Application of Two-Dimensional Gel Electrophoresis To Interrogate Alterations in the Proteome of Gentically Modified Crops. 3. Assessing Unintended Effects

ARTICLE *in* JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY · APRIL 2006

Impact Factor: 2.91 · DOI: 10.1021/jf052358q · Source: PubMed

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

J. Agric. Food Chem. **2006,** *54,* 2169−2177

Application of Two-Dimensional Gel Electrophoresis To Interrogate Alterations in the Proteome of Gentically Modified Crops. 3. Assessing Unintended Effects

MARTIN C. RUEBELT,*,§ MARKUS LIPP,§ TRACEY L. REYNOLDS,§ JON J. SCHMUKE,§ JAMES D. ASTWOOD,§ DEAN DELLAPENNA,# KARL-HEINZ ENGEL,† AND KLAUS-DIETER JANY‡

Product Safety Center, Monsanto Company, 800 North Lindbergh Boulevard, St. Louis, Missouri 63167; Department of Biochemistry and Molecular Biology, 212 Biochemistry, Michigan State University, East Lansing, Michigan 48824; Center of Food and Life Sciences, Chair of General Food Technology, Technical University of Munich, Am Forum 2, 85350 Freising-Weihenstephan, Germany; and Federal Research Centre for Nutrition and Food, Haid-und-Neu-Strasse 9, 76131 Karlsruhe, Germany

The current procedures to assess the safety of food and feed derived from modern biotechnology include the investigation of possible unintended effects. To improve the probability of detecting unintended effects, profiling techniques such as proteomics are currently tested as complementary analytical tools to the existing safety assessment. An optimized two-dimensional gel electrophoresis (2DE) method was used as a proteomics approach to investigate insertional and pleiotropic effects on the proteome due to genetic engineering. Twelve transgenic *Arabidopsis thaliana* lines were analyzed by 2DE, and their seed proteomes were compared to that of their parental line as well as to 12 *Arabidopsis* ecotype lines. The genetic modification of the *Arabidopsis* lines, using three different genes and three different promoters, did not cause unintended changes to the analyzed seed proteome. Differences in spot quantity between transgenic and nontransgenic lines fell in the range of values found in the 12 *Arabidopsis* ecotype lines or were related to the introduced gene.

KEYWORDS: Two-dimensional gel electrophoresis; *Arabidopsis thaliana*; seed proteome; unintended effect; substantial equivalence; GMO

INTRODUCTION

The assessment of the safety of genetically modified (GM) crops aims to address both intended and unintended effects as a result of the genetic modification. Unintended effects can be the result of secondary or pleiotropic effects of the transgene expression and insertional effects resulting from transgene integration into plant genomes (*1*−*3*). The introduction of exogenous DNA sequences into the plant genome is a random process leading to physical disruption in the genome and possible inactivation of endogenous genes. Activation of silent genes and formation of fusion proteins by transcriptional read-through processes are also possible (*1*), although these are routinely minimized during the process of transgenic product development by sequencing the insertion site of transgenes and selecting those that are not in known genes or inserted near promoters. In addition, the introduced gene(s), the gene product, or the changed biochemical pathway may interact with the regulation of other genes or biochemical pathways. Worldwide, there has been no verifiable unintended toxic or nutritionally harmful effects resulting from the cultivation and consumption of products from GM crops (*4*, *5*) after 10 years of consumption. The occurrence of unintended effects is not unique to the application of recombinant DNA techniques. It is a phenomenon that occurs frequently in traditional breeding, for example, due to hybridization [potato breeding lines with novel, toxic glycoalkaloids (*6*)], natural genetic recombination, natural chromosomal rearrangements (translocations and inversions), activity of transposable elements in plant genomes, cell fusion, or chemical and radiation induced mutations (*1*, *7*).

To improve the probability of detecting unintended effects, profiling techniques such as proteomics are currently tested as analytical tools complementary to the existing safety assessment (*8*−*10*). Profiling technologies allow the simultaneous measurement and comparison of thousands of plant components without prior knowledge of their identity. The combination of these nontargeted approaches is considered to facilitate a more comprehensive approach than the targeted methods and, thus, provide additional opportunities to identify unintended effects. Three main cell constituent groups are targeted by these

* Corresponding author (e-mail m.ruebelt@gmail.com).
§ Monsanto Company.
# Michigan State University.
† Technical University of Munich.
‡ Federal Research Centre for Nutrition and Food.

technologies: RNA, metabolites, and proteins. Because proteins are direct products of gene transcription and translation, they are ideally suited for the detection of changes in the genome (e.g., insertional mutation), in gene regulation (pleiotropic effect), or in biochemical pathways (direct or pleiotropic effect) of a genetically modified plant. The expected changes in the plant proteome are similar to naturally occurring mutations (*11*) and can lead to changes in the net charge (isoelectric point), the molecular weight, and/or the quantity of the affected protein.

To investigate insertional and pleiotropic effects due to genetic engineering, the seed proteomes of 12 transgenic (TG) *A. thaliana* lines were analyzed by two-dimensional electrophoresis (2DE) according to the method of Ruebelt et al. (*12*) and compared to the seed proteome of the wild-type line. The transgenic *A. thaliana* lines contain an inserted β-glucuronidase (*gus*) gene, an inserted *p*-hydroxyphenylpyruvate dioxygenase (*hppd*) gene, or an inserted γ-tocopherol-methyltransferase (γ-TMT) gene. These lines were chosen because they represent two different strategies: (i) no change to an endogenous metabolic pathway (*gus* gene) and (ii) change of an endogenous biochemical pathway (*hppd* and γ-TMT genes). The data generated from the previous study on natural variability (*11*) are used to discuss the relevance of potential differences between the transgenic lines and their nontransgenic parental lines. This study will address the feasibility of proteomics technology to identify unintended or intended changes in the seed proteome due to genetic engineering.

MATERIALS AND METHODS

Plant Material. Seeds for six transgenic *Arabidopsis* lines (T3 generation) containing an inserted β-glucuronidase (*gus*) gene and their parental line (WT Col-0) were provided by Monsanto Co., St. Louis, MO. The transgenic *Arabidopsis* events were obtained by *Agrobacterium*-mediated transformation with a T-DNA containing the β-glucuronidase (*gus*) gene and the neomycin phosphotransferase II (*nptII*) gene for antibiotic selection. The GUS construct contains the enhanced 35S promoter, derived from cauliflower mosaic virus (E35S), and a transcription terminator sequence (E9). The *nptII* gene is driven by a nopaline synthase promoter (pNOS) and is followed by the 3′ nontranslated region of the nopaline synthase gene from *Agrobacterium tumefaciens* (3′NOS). The lines have been advanced as single seeds. The six transgenic lines were selected because they contain one gene copy, are homozygous for the transgene, and have different transgene expression rates in leaf tissues.

Seeds for six tocopherol-enhanced transgenic *Arabidopsis* lines and their parental line (Col-0) were obtained from Michigan State University (D.D.). The transgenic *A. thaliana* lines contain an inserted *p*-hydroxyphenylpyruvate dioxygenase (*hppd*) gene, or an inserted γ-tocopherol-methyltransferase (*gtmt*) gene and have been previously described (*13*, *14*). The expression of the transgene *hppd* is controlled by a constitutive cauliflower mosaic virus (CaMV) 35S promoter (35S:HPPD-2 and -3) and by a seed-specific promoter DC3 (DC3:HPPD-3 and -8). The expression of the transgene *gtmt* is controlled in both TG lines (35S:gTMT-18 and -49) by the CaMV 35S promoter.

