

ĐÁNH GIÁ MỐI QUAN HỆ DI TRUYỀN CỦA MỘT SỐ GIỐNG GỪNG (*Zingiber spp.*) Ở VIỆT NAM DỰA TRÊN CHỈ THỊ PHÂN TỬ DNA BARCODE

Nguyễn Trường Giang, Mai Thị Kim Ngọc, Phan Quang Hương, Huỳnh Hữu Đức*

Trung tâm Công nghệ Sinh học Thành phố Hồ Chí Minh

TÓM TẮT

Hiện nay, các nghiên cứu về phân loại các loài gừng ở Việt Nam chủ yếu dựa vào các chỉ thị hình thái. Các phương pháp dựa trên hình thái này cần phải có đầy đủ thông tin dữ liệu về cấu tạo và cấu trúc hoa mà trong nhiều trường hợp rất khó khăn và hạn chế. Phương pháp này rất khó áp dụng ở các loài trong chi Gừng do chúng có nhiều đặc điểm hình thái khá giống nhau và khó ra hoa nên khó khăn cho việc phân loại. Trong nghiên cứu này, chúng tôi tiến hành đánh giá 6 mẫu gừng và kết hợp với 2 mẫu nghệ làm nhóm ngoài để đánh giá mối quan hệ di truyền bằng chỉ thị phân tử DNA barcode dựa trên vùng *matK*, *rbcL* trên cDNA và ITS trong nhân. Kết quả cho thấy tỷ lệ khuếch đại vùng *matK*, *rbcL*, ITS với 4 cặp mồi tương ứng *matK390-F* và *matK1326-R*, *rbcLa-F* và *rbcLa-R*, ITS1F và ITS 1R, ITS-1F và ITS-1R lần lượt là 100%, 87,5%, 87,5%, 87,6%. Mức độ tương đồng di truyền giữa các loài trong chi gừng với các trình tự tham khảo từ 97 - 100% và mức độ bao phủ từ 99 - 100%. Dựa trên cây phát sinh loài cho thấy gừng và nghệ thuộc hai nhóm khác nhau, gừng được chia làm 2 loài chính là *Zingiber zerumbet* và *Zingiber officinale*. Trên cơ sở đó chúng tôi đã đánh giá được mối quan hệ của 6 mẫu gừng và 2 mẫu nghệ. Các kết quả này có thể ứng dụng trong việc đánh giá mối quan hệ di truyền, nhận diện các loài thuộc họ Zingiberaceae, từ đó mở ra khả năng trong việc xác định các sản phẩm dược liệu có nguồn gốc từ gừng.

Từ khóa: DNA barcode, Zingiber, chỉ thị phân tử.

MỞ ĐẦU

Gừng là một loại thảo dược được biết từ rất lâu với nhiều công dụng như làm thuốc trong chữa trị bệnh và làm gia vị trong các món ăn. Thân ngầm (thân rễ) có thể được sử dụng ở dạng tươi, dạng bột, sấy khô, hoặc ở dạng dầu hay nước ép. Gừng là một chi của họ Zingiberaceae, cùng họ với bạch đậu khấu, nghệ và riềng. Gừng được sử dụng rộng rãi trên toàn thế giới để điều trị các chứng bệnh chán ăn, buồn nôn và nôn sau phẫu thuật, nhiễm trùng đường hô hấp trên, viêm phế quản, ho, viêm khớp và đau cơ. Ở một số nơi trên thế giới, nước gừng được bôi cho lên da để điều trị bỏng. Gừng cũng được sử dụng như một hương liệu trong công nghiệp thực phẩm và đồ uống, làm gia vị và hương liệu trong nấu ăn, và hương thơm trong xà phòng và mỹ phẩm. Gừng được sử dụng từ rất lâu ở khu vực Nam Á, cả ở dạng tươi và khô (Ali *et al.*, 2008; Shakya, 2015). Họ Gừng ở Việt Nam khá đa dạng, cho đến nay đã biết được khoảng hơn 140 loài thuộc 20 chi và gần 100 loài, tương đương với đa dạng họ Gừng ở Thái Lan, cao hơn so với họ Gừng ở Lào và Campuchia. Trong đó nhiều cây có giá trị như riềng (*Alpinia officinarum*), Nghệ (*Curcuma domestica* Val. hay *Curcuma longa* L.), gừng (*Zingiber officinale* Rosc), gừng gió (*Zingiber zerumbet*) (Phạm Hoàng Hộ, 1993).

