The Plant Journal (2005) **41**, 791–800 doi: 10.1111/j.1365-313X.2005.02350.x

The stability of the Arabidopsis transcriptome in transgenic plants expressing the marker genes *nptII* and *uidA*

Souad El Ouakfaoui and Brian Miki*
Bioproducts and Bioprocesses, Research Branch, Agriculture and Agri-Food Canada, Ottawa, ON K1A 0C6, Canada

Received 18 November 2004; accepted 25 November 2004.
*For correspondence (fax 613 759 1701; e-mail mikib@agr.gc.ca).

Summary

The ATH1 Arabidopsis GeneChip from Affymetrix was used to search for transcriptome changes in Arabidopsis associated with the strong expression of transgenes regulated by constitutive promoters. The insertion and expression of the commonly used marker genes, *uidA* and *nptII*, did not induce changes to the expression patterns of the approximately 24 000 genes that were screened under optimal growth conditions and under physiological stress imposed by low temperatures. Approximately 8000 genes (35% of the Arabidopsis genome) underwent changes in gene expression in both wild-type and transgenic plants under abiotic stresses such as salt, dehydration, cold, and heat. This study provides detailed information on the extent of non-targeted or pleiotropic effects of transgenes on plants and shows that the transgenic and non-transgenic plants were equivalent in their global patterns of transcription. This information may help to extend our understanding and interpretation of the principle of substantial equivalence which is used as a first step in the biosafety evaluation of transgenic crops.

Keywords: Arabidopsis, pleiotropic effects, stress response, substantial equivalence, transgene, transcriptome, microarray analysis.

Introduction

The generation of plant transformation technologies created new experimental opportunities to study the genetic control of plant traits. This was quickly followed by the development and commercialization of transgenic crops without a full understanding of the unintended effects on global gene expression by the transgenes. Most of the research addressed the effects of transgenes on targeted processes such as pathways that control quality, productivity, and biosafety. A broader understanding of gene expression that may be altered inadvertently by the insertion and/or expression of transgenes has not been explored (Kuiper *et al.*, 2002). Only recently has the technology been developed to monitor the crosstalk that occurs between signaling pathways that control the fundamental biological processes associated with plant growth, development, and the safety of foods (van Hal *et al.*, 2000).

Some of these questions may be addressed by technologies developed in the areas of genomics. The assessment of microarray technologies has been identified as an important first step (Kuiper *et al.*, 2001). It provides a mechanism for profiling targeted and non-target changes to gene expression. The information gained is very important for adding to our understanding of 'substantial equivalence'. This concept is currently used by all jurisdictions as a first step in evaluating whether a transgenic crop is as safe as its non-transgenic progenitor for production and consumption (Kuiper *et al.*, 2001, 2002).

With the complete sequencing of model plant genomes such as Arabidopsis and rice, it is now possible to evaluate the pleiotropic effects of transgene insertions on global gene expression without omission of a large proportion of the expressed genes. Affymetrix (Santa Clara, CA, USA) produces the ATH1 genome array with about 24 000 Arabidopsis genes represented (Redman *et al.*, 2004; Zhu, 2003). This chip and an earlier version, with about 8000 gene sequences, proved very useful in profiling the changes to gene expression in Arabidopsis induced by nutrients (Maruyama-Nakashita *et al.*, 2003; Wang *et al.*, 2003) and stresses such as salt, osmotic, and cold (Kreps *et al.*, 2002). The Arabidopsis GeneChip therefore represents an ideal vehicle to identify the effects of gene insertions on global gene expression.

Our research objectives were threefold: (i) We wanted to determine the number of genes with altered expression levels induced by the insertion of selectable marker genes (*nptII*) and reporter genes (*uidA*, *uidA-gfp*) that are not believed to alter biological processes in plants (Jefferson *et al.*, 1987; Nap *et al.*, 1992; Stewart, 2001). (ii) We wanted to compare the number of genes affected by transgene insertions with the number of genes affected by common abiotic stresses such as heat, cold, salt, and drought. (iii) Finally, we wanted to determine if the presence of selectable marker genes altered the plant responses to abiotic stresses, as indicated by changes to the global expression pattern which may not have been observed under ideal conditions. As previous studies (Kreps *et al.*, 2002) have demonstrated that plants are continuously in a state of adaptation, a careful experimental plan was required to reduce natural variation in gene expression profiles related to developmental stage, environment stresses, and growth conditions.

Results and discussion

The experimental design

The study of transgene-induced, global changes to transcript profiles in plants requires a carefully planned experimental design that will isolate the influence of the specific transgene insertions under investigation from unrelated effects. The scientific literature shows that a large proportion of the Arabidopsis genome undergoes changes in transcript profiling in response to common environmental factors such as light (30%: Ma *et al.*, 2001), abiotic and biotic stresses (30%: Kreps *et al.*, 2002; 25–30%: Maleck *et al.*, 2001, respectively), oxygen (Klok *et al.*, 2002) or nutrients levels (Maruyama-Nakashita *et al.*, 2003; Thimm *et al.*, 2001; Wang *et al.*, 2003). Furthermore, transcript profiles vary with diurnal and circadian-related changes (Schaffer *et al.*, 2001) and hormone-induced shoot development (Che *et al.*, 2002). It is therefore highly likely that the background level of variation in transcript profiles will be very high among individual plants and could mask the influence of the transgenes unless highly controlled experimental conditions are implemented. As described in Experimental procedures care was taken to control the environmental conditions in growth cabinets and hydroponic cultures using Arabidopsis plantlets that were undergoing vegetative growth (i.e. pre-bolting). Furthermore, the plantlets that were being compared were all grown together in a random pattern and entire plantlets were sampled at the same growth stage to capture changes that may occur in all of the organs. Potential variations due to external and developmental factors were therefore removed to the extent that was possible.

The transgenic lines that were selected for this study expressed bacterial marker genes coding for: (i) neomycin phosphotransferase II (NPTII), (ii) NPTII and β-glucuronidase (GUS) (Malik *et al.*, 2002), and (iii) hygromycin phosphotransferase (HPH) and a GUS/green fluorescent fusion protein (GUS-GFP). High levels of GUS activity were confirmed for the last two lines (data not shown) by histochemical staining for GUS activity (Jefferson *et al.*, 1987) and by the microarray signal log ratio (SLR) (see Supplementary Material). The genes are bacterial in origin and are not known to confer phenotypic alterations or cytotoxicity to plants (Jefferson *et al.*, 1987; Nap *et al.*, 1992; Stewart, 2001). To eliminate variation originating from differences among individual plants, RNA samples were collected from a total of 30 plantlets per line.

