Plant Molecular Biology **30**: 839–844, 1996.

Short communication

Isolation and characterization of an auxin-inducible glutathione *S*-transferase gene of *Arabidopsis thaliana*

Dianne A.M. van der Kop[1], Monique Schuyer[1], Ben Scheres[2], Bert J. van der Zaal[1] and Paul J.J. Hooykaas[1]

[1] *Institute of Molecular Plant Sciences, RUL-TNO Centre for Phytotechnogy, Clusius Laboratory, Leiden University, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands;* [2] *Department of Molecular Cell Biology, University of Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands*

Received 13 July 1995; accepted 20 November 1995

Key words: *Arabidopsis thaliana*, auxin, gene expression, glutathione *S*-transferase

Abstract

Genes homologous to the auxin-inducible *Nt103* glutathione *S*-transferase (GST) gene of tobacco, were isolated from a genomic library of *Arabidopsis thaliana*. We isolated a λ clone containing an auxin-inducible gene, *At103–1a*, and part of a constitutively expressed gene, *At103–1b*. The coding regions of the *Arabidopsis* genes were highly homologous to each other and to the coding region of the tobacco gene but distinct from the GST genes that have been isolated from arabidopsis thusfar. Overexpression of a cDNA clone in *Escherichia coli* revealed that the AT103–1A protein had GST activity.

The effects of auxin on the cellular differentiation in plants are likely to be due, at least in part, to changes in the pattern of gene expression. Recently it was found that some of the auxin-regulated genes encode proteins with glutathione *S*-transferase (GST) activity [4, 6, 15] that can also be induced by other hormones, heavy metals or environmental stress [9]. The coding sequence of one of these genes, the tobacco *Nt103–1* gene, represented by the cDNA clone pCNT103 [19], was used as a probe to isolate homologous genes from *Arabidopsis thaliana*.

After screening a genomic library at low stringency with pCNT103, recombinant clones were isolated. In three clones a 1 kb *Hin*dIII fragment hybridized with the pCNT103 probe. One clone containing the 1 kb fragment was designated λ103–1. The 1 kb *Hin*dIII fragment of λ103–1 was subcloned in pBlueScript SK$^+$, giving rise to the plasmid pSK103. This plasmid was used as a probe in DNA hybridization experiments. Southern blot analysis at high stringency on *Arabidopsis* genomic DNA digested with *Bam*HI, *Eco*RI or *Hin*dIII revealed one strongly hybridizing fragment of 11 kb, 6 kb and 1 kb, respectively. Also a weakly hybridizing fragment was present in these lanes, indicating that probably a second related gene was present in the genome.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Necleotide Sequence Databases under the accession number X89216.

The gene that hybridized strongly to the pSK103 probe was designated *At103–1a* while the gene that hybridized weakly to pSK103 was designated *At103–1b*. Digestion of the genomic DNA with *Xho*I revealed two strongly and two weakly hybridizing bands, indicating that internal *Xho*I restriction sites were present in the genes (Fig. 1A). We repeated the Southern blot analysis using λ103–1 DNA for the digestions and pSK103 as a probe. Fragments of 1 kb (*Hin*dIII), 6 kb

