

Nghiên cứu phát triển chỉ thị phân tử SSR đặc hiệu họ gen NBS-LRR cho chanh leo *Passiflora edulis* Sims

Trần Đức Trung^{1*}, Nguyễn Thị Thu Hằng¹, Đào Thu Hương², Tạ Hồng Linh³, Nguyễn Thị Lan Hoa¹,
Phạm Thanh Huyền¹, Trần Thị Huệ Hương³, Nguyễn Hoàng¹

¹Viện Dược liệu, 3B Quang Trung, phường Trảng Tiền, quận Hoàn Kiếm, Hà Nội, Việt Nam

²Viện Ứng dụng Công nghệ Tế bào Mesocells, 204 Nguyễn Tuân, phường Nhân Chính, quận Thanh Xuân, Hà Nội, Việt Nam

³Viện Khoa học Nông nghiệp Việt Nam, xã Vĩnh Quỳnh, huyện Thanh Trì, Hà Nội, Việt Nam

Ngày nhận bài 7/11/2023; ngày chuyển phản biện 10/11/2023; ngày nhận phản biện 1/12/2023; ngày chấp nhận đăng 4/12/2023

Tóm tắt:

Thành tựu của công nghệ giải trình tự gen đã tạo nguồn dữ liệu phong phú và giá trị cho các phân tích tin sinh học nhằm phát triển chỉ thị phân tử phục vụ nghiên cứu di truyền và chọn tạo giống ở các loài cây trồng, trong đó có cây chanh leo (*Passiflora edulis* Sims). Trong nghiên cứu này, dựa trên dữ liệu trình tự họ gen kháng NBS-LRR đã được xác định từ hệ gen chanh leo, phân tích *in-silico* lựa chọn và thiết kế chỉ thị SSR đặc hiệu cho các locus microsatellite tiềm năng cũng như phân tích thực nghiệm kiểm chứng tính đa hình của các chỉ thị SSR đã được triển khai. Tổng số 497 locus microsatellite đã được xác định từ trình tự 161 gen NBS-LRR, từ đó 383 chỉ thị SSR đặc hiệu họ gen NBS-LRR được thiết kế cho chanh leo. Qua sàng lọc thực nghiệm 46 chỉ thị SSR đặc hiệu cho 35 gen NBS-LRR tiềm năng có biểu hiện gen liên quan đến các stress sinh học và phi sinh học đã xác định được 25 chỉ thị tạo băng alen ổn định và đa hình giữa các mẫu giống chanh leo khi phân tích PCR-PAGE. Các chỉ thị SSR đặc hiệu gen kháng NBS-LRR mới này là công cụ bổ sung cho các nghiên cứu đánh giá nguồn gen, lập bản đồ di truyền và chọn tạo giống chanh leo kháng bệnh bằng chỉ thị phân tử.

Từ khóa: chỉ thị SSR, họ gen NBS-LRR, *Passiflora edulis* Sims, PCR-PAGE, phân tích *in silico*.

Chỉ số phân loại: 1.6

1. Đặt vấn đề

Chanh leo (*Passiflora edulis* Sims) là loại cây ăn quả có giá trị kinh tế cao, xuất xứ từ Nam Mỹ và hiện đã được trồng phổ biến ở các vùng nhiệt đới và á nhiệt đới [1]. Tại Việt Nam, cả hai dạng chanh leo quả tím (*P. edulis* f. *edulis*) và quả vàng (*P. edulis* f. *flavicarpa* Deg.) cũng như dạng lai giữa chúng hiện đang được phát triển sản xuất tại nhiều địa phương như các tỉnh Tây Bắc, Nghệ An và Tây Nguyên. Ngoài giá trị sử dụng là cây ăn trái (ăn tươi hoặc chế biến nước quả), quả và lá chanh leo chứa một số hợp chất tự nhiên có lợi cho sức khỏe, do đó còn được sử dụng làm nguyên liệu sản xuất các sản phẩm chăm sóc sức khỏe [2]. Mặc dù có tiềm năng lớn về chuỗi sản phẩm cũng như nhu cầu thị trường nhưng sản xuất chanh leo còn bị hạn chế bởi tính mẫn cảm với nhiều loại bệnh hại, đặc biệt do virus và vi khuẩn gây ra [3]. Cũng giống như các loại cây trồng khác, việc phát triển các giống chanh leo mang gen quy định tính kháng bệnh hại được xem là giải pháp căn cơ nhất để phát triển sản xuất bền vững. Tuy nhiên, nền tảng di truyền về gen kháng bệnh và các chỉ thị đặc hiệu gen kháng ở chanh leo hiện vẫn còn hạn chế.

Tính kháng bệnh hại ở thực vật được quy định bởi các gen kháng R, phần lớn trong đó là các gen thuộc họ NBS-

LRR (Nucleotide binding site - Leucine rich repeats) [4]. Họ gen NBS-LRR ở các loài cây trồng có thể được xác định và phân lập thông qua nghiên cứu lập bản đồ gen [5], giải trình tự chọn lọc vùng gen kháng RenSeq [6] và phổ biến là ứng dụng tin sinh học trong phân tích dữ liệu hệ gen [7]. Gần đây, trình tự hệ gen được giải mã ở cấp độ nhiễm sắc thể [8] đã tạo cơ sở cho phân tích *in-silico* xác định 161 gen hoàn chỉnh thuộc họ gen NBS-LRR ở chanh leo (dữ liệu chưa công bố).

