

Anthocyanin galactosyltransferase from *Aralia cordata*, cDNA cloning and characterization

Shigeyuki NAGASHIMA, Aya OKAMOTO, Hideyuki SUZUKI, Yoshihisa ASADA,
Toshiya KONDO and Takafumi YOSHIKAWA*

*School of Pharmaceutical Sciences, Kitasato University 5-9-1 Shirokane,
Minato-ku, Tokyo 108-8641, Japan*

*Corresponding author E-mail address: yoshikawat@pharm.kitasato-u.ac.jp

Received 7 January 2004; accepted 2 April 2004

Abstract

The UDP-galactose: anthocyanin 3-*O*-galactosyltransferase (ACGaT) cDNA was cloned from cell-suspension cultures of *Aralia cordata* Thunb. (Araliaceae). The cDNA was isolated from a cDNA library using a galactosyltransferase cDNA from *Vigna mungo*. The cDNA contained an open reading frame encoding 452 amino acids with a calculated molecular mass of 50.0 kDa. The deduced amino acid sequence of ACGaT cDNA included a plant secondary product glycosyltransferase signature sequence and was similar to the flavonoid 3-*O*-glycosyltransferases from petunia and grape. The enzymatic activity of the recombinant ACGaT expressed in *Escherichia coli* was detected toward anthocyanins and flavonols as an aglycon and UDP-galactose as a sugar donor. The reaction product for quercetin and UDP-galactose catalyzed by the recombinant ACGaT was identified as quercetin 3-*O*- β -galactoside using nuclear magnetic resonance spectroscopy.

Accession number: AB103471

UGT number: UGT78A2

Key words: *Aralia cordata*, UDP-galactosyltransferase, heterologous expression

Abbreviations

ACGaT, *Aralia cordata* UDP-galactose: anthocyanin 3-*O*-galactosyltransferase; rACGaT, recombinant ACGaT; UDP, uridine diphosphate.

Introduction

Anthocyanins are one of the most important water-soluble pigments and glycosylation is the major modification reaction of these compounds. The glycosylation reaction is catalyzed by glycosyltransferases which ubiquitously exist in mammals and plants. Various plant glycosyltransferases have been cloned and characterized. Sugar acceptors include a wide range of compounds such as flavonoids, anthocyanins, terpenoids, sterols and thiohydroximates. On the other hand, sugar donors are mainly limited to UDP-glucose, UDP-galactose and UDP-glucuronic acid; the other UDP-sugars are minor donors (Vogt and Jones, 2000).

Aralia cordata has long been used as a source for

crude drugs and for food; this plant is known to include anthocyanins in seeds and young buds. We have already reported that cyanidin 3-*O*-xylosylgalactoside is the main anthocyanin of *A. cordata* cell-suspension cultures (Sakamoto *et al.*, 1993). In this paper, we report the cDNA cloning of anthocyanin 3-*O*-galactosyltransferase from *A. cordata* cells and expression in *E. coli*. The enzymatic reaction product using quercetin and UDP-galactose was identified as quercetin 3-*O*-galactoside based on the ^1H - and ^{13}C -NMR spectra.

Materials and Methods

Plant material

Callus tissue of *Aralia cordata* Thunb. was initiated from young leaves and stems (Sakamoto *et al.*, 1994) and subcultured on Murashige and Skoog medium (Murashige and Skoog, 1964) containing 1 mg l⁻¹ 2,4-dichlorophenoxyacetic acid and 0.1 mg l⁻¹ kinetin, 3% (w/v) sucrose and 70 ml l⁻¹ coconut water. To produce the cell-suspension cultures,

approximately 4 g of three-week-old callus tissue were transferred into 125 ml of liquid medium in a 500-ml Erlenmeyer flask and cultured at 25°C in the dark at 80 rpm on a reciprocal shaker. All samples were frozen in liquid nitrogen and stored at -80°C until use.

