

TÓM TẮT

PROMOTER VÀ ỨNG DỤNG TRONG CÔNG NGHỆ GEN THỰC VẬT

Lê Thị Thu Hiền, Trần Thị Phương Liên, Nông Văn Hải

Viện Công nghệ Sinh học

TÓM TẮT

Thuật ngữ ‘promoter’ được sử dụng để chỉ vùng trình tự nucleotide trong genome nằm ngược dòng về phía đầu 5’ của vị trí khởi đầu phiên mã gen (transcription start site - TSS) - vị trí được xác định là +1. Promoter có cấu trúc rất phức tạp và chứa nhiều yếu tố đặc trưng tham gia điều khiển quá trình phiên mã nhờ khả năng đảm bảo các vị trí nhận biết cho các protein tham gia vào việc điều khiển biểu hiện gen. Yếu tố quan trọng nhất là hộp TATA. Đây là vị trí gắn cho các yếu tố khởi đầu phiên mã và thường nằm ở vị trí khoảng -30 bp. Một yếu tố phổ biến khác là hộp CCAAT. Ở thực vật, hộp CCAAT có thể có trình tự khác nhau và còn được gọi là hộp AGGA. Ngoài hộp TATA và CCAAT, hầu hết promoter của sinh vật nhân chuẩn còn chứa thêm các yếu tố điều hòa khác là các trình tự DNA nằm ở phía trước TSS, được biết đến rộng rãi như là các vùng tăng cường. Những trình tự này rất đa dạng về chiều dài, có thể dao động từ khoảng -100 bp đến 100 bp hoặc ngược lại xa về phía đầu 5’ của TSS. Theo khả năng biểu hiện gen, promoter được chia làm hai loại chính: promoter không đặc hiệu hay còn gọi là promoter cơ định và promoter đặc hiệu. Trong lĩnh vực công nghệ gen thực vật, các promoter được sử dụng rất rộng rãi để điều khiển biểu hiện các gen ngoại lai trong cây trồng biến đổi gen. Các promoter được sử dụng phổ biến nhất bao gồm CaMV35S, CaMV19S và nopaline synthase (NOS) promoter. Ngoài ra, rất nhiều nhóm nghiên cứu đã nỗ lực phân lập và thiết kế các promoter mới, đặc hiệu có khả năng điều khiển biểu hiện gen ở mức độ cao hoặc đặc hiệu để ứng dụng trong công nghệ gen thực vật. Bên cạnh đó, vấn đề xây dựng cơ sở dữ liệu về các trình tự promoter và chức năng của chúng cùng các chương trình phân mềm thiết kế promoter cũng được xây dựng và đưa vào ứng dụng. Ở Việt Nam, việc phân lập các promoter có giá trị để thiết kế vector chuyển gen thực vật đang được quan tâm nghiên cứu.

Từ khóa: Biểu hiện gen, công nghệ gen thực vật, promoter, yếu tố điều hòa, yếu tố phiên mã

KHÁI NIỆM VỀ PROMOTER

Promoter là một thành phần quan trọng trong cấu trúc của tất cả các gen và đóng vai trò then chốt trong việc biểu hiện gen. Từ những nghiên cứu mở đầu về promoter trên vi khuẩn vào cuối những năm 70, đến nay, nhiều promoter từ các gen ở nhiều loài khác nhau đã được phân lập, xác định các vùng chức năng. Khái niệm về promoter được hiểu rộng và sâu sắc hơn.

Promoter, bao gồm vùng trình tự nucleotide nằm về phía đầu 5’ của vị trí khởi đầu phiên mã gen (transcription start site - TSS), đóng vai trò điều khiển quá trình phiên mã nhờ khả năng đảm bảo các vị trí nhận biết cho các protein tham gia vào việc điều khiển biểu hiện gen như enzyme RNA polymerase, các yếu tố phiên mã (transcription factor - TF) (TF là các protein tham gia điều khiển mức độ biểu hiện gen thông qua việc gắn vào các trình tự DNA đặc hiệu (các yếu tố *cis*) trong vùng

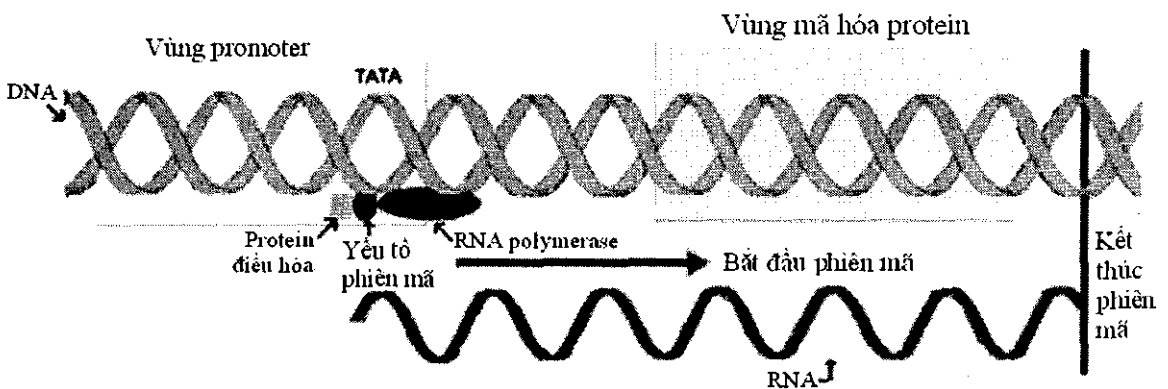
promoter của các gen đích và làm tăng hoặc giảm mức độ biểu hiện của chúng)). Tùy thuộc vào khoảng cách từ TSS, các thuật ngữ ‘promoter gần’ hay ‘promoter tối thiểu’ (khoảng vài trăm nucleotide quanh vùng TSS) (proximal promoter hay minimal promoter) và ‘promoter xa’ (hàng nghìn nucleotide ở phía trên TSS) (distal promoter) có thể được sử dụng. Cả hai loại promoter gần và xa đều chứa rất nhiều yếu tố tham gia vào quá trình điều hòa phiên mã đặc hiệu tế bào, mô, cơ quan giai đoạn phát triển và môi trường (<http://www.freepatentsonline.com/6987179.html>; Shahmuradov *et al.*, 2003; Rombauts *et al.*, 2003). Trong đó, vùng promoter gần (vùng promoter trung tâm) được xem là đảm bảo sự gắn kết phức hợp RNA polymerase II vào vị trí chính xác và tác động trực tiếp đến mức độ phiên mã (Nikolov, Burley, 1997; Berk, 1999). Vùng này bao gồm các yếu tố như hộp TATA và các hộp khởi đầu (Initiator boxes) - là những vị trí gắn cho các protein liên kết hộp TATA và các TF khác đặc hiệu cho RNA polymerase II (Featherstone, 2002). Vùng promoter

xa được xem là có chứa các yếu tố điều hòa sự biểu hiện đặc hiệu không gian và thời gian (Tjian, Maniatis, 1994; Fessele *et al.*, 2002).

PROMOTER, VAI TRÒ ĐIỀU KHIỂN BIỂU HIỆN GEN VÀ CẤU TRÚC PHÂN TỬ

Về cơ bản, promoter đóng vai trò như một công tắc bật gen. Mỗi gen đều cần một promoter để hoạt động hoặc bắt đầu biểu hiện. Nhưng promoter không chỉ đơn giản như một công tắc điện, ví dụ, chỉ có hai chiều, bật hoàn toàn hoặc tắt hoàn toàn. Thay vào đó, promoter có rất nhiều module khác nhau hoạt động như các cảm biến (sensor) và cho phép chúng có thể phản ứng, theo những cách mà hiện nay chúng ta chưa hiểu được hoàn toàn, đối với những tín hiệu khác nhau từ những gen khác hoặc từ môi trường, định rõ khi nào thì khởi động và khởi động ở đâu, với cường độ và thời gian như thế nào. Trong những điều kiện nhất định, promoter có thể không hoạt động và khi đó thì gen hoàn toàn không được khởi động. Nói chung, vai trò của promoter ở một gen bình thường trong một sinh vật là cho phép gen hoạt động đúng trong các chu trình điều hòa hết sức phức tạp của sinh vật.

Về mặt cấu trúc, promoter có cấu trúc phức tạp và chứa rất nhiều yếu tố đặc trưng có chức năng gắn/liên kết với các yếu tố protein tham gia vào quá trình phiên mã gen (Hình 1). Các yếu tố điều hòa (regulatory elements – RE) nằm trên promoter và cùng một sợi với vùng mang mã của gen được gọi là các yếu tố *cis* (*cis* elements) (Buchanan *et al.*, 2000). Mỗi yếu tố được đặt tên trên cơ sở trình tự nucleotide của chúng. Yếu tố *cis* cơ bản nhất là **hộp TATA (TATA box)**, được tìm thấy ở hầu hết các gen của sinh vật nhân chuẩn, nằm ở vị trí khoảng –30 (có nghĩa là 30 nucleotide trước TSS, vị trí được xác định là +1). Ở sinh vật nhân sơ, hộp TATA thường nằm ở quanh vị trí –10. Hộp này chịu trách nhiệm giúp cho quá trình gắn RNA polymerase chính xác vào TSS. Một gen vẫn có thể gắn RNA polymerase vào promoter với một hộp TATA cải biến nhưng hiệu quả của quá trình khởi đầu phiên mã có thể bị ảnh hưởng rất nhiều. Hầu hết các gen cảm ứng (inducible genes) thường có chứa hộp TATA và ít nhất là hai yếu tố *cis*. Trong khi đó, các “gen giữ nhà” (“housekeeping gene”) có các yếu tố *cis* ít đa dạng hơn và thậm chí có thể không chứa hộp TATA. Hộp TATA thường nằm ngay cạnh một yếu tố *cis* cơ bản khác là **hộp CAAT (CAAT box)** cung cấp vị trí gắn cho RNA polymerase.



Hình 1. Cấu trúc của một gen điển hình với vùng promoter (Nguồn NRC/PBI Promoter Technology).

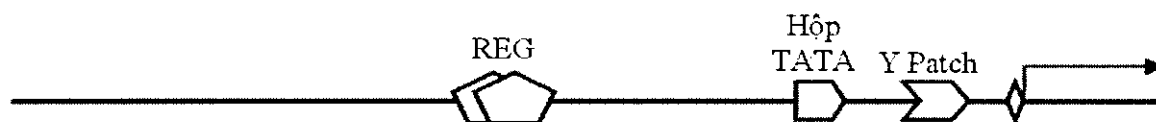
Ở thực vật, hộp TATA thường nằm ở vị trí từ –35 đến –25 bp trước TSS (Breathnach, Chambon, 1981). Hộp CCAAT nằm ở vị trí giữa –70 và –100 bp trước TSS. Hộp này có thể có trình tự khác nhau và được gọi là hộp AGGA. Ngoài các trình tự đặc trưng như hộp TATA và CCAAT, hầu như tất cả các promoter của sinh vật nhân chuẩn được nghiên cứu cho đến nay có chứa thêm các yếu tố điều hòa khác

là các trình tự DNA nằm ở phía trước TSS, tham gia điều hòa hoạt động của gen (Benoist, Chambon, 1981; Gruss *et al.*, 1981; Khoury, Gruss, 1983). Những trình tự điều hòa trước TSS được biết đến rộng rãi như là các vùng tăng cường (enhancers) hay các trình tự hoạt hóa ngược dòng (upstream activating sequences). Những trình tự này rất đa dạng về chiều dài, có thể dao động từ khoảng –100

bp đến 100 bp hoặc ngược khá xa về phía đầu 5' của TSS với những dạng cấu trúc là các trình tự lặp lại. Tuy nhiên, các vùng tăng cường cũng có thể được tìm thấy ở vùng 3' không được dịch mã và thậm chí ở cả các intron (Buchanan *et al.*, 2000). Mỗi vùng tăng cường có thể có chức năng độc lập như là một yếu tố *cis* và nằm trên cùng một sợi với các gen được tác động. Vùng tăng cường có thể có vai trò điều hòa biểu hiện đặc hiệu, điều khiển gen biểu hiện ở những mô hoặc cơ quan cụ thể. Vai trò này thường vẫn được duy trì khi vùng tăng cường được cắt và đưa vào một gen khác để điều khiển biểu hiện một gen mới trong cơ thể chủ.

