10918 J. Agric. Food Chem. 2010, 58, 10918–10928
DOI:10.1021/jf102208a

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

Seed Coat Phenolics and the Developing Silique Transcriptome of *Brassica carinata*

Xiang Li,[†,‡] Neil Westcott,[†] Matthew Links,[†] and Margaret Y. Gruber*[,†]

[†]Agriculture and Agri-Food Canada, Saskatoon Research Centre, 107 Science Place, Saskatoon, Saskatchewan, S7N 0X2, Canada, and [‡]College of Plant Sciences, Jilin University, Changchun 130062, China

Structures for nine compounds were elucidated in seed coats of two genetically related *Brassica carinata* lines. The yellow-seeded line accumulated monomeric kaempferols, phenylpropanoids, and lignans, while extractable and unextractable proanthocyanidins and a high-performance liquid chromatography peak containing polymeric-like quercetin/lignan structures were strongly reduced. The brown-seeded line accumulated large amounts of both types of proanthocyanidins (extractable and unextractable), as well as phenylpropanoids and lignans equivalent to the amounts in the yellow-seeded seed coats, but the brown-seeded seed coats lacked kaempferols. A *Brassica napus* 15K oligoarray experiment indicated that yellow-seeded siliques had more extreme gene expression changes and a 2.4-fold higher number of upregulated genes than brown-seeded siliques, including a host of transcription factors and genes with unknown function. Transcripts for six flavonoid genes (*CHS*, *F3H*, *FOMT*, *DFR*, *GST*, and *TTG1*) were lower and two (*F3′H* and *FLS*) were higher in yellow-seeded siliques, but expression of *CHI*, *PAP1*, and phenylpropanoid genes was unchanged.

KEYWORDS: Flavonoids; lignans; phenylpropanoids; *B. carinata* developing silique transcriptome; *B. napus* 15K microarray

INTRODUCTION

Rapeseed and canola occupy the second position in seed oil production volume in the world after soybean (*1*). *Brassica carinata* A. Braun (Ethiopian mustard) is a less well-known *Brassica* crop, which is used in Ethiopia and Zambia as both a leaf vegetable and an oilseed (*2*). Elsewhere, *B. carinata* is under development as a biorefinery crop (*3*). Cultivars of *B. carinata* currently grown for oilseed production are brown-seeded lines, but the yellow-seeded phenotype has decreased fiber, lignin, and proanthocyanidins (PAs) in the seeds compared to the brown-seeded phenotype (*4*, *5*). This yellow-seeded phenotype is inherited as a monogenic, dominant trait (*6*). Dominance is rare among yellow-seeded breeding germplasm in other related crop *Brassica* species (*7*). Usually (but not always) yellow-seeded seed is specified by two or three recessive genes (*8*, *9*).

Flavonoids belong to a group of plant natural products with variable phenolic structures and play important roles in protection against biotic and abiotic stress (*10*). They are well-known for their positive effect on health, including antioxidant and antitumor properties (*11*). Anthocyanins and PAs are two important plant pigments, which share common flavonoid intermediates until the formation of anthocyanidins (**Figure 1**). Previously, we reported the correlation of PA pigment reduction with the reduction of dihydroflavonol reductase (DFR) transcripts and a rise in flavonoid content in the seed coat of a yellow-seeded *B. carinata* line compared to a genetically related brown-seeded line (*5*) (**Figure 1**). We also determined specific metabolites and transcriptome profiles that protected brown-seeded seedlings

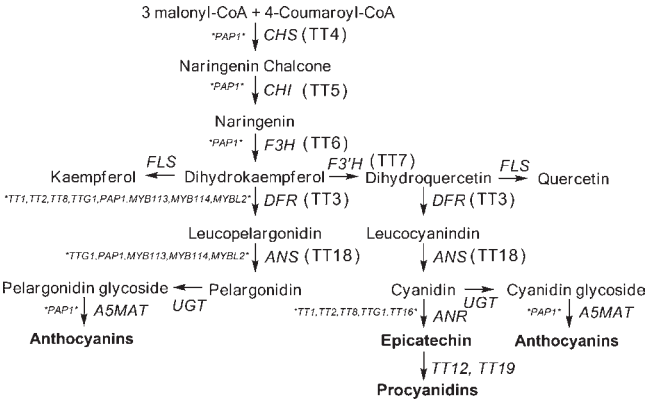

Figure 1. Flavonoid biosynthetic pathways: chalcone synthase (*CHS*), chalcone isomerase (*CHI*), flavanone 3β-hydroxylase (*F3H*), flavonoid 3′-hydroxylase (*F3′H*), flavonol synthase (*FLS*), dihydroflavanol reductase (*DFR*), leucoanthocyanidin dioxygenase or anthocyanin synthase (*LDOX* or *ANS*), banyuls or anthocyanidin reductase (*BAN* or *ANR*), UDP-glycosyltransferase (*UGT*). anthocyanidin 5-methyl-acyl-transferase (*A5MAT*), multi-drug and toxic efflux transporter (*MATE*, TT12), and glutathione *S*-transferase (*GST*, TT19). Transparent Testa biochemical loci from *Arabidopsis* are indicated in parentheses. Regulatory genes known in the *Arabidopsis* literature are indicated in small font between asterisks at the loci that they are known to affect. The 3′5′-hydroxylated branch of the pathway leading to myricetin and trihydroxylated PAs is not shown.

*To whom correspondence should be addressed. Telephone: 306-956-7263. Fax: 306-956-7247. E-mail: margie.gruber@agr.gc.ca.

 Published on Web 10/06/2010

from exposure to lithium chloride and allowed them to accumulate this toxic metal salt to relatively high levels compared to the yellow-seeded seedlings (*12*). In the present study, a more detailed analysis of seed coat phenolic composition and a microarray of developing silique gene expression were undertaken for the two genetically related lines. The outcome will aid in our understanding of how to use *B. carinata* to the fullest advantage in a bioproduct and bioprocess economy and in plant breeding for higher seed quality.

MATERIAL AND METHODS

General Experiments and Plant Line Development. All chemicals used in this study were obtained from Sigma Chemical Co. (St Louis, MO). Solvents were HPLC-grade. The *p*-dimethylaminocinnamaldehyde (DMACA) reagent, consisting of 2% DMACA dissolved in 1.5 N sulfuric acid (*13*), was used to detect PAs histochemically by soaking dissected seed coat tissues for 30−60 min after imbibition, and then excess reagent was removed by washing for a short time with water. The purple/black color of the stained seed coat was very stable for about 12−24 h.