Cloning of ACGaT cDNA

Standard recombinant DNA manipulation was according to Sambrook and Russell (2001). All kits required for a mRNA preparation and a cDNA library construction were purchased from Amersham Biosciences. The total RNA was isolated from 140 g of 3-week-cultured cell suspension cultures according to Shirzadegan *et al.* (1991). A poly (A)⁺ RNA was prepared using a mRNA Purification Kit. The cDNA library was constructed using a TimeSaver cDNA Synthesis Kit and a Lambda Packaging Kit according to the manufacturer's instructions. Approximately 3.5×10^5 non-amplified plaques were screened with the ³²P-labeled cDNA insert of the VmUF3GT from *Vigna mungo* (Mato *et al.*, 1998). Hybridization on Hybond N⁺ (Amersham Biosciences) was carried out at 65°C for 14 h in 6 × SSPE, 5 × Denhardt's solution and 0.1 mg l⁻¹ salmon sperm DNA. A final washing of the membranes was conducted at 65°C in 2 × SSPE and 0.1% SDS. One positive plaque was isolated and subjected to *in vivo* excision to produce the plasmid clone pACGaT. The nucleotide sequences were determined using a DNA sequencer DSQ-2000L (Shimadzu, Kyoto, Japan). The nucleic acid sequence was deposited in DDBJ and the UGT name was provided by the UGT Nomenclature Committee (Mackenzie *et al.*, 1997).

Heterologous expression in E. coli

The open reading frame of the ACGaT cDNA was cloned into an *Nde*I site of a pET-21d vector (Novagen). The recombinant ACGaT (rACGaT) was expressed in *E. coli* strain BL21 (DE3). *E. coli* transformants were cultured in 10 ml of an LB medium with 50 mg l⁻¹ carbenicillin at 37°C. Isopropyl 1-thio-β-D-galactopyranoside was added to a final concentration of 1 mM when the OD₆₀₀ reached approximately 0.7. rACGaT was induced for 3 h at 37°C. All subsequent operations were carried out at 4°C. Cells were collected by centrifugation and resuspended in 1 ml of an extraction buffer (200 mM Tris-HCl, pH 7.5, containing 1 mM 2-mercaptoethanol and 0.5 mM EDTA). After sonication with an Astrason sonicator, cell debris was removed by centrifugation and the supernatant was used as the recombinant enzyme solution for an assay.

Enzyme assay

The reaction mixture contained 100 mM Tris-HCl buffer (pH 7.5), 1 mM UDP-galactose, 0.25 mM substrate and 80 μl of the rACGaT enzyme solution in a final volume of 200 μl. The mixture was incubated at 30°C for 30 min and the reaction was terminated by the addition of 200 μl of methanol. After centrifugation at 10,000g for 10 min, a 100-μl aliquot of the supernatant was subjected to the HPLC analysis to detect the enzymatic reaction products. The HPLC system was composed of a 600E pump and a 996 Photodiode Array Detector (Waters) with a Mightysil RP-18 column (4.6 mm i.d., 150 mm long; Kanto Chemical, Tokyo, Japan). Substrates and reaction products were separated by isocratic elution of CH₃CN-60 mM phosphoric acid (20:80) at a flow rate of 1.0 ml min⁻¹. The eluent was monitored by absorption at 274 nm.

To compare enzymatic activities toward various substrates, reactions were carried out using UDP-[¹⁴C]-sugars as a sugar donor. The reaction mixture contained 100 mM Tris-HCl buffer (pH 7.5), 1 mM UDP-sugar, 5 mM substrate, 1.85 kBq of UDP-[U-¹⁴C]-sugar and 100 μl of rACGaT enzyme solution in a final volume of 600 μl. The mixture was incubated at 30°C for 3 h and extracted twice with 500 μl of water saturated 1-butanol. The radioactivity in a 600-μl aliquot of the 1-butanol extract was measured using a liquid scintillation counter (Beckman LS6000IC).

Identification of the enzymatic reaction product of quercetin by nuclear magnetic resonance (NMR) spectroscopy

To identify the enzymatic reaction product of quercetin, a scaled-up reaction was carried out. The reaction mixture contained 100 mM Tris-HCl (pH 7.5), 0.5 mM UDP-galactose, 0.25 mM quercetin and 26 ml of rACGaT enzyme solution in a final volume of 65 ml. The reaction mixture was incubated at 30°C for 36 h and extracted twice with 100 ml 1-butanol. After evaporation of the 1-butanol fraction, quercetin 3-O-galactoside was purified by LH-20 column chromatography. The purified quercetin 3-O-galactoside was dissolved in methanol-d₄; the ¹H- and ¹³C- NMR spectra were recorded on a Varian XL-400 spectrometer.