The T3 generation of the 6 transgenic GUS lines was grown side-by-side with the parental line (Col-0) and 12 *Arabidopsis* ecotypes (previously discussed in ref *11*). The 6 tocopherol-enhanced transgenic *Arabidopsis* lines were grown side-by-side with the parental line (Col-0) and the interexperimental control line (Col-0). The parental line and the interexperimental control line both represent the same *Arabidopsis* ecotype, Columbia (Col-0), but originate from two different laboratories. Plants were grown in individual 2.5-in. pots (Metro-Mix 200 soil; Hummert International, Earth City, MO) in growth chambers at 20 °C and 70% relative humidity with 16 h of light (150−200 μeinstein/s/m^2). For each line, 10−12 replicates were planted and randomly distributed in the growth chamber to limit the influence of environmental factors. The plants were watered and fertilized (100 ppm of Peter's 20:20:20) twice weekly via subirrigation. Seeds were harvested after complete maturity of all seeds on a plant. To avoid seed deterioration, seeds were placed in freezer bags containing desiccant (Drierite Anhydrous Calcium Sulfate) and stored at 4 °C.

GUS Expression Analysis. DNA and RNA were extracted from *Arabidopsis* leaf tissue using the ABI 6700 nucleotide purification system per the manufacturer's recommendations (Applied Biosystems, Foster City, CA). Copy number of the transgenic insertion in the genomic DNA was determined using quantitative Taqman PCR. An assay designed to amplify a region of the E9 3′utr was designed and validated to known Southern-confirmed controls of one, two, and four copies. The primers used in the reaction were a forward primer, 5′-CAACGTTCGTCAAGTTCAATGC, a reverse primer, 5′-TGCCATAATACTCGAACTCAGTAGGA, and the labeled Taqman probe, 5′-6fam-TCAGTTTCATTGCGCACACACCAGAA-tamra. The sequence for ctpA2 DNA was used as an endogenous control. The primers used for ctpA2 gene were a forward primer, 5′-TGGTTGTTGTATAGGTCGGTGTAAC, a reverse primer, 5′-ATCCACAGAACGCCTCTTCATC, and the labeled Taqman probe, vic-CATCCATTGCCAAAGTCGTTTCCGAA-tamra. The multiplex reaction consisted of 100 nM/100 nM primers/probe for the E9 region and 300 nM/100 nM for the primers/probe amplifying the ctpA2 control. PCR cycle conditions were 2 min at 50 °C, 10 min at 95 °C, and 40 cycles of 95 °C, 15 s, 60 °C, 1 min. ABI Universal PCR Mastermix (Applied Biosystems) was used for the reaction. The ΔΔCT method was used to quantify the PCR-amplified products and calibrate them to the one-copy control DNA.

Similarly, RNA purified from leaf and seed tissue was used to assess expression of the GUS transcript. The same E9 assay reagents used for the DNA assay were used in the expression assay. The ribosomal 18S RNA level was used as an endogenous control. The primers used for 18S control were a forward primer, 5′-CGTCCCTGCCCTTTGTACAC, a reverse primer, 5′-CGAACACTTCACCGGATCATT, and the labeled Taqman probe, vic-CCGCCCGTCGCTCCTACCGATtamra. The multiplex reaction consisted of 100 nM/100 nM primers/probe for the E9 region and 50 nM/50 nM for the primers/probe amplifying the 18S control. PCR cycle conditions were 30 min at 48 °C, 10 min at 95 °C, and 40 cycles of 95 °C, 15 s, 60 °C, 1 min. ABI One-step RT Mastermix (Applied Biosystems) was used for the reaction. Again, the ΔΔCT method was used to calculate a relative expression value of the PCR-amplified products and calibrate them to a known positive control *Arabidopsis* line expressing the GUS protein previously characterized by histochemical analysis.

Phenotypic Analysis and Methodology. Four phenotypic traits, that is, first flowering date (FFD), rosette diameter (RD), seed yield, and seed protein content, were assayed. The FFD is the number of days from the date of planting until the opening of the first flower and was assayed by daily inspection of the plants. The RD is the diameter (centimeters) of the leaf rosette at the time of first flowering. The seed yield is the amount of harvested seeds for one plant. Leaf and stem morphology was visually assessed by the overall shape, length, thickness, and pubescence. The seed protein content was determined using the FlashEA 1112 protein analyzer. The protein content was calculated with Eager 300 software using the protein factor of 6.25. The analysis was performed in replicates of the pooled seed samples, also used for the 2DE analysis. Total tocopherol contents and tocopherol compositions were assayed in seeds from the enhanced tocopherol transgenic lines and their control lines (WT and WT-P). The tocopherol analysis of the seeds was performed as described in ref *13*. Total tocopherol and composition are represented as the mean ± standard deviation (SD) of four measurements of a pooled seed sample of six plants. A statistically significant difference was evaluated with the statistical software program JMP v. 5 (SAS Institute Inc., Cary, NC) using the Tukey−Kramer simultaneous pairwise comparison procedure with a significance level of 5%.

Sample Preparation. Protein extracts were prepared as described in ref *12*. In brief, seeds harvested from six to eight plants were pooled and ground with a paint shaker-like device. Proteins were extracted with an extraction buffer containing 7 M urea, 2 M thiourea, 0.75% (w/v) CHAPS, 0.75% (v/v) Triton X-100, 100 mM DTT, 1% (v/v) carrier ampholytes stock, 20% (v/v) 2-propanol, and protease inhibitor

cocktail Complete (Roche, Mannheim, Germany) for 1 h while shaking at room temperature. After centrifugation, the supernatants were stored in aliquots at −80 °C until analysis. Protein concentration was estimated using the Bio-Rad protein assay with bovine serum albumin (BSA) as a standard.

Two-Dimensional Gel Electrophoresis. 2DE and gel staining were performed according to the method of ref *12*. Briefly, protein extracts were diluted in rehydration buffer containing 7 M urea, 2 M thiourea, 0.75% (w/v) CHAPS, 0.75% (v/v) Triton X-100, 100 mM DTT, 0.3% (v/v) carrier ampholytes stock, 10% (v/v) 2-propanol, 12.5% (v/v) water-saturated isobutanol, protease inhibitor cocktail Complete (Roche), and a trace of bromophenol blue. Nonlinear immobilized pH gradient gel strips (IPG) with nonlinear pH 3−10 gradients (13 cm, GE Healthcare) were rehydrated using 230 μL of diluted sample (150 μg of total protein). The IEF was carried out using a Bio-Rad PROTEAN IEF cell with a controlled cell temperature of 20 °C to a total of 35000 Vh. The IPG strips were equilibrated first for 10 min in 50 mM Tris-HCl (pH 8.8), 6 M urea, 30% (w/v) glycerol, 2.3% (w/v) SDS, 1% (w/v) DTT, and bromophenol blue and then for another 10 min in the same solution except DTT was replaced with 4% (w/v) iodoacetamide. The second dimension was run in a Bio-Rad Criterion Dodeca cell system in 8−16% Tris-HCl linear gradient Criterion gels. The gels were run for the first 15 min at 130 V and then at 180 V until the tracking dye reached the bottom of the gel. The gels were stained with colloidal Coomassie Brilliant Blue (CBB) solution according to method of Neuhoff et al. (*15*). The CBB staining solution was prepared fresh by mixing 4 parts of 0.1% (w/v) CBB G-250 in 2% (w/v) phosphoric acid and 11% w/v ammonium sulfate with 1 part of methanol. The gels were incubated in this solution for 3 days at room temperature on an orbital rotator. Imaging of the stained proteins was performed at a scan resolution of 36.3 × 36.3 μm using the GS-800 calibrated densitometer (Bio-Rad Laboratories, Hercules, CA).

