

Biên soạn: **TS. HUỖNH KỶ** (Chủ biên)
TS. NGUYỄN CHÂU THANH TÙNG - TS. NGUYỄN LỘC HIỂN
ThS. HUỖNH NHƯ ĐIỂN - TS. PHẠM THỊ BÉ TỬ
KS. LÊ THỊ HỒNG THANH - KS NGUYỄN VĂN MẠNH
KS. VĂN QUỐC GIANG - KS. TRẦN IN ĐỒ
KS. CHUNG TRƯỞNG QUỐC KHANG

CHỈ THỊ GEN CHỨC NĂNG TRONG CHỌN GIỐNG LÚA



NHÀ XUẤT BẢN ĐẠI HỌC CẦN THƠ
2021

LỜI GIỚI THIỆU

Đồng bằng sông Cửu Long (ĐBSCL) có thể được xem là cái nôi của nền nông nghiệp của Việt Nam. Vì với chỉ khoảng 4 triệu hecta đất nông nghiệp, ĐBSCL đã cung cấp khoảng 50% sản lượng lương thực cho cả nước. Trong đó, cây lúa được xem là cây chủ lực vì chúng chiếm diện tích lớn đất nông nghiệp trong sản xuất và đang gánh vác một trọng trách rất lớn cho an ninh lương thực quốc gia và còn cung cấp gạo cho thế giới. Tuy nhiên thị hiếu về gạo trên thế giới luôn luôn thay đổi, do đó các nhà chọn giống luôn phải song hành chọn tạo ra những giống lúa phù hợp với thị hiếu người tiêu dùng.

Lịch sử của chọn tạo giống lúa đã xuất từ cách đây hơn 8.000 năm, khi đó nhằm tìm nguồn lương thực cung cấp, con người cổ đại đã thu lượm và chọn lọc các giống lúa hoang dại để thuần hóa, tiến trình này dần dần làm các giống lúa được thuần hóa hạn chế khả năng chống chịu trước điều kiện bất lợi khi tồn tại theo kiểu hoang dại. Vào thời đó thì người cổ đại chọn giống bằng cách cảm nhận các đặc tính mong muốn và chọn lọc, tiến trình kéo dài qua hàng ngàn năm để chọn lọc được các giống mang đặc tính tốt. Do đó tiến trình chọn lọc đã dần dần thu hẹp nguồn gen trong tự nhiên. Tuy nhiên trong phương thức canh tác của người cổ đại như du nhập giống từ nhiều nguồn khác nhau, sự giao phấn trong tự nhiên đã giúp đa dạng hóa nguồn gen, và với sự lựa chọn có ý thức hay vô thức đã tạo ra trên 150.000 giống cây trồng trên thế giới.

Nền tảng của phương thức chọn giống hiện đại được hình thành từ năm 1700 – 1900. Từ khi phát hiện ra giới tính trên cây trồng vào năm 1694 và nhận thấy được ra sự lai tạo ở thực vật vào năm 1719, chính hiện tượng này đã tạo ra rất nhiều các biến dị trong tự nhiên. Sau này cơ sở khoa học của việc chọn giống cây trồng đã được nâng cao rất nhiều trong thế kỷ 20. Thêm vào đó, những bước đột phá khoa học kỹ thuật vào thế kỷ thứ 20 cũng góp phần cho việc phát triển giống mới, một trong những khám phá về DNA là nền tảng của sự di truyền và việc phát hiện cấu trúc của vật chất di truyền đã giúp cho ngành chọn giống bước vào kỷ nguyên mới. Dựa vào nền tảng DNA là vật chất di truyền, các nhà chọn giống đã vận dụng những kỹ thuật hiện đại giúp cho công tác chọn giống được chính xác và nhanh chóng hơn, như kỹ thuật di truyền (chuyên gen), ứng dụng hỗ trợ chỉ thị phân tử trong chọn giống (MAS), trong đó một trong những phương cách chọn giống chính xác và hiệu quả nhất là sử dụng gen chức năng làm dấu chỉ thị phân tử (FM). Trong khuôn khổ quyển sách này nhóm tác giả hy vọng giới thiệu độc giả những kiến thức cơ bản về dấu chỉ thị phân tử là gen chức năng và những ứng dụng thực tế của nó trong chọn giống cây trồng.

Quyển sách này mục đích chính làm tài liệu tham khảo cho sinh viên chuyên ngành Nông nghiệp, Công nghệ Sinh học và Sinh học ứng dụng còn có thể làm tài liệu hướng dẫn cho các Viện trường, trung tâm giống dùng làm tài liệu nghiên cứu giúp công tác chọn giống chính xác hơn nhờ ứng dụng các kỹ thuật hiện đại đã trình bày trong quyển sách này.

Trong quyển tài liệu tham khảo này, nhóm tác giả đã tham khảo nhiều tài liệu chuyên ngành đã được xuất bản hoặc công bố, hầu hết các kết quả nghiên cứu trên đối tượng cây lúa và hình ảnh minh họa là của chính nhóm tác giả (trong phần ứng dụng). Trong quá trình biên soạn, mặc dù đã có nhiều cố gắng nhưng chắc chắn còn những thiếu sót nhất định, rất mong nhận được sự đóng góp chân thành của quý Thầy Cô, các bạn đồng nghiệp, sinh viên và những đọc giả quan tâm có cơ hội sử dụng hoặc tham khảo quyển sách này để có cơ cập nhật và hoàn thiện hơn.