Genetically related *B. carinata* lines were obtained by single-seed descent from backcross (BC_1) lines with yellow- and brown-seeded seeds from a cross between an Ethiopian yellow-seeded mustard accession PGRC/E 21164 and a brown-seeded Ethiopian mustard accession S67 and a backcross between the dominant F_1 yellow-seeded seed and the parental brown-seeded line (*14*). Developing siliques and bulked seed increases were obtained by planting S_3 seeds (generation 3) in a soilless potting mixture composed mainly of sphagnum moss and slower release fertilizer at pH 5.8 (Redi-Earth, Grace and Co., Canada). Pots with planted seeds were placed in a controlled environment chamber (Conviron, Winnipeg, Canada) with an 18 h photoperiod (22 °C dark and 20 °C light) under fluorescent and incandescent lighting (320−510 μC) tested with a model LI-185B luminometer (Licor Bioscience, Lincoln, NB). Seeds were crushed in a Reliance grinder (Baldor, CA), and seed coats were separated from the meal in a homemade aspirator at Plant Gene Resources Canada (Saskatoon, Saskatchewan, Canada).

Flavonoids and PA Analysis. Small-scale extracts were prepared from 200 mg of seed coat ground in liquid N_2 with 10 mL of acetone/water (70:30, v/v) for 10 min. Ground samples were extracted 3 times for 24 h in the dark with 50 mL of acetone/water (70:30, v/v). The extracts were combined and evaporated to dryness at 35 °C under vacuum, and the residue was redissolved to 10 mg/mL in methanol/water (50:50, v/v) and used for analysis of flavonoids, lignans, and phenylpropanoids. Liquid chromatography−mass spectrometry (LC−MS) was performed on a Zorbax C_{18} column (150 × 4.6 mm, 5 μm inner diameter, Mississauga, Ontario, Canada) using Agilent 1100 high-performance liquid chromatography (HPLC) coupled to a photodiode array detector and an "API Qstar XL" pulsar hybrid LC−MS/MS system (Applied Biosystems) in the electrospray index (ESI) mode. Compounds **1**−**18**, which were detected in small-scale extracts, were used as a basis for large-scale extractions and identification of nine compounds from yellow-seeded seed coats (**1**, **7**, **8**, **10**, **13**−**16**, and **18**).

For large-scale extracts, powdered yellow-seeded *B. carinata* seed coats (150 g) were extracted 3 times with 70% MeOH (500 mL) at room temperature to purify flavonoids and phenolics present in the yellow-seeded seed coats but absent from brown-seeded seed coat. The extract was concentrated to give a brown residue (8 g), which was eluted into 96 tubes on an open Sephadex LH-20 chromatography column (400 g of dry weight, 80 × 5 cm) with 70% MeOH (4 L) over 90 min and combined to yield five fractions. Fraction 1 (tubes 1−20) showed four major spots on silica gel SilG/U254 thin-layer chromatography (TLC) plates (0.20 mM, Macherey-Nagel, Bethlehem, PA) at λ_{254} when detected using 0.5% anisaldehyde in 10% H_2SO_4. Fraction 1 was applied to an open silica gel column (Silica 60M, 200−300 mesh, GE Healthcare, Piscataway, NJ) and chromatographed with chloroform/methanol (3:1). Fractions were collected and detected by TLC, and then the eluent (F1-3) was separated by semi-preparative HPLC−ultraviolet (UV) on a 250 × 25 mm, inner diameter 5 μm, Zorbax C_{18} column (Agilent, Mississauga, Ontario, Canada) with a gradient from 20% aqueous MeOH to 100% MeOH to produce compounds **1** (3 mg), **8** (56 mg), and **10** (4 mg). Fraction 2 (tubes 21−40), showing five spots by identical TLC conditions, was reapplied to Sephadex LH-20 (200 g of dry weight) and eluted with 70% acetone (1.5 L), and then subfraction F2-2 was purified by semi-preparative HPLC−UV as above to produce compounds **7** (9 mg), **14** (28 mg), and **15** (2 mg). Fraction 3 (tubes 41−60) and fraction 4 (tubes 61−80) were directly applied as concentrates to the semi-preparative HPLC as above to recover compounds **13** (31 mg) and **18** (22 mg). Compound **16** (3 mg) was isolated from fraction 5 (tube 81−96) on the open silica gel column eluted with a step gradient of chloroform/methanol [nine 1 L steps from 9:1 (v/v) to 1:1 (v/v)].

Purified compounds were confirmed by LC−MS/MS as above and nuclear magnetic resonance (NMR) as follows. Purified compounds were dried under nitrogen and dissolved in 500 μL of freshly opened dimethylsulfoxide (DMSO)-d_6 [99.96 + 0.03 tetramethylsilane (TMS)], and ^{1}H NMR and heteronuclear multiple-bond correlation (HMBC) spectra were measured with a Bruker Avance 500 NMR spectrometer equipped with a Bruker 5 mm inverse triple-resonance TXI probe (Bruker Biospin, Germany). Chemical shifts (δ) were expressed in parts per million (ppm), and coupling constants (J) are reported in Hertz (Hz). Compound structures were identified by a comparison of HPLC retention time, UV spectroscopic parameters, and MS/MS fragmentation patterns to those of authentic standards (where available) and a comparison of the NMR data to literature values (*15*−*25*).

Determination of Extractable and Unextractable PA by BuOH/HCl Hydrolysis. The butanol/HCl assay was used to quantify the total amount of extractable PAs in *B. carinata* seed coat according to Naczk et al. (*26*). In each tube, 0.1 mL (10 mg/mL) of the small-scale 70% acetone extract was incubated for 75 min at 95 °C with 2 mL of *n*-BuOH/HCl reagent [95:5 (v/v) with 0.1 mL of $FeSO_4$ in 2 M HCl]. After cooling in the dark and centrifugation, absorbance of red anthocyanidins in the supernatant (λ_{max} = 525 nm) was determined after subtraction of the non-PA-related background scan. Samples were measured against a blank of *n*-BuOH/HCl reagent, and the value was calculated using PA B_2 as a standard (Sigma). Unextractable PA was measured by heating the solid residue from the extractable PA method 3 times with 2 mL of freshly prepared *n*-BuOH/HCl reagent as above (*27*).