Chỉ thị phân tử là công cụ phổ biến trong các nghiên cứu đánh giá đa dạng di truyền tính kháng và lập bản đồ phân tử phục vụ công tác chọn tạo giống kháng. Dựa trên nền tảng phản ứng nhân gen PCR và khả năng phân biệt các alen của một locus cụ thể, các loại chỉ thị phân tử có thể phân thành nhóm các chỉ thị trội (bao gồm RAPD, ISSR, SRAP, một số chỉ thị dựa trên gen nhảy...) và nhóm các chỉ thị đồng trội (bao gồm SSR, InDel, SNP...) [9]. Sự phát triển vượt bậc của công nghệ giải trình tự hiện nay đã tạo ra lượng dữ liệu phong phú về thông tin hệ gen các loài thực vật, tạo điều kiện cho các nghiên cứu xác định các microsatellite và phát triển, ứng dụng chỉ thị SSR. Dựa trên biến động trình tự của locus microsatellite - bản chất là các đột biến thay đổi số lượng motif lặp đơn (SSR) trong quá trình nhân bản và tái tổ

*Tác giả liên hệ: Email: ductrung83@gmail.com

Development of NBS-LRR gene-specific microsatellite markers for passion fruit *Passiflora edulis* Sims

Duc Trung Tran^{1*}, Thi Thu Hang Nguyen¹,
Thu Huong Dao², Hong Linh Ta³, Thi Lan Hoa Nguyen¹,
Thanh Huyen Pham¹, Thi Hue Huong Tran³, Hoang Nguyen¹

¹National Institute of Medicinal Materials,
3B Quang Trung Street, Trang Tien Ward, Hoan Kiem District, Hanoi, Vietnam

²Mesocells Cell Technology Application Research Institute,
204 Nguyen Tuan Street, Nhan Chinh Ward, Thanh Xuan District, Hanoi, Vietnam

³Vietnam Academy of Agricultural Sciences,
Vinh Quynh Ward, Thanh Tri District, Hanoi, Vietnam

Received 7 November 2023; revised 1 December 2023; accepted 4 December 2023

Abstract:

The chromosome-level genome assembly of passion fruit (*Passiflora edulis* Sims) has facilitated computational analyses in developing novel informative SSR markers for genetics and breeding research. In this study, *in-silico* analysis was conducted to discover potential microsatellite loci and to design specific primers from the sequences of the characterised nucleotide-binding site - leucine-rich repeats (NBS-LRR) gene family of passion fruit, and a genotyping experiment was performed to validate the novel simple sequence repeat (SSR) markers. In total, 497 microsatellite loci were detected from 161 passion fruit NBS-LRR genes, of which 383 loci were qualified for primer generation. Forty-six SSR markers were selected for polymorphism verification, targeting 35 NBS-LRR genes that showed potential gene expression in response to biotic and abiotic stresses. From that, 25 SSR markers showed consistent polymorphic alleles among different passion fruit genotypes *via* PCR-PAGE analysis. These novel NBS-LRR-specific SSR markers might provide valuable tools for fortifying research on germplasm evaluation, genetic mapping, and molecular breeding for disease-resistant passion fruit varieties.

Keywords: *in silico* analysis, NBS-LRR gene family, *Passiflora edulis*, PCR-PAGE, SSR marker.

Classification number: 1.6

hợp trong hệ gen giữa các cá thể [10], chỉ thị SSR có nhiều ưu điểm nổi trội so với các loại chỉ thị khác nhờ số lượng lớn và phân bố đồng đều trong hệ gen, tính đặc hiệu và đa hình cao, tính ổn định và khả năng tái lập cao trong thực hành và không yêu cầu phức tạp về trang thiết bị phân tích. Một số chỉ thị phân tử SSR đã được phát triển cho dạng chanh leo quả vàng [11, 12]. Tuy nhiên, số lượng các chỉ thị này còn khiêm tốn so với kích thước hệ gen chanh leo và đặc biệt chưa có chỉ thị đặc hiệu họ gen NBS-LRR được phát triển cho dạng chanh leo quả tím phổ biến ở nước ta. Nhằm thiết lập bộ công cụ phân tử cho nghiên cứu đánh giá nguồn gen và chọn tạo giống kháng bệnh, nghiên cứu này sử dụng trình tự họ gen NBS-LRR xác định được từ hệ gen chanh leo tím để phát triển bộ chỉ thị SSR mới và kiểm chứng tính đa hình của các chỉ thị này trên một số mẫu giống chanh leo.

2. Vật liệu và phương pháp nghiên cứu

2.1. Vật liệu

Mẫu lá non từ 10 mẫu giống chanh leo được thu thập từ vườn lưu giữ mẫu giống của Viện Nghiên cứu Chanh leo Nafoods (Quê Phong, Nghệ An) (bảng 1).

Bảng 1. Danh sách 10 mẫu giống chanh leo sử dụng cho nghiên cứu.

Thứ tự	Ký hiệu mẫu giống	Thứ tự	Ký hiệu mẫu giống
1	CKL	6	CĐĐ
2	H4.115	7	ĐN1
3	TC(14-5)-51	8	TC(14-5)-45
4	TC(14-5)-7	9	H5-27
5	H14.7	10	(14-11)(16-6).20

Trình tự gen hoàn chỉnh bao gồm 2.000 bp trình tự ngoại vi từ đầu 5' (tính từ vị trí bắt đầu phiên mã TSS - Transcription start site) và đầu 3' (tính từ vị trí poly-A) của 161 gen NBS-LRR xác định từ hệ gen chanh leo tím (dữ liệu chưa công bố).

2.2. Phương pháp nghiên cứu

Xác định các locus microsatellite: Trình tự DNA của 161 gen NBS-LRR được sử dụng làm tham chiếu để phân tích tìm kiếm các microsatellite bằng công cụ MISA-web (<https://webblast.ipk-gatersleben.de/misa>) [13]. Tiêu chí để xác định các microsatellite là chiều dài motif 1, 2, 3, 4, 5 và 6 nucleotide với số lần motif lặp liên tục tương ứng là 10, 6, 5, 4, 4 và 4. Đối với các microsatellite phức hợp, các microsatellite thành phần không cách xa nhau quá 100 bp.

