

NGHIÊN CỨU ĐA DẠNG DI TRUYỀN NGUỒN GEN CAO SU ĐANG ĐƯỢC BẢO TỒN Ở VIỆT NAM

Vũ Văn Trường¹, Huỳnh Đức Định¹, Nguyễn Thị Thảo¹,

Trần Thanh¹, Huỳnh Thị Minh Tâm¹, Ronan Rivallan²,

Huỳnh Văn Biết³, Le Guen Vincent²

TÓM TẮT

Cao su thiên nhiên là nguyên liệu chiến lược cho nhiều ngành công nghiệp và chỉ được sản xuất bằng việc trồng cây cao su (*Hevea brasiliensis* Muell. Arg.). Chọn tạo giống loài này chủ yếu tập trung trên những giống đã cải tiến ở châu Á và châu Phi, nhưng cũng dựa vào đa dạng di truyền có trong quỹ gen hoang dại. Nghiên cứu nhằm đánh giá đa dạng di truyền của các mẫu giống từ bang Rondonia (Brazil) đang bảo tồn tại Việt Nam với 15 chỉ thị SSR. Toàn bộ 1.127 mẫu giống được phân tích đã phát hiện 200 mẫu có quan hệ gần gũi di truyền 100% từ 85 cặp và 10 bộ ba, cũng như 7 mẫu không thuộc về nguồn gen Rondonia. Số lượng allele khá phong phú trên các vị trí SSR (15 - 47 allele) và tần số allele giữa các nhóm di truyền là khác biệt; tổng số 467 allele từ 11 nhóm di truyền và 458 allele từ nguồn gen Rondonia. Đa dạng di truyền cao ở tất cả các nhóm, với chỉ số đa dạng (He = 0,78) cao hơn chỉ số đa dạng quan sát (Ho = 0,66) và giữa các cá thể trên các nhóm di truyền chưa xảy ra cận huyết. Nguồn gen Rondonia có đa dạng di truyền hơn Wickham và Mato Grosso; giữa Wickham và Mato Grosso gần gũi nhau với khác biệt di truyền (Fst) lớn hơn các nhóm di truyền khác, trong khi giữa các nhóm trong nguồn gen Rondonia có sự khác biệt di truyền rất thấp. Quỹ gen cao su Việt Nam có tiềm năng di truyền rất lớn, biến dị di truyền bên trong cá thể là 74%, nhưng giữa các nhóm vẫn có sự khác biệt di truyền và cấu trúc di truyền phù hợp với vị trí được sưu tập.

Từ khóa: Bảo tồn, cây cao su, cấu trúc di truyền, đa dạng di truyền, khác biệt di truyền, nguồn gen.

1. GIỚI THIỆU

Cây cao su có nguồn gốc từ rừng nhiệt đới thuộc lưu vực sông Amazon và phân bố ở 9 quốc gia Nam Mỹ (Compagnon, 1986). Chi *Hevea* thuộc họ Euphorbiaceae gồm 11 loài đều là nhị bội (2n = 36) và có thể giao phấn lẫn nhau (Ong, 1979). Trong số các loài, *Hevea brasiliensis* là loài quan trọng nhất để sản xuất mủ cao su và cung cấp hơn 98% sản lượng mủ (Backhaus, 1985; Bowers, 1990). Sản lượng mủ cao su trên thế giới đã tăng gấp 3 lần trong vòng ba thập kỷ, từ 4,2 triệu tấn năm 1985 lên 13,2 triệu tấn năm 2015 (<http://www.fao.org/faostat/>). Sự gia tăng sản lượng nhanh chóng là do tăng diện tích và tăng năng suất, trong đó giống vẫn là yếu tố chính. Hiện nay, Việt Nam trở thành một trong những nước dẫn đầu về năng suất và đứng thứ ba về sản lượng mủ cao su (ANRPC, 2018).

Chọn tạo giống cao su chủ yếu tập trung trên những giống đã được cải tiến và phát triển hơn một thế kỷ ở châu Á và châu Phi, nhưng cũng dựa vào đa dạng di truyền có trong các bộ sưu tập quỹ gen hoang dại. Nguồn gen cao su ban đầu đưa vào Đông Nam Á từ bộ sưu tập của Wickham tại Brazil năm 1876 (Dean, 1987; Webster và Baulkwill, 1989). Từ năm 1945 - 1982, có khoảng 10 bộ sưu tập quỹ gen từ rừng Amazon được chuyển đến các nước trồng cao su. Từ năm 1984, các nước thành viên của Hiệp hội Nghiên cứu và Phát triển Cao su quốc tế (IRRDB) đã tiếp nhận nguồn gen IRRDB'81 để mở rộng vốn di truyền cho công tác cải tiến giống.

Nguồn gen IRRDB'81 gần như đã được đánh giá về mặt nông học ở nhiều nước (Clement-Demange và ctv, 1998; Ong và ctv, 1995). Về mặt di truyền, ban đầu được đánh giá với chỉ thị isozyme (Chevallier, 1988; Lai và ctv, 2012), sau đó là các chỉ thị phân tử như RAPD (Lai và ctv, 2009; Low và ctv, 1996), SSR (Le Guen và ctv, 2009; Souza và ctv, 2015) hoặc SNP (Chanroj và ctv, 2017). Ở các trung tâm bảo tồn quỹ gen cao su thường lưu giữ số lượng mẫu giống rất lớn, nhưng đánh giá di truyền chỉ với một vài mẫu cho mỗi vùng địa lý. Tất cả các nghiên cứu về đa dạng di truyền là chia các quần thể thành từng cụm

¹ Viện Nghiên cứu Cao su Việt Nam

² Trung tâm Hợp tác quốc tế Nghiên cứu Phát triển Nông nghiệp Pháp (CIRAD)

³ Trường Đại học Nông Lâm thành phố Hồ Chí Minh
Email: truongriv@gmail.com

riêng biệt theo vùng địa lý được sưu tập gồm Acre, Mato Grosso và Rondonia và cuối cùng là quan hệ gần gũi di truyền với nguồn gen Wickham (Le Guen và ctv, 2009; Souza và ctv, 2015).