Microarray analysis using the Affymetrix Arabidopsis ATH1 genome GeneChip provided quantitative analysis of approximately 24 000 Arabidopsis genes in a highly reproducible and reliable manner (Redman *et al.*, 2004). Meaningful increases or decreases in RNA levels compared with the wild-type control levels were determined using the Affymetrix Microarray Suite software (MAS 5.0) (see 'Statistical Algorithms Reference Guide' at http://www.affymetrix.com for more details). Briefly, the algorithm assesses probe pair saturation, calculates a change in *P*-value and assigns an increase (I), marginal increase (MI), no change (NC), decrease (D), or marginal decrease (MD) to each probe set. The log scale used in this algorithm is base 2; thus, a twofold increase or decrease in the level of a given transcript corresponds to an SLR of 1 or −1, respectively. An SLR of zero corresponds to no change in the expression level. For this study, data was filtered by considering only genes with P call and an SLR >1 (i.e. twofold change) in order to compare our results with those reported in other studies (Kreps *et al.*, 2002; Seki *et al.*, 2002; Wang *et al.*, 2003).

Global gene expression patterns of transgenic plantlets under unstressed growth conditions

Expression of the marker genes, *nptII*, and the reporter, *uidA* or the selectable marker, *hph*, and the reporter *uidA-gfp*, did not confer visible phenotypes on the transgenic plantlets (Table 1) and revealed a small number of genes with altered expression levels (39, 86 and 180 genes, respectively; Table 2). These represented a very small proportion of the Arabidopsis genome (0.17–0.8%; Table 2; Figure 1).

Table 1 Transgene composition and phenotypes of Arabidopsis lines

Genes	Source	Phenotype
35S-nptII-nos	*Escherichia coli Tn5*	None
35S-uidA-nos, *35S-nptII-nos*	*E. coli*, *E. coli Tn5*	None
35S-uidA/gfp-nos, *35S-hph-nos*	*E. coli/A. victoria*, *E. coli Tn5*	None

Table 2 Summary of genes with altered expression in transgenic Arabidopsis lines or in wild-type Arabidopsis under abiotic stresses

	Number of altered genes[a]			Signal log ratio[b]		Percentage of the ATH1 Arabidopsis genome array (22 500 probe sets)
	Total	Increased (I)	Decreased (D)	Maximum increase	Maximum decrease	
Transgenes						
35S-nptII-nos	39	31	8	+5.9	−1.4	0.17
35S-uidA-nos, *35S-nptII-nos*	86	52	34	+5.6	−2.9	0.38
35S-uidA/GFP-nos, *35S-hph-nos*	180	172	8	+4.7	−1.8	0.80
Abiotic stresses						
All four stresses	8084	–	–	–	–	35.42
Salt	4406	2096	2310	+11.2	−8.1	19.58
Drought	3790	1734	2056	+12.0	−9.3	16.84
Cold	3178	1329	1849	+8.8	−7.8	14.12
Heat	1080	518	562	+8.4	−6.5	4.80

Genes are listed in Supplements 1 and 2a.
[a]Number of genes with increased (I) or decreased (D) calls and a signal log ratio greater than 1.
[b]Signal log ratio estimates the magnitude and direction of change of a transcript level.

To determine the background level of 'noisy' genes in our experiments, three independent experiments were performed using four independent transgenic lines that express the *npt II* and *uidA* genes. The lines contained a total of 11 different T-DNA insertions (Table 3). The microarray analyses were performed using MAS 5 normalized data and pairwise comparisons of each transgenic line relative to wild-type across the three experiments were analyzed. The number of genes that changed expression levels by at least 2.5-fold in the transgenic lines relative to wild-type ranged

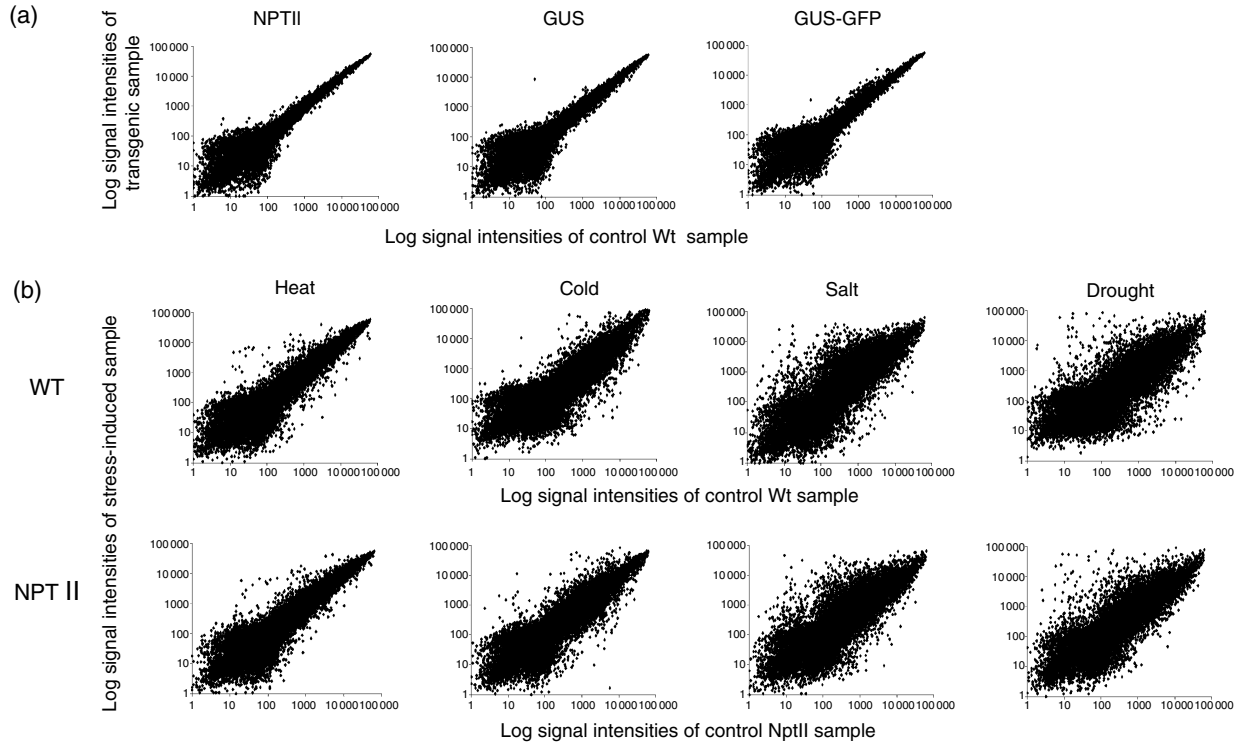

Figure 1. Scatter plots of normalized signal intensities resulting from Affymetrix ATH1 microarray analysis contrasting transgenic Arabidopsis plants (a) with wild-type Arabidopsis plants under abiotic stresses (b).
(a) Changes to global expression in transgenic Arabidopsis lines expressing NPTII; NPTII/GUS; HPH/GUS-GFP compared with wild-type Arabidopsis.
(b) Changes to global expression in wild-type (WT) and transgenic NPTII (NPTII) Arabidopsis plants under heat, cold, salt, and drought stresses compared with their respective controls (WT-unstressed control and transgenic NPTII-unstressed control).

