Lack of repeatable differential expression patterns between MON810 and comparable commercial varieties of maize

Anna Coll[1†], Anna Nadal[2†], Montserrat Palaudelmàs[3], Joaquima Messeguer[3], Enric Melé[3], Pere Puigdomènech[2], and Maria Pla[1*]

[1] INTEA, Universitat de Girona. Campus Montilivi, EPS-I; 17071 Girona, Spain

[2] Departament Genètica Molecular, Centre de Recerca en Agrigenòmica. CSIC-IRTA-UAB. Jordi Girona, 18; 08034 Barcelona, Spain

[3] Departament Genètica Vegetal, Centre de Recerca en Agrigenòmica. CSIC-IRTA-UAB. Ctr. Cabrils; 08348 Barcelona, Spain

[*]Corresponding author: Maria Pla. Institut de Tecnologia Agroalimentària (INTEA), Universitat de Girona. Campus Montilivi, EPS-I; 17071 Girona, Spain. Tel. +34 972 419852; Fax. +34 972 418399; e-mail. maria.pla@udg.edu

[†] AC and AN equally contributed to this work

ABSTRACT

The introduction of genetically modified organisms (GMO) in most countries follows strict regulations to assure that only products that have been safety tested in relation to human health and the environment are marketed. Thus, GMOs must be authorized before use. Complementing more targeted approaches, profiling methods can assess possible unintended effects of transformation. We used microarrays to compare the transcriptome profiles of widely commercialized maize MON810 varieties and their non-GM near-isogenic counterparts.

The expression profiles of MON810 seedlings are more similar to those of their corresponding near-isogenic varieties than are the profiles of other lines produced by conventional breeding. However, differential expression of ~1.7 and ~0.1% of transcripts was identified in two variety pairs (AristisBt/Aristis and PR33P67/PR33P66) that had similar *cryIA(b)* mRNA levels, demonstrating that commercial varieties of the same event have different similitude levels to their near-isogenic counterparts without the transgene.

In the tissues, developmental stage and varieties analyzed, we could not identify any gene differentially expressed in all variety-pairs. However, a small set of sequences were differentially expressed in various pairs. Their relation to the transgenesis could not be ruled out, although this is likely to be modulated by the genetic background of each variety.

Key words: GMO (Genetically Modified Organism), MON810, maize, transcriptome, unintended effects, expression profile

ABBREVIATIONS

cDNA: complementary DNA

CRM: certified reference material

E: efficiency

EBI: European Bioinformatics Institute

EFSA: European Food Safety Authority

EU: European Union

FAO/WHO: Food and Agriculture Organization / World Health Organization

GM: Genetically Modified

GMO: Genetically Modified Organism

IRMM: Institute for Reference Materials and Measurements

ISAAA: International Service for the Acquisition of Agri-biotech Applications

mRNA: messenger RNA

OECD: Organisation for Economic Co-operation and Development

real-time RT-PCR: reverse transcription - real-time polymerase chain reaction

RMA: Robust Multichip Average

rRNA : ribosomal RNA

V2: vegetative two-leaf stage

INTRODUCTION

Genetically modified (GM) crops are subjected to different legislation worldwide to cover aspects of consumer safety and protection. A number of publications (including work performed by the developing companies) show the equivalence of transformed and non-transformed lines of the same species [see reviews in (Cellini et al. 2004; Shewry et al. 2007)]. Risk assessment for approval of new GM crops include field, animal nutrition and basic chemical composition studies (e.g. nutrient, anti-nutrient, allergens) [see the guidelines of the OECD, http://www.oecd.org), the EFSA (2004), the FAO/WHO and the Codex (FAO/WHO, 2001; FAO / WHO, 2002)]. In addition, the sequence of the insert is analyzed as well as the copy number, insertion site, transgene levels of expression and protein accumulation and expected direct and indirect consequences of a functional transgenic protein.

However, targeted approaches have detected some unpredicted differences between transgenic and conventional lines. Saxena and Stotzky (2001) described higher lignin levels in insect resistant transgenic maize than in conventional isogenic lines, and Poerschmann et al (2005) also observed differences in lignin composition. As a consequence, the need for an in-depth study of any unexpected differences among GM and conventional lines by profiling techniques has been suggested by various authors (Cellini et al. 2004; Kok and Kuiper 2003; Millstone et al. 1999) and is currently the focus of a number of research projects. Gene expression profiling technologies are powerful tools to substantially increase the number of targets which can be simultaneously analyzed and allows the study of transcriptional re-programming in various plant species, triggered by a variety of factors. See as examples (Jia et al. 2006; Soitamo et al. 2008; Walia et al. 2006; Zhou et al. 2007). With these technologies detailed information has also been obtained on non-targeted effects of transgenes in several plant species including *Arabidopsis thaliana*, potato, rice and wheat. In these cases unintended variation did not significantly alter overall gene expression and falls within the range of natural variation of landraces and varieties (Baudo et al. 2006; Dubouzet et al. 2007; El Ouakfaoui and Miki 2005; Gregersen et al. 2005; Kristensen et al. 2005), supporting the consideration of transgenic plants as substantially equivalent to non-transformed plants. Proteomic and metabolomic approaches are generally in agreement (Baker et al. 2006; Catchpole et al. 2005; Ioset et al. 2007; Ruebelt et al. 2006)..

Maize is the second most widespread GM crop, after soybean, with a global area of 35,2 million Ha in 2007 [ISAAA, (James 2007)]. In the European Union (EU), apart from Romania, maize is the only GM crop cultured and MON810 (YieldGard®) the single event. Nearly 110,000 ha were grown in 2007 (GMO Compass, http://www.gmo-compass.org/eng/agri_biotechnology/gmo_planting/). MON810 contains a plant expression cassette with the cauliflower mosaic virus 35S promoter and *hsp70* maize intron sequences driving the expression of a synthetic *cryIA(b)* gene. The *cryIA(b)* gene codes for a delta-endotoxin which acts as a potent and highly specific insecticide (van Rie et al. 1989). The approved event has a single copy of the expression cassette with a 3'-truncation partially affecting the coding sequence and resulting in the deletion of the nopaline synthetase terminator (Hernández et al. 2003).

Transgenes are introduced into different commercial varieties through breeding programs to produce the GM plant containing the new traits resulting from transformation. MON810 is one example of this situation. Efficient transformation of maize is commonly achieved using specific inbred lines for efficient regeneration. These include A188 and crosses between A188 and other inbred lines (Ishida et al. 1996), but, as A188 is very poor agronomically, for marketing, lines that contain the transformation event are crossed with several diverse conventional (non-GM) plants to introduce the transgenic insert in selected lines from different breeding programs (Holst-Jensen et al. 2006). This implies that genetic differences between transgenic and near-isogenic varieties may not solely rely on the presence of the transgene but possibly other genomes (besides the near-isogenic line) used for breeding could also contribute to some extent to the final genome of each commercial transgenic variety. Recipient varieties can be largely divergent genetically and are adapted to diverse geographical and climatologic conditions and final uses of the culture (e.g. food or feed) (Serra et al. 2006). In consequence, a number of commercial GM varieties displaying different agronomic properties can be obtained from one single event. In March 2007, 47 MON810 varieties were inscribed in the Common EU Catalogue of Varieties of Agricultural Plant Species and can now be marketed and grown in Member States (GMO Compass, http://www.gmo-compass.org/eng/gmo/db/).

MATERIALS & METHODS

Plant material

Seeds from the following MON810 varieties (company, date of authorization in the Spanish official publication BOE) were used: Aristis Bt (Nickerson Sur / Senasa, 11/03/2003, now commercialized by Limagrain Ibérica), Beles Sur (Limagrain Ibérica, 07/09/2006), DKC6575 (DeKalb, Monsanto Agricultura, 11/03/2003), Helen Bt (Advanta, 11/08/2005, now commercialized by Limagrain Ibérica) and PR33P67 (Pioneer Hi-Bred, 11/03/2003). Their corresponding near-isogenic varieties (Aristis, Sancia, Tietar, Helen and PR33P66) from the same companies were used as well.