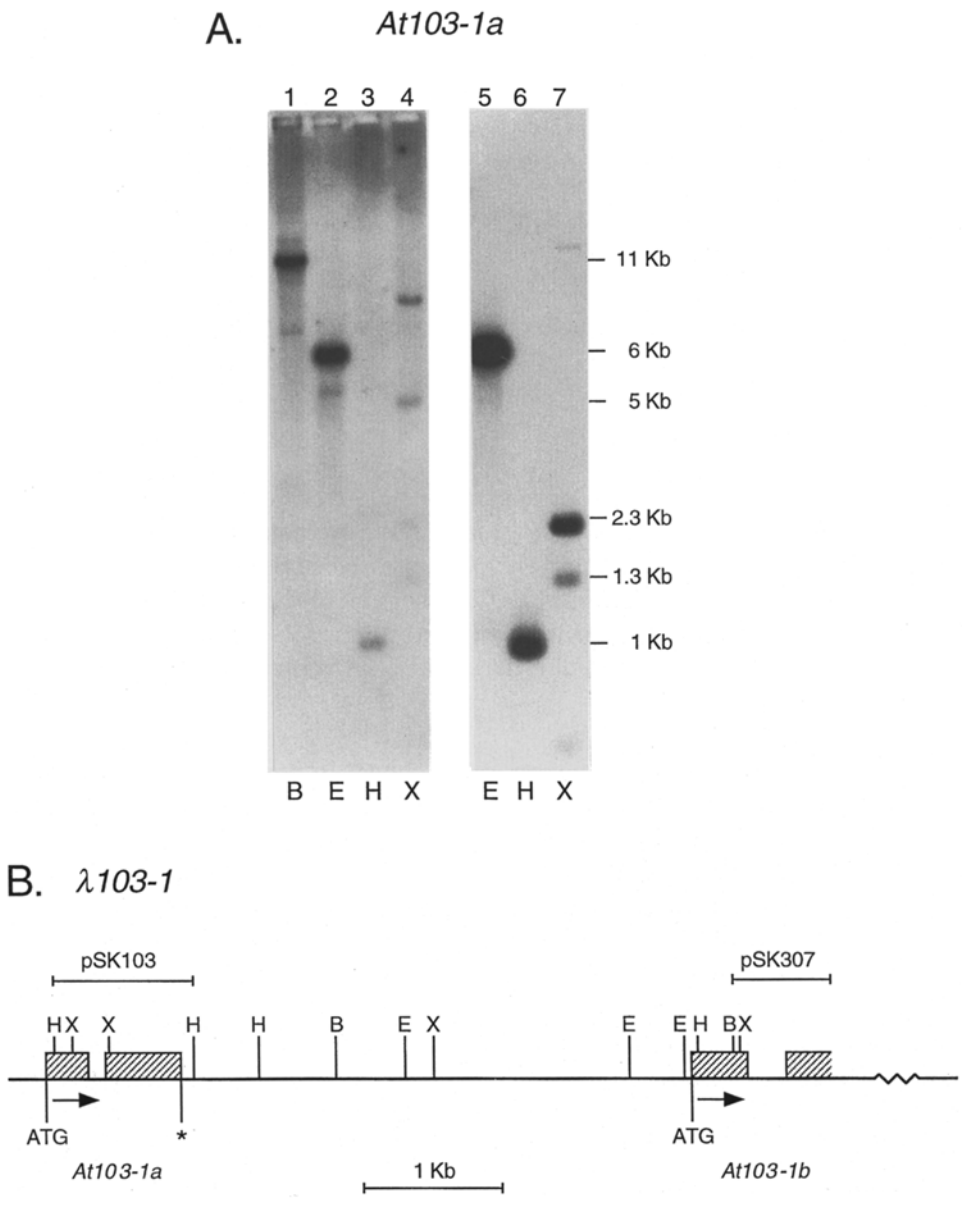

Fig. 1. A. Southern blot analysis of genomic DNA of *Arabidopsis thaliana* digested with *Bam*HI (lane 1), *Eco*RI (lane 2), *Hin*dIII (lane 3) or *Xho*I (lane 4) and λ103–1 DNA digested with *Eco*RI (lane 5), *Hin*dIII (lane 6) or *Xho*I (lane 7) hybridized to the 1 kb *Hin*dIII fragment of *At103–1a* (pSK103). B. Restriction map of λ103–1. The coding region of the genes with the start (ATG) and stop (*) codons are indicated as well as the exons and introns of the genes. The position of the probes that were used in the Southern and northern analysis are indicated as well (B, *Bam*HI; E, = *Eco*RI; H, *Hin*dIII; X, *Xho*I).