NHÓM TÁC GIẢ

LỜI CẢM ƠN

Chân thành cảm ơn Dự án nâng cấp Trường Đại học Cần Thơ VN14-P6 bằng nguồn vốn vay ODA từ chính phủ Nhật Bản đã tài trợ kinh phí để biên soạn tập sách này.

Để biên soạn hoàn chỉnh cuốn sách này xin chân thành cảm ơn các bạn sinh viên ngành Công nghệ Giống Cây trồng K44, ngành Nông học K43, 44 khoa Nông nghiệp, Trường Đại học Cần Thơ đã hỗ trợ các kết quả phân tích chỉ tiêu chất lượng trên lúa bằng kỹ thuật sinh học phân tử và sinh hóa. Ngoài ra, sự đóng góp rất quan trọng từ quý Thầy Cô Bộ môn Di truyền và Chọn giống Cây trồng cho tập sách này, xin chân thành cảm ơn.

NHÓM TÁC GIẢ

MỤC LỤC

Chương 1. CHỈ THỊ PHÂN TỬ LÀ GEN CHỨC NĂNG	1
1.1 GIỚI THIỆU	1
1.2 LỊCH SỬ PHÁT TRIỂN CỦA CÁC CHỈ THỊ PHÂN TỬ	2
1.3 DẤU CHỈ THỊ CHỨC NĂNG	4
1.4 ƯU ĐIỂM CỦA CHỈ THỊ PHÂN TỬ FM SO VỚI CÁC DẤU CHỈ THỊ KHÁC	8
1.5 FM _s TRONG NHÂN GIỐNG CÂY TRỒNG CHÍNH XÁC	8
1.5.1 Sàng lọc quần thể và đa dạng di truyền	9
1.5.2 Chọn giống dưới sự hỗ trợ của chỉ thị phân tử MAS	10
1.5.3 Tháp Gen	12
1.5.4 Chọn lọc bộ gen	13
1.6 FM _s CẢI THIÊN CÁC ĐẶC TÍNH NÔNG HỌC, ĐẶC TÍNH CHẤT LƯỢNG VÀ KHẢ NĂNG CHỐNG CHỊU STRESS	13
1.6.1 Chỉ thị phân tử FMs cho các đặc tính nông học	14
1.6.2 Chỉ thị phân tử FM dùng trong chọn giống các đặc tính chất lượng	15
1.6.3 FMs dùng trong chọn lọc giống kháng sinh học	16
1.6.4 FMs dùng trong chọn giống chống chịu yếu tố phi sinh học	17
TÀI LIỆU THAM KHẢO	18
Chương 2. CHỈ THỊ GEN CHỨC NĂNG TRONG ĐÁNH GIÁ CHẤT LƯỢNG GIỐNG LÚA NHẬP NỘI	30
2.1 GIỚI THIỆU	30
2.2 PHƯƠNG TIỆN PHƯƠNG PHÁP	32
2.2.1 Vật liệu thí nghiệm	32
2.2.2 Phương pháp phân tích	33
2.3 KẾT QUẢ THẢO LUẬN	37
2.3.1 Đặc tính chiều dài và hình dạng hạt gạo	37
2.3.2 Hàm lượng amylose	40
2.3.3 Nhiệt trở hồ	43
2.3.4 Độ bền thể gel	44
2.3.5 Phân tích mùi thơm	46
2.4 KẾT LUẬN	47
TÀI LIỆU THAM KHẢO	47

Chương 3. CHỈ THỊ GEN CHỨC NĂNG ĐÁNH GIÁ CHẤT LƯỢNG CỦA 20 GIỐNG LÚA RÂY Ở HAI VÙNG SINH THÁI	50
3.1 ĐẶT VẤN ĐỀ	50
3.2 VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU	52
3.2.1 Vật liệu nghiên cứu	52
3.2.2 Phương pháp nghiên cứu	52
3.3 KẾT QUẢ VÀ THẢO LUẬN	55
3.3.1 Điều kiện môi trường ảnh hưởng đến chiều dài hạt gạo	55
3.3.2 Ảnh hưởng của vùng sinh thái lên hàm lượng amylose (AC) trong hạt	57
3.3.3 Đánh giá mùi thơm ở gạo	60
3.4 KẾT LUẬN VÀ ĐỀ NGHỊ	61
3.4.1 Kết luận	61
3.4.2 Đề nghị	61
TÀI LIỆU THAM KHẢO	61

Chương 1

CHỈ THỊ PHÂN TỬ LÀ GEN CHỨC NĂNG

Huỳnh Kỳ*, Nguyễn Châu Thanh Tùng, Nguyễn Văn Mạnh,
Trần In Đô, Chung Trương Quốc Khang

Bộ môn Di truyền và Chọn giống Cây trồng, Khoa Nông nghiệp,
Trường Đại học Cần Thơ

* Tác giả chịu trách nhiệm: Huỳnh Kỳ (Email: hky@ctu.edu.vn)

