

BÀI TỔNG QUAN

UBIQUITIN, CẤU TRÚC, CHỨC NĂNG VÀ KHẢ NĂNG SỬ DỤNG PROMOTER UBIQUITIN TRONG CÔNG NGHỆ GEN THỰC VẬT

Trần Thị Phương Liên¹, Chu Hoàng Hà¹, Nông Văn Hải²

¹Viện Công nghệ sinh học, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

²Viện Nghiên cứu hệ gen, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

TÓM TẮT

Ubiquitin, protein có khối lượng phân tử nhỏ, bảo thủ, có mặt trong hầu hết các loài mô tế bào nhân thực. Trong cấu trúc gen, *ubiquitin* có hai dạng: *polyubiquitin* và gen *ubiquitin* protein mơ rỗng (còn có tên gọi là gen *ubiquitin* liên kết). Trong mỗi loài có ít nhất một *polyubiquitin* locus chứa các đoạn lặp lại 228 bp mã hóa cho ubiquitin đơn 76 amino acid của polypeptide này. Số lượng các đoạn lặp lại có thể khác nhau giữa các locus trong một loài và giữa các loài. Ubiquitin tham gia nhận biết các protein, polypeptide bị biến tính hoặc cần loại bỏ thông qua quá trình ubiquitin hóa nhờ hệ thống ubiquitin-proteasome. Có 3 loại enzyme tham gia vào quá trình phân giải protein này. Đó là enzyme hoạt hóa ubiquitin E1, enzyme tiếp hợp với ubiquitin E2 và enzyme gắn với ubiquitin E3 (E3 ligase). Sự ubiquitin hóa tham gia điều hòa chức năng nhiều loại protein trong tế bào, đóng vai trò quan trọng trong quá trình sinh trưởng và phát triển của cây cũng như khả năng kháng bệnh và chống chịu với các yếu tố cực đoan trong môi trường. Sự đa dạng về chức năng của ubiquitin đã tạo tiền đề sử dụng promoter của loài gen này như một loại promoter có định hướng hiệu trong công nghệ gen thực vật. Các nghiên cứu cho thấy promoter *ubiquitin* có khả năng biểu hiện gen trong thực vật cao hơn promoter 35S CAMV nhiều lần. Đối với cây một lá mầm khả năng sử dụng promoter này cao hơn đối với cây hai lá mầm. Trong bài này, các nghiên cứu về ubiquitin, các gen mã hóa cho chúng, giá thiết về các chức năng của chúng và tiếp theo đó là các nghiên cứu phân lập, sử dụng promoter của gen *ubiquitin* một hướng ứng dụng đầy triển vọng trong công nghệ gen thực vật sẽ được tổng quan và phân tích.

Từ khóa: ubiquitin, hệ ubiquitin-proteasome, E3 ligase, cấu trúc promoter, biểu hiện gen ở thực vật

MỞ ĐẦU

Ubiquitin là protein bảo thủ với khối lượng phân tử nhỏ 8,5 kDa được tồn tại trong hầu hết các tế bào nhân thực. Hiện tượng ubiquitin hóa hay quá trình ubiquitin hóa (ubiquitination) được định nghĩa là quá trình gắn các đoạn ubiquitin lên một phân tử protein, quá trình này có vai trò trong điều khiển hoạt động của các protein và phân hủy của chúng. Những nghiên cứu sâu về ubiquitin cho thấy chúng tương tác với histon cũng như nhiều loại protein khác nhau và tham gia vào quá trình phân hủy protein nội bào bên ngoài lysosome. Để khởi đầu cho sự phân hủy, ubiquitin cần nhận biết cơ chất - protein cần loại bỏ, và chuyển cho proteasome phân giải thành peptide và amino acid (aa). Quá trình này được thực hiện nhờ hệ thống ubiquitin-proteasome - UPS (Ubiquitin proteasome system). Năm 2004, ba nhà nghiên cứu Aaron Ciechanover, Avram Herskko và Irwin Rose đã được nhận giải Nobel hóa học về phát minh ra sự tương tác giữa ubiquitin với protein cần được phân hủy - sự nhận biết đặc hiệu khon đầu cho hiện tượng

ubiquitin hóa http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2004/

Sự ubiquitin hóa cho phép điều hòa chức năng của nhiều loại protein khác nhau, mở khóa các protein có vai trò quan trọng trong tế bào, hoạt hóa các yếu tố phiên mã (transcription factor - TF), khởi đầu cho nhiều quá trình biểu hiện gen, cho sự phân chia tế bào, sao chép DNA cũng như duy trì sự sống trong tế bào. Trong thực vật, các protein nội bào cũng được điều hòa thông qua sự phân giải chúng nhờ UPS. Sự ubiquitin hóa điều hòa quá trình sinh trưởng và phát triển, điều hòa các hormone sinh trưởng thực vật, quá trình phân bào cũng như phản ứng với các yếu tố cực đoan.

Vì ubiquitin có mặt ở mọi tế bào, mô của cơ thể thực vật và có chức năng đa dạng, việc nghiên cứu sâu về gen mã hóa cho ubiquitin, tìm kiếm khả năng sử dụng promoter của chúng đã được tiến hành từ nhiều năm nay. Trong bài này, chúng tôi tổng quan các nghiên cứu về ubiquitin, các gen mã hóa cho chúng, giá thiết về các chức năng của chúng và tiếp

theo đó là các nghiên cứu phân lập và sử dụng promoter của gen *ubiquitin* trong biến nạp gen ở thực vật.

Ubiquitin, cấu trúc của gen mã hóa ubiquitin.

Ubiquitin được tách chiết lần đầu tiên vào năm 1975 từ tuyến ức bò (Goldstein *et al.*, 1975). Nhưng trong thời gian đó, chức năng của chúng vẫn còn là vấn đề bí ẩn. Sau này, ubiquitin đã được tìm thấy trong nhiều loại mô tế bào của các cơ thể nhân thực khác nhau. Tiếp theo, Busch và cộng sự đã phát hiện thấy dạng liên kết đầu tiên của loại protein này với histon H2A và gọi phức hệ này là protein A24 trong nhân tế bào (Goldknopf, Busch, 1977). Năm 1978, Ciechanover và cộng sự nghiên cứu trên dịch chiết reticulocytes (tế bào hồng cầu chưa trưởng thành) đã tìm thấy một loại protein bền nhiệt có liên quan sự phân hủy protein phụ thuộc vào năng lượng ATP và sau này gọi chúng là APF1 (active principle of fraction 1). Protein được xác định với khối lượng phân tử khoảng 9000 Da, Wilkinson và cộng sự xác nhận protein này là ubiquitin. Theo tiếng Latinh, "ubique" có nghĩa là mọi nơi (everywhere) (Ciechanover *et al.*, 1978, Wilkinson *et al.*, 1980). Ubiquitin còn được xếp vào nhóm protein sốc nhiệt (heat-shock protein) theo đặc tính phản ứng của chúng khi nghiên cứu protein này trong các sợi nguyên bào phôi gà (Bond, Schlesinger, 1985). Trong nguyên sinh chất, ubiquitin tham gia vào quá trình phân hủy protein bên ngoài lysosome (nonlysosomal proteolysis) (Ciechanover *et al.*, 1981, Pickart, Eddins, 2004; Moon *et al.*, 2004).

Ubiquitin được phân lập từ nhiều loài khác nhau, nhưng trình tự 76 aa đều có cấu trúc bảo thủ. Giữa nấm men, thực vật và động vật, ubiquitin chỉ khác nhau bởi 4 gốc aa (Sharp, Li, 1987). Trình tự ubiquitin 76 aa ở người như sau.



