

# Phân tích in silico các gene *galactinol synthase* ở cây đậu Cove (*Phaseolus vulgaris* L.)

• Cao Phi Bằng

Trường Đại học Hùng Vương, Phú Thọ

( Bài nhận ngày 01 tháng 04 năm 2015, nhận đăng ngày 21 tháng 11 năm 2016)

## TÓM TẮT

*Galactinol synthase* (GolS, EC 2.4.1.123) là enzyme thuộc họ glycosyltransferase 8 (GT8) nhóm xúc tác sự galactosyl hóa myo-inositol để tạo thành galactinol, tiền chất của raffinose, đường tan có vai trò quan trọng trong tính chống chịu ở thực vật. Chúng tôi đã xác định và phân tích năm gene GolS trong hệ gene của cây đậu Cove. Các gene này có từ hai tới ba intron. Các protein suy diễn có từ 326 tới 339 amino acid. Các protein GolS của cây đậu Cove có mang đầy đủ các motif bảo thủ đặc trưng đã biết cho các GolS, có tính axit, ưa nước yếu. Chỉ một hiện tượng nhân gene

sau quá trình biệt hóa loài được phát hiện ở cây đậu Cove. Các gene PvGolS biểu hiện ở cả các cơ quan sinh dưỡng và cơ quan sinh sản với mức độ khác nhau. Khi so sánh mức độ biểu hiện của ba gene, PvGolS2 biểu hiện yếu nhất, PvGolS1 biểu hiện mạnh nhất ở lá và thân, trong khi PvGolS3 biểu hiện cao nhất ở rễ và nốt sần. Lần đầu tiên phát hiện *Rhizobium* có ảnh hưởng tới sự biểu hiện các gen GolS. Ở mô sinh sản, gene PvGolS1 biểu hiện ở hầu hết các mô sinh sản được nghiên cứu. Mức độ biểu hiện của gene này tăng dần theo sự phát triển của cơ quan.

**Từ khóa:** galactinol synthase (GolS), biểu hiện gene, cây di truyền, đặc trưng của gene, đậu Cove

## MỞ ĐẦU

Cây đậu Cove (*Phaseolus vulgaris* L.) là cây thuộc họ Đậu, có khả năng cố định đạm. Hạt đậu Cove có nhiều protein, năng lượng, xơ, khoáng và vitamin nên có giá trị dinh dưỡng cao và tốt cho sức khỏe [3]. Do có giá trị lớn, cây đậu Cove được trồng rộng rãi trên thế giới. Gần đây, hệ gene của cây đậu Cove đã được giải trình tự vào năm 2014 nhằm thúc đẩy các nghiên cứu về cây đậu Cove ở mức phân tử. Phiên bản hệ gene gần đây của cây đậu Cove có kích thước 472,5 Mb [14].

Galactinol synthase (GolS, EC 2.4.1.123) còn được gọi là inositol-3- $\alpha$ -galactosyltransferase. Các galactinol synthase xúc tác cho sự galactosyl hóa myo-inositol để tạo thành galactinol (*O*- $\alpha$ -D-galactopyranosyl-[1 $\rightarrow$ 1]-L-myoinositol).

Galactinol là tiền chất của họ các đường tan có tên gọi raffinose. Sự tích lũy raffinose ở thực vật được báo cáo có liên quan tới tính chống chịu các điều kiện bất lợi, như là lạnh, nóng và khô hạn [12, 17, 22]. Ở hạt, sự tích lũy raffinose xảy ra đồng thời với sự phát triển của sự giảm hàm lượng nước trong quá trình chín của hạt [2], đồng thời, hàm lượng raffinose có mối tương quan thuận chiều với khả năng sống của hạt trong quá trình dự trữ [1].