(*Eco*RI) and 2.3 and 1.3 (*Xho*I) were detected. These corresponded in size to the genomic fragments that hybridized strongly to the same probe.

Further analysis of λ103–1 revealed that the promoter region and part of the coding region of the *At103–1b* gene were also located on this clone (unpublished results). The *At103–1* genes were arranged in a tandem repeat at a distance of ca. 3.5 kb in the genome (Fig. 1B). Clusters of auxin-regulated genes have been reported earlier. In soybean five different SAUR genes were found on one phage clone [14]. Also a cluster of two ethylene-responsive GST genes was detected in carnation [10].

Gene structure and organization

The 1 kb *Hin*dIII fragment of λ103–1 was sequenced to obtain information on the coding region of the *At103–1a* gene (Fig. 2). The promoter region of this gene was subsequently cloned via IPCR [3] and then sequenced. The sequence we obtained from the *At103–1a* gene starts at a *Xmn*I site present 376 bp upstream of the ATG initiation codon, and ends at a *Hin*dIII site 170 bp downstream of the stop codon. The coding region of the gene is interrupted by an 118 bp intron starting at position 319 downstream of the ATG. The location of the splice site was confirmed after sequencing the cDNA corresponding to the *At103–1a* gene (results not shown). The position

```
-376  GAATTAATTCTTGACGAAGCATGGGGTTGTCTTCAGAAAATGAATGGATCAAGTCAAGGTACACATGCTCCAGGGATAAGGCAAGGTTAGGAATTAGGAC

-276  ACATCTCCACTTACTAAAAAAGAGATAAAAAAAAATTGTATAGGGAACGTTATAAATATGTTGTAAAGTCAACATCTGTTTCCTTCTAGACTCTTCGCAT

-176  TTACATCACACTGCCGACCATATAAAACGGCAAAGTTCGTCGTCGTTTTATCACAAGACCATCAACACCATAAGGCTATAAATCCAAGCTAAAAGGTAGT

                                                    start
 - 76 GATTAACTCCACAAAACCAGAAAAACTACATTTCTAACATATAGAAGAAACAGAGAAAAAGAGAGAGAGACCCCTAATGGCTGAGAAAGAAGAAGTGAAG
                                                      M  A  E  K  E  E  V  K

  25  CTTTTGGGGATATGGGCGAGCCCTTTTAGCCGTCGGGTCGAGATGGCTCTCAAACTCAAAGGCATACCGTACGAGTACGTGGAAGAGATACTGGAGAACA
       L  L  G  I  W  A  S  P  F  S  R  R  V  E  M  A  L  K  L  K  G  I  P  Y  E  Y  V  E  E  I  L  E  N

 125  AAAGCCCTTTGCTTCTTGCTCTTAACCCTATTCACAAGAAAGTCCCTGTTCTTGTCCACAATGGTAAAACCATTCTCGAGTCTCATGTGATTCTTGAATA
       K  S  P  L  L  L  A  L  N  P  I  H  K  K  V  P  V  L  V  H  N  G  K  T  I  L  E  S  H  V  I  L  E  Y
                                                                                                    intron
 225  CATCGATGAGACTTGGCCACAAAATCCAATTCTCCCTCAAGATCCTTATGAAAGATCCAAAGCTCGTTTCTTTGCTAAACTCGTCGATGAACAGgtaatt
         I  D  E  T  W  P  Q  N  P  I  L  P  Q  D  P  Y  E  R  S  K  A  R  F  F  A  K  L  V  D  E  Q

 325  gaattggttcaaaattgcatgtcaaataataaacaattggttctgctttgttaatttatcaaacaagtaatttttctattaacattagcgattatatgtct

            intron
 425  ctgtcattgtagATTATGAACGTGGGGTTTATATCAATGGCAAGAGCAGACGAGAAAGGAAGAGAAGTTTTAGCCGAGCAGGTAAGAGAACTGATTATGT
                  I  M  N  V  G  F  I  S  M  A  R  A  D  E  K  G  R  E  V  L  A  E  Q  V  R  E  L  I  M

 525  ATCTCGAGAAAGAACTTGTCGGAAAAGATTACTTCGGAGGCAAGACTGTCGGATTCTTGGACTTTGTCGCCGGAAGTTTAATTCCGTTTTGTTTGGAGAG
       Y  L  E  K  E  L  V  G  K  D  Y  F  G  G  K  T  V  G  F  L  D  F  V  A  G  S  L  I  P  F  C  L  E  R

 625  AGGTTGGGAAGGAATAGGATTGGAAGTGATTACAGAGGAGAAGTTTCCAGAGTTCAAGAAGATGGGTTAGGAATTTGGAGAAGGTTGAGATTGTTAAAGAT
         G  W  E  G  I  G  L  E  V  I  T  E  E  K  F  P  E  F  K  R  W  V  R  N  L  E  K  V  E  I  V  K  D
                                                                  stop
 725  TGTGTTCCACCAAGAGAGGAACATGTAGAACACATGAACTATATGGCAGAGAGAGTGAGATCTTCTTAAgaaaacaaatcatgtttagttcttgatcatg
       C  V  P  P  R  E  E  H  V  E  H  M  N  Y  M  A  E  R  V  R  S  S  *

 825  caatgtttgtatggttatgttgttgtttattttattgaatatctttgtatgttgtgtggttgagaagtgaggttttatcatcatctctcacgttatctta

 925  tttggtcccagccactatttagaattaatggtaaagctt 963
```