Kiểm tra locus microsatellite và thiết kế chỉ thị đặc hiệu: Vị trí các locus microsatellite được chú giải trên trình tự 161 gen NBS-LRR của chanh leo để kiểm chứng trước khi thiết kế các bộ mồi bằng công cụ Primer 3 tích hợp trên phần mềm Geneious v11. Các thông số thiết kế mồi bao

gồm: chiều dài trình tự 18-23 nucleotide (tối ưu là 20), hàm lượng GC 40-60%, nhiệt độ gắn mỗi tối ưu 55°C, sản phẩm nhân gen PCR dự kiến 150-300 bp. Các cặp mồi SSR đáp ứng yêu cầu được lựa chọn và tổng hợp tại Phù Sa Genomic.

Tách chiết DNA tổng số: Các mẫu lá non sau khi thu thập được xử lý, tách chiết DNA và đánh giá độ nguyên vẹn, độ tinh sạch và nồng độ theo quy trình đã áp dụng cho chanh leo [14].

Phản ứng nhân gen PCR: Phản ứng khuếch đại PCR sử dụng kit GoTaq Green Master Mix (Promega) được chuẩn bị theo hướng dẫn của nhà cung cấp với 0,5 µM mỗi mỗi loại, 25 ng DNA tổng số trong tổng thể tích 20 µl. Phản ứng khuếch đại PCR thực hiện trên máy luân nhiệt Mastercycler X50s (Eppendorf) sử dụng chu trình nhiệt bao gồm 1 chu kỳ biến tính ban đầu ở 94°C trong 5 phút, 35 chu kỳ gồm các bước biến tính ở 94°C trong 30 giây, bắt cặp mồi - khuôn ở nhiệt độ tối ưu cho từng mồi trong 30 giây, tổng hợp chuỗi ở 72°C trong 30 giây, và 1 chu kỳ tổng hợp kết thúc ở 72°C trong 4 phút. Đối với một số chỉ thị SSR có tính đặc hiệu thấp (thể hiện qua băng sản phẩm khuếch đại không đặc hiệu), chu trình phản ứng Touchdown-PCR được áp dụng theo hướng dẫn của D.J. Korbie và cs (2008) [15].

Phân tích kết quả: Sản phẩm phản ứng khuếch đại PCR được phân giải trên gel agarose 1,5% trong đệm TBE 1X ở hiệu điện thế 7V/cm trong 1 giờ (xác định nhiệt độ gắn mỗi tối ưu T_a) hoặc polyacrylamide không biến tính 8% trong đệm TAE 1X ở hiệu điện thế 12 V/cm trong 3 giờ (đánh giá tính đa hình). Bản gel sau đó được nhuộm ethidium bromide (0,5 µg/ml) trước khi chụp ảnh dưới ánh sáng UV bằng hệ thống UVP Transilluminator PLUS (Analytik Jena). Các alen được ghi nhận dựa trên kích thước thang DNA chuẩn 50 bp DNA Ladder (Thermo Scientific) hoặc ΦX174 DNA/HinfI Marker (Thermo Scientific).

2.3. Thời gian và địa điểm nghiên cứu

Nghiên cứu được triển khai từ tháng 6/2022 đến tháng 9/2023 tại Viện Dược liệu và Viện Khoa học Nông nghiệp Việt Nam (Hà Nội).

3. Kết quả

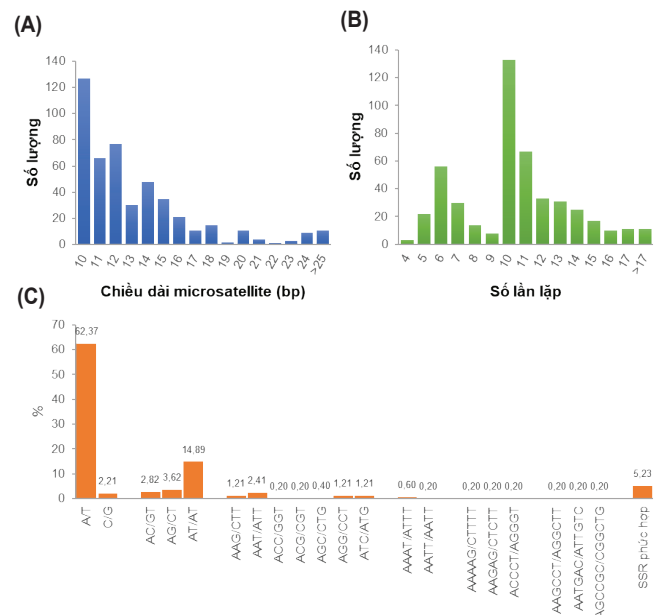
3.1. Phân bố microsatellite trên trình tự các gen họ NBS-LRR ở chanh leo

Tổng số 497 microsatellite đã được phát hiện từ 139/161 gen NBS-LRR ở chanh leo bằng công cụ MISA-web, số lượng trung bình 3,08 microsatellite/gen và mật độ trung bình 0,23 microsatellite/kb. Trong số này, 114 gen NBS-LRR (70,8%) có nhiều hơn một microsatellite (bảng 2). Trong số các microsatellite đã xác định, các motif có độ

dài 1 (mono-), 2 (di-), và 3 (tri-nucleotide) chiếm tỷ lệ nhiều nhất, lần lượt tương ứng 64,59, 21,33 và 6,84%. Các microsatellite của motif có độ dài 4 (tetra-), 5 (penta-) và 6 (hexa-nucleotide) có số lượng ít nhất. Bên cạnh đó có 26 (5,23%) microsatellite dạng phức hợp đã được xác định trên các gen NBS-LRR phân tích (bảng 2 và hình 1A).

Bảng 2. Thông tin các microsatellite xác định trên họ gen NBS-LRR ở chanh leo.