Hình 1 Cấu trúc locus *ubiquitin* (Sharp, Li, 1987)

Trong thực vật, số đoạn ubiquitin đơn, chứa 76 aa, lặp lại này khác nhau giữa các gen trong cùng một loại cây. Nghiên cứu ở *Arabidopsis* (Mỹ) đã phân lập được 5 gen *UBQ3*, *UBQ4*, *UBQ10*, *UBQ11* và *UBQ14*, số đoạn lặp lại là 3 đoạn đối với *UBQ11*,

MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDK
EGIPPDQQRLLFAGKQLEDGRTLSDYNIQKEST
LHLVLRLRGG

Gen *ubiquitin*: Ho gen *ubiquitin* có hai dạng gen với cấu trúc khác nhau: *polyubiquitin* và gen *ubiquitin* protein mở rộng (*ubiquitin extension protein gene*). còn có tên gọi khác là gen *ubiquitin* liên kết (*ubiquitin fusion gene*). Người ta tìm thấy trong mỗi loài ít nhất một *polyubiquitin* locus, có chứa các đoạn lặp lại 228 bp liên tiếp nhau, mã hóa cho ubiquitin đơn 76 aa của polypeptide này. Trong vùng mã hóa này không phát hiện thấy có intron giữa các đoạn lặp lại này. Gen *ubiquitin* protein mở rộng bao gồm 2 loại, mỗi loại chứa một đoạn ubiquitin đơn và đoạn mã hóa cho polypeptide ribosome L40 hoặc S27 (hoặc còn gọi là gen *ubiquitin* liên kết *uf80*, *uf52*). Protein đã thuần thực là ubiquitin dạng đơn được tách ra từ chuỗi polypeptide này (Catic, Ploegh, 2005, Sullivan *et al.*, 2003). Sharp và cộng sự (1987) nghiên cứu cấu trúc của vùng mã hóa cho *polyubiquitin* của gen *polyubiquitin* ở các loài khác nhau (Hình 1). Họ thấy rằng số đoạn ubiquitin đơn lặp lại có thể dao động trong khoảng từ 1 đến 9 (ở người), cho đến 50 ở *Trypanosoma cruzi*. Các locus có thể khác nhau ở gốc amino acid cuối cùng không nằm trong thành phần đoạn lặp lại. Chúng nằm gần kề với codon kết thúc. Đó có thể là GTC (Val) ở người; TAT (Tyr) ở gà, hay AAC (Asn) ở nấm men. Tan và cộng sự (1993) phân tích trên 40 gen *ubiquitin* của cả 2 nhóm cũng cho thấy sự đa dạng của gene này giữa các đoạn lặp lại cả trong một locus cũng như giữa các locus khác nhau trong một loài. Tuy trình tự amino acid bảo thủ, nhưng trình tự nucleotide có mức độ đa dạng nhất định. Độ đồng nhất giữa các đoạn lặp lại trong cùng một gene, giữa các gen trong cùng một loài, giữa các loài dao động trong khoảng 50-98% (Sun *et al.*, 1997, Tan *et al.*, 1993).

là 4 đoạn đối với *UBQ3* và *UBQ14*, là 5 đoạn đối với *UBQ4*, và là 6 đoạn đối với *UBQ10*. Số gen này với gen cùng loại có gốc từ Đức, Ba Lan, Nga, Pháp, cho thấy 3 gen *UBQ4*, *UBQ10*, và *UBQ14* không thấy có sự khác biệt về số đoạn lặp lại. Còn hai gen *UBQ3*,

UBQ11 có số đoạn lặp lại dao động trong khoảng 3-6 đoạn phụ thuộc vào các dòng cây có nguồn gốc khác nhau. Nghiên cứu trên cây đậu Hà Lan đã phân lập được 4 gen *ubiquitin PUB11*, *PUB12* *PUB13* *PUB14* Số đoạn lặp lại từ 5-6 đoạn. Ở đậu tương, trong 9 gen *polyubiquitin*, gen *Gmubi1*, *Gmubi2*, *Gmubi3* và *Gmubi8* có số đoạn lặp lại là 4, còn *Gmubi4*, *Gmubi6* có số đoạn lặp lại này là 7. Trong khi đó, *Gmubi5*, *Gmubi7* và *Gmubi9* có số đoạn lặp lại tương ứng là 6, 5 và 2. Cây hoa hướng dương có 2 gen với số đoạn lặp lại là 6 đoạn. Gốc amino acid cuối cùng có thể là AAG (Lys) ở đại mạch, CAG (Gln) ở ngô và lúa, TTT (Phe) ở đậu tương. Riêng cà chua có thêm 2 amino acid. Asp và Leu, còn đậu Hà Lan - Phe và Iso. Số locus trong từng loài cũng dao động từ 2 đến 10 locus (Binet *et al.*, 1991, Callis *et al.*, 1990, Sun *et al.*, 1997, Christensen *et al.*, 1992, Hernandez-Garcia *et al.* 2010, Rollfinke, Pfützner, 1994, Xia, Mahon, 1998).

Các cơ thể tiền nhân không có dạng phân tử protein nào giống như ubiquitin. Các protein có thể tương tác với với các protein khác như dạng ubiquitin cũng không phát hiện được. Tuy nhiên, vì khuẩn có hai protein, có thể là tổ tiên của ubiquitin, trong cấu trúc phân tử của chúng có những đoạn gấp khúc theo cách thức như ubiquitin. Đó là ThiS và MoaD (Pickart, Eddins, 2004).

Hệ ubiquitin - proteasome

Trong các cơ thể sống, protein phân hủy theo hai con đường: Protein ngoại bào được các enzyme thủy phân protein phân hủy trong lysosome hoặc các bào quan tương tự. Các enzyme đó là các protease, cysteine protease, serine protease, carboxypeptidase, aminopeptidase Trong nguyên sinh chất và trong nhân tế bào, các protein nội bào và polypeptide cần đào thải hoặc bị biến tính thường bị phân giải trong proteasome với sự tham gia của một loại protein với khối lượng phân tử nhỏ - ubiquitin. Sự phát hiện ra ubiquitin như chất đánh dấu dẫn gọi protein cần được phân hủy trong nguyên sinh chất đã mở ra một cách nhìn mới trong quá trình đào thải protein trong tế bào.

Sự ubiquitin hóa

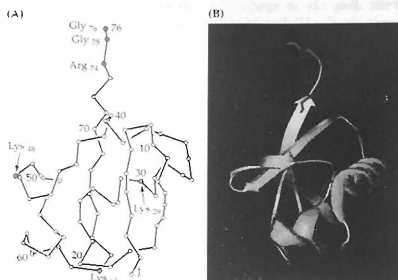
Trong chuỗi 76 aa của ubiquitin có một số vị trí quan trọng như Gly76 ở đầu C, còn các gốc Lys29, Lys48 và Lys63 tham gia vào tạo chuỗi polyubiquitin trong quá trình ubiquitin hóa (Hình 2).

Quá trình tương tác giữa ubiquitin và protein có sự tham gia hoạt hóa của nhiều enzyme. Chúng được vận chuyển theo quy trình phức tạp bao gồm ba bước chính sử dụng năng lượng ATP. Trong đó, có ít nhất 3 loại enzyme tham gia vào quá trình này: Enzyme hoạt hóa ubiquitin (ubiquitin-activating enzyme) E1; enzyme tiếp hợp với ubiquitin (ubiquitin-conjugating enzyme) E2 và enzyme gắn với ubiquitin (ubiquitin-ligating enzyme) E3 (E3 ligase). Enzyme E1 được xác định, tách chiết và tinh sạch lần đầu tiên từ dịch chiết reticulocytes ở dạng homodimer với khối lượng phân tử 210 kDa (Ciechanover *et al.*, 1981). Lần lượt các enzyme E2, E3 được tách chiết, tinh sạch và xác định khối lượng phân tử tương ứng là 35 kDa và 300 kDa (Hershko *et al.*, 1983). Quá trình có thể mô tả tổng quát theo các bước được tiến hành theo hình 3 (Hershko, Ciechanover, 1998, Pickart, Eddins, 2004, Moon *et al.*, 2004).

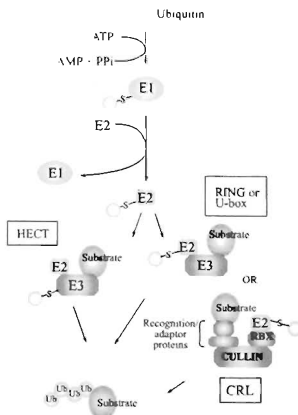
Gắn E1 Đầu tiên, nhóm COOH đầu C (Gly76) của ubiquitin gắn covalent với nhóm SH của gốc Cys trong enzyme hoạt hóa ubiquitin E1 sử dụng năng lượng ATP. Thường mỗi loại ubiquitin chỉ có duy nhất một enzyme E1 tương ứng.

Gắn E2 Tiếp theo, ubiquitin được enzyme E1 chuyển sang nhóm SH của gốc Cys trong enzyme tiếp hợp với ubiquitin - E2, đồng thời giải phóng ra enzyme E1. Số lượng E2 có thể nhận biết được mỗi E1 hơn gấp nhiều lần.