Ở thực vật, promoter của nhiều gen cảm ứng bởi môi trường còn mang một số yếu tố *cis* là các trình tự đặc trưng khác như hộp G (G box). Hộp G được phát hiện lần đầu tiên như là một trình tự 11 bp (5'-C/AACACGTGGCA-3') nằm ở phía trước của rất nhiều gen mã hóa cho subunit nhỏ của ribulose biphosphate carboxylase (Giuliano *et al.*, 1988). Từ đó đến nay, các yếu tố *cis* chứa hộp G hoặc các tiết tấu trình tự (motif) liên quan với lõi hexanucleotide CACGTG đã được xác định trong các promoter của nhiều gen không liên quan, như gen được cảm ứng bởi ánh sáng, bởi abscisic acid (ABA), khi bị tổn thương, trong điều kiện kị khí (Guiltinan *et al.*, 1990). Motif Iwt và PA, hai motif liên quan đến hộp G, đã được phát hiện có liên quan đến tính biểu hiện đặc hiệu mô của promoter (US patent 6987179). Trong đó, motif Iwt liên quan đến biểu hiện đặc hiệu phôi, motif PA liên quan đến biểu hiện đặc hiệu cao ở rễ, biểu hiện mức độ thấp ở lá và không biểu hiện ở hạt. Nhiều nghiên cứu cho thấy một số protein chiết xuất từ nhân thực vật có mang một số hoạt tính gắn đặc trưng cho hộp G và các trình tự liên quan. Một số dòng cDNA mã hóa cho các protein tương tác đặc hiệu với các yếu tố *cis* chứa trình tự

CACGTG đã được phân lập: EmBP-1 và HBP-la ở lúa mì, TAF-1 ở thuốc lá, CPRF - 1,2,3 ở cây mùi tây, GBF - 1,2,3 ở *Arabidopsis* (US patent 5723751). Ngoài ra, các vị trí gắn tương tự GT-1 (GT-1 like binding sites) (nhóm các trình tự bảo thủ GGT^A/T^A) cũng được tìm thấy ở vùng ngược dòng rất xa về phía đầu 5' của promoter ở *Arabidopsis* plastocyanin và dường như liên quan đến sự hoạt hóa quá trình phiên mã khi có ánh sáng (Fischer *et al.*, 1994). Một điểm đặc trưng khác về trình tự của các promoter thực vật là sự có mặt của các trình tự cho phép tạo Z-DNA. Z-DNA là DNA xoắn trái xuất hiện do những sự lặp lại của dinucleotide GC hoặc AC. Việc hình thành dạng Z được xem là có tác động làm giảm mức độ phiên mã (US patent 6987179). Gần đây, Yamamoto và đồng tác giả (2007) có phân tích promoter thực vật theo hướng chứa 3 nhóm các trình tự DNA đặc biệt: pyrimidine patch (Y patch) (thường ở vị trí -100 đến -13), hộp TATA (vị trí từ -50 đến -20) và nhóm yếu tố điều hòa (REG - regulatory element group) (vị trí -20 đến -400) (Hình 2). Trong đó, Y patch hay còn gọi là vùng pyrimidine thường có trình tự TCTCTC hoặc CTCCTC hay TTCTTC. Chức năng sinh học của Y patch chưa được nghiên cứu rõ nhưng đây là thành phần chung của promoter thực vật (không có trong promoter động vật). Y patch được các tác giả này phát hiện thấy khi nghiên cứu genome của *Arabidopsis*. REG thường là các trình tự ngắn 5 - 15 bp trên promoter. Nhóm này bao gồm trên 200 loại trình tự, phần lớn là các yếu tố *cis* và là các vị trí nhận biết của các protein tham gia điều khiển biểu hiện gen. Có thể nói, những nghiên cứu về promoter, cấu trúc chính xác và vai trò của từng yếu tố trong điều hòa phiên mã, hoặc cách thức sử dụng các yếu tố này ở dạng độc lập, rất có giá trị ứng dụng trong việc cải biến các promoter hoạt động hoặc để thiết kế các promoter mới.



Hình 2. Mô hình promoter thực vật theo Yamamoto và đồng tác giả (2007).

PHÂN LOẠI PROMOTER

Promoter có thể được chia làm hai loại chính: promoter không đặc hiệu hay còn gọi là promoter cơ định (constitutive promoter) và promoter đặc hiệu với các yếu tố *cis* chuyên biệt (specific promoter).

Promoter cơ định

Loại promoter này tham gia điều khiển biểu hiện gen ở tất cả hoặc hầu như tất cả các loại mô khác nhau trong tất cả hay hầu như mọi giai đoạn phát triển của sinh vật. CaMV35S promoter điều khiển sự

biểu hiện RNA 35S được phân lập từ virus khảm súp lơ (cauliflower mosaic virus - CaMV) là một ví dụ điển hình của nhóm promoter này. CaMV là một pararetrovirus của các thực vật thuộc họ cải. Genome của virus này là một phân tử DNA vòng sợi kép có kích thước 8 kb với ba quãng ngắt là sợi đơn. Hai transcript RNA chính (19S và 35S) và 6 khung đọc mở được mã hóa bởi DNA. Sự phiên mã xảy ra từ tiểu nhiễm sắc thể vòng và không tiếp hợp trong nhân của tế bào thực vật và DNA virion được tổng hợp trong tế bào chất thông qua sự phiên mã ngược của 35S RNA transcript (Turner *et al.*, 1996; Kohli *et al.*, 1999). CaMV35S promoter là trình tự có kích thước vào khoảng 350 bp nằm ở phía đầu của 35S transcript (-343 đến +8, với vị trí Cap ở +1), trong đó có khoảng 250 bp gói lên 3% đoạn kết thúc của gen VI, khung đọc mở cuối cùng trong 6 khung đọc mở lớn. Có 3 vùng trên promoter, promoter trung tâm chứa hộp TATA (vị trí từ 46 đến +8) và 2 vùng chính khác có chức năng tăng cường. Vùng A (-90 đến -46) chủ yếu cần cho biểu hiện ở rễ và vùng B (-343 đến 90) cần cho biểu hiện ở lá (Benfey, Chua, 1990; Hohn, Futterer, 1992; Ho *et al.*, 1999). Tiểu vùng của vùng B (B1 đến B5) có thể được nhận dạng dựa trên những tương tác khác biệt với những nhân tổ phiên mã khác nhau. Tuy nhiên, vùng B hoàn chỉnh cho phép biểu hiện đa dạng hơn so với trường hợp kết hợp các tiểu vùng. Vai trò của các vùng khác nhau trên CaMV35S promoter trong việc gây bệnh của CaMV mới được nghiên cứu trong thời gian gần đây (Noad *et al.*, 1997; Turner *et al.*, 1996). Các nghiên cứu cho thấy việc loại bỏ hộp TATA làm mất khả năng gây nhiễm. Tuy nhiên, khi loại vùng phía trước giữa vùng tăng cường (từ -207 đến -56) lại không bị ảnh hưởng trong khi loại bỏ hoàn toàn đoạn này sẽ làm mất khả năng gây nhiễm. Hai vùng tăng cường tách biệt liên quan đến khả năng gây nhiễm đã được xác định là (-207 đến -150) và (-95 đến -56). Vùng tăng cường thậm chí có thể hoạt động theo hướng ngược chiều.

Promoter đặc hiệu

Sự biểu hiện đặc hiệu là sự biểu hiện các sản phẩm của gen giới hạn ở một hoặc một vài mô (hạn chế về không gian - spatial limitation) và/ hoặc ở một hoặc vài giai đoạn phát triển (đặc hiệu thời gian - temporal limitation). Cần nhấn mạnh rằng không có một promoter hoàn toàn đặc hiệu: các promoter dường như hoạt động ở một số mô, trong khi ở một số mô khác không hoặc chỉ hoạt động yếu. Hiện tượng này được biết đến như là sự biểu hiện yếu.

Promoter đặc hiệu mô (tissue specific promoter): Các loại promoter đặc hiệu mô chỉ điều khiển biểu hiện gen ở những mô nhất định, ví dụ promoter đặc hiệu rễ (Conkling *et al.*, 1990), promoter đặc hiệu bó mạch ở thực vật (Wang, 1992), promoter đặc hiệu hoa (Robert, 1998). Thuộc nhóm promoter này có promoter điều khiển biểu hiện gen mã hóa sucrose synthase. Sucrose synthase là một trong những enzyme quan trọng tham gia vào mọi hoạt động sống của tế bào. Ở lúa, người ta đã phát hiện được ba gen cùng mã hóa cho sucrose synthase (Rsuc1, Rsuc2, Rsuc3) (Yu *et al.*, 1992). Tuy nhiên, mức độ biểu hiện của các gen này ở các mô là khác nhau. Gen Rsuc2 không có tính đặc hiệu mô cao. Gen Rsuc3 đặc hiệu cao ở hạt (Huang *et al.*, 1996). Chỉ có gen Rsuc1 biểu hiện đặc hiệu ở bó mạch, rễ, mầm (Wang, 1992). Kích thước của gen Rsuc1 là 8167 bp, trong đó kích thước của vùng 5' tương đối lớn, chiếm khoảng 3 kb. Promoter của Rsuc1 chứa các trình tự đặc trưng của một promoter như hộp TATA, hộp CCAAT... và các trình tự đặc hiệu, quyết định sự biểu hiện gen đặc hiệu bó mạch như hộp ASL, hộp GATA... (Yin *et al.*, 1997).

Promoter cảm ứng với môi trường (environmentally inducible promoter): Trong các điều kiện cực đoan (stress) như hạn, nhiệt độ quá cao hoặc quá thấp, oxy hóa cao... các gen chống chịu thường biểu hiện với tốc độ rất nhanh. Vì vậy, vấn đề khởi động sự phiên mã cũng như cấu trúc promoter của chúng và các protein tham gia vào việc điều khiển biểu hiện gen thông qua việc nhận biết và hoạt hóa quá trình khởi động phiên mã được đặc biệt quan tâm.