Mục	Giá trị
Số gen NBS-LRR phân tích	161
Tổng chiều dài trình tự phân tích (nu)	2.180.487
Tổng số microsatellite xác định được	497
- Số gen NBS-LRR chứa SSR	139
- Số gen NBS-LRR có nhiều hơn 1 microsatellite	114
Phân loại microsatellite theo kích thước motif	Số lượng (tỷ lệ %)
Mono-nucleotide	321 (64,59%)
Di-nucleotide	106 (21,33%)
Tri-nucleotide	34 (6,84%)
Tetra-nucleotide	4 (0,81%)
Penta-nucleotide	3 (0,6%)
Hexa-nucleotide	3 (0,6%)
Microsatellite phức hợp	26 (5,23%)



Hình 1. Thông số các microsatellite xác định trên trình tự họ gen NBS-LRR ở chanh leo. (A) Phân bố các microsatellite theo chiều dài; (B) Phân loại các microsatellite theo số lần lặp của motif; (C) Phân bố các microsatellite tính theo trình tự của từng motif.

Theo M.K Biswas và cs (2020) [16], các microsatellite được phân loại thành hai nhóm dựa trên chiều dài trình tự, bao gồm nhóm I (>20 bp) và nhóm II (≤20 bp). Ở họ gen NBS-LRR chanh leo, số microsatellite thuộc nhóm II (443) nhiều gấp 11,6 lần nhóm I (38) (hình 1A). Số motif lặp lại của các microsatellite biến thiên từ 4 đến hơn 17 lần, trong đó các microsatellite chứa các motif lặp lại 10 lần (133) và 11 lần (67) chiếm tỷ lệ cao nhất (chủ yếu là motif 1-nucleotide) (hình 1B). Về trình tự, nhóm motif A/T chiếm đa số trong các motif 1-nucleotide, trong khi đó motif AT/AT phổ biến nhất trong số các motif 2-nucleotide và motif AAT/ATT thường gặp trong số các motif 3-nucleotide của các microsatellite xác định được ở họ gen NBS-LRR ở chanh leo (hình 1C).

3.2. Thiết kế chỉ thị phân tử SSR đặc hiệu họ gen NBS-LRR ở chanh leo

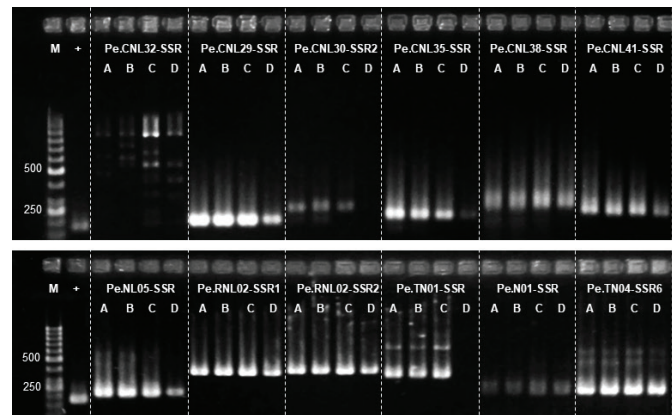
Các locus microsatellite có vị trí quá gần hai đầu trình tự gen NBS-LRR tham chiếu không đủ khoảng cách để thiết kế mỗi hoặc có thành phần trình tự hai vùng ngoại vi không đáp ứng yêu cầu thiết kế mỗi sẽ bị loại bỏ. Tổng số 405 cặp mỗi đã được thiết kế cho các microsatellite trên họ gen NBS-LRR ở chanh leo, trung bình 2,9 cặp mỗi cho mỗi gen được xác định mang microsatellite. Các cặp mỗi có sản phẩm khuếch đại dự kiến từ 175 đến 575 bp (các cặp mỗi đặc hiệu cho SSR phức hợp). Rà soát tính đặc hiệu thông qua so sánh với trình tự hệ gen chanh leo bằng công cụ BLASTn cho thấy chỉ có 22 cặp mỗi SSR (5,43%) có nhiều hơn 1 kết quả truy vấn. Các cặp mỗi này có khả năng khuếch đại nhiều hơn 1 locus nên bị loại bỏ. Các cặp mỗi còn lại (383 cặp mỗi, 94,57%) được xác định là các chỉ thị SSR tiềm năng đặc hiệu cho họ gen kháng NBS-LRR ở chanh leo.

3.3. Kiểm chứng tính hiệu dụng của bộ chỉ thị SSR đặc hiệu họ gen NBS-LRR ở chanh leo

Dựa trên bộ dữ liệu RNA-Seq về tính chịu lạnh [17] và phản ứng với bệnh hại [18] của các mẫu giống chanh leo, phân tích biểu hiện gen đã xác định được 35 gen NBS-LRR tiềm năng có mức độ biểu hiện gen khác biệt và tương quan với khả năng chống chịu các điều kiện bất thuận vô sinh và hữu sinh (kết quả chưa công bố). Tập trung vào nhóm gen NBS-LRR tiềm năng này, 46 chỉ thị SSR đặc hiệu cho 35 gen đã được lựa chọn để khảo sát kiểm chứng tính đặc hiệu và tính đa hình trên bộ 10 mẫu giống chanh leo.