Image Analysis. The scanned images of the 2DE gels were processed and analyzed with PDQuest 2-DE gel analysis software v. 7.1 (Bio-Rad Laboratories). All images were processed with the following software settings for spot detection and background subtraction: sensitivity, 40; size scale, 3; min peak, 400; power mean, 3 × 3; floater, 97; speckles filter. Spots detected by the software program were manually verified. False-positive spots (e.g., artifacts and multiple spots in a cluster) were manually removed; false-negative spots (obviously missed spots with OD > LOD) were added to the images. A spot was considered to be reproducibly present/absent when it was present/absent in all three replicate gels of one extraction. To compare spots across gels, a match set was created from the images of the gels in an experiment. A standard gel (master) was generated of the image with the greatest number of spots. Spots reproducibly present in a match set member but not present in the image with the most spots were manually added to the standard gel. The automated matching tool of the PDQuest software package was used to match spots across the gels. All spots matched by the software program were manually verified. The spots were quantified by 2D Gaussian modeling. Spot quantities of all gels were normalized to remove non-expression-related variations in spot intensity, so the raw quantity of each spot in a gel was divided by the total quantity of all the spots in that gel that have been included in the standard. All statistical analyses were performed with the statistical software packages JMP v. 5 or SAS v. 9 (SAS Institute Inc.). A one-way analysis of variance (ANOVA) with protected least significant difference (LSD) was performed for each spot to identify significant differences in spot quantities of 2-fold and greater with $P < 0.05$. To avoid overestimation of quantitative differences due to inaccurate quantification of poorly resolved protein spots, all spots with the maximum value and an average spot quantity ≤40 were excluded from the data set. In addition, the detected statistically significant difference ($P < 0.05$) must have a power above 80% (*12*).

Protein Identification by Matrix-Assisted Laser Desorption Ionization−Time of Flight Mass Spectrometry (MALDI-TOF MS). Protein spots were excised from the stained 2DE gels using Bio-Rad's spot cutter and placed into siliconized microcentrifuge tubes. Proteins were in-gel digested with trypsin according to the published procedure (*16*), except that the alkylation step was omitted, having been included in the 2DE gel procedure. Briefly, the gel pieces were washed three times with 100 μL of 50 mM ammonium bicarbonate for 15 min at room temperature, dehydrated by the addition of 100 μL of acetonitrile [50% (v/v) final concentration], and incubated for an additional 15 min at room temperature. The gel pieces were dried to completion for 1 h in a Speed-vac. Digestion was performed by incubating each gel piece in 30 μL of trypsin solution for 16 h with shaking at 37 °C. The trypsin solution was prepared by diluting sequencing-grade modified trypsin (Promega Co., Madison, WI) in 25 mM ammonium bicarbonate to a final concentration of 33 μg/mL. Peptides were extracted at room temperature by one change of 5% (v/v) trifluoroacetic acid and three changes of 5% (v/v) trifluoroacetic acid in 50% (v/v) acetonitrile (30 min for each change) at room temperature and dried down. The mass spectra of the tryptic digests were acquired with a Voyager-DE Pro MALDI-TOF mass spectrometer (Applied Biosystems) equipped with a nitrogen laser ($\lambda = 337$ nm). The samples were reconstituted in 5 μL of 0.1% (v/v) trifluoroacetic acid. In some cases, the samples were desalted with ZipTip C_{18} (Millipore, Billerica, MA) according to the instruction manual before the MALDI-TOF MS analysis. Peptides were cocrystallized 1:2 (v/v) with matrices consisting of saturated α-cyano-4-hydroxycinnamic acid prepared in 60% (v/v) acetonitrile/36% (v/v) methanol/4% (v/v) water. The spectra were either internally calibrated using known trypsin autocatalytic fragment peaks or externally calibrated using a standard peptide mixture. Monoisotopic peptide masses were assigned and searched against the NCBInr database using MS-Fit (*17*, *18*) and against the Swiss-Prot/TrEMBL databases using PeptIdent (*19*, *20*) to identify the proteins. The search parameters were set up as follows: cysteine as carbamidomethyl-cysteine; maximum allowed peptide mass error, 0.5 Da; consideration of one incomplete cleavage per peptide; minimum number of matched peptides, 4; and searching range, within the experimental p*I* value ± 1 pH unit and experimental molecular mass (MM) ± 20%.

Table 1. Phenotypic Measurements of the WT Col-0 and the TG Lines (Means ± SD)

name	*N*	FFD[a] (days)	RD[b] (cm)	seed yield (mg)	protein[c] (% fw[d])
Col-0	8	31 ± 1	7.4 ± 1.0	722 ± 92	24.7
TG-1	10	31 ± 1	9.1 ± 0.8*	711 ± 220	23.4
TG-2	10	31 ± 4	8.4 ± 0.5	641 ± 184	23.9
TG-3	12	32 ± 4	7.7 ± 0.8	583 ± 131	24.3
TG-4	11	31 ± 4	8.7 ± 0.9*	593 ± 200	23.2
TG-5	11	31 ± 2	9.1 ± 0.7*	596 ± 136	24.0
TG-6	12	32 ± 5	9.5 ± 1.4*	540 ± 184	25.4

[a] Number of days from the date of planting until the opening of the first flower. [b] Rosette diameter at the time of first flowering. The asterisk indicates significant difference ($P < 0.05$) from Col-0. [c] Protein content of pooled seed samples of six to eight plants; $N = 2$. [d] Fresh weight.

RESULTS AND DISCUSSION

GUS Lines. The T3 generation of 6 transgenic (TG) *A. thaliana* lines was grown side-by-side with the parental line (Col-0) and 12 ecotypes (*11*) under controlled environmental conditions. The transgenic *Arabidopsis* events contain one copy of the β-glucuronidase (*gus*) gene. The transgene is driven by the enhanced 35S promoter, derived from cauliflower mosaic virus (E35S), and is followed by a transcription terminator sequence (E9). The *gus* gene is a prokaryotic gene and not normally found in plant tissues (*21*, *22*). A physiological effect of the introduced gene is, therefore, not anticipated.

The six transgenic lines were selected because they contain one gene copy, are homozygous, and have different transgene expression rates in leaf tissues (**Table 3**). The expression rates were assayed by reverse transcription Polymerase Chain Reaction (RT-PCR) to determine the amount of E9-mRNA in relation to the expression rate of a transgenic control line with known expression rate (**Table 3**). The relative levels of *GUS* expression

Table 2. Qualitative Seed Proteome Comparison between WT Line and TG Lines (Means ± SD)

line	detected spots[a]	reproducible spots[b]	spots matched to WT Col-0	spots reproducibly absent[c]
Col-0	463 ± 10	429	429	1 (SSP 2814)
TG-1	472 ± 30	440	420	0
TG-2	465 ± 15	434	417	0
TG-3	472 ± 9	444	415	1 (SSP 2814)
TG-4	477 ± 6	447	421	0
TG-5	458 ± 13	426	416	0
TG-6	470 ± 19	442	419	0

[a] False-positive spots due to edge effects, focusing problems, or dust were excluded. Values are means ± SD. [b] Spot must be detected in all three replicate gels. [c] Spot must be absent in all three replicate gels.

Table 3. GUS Expression Measured in Leaf and Seed Tissue of T2 and T3 Generation Plants

	GUS expression (E9-mRNA) in			GUS protein
line	T2 generation (leaf tissue) av rel exp[a]	T3 generation (leaf tissue) av rel exp[a]	T3 generation (seed tissue) av rel exp[a]	T3 generation (seed tissue) IOD[b]
WT Col-0	0.00	0.00	0.00	0
TG-1	1.78	4.64	0.01	36
TG-2	2.95	1.88	1.37	4172
TG-3	5.45	0.28	0.00	0
TG-4	2.12	2.58	2.29	1745
TG-5	2.49	30.30	0.61	910
TG-6	11.88	34.75	0.75	3350

[a] Data are expressed as average relative expression of E9-mRNA to a transgenic control line with known expression rate. [b] Integrated optical density.

ranged from 1.8 to 11.9 times the levels in the control line with the highest levels seen in TG-6. Therefore, insertion of a single gene cassette can result in huge variations in transgene expression. The different levels of transgene expression may be due to differences in the integration site of the introduced *gus* gene, referred to as position effect (*23*−*25*).