Results and Discussions

Cloning, nucleotide sequencing and heterologous expression of ACGaT cDNA

ACGaT cDNA was cloned by a screening of approximately 3.5×10^5 plaques of the *A. cordata* cDNA library using *Vigna mungo* flavonoid 3-O-

galactosyltransferase cDNA (Mato *et al.*, 1998) as a probe. ACGaT cDNA was 1,584 bp in length with a poly (A)⁺ tail and contained a 1,356 bp open reading frame encoding a polypeptide of 452 amino acids. The calculated molecular mass was 50.0 kDa and the isoelectric point was 6.0. The deduced amino acid sequence has a plant secondary product glycosyltransferase consensus sequence (PSPG box; Hughes and Hughes, 1994), which is thought to be the UDP binding domain (Fig. 1). The deduced amino acid sequence of ACGaT cDNA was 61% identical to flavonoid 3-*O*-galactosyltransferase from petunia (Miller *et al.*, 1999) and 48 to 53% to flavonoid 3-*O*-glucosyltransferase from grape (Kobayashi *et al.*, 2001) and perilla (Gong *et al.*, 1997) by a BLAST program (Zhang and Madden, 1997) without low complexity sequences. Identity with *Vigna mungo* flavonoid 3-*O*-galactosyltransferase cDNA, used as a probe to ACGaT, was 45%. In a molecular phylogenetic tree based on deduced amino acid sequences (Fig. 2), flavonoid 3-*O*-glycosyltransferases are combined into a single cluster (Hirotani *et al.*, 2000, Vogt and Jones,

2000, Taguchi *et al.*, 2001). Because the deduced amino acid sequence of ACGaT cDNA was located in the cluster of the flavonoid 3-*O*-glycosyltransferases, ACGaT cDNA is thought to encode a glycosyltransferase that transfer a sugar to the 3-hydroxyl group of a flavonoid. To confirm the substrate specificity of ACGaT, the open reading frame of ACGaT was cloned into the pET-21d and expressed. rACGaT was assayed using cyanidin, pelargonidin, quercetin, luteolin, naringenin, taxifolin and (+)-catechin as an aglycon, and UDP-[¹⁴C]-galactose and UDP-[¹⁴C]-glucose as a sugar donor (Table 1). It was shown that rACGaT was specific to UDP-galactose rather than to UDP-glucose and that cyanidin was the best sugar acceptor followed by quercetin and pelargonidin. The others were not used as an aglycone. Subsequently, the reaction mixture using quercetin and UDP-galactose was separated by reversed-phase HPLC. Quercetin was converted to quercetin 3-*O*-galactoside (Fig. 3). These results indicate that ACGaT encodes an anthocyanin 3-*O*-galactosyltransferase.