Gắn E3 trước hết, những protein cần bị phân giải, protein cơ chất, được gắn với enzyme E3-enzyme gắn với ubiquitin. Tiếp theo, enzyme E2 và E3 tương tác tạo phức E2-E3 để thúc đẩy ubiquitin gắn với protein cơ chất cần được phân giải. Số lượng E3 trong tế bào nhiều hơn E2 phụ thuộc vào protein cơ chất cần phân giải phức này tạo khả năng để nhóm COOH - ubiquitin chuyển sang gắn với nhóm NH₂ của gốc Lys của protein cơ chất. Việc chuyển ubiquitin được thực hiện liên tiếp nhiều lần lặp lại các bước trên để tạo thành chuỗi polyubiquitin: gốc COOH ở đầu C của ubiquitin tiếp theo gắn với gốc NH₂ của Lys48 của phân tử ubiquitin trước nó. Số lượng ubiquitin chiếm 4 phân tử. Phức hệ E2-E3-polyubiquitin-protein phân ly, tách E2, E3 ra khỏi chuỗi polyubiquitin- protein cơ chất. Chuỗi polyubiquitin-protein cơ chất thúc đẩy việc chuyển protein này từ ubiquitin sang proteasome để phân giải thành aa



Hình 2. Cấu trúc phân tử Ubiquitin. **A** Cấu trúc không gian với các vị trí đặc hiệu. Lys29, Lys48, Lys63 và Gly76. **B.** Mô hình nbbon (Buchanan *et al.*, 2000)



Hình 3. Sự hình thành chuỗi polyubiquitin và vận chuyển protein cơ chất với sự tham gia của các enzyme: enzyme hoạt hóa ubiquitin (ubiquitin-activating enzyme) - E1; enzyme hiệp hợp với ubiquitin (ubiquitin-conjugating enzyme) - E2 và enzyme gắn với ubiquitin (ubiquitin-ligating enzyme) - E3 sử dụng năng lượng ATP (Dreher, Callis, 2007)

Sự đặc hiệu của enzyme E1 đối với E2 và đối với E3 được thể hiện trong cơ thể sinh vật bằng sự tương quan số lượng của chúng. Trong các cơ thể nhân thực có thể có một vài dạng enzyme E1, nhưng có nhiều E2 và rất nhiều E3. Ví dụ trong tế bào nấm men, mỗi loại ubiquitin có duy nhất một enzyme E1 đặc hiệu tương ứng, nhưng có đến 11 enzyme E2 có thể nhận biết được phức hệ E1-ubiquitin này và 20 loại enzyme E3 có thể tương tác với các E2 kể trên. Trong tế bào động vật có vú, có một vài E1, vài chục E2 và vài trăm E3. Số lượng E3 nhiều như vậy là do sự đa dạng của cơ chất protein cần bị phân giải. E3 cần có cấu trúc phù hợp để nhận biết và tương tác với các dạng protein cơ chất này (Pickart, Eddins, 2004). Cho đến 2009, người ta xác định được ở đậu tương (*Glycine max*) có đến 1519 gen mã hóa cho quá trình phân giải protein theo hệ ubiquitin/26S proteasome, chiếm khoảng 5% tổng số gen của cây này, nó lớn hơn so với bất kỳ cơ thể động vật nào khác. Trong đó, E1 có 6 gen, E2 có 107 gen còn E3 có 1406 gen. Điều đó chứng tỏ nhiều quá trình hình thành, phát triển và trao đổi chất trong thực vật phụ thuộc vào sự ubiquitin hóa này. Trong số gen đó, có đến 92,5% mã hóa cho các enzyme E3 (Moon *et al.*, 2004, Schwechheimer & Schwager, 2004, Du *et al.*, 2009).

Enzym E3 ligase

Ở thực vật có nhiều nhóm E3 ligase: HECT (Homologous to E6AP carboxy terminus), RING (Really interesting new gene), F-box, SKP (S-phase kinase-associated protein), BTB (Broad-complex, Tramtrack, Bric-a-Brac), Cullin, U-box, RBX (RING box), DDB (DNA Damage Binding) (Moon *et al.*, 2004, Dreher *et al.*, 2007)

HECT (Homologous to E6AP carboxy terminus) là một nhóm E3 nhỏ ở thực vật, ở *Arabidopsis* có 7 đại diện của nhóm này, trong khi đó ở người có trên 50 gen thuộc loại này. Đây là những protein có khối lượng phân tử từ 100 đến 400 kDa. Vùng hoạt tính HECT khoảng 350 aa có chứa điểm gắn với ubiquitin và điểm gắn với enzym E2.

RING (Really interesting new gene) là một nhóm protein lớn ở thực vật, chúng có vùng gần với ion Zn hoặc còn gọi là vùng đầu RING (RING finger) như một đặc điểm quan trọng của protein dạng này. Vùng này có khoảng 70 aa chứa ion kẽm để ổn định cấu trúc và là điểm gắn với enzyme E2. Ở *Arabidopsis* có đến trên 465 gen mã hóa cho protein dạng này. RING có thể tồn tại ở dạng đơn lẻ chỉ có một tiểu đơn vị hoặc dạng đa tiểu đơn vị. Dạng đơn lẻ có thể ví dụ như COP1 (Constitutive Photomorphogenesis 1), SINAT5 (SEVEN IN ABSENTIA IN ARABIDOPSIS THALIANA 5) và ARCI (Arim repeat containing 1). Dạng đa tiểu đơn vị gồm có SCF, CUL3-BTB, CRL (Cullin RING ligase) và APC (anaphase -promoting complex). Trong các dạng đa tiểu đơn vị, nhóm SCF được nghiên cứu sâu nhất. Tên của nhóm protein này lấy từ tên của 3 trong số 4 tiểu đơn vị cấu thành: SKP1, Cullin protein và F-box. Tiểu đơn vị thứ 4 chính là đầu RING protein.

Ngoài RING finger và HECT, các nhóm như F-box, SKP (S-phase kinase-associated protein), BTB (Broad-complex, Tramtrack, Bric-a-Brac), Cullin, U-box, RBX (RING box), DDB (DNA Damage Binding) có thể kết hợp cùng hoạt động hỗ trợ trong một số trường hợp, tạo dạng đa tiểu đơn vị.

Bảng 1. Số liệu thống kê về UPS ở thực vật (Du *et al.*, 2009)

Tên loài	E1	E2	E3				Tổng số
			F-box	RING	HECT	Nhóm khác	
<i>Arabidopsis</i>	2	47	654	465	7	179	1305
Lúa	6	49	702	378	8	236	1332
Cây dương	6	70	335	399	7	210	951
Đậu tương	6	107	418	725	10	253	1406
Nho	3	45	153	330	9	137	629
Cỏ Medicago	2	40	539	294	9	127	969
Ngô	9	59	325	401	21	285	1032

Cơ sở dữ liệu database về hệ thống UPS đã hình thành trên kết quả nghiên cứu các Ngân hàng genome thực vật chính. Đây là 7 đối tượng thực vật: *Arabidopsis*, Lúa, cây dương (Poplar), Đậu tương, Nho, cỏ Medicago, Ngô (Du et al., 2009). Cơ sở dữ liệu đối với *Arabidopsis* là TIDR; đối với lúa – Rice Genome Annotation Resource, đối với đậu tương – Soybean Genome Project; đối với ngô – Maize Genome Sequence project, đối với cỏ Medicago – Medicago Genome Sequence Consortium, đối với cây dương: Populus genome project, đối với Nho: Genoscope. Trên 8000 gen đã được nghiên cứu phân bố thành 11 nhóm UPS, trong đó có 1 nhóm E1, 1 nhóm E2 và 9 nhóm E3. Các dữ liệu này có thể truy cập trên địa chỉ <http://bioinformatics.cau.edu.cn/plantUPS/>. Một số dữ liệu thông kê về các enzyme E1, E2 và E3 với một số nhóm diễn hình được trình bày trên bảng 1. Trong số đó, genome của các loài Lúa, Ngô, Đậu tương và *Arabidopsis* được nghiên cứu, đọc trình tự tương đối hoàn chỉnh nên số liệu tương đối đồng đều: số lượng gen E3 có trên 1000 gen; từ 1023 gen của ngô đến 1406 gen ở đậu tương. Các loài khác số liệu này hạn chế hơn.

Proteasome

Một tế bào của người chứa khoảng 30.000 proteasomes. Nghiên cứu phát hiện ra bề mặt phân hủy protein trong tế bào được tiến hành từ những năm 80 của thế kỷ trước. Proteasome có thể phá vỡ tất cả các protein dài 7-9 aa. Đây là phức hệ protein với khối lượng phân tử trên 1.5 MDa, tồn tại ở hai dạng: 20S và 26S. Proteasome 20S cấu tạo từ 4 vòng liên kết các tiểu đơn vị có dạng hình trụ. Mỗi vòng cấu tạo từ 7 chuỗi polypeptide. Tâm hoạt tính protease nằm bên trong rãnh của các vòng. Phức hệ protein điều hòa 19S gắn với hai đầu của proteasome 20S tạo thành proteasome 26S sử dụng năng lượng ATP. Tiếp theo, chuỗi polyubiquitin chuyển protein đích đến proteasome và đưa vào bên trong lõi của phức hệ 26S. Thế hình khối 26S có hai nắp đậy ngăn các protein khác vào proteasome và không phân giải các protein khác của tế bào. Protein 19S ở thực vật cấu tạo từ 9 tiểu đơn vị nhỏ. Chức năng của chúng là nhận biết cơ chất protein do chuỗi ubiquitin chuyển đến và giải phóng chuỗi ubiquitin sau đó. Trong quá trình phân giải này có sự tham gia của nhiều protein – sản phẩm biểu hiện của trên 100 gen.