Việt Nam bắt đầu nhập nguồn gen IRRDB'81 từ năm 1984 chủ yếu từ Malaysia và một phần từ Côte d'Ivoire (Vo và ctv, 1994). Tổng số 2.976 mẫu giống IRRDB'81 đang được bảo tồn ở dạng vườn nhân chồi ghép (dòng vô tính) từ những cây ban đầu sưu tập. Cho đến nay, nguồn gen cao su Amazon hoang dại ở Việt Nam đã được đánh giá về mặt nông học, nhưng về mặt di truyền chỉ một phần rất nhỏ đã được đánh giá với chỉ thị RAPD (Lai và ctv, 2009). Nghiên cứu này nhằm đánh giá một cách hệ thống về sự khác biệt di truyền ở mức độ phân tử trên toàn bộ nguồn gen Rondonia đang được bảo tồn tại Việt Nam.

2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Bảng 1. Danh sách mẫu giống cao su theo nguồn gốc địa lý và Trung tâm Bảo tồn quỹ gen

Nguồn gốc giống	Vùng địa lý	Ký hiệu	Trung tâm bảo tồn		Tổng số
			CIRAD	Việt Nam	
Rondonia	Sưu tập 1974	RO		18	18
Rondonia 1981	Ariquemes	RO/A	1	152	153
Rondonia 1981	Calama	RO/C	1	210	211
Rondonia 1981	Costa Marques	RO/CM		251	251
Rondonia 1981	Jaru	RO/J		113	113
Rondonia 1981	Ji-Parana	RO/JP		144	144
Rondonia 1981	Ouro Preto	RO/OP		27	27
Rondonia 1981	Pimenta Bueno	RO/PB		145	145
Acre		AC	14		14
Mato Grosso		MT	9		9
Wickham		W	8	34	42
Tổng cộng			33	1.094	1.127

Nguồn gen Rondonia 1981 thuộc bộ sưu tập IRRDB'81 tại bang Rondonia của Brazil.

2.2. Các chỉ thị SSR

Để phân tích kiểu gen cho bộ mẫu giống cao su, 15 chỉ thị SSR được chọn lọc để phân tích là rất đa dạng và đã được xác định trên 12 nhiễm sắc thể khác nhau của bộ gen. Tất cả các chỉ thị SSR này đều đã

2.1. Vật liệu giống

Tổng số 1.127 mẫu giống đưa vào nghiên cứu từ nhiều bộ sưu tập, vùng địa lý và Trung tâm bảo tồn quỹ gen (Bảng 1). Chi tiết về số lượng và nguồn gốc giống như sau:

- 1.060 mẫu giống Rondonia được bảo tồn trong vườn nhân lưu trữ (SNLK86) tại Viện Nghiên cứu Cao su Việt Nam (RRIV), trong đó 1.042 mẫu thuộc bộ sưu tập IRRDB'81 và 18 mẫu thuộc bộ sưu tập 1974.

- 67 mẫu giống đối chứng, trong đó 34 dòng vô tính Việt Nam xuất phát từ nguồn gen Wickham; 33 mẫu giống được bảo tồn tại Trung tâm Nghiên cứu của CIRAD ở Kourou (Guiana - Pháp), các mẫu giống của CIRAD là mẫu DNA.

được sử dụng để đánh giá cho các quần thể cây cao su trước đó (Le Guen và ctv, 2009; Souza và ctv, 2015; Lespinasse và ctv, 2000a hoặc Le Guen và ctv, 2011). Các chỉ thị SSR được mô tả chi tiết trong cơ sở dữ liệu NCBI (Bảng 2).

Bảng 2. Đặc điểm của 15 chỉ thị SSR được chọn lọc để phân tích kiểu gen

Tên chỉ thị	Ký hiệu mẫu trong NCBI	Motif lặp lại	Kích thước allele (bp)	Định vị trên nhiễm sắc thể	Nghiên cứu đã sử dụng
A2365	AY486666	CT	188 - 229	4	[1], [2]
A2368	AY486668	CT	191 - 279	2	[1], [2]
A2387	AY486690	CT	187 - 227	17	[1], [2]
A2406	AY486697	CT	114 - 178	6	[1], [2]
A2508	AY486740	CT	238 - 317	13	[1], [2]

A2684	AY486821	CT	238 – 315	1	[1], [2]
A2689	AY486822	CT	84 - 120	9	[1]
A2736	AY486841	CT	102 – 148	11	[1], [2]
A31	AF383940	GA	124 - 180	3	[4]
BAC55B02	DQ115609	CT	147 - 175	8	[1], [2]
M574	AF221706	(TA) (GA)	210 - 270	15	[3]
MnSod	G73377	CT	178 - 244	11	[3]
T2083	AY486904	CA	280 - 308	12	[1], [2]
TA2163	AY486617	(CT) (CA)	193 - 248	15	[1], [2]
TAs2558	AY486760	(CT) (GA)	212 - 269	3	[1], [2]

[1] Le Guen và ctv (2009), [2] Souza và ctv (2015), [3] Lespinasse và ctv (2000a), [4] Le Guen và ctv (2011).

2.3. Phương pháp nghiên cứu

2.3.1. Lấy mẫu và chiết xuất DNA

Mẫu lá non ở giai đoạn màu tím đồng của tất cả các mẫu giống được thu thập trên vườn nhân lưu trữ (SNLK86) và DNA được chiết xuất tại phòng thí nghiệm của Viện Nghiên cứu Cao su Việt Nam (RRIV). DNA tổng số được phân lập từ 200 mg mẫu lá theo phương pháp được điều chỉnh từ quy trình của Murray và Thompson (1980). Mẫu lá được nghiền trong nitor lỏng và DNA được chiết xuất ở 65°C trong 45 phút với dung dịch đệm trên nền CTAB, pH 8,0; sau đó thêm dung dịch phenol:chloroform:isoamyl alcohol (25:24:1). RNA được loại bỏ ở giai đoạn lỏng bằng cách ủ với RNase A ở 37°C trong 1 giờ. Thêm Isopropanol lạnh để kết tủa DNA, mẫu DNA cuối cùng được hòa tan trong dung dịch TE 1X. Chất lượng DNA được kiểm tra bằng điện di trên gel agarose 1% và nồng độ xác định bằng máy đo Nanophotometer® P330 (Implen, Schatzbogen, Đức).