Phenotypic Comparison. Four phenotypic parameters were assayed in the transgenic and wild-type (WT) *Arabidopsis* plants (**Table 1**). The FFD [measured at Boyes's growth stage 6.00 (*26*)], seed yield, and protein content do not show significant differences ($P > 0.05$) between Col-0 and the transgenic lines. A statistically significant difference ($P < 0.05$) in RD was observed between Col-0 and all transgenic lines except TG-2 and TG-3. The leaf rosettes of the transgenic lines TG-1, TG-4, TG-5, and TG-6 were on average 24% larger than the leaf rosettes of the WT. The line TG-6 had the largest RD (9.5 cm) compared to WT with a 7.4 cm RD. However, the differences in RD are in the range of natural variability (7.4−13.5 cm) observed among 12 *Arabidopsis* ecotypes in a previous study (*11*). All of the transgenic lines demonstrated very similar phenotypes compared to the WT Col-0 with respect to their leaf and stem morphology as assessed visually for overall shape, length, thickness, and pubescence.

Comparison of 2DE Patterns. Seeds harvested from eight individual replicate plants were pooled to one representative genotype sample to minimize the influence of environmental and genetic variation within a line. Using an optimized 2DE technique (*12*), the seed protein profiles of the six transgenic *Arabidopsis* lines were compared to the seed protein profile of the parental line Columbia (Col-0) (**Table 2**). The number of detected spots varied from 458 ± 13 protein spots for the line TG-5 to 477 ± 6 protein spots for the line TG-4, which is a very narrow range of spot variation in comparison to the range observed in 12 *Arabidopsis* ecotypes (*11*). The number of detected spots for the Col-0 fell into this range with 462 ± 10 resolved spots. The majority of the spots were reproducible (resolved in all three replicate gels) and varied between 426 spots for TG-5 and 447 spots for TG-4. The protein profiles of the seven lines were very similar as 93% (TG-3) to 98% (TG-5) of the reproducible spots of a line were matched to the parental wild-type Col-0 and 97% (414 spots) of the wild-type spots were matched to all lines. One spot (SSP 2814) in the lines TG-1, TG-2, TG-4, TG-5, and TG-6 was consistently absent in the 2DE gels of the line TG-3 and the WT line Col-0. **Figure 1** shows representative 2DE gels for WT Col-0 and for the line TG-2 as an example of the comparison. The gel region indicating the difference between the two lines is enlarged for both profiles. Protein spot SSP 2814 is absent in the 2DE pattern of WT Col-0 and present in the 2DE pattern of TG line TG-2. All other reproducibly present spots of one of the lines were also present in at least one replicate gel of all other samples, that is, no additional reproducible qualitative difference was detected between any of the lines.

Protein spot SSP 2814 was excised and digested in-gel with trypsin. The tryptic peptides were subjected to MALDI-TOF MS analysis. The measured peptide masses were searched against the NCBInr database using MS-Fit (*17*, *18*) and against the Swiss-Prot/TrEMBL databases using PeptIdent (*19*, *20*). To consider a protein unambiguously identified, at least five peptide masses need to be matched to the protein and 15% of the protein sequence needs to be covered (*27*). Twenty-three peptides matched the theoretical tryptic peptide masses for β-glucuronidase (GUS) ranging from position 26 to 602. When compared to the full GUS sequence, this covered 36% of the amino acid sequence of GUS. The GUS protein has a theoretical isoelectric point (p*I*) of 5.13 and a theoretical molecular mass (MM) of 70 kDa. The experimental p*I* and MM of spot SSP 2814 were found to be 5.1 and 73 kDa, respectively. Hence, the protein spot SSP 2814 was identified as GUS according to its peptide mass fingerprint (PMF) and its isoelectric point and molecular mass.

As expected, the GUS protein was detected in the 2DE gels of TG-1, TG-2, TG-4, TG-5, and TG-6 but not in WT Col-0 (**Figure 2**). The line TG-3 had >5-fold expression of the GUS construct but did not have detectable levels of GUS protein. The spot quantity of GUS varied among the transgenic lines, where GUS was reproducibly detected, and increased in the following order: TG-1 [<0.01% of total integrated optical density (IOD)], TG-5 (0.10% of total IOD), TG-4 (0.18% of total IOD), TG-6 (0.37% of total IOD), and TG-2 (0.44% of total IOD). Although the line TG-3 had the second highest expression of *gus* (E9-mRNA) in leaves of T2 generation plants (**Table 3**), neither the GUS protein (**Figure 2**; **Table 3**) nor the *gus* transcript (E9-mRNA) was detected in the seed of the T3 generation of this line. The 2DE analysis of the progenitor (T2 generation) seeds confirmed the absence of GUS on the 2DE gels. The *gus* gene is driven by the constitutive CaMV 35S promoter (*28*), and GUS expression should occur in both leaf and seed tissue. A possible explanation for the leaf-specific expression in TG-3 is that the insertion site is restricting expression of GUS to the leaves. There is no linear correlation between the GUS mRNA in leaves ($R^2 = 0.095$, $P = 0.50$) or seeds ($R^2 = 0.381$, $P = 0.14$) and the actual GUS protein in seeds. The poor correlation between mRNA and protein has been demonstrated previously (*29*; in yeast, refs *30*−*33*) and is thought to be due to different turnover rates of mRNA and protein, alternative splicing, and post-translational modification.

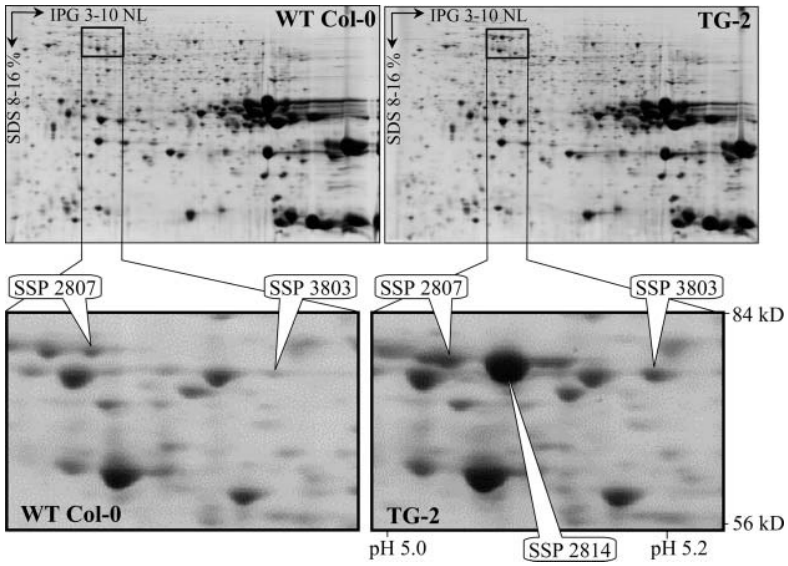

Figure 1. Head-to-head comparison between seed protein profiles of WT Col-0 and TG lines. Representative 2DE gels of WT Col-0 and TG line TG-2 are shown with enlargements of the gel regions of interest.

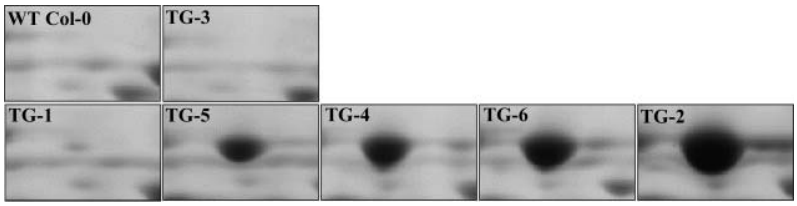

Figure 2. Differences in GUS expression among the six transgenic lines compared to WT.