Nhiều nghiên cứu đã chỉ ra rằng GolS có vai trò quan trọng, làm tăng tính chống chịu với nhiều điều kiện bất lợi ở thực vật [10, 17, 21, 22]. Sự biểu hiện và hoạt tính của galactinol synthase trong trả lời với các điều kiện bất lợi đã được báo cáo ở nhiều thực vật như *A. thaliana* [12, 17], cây Đan sâm [20], cây cà phê [4] và cây dương [21]. Tuy

nhiên, đến nay chưa có công trình nghiên cứu nào được thực hiện về các GolS của cây đậu Cove.

Trong công trình này, chúng tôi có mục tiêu xác định các gene mã hóa GolS trong hệ gene của cây đậu Cove, phân tích các đặc tính hóa-lí và cấu trúc của các gene *GolS* ở loài cây này. Chúng tôi cũng nghiên cứu sự biểu hiện của các gene *GolS* của loài cây này. Những kết quả nghiên cứu này bước đầu cung cấp các thông tin khoa học có ý nghĩa về nghiên cứu chức năng của các GolS của loài cây có hạt chứa dầu quan trọng này.

## VẬT LIỆU VÀ PHƯƠNG PHÁP

### Cơ sở dữ liệu

Trình tự hệ gene của cây đậu Cove được lấy từ Schmutz et al (2014) [14]. Dữ liệu RNA-seq của cây đậu Cove được lấy từ O'Rourke et al (2014) [13]. Các trình tự GolS của cây *Arabidopsis* được lấy từ Sengupta và et al (2012) [15].

### Xác định các gene thuộc họ GolS ở cây đậu Cove

## KẾT QUẢ VÀ THẢO LUẬN

### Xác định họ gene *GolS* ở cây đậu Cove

**Bảng 1.** Các gene thuộc họ GolS của cây đậu Cove và đặc điểm của chúng

Gene	Tên locus	Kích thước gene/CDS (bp)	Chiều dài protein (aa)	Khối lượng protein (kD)	pI	GR AVY	NST	Số lượng intron
<i>PvGolS1</i>	Phvul.001G21 5300	1529/1020	339	38,81	5,25	-0,16	1	2
<i>PvGolS2</i>	Phvul.001G22 3700	2382/981	327	37,70	5,92	-0,32	1	3
<i>PvGolS3</i>	Phvul.007G20 3400	1269/984	326	37,35	5,85	-0,22	7	3

Nhờ sử dụng GolS của cây *A. thaliana* làm khuôn dò để tìm kiếm các gene mã hóa GolS trong hệ gene của cây đậu Cove, chúng tôi tìm thấy ba gene mã hóa cho các GolS (Bảng 1). Như vậy, họ *GolS* của cây đậu Cove nhỏ hơn hai lần so với của

Chúng tôi sử dụng các protein GolS của cây *Arabidopsis* [15] làm khuôn dò để tìm kiếm các gene tương đồng trên dữ liệu nucleotide hệ gene của cây đậu Cove nhờ chương trình TBLASTN.

### Xây dựng cây phả hệ

Trình tự protein suy diễn GolS của cây đậu Cove và của *A. thaliana* được sắp dãy bằng MAFFT [7], cây phả hệ được xây dựng nhờ phần mềm MEGA 5 [18].

### Phân tích các đặc điểm hóa - lí

Các đặc điểm vật lí, hóa học của các gene cũng như protein được phân tích bằng các công cụ của ExPASy [5]. Cấu trúc exon/intron được xây dựng nhờ GSDS 2.0 [6].