2.3.2. Phân tích các chỉ thị SSR

Khuếch đại các chuỗi SSR được thực hiện tại phòng thí nghiệm của CIRAD (Montpellier - Pháp). Phản ứng PCR được sử dụng với hệ thống 3 cặp mồi gắn nhãn huỳnh quang theo quy trình của Schuelke (2000). Tổng thể tích cho phản ứng PCR gồm 10 µL, trong đó 25 ng DNA khuôn, 0,2 µM mỗi cặp mồi, 200 µM deoxynucleotide (dNTP), 2 mM MgCl₂, 1x PCR buffer và 1U Taq DNA polymerase. Chạy phản ứng PCR bao gồm các bước, bước biến tính ban đầu ở 94°C trong 5 phút, tiếp theo là 10 chu kỳ với nhiệt độ giảm 0,5°C trong khi bắt cặp cho mỗi chu kỳ từ 55°C xuống còn 50°C (94°C trong 45 giây, 55°C trong 1 phút, 72°C trong 1 phút 15 giây), 25 chu kỳ với nhiệt độ bắt cặp mồi ở 50°C (94°C trong 45 giây, 50°C

trong 1 phút, 72°C trong 1 phút) và bước kéo dài cuối cùng ở 72°C trong 30 phút. Sản phẩm PCR được biến tính với formamide và được phân tích điện di trên máy đọc trình tự ABI 3500 (Applied Biosystems, Foster City, CA, USA). Tất cả các allele được ghi nhận bằng phần mềm GeneMapper® v4.1.

2.3.3. Phân tích các thông số di truyền

Toàn bộ mẫu giống phân tích kiểu gen thành công với 15 chỉ thị SSR được đưa vào kiểm tra để phát hiện quan hệ thân thuộc di truyền. Trước tiên, phân tích khoảng cách di truyền dựa vào ma trận sai biệt và xây dựng cây phả hệ bằng phần mềm DARWIN v6.0.14 (Perrier và Jacquemoud-Collet, 2006).

Tiếp theo, các bước phân tích sâu hơn được thực hiện trên các mẫu của cùng một nhóm di truyền (đã loại bỏ các mẫu dư thừa). Phần mềm GenAlex (Peakall và Smouse, 2012) được sử dụng để xác định giá trị của các thông số di truyền gồm dị hợp tử kỳ vọng (He), dị hợp tử quan sát (Ho), hệ số nội phối (Fst) theo công thức của Nei (1973) với xác suất (P) dựa trên 999 hoán vị và chỉ số khác biệt di truyền giữa các cặp quần thể với xác suất liên quan. Phân tích thành phần chính (PCoA) để đưa ra một đồ thị về khoảng cách đa chiều giữa các quần thể từ ma trận khoảng; phân tích phương sai phân tử (AMOVA) để xác định quan hệ di truyền tồn tại bên trong và giữa các cá thể trong cùng một nhóm di truyền.

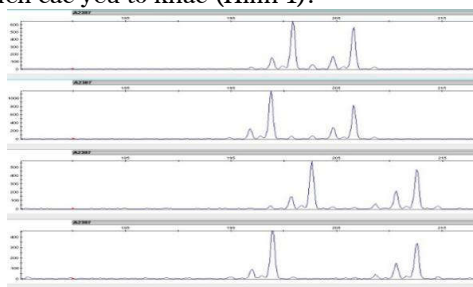
3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Xác định quan hệ thân thuộc di truyền giữa các mẫu giống thông qua chỉ thị SSR

Các nguồn gen cao su được bảo tồn ở Việt Nam trong vườn nhân lưu trữ đã kéo dài hơn 35 năm mà chưa được truy xuất nguồn gốc hoặc kiểm soát về mặt di truyền. Cho đến nay, ở Việt Nam chỉ có một vài nghiên cứu về di truyền với các chỉ thị phân tử và

chỉ giới hạn vài chục mẫu giống (Lai và ctv, 2009). Do đó, các nghiên cứu trước chưa phản ánh được đầy đủ về mức độ tin cậy cho toàn bộ các nguồn gen cao su có trong bộ sưu tập quỹ gen. Đây là nghiên cứu đầu tiên được đánh giá một cách toàn diện nhất về mối quan hệ và khác biệt di truyền ở mức độ phân tử của những cây cao su có nguồn gốc từ bang Rondonia (Brazil) được bảo tồn tại Việt Nam. Bên cạnh đó, tính độc đáo của nghiên cứu này là số lượng mẫu được phân tích khá lớn cho mỗi vị trí sưu tập tại bang Rondonia, từ 27 đến 251 mẫu giống (Bảng 1). Trong khi, những nghiên cứu trước, số lượng mẫu được phân tích cho mỗi vị trí của bang Rondonia chỉ từ 2 đến 23 mẫu (Chevallier, 1988; Lai và ctv, 2009; Le Guen và ctv, 2009).