Table 4. Quantitative Comparison of the Seed Proteomes of Col-0 and the Six TG Lines

		spots quantitatively different[a]	
line vs WT-P	compared spots	2-fold	≥3-fold
TG-1	420	0	0
TG-2	417	1	2 (4- and 6-fold)
TG-3	415	2	0
TG-4	421	2	1 (3-fold)
TG-5	416	1	0
TG-6	419	2	2 (4- and 6-fold)

[a] Power of difference must be >80% and mean spot quality >40.

Table 5. Quantities of the Significantly Different Spots in the Context of Natural Variation[a]

SSP	ecotypes (*11*)	WT Col-0	***TG-2***	TG-3	TG-4	TG-5	TG-6
2406	72–503	324		589			
2807	105–743	169	***1059***				823
3210	353–3768	1770			2741		
3803	41–202	83	***356***				304
6409	114–276	180	103	82	76	93	86
8103	117–2118	578			194		236

[a] Values are the mean ($N = 3$) normalized integrated optical density (IOD) of the spots.

Quantitative Comparison of 2DE Patterns. To investigate differences in spot quantities between the wild-type and the transgenic lines, the spot quantities of the 429 reproducibly present spots in Col-0 were compared to the spot quantities of the transgenic lines. A statistical significance of ≥2-fold was set for quantitative differences. **Table 4** summarizes the quantitative comparison of the seed proteomes of Col-0 and the six transgenic lines. In five of the six transgenic lines <1% of the spots varied significantly ($P < 0.05$) in spot quantity, whereas the proteome comparison of TG-1 did not reveal any differences in spot quantity relative to WT. The transgenic line TG-6 had, with four spots (0.95%), the highest number of significantly different spots. The differences in spot quantity varied between 2- and 6-fold. To evaluate any relevance of those differences, the spot quantities of these spots were set in context of natural variation of protein quantities found in 12 *Arabidopsis* ecotypes (**Table 5**). The data from an ecotype comparison (*11*) were included in the data set for the significance test. **Table 5** shows the ranges of spot quantity among the 12 ecotypes, the spot quantity of the WT Col-0, and the values of the significantly different transgenic lines. Eleven of the 13 different spots (from **Table 4**) are not significantly different from the values found in natural variation of the 12 ecotypes using a one-way ANOVA with protected LSD ($P < 0.05$). Two of the three significant spots of TG-2 exceeded the range of natural variation. Both spots are located very close to the GUS protein (SSP 2814, pI = 5.1; MM = 73 kDa) on the 2DE gel (**Figure 1**). The two protein spots SSP 2807 (pI = 5.1; MM = 76 kDa) and SSP 3803 (pI = 5.2; MM = 71 kDa) of TG-2 were subjected to MALDI-TOF MS analysis. The measured peptide masses of the protein spots were searched against Swiss-Prot databases using PeptIdent (*19*, *20*). Two different proteins were identified in spot SSP 2807; the *Arabidopsis* endogenous luminal binding protein (pI = 5.08; MM = 71 kDa) with 13 matched peptides and 21.2% sequence coverage and β-glucuronidase (pI = 5.24; MM = 68 kDa) with 11 matched peptides and a sequence coverage of 18.2%. Spot SSP 3803 also comprised more than one protein. β-Glucuronidase (pI = 5.24; MM = 68 kDa) was identified with nine matched peptides and a sequence coverage of 16.7%. Five of the remaining masses match the computed

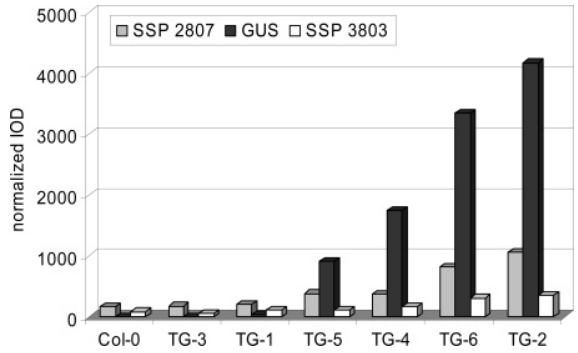

Figure 3. Correlation between the spot quantities of GUS, spot SSP 2807, and spot SSP 3803.

Table 6. Description of Control and Transgenic Lines

line	description
WT	interexperimental control line, WT Columbia (Col-0)
WT-P	parental line of TG lines, WT Columbia (Col-0)
35S:HPPD-2 35S:HPPD-3	*p*-hydroxyphenylpyruvate dioxygenase (*hppd*) gene driven by CaMV 35S promoter (35S); kanamycin resistant; (*14*)
DC3:HPPD-3 DC3:HPPD-8	*p*-hydroxyphenylpyruvate dioxygenase (*hppd*) gene driven by seed-specific DC3 promoter (DC3); hygromycin resistant; (*14*)
35S:gTMT-18 35S:gTMT-49	γ-tocopherol-methyltransferase (γTMT) gene driven by CaMV 35S promoter (35S); kanamycin resistant; (*13*)

tryptic peptide masses of *Arabidopsis* endogenous heat shock protein (HSP81) and cover 9.3% of the protein's sequence. The increase in spot quantity of spot SSP 2807 and SSP 3803 may be due to fragments of GUS migrating to the same position as the endogenous proteins of *Arabidopsis*. To support this assumption the spot quantities of spot SSP 2807 and spot SSP 3803 were compared to the spot quantity of the GUS spot (**Figure 3**). There is a strong correlation between the GUS quantity and the quantities of spot SSP 2807 (correlation coefficient = 0.975, $P = 0.0046$) and spot SSP 3803 (correlation coefficient = 0.973, $P = 0.0054$), indicating that SSP 2807 and SSP 3803 contain GUS fragments.

Enhanced Tocopherol Lines. In this portion of the study, the goal was to investigate changes in the protein pattern associated with transgenic lines containing an altered endogenous metabolic pathway. The transgenic lines overexpress one of two tocopherol biosynthetic pathway enzymes, the *Arabidopsis p*-hydroxyphenylpyruvate dioxygenase (*hppd*) gene or the *Arabidopsis* γ-tocopherol-methyltransferase (γTMT) gene. The enzyme HPPD (EC 1.13.11.27) catalyzes the conversion of *p*-hydroxyphenylpyruvate to homogentisic acid (HGA), the aromatic precursor for the biosynthesis of tocopherol and plastoquinone. The enzyme γTMT (EC 2.1.1.95) catalyzes the methylation of γ- and δ-tocopherol to yield α- and β-tocopherol, respectively (*34*). The six transgenic lines have previously been described (*13*, *14*) and were homozygous for the transgene (*35*). **Table 6** shows the lines and their descriptions. The WT-P line is the parental line of the transgenic lines. The WT line is an interexperimental control line used in the previous experiments with the transgenic GUS lines and in the natural variability study (*11*). Both control lines (WT-P and WT) have the same genetic background, *A. thaliana* ecotype Col-0, but originate from two different laboratories. Differences between these two lines represent natural variation within the ecotype Col-0.

Phenotypic Comparison. The two wild-type and the six transgenic *A. thaliana* lines were grown side-by-side in an environmentally controlled growth chamber. Six phenotypic traits (FFD, RD, seed yield, protein seed content, total tocopherol contents, and the tocopherol compositions) were assayed (**Table 7**).