Table 3 Transgenic lines expressing *nptII* and *uidA* genes

Lines	Number of inserts[a]	Expression level[b]
GUS1	1	High
GUS2	3	Low
GUS3	4	Medium
GUS4	3	Medium

[a]The number of inserts was determined by Southern blots probed with a ^{32}P-radiolabeled GUS cDNA fragment. Three restriction enzymes *Eco*RI, *Bam*HI and *Hin*dIII were used separately to digest approximately 5 μg of DNA from T2 seedlings.
[b]The expression level was estimated based on the GUS staining assay results and on the signal intensities obtained for the AFFX-r2-At-U12639 probe set that represents GUS fusion vector pBI101 T-DNA region, used as control sequence on the ATH1 Genechip.

from 18 to 119 genes (Figure 2); however, none of them changed in a reproducible manner in all of the experiments or in all of the lines. We confirmed our results by filtering the genes showing non-significant expression across all samples (15 chips), using ANOVA statistical techniques. At a significance level of 0.01, only one probe set changed among the transgenic lines: AFFX-r2-At-U12639. It represents the GUS fusion vector pBI101 T-DNA region used as a control sequence on the ATH1 GeneChip. The experimental design eliminated potential variation arising from one experiment to another or specific position effects generated in the individual lines. The level of variation in transgenic lines was therefore non-specific and reflected the biological variability inherent to our system. Among the noisy genes, no single gene was specific to any of the four transgenic lines compared with WT.

It is apparent that non-targeted or pleiotropic effects on global gene expression could not be observed with T-DNA insertions of the common selectable marker gene, *nptII* and the reporter gene, *uidA*. We cannot make the same conclusion for the *uidA-gfp* and *hpt* genes without more detailed investigations. The overall data indicates that the stable insertion of T-DNA into the Arabidopsis genome alone did not cause detectable pleiotropic effects among the genes represented on the ATH1 microarray. The possibility remains that small changes would be detected by studying the ORFeome with tiling arrays (Yamada *et al.*, 2003).

Global gene expression patterns of wild-type plantlets under abiotic stresses

In striking contrast to the limited changes induced by transgenes, changes in the global profile of gene expression induced by the abiotic stresses were orders of magnitude greater. As shown in Table 2 and Figure 1, salt, drought, and cold stresses induced changes to the greatest number of genes (4406, 3790, and 3178 genes, respectively); whereas, heat induced changes to a smaller number (1080 genes). Although there was a similar distribution of upregulated genes and downregulated genes, the repressed genes were generally predominant (Table 2).

In total, 8084 genes responded to one or more of the stresses and this represents approximately 35% of the Arabidopsis genome (Table 2). Kreps *et al.* (2002) reported that approximately 30% of the Arabidopsis genome responded to similar abiotic stresses. Their study differed from ours in that they used an earlier version of the Arabidopsis GeneChip from Affymetrix with approximately 8100 genes and they used different growth and sampling conditions. Despite these differences our results are similar. We have confirmed almost all the genes identified by them and others

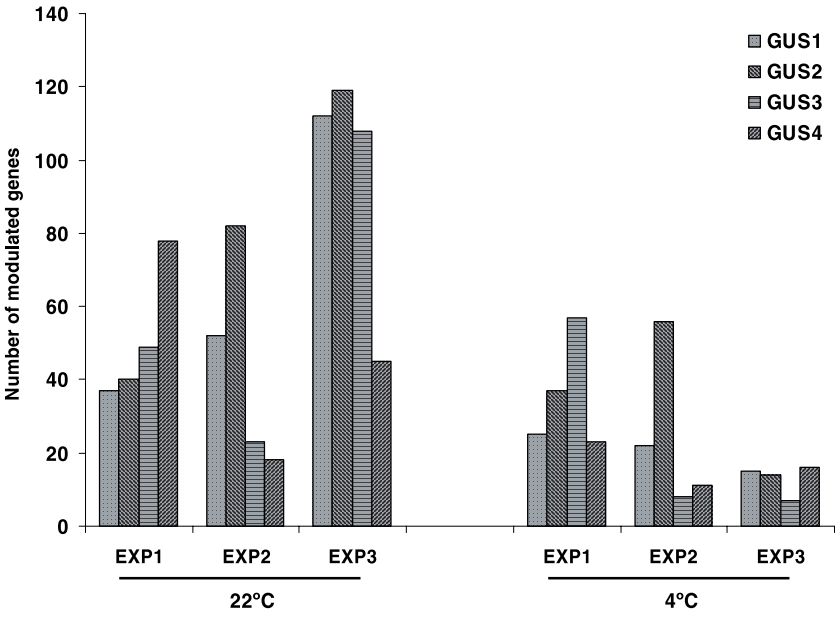

Figure 2. Genes modulated in transgenic plants expressing NPTII and GUS relative to wild-type under optimal growth conditions and under cold stress.
The numbers of modulated genes were obtained from paired comparisons of transgenic lines (GUS1, GUS2, GUS3, and GUS4) relative to wild-type Arabidopsis, under optimal growth conditions (22°C) and under cold stress conditions (4°C) across three experiments (EXP1–3). Full lists of genes are given in Supplement 3.

except for genes that are involved in the initiation of the stress responses (Kreps *et al.*, 2002; Seki *et al.*, 2002). Our study examines more stable changes to gene expression profiles that occur after the initial stress response. Table 4 contains a set of genes previously reported as markers of stress responses (Kreps *et al.*, 2002; Seki *et al.*, 2002) and shows the expression signals observed in this study. The use of the ATH1 GeneChip has allowed us to identify many more indicators of stress, especially downregulated genes, and these are listed in Supplements 2a and 2b. Thus, stress induces reprogramming of a large proportion of the transcriptome.

Influence of NPTII and GUS on the stress response in transgenic plants

Although the insertion of marker genes had little effect on the global patterns of expression in Arabidopsis under optimal growth conditions, it was important to test if differences might appear in the transcriptional reprogramming that accompanies adaptation to abiotic stresses. For this purpose, we have compared the profiles of transgenic NPTII plantlets against WT control plantlets under salt, drought, cold, and heat stresses. Abiotic stresses were applied for durations of 48 h to both WT and NPTII seedlings. During this time, growth continued and could be measured by changes in root length and petiole length (Figure 3a). Under all of the stresses, but particularly cold, salt and drought, the extent of root growth decreased relative to controls. Petiole growth was particularly retarded under salt stress; whereas, heat stress noticeably enhanced the growth of petioles. The presence of the *nptII* gene did not alter these phenotypes and no other obvious differences were observed between WT and NPTII lines at the time of harvesting under abiotic stress. This was confirmed in three independently repeated experiments (data not shown).