Đến nay, các nghiên cứu về phân loại các loài gừng ở Việt Nam chủ yếu dựa vào chỉ thị hình thái. Phương pháp này trong nhiều trường hợp còn gặp nhiều khó khăn và hạn chế, đặc biệt các loài trong chi Gừng có nhiều đặc điểm hình thái khá giống nhau nên khó khăn cho việc phân loại. Trên thế giới, ngoài việc sử dụng các chỉ thị hình thái, các nhà nghiên cứu cũng đã sử dụng các chỉ thị phân tử như RFLP, RAPD, ALFP, SSR, ISSR, SCAR để phân loại, giám định một số loài gừng. Như vậy, việc kết hợp giữa chỉ thị hình thái và chỉ thị phân tử DNA sẽ làm tăng độ chính xác và nhanh chóng xác định được sự khác biệt giữa các sinh vật. DNA barcode là một phương pháp mới để xác định nhanh chóng bất kỳ loài nào dựa trên trình tự DNA được tách chiết từ một mẫu mô nhỏ của sinh vật. Như một công cụ nghiên cứu cho các nhà phân loại, DNA barcode hỗ trợ trong việc xác định các loài bằng việc mở rộng khả năng xác định loài đó qua các giai đoạn lịch sử sự sống của sinh vật. DNA barcode đang được sử dụng để giải quyết câu hỏi về sinh thái và tiến hóa cơ bản từ động vật và nấm nguyên sinh vật, tảo và thực vật. Các kết quả của nghiên cứu gần đây, cho thấy các chỉ thị phân tử DNA barcode của thực vật được lựa chọn là vùng *rbcL* và *matK* cộng với hai khu vực bổ sung *trnH-psbA* và vùng ITS đã cho thấy mức độ phù hợp cho sự phân biệt loài đối với thực vật (Paula, 2013; Jalil *et al.*, 2015; Kress *et al.*, 2015). Các nhà khoa học đã xác định được nhiều đoạn DNA đặc hiệu được sử dụng làm DNA barcode. Mỗi đoạn DNA barcode có những đặc trưng riêng và có khả năng phân biệt sinh vật ở các mức độ loài và dưới loài. Việc lựa chọn những đoạn DNA đặc trưng để làm chỉ thị phân tử DNA barcode là phụ thuộc vào từng nhóm sinh vật cụ thể và mục đích nghiên cứu (Kress, Erickson, 2008).

Hiện nay, một số sản phẩm dược liệu được làm từ nguồn cây dược liệu trồng thương mại, những sản phẩm này có thể trộn từ nhiều nguồn dược liệu khác nhau. Điều này ảnh hưởng đến việc kiểm tra chất lượng của sản phẩm

có nguồn gốc từ cây dược liệu. Trên một số đối tượng như sâm Hàn Quốc và Sâm Mỹ có đặc điểm hình thái tương tự nhau nhưng hiệu quả dược liệu khác nhau. Việc xác định hai loài này trong sản phẩm dược liệu rất khó khăn vì chúng ở dạng chế biến như Hồng Sâm, bột nhân sâm, lát cắt nhỏ, viên nén, dạng lỏng. Do đó việc xác định, truy xuất nguồn vật liệu các thương phẩm trên rất cần thiết. Có thể sử dụng các phương pháp như sắc ký lỏng cao áp, sắc ký khí, quang phổ,... để xác định. Tuy nhiên, gặp một số hạn chế bởi sự chuyển hóa các hợp chất thứ cấp trong sâm bị ảnh hưởng các yếu tố khác nhau như điều kiện sinh trưởng, quá trình sản xuất, đồng thời chi phí phân tích cao. Phương pháp dựa trên chỉ thị phân tử có nhiều ưu thế trong việc xác định, chỉ cần một lượng mẫu nhỏ và hiệu quả về thời gian, chi phí. Phương pháp này có thể áp dụng cho bất kỳ loại mô thực vật nào và sản phẩm đã qua chế biến với kết quả ổn định, nhanh chóng, tin cậy (Ali *et al.*, 2008; Jalil *et al.*, 2015; Shakya, 2015).

Hai loài *Zingiber officinale* và *Zingiber zerumbet* có đặc điểm hình thái dễ dàng phân biệt với nhau và có sự khác nhau về các chất có hoạt tính sinh học có trong chúng, tuy nhiên các sản phẩm dược liệu có nguồn gốc từ hai loài này rất khó phân biệt nếu không dựa trên phân tích di truyền. Bên cạnh đó, việc nâng cao chất lượng giống cây gừng còn hạn chế trong việc đánh giá và chọn lọc dựa trên đặc điểm di truyền do có sự biến đổi di truyền trong tự nhiên. Do đó việc phân tích đa dạng và xác định di truyền rất quan trọng trong quá trình phát triển giống gừng. Việc ứng dụng chỉ thị phân tử để đánh giá và xây dựng cơ sở dữ liệu di truyền ban đầu rất cần thiết từ đó ứng dụng cho việc xác định đánh giá, xác định các sản phẩm dược liệu có nguồn gốc từ gừng và một số loài cây khác.

Trong nghiên cứu này, chúng tôi phát triển ứng dụng chỉ thị phân tử để phân biệt sản phẩm thương mại có nguồn gốc từ gừng, bước đầu tiến hành đánh giá mối quan hệ di truyền của một số loài trong chi *Zingiber* như gừng (*Zingiber officinale*), gừng gió (*Zingiber zerumbet*), xây dựng cơ sở dữ liệu di truyền từ đó ứng dụng trong việc xác định các sản phẩm dược liệu có nguồn gốc từ gừng và đánh giá các chất có hoạt tính sinh học có trong hai loài *Zingiber officinale* và *Zingiber zerumbet*. Đồng thời, kết hợp với một số loài trong chi *Curcuma* để làm nhóm ngoài và xem mối quan hệ di truyền của một số loài trong hai chi này. Việc ứng dụng chỉ thị phân tử trong công tác xác định nguồn gốc sản phẩm dược liệu góp phần nâng cao hiệu quả sử dụng.