Fig. 2. The genomic sequence of the *At103–1a* gene. The deduced amino acid sequence is shown below the nucleotide sequence. The presumptive TATA box, the translational start and the translational stop are represented in bold. Distances are given in base pairs with respect to the translational start codon (+1). The intron is indicated in small characters. The *as-1*-like element [13] is doubly underlined. The *I* box [8] is singly underlined. The start of the cDNA sequence of pCAT103–1A is indicated by the open arrow.

of the intron was identical in *Arabidopsis* and tobacco. The exons harbour an open reading frame specifying a protein of 224 amino acids. The AT103–1A protein was 67% similar and 48% identical to the tobacco NT103 protein (GST1–1). There was less homology with other auxin-inducible tobacco proteins, NT107 and NT114 (Table 1), indicating that indeed the tobacco NT103 protein was the best possible homolog of the *Arabidopsis* AT103–1A protein.

The coding region of *At103–1b* was isolated from the genomic DNA via IPCR [3], cloned in pBlueScript SK$^+$ leading to plasmid pSK307, and sequenced (results not shown). The encoded protein AT103–1B was 78% identical to the AT103–1A protein, and 43% identical to the tobacco NT103 protein.

The protein encoded by the *At103–1a* gene clearly belongs to the family of type III GSTs and shows homology with other members of this family (NT103, GMHSP26 from soybean [2] and PRP1 from potato [16]) ranging from 46 to 48%. This homology was significantly higher than that with GSTs that were identified in *Arabidopsis* so far, where amino acid identity is only 23% to 27% (Table 1).

The structure of the *At103–1a* gene is very similar to the structure of the tobacco *Nt103* genes [19]. Like the *Nt103* [18], as well as the *prp1* gene [16] and the *Gmhsp26-A* gene [2] the gene contains two exons interrupted by one intron. The structure of other *gst* genes isolated thusfar is completely different. The *gst* genes of wheat [7], *Silene cucubalus* [12] and maize [17] contain three exons interrupted by two introns. The carnation genes *gst1* and *gst2* contain ten exons interrupted by nine introns [10]. Therefore also structural features support the assumption that different classes of *gst* genes exist. The one intron containing *gst* genes may define a specific function in the cell which is different from that determined by other *gst* genes.