Microarray analysis showed that the transgenic NPTII plantlets underwent extensive changes in gene expression under abiotic stress similar to WT plantlets. The changes included a total of 7451 genes representing 33% of the Arabidopsis genome. More specifically, this included 4316, 3791, 2530, and 1069 genes altered by salt, drought, cold, and heat, respectively (Supplement 2b). When compared with the changes in WT plantlets that were grown in parallel,

Table 4 Examples of stress marker genes modulated in transgenic and wild-type Arabidopsis under abiotic stresses

Transgenic lines[a]			Wild-type[a]					
NPTII	GUS	GUS-GFP	Salt	Drought	Cold	Heat	locus ID	Gene name
−0.6 (D)	−1.2 (D)	−0.7 (D)		−0.7 (D)	4.4 (I)	−2 (D)	At2g42530	COR15b
−1.1 (D)	−1.3 (D)	−1.1 (D)	2.1 (I)	5.8 (I)	6.2 (I)	−1.2 (D)	At2g42540	COR15a
−1 (D)	−1.5 (D)	−1.1 (D)	3.3 (I)	5.2 (I)	5.4 (I)	−0.7 (D)	At5g52310	COR 78 (RD29A)
			3.9 (I)	7.7 (I)	3.3 (I)		At5g52300	COR 65 (RD29B)
				2.5 (I)	6.2 (I)		At4g14690	Light-induced protein
			5.6 (I)	1.3 (I)	1.1 (I)	0.6 (I)	At2g03760	Putative steroid sulfotransferase
	−1.5 (D)		2.8 (I)	1.5 (I)	2.2 (I)		At4g25480	DREB1A (CBF3)
	−2.3 (D)	1.3 (I)				0.4 (MI)	At4g25490	DREB1B (CBF1)
−0.5 (D)	−1 (D)	1.3 (I)	3.7 (I)	4.4 (I)	1 (I)		At5g05410	DREB2A
				2 (I)			At5g18450	
			1 (I)				At1g75490	
	−1.2 (D)	1.4 (I)	1.4 (I)			0.8 (I)	At4g25470	DREB1C
	−0.6 (D)		1.8 (I)	3.3 (I)	2.8 (I)	0.5 (I)	At1g20450	LEA
			2.7 (I)	4.8 (I)		0.7 (I)	At2g23110	
	−0.3 (D)		1.8 (I)	3.7 (I)	3 (I)	0.5 (I)	At1g01470	
			6.8 (I)	10.7 (I)			At5g06760	
			5.7 (I)	10.3 (I)			At1g52690	
−0.3 (D)	−0.4 (D)		0.9 (I)	3.2 (I)	2.9 (I)	0.6 (I)	At1g20440	
			4.9 (I)	9.3 (I)	1.6 (I)		At5g66400	
			1 (I)	1.5 (I)	−1.1 (D)		At4g02380	
			4.4 (I)	10.5 (I)			At5g66780	
−0.5 (D)	−0.8 (D)	−0.3 (D)	1.4 (I)	2.6 (I)	3.2 (I)	−0.5 (D)	At5g15960	KIN1
			−2 (D)				At1g59540	KIN2
			1.1 (I)	2.5 (I)		0.9 (I)	At3g18280	Non-specific lipid-transfer protein
0.9 (I)	0.4 (I)		7.2 (I)	8.4 (I)	1.9 (I)	3.6 (I)	At5g59320	
5.9 (I)	5.6 (I)		10.5 (I)	11.3 (I)		3.3 (I)	At5g59310	
			6.5 (I)	7.5 (I)	1.4 (I)	2.9 (I)	At2g37870	
			−5.2 (D)	−5.9 (D)	−7.8 (D)		At1g74670	GAST1
			−2 (D)	−2.9 (D)		−1.2 (D)	At1g63100	Scarecrow

[a]Values are signal log ratio resulting from the Affymetrix ATH1 Arabidopsis microarray analysis of Arabidopsis transgenic lines or wild-type plants under abiotic stresses. Change: increased (I), decreased (D), marginal increase (MI), as determined by MAS 5.0 analysis.