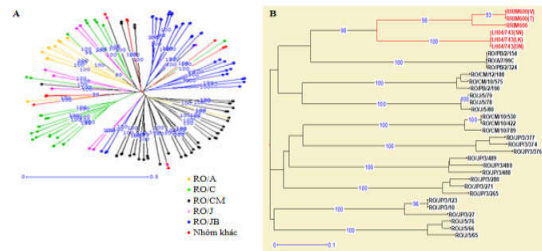
Trước tiên, toàn bộ dữ liệu được kiểm tra trên phần mềm GeneMapper®v4.1, những mẫu phân tích kiểu gen thành công với chỉ thị SSR được ghi nhận có các giá trị đỉnh allele cao nhất. Dữ liệu cũng được so sánh với đối chứng PB 260 và RRIM 600 cho mỗi lần chạy trên máy đọc trình tự ABI 3500 để đảm bảo dữ liệu phù hợp giữa các lần chạy. Một số trường hợp bất thường đã được ghi nhận như chiều dài sản phẩm khuếch đại có biến thể nhỏ 1 bp hoặc 2 bp và số lượng allele bất thường gồm (a) trên cùng một mẫu có nhiều chỉ thị SSR (từ 2 trở lên); (b) một hoặc hai chỉ thị trên một mẫu có trên 2 allele. Cho dù bất kỳ nguyên nhân nào về số lượng allele bất thường, cây cao su vẫn được xem là loài lưỡng bội và không thể tồn tại thể tam bội hoặc tứ bội vì các phần mềm phân tích di truyền không thể xử lý cho nhiều mức độ bội thể trong cùng một bộ dữ liệu. Do đó, tất cả những mẫu bất thường bị loại, chỉ giữ lại 1.127 mẫu giống có hai đỉnh allele với chiều cao lớn hơn và không xem xét đến các yếu tố khác (Hình 1).



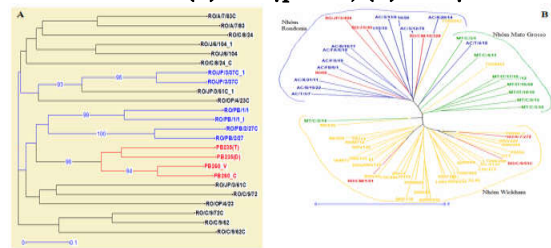
Hình 1. Hình ảnh minh họa phân tích điện di trên máy đọc trình tự ABI 3500 cho 4 mẫu giống cao su dị hợp tử với chỉ thị A2387

Một ma trận sai biệt và cây phát sinh loài được tính toán bằng phần mềm DARWIN để phát hiện quan hệ thân thuộc di truyền giữa các mẫu giống.

Tổng số 1.127 mẫu được đưa vào kiểm tra đã xác định được 200 mẫu giống có quan hệ gần gũi di truyền với chỉ số bootstrap là 100% từ 85 cặp và 10 bộ ba (Hình 2A và 2B). Ngoài ra, 11 cặp mẫu có cùng tên giống cũng được kiểm tra, nhưng chỉ xác định được ba cặp mẫu có quan hệ di truyền giống nhau (Hình 3A). Tương tự, đối với 7 mẫu giống được cho là không thuộc về nguồn gen Rondonia mà chúng thuộc về các nguồn di truyền khác, trong đó bốn mẫu thuộc nguồn gen Acre và 3 mẫu thuộc nguồn gen Wickham (Hình 3B).



Hình 2. Quan hệ di truyền giữa các mẫu giống thông qua cây phả hệ được xây dựng từ phần mềm DARWIN. (A) 85 cặp mẫu; (B) 10 bộ ba



Hình 3. Quan hệ di truyền giữa các mẫu giống thông qua cây phả hệ được xây dựng từ phần mềm DARWIN. (A) 11 cặp mẫu có cùng tên giống; (B) 7 mẫu giống của nguồn gen Rondonia

Xem xét về mặt di truyền, những cá thể có quan hệ gần gũi di truyền với nhau được cho là dư thừa và bị loại ra khỏi bộ dữ liệu gồm có 105 mẫu giống từ 85 cặp và 10 bộ ba; 7 mẫu giống không thuộc về nguồn gen Rondonia cũng bị loại. Do đó, 1.015 mẫu được xem là đúng giống vì không trùng lặp về mặt di truyền trong tổng số 1.127 mẫu được phân tích với tỷ lệ bảo tồn đạt 90%. Tất cả những sai sót về tên giống có thể do nhiều nguyên nhân hoặc nguồn gen nhập vào Việt Nam từ năm 1984 hoặc thu thập mẫu lá và phân tích mẫu tại CIRAD. Tuy nhiên, giá trị của bộ dữ liệu đạt được khi phân tích kiểu gen với 15 chỉ thị SSR rất đáng tin cậy. Như vậy, bộ dữ liệu cuối cùng tiếp tục được phân tích sâu hơn gồm 951 mẫu có nguồn gốc từ bang Rondonia và 62 mẫu từ các nhóm di truyền khác là Acre, Mato Grosso và Wickham.

3.2. Mức độ đa hình trên các vị trí SSR

Đa hình của các chỉ thị SSR đạt được từ bộ mẫu của 11 nhóm di truyền gồm 1.013 cá thể, số lượng allele khá phong phú trên tất cả các vị trí SSR với tổng số 467 allele được phát hiện, trung bình là 14,1 allele và dao động từ 15 đến 47 allele. Mức độ đa dạng di truyền cao của các nhóm di truyền trên các vị trí SSR với dị hợp tử kỳ vọng ($H_e = 0,78$) lớn hơn dị hợp tử quan sát ($H_o = 0,66$), do đó tần số allele giữa các nhóm di truyền là không đồng nhất và hệ số nội phối bên trong quần thể (F_{st}) có sự khác biệt rất ý nghĩa thống kê (Bảng 3). Tương tự, đối với nguồn gen Rondonia hoang dại được bảo tồn ở Việt Nam gồm 951 cá thể, tổng số 458 allele đã phát hiện với trung bình là 16,2 allele; đa dạng di truyền cao của tất cả các nhóm di truyền trên các vị trí SSR, với dị hợp tử kỳ vọng (H_e) đạt 0,80 và dị hợp tử quan sát (H_o) đạt 0,65 (Bảng 3).