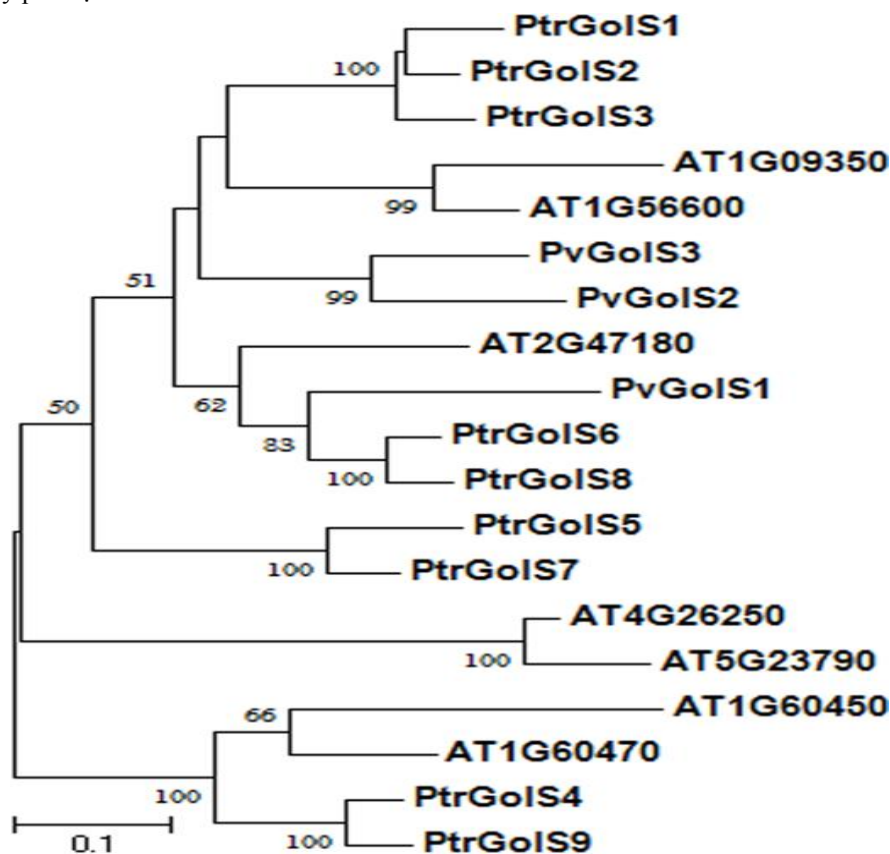
### Nghiên cứu sự biểu hiện gene

Sự biểu hiện của các gene được nghiên cứu nhờ phân tích kết quả RNA-seq của cây đậu Cove [13].

cây *A. thaliana* (có 7 gene) [15] và nhỏ hơn ba lần so với của cây dương (có 9 gene) [21]. Các protein suy diễn của các gene này đều mang vùng bảo thủ đặc trưng cho họ GT8 (họ glycosyltransferase 8) [15]. Trong trình tự của chuỗi peptide, các GolS



Phân tích cây phả hệ



**Hình 3.** Cây phả hệ được xây dựng từ các GolS của cây đậu Cove (Pv), cây dương (Ptr) và cây *A. thaliana* (At)

Cây phả hệ được thiết lập từ các protein GolS của cây đậu Cove, cây dương và *A. thaliana* (hình 3) gồm ba nhánh, trong đó chỉ một nhánh có các GolS của cả ba loài (nhưng giá trị bootstraps không cao), một nhánh khác chỉ mang hai gene của cây *A. thaliana*, nhánh còn lại mang các GolS của hai loài dương và *A. thaliana*. Kết quả phân tích cây phả hệ gợi ý rằng gen mã hóa *GolS* thuộc về một họ đa gene nhỏ trước quá trình biệt hóa giữa các loài này. Chúng tôi cũng nhận thấy rằng ở cây đậu Cove chỉ có một hiện tượng nhân gene sau quá trình biệt hóa loài (*PvGolS2* và *PvGolS3*). Trong khi đó, ở cây *A.*

*thaliana* có ba sự kiện nhân gene và có tới 5 sự kiện nhân gene ở cây dương [15, 21].

**Phân tích sự biểu hiện gene**

O'Rourke et al đã giải trình tự tập hợp các ARN thông tin tách từ các loại mô nghiên cứu của cây đậu Cove ở các thời điểm khác nhau và trong các điều kiện trồng khác nhau (bón phân, ủ với các *Rhizobium* hữu hiệu hoặc vô hiệu) [13]. Chúng tôi phân tích sự biểu hiện của các *GolS* của cây đậu Cove từ dữ liệu RNAseq của các mô sinh dưỡng và mô sinh sản riêng rẽ.