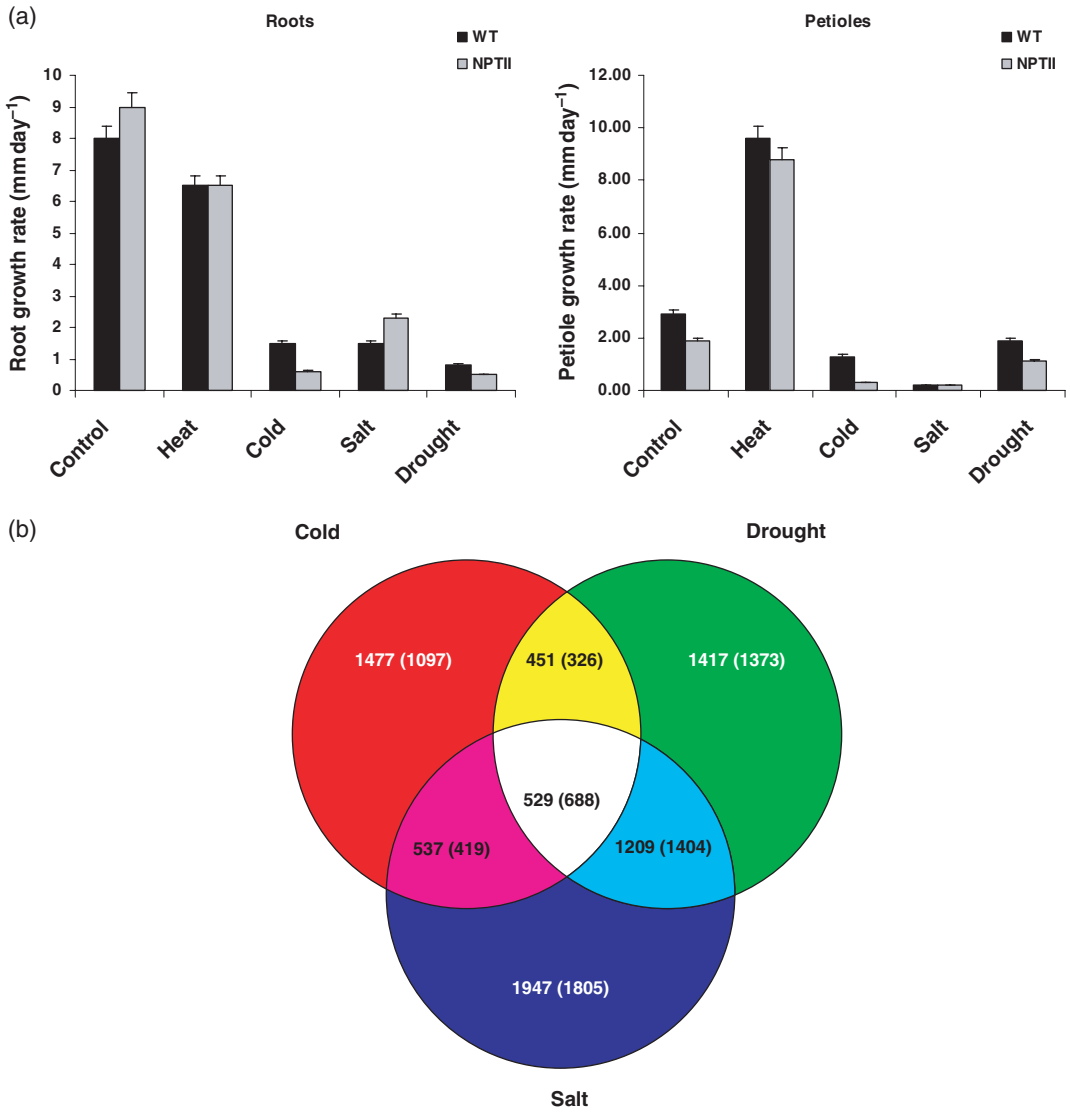

Figure 3. Comparison of wild-type (WT) and transgenic NPTII Arabidopsis plants under abiotic stresses.
(a) Comparison of root and petiole growth rate in wild-type (WT) and transgenic (NPTII) Arabidopsis seedlings. The mean and standard deviation were measured for 10 seedlings under each treatment.
(b) Venn diagram showing the crosstalk among the responses to cold, salt and drought stresses in wild type (WT) and transgenic NPTII plants. The numbers of genes with altered expression in WT plants are indicated and compared with those in NPTII-expressing plants shown in parentheses.

several differences were found in both the numbers of genes as well as the specific genes responding to the stresses (Supplements 2a and 2b). Generally, the number of genes altered in the NPTII plantlets was lower than those altered in the comparable WT plants, except for genes altered by salt stress and those in common with salt and drought stresses (Figure 3b). Salt and drought stresses were also the stresses that share the highest number of common genes indicating the greatest level of crosstalk among the processes involved. Heat stress was omitted in the Venn diagram so that comparisons could be made with results of WT plants in other studies where similar patterns were observed (Seki *et al.*, 2002).

Although there appeared to be a large number of stress-induced changes to expression that were specific to NPTII or to WT plants under all the four stresses (1345 and 1978, respectively), the differences were not reflected in the overall pattern of distribution of the genes among the different functional classes (Figure 4). By this criterion the WT and transgenic NPTII lines were indistinguishable. Many of the stress-responsive genes in NPTII transgenic plants were altered by small amounts relative to WT plants. We hypothesized that the observed differences reflect inherent variability in the experiment.

To study the inherent variability more closely, the same four transgenic lines with the *nptII* and *uidA* genes (Table 3)

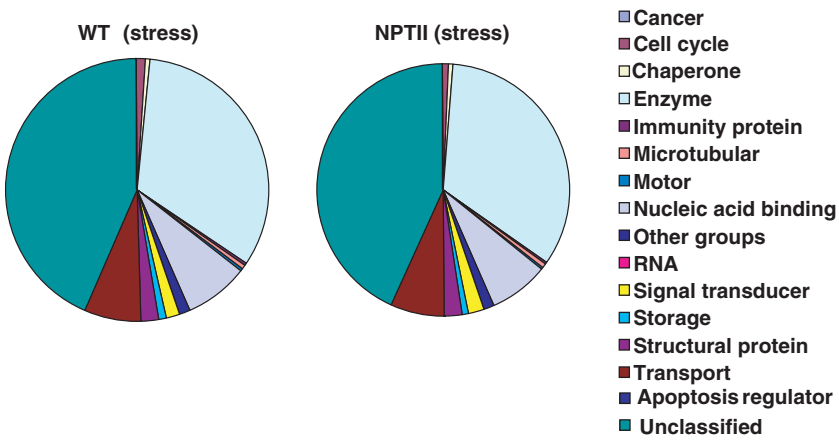

Figure 4. Distribution of the genes with altered expression in wild-type (WT) and transgenic NPTII (NPTII) Arabidopsis plants under stress among the functional classes ontology according to NetAffx (Liu *et al.*, 2003). The total number of genes responding to the four stresses (heat, cold, salt, and drought) is included in the analysis.

were examined in three independent experiments under cold stress at 4°C. The number of genes with altered expression levels in the four transgenic lines relative to the wild-type plants ranged from seven to 57 genes (Figure 2). None of these changes occurred in all lines and all experiments. Moreover, all were eliminated by filtering the genes for non-significant expression across samples using ANOVA either at $P < 0.001$ or at $P < 0.01$. The NPTII/GUS lines responded to cold stress with the same range of changes in gene expression as wild-type (Figure 5). The response is shared by all the lines and the same extent of variation was observed among the transgenic lines when compared with each other or when compared with wild-type plants (Table 5). The data showed that specific non-targeted or pleiotropic changes in expression could not be attributed solely to T-DNA insertion with the *nptII* and *uidA* genes even under cold stress. It is likely to be similar under the other stresses; however, this would need to be confirmed experimentally.

As stress-induced changes to expression profiles include a large proportion of the genome (35% of expressed genes) and stimulate crosstalk among a number of signal transduction pathways it was anticipated that variation in some transgenic lines would be generated by unpredictable factors or events. For example, the number of T-DNA insertions and the location of the insertion sites varied among the lines and could generate different local disturbances in the genome. In Arabidopsis, T-DNA inserts randomly throughout the genome and will create local rearrangements at the genomic insertion site in the majority of events (Forsbach *et al.*, 2003). As the gene density is so high in Arabidopsis (Arabidopsis Genome Initiative, 2000) it is possible that these disturbances can alter gene expression in an insertion-specific manner and cause subsequent downstream effects. In plant species with larger genomes and more intergenic DNA we predicted that this may be less of a consideration. Furthermore, a variety of epigenetic interactions may occur between transgenes and resident genes depending on the specific construct (Qin *et al.*, 2003). Complex inverted repeats can trigger RNA-mediated gene silencing through chromatin remodeling (Mette *et al.*, 2000). Any of these factors could generate variation in individual