Mức độ đa hình và đa dạng di truyền cao của các nhóm di truyền trên các vị trí chỉ thị SSR được phát

hiện từ nguồn gen cao su của Việt Nam. Nghiên cứu này khá tương đồng với kết quả từ những nghiên cứu trước đó khi phân tích trên các nguồn di truyền Amazon hoang dại với số lượng mẫu giống lớn và đa dạng. Theo Souza và ctv (2015) sử dụng 13 chỉ thị SSR để phân tích cho 1.117 mẫu giống từ 9 nhóm di truyền của 7 bộ sưu tập tại Brazil, tổng số 408 allele được phát hiện với trung bình trên tất cả các vị trí là 14,5 allele, dị hợp tử (H_o và H_e) là 0,64 và 0,76; Le Guen và ctv (2009) sử dụng 18 chỉ thị SSR để phân tích cho 307 cá thể từ 19 vị trí sưu tập, giá trị đa hình khá cao với trung bình đạt 21,7 allele và trung bình dị hợp tử quan sát (H_o) đạt 0,59 trên cả các vị trí SSR; cũng theo Lekawipat và ctv (2003), số lượng allele trung bình trên 12 vị trí SSR là 14 allele từ 108 mẫu giống được phân tích. Như vậy, 15 chỉ thị SSR này được cho là rất tốt và đáng tin cậy để đánh giá về đa dạng di truyền cho các nguồn gen cao su ở Việt Nam.

Bảng 3. Đa hình và đa dạng di truyền của các nguồn gen trên các vị trí chỉ thị SSR được phân tích từ bộ mẫu giống của 11 nhóm di truyền

Chỉ thị SSRs	Na	Số allele	H_o	H_e	F_{st}
A2365	13,5	34	0,50	0,74	0,12***
A2368	11,8	21	0,60	0,79	0,12***
A2387	8,5	17	0,66	0,75	0,10***
A2406	18,7	42	0,68	0,84	0,09***
A2508	18,0	42	0,84	0,85	0,06***
A2684	17,7	39	0,71	0,84	0,11***
A2689	13,3	31	0,63	0,79	0,14***
A2736	12,5	22	0,76	0,84	0,07***
A31	15,9	29	0,86	0,87	0,06***
BAC55B02	7,5	15	0,39	0,64	0,13***
M574	19,2	42	0,79	0,88	0,07***
MnSod	20,5	47	0,84	0,88	0,07***
T2083	3,6	16	0,29	0,30	0,18***
TA2163	15,3	35	0,44	0,82	0,11***
TAs2558	16,0	35	0,84	0,87	0,06***
Trung bình	14,1		0,66	0,78	0,09
Sai số chuẩn (SE)			0,05	0,04	0,01
Khoảng tin cậy			[0,55-0,75]	[0,69-0,84]	[0,08-0,11]
Tổng số allele		467			
Quần thể Rondonia*	16,2	458	0,65	0,80	0,06

- Na: Số allele trung bình trên mỗi vị trí SSRs; H_o : Dị hợp tử quan sát; H_e : Dị hợp tử kỳ vọng; F_{st} : Hệ số nội phối bên trong quần thể với xác suất $P(F_{st}) \leq 0,001$.

3.3. Đánh giá đa dạng di truyền của các quần thể cây cao su

Đa dạng di truyền được đánh giá trên bộ mẫu gồm 1.013 cá thể từ 11 nhóm di truyền khác nhau,

trong đó 951 cá thể thuộc 8 nhóm thuộc nguồn gen Rondonia và 62 mẫu giống thuộc ba nhóm di truyền là Acre, Mato Grosso và Wickham. Tỷ lệ đa hình trên vị trí chỉ thị SSR của các nhóm di truyền đạt 98,8%, nhưng 8 nhóm của nguồn gen Rondonia đạt 100% và nguồn gen Acre cũng có tỷ lệ đa hình đạt 100%. Trong khi tỷ lệ đa hình trên các vị trí chỉ thị SSR của nguồn gen Mato Grosso và Wickham là 93,9% (Bảng 4). Ở các nhóm di truyền, số lượng allele khá phong phú với trung bình chung trên tất cả các nhóm đạt 14,1 allele và dao động từ 6,7 đến 22,0 allele; các nhóm di truyền của nguồn gen Rondonia có số lượng allele dao động từ 11,0 đến 22,0 với trung bình đạt 16,2 allele, trong khi số lượng allele thấp hơn ở nguồn gen Mato Grosso và Wickham tương ứng 6,7 allele và 9,2 allele. Bên cạnh đó, tổng số 99 allele cá thể (allele hiếm) được phát hiện đều có mặt trên tất cả các nhóm di truyền, chủ yếu allele hiếm tập trung ở các nhóm của nguồn gen Rondonia và số lượng allele hiếm phong phú nhất là nhóm di truyền RO/C (27 allele) và RO/CM (28 allele). Các nguồn gen cao su ở Việt Nam có đa dạng di truyền cao với dị hợp tử kỳ vọng ($H_e = 0,78$) cao hơn dị hợp tử quan sát ($H_o = 0,66$) ở tất cả các nhóm, do đó chỉ số cố định (F) là dương. Chỉ số cố định hay hệ số nội phối bên trong các nhóm di truyền là dương đã cho thấy rằng giữa các cá thể trong các quần thể có sự giao phối ngẫu nhiên và giữa các cá thể trong các nhóm di truyền chưa xảy ra hiện tượng cận huyết, ngay cả nguồn gen Wickham đã qua quá trình chọn tạo giống lâu dài mặc dù chỉ số (F) rất nhỏ 0,02 (Bảng 4).

