

NGHIÊN CỨU SỰ ĐA DẠNG TRONG HỆ GEN CỦA MỘT SỐ GIỐNG CHÈ ĐẶC SẢN TRỒNG TẠI THÁI NGUYÊN BẰNG KỸ THUẬT SSR

TỔNG QUAN

Chè là thứ nước uống tự nhiên lâu đời nhất, được tiêu thụ rộng rãi với chi phí thấp, được trồng chủ yếu ở các nước nhiệt đới của Châu Á (Ấn Độ, Sri Lanka, China, Indonesia, Việt Nam..), Châu Phi (Kenya, Uganda, Malawi..) và Châu Mỹ la tinh (Argentina). Cây chè có tên khoa học là *Camellia sinensis* (L) O. Kuntze, thuộc chi chè (*Camellia*), họ chè (*Theaceae*), bộ chè (*Theales*), lớp ngọc lan (*Dicotyledonea*), ngành ngọc lan (*Angiospermae*) (Min et al., 2007). Theo Wight (1962), chi chè (*Camellia*) được phân loại thành 3 loài đó là chè Trung Quốc (*Camellia sinensis* - China), Ấn Độ (*Camellia assamica* - Assam) và chè trung gian giữ a chè Trung Quốc và Ấn Độ (*Camellia assamica lasiocalyx* - Campod). Ba loài chè này là nguồn gốc cho hầu hết các loài chè trồng ở các nước trên thế giới hiện nay. Theo thống kê của Min và đồng tác giả (2007) có khoảng 120 loài chè được trồng ở các nước châu Á, Trung Quốc có tới 97 loài, trong đó chè *Camellia sinensis* (L) O. Kuntze được phân thành 4 thứ khác nhau, đó là thứ chè *Camellia sinensis* var *sinensis*, *Camellia sinensis* var *pubilimba*, *Camellia sinensis* var. *Assamica* và *Camellia sinensis* var. *dehungensis* (Min et al., 2007). Đặc điểm về giống và chu kỳ sống của chè phức tạp tạo nên một số hạn chế cho chọn tạo và cải tiến giống theo phương pháp truyền thống. Việc phân biệt giữa các thứ chè thuộc chè China, Assam và Compod gặp khó khăn do tính chất không đồng nhất của chè (Visser, 1996). Hơn nữa, các mô tả về hình thái không phản ánh các biến đổi về mặt di truyền trong cây trồng mà chỉ cho thấy khi mô tả về sinh hóa và sinh lý (Sharman et al., 2009). Do đó, việc sử dụng các chỉ thị phân tử ở mức độ DNA được các nhà khoa học quan tâm sử dụng để phát hiện các biến đổi về mặt di truyền.

Trên thế giới đã có nhiều nghiên cứu đa dạng cây chè ở mức độ phân tử. Genome chè đã được nghiên cứu bằng cách sử dụng các chỉ thị phân tử như chỉ thị RAPD {Chen et al., 2002, Francis et al., 1997, Li et al., 2005, Lin et al., 2005, Wachira et al., 2001}; chỉ thị AFLP {Ji et al., 2009, Mishra et al., 2009, Paul et al., 1997, Wachira et al., 2001}; chỉ thị RFLP {Matsumoto et al., 1994}; chỉ thị SSR {Sharman et al., 2009, Ma et al., 2010, Sharma et al., 2011, Sahu et al., 2012}. Trong đó SSR (single sequence region – SSR) còn được gọi là DNA vệ tinh là một đoạn DNA có sự lặp lại của một trình tự nucleotide nào đó. Hiện tượng tồn tại các SSR trong cơ thể sinh vật nhân chuẩn là khá phổ biến. Tuy nhiên, tùy từng loài mà số lượng nucleotide trong mỗi đơn vị lại có thể thay đổi từ một đến hàng chục và số lượng đơn vị lặp lại có thể biến động từ 2 đến hàng ngàn lần hoặc nhiều hơn. Các đoạn lặp lại hai, ba, bốn nucleotide phân bố phổ biến trong genome của các loài động thực vật. Kỹ thuật PCR khuếch đại đoạn SSR (PCR-SSR) với mỗi thiết kế dựa trên trình tự genome của từng loài nên chỉ thị SSR cho kết quả nhanh và cao. Chỉ thị phân tử SSR với nhiều đặc tính như độ đa dạng cao, phân bố rộng rãi trong genome, di truyền theo quy luật Mendel, biểu hiện đồng hợp trội. Trình tự SSR được xác định từ trình tự DNA genome và cDNA (EST – Expression Sequence Tag). Chỉ thị SSR có nguồn gốc từ cDNA/EST (EST-SSR) có thể giúp giảm chi phí và thời gian xây dựng chỉ thị Do các đoạn cDNA/ EST liên quan trực tiếp đến các gene chức năng, nên SSR phát triển từ nguồn dữ liệu này có thể sẽ hữu ích hơn (Sharman et al., 2009).