DNA Microarray Analysis. A 15K *Brassica napus* microarray used in the experiment was spotted at the Microarray and Proteomics Facility, University of Alberta, Edmonton, Alberta, Canada, using 50-mer *B. napus* oligonucleotide sequences based on expressed sequence tag (EST) deposits at the Saskatoon Research Centre, Agriculture and Agri-Food Canada. Total RNA from 90-day-old developing siliques (22 days after pollination) of the yellow- and brown-seeded *B. carinata* lines was extracted using a commercial RNAEasy mini kit (Qiagen, Valencia, CA). Three independent silique RNA extractions were collected for each line, and the two contrasting RNA sets were prepared as Cy5- and Cy3- (reactive water-soluble fluorescent dyes of the cyanine dye family) labeled cDNA probe pairs. RNA amplification, labeling with Cy3- or Cy5-dCTP dyes (GE Healthcare, Buckingamshire, U.K.), and probe fragmentation were carried out using an Ambion AmnoAllyl MessageAmp II RNA amplification kit according to the instructions of the manufacturer (Ambion, Austin, TX). A dye swap (Cy3/Cy5) experiment was performed for each biological replicate. The *B. napus* oligoarray was hybridized with the Cy5- and Cy3-labeled probe pairs at 65 °C in a solution of 25% formamide, 5× SSC (150 mM sodium chloride and 15 mM sodium citrate), 0.1% sodium dodecyl sulfate (SDS), and 0.1 mg/mL sonicated salmon sperm DNA at 65 °C for 17 h in a MAUI hybridization station (BioMicro Systems, Salt Lake City, UT). Labeling, hybridization, and post-hybridization washing were conducted according to directions in the CyScribe post-labeling kit (GE Healthcare, Piscataway, NJ). After the post-hybridization washes, slides were scanned with the Genepix 4000 (Axon, CA). Image analysis and feature extractions were performed with ArrayPro analyzer software (Media Cybernetics, Inc., Bethesda, MD). The intensity of each spot at $\lambda_{546\ nm}$ (Cy5) and $\lambda_{647\ nm}$ (Cy3) was transformed into a yellow-seeded/brown-seeded ratio. Initial data processing was performed using tools available in a BASE database (*12*). The filtered data were analyzed using Gene-Spring, version 6.1 (Silicon Genetics, Redwood City, CA). Transcripts showing increased or reduced expression were highlighted on the array, which previously had been annotated using the *Arabidopsis* genome using BLASTn and gene ontology. Cluster analysis of *Arabidopsis* Transparent Testa loci and *Arabidopsis* loci homologous to

array genes alerted by the *B. carinata* transcriptome was conducted using ClusterX.

Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR). RNA aliquots from the microarray experiment were used in RT reactions with SuperScript III First-Strand Synthesis SuperMix (Invitrogen, Carlsbad, CA) according to the instructions of the manufacturer. Gene *BNACT2* was chosen as the endogenous reference gene as described in Li et al. (*28*). Primer sequences for qRT-PCR were designed using online Perfect Oligo Design software provided by Invitrogen based on *Brassica* EST and homologous *Arabidopsis* cDNA sequences as described (*12*) (**Table 1**). The qRT-PCR mixtures contained 8 μL of diluted *B. carinata* cDNA (or 8 μL for control reactions), 10 μL of 2× SYBR Green qPCR Master Mix (Invitrogen), and 200 nM of each gene-specific primer in a final volume of 20 μL. The qRT-PCR reactions were performed using a StepOnePlus Real-Time PCR system (Applied Biosystems) as described (*28*) under the following conditions: 2 min at 50 °C, followed by 2 min at 95 °C, and 40 cycles of 95 °C (15 s) and 62 °C (30 s) in a 96-well optical reaction plate (Bio-Rad Laboratories, Hercules, CA). For each pair of primers, gel electrophoresis and melting curve analyses were performed to ensure that only a single PCR amplicon of the expected length and melting temperature was generated. Each sample was assayed in triplicate, and data were analyzed using the Step-One Software, version 2.0 (Applied Biosystems). The level of each mRNA was calculated using the mean threshold cycle (Ct) value and normalized to that of the reference gene *BnACT2*. All results were shown as the means of at least three biological replicates (RNA extracts) with corresponding standard deviations (SDs).

Table 1. qRT-PCR Primers Used in This Study

genes[a]	primers
BN24435-AT5G65690-PPC	forward: 5′-AGATGAAGACGCAAGGTGCT-3′ reverse: 5′-AACCTCCCAGCTTCAACAGA-3′
BN25290-AT5G45780-LRT	forward: 5′-CGGGTCACAGAACTCATCCT-3′ reverse: 5′-GAGGTTGAAGAGCGAGTTGG-3′
BN24247-AT5G46900-LTP	forward: 5′-ACCGCAAACCCACTTGTAAA-3′ reverse: 5′-GCCTTCAAAGCGGTACAGAG-3′
BN26434-AT3G19010-FLS	forward: 5′-TCCCATCCCATGATACCAAC-3′ reverse: 5′-TGAAGCGATCAATCTGGATG-3′
BN17561-AT5G07990-F3′H	forward: 5′-CGGTTACGGACGATTCAGTT-3′ reverse: 5′-GGATGCACAACCAAAGGAAC-3′
BN17805-AT2G03740-LEA	forward: 5′-TGATCAACGGAAGCAAGAGA-3′ reverse: 5′-TGGGGTCGTTCTTTGATTTT-3′
BN21104-AT2G41070-bZIP12	forward: 5′-CCCACCAGAAGAAGGAACAG-3′ reverse: 5′-ACCGAGTGTAGGCTGCTTGT-3′
BN23926-AT1G33700-unknown	forward: 5′-GATGTTACCGGATGGACGAG-3′ reverse: 5′-TGCCTCATAGATTCCCCTTG-3′
BN25866-AT1G25460-DFR	forward: 5′-GAAGGCACTTCTCGACTTGG-3′ reverse: 5′-TCAAAGCTTCCCTCCACTGT-3′
BN13710-AT5G54160-FOMT	forward: 5′-AATCATGCTCGACCGTATCC-3′ reverse: 5′-GGACACACCATCCTCGTTCT-3′
BN15497-AT3G51240-F3H	forward: 5′-TCGACGATGTTGGTGAGAAA-3′ reverse: 5′-CAGGAGGTAACGCGAAGAAG-3′
BNACT2	forward: 5′-CATCGGTGCTGAGAGATTCA-3′ reverse: 5′-CACTGAGCACGATGTTACCG-3′

[a] Full names of the genes are found in **Tables 2** and **3**.