In the promoter of *At103–1a*, the putative TATA box was positioned 100 bp upstream of the initiation codon. The promoter did not show extensive homology to the promoter of the *Nt103* gene. However, typically a sequence related to an *as-1* like element [13], as present in the promoters of genes belonging to the auxin-inducible *Nt103* gene family, was present in the promoter of the *At103–1a* gene as well (Fig. 2). Recently we found that these *as-1*-like elements were sufficient to mediate the auxin-responsive transcriptional activation [5]. Also a sequence with homology to a light-responsive element (*I* box) [8] was present in the promoter of the Arabidopsis gene. No *as-1* like or *I* box elements were present in the promoter of the *At103–1b* gene.

Table 1. Percentage of amino acid identity of the protein encoded by the *At103–1a* gene compared to the proteins encoded by the tobacco *Nt103*, [19] *Nt107* and *Nt114* [12] genes, the *Gmhsp26-A* gene of soybean [2] and the *prp1* gene of potato [16]. Also proteins encoded by different GSTs of arabidopsis were compared to the AT103–1A protein. These protein sequences were dervived from cDNAs or genes gst2 [20], PM239x14 [1], ERD11 and ERD13 [11].

NT103 (GST1–1)	48
NT114	38
NT107 (GST2–1)	37
GMHSP26	46
PRP1A	46
ERD13	27
PM239 x 14	26
GST2	23
ERD11	23

Determination of the GST enzyme activity

Based on the homology to GST proteins, it was interesting to test if the protein encoded by the *At103–1a* gene also showed GST activity. In the cDNA clone pCAT103–1A the open reading frame (ORF) of the *At103–1a* gene was cloned in frame with the *lacZ* ORF of the pSK$^+$ expression vector. This resulted in the expression of the AT103–1A protein as a fusion protein in *E. coli*. GST activity was measured as described earlier [4]. The *E.coli* cell extract with the protein encoded by the arabidopsis pCAT103–1A clone showed significant GST activity (1.9 $\pm$ 0.1) compared to the background activity found in extracts from cells containing the empty pSK$^+$ vector (0.1 $\pm$ 0.0). The positive control clone pCNT103 [4]

also provided *E.coli* with significant GST activity (5.6 ± 1.3).

mRNA expression of the At103–1 *genes*

The expression of the *At103–1* genes was studied in seedlings using northern analysis. Total RNA was isolated from the green parts and from the roots of 2-week-old seedlings. The mRNA hybridizing to pSK103, was constitutively present in the green parts of the plants. This may relate to the presence of an *I* box element [8] that was found in the promoter of the *At103–1a* gene (Fig. 2).

In the roots the mRNA was present at a much lower level. Therefore, induction by auxin and other compounds was tested in roots. Figure 3A shows that the mRNA hybridizing to pSK103 was induced by the auxins 2,4-D and NAA, but not by IAA, in roots. The failure of IAA to induce the mRNA could be due to the rapid degradation of IAA in liquid medium. The presence of abscisic acid (ABA) also led to induction of the mRNA hybridizing to pSK103 in roots, but the concentration of hormone needed was 100 times higher than the 2,4-D concentration required for induction. Presence of the cytokinin kinetin at high concentrations also led to a clear increase in the steady state level of mRNA. Other compounds like the inactive auxin 3,5-dichlorophenoxyacetic acid (3,5-D), the auxin transport inhibitor 2,3,5-triiodobenzoic acid (TIBA), glutathione (GSH), gibberellic acid (GA_3), Cu^{2+}, benzoic acid (BA) and salicylic acid (SA) did not induce the mRNA although they were used at concentrations considerably higher than the 2,4-D concentration used.

The mRNA hybridizing to pSK307 was constitutively expressed in green parts as well as in roots (Fig. 3B). Presence of any of the compounds mentioned above did not lead to a significant increase in the steady state level of this mRNA. Invariably, the mRNA levels were compared to those hybridizing with the constitutive *Ubq3ds* gene (Fig. 3C).