Trong những năm gần đây, những phân tích cDNA/EST của chè đã tạo ra nguồn thông tin khá lớn. Nhiều chỉ thị phân tử SSR được phát triển từ cơ sở dữ liệu cDNA/ EST. Năm 2009, nhóm nghiên cứu của Sharma đã thống kê có 2181 đoạn EST từ cơ sở dữ liệu NCBI có 1223 gen có

trình tự SSR từ 2-34 nucleotide. Trong 109 gen phân tích nghiên cứu có chứa 120 SSR được xác định. Phân tích dữ liệu motif SSR trong các gen này đã phát hiện 61 lặp lại 2 (50.8%), 37 lặp lại 3 (30.8%), 8 lặp lại 4 (6.67%), 9 lặp lại 5 (7.5%) và 5 lặp lại 6 (4.16%) (Sharman et al., 2009). Theo thống kê của Ma và cộng sự (2010), ở chè có 6899 đoạn EST được công bố trên ngân hàng gen và có thêm 74 chỉ thị SSR – EST mới được nghiên cứu thực nghiệm (Ma et al., 2010). Nhóm nghiên cứu của Sharma (2011) đã phát hiện thêm được 112 chỉ thị SSR mới (Sharma et al., 2011). Năm 2012, nhóm tác giả Sahu thống kê được 12852 đoạn EST ở chè đã công bố, theo tính toán lý thuyết 1636 đoạn EST phân tích chứa 2371 SSR (Sahu et al., 2012). Khi so sánh các đoạn EST ở chè với các gen đã biết chức từ các loài khác cho thấy hầu hết các chỉ thị SSR-EST được nghiên cứu cho đến nay được cho rằng có liên quan đến các quá trình sinh học, thành phần tế bào và chức năng phân tử ở chè. Các chỉ thị SSR-EST liên quan đến quá trình sinh học ở chè bao gồm: quá trình trao đổi chất, protein đáp ứng lại stress, tổ chức tế bào. Có 3 dạng chỉ thị SSR-EST liên quan đến thành phần tế bào là các đoạn SSR mã hóa protein tham gia cấu tạo màng tế bào, nội bào và ngoại bào. Khi phân tích dưới dạng chức năng phân tử, các chỉ thị SSR-EST dựa trên các đoạn DNA mã hóa protein cấu trúc, protein bám, protein xúc tác (Sharman et al., 2009, Ma et al., 2010).