NGUYÊN VẬT LIỆU VÀ PHƯƠNG PHÁP

Vật liệu nghiên cứu

06 mẫu giống Gừng (*Zingiber sp.*), 02 mẫu giống nghệ (*Curcuma sp.*): thân rễ (củ) được thu thập ở các tỉnh khác nhau. Trong đó, nghệ được sử dụng làm nhóm ngoài trong nghiên cứu này.

Bảng 1. Danh sách 06 giống gừng và 02 giống nghệ

| TT | Kí hiệu | Tên khoa học | Địa điểm thu thập |
|----|---------|---------------------|--|
| 01 | G01 | <i>Zingiber sp.</i> | Huyện Cần Giuộc - tỉnh Long An |
| 02 | G02 | <i>Zingiber sp.</i> | Trung tâm Công nghệ Sinh học TP. Hồ Chí Minh - quận 12 - TP. HCM |
| 03 | G03 | <i>Zingiber sp.</i> | Huyện Nhơn Trạch - tỉnh Đồng Nai |
| 04 | G04 | <i>Zingiber sp.</i> | TP. Bảo Lộc - tỉnh Lâm Đồng |
| 05 | G05 | <i>Zingiber sp.</i> | TP. Đà Lạt - tỉnh Lâm Đồng |
| 06 | G06 | <i>Zingiber sp.</i> | Mã Pì Lèng - huyện Đồng Văn - tỉnh Hà Giang |
| 07 | N01 | <i>Curcuma sp.</i> | Mã Pì Lèng - huyện Đồng Văn - tỉnh Hà Giang |
| 08 | N02 | <i>Curcuma sp.</i> | Sông Nậm - tỉnh Hà Giang |

Ly trích DNA tổng số

Tách chiết DNA tổng số dựa trên quy trình CTAB cơ bản (Doyle, 1987) được cải tiến. Chuẩn bị 0,5 g mẫu, thân rễ (củ). Nghiền mẫu trong nitor, bổ sung 4 µl RNAse, ủ ở nhiệt độ 65°C, trong thời gian 60 phút. Ly tâm 14,000 vòng/10 phút, thu dịch nổi sang tube mới. Bổ sung 500 µl Chloroform/Isoamyl alcohol (24:1), đảo đều, ly tâm 14,000 vòng/10 phút, chuyển phần dịch nổi sang tube 1,5 mới, loại bỏ phần kết tủa, lặp lại 2 lần. Thu được dịch nổi chứa DNA tổng số. Tủa DNA ở các điều kiện ½ thể tích NaCl 5M + 1:1 thể tích ethanol 99% để 1 giờ trên đá, ly tâm 13,000 vòng/10 phút, loại bỏ dịch nổi và thu kết tủa. Rửa DNA với 500 µl ethanol 70%, đảo đều, ly tâm 13,000 vòng/5 phút, loại bỏ dịch nổi, thu kết tủa. Để khô ở nhiệt độ phòng, hòa tan kết tủa DNA bằng 100 µl TE. Bảo quản -20°C. Kiểm tra chất lượng của các DNA bằng chạy điện di gel agarose 0,8% và kiểm tra nồng độ ở bước sóng 260 nm, độ tinh sạch ở bước sóng 280 nm trên máy nanodrop (hút 1 µl DNA tổng số cho vào đầu dò của máy, tiến hành đo theo các thông số đã cài đặt trên máy).

PCR và điện di

Chọn lọc một số chỉ thị phân tử DNA barcode: *matK*, *rbcL* và ITS để đánh giá mối quan hệ di truyền của một số loài gừng thuộc chi gừng ở Việt Nam.

Thực hiện phản ứng PCR khuếch đại vùng *rbcL*, *matK*, ITS với các cặp mồi tương ứng (bảng 1). Thành phần phản ứng cho 20 µl chứa 0,2 mM dNTP mix, enzyme dreamTaq polymerase 1 U, dream taq buffer 10 X, 0,4 µM

mồi xuôi, 0,4 μ M mỗi ngược, nước cất 2 lần khử trùng, 1 μ l mẫu. Chu kỳ PCR: 1 chu kỳ tiền biến tính 95° C trong 1 phút; 30 chu kỳ: 95°C trong 1 phút; 95°C trong 15 giây; T_a trong 30 giây; 72°C trong 40 giây; 1 chu kỳ hoàn thành 72°C trong 10 phút. Sản phẩm PCR được phát hiện bởi điện di trên gel agarose 1,2%, nhuộm với ethidium bromide và chụp hình trên máy geldoc.