Gene	YL	L5	LF	LE	LI	YS	ST	RT	YR	R5	RF	RE	RI	N5	NE	NI	Chú thích
<i>PvGolS1</i>	5	7	25	63	114	111	13	0	9	14	7	6	6	1	2	0	≥30
<i>PvGolS2</i>	0	0	1	1	0	4	0	0	1	1	0	2	3	0	0	0	15
<i>PvGolS3</i>	5	3	4	11	19	2	2	3	1	15	3	17	15	25	3	2	0

**Hình 4.** Sự biểu hiện của các gene *Gols* của cây đậu Cove trong các mô nghiên cứu, (YL = Mô lá của cây ở giai đoạn lá chết thứ 2 hoàn chỉnh của cây đã được bón phân; L5 = Mô lá của cây đã được ủ với *Rhizobium* hữu hiệu 5 ngày; LF = Mô lá của cây được bón phân 21 ngày; LE = Mô lá của cây được ủ 21 ngày với *Rhizobium* hữu hiệu; LI = Mô lá của cây được ủ 21 ngày ủ với *Rhizobium* vô hiệu; YS = Tất cả các lông thân phía trên lá mầm của cây ở giai đoạn lá chết thứ hai; ST = Chồi bao gồm mô phân sinh đỉnh của cây ở giai đoạn lá chết thứ hai; RT = Chóp rễ thu từ cây được tưới phân ở giai đoạn hai lá chết; YR = Rễ hoàn chỉnh thu từ cây được tưới phân ở giai đoạn hai lá chết; R5 = Rễ hoàn chỉnh loại bỏ nốt sần trước cố định ở thời điểm 5 ngày; RF = Rễ hoàn chỉnh từ cây được bón phân 21 ngày; RE = Rễ hoàn chỉnh thu từ cây có nốt sần cố định hoạt động sau 21 ngày ủ; RI = Rễ hoàn chỉnh thu từ cây có nốt sần cố định bất hoạt sau 21 ngày ủ; N5 = Nốt sần trước cố định (hữu hiệu) được thu sau 5 ngày ủ; NE = Nốt sần cố định hữu hiệu được thu sau 21 ngày ủ; NI = Nốt sần cố định vô hiệu được thu sau 21 ngày ủ; đơn vị tính: RPKM (Số bản mã phiên (ARN)/1000 nucleotide của mã phiên/Một triệu bản mã phiên).

Gene	FY	PY	PH	P1	P2	SH	S1	S2	Chú thích
<i>PvGolS1</i>	11	0	5	23	32	0	48	122	≥30
<i>PvGolS2</i>	0	0	2	0	0	0	1	1	15
<i>PvGolS3</i>	3	0	3	1	8	1	0	0	0

**Hình 5.** Sự biểu hiện của các gene *Gols* của cây đậu Cove trong các mô nghiên cứu, (FY = Hoa non; PY = Vô quả non ở giai đoạn 1-4 ngày sau rụng hoa, mẫu chứa các phôi đang phát triển ở giai đoạn hình cầu; PH = Vô của quả dài gần 9 cm, hạt ở giai đoạn hình tim; P1 = Vô của quả dài gần 10-11 cm, hạt ở giai đoạn 1; P2 = Vô của quả dài gần 12-13 cm, hạt ở giai đoạn 2; SH = Hạt ở giai đoạn hình tim, dài 3-4 mm, nặng gần 7 mg; S1 = Hạt ở giai đoạn 1, dài 6-7 mm, nặng gần 50 mg; S2 = Hạt ở giai đoạn 2, dài 8-10 mm, nặng gần 140-150 mg; đơn vị tính: RPKM (Số bản mã phiên (ARN)/1000 nucleotide của mã phiên/Một triệu bản mã phiên).