Hiện nay, Việt Nam là một trong 10 quốc gia đứng đầu thế giới về diện tích và sản lượng chè, đứng thứ 8 về xuất khẩu chè, cả nước có 34 tỉnh thành sản xuất chè. Các giống chè được trồng chủ yếu là chè Trung du, chè Shan và các giống chè mới. Các giống chè được trồng ở Việt Nam có thể được chọn lọc dựa trên các đặc điểm hình thái, phương pháp lai tạo, gây đột biến bằng bức xạ, hóa chất và phương pháp nhân giống vô tính bằng cách ghép, giâm cành (Hoàng Văn Chung, 2012). Ở nhiều địa phương chè còn được trồng tự hạt. Do đó, nguồn genome chè rất đa dạng và phong phú. Các công trình nghiên cứu về cây chè chủ yếu đi sâu nghiên cứu đặc tính hoá sinh, đặc điểm hình thái, giải phẫu lá, thân, đặc điểm sinh trưởng, phát triển, năng suất, chất lượng và chọn tạo giống chè bằng phương pháp truyền thống. Việc ứng dụng các kỹ thuật sinh học phân tử vào việc đánh giá hệ gen của cây chè trong chọn tạo giống cây trồng còn là vấn đề mới mẻ. Năm 2004, nhóm tác giả Nguyễn Minh Hùng, Đinh Thị Phòng đã sử dụng kỹ thuật RAPD để nghiên cứu tính đa hình của một số dòng chè đột biến. Các tác giả đã sử dụng 9 cặp mồi ngẫu nhiên để nghiên cứu trên 14 dòng chè đột biến, kết quả có 94 phân đoạn DNA được tạo ra trong đó 85 phân đoạn đa hình (Nguyễn Minh Hùng et al., 2004). Năm 2010, Nguyễn Thị Thu Hương và đtg đã sử dụng kỹ thuật này để phân tích sự đa dạng trình tự genome ở các dòng chè Shan (Nguyễn Thị Thu Hương et al., 2010). Nhóm nghiên cứu này đã bước đầu nghiên cứu sự đa dạng genome một số giống chè trồng ở xã Tân Cương - vùng chè đặc sản chè nổi tiếng của Tỉnh Thái Nguyên cũng như của nước ta (Bảo Lâm 2011, Thanh Thuy 2012, Hoàng Thị Thu Yến et al., 2012).

MỤC TIÊU

- Xác định được mức độ đa dạng trình tự SSR trong genome chè
- Xác định được khoảng cách di truyền của một số giống chè đặc sản trồng ở Thái Nguyên

NỘI DUNG

1) Thu thập một số giống chè đặc sản trồng tại Thái Nguyên

Thu thập các giống chè được đánh giá có chất lượng tốt và là đặc sản được trồng tại một số vùng trồng chè ở Thái Nguyên

2) Sàng lọc các cặp mồi SSR cho chè và xác định mồi SSR đặc hiệu

- Thực hiện kỹ thuật PCR-SSR với nhiều cặp mồi để đánh giá sự đa dạng trong trình tự genome của chè
- Thực hiện kỹ thuật PCR-SSR với các cặp mồi liên quan đến các protein cấu trúc, protein tham gia vào quá trình trao đổi chất, quá trình sinh trưởng và khả năng chống chịu ở chè ở chè
- 3) Phân tích hệ số đa dạng ở mức độ DNA và mối quan hệ di truyền giữa các giống
 - Sử dụng phần mềm sinh học phân tử NTSYS version 2.0 để phân tích kết quả của PCR – SSR để chỉ ra được mối quan hệ di truyền giữa các giống nghiên cứu
- 4) Phân lập, tách dòng và so sánh trình tự đoạn SSR liên quan đến sự mã hóa protein tham gia vào con đường trao đổi chất ở chè.
 - Tách dòng và xác định trình tự đoạn SSR
 - Phân tích trình tự thu được bằng cách sử dụng phần mềm sinh học phân tử Blast, bioedit...

PHƯƠNG PHÁP NGHIÊN CỨU

a) Điều tra các giống chè đặc sản được trồng ở Thái Nguyên

b) Phương pháp thu mẫu lá chè

c) Phương pháp sinh học phân tử

- Kỹ thuật tách chiết DNA tổng số từ lá chè
- Kỹ thuật điện di
- Kỹ thuật PCR-SSR
- Kỹ thuật tách dòng phân tử
- Kỹ thuật xác định trình tự

d) Phương pháp phân tích số liệu

- Phần mềm NTSYS version 2.0
- Phần mềm Blast
- Phần mềm Bioedit

HIỆU QUẢ KTXH

Sản phẩm khoa học:

- Số bài báo khoa học đăng trên tạp chí trong nước: 3
- Đăng ký 2 trình tự gen trên Genbank

Sản phẩm đào tạo:

- Số nhóm sinh viên NCKH: 2
- Thạc sỹ: 1

Sản phẩm ứng dụng:

- Thống kê được sự đa dạng về các dòng chè trồng tại tỉnh thái nguyên
- Nghiên cứu được sự đa dạng di truyền genome của các dòng chè thu được bằng kỹ thuật sinh học phân tử từ đó góp phần đánh giá mối quan hệ di truyền và làm chỉ thị để đánh giá chọn lọc các dòng chè có chất lượng tốt

ĐƠN VỊ SỬ DỤNG