TÓM LƯỢC

Ngày nay nhờ những tiến bộ vượt trội trong kỹ thuật sinh học phân tử đã giúp giải mã thông tin toàn bộ bộ gen thông qua kỹ thuật giải trình tự thế hệ mới, từ những dữ liệu đó giúp cho việc phát triển cây trồng ngày càng nhanh và chính xác hơn. Tuy nhiên việc xác định gen mang chỉ thị chức năng (FM) hiện diện trong bộ gen có liên quan nhiều đến sự biến đổi kiểu hình của cây là một thách thức lớn. Do đó, nhằm tìm ra được gen chức năng cần tiếp cận các bước thông qua phiên mã nhằm tìm ra biến thể của các gen chức năng, tái tổ hợp tương đồng, lập bản đồ liên kết và sử dụng các alleles đa hình để xác định chỉ thị phân tử FM cho mục tiêu chọn tạo giống, chẳng hạn như đặc điểm nông học, kháng stress sinh học và phi sinh học. Một trong những đặc điểm nổi bật của FM so với chỉ thị phân tử khác được sử dụng trong tạo giống cây trồng là sự liên kết gen chặt chẽ với một kiểu hình. Do đó, FM có thể tạo điều kiện thuận lợi cho việc chọn lọc trực tiếp gen liên quan đến đặc điểm kiểu hình, gia tăng hiệu quả chọn lọc trong chọn giống. Trong chương sách này, chúng tôi tóm lược các phương pháp mới nhất trong phát triển chỉ thị phân tử FM và ứng dụng trong phương pháp chọn lọc hỗ trợ đánh dấu (MAS) trong chọn giống, giúp chọn chính xác các giống mang đặc điểm nông học và tính trạng chất lượng cũng như trong chọn các giống cây trồng kháng stress sinh học và phi sinh học.

1.1 GIỚI THIỆU

Theo dữ liệu dự đoán sự bùng nổ dân số vào năm 2030 thì nhu cầu lương thực cần cung cấp tăng 50%, nhưng nguồn gen và môi trường canh tác ngày càng thu hẹp, do đó các nhà chọn giống gặp rất nhiều khó khăn trong việc chọn tạo giống vừa tăng sản lượng mà đáp ứng được sự tác động của môi trường (Frona *et al.*, 2019). Hơn nữa, nguồn gen ngày càng bị hẹp dần do trong thời gian dài công tác nghiên cứu và hoạt động trồng trọt tập trung chủ yếu vào canh tác các giống cây trồng cao sản, nhằm rút ngắn thời gian và tăng sản lượng (Chen *et al.*, 2014). Trong hoàn cảnh đó, các nhà chọn tạo giống cây trồng phải tìm ra và khai thác nguồn gen từ thông tin bộ gen. Việc áp dụng công nghệ giúp nhà chọn giống tìm ra cách tiếp cận và giải quyết các thách thức an ninh lương thực toàn cầu, phát triển bền vững (Oliver, 2014).

Từ xưa đến nay chọn giống theo phương pháp truyền thống là cho hai cá thể mang các kiểu gen ưu việt được thể hiện qua kiểu hình và chọn lọc lại các cá thể con cái mang các kiểu hình ưu việt đó. Trong thực tế, chọn giống cây trồng là chọn lọc các sản phẩm của sự tái tổ hợp kiểu gen và dưới tác động của điều kiện môi trường. Thực vậy, sự tái tổ hợp dẫn đến trao đổi vật chất di truyền là cơ sở dùng để cải tiến và phát triển giống cây trồng mới phù hợp với nhu cầu xã hội và môi trường tự nhiên. Sự trao đổi vật chất di truyền từ việc lai tạo truyền thống, truyền lại cho thế hệ sau qua các phép lai và chọn lọc các cá thể mang tính trạng mong muốn. Theo các phương thức chọn giống hiện nay, ngoài các đặc điểm hình thái được chọn còn bao gồm dấu hiệu di truyền như tế bào học và dấu hiệu sinh hóa, vì vậy các nhà chọn giống cần hơn 12 năm để phát triển một giống mới. Tuy nhiên, khi có thông tin về gen mục tiêu và các biến thể allele, dựa vào các biến thể allele của gen có thể phát triển thành dấu chỉ thị phân tử dùng nhận diện và theo dõi sự di truyền ở các thế hệ lai tiếp theo. Ngày nay, phương thức chọn giống truyền thống được thay thế bằng chọn giống với sự hỗ trợ của dấu chỉ thị phân tử nhờ vào phương pháp tiếp cận mới thông qua thông tin từ bộ gen (Gupta *et al.*, 2005; Collard *et al.*, 2005; Kumar *et al.*, 2018). Với phương thức chọn giống này sẽ không bị ảnh hưởng bởi các yếu tố môi trường, giai đoạn thu thập mẫu và các giai đoạn phát triển của cây (Winter *et al.*, 1995), như vậy bằng phương thức này giúp rút ngắn thời gian và chính xác hơn trong công tác chọn giống.