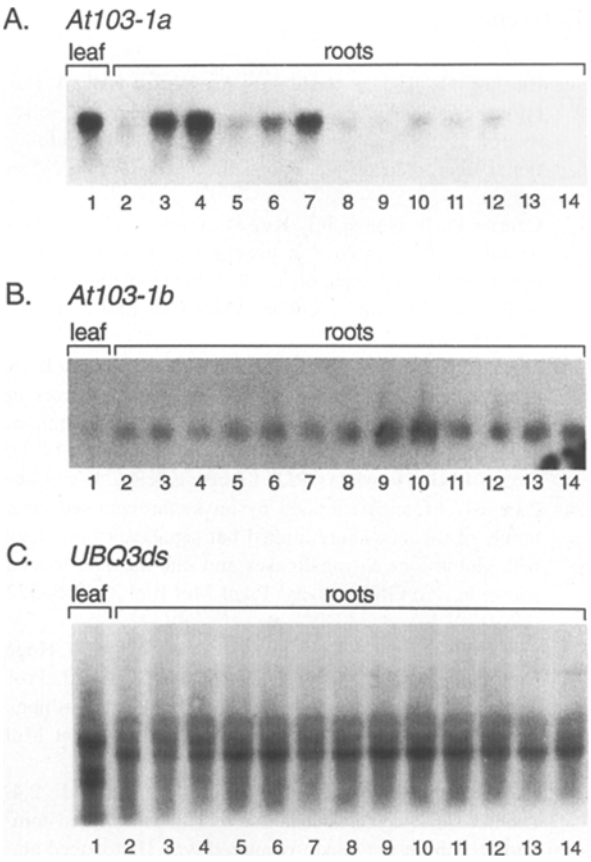

Fig. 3. The mRNA expression in green parts (lane 1) and roots (lane 2) of *Arabidopsis thaliana*. Roots treated with 1 μM 2,4-D (lane 3), 10 μM NAA (lane 4), 1 μM IAA (lane 5), 100 μM kinetin (lane 6), 100 μM ABA (lane 7), 100 μM GA_3 (lane 8), 100 μM benzoic acid (lane 9), 10 μM 3,5-D (lane 10), 100 μM TIBA (lane 11), 1 mM GSH (lane 12), 10 μM $CuSO_4$ (lane 13), 100 μM SA (lane 14). A. Hybridized with the 1 kb *Hin*dIII fragment containing *At103–1a* (pSK103). B. Hybridized to the *Cla*I fragment containing *At103–1b* (pSK307). C. Hybridized to the control probe UBQ3ds.

Acknowledgements

We thank Diane Jofuku and Jack Okamuro (University of California, USA) for the genomic library. We are also grateful to Judy Brusslan and Elaine Tobin (University of California, Los Angeles) for the pUBQ3ds plasmid. The research described in this article was financially supported by the Netherlands Organization for Applied Scientific Research (TNO) and by the European Communities' BIOTECH Programme, as part of the Project of Technological Priority 1993–1996.