Gen chịu hạn

Các protein có chức năng điều khiển sự biểu hiện gen và có thể có chức năng trong việc phản ứng đối với tình trạng hạn như protein kinase và TF. Promoter của gen mã hóa cho TF có chứa motif gắn với DNA (DNA binding motif) (bZIP, MYB, MYC, EREBP/AP2...) được hoạt hóa bởi tác động của các điều kiện cực đoan. Tiếp theo đó, chúng đóng vai trò điều hòa các gen chức năng dưới tác động của những điều kiện này. Phần lớn các gen cảm ứng với điều kiện khô hạn đều cảm ứng với điều kiện mặn. Một trong số gen này còn được cảm ứng bởi điều kiện lạnh. Phụ thuộc vào phản ứng với ABA, các yếu tố điều hòa trên những promoter của các gen này được chia thành 2 loại như sau (Shinozaki, Yamaguchi-Shinozaki, 1999):

Các yếu tố điều hòa phụ thuộc vào ABA: trên

promoter có trình tự để các protein MYB, MYC nhận biết. Ở *Arabidopsis*, ATMYB2 nhận biết trình tự TGGTTAG trên promoter để hoạt hóa gen *rd22* (dehydration - responsive gene 22) trong điều kiện mất nước (Abe *et al.*, 1997). Vùng gắn DNA - MYB (MYB - DNA binding domain) được biết đến lần đầu tiên ở dạng *v-myb* oncogene (gen gây ung thư) của virus gây bệnh ung thư tủy xương (*Myeloblastosis*) ở chim. Đây là một nhóm TF rất đa dạng ở các loài thuộc sinh vật nhân chuẩn (Braun, 1999). MYB protein đặc trưng bởi sự có mặt của 2 hoặc 3 MYB motif, trong đó mỗi motif này chứa cấu trúc xoắn chuyển xoắn (helix-turn-helix) với 3 vị trí gốc Trp. Ở động vật, MYB protein có chứa 3 motif MYB được ký hiệu là R1, R2, R3, trong khi thực vật chỉ chứa 2 motif MYB là R2, R3.

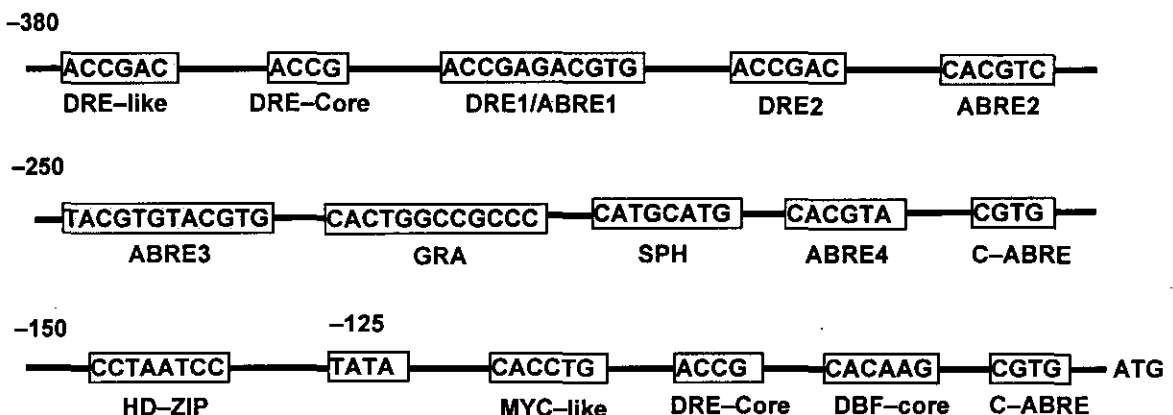
MYC protein được đặc trưng bởi cấu trúc vùng cơ bản xoắn vòng xoắn (helix-loop-helix - bHLH). bHLH có thể nhận biết được trình tự dạng CANNTG trên promoter (N - nucleotide bất kỳ). Nhưng kết quả nghiên cứu promoter *rd22* của *Arabidopsis* trên cây thuốc lá chuyển gen cho thấy: protein *rd22BP1* chỉ nhận biết trình tự CACATG đầu tiên phản ứng với hiện tượng mất nước, còn trình tự CACATG thứ hai trong vùng này lại có chức năng ức chế (Abe *et al.*, 1997). AtMYC2 nhận biết được trình tự CATGTG trên promoter của các gen liên quan đến tổn thương cơ học hoặc kháng bệnh của *Arabidopsis* (Lorenzo, 2004).

Trong một số trường hợp, gen có chứa yếu tố phản ứng với ABA (ABRE - ABA responsive element) - PyACGTGGC trong vùng promoter. ABRE lần đầu tiên được phân lập từ promoter của gen *Em* ở lúa mì và từ *rab* (abscisic acid responsive gene) của lúa gạo. Protein gắn với *Em* (EmBP-1 - *Em* binding protein) là dạng bZIP (leucine zipper

motif) protein. Trình tự ACGT là điểm gắn đặc hiệu của bZIP protein. Tuy nhiên, làm thế nào ABA hoạt hóa được bZIP protein để gắn với ABRE và khởi đầu sự phiên mã vẫn còn chưa được rõ (Shinozaki, Yamaguchi-Shinozaki, 1999). Trình tự ACGT cũng là vùng trung tâm của hộp G trên promoter của các gen phản ứng với ánh sáng.

Các yếu tố điều hòa không phụ thuộc vào ABA:
DRE - yếu tố phản ứng sự mất nước (dehydration response element) là tên của trình tự dài 9 bp: TACCGACAT. Trong đó, vùng quan trọng nhất của DRE motif là vùng lặp lại giàu C (CRT - C repeat) - CCGAC. Các yếu tố protein gắn với trình tự DRE (DRE binding proteins - DREBs) ở vùng promoter đã được nghiên cứu và phân lập. Những protein này chứa các vùng bảo thủ như các protein EREBP (tham gia biểu hiện ethylene) và AP2 (tham gia vào quá trình biểu hiện kiểu hình của thực vật có hoa). Các protein này còn được ký hiệu là nhân tố gắn với CRT (CRT binding factor - CBF). Chúng phản ứng rất nhanh với trạng thái cực đoan, trước khi xuất hiện ABA ở lượng đáng kể. Ví dụ như promoter của *rd29A* (dehydration - response gene 29A) ở *Arabidopsis* (Liu *et al.*, 1998).

Trong promoter của gen *rab17* ở ngô có 9 dạng vị trí nhận biết cho các TF với trình tự từ 4 - 12 bp khi phản ứng với ABA, điều kiện mất nước và áp suất thẩm thấu cao: DRE1; ABRE1; DRE2; ABRE2; ABRE3a/3b; GRA (GC - rich RAB activator); SPH; ABRE4; TATA (Hình 3). Tuy vậy, tác động của mỗi yếu tố đều khác nhau: Phần lớn, chúng đều tác động ở phôi, riêng ABRE4, DRE2 và GRA tác động ở mầm và lá non. Promoter các gen *rab* của lúa gạo và lúa miến cũng có những vùng tương tự (Buchanan *et al.*, 2004).



Hình 3. Promoter của gen *rab17* ở ngô (Buchanan *et al.*, 2004).

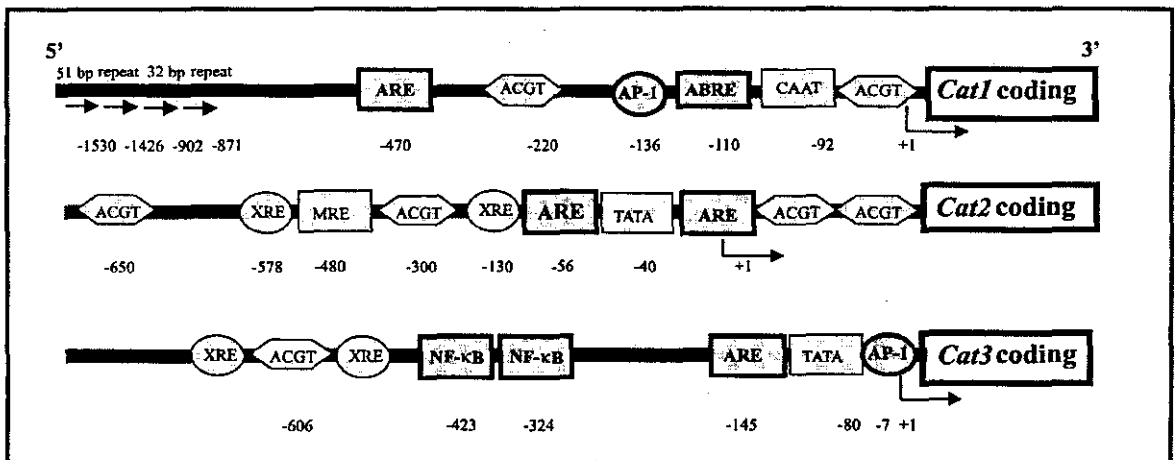
Gen chịu nhiệt

Việc biểu hiện của nhiều protein sốc nhiệt (heat shock protein - HSP) được điều khiển bằng các yếu tố phiên mã trong điều kiện sốc nhiệt - heat shock transcription factor (HSF). HSF nhận biết trình tự bảo thủ của yếu tố sốc nhiệt (heat shock element - HSE) trên promoter để hoạt hóa sự phiên mã của gen này. HSE trên promoter của HS (heat shock) gen có trình tự DNA bảo thủ như nGAAn và các trình tự tương tự nhưng thay đổi chiều khác nhau (ví dụ: 5'-CTnGAAnnTTCnAG-3') (Amin *et al.*, 1988). Ở thực vật và động vật, có hai vị trí HSE trên promoter. Cơ chế hoạt động của HSF đã được nghiên cứu sâu ở *Arabidopsis*. Trong trạng thái bình thường, ATHSF-1 ở dạng monomer, khi bị kích hoạt do điều kiện cực đoan, HSF tạo thành trimer và gắn với HSE trên promoter của HS gen và bị phosphoryl hóa làm tăng sự phiên mã của gen để tạo ra HSP70. HSP70 tạo phức với HSF và phân ly HSF ra khỏi DNA. Sau đó, phức HSP70-HSF lại phân ly tiếp ra thành từng thành phần riêng biệt: HSF về dạng monomer và HSP70. HSF gen ở *Arabidopsis* (ATHSF1) đã được tạo dòng và chuyển vào cây làm cho HSF ở cây luôn ở trạng thái hoạt hóa ngay cả khi không xử lý nhiệt và cây có thể chịu được nhiệt độ 54°C (Lohmann *et al.*, 2004).

Gen chống chịu oxy hóa

Tất cả các tế bào sống đều gặp phải các dạng oxy hoạt hóa - ROS (reactive oxygen species) và phản ứng với hiện tượng này trong hoạt động sống

bình thường cũng như khi gặp phải các điều kiện cực đoan như ánh sáng, hạn, tổn thương cơ học, bị bệnh... Cách đây gần 30 năm, trong vi khuẩn, người ta đã phát hiện ra khoảng 30 protein, trong đó có protein điều hòa OxyR, phản ứng với H_2O_2 . OxyR còn điều hòa sự biểu hiện nhanh của 9 - 12 protein nữa. Dạng tetramer của OxyR là thành viên của họ LysR - hoạt hóa sự phiên mã. Chúng tồn tại ở hai dạng oxy hóa và khử và chỉ có dạng oxy hóa mới có hoạt tính hoạt hóa đó. Tương tự, SoxRS, gen phản ứng với O_2^- , cũng hoạt hóa quá trình phiên mã này. Trong tế bào sinh vật nhân chuẩn không có protein giống như vậy, nhưng có một loạt các yếu tố điều hòa sự biểu hiện này. Trong các hệ thống của động vật có vú có các yếu tố: yếu tố nhân kB (nuclear factor kB - NF-kB) và protein hoạt hóa (activator protein - AP-1) đóng vai trò điều hòa. Yếu tố phản ứng chống oxy hóa (Antioxidant responsive element - ARE) có trình tự nhận biết trong vùng promoter - TGACTCA- của các gen glutathione S-transferase, Mn-SOD tạo ra sự có mặt của các gen này trong điều kiện cực đoan oxy hóa. Những yếu tố này cũng tìm thấy trong vùng promoter của các gen chống oxy hóa ở thực vật (Scandalios, 2005). Vai trò của các yếu tố này không chỉ đơn thuần hoạt hóa các gen chống oxy hóa mà còn hoạt hóa các gen chống chịu trong các điều kiện cực đoan khác như hạn, lạnh... Vùng promoter của 3 gen *Cat* (gen mã hóa Catalase của ngô) được nghiên cứu rất kỹ. Chúng có những yếu tố hoạt hóa phiên mã như ARE, AP-1, NF-kB, abscisic acid - responsive element 2 (ABRE) - ACGT (Hình 4).