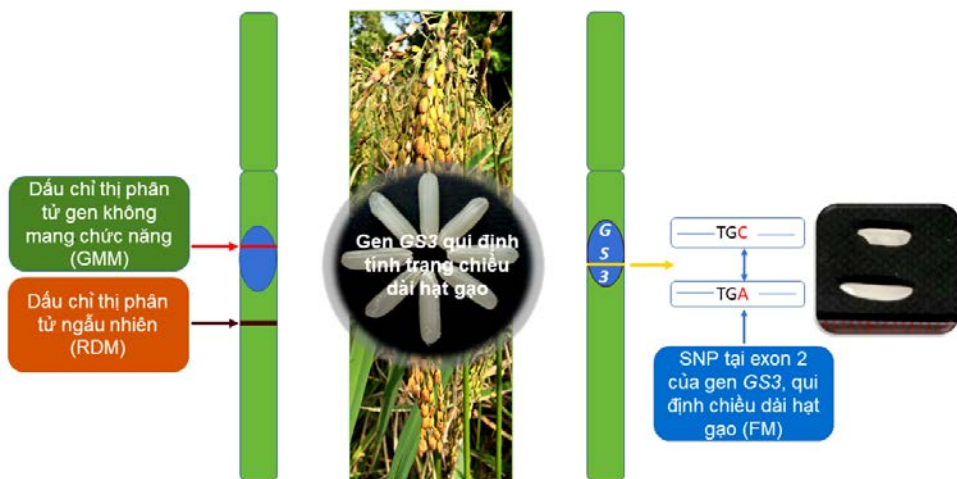
Nhằm đáp ứng nhu cầu trên, chọn giống chuẩn xác là phương pháp chọn giống cây trồng mang tính trạng mong muốn và được nhận diện bằng dấu chỉ thị phân tử gen chức năng (FM), chỉ thị này có nguồn gốc trực tiếp từ bộ gen, biết chính xác vùng gen kiểm soát tính trạng đó (Conner *et al.*, 2004). Do biết chính xác gen quy định tính trạng mong muốn và khi có sự xuất hiện biến thể của gen dẫn đến sự biến đổi kiểu hình. Từ những phát hiện đó, chỉ thị phân tử FM được phát triển và dùng trong chọn giống một cách hiệu quả và chính xác hơn (Bohra *et al.*, 2019). Bên cạnh đó, FM còn giải thích mối quan hệ giữa phiên mã, gen, protein và kiểu hình. Ngoài việc dùng trong chọn giống cây trồng, dấu chỉ thị phân tử gen chức năng FM còn dùng trong nghiên cứu phát sinh loài, cấu trúc và đa dạng quần thể nhằm hiểu rõ hơn về sự tiến hóa của bộ gen (Dhutmal *et al.*, 2018; Amom *et al.*, 2017).

1.2 LỊCH SỬ PHÁT TRIỂN CỦA CÁC CHỈ THỊ PHÂN TỬ

Vào thế kỷ thứ 20, việc khám phá ra gen đã giúp hiểu rằng bộ gen là thông tin di truyền và là cơ sở tạo ra sinh giới, trong đó bao gồm cả cây trồng. Từ cơ sở về thông tin di truyền cùng với tiến bộ của khoa học kỹ thuật nên nhiều dấu chỉ thị phân tử được tạo ra và dùng nhận diện tính trạng số lượng (QTL), lập bản đồ di truyền (Amom *et al.*, 2017). Nhiều dấu chỉ thị phân tử được tạo ra như: đa hình về chiều dài DNA bằng cắt giới hạn (RFLP)

(Botstein *et al.*, 1980), đa hình DNA dựa vào khuếch đại ngẫu nhiên (RAPD) (William *et al.*, 1990; Welsh *et al.*, 1990), đa hình khuếch đại chiều dài DNA (AFLP) (Vos *et al.*, 1995), microsatellite hay trình tự lặp lại giản đơn (SSR) (Tautz *et al.*, 1989; Litt *et al.*, 1989), trình tự vùng DNA khuếch đại (SCARs) (Paren *et al.*, 1993), trình tự DNA khuếch đại được cắt ra từ gel (CAPS) (Konieczny *et al.*, 1993), đa hình đơn nucleotide (SNP) (Gupta *et al.*, 2001). Hầu hết các dấu chỉ thị phân tử là ngẫu nhiên (RDMs), chúng có thể mất đi trong quá trình tái tổ hợp. Ngược lại, nếu dấu chỉ thị là gen quy định tính trạng cụ thể, khi đó dấu chỉ thị DNA (FM) nằm trong gen hay liên kết chặt với gen thì dấu phân tử đó được tạo ra và có thể sử dụng trong cải tiến giống cây trồng. Trong trường hợp này dấu chỉ thị được gọi là chỉ thị phân tử gen chức năng (FMs) rất thuận tiện trong chọn giống.

FMs được phân loại trong nhóm chỉ thị phân tử DNA, có nguồn gốc dựa vào trình tự DNA chức năng đặc trưng đã được kiểm chứng (Andersen *et al.*, 2003). Vì vậy SNP hay InDels như là chỉ thị FMs rất hữu ích trong chọn giống cây trồng khi so sánh giữa chỉ thị RDMs và chỉ thị không mang gen chức năng (GMMs) (Hình 1.1). Chỉ thị GMMs nằm trong một gen chức năng, có thể không liên kết đến kiểu hình mong đợi, vì vậy dẫn đến chọn lọc sai trong MAS. FMs được gọi là chỉ thị phân tử “chuẩn xác”, ngược lại chỉ thị RDMs là chỉ thị phân tử “không chuẩn” như RFLPs, AFLPs, hay SSR (Varshney *et al.*, 2005). Thật vậy, chương sách này, chúng tôi giới thiệu sự hiệu quả của việc sử dụng chỉ thị phân tử FMs dùng trong chọn giống lúa.