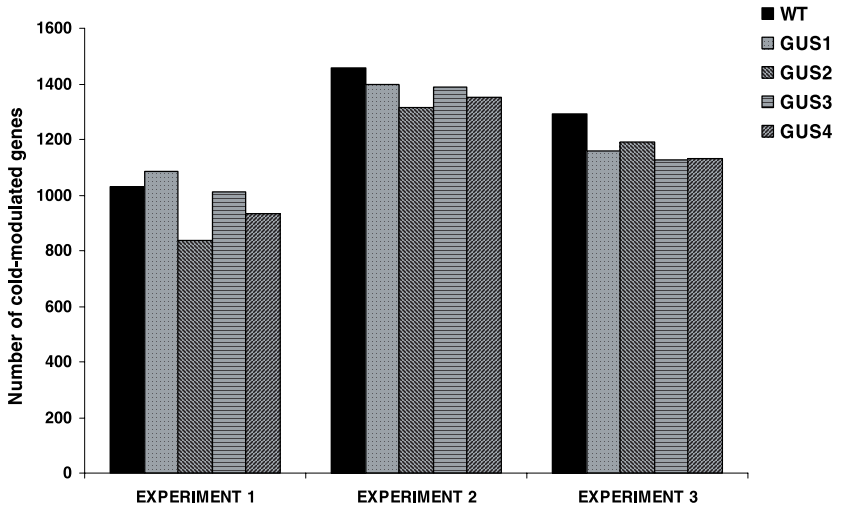

Figure 5. Transgenic plants response to cold stress.
Number of modulated genes obtained from paired comparisons of wild-type and four NPTII-GUS transgenic lines, grown under cold stress conditions (4°C) relative to controls grown under optimal growth conditions (22°C).

Table 5 Selected cold-responsive genes modulated in wild-type (WT) and transgenic lines expressing *nptII* and *uidA* genes (GUS1-4)

		Log_2 (expression ratio treated/control)[a]					
Probe set	Locus	WT	GUS1	GUS2	GUS3	GUS4	Gene description
245306_at	At4g14690	4.74 ± 0.11	5.22 ± 0.29	5.27 ± 0.48	4.69 ± 0.36	5.08 ± 0.21	Chlorophyll A–B binding family protein
263497_at	At2g42540	3.63 ± 0.47	4.2 1 ± 0.86	4.29 ± 0.81	4.16 ± 0.72	4.83 ± 0.53	Cold-regulated protein (cor15a)
248337_at	At5g52310	3.10 ± 0.20	3.47 ± 1.07	3.40 ± 0.79	3.79 ± 1.05	4.00 ± 0.31	Desiccation-responsive protein 29A (RD29A)
252102_at	At3g50970	2.77 ± 0.07	3.23 ± 0.11	3.05 ± 0.16	3.15 ± 0.17	3.02 ± 0.03	Low-temperature-induced protein LTI30 (LTI30)
260556_at	At2g43620	3.59 ± 1.12	3.41 ± 0.80	3.48 ± 0.77	3.28 ± 0.59	3.39 ± 0.30	Chitinase putative
263495_at	At2g42530	3.19 ± 0.27	3.00 ± 0.81	3.19 ± 0.96	3.23 ± 0.91	3.98 ± 0.71	Cold-regulated protein (cor15b)
264436_at	At1g10370	2.83 ± 0.78	3.03 ± 1.00	3.02 ± 1.08	2.83 ± 0.66	2.61 ± 0.65	Glutathione S-transferase putative (ERD9)
266141_at	At2g02120	3.02 ± 0.47	2.94 ± 0.28	2.83 ± 0.16	2.70 ± 0.27	3.02 ± 0.03	Plant defensin-fusion protein putative (PDF2.1)
254085_at	At4g24960	2.05 ± 0.20	2.19 ± 0.06	1.94 ± 0.08	1.98 ± 0.23	2.66 ± 0.02	ABA-responsive protein (HVA22d)
259789_at	At1g29395	2.12 ± 0.28	2.23 ± 0.16	2.51 ± 0.12	2.47 ± 0.23	2.91 ± 0.52	Stress-responsive protein putative
263517_at	At2g21620	1.61 ± 0.18	1.60 ± 0.10	1.64 ± 0.12	1.56 ± 0.13	1.70 ± 0.13	Responsive to dessication protein (RD2)
259516_at	At1g20450	1.97 ± 0.17	2.17 ± 0.30	2.10 ± 0.13	1.91 ± 0.15	2.05 ± 0.06	Dehydrin (ERD10)
259426_at	At1g01470	1.66 ± 0.06	1.63 ± 0.22	1.63 ± 0.23	1.71 ± 0.25	1.77 ± 0.16	Late embryogenesis abundant protein putative (LEA)
256226_at	At1g56280	−2.25 ± 0.07	−2.11 ± 0.27	−1.96 ± 0.17	−2.13 ± 0.16	−2.09 ± 0.23	Drought-responsive family protein
264435_at	At1g10360	−1.69 ± 0.15	−1.85 ± 0.07	−1.51 ± 0.13	−1.66 ± 0.04	−1.56 ± 0.14	Glutathione S-transferase putative
266460_at	At2g47930	−1.87 ± 0.30	−1.89 ± 0.06	−1.68 ± 0.13	−1.98 ± 0.11	−1.80 ± 0.26	Hydroxyproline-rich glycoprotein family protein
260221_at	At1g74670	−7.53 ± 0.61	−8.00 ± 1.36	−6.32 ± 0.14	−8.98 ± 1.61	−10.03 ± 0.23	Gibberellin-responsive protein putative (GAST1)

[a]Expression ratio mean of three biological replicates. Values are log ratio resulting from the Affymetrix ATH1 Arabidopsis microarray analysis of Arabidopsis transgenic lines or wild-type plants under cold stress.

transgenic lines and would probably be found if enough lines were examined but none were uncovered in the four independent transgenic lines in this study generated with simple T-DNA constructs.

Previous studies (Kreps *et al.*, 2002) have demonstrated that caution must be exercised when interpreting the profiles generated in microarray studies. This is because of the tremendous experimental variability and therefore the difficulty in selecting the appropriate times, tissues or conditions for backgrounds needed to observe the specific changes related to specific stimuli. The experimental design in this study allowed us to filter out such effects and to show that changes in gene expression specifically due to non-target effects of transgene insertions were difficult to find even under conditions of physiological stress that were used to induce major transcriptional reprogramming of a large proportion of the genome.

Conclusions and perspective

Under controlled hydroponic growth conditions, transgenic Arabidopsis plantlets expressing marker genes such as *nptII* and *uidA*, were phenotypically similar to untransformed wild-type seedlings and demonstrated no measurable alterations in global gene expression patterns. The data implies very little functional disturbance to the genomes of transgenic plants by the insertion of simple T-DNA constructs. This contrasts with the large number of genes that changed in expression levels in WT or transgenic plantlets under abiotic stresses (heat, cold, salt, and drought). The stress response was not affected by the T-DNA insertions of the *nptII* and *uidA* genes. Our results demonstrate the stability of the Arabidopsis transcriptome in transgenic plants.