Hình 4. Promoter của gen mã hóa enzyme catalase ở ngô (Scandalios, 2005).

Phản ứng với ánh sáng

Nghiên cứu vùng promoter của các gen *rbcS* (ribulose - 1,5-bisphosphate carboxylase small subunit) ở cây hai lá mầm cho thấy chúng có trên 30 vùng bảo thủ. Trong đó, những vùng phản ứng với ánh sáng chính là hộp G, hộp I, GT-1 hay còn gọi là hộp II.... Hộp G là vùng trên promoter phản ứng với ánh sáng, có trình tự CACGTG. Hộp I có trình tự nhận biết là GATAAGR và trình tự ngược chiều của nó - YCTTATC được chú ý đặc biệt. Đột biến vùng này làm giảm rõ rệt mức độ biểu hiện của gen *rbcS* khi có ánh sáng (Weeks *et al.*, 2007).

Ở bèo tấm, các nhà khoa học đã phân lập được SSU5A; SSU5B; SSU13 promoter của gen *rbcS*.

Promoter của các gen *rbcS* ở bèo tấm tương đồng đến trên 60%. Promoter của các gen *rbcS* này có 3 vùng bảo thủ chính là X, Y, Z với các trình tự tương ứng là: X - GATAAGGATGTAATCC; Y - GTGGAGGAG; Z - GATCGAACGTGGCAGCCGTAGAA. Hộp TATA nằm trong vùng từ -300 đến 0. Trong đó, vùng X có chứa vùng bảo thủ GATAAG giống như hộp I của các gen *rbcS* của cây hai lá mầm. Gây đột biến trong vùng -480 đến 0 trên promoter sẽ làm giảm khả năng biểu hiện của gen này. Nghiên cứu LRF-1 (light-regulated DNA binding factor) ở bèo tấm cho thấy chúng ảnh hưởng đến khả năng này ở các SSU promoter của bèo tấm khác nhau. SSU5B chịu tác động mạnh nhất (Buzby *et al.*, 1990).

Bảng 1. Một số trình tự nhận biết của TF trên promoter của thực vật.

Gen	Yếu tố <i>cis</i>	Trình tự
Rab17 - ngô (Buchanan <i>et al.</i> , 2004)	DRE1	CACCGACG
	ABRE1	CACGTGCC
	DRE2	CACCGACG
	ABRE2	ACACGTCC
	ABRE3a/3b	GTACGTGTACGTG
	GRA	CACTGGCCGCC
	SPH	CATGCATG
	ABRE4	GCCACGTA
<i>rbcS</i> - cây hai lá mầm (Weeks <i>et al.</i> , 2007)	Hộp G	CACGTG
	Hộp I	GATAAGR
	GT-1	TTAACC
<i>rbcS</i> - bèo tấm (Buzby <i>et al.</i> , 1990)	X	GATAAGGATGTAATCC
	Y	GTGGAGGAG
	Z	GATCGAACGTGGCAGCCGTAGAA
Em - lúa mì	bZIP	ACGT

Phản ứng với điều kiện thiếu oxy

Thực vật có thể chịu được với điều kiện hạn chế oxy khi xảy ra lụt hoặc khi bị ngập nước hoàn toàn và có thể khởi động chu trình Krebs chuyển quá trình hô hấp sang lên men kèm với việc biểu hiện các gen phản ứng với điều kiện kỵ khí mã hóa cho các enzyme liên quan đến quá trình glycolysis và lên men. Mohanty và đồng tác giả (2005) đã phân tích 13 promoter của các gen kỵ khí và thấy chúng có một số motif quan trọng, có mặt ở hầu hết các promoter này.

Như đã thảo luận, hộp TATA có mặt ở rất nhiều gen nhưng motif AA(A/T)CAAA cũng có vai trò chính ở các gen phản ứng với các điều kiện bất lợi ở thực vật. Mười tám motif còn lại rất khác so với các motif đã được phát hiện ở các gen phản ứng với điều kiện thiếu oxy (gen kỵ khí).

Những kết quả nghiên cứu promoter và TF ở trên cho thấy, TF có khả năng kích hoạt khởi động phiên mã cùng một lúc nhiều gen có chức năng phản ứng với cùng một loại điều kiện cực đoan bằng cách nhận biết các trình tự đặc hiệu trong vùng promoter

này. Đây là hướng biến nạp gen liên quan đến tính chống chịu các điều kiện cực đoan như hạn, lạnh...

có triển vọng và đang được nhiều nhà khoa học quan tâm nghiên cứu và đưa vào ứng dụng.

Bảng 2. Tần suất xuất hiện các motif đã xác định trong promoter của các gen phản ứng với lạnh, hạn, nồng độ muối cao và ABA ở lúa (Mohanty *et al.*, 2005).

Motif	Tần suất xuất hiện (%)	Motif	Tần suất xuất hiện (%)
(A/C)AACTT	82	TTTCTTTAT	71
AA(A/T)CAAA	71	AA(A/G)AAAAA	71
TATATC	76	AACTTTG	59
ATATAA	71	AACAAAA(A/T)GA	47
ACTCTTT	76	C(A/T)AAAACAAA	47
TTAAAAA	82	TAAATATAGAT	65
A(G/T)CCATG	65	TTTAAAGA	76
CTCTT(C/T)CA	65	AATCAATTC	65
AAATT(A/G)TT	71	CTATATAA	76
ATATAATA	76	AAGAAAA(A/T)	59

PROMOTER VÀ ỨNG DỤNG TRONG CÔNG NGHỆ GEN THỰC VẬT

Genome của từng sinh vật được sắp xếp theo cách nhất định và tương đối bảo thủ giữa các loài. Nhờ vậy, các cá thể trong cùng một loài có thể lai với nhau. Mỗi loài duy trì sự toàn vẹn và ổn định di truyền bởi vì các rào cản sinh học ngăn cản các loài có khoảng cách di truyền xa nhau lai hoặc trao đổi vật chất di truyền với nhau. DNA ngoại lai thường bị phá hủy hoặc bất hoạt. Sử dụng các kỹ thuật lai tạo giống cùng loài, những sinh vật có các đặc tính riêng biệt có thể được tạo ra. Tuy nhiên, những kỹ thuật này đòi hỏi nhiều thời gian và có thể phải trải qua nhiều thế hệ mới có được những tính trạng mong muốn và loại bỏ những đặc tính không cần thiết. Công nghệ DNA tái tổ hợp được sử dụng để biến đổi cây trồng, vật nuôi bằng cách phân lập các gen có giá trị từ động vật, thực vật và vi sinh vật, thiết kế và chuyển chúng vào bộ gen của sinh vật nhận (gen có thể của các loài vốn không có quan hệ họ hàng) và nhanh chóng tạo ra các sinh vật biến đổi gen (Genetically Modified Organisms - GMOs) mang những đặc tính mong muốn... (Chrispeels, Sadava, 2002; Mackenzie *et al.*, 2003). Trong đó, gen được chuyển và biến đổi theo cách không thể xảy ra trong quá trình tiến hóa tự nhiên cũng như qua lai chọn giống (ví dụ, giữa các loài khác nhau và giữa động vật - thực vật và vi sinh vật). Rõ ràng, với sự can thiệp của con người, quá trình chọn tạo giống đã

vượt qua các rào cản tự nhiên (như việc chuyển một hoặc nhiều gen giữa các cơ thể sinh vật không có quan hệ di truyền).

Về cơ bản, việc chuyển gen ngoại lai vào sinh vật nhận để tạo GMOs cần phải tiến hành một số bước bắt buộc, trong đó có bước phân lập các gen riêng lẻ kiểm soát các tính trạng cụ thể, nhân chúng lên và gắn một promoter vào trước gen này nhằm đảm bảo chúng có thể hoạt động tốt trong các sinh vật đích (nếu không có bước ghép với promoter thì gen ngoại lai cần chuyển sẽ không hoạt động). Promoter và gen được điều khiển biểu hiện dính kèm tạo thành một "cassette biểu hiện gen" ("gene-expression cassette") hay "kết cấu gen" ("gene construct"). Gen và promoter có thể được phân lập từ các nguồn sinh vật khác nhau như từ thực vật, động vật, vi khuẩn và virus. Một số kết cấu gen biểu hiện được kết nối với nhau thành chuỗi trong vector, một hoặc vài gen trong số đó sẽ là gen kháng kháng sinh, giúp cho việc chọn lọc các tế bào tái tổ hợp. Sau đó, các kết cấu gen này được chuyển vào trong sinh vật đích.

Tìm kiếm và phân lập các loại promoter dạng cơ định từ các nguồn thực vật, virus và vi khuẩn gây bệnh thực vật... là một trong những hướng chính trong tạo giống cây trồng biến đổi gen. Nhờ các promoter cơ định, gen quan tâm có giá trị được biểu hiện ở rất nhiều mô, cơ quan khác nhau. Trong một số trường hợp, chúng có thể là nguyên nhân gây hại

môi trường. Ví dụ, áp lực đối với môi trường sẽ tăng cao khi promoter dạng cơ định điều khiển sản sinh protein độc tố ở những bộ phận khác nhau của thực vật, nơi sâu hại không tấn công vào. Để bảo vệ khả năng sinh sản và phát triển bình thường của quần thể côn trùng có ích, protein độc tố cần được điều khiển biểu hiện ở những mô thực vật là nguồn thức ăn của côn trùng gây hại (Sardana, Altosaar, 1996). Vì vậy, bên cạnh việc phân lập và đưa vào sử dụng các promoter cơ định, các nhà công nghệ gen thực vật cũng tập trung thiết kế các promoter mới có các hoạt tính điều khiển biểu hiện gen đặc hiệu mô, ở những giai đoạn phát triển nhất định của cây, hay dưới sự tác động của từng điều kiện môi trường hoặc chỉ biểu hiện ở nơi tế bào, mô bị tổn thương. Với những promoter như vậy, người ta có thể biểu hiện các tính trạng quan tâm ở những vị trí là mô hay cơ quan mong muốn hoặc trong những giai đoạn phát triển nhất định của cây. Nhờ vậy, có thể kiểm soát chính xác sự biểu hiện gen quan tâm được chuyển vào nhằm tăng cường tính hữu ích của cây trồng biến đổi gen (Sardana, Altosaar, 1996).