Hình 1.1 Mô hình xây dựng chỉ thị phân tử là gen chức năng (FM) dựa trên biến thể allele ở exon 2 của gen GS3, trong đó GMM là dấu phân tử không mang gen chức năng và RDM dấu chỉ thị phân tử ngẫu nhiên

1.3 DẤU CHỈ THỊ CHỨC NĂNG

Chỉ thị phân tử FMs nằm trong vùng gen được mã hóa thường liên kết trực tiếp 1 kiểu hình, được biết chức năng và sự đa dạng trong kiểu hình (Andersen *et al.*, 2003). FMs còn được biết như dấu phân tử nhận diện chính xác kiểu gen và kiểu hình. Kiểu gen chỉ bị ảnh hưởng khi trình tự qui định chức năng đặc trưng bị thay đổi, FMs chỉ có thể chỉnh sửa khi những alleles có lợi hiện diện trong quần thể (Varshney *et al.*, 2005). Khi đa hình xảy ra ở gen mong muốn và được sửa chữa bằng alleles thích hợp trong chọn lọc quần thể, khi đó FMs được sử dụng trong chọn lọc tính trạng phức tạp (Bagge *et al.*, 2007). Vì thế chỉ thị phân tử FMs có tính hiệu quả, nhanh chóng và sử dụng trong chọn lọc đa hình alleles với sự chính xác cao vì chúng không bị tác động bởi sự tái tổ hợp.

Muốn phát triển dấu chỉ thị phân tử FMs cần phải nhận diện được gen mong muốn ảnh hưởng đến đặc điểm hình thái, trình tự gen, và đặc điểm chức năng (Amom *et al.*, 2017). Để xác định được chức năng của gen thì có rất nhiều phương pháp khác nhau như nghiên cứu sự biểu hiện gen, lập bản đồ tạo dòng, bản đồ QTL, chuyển vị gen (Borevitz *et al.*, 2003; Salgotra *et al.*, 2014) (Hình 1.2). Kỹ thuật chuyển gen biểu hiện vượt mức (overexpression) và giảm sự biểu hiện của gen (knockdown) cũng được dùng để nghiên cứu chức năng của gen mục tiêu. Bước tiếp theo của quá trình phát triển chỉ thị phân tử FM là nghiên cứu các dạng biến thể allelic của gen mục tiêu. Khi đó, trình tự các allele biến thể kiểu gen liên quan đến kiểu hình tương ứng và nhận thấy được sự thể hiện giữa kiểu gen và kiểu hình (Thornsberry *et al.*, 2001). Việc phát triển chỉ thị phân tử gen chức năng FM yêu cầu phải biết được trình tự của alleles/gen liên kết đến kiểu hình được biểu hiện. Sự đa hình giữa các alleles của gen mục tiêu do việc mất hoặc thêm nucleotide (InDels), SNPs, sự khác nhau về số lượng và lặp lại đặc trưng của SSR, sự mất đi một đoạn hay mất cả gen (Borevitz *et al.*, 2003). Thực vậy, chỉ thị phân tử FMs có thể dựa trên nhiều vùng gen mã hóa (exon) và không mã hóa (intron, UTR) (McCouch *et al.*, 1997).

Trải qua hai thập kỷ với sự phát triển của nhiều phương pháp giải trình tự nucleotide, cho đến hôm nay đã có phương pháp giải trình tự thế hệ mới (NGS), nhiều QTLs được nhận diện và bản đồ di truyền được xây dựng dựa vào các dòng tái tổ hợp (RILs), dòng đẳng gen (NILs), đa bội thể (DHs),.... Phát triển bản đồ di truyền dựa vào lai tạo giữa các cặp bố mẹ có kiểu hình tương phản cần nhiều thời gian và kinh phí. Để vượt qua các khó khăn trên, một phương thức tiếp cận khác là lập bản đồ di truyền tương quan dùng nhận diện các QTL có trong tự nhiên hoặc từ nguồn gen lai tạo. Bản đồ di truyền tương quan (AM) là công cụ quan trọng trong nhận diện đa hình gần hoặc trong gen, mà gen thể hiện kiểu hình khác nhau giữa các kiểu gen (Soto-Cerda

et al., 2012). Liên kết không ổn định (Linkage Disequilibrium ~LD) có xu hướng duy trì qua nhiều thế hệ giữa các loci có kiểu gen liên kết với nhau, có thể hữu ích trong chọn lọc (MAS) (Rostoks *et al.*, 2006). Bản đồ tương quan giúp ích cho nghiên cứu đa dạng kiểu gen mục tiêu, một khi quần thể nghiên cứu càng đa dạng và càng lớn thì mối tương quan càng lớn (Bresseghele *et al.*, 2006). Khi có sự kết hợp giữa kiểu gen SNPs nằm trên các vùng LD với tính ổn định của LD thì những SNPs này rất hữu ích cho sự phát triển của chỉ thị phân tử FM và hỗ trợ tốt cho chọn giống theo phương pháp hồi giao (MABB) (Simko *et al.*, 2009).