Seeds of all five GMO varieties were analyzed to confirm they were MON810. Powdered certified reference material (CRM) for GM maize line MON810 (ref#ERM-BF413A,B,D,F) used as control, was purchased from the Institute for Reference Materials and Measurements (IRMM, Geel, Belgium), commercialised by Fluka (Fluka-Riedel, Geel, Belgium). Genomic DNAs were isolated from 0.2g of plant material using the Nucleospin food kit (Macherey-Nagel Int, Easton, PA) and subsequently subjected to event specific real-time PCR (Hernández et al. 2003) using *hmg* as the endogenous control (Hernández et al. 2005).

Seeds from these maize lines were surface sterilized and germinated in in vitro conditions. Three seeds were sown in glass tins containing 100 ml MS medium (Murashige and Skoog 1962) supplemented with 3% sucrose and 0.7% agar previously sterilized by autoclaving for 20 min at 121 ºC. Seeds were incubated in a in vitro culture chamber at 25± 1ºC with a photoperiod of 16h light/ 8h dark under fluorescent Sylvania Cool White lamps. All plants were simultaneously grown in the same in vitro culture chamber; and glass tins were randomly placed in the plot.

Maize plantlets were harvested at the vegetative two-leaf stage (V2) at the same time of the day, immediately frozen in liquid nitrogen and stored at -80ºC. Each sample consisted of 2 leaves of each of 3 plantlets, without lesions. Three biological replicates were sampled per maize variety.

Total RNA extraction

Total RNA was extracted using a protocol based on the Trizol reagent (Invitrogen Life Technologies, Carlsbad, CA, USA) and purified with the Qiagen RNeasy MiniElute Cleanup Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

RNA concentration was quantified by UV absorption at 260nm using a NanoDrop ND1000 spectrophotometer (Nanodrop technologies, Wilmington, DE, USA). Integrity and purity of the RNA samples were determined by agarose gel electrophoresis and OD 260/280 nm absorption ratios (mean and standard deviation [SD] = 2.06±0.02). All RNA samples had appropriate values.

Microarray hybridization and analyses

The GeneChip® Maize Genome Array (Affymetrix, Santa Clara, CA, USA) was used to search for transcriptome differences between Mon810 transgenic maize varieties and the corresponding isogenic varieties (Aristis Bt vs. Aristis; and PR33P67 vs. PR33P66). The Maize GeneChip has 17,555 probe sets to analyze approximately 14,850 *Zea mays* transcripts, which represent 13,339 genes. It provides comprehensive coverage of over 100 cultivars present in the NCBI UniGene data set (http://www.affymetrix.com/products/arrays/specific/maize.affx).

Three GeneChips were employed to analyze 3 independent replicates per variety. Hybridization and statistical analysis were performed at the Unidad de Genómica, Parque Científico de Madrid. Briefly, the integrity of RNA samples was assessed by capillary electrophoresis using a Bioanalyser 2100 (Agilent Technologies, Palo Alto, California, USA). From 5μg of each RNA sample, double-stranded DNA was synthesized using the One-cycle cDNA Synthesis Kit (Affymetreix) according to the eukaryotic sample processing protocol. The complementary DNA (cDNA) was used as template for in vitro transcription using the GeneChip IVT Labeling Kit (Affymetreix), yielding biotin labeled cRNA. Following cleanup and spectrophotometric quantification, 15μg of the biotinylated target cRNA was fragmented into short sequences (around 100 nt) and used to hybridize to GeneChip Maize Genome Array (Affymetreix) in the GeneChip Hybridization Oven 640 (Affymetreix) for 16 h at 45ºC. Chips were subsequently washed and fluorescently labeled with phycoerythrin using the antibody amplification step in the GeneChip® Fluidics Station 450, and fluorescence was quantified using the GeneChip® 3000 scanner device. The Robust Multichip Average (RMA) software (Irizarry et al. 2003) was

used to extract the data. It includes background adjustment, quantile normalization and summarization. The Venn diagrams were performed using Applet Draws Venn Diagrams (http://theory.cs.uvic.ca/venn/EulerianCircles/).

Reverse transcription and real-time PCR amplifications

The expression of 40 sequences was assayed by reverse transcription - real-time polymerase chain reaction (real-time RT-PCR) to confirm the results of the microarray and for further expression analyses in different maize varieties. Three maize housekeeping genes and the *cryIA(b)* transgene were also analyzed. Reverse transcription was performed on 500ng total RNA, previously treated with Turbo DNase (Ambion, Austin, TX, USA) using 50U of MultiScribe Reverse Transcriptase (Applied Biosystems, Foster City, CA, USA) and random hexamer primers (Applied Biosystems) according to the manufacturer's protocol. For each sample, cDNA was prepared at least in duplicate and the 40 sequences were analyzed with all cDNA preparations. The absence of remaining DNA targets was demonstrated by real-time PCR analyses (see below) of DNase-treated RNA samples.

Real-time PCR assays targeting all 40 sequences selected from the microarrays were developed based on SYBR Green technology. PCR primers were designed using the Beacon Designer 7.0 software (Premier Biosoft International, Palo Alto, CA, USA) and targeted the sequences used for generation of the GeneChip® Maize Genome Array. The same software was used to design a real-time PCR assay targeting 18S ribosomal RNA (rRNA, GenBank Accession # M82384) that was used as a housekeeping gene control. Real-time PCR assays targeting the housekeeping genes β-actin and α-tubulin were developed at Consorci CSIC-IRTA (manuscript in preparation). The *cryIA(b)* assay was designed with TaqMan® technology using the Primer Express™ 3.0 software (Applied Biosystems). The BLAST-N v.2.2.6 tool (National Centre for Biotechnology Information, www.ncbi.nlm.nih.gov) was used to confirm that only the target sequence was recognized. The oligonucleotides, shown in Table 1, were purchased from MWG Biotech AG (Germany). *Insert Table 1*

After optimization of the primers concentrations, SYBR Green QPCR assays were performed in a 20μl reaction volume containing 1X SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA, USA); 300 nM primers [except for *ai8* (50 nM); *ar4* (600 nM); and *ar5, ar8, ar10,*

pr4 and *pr5* (900 nM)] and 1μl cDNA. A two-step experimental run protocol was used: (1) denaturation program (10min at 95ºC); amplification and quantification program (50 repeats of 15s at 95ºC and 1min at 60ºC); and (2) melting curve program (60-95ºC with a heating rate of 0.5ºC/s). The specificity of the PCR was demonstrated by melting curve analysis, which gave single peaks with no primer-dimer peaks or artefacts. The *cryIA(b)* TaqMan real-time PCR assay was optimized as previously described (Rodríguez-Lázaro et al. 2004). TaqMan PCR core reagents (Applied Biosystems, Foster City, CA, USA) were used in a 20μl reaction volume containing 1X PCR TaqMan buffer A (including 5-carboxy-Xrhodamine [ROX] as a passive reference dye); 6mM MgCl2; 200 μm each dATP, dCTP, and dGTP; 400 μm dUTP; 300nM primers; 150nM probe; 1U AmpliTaq Gold DNA polymerase; 0.2U AmpErase uracil N-glycosylase and 1μl cDNA template. Reactions were run with the following program: 2 min at 50ºC, 10 min at 95ºC, and 50 cycles of 15 s at 95ºC and 1 min at 60ºC. All reactions were run on a 7500 Fast Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) and performed in triplicate or duplicate. Linearity (R^2) and efficiency ($E= 10^{[-1/slope]}$), (Rasmussen 2001) of each reaction were compared to the accepted values. The suitability of the housekeeping genes as internal standards was confirmed in our samples through the geNORM v3.4 statistical algorithm, with *M* values below 0.5 in all cases.

Bioinformatics expression analysis

Normalization data and statistical analyses (t-test) were performed using the Genex software v.4.3.1 (MultiDAnalyses). The Benjamini and Hochberg False Discovery Rate multiple testing correction was applied (Benjamini and Hochberg 1995).