Bảng 2. Danh sách các cặp mồi sử dụng trong nghiên cứu

| STT | Tên mồi | Trình tự mồi (5' – 3') | Nhiệt độ bắt cặp (T _a) (°C) | Kích thước dự kiến (bp) |
|-----|--------------------|----------------------------|---|-------------------------|
| 1 | ITS-1F | ATGCGATACTTGGTGTGAAT | 53 | 700 |
| | ITS-1R | TCCTCCGCTTATTGATATGC | | |
| 2 | ITS-1F | TCCGTAGGTGAACCTGCGG | 55 | |
| | ITS-4R | TCCTCCGCTTATTGATATGC | | |
| 3 | <i>matK</i> 390-F | CGATCTATTCAATCAATATTC | 50 | 900 |
| | <i>matK</i> 1326-R | TCTAGCACACGAAAGTCGAAG | | |
| 4 | <i>rbcLa</i> -F | ATGTCACCACAAACAGAGACTAAAGC | 55 | 600 |
| | <i>rbcLa</i> -F | GTAAAATCAAGTCCACCRGC | | |

(Asahina et al., 2010, Takamiya et al., 2011)

Phân tích và đánh giá mối quan hệ di truyền dựa trên trình tự DNA vùng gen của một số loài thuộc chi gừng

Sản phẩm PCR vùng *matK* được chọn tiến hành giải trình tự bằng máy AB3500 tại trung tâm Công nghệ Sinh học Thành phố Hồ Chí Minh.

Hiệu chỉnh trình tự bằng phần mềm ATGC, SeaView. Sau khi hiệu chỉnh các trình tự và kiểm tra các sai lệch. Các trình tự được kiểm tra bằng công cụ BLAST (Basic Local Alignment Search Tool) để tìm các trình tự tương đồng có sẵn trên cơ sở dữ liệu Genbank. Sử dụng phần mềm MEGA (The Molecular Evolution Genetics Analysis) để xây dựng cây phát sinh loài với hệ số boottrap 1000.

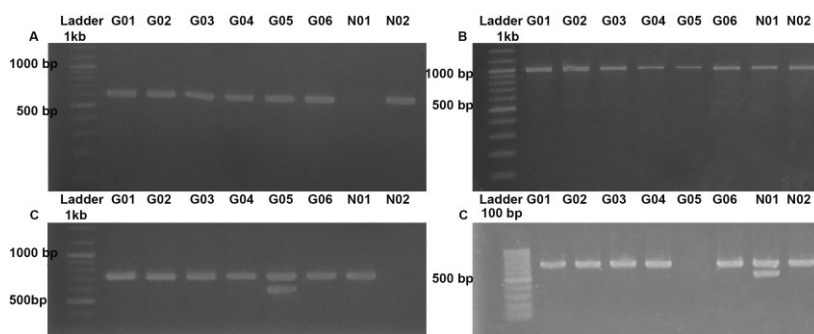
KẾT QUẢ VÀ THẢO LUẬN

Ly trích DNA tổng số

DNA tổng số sau khi tách chiết được định lượng và kiểm tra độ tinh sạch bằng máy Nano Drop cho kết quả tỉ lệ OD₂₆₀/OD₂₈₀ của 08 mẫu nằm trong khoảng 1,7 - 2,0 cho thấy DNA tổng số thu được là tinh sạch. Nồng độ DNA tổng số ở các mẫu nằm trong khoảng 121 - 189 ng/ μ l đảm bảo cho phép sử dụng cho các bước tiếp theo.

PCR và điện di

Kết quả khuếch đại các vùng DNA barcode như vùng *rbcL*, *matK*, ITS với các mồi tương ứng *rbcLa*-F và *rbcLa*-R, *matK*-390F và *matK*-1326R, ITS-1F và ITS-1R, ITS-1F và ITS-4R như sau:



Hình 1. Kết quả điện di sản phẩm PCR: A/ vùng *rbcL*, B/ vùng *matK*, C/ vùng ITS của 06 mẫu gừng và 02 mẫu nghệ trên gel agarose 1,2%

Kết quả khuếch đại vùng *matK* được kiểm tra trên gel garose 1,2%, cho thấy tỉ lệ khuếch đại đạt 100% và kích thước vùng gen mục tiêu khoảng 1000 bp. Vùng *rbcL* được khuếch đại với tỉ lệ khuếch đại 87,5%, các mẫu xuất hiện vạch có kích thước khoảng 600 bp. Đối với vùng ITS khuếch đại với cặp mồi ITS 1-F - 1R, tỉ lệ khuếch đại là 87,5%, sản phẩm khuếch đại có kích thước khoảng 700 bp - 800 bp, trong đó mẫu G05 xuất hiện 2 vạch sản phẩm. Vùng ITS khuếch đại với cặp mồi ITS 1F - 4R cho thấy tỉ lệ khuếch đại là 87,5%. Sản phẩm xuất hiện vạch có kích thước khoảng 700 bp - 800 bp, trong đó mẫu N01 xuất hiện 2 vạch sản phẩm. Vùng *matK* nằm trong vùng intron của vùng *trnK* trên lục lạp, chỉ có một vùng *matK* với kích thước từ 600 - 1500 bp được sử dụng cho