Ứng dụng của chỉ thị phân tử là gen chức năng

- Hỗ trợ trong công tác chọn lọc (MAS)
- Hỗ trợ trong công tác lai hồi giao (MABC)
- Hỗ trợ trong công tác chọn lọc tuần hoàn
- Chọn lọc kiểu gen

Hình 1.2 Các phương pháp phát triển chỉ thị phân tử gen chức năng (FM)

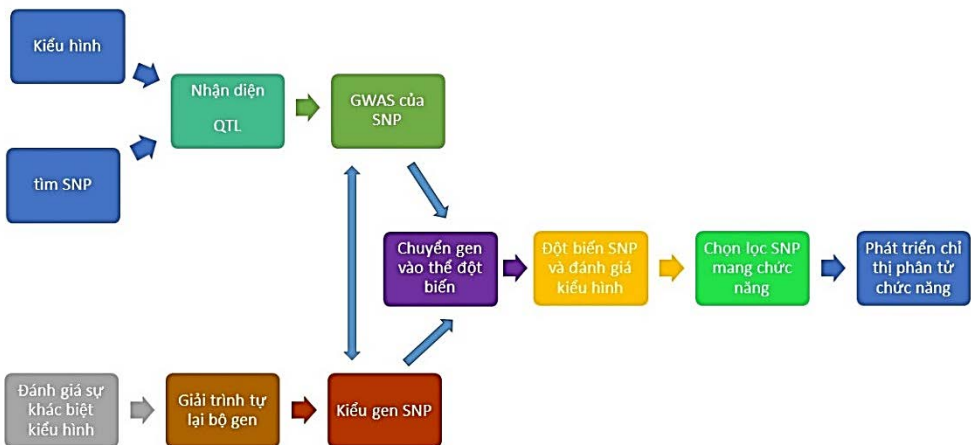
Với công nghệ giải trình tự thế hệ mới (NGS), số lượng bộ gen được giải trình tự ngày càng nhiều dẫn đến nhiều SNP được tìm ra, từ thông tin đó nhiều bản đồ gen có độ phân giải cao được thiết lập (Huang *et al.*, 2017). Hơn nữa, các nghiên cứu liên kết toàn bộ bộ gen (GWAS) hiện là một giải pháp thay thế khả thi cho việc lập bản đồ QTL giúp phân tích các tính trạng số lượng quan trọng. GWAS có thể cung cấp thông tin đánh giá mẫu đại diện mang vùng gen mục tiêu. GWAS tận dụng khối gen liên kết có mật độ cao trải dài toàn bộ bộ gen của một kiểu gen để xác định các vùng gen liên quan đến đặc điểm kiểu hình. Gần đây, GWAS được ứng dụng thành công để xác định các vùng gen liên quan đến tính trạng quan trọng ở lúa, lúa mạch, bắp, lúa mì và các loài cây trồng khác (Huang *et al.*, 2017). Ví dụ, GWAS sử dụng trong một tập hợp bản đồ liên kết xác định SNP liên quan đến khả năng chống chịu mặn trên cây lúa (Simko *et al.*, 2009).

Thêm vào đó, những tiến bộ trong công nghệ giải trình tự thế hệ mới dẫn đến chi phí giải trình tự DNA giảm, hiện nay việc xác định kiểu gen dựa vào trình tự (GBS) của các loài thực vật vì chúng có sự đa dạng di truyền cao và bộ gen lớn (Elshire *et al.*, 2011), phương pháp mới này hỗ trợ rất hiệu quả cho chọn giống. Kỹ thuật GBS liên quan đến việc sử dụng các enzym giới hạn để giảm độ phức tạp của bộ gen, theo sau là giải trình tự bằng kỹ thuật NGS. GBS được xây dựng dựa vào những biến thể SNP hiện diện trong phần lớn bộ gen, phương pháp này giúp giảm nhiều về chi phí (Elshire *et al.*, 2011, Sonah *et al.*, 2013). Do đó, các SNP hiện diện trong bộ gen có thể được sử dụng trong các nghiên cứu chọn lọc bộ gen, GWAS và đa dạng di truyền. Ví dụ như SNP có nguồn gốc từ GBS liên kết với alen chức năng E3Ha cho sự trưởng thành ở đậu nành đã được xác định (Tardivel *et al.*, 2014).

Tương tự việc giải trình bộ gen, việc giải trình tự hệ gen phiên mã (RNA-Seq hay Transcriptome) cũng cung cấp một lượng lớn thông tin về các mối liên hệ giữa gen và tính trạng được thể hiện thông qua việc giải trình tự bộ gen trên mô hình và không có mô hình. Tập hợp thông tin từ việc so sánh các biến thể trong hệ gen giữa các mẫu thí nghiệm, cho phép phát hiện gen và phát triển chúng thành chỉ thị phân tử FM cũng như nghiên cứu sự biểu hiện gen và so sánh mức độ biểu hiện giữa các cá thể (Garg *et al.*, 2013). RNA-Seq có thể cho phép xây dựng dấu chỉ thị phân tử, bao gồm FM ở các loài thực vật chưa có bộ gen tham chiếu, chưa được giải trình tự toàn bộ bộ gen (Chen *et al.*, 2013). RNA-Seq đã được ứng dụng thành công trong nghiên cứu sinh vật từ nấm men cho đến thực vật. Thêm vào đó, với trình tự của hệ gen phiên mã cũng giúp phát hiện SNP là một phần quan trọng trong nghiên cứu di truyền phân tử vì các locus của SNP có thể được dùng xây dựng bản đồ di truyền mật độ cao GWAS (Novaes *et al.*, 2008; Guo *et al.*, 2003; Duran *et al.*, 2009; Li *et al.*, 2016; Park *et al.*, 2019; Huang *et al.*, 2020).