RESULTS

The experimental design

Commercial varieties of maize MON10 have the same transgene in different genetic backgrounds. The main objective of this study was to compare gene expression profiles of MON810 and comparative varieties which do not contain MON810. We initially used microarrays to compare the transcriptome patterns of MON810 and near-isogenic varieties for

2 different GM vs. near-isogenic pairs, Aristis Bt vs. Aristis and PR33P67 vs. PR33P66. These varieties were selected to represent phenotypic diversity on the basis of previous agronomic studies in the region of Girona, Spain (Serra et al. 2006): MON810 and near-isogenic pairs showed consistent differences for a number of agronomic parameters. Aristis Bt had different stay-green and number of files per cob compared to Aristis; and PR33P67 and PR33P66 differed in grain humidity and percentage of broken plants at harvest. After validation of microarray data, the differential gene expression pattern of Aristis Bt vs. Aristis was compared to the one of PR33P67 vs. PR33P66. In addition, the two conventional (Aristis vs. PR33P66) and the two MON810 (Aristis Bt vs. PR33P67) lines were also compared.

Our study was extended to other MON810/near-isogenic pairs using real-time RT-PCR on a number of selected transcripts. Helen Bt/Helen, Beles Sur/Sancia and DKC6535/Tietar GM/non-GM pairs were selected to represent commonly used varieties obtained by different seed companies through specific breeding programs. The A188 line was also included in the study.

Environmental factors are known to cause considerable transcriptome changes in plants. The study of changes that may be related to the transgenic character in transcript profiles requires careful experimental design to avoid the effect of unrelated factors. Abiotic and biotic stress, light and nutrient levels cannot be standardized in agricultural or experimental fields. Therefore, although MON810 maize is of major agricultural interest, our approach was based on in vitro cultured plantlets under highly controlled experimental conditions, based on 3 biological replicates independently analyzed in 3 microarrays for each variety (6 microarrays per GM/near-isogenic pair). Real-time RT-PCR analyses were carried out following the same design.

Analysis of microarray data

Microarray data are available at the European Bioinformatics Institute (EMBL-EBI) ArrayExpress repository database under accessions E-MEXP-1464 and E-MEXP-1465. They were independently analyzed for each GM/non-GM pair: Aristis Bt was compared to Aristis;

and PR33P67 was compared to PR33P66. Both for Aristis Bt/Aristis and PR33P67/PR33P66 pairs the data obtained in the 3 replicates were collectively analyzed using the RMA software for gene expression summary values. The estimated $\log_2$-fold changes and log odds values for differential expression produced by the T-test function of the data are shown in Figure 1. The data was subsequently filtered by considering only probes with p-values < 0.05 and at least a twofold increase or decrease in the level of a given transcript. *Insert Figure 1*

Both for Aristis Bt/Aristis and PR33P67/PR33P66 pairs our results revealed a number of genes with altered expression levels. A total of 307 probes (equivalent to 282 genes) were differentially expressed in Aristis Bt and Aristis shoots. This corresponded to approximately 1.7% of probes (or 2.1% genes) assayed. The plot was symmetrically ordered, with 150 probes over-expressed in Aristis Bt and 157 probes down-regulated in GM plantlets. A total of 29 probes displayed differential expression ratios above 5-fold with the highest ratios around 10-fold. Filtering the data for annotated genes gave 67 differentially expressed genes (corresponding to 78 probes), 30 over-expressed and 37 down-regulated in Aristis Bt compared to Aristis (Table 2 shows the annotations of the sequences selected for further analyses). *Insert Table 2*

In contrast, 25 probes (equivalent to 24 genes) were differentially expressed in PR33P67 and PR33P66 leaves, equivalent to around 0.14% of probes (or 0.18% genes) assayed. The plot showed as few as 6 probes over-expressed in PR33P67 and 19 probes down-regulated in GM plantlets. Only 4 down-regulated probes gave higher than 5-fold differential expression, with the highest ratios around 10-fold. Ratios of up-regulation in GM plantlets was from 2 to 2.5-fold. Filtering the data for annotated genes resulted in only 5 differentially expressed genes (corresponding to 5 probes); all of them down-regulated in PR33P67 compared to PR33P66 (see Table 2).

Validation of the microarray data

Real-time PCR has become the most commonly used method for validating microarray data and for gene expression analyses of small sets of genes. After filtering probes with signal intensity differences below 200 fluorescence units, 40 sequences were selected for validation purposes. The Aristis Bt vs. Aristis array results were validated using 30 sequences that

corresponded to those displaying differential expression levels above 5-fold (9 induced and 7 repressed), the best Student test p scores (3 induced and 3 repressed) and 8 randomly selected sequences displaying a minimum of 2-fold differential expression (4 induced and 4 repressed). Similarly, the PR33P67 vs. PR33P66 array results were validated using 13 sequences selected with the same criteria (2 repressed more than 5-fold and 1 induced and 10 repressed down to 2-fold, which were the majority of sequences remaining after filters were applied) (Figure 1 and Table 2). This corresponded to approximately 20% (Aristis Bt/Aristis) and 90% (PR33P67/PR33P66) possible sequences after application of filters. Real-time PCR assays were designed and optimized to target each selected sequence. All assays produced unique amplicons, demonstrated by a single sharp peak in the first derivative plot of dissociation curve analysis. Linearity values were above 0.9; mean R^2 = 0.92 ± 0.04 and efficiency values, above 0.9; mean E = 0.92 ± 0.04.

At least 2 or 3 housekeeping genes need to be used as internal standards for normalization of expression signals (Vandesompele et al. 2002). Here 3 of the most commonly used genes were selected (Jain et al. 2006), 18S rRNA, β-actin and α-tubulin. The stability of expression of these housekeeping genes was assessed in our samples. Total RNAs extracted from biological triplicates of Aristis and Aristis Bt; PR33P67 and PR33P66 plantlets were reverse transcribed with random hexamers and the resulting cDNAs were assayed for 18S rRNA, β-actin and α-tubulin by real-time PCR. Application of the geNORM v3.4 statistical algorithm showed that all 3 housekeeping genes displayed stability measures (*M*) below 0.5, making them suitable internal standards for gene expression under our experimental conditions.

Messenger RNA levels of the 30 selected sequences (*ar1 – ar14* and *ai1 – ai16*) were assessed in biological triplicates of Aristis Bt and Aristis leaflet samples by real-time RT-PCR and the 3 housekeeping genes were used to normalize the data (Table 2). The expression profile in the microarray experiments was confirmed by real-time PCR for all selected genes (mean p = 0.011 ± 0.011, see Table 2) except for *ai5*. Similarly, the 12 selected sequences (*pr1 – pr9*, *ar6*, *ar10* and *pi1*) were assessed in triplicates of PR33P67 and PR33P66 samples by real-time RT-PCR and the data normalized using the 3 housekeeping genes. The expression profile showed in the microarray experiments was confirmed by real-time PCR for all selected genes (mean p = 0.010 ± 0.007) except for *pr2*. Both, *ai5* and *pr2* were excluded from further

316 experiments. According to these results, the degree of coincidence between the microarrays
317 and the real-time RT-PCR was 97 % and 92 % of the sequences assayed, which is within the
318 expected range (Dallas et al. 2005). Additional experiments confirmed the reliability of the
319 microarray results (see below) in which not only genes showing differential expression levels
320 above 5-fold were confirmed but also those displaying down to 2-fold values.

321 The two MON810/near-isogenic pairs exhibited different numbers of differentially expressed
322 sequences. Around 1.7% sequences in Aristis Bt vs. Aristis shoots, sampled at the same
323 developmental stage, were differentially expressed. Under the highly controlled experimental
324 conditions, potential variation due to external and developmental factors was eliminated as far
325 as possible. With PR33P67 and PR33P66, under the same conditions, differential expression
326 was around 10-fold less (Figure 2). Comparison of the transcriptome profiles of Aristis and
327 PR33P66 using the same filtering criteria and the RMA tool showed that they differed in around
328 4% of transcripts (309 up-regulated and 384 down-regulated sequences). As expected, Aristis
329 Bt vs. PR33P67 transgenic varieties differed in around 5% of the sequences (386 up-regulated
330 and 446 down-regulated sequences). These results place the numbers of sequences
331 differentially expressed in GM compared to near-isogenic varieties far below those with altered
332 expression levels comparing conventional varieties (Table 3). *Insert Table 3 and Figure 2*

333 **Comparison of differential transcriptome patterns between the Aristis Bt / Aristis and**
334 **PR33P67 / PR33P66 pairs**

335 To assess the significance of the differential expression patterns between each MON810 variety
336 and its corresponding near-isogenic counterpart we initially compared the transcriptome
337 results obtained for the two pairs of varieties. From the 150 sequences over-expressed in
338 Aristis Bt vs. Aristis, none were significantly different (T-test p values < 0.05) above a 2-fold
339 ratio between PR33P67 and PR33P66. Equally, none of the 6 probes with significant over-
340 expression in PR33P67 vs. PR33P66 were over-expressed in Aristis Bt vs. Aristis (T-test p
341 values < 0.05, at least, 2-fold ratio). This indicates that the over-expressed sequences in
342 MON810 vs. near-isogenic conventional varieties were variety-specific.