Statistical Analysis. Phytochemical data were expressed as mean ± standard error on samples developed from three independent extractions, and calculated data was statistically analyzed [analysis of variation (ANOVA)] for least significance differences (LSD) at $p < 0.05$ using SAS 8.0 (SAS Institute, Inc., Cary, NC). For microarray analysis, background-corrected log ratio intensity values were scaled to have similar distribution and consistency across and among arrays. The calculation ANOVA model, $\log(y_{ijkgr}) = \mu + A_i + D_j + T_k + G_g + AG_{ig} + DG_{jg} + TG_{kg} + \varepsilon_{ijkr}$, was employed to detect differentially expressed genes using the normalized data according to Li et al. (*28*), where $\log(y_{ijkgr})$ represents the background-corrected and normalized natural logarithm of the intensity of the rth replicate of gene g on array i, with dye j and treatment/condition k, μ represents the average natural logarithm of gene intensity over all of the genes, arrays, and dyes, A, D, T, and G represent the array, dye, treatment, and main gene effects, respectively, and ε_{ijkr} represents normal distribution with a mean of 0.

RESULTS

Extractable and Unextractable PAs in Seed Coats. BuOH/HCl hydrolysis and DMACA histochemical staining were used to compare extractable and non-extractable PAs in the seed coats of *B. carinata*. Seed coats of the brown-seeded line accumulated 10-fold higher levels of extractable and non-extractable PAs than those of the yellow-seeded line (panels **A**−**C** of **Figure 2**). These differences could be seen readily when seed coats from the brown-seeded line changed from a red−brown color to a purple−black color when stained with DMACA (**Figure 2E**). DMACA stimulates this color change when it binds to flavan-3′,4′-diols, including PA and its precursors. In contrast, seed coat tissue from the yellow-seeded line remained a light yellow color, except for very small intense purple−black spots scattered throughout the seed coat (indicated by arrows in **Figure 2D**).

Flavonoids, Phenylpropanoids, and Lignans in *B. carinata* Seed Coats. LC-time-of-flight (TOF)-MS of 70% acetone extracts of *B. carinata* seed coats showed 13 distinct peaks in the brown-seeded line, while the yellow-seeded line had 5 additional peaks (7, 9, 13, 15, and 18) not reported previously in *B. carinata* (**Figure 3**). Semi-preparative extraction in 70% MeOH led to the isolation of

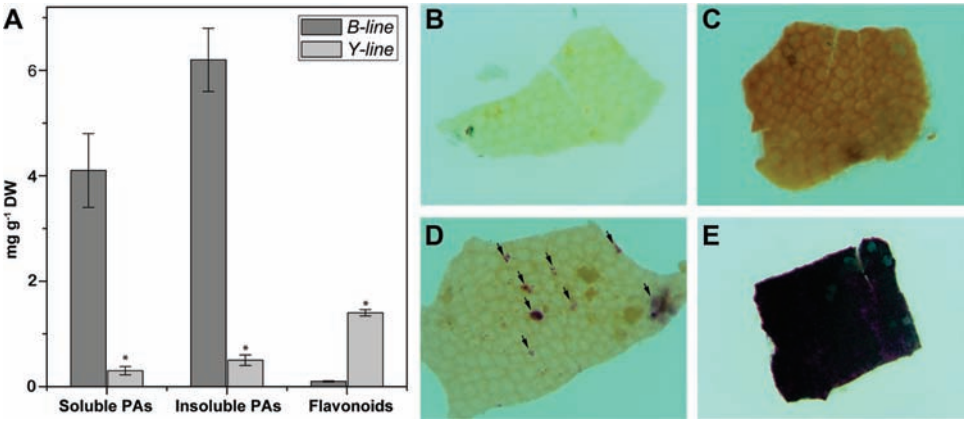

Figure 2. Analysis of extractable and non-extractable PAs and flavonoids in seed coats of brown- and yellow-seeded *B. carinata*. (A) Quantification of PAs and total flavonoids. Bars represent the standard error of the means for three independent extractions. Duncan's new multiple range tests were used to indicate significant differences (∗) between yellow- and brown-seeded seed coats at $p < 0.05$. (B) Unstained yellow-seeded seed coat. (C) Unstained brown-seeded seed coat. (D) DMACA-stained yellow-seeded seed coat. Arrows indicate DMACA-stained PA spots scattered over the pale tissue. (E) DMACA-stained brown-seeded seed coat.

9 of these compounds, with structures identified by retention time, UV spectroscopic parameters, MS/MS fragmentation patterns, and NMR data compared to commercial standards or from values published by other laboratories (*15*−*25*). Our results indicated that the yellow-seeded seed coat extracts accumulated flavonoids, phenylpropanoids, and lignans, while the brown-seeded seed coat extracts accumulated only phenylpropanoids and lignans (**Figures 3** and **4**). Flavonoids newly accumulated in the yellow-seeded seed coats included kaempferol 3-sophoroside 7-rhamnoside (peak 7), kaempferol 3-glucoside 7-rhamnoside (peak 13), kaempferol 3-glucoside 7-xyloside (peak 15), and kaempferol 3-xyloside 7-rhamnoside (peak 18) (**Figure 4**). Phenylpropanoids and identifiable lignans, which were unchanged in both seed coat types, were determined to be 1,2-disinapoylgentiobiose (peak 14), 1,2-disinapoylglucose (peak 16), pinoresinol diglucopyranoside (peak 1), 7*S*,8*R*,8′*R*-(−)-lariciresinol-4,4′-bis-*O*-*β*-D-glucopyranoside (peak 8), and 3,3′,4,4′,9-pentahydroxy-7,9′-epoxylignan 3,3′-dimethyl ether, 4-glucopyranoside (peak 10).