This finding validates the numerous experiments that use transgenic plants for the functional analysis of unknown genes on faith that non-target effects of T-DNA with marker genes are not contributing to the results. It indicates that transgenic plants generated with simple T-DNA constructs containing common marker genes are fundamentally equivalent to non-transgenic plants. This in turn implies that differences related to the specific sequences that are inserted or functional changes to the genome originating at the site of T-DNA insertion can be revealed through rigorously designed microarray experiments.

Experimental procedures

Transgenic lines

Wild-type *Arabidopsis thaliana* variety Columbia and three transgenic T2 lines created by *Agrobacterium*-mediated transformation were examined. Plant transformation and selection were performed as previously described (Malik *et al.*, 2002; Wu *et al.*, 2002). The lines were transformed with genes coding for neomycin phosphotransferase II, NPTII (*nptII*), alone or with either β-glucuronidase, GUS (*uidA*). For the line expressing a GUS fusion with the green fluorescent protein GUS-GFP (*uidA-gfp*), the pCAMBIA1303 expressing a gusA-mgfp5-His6 fusion was transformed into Arabidopsis as described (Clough and Bent, 1998) and the hygromycin phosphotransferase gene (*hph*) was used as the selectable marker. T1 seeds were harvested, dried at 25°C, and germinated on sterile medium containing 40 μg ml^{-1} hygromycin to select the transformants. Surviving T1 plantlets were transferred to soil to set T2 seeds. All of the genes were driven by the 35S promoter. The lines were chosen because each expressed high levels of the transgenes,

as indicated by Northern blot analysis (data not shown). The experiment was performed with 14 Genechips: 10 chips for wild-type and NPTII seedlings under four abiotic stresses and four chips for wild-type and transgenic lines NPTII, NPTII/GUS, HPH/GUS-GFP.

Four T2 independent transgenic lines expressing *nptII* and *uidA* genes were used to set the variation baseline in our experimental model. All were characterized by Southern blots and GUS expression was assessed using the histochemical staining assay (Jefferson *et al.*, 1987). Three restriction enzymes profiles were obtained by digesting 5 μg DNA from T2 seedlings separately with *Eco*RI, *Bam*HI and *Hin*dIII. The number of inserts was determined by Southern blots probed with a ^{32}P-radiolabeled GUS cDNA fragment. The T-DNA insertion numbers ranged from 1 to 4, for a total of 11 insertions (Table 3). Three independent experiments were conducted using the four transgenic lines along with wild-type Arabidopsis as a control. RNA was extracted from a total of 30 seedlings of each line grown under optimal growth conditions (22°C) or under cold stress conditions (4°C). Each biological replicate was hybridized to a Genechip for a total of 30 chips (five biological samples, three biological replicate of each, grown under two conditions 22 and 4°C).

Germination and growth of Arabidopsis

Seeds from both wild-type and transgenic Arabidopsis lines were sterilized and germinated on solid agarose medium as described previously (Wu *et al.*, 2002). The seedlings were grown in a Conviron I23L incubator (Conviron, Winnipeg, Canada) for 2 weeks at 22°C, under a mixture of cool-white fluorescent and incandescent lights and a photoperiod of 10 h light/14 h dark to favor vegetative growth. They were then transferred to 200 ml hydroponic cultures, 22°C, in Magenta boxes by floating five seedlings per box on half-strength MS medium (Murashige and Skoog, 1962) using Parafilm with holes punched out to support the plantlets. The boxes were randomly placed in the incubator to minimize experimental errors due to variation in environmental conditions and the plantlets were allowed to adapt to hydroponic growth for 1 week.

The Arabidopsis plantlets were next subjected to varying conditions of abiotic stress by transferring the Parafilm rafts with plantlets to new media. All of the stress treatments were initiated and completed at 10 AM to avoid variation due to circadian clock-related genes. The seedlings were used prior to the bolting stage to avoid the developmentally regulated genes that undergo major changes at this stage of development. The sample size for each treatment was 30 plantlets distributed among six Magenta. The duration of the abiotic stress was 48 h in each case. Drought stress was imposed by preventing contact with the liquid media and the extent of dehydration was monitored by measuring the relative water content (RWC) (Ascenzi and Gantt, 1999). Salt stress was imposed by including 250 mM NaCl in the medium. Heat and cold stresses were imposed at 32 and 4°C, respectively.

Transgenic plantlets were also transferred to new media and allowed to grow for 48 h under normal (unstressed) growth conditions without any additional stress.

Sample preparation

The complete plantlets of each sample were harvested, immediately frozen in liquid nitrogen and stored at −80°C. Total RNA was further extracted using TRIZOL reagent (Invitrogen Life Technologies, Burlington, Canada) and purified using the RNA-easy total RNA cleanup protocol (Qiagen, Valencia, CA, USA). The integrity of the purified RNA was assessed by formaldehyde agarose gel and quantified by absorbance at 260 nm.

Microarray hybridization and analyses

The hybridization and the statistical analysis were performed on a total of 44 ATH1 Arabidopsis GeneChips (Affymetrix) at McGill University and Genome Québec Innovation Centre DNA Microarray Laboratory facility (Montréal, Québec).

The integrity of the RNA samples was assessed by running aliquots of samples on RNA 6000 Nano LabChip (Agilent) using the 2100 bioanalyzer (Agilent). Probe synthesis, hybridization and scanning were performed according to the manufacturer's instructions (Affymetrix). Specifically bound probes were detected by the Agilent GeneArray scanner 2500 (Agilent Technologies, Ville St_Laurent, Canada) and the scanned images were then analyzed using the Microarray Analysis Suite version 5.0 (MAS 5.0) (Affymetrix). The resulting data was filtered using a cut-off of twofold so that our results could be compared with results reported in other studies that also used the Affymetrix Genechip (Kreps *et al.*, 2002). The Venn diagrams were performed using GeneSpring version 5.0 (Silicon Genetics, Redwood City, CA, USA).

To assess the inherent noise to our system, we applied one-way ANOVA statistical analysis involving paired fold-change comparisons (performed at http://genomequebec.mcgill.ca/services/chip.php). Filtration criteria were a fold-change of at least 2.5 and a signal intensity change of 200. In addition, only genes with P calls were considered in theses results. We also used Genowiz™ (Ocimum Biosolutions, Indianapolis, IN, USA) to analyze our data.

Acknowledgements

The authors are grateful to André Ponton, Genome Québec and the Montréal General Hospital, for performing the microarray analysis under contract to the Research Branch, Agriculture and Agri-Food Canada (AAFC); AAFC colleagues, Serge Laberge, Jim Brandle and Thérèse Ouellet and for reviewing the manuscript prior to submission. Special thanks to Mylène Thériault and Chantal Duval for technical assistance. The work was supported through a research contract to AAFC from the Plant Health and Production Division, Plant Biosafety Office, Canadian Food Inspection Agency. ECORC publication number 03_325.