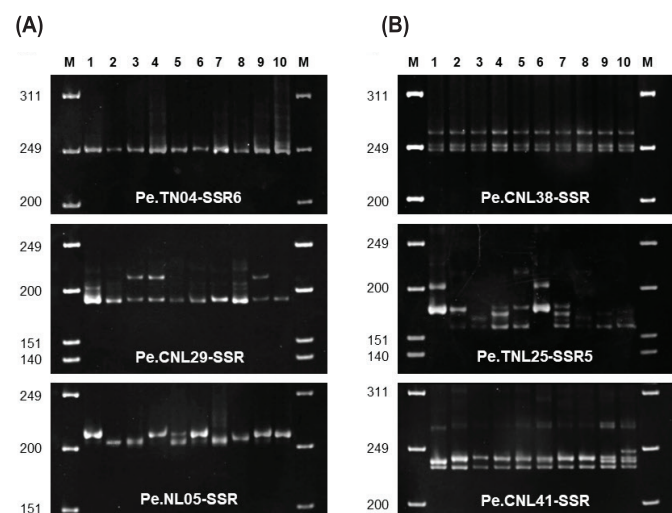
Khả năng bắt cặp đặc hiệu của các chỉ thị SSR với DNA khuôn của mẫu giống chanh leo H14.7 được đánh giá ở các nhiệt độ gắn mỗi T_a khác nhau. Kết quả khảo sát nhiệt độ gắn mỗi ghi nhận giá trị T_a tối ưu khá cao (60~62°C) cho hầu hết các chỉ thị SSR (hình 2). Đối với một số chỉ thị SSR có độ đặc hiệu kém ngay cả ở T_a cao nhất, phản ứng

Touchdown-PCR đã được thử nghiệm và cho thấy hiệu quả giảm thiểu sản phẩm khuếch đại không đặc hiệu mà không cần sử dụng các chất phụ gia bổ sung khác. Kết hợp cả hai phương pháp tối ưu, đã xác định được 32/46 chỉ thị SSR nhân bản đặc hiệu vùng trình tự mục tiêu từ 27 gen NBS-LRR tiềm năng ở chanh leo (bảng 3).



Hình 2. Kết quả khảo sát xác định nhiệt độ gắn mỗi tối ưu (T_a) của một số cặp mỗi SSR đặc hiệu gen NBS-LRR ở chanh leo. M: 50 bp DNA Ladder; +: đối chứng dương tại $T_a=54^\circ\text{C}$. Dải nhiệt độ gắn mỗi khảo sát: A: 50°C , B: 54°C , C: 58°C , D: 62°C . Giá trị T_a của các cặp mỗi được trình bày tại bảng 3.

Dựa trên giá trị T_a tối ưu, tính đa hình của 32 chỉ thị SSR đặc hiệu gen NBS-LRR tiềm năng tiếp tục được đánh giá thông qua phân tích PCR-PAGE trên 10 mẫu giống chanh leo. Kết quả cho thấy, về đa dạng số lượng alen của các locus microsatellite có 23/32 (71,87%) chỉ thị SSR là chỉ thị 2 alen (hình 3A) và 9/32 (28,13%) là chỉ thị đa alen (hình 3B). Trong khi đó, về tính đa hình, có 7/32 (21,88%)



Hình 3. Kết quả đánh giá kiểu gen của 10 mẫu giống chanh leo bằng các chỉ thị đặc hiệu gen kháng NBS-LRR. M: - Φ X174 DNA/HinfI marker; 1-10: thứ tự các mẫu giống chanh leo (danh sách theo bảng 1). Giá trị T_a của các cặp mỗi được trình bày tại bảng 2.

Bảng 3. Thông tin các chỉ thị SRR đặc hiệu gen NBS-LRR khuếch đại thành công ở các mẫu giống chanh leo khảo sát.