Phản cắt ubiquitin

Chuỗi polyubiquitin không bị phân giải cùng với protein cơ chất. Chúng bị phân cắt bằng một loại

enzyme đặc hiệu: deubiquitylating enzyme (DUBs). Có nhiều loại DUBs với các vùng đặc hiệu có cấu trúc khác nhau để có thể kiểm soát hoặc điều hòa các quá trình tạo mới, chuyển hóa, hoặc chỉnh sửa ubiquitin khác nhau trong tế bào. Tất cả các DUB đều là cysteine protease, chúng phân giải mỗi liên kết gốc amin ngay sau gốc COOH của Gly76. Ở nấm men có ít nhất 17 loại DUB. Các loài nhân thực khác có thể nhiều hơn. Phụ thuộc vào kích thước phân tử, sự đồng nhất trình tự amino acid trong phân tử và tâm hoạt tính, DUB có thể chia thành 2 nhóm: UCH (Ubiquitin COOH-terminal hydrolyscs) và UBP (ubiquitin-specific protease). UCH thường là các enzyme nhỏ khoảng 20-30 kDa. Chúng loại bỏ đoạn peptide nhỏ cuối đuôi COOH của phân tử ubiquitin. Chúng thường có tâm hoạt tính là Cys, His, và Asp. Trong một số trường hợp có thêm Glu. UBP là các enzyme lớn với khối lượng phân tử khoảng 100 kDa. UBP có chức năng cắt mỗi liên kết Ubi-Ubi hoặc Ubi-Protein, cũng như chia cắt chuỗi polyubiquitin mới tổng hợp thành ubiquitin dạng đơn. Tâm hoạt tính của chúng thường có Cys, His và Aspartate (Glickman, Ciechanover, 2002, Schwechheimer, Schwager, 2004).

Chức năng của hệ ubiquitin-proteasome trong tế bào thực vật

Ở động vật và người, người ta đã giá định nhiều chức năng của UPS này. UPS loại bỏ các protein bị biến tính hoặc tạo cấu trúc không đúng trong cả nguyên sinh chất và nhân tế bào. Và qua đó chúng tham gia vào quá trình phân bào, sửa chữa DNA, phản ứng bảo vệ tế bào và cơ thể, phản ứng miễn dịch trong tế bào, apoptosis – hiện tượng hoại tử tại chỗ được lập trình trước. Chỉ một lỗi nhỏ trong quá trình ubiquitin hoá có thể gây ra bệnh hiểm nghèo cho người. UPS loại bỏ protein ức chế hoặc coi trời cho TF nhằm kích hoạt một quá trình hoạt động cụ thể trong tế bào, ví dụ như phức hợp protein gọi là "anaphase – promoting complex" (APC); p53, NF- κ B (nuclear factor kappaB) (Glickman & Ciechanover, 2002).

Trong tế bào thực vật, tương tự như đối với tế bào động vật, người và các tế bào nhân thực khác, hệ thống UPS, với chức năng phân hủy protein biến tính, tham gia vào nhiều chu trình sống trong tế bào, phản ứng với ánh sáng và các điều kiện môi trường bất lợi. Đặc biệt UPS điều hòa các hormone sinh trưởng chịu trách nhiệm về các quá trình sinh trưởng và phát triển trong tế bào cũng như toàn bộ cá thể thực vật. Chúng hoạt động trong cả nguyên sinh chất và nhân tế bào:

- UPS tham gia vào quá trình phát triển tế bào, phân bào, phiên mã, dịch mã, sửa chữa DNA. Chúng còn tham gia quá trình phát triển hoa, hình thành cánh, đài hoa, nhị hoa với sự tham gia của nhiều gen. Ví dụ như *UFO* (UNUSUAL FLORAL ORGANS) mã hóa cho F-box protein điều hòa quá trình phiên mã gen *APETALA3*. Đây là TF của các gen tổng hợp cánh hoa và nhị hoa. SCF còn điều khiển sự không tương thích đối với bào tử, hạt phấn. Chỉ với điều kiện các yếu tố protein trong tế bào nhận biết được sự tương thích từ cấu trúc protein đặc thù của hạt phấn, hạt phấn mới phát triển được (Sullivan *et al.*, 2003; Moon *et al.*, 2004).

- UPS tham gia vào quá trình điều hòa các hormone sinh trưởng thực vật. Các enzyme E3 ligases đặc biệt tích cực tham gia trong việc nhận truyền tín hiệu hormone, phân hủy yếu tố phiên mã đặc hiệu cho hormone và điều hòa quá trình sinh tổng hợp hormone. Ngày nay, 10 loại hormone thực vật được ghi nhận là auxin, cytokinin, gibberillin, abscisic acid (ABA), ethylene, jasmonic acid (JA), brassinosteroid, salicylic acid, nitric oxide và strigolactone. Các đại diện của nhóm SCF điều khiển sự tổng hợp các chất hormone sinh trưởng trong thực vật thông qua việc phân giải các protein kìm hãm hoặc hoạt hóa phiên mã. Một số ví dụ được mô tả sau đây. Nhóm SCF^{TR1-AL1} điều hòa sự nhận biết tín hiệu auxin, trong đó, có các protein tín hiệu auxin (auxin-signaling F-box protein -AFP). Đối với JA, SCF^{COI} tham gia vào đường dẫn truyền tín hiệu JA. COI1 (coronatine-insensitive1) là SCF chứa ASK1, CUL1 và RBX1. SCF^{COI} chịu trách nhiệm giải phóng protein điều hòa phiên mã JA là JAZ (jasmonate ZIM-domain protein). Đối với ethylene, các F-box ETP1/2 (EIN2 TARGETING PROTEIN ½) thúc đẩy quá trình đào thải yếu tố EIN2 (ETHYLEN INSENSITIVE 2) khi thiếu ethylene. Ngoài ra, các F-box EBF1 (EIN3 Binding F-box 1) điều hòa yếu tố EIN3. Đối với ABA, hai loại RING E3 thường thấy là AIP2 (ABI3-Interacting protein) và KEG (Keep on Going) điều hòa hoạt động của hai TF - ABI3 (ABA Insensitive 3) và ABI5 (Moon *et al.*, 2004; Santner, Estelle, 2010; Sullivan *et al.*, 2003; Liu & Stone, 2011).

Trong trường hợp điều khiển sự sinh tổng hợp gibberellin bằng protein DELLA có sự tham gia của SCF DELLA là họ protein có chức năng làm TF có tận cùng đầu N là chuỗi aa DELLA, vùng tín hiệu

định vị nhân và vùng hoạt tính. Trong *Arabidopsis* có 5 protein thuộc dạng này. DELLA điều khiển bằng cách ức chế sự tổng hợp gibberillin. Khi không có gibberillin trong tế bào, DELLA ở dạng tự do và ức chế sự sinh tổng hợp chất này. Khi có chất hoạt hóa gibberellin, chất này tác động với GIP1 (GIBBERELLIN INSENSITIVE DWARF1), thúc đẩy DELLA đến với SCF-E3 và bị phân giải trong hệ UPS. DELLA bị phân giải, quá trình tổng hợp gibberellin không còn bị ức chế và tạo điều kiện tổng hợp nhiều gibberellin trong tế bào. Ở lúa, mới phát hiện một gen mã hóa cho protein dạng DELLA là SLR1 (Slender rice 1) (Zeng *et al.*, 2002; Wang, Deng, 2011).

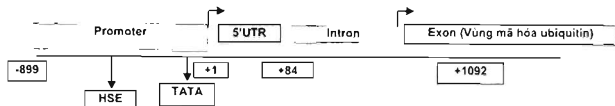
Khả năng sử dụng ubiquitin promoter trong công nghệ gen thực vật.

Sự đa dạng về chức năng của ubiquitin, khả năng biểu hiện phản ứng với các điều kiện cực đoan nhất là nhiệt độ, đã thu hút sự chú ý của nhiều nhà nghiên cứu trong lĩnh vực chuyển gen thực vật. Từ nhiều năm nay, ngoài promoter *CAMV 35S*, việc tìm kiếm các promoter mới có khả năng biểu hiện cao trong các loại mô thực vật khác nhau luôn được xúc tiến mạnh mẽ.

