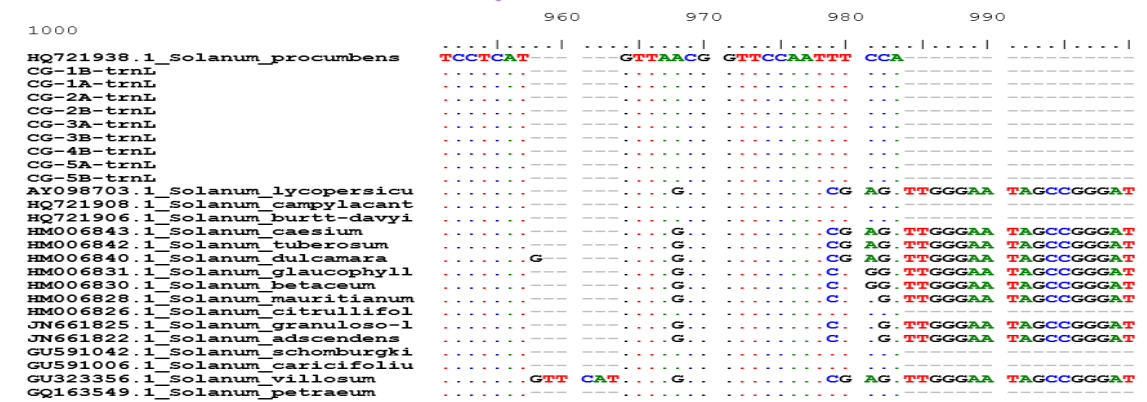
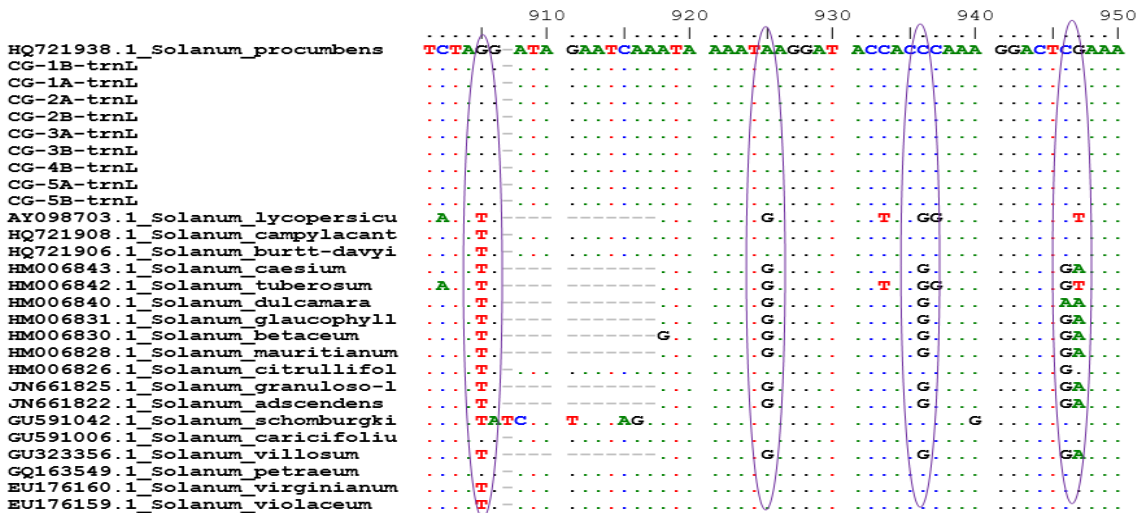
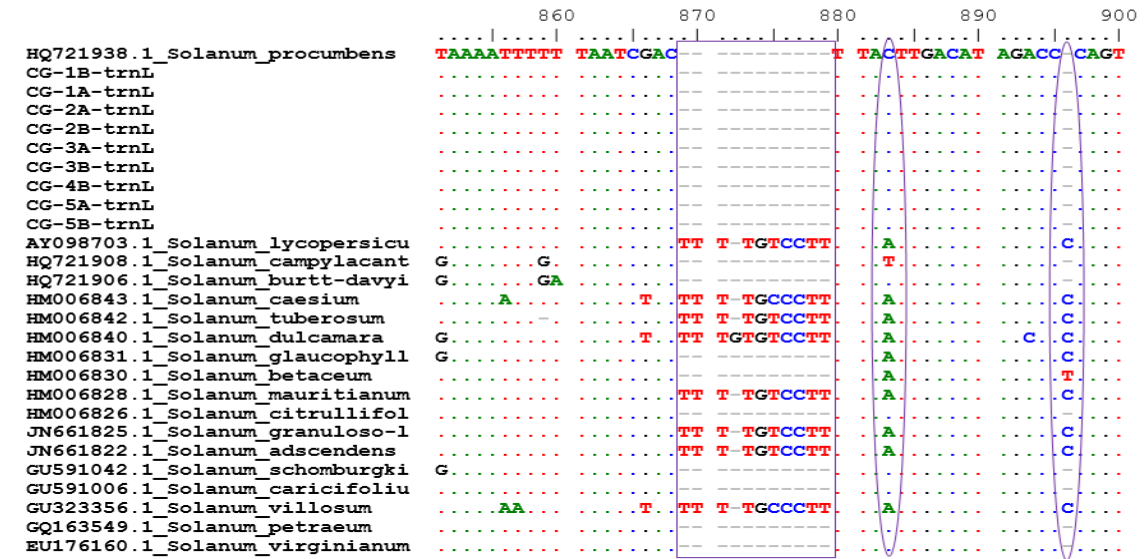
Bảng 3 cho thấy khoảng cách di truyền của Cà gai leo với 19 loài được so sánh thuộc chi *Solanum* là rất thấp và dao động từ 0,0082 (với