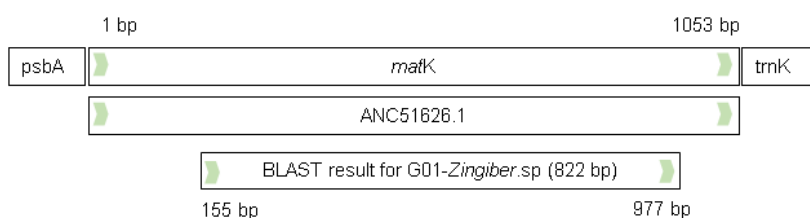
mã vạch DNA và có mặt hầu hết trong thực vật, có tính phổ quát nên được sử dụng như một chỉ thị trong nghiên cứu mối quan hệ giữa các loài. Vùng *rbcl* là vùng gen nằm trong lục lạp có thể nhân bản một cách dễ dàng ở nhiều loài thực vật. Ở các loài khác nhau kích thước của gene này cũng khác nhau (Paula, 2013). Vùng DNA ribosomal trong nhân (ITS1, 5.8S, ITS2) được coi là một trong những mã vạch DNA dễ dàng của khuếch đại, và có tính chuyên biệt đủ để phân biệt các loài. Kết quả trên cho thấy, tỷ lệ khuếch đại từng vùng gen với các cặp cặp mỗi khác nhau và trong cùng một gen nhưng sử dụng cặp cặp mỗi khác nhau có tỷ lệ khuếch đại khác nhau. Đối với cặp cặp mỗi ITS1-F và ITS1-R không khuếch đại được mẫu N02, cặp mỗi ITS1-F và ITS4-R không khuếch đại được ở mẫu G05. Trong số các vùng gen trên lục lạp, *matK* phát triển nhanh chóng và biến đổi cao, đã được nhóm nghiên cứu thực vật của Hiệp hội mã vạch sự sống (Consortium for the Barcode of Life - CBOL) đề xuất như là một locus cho DNA barcode. Khuếch đại và sắp xếp của vùng *matK* là khó khăn do biến đổi cao trình tự trong các vị trí liên kết cặp mỗi (Group *et al.*, 2009). Từ kết quả tỷ lệ khuếch đại trên có thể cho thấy vùng *matK* là một ứng cử viên tốt trong việc đánh giá mối quan hệ di truyền cho các mẫu cây thuộc họ Zingiberaceae. Nghiên cứu trước đây cũng cho thấy khả năng xác định giữa các vùng gen khác nhau trên một số loài trong họ Zingiberaceae có sự khác nhau rõ rệt. Vùng gen *rbcl* chỉ có khả năng phân biệt giữa các mẫu trong cùng một loài là 0,01% và khác loài là 0,9% trong khi đó vùng gen *matK* cho khả năng phân biệt trong cùng loài là 0,06% và khác loài là 2,4% (Vinitha *et al.*, 2014). Trong nghiên cứu này, vùng *matK* được sử dụng để giải trình tự và phân tích các biến thể và phát sinh loài.

Phân tích và đánh giá mối quan hệ di truyền dựa trên trình tự DNA vùng gen của một số loài thuộc chi gừng

Sản phẩm PCR của vùng *matK* khuếch đại bởi cặp mỗi *matK*-390F và *matK*-1326R của 06 mẫu gừng và 02 mẫu nghệ được giải trình tự DNA bằng máy AB3500. Trình tự DNA thu được hiệu chỉnh bằng phần mềm ATGC và Seaview và được BLAST trên ngân hàng NCBI để đánh giá sự tương đồng giữa các mẫu nghiên cứu, xây dựng cây phát sinh loài bằng phần mềm MEGA 7.0 (The Molecular Evolution Genetics Analysis) dựa trên thuật toán Construct/Test Maximum Likelihood tree (ML) với hệ số bootstrap 1000. Kết quả phân tích được trình bày ở bảng 3, bảng 4, hình 2, hình 3.

Bảng 3. Kết quả BLAST trình tự vùng *matK* của 6 mẫu gừng và 2 mẫu nghệ trong nghiên cứu với cơ sở Genbank

| Mẫu | Trình tự tham chiếu | Độ bao phủ (%) | Độ tương quan (%) |
|-----|--|----------------|-------------------|
| G01 | KU145125.1: <i>Zingiber officinale</i> | 100% | 97,93% |
| G02 | HM367680.1: <i>Zingiber zerumbet</i> | 99% | 100% |
| G03 | KU145125.1: <i>Zingiber officinale</i> | 99% | 99% |
| G04 | KU145125.1: <i>Zingiber officinale</i> | 99% | 99% |
| G05 | KU145125.1: <i>Zingiber officinale</i> | 99% | 99% |
| G06 | KU145125.1: <i>Zingiber officinale</i> | 99% | 99% |
| N01 | KX832044.1: <i>Curcuma longa</i> | 100% | 97% |
| N02 | KF601574.1: <i>Curcuma roscoeana</i> | 99% | 99% |