TCTTCACTAGTCACTAGACTACTATACAACCTGGATATATATATATGTATGTATTATATCCCTTA	65
TTTATTATAAGTTAATTCATCTGTGCTAGTTACCTATCTCTCTCTGGAATTCGAGAATTAATCAGCAGGTGGAATAATTAATTATTA	155
ATGGGGAGTTTCAGCGGAGCCACATGTGGGGTTCCTTGCCTTCCCATTGCTACACACGCGGTCTCCTCCTGGGCCTAGTCCGTAGGCTA	245
M G S S A E P H V G V L A F P F A T H A G L L L G L V R R L	30
GCCGCGCTGCCCCAATCACTCAATTTCTCTTCTATAGCACCGCTGCATCCAAACCGTTCATTATTTTCATATCCAAATTTCTCCTTATTTCT	335
A A A A P N V N F S F Y S T A A S N R S L F S Y P N S P Y S	60
AATGTAATACCTACGATGTCTCCGATGGCTACAGAGGGTACGTGTTTTCCGGGAAGCCACAAGAGGATATTAACCTTGTTCCTTGACC	425
N V I P Y D V S D G V P E G Y V F S G K P Q E D I N L F L T	90
GTTGCCTCCGATGAATTTAAGAGAGGCTTGGAAAAAGCCGCGGTGGACTCCGGTAGGAAGATCACCTGTTTGGTGGCGGACGCTTTTFTA	515
V A S D E F K R G L E K A A V D S G R K I T C L V A D A F L	120
TGGTTTTCCGGCGATTGGCCGAACAAATTCGTGTGCCATGGGTCCCACTTTGGACTTCCGGAGCTTGCTCGCTTTGATCCACGTTTAC	605
W F S G D L A E Q I R V P W V P L W T S G A C S L S I H V Y	150
ACCGATCTTATCAGGCAAACCTGTTGGACTTGGTGGTATTGAAGGACGTATGGATGAGATTCTGACGTTTCCCGGATTTTCTGAATTA	695
T D L I R Q T V G L G G I E G R M D E I L T F I P G F S E L	180
CGGCTCGGTGACTTACCCGGTGGAGTCTTTTCGGTAACTTGAATACACCTTCTCAATAATGCTACATAAAATGGGACAACTCTACCC	785
R L G D L P G G V L F G N L E S P F S I M L H K M G Q T L P	210
CGGGCAGCTGCCGTTCCCAITTAACSTCTCGAAGAGCTTGATCCTGATCATGAAAGATATAAAATCAAAGTTCAAAGAAATCTCTCAAC	875
R A A A V P F E E L D P D L M K D I K S K F K I L N	240
GTCGGTCCCTTCAACCTAACATCGCCACCACCATCATCAAACCTCCGACGAACACGGCTGCATACCGTGGTGTAGACAACCAAAACCTTAA	965
V G P F N L T S P P P S S N S D E H G C I P W L D N Q N P K	270
TCGGTAGCGTATATGCTTCCGAAACCGTAGCAACGCCGCCGAATGAGCTAGTTTTCATTAGCTGAGGCGCTAGAAGAAAGTGGTACT	1055
S V A Y I A F G T V A T P P P N E L V S L A E A L E E S G T	300
CCATTTCTTTGGTCTTTAAAGGACAATTTCAAAAATCATTTACCAAAAGGGTTTTTAGAAAGGAATAGTAAAAGTGGGAAAATTTGTGGCA	1145
P F L W S L K D N F K N H L P K G F L E R N S K S G K I V A	330
TGGGCACCCAAATACAAGTTTGTGTCACATGATGCGGCTGGAGTTGTTATAACGCATGGCGGATGGAATTCGGTGGTGGAGAGCATTGCG	1235
W A P Q I Q V L S H D A V G V V I T H G G W N S V V E S I A	360
GCCGGTGTGCCAGTGATATGTAGGCGTCTTFCGGAGATCATCACATAAATACGTGGATGGTTGAAAATGTATGAAAATTTGGTGTGAGA	1325
A G V P V I C R P F F G D H H I N T W M V E N V W K I G V R	390
ATTGAAGTGGGGTTTTTCAAGAAGCTGGTACTATGAATGCGCTTGAACAAGTACTTTTTGTCACAAGAAAAGGGGAAGAAATTTGAAAGAG	1415
I E G G V F T R T G T M N A L E Q V L L S Q E K G K K L K E	420
CAAATCACAGTGTTTAAGGAGCTTGTGTTGAAGGCGTTGGACCTAATGGCAGCTCTACTCAAATTTTAAACGCTACTAGAAGTGATT	1505
Q I T V F K E L A L K A V G P N G S S T Q N F K R L L E V I	450
ACAACCTAATTTTTGTGGCATATGCACCTAGCTAGCTTGTGATGAATGAAGTTACCTTTTGTATTAAAAAATAAAAAA	1584
T T *	452

Fig. 1 Nucleotide and deduced amino acid sequences of ACGaT. Underlining indicates a plant secondary product glycosyltransferase consensus sequence (Hughes and Hughes, 1994). Asterisks indicate stop codons.