343 Many probes displaying down-regulation in one GM variety vs. its near-isogenic counterpart
344 (142 sequences for Aristis Bt vs. Aristis and 6 in PR33P67 vs. PR33P66) were equally expressed

345 in PR33P67 and PR33P66 or Aristis Bt and Aristis, respectively (T-test p values > 0.05, less
346 than 2-fold ratios). Fourteen sequences were more than 2-fold down-regulated in both Aristis
347 Bt vs. Aristis and PR33P67 vs. PR33P66 ($p<0.05$) (Figure 3): just over 9 % of the down-
348 regulated sequences in Aristis Bt vs. Aristis and, remarkably, almost 74 % of those in PR33P67
349 vs. PR33P66. Therefore, 14 out of the 19 sequences down-regulated in the PR33P67/PR33P66
350 pair appear to be a subset of those down-regulated in Aristis Bt/Aristis (which had 150
351 differentially expressed sequences). Four of these sequences were annotated. They correspond
352 to sulphur starvation induced isoflavone reductase-like IRL (*pr6*), adenosine 5'-phosphosulfate
353 reductase 1 (*ar10*), homocysteine S-methyltransferase-1 (pr3) and sulphate transporter ST1
354 (not suitable for real-time RT-PCR analysis according to filtering applied). They appear to be
355 related to S metabolism or transport. *Insert Figure 3*

356 **Differential expression of selected genes in other commercial MON810 vs. near-**
357 **isogenic varieties**

358 The Mon810 transgenic modification has been introduced into many different varieties. The
359 relevance of the differential expression patterns observed between transgenic varieties and
360 the corresponding non-GM near-isogenic lines was assessed by analyzing the expression of
361 the sequences regulated both in Aristis Bt/Aristis and PR33P67/PR33P66 in other MON810 and
362 non-GM varieties, using real-time RT-PCR. Ten out of 14 probes displaying down-regulation in
363 the 2-variety pairs successfully went through the filtering process (see above) and therefore 9
364 real-time RT-PCR assays were used: note that 2 probes corresponded to the same gene (see
365 Table 2).

366 DKC6575, Beles Sur, Helen Bt, Aristis Bt and PR33P67 (MON810) and Tietar, Sancia, Helen,
367 Aristis and PR33P66 (non-GM near-isogenic counterparts) were selected as representative of
368 those authorized in the EU, being produced through independent breeding programs. For each
369 of the 10 varieties, seedlings were grown in vitro and 3 replicates were sampled. The thirty
370 samples were analyzed by real-time RT-PCR using the same 3 internal controls (previously
371 validated in these samples). The results were analyzed using GenEX software. Pair wise
372 comparison of each MON810 variety to its near-isogenic counterpart gave a complex pattern
373 (Table 4). As we have shown, all sequences were regulated in Aristis Bt/Aristis and

PR33P67/PR33P66 pairs. In DKC6575/Tietar there was down-regulation of around 70% of the analyzed sequences, in Beles Sur/Sancia this was around 10% whereas in Helen Bt/Helen none of these 9 sequences were differentially regulated. These results further indicate that different variety pairs have different levels of similitude. *Insert Table 4*

We also assessed the expression pattern of sequences regulated either in Aristis Bt/Aristis or PR33P67/PR33P66, in DKC 6575/Tietar, Beles Sur/Sancia and Helen Bt/Helen. None of the 30 sequences was regulated in 4 pairs and most were exclusively regulated in one single pair. As with the results shown in Table 4, the DKC6575/Tietar pair showed the highest level of gene regulation and only one sequence was identified that was differentially expressed in Helen Bt/Helen.

Using our experimental approach of transcriptomic comparisons of in vitro grown MON810 and near-isogenic leaves no sequence was found to be differentially expressed in all variety pairs tested.

For control purposes, all our GM samples were analyzed to compare the levels of expression of the transgene in the different varieties. A real-time RT-PCR assay was developed and optimized targeting the *cryIA(b)* coding region with performance values of R^2 = 0.998, E = 0.89. Statistical analyses of the results normalized with 18S rRNA, α-tubulin and β-actin messenger RNA (mRNA) levels (ANOVA and Tukey test, $p<0.05$) indicated all 5 varieties expressed similar levels of transgenic mRNA (significance level, 0.497), so discounting any differential expression pattern to be attributable to different *cryIA(b)* mRNA levels among varieties.

DISCUSSION

Maize transgenic event MON810 is widely grown and commercialised as different varieties, which are genetically diverse but they all harbour the same insert at the same chromosomal position. The aim of the present study was to investigate possible transcriptome differences between MON810 commercial varieties and near-isogenic non-GM counterparts. This approach could shed light on possible effects of the transgene and its possible modulation by the genetic

402 background of each GM variety. Authorized events such as MON810 have been submitted to
403 rigorous selection by the developing companies, such that unforeseen effects of the
404 transformation are expected to be minimal.

405 Different transcriptome profiles were found in Aristis Bt vs. Aristis and PR33P67 vs. PR33P66
406 leaves of seedlings grown under experimental conditions which would limit potential variations
407 due to external and developmental factors. As expected, the differentially expressed sequences
408 were a low percentage of the transcriptomes, around 1.7% and 0.1%, respectively. This low
409 number of differentially expressed genes of the GM and their near-isogenic varieties contrasts
410 with the levels of divergence calculated between Aristis and PR33P66 non-GM lines obtained
411 through conventional breeding, with values around 4%, considering sequences with $p<0.05$
412 and rates above 2-fold. Thus, the inserted *cryIA(b)* transgene does not involve consistent
413 major transcriptome modifications, which is further evidence of the equivalence of MON810
414 and non-GM samples.

415 Plant varieties have a wide degree of diversity due in part to the genetic fluidity of plant
416 genomes (Parrott 2005), with extensive variation within a species. Our results are in
417 agreement with a number of studies reporting that unintended variation between GM and non-
418 GM plants has very little impact, particularly when compared to the large differences observed
419 between lines produced by conventional breeding approaches (Baudo et al. 2006; Catchpole
420 et al. 2005; Ioset et al. 2007; Lehesranta et al. 2005; Shepherd et al. 2006) . Different reviews
421 provide an overview of comparative safety assessment of conventional breeding and GM crops
422 (Bradford et al. 2005; Cellini et al. 2004; Chassy et al. 2008; Kok et al. 2008). The occurrence
423 of unintended effects is not a phenomenon specific to genetic modification. Long-accepted
424 plant breeding methods for incorporating new diversity into crop varieties include the use of
425 techniques known to cause genome alteration (e.g. interspecies crosses, tissue culture,
426 chemical or irradiation mutagenesis and the use of transposons). Pleiotropic effects are
427 routinely considered through the process of conventional breeding and varieties having
428 undesired phenotypic traits are discarded. GMOs are less divergent to comparable non-GM
429 lines than varieties obtained by these breeding strategies. As an example, Batista et al (2008)
430 showed that γ-irradiated rice plants had more transcriptomic changes than GMO when
431 compared to the corresponding wild type.