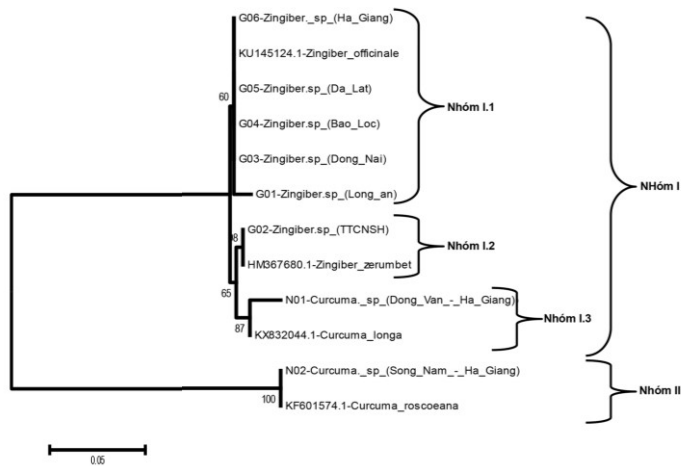
Hình 2. Cấu trúc vùng gen *matK* của mẫu gừng G1 khi phân tích trên NCBI bằng công cụ BLAST

Phân tích trình tự DNA vùng *matK* của 6 mẫu gừng và 2 mẫu nghệ thu thập với thông tin có sẵn trên NCBI bằng công cụ BLAST cho thấy mức độ bao phủ đạt 99-100% và mức độ tương đồng cao 97-100%. Điều này cho thấy hệ gene lục lạp ở gừng có tính bảo tồn cao giữa các cá thể ở các khu vực cách xa nhau về mặt địa lý. Hơn nữa, theo bản chất của hệ gene lục lạp thì đây là hệ gene ít có những biến động giữa các cá thể thuộc các loài gần nhau. Kích thước vùng gen *matK* của chi gừng nằm từ 1 – 1053 bp, khi so sánh trình tự gen vùng *matK* của mẫu gừng G01 trên Genbank với trình tự tham chiếu có mã ANC51626.1 cho thấy vị trí của trình tự nghiên cứu nằm từ vị trí 155 bp đến vị trí 977 bp.

Kết quả phân tích trình tự DNA vùng gen *matK* của 6 mẫu gừng và 2 mẫu nghệ bằng phần mềm MEGA 7.0 cho thấy có 29/926 vị trí biến đổi và 879/926 vị trí bảo tồn, chứng tỏ vùng *matK* của chi gừng có tính bảo tồn cao. Ở các vị trí sai khác giữa các mẫu gừng cho thấy mẫu gừng G02 có 10 vị trí sai khác so với các mẫu gừng còn lại, mẫu gừng G01 có 5 vị trí sai khác so với các mẫu còn lại. Bốn mẫu gừng G03, G04, G05, G06 chỉ sai khác ở hai vị trí là 924 và 925. Đối với hai mẫu nghệ cho thấy có sự sai khác ở 18 vị trí.

Bảng 4. Vị trí biến đổi các nu của 8 trình tự nghiên cứu

| TT | Vị trí | Mẫu | | | | | | | | TT | Vị trí | Mẫu | | | | | | | |
|----|--------|-----|-----|-----|-----|-----|-----|-----|-----|----|--------|-----|-----|-----|-----|-----|-----|-----|-----|
| | | G01 | G02 | G03 | G04 | G05 | G06 | N01 | N02 | | | G01 | G02 | G03 | G04 | G05 | G06 | N01 | N02 |
| 1 | 89 | T | G | T | T | T | T | T | T | 16 | 302 | T | C | T | T | T | C | C | |
| 2 | 99 | A | G | A | A | A | A | G | G | 17 | 396 | A | G | A | A | A | A | G | |
| 3 | 113 | T | T | T | T | T | T | T | C | 18 | 472 | G | A | G | G | G | G | G | |
| 4 | 134 | A | A | A | A | A | A | A | T | 19 | 476 | C | T | C | C | C | C | C | |
| 5 | 135 | C | C | C | C | C | C | C | T | 20 | 522 | A | A | A | A | A | A | T | |
| 6 | 136 | A | A | A | A | A | A | A | T | 21 | 524 | C | C | C | C | C | C | G | |
| 7 | 137 | T | T | T | T | T | T | T | C | 22 | 685 | A | A | A | A | A | A | G | |
| 8 | 140 | A | A | A | A | A | A | A | T | 23 | 715 | T | T | T | T | T | C | C | |
| 9 | 144 | C | C | C | C | C | C | C | T | 24 | 724 | G | G | G | G | G | A | A | |
| 10 | 145 | A | A | A | A | A | A | A | G | 25 | 750 | T | T | T | T | T | C | C | |
| 11 | 153 | T | C | C | C | C | C | C | T | 26 | 835 | C | G | G | G | G | G | G | |
| 12 | 184 | A | A | A | A | A | C | C | A | 27 | 836 | C | A | A | A | A | A | A | |
| 13 | 230 | G | A | G | G | G | G | G | G | 28 | 924 | - | - | C | C | T | - | - | |
| 14 | 231 | A | G | A | A | A | A | A | G | 29 | 925 | - | - | T | T | C | - | - | |
| 15 | 238 | A | T | T | T | T | T | T | T | | | | | | | | | | |