Cây ngô: Lần đầu tiên, Christensen và cộng sự tiến hành nghiên cứu phân lập để ứng dụng ubiquitin promoter vào khoảng những năm 90 thế kỷ trước. Trong số 8-10 gen được dự đoán từ kết quả lai Southern, hai gen *polyubiquitin* được phân lập từ Ngân hàng genome ngô. Charon Đồng thời, các tác giả phân lập gen này từ cDNA của ngô. Cả hai gen *Ubi1*, *Ubi2* đều chứa promoter và intron trước vùng mã hóa protein. Với mỗi gen, vùng mã hóa cho polyubiquitin với 7 lần lặp lại và có amino acid cuối là CAG (Gln). Hai gen khác nhau bởi vùng promoter, trong đó promoter của gen *Ubi1* được nghiên cứu kỹ và sử dụng để thiết kế vector chuyển gen (Hình 4). Vùng này khoảng 2 kb, chứa promoter 0,9 kb, vùng 5' không dịch mã (5' Untranslated Region - 5' UTR) (exon1) 82 bp và intron gần 1,1 kb. Trong đoạn promoter có *cis*-element (TAATAATA) và yếu tố sốc nhiệt (heat-shock element HSE) (GGA TTC). So sánh khả năng biểu hiện gen chỉ thị *cat* (chloramphenicol acetyltransferase) của hai cấu trúc pUbi-CAT và p35S-CAT cho thấy khả năng biểu hiện của promoter *ubiquitin* cao gấp 10 lần so với promoter *CAMV35S* ở ngô (Christensen *et al.*, 1992).

Bảng 2. Một số nhóm E3 và chức năng của chúng (Dreher, Callis, 2007)

E3	Loài	Đường dẫn truyền	Cơ chất dự đoán
SCF F-box			
AFB1, AFB2, AFB3, AFB5	<i>A. thaliana</i>	Tín hiệu auxin	Auxin/IAA với vùng II
ACRE189	<i>N. tabacum</i>	Bảo vệ thực vật	
CEGENDUO	<i>A. thaliana</i>	Hình thành rễ phụ	
COI1	<i>A. thaliana</i>	Tín hiệu JA	HDAC6, RUBISCO
EBF1, EBF2	<i>A. thaliana</i>	Tín hiệu ethylene	EIN3, EIL1
SLY1/GID2	<i>A. thaliana/O. sativa</i>	Tín hiệu GA	DELLA protein RGA, GA1, RGL1, RGL2, RGL3 (Ar), SLR1 (Os), SLN1 (Hv)
SNE	<i>A. thaliana</i>	Tín hiệu GA	DELLA protein: RGA, GA1, RGL1, RGL2, RGL3
SNE	<i>A. thaliana</i>	Tín hiệu GA	
SON1	<i>A. thaliana</i>	Bảo vệ thực vật	
TIR1	<i>A. thaliana</i>	Tín hiệu auxin	Auxin/IAA với vùng II
TLP9	<i>A. thaliana</i>	Tín hiệu ABA	
VirF	<i>A. tumefaciens</i>	Nhiễm vi khuẩn	VirE2, VIP1
BTB			
ARIA	<i>A. thaliana</i>	Tín hiệu ABA	ABF2
ETO1, EOL1, EOL2	<i>A. thaliana</i>	Tổng hợp ethylene	ACC synthase 5 (ACS5)
GMPOZ	<i>H. vulgare</i>	Tín hiệu GA và ABA	
NPR1/NIM1	<i>A. thaliana</i>	Bảo vệ thực vật	
U-box			
ACRE276/PUB17	<i>N. tabacum/A. thaliana</i>	Bảo vệ thực vật	
CHIP	<i>A. thaliana</i>	Tín hiệu ABA	Protein phosphatase 2A subunit
CMPG1/ELI17 CMPG1/ACRE74 Cmpg1 PUB20/PUB21	<i>P. crispum, N. tabacum, S. lycopersicum, A. thaliana</i>	Bảo vệ thực vật	
PUB5, PUB12	<i>A. thaliana</i>	Bảo vệ thực vật	
PUB27, PUB28, PUB29	<i>A. thaliana</i>	Tín hiệu GA?	
SPL11	<i>O. sativa</i>	Bảo vệ thực vật	
RING			
ACRE132	<i>N. tabacum</i>		
AIP2, ATL43	<i>A. thaliana</i>	Tín hiệu ABA	
ATL2, ATL6	<i>A. thaliana</i>	Bảo vệ thực vật	
BRH1	<i>A. thaliana</i>	Tín hiệu BR	
EL5	<i>O. sativa</i>	Bảo vệ thực vật	
RHA1b, RHA3b, RMA	<i>A. thaliana</i>	Bảo vệ thực vật	
RIN2/RIN3	<i>A. thaliana</i>	Bảo vệ thực vật	
SINAT5, XBAT32	<i>A. thaliana</i>	Tín hiệu auxin	
XERICO	<i>A. thaliana</i>	Tổng hợp ABA	
HECT: UPL3	<i>A. thaliana</i>	Tín hiệu GA	
APC HOBBIT	<i>A. thaliana</i>	Tín hiệu auxin	



Hình 4. Cấu trúc hệ thống điều khiển gen ubiquitin ở thực vật (mô hình gen ubiquitin của ngô) HSE (heat-shock element) – yếu tố sốc nhiệt

Tiếp sau đó, các nhà nghiên cứu Nhật Bản sử dụng *ZmUbi1* promoter đã biểu hiện thành công gen *bar* (kháng thuốc diệt cỏ) trong cấu trúc pUBA khi biến nạp cấu trúc này vào lúa gạo (Toki *et al.*, 1992). Nghiên cứu này đã mở ra cơ hội sử dụng promoter mới trong công nghệ gen thực vật

Từ đó trở đi, trên thực vật đã phát hiện ra một số lượng đáng kể nhóm promoter ubiquitin có khả năng biểu hiện mạnh ở thực vật. Ngoài ngô, polyubiquitin promoter được phân lập và xác định đặc tính từ *Arabidopsis* (Callis *et al.*, 1990), cây hương dương (Binet *et al.*, 1990), khoai tây (Garbarino *et al.*, 1995), cà chua (Rollfinke *et al.*, 1998), cây thuốc lá (Plesse *et al.*, 2001), lúa (Wang *et al.*, 2003, Lu *et al.*, 2008), mía đường (Wei *et al.*, 2003) và đậu tương (Chiera *et al.*, 2007).

Khoai tây: Trên cây củ, đại diện là khoai tây, các nghiên cứu về promoter ubiquitin cũng được xúc tiến mạnh. Garbarino và cộng sự đã phân lập ubiquitin promoter của gen *ubi7* từ thư viện genome khoai tây *FIXII* (*Solanum tuberosum*). Trước đây, cDNA của gen này đã được phân lập trong điều kiện cảm ứng với nhiệt độ cao, JA và ABA. Đoạn này chứa promoter và 5'UTR dài 1156 bp và intron 569 bp. Trong đoạn promoter có cis-element (TATAAAA), G-box và vùng giàu GT. Đoạn promoter của gen *ubi7* đã được thiết kế vào cấu trúc gắn với gen chỉ thị *GUS*-NOS ter. Các dòng khoai tây chuyển gen cho thấy khả năng biểu hiện của gen *GUS* cao gấp khoảng 10 lần so với cấu trúc không có intron (Garbarino *et al.*, 1995).

Gần đây, Rockhold và cộng sự đã phân lập được hai gen ubiquitin *bul409* và *bul427* từ loài khoai tây đại *Solanum bulbocastanum*. Gen *bul409* và *bul427* có số đoạn ubiquitin đơn lặp lại tương ứng là 6 và 7 đoạn. Gen *bul409* có độ tương đồng của vùng 5'UTR, vùng mã hóa polyubiquitin và 3'UTR trên 95% so với gen *ubi7* của khoai tây. Riêng vùng promoter có độ tương đồng thấp hơn khoảng 80%. Còn gen *bul409* có độ tương đồng của vùng mã hóa

polyubiquitin đến 96% so với gen *LEPUB* của cà chua. Riêng vùng promoter có độ tương đồng thấp hơn 80%. Trong cấu trúc cùng với gen *GUS* để chuyển vào khoai tây, vùng điều khiển của gen *bul409* chứa 771 bp (trong đó intron chiếm 536 bp), còn vùng điều khiển của gen *bul427* bao gồm promoter và 5'UTR dài 1278 bp và intron 1555 bp. So với cấu trúc sử dụng promoter 35S, hoạt tính gen *gus* của các dòng sử dụng ubiquitin promoter tăng gấp nhiều lần (Rockhold *et al.*, 2008).