The RD [measured at Boyes's growth stage 6.00 (*26*)], seed yield, and protein content do not show significant differences ($P > 0.05$) between the transgenic lines and the parental line WT-P. A statistically significant difference ($P < 0.05$) in FFD was observed between WT-P and the transgenic lines DC3:HPPD-3 and DC3:HPPD-8. Plants of the transgenic lines DC3:HPPD-3 and DC3:HPPD-8 started to flower, on average, 2 days later than WT-P. Although this difference is statistically significant, it is within the range of natural variation of flowering time for the ecotype Col-0 as WT plants started to flower 3 days later than WT-P. There were no observed differences in leaf and stem morphology, assessed by overall shape, length, thickness, and pubescence. Five of the six transgenic lines demonstrated the expected phenotype regarding total tocopherol level and tocopherol composition as previously described (*13*, *14*). The seed total tocopherol contents were increased 14% in 35S:HPPD-2, 17% in 35S:HPPD-3, and 9% in DC3:HPPD-8 above wild-type (WT-P) level (**Table 7**). The increases of total tocopherol in the transgenic lines over the WT-P are due to increases of γ- and δ-tocopherol (**Table 7**). The α- and β-tocopherol contents of these three transgenic lines were not significantly different ($P > 0.05$) from the values for WT-P. The α- and β-tocopherol levels of the transgenic lines 35S:gTMT-18 and -49 were increased 36- and >17-fold, respectively, at the expense of γ- and δ-tocopherol (**Table 7**), whereas the total seed tocopherol levels were not significantly different ($P > 0.05$) from the wild-type (WT-P). The observed changes in the seed tocopherol compositions of the γTMT-

Table 7. Phenotypic Measurements of the WT Col-0 and TG Lines (Mean ± SD)

line	FFD[a] (days), $N = 10$	RD[b] (cm), $N = 10$	seed yield (mg), $N = 10$	protein (% fw[c]), $N = 2$[d]	total tocopherol[e] (ng/mg pf fw), $N = 4$[d]	α-tocopherol[e] (ng/mg of fw), $N = 4$[d]	β-tocopherol (ng/mg of fw), $N = 4$[d]	γ-tocopherol[e] (ng/mg of fw), $N = 4$[d]	δ-tocopherol[e] (ng/mg of fw), $N = 4$[d]
WT	39 ± 0[f]	10.8 ± 0.4[f]	721 ± 101[g]	22.7	378 ± 10	8.98 ± 0.19	< 1	350 ± 8	18.5 ± 1.0*
WT-P	36 ± 1	9.0 ± 1.0	567 ± 194	23.3	370 ± 5	8.86 ± 0.85	<1	338 ± 4	22.7 ± 1.1
35S:HPPD-2	35 ± 1	8.5 ± 1.0	532 ± 194	22.9	420 ± 8*	9.38 ± 0.42	<1	383 ± 6*	27.4 ± 1.6*
35S:HPPD-3	36 ± 1	8.5 ± 1.0	581 ± 175	23.4	432 ± 6*	9.12 ± 0.39	<1	394 ± 5*	28.6 ± 2.0*
DC3:HPPD-3	38 ± 2*	9.0 ± 1.0	572 ± 177	23.5	368 ± 4	8.38 ± 0.43	<1	338 ± 5	22.2 ± 0.5
DC3:HPPD-8	38 ± 2*	9.5 ± 1.0	553 ± 208	22.9	404 ± 6*	8.77 ± 0.46	<1	370 ± 5*	25.7 ± 1.2*
35S:gTMT-18	37 ± 2	9.0 ± 1.5	633 ± 242	23.6	371 ± 4	315 ± 3*	20.3 ± 1.6	33.7 ± 0.5*	2.0 ± 0.08*
35S:gTMT-49	37 ± 1	10.0 ± 1.0	562 ± 264	22.2	367 ± 7	315 ± 6*	16.6 ± 0.3	34.4 ± 0.8*	1.6 ± 0.04*

[a] Number of days from the date of planting until the opening of the first flower. The asterisk indicates significant difference ($P < 0.05$) from WT-P. [b] Rosette leaf diameter at the time of first flowering. [c] Fresh weight. [d] Pooled seed samples of six to eight plants. [e] The asterisk indicates significant difference ($P < 0.05$) from WT-P. [f] $N = 2$. [g] $N = 6$.

Table 8. Qualitative Comparison of the Seed Proteomes of WT Lines and TG Lines

line	detected spots[a]	reproducible spots[b]	spots matched to WT-P	spots reproducibly absent[c]
WT	747 ± 9	708	702	0
WT-P	736 ± 10	706	706	0
35S:HPPD-2	729 ± 10	702	700	0
35S:HPPD-3	737 ± 12	702	701	0
DC3:HPPD-3	747 ± 15	704	703	0*
DC3:HPPD-8	750 ± 27	704	702	0
35S:gTMT-18	730 ± 13	700	699	0
35S:gTMT-49	733 ± 17	705	704	0

[a] False-positive spots due to edge effects, focusing problems, or dust were excluded. Values are means ± SD. [b] Spot must be detected in all three replicate gels. [c] Spot must be absent in all three replicate gels. *, spot SSP 9003 was absent in all three replicate gels but present in the fourth (control) replicate gel.

Table 9. Quantitative Comparison of Seed Proteomes of the Six TG Lines and Their Parental Line WT-P

line vs WT-P	compared spots	spots quantitatively different[a] 2-fold	spots quantitatively different[a] ≥3-fold
35S:HPPD-2	708	6	0
35S:HPPD-3	707	4	0
DC3:HPPD-3	707	3	0
DC3:HPPD-8	708	4	0
35S:gTMT-18	707	4	0
35S:gTMT-49	707	2	0

[a] Power of difference must be >80% and mean spot quality >40.

overexpressing lines were expected, as γ- and δ-tocopherol are substrates for γ-TMT (*34*).

The transgenic line DC3:HPPD-3 did not show any changes in total tocopherol or tocopherol composition compared to WT-P. Tsegaye et al. (*14*) found an increase in tocopherol content of 28% for the transgenic line DC:HPPD-3 compared to the wild-type line. Variations in plant growth and/or seed harvesting conditions may be possible explanations. The nature of this observation was not further investigated.

Comparison of 2DE Patterns. Using an optimized 2DE technique, the seed protein profiles of the six transgenic *Arabidopsis* lines were compared to the seed protein profiles of the parental line Columbia (Col-0) (**Table 8**). The number of detected spots varied from 729 ± 10 protein spots for 35S:HPPD-2 to 750 ± 27 protein spots for DC3:HPPD-8. The majority of the spots found in all lines were reproducible spots (resolved in all three replicate gels) and varied between 700 spots for 35S:gTMT-18 and 708 spots for WT. The protein profiles of the eight lines were very similar, as 99.1% (WT) to 99.9% (35S:gTMT-49) of the reproducible spots of a line were matched to the parental wild-type WT-P and >97% (686 spots) of the reproducible spots were matched to all of the lines. One reproducible spot (SSP 9003) resolved for the WT-P was absent in the three 2DE replicate gels of the DC3:HPPD-3. However, the protein spot was not scored as reproducibly absent because it is a very faint spot that is negatively affected by focusing problems of the neighboring, very abundant spot SSP 8017. In addition, the spot was unambiguously resolved in the fourth 2DE replicate gel of the TG line DC3:HPPD-3. Therefore, all reproducibly present spots of one of the lines were present in at least one replicate gel of all other samples; that is, no reproducible qualitative difference was detected between any of the lines.

Quantitative Comparison of 2DE Patterns. To investigate differences in spot quantities between parental and transgenic lines, the quantities of all spots reproducibly present in at least one line and matched to WT-P were examined (**Table 9**). About 708 spots were compared, and fewer than 1% of the spots (6 spots) varied significantly ($P < 0.05$) by ≥2-fold in spot quantity. The transgenic line 35S:HPPD-2 had, with 6 spots, the highest number of spots significantly different in spot quantity. The differences in spot quantity were <3-fold for all significantly different spots.