Ngoài các promoter nguyên vẹn, một số nghiên cứu cũng tập trung tạo các promoter mới thông qua việc kết hợp các trình tự điều hòa của một promoter với các phần của một promoter khác (Fluhr *et al.*, 1986; Ellis *et al.*, 1987; US patent 6987179). Những kết cấu như vậy được tạo ra nhờ đưa một trình tự điều hòa khác loại vào một promoter hoạt động có mang một phần hoặc toàn bộ trình tự điều hòa của chúng (Ellis *et al.*, 1987; Strittmatter, Chua, 1987). Bên cạnh đó, các promoter cải biến cũng được tạo ra bằng cách thêm một trình tự điều hòa khác loại vào phía đầu 5' của một promoter không hoàn chỉnh và bất hoạt, có nghĩa là promoter chỉ chứa hộp TATA và trong một số trường hợp là các yếu tố CCAAT (Fluhr *et al.*, 1986; Strittmatter, Chua, 1987; Aryan *et al.*, 1991). Những sự kết hợp như vậy có thể tạo ra các promoter với những hoạt tính mới.

Trong công nghệ gen thực vật, có thể nói, một trong những thành công sớm nhất và có giá trị nhất đối với lĩnh vực biểu hiện protein thực vật chính là việc sử dụng các promoter cơ định có nguồn gốc từ virus và vi khuẩn *Agrobacterium* làm công cụ biểu hiện gen ngoại lai hiệu quả trong cây trồng biến đổi gen. Rất nhiều promoter như vậy đã được sử dụng rộng rãi trong công nghệ gen thực vật và vẫn tiếp tục là promoter được chọn lựa trong các nghiên cứu biểu hiện nhanh, đơn giản với rủi ro thấp. Đến nay, rất nhiều loại promoter được sử dụng để điều khiển biểu hiện các gen có giá trị trong cây trồng

(Schledzewski, Mendel, 1994). Các nghiên cứu chuyển gen thực vật thường sử dụng các promoter: CaMV35S, CaMV19S, nopaline synthase (NOS), Adh, Ubiquitin của ngô, tubullin, actin, cab, pepcarboxylase (PEPC), promoter của gen mã hóa sucrose synthase (Yang, Russell, 1990); promoter đặc hiệu hạt phấn (Bp10 gene promoter) (Odell *et al.*, 1985; Ebert *et al.*, 1987; Sullivan *et al.*, 1989; McElroy *et al.*, 1990; Albani *et al.*, 1992; Wang *et al.*, 1992; Walden, Wingender, 1995; Sardana *et al.*, 1996; Cheng *et al.*, 1998; Datta *et al.*, 1998).

CaMV35S promoter và các promoter lai

CaMV35S promoter được sử dụng rộng rãi nhờ khả năng hoạt động mạnh ở nhiều loại mô của nhiều loài cây. Chúng có khả năng điều khiển biểu hiện các RNAs khác nhau ở cả cây một lá mầm và hai lá mầm bao gồm thuốc lá, đậu tương, cà rốt, ngô và lúa (Fromm *et al.*, 1986; Ou-Lee *et al.*, 1986). Promoter này là promoter dạng cơ định, hoạt động mạnh ở cả các hệ thống biểu hiện tạm thời và các tế bào chuyển gen thực vật ổn định với mức độ biểu hiện tăng hơn 30 lần so với promoter của gen mã hóa nopaline synthase (NOS) (Sanders *et al.*, 1987). Laporte *et al.* (2001) cũng đã nghiên cứu và chứng tỏ ảnh hưởng quan trọng của CaMV35S promoter khi biểu hiện gen mã hóa sucrose phosphate synthase ở cà chua. Sản lượng cà chua thường không thay đổi khi sử dụng ribulose-1,5-bisphosphate carboxylase-oxygenase small subunit (rbcS) promoter, trong khi sản lượng tăng tới 80% khi sử dụng CaMV35S promoter để điều khiển biểu hiện gen này. Phân tích huỳnh quang và hóa học lúa chuyển gen *gusA* dưới sự điều khiển của CaMV35S promoter cho thấy, promoter này điều khiển biểu hiện gen chủ yếu ở quanh các bó mạch chồi lá, rễ và hạt. Tuy nhiên, không thấy sự biểu hiện của *gusA* ở các cơ quan từ của hoa. Hoạt tính của *gusA* ở rễ và lá lúa khá cao, tương đương hoặc cao hơn ở lá thuốc lá có chứa cấu trúc gen này (Bảng 3) (Smith, Hood, 1995).

Trên cơ sở CaMV35S promoter, rất nhiều promoter lai hoặc kết hợp có khả năng tăng cường mức độ biểu hiện gen đã được thiết kế (Trần Thị Phương Liên *et al.*, 1992) và đưa vào sử dụng như promoter 35S kép (Kay *et al.*, 1987; Fang *et al.*, 1989), promoter lai với vùng trung tâm của CaMV19S promoter và vùng tăng cường phía trên của CaMV35S (Odell *et al.*, 1988), promoter kết hợp giữa CaMV35S promoter và các yếu tố tổng hợp mannopine (Comai *et al.*, 1990) hoặc với các yếu tố của Adh1 và ocs promoter để biểu hiện trong cây

một lá mầm (Last *et al.*, 1991).

NOS promoter

NOS promoter có nguồn gốc từ Ti-plasmid của vi khuẩn *A. tumefaciens* và thường được sử dụng để biểu hiện vùng mang mã của neomycin phosphotransferase II của vi khuẩn kháng lại kanamycin trong chọn lọc các tế bào biến nạp (An *et al.*, 1986).

Ubiquitin promoter của ngô và actin promoter của lúa

Ubiquitin promoter của ngô và actin promoter của lúa/ các trình tự intron thường được sử dụng để biểu hiện ở cây một lá mầm (MacElroy *et al.*, 1990). Trong đó, Ubiquitin được thử nghiệm điều khiển biểu hiện gen trong ngô, thuốc lá và lúa (Christensen *et al.*, 1992; 1996; Cornejo *et al.*, 1993; Toki *et al.*, 1992; Sardana *et al.*, 1996). Actin-1 promoter của lúa được thử nghiệm biểu hiện ở một vài giống lúa (*indica* và *japonica*) (Datta *et al.*, 1998). Trong nghiên cứu này, mức độ biểu hiện của protein CryIA(b) ở lá và thân dưới sự điều khiển của CaMV35S promoter và actin promoter khác nhau không đáng kể.

Adh1 promoter

Adh1 promoter điều khiển *gusA* biểu hiện khá mạnh ở rễ và các mô của lúa. Ngoài ra, ở phôi và nội nhũ của hạt cũng có sự biểu hiện đáng kể của *gusA*. Trong khi ở thuốc lá, Adh1 promoter chỉ hoạt động khi vùng tăng cường của gen tổng hợp octopin hoặc CaMV35S promoter được đưa vào trước promoter này. Bảng 3 cho thấy khả năng hoạt động của CaMV35S promoter và Adh1 promoter ở các loại mô khác nhau của lúa (Shimamoto, 1991).

PEPC promoter

PEPC là promoter đặc hiệu mô màu xanh (green tissue), có khả năng điều khiển biểu hiện lượng lớn protein CryIA(b) trong lá và thân của lúa biến đổi gen (Datta *et al.*, 1998).

Promoter đặc hiệu mô và cảm ứng bởi ánh sáng

Ngoài ra, một promoter đặc hiệu mô và cảm ứng bởi ánh sáng được xem là có mức độ hoạt động cao như CaMV35S promoter trong các thực vật biến đổi gen là các promoter của gen mã hóa cho protein liên kết chlorophyll a/b và tiểu đơn vị của ribulose bisphosphate carboxylase (rbcS) (Jones *et al.*, 1985; Fluhr *et al.*, 1986). Các yếu tố DNA tham gia vào cảm ứng ánh sáng với mức độ cao ở lá của hai gen rbcS đã được phân lập (Fluhr *et al.*, 1986; Timko *et al.*, 1985). Các yếu tố này được xem là các yếu tố tương tự vùng tăng cường (enhancer) nhờ khả năng điều khiển biểu hiện ở mức độ cao trên các đoạn promoter có hoạt tính yếu khi được đặt theo hướng thích hợp ở vị trí đầu 5' (từ -145 hoặc -46). Khi được thử nghiệm đặt ở vị trí đầu 3', một yếu tố đã không hoạt động hiệu quả chứng tỏ chúng không mang đầy đủ các đặc tính của những vùng tăng cường đã được mô tả trước đây (Timko *et al.*, 1985). Vùng tăng cường đầu tiên, có hoạt tính không phụ thuộc hướng ở khoảng cách xa hơn 1 kb về phía đầu 5' hoặc 3' tính từ vị trí khởi đầu phiên mã, đã được phân lập ở các vùng promoter của virus động vật biểu hiện mức độ cao như vùng promoter SV40. Tương tự, CaMV35S promoter có mang yếu tố tăng cường thực vật, đó là đoạn trình tự dài 338 bp phân lập từ vùng phía trước của hộp TATA. Đoạn trình tự này có thể tăng cường mức độ biểu hiện của promoter khác có hoạt tính thấp lên tới mức độ đã quan sát được ở CaMV35S nguyên vẹn (Odell *et al.*, 1988).

Bảng 3. Mức độ biểu hiện của CaMV35S promoter - *gusA* và Adh1 promoter - *gusA* ở mô lúa chuyển gen (Shimamoto, 1991).

	Mô	CaMV35S promoter - <i>gusA</i>	Adh1 promoter - <i>gusA</i>
Rễ		++	+++
Lá		+++	+
Hoa	Bao phấn	-	++
	Vòi nhị	-	+++
	Phấn hoa	-	++
	Bầu nhụy	-	-
Hạt	Phôi	++	+++
	Nội nhũ	++	++

Bảng 4. Các promoter được sử dụng trong những cây trồng biến đổi gen đã được phê chuẩn (http://www.bats.ch/bats/publikationen/1997-2_gmo/4.1.2_promoters.php).

Promoter được sử dụng	Cây trồng	Số sản phẩm xuất hiện
Ví khuẩn	Ngô, bông, đu đủ, khoai tây, đậu tương	10
CaMV35S	Ngô, bông, cải dầu, đu đủ, khoai tây, dưa hấu, cà chua	16
CaMV35S, AMV-5'-utr	Đậu tương	1
CaMV35S, CMV-5'-utr	Đu đủ, dưa hấu	3
CaMV35S, Cab22R-5'-utr	Cà chua	1
CaMV35S, IVS2-utr	Ngô	1
CaMV35S, IVS6-utr	Ngô	1
CaMV35S, IVS6-utr	Ngô	1
CaMV35S, m-HSP70-utr	Ngô	1
CaMV35S promoter tăng cường	Bông, khoai tây, đậu tương	1
CaMV35S promoter kép	Cà chua	4
Als1	Bông	1
BcNa	Cải dầu	1
CDPK	ngô	1
E8	Cà chua	1
FMV (CMoVb)	Bông, cải dầu	2
FMV, HSP70	Cà chua	1
Mas 5'	Cà chua	1
NOS	Rau riếp xoắn, cải dầu, đu đủ, cà chua	6
NOS, kép	Cà chua	1
PEPC	Ngô	1
SsuAra	Rau riếp xoắn, cải dầu, khoai tây	3
SsuHel (HelSsu)	Thuốc lá	1
TA29	Rau riếp xoắn, ngô, cải dầu	3
nda	Bông	1

Promoter đặc hiệu quả

Promoter đặc hiệu quả thường được sử dụng làm chậm quá trình chín và kéo dài thời gian chín của quả thông qua kỹ thuật di truyền. Các promoter này có giá trị để biểu hiện chính xác gen quan tâm trong các loại quả ở một giai đoạn phát triển nhất định. Đây là một quá trình cần thiết nhằm tạo cây trồng biến đổi gen (như cà chua, chuối, đu đủ và xoài) có khả năng chín chậm mà không cần sử dụng ethylene.