Thứ tự	Tên chỉ thị	Trình tự (5'-3')	Motif lặp	Ta (°C)	Kích thước alen (bp)	Phân loại chỉ thị		
						Hai alen	Đa alen	Đa hình
1	Pe.CN06-SSR	F: GATAAAAATTCCTTCCTCGACC R: CTACTTTTCATGCCTCGAATG	(CGACCA) ₃	56	175-260		X	X
2	Pe.CNL01-SSR1	F: ATGATATTGGGGATAGTAGGGA R: CCTGAGTAAAAGTCGATAAGGA	(G) ₁₃ ---(AT) ₉	56	413-427		X	X
3	Pe.CNL01-SSR5	F: GTTTGATTGGTGAAGGAGATAG R: ACTACCATCCACTCTAAGACTA	(AT) ₁₈	58	250-260	X		X
4	Pe.CNL05-SSR	F: CAAGGAAGTAACAGCGATTATG R: GCCAATTCATCCCTTGATC	(AT) ₇	58	300-310	X		X
5	Pe.CNL09-SSR	F: CTTGCCTTCCATTTAGCTTAAAG R: GATCCTCGTGACCACTATTAAC	(GAA) ₅	60	412-427	X		X
6	Pe.CNL16-SSR2	F: CAAATGAGACAGAAATTCGTGG R: CAGCTTCATTCGTCTTCATC	(AC) ₉	60	250-262	X		X
7	Pe.CNL16-SSR3	F: GCCCCAAGATATATTGGAATTG R: ATACAGGTCGGATTGATGAATG	(TTTTA) ₃ ---(TCTTA) ₃	60	206-225		X	X
8	Pe.CNL20-SSR	F: ATAAAACAGGGGCTAGATATG R: TCAGGTCGGGAAAAATCTTATG	(CCCGA) ₃	58	180-230	X		X
9	Pe.CNL22-SSR	F: AAGGAGTTTGGATCATCGTTTG R: ACTACCATCCACTCTAAGACTA	(TA) ₈	56	160-275		X	X
10	Pe.CNL29-SSR	F: GAGGGATCTTATGGAGGTTAAT R: CCGAAGACCTATCATAAGAACA	(AG) ₆	62	189-220	X		X
11	Pe.CNL30-SSR2	F: CAAGCAACTTAAGTGAAGAAC R: ACTGAGTCATCTTATCCCATC	(AAGCTC) ₃	58	240-270	X		
12	Pe.CNL35-SSR	F: CTTACAGCAAAAAGAAAGCA R: GGCTGTGGTGTGTGATATATA	(T) ₁₃	59*	209-239		X	X
13	Pe.CNL38-SSR	F: GAGAGAACGTTCAATTTAGTGG R: TTCATCCTTTAGTGGGTACAT	(TTAA) ₅	60*	248		X	
14	Pe.CNL41-SSR	F: TCGAGGACCAAGTAAAGTTAAG R: AAGACAGAAGACACAATAGCTC	(CA) ₆ ---(TA) ₆	60*	237-245		X	X
15	Pe.N01-SSR	F: GTGGAATTTGGAGTACTGTTT R: CATGCTCTAATATCGGATTGTG	(AAAAT) ₃ ---(ATA) ₆	45	237-255	X		
16	Pe.NL05-SSR	F: GGGCTCATATACATAAACTG R: CCTCCTATTGTTTTCCCTTTTC	(GAACCT) ₃	60	210-220	X		X
17	Pe.RNL02-SSR	F: GACCTCTTAAACCTCACAATG R: TCAAGTATATAGAGAGAGGCGA	(CAT) ₆	62	413-450	X		X
18	Pe.RNL02-SSR1	F: CCTTCTCTTTGTTGGTATTGTC R: GAACAGAACGTACGATTTCATC	(TA) ₆	62	371-417	X		X
19	Pe.TN01-SSR	F: GGCATTGAAGTCTAATAGAAGC R: TTTAACAGGGAGAGGAAAGATC	(A) ₁₅	58	337-382		X	X
20	Pe.TN04-SSR6	F: ATTATAACACCCCTTAACCGGTC R: GGTTGTTATGCATGATAGGTC	(GAACC) ₃	58	246	X		
21	Pe.TNL09-SSR1	F: TCACTATAAAACCCTCTTCACC R: GGAAGTGAATAGGAAGGAGTTT	(TC) ₆	50*	212-216	X		X
22	Pe.TNL14-SSR	F: CCTCAATGTAAGACAAGGAAGA R: TAATCCTCAACTCGATCTAACC	(AT) ₁₄ ---(ATATA) ₃	55	200	X		X
23	Pe.TNL19-SSR	F: GGAAGCTTTAATTCATCACCG R: CTGTGTTGGTATACAGTTTCGG	(ATA) ₈	46	204	X		
24	Pe.TNL25-SSR5	F: GGTTTTGGGCAATTTGTCATAC R: AGGATATGGACCATACACAATC	(AT) ₇	58*	175-200		X	X
25	Pe.TNL25-SSR6	F: CTACAGGATTACGTCAACTTC R: ATCAGATATCGCGACTATACAG	(GAACAAG) ₂ --- (ACAAGCA) ₂	58	221	X		
26	Pe.TNL27-SSR2	F: CAGTTGTTGGTATTGGGTAGAA R: GGATACCTTCTCGAGTTTAGT	(T) ₁₆	58	214-232	X		X
27	Pe.TNL27-SSR3	F: ACAAAAATGAGTTCGTGCTG R: TGGCAGACAAAAAGATAGTG	(TCG) ₈	58	140-220	X		X
28	Pe.TNL28-SSR	F: TATATAGGGCAGGATTTTCACC R: ATTCATACATCTCAGCTCTCTC	(CT) ₇ ---(TA) ₆	56	175-185		X	X
29	Pe.TNL34-SSR	F: GCTAACAGAAGACATAAGTTCC R: CAGTGATATTCTACCAAGACCA	(T) ₁₄	52*	210-220	X		X
30	Pe.TNL39-SSR	F: CTCTTCGTGGAATCAATTACC R: CTATTTTCAACGTTTCACTCCC	(A) ₁₄	52	240-250	X		
31	Pe.TNL50-SSR5	F: CTGACCAAATCTGAAATGTCTC R: TGGTGTGGAAAGAAGATACTAG	(CAT) ₅	40	200 -235	X		X
32	Pe.TNL55-SSR	F: GCTAACAGAAGACATAAGTTCC R: CAGTGATATTCTACCAAGACCA	(T) ₁₄	52	213-220	X		X

*: chỉ thị áp dụng kỹ thuật Touch down - PCR.

chỉ thị đơn hình (ví dụ chỉ thị 2 alen Pe.TN04-SSR6 và chỉ thị đa alen Pe.CNL38-SSR, hình 3) và có 25/32 (78,12%) chỉ thị SSR cho đa hình (ví dụ các chỉ thị Pe.CNL29-SSR, Pe.NL05-SSR, Pe.TNL25-SSR5, và Pe.CNL41-SSR, hình 3) trên bộ 10 mẫu giống chanh leo khảo sát.

4. Bàn luận

Các yếu tố DNA lặp, bao gồm DNA lặp liên tục (tandem repeat) và DNA lặp rải rác (interspersed repeat), có đóng góp quan trọng vào sự biến thiên về kích thước và phức tạp về cấu trúc của hệ gen thực vật. Microsatellite (kích thước motif lặp <10 nucleotide) cùng với minisatellite (motif lặp dài hơn 10 nucleotide) là hai phân lớp của nhóm DNA lặp liên tục, đặc trưng bởi tính không ổn định với tỷ lệ đột biến từ 10^{-3} đến 10^{-6} qua mỗi thế hệ tế bào, cao gấp 10 lần tỷ lệ đột biến điểm xảy ra trong hệ gen. Đây cũng chính là cơ sở tạo nên tính đa hình cao của các microsatellite trong hệ gen thực vật [19]. Nghiên cứu trên các loài cây trồng cho thấy các microsatellite phân bố không đồng đều, chủ yếu tập trung ở các vùng gen không mã hóa do có tiềm năng đột biến cao [20, 21]. Tuy nhiên ở nghiên cứu này, số lượng microsatellite xác định được trên các gen NBS-LRR và vùng trình tự phụ cận là tương đối lớn (trung bình 3,08 microsatellite/gen), điều này có thể phản ánh sự đa dạng của họ gen NBS-LRR ở thực vật, thể hiện ở tính lặp lại và cấu trúc cụm gen (cluster) hình thành trên các nhiễm sắc thể [7]. Tùy theo nhóm loài, thành phần và tỷ lệ các motif microsatellite ghi nhận được là khác nhau [19]. Đối với họ gen NBS-LRR ở chanh leo, các nhóm motif A/T và AT/AT chiếm số lượng đa số (tỷ lệ lần lượt là 62,37% và 14,89%), tương ứng với kết quả xác định microsatellite trước đó từ dữ liệu giải trình tự thư viện tách dòng (BAC-end sequencing) [22] và dữ liệu trình tự hệ gen chưa hoàn chỉnh (draft-genome) [11] ở chanh leo vàng.