*S. violaceum*) đến 0,0399 (với *S. dulcamara*). Khoảng cách giữa các mẫu phân tích so với trình tự tham chiếu *S. procumbens* HQ721938.1 là 0.

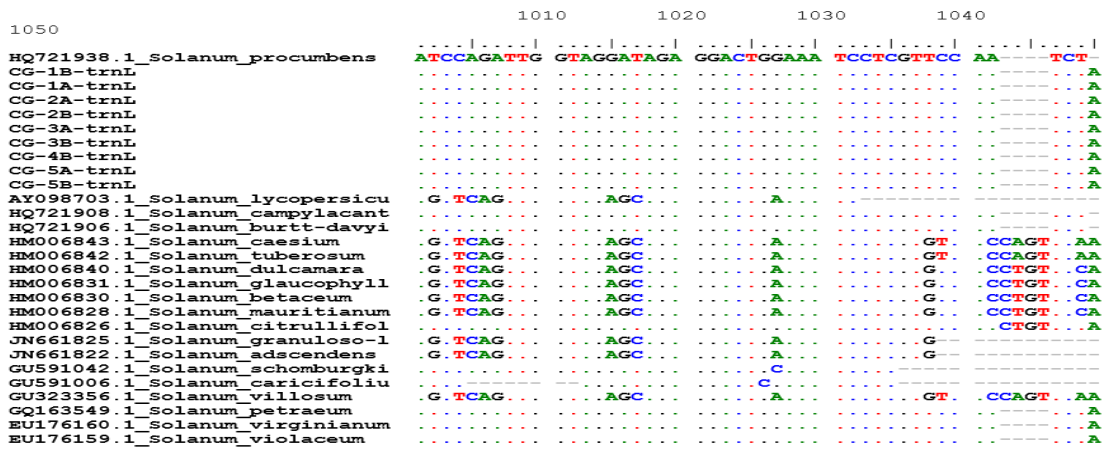
Dựa vào các trình tự đã xác định và các trình tự đoạn gen trên GenBank, chúng tôi đã tiến hành phân tích định danh loài dựa trên phương pháp xây dựng cây phân loại. Quan sát biểu đồ hình cây về mối quan hệ phát sinh loài và trên cơ sở so sánh trình tự các vùng gen *trnL* ở các mẫu nghiên cứu (Hình 3) cho thấy các mẫu này đều có quan hệ gần gũi với loài *S. procumbens* với giá trị khoảng cách di truyền bằng 0. Điều này cho thấy các mẫu nghiên cứu đều thuộc loài Cà gai leo (*S. procumbens*).











Hình 2. Đặc điểm một số đa hình nucleotide của *S. procumbens* với các loài khác trong chi *Solanum*.

Bảng 3. Khoảng cách di truyền giữa các mẫu dựa trên đoạn *trnL-trnF*.