The two MON810 / near-isogenic pairs tested by microarrays displayed different levels of similarity regarding the number of probes and the identity of most sequences showing differential expression. Therefore, MON810 varieties exist (e.g. PR33P67) with more limited transcriptome divergence than their near-isogenic counterpart compared to others (e.g. Aristis Bt). Aristis Bt and PR33P67 showed similar levels of *cryIA(b)* expression under our experimental conditions, indicating that different transgene mRNA levels were not the cause of the different patterns observed in these varieties. Up-regulated sequences appear residual in the more conserved pair (PR33P67/PR33P66, 0.01% of the analyzed sequences) and are up-regulated in just one of these pairs (PR33P67 vs. PR33P66 or Aristis Bt vs. Aristis). The same pattern was observed for most down-regulated sequences, indicating variety-specific regulation. The Aristis Bt/Aristis pair has a leaf phenotype difference (stay-green) and had a greater number of differentially expressed genes in leaf as compared to the PR33P67/PR33P66 pair, which has a root phenotype difference (lodging). In other systems, differences between controls and specific GM lines have often been observed but they also appear to be random and not associated with any specific construct (Baudo et al. 2006; El Ouakfaoui and Miki 2005).

A narrow set of sequences was down-regulated in Aristis Bt/Aristis and PR33P67/PR33P66. Three of them were in silico located on 3 different chromosomes (www.maizegdb.org, Maize Genetics and Genomics Database), demonstrating they do not all belong to a single genome portion physically linked to the transgene. Their expression patterns were analyzed in three other MON810 / near isogenic variety pairs. Although all tested GM varieties had similar *cryIA(b)* mRNA levels we could not identify any sequence consistently repressed in all the MON810 varieties, which suggested that the MON810 transgene does not directly influence the regulation of these sequences. There was no regulation of these sequences in Helen Bt/Helen under our experimental conditions; and only minimal regulation in Beles Sur/Sancia. These results further support that, as observed for Aristis Bt/Aristis and PR33P67/PR33P66, different MON810 / near-isogenic pairs have different levels of divergence, even though we cannot rule out that another set of sequences are differentially regulated in other variety pairs, plant tissues or developmental stages.

Some of the analyzed comparative varieties (e.g. PR33P67/PR33P66) were clearly near-isolines; whereas others (Aristis Bt/Aristis) seemed to be more distantly related. The different

462 companies performed different backcrossings to introduce the MON810 character into
463 commercial varieties. As a result of this process portions of the genome where the MON810
464 transformation occurred and other genomes used along breeding (other than the near-isogenic
465 line) are likely to remain in the transgenic varieties. Variety specific regulation of some
466 sequences might be attributed to these remaining genome portions. We analyzed the
467 expression levels in A188 of the variety specific regulated sequences. There were similar levels
468 of expression, around 1/3 of the sequences in A188 and Aristis Bt, and they were both different
469 from Aristis. These results suggested that the observed changes might derive from portions of
470 conventional genomes used to obtain the transgenic commercial variety. In addition, they
471 proved that these expression levels fall within the range of natural differences between maize
472 varieties. It should be kept in mind that the final pattern of expression of MON810 varieties
473 could also be affected by processes not directly linked to the transgene such as the in vitro
474 culture of transformed cells (Filipecki and Malepszy 2006; Larkin and Scowcroft 1981).

475 Our results are consistent with agronomic differences between the variety pairs previously
476 observed over a number of seasons (Serra et al. 2006), and with recent publications showing
477 differences between particular pairs of MON810 / near isogenic varieties. Statistical differences
478 have been reported in enantiomeric amino acid composition of Aristis Bt/Aristis (% D content
479 of Arg, Ser, and Asp) and PR33P67/PR33P66 (% D content of Arg, Ser, and Ala) but not of
480 Tietar Bt and Tietar (Herrero et al. 2007). Similarly, unexpected metabolic variations involving
481 the primary nitrogen pathway were observed when comparing La73-Bt (MON810) and La73
482 (non-GM) (Manetti et al. 2006). Our results suggest that a comparison between only one
483 variety pair is not useful to infer the effects of the transgene on the general gene expression
484 of maize plants.

485

486 These studies show that gene expression profiles in leaves of MON810 seedlings grown under
487 controlled conditions are more similar to those of near-isogenic varieties than are the profiles
488 of lines produced by conventional breeding. This supports the possibility of producing
489 transgenic maize lines which are substantially equivalent to non-GM counterparts at the level
490 of transcriptomics. Our results show different levels of divergence between various GM and

491 near-isogenic pairs, and suggest the genetic background of each variety influences the
492 divergence. In the tissues, developmental stage and varieties analyzed, we have not identified
493 any gene consistently regulated that can be attributed to the presence of the MON810
494 characteristic.

495

496 **ACKNOWLEDGEMENTS**

497 We specially thank Prof. E. Montesinos (INTEA, UdG) for critically reading of the manuscript.
498 We thank R. Collado (UdG) and J.M. García-Cantalejo (Parque Científico de Madrid) for
499 technical assistance; T. Esteve (CRAG), J. Serra and J. Salvia (E.E.A. Mas Badia) for valuable
500 suggestions. This work was financially supported by the Spanish MEC project with ref.
501 AGL2007-65903/AGR. AC received a studentship from the Generalitat de Catalunya (2005FI
502 00144).

503

504 **REFERENCES**

505

506 Baker JM, Hawkins ND, Ward JL, Lovegrove A, Napier JA, Shewry PR, Beale MH (2006) A
507 metabolomic study of substantial equivalence of field-grown genetically modified
508 wheat. Plant Biotechnol J 4:381-392

509 Batista R, Saibo N, Lourenco T, Oliveira MM (2008) Microarray analyses reveal that plant
510 mutagenesis may induce more transcriptomic changes than transgene insertion. Proc
511 Natl Acad Sci USA 105:3640-3645

512 Baudo MM, Lyons R, Powers S, Pastori GM, Edwards KJ, Holdsworth MJ, Shewry PR (2006)
513 Transgenesis has less impact on the transcriptome of wheat grain than conventional
514 breeding. Plant Biotechnol J 4:369-380

515 Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful
516 approach to multiple testing. J Royal Statistical Soc B 57:289-300

517 Bradford KJ, Van Deynze A, Gutterson N, Parrott W, Strauss SH (2005) Regulating transgenic
518 crops sensibly: lessons from plant breeding, biotechnology and genomics. Nature
519 Biotechnology 23:439-444

520 Catchpole GS, Beckmann M, Enot DP, Mondhe M, Zywicki B, Taylor J, Hardy N, Smith A, King
521 RD, Kell DB, Fiehn O, Draper J (2005) Hierarchical metabolomics demonstrates
522 substantial compositional similarity between genetically modified and conventional
523 potato crops. Proc Natl Acad Sci USA 102:14458-14462

524 Cellini F, Chesson A, Colquhoun I, Constable A, Davies HV, Engel KH, Gatehouse AM,
525 Karenlampi S, Kok EJ, Leguay JJ, Lehesranta S, Noteborn HP, Pedersen J, Smith M
526 (2004) Unintended effects and their detection in genetically modified crops. Food Chem
527 Toxicol 42:1089-1125

528 Chassy B, Egnin M, Gao Y, Glenn K, Kleter GA, Nestel P, Newell-McGloughlin M, Phipps RH,
529 Shillito R (2008) Nutritional and safety assessments of foods and feeds nutritionally
530 improved through biotechnology: case studies. Comp Rev Food Sci Food safety 7:65-
531 74

532 Dallas PB, Gottardo NG, Firth MJ, Beesley AH, Hoffmann K, Terry PA, Freitas JR, Boag JM,
533 Cummings AJ, Kees UR (2005) Gene expression levels assessed by oligonucleotide
534 microarray analysis and quantitative real-time RT-PCR - how well do they correlate?
535 BMC Genomics 6

536 Dubouzet JG, Ishihara A, Matsuda F, Miyagawa H, Iwata H, Wakasa K (2007) Integrated
537 metabolomic and transcriptomic analyses of high-tryptophan rice expressing a mutant
538 anthranilate synthase alpha subunit. J Exp Bot 58:3309-3321

539 EFSA (2004) Guidance document of the Scientific Panel on Genetically Modified Organisms for
540 the risk assessment of genetically modified plants and derived food and feed. EFSA J
541 99:1-94

542 El Ouakfaoui S, Miki B (2005) The stability of the *Arabidopsis* transcriptome in transgenic plants
543 expressing the marker genes nptII and uidA. Plant J 41:791-800

544 FAO/WHO (2001) Evalutaion of allergenicity of Genetically Modified Foods. In: Report of a Joint
545 FAO/WHO Expert Consultation on Foods Derived from Biotechnology. Food and
546 Agriculture Organisation of the United Nations. Available from:
547 http://www.who.int/foodsafety/publications/biotech/en/ec_jan2001.pd
548 Accessed 15 Nov 2007