So với các loại chỉ thị phân tử khác, chỉ thị đồng trội SSR cho kết quả đánh giá kiểu gen ổn định và do đó mức độ tin cậy của các phân tích đa dạng di truyền, xác định đặc trưng di truyền của cá thể/loài hay lập bản đồ di truyền dựa trên dữ liệu SSR cũng cao hơn. Chính vì vậy, chỉ thị SSR được ứng dụng phổ biến trong các nghiên cứu thuộc lĩnh vực sinh học phân tử và di truyền ở nhiều đối tượng cây trồng khác nhau [22]. Đối với chanh leo, nghiên cứu xác định các gen quy định tính kháng bệnh hại và chỉ thị phân tử đặc hiệu hoặc liên kết gen kháng làm nền tảng phục vụ chọn tạo giống chanh leo kháng bệnh là hướng nghiên cứu được quan tâm. Một số nghiên cứu lập bản đồ di truyền ở

chanh leo đã được thực hiện nhằm xác định các locus quy định tính kháng nhưng chưa thành công do số lượng hạn chế của các chỉ thị phân tử so với kích thước hệ gen chanh leo không đảm bảo độ phân giải của vùng chứa locus mục tiêu [1]. Để làm phong phú thêm các chỉ thị đa hình cho chanh leo, E.J. Olivera và cs (2005) [12] đã phát triển 10 chỉ thị SSR từ thư viện tách dòng hệ gen, trong khi đó S. Araya và cs (2017) [11] thông qua phân tích bộ dữ liệu trình tự hệ gen chưa hoàn chỉnh đã phát triển 42 chỉ thị SSR khác. Trong nghiên cứu này, 383 chỉ thị SSR đặc hiệu họ gen NBS-LRR ở chanh leo đã được thiết kế thông qua các phân tích *in-silico*. Kết quả kiểm chứng thực nghiệm 32 chỉ thị SSR trên bộ 10 mẫu giống chanh leo cho thấy, mặc dù đã sàng lọc tính đặc hiệu bằng công cụ BLAST cũng như tối ưu nhiệt độ gắn mồi T_a , một số chỉ thị SSR vẫn có khả năng khuếch đại các locus khác ngoài vùng mang microsatellite mục tiêu. Theo E. Guichoux và cs (2011) [23], điều kiện PCR bao gồm nhiều yếu tố ảnh hưởng và không hoàn toàn nghiêm ngặt có thể là nguyên nhân của hiện tượng này. Ngoài ra, do bản chất đa dạng của họ gen NBS-LRR, việc phát sinh bản sao của một gen NBS-LRR trong hệ gen hoàn toàn có thể xảy ra ở một số nguồn gen ở các điều kiện áp lực tiến hóa bệnh hại khác nhau, như là một cơ chế đảm bảo nguồn alen gen kháng phong phú trong quá trình tiến hóa gen kháng - gen gây bệnh mang tính tuần tự kế thừa của mô hình gen-đối-gen [24]. Ngoài ra, tỷ lệ các chỉ thị SSR đa hình đạt 78,12% phần nào đã phản ánh sự đa dạng alen tương đối cao của các gen NBS-LRR ở các mẫu giống chanh leo, khẳng định tiềm năng ứng dụng trong đánh giá đa dạng di truyền gen kháng và nghiên cứu lập bản đồ di truyền, chọn tạo giống chanh leo bằng chỉ thị phân tử trong thời gian tới.

5. Kết luận

Nghiên cứu này đã xác định các locus microsatellite và thiết kế được 383 chỉ thị SSR đặc hiệu gen NBS-LRR ở chanh leo. Trong số này, 25 chỉ thị đặc hiệu cho các gen NBS-LRR tiềm năng có biểu hiện liên quan đến phản ứng stress sinh học và phi sinh học đã được kiểm chứng thực nghiệm về tính đặc hiệu và tính đa hình trên bộ mẫu giống chanh leo. Các chỉ thị đều khuếch đại được locus microsatellite với kích thước như thiết kế và có sự đa dạng về số lượng các alen, phản ánh đúng bản chất di truyền của họ gen NBS-LRR. Các chỉ thị SSR đặc hiệu gen NBS-LRR sẽ là công cụ bổ sung có giá trị cho công tác đánh giá nguồn gen kháng và nghiên cứu chọn tạo giống chanh leo kháng bệnh hại bằng chỉ thị phân tử trong thời gian tới.