Ở các mô sinh dưỡng, cả ba *Gols* biểu hiện trong ít nhất một loại mô nhưng chỉ duy nhất *PvGolS3* biểu hiện ở tất cả các mô và các điều kiện thí nghiệm khác nhau. Gene *PvGolS2* biểu hiện yếu hoặc không biểu hiện ở nhiều mô và điều kiện khác nhau. Ở các mô trên mặt đất (bộ phận khí sinh), gene *PvGolS1* biểu hiện mạnh nhất, ngược lại gene *PvGolS3* biểu hiện cao nhất trong số ba gene ở các mô phía dưới mặt đất. Việc ủ với các *Rhizobium* có hiệu ứng dương đối với sự biểu hiện của cả hai gene (*PvGolS1* và *PvGolS3*) ở mô lá, hiệu ứng dương của *Rhizobium* đối với sự biểu hiện gene ở mô rễ chỉ được quan sát ở gene *PvGolS2* và *PvGolS3* (Hình 4).

Sự biểu hiện của các gene *Gols* của cây đậu Cove ở các mô sinh sản được thể hiện ở Hình 5. Gene *PvGolS2* vẫn là gene biểu hiện yếu nhất, chỉ biểu hiện ở vô quả giai đoạn PH và hạt ở giai đoạn S1 và S2. Gene *PvGolS3* biểu hiện mạnh hơn so

với *PvGolS2* ở hoa và quả, ngoài ra gene này còn biểu hiện ở cả hoa non (FY), vô quả giai đoạn P1 và P2 và hạt ở giai đoạn S1. Gene *PvGolS1* biểu hiện ở hầu hết các mô sinh sản được nghiên cứu (trừ vô quả ở giai đoạn PY và hạt ở giai đoạn SH) và biểu hiện mạnh nhất trong số ba gene. Mức độ biểu hiện của gene này có xu hướng tăng dần theo sự phát triển của cơ quan, đặc biệt ở hạt nơi quá trình mất nước diễn ra dần khi hạt trưởng thành, đồng thời các cơ chế chống mất nước bảo vệ phôi cũng diễn ra mạnh mẽ [8].

Như vậy, các gene *Gols* của cây đậu cove biểu hiện ở cả các mô sinh dưỡng và mô sinh sản. Đã có nhiều báo cáo về sự biểu hiện của các gene *Gols* ở mô lá, mô rễ ở các loài thực vật khác, đặc biệt trong điều kiện bất lợi của môi trường [4, 12, 16, 17, 20, 21, 22]. Tuy nhiên, lần đầu tiên sự biểu hiện của các *Gols* dưới sự ảnh hưởng của các *Rhizobium* được giới thiệu trong công trình này.

Hai trong ba gene *GolS* của đậu Cove biểu hiện ở mô hoa, giống với sự biểu hiện của *CaGolS3* ở cây cà phê [4]. Sự biểu hiện các *GolS* ở mô hạt cũng đã được báo cáo [1, 15, 17, 19]. Các *GolS* được cho là tham gia vào cơ chế chống mất nước trong quá trình phát triển của hạt [9, 10].

#### KẾT LUẬN

Trong công trình này, chúng tôi đã phân tích tổng số ba gene mã hóa *GolS* ở cây đậu Cove. Các protein suy diễn *GolS* của cây đậu Cove có các motif bảo thủ đặc trưng cho các *GolS* đã được biết.

Các gene mã hóa *GolS* không liên tục, có hai hoặc ba intron. Các protein suy diễn có từ 322 tới 340 amino acid, có tính acid, ưa nước yếu. Chỉ một hiện tượng nhân gene *GolS* sau quá trình biệt hóa loài được phát hiện ở cây đậu Cove. Các gene *PvGolS* biểu hiện ở cả mô sinh dưỡng và mô sinh sản của cây đậu Cove.