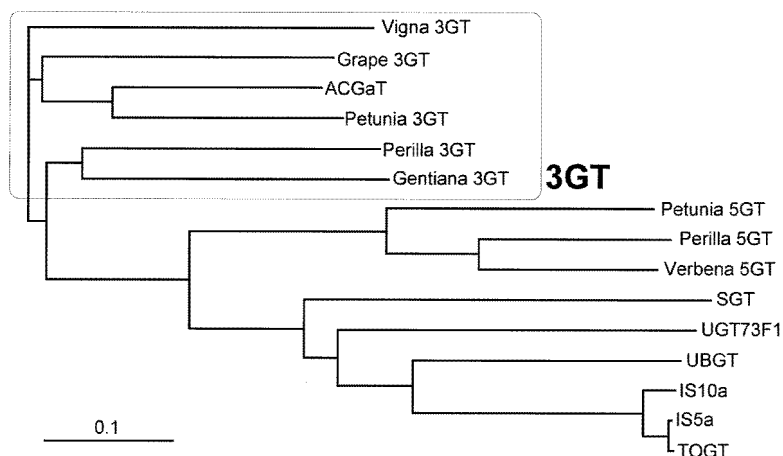


Fig. 2 Molecular phylogenetic tree of plant-derived glycosyltransferases. The tree was constructed by the neighbor-joining method using a ClustalX program. Nodes in the box indicate the flavonoid 3-*O*-glycosyltransferases. GenBank accession numbers: TOGT (AAK28303; *Nicotiana tabacum*); IS5a (AAB36653; *N. tabacum*); IS10a (AAB36652; *N. tabacum*); UBG (BAA83484; *Scutellaria baicalensis*); UGT73F1 (BAC78438, *Glycyrrhiza echinata*); SGT (AAB48444; *Solanum tuberosum*); Vigna 3GT (BAA36972; *Vigna mungo*); Petunia 3GT (AAD55985; *Petunia hybrida*); Grape 3GT (AAB81683; *Vitis vinifera*); Gentiana 3GT (BAA12737; *Gentiana triflora*); Perilla 3GT (BAA19659; *Perilla frutescens*); Verbena 5GT (BAA36423; *Verbena hybrida*); Perilla 5GT (BAA36422; *P. frutescens*); Petunia 5GT (BAA89009, *P. hybrida*).

Table 1 Substrate specificity of rACGaT measured by glycosylation of the indicated substrates. Radioactivity was not detected in the butanol extract of the control reaction using an empty vector. n.d.: not detected.

	relative activity (%)	
	UDP-galactose	UDP-glucose
cyanidin	100	23
pelargonidin	20	n.d.
quercetin	77	29
luteolin	n.d.	n.d.
naringenin	n.d.	n.d.
taxifolin	n.d.	n.d.
(+)-catechin	n.d.	n.d.

Identification of the enzymatic reaction product by ^1H - and ^{13}C -NMR spectra

To clarify the enzymatic reaction product of rACGaT with quercetin and UDP-galactose, a scaled-up reaction was carried out. The enzymatic reaction product was purified by LH-20 column chromatography. The product showed a pseudomolecular ion peak at $(\text{M}+\text{H})^+$ at m/z 465 in the FAB-MS spectrum. In the ^1H -NMR spectrum, an anomeric proton signal was observed at δ 5.14 (1H, d, $J=8.0\text{Hz}$, and Gal-1-H); the sugar moiety was confirmed to be β -D-galactopyranoside based on the coupling constants of each proton signal. The signals due to the aglycon part were observed at δ

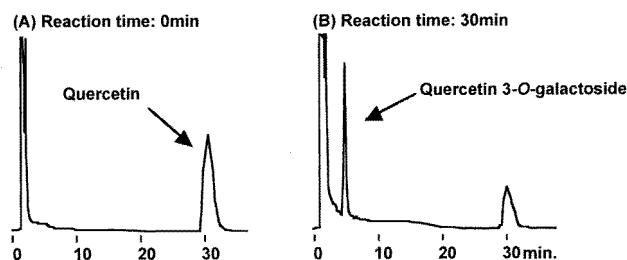


Fig. 3 HPLC analysis of the enzymatic reaction product of rACGaT. **A** The elution profile of the reaction product at time point 0 min. The reaction was carried out using quercetin and UDP-galactose at 30°C for 30 min. **B** The elution profile of the reaction mixture at time point 30 min.

6.18 (1H, d, $J=2.0\text{Hz}$, 6-H), 6.37 (1H, d, $J=2.0\text{Hz}$, 8-H), 6.86 (1H, d, $J=8.5\text{Hz}$, 5'-H), 7.58 (1H, dd, $J=8.5, 2.0\text{Hz}$, 6'-H) and 7.84 (1H, dd, $J=2.0\text{Hz}$, 2'-H). In order to elucidate the linkage position of the β -D-galactopyranosyl moiety, the ^{13}C -NMR spectrum was obtained. Comparison of the ^{13}C -NMR spectra of the product with quercetin revealed glycosylation shifts of aglycone carbons for the product at C-2 (+10.6), C-3 (-1.3) and C-4 (+2.1) indicating that the galactosyl moiety in the product was attached to the 3-position of quercetin (Markham and Chari, 1982). The structure of the product was confirmed to be quercetin 3-*O*-galactopyranoside.

In conclusion, we have isolated UDP-galactose:

anthocyanin 3-*O*-galactosyltransferase cDNA from *A. cordata* cultured cells. Although some galactosyltransferases have been reported, the enzymatic product was not confirmed precisely. In this report, we demonstrated that rACGaT transfers galactose to the 3-*O*-position of quercetin by the H^1 - and ^{13}C - NMR spectra. Knowing with confidence, the enzymatic reaction product of a plant glycosyltransferase is a critical point for the further analysis of plant glycosyltransferases. This clone, ACGaT, would be a useful tool to analyze the substrate specificity and the regio-specificity of plant glycosyltransferases.