Hình 3. Cây phát sinh loài dựa trên trình tự DNA vùng *matK* của 6 mẫu gừng, 2 mẫu nghệ và 4 trình tự tham chiếu trên NCBI

Cây phát sinh loài xây dựng dựa trên vùng gen *matK* cho các mẫu nghiên cứu với hệ số tương đồng di truyền là 0,81 chúng tôi chia làm 2 nhóm chính: nhóm gừng với hệ số tương đồng di truyền là 0,71 và nhóm nghệ với hệ số tương đồng di truyền là 0,69; điều này cho thấy giữa các mẫu trong mỗi nhóm trên có sự khác nhau về mặt di truyền. Nhóm gừng được chi làm 3 nhóm nhỏ là nhóm I.1, I.2 và I.3, trong đó nhóm I.3 là mẫu nghệ N02 với hệ số tương đồng di truyền là 0,87. Khi so sánh sự sai khác ở 29 vị trí sai khác của 2 mẫu nghệ và 6 mẫu gừng cho thấy mẫu nghệ N01 có sự khác biệt về mặt di truyền so với các mẫu gừng. Nhóm I.1 gồm các mẫu gừng thu thập tại Đà Lạt, Bảo Lộc, Lâm Đồng, Long An và Hà Giang và nằm cùng với *Zingiber officinale* với hệ số tương đồng di truyền là 0,97. Nhóm I.2 gồm mẫu gừng thu thập tại Tp Hồ Chí Minh và nằm cùng với *Zingiber zerumbet* với hệ số tương đồng di truyền là 0,99. Nhóm II gồm mẫu nghệ N02 được thu thập tại Hà Giang và nằm cùng với *Curcuma roscoeana* với hệ số tương đồng di truyền là 0,94. *Zingiber zerumbet* là một loài thuộc chi *Zingiber* nhưng không có mối quan hệ gần gũi đối với loài *Zingiber officinale* nên được nằm ở 1 nhóm riêng trong chi *Zingiber*. Từ kết quả trên cho thấy sử dụng trình tự DNA vùng gen *matK* có hiệu quả trong việc đánh giá mối quan hệ di truyền và xác định loài trên mẫu gừng và nghệ.

Kết quả phân tích trình tự DNA của các mẫu gừng từ loài *Zingiber officinale* và *Zingiber zerumbet* có sự khác nhau ở các vị trí 89, 99, 153, 230, 231, 238, 302, 396, 472, 476, 835 và 836. Khi phân tích các mẫu trong cùng loài *Zingiber officinale* cho thấy giữa các mẫu có sự tương đồng cao, tuy nhiên đối với mẫu G02 có sự sai khác trình tự DNA ở vị trí 238 và 838 so với các mẫu G03, G04, G05. Kết quả này cho thấy mặc dù các mẫu G01-05 đều thuộc loài *Zingiber officinale* nhưng có sự khác biệt nhất định giữa giống G02 và các giống còn lại. Bên cạnh đó, mặc dù các mẫu giống gừng thu thập từ nhiều vùng địa lý khác nhau nhưng lại có sự giống nhau ở một mức độ nhất định về mặt di truyền. Các kết quả này cho thấy việc thu thập và đánh giá di truyền các giống trồng/thương mại dựa trên chỉ thị phân tử DNA barcode có thể giúp nhận diện mẫu ở mức độ loài, đồng thời cũng có thể giúp nhận diện ở mức độ dưới loài trong một số trường hợp. Ngoài ra, kết quả nghiên cứu cũng cho thấy các giống gừng trồng/thương mại thu thập có thể có nguồn gốc khác nhau và vị trí địa lý không phải là yếu tố quyết định đến đặc điểm di truyền của một giống mà sự di thực hay mục đích khai thác thương mại của giống có thể làm ảnh hưởng đến sự phân bố của loài/giống đó trong tự nhiên. Từ kết quả nghiên cứu di truyền ban đầu này, việc nghiên cứu, đánh giá bằng các chỉ thị phân tử khác như SSR, SNP, các hoạt chất có hoạt tính sinh học giữa các mẫu giống thu thập có thể giúp hiểu rõ hơn về sự phân bố cũng như khả năng khai thác dược chất kết hợp bảo tồn và phát triển các giống gừng ở Việt Nam.