	HQ72193	AY09870	HQ7219	HQ72190	HMD068	HMD06	HMD068	HMD068	HMD068	HMD068	JN66182	JN66182	GU5910	GU5910	GU3233	GQ1635	EU1761	EU17615	
HQ721938.1 <i>Solanum procumbens</i>																			
AY098703.1 <i>Solanum lycopersicum</i>	0.0333																		
HQ721908.1 <i>Solanum campylacanthum</i>	0.0104	0.0387																	
HQ721906.1 <i>Solanum burtt-davyi</i>	0.0104	0.0376	0.0084																
HMD06843.1 <i>Solanum caesium</i>	0.0340	0.0136	0.0392	0.0382															
HMD06842.1 <i>Solanum tuberosum</i>	0.0350	0.0042	0.0392	0.0382	0.0092														
HMD06840.1 <i>Solanum dulcamara</i>	0.0401	0.0198	0.0432	0.0422	0.0163	0.0193													
HMD06831.1 <i>Solanum glaucophyllum</i>	0.0338	0.0169	0.0359	0.0369	0.0154	0.0154	0.0184												
HMD06830.1 <i>Solanum betaceum</i>	0.0317	0.0158	0.0370	0.0380	0.0144	0.0144	0.0195	0.0071											
HMD06828.1 <i>Solanum mauritianum</i>	0.0370	0.0209	0.0422	0.0412	0.0204	0.0194	0.0233	0.0154	0.0144										
HMD06826.1 <i>Solanum citrullifolium</i>	0.0084	0.0337	0.0148	0.0148	0.0350	0.0339	0.0411	0.0306	0.0285	0.0348									
JN661825.1 <i>Solanum granuloso-leprosum</i>	0.0362	0.0219	0.0415	0.0404	0.0187	0.0177	0.0248	0.0167	0.0157	0.0010	0.0344								
JN661822.1 <i>Solanum adscendens</i>	0.0342	0.0199	0.0395	0.0385	0.0167	0.0156	0.0239	0.0147	0.0136	0.0114	0.0324	0.0124							
GU591042.1 <i>Solanum schomburgkii</i>	0.0158	0.0333	0.0196	0.0200	0.0322	0.0311	0.0373	0.0288	0.0288	0.0342	0.0160	0.0352	0.0332						
GU591006.1 <i>Solanum caricifolium</i>	0.0117	0.0358	0.0202	0.0202	0.0346	0.0335	0.0419	0.0333	0.0312	0.0366	0.0140	0.0377	0.0357	0.0213					
GU323356.1 <i>Solanum villosum</i>	0.0385	0.0179	0.0439	0.0428	0.0041	0.0134	0.0205	0.0197	0.0187	0.0247	0.0396	0.0230	0.0210	0.0368	0.0393				
GQ163549.1 <i>Solanum petraeum</i>	0.0031	0.0344	0.0115	0.0115	0.0348	0.0358	0.0409	0.0346	0.0325	0.0378	0.0094	0.0373	0.0353	0.0148	0.0128	0.0394			
EU176160.1 <i>Solanum virginianum</i>	0.0063	0.0354	0.0104	0.0104	0.0358	0.0368	0.0419	0.0356	0.0335	0.0388	0.0105	0.0383	0.0363	0.0179	0.0160	0.0404	0.0072		
EU176159.1 <i>Solanum violaceum</i>	0.0084	0.0367	0.0063	0.0063	0.0381	0.0381	0.0410	0.0347	0.0358	0.0411	0.0126	0.0406	0.0386	0.0180	0.0182	0.0427	0.0092	0.0082	
CG-trnL	0.0000	0.0334	0.0105	0.0105	0.0338	0.0348	0.0399	0.0336	0.0315	0.0367	0.0084	0.0362	0.0342	0.0159	0.0117	0.0383	0.0031	0.0061	0.0082

Nghiên cứu dựa trên các vùng chỉ thị mã vạch được sử dụng nhiều do thuận lợi cho phân loại học hệ thống học phân tử và phát sinh loài. Nhờ cách lấy mẫu đa dạng có thể từ lá, thân, củ v.v cũng như mẫu có thể tươi hoặc khô. Đoạn *trnL-trnF* được biết là vùng có nhiều biến đổi của DNA lục lạp, vùng này đã được sử dụng rộng rãi, trong các nghiên cứu phát sinh giữa các loài có liên quan chặt chẽ, trong nghiên cứu hệ thống, tiến hóa thực vật hoặc để xác định loài. Với vùng I intron duy nhất trong DNA lục lạp, việc so sánh nhiều trình tự các vùng *trnL* intron có thể cho phép thiết kế các đoạn mồi đa năng được gắn