Promoter đặc hiệu bó mạch

Gần đây, promoter điều khiển gen mã hóa sucrose sythase ở cây lúa đặc biệt được quan tâm

ngiên cứu và ứng dụng do khả năng điều khiển đặc hiệu trong bó mạch (Yang, Russell, 1990). Shin *et al.* (1994) đã sử dụng đoạn promoter này để điều khiển quá trình biểu hiện gen mã hóa beta-glucuronidase và gen mã hóa lectin ở cây hoa diêm tuyết trong cây thuốc lá. Năm 1998, Rao và đồng tác giả đã sử dụng đoạn promoter này để điều khiển quá trình biểu hiện gen mã hóa lectin ở cây hoa diêm tuyết trong cây lúa cho hiệu quả diệt côn trùng cao.

Kết quả điều tra việc sử dụng các promoter trong những sản phẩm cây trồng biến đổi gen đã phê chuẩn cho thấy, CaMV35S promoter và các promoter có nguồn gốc từ CaMV35S (promoter lai hoặc kép) được sử dụng trong rất nhiều sản phẩm (có mặt trong

22 trên tổng số 28 sản phẩm điều tra). Các promoter này có thể chỉ khác nhau rất ít khi được thêm các vùng 5' không dịch mã đảm nhận chức năng điều hòa (Bảng 4). Trên tổng số 28 sản phẩm chuyển gen đã được phê chuẩn thì NOS promoter có mặt ít nhất trong 7 sản phẩm. Ngoài CaMV35S và NOS promoter, bốn sản phẩm mang gen ngoại lai được điều khiển bởi nhiều loại promoter, từ ribulose-1,5-bisphosphate carboxylase promoter có nguồn gốc từ nhiều loài thực vật. TA29 promoter là promoter đặc hiệu mô và đặc hiệu theo giai đoạn phát triển được

phân lập từ thuốc lá, tham gia điều khiển một hoặc một vài gen trong 3 sản phẩm đã phê chuẩn. Bảy loài cây trồng biến đổi gen có mang các promoter khác, trong khi 10 loài có mang các gen được điều khiển bởi promoter vi khuẩn. Một số promoter hiện mới chỉ được sử dụng trong một loại sản phẩm. Tuy nhiên, việc phát hiện các trình tự của những promoter này thường không phù hợp cho xác định GMOs bởi vì rất nhiều promoter, ví dụ, PEPC và CDPK, hai promoter đặc hiệu mô, vốn có nguồn gốc từ các giống cây trồng nông nghiệp.

Bảng 5. Các cơ sở dữ liệu về các trình tự promoter và các yếu tố điều hòa *cis* (Rombauts *et al.*, 2003).

Cơ sở dữ liệu	Mô tả	Web Sites
TRANSFAC	Cơ sở dữ liệu về TF	http://transfac.bgf.de/TRANSFAC/
TFD	Cơ sở dữ liệu về TF	http://www.tfdg.com/Pages/tfddata.html
TRRD	Cơ sở dữ liệu về các vùng điều hòa phiên mã	http://www.mgs.bionet.nsc.ru/mgs/dbases/trrd4/
PlantCARE	Các yếu tố điều hòa <i>cis</i> ở thực vật	http://sphinx.rug.ac.be:8080/PlantCARE/
PLACE	Các yếu tố điều hòa <i>cis</i> ở thực vật	http://www.dna.affrc.go.jp/htdocs/PLACE/
RegulonDB	Cơ sở dữ liệu về các yếu tố điều hòa phiên mã ở <i>Escherichia coli</i>	http://www.cifn.unam.mx/Computational_Genomics/regulondb/
SCPD	Cơ sở dữ liệu về promoter của <i>Saccharomyces cerevisiae</i>	http://cgsigma.cshl.org/jian
EPD	Cơ sở dữ liệu về promoter của sinh vật nhân chuẩn	http://www.epd.isb-sib.ch/

CƠ SỞ DỮ LIỆU VỀ CÁC TRÌNH TỰ CỦA PROMOTER

Việc xây dựng những cơ sở dữ liệu về các trình tự promoter và chức năng của chúng rất có giá trị về khoa học và thực tiễn. Bên cạnh đó, chương trình phần mềm thiết kế promoter nếu có cũng sẽ hỗ trợ rất nhiều cho công nghệ gen thực vật. Đây sẽ là nguồn thông tin quan trọng để các nhà khoa học sử dụng khi phân lập hoặc thiết kế một promoter đáp ứng những yêu cầu cụ thể trong tạo giống cây trồng biến đổi gen. Thay vì phải mất hàng tháng, hay hàng năm để phân lập và xác định chức năng của promoter, các nhà khoa học đơn giản chỉ cần đăng nhập vào cơ sở dữ liệu và chương trình, sau đó đưa ra các lựa chọn để thu thập thông tin tham khảo quan trọng. Hiện nay, có một số cơ sở dữ liệu cung cấp thông tin về các yếu tố *cis* tham gia kiểm soát quá trình phiên mã thông qua việc gắn kết với các yếu tố nhân tương ứng. Đó là các cơ sở TRANSFAC (Wingender *et al.*, 2001), TRRD (Kolchanov *et al.*, 2002), ooTFD (Ghosh, 2000), COMPEL (Kel-

Margoulis *et al.*, 2002), PlantCARE (Lescot *et al.*, 2002), PLACE (Higo *et al.*, 1999) và RegSite (<http://softberry.com>)... (Bảng 5). Ba cơ sở dữ liệu sau là bộ sưu tập các yếu tố phiên mã thực vật. Cơ sở dữ liệu promoter của sinh vật nhân chuẩn (the Eukaryotic Promoter Database - EPD) chỉ sưu tập các trình tự promoter Pol II của sinh vật nhân chuẩn (Praz *et al.*, 2002). Trong đó, bản công bố mới nhất bao gồm 1402 trình tự, chủ yếu là các promoter từ động vật, trong đó chỉ có 200 có nguồn gốc từ các loài thực vật. Trên đối tượng thực vật, Đại học London đã phối hợp với Softberry Inc. của Hoa Kỳ xây dựng một cơ sở dữ liệu PlantProm DB với các trình tự promoter thực vật từ rất nhiều loài khác nhau (<http://mendel.cs.rhul.ac.uk>; <http://www.softberry.com>). Bản công bố đầu tiên (2002.01) của PlantProm DB có chứa 305 trình tự trong đó có 71, 220 và 14 promoter tương ứng từ cây một lá mầm, hai lá mầm và các thực vật khác. Các trình tự DNA bao gồm vùng promoter với TSS và các yếu tố đặc trưng của promoter như hộp TATA, hộp CCAAT và TSS-motif (Shahmuradov *et al.*,

2003). Ngoài ra, một vài cơ sở dữ liệu cung cấp các thông tin về yếu tố *cis* của những loài cụ thể, ví dụ cơ sở dữ liệu AGRIS (Davuluri *et al.*, 2003) và AthaMap (Steffens *et al.*, 2004; Bülow *et al.*, 2006).

Bảng 6 cung cấp một số chương trình phân tích, dự đoán trình tự promoter như PromoterInspector (Scherf *et al.*, 2000); McPromoter (Ohler *et al.*, 2001).

Bảng 6. Các chương trình dự đoán promoter (Rombauts *et al.*, 2003).

Chương trình	Web Sites
MCPromoter MM	http://gene.mit.edu/McPromoter.html
PromoterInspector	http://www.gsf.de/biodv/ http://www.genomatix.de/cgi-bin/promoterinspector/promoterinspector.pl
FunSiteP	http://transfac.gbl.de/programs/funsitep/fsp.html
Dragon Promoter Finder	http://sdmc.lit.org.sg/promoter/promoter1_3/DPFV13.htm
CONPRO	http://stl.bioinformatics.med.umich.edu/conpro/
Core-promoter	http://argon.cshl.org/genefinder/CPROMOTER/human.htm
WWW PromoterScan	http://bimas.dcrf.nih.gov/molbio/proscan/
Promoter 2.0	http://www.cbs.dtu.dk/services/promoter/
NNPP	http://www.fruitfly.org/seq_tools/promoter.html

TÌNH HÌNH NGHIÊN CỨU Ở VIỆT NAM

Ở Việt Nam, với sự đầu tư của nhà nước trong xây dựng nguồn nhân lực và phòng thí nghiệm, chúng ta đã làm chủ được các kỹ thuật cơ bản và triển khai được một số nghiên cứu làm cơ sở để tạo cây trồng biến đổi gen. Từ năm 2001, Chương trình Công nghệ sinh học cấp Nhà nước đã đầu tư một số đề tài nghiên cứu tạo các cây trồng biến đổi gen. Bắt đầu từ năm 2006, Chương trình Công nghệ sinh học nông nghiệp do Bộ Nông nghiệp và Phát triển nông thôn chủ trì hàng năm cũng phê duyệt một số đề tài liên quan đến tạo cây trồng biến đổi gen. Nhìn chung, các hướng nghiên cứu chủ đạo liên quan đến tạo cây trồng biến đổi gen bao gồm phân lập, tuyển chọn các gen quý có giá trị ứng dụng cao từ nguồn tài nguyên sinh vật Việt Nam để chuyển vào sinh vật nhận nhằm tạo nên những giống ưu việt, thiết kế các vector mang gen có giá trị, hoàn thiện các phương pháp chuyển gen và tiến hành chuyển gen có giá trị vào những đối tượng cây trồng với nhiều nguồn gen khác nhau. Về mặt kỹ thuật, các công trình chuyển gen thường được triển khai sử dụng các vector được nước ngoài thiết kế sẵn mang gen chỉ thị hay chỉ dừng ở một số gen quan tâm (gen *cry* có hoạt tính diệt côn trùng, gen *bar* kháng thuốc diệt cỏ) với các promoter thông dụng như CaMV35S promoter và NOS promoter (Nguyễn Hữu Tâm *et al.*, 1998; Đỗ Tiến Phát *et al.*,

2005). Có thể nói, chỉ một số rất ít các promoter có giá trị được quan tâm nghiên cứu. Tại Viện Công nghệ sinh học, promoter của gen mã hóa sucrose synthase ở giống lúa C71 đã được tách dòng, xác định trình tự và sử dụng để thiết kế vector chuyển gen thực vật. Bên cạnh các trình tự đặc trưng, promoter này còn mang các yếu tố điều khiển biểu hiện gen đặc hiệu bó mạch như hộp ASL, hộp II, hộp GATA (Lê Thị Thu Hiền *et al.*, 2001; 2002). Hiện nay, trong khuôn khổ của các đề tài thuộc Chương trình Công nghệ sinh học Nông nghiệp, song song với những nghiên cứu về gen có giá trị, rất nhiều promoter đặc hiệu từ các nguồn tài nguyên sinh vật Việt Nam đang được triển khai phân lập và đưa vào ứng dụng góp phần phát triển hướng nghiên cứu tạo giống cây trồng biến đổi gen có giá trị trong nông, lâm nghiệp và y dược.