549 FAO/WHO (2002) Report of the Third Session of the Codex Ad Hoc Intergovernmental Task
550 Force on Foods Derived from Biotechnology (ALINORM 01/34). In: Codex Ad Hoc
551 Intergovernmental Task Force on Foods Derived from Biotechnology. Food and
552 Agriculture Organisation of the United Nations. Available from:
553 ftp://ftp.fao.org/codex/alinorm03/Al03_34e.pdf
554 Accessed 15 Nov 2007

555 Filipecki M, Malepszy S (2006) Unintended consequences of plant transformation: a molecular
556 insight. J Appl Genet 47:277-286

557 Gregersen PL, Brinch-Pedersen H, Holm PB (2005) A microarray-based comparative analysis
558 of gene expression profiles during grain development in transgenic and wild type
559 wheat. Transgenic Res 14:887-905

560 Hernández M, Esteve T, Pla M (2005) Real-time PCR based methods for quantitative detection
561 of barley, rice, sunflower and wheat. J Agric Food Chem 53:7003-7009

562 Hernández M, Pla M, Esteve T, Prat S, Puigdomènech P, Ferrando A (2003) A specific real-time
563 quantitative PCR detection system for event MON810 in maize YieldGard based on the
564 3'-transgene integration sequence. Transgenic Res 12:179-189

565 Herrero M, Ibanez E, Martin-Alvarez PJ, Cifuentes A (2007) Analysis of chiral amino acids in
566 conventional and transgenic maize. Anal Chem 79:5071-5077

567 Holst-Jensen A, De Loose M, Van den EG (2006) Coherence between legal requirements and
568 approaches for detection of genetically modified organisms (GMOs) and their derived
569 products. J Agric Food Chem 54:2799-2809

570 Ioset JR, Urbaniak B, Ndjoko-Ioset K, Wirth J, Martin F, Gruissem W, Hostettmann K, Sautter
571 C (2007) Flavonoid profiling among wild type and related GM wheat varieties. Plant
572 Mol Biol 65:645-654

573 Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP (2003)
574 Exploration, normalization, and summaries of high density oligonucleotide array probe
575 level data. Biostatistics 4:249-264

576 Ishida Y, Saito H, Ohta S, Hiei Y, Komari T, Kumashiro T (1996) High efficiency transformation
577 of maize (*Zea mays* L.) mediated by *Agrobacterium tumefaciens*. Nat Biotechnol
578 14:745-750

579 Jain M, Nijhawan A, Tyagi AK, Khurana JP (2006) Validation of housekeeping genes as internal
580 control for studying gene expression in rice by quantitative real-time PCR. Biochem
581 Biophys Res Commun 345:646-651

582 James C (2007) Global Status of Commercialized Biotech/GM Crops: 2007. ISAAA Briefs 37.
583 ISAAA: Ithaca, NY.

584 Jia JP, Fu JJ, Zheng J, Zhou X, Huai JL, Wang JH, Wang M, Zhang Y, Chen XP, Zhang JP, Zhao
585 JF, Su Z, Lv YP, Wang GY (2006) Annotation and expression profile analysis of 2073
586 full-length cDNAs from stress-induced maize (*Zea mays* L.) seedlings. Plant Journal
587 48:710-727

588 Kok EJ, Keijer J, Kleter GA, Kuiper HA (2008) Comparative safety assessment of plant-derived
589 foods. Regul Toxicol Pharmacol 50:98-113

590 Kok EJ, Kuiper HA (2003) Comparative safety assessment for biotech crops. Trends Biotechnol
591 21:439-444

592 Kristensen C, Morant M, Olsen CE, Ekstrom CT, Galbraith DW, Moller BL, Bak S (2005)
593 Metabolic engineering of dhurrin in transgenic *Arabidopsis* plants with marginal
594 inadvertent effects on the metabolome and transcriptome. Proc Natl Acad Sci USA
595 102:1779-1784

596 Larkin PJ, Scowcroft WR (1981) Somaclonal variation - a novel source of variability from cell
597 cultures for plant improvement. Theor Appl Genet 60:197-214

598 Lehesranta SJ, Davies HV, Shepherd LV, Nunan N, McNicol JW, Auriola S, Koistinen KM,
599 Suomalainen S, Kokko HI, Karenlampi SO (2005) Comparison of tuber proteomes of
600 potato varieties, landraces, and genetically modified lines. Plant Physiol 138:1690-
601 1699

602 Manetti C, Bianchetti C, Casciani L, Castro C, Di Cocco ME, Miccheli A, Motto M, Conti F (2006)
603 A metabonomic study of transgenic maize (*Zea mays*) seeds revealed variations in
604 osmolytes and branched amino acids. J Exp Bot 57:2613-2625

605 Millstone E, Brunner E, Mayer S (1999) Beyond 'substantial equivalence'. Nature 401:525-526

606 Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco
607 tissue cultures. Physiologia Plantarum 15:473-497

608 Parrott W (2005) The nature of change: towards sensible regulation of transgenic crops based
609 on lessons from plant breeding, biotecnology and genomics. In: Proceedings of the
610 17th North American Biothecnology Council, Nahville, Tenn., June 27-29 2005.
611 Available from:
612 http://nabc.cals.cornell.edu/pubs/nabc_17/parts/NABC17_Banquet_1.pdf
613 Accessed 25 March 2008

614 Poerschmann J, Gathmann A, Augustin J, Langer U, Gorecki T (2005) Molecular composition
615 of leaves and stems of genetically modified bt and near-isogenic non-bt maize -
616 Characterization of lignin patterns. J Environ Qual 34:1508-1518

617 Rasmussen R (2001) Quantification on the LightCycler. In: Meuer S, Wittwer C, Nakagawara
618 K (ed) Rapid cycle real-time PCR, methods and applications. Springer-Verlag, Berlin.

619 Rodríguez-Lázaro D, Hernández M, Scortti M, Esteve T, Vázquez-Boland JA, Pla M (2004)
620 Quantitative detection of *Listeria monocytogenes* and *Listeria innocua* by real-time

PCR:assessment of hly, iap, and lin02483 targets and AmpliFluor technology. Appl Environ Microbiol 70:1366-1377

Ruebelt MC, Lipp M, Reynolds TL, Schmuke JJ, Astwood JD, DellaPenna D, Engel KH, Jany KD (2006) Application of two-dimensional gel electrophoresis to interrogate alterations in the proteome of gentically modified crops, 3-Assessing unintended effects. J Agric Food Chem 54:2169-2177

Saxena D, Stotzky G (2001) Bt corn has a higher lignin content than non-Bt corn. Am J Bot 88:1704-1706

Serra J, López A, Salvia J (2006) Varietats de blat de moro genèticament modificades (GM), amb resistència als barrinadors:productivitat i altres paràmetres agronòmics. Dossier Tècnic 10:13-18

Shepherd LV, McNicol JW, Razzo R, Taylor MA, Davies HV (2006) Assessing the potential for unintended effects in genetically modified potatoes perturbed in metabolic and developmental processes. Targeted analysis of key nutrients and anti-nutrients. Transgenic Res 15:409-425

Shewry PR, Baudo M, Lovegrove A, Powers S, Napier JA, Ward JL, Baker JM, Beale MH (2007) Are GM and conventionally bred cereals really different? Trends Food Sci Tech 18:201-209

Soitamo AJ, Piippo M, Allahverdiyeva Y, Battchikova N, Aro EM (2008) Light has a specific role in modulating *Arabidopsis* gene expression at low temperature. BMC Plant Biol. Doi:10.1186/1471-2229-8-13

van Rie J, Jansens S, Hofte H, Degheele D, Mallaert HV (1989) Specificity of Bacillus thuringiensis delta-endotoxins: importance of specific receptors on the brush border membrane of the mid-gut of target insects. Eur J Biochem 186:239-247

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3:1-12

Walia H, Wilson C, Wahid A, Condamine P, Cui X, Close TJ (2006) Expression analysis of barley (*Hordeum vulgare* L.) during salinity stress. Funct Integr Genomics 6:143-156

Zhou J, Wang X, Jiao Y, Qin Y, Liu X, He K, Chen C, Ma L, Wang J, Xiong L, Zhang Q, Fan L, Deng XW (2007) Global genome expression analysis of rice in response to drought and high-salinity stresses in shoot, flag leaf, and panicle. Plant Mol Biol 63:591-608

Table 1. Selected sequences and oligonucleotides used. For housekeeping genes, only those developed in this work are shown.