Đậu tương: Nghiên cứu phân lập và biểu hiện promoter ubiquitin *Gmubi* từ gen *SUB13* của đậu tương đã được tiến hành. Chiera và cộng sự đã sử dụng promoter này để biểu hiện gen *gfp* (green fluorescent protein) trên mô lá mầm hạt đậu (trắng và cho thấy cấu trúc promoter *Gmubi3* (923 bp), trong đó có intron (595 bp) có khả năng biểu hiện cao gấp 5 lần so với promoter *CAMV 35S*, cao gấp 2 lần so với promoter này không chứa intron (Chiera *et al.*, 2007). Nhóm tác giả này lại khẳng định khả năng biểu hiện gen *gfp* của promoter *Gmubi3* trên mô phôi mầm của đậu tương. Trong cây biến nạp gen, GFP được biểu hiện trong nhiều loại mô tế bào khác nhau: mô biểu bì, lá, hoa, chóp rễ, hạt phôi và hạt đang trưởng thành (Hernandez-Garcia *et al.*, 2009). Và mở rộng hơn, sử dụng các dữ liệu trên Ngân hàng Gen, nhóm tác giả này đã nghiên cứu phân lập 10 gen *Gmubi* và 10 gen *GmERF* (*G* max Ethylene Response Factor) và so sánh khả năng biểu hiện gen *gfp* trên lá mầm, rễ cây đậu trắng. Kết quả cho thấy 7 trong số 10 promoter *Gmubi* có khả năng làm tăng hoạt tính GFP từ 2 đến 7 lần so với promoter *CAMV 35S* trên lá mầm. Trong đó, 4 trong số 10 promoter *GmERF* chỉ làm tăng từ 1.5 đến 2.2 lần so với promoter *CAMV 35S* này. Còn ở rễ 6 trong số 10 promoter *Gmubi* có khả năng làm tăng hoạt tính GFP từ 2 đến 4 lần so với promoter *CAMV 35S*. Trong đó, các promoter *GmERF* chỉ làm tăng từ 1.4 đến 1.7 lần so với promoter *CAMV 35S* này. Promoter *Gmubi3* cho kết quả cao nhất. Trong số 10 gen *Gmubi* phần lớn đều có intron trong vùng điều

khien, tuy nhiên sự khác biệt rõ rệt giữa cấu trúc có intron và khuyết intron chưa phát hiện thấy (Hernandez-Garcia *et al.*, 2010).

Cây lúa Trong các cây lương thực, cây lúa được quan tâm đặc biệt. Wang và cộng sự đã phân lập 4 gen *ubiquitin* từ cây lúa, trong đó có 2 gen *RUBQ1* và *RUBIQ2* là *polyubiquitin*. Trong đoạn 3184 bp phân lập được của gen *RUBQ1*, đoạn promoter và 5'-UTR dài 920 bp, còn intron 782 bp. Còn gen *RUBQ2* bao gồm 4442 bp, trong đó promoter và 5'-UTR dài 1824, còn intron 926 bp. Kết quả nghiên cứu biểu hiện gen *gus* cho thấy promoter của *RUBIQ2* có khả năng biểu hiện cao hơn và hơn 10-20 so với 35S CAMV. Promoter này ưu việt hơn so với *ZnUbi1* của ngô biểu hiện trên cây lúa (Wang *et al.*, 2003). Nhóm nghiên cứu của Sivamani đã phân lập gen *rub13* promoter: 808 bp, vùng 5' không dịch mã (exon 1): 67bp, intron: 1140 và exon2 có 1140 bp mã hóa cho polyubiquitin với 5 đoạn lặp lại. Khả năng biểu hiện gen *gus* của promoter này có intron cao hơn khoảng 20 lần so với promoter không chứa intron (Sivamani, Qu 2006). Gần đây nhất, Bhattacharyya và cộng sự đã phân lập *OsUbi1* promoter từ cây lúa. Đoạn gen bao gồm promoter 678 bp, vùng 5' UTR (exon 1): 74 bp, intron: 787 bp và một đoạn exon 2237 bp được gắn vào vector với gen *gus* để nghiên cứu khả năng biểu hiện. So với *ZnUbi1*, promoter này có khả năng biểu hiện tốt hơn trên mô cây lúa (Bhattacharyya *et al.*, 2012).

Các cây trồng khác: Ngoài các cây lương thực, cây trồng có giá trị kinh tế cao, việc tìm kiếm các ubiquitin promoter từ các loài khác nhau vẫn đang được tiếp tục. Đặc biệt với sự bùng nổ của năng lượng sinh học, một số loại cây một lá mầm như cây có đang được xúc tiến mạnh mẽ. Hai gen *PvUbi1* và *PvUbi2* được phân lập từ một loài cỏ ở Bắc Mỹ (*Panicum virgatum* L.). *PvUbi1* có chứa 607 bp promoter, 5' không dịch mã (UTR) 93 bp (non-coding exon) và 1291 bp intron, đoạn mã hóa polyubiquitin dài 918 bp với 4 đoạn lặp lại nối tiếp nhau và 191 bp đầu 3' UTR. Còn gen *PvUbi2* chứa 692 bp promoter, 5' UTR dài 97 bp (non-coding exon) và 1072 bp intron, đoạn mã hóa polyubiquitin dài 1146 bp với 5 đoạn lặp lại và cuối cùng là 183 bp 3' UTR. Tuy nhiên không phát hiện thấy HSE trong vùng promoter của hai gen này (Mann *et al.*, 2011). Từ loại cây một lá mầm không thuộc nhóm cây lương thực *Gladiolus* đã phân lập gen được gen *GUBIQ1*. Gen này chứa vùng promoter dài 1898 bp, bao gồm promoter 600 bp, 5' không dịch mã 80 bp, và intron 1218 bp. Các thử nghiệm biểu hiện gen *gus*

và gfp sử dụng promoter này trên các đối tượng cây mô hình khác nhau *Arabidopsis*, thuốc lá, *Gladiolus* cho thấy chúng biểu hiện trên tất cả các đối tượng trên và khả năng biểu hiện cao hơn nhiều lần so với promoter không chứa intron (Kamo *et al.*, 2012).

Bèo tấm là cây một lá mầm với khả năng sinh trưởng và tổng hợp protein cao. Với tốc độ sinh sản rất nhanh, bèo tấm có thể nhân đôi trọng lượng trong vòng 24 – 72 giờ. Một số giống bèo tấm có tốc độ tăng sinh khối cao hơn hẳn so với các loài thực vật khác. Mặt khác hàm lượng protein của bèo tấm có thể đạt từ 6.8 – 45 % trọng lượng khô tùy theo loài, tức là tỷ lệ tương đương với đậu tương, và có triển vọng sử dụng để sản xuất protein tái tổ hợp. Nghiên cứu phân lập promoter *SpUbi* bèo tấm *Spirodela polyrrhiza* và *Lemna aequinoctialis* đã được một tác giả tiến hành. Trên promoter cả hai loại này đều có các trình tự tương ứng với HSE và cis-element (TATAAATA), G-box. Trên bèo tấm, promoter *LaUbi* này có khả năng biểu hiện cao hơn promoter *CAMV 35S* (Dickey *et al.*, 2007).

Đa số các promoter này điều khiển ở mức độ cao hơn so với promoter *CAMV35S* trong quá trình biểu hiện ở các loài thực vật được chuyển gen. Do gen *ubiquitin* được biểu hiện ở hầu hết các mô thực vật trong điều kiện tự nhiên, nên promoter *ubiquitin* đã được sử dụng để biểu hiện gen trong các loại thực vật chuyển gen đặc biệt là giữa các loài gần gũi, kể cả cây một lá mầm và hai lá mầm. Nhiều nghiên cứu trên các đối tượng cây trồng khác nhau như *Arabidopsis*, lúa, ngô... cho thấy: Sự biểu hiện mạnh của promoter *ubiquitin* được chi phối bởi sự xuất hiện của các đoạn intron, những đoạn chủ yếu định vị trong vùng 5' UTR của gen *ubiquitin*. Các đoạn intron gần với bộ 3 khởi đầu dịch mã là những nhân tố *cis* quan trọng, là nguyên nhân gây ra sự tăng cường gián tiếp có liên quan đến intron (intron-mediated enhancement – IME) của quá trình biểu hiện gen ở thực vật. Đối với cây hai lá mầm, intron có thể làm tăng khả năng biểu hiện lên từ 2 đến 20 lần, nhưng đối với cây một lá mầm có thể trên 20 lần. Trong một số thử nghiệm, người ta còn sử dụng một đoạn gen mã hóa cho ubiquitin từ 27-237 bp ngay gần kề với intron thiết kế cùng trong vector chuyển gen và cho thấy khả năng biểu hiện gen cao hơn so với cấu trúc chỉ có promoter và intron. Về phản ứng với nhiệt độ, trong promoter *ZnUbi* có HSE và có phản ứng biểu hiện gen của chúng ở nhiệt độ cao là đáng ghi nhận. Trong phần lớn các promoter *ubiquitin* đều có yếu tố này, tuy nhiên một số promoter *ubiquitin* không có hiện diện của HSE, vì

du như *PvUbi* (Callis *et al.*, 1990, Chiera *et al.*, 2007, Hernandez-Garcia *et al.*, 2012, Kamo *et al.*, 2012, Maas *et al.* 1991, Mann *et al.* , 2011, Lu *et al.*, 2008, Rose, 2004, Sivamani *et al.* , 2006).