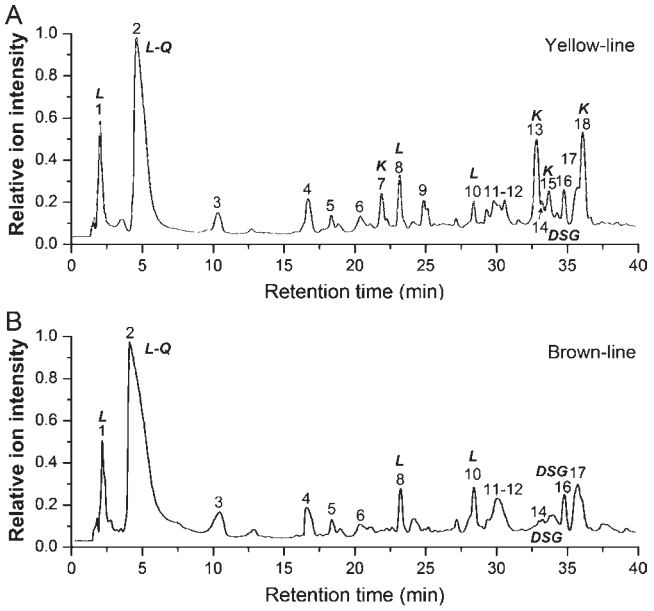

Figure 3. Representative LC−MS total ion chromatogram of 70% acetone extract of brown- and yellow-seeded *B. carinata* seed coat. K, kaempferol glycosides; L, lignans; S, sinapoylglucosides; L−Q, lignan quercetin mixture. LC−MS/MS and UV spectroscopic data for labeled peaks were consistent with literature values (*15*−*25*). Peak 1, pinoresinol diglucopyranoside (L); peak 2, mixture of lignin fragments and quercetin derivatives with sizes ranging between 800 and 1600 Da (L−Q); peaks 3−6, not elucidated; peak 7, kaempferol 3-sophoroside 7-rhamnoside (K); peak 8, 7*S*,8*R*,8′*R*-(−)-lariciresinol-4,4′-bis-*O*-glucopyranoside (L); peak 9, not elucidated; peak 10, 3,3′,4,4′,9-pentahydroxy-7,9′-epoxylignan 3,3′-dimethyl ether, 4-glucopyranoside (L); peaks 11 and 12, not elucidated; peak 13, kaempferol 3-glucoside 7-rhamnoside (K); peak 14, 1,2-disinapoylgentiobiose (S); peak 15, kaempferol 3-glucoside 7-xyloside (K); peak 16, 1,2-disinapoylglucose (S); peak 17, not elucidated; peak 18, kaempferol 3-xyloside 7-rhamnoside (K).

Both yellow- and brown-seeded seed coats accumulated large amounts of a single HPLC peak (peak 2). This major peak contained a mixture of compounds indicated by its UV spectroscopic properties of lignan fragments and quercetin derivatives and was composed of mixed species with molecular sizes ranging between 800 and 1600 Da after LC−MS analysis. The peak was 0.4-fold lower in yellow-seeded seed coats than in brown-seeded tissues and could not be separated into individual components using either C_{18} or C_8 columns. Hence, the full identity of peak 2 components was not confirmed by MS/MS.

Expression Analysis Using Microarray and Q-PCR Analysis of Developing Siliques. A comparison of gene expression in *B. carinata* developing siliques 22 days after pollination (DAP) was conducted using a 15K gene *B. napus* oligoarray developed by the Saskatoon Research Centre. A total of 1316 genes showed statistically significant upregulation, while 832 genes showed statistically significant downregulation. *B. carinata* genes determined by the array to be up- or downregulated ≥2-fold in the yellow-seeded line relative to the brown-seeded line were catalogued and placed into nine categories based on gene function annotated from the closest *Arabidopsis* homologue (**Tables 2** and **3**). Although strong differences between the two lines were noted for individual genes, the numbers of genes with differential expression in the yellow-seeded siliques relative to the brown-seeded siliques were quite similar in the categories of defense, pathogenesis, hormone-related, and aging (10 versus 7, respectively), primary metabolism (12 versus 9, respectively), secondary metabolism (15 versus 13, respectively), transport (3 versus 3, respectively), and others with some type of functional identity (37 versus 41, respectively). In contrast, a 2.5-fold higher number of signal transduction genes (13 versus 5, respectively), transcription factors (19 versus 8, respectively), and proteins with unknown functions (34 versus 14, respectively) were upregulated and downregulated in the yellow-seeded siliques relative to the brown-seeded siliques. Upregulated genes also had more extreme expression profiles than downregulated genes in the yellow-seeded siliques relative to brown-seeded siliques. Q-PCR of 11 of these genes was consistent with the microarray data (**Figure 5**). Expression of genes related to *Brassica*-specific compounds was mainly stable. Only a myrosinase-associated protein BN20003 (homologous to At1g54000) was

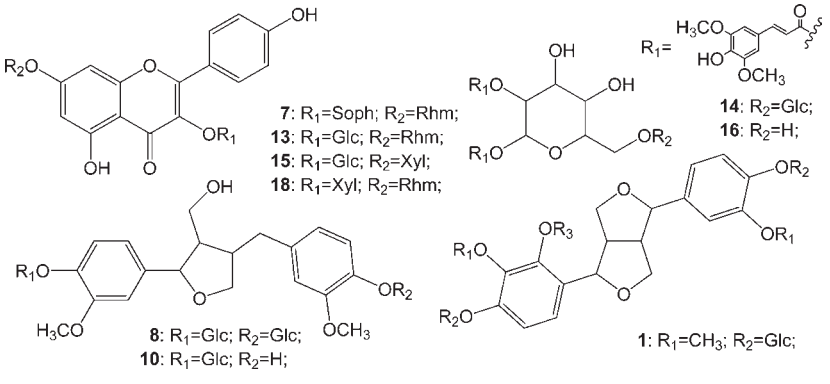

Figure 4. Structures of flavonoids, lignans, and phenylpropanoids found in yellow-seeded *B. carinata* seed coat. Soph, sophoroside; Rhm, rhamnoside; Glc, glucoside; Xyl, xyloside.