Accession Number	Code	Forward Primer (5'-3')	Reverse Primer (5'-3')
DOWN-REGULATED SEQUENCES			
BM379705	*ar1*	CTGAGCGGTCATCGGTGTG	GAGGGACATACAATACAACAAGCC
AF056326.1	*ar2*	GTCTCGAGCAATGCCATCCT	CACATACTTGATGACAACGACATGA
CO528265	*ar3*	GGCGTCCTCCATCCAATCC	ATTACTTCTTTCTGCGTGCTACTG
CF623731	*ar4*	TGTCAGTTAAATCACACACTCCAG	CAGCACAGCAAGAGCATTCG
AF133840.1	*ar5*	GCCAATCAGGAGGTGGATCG	CTGGATCAGGATGTCCGACTTC
AY108935.1	*ar6*	CCATTGCCATGTCGTCTTCAG	TCACATCCATATCCATGCTTACAC
AI666020	*ar7*	TAACCCAACCCAACGACATACC	ACATCTGTTCAGTCTACGCCTAG
CF624123	*ar8*	AGCGGTGCTGGTCCCAAG	ACTCATCTCACAAACAACCTTCAG
CK827218	*ar9*	TTCCCTGCCATGATTTTGGTCTC	CGTCTTCCCACTGGATACCCTAG
BM382651	*ar10*	TCTCTGTCAGTCTGTATGATCTTATGTTG	CAATCATTTTGCAGTTACAAAAGCTACA
X54076.1	*ar11*	GATCGTGGCTCTGAAGATGTGG	AACACGCACACCAGAAGCAC
CF632382	*ar12*	AAGCAGCCGTACAAGTTCTCC	TCTTCACCGTGTCGTAAAGCG
AY105790.1	*ar13*	GCAGTTTATCACCACAGAGAAGC	GACACAGACCTGGAGTACGAAC
BM896110	*ar14*	AGCGGCACAACGGGTCTG	GGCGAGTTCTCAAAGCAGTGG
AF297046.1	*pr1*	AATATACTGTTGCGTGTTCTCCTG	GGTTGTATCTCCAAGTTGGATAGC
CA404367	*pr2*	AGTTTTGTATGCTGTGGTTGCTG	AAACGGACACCCAATAGTAGAGC
AF297044.1	*pr3*	ATTTAGGAAGCAAACCAAGAAGGC	TGACCCAAGCACTCAACCG
CO518420	*pr4*	GCAGCAATTCCACTAACCGC	AATTAACCTAAGACATCCCAATTTCCT
CF635310	*pr5*	ATAAGTACCTTTGGATCGAAGAGC	GCTACCTGCTGAGGGAACG
U33318.1	*pr6*	GGAGGAAGCCGTGCTCAAG	GATCTCGAAGCCCGTCTGC
CD438478	*pr7*	GGCAAAGAGGTCGTCTGGAG	AATGGAGCCGTAGCCTGAATAG
AW927712	*pr8*	GTATGTCATCGCCGATAAAACCG	CAGCTCTACACACCCGTCATC
CK144500	*pr9*	CCACACAACACTCCGACCAC	GTACCGTCAGGATAGCAGATTTC
UP-REGULATED SEQUENCES			
AF057184.1	*ai1*	CAAGTGCTCCGCCGACTG	AGGGTCCGACTCCACAAGC
BI431120	*ai2*	TGAAGTGGTTGCTAAAGAGGACTC	ACACCTTTGTTGCGGAGACG
U17351.1	*ai3*	TGTGTCGTGTTCGTGTAGCA	CTTATTCGTTCTGACAGCAGCAG
BM335222	*ai4*	ACCACAACAGCAATCCTTCAAC	AGCAGACTCCTCTTCAGAAACG
BQ539064	*ai5*	GGATCACCCTCATGCTACCG	GCTTCACGCTGTCTTCAATGG
U17350.1	*ai6*	AGCGTCTTATCTTAATTGCCTTGTT	GAGCACAGGTGTGGCATGT
CK371178	*ai7*	CAAGGAGGAGATCAGGGTGGAG	TGTACGCCGCCGAGATGC
BM378406	*ai8*	TGGAAGCACACACCGAGAGG	GGTCGTGTGGTGCTGCTC
CF638013	*ai9*	GGGCAGTGGGCGTCCTTC	ACTTGCGTTGGTAGTGTATCCG
CD219268	*ai10*	TCTCGCAATTCAGTACCGTCAAG	TCCTTCTCAGCAGCCTCGTG
AY639018.1	*ai11*	CTGCTAGTCGTGTTGAAAATCTCG	GCATTTCCACCGCCAAACAG
M33103.1	*ai12*	CGACCGACAGGACCGATT	TGGCGAGGAGGTCTATCCA
D45402.1	*ai13*	TCTCTACCGTGTCCGCAGTC	GCCTAGCAAGCCAAACATATTACC
CK985533	*ai14*	GGACACGCCACCGAGCAG	GAAGCCCTCCGACGACTTG
CD435044	*ai15*	CCAAGCCGTGAAGACACTCTG	CACACAATTTCTACTCTTGACTAGATACTC
CO519322	*ai16*	TCGTCTTCTGCGTGTAATGTCTC	CTCCATGCTTTCCCTGATCTCTAC
CF625331	*pi1*	CAACCTCTGTTTCACACCGTAC	CGCTGCGACGACATCGG
HOUSEKEEPING GENES			
M82384	*h1*	AGAAACGGCTACCACATCCAA	CTACCTCCCCGTGTCAGGATT

680 **Table 2.** Validation of microarray data. A number of sequences (30 for Aristis Bt/Aristis and
681 12 for PR33P67/PR33P66 pairs) with more than 2-fold regulation on microarray experiments
682 ($p < 0.05$) were verified by real-time RT-PCR. Variation folds (microarrays) and significance
683 levels are indicated for each sequence and technique. Values with statistical significance are
684 highlighted. Candidate sequences regulated in the two variety pairs are also highlighted.

DOWN-REGULATED SEQUENCES

Annotation (where available)	Candidate sequences	Aristis Bt vs. Aristis			PR33P67 vs. PR33P66		
		Microarray		RT-QPCR	Microarray		RT-QPCR
		x-fold	T-test p-value	T-test p-value	x-fold	T-test p-value	T-test p-value
	ar1	0.11	0.040	0,030	0.15	0.060	
Myo-inositol 1-phosphate synthase	*ar2*	0.12	0.002	0,005	0.51	0.058	
Phosphoenolpyruvate carboxylase kinase 1	*ar3*	0.12	0.000	0,005	1.30	0.370	
	ar4	0.13	0.007	0,007	1.20	0.800	
Heat shock protein 101	*ar5*	0.18	0.000	0,006	0.80	0.510	
	ar6	0.18	0.012	0,015	0.25	0.040	0,007
	ar7	0.19	0.011	0,006	0.80	0.430	
	ar8	0.22	0.004	0,016	0.82	0.500	
Invertase	*ar9*	0.23	0.002	0,006	0.59	0.030	
Adenosine 5'-phosphosulfate reductase 1	*ar10*	0.26	0.004	0,007	0.34	0.020	0,000
18kDa heat shock protein	*ar11*	0.33	0.050	0,048	0.82	0.220	
Cinnamoyl CoA reductase	*ar12*	0.34	0.000	0,010	0.85	0.441	
	ar13	0.23	0.000	0,006	1.16	0.554	
	ar14	0.27	0.000	0,006	0.60	0.060	
Homocysteine S-methyltransferase-3	*pr1*	0.89	0.124		0.46	0.003	0,007
	pr2	0.70	0.005		0.45	0.001	0.169
Homocysteine S-methyltransferase-1	*pr3*	0.43	0.002		0.43	0.011	0,008
	pr4	0.26	0.012		0.40	0.048	0,011
	pr4	0.30	0.034		0.40	0.037	0,011
	pr5	0.32	0.004		0.37	0.011	0,007
Sulfur starvation induced isoflavone reductase-like IRL	*pr6*	0.26	0.001		0.25	0.032	0,008
	pr7	0.22	0.000		0.25	0.000	0,022
	pr8	0.24	0.039		0.20	0.020	0,008
	pr9	0.35	0.001		0.12	0.034	0,009