Các loại promoter *ubiquitin* của ngô, lúa được thử nghiệm sử dụng để biểu hiện các gen có giá trị kinh tế trong các loại cây trồng khác nhau: như gen *CryIac*, gen zeatin O-glucosylation (*-ZOG1*), gen *ch11*... (Rodo *et al.*, 2008; Sripriya *et al.*, 2008, Wu *et al.*, 2008, Loza-Rubio *et al.*, 2008). Promoter *ZmUbi1* còn được sử dụng rộng rãi để nghiên cứu chức năng các protein liên quan đến tính chịu hạn ở một số cây lương thực một lá mầm như đối với gen *TPSP* liên quan đến trao đổi chất trehalose trên cây lúa; gen *DREB1A/CBF3* mã hóa TF liên quan đến tính chịu hạn ở cây lúa, gen *HVA1* mã hóa cho LEA protein của đại mạch biểu hiện trên cây lúa mì... (Umezawa *et al.*, 2006). Tuy nhiên, các promoter *ubiquitin* từ các giống cây trồng khác chủ yếu vẫn đang còn trong giai đoạn nghiên cứu thử nghiệm biểu hiện gen *gus* hoặc *gfp*. Khả năng biểu hiện vượt trội của các loại promoter *ubiquitin* là tiền đề để sử dụng chúng ngày càng nhiều trong kỹ thuật biến nạp gen ở thực vật bậc cao

Ở Việt Nam, nghiên cứu tạo cây chuyển gen mang những tính trạng mong muốn được tiến hành trong một số phòng thí nghiệm tiên tiến trong nước. Promoter chủ yếu được sử dụng là promoter *CAMV 35S*. Các loại promoter khác như *ubiquitin* của ngô (*ZmUbi1*), actin của lúa đang được dần dần đi vào thử nghiệm thiết kế vào các vector và chuyển vào cây một lá mầm như ngô, lúa (Nông Văn Hai, 2010). Trên cây lúa, gen *cryIA(c)* dưới sự điều khiển của promoter *ZmUbi1* đã được chuyển vào giống C71 thông qua vi khuẩn *A. tumefaciens*. Hiện nay đã nhận được dòng lúa biến đổi gen C67T311-1 biểu hiện tính kháng sâu đục thân (*Scorophaga incertulas*) cao (Đỗ Xuân Đồng *et al.*, 2009). Bước đầu, cấu trúc có promoter *ZmUbi1* đã được sử dụng để chuyển được gen *cryIA(c)* vào cây ngô (Nguyễn Văn Đồng *et al.*, 2010). Ngoài ra, việc phát triển tìm kiếm các promoter *ubiquitin* trên các cây trồng khác cũng được tiến hành. Trên bèo tấm, trình tự đoạn điều khiển gen *SpUbi* của *Spirodela polyrrhiza* DB2 có đoạn promoter và 5'UTR có chiều dài là 1033 bp, intron 980 bp, còn đối với gen *LaUbi* của *Lemma acuinotialis* DB1 có đoạn promoter và 5'UTR có chiều dài là 964 bp, intron 1104 bp (Trần Thị Phương Liên *et al.*, 2009). Các cấu trúc vector và các promoter phân lập được trên cây trồng ở nước ta vẫn đang được tiếp tục nghiên cứu khảo sát

Tóm lại, với sự phát triển mạnh mẽ của sinh học hiện đại, ubiquitin cũng như quá trình phân giải protein của hệ UPS ngày càng được làm sáng tỏ. Trong 3 nhóm enzyme E1, E2 và E3 tham gia vào quá trình này, số lượng E3 nhiều và đa dạng nhất. Nhiều enzyme E3 đã được phân lập và xác định tính chất, chức năng trong nhiều quá trình sống của tế bào. Vai trò điều hòa nhiều quá trình sinh trưởng phát triển, phản ứng với các yếu tố cực đoan trong tế bào của UPS được phát hiện một cách toàn diện và đánh giá tổng thể.

Sự đa dạng về chức năng cũng như sự tồn tại của ubiquitin trong tế bào thực vật đã mở ra tiềm năng sử dụng promoter của chúng trong chuyển gen thực vật. Trên cơ sở nghiên cứu sâu về gen *ubiquitin*, việc phân lập ngày càng nhiều promoter của gen *polyubiquitin* trên các đối tượng cây trồng khác nhau, từ cây lương thực thực phẩm đến cây có giá trị phát triển năng lượng sinh học càng được xác tín mạnh. Các promoter *ubiquitin* có khả năng biểu hiện gen cao so với các loại promoter thông dụng khác, đặc biệt đối với cây một lá mầm. Đây là hướng nghiên cứu ứng dụng đầy triển vọng trong tương lai. *Lời cảm ơn:* Công trình được tổng quan trên một số hướng nghiên cứu về promoter của các đề tài thuộc Chương trình Công nghệ sinh học phục vụ Nông nghiệp 2006-2011

TÀI LIỆU THAM KHẢO

- Đỗ Xuân Đồng, Chu Hoàng Hà, Lê Trần Bình (2009) Dòng lúa biến đổi gen C67T311-1 mang gen *cryIac* biểu hiện tính kháng sâu đục thân (*Scorophaga incertulas*) cao *Tạp Chí Nông Nghiệp và Phát Triển Nông Thôn* 1/2009 9-12
- Nguyễn Văn Đồng, Phạm Thị Lý Thu, Trần Minh Thu, Lê Thanh Nga, Lê Thị Thu Vê, Lê Thị Thu Hiền, Nông Văn Hai, Lê Huy Hàm (2010) Nghiên cứu khả năng tiếp nhận gen kháng sâu (*cryIac*) thông qua vi khuẩn *Agrobacterium tumefaciens* của một số dòng ngô Việt Nam *Tạp chí Khoa học và Công nghệ nông nghiệp Việt Nam*, 6(19) 36-41
- Nông Văn Hai (2010) Phân lập các gen có giá trị kinh tế của cây trồng nông lâm nghiệp Việt nam, thiết kế vector, tạo chủng *Agrobacterium* phục vụ cho tạo giống cây trồng chuyển gen. Đề tài thuộc Chương trình Công nghệ Sinh học Nông nghiệp (2006-2010).
- Trần Thị Phương Liên, Hà Hồng Hạnh, Lê Huy Hàm (2009). Nghiên cứu phân lập các yếu tố điều khiển biểu hiện gen đặc hiệu từ bèo tấm *Spirodela polyrrhiza* DB2 *Tạp chí Công nghệ Sinh học* 7(1), 53-58
- Bhattacharyya J, Chowdhury AH, Ray S, Jha JK, Das S, Gayen S, Chakraborty A, Mitra J, Maiti MK, Basu A, Sen