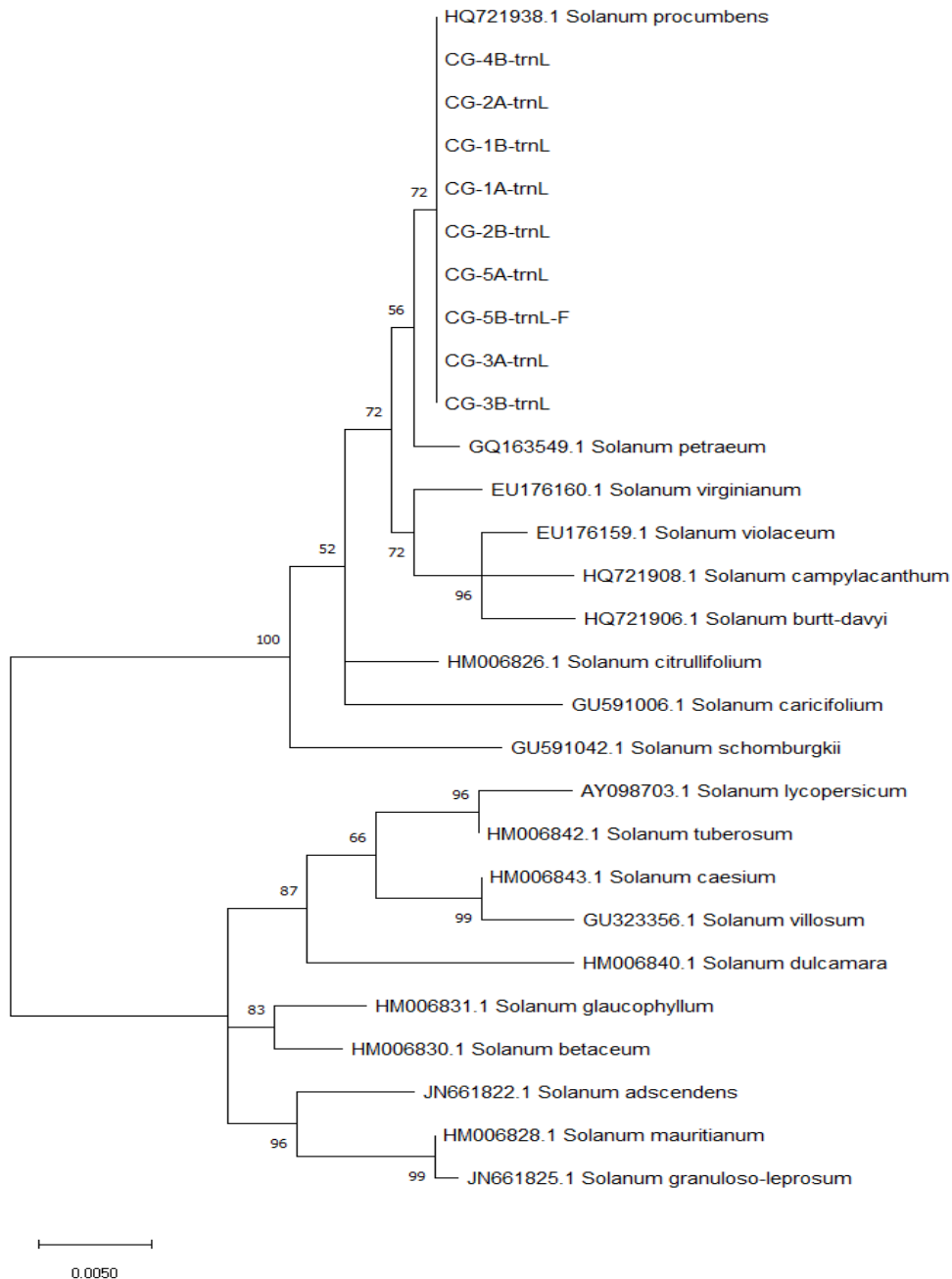
trong các vùng được bảo thủ và khuếch đại vùng biến đổi ngắn ở giữa hai gen *trnL* và *trnF* (Pierre *et al.*, 2007).