KẾT LUẬN

Khuếch đại thành công các vùng *matK*, *rbcL*, ITS với 4 cặp cặp mỗi tương ứng *matK*390-F và *matK*1326-R, *rbcLa*-F và *rbcLa*-R, ITS1F và ITS 1R, ITS-1F và ITS-1R với tỉ lệ khuếch đại lần lượt là 100%, 87,5%, 87,5%, 87,6%. Đánh giá và phân tích được mối quan hệ di truyền giữa các loài trong chi gừng với chỉ số tương đồng di truyền khoảng 97 - 100% và chỉ số bao phủ từ 99 - 100%. Kết quả phân loại di truyền và định danh cho thấy mẫu gừng G01, G03, G04, G05, G06 thuộc loài *Zingiber officinale* và mẫu G02 thuộc loài *Zingiber zerumbet*. Mẫu nghệ N01 thuộc loài *Curcuma longa* và mẫu N02 thuộc loài *Curcuma roscoeana*. Kết quả này có thể được ứng dụng trong nhận diện cây dược liệu, các sản phẩm dược liệu có nguồn gốc thuộc họ gừng.

Lời cảm ơn: Các tác giả xin chân thành cảm ơn Trung tâm Công nghệ Sinh học Thành phố Hồ Chí Minh đã hỗ trợ một phần kinh phí và cơ sở vật chất để thực hiện nghiên cứu này.

TÀI LIỆU THAM KHẢO

- Ali BH, Blunden G, Tanira MO, Nemmar A (2008). Some phytochemical, pharmacological and toxicological properties of ginger (*Zingiber officinale* Roscoe): a review of recent research. *Food Chem Toxicol* 46: 409-420.
- Asahina H, Shinozaki J, Masuda K, Morimitsu Y, Satake M (2010). Identification of medicinal *Dendrobium* species by phylogenetic analyses using *matK* and *rbcL* sequences. *J Nat Med* 64: 133-138.
- Doyle JJ (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19: 11-15.
- Group CPW, Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, Chase MW, Cowan RS, Erickson DL (2009). A DNA barcode for land plants. *Proc Acad Sci USA* 106: 12794-12797.
- Jailil M, Annuar MSM, Tan BC, Khalid N (2015). Effects of selected physicochemical parameters on zerumbone production of *Zingiber zerumbet* Smith cell suspension culture. *Evid Based Complement Alternat Med* 2015: 1-7.
- Kress WJ, Erickson DL (2008). DNA barcoding: A windfall for tropical biology?. *Biotropica* 40: 405-408.
- Kress WJ, García-Robledo C, Uriarte M, Erickson DL (2015). DNA barcodes for ecology, evolution, and conservation. *Trends Ecol Evol* 30: 25-35.
- Paula S (2013). A comparison of the efficiency of DNA barcode regions in a small and a large genus. *PhD dissertation, Department of Genetics, University of Free State*.
- Phạm Hoàng Hộ (1993). *Cây cỏ Việt Nam*, tập 1. NXB Trẻ Thành phố Hồ Chí Minh.
- Shakya SR (2015). Medicinal uses of ginger (*Zingiber officinale* Roscoe) improves growth and enhances immunity in aquaculture. *Int J Chem Stud* 3: 83-87.
- Takamiya T, Wongsawad P, Tajima N, Shioda N, Lu JF, Wen CL, Wu JB, Handa T, Iijima H, Kitanaka S (2011). Identification of *Dendrobium* species used for herbal medicines based on ribosomal DNA internal transcribed spacer sequence. *Biol Pharm Bull* 34: 779-782.
- Vinitha MR, Kumar US, Aishwarya K, Sabu M, Thomas G (2014). Prospects for discriminating Zingiberaceae species in India using DNA barcodes. *J Integr Plant Biol* 56: 760-773.

ASSESSMENT OF GENETIC RELATIONSHIP OF GINGERS (*Zingiber* spp.) IN VIETNAM BASED ON DNA BARCODE MARKERS

Nguyen Truong Giang, Mai Thi Kim Ngoc, Phan Quang Huong, Huynh Huu Duc*

Biotechnology Center of Ho Chi Minh City

SUMMARY

Currently, many studies on the classification of *Zingiber* in Vietnam were based primarily on morphological markers. These morphological methods need to have sufficient data on the flower structure which in many cases are very difficult and limited. These methods also were very difficult to apply in the ginger species because they have many similar morphological characteristics and are difficult to flower, so difficult to classify. In this study, we evaluated 6 ginger samples and 2 turmeric samples is out group to evaluate the genetic relationship by DNA barcode marker based on *matK* region, *rbcL* on cDNA and ITS in the nucleus. The results showed that the ratio of amplified regions of *matK*, *rbcL*, ITS corresponding 4 primer pairs (*matK*-390F and *matK*-1326R, *rbcLa*-F and *rbcLa*-R, ITS-1F and ITS-1R, ITS-1F and ITS-1R) is 100%, 87.5%, 87.5%, 87.6% respectively. The genetic homology among *Zingiber* with reference sequence of 97-100% and coverage rates of 99 - 100%. Based on the phylogenetic tree, the ginger and turmeric species were divided into two different groups and *Zingiber* is divided into two main species, including *Zingiber zerumbet* and *Zingiber officinale*. The results lead to the ability to relationship analysis and identification among Zingiberaceae species by DNA barcode marker and offer opportunities for identification medicinal produce had source from *Zingiber*.

Keywords: DNA barcode, molecular marker, *Zingiber*.

* Author for correspondence: Tel: +84-967.137.046, Email: huuduchuyh82@gmail.com