So sánh các thông số di truyền từ các nguồn gen cao su ở Việt Nam với những nghiên cứu trước đó, Souza và ctv (2015), phân tích 1.117 mẫu giống từ 9 quần thể khác nhau với 13 chỉ thị SSR đã phát hiện

số lượng allele đa hình trên các quần thể là 14,5 allele, dị hợp tử ($H_o = 0,64$ và $H_e = 0,76$) và nhận thấy nguồn gen hoang dại có đa dạng di truyền hơn nguồn gen đã qua chọn tạo, số lượng allele hiếm có ở nguồn gen Amazon mà không có trên các dòng vô tính Wickham; Le Guen và ctv (2009) đã sử dụng 15 chỉ thị SSR để phân tích 220 cá thể từ 14 quần thể, mức độ đa hình cao trên các vị trí SSR là 7,5 allele, trung bình dị hợp tử ($H_o = 0,63$ và $H_e = 0,74$). Ngoài ra, nhiều nghiên cứu cũng đã chứng minh nguồn gen Amazon hoang dại đa dạng di truyền hơn nguồn gen Wickham. Theo Lekawipat và ctv (2003), sử dụng 12 chỉ thị SSR để phân tích 40 cá thể Wickham và 68 cá thể Amazon đã cho thấy nguồn gen hoang dại đa hình hơn Wickham và chia các mẫu giống thành ba nhóm mà chúng phụ thuộc vào vị trí địa lý được sưu tập; Oktavia và ctv (2017) đã phát hiện nguồn gen IRRDB'81 đa dạng di truyền hơn Wickham và cho rằng nguồn gen IRRDB'81 có thể được sử dụng để làm tăng vốn di truyền cho công tác cải tiến giống cao su. Một nghiên cứu khác tại Việt Nam, phân tích nguồn gen IRRDB'81 gồm 59 mẫu giống từ ba quần thể gồm Acre, Mato Grosso và Rondonia với chỉ thị RAPD đã cho thấy dị hợp tử là rất thấp chỉ đạt 0,20 (Lai và ctv, 2009). Vì vậy, bộ sưu tập quỹ gen cao su ở Việt Nam đã phát hiện có đa dạng di truyền cao khi phân tích với chỉ thị SSR, nguồn gen Rondonia đa dạng di truyền hơn nguồn gen Wickham và Mato Grosso, các nhóm di truyền trong nguồn gen Rondonia có số lượng allele hiếm rất phong phú và giữa các cá thể trong các nhóm di truyền chưa xảy ra hiện tượng cận huyết. Do đó, nguồn gen Rondonia là rất quan trọng để cải tiến giống cao su trong tương lai.

Bảng 4. Đánh giá đa dạng di truyền của các quần thể cây cao su dựa trên 15 chỉ thị SSR

Quần thể	Na	Ho	He	F	Pa	Đa hình trên các vị trí SSR (%)
RO	11,0	0,69	0,81	0,15	1	100
RO/A	16,4	0,63	0,79	0,20	9	100
RO/C	22,0	0,64	0,84	0,23	27	100
RO/CM	20,5	0,67	0,78	0,14	28	100
RO/J	16,5	0,69	0,83	0,17	13	100
RO/JP	15,3	0,65	0,79	0,19	5	100
RO/OP	11,7	0,63	0,80	0,22	4	100
RO/PB	15,9	0,62	0,78	0,21	7	100
AC	10,3	0,70	0,79	0,11	3	100
MT	6,7	0,62	0,67	0,08	1	93,3
W	9,2	0,68	0,69	0,02	1	93,3

Trung bình	14,1	0,66	0,78	0,16		98,8
Sai số chuẩn (SE)	0,55	0,02	0,01	0,02		0,81
Tổng allele cá thể					99	
Quần thể Rondonia*	16,2	0,65	0,80	0,19		

Na: Số allele trung bình; Ho: Dị hợp tử quan sát; He: Dị hợp tử kỳ vọng; F: Chỉ số cố định; Pa: Số allele cá thể.

3.4. Khác biệt di truyền giữa các quần thể trong bộ sưu tập quỳ cao su ở Việt Nam

Đã tìm hiểu về sự khác biệt di truyền ở mức độ phân tử giữa các quần thể, đặc biệt là các nhóm di truyền của nguồn gen Rondonia được bảo tồn ở Việt Nam. Phân tích phương sai phân tử (AMOVA) đã phát hiện biến dị di truyền chủ yếu xảy ra bên trong cá thể (73,6%) mà không phải là giữa các cá thể (19%) và giữa các nhóm di truyền (7,4%) (Hình 4A), nhưng giữa các nhóm cũng có sự khác biệt di truyền. Ở cây cao su, quan hệ di truyền có liên quan chặt chẽ với vị trí địa lý giữa các quần thể và được đánh giá thông qua sự khác biệt di truyền. Khác biệt di truyền giữa các quần thể được kiểm tra dựa trên 1.000 hoán vị ngẫu nhiên, kết quả đã cho thấy giữa các cặp có sự

khác biệt di truyền rất ý nghĩa ở mức $P \leq 0,001$; nguồn gen Mato Grosso và Wickham khác biệt di truyền lớn hơn các nhóm di truyền khác và cũng như với nguồn gen Acre, dao động từ 0,07 đến 0,10. Trong cùng nguồn gen Rondonia, giữa các nhóm có quan hệ gần gũi nhau với khác biệt di truyền là rất thấp từ 0,02 đến 0,06 (Bảng 5). Ngoài ra, phân ly giữa các nhóm di truyền (PCoA) đã chia toàn bộ các quần thể thành hai cụm, một cụm là nguồn gen Mato Grosso và Wickham, một cụm khác là các nhóm của nguồn gen Rondonia và Acre. Điều đó có nghĩa là giữa Mato Grosso và Wicham có quan hệ gần gũi di truyền hơn so với các nguồn gen khác, giữa các nhóm của nguồn gen Rondonia có mối quan hệ gần gũi di truyền với nhau (Hình 3B).

Bảng 5. Sự khác biệt di truyền (Fst) giữa các cặp quần thể thông qua phân tích AMOVA

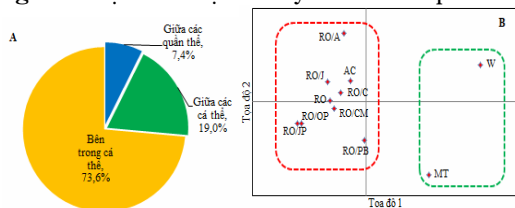
Quần thể	AC	MT	RO	RO/A	RO/C	RO/CM	RO/J	RO/JP	RO/OP	RO/PB
MT	0,09									
RO	0,04	0,08								
RO/A	0,06	0,09	0,04							
RO/C	0,05	0,08	0,04	0,05						
RO/CM	0,04	0,08	0,03	0,05	0,04					
RO/J	0,04	0,08	0,03	0,03	0,03	0,03				
RO/JP	0,06	0,10	0,04	0,06	0,04	0,03	0,03			
RO/OP	0,05	0,09	0,03	0,06	0,04	0,03	0,03	0,02		
RO/PB	0,06	0,07	0,04	0,05	0,04	0,04	0,04	0,05	0,05	
W	0,07	0,08	0,07	0,08	0,07	0,07	0,07	0,10	0,09	0,07

Xác suất $P \leq 0,001$ tương ứng giữa các cặp quần thể dựa trên 1.000 hoán vị.