Trong nghiên cứu này, chúng tôi đã tối ưu việc tách chiết DNA và nhân đoạn bằng PCR dựa vào phân tích vùng *trnL-trnF* từ các mẫu lá khô, đã xác định đúng mẫu Cà gai leo. Cùng với những nghiên cứu đã có như nghiên cứu về đoạn *trnL-trnF* ở cây Huyền sâm, Hoàng liên gai, Xạ đen (Bùi Mạnh Minh *et al.*, 2018; Le Thi Thu Hien *et al.*, 2018; Nguyen Thuy Linh *et al.*, 2017) cũng cho thấy đây là chỉ thị



hữu ích trong nhận diện loài và kết quả cho thấy có thể sử dụng đoạn này để phân tích trên các mẫu khác chưa xác định. Các số liệu trình tự từ nghiên cứu sẽ góp phần cung cấp thêm thông tin cho những nghiên cứu đa dạng di

truyền về cây thuốc của Việt Nam cũng như khẳng định về tính hiệu quả khi sử dụng các chỉ thị mã vạch cho thực vật khi cần định loại với những mẫu vật khô không thể nhận biết hình thái.



**Hình 3.** Biểu đồ hình cây về mối quan hệ phát sinh loài dựa trên trình tự nucleotide vùng gen *trnL-trnF* được xây dựng bằng phương pháp Maximum-Likelihood.

## KẾT LUẬN

Trong nghiên cứu này, các DNA vùng gen *trnL-trnF* dài 1075 bp từ mẫu lá khô cần kiểm định Cà gai leo đã được nhân bản và xác định trình tự. Kết quả phân tích và so sánh cho thấy các mẫu thu thập có độ tương đồng 100% với trình tự tham chiếu đoạn gen *trnL-trnF* của Cà gai leo *Solanum procumbens*. Khi so sánh trình tự *trnL-trnF* của các mẫu nghiên cứu với các trình tự *trnL-trnF* của một số loài trong chi Cà (*Solanum*) cho thấy biên độ dao động khoảng cách di truyền thấp, chỉ từ 0.0082–0.0399. Các số liệu thu được từ nghiên cứu đã cung cấp thêm thông tin cho những nghiên cứu đa dạng di truyền về cây thuốc của Việt Nam cũng như chi Cà.

**Lời cảm ơn:** Nghiên cứu được sự hỗ trợ kinh phí và trang thiết bị của Viện Nghiên cứu hệ gen.

## TÀI LIỆU THAM KHẢO

- Hoàng Thị Sáu, Phạm Thị Lý, Trần Thị Mai (2016) Nghiên cứu một số biện pháp kỹ thuật trồng cây Cà gai leo tại Thanh Hoá. *Tạp chí Khoa học, Trường Đại học Hồng Đức*, 30: 79–89.
- Hoàng Kim Toàn, Lê Văn Tinh, Trần Thị Thu Giang, Trần Đăng Hòa, Lê Như Cương, Nguyễn Đình Thi (2018) Nghiên cứu một số biện pháp kỹ thuật nhân giống từ hạt cây Cà gai leo (*Solanum procumbens*). *Tạp chí Khoa học, Trường Đại học Huế: Khoa học Tự nhiên*, 127(1C): 59–170.
- Phùng Thị Thu Hà, Phạm Thị Huyền Trang, Nguyễn Hữu Cường (2017) Đặc điểm thực vật học và một số biện pháp kỹ thuật trồng Cà gai leo tại Gia lâm, Hà Nội. *Tạp chí Khoa học Nông nghiệp Việt Nam*, 15(6): 146–154.
- Le Hoang Pham, Michael Bohme, Ina Pinker (2016) Solanaceae diversity in Vietnam: a preliminary taxonomic inventory for conservation and utilization. *Agric For*, 62(4): 45–55.
- Bennett JR (2008) Revision of *Solanum* section Regmanda (Solanaceae). *Edinb J Bot*, 65(1): 69–112.
- Pham LH, Bohme M, Pinker I (2016) Solanaceae diversity in Vietnam: a preliminary taxonomic inventory for conservation and utilization. *Agric For*, 62(4): 45–55.
- Levin RA, Myers NR, Bohs L (2006) Phylogenetic relationships among the “spiny Solanums” (*Solanum* subgenus *Leptostemonum*, Solanaceae). *Am J Bot*, 93(1): 157–169.
- Feng S, Jiang M, Shi Y, Jiao K, Shen C, Lu J, Ying Q, Wang H (2016) Application of the Ribosomal DNA ITS2 Region of *Physalis* (Solanaceae): DNA Barcoding and Phylogenetic Study. *Front Plant Sci*, 7: 1047.
- Kress WJ (2017) Plant DNA barcodes: Applications today and in the future. *J Syst Evol*, 55(4): 235–410. Review.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol*, 35(6): 1547–1549.
- Olmstead RG, Bohs L, Migid HA, Santiago-Valentin E, Garcia VF, Collier SM (2008) A molecular phylogeny of the Solanaceae. *Mol Phylogenet*, 57(4): 1159–1181.
- Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermet T, Corthier G, Brochmann C, Willerslev E (2007) Power and limitations of the chloroplast *trnL* (UAA) intron for plant DNA barcoding. *Nucleic Acids Res*, 35(3): e14.
- Bakker FT, Culham A, Gomez-Martinez R, Carvalho J, Compton J, Dawtrey R, Gibby M (2000) Patterns of Nucleotide Substitution in Angiosperm cpDNA *trnL* (UAA)–*trnF* (GAA) Regions. *Mol Biol Evol*, 17(8): 1146–1155.
- Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price RA, Hills HG, Qiu Y, Kron KA, Rettig JH, Conti E, Palmer JD, Manhart JR, Sytsma KJ, Michaels HJ, Kress WJ, Karol KG, Clark WD, Hedren M, Gaut BS, Jansen RK, Kim KJ, Wimpee CF, Smith JF, Furnier GR, Strauss SH, Xiang QY, Plunkett GM, Soltis PS, Swensen SM, Williams SE, Gadek PA, Quinn CJ, Eguiarte LE, Golenberg E, Learn GH, Graham SW, Barrett SCH, Dayanandan S, Albert VA (1993) Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcl*. *Ann. Mo Bot Gard*, 80: 528–580.
- Clegg MT (1993) Chloroplast gene sequences and the study of plant evolution. *PNAS*, 90: 363–367.