To evaluate the relevance of those differences, the spot quantities of these spots were compared first to the spot quantities of the other WT line grown with the WT-P and the transgenic plants using a one-way ANOVA and then to the spot quantity ranges based on environmental variation (**Table 10**). Both wild-type lines, WT and WT-P, represent the same *Arabidopsis* ecotype, Columbia (Col-0), but originate from two different laboratories. Differences in their proteome represent natural variation within the ecotype Col-0. **Table 10** shows the significantly different spot quantities of the transgenic lines in the context of the spot quantities of the two wild types (WT and WT-P) as well as the spot quantity ranges spanned by the 12 *Arabidopsis* ecotypes (*11*). Sixteen of the 21 significantly different spots (transgenic lines vs WT-P) are not significantly different from the spot quantities of WT (Col-0) using a one-way ANOVA. The other five significantly different spots (boldface, **Table 10**) fall in the range of natural variation (ecotypes). Hence, it can be concluded that the differences between the transgenic lines and the parental wild-type WT-P did not exceed the range of natural variation.

Contrary to the transgenic protein GUS, which was easily detected in five of the six transgenic GUS lines, neither HPPD (pI = 5.74; MM = 49 kDa) nor gTMT (p = 5.81; MM = 33 kDa) was detected in the 2DE patterns of the transgenic lines

Table 10. Quantities of the Significantly Different Spots in the Context of Natural Variation[a]

SSP	ecotypes (*11*)	WT (Col-0)	WT-P (Col-0)	35S: HPPD-2	35S: HPPD-3	DC3: HPPD-3	DC3: HPPD-8	35S: gTMT-18	35S: gTMT-49
0204	82–187	171	119				217	217	204
0302	481–799	666	470				889		
0308	53–140	122	76			142	141		
1106	0–1919	3037	2414					3306	
4603	100–208	159	96	169	156	164		160	
6104	0–882	489	428		645				
8302	0–202	198	237	103	111		110		
8414	*a*	236	330	205					
8611	0–2772	408	319	**189**				**210**	**188**
8618	0–1762	183	88	**45**					
8708	0–438	196	232	**102**					

[a] Values are the mean (N = 3) normalized IOD of the spots. [b] No quantitative data are available because of low spot qualities.

as a significant difference compared to the parental wild type (WT-P). Possible explanations may be (i) that the transgenic proteins are not readily distinguishable from endogenous proteins in the 2DE protein pattern, (ii) low protein levels, and/or (iii) low protein extractability. Unlike GUS, which has not been found in plant species (*21*, *22*), HPPD and gTMT are endogenous *Arabidopsis* proteins, which were overexpressed in the transgenic lines. The proteins will most likely comigrate to the same position in the 2DE gel as the nontransgenic proteins and not appear as additional protein spots in the 2DE protein pattern, such as GUS in the TG GUS lines. Therefore, the transgenic proteins would not be identified as new proteins but only as changes in spot quantity. Although Tsegaye et al. (*14*) demonstrated a HPPT protein increase in seeds of 3.5-fold (35S:HPPD-2) to 17-fold (DC3:HPPD-8) relative to wild type by immunoblot analysis, the protein amount of HPPD is probably still below the limit of quantification of the 2DE method (staining with colloidal CBB) due to the low abundance of HPPD in plant tissues (*36*). No protein data have been published on gTMT of the transgenic lines used. However, unlike the cytosolic HPPD (*36*−*38*), gTMT is a membrane-bound protein (*34*, *39*) and, therefore, most likely not readily extractable.

None of the marker gene products was detected on the 2DE gels. It is presumed that the expression of the marker genes is below the limit of detection (LOD) of the developed method. The same was observed by Corpillo et al. (*40*). They were only able to detect the product of their marker gene (*nptII*) with immunoblotting of a 2DE gel using a primary antibody against the protein NPT II.

Conclusion. The applicability of proteomics to investigate differences in the plant proteome due to genetic engineering was explored using *A. thaliana* as a model organism. Differences in the proteome were evaluated in the context of natural variability. On the basis of the changes detected for the proteins surveyed, the genetic modification of *Arabidopsis* using three different genes and three different promoters did not result in any phenotypic or seed proteome differences exceeding the natural variation other than the intended differences due to the introduction of the transgene. The process of transformation seems not to have caused insertional or pleiotropic changes to the analyzed seed proteome. Not much change was seen here that would inform a safety assessment. Differences in spot quantity between transgenic and nontransgenic lines fell in the range of natural variation or were part of the intended effect. A similar 2DE-based proteomics approach was utilized by Corpillo et al. (*40*) and Lehesranta et al. (*41*) to compare the protein profiles of transgenic tomato (*40*) and potato (*41*) lines with their nontransgenic counterparts. They also did not find any protein changes due to genetic modification.

This study demonstrated that 2DE can be utilized to reliably analyze the seed proteome of transgenic *A. thaliana*. However, it has to be noted that the presented 2DE method is limited to a subset of the seed proteome (*12*). During the course of the study, it became evident that a critical data analysis needs to take into consideration the analytical and natural variability of the proteome. The latter is essential for the evaluation of potential insertional and pleiotropic effects in comparing transgenic and nontransgenic plants. Thus, a proteome analysis should comprise the following steps: (i) method validation, (ii) generation of baseline data for the natural variation, and (iii) head-to-head comparison between transgenic and nontransgenic plants in the context of the established analytical and natural variation. The proteomics approach described for *A. thaliana* promises to be useful for the analysis of the proteome in other plant species including crop plants. The method could be applied to classification of ecotypes. In the context of safety assessment it could be used as a screen for global changes in protein profiles that could be taken as a signal for further investigation. However, it should be kept in mind that the detection of changes in protein profiles does not present a safety issue per se; the relevance of such changes for food safety would have to be assessed by subsequent elucidation of the nature of the proteins affected. The overwhelming majority of proteins within a species are harmless, and many species will lack toxic, antinutritional, or allergenic proteins entirely. It is suspected that profiling methodologies will add modest, if any, incremental benefit to the evaluation of endogenous toxins, antinutrients, and allergens for which routinely applied traditional analytical methods are available (*42*). Indeed, nonproteinaceous compounds that arise through plant breeding may have greater relevance for food toxicology and allergenicity (*43*). At this time the adoption of proteomics as a compulsory part of a regulatory assessment procedure is premature. Future studies are needed before any such conclusion can be drawn.

ACKNOWLEDGMENT

We thank Daniel Free for assistance with plant cultivation and Margaret A. Nemeth for statistical expertise.

LITERATURE CITED

(1) Conner, A. J.; Jacobs, J. M. Genetic engineering of crops as potential source of genetic hazard in the human diet. *Mutat. Res.* **1999**, *443*, 223−234.

(2) Conner, A. J.; Jacobs, J. M. E. Food risks from transgenic crops in perspective. *Nutrition* **2000**, *16*, 709−711.

(3) Pedersen, S.; Eriksen, F. D.; Knudsen, I. Toxicity and food safety of genetically engineered crops. In *Safety of Genetically Engineered Crops*; Custers, R., Ed.; Jo Bury, VIB: Zwijnaarde, Belgium, 2001; pp 27−60.

(4) *GM Science Review Panel—UK GM Science Review Panel: Second Report*; an open review of the science relevant to GM crops and food based on interests and concerns of the public; Jan 22, 2004; http://www.gmsciencedebate.org.uk/report/.

(5) Society of Toxicology. The safety of genetically modified foods produced through biotechnology. *Toxicol. Sci.* **2003**, *71*, 2−8.

(6) van Gelder, W. M. J.; Scheffer, J. J. C. Transmission of steroidal glycoalkaloids from *Solanum vernei* to the cultivated potato. *Phytochemistry* **1991**, *30*, 165−168.

(7) FAO/WHO. *Report of the third session of the Codex ad hoc Intergovernmental Task Force on Foods Derived from Biotechnology*, Yokohama, Japan, March 4−8, 2002; Codex Alimentarius Commission, FAO: Rome, Italy, 2002.

(8) FAO/WHO. *Safety Aspects of Genetically Modified Foods of Plant Origin*; joint FAO/WHO consultation on foods derived from biotechnology, Geneva, Switzerland, May 29−June 2, 2000; Food and Agriculture Organization of the United Nations and World Health Organization: Rome, Italy, 2000.