Hầu hết các nghiên cứu đã chứng minh về mối quan hệ thân thuộc di truyền giữa các quần thể cây cao su có liên quan chặt chẽ với vị trí địa lý; đối với nguồn gen Amazon hoang dại, biến dị di truyền chủ yếu xảy ra bên trong cá thể, nhưng giữa các quần thể cũng có sự khác biệt di truyền. Souza và ctv (2015), biến dị di truyền bên trong cá thể là 73%, giữa các cá thể là 21% và giữa các quần thể là 7%. Cũng theo Souza và ctv (2015), toàn bộ 1.117 mẫu giống từ 9 nhóm di truyền của 7 bộ sưu tập được chia thành hai cụm, một cụm là những giống trồng thuộc nguồn gen Mato Grosso và Wickham, một cụm là các mẫu thuộc nguồn gen hoang dại với khác biệt di truyền thấp giữa hai cụm ($Gst = 0,018$) và cho rằng do chúng ở các lưu vực sông Amazon khác nhau, kết

quả đã chứng thực cho nghiên cứu trước đó của Le Guen và ctv (2009). Theo Le Guen và ctv (2009) đã phân tích 220 cá thể từ 14 quần thể, kết quả đã chứng minh sự phân chia giữa các quần thể như Acre, Rondonia và Mato Grosso là phù hợp với ranh giới của lưu vực sông Amazon và đã giải thích về sự khác nhau giữa các quần thể là do vị trí địa lý bên trong mạng lưới thủy văn và sự cách ly giữa các quần thể ở các lưu vực khác nhau với hệ số khác biệt di truyền thấp ($Fst = 0,12$) giữa các nhóm. Ngoài ra, Lekawipat và ctv (2003) phân tích từ 40 dòng vô tính Wickham và 68 mẫu giống IRRDB'81, kết quả cũng đã chứng minh nguồn gen hoang dại đa hình hơn các dòng vô tính Wickham và chia thành ba cụm tùy thuộc vào nguồn gốc địa lý của vùng sưu tập như Acre,

Rondonia và Mato Grosso. Trước đó, nguồn gen IRRDB'81 của Việt Nam được phân tích với chỉ thị RAPD, biến dị di truyền xảy ra bên trong cá thể 86% và 14% giữa các quần thể (Lai và ctv, 2009). Theo đó, các bộ sưu tập quỹ gen cao su ở Việt Nam hiện đang sở hữu tiềm năng di truyền rất lớn vẫn chưa khai thác, quan hệ di truyền giữa các nhóm di truyền có liên quan chặt chẽ với vị trí địa lý được sưu tập; nguồn gen Mato Grosso và Wickham có mối quan hệ di truyền gần gũi hơn nguồn gen Rondonia, giữa các nhóm của nguồn gen Rondonia có quan hệ di truyền gần gũi với sự khác biệt di truyền là rất thấp.



Hình 4. Biến dị di truyền và sự phân ly giữa nhóm di truyền. (A) tỷ lệ biến dị di truyền được phân tích AMOVA; (B) phân bố của các nhóm di truyền dựa trên phân tích PCoA

4. KẾT LUẬN

- Chỉ thị SSR đã xác định được 200 mẫu có quan hệ thân thuộc di truyền (100%) với 85 cặp và 10 bộ ba; 7 mẫu không thuộc nguồn gen Rondonia trong số 1.127 mẫu giống được phân tích.

- Các nguồn gen cao su ở Việt Nam có đa dạng di truyền cao, nguồn gen Rondonia đa dạng di truyền hơn Wickham và Mato Grosso; số lượng allele hiếm khá phong phú trên các nhóm của nguồn gen Rondonia và giữa các cá thể trong tất cả các nhóm di truyền chưa xảy ra cận huyết.

- Quỹ gen cao su ở Việt Nam đang sở hữu tiềm năng di truyền rất lớn với biến dị di truyền bên trong cá thể là 74%, quan hệ thân thuộc có liên quan đến vị trí địa lý của các bộ sưu tập.

- Nguồn gen Mato Grosso và Wickham có quan hệ gần gũi di truyền hơn các nguồn gen khác và giữa các nhóm trong cùng nguồn gen Rondonia là gần gũi di truyền với nhau.

TÀI LIỆU THAM KHẢO

1. ANRPC (2018). Natural Rubber Trends & Statistics. Vol.10, No.6, August 2018. <http://anrpc.org/>
2. Backhaus R. A. (1985). Rubber formation in plants a mini-review. Israel J. Bot. 34: 283 - 293.

3. Bowers J. E. (1990). Natural rubber producing plants for the United States. Proc. Natl. Acad. Sci. U.S.A. 100: 1352 - 1357.

4. Chanroj V., Rattanawong R., Phumichai T., Tangphatsornruang S. and Ukoskit K. (2017). Genome-wide association mapping of latex yield and girth in Amazonian accessions of *Hevea brasiliensis* grown in a suboptimal climate zone. Genomics 109:475-484.

5. Chevallier M. H. (1988). Genetic variability of *Hevea brasiliensis* germplasm using isozyme markers. Journal of Natural Rubber Research, 3:42-53.

6. Clement-Demange A., Legnate H., Chapuset T., Pinard F. and Seguin M. (1998). Characterisation and use of the IRRDB germplasm in Ivory Coast and French Guyana: status in 1997. Symposium on natural rubber (*Hevea brasiliensis*) Volume 1 General, soils and fertilization, and breeding and selection sessions, Ho Chi Minh City, Vietnam, 14-15 october 1997. IRRDB, Brickendonbury (UK).