Lời cảm ơn: Công trình được thực hiện dưới sự hỗ trợ kinh phí của các đề tài thuộc Chương trình Công nghệ sinh học nông nghiệp thuộc Bộ Nông nghiệp và Phát triển nông thôn do Phòng Công nghệ ADN Ứng dụng, Viện Công nghệ Sinh học chủ trì hoặc tham gia.

TÀI LIỆU THAM KHẢO

Abe H, Yamaguchi-Shinozaki K, Urao T, Iwasaki T, Hosokawa D, Shinozaki K (1997) Role of *Arabidopsis* MYC and MYB homologs in drought - and abscisic

acid - regulated gene expression. *Plant Cell* 9: 1859-1868.

Albani D, Sardana R, Robert LS, Altosaar I, Arnison PG, Fabijanski SF (1992) A *Brassica napus* gene family which shows sequence similarity to ascorbate oxidase is expressed in developing pollen. Molecular characterization and analysis of promoter activity in transgenic tobacco plants. *Plant J* 2: 331-342.

Amin J, Anathan J, Voellmy R (1988) Key features of heat shock regulatory elements. *Mol Cell Biol* 8: 3761-3769.

An G (1986) Development of plant promoter expression vectors and their use for analysis of differential activity of nopaline synthase promoter in transformed tobacco cells. *Plant Physiol* 81: 86-91.

Arayan AP, An G, Okita TW (1991) Structural and functional analysis of promoter from gliadin, an endosperm-specific storage protein gene of *Triticum aestivum*. *Mol Gen Genet* 225: 67-71.

Benfy PN, Chua NH (1990) The cauliflower mosaic virus 35S promoter: combination regulation of transcription in plants. *Science* 250: 959-66.

Benoist C, Chambon P (1981) *In vivo* sequence requirements of the SV40 early promoter region. *Nature* 290: 304-310.

Berk AJ (1999) Activation of RNA polymerase II transcription. *Curr Opin Cell Biol* 11: 30-38.

Braun EL, Grotewold E (1999) Newly discovered plant *c-myc*-like genes rewrite the evolution of the plant *myb* gene family. *Plant Physiol* 121: 21-24.

Breathnach R, Chambon P (1981) Organization and expression of eucaryotic split genes coding for proteins. *Ann Rev Biochem* 50: 349-383.

Buchanan CD, Klein PE, Mullet JE (2004) Phylogenetic analysis of 5' - noncoding regions from the ABA- responsive *rab* 16/17 gene family of sorghum, maize and rice provides insight into the composition, organization and function of cis-regulatory modules. *Genetics* 168: 1639-1654.

Buzby JS, Yamada T, Tobin EM (1990) A light-regulated DNA-binding activity interacts with a conserved region of a *Lemna gibba* *rbcS* promoter. *Plant Cell* 2: 805-814.

Bülow L, Steffens NO, Galuschka C, Shindler M, Hehl R (2006) AthaMap: from in silico data to real transcription factor binding sites. *Silico Biol* 6: 23.

Cheng X, Sardana R, Kaplan H, Altosaar I (1998) *Agrobacterium* - transformed rice plants expressing

synthetic *cryIA(b)* and *cryIA(c)* genes are highly toxic to striped stem borer and yellow stem borer. *Proc Natl Acad Sci USA* 95: 2767-2772.

Chrispeels MJ, Sadava DE (2002) Plants, Genes, and Crop Biotechnology. Jones and Bartlett publishers, Sudbury, Massachusetts, USA.

Christensen AH, Quail PH (1996) Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res* 5: 213-218.

Christensen AH, Sharrock RA, Quail PH (1992) Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol Biol* 18: 675-689.

Comai L, Moran P, Maslyar D (1990) Novel and useful properties of a chimeric plant promoter combining CaMV 35S and MAS elements. *Plant Mol Biol* 15(3): 373-381.

Conkling MA, Cheng CL, Yamamoto YT, Goodman HM (1990) Isolation of Transcriptionally Regulated Root-Specific Genes from Tobacco. *Plant Physiol* 93: 1203-1211.

Cornejo MJ, Luth D, Blankenship KM, Blankenship KM, Anderson OD, Blechl AE (1993) Activity of a maize ubiquitin promoter in transgenic rice, *Plant Mol Biol* 23(3): 567-81.

Datta K, Vasquez A, Tu J, Torrizo L, Alam MF, Oliva N, Abrigo E, Khush GS, Datta SK (1998) Constitutive and tissue-specific differential expression of the *cryIA(b)* gene in transgenic rice plants conferring resistance to rice insect pest. *Theor Appl Genet* 97: 20-30.

Davuluri RV, Sun H, Palaniswamy SK, Matthews N, Monila C, Kurtz M, Grotewold E (2003) AGRIS: *Arabidopsis* gene regulatory information serve, an information resource of *Arabidopsis* cis-regulatory elements and transcription factors. *BMC Bioinformatics* 4: 25.

Đỗ Tiến Phát, Nguyễn Xuân Tài, Đinh Thị Phòng, Chu Hoàng Hà (2005) Chuyển gen *gus* vào cây bông (*Gossypium hirsutum* L.) thông qua *Agrobacterium tumefaciens*. *Hội nghị Khoa học toàn quốc 2005, Công nghệ sinh học trong nghiên cứu cơ bản*: 293-296.

Ebert PR, Ha SB, An G (1987) Identification of an essential upstream element in the nopaline synthase promoter by stable and transient assays. *Proc Natl Acad Sci USA* 84(16): 5745-5749.

Ellis JG, Llewellyn DJ, Dennis ES, Peacock WJ (1987)

Maize Adh-1 promoter sequences control anaerobic regulation: addition of upstream promoter elements from constitutive genes is necessary for expression in tobacco. *EMBO J* 6: 11-16.

Fang RX, Nagy F, Sivabramaniam S, Chua NH (1989) Multiple cis regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. *Plant Cell* 1: 141-150.

Featherstone (2002) Coactivator in transcription initiation: here are your orders. *Curr Opin Genet Dev* 12: 149-155.

Fessele S, Maier H, Zischek C, Nelson PJ, Werner T (2002) Regulatory context is a crucial part of gene function. *Trends Genet* 18: 60-63.

Fischer U, Weisbeek P, Smekens S (1994) Identification of potential regulatory elements in the far-upstream region of the Arabidopsis thaliana plastocyanin promoter. *Plant Mol Biol* 26(3): 873-886.

Fluhr R, Kuhlemeier X, Nagy F, Chua NH (1986) Organ-specific and light-induced expression of plant genes. *Science* 232: 1106-1112.

Fromm ME, Taylor LP, Walbot V (1986) Stable transformation of maize after gene transfer by electroporation. *Nature* 319: 791-793.

Ghosh D (2000) Object oriented transcription factors database (ootFD). *Nucl Acids Res* 28: 308-310.

Giuliano G, Pichersky E, Malik VS, Timko MP, Scolnik PA, Cashmore AR (1988) An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. *Proc Natl Acad Sci USA* 85(19): 7089-7093.

Gruss P, Dhar R, Khoury G (1981) Simian virus 40 tandem repeated sequences as an element of the early promoter. *Proc Natl Acad Sci USA* 78: 943-947.

Guitinan MJ, Marcotte WR Jr, Quatrano RS (1990) A plant leucine zipper protein that recognizes an abscisic acid response element. *Science* 250(4978): 267-271.

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database. *Nucl Acids Res* 27: 297-300.

Ho M-W, Ryan A, Cummins J (1999) Cauliflower mosaic viral promoter - A recipe for disaster? *Microb Ecol Health Dis* 11: 194-197.

Hohn T, Fütterer J (1992) Transcriptional and translation control of gene expression in cauliflower mosaic virus. *Curr Opin Genet Develop* 2: 906.

http://www.bats.ch/bats/publikationen/1997-2_gmo/4.1.2_promoters.php.

Huang JM, Chen JT, Yu WP, Shyur LF, Wang AY, Sung HY, Lee PD, Su JC (1996) Complete structures of three rice sucrose synthase isogenes and differential regulation of their expressions. *Biosci Biotechnol Biochem* 60: 233-239.

Hudspeth RL, Grula JW, Dai Z, Edwards GE, Ku MS (1992) Expression of Maize Phosphoenolpyruvate Carboxylase in Transgenic Tobacco: Effects on Biochemistry and Physiology. *Plant Physiol* 98(2): 458-464.

Johnson PF, McKnight SL (1989) Eukaryotic transcriptional regulatory proteins. *Annu Rev Biochem* 58: 799-839.

Jones JDG, Dunsmuir P, Bedbrook J (1985) High level expression of introduced chimaeric genes in regenerated transformed plants. *EMBO J* 4: 2411-2418.

Kay R, Chan A, Daly M, McPherson J (1987) Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science* 236: 1299-1302.

Kel-Margoulis OV, Kel AE, Reuter I, Deineko IV, Wingender E (2002) TRANSCOMP: a database on composite regulatory elements in eukaryotic genes. *Nucl Acids Res* 30: 332-334.

Khoury G, Gruss P (1983). Enhancer elements. *Cell* 33(2): 313-314.

Kohli A, Griffiths S, Palacios N, Twyman RM, Vain P, Laurie DA, Christou P (1999) Molecular characterization of transforming plasmid rearrangements in transgenic rice reveals a recombination hotspot in the CaMV 35S promoter and confirms the predominance of microhomology mediated recombination. *Plant J* 17, 591-601.

Kolchanov NA, Ignatieva EV, Ananko EA, Pdkolodny NL, Naumochkin AN, Romashchenko AG (2002) Transcription regulatory regions database (TRRD): its status in 2002. *Nucl Acids Res* 30: 312-317.

Laporte MM, Galagan JA, Prasch AL, Vanderveer PJ, Hanson DT, Shewmaker CK, Sharkey TD (2001) Promoter strength and tissue specificity effects on growth of tomato plants transformed with maize sucrose-phosphate synthase. *Planta* 212(5-6): 817-822.

Last DI, Brettell DA, Chamberlain AM, Chaudhury AM, Larkin PJ, Marsh EL, Peacock WJ, Dennis ES (1991) pEmu: an improved promoter for gene expression in cereal cells. *Theor Appl Genet* 81: 581-588.

Lê Thị Ánh Hồng (2006) Nghiên cứu ứng dụng công nghệ sinh học trong chọn tạo giống kháng sâu bệnh và có chất lượng cao ở một số cây có củ. *Báo cáo tổng kết*

khoa học kỹ thuật để tài KC.04.22: Bộ Nông nghiệp và Phát triển Nông thôn, Hà Nội.

Lê Thị Thu Hiền, Đinh Duy Kháng, Lê Trần Bình, Nông Văn Hải (2002) Thiết kế vectơ mang gen độc tố *cryIA(c)* dưới sự điều khiển của đoạn khởi động của gen tổng hợp đường phân lập từ giống lúa C71. *Tạp chí Khoa học và Công nghệ* 40: 135-141.

Lê Thị Thu Hiền, Lê Trần Bình, Đinh Duy Kháng, Nông Văn Hải (2001) Phân tích trình tự của đoạn điều khiển của gen tổng hợp đường phân lập từ giống lúa C71. *Tạp chí Sinh học* 23(2): 45-50.

Lê Trần Bình (2005) Nghiên cứu áp dụng công nghệ gen để tạo cây chuyển gen nâng cao sức chống chịu đối với sâu bệnh và ngoại cảnh bất lợi. *Báo cáo tổng kết khoa học kỹ thuật để tài KC.04.13*: Viện Khoa học và Công nghệ Việt Nam, Hà Nội.

Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts P (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucl Acids Res* 30: 325-327.

Liu Q, Sakuma Y, Abe H, Kasuga M, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain, separate two cellular signal transduction pathways in drought and low temperature responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell* 10: 1391-1406.

Lohmann C, Eggert-Schumacher G, Wunderlich M, Schoffl F (2004) Two different heat shock transcription factors regulate immediate early expression of stress genes in *Arabidopsis*. *Mol Genet Genomics* 271(1): 11-21.

Lorenzo O, Chico JM, Sanchez-Serrano JJ, Solano R (2004). Jasmonate-insensitive1 encodes a MYC transcription factor essential to discriminate between different jasmonate-regulated defense responses in *Arabidopsis*. *Plant Cell* 16: 1938-1950.

Mackenzie R, Burhenne-Guilmin F, La Vina AGM, Werksman JD (2003) *An explanatory guide to the Cartagena Protocol on Biosafety*. IUCN Environmental Policy and Law Paper N^o46.

McElroy D, Zhang W, Cao J, Wu R (1990) Isolation of an efficient actin promoter for use in rice transformation. *Plant Cell* 2: 163-171.

Mohanty B, Krishnan SPT, Swarup S, Bajic V (2005) Detection and preliminary analysis of motifs in promoters of anaerobically induced genes of different plant species. *Ann Bot* 96: 669-681.

Nguyễn Hữu Tâm, Lê Tấn Đức, Nguyễn Thị Thanh,

Nguyễn Văn Uyển (1998) Sử dụng vi khuẩn *Agrobacterium tumefaciens* để chuyển gen kháng hygromycin và gen *gusA* vào lúa *indica* (*Oryza sativa* L.). *Tóm tắt Báo cáo Hội nghị toàn quốc lần thứ nhất về Công nghệ Sinh học cây lúa*: 48-49.

Nikolov DB, Burley SK (1997) RNA polymerase II transcription initiation: a structural view. *Proc Natl Acad Sci USA* 94: 15-22.

Noad RJ, Turner DS, Covey SN (1997) Expression of functional elements inserted into the 35S promoter region of infectious cauliflower mosaic virus replicons. *Nucl Acids Res* 25, 1123-1129.

NRC/PBI Promoter Technology. <http://pbi.nrc-cnrc.gc.ca/en>.

Odell JT, Knowlton S, Lin W, Mauvais CJ (1988) Properties of an isolated transcription stimulating sequence derived from the cauliflower mosaic virus 35S promoter. *Plant Mol Biol* 10: 263-272.

Odell JT, Nagy F, Chua N-H (1985) Identification of sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature* 313: 810-812.

Ou-Lee TM, Turgeon R, Wu R (1986) Expression of a foreign gene linked to either a plant-virus or a *Drosophila* promoter, after electroporation of a protoplasts of rice, wheat, and sorghum. *Proc Natl Acad Sci USA* 83: 6815-6819.

Praz V, Périer R, Bonnard C, Bucher P (2002) The eukaryotic promoter database, EPD: new entry types and links to gene expression data. *Nucl Acids Res* 30: 322-324.

Rao KV, Rathore KS, Hodges TK, Fu X, Stoger E, Sudhakar D, Williams S, Christou P, Bharathi M, Bown DP, Powell KS, Spence J, Gatehouse AM, Gatehouse JA (1998) Expression of snowdrop lectin (GNA) in transgenic rice plants confers resistance to rice brown planthopper. *Plant J* 15(4): 469-477.

Robert L (1998) Flower-specific promoters. *PBI Bulletin*. <http://www.pbi.nrc.ca/cgi-bin/bulletin/may1998.sh?promoters.html>.

Rombauts S, Florquin K, Lescot M, Marchal K, Rouzé, Van de Peer Y (2003) Computational approaches to identify promoters and cis-regulatory elements in plant genomes. *Plant Physiol* 132: 1162-1176.

Sanders PR, Winter JA, Zarnason AR, Roger SG, Fraley RT (1987) Comparison of cauliflower mosaic virus 35S and nopaline synthase promoters in transgenic plants. *Nucl Acids Res* 15: 1543-1558.

Sardana R, Altosaar I (1996) Identification and use of specific promoters and synthetic *Bacillus thuringiensis*

toxin genes in rice biotechnology. *Rice Genet* III: 723-729.

Sardana R, Dukiandjiev S, Giband M, Cheng X, Cowan K, Sauder C, Altosaar I (1996) Construction and rapid testing of synthetic and modified toxin gene sequences CryIA(b&c) by expression in maize endosperm culture. *Plant Cell Rep* 15: 677-681.

Scandalios JG (2005) Oxidative stress: molecular perception and transduction of signals triggering antioxidant gene defenses. *Braz J Med Biol Res* 38: 995-1014.

Schledzewski K, Mendel R (1994) Quantitative transient gene expression: Comparison of the promoters for maize polyubiquitin 1, rice actin 1, maize derived Emu and CaMV 35S in cells of barley, maize and tobacco. *Transgenic Res* 3: 249-255.

Shahmuradov IA, Gammernan AJ, Hancock JM, Bramley PM, Solov'yev VV (2003) PlantProm: a database of plant promoter sequences. *Nucl Acids Res* 31(1): 114-117.

Shimamoto K (1991) Transgenic rice plants. In: *Molecular approaches to crop improvement*, Dennis ES, Llewellyn DJ (Eds.), Springer-Verlag, Wien, New York, 1-15.

Shin DL, Podila GK, Huang Y, Karnosky DF (1994) Transgenic larch expressing genes for herbicide and insect resistance. *Can J For Res* 24: 2059-2067.

Shinozaki K, Yamaguchi-Shinozaki K (1999) *Molecular responses to cold, drought, heat and salt stress in higher plants*. RG Landes company, Austin, USA: 11-28.

Smith RH, Hood EE (1995) *Agrobacterium tumefaciens* transformation of monocotyledons. *Crop Science* 35(2): 301-309.

Steffens NO, Galuschka C, Schindler M, Bülow L, Hehl R (2004) AthaMap: an online resource for in silico transcription factor binding sites in the *Arabidopsis thaliana* genome. *Nucl Acids Res* 32: D368-372.

Strittmatter G, Chua NH (1987) Artificial combination of two cis-regulatory elements generates a unique pattern of expression in transgenic plants. *Proc Natl Acad Sci USA* 84: 8986- 8990.

Sullivan TD, Christensen AH, Quail PH (1989) Isolation and characterization of a maize chlorophyll a/b binding protein gene that produces high levels of mRNA in the dark. *Mol Gen Genet* 215: 431-440.

Timko MP, Kausch AP, Castresana C, Fassler J, Herrera-Estrella L, Van de Broeck G, Van Montagu M, Schell J, Cashmore AR (1985) Light regulation of plant

gene expression by an upstream enhancer-like element. *Nature* 318: 579-582.

Tjian R, Maniatis T (1994) Transcriptional activation: a complex puzzle with a few easy pieces. *Cell* 77: 5-8.

Toki S, Takamatsu S, Nojiri C, Ooba S, Anazi H, Iwata M, Christensen AH, Quail PH, Uchimiya H (1992) Expression of a maize ubiquitin gene promoter-bar chimeric gene in transgenic rice plants. *Plant Physiol* 100: 1503-1507.

Trần Thị Phương Liên, Lê Xuân Tú, Rumar SE (1992) Nhân đôi chuỗi promoter 35S virus khảm súp lơ. *Tap chí Sinh học* 14(3): 21-23.

Turner DS, McCallum DG, Covey SN (1996) Roles of the 35S promoter and multiple overlapping domains in the pathogenicity of the pararetrovirus cauliflower mosaic virus. *J Virol* 70, 5414-5421.

US patent 6987179. Constitutive plant promoters. <http://www.freepatentsonline.com/6987179.html>.

Walden R, Wingender R (1995) Gene-transfer and plant regeneration techniques. *Trends Biotechnol* 13: 324-331.

Wang Y, Zhang W, Cao J, McEhoy D, Wu R (1992) *Mol Cell Biol* 12: 3399-3406.

Weeks KE, Chuzhanova NA, Donnison IS, Scott IM (2007) Evolutionary hierarchies of conserved blocks in 5'-noncoding sequences of dicot *rbcS* genes. *BMC Evol Biol* 7: 51. <http://www.biomedcentral.com/1471-2148/7/51>.

Wingender E, Chen X, Fricke E, Geffers R, Hehl R, Liebich I, Krull M, Matys V, Michael H, Ohnhäuser R, Prüß M, Schacherer F, Thiele S, Urbach S (2001) The TRANSFAC system on gene expression regulation. *Nucl Acids Res* 29: 281-283.

Yamamoto YY, Ichida H, Matsui M, Obokata J, Sakurai T, Satou M, Seki M, Shinazaki K, Abe T (2007) Identification of plant promoter constituents by analysis of local distribution of short sequences. *BMC Genomics* 8: 67. <http://www.biomedcentral.com/1471-2164/8/67>.

Yang NS, Russell D (1990) Maize sucrose synthase-1 promoter directs phloem cell-specific expression of GUS gene in transgenic tobacco plants. *Proc Natl Acad Sci USA* 87: 4144-4148.

Yin Y, Chen L, Beachy R (1997) Promoter elements required for phloem-specific gene expression from the RTBV promoter in rice. *Plant J* 12: 1179-1188.

Yu WP, Wang AY, Yuang RH, Sung HY, Su JC (1992) Isolation and sequences of rice sucrose synthase cDNA and genomic DNA. *Plant Mol Biol* 18: 139-142.

PROMOTERS AND THEIR APPLICATION IN PLANT GENE TECHNOLOGY

Le Thi Thu Hien, Tran Thi Phuong Lien, Nong Van Hai*

Institute of Biotechnology

SUMMARY

The term 'promoter' is used to designate a region in the genome sequence upstream, or 5' end, of a gene transcription start site (TSS), which is defined as position +1. Promoters have complex structures and contain various distinct sequence elements that function in the recruitment of protein factors facilitating the transcription of the protein-coding region of the gene. One essential element is the TATA box that serves as a binding place for transcription initiation factors and is typically located around -30 bp upstream of the TSS. Another common element is CCAAT box. In plants, the CCAAT box may have a different consensus sequence and has been termed as the AGGA box. Besides TATA and CCAAT boxes, virtually all eukaryotic promoters contain additional upstream DNA sequences that are variously known as enhancers or upstream activating sequences. Such sequences are variable in length and may extend from around 100 to 1,000 bp or more upstream of the TSS. A promoter can either be constitutive or spatial/ temporal specific according to its ability to express the gene that it controls. In the field of plant gene technology, promoters have been widely used for the expression of exogenous genes in transgenic plants. The most famous are the CaMV35S, CaMV19S and nopaline synthase (NOS) promoters. To date, many attempts have been made to isolate and construct specific or novel promoters with high level of expression or with precisely controlled activities for plant genetic engineering applications. Besides that, a number of databases with information on promoter sequences and their functions and of promoter design software programs have been developed and applied. In Vietnam, the isolation and characterization of promoters from natural resources for construction of plant expression vectors are under development.

Keywords: *Gene expression, plant gene technology, promoter, regulatory element, transcription factor*

* Author for correspondence: Tel/Fax: 84-4-8363222; E-mail: vhnong@ibt.ac.vn