Table 2. Genes Upregulated ≥2-Fold in Yellow-Seeded *B. carinata* Developing Siliques Compared to a Brown-Seeded Line

oligo ID	*Arabidopsis* homologue, putative function	fold change[a]	*p* value[b]	oligo ID	*Arabidopsis* homolgue, putative function	fold change[a]	*p* value[b]
			Cell Rescue, Pathogenesis, Cell Wall, Hormone-Related, Defense, and Aging				
BN25488	At4g32280, auxin-responsive AUX/IAA family protein	5.24	0.0326	BN20290	At3g13550, ubiquitin-conjugating enzyme (COP10)	4.07	0.0272
BN12302	At3g07840, polygalacturonase	2.99	0.0071	BN20738	At5g59720, 18.1 kDa class I heat-shock protein (HSP18.1-CI)	2.50	0.0077
BN12645	At1g32170, xyloglucan:xyloglucosyl transferase	2.46	0.0061	BN27646	At1g57590, pectinacetylesterase	2.39	0.0068
BN11206	At4g05320, polyubiquitin (UBQ10) (SEN3) senescence-associated protein	2.37	0.0193	BN24466	At3g24020, disease resistance-responsive family protein	2.22	0.0098
BN13585	At1g57590, pectinacetylesterase	2.19	0.0081	BN27019	At5g08350, GRAM domain-containing protein/ABA-responsive protein-related	2.14	0.0200
			Signal Transduction				
BN24435[c]	At5g65690, phosphoenolpyruvate carboxykinase (**PPC**)	17.03	0.0002	BN25290[c]	At5g45780, leucine-rich repeat transmembrane protein kinase (**LRRT**)	9.42	0.0364
BN20275	At5g53080, kinesin light chain related	4.08	0.0015	BN24868	At1g14380, calmodulin-binding family protein	3.40	0.0188
BN22525	At5g44100, casein kinase	3.19	0.0157	BN25491	At3g60720, receptor-like protein kinase	3.16	0.0090
BN22068	At1g66980, protein kinase family protein	3.11	0.0416	BN26515	At5g10900, calcineurin-like phosphoesterase family protein	3.04	0.0211
BN21493	At4g11310, cysteine proteinase	2.67	0.0082	BN24994	At1g12040, leucine-rich repeat family protein	2.66	0.0063
BN25202	At4g37840, hexokinase	2.50	0.0072	BN26905	At1g16260, protein kinase family protein	2.36	0.0035
BN20994	At1g01140, CBL-interacting protein kinase 9 (CIPK9)	2.21	0.0051	BN23190	At2g02950, phytochrome kinase substrate 1 (PKS1)	2.10	0.0074
BN18051	At3g21930, receptor-like protein kinase-related	2.08	0.0073				
			Primary Metabolism and Lipid Genes				
BN24247[c]	At5g46900, protease inhibitor/seed storage/lipid transfer protein (**LTP1**)	17.81	0.0013	BN24248	At5g46910, protease inhibitor/seed storage/lipid transfer protein (LTP)	3.55	0.0055
BN23138	At1g14030, ribulose-1,5 bisphosphate carboxylase oxygenase	2.89	0.0041	BN11195	At2g31320, poly(ADP-ribose) polymerase	2.81	0.0043
BN17773	At2g19010, GDSL-motif lipase/hydrolase family protein	2.76	0.0290	BN12198	At1g20130, GDSL-like lipase/acylhydrolase	2.67	0.0027
BN13344	At2g10940, protease inhibitor/seed storage/lipid transfer protein (LTP)	2.60	0.0025	BN23418	At5g07520, glycine-rich protein (GRP18) oleosin	2.47	0.0082
BN20530	At5g40420, glycine-rich protein/oleosin	2.38	0.0011	BN27225	At2g07687, cytochrome *c* oxidase	2.17	0.0016
BN22754	At5g03610, GDSL-motif lipase/hydrolase family protein	2.15	0.0008	BN11159	At4g23600, coronatine-responsive tyrosine aminotransferase	2.12	0.0380
			Secondary Metabolism				
BN15162	At5g49360, glycosyl hydrolase family 3 protein	3.57	0.0062	BN24594	At3g25770, allene oxide cyclase	2.91	0.0183
BN21263	At1g30000, glycoside hydrolase family 47 protein	2.85	0.0015	BN24468	At1g47600, glycosyl hydrolase family 1 protein	2.71	0.0055
BN25525	At1g35513, isochorismate synthase-related	2.67	0.0354	BN18584	At3g26840, esterase/lipase/thioesterase family protein	2.66	0.0094
BN26434[c]	At3g19010, 2OG-Fe^{II} oxygenase family protein similarity to flavonol synthase (**FLS**)	2.57	0.0078	BN20826	At4g10490, 2OG-Fe^{II} oxygenase family protein	2.52	0.0017
BN13729	At2g43910, thiol methyltransferase	2.42	0.0139	BN23299	At2g19070, anthranilate *N*-hydroxycinnamoyl/benzoyltransferase	2.39	0.0019
BN17561[c]	At5g07990, flavonoid 3′-monooxygenase/flavonoid 3′-hydroxylase (**F3′H**, ***TT7***)	2.35	0.0382	BN18140	At5g65205, short-chain dehydrogenase/reductase (SDR) family protein	2.32	0.0233
BN25322	At5g44630, terpene synthase/cyclase family protein	2.26	0.0075	BN27228	At2g37540, short-chain dehydrogenase/reductase (SDR)	2.17	0.011
BN13932	At4g35670, glycoside hydrolase family 28 protein	2.12	0.0072				
			Cytochrome P450				
BN21432	At1g01190, cytochrome P450	3.28	0.0050	BN24671	At2g21910, cytochrome P450	3.17	0.0335
BN27227	At1g13080, cytochrome P450	2.12	0.0277				
			Transcription Factor				
BN26979	At2g24650, transcriptional factor B3 family protein	4.88	0.0285	BN24410	At1g69810, WRKY family transcription factor	3.26	0.0075
BN18693	At5g10030, bZIP family transcription factor (OBF4)	3.03	0.0272	BN13133	At5g37280, zinc finger (CCCH-type/C_3HC_4-type RING finger) family protein	2.97	0.0035
BN12727	At3g08505,, zinc finger (CCCH-type/C_3HC_4-type RING finger) family protein	2.95	0.0020	BN18807	At1g68190, zinc finger (B-box type) family protein	2.87	0.0008
BN18034	At2g45540, WD-40 repeat family protein	2.80	0.0123	BN21168	At1g29800, zinc finger (FYVE type) family protein	2.75	0.0039
BN22451	At4g25440, WD-40 repeat family protein	2.74	0.0029	BN22680	At4g34410, AP2 domain-containing transcription factor	2.70	0.0116
BN24368	At4g21850, similarity to pilin-like transcription factor	2.68	0.0044	BN16004	At5g05610, PHD finger family protein	2.65	0.0044
BN16150	At2g16910, basic helix−loop−helix (bHLH) family protein	2.52	0.0056	BN18929	At5g59780, myb family transcription factor (MYB59)	2.47	0.0022