7. Compagnon P. (1986). Le Caoutchouc naturel: Biologique - Culture - Production. Techniques agricoles et productions tropicales. Maisonneuve et Larose, Paris, France, pp. 12-13.

8. Dean W. (1987). Brazil and the struggle for rubber. Cambridge University Press, Cambridge.

9. Lai V. L., Tran T., Vu T. Q., Le M. T. (2009). Genetic diversity of *Hevea* IRRDB'81 collection assessed by RAPD markers. Molecular biotechnology 42:292-298. Doi: 10.1007/s12033-009-9159-7.

10. Lai V. L., Tran T., Le T. T. T., Vu V. T., Huynh B. L., Le L. T. (2012). *Hevea* germplasm in Vietnam: conservation, characterization, evaluation and utilization. In: Caliskan M (ed) Genetic diversity in plants. In Tech, Shanghai, pp 433-456.

11. Le Guen V., Doare F., Weber C. and Seguin M. (2009). Genetic structure of Amazonian populations of *Hevea brasiliensis* is shaped by hydrographical network and isolation by distance. Tree Genetics and Genomes, 5: 673-683. doi: 10.1007/s11295-009-0218-9.

12. Le Guen V., Garcia D., Doaré F., Mattos C. R. R., Condina V., Couturier C. (eds). (2011). A rubber tree's durable resistance to *Microcyclus ulei* is conferred by a qualitative gene and a major quantitative resistance factor. Tree Genet. Genomes 7, 877-889. doi: 10.1007/s11295-011-0381-7.

13. Lekawipat N., Teerawatanasuk K., Rodier-Goud M., Seguin M., Vanavichit A., Toojinda T. and Tragoonrung S. (2003). Genetic Diversity Analysis of Wild Germplasm and Cultivated Clones of *Hevea brasiliensis* Muell. Arg. by Using Microsatellite Markers. *Journal of rubber research* 6: 36-47.
14. Lespinasse D., Rodier-Goud M., Grivet L., Leconte A., Legnaté H., Seguin M. (2000a). A saturated genetic linkage map of rubber tree (*Hevea* spp.) based on RFLP, AFLP, microsatellite and isozyme markers. *Theor. Appl. Genet.* 100:127-138.
15. Low F. C., Safiah A., Hafsah J. and Tan H. (1996). Recent advances in the development of molecular markers for *Hevea* studies. *J Nat Rubber Res* 11:32-44.
16. Murray M. G. and Thompson W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* 8:4321-4326.
17. Nei M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of National Academy of Science, USA* 70:3321-3323.
18. Oktavia F., Kuswanhadi, Dinarty D., Widodo and Sudarsono (2017). Genetic diversity and population structure of IRRDB 1981 and Wickham rubber germplasm based on EST-SSR. *AGRIVITA Journal of Agricultural Science*, 39(3), 239-251. <http://doi.org/10.17503/agrivita.v39i3.881>.
19. Ong S. H. (1979). Cytotaxonomic investigation of the genus *Hevea*. PhD thesis, University of Malaysia.
20. Ong S. H., Othman R. and Benong M. (1995). Status report on the 1981 *Hevea* germplasm collection 1995. Symposium on Physiological and Molecular Aspects of the Breeding of *Hevea brasiliensis*, Penang, Malaysia 6-7 November 1995, pp 95-105.
21. Peakall R. and Smouse P. E. (2012). Genalex 6.5: Genetic analysis in Excel. Population genetic software for teaching and research an update. *Bioinformatics*. 28:2537-2539. PMID: 22820204.
22. Perrier X. and Jacquemoud-Collet J. P. (2006). Darwin software. 5.0 edn. <http://darwin.cirad.fr/darwin>
23. Schuelke M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nat Biotechnol.* 18: 233-234. PMID: 10657137.

GENETIC DIVERSITY OF *HEVEA* GERmplasm CONSERVED IN VIETNAM

Vu Van Truong, Huynh Duc Dinh, Nguyen Thi Thao,

Tran Thanh, Huynh Thi Minh Tam, Ronan Rivallan, Huynh Van Biet, Le Guen Vincent

Summary

Natural rubber is a strategic raw material in many industries and is almost exclusively produced by cultivating rubber trees (*Hevea brasiliensis* Muell. Arg.). Breeding of this species focuses mainly on improved varieties in Asia and Africa; concurrently, it also relies on genetic diversity of wild germplasm. This study aims to assess the genetic diversity of accessions from the state of Rondonia in Brazil and at present they are being preserved in Vietnam with 15 SSR markers. A total of 1,127 accessions were genotyped to identify 200 accessions with 100% genetic relation from 85 pairs and 10 triplets of genotypes, as well as to identify 7 accessions which do not belong to the Rondonia population. As a result, the number of allelic richness on SSR sites ranged from 15 to 47 alleles and the allele frequency among genetic groups was different; a total of 467 allele were from 11 genetic groups and 458 allele were from Rondonia population. The level of high genetic diversity was in all groups, the expected heterozygosity ($H_e = 0.78$) was higher than observed heterozygosity ($H_o = 0.66$) in all cases, and the inbreeding depression did not occur among individuals. The Rondonia population was more genetically diverse than Wickham and Mato Grosso populations; the Wickham and Mato Grosso populations had a closely genetic relationship, and genetic differentiation (F_{st}) of these populations was higher than that of other genetic groups while there was very little genetic differentiation among Rondonia subpopulations. The *Hevea* germplasm in Vietnam had a great genetic potential with most of the genetic variance occurred within accessions (74%), but the genetic differentiation among the populations was significant, and the genetic structure was relevant to collected locations.

Keywords: Conservation, genetic differentiation, genetic diversity, genetic resource, genetic structure, rubber tree.

Người phản biện: PGS.TS. Lê Tuấn Nghĩa

Ngày nhận bài: 20/5/2020

Ngày thông qua phản biện: 22/6/2020

Ngày duyệt đăng: 29/6/2020