*Lời cảm ơn:* Công trình này được hoàn thành với sự hỗ trợ kinh phí từ chương trình nghiên cứu khoa học cơ bản của Trường Đại học Hùng Vương, tỉnh Phú Thọ.

## In silico analysis of *galactinol synthase* genes in common bean (*Phaseolus vulgaris* L.)

• Cao Phi Bang

Hung Vuong University –Phu Tho

#### ABSTRACT

*Galactinol synthases (GolS, EC 2.4.1.123) belong to the glycosyltransferase 8 family. These enzymes catalyzes galactosylation of myo-inositol to form galactinol, precursor of raffinose which is soluble sucrose playing an important role in stress tolerance in plants. We identified and analyzed a total of three GolS encoded genes in the whole genome of common bean. These genes possessed two or three introns. The predicted proteins contained from 322 to 340 amino acids. The common bean GolS proteins contained all conserved motifs that were characterized for other previously known GolSs. These proteins were acidic and weakly hydrophilic. The phylogenetic tree constructed from GolS proteins of Arabidopsis, poplar and common bean showed that only one*

*gene duplication event was detected in the common bean genome. Common bean GolS genes expressed in all vegetative and reproductive tissues. However, the transcript abundances of each gene varied in different tissues. When comparing the expression levels of these three genes as defined RNAseq, we found that the expression level of PvGolS2 was the lowest, PvGolS1 expressed the strongest leaves and stem, while PvGolS3 expressed in roots and nodules. We discovered that Rhizobium incubation affected on the expression of GolS genes for the first time. In the reproductive tissues, PvGolS1 expressed in most studied tissues. Its expression level correlatively increased with the tissue development.*