Acknowledgements

The authors thank Dr. M. Mato (Akita Agricultural Experiment Station, Japan) for supplying the galactosyltransferase cDNA used for the probe. This research was funded by a Kitasato University Research Grant for Young Researchers.

References

- Hirofani, M., Kuroda, R., Suzuki, H., Yoshikawa, T., 2000. Cloning and expression of UDP-glucose: flavonoid 7-*O*-glucosyltransferase from hairy root cultures of *Scutellaria baicalensis*. *Planta*, **210**: 1006-1013.
- Hughes, J., Hughes, M. A., 1994. Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta* Crantz) cotyledons. *DNA Seq.*, **5**: 41-49.
- Gong, Z., Yamazaki, M., Sugiyama, M., Tanaka, Y., Saito, K., 1997. Cloning and molecular analysis of structural genes involved in anthocyanin biosynthesis and expressed in a forma-specific manner in *Perilla frutescens*. *Plant Mol. Biol.*, **35**: 915-927.
- Kobayashi, S., Ishimaru, M., Ding, C. K., Yakushiji, H., Goto, N., 2001. Comparison of UDP-glucose: flavonoid 3-*O*-glucosyltransferase (UFGT) gene sequences between white grapes (*Vitis vinifera*) and their sports with red skin. *Plant Sci.*, **160**: 543-550.
- Mackenzie, P. I., Owens, I. S., Burchell, B., Bock, K.W., Bairoch, A., Bélanger, A., Fournel-Gigleux, S., Green, M., Hum, D. W., Iyanagi, T., Lancet, D., Louisor P., Magdalow, J., Chowdhury, J. R., Ritter, J. K., Schachter, H., Tephly, T. R., Tipton, K.F., Nebert, D. W., 1997. The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics*, **7**: 255-269.
- Markham, K. R., Chari, V. M. 1982. Carbon-13 NMR spectroscopy of flavonoids. In: Harborne, J. B., Marby, T. J. (Eds.): *The Flavonoids: Advances in Research*, pp. 19-134. Chapman and Hall, London.
- Mato, M., Ozeki, Y., Itoh, Y., Higeta, D., Yoshitama, K., Teramoto, S., Aida, R., Ishikura, N., Shibata, M., 1998. Isolation and characterization of a cDNA clone of UDP-glucose: flavonoid 3-*O*-galactosyltransferase (UF3GaT) expressed in *Vigna mungo* seedlings. *Plant Cell Physiol.*, **39**: 1145-1155.
- Miller, K. D., Guyon, V., Evans, J. N. S., Shuttleworth, W. A., Taylor, L. P., 1999. Purification, cloning and heterologous expression of a catalytically efficient flavonol 3-*O*-galactosyltransferase expressed in the male gametophyte of *Petunia hybrida*. *J. Biol. Chem.*, **274**: 34011-34019.
- Murashige, T., Skoog, F., 1962. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.*, **15**: 473-497.
- Sakamoto, K., Iida, K., Sawamura, K., Hajiro K., Asada, Y., Yoshikawa, T., Furuya, T., 1993. Effects of nutrients on anthocyanin production in cultured cells of *Aralia cordata*. *Phytochemistry*, **33**: 357-360.
- Sakamoto, K., Iida, K., Sawamura, K., Hajiro K., Asada, Y., Yoshikawa, T., Furuya, T., 1994. Anthocyanin production in cultured cells of *Aralia cordata* Thunb. *Plant Cell Tissue Organ Cult.*, **36**: 21-26.
- Sambrook, J., Russel, D. W., 2001. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Shirzadegan, M., Christie, P., Seemann, J. R., 1991. An efficient method for isolation of RNA from tissue cultured plant cells. *Nucleic Acids Res.*, **19**: 6055.
- Taguchi, G., Yazawa, T., Hayashida, N., Okazaki, M., 2001. Molecular cloning and heterologous expression of novel glucosyltransferases from tobacco cultured cells that have broad substrate specificity and are induced by salicylic acid and auxin. *Eur. J. Biochem.*, **268**: 4086-4094.
- Vogt, T., Jones, P., 2000. Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trend. Plant Sci.*, **5**: 380-386.
- Zhang, J., Madden T. L., 1997. PowerBLAST: a new network BLAST application for interactive or automated sequence analysis and annotation. *Genome Res.*, **7**: 649-656.