(9) *EC Scientific Steering Committee Guidance Document for the Risk Assessment of Genetically Modified Plants and Derived Food and Feed* (March 6−7, 2003); European Commission: Brussels, Belgium, 2003; http://europa.eu.int/comm/food/fs/sc/ssc/out327_en.pdf.

(10) *National Research Council Safety of Genetically Engineered Foods: Approaches to Assessing Unintended Health Effects*; National Academies Press: Washington, DC, 2004; http://books.nap.edu/catalog/10977.html.

(11) Ruebelt, M. C.; Lipp, M.; Reynolds, T. L.; Astwood, J. D.; Engel, K.-H.; Jany, K.-D. Application of two-dimensional gel electrophoresis to interrogate alterations in the proteome of genetically modified crops. 2. Assessing natural variability. *J. Agric. Food Chem.* **2006**, *54*, 2162−2168.

(12) Ruebelt, M. C.; Leimgruber, N. K.; Lipp, M.; Reynolds, T. L.; Nemeth, M. A.; Astwood, J. D.; Engel, K.-H.; Jany, K.-D. Application of two-dimensional gel electrophoresis to interrogate alterations in the proteome of genetically modified crops. 1. Assessing analytical validation. *J. Agric. Food Chem.* **2006**, *54*, 2154−2161.
(13) Collakova, E.; DellaPenna, D. Homogentisate phytyltransferase activity is limiting for tocopherol biosynthesis in *Arabidopsis*. *Plant Physiol.* **2003**, *131*, 632−642.
(14) Tsegaye, Y.; Shintani, D.; DellaPenna, D. Overexpression of the enzyme *p*-hydroxyphenolpyruvate dioxygenase in Arabidopsis and its relation to tocopherol biosynthesis. *Plant Physiol. Biochem.* **2002**, *40*, 913−920.
(15) Neuhoff, V.; Arold, N.; Taube, D.; Ehrhardt, W. Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* **1988**, *9*, 255−262.
(16) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **1996**, *68*, 850−858.
(17) Clauser, K. R.; Baker, P.; Burlingame, A. L. Role of accurate mass measurement (± 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal. Chem.* **1999**, *71*, 2871−2882.
(18) UCSF Mass Spectrometry Facility MS-Fit. Sept 2004; http://prospector.ucsf.edu/ucsfhtml4.0/msfit.htm.
(19) Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **2003**, *31*, 3784−3788.
(20) ExPASy. Expert Protein Analysis System Proteomics tools: PeptIdent; Sept 2004; http://www.expasy.org/tools/peptident.html.
(21) Jefferson, R. A. Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol. Biol. Rep.* **1987**, *5*, 387−405.
(22) Jefferson, R. A.; Kavanagh, T. A.; Bevan, M. W. GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **1987**, *6*, 3901−3907.
(23) Peach, C.; Velten, J. Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. *Plant Mol. Biol.* **1991**, *17*, 49−60.
(24) Dean, C.; Jones, J.; Favreau, M.; Dunsmuir, P.; Bedbrook, J. Influence of flanking sequences on variability in expression levels of an introduced gene in transgenic tobacco plants. *Nucleic Acids Res.* **1988**, *16*, 9267−9283.
(25) Van Leeuwen, W.; Mlynarova, L.; Nap, J. P.; van der Plas, L. H.; van der Krol, A. R. The effect of MAR elements on variation in spatial and temporal regulation of transgene expression. *Plant Mol. Biol.* **2001**, *47*, 543−554.
(26) Boyes, D. C.; Zayed, A. M.; Ascenzi, R.; McCaskill, A. J.; Hoffman, N. E.; Davis, K. R.; Gorlach, J. Growth stage-based phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants. *Plant Cell* **2001**, *13*, 1499−1510.
(27) Mann, M.; Hendrickson, R. C.; Pandey, A. Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem.* **2001**, *70*, 437−473.
(28) Odell, J. T.; Nagy, F.; Chua, N. H. Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature* **1985**, *313*, 810−812.
(29) Anderson, L.; Seilhamer, J. A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* **1997**, *18*, 533−537.
(30) Washburn, M. P.; Koller, A.; Oshiro, G.; Ulaszek, R. R.; Plouffe, D.; Deciu, C.; Winzeler, E.; Yates, J. R., 3rd. Protein pathway and complex clustering of correlated mRNA and protein expression analyses in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3107−3112.
(31) Griffin, T. J.; Gygi, S. P.; Ideker, T.; Rist, B.; Eng, J.; Hood, L.; Aebersold, R. Complementary profiling of gene expression at the transcriptome and proteome levels in *Saccharomyces cerevisiae*. *Mol. Cell. Proteomics* **2002**, *1*, 323−333.
(32) Gygi, S. P.; Rochon, Y.; Franza, B. R.; Aebersold, R. Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **1999**, *19*, 1720−1730.
(33) Futcher, B.; Latter, G. I.; Monardo, P.; McLaughlin, C. S.; Garrels, J. I. A sampling of the yeast proteome. *Mol. Cell. Biol.* **1999**, *19*, 7357−7368.
(34) Shintani, D.; DellaPenna, D. Elevating the vitamin E content of plants through metabolic engineering. *Science* **1998**, *282*, 2098−2100.
(35) DellaPenna, D. Personal communication: description to tocopherol enhanced transgenic *Arabidopsis* lines, 2003.
(36) Garcia, I.; Rodgers, M.; Lenne, C.; Rolland, A.; Sailland, A.; Matringe, M. Subcellular localization and purification of a *p*-hydroxyphenylpyruvate dioxygenase from cultured carrot cells and characterization of the corresponding cDNA. *Biochem. J.* **1997**, *325*, 761−769.
(37) Norris, S.; Barrette, T.; DellaPenna, D. Genetic dissection of carotenoid synthesis in arabidopsis defines plastoquinone as an essential component of phytoene desaturation. *Plant Cell* **1995**, *7*, 2139−2149.
(38) Garcia, I.; Rodgers, M.; Pepin, R.; Hsieh, T.; Matringe, M. Characterization and subcellular compartmentation of recombinant 4-hydroxyphenylpyruvate dioxygenase from *Arabidopsis* in transgenic tobacco. *Plant Physiol.* **1999**, *119*, 1507−1516.
(39) Arango, Y.; Heise, K. P. Tocopherol synthesis from homogentisate in *Capsicum annuum* L. (yellow pepper) chromoplast membranes: evidence for tocopherol cyclase. *Biochem. J.* **1998**, *336* (Part 3), 531−533.
(40) Corpillo, D.; Gardini, G.; Vaira, A. M.; Basso, M.; Aime, S.; Accotto, G. P.; Fasano, M. Proteomics as a tool to improve investigation of substantial equivalence in genetically modified organisms: the case of a virus-resistant tomato. *Proteomics* **2004**, *4*, 193−200.
(41) Lehesranta, S. J.; Davies, H. V.; Shepherd, L. V.; Nunan, N.; McNicol, J. W.; Auriola, S.; Koistinen, K. M.; Suomalainen, S.; Kokko, H. I.; Karenlampi, S. O. Comparison of tuber proteomes of potato varieties, landraces, and genetically modified lines. *Plant Physiol.* **2005**, *138*, 1690−1699.
(42) Astwood, J. D.; Fuchs, R. L. Status and safety of biotech crops. In *Agrochemical Discovery: Insect, Wood and Fungal Control*; Baker, D. R., Umetsu, N. K., Eds.; ACS Symposium Series 774; American Chemical Society: Washington, DC, 2001; Chapter 14, pp 152−164.
(43) Franck-Oberaspach, S. L.; Keller, B. Consequences of classical and biotechnological resistance breeding for food toxicology and allergencity. *Plant Breeding* **1997**, *116*, 1−17.

Received for review September 23, 2005. Revised manuscript received January 19, 2006. Accepted January 20, 2006.

JF052358Q