Table 2. Continued

oligo ID	*Arabidopsis* homologue, putative function	fold change[a]	*p* value[b]	oligo ID	*Arabidopsis* homolgue, putative function	fold change[a]	*p* value[b]
BN18157	At3g57230, MADS-box transcription factor DEFH125	2.29	0.0342	BN25629	At5g11590, AP2 domain-containing transcription factor	2.28	0.0142
BN18946	At2g20570, golden2-like transcription factor (GLK1)	2.25	0.0095	BN13215	At5g06420, zinc finger (CCCH-type/C_3HC_4-type RING finger) family protein	2.19	0.0006
BN25824	At3g21330, basic helix−loop−helix (bHLH) family protein	2.17	0.0134				
			Transport Facilitation				
BN25519	At5g57090, auxin transport protein (EIR1)	3.56	0.0019	BN25053	At2g42940, DNA-binding family protein contains an AT hook motif	2.70	0.0148
BN14042	At1g16820, vacuolar ATP synthase catalytic subunit related	2.34	0.0024	BN18410	At4g35440, voltage-gated chloride channel family protein	2.33	0.0104
BN14292	At3g28510, AAA-type ATPase family protein	2.28	0.0037				
			Others				
BN17167	At1g74210, glycerophosphoryl diester phosphodiesterase	5.99	0.0150	BN24171	At1g80240, expressed protein	5.18	0.0090
BN19033	At5g01300, phosphatidylethanolamine-binding family protein	3.80	0.0068	BN18562	At1g28190, expressed protein	3.60	0.0364
BN18974	At1g11440, expressed protein	3.57	0.0120	BN27348	At5g10770, chloroplast nucleoid DNA-binding protein	3.52	0.0372
BN25500	At1g34245, expressed protein	3.51	0.0085	BN27257	At1g10000, expressed protein	3.40	0.0165
BN24765	At2g23000, serine carboxypeptidase S10 family protein	3.37	0.0050	BN20788	At1g15825, hydroxyproline-rich glycoprotein	3.28	0.0025
BN15667	At4g11960, expressed protein	3.25	0.0033	BN20860	At2g31890, expressed protein	3.23	0.0072
BN19525	At3g03720, amino acid permease family protein	3.14	0.0202	BN20655	At1g23170, expressed protein	3.05	0.0024
BN25609	At1g35350, EXS family protein	3.02	0.0055	BN19182	At3g10040, expressed protein	3.01	0.0038
BN27466	At5g32620, expressed protein	3.00	0.0057	BN16007	At2g40430, expressed protein	2.99	0.0122
BN24469	At2g44220, expressed protein	2.92	0.0037	BN12598	At4g32760, VHS domain-containing protein	2.91	0.0007
BN15559	At3g11840, U-box domain-containing protein	2.87	0.0202	BN27332	At4g13430, aconitase family protein	2.84	0.0083
BN10288	At5g35680, eukaryotic translation initiation factor 1A	2.83	0.0008	BN23293	At3g26910, hydroxyproline-rich glycoprotein family protein	2.80	0.0187
BN24284	At2g24850, aminotransferase	2.74	0.0042	BN12471	At3g12000, S-locus-related protein SLR1	2.74	0.0017
BN15489	At1g17710, expressed protein	2.74	0.0313	BN16634	At3g16400, jacalin lectin family protein	2.73	0.0079
BN25453	At2g37980, expressed protein	2.70	0.0014	BN22808	At5g22330, TATA box-binding protein-interacting protein related	2.70	0.0001
BN20218	At3g06980, DEAD/DEAH box helicase	2.70	0.0143	BN18562	At1g28190, expressed protein	2.70	0.0095
BN18374	At2g46860, inorganic pyrophosphatase	2.67	0.0011	BN27550	At2g15815, expressed protein	2.66	0.0046
BN18623	At5g24316, proline-rich family protein	2.66	0.0086	BN25098	At1g31885, major intrinsic family protein	2.65	0.0005
BN25413	At3g29010, expressed protein	2.64	0.0080	BN20972	At1g28410, expressed protein	2.59	0.0225
BN27513	At2g20310, expressed protein	2.57	0.0034	BN20859	At2g31890, expressed protein	2.56	0.0024
BN25717	At1g50890, expressed protein	2.54	0.0195	BN20945	At4g22840, bile acid/sodium symporter family protein	2.52	0.0380
BN12678	At3g19440, pseudo-uridine synthase family protein	2.49	0.0025	BN16082	At5g60180, pumilio/Puf RNA-binding domain-containing protein	2.47	0.0033
BN25500	At1g34245, expressed protein	2.46	0.0045	BN21283	At5g52590, RabGAP/TBC domain-containing protein	2.44	0.0002
BN21735	At3g56290, expressed protein	2.38	0.0047	BN25640	At1g13540, expressed protein	2.35	0.0185
BN17898	At2g25110, MIR domain-containing protein	2.34	0.0017	BN25581	At2g21950, SKP1-interacting partner 6 (SKIP6)	2.33	0.0040
BN24171	At1g80240, expressed protein	2.30	0.0279	BN26936	At5g08710, regulator of chromosome condensation (RCC1)	2.30	0.0088
BN22490	At4g38550, expressed protein	2.29	0.0091	BN18016	At5g44320, eukaryotic translation initiation factor 3 subunit 7	2.29	0.0098
BN22785	At5g54850, expressed protein	2.24	0.0134	BN20622	At5g07200, gibberellin 20-oxidase	2.23	0.0239
BN10988	At1g29330, ER lumen protein-retaining receptor (ERD2)	2.23	0.0071	BN12468	At1g60770, pentatricopeptide (PPR) repeat-containing protein	2.22	0.0092
BN11330	At3g60310, expressed protein	2.22	0.0283	BN27508	At5g06280, expressed protein	2.20	0.0323
BN11461	At5g10710, expressed protein	2.18	0.0267	BN17858	At3g03760, LOB domain protein 20	2.17	0.0451
BN20655	At1g23170, expressed protein	2.16	0.0208	BN21276	At1g08570, thioredoxin family protein	2.16	0.0078
BN25579	At3g55240, expressed protein	2.15	0.0146	BN10310	At5g44540, tapetum-specific protein related	2.15	0.0074
BN24762	At4g15930, dynein light chain	2.11	0.0021	BN24376	At2g47540, pollen Ole e 1 allergen and extensin family protein	2.11	0.0002
BN16007	At2g40430, expressed protein	2.10	0.0043	BN19196	At1g05510, expressed protein	2.08	0.0064