TÀI LIỆU THAM KHẢO

- [1] C.B.M.C. Silva, F.G. Faleiro, O.N.D. Jesus, et al. (2018), “Passion fruit (*Passiflora* spp.) breeding”, *Advances in Plant Breeding Strategies: Fruit*, Springer International Publishing, **3**, pp.929-951.
- [2] X. He, F. Luan, Y. Yang, et al. (2020), “*Passiflora edulis*: An insight into current researches on phytochemistry and pharmacology”, *Frontiers in Pharmacology*, **11**, DOI: 10.3389/fphar.2020.00617.
- [3] D.H. Do, Y.H. Chong, V.C. Ha, et al. (2021), “Characterization and detection of passiflora mottle virus and two other potyviruses causing passion fruit woodiness disease in Vietnam”, *Phytopathology*, **111**(9), pp.1675-1685, DOI: 10.1094/PHYTO-10-20-0481-R.
- [4] M.A. Gururani, J. Venkatesh, C.P. Upadhyaya, et al. (2012), “Plant disease resistance genes: Current status and future directions”, *Physiol. Mol. Plant Pathol.*, **78**, pp.51-65, DOI: 10.1016/j.pmpp.2012.01.002.
- [5] J. Niu, N. Guo, J. Sun, et al. (2017), “Fine mapping of a resistance gene RpsHN that controls *Phytophthora sojae* using recombinant inbred lines and secondary populations”, *Frontiers in Plant Sciences*, **8**, DOI: 10.3389/fpls.2017.00538.
- [6] F. Jupe, K. Witek, W. Verweij, et al. (2013), “Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations”, *Plant Journal*, **76**(3), pp.530-544, DOI: 10.1111/tj.12307.
- [7] T.D. Trung, T.H. Linh (2021), “*In-silico* genome-wide identification and characterisation of NBS-LRR gene family in banana *Musa acuminata*”, *Vietnam Journal of Science and Technology - MOST*, **63**(12), pp.52-58, DOI: 10.31276/VJST.63(12).52-58 (in Vietnamese).
- [8] D. Ma, S. Dong, S. Zhang, et al. (2021), “Chromosome-level reference genome assembly provides insights into aroma biosynthesis in passion fruit (*Passiflora edulis*)”, *Molecular Ecology Resources*, **21**(3), pp.955-968, DOI: 10.1111/1755-0998.13310.
- [9] S. Amiteye (2021), “Basic concepts and methodologies of DNA marker systems in plant molecular breeding”, *Heliyon*, **7**(10), DOI: 10.1016/j.heliyon.2021.e08093.
- [10] D. Tautz (1989), “Hypervariability of simple sequences as a general source for polymorphic DNA markers”, *Nucleic Acids Res.*, **17**(16), pp.6463-6471, DOI: 10.1093/nar/17.16.6463.
- [11] S. Araya, A.M. Martins, N.T.V. Junqueira, et al. (2017), “Microsatellite marker development by partial sequencing of the sour passion fruit genome (*Passiflora edulis* Sims)”, *BMC Genomics*, **18**(1), DOI: 10.1186/s12864-017-3881-5.
- [12] E.J. Olivera, J.G. Pádua, M.I. Zucchi, et al. (2005), “Development and characterization of microsatellite markers from the yellow passion fruit (*Passiflora edulis* f. *flavicarpa*)”, *Mol. Ecol. Notes*, **5**(2), pp.331-333, DOI: 10.1111/j.1471-8286.2005.00917.x.
- [13] S. Beier, T. Thiel, T. Münch, et al. (2017), “MISA-web: A web server for microsatellite prediction”, *Bioinformatics*, **33**(16), pp.2583-2585, DOI: 10.1093/bioinformatics/btx198.
- [14] T.D. Trung, N.V. Viet, B.Q. Dang (2022), “Genetic diversity and population structure of passion fruit (*Passiflora edulis* Sim.) accessions revealed by SRAP and ISSR markers”, *Vietnam Journal of Agricultural Sciences*, **20**(4), pp.425-435.
- [15] D.J. Korbie, J.S. Mattick (2008), “Touchdown PCR for increased specificity and sensitivity in PCR amplification”, *Nat. Protoc.*, **3**(9), pp.1452-1456, DOI: 10.1038/nprot.2008.133.
- [16] M.K. Biswas, M. Bagchi, U.K. Nath, et al. (2020), “Transcriptome wide SSR discovery cross-taxa transferability and development of marker database for studying genetic diversity population structure of *Lilium* species”, *Scientific Reports*, **10**(1), DOI: 10.1038/s41598-020-75553-0.
- [17] Y. Wu, W. Huang, Q. Tian, et al. (2021), “Comparative transcriptomic analysis reveals the cold acclimation during chilling stress in sensitive and resistant passion fruit (*Passiflora edulis*) cultivars”, *PeerJ*, **9**, DOI: 10.7717/peerj.10977.
- [18] F. Ma, H. Zhou, Y. Xu, et al. (2023), “Comprehensive analysis of bZIP transcription factors in passion fruit”, *iScience*, **26**(4), DOI: 10.1016/j.isci.2023.106556.
- [19] R. Gemayel, J. Cho, S. Boeynaems, et al. (2012), “Beyond junk-variable tandem repeats as facilitators of rapid evolution of regulatory and coding sequences”, *Genes (Basel)*, **3**(3), pp.461-480, DOI: 10.3390/genes3030461.
- [20] G. Bhattarai, A. Shi, D.R. Kandel, et al. (2021), “Genome-wide simple sequence repeats (SSR) markers discovered from whole-genome sequence comparisons of multiple spinach accessions”, *Scientific Reports*, **11**(1), DOI: 10.1038/s41598-021-89473-0.
- [21] S.R. McCouch, L. Teytelman, Y. Xu, et al. (2002), “Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.)”, *DNA Res.*, **9**(6), pp.199-207, DOI: 10.1093/dnares/9.6.199.
- [22] M.L. Vieira, L. Santini, A.L. Diniz, et al. (2016), “Microsatellite markers: What they mean and why they are so useful”, *Genet. Mol. Biol.*, **39**(3), pp.312-328, DOI: 10.1590/1678-4685-GMB-2016-0027.
- [23] E. Guichoux, L. Lagache, S. Wagner, et al. (2011), “Current trends in microsatellite genotyping”, *Mol. Ecol. Resour.*, **11**(4), pp.591-611, DOI: 10.1111/j.1755-0998.2011.03014.x.
- [24] E.J. Andersen, S. Ali, E. Byamukama, et al. (2018), “Disease resistance mechanisms in plants”, *Genes*, **9**(7), DOI: 10.3390/genes9070339.