Thật vậy, sự xuất hiện các biến thể SNPs ở gen chức năng đã giúp ích trong việc nâng cao hiệu quả trong chọn giống. Thông qua các biến thể SNPs xảy ra ở gen chức năng, ngày càng nhiều chỉ thị phân tử FM liên quan đến đặc tính nông học quan trọng đã xác định được các tính trạng, cung cấp nguồn gen quý cho tạo giống. Việc giải lại trình tự bộ gen và việc xác định biến thể SNP có liên kết chặt chẽ với kiểu hình là hai chiến lược quan trọng được sử dụng trong GWAS để xác định SNPs chức năng và sự phát triển dấu chỉ thị phân tử FM (Hình 1.3). Lập bản đồ quần thể (RIL, DH, phân ly, vv...) sử dụng xác định QTL thích hợp, trong đó trình tự bộ gen các dòng khác nhau tạo ra đủ số SNP. Trong đó SNP xuất hiện tại các QTL được coi là SNP GWAS, vì các gen tiềm năng tại locus QTL được dự đoán theo phân tích GWAS. Mặt khác, trong lập bản đồ tương quan, nguồn gen đa dạng rất hữu ích và việc xác định kiểu gen SNP trên cơ sở giải mã lại bộ gen cung cấp một công cụ tốt để phát hiện SNP trong bộ sưu tập nguồn gen lớn. Việc so sánh các SNP GWAS từ quần thể giúp xác định các SNP chức năng có liên quan đến tính trạng kiểu hình.

Tóm lại, bất kể biến thể DNA có nguồn gốc từ đâu, một khi chỉ thị phân tử FM được phát triển, bước tiếp theo là đánh giá chức năng của các chỉ thị được xây dựng liên kết với gen được quan tâm (Kage *et al.*, 2016). Việc đánh giá các chỉ thị phân tử FM mới được phát triển về chức năng có thể được thực hiện bằng các nghiên cứu biểu hiện gen, bao gồm cả việc làm im lặng gen do virus gây ra (VIGS) và các phân tích loại bỏ hoặc làm giảm sự biểu hiện gen như đã nêu ở trên (Rodenburg *et al.*, 2018; Burch-Smith *et al.*, 2004; Tadege *et al.*, 2005).



Hình 1.3 Sơ đồ nghiên cứu tương tác bộ gen (GWAS) dùng để nhận diện các biến thể SNPs và xây dựng nên chỉ thị phân tử là gen chức năng (FM)

1.4 ƯU ĐIỂM CỦA CHỈ THỊ PHÂN TỬ FM SO VỚI CÁC DẤU CHỈ THỊ KHÁC

RDMs là dấu chỉ thị DNA phổ biến nhất được sử dụng trong chọn lọc gián tiếp. RDMs có nguồn gốc từ sự đa hình của trình tự DNA trong vùng gen lân cận gen mục tiêu, dẫn đến việc tiên đoán không chính xác về tính trạng được quan tâm. Tuy nhiên, vấn đề xảy ra khi sử dụng RDMs có thể khắc phục được bằng chỉ thị phân tử FMs với sự tương ứng kiểu hình 100%. Khi tác động di truyền quy định chức năng dựa trên các trình tự gen, chỉ thị phân tử FM được thiết lập dựa vào allele trong một số gen gốc mà không cần hiệu chỉnh (Cakir *et al.*, 2010; Gupta *et al.*, 2010). Điều này có lợi cho việc ứng dụng các dấu chỉ thị phân tử, đặc biệt là trong chọn giống cây trồng, giúp lựa chọn vật liệu di truyền ban đầu để lập một hệ thống quần thể phân ly gồm những cây được chọn cũng như tiếp tục chọn lọc các dòng tiên tiến (Andersen *et al.*, 2003).

Chỉ thị phân tử FMs nằm ở tại các gen mong muốn và được liên kết trực tiếp với đặc điểm hình thái, vì thế có thể sử dụng với độ tin cậy và hiệu quả cao dựa vào allele được nhận diện chính xác dùng trong chương trình chọn giống (Simko *et al.*, 2009). Chỉ thị phân tử FMs cũng làm giảm nguy cơ mất hoặc chọn lọc sai các cá thể được chọn khi sử dụng phương pháp chọn giống MAS (Rodenburg *et al.*, 2018; Lau *et al.*, 2015). Trong khi đó, chọn giống bằng MAS, QTL được xác nhận là cần thiết dùng để đánh giá sự khác biệt nguồn gốc di truyền của chúng. Tuy nhiên, chỉ thị phân tử FMs thì sử dụng trực tiếp mà không cần thiết phải kiểm chứng (Ingvarlsen *et al.*, 2008; Rodenburg *et al.*, 2018; Nawaz *et al.*, 2017). FMs tạo điều kiện cho việc lựa chọn đặc điểm kiểu hình, đặc biệt cho phép các nhà nhân giống các tổ hợp lai hiếm trong một quần thể lớn (Edmeades *et al.*, 2004). Ngoài ra, FM rất hữu ích trong việc sàng lọc ra các allele trong tự nhiên cũng như quần thể trong nhân giống, sửa chữa các allele mong muốn trong quần thể, xây dựng liên kết haplotype FM và tổ hợp các allele chỉ thị phân tử FM liên kết với các tính trạng phức tạp (Andersen *et al.*, 2005).