Supplementary Material

The following material is available from http://www.blackwellpublishing.com/products/journals/suppmat/TPJ/TPJ2350/TPJ2350sm.htm

Supplement 1. Lists of genes modulated in transgenic lines NPTII, GUS and GUS-GFP.

Supplement 2a. Lists of genes modulated in wild-type Arabidopsis under abiotic stresses.

Supplement 2b. Lists of genes modulated in NPTII-expressing Arabidopsis under abiotic stresses.

Supplement 3. Lists of noisy genes modulated in NPTII-GUS transgenic lines relative to wild-type Arabidopsis plants under optimal growth conditions (22°C) and under cold (4°C).

References

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**, 796–815.

Ascenzi, R. and Gantt, J.S. (1999) Molecular genetic analysis of the drought-inducible linker histone variant in *Arabidopsis thaliana*. *Plant Mol. Biol.* **41**, 159–169.

Che, P., Gingerich, D.J., Lall, S. and Howell, S.H. (2002) Global and hormone-induced gene expression changes during shoot development in Arabidopsis. *Plant Cell*, **14**, 2771–2785.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743.

Forsbach, A., Schubert, D., Lechtenberg, B., Gils, M. and Schmidt, R.A. (2003) Comprehensive characterization of single-copy T-DNA insertions in the *Arabidopsis thaliana* genome. *Plant Mol. Biol.* **52**, 161–176.

van Hal, N.L.W., Vorst, O., van Houwelingen, A.M.M.L., Kok, E.J., Peijnenburg, A., Aharoni, A., van Tunen, A.J. and Keijer, J. (2000) The application of DNA microarrays in gene expression analysis. *J. Biotechnol.* **78**, 271–280.

Jefferson, R.A., Kavanaugh, T.A. and Bevan, M.W. (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**, 3901–3907.

Klok, E.J., Wilson, I.W., Wilson, D., Chapman, S.C., Ewing, R.M., Somerville, S.C., Peacock, W.J., Doferus, R. and Dennis, E.S (2002) Expression profile analysis of the low-oxygen response in Arabidopsis root cultures. *Plant Cell*, **14**, 2481–2494.

Kreps, J.A., Wu, Y., Chang, H.-S., Zhu, T., Wang, X. and Harper, J.F. (2002) Transcriptome changes for Arabidopsis in response to salt, osmotic and cold stress. *Plant Physiol.* **130**, 2129–2141.

Kuiper, H.A., Kleter, G.A., Noteborn, H.P.J.M. and Kok, E.J. (2001) Assessment of the food safety issues related to genetically modified foods. *Plant J.* **27**, 503–528.

Kuiper, H.A., Kleter, G.A., Noteborn, H.P.J.M. and Kok, E.J. (2002) Substantial equivalence – an appropriate paradigm for the safety assessment of genetically modified foods? *Toxicology*, **181–182**, 427–431.

Liu, G, Loraine, A.E., Shigeta, R., Cline, M., Cheng, J., Valmeekam, V., Sun, S., Kulp, D. and Siani-Rose, M.A. (2003) NetAffx: Affymetrix probesets and annotations. *Nucleic Acids Res.* **31**, 82–86.

Ma, L., Li, J., Qu, L., Hager, J., Chen, Z., Zhao, H. and Deng, X.W. (2001) Light control of Arabidopsis development entails coordinated regulation of genome expression and cellular pathways. *Plant Cell*, **13**, 2589–2607.

Maleck, K., Levine, A., Eulgem, T., Morgan, A., Schmid, J., Lawton, K.A., Dangl, J.L. and Dietrich, R.A. (2001) The transcriptome of *Arabidopsis thaliana* during systemic acquired resistance. *Nat. Genet.* **26**, 403–409.

Malik, K., Wu, K., Li, X.-Q. *et al.* (2002) Constitutive gene expression system derived from the *tCUP* cryptic promoter elements. *Theor. Appl. Genet.* **105**, 505–514.

Maruyama-Nakashita, A., Inoue, E., Watanabe-Takahashi, A., Yamaya, T. and Takahashi, H. (2003) Transcriptome profiling of sulfur-responsive genes in Arabidopsis reveals global effects of sulfur nutrition on multiple metabolic pathways. *Plant Physiol.* **132**, 597–605.

Mette, M.F., Aufsatz, W., van der Winden, J., Matzke, M.A. and Matzke, A.J. (2000) Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J.* **19**, 5194–5201.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassay with tobacco tissue cultures. *Physiol. Plant.* **15**, 473–497.

Nap, J.-P., Bijvoet, J. and Stiekema, W.J. (1992) Biosafety of kanamycin-resistant transgenic plants. *Transgenic Res.* **1**, 239–249.

Qin, H., Dong, Y. and von Armin, A.G. (2003) Epigenetic interactions between Arabidopsis transgenes: characterization in light of transgene integration sites. *Plant Mol. Biol.* **52**, 217–231.

Redman, C.J., Haas, B.J., Tanimoto, G. and Town, C.D. (2004) Development and evaluation of an Arabidopsis whole genome Affymetrix probe array. *Plant J.* **38**, 545–561.

Schaffer, R., Landgraf, J., Accerbi, M., Simon, V., Larson, M. and Wisman, E. (2001) Microarray analysis of diurnal and circadian-regulated genes in Arabidopsis. *Plant Cell*, **13**, 113–123.

Seki, M., Narusaka, M., Ishida, J. *et al.* (2002) Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. *Plant J.* **31**, 279–292.

Stewart, C.N. (2001) The utility of green fluorescent protein in transgenic plants. *Plant Cell Rep.* **20**, 376–382.

Thimm, O., Essigman, B., Kloska, S., Altman, T. and Buckhout, T.J. (2001) Response of Arabidopsis to iron deficiency stress as revealed by microarray analysis. *Plant Physiol.* **127**, 1030–1043.

Wang, R., Okamoto, M., Xing, X. and Crawford, N.M. (2003) Microarray analysis of the nitrate response in Arabidopsis roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. *Plant Physiol.* **132**, 446–567.

Wu, K., Tian, L., Malik, K., Hollingsworth, J., Brown, D. and Miki, B. (2002) Ectopic expression of tomato Pti4 gene activates ethylene response in *Arabidopsis thaliana*. *Plant Physiol.* **128**, 30–37.

Yamada, K., Lim, J., Dale, J.M. *et al.* (2003) Empirical analysis of transcriptional activity in the Arabidopsis genome. *Science*, **302**, 842–846.

Zhu, T. (2003) Global analysis of gene expression using GeneChip microarrays. *Curr. Opin. Plant Biol.* **6**, 418–425.