References

1. Bartling D, Radzio R, Steiner U, Weiler EW: A glutathione *S*-transferase with glutathione-peroxidase activity from *Arabidopsis thaliana*. Molecular cloning and functional characterization. Eur J Biochem 216: 579–586 (1993).
2. Czarnecka E, Nagao RT, Key JL, Gurley WB: Characterization of *Gmhsp26-A*, a stress gene encoding a divergent heat shock protein of soybean: heavy-metal-induced inhibition of intron processing. Mol Cell Biol 8: 1113–1122 (1988).
3. Does MP, Dekker BMM, de Groot JA, Offringa R: A quick method to estimate the T-DNA copy number in transgenic plants at an early stage after transformation, using inverse PCR. Plant Mol Biol 17: 151–153 (1991).
4. Droog FNJ, Hooykaas PJJ, Libbenga KR and van der Zaal EJ: Proteins encoded by an auxin-regulated gene family of tobacco share limited but significant homology with glutathione *S*-transferases and one member indeed shows *in vitro* GST activity. Plant Mol Biol 21: 965–972 (1993).
5. Droog FNJ, Spek A, van der Kooy A, de Ruyter A, Hoge H, Libbenga KR, Hooykaas PJJ, van der Zaal EJ: Promoter analysis of the auxin-regulated tobacco glutathione *S*-transferase genes *Nt103–1* and *Nt103–35*. Plant Mol Biol, 29: 413–419 (1995).
6. Droog FNJ, Hooykaas PJJ, van der Zaal BJ: 2,4-Dichlorophenoxyacetic acid and related chlorinated compounds inhibit two auxin-regulated type-III tobacco glutathione *S*-transferases. Plant Physiol 107: 1139–1146 (1995).
7. Dudler R, Hertig C, Rebmann G, Bull J, Mauch F: A pathogen-induced wheat gene encodes a protein homologous to glutathione *S*-transferases. Mol Plant-Microbe Interact 4: 14–18 (1991).
8. Giuliano G, Pichersky E, Malik VS, Timko MP, Scolnik PA, Cashmore AR: An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. Proc Natl Acad Sci USA 88: 7089–7093 (1988).
9. Hagen G: The control of gene expression by auxin. In: Davies PJ (eds) Plant Hormones: Physiology, Biochemistry and Molecular Biology, p. 228–245 Kluwer Academic Publishers, Dordrecht, the Netherlands (1995).
10. Itzhaki H, Woodson WR: Characterization of an ethylene-responsive glutathione *S*-transferase gene cluster in carnation. Plant Mol Biol 22: 43–58 (1993).
11. Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K: Characterization of two cDNAs (ERD11 and ERD13) for dehydration-inducible genes that encode putative glutathione *S*-transferases in *Arabidopsis thaliana* L. FEBS Lett 335: 189–192 (1993).
12. Kutchan TM, Hochberger A: Nucleotide sequence of a cDNA encoding a constitutively expressed glutathione *S*-transferase from cell suspension cultures of *Silene cucubalus*. Plant Physiol 99: 789–790 (1992).
13. Lam E, Benfey PN, Gilmartin PM, Fang R-X, Chua N-H: Site-specific mutations alter *in vitro* factor binding and change promoter expression pattern in transgenic plants. Proc Natl Acad Sci USA 86: 7890–7894 (1989).
14. McClure BA, Hagen G, Brown CS, Gee MA, Guilfoyle TJ: Transcription, organisation and sequence of an auxin-regulated gene cluster in soybean. Plant Cell 1: 229–239 (1989).
15. Takahashi Y, Nagata T: *parB*: an auxin-regulated gene encoding glutathione *S*-transferase. Proc Natl Acad Sci USA 89: 56–59 (1992).
16. Taylor JL, Fritzemeier K-H, Häuser I, Kombrink E, Rohwer F, Schröder M, Strittmatter G, Hahlbrock K: Structural analysis and activation by fungal infection of a gene encoding a pathogenesis-related protein in potato. Mol Plant-Microbe Interact 3: 72–77 (1990).
17. Wiegand Rl, Shah DM, Mozer TJ, Harding EI, Diaz-Collier J, Saunders C, Jaworski EG, Tiemeier DC: Messenger RNA encoding a glutathione *S*-transferase responsible for herbicide tolerance in maize is induced in response to safener treatment. Plant Mol Biol 7: 235–243 (1986).
18. van der Zaal EJ, Memelink J, Mennes AM, Ouint A, Libbenga KR: Auxin-induced mRNA species in tobacco cell cultures. Plant Mol Biol 10: 145–157 (1987).
19. van der Zaal EJ, Droog FNJ, Boot CJM, Hensgens LAM, Hoge JHC, Schilperoort RA, Libbenga KR: Promoters of auxin-inducible and root tip-specific expression. Plant Mol Biol 16: 983–998 (1991).
20. Zhou J, Goldsbrough PB: An arabidopsis gene with homology to glutathione *S*-transferase is regulated by ethylene. Plant Mol Biol 22: 517–523 (1993).