UP-REGULATED SEQUENCES

	Candidate sequences	Aristis Bt vs. Aristis			PR33P67 vs. PR33P66		
		Microarray		RT-QPCR	Microarray		RT-QPCR
		x-fold	T-test p-value	T-test p-value	x-fold	T-test p-value	T-test p-value
Trypsin inhibitor	*ai1*	25.30	0,014	0,000	1.53	0.007	
	ai2	9.53	0,000	0,027	0.84	0.503	
Thiamine biosynthetic enzyme (thi1-2)	*ai3*	8.05	0,004	0,022	0.74	0.698	
	ai4	7.72	0,000	0,008	1.02	0.940	
	ai5	6.83	0,001	0.278	0.69	0.070	
Thiamine biosynthetic enzyme (thi1-1)	*ai6*	6.69	0,002	0,018	0.88	0.784	
	ai7	6.18	0,003	0,016	1.00	0.999	
	ai8	5.99	0,010	0,007	0.82	0.713	
	ai9	5.73	0,005	0,008	1.85	0.034	
	ai10	4.72	0,000	0,000	0.75	0.643	
Sucrose transport protein (SUT2)	*ai11*	3.87	0,000	0,009	1.20	0.263	
Catalase-3 (AA 1-495)	*ai12*	2.79	0,003	0,027	1.07	0.574	
Cysteine proteinase	*ai13*	2.25	0,000	0,007	0.98	0.634	
	ai14	2.81	0.000	0,000	1.10	0.645	
	ai15	2.60	0.000	0,010	0.92	0.048	
	ai16	2.88	0.000	0,000	1.28	0.004	
	pi1	0.54	0.225		2.38	0.000	0,027

685

686

687 **Table 3.** Numbers and percentages of statistically significant differentially expressed
688 sequences in pair wise comparisons of 2 MON810 and 2 near-isogenic maize varieties.

689

Lines used for comparison	No.	%
690 Aristis Bt (MON810) / Aristis (non-GM)	307	1.75
691 PR33P67 (MON810) / PR33P66 (non-GM)	25	0.14
692 Aristis (non-GM) / PR33P66 (non-GM)	693	3.94
Aristis Bt (MON810) / PR33P67 (MON810)	832	4.74

693

694

695 **Table 4.** Expression patterns of sequences selected to show down-regulation both in Aristis
696 Bt/Aristis and PR33P67/PR33P66 in 5 MON810 commercial varieties vs. their corresponding
697 non-GM isogenic lines. P-values obtained by paired comparison of each GM vs. its near-isogenic
698 variety pair are indicated. Significant values ($p<0.05$) indicating down-regulation are
699 highlighted.

	pr6	***pr3***	***pr5***	***pr7***	***pr9***	***ar6***	***ar10***	***pr4***	***pr8***
Aristis Bt / Aristis	0.010	0.047	0.009	0.055	0.018	0.018	0.008	0.037	0.047
PR33P67 / PR33P66	0.011	0.011	0.009	0.004	0.009	0.010	0.002	0.014	0.009
DKC 6575 / Tietar	0.034	0.007	0.008	0.026	0.015	0.012	0.057	0.092	0.080
Beles Sur / Sancia	0.042	0.279	0.309	0.150	0.178	0.590	0.447	0.613	0.512
Helen Bt / Helen	0.739	0.599	0.787	0.208	0.785	0.706	0.751	0.517	0.810

700

701

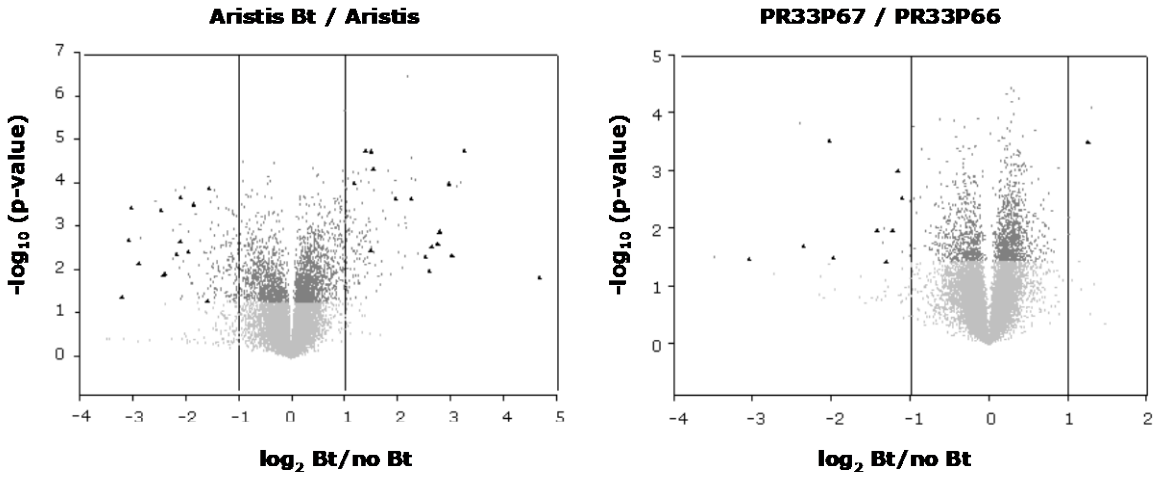
FIGURE LEGENDS

Figure 1. Changes in gene expression in MON810 vs. near-isogenic maize lines Aristis Bt vs. Aristis and PR33P67 vs. PR33P66. Each point represents one gene in the maize Affymetrix microarray. The log odds for differential expression of all genes, estimated from the RMA analysis of the data were plotted against the estimated log2 fold changes. Thus, a twofold increase or decrease in the level of a given transcript corresponds to 1 or -1, respectively. Bold, sequences further analyzed by real-time RT-PCR.

Figure 2. Venn diagrams representing differential gene expression in Aristis Bt/Aristis, PR33P67/PR33P66 and Aristis/PR33P66 based on microarray analysis. To allow clear visualization of the data, the areas representing differentially expressed sequences ($p < 0.05$, at least, 2-fold difference) are proportional to the amount of sequences; and those representing sequences with similar expression values were drawn 10-fold smaller than it would correspond.

Figure 3. Schematic representation of differential gene expression in Aristis Bt/Aristis and PR33P67/PR33P66 based on microarray analysis. Differentially expressed sequences ($p < 0.05$, at least, 2-fold difference) are represented in green (down-regulated in MON810 varieties) and red (up-regulated in MON810 varieties). Sequences displaying similar expression values in GM and non-GM varieties (around 99% sequences analyzed) are not included.

Figure 1


Aristis Bt / Aristis
-log10 (p-value)
log2 Bt/no Bt
PR33P67 / PR33P66
-log10 (p-value)
log2 Bt/no Bt

737 **Figure 2**

738

739

740

741

742

743

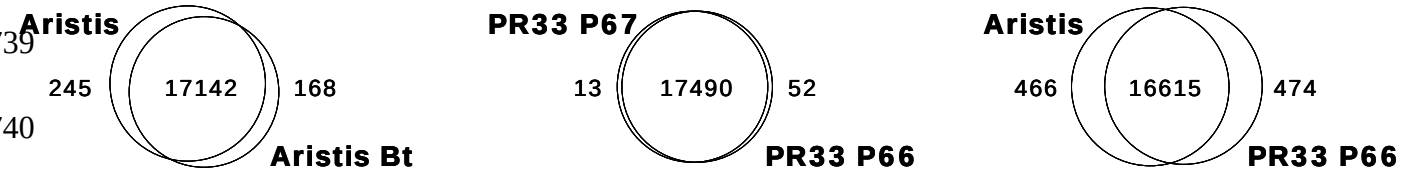
744

745

746

747

748



749 **Figure 3**

750

751