- SK (2012) Native polyubiquitin promoter of rice provides increased constitutive expression in stable transgenic rice plants. *Plant Cell Rep* 31:271-279
- Binet MN, Lepetit M, Weij JH, Tessier LH (1990) Analysis of a sunflower polyubiquitin promoter by transient expression. *Plant Sci* 79: 87-94
- Bond U, Schlesinger MJ (1985) Ubiquitin is a heat shock protein in chicken embryo fibroblasts. *Mol Cell Biol* 5: 949-956.
- Callis J, Raasch J A, Vierstra R D (1990) Ubiquitin extension proteins of *Arabidopsis thaliana* – structure, localization, and expression of their promoters in transgenic tobacco. *J Biol Chem* 265: 12486-12493.
- Catic A, Ploegh HL (2005) Ubiquitin – conserved protein or selfish gene. *Trends in Biochem Sci* 30 (11): 600-604.
- Chiera JM, Bouchard RA, Dorsey SL, Park EH, Buenrostro – Nava MT, Ling PP, Finer JJ (2007). Isolation of two highly active soybean promoters and their characterization using a new automated image collection and analysis system. *Plant Cell Rep* 26: 1501-1509
- Christensen AH, Sharrack R, 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
- Ciechanover A, Heller H, Katz-Elzoun R, Hershko A (1981) Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system. *Proc Nat Acad Sci USA* 78:761-765
- Ciechanover A, Hod Y, Hershko A (1978) A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochem Biophys Res Commun* 81: 1100-1105.
- Dickey L, Cox K, Peele C (2007) Expression control elements from the Lemnaceae family US Patent Appl 20070180583
- Dreher K, Callis J (2007) Ubiquitin, hormones and biotic stress in plants. *Annals of Botany* 99: 787-822
- Du Z, Zhou X, Li L, Su Z. (2009) PlantsUPS: a database of plants ubiquitin proteasome system. *BMC Genomics* 10: 227
- Garbarino JE, Oosumi T, Belknap WR (1995) Isolation of a polyubiquitin promoter and its expression in transgenic potato plants. *Plant Physiol* 109: 1371-1378.
- Glickman MH, Ciechanover A (2002) The ubiquitin – proteasome proteolytic pathway destruction for the sake of construction. *Physiol Rev* 82: 373-428.
- Goldknopf IL, Busch H (1977) Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24. *Proc Natl Acad Sci USA* 74: 864-868.
- Goldstein G, Scheid M, Hammerling U., Boye EA, Schlesinger DH, Niall HD (1975) Isolation of a polypeptide that has lymphocyte-differentiating properties and its probably represented universally in living cells. *Proc Natl Acad Sci USA* 72: 11-15.
- Hernandez-Garcia CM, Bouchard AR, Rushton PJ, Jones ML, Chen X, Timko MP, Finer JJ (2010) High level transgenic expression of soybean (*Glycine max*) *GmERF* and *Gmubr* gene promoters isolated by a novel promoter analysis pipeline. *BMC Plant Biology* 2010: 10: 237.
- Hernandez-Garcia CM, Martinelli AP, Bouchard AR, Finer JJ (2009) A soybean (*Glycine max*) polyubiquitin promoter gives strong constitutive expression in transgenic soybean. *Plant Cell Rep* 28: 837-849.
- Hershko A, Ciechanover A (1998) The ubiquitin system. *Annu Rev Biochem* 67: 425-479.
- Hershko A, Heller H, Elias S, Ciechanover A (1983) Components of ubiquitin-protein ligase system. *J Biol Chem* 258: 8206-8214
- Kamo K, Kim AY, Park SH, Joung YH (2012) The 5'UTR-intron of the *Gladiolus* polyubiquitin promoter GUBO1 enhances translation efficiency in *Gladiolus* and *Arabidopsis*. *BMC Plant Biology* 2012: 12-79.
- Liu H, Stone SL (2011) E3 ubiquitin ligases and abscisic acid signaling. *Plant Signal Behav* 6(3): 344-348.
- Loza-Rubio E, Rojas E, Gómez L, Olivera MT, Gómez-Lim MA (2008) Development of an edible rabies vaccine in maize using the Vnukovo strain. *Dev Biol (Basel)* 131: 477- 482
- Lu J, Sivamani E, Li X, Qu R (2008) Activity of the 5' regulatory regions of the rice polyubiquitin *rub1.3* gene in transgenic rice plants as analyzed by both GUS and GFP reporter genes. *Plant Cell Rep* 27: 1587-1600.
- Maas C, Laufs J, Grant S, Korfhage C, Werr W (1991) The combination of a novel stimulatory element in the first exon of the maize *shrunken-1* gene with the following intron 1 enhances reporter gene expression up to 1000-fold. *Plant Mol Biol* 16:199-207.
- Mann DGJ, King ZR, Liu W, Joyce BL, Percifield RJ, Hawkins JS, LaFayette PR, Arieli BJ, Burris JN, Mazarei M, Benneisen JL, Parrott WA, Charles N Stewart CN (2011) Switchgrass (*Panicum virgatum* L.) polyubiquitin gene (*PvUbl1* and *PvUbl2*) promoters for use in plant transformation. *BMC Biotechnol* 2011, 11: 74
- McElroy D., Chamberlain D. A., Moon E., Wilson K. J. (1995) Development of gusA reporter gene constructs for cereal transformation, availability of plant transformation vectors from the CAMBIA molecular genetics resource service. *Mol Breed* 1: 27-37.

- Moon J, Parry G, Estelle M (2004) The ubiquitin – proteasome pathway and plant development. *Plant Cell* 16: 3181-3195
- Pickart CM, Eddius MJ (2004) Ubiquitin: structures, functions, mechanisms. *Biochim Biophys Acta* 1695: 55-72.
- Pineda Rodó A, Brugiere N, Vankova R, Malbeck J, Olson JM, Haines SC, Martin RC, Habben JE, Mok DW, Mok MC (2008) Over-expression of a zeatin O-glucosylation gene in maize leads to growth retardation and tasselseed formation. *J Exp Bot* 59(10): 2673-86.
- Rockhold DR, Chang S, Taylor N, Allen PV, McCue KF, Belknap WR (2008) Structure of two *Solanum bulbocastanum* polyubiquitin genes and expression of their promoter in transgenic potatoes. *Am J Pot Res* 85: 219-226.
- Rollfinke I K, Silber M V, Pfizner U M (1998) Characterization and expression of a heptaubiquitin gene from tomato. *Gene* 211: 267-276
- Rollfinke IK, Pfizner UM (1994) Structure of a heptaubiquitin gene from tomato. *Plant Physiol* 184: 299-300.
- Santner A, Estelle M (2010) The ubiquitin-proteasome system regulates plant hormone signaling. *Plant J* 61(6): 1029-1040.
- Schwechheimer C, Schwager K (2004) Regulated proteolysis and plant development. *Plant Cell Rep* 23 (6): 353 - 364.
- Sharp PM, Li WH (1987) Ubiquitin genes as a paradigm of concerted evolution of tandem repeats. *J Mol Evol* 25: 58-64.
- Sivamani E, Qu R (2006) Expression enhancement of a rice polyubiquitin gene promoter. *Plant Mol Biol* 60, 225-239
- Sullivan JA, Shirasu K, Deng XW (2003) The diverse roles of ubiquitin and the 26S proteasome in the life of plants. *Nat Rev Genet* 4, 948-958
- Sun CW, Griffen S, Callis J (1997) A model for the evolution of polyubiquitin genes from the study of *Arabidopsis thaliana* ecotypes. *Plant Mol Biol* 34: 745-758.
- Tan Y, Bishoff ST, Riley MA (1993) Ubiquitins revisited: further examples of within and between –locus concerted evolution. *Mol Phylog Evol* 2(4): 351-360
- Umezawa T, Fujita M, Fujita Yamaguchi-Shinozaki K, Shinozaki K (2006) Engineering drought tolerance in plant: discovering and tailoring genes to unlock the future. *Curr Opin Plant Biol* 17:113-122
- Wang F, Deng XW (2011) Plant ubiquitin-proteasome pathway and its role in gibberellin signaling. *Cell Res* 21: 1286-1294.
- Wang J, Qard J H (2003) Rice ubiquitin promoters: deletion analysis and potential usefulness in plant transformation systems. *Plant Cell Rep* 22: 129-134
- Wei H R, Wang M L, More PH, Albert HH (2003). Comparative expression analysis of two sugarcane polyubiquitin promoters and flanking sequences in transgenic plants. *Plant Physiol* 160: 1241-1251
- Wilkinson KD, Urban MK, Haas AL (1980) Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. *J Biol Chem* 255: 7529-7532
- Wu G, Cui H, Shu Q, Ye G, Xia Y, Gao M, Altosaar I, Li Y (2008) Transcriptional silencing and developmental reactivation of *cre1Ab* gene in transgenic rice. *Sci China Life Sci* 45(1): 68-78

UBIQUITINS: STRUCTURE, FUNCTION AND THE POTENTIAL USE OF THEIR PROMOTER IN PLANT GENETIC ENGINEERING.

Tran Thi Phuong Lien^{1,*}, Chu Hoang Ha¹, Nong Van Hai²

¹Institute of Biotechnology, Vietnam Academy of Science and Technology,
²Institute of Genome Research, Vietnam Academy of Science and Technology.

SUMMARY

Ubiquitin, the small conserved protein, is present in most types of tissues in eukaryotes. In their structure, ubiquitin genes have two forms, polyubiquitins and ubiquitin extension protein genes (also known as ubiquitin fusion genes). Each species has at least one polyubiquitin locus containing the tandem repeats of ubiquitin-coding monomer containing 228 bp for 76 amino acids. The number of tandem repeats can vary between loci within species and between species. Ubiquitin involves in proteolysis recognizing denatured proteins and polypeptides, or polypeptides need to be removed through the specific ubiquitination by the ubiquitin-proteasome system. There are three types of enzymes involved in this process: ubiquitin-activating enzyme E1, ubiquitin-conjugating enzyme E2 and ubiquitin-ligating enzyme E3 (E3 ligase). The ubiquitination

* Author for correspondence: Tel: +84-4-37918003, E-mail: tplien@ibt.ac.vn

regulates many protein function in cells, plays an important role in the growth and development of plants as well as disease resistance and tolerance to different stresses in the environment. The diversity of functions of ubiquitin has created the permission to use the ubiquitin promoter as an effective constitutive promoter in plant genetic engineering. The promoters showed higher level of gene expression in plants than CAMV35S promoter in times. For monocots this level has been higher than that in dicotyledonous plants. In this article, our knowledge about ubiquitins, the assumption of their functions, the encoding genes and the subsequent isolation and characterisation of ubiquitin promoter, prospects of their application in plant genetic engineering will be analyzed and discussed.

Keywords *ubiquitin, ubiquitin-proteasome system, E3 ligase, promoter structure, gene expression into plants*