**Keywords:** galactinol synthase (*GolS*), gene expression, gene characterization, phylogeny, common bean

## TÀI LIỆU THAM KHẢO

- [1]. I.L. Bernal, A.C. Leopold., Changes in soluble carbohydrates during seed storage. *Plant Physiol.*, 98, 3, 1207–1210 (1992).
- [2]. M. Black, F. Corbineau, H. Gee, D. Come, Water content, raffinose, and dehydrins in the induction of desiccation tolerance in immature wheat embryos. *Plant Physiol.*, 120, 2, 463–471 (1999).
- [3]. G.E. de Almeida Costa., K. da Silva Queiroz-Monici, S.M. Pissini Machado Reis., A.C. de Oliveira., Chemical composition, dietary fibre and resistant starch contents of raw and cooked pea, common bean, chickpea and lentil legumes. *Food Chem.*, 94, 3, 327–330 (2006).
- [4]. T.B. Dos Santos, G.I. Budzinski, C.J. Marur, C.L. Petkowicz., L.F. Pereira, L.G. Vieira, Expression of three galactinol synthase isoforms in *Coffea arabica* L. and accumulation of raffinose and stachyose in response to abiotic stresses. *Plant Physiol. Biochem.*, 49, 4, 441–448 (2011).
- [5]. E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, R. Marc, M.R. Wilkins., R.D. Appel., A. Bairoch, Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook*, 571–607, Springer (2005).
- [6]. Guo A. Y., Zhu Q. H., Chen X., Luo J. C., GSDS: a gene structure display server. *Yi Chuan*, 29, 8, 1023–1026 (2007).
- [7]. K. Katoh, M.D. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.*, 30, 4, 772–780 (2013).
- [8]. Leprince O., Hendry G. A. F., McKersie B. D., The mechanisms of desiccation tolerance in developing seeds. *Seed Sci. Res.*, 3, 04, 231–246 (1993).
- [9]. X. Li, J. Zhuo, Y. Jing, X. Liu, X. Wang, Expression of a galactinol synthase gene is positively associated with desiccation tolerance of *Brassica napus* seeds during development. *J. Plant. Physiol.*, 168, 15, 1761–1770 (2011).
- [10]. J.J.J. Liu., C.D. Krenz, F.A. Galvez, B.O. de Lumen, Galactinol synthase (GS): increased enzyme activity and levels of mRNA due to cold and desiccation. *Plant Sci.*, 134, 1, 11–20 (1998).
- [11]. J.J. Liu., W. Odegard, B.O. de Lumen., Galactinol synthase from kidney bean cotyledon and zucchini leaf. Purification and N-terminal sequences. *Plant Physiol.*, 109, 2, 505–511 (1995).
- [12]. A. Nishizawa, Y. Yabuta, Shigeoka S., Galactinol and raffinose constitute a novel function to protect plants from oxidative damage. *Plant Physiol.*, 147, 1251–1263 (2008).
- [13]. J. O'Rourke, L. Iniguez, F. Fu, B. Bucciarelli, S.S. Miller, A.S. Jackson, P.E. McClean, J. Li, X. Dai, P.X. Zhao., G. Hernandez, C.P. Vance, An RNA-Seq based gene expression atlas of the common bean. *BMC Genomics*, 15, 1, 866 (2014).
- [14]. J. Schmutz, P.E. McClean., S. Mamidi, G. A. Wu, S.B. Cannon., J. Grimwood, J. Jenkins, S. Shu, Q. Song, C. Chavarro, M. Torres-Torres , V. Geffroy, S.M. Moghaddam., D. Gao, B. Abernathy, K. Barry, M. Blair, M.A. Brick, M. Chovatia, P. Gepts, D.M. Goodstein., M. Gonzales, U. Hellsten, L.D. Hyten, G. Jia, J.D. Kelly., D. Kudrna, R. Lee, M.M. Richard, P.N. Miklas., J.M. Osorno., J. Rodrigues, V. Thareau, C.A. Urrea., M. Wang, Y. Yu, M. Zhang, R.A. Wing., P.B. Cregan, D.S. Rokhsar , S.A. Jackson., A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.*, 46, 7, 707–713 (2014).
- [15]. S. Sengupta, S. Mukherjee, S. Parween, A.L. Majumder., Galactinol synthase across evolutionary diverse taxa: functional preference for higher plants? *FEBS Lett.*, 586, 10, 1488–1496 (2012).
- [16]. T.P. Smith, M.T. Kuo, G.C. Crawford, Purification and characterization of galactinol synthase from mature zucchini squash leaves. *Plant Physiol.*, 96, 3, 693–698 (1991).
- [17]. T. Taji, C. Ohsumi, S. Iuchi, M. Seki, Kasuga M., Kobayashi M., Yamaguchi-Shinozaki K.,

- Shinozaki K., Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. *Plant J.*, 29, 4, 417–426 (2002).
- [18]. K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.*, 28 2731–2739 (2011).
- [19]. T. Ueda, P.M. Coseo, J.T. Harrell, L.R. Obendorf, A multifunctional galactinol synthase catalyzes the synthesis of fagopyritol A1 and fagopyritol B1 in buckwheat seed. *Plant Sci.*, 168, 3, 681–690 (2005).
- [20]. D. Wang, W. Yao, Y. Song, W. Liu, Z. Wang, Molecular characterization and expression of three galactinol synthase genes that confer stress tolerance in *Salvia miltiorrhiza*. *J. Plant Physiol.*, 169, 18, 1838–1848 (2012).
- [21]. J. Zhou, Y. Yang, J. Yu, L. Wang, X. Yu, M. Ohtani, M. Kusano, K. Saito, T. Demura, Zhuge Q., Responses of *Populus trichocarpa* galactinol synthase genes to abiotic stresses. *J. Plant Res.*, 127, 2, 347–358 (2014).
- [22]. C. Zhuo, T. Wang, S. Lu, Y. Zhao, X. Li, Guo Z., A cold responsive galactinol synthase gene from *Medicago falcata* (MfGolS1) is induced by myo-inositol and confers multiple tolerances to abiotic stresses. *Physiol. Plant.*, 149, 1, 67–78 (2013).