Bui Manh Minh, Vu Anh Tuan, Vu Phuong Nhung, Pham Quang Cu, Nguyen Dang Ton, Huynh Thi Thu Hue (2018) DNA barcoding, an approach for molecular identification of Huyen Sam (*Scrophularia L.*) samples collected at Northern Vietnam. *Vietnam J Sci Technol*, 60(2): 56–64.

Le Thi Thu Hien, Nguyen Nhat Linh, Pham Le Bich Hang, Nguyen Phuong Mai, Ha Hong Hanh, Huynh Thi Thu Hue, Ha Van Huan (2018) Developing DNA

Barcodes for Species Identification of Berberis and Dysosma Genera in Vietnam. *Int J Agric Biol*, 20(5): 1097–1106.

Thuy Linh Nguyen, Thi Hang Pham, Van Truong Do, Thi Thu Hue Huynh (2017) Evaluating the systematic position of *Ehretia asperula* Zoll. & Moritzi based on ITS1, *matK* and *trnL-trnF* DNA sequences. *Vietnam J Sci Technol*, 59(4): 61–65.

## **ANALYZING THE PLASTID *trnL-trnF* SEQUENCES OF VIETNAMESE *SOLANUM PROCUMBENS* LOUR. PLANTS**

**Huynh Thi Thu Hue<sup>1,2</sup>, Nguyen Thi Thanh Hoa<sup>1</sup>, Le Thi Thu Hien<sup>1</sup>, Nguyen Dang Ton<sup>1,2\*</sup>**

<sup>1</sup>*Institute of Genome Research, Vietnam Academy of Science and Technology*

<sup>2</sup>*Graduate University of Science and Technology, Vietnam Academy of Science and Technology*

### **SUMMARY**

*Solanum* is a genus which belongs to Solanaceae family, includes many species of crops or medicinal plants as eggplant, tomato, potato, twoleaf nightshade, etc. Number of species used as medicinal herbs of *Solanum* genus in Vietnam is up to 41, one of those is *Solanum procumbens* Lour., which is often used to treat hepatitis or cirrhosis. The needs of planting this plant is rising up, but it still lacks of studies based on DNA barcode of *Solanum*, in order to prevent the confusion of the raw materials in dry form of medicinal plants. This study provides sequencing and analysis the DNA barcode region *trnL-trnF*, in order to support further researches about identification, classification and evolution of *Solanum procumbens* species. The DNA covering the *trnL-trnF* region was amplified, sequenced and the nucleotide sequences from 10 samples of dry leaves *Solanum procumbens* species were analyzed with the reference sequences from GenBank. The results show that the similarity between these samples and others in previous publications of *Solanum procumbens* Lour. is 100%. Moreover, the estimated genetic distance on *trnL-trnF* region of the Vietnamese samples compared to few other species in *Solanum* genera varies between 0,0082 to 0,0399. The data from this research provides more information for any further studies about genome biodiversity of the *Solanum* genus as well as Vietnamese medicinal plants.

**Keywords:** gene analysis, genetic distance, *trnL-trnF*, *Solanum procumbens*