1.5 FM_s TRONG NHÂN GIỐNG CÂY TRỒNG CHÍNH XÁC

Chi phí giải trình tự gen ngày càng giảm, điều này đã tạo điều kiện thuận lợi cho sự phát triển chỉ thị phân tử FMs. Do chỉ thị phân tử FMs có nhiều ưu điểm hơn RDM, FMs ngày càng được sử dụng nhiều hơn trong phương pháp chọn giống MABB về tính trạng chất lượng và đặc tính tốt của cây trồng như chống chịu stress hay kháng sâu bệnh hại và năng suất cao (Edmeades *et al.*, 2004; Yang *et al.*, 2015). Trong chọn giống, chỉ thị phân tử FMs được sử dụng sàng lọc nguồn gen, phân tích đa dạng di truyền, MAS, MABB, ứng dụng chỉ thị phân tử hỗ trợ chọn lọc trong chọn giống (MARS) và chọn

lọc bộ gen (GS) nhằm cải thiện các tính trạng quan trọng (Salgotra *et al.*, 2014; Xu *et al.*, 2005; Yi *et al.*, 2009; Jin *et al.*, 2010).

1.5.1 Sàng lọc quần thể và đa dạng di truyền

Nguồn gen thực vật (PGR) là nguyên liệu cơ bản cần thiết để cải thiện cây trồng, nâng cao năng suất, cải thiện chất lượng, giảm dịch bệnh và gây hại của côn trùng cũng như các yếu tố stress phi sinh học. Trước khi xuất hiện phương pháp mới về tiếp cận bộ gen, các dấu chỉ thị di truyền chính chỉ dùng trong đánh đa dạng di truyền của nguồn gen dựa vào các đặc điểm hình thái khác nhau (Vos *et al.*, 1995). Với tiến bộ của kỹ thuật giải trình tự gen đã cho phép phát triển nhiều loại dấu chỉ thị phân tử khác nhau như RDBMs, GMs và FMs (Yu *et al.*, 2002). Để chọn nguồn gen có sự biến đổi allele đối với tính trạng kiểu hình cụ thể, chỉ thị phân tử FMs liên kết trực tiếp với gen mục tiêu và có thể sử dụng trực tiếp trong các chương trình chọn giống (Nadeem *et al.*, 2018). Chỉ thị phân tử FMs cho phép xác định đặc tính và chọn lọc các kiểu gen mong muốn với độ chính xác cao. Khi so sánh với chỉ thị phân tử RDMs, chỉ thị phân tử FMs dùng để xác định trực tiếp các kiểu gen quan trọng về mặt nông học từ nguồn gen được thuần hóa, giống đang canh tác, cây hoang dã và nguồn gen thực vật. Các tính trạng mục tiêu được chọn lọc trong nguồn gen bằng chỉ thị phân tử FMs có thể được sử dụng trong các chương trình phát triển cây trồng mới. Điều này, giúp cho nhà chọn giống cây trồng phát triển nguồn gen mới bằng cách đưa các gen mong muốn sang bộ gen cây trồng đang cần có gen đó (Xu *et al.*, 2004). Bên cạnh đó, Liên minh quốc tế Bảo vệ giống Cây trồng mới (UPOV) cũng đồng ý sử dụng chỉ thị phân tử FMs để xác định cụ thể các đặc điểm kiểu gen và dùng để đánh giá sự khác biệt giữa giống, tính đồng nhất và đặc tính ổn định (DUS) của giống (Jamali *et al.*, 2018).

Thông tin về biến thể di truyền có trong nguồn gen là cơ sở dùng để cải thiện và phát triển giống cây trồng (Hodgkin *et al.*, 2001). Thông tin đa dạng di truyền của PGR là điều cần thiết và đầu tiên cho chọn giống, vì ngày càng nhiều PGR đã được đánh giá đặc điểm của chúng dựa trên các tính trạng kiểu hình, nhưng rất ít nguồn gen được đánh giá ở cấp độ phân tử. Với kỹ thuật hiện đại giải trình tự rất nhiều bộ gen của các loài, từ cơ sở đó đã xây dựng được nhiều chỉ thị phân tử dùng trong phân tích đa dạng di truyền, nhận diện giống, chọn lọc chính xác cá thể F1 và dùng trong bảo hộ các giống cây trồng (Kaisoon *et al.*, 2008; Salgotra *et al.*, 2015). Để xác định kiểu gen tiềm năng và đa dạng dùng cho chọn giống, cần có một nguồn gen di truyền lớn (Collard *et al.*, 2005; Jiang *et al.*, 2013). Trong kỷ nguyên thông tin di truyền từ các bộ gen, những thông tin này đã giúp cho việc xây dựng các chỉ thị phân tử như RDM, GMM, FM dùng đánh giá sự biến đổi di truyền có sẵn trong nguồn gen (Singh *et al.*, 2019). Trong đó, tính chuyên biệt của chỉ thị phân tử FM