[a] Fold change of gene expression in yellow-seeded developing siliques relative to brown-seeded developing siliques when tested on an Agriculture Canada 15K *B. napus* microarray. Data were expressed as the mean of three individually replicated experiments with independent RNA extractions. [b] *p* values indicate probability of differential expression between yellow- and brown-seeded samples at a significance level of 0.05. [c] Bold font indicate genes analyzed by Q-PCR in **Figure 5**.

Table 3. Genes Downregulated ≥2-Fold in Yellow-Seeded *B. carinata* Developing Siliques Relative to a Brown-Seeded Line

oligo ID	*Arabidopsis* homologue, putative function	fold change[a]	*p* value[b]	oligo ID	*Arabidopsis* homologue, putative functions	fold change[a]	*p* value[b]
			Cell Rescue, Pathogenesis, Cell Wall, Hormone-Related, Defense, and Aging				
BN17805[c]	At2g03740, late embryogenesis abundant domain-containing protein (**LEA**)	3.42	0.0404	BN10675	At1g75030, pathogenesis-related thaumatin family protein	2.57	0.0102
BN14490	At1g75830, plant defensin-fusion protein, putative (PDF1.1)	2.51	0.0015	BN10431	At5g66400, dehydrin (RAB18)	2.18	0.0482
BN27158	At5g06410, DNAJ heat-shock N-terminal domain-containing protein	2.05	0.0023	BN20580	At5g07190, embryo-specific protein 3	2.01	0.0302
BN19121	At4g30280, xyloglucan:xyloglucosyl transferase	2.00	0.0406				
			Signal Transduction				
BN19606	At4g38470, protein kinase family protein	2.46	0.0066	BN13516	At2g41410, calmodulin	2.33	0.0145
BN18496	At4g19110, protein kinase	2.05	0.00251	BN21687	At1g07390, leucine-rich repeat family protein	2.04	0.0076
BN24856	At2g20630, protein phosphatase 2C	2.00	0.006				
			Primary Metabolism				
BN10912	At3g52130, protease inhibitor/seed storage/lipid transfer protein (LTP)	2.75	0.0078	BN21214	At4g35260, isocitrate dehydrogenase subunit 1	2.55	0.0261
BN24865	At2g42790, citrate synthase	2.42	0.0181	BN24095	At4g39210, glucose-1-phosphate adenylyltransferase	2.34	0.0054
BN12219	At2g38530, non-specific lipid transfer protein 2 (LTP2)	2.41	0.0067	BN23985	At1g60730, aldo/keto reductase family protein	2.32	0.0056
BN17746	At5g12200, dihydropyrimidinase/DHPase/dihydropyrimidine amidohydrolase/hydantoinase (PYD2)	2.27	0.0004	BN12223	At2g38540, non-specific lipid transfer protein 1 (LTP1)	2.22	0.0119
BN22122	At3g22200, 4-aminobutyrate aminotransferase	2.18	0.0093	BN21499	At5g50240, protein-L-isoaspartate *O*-methyltransferase	2.12	0.0019
			Secondary Metabolism				
BN26739	At5g05870, UDP-glucuronosyl/UDP-glucosyl transferase	2.81	0.0130	BN13774	At1g24170, glycosyl transferase family 8 protein	2.63	0.0019
BN23667	At5g60700, glycosyltransferase	2.44	0.0011	BN15497[c]	At3g51240, naringenin 3-dioxygenase/flavanone 3-hydroxylase (**F3H**, ***TT6***)	2.31	0.0031
BN20175	At1g18650, glycosyl hydrolase family protein	2.25	0.0015	BN15240	At4g16765, oxidoreductase, 2OG-Fe^{II} oxygenase	2.24	0.0065
BN25866[c]	At1g25460, dihydroflavonol 4-reductase, also annotated as a cinnamoyl CoA reductase (**DFR**)	2.23	0.0040	BN23297	At2g19070, anthranilate *N*-hydroxycinnamoyl/benzoyltransferase	2.19	0.0060
BN13710[c]	At5g54160, flavanol (quercetin) 3-*O*-methyltransferase, also annotated as a potential phenolic acid OMT (**FOMT**)	2.18	0.0080	BN25736	At5g17220, glutathione *S*-transferase *TT19*	2.12	0.0083
BN20789	At5g13930, chalcone synthase *TT4*	2.02	0.0061				
			Transcription Factor				
BN21104[c]	At2g41070, basic leucine zipper transcription factor (**bZIP12**)	6.36	0.0030	BN27589	At5g41580, zinc finger (MIZ type) family protein	2.34	0.0004
BN17280	At5g52660, myb family transcription factor	2.16	0.0442	BN23147	At2g41940, zinc finger (C_2H_2 type) family protein	2.11	0.0047
BN23889	At3g50060, myb family transcription factor	2.04	0.0095	BN19686	At2g40200, basic helix−loop−helix (bHLH) family protein	2.02	0.0027
BN16052	At2g31220, basic helix−loop−helix (bHLH) family protein	2.02	0.0284	BN14987	At5g24520, Transparent Testa Glabra 1 protein (*TTG1*)	2.02	0.0109
			Transport Facilitation				
				BN23805	At1g03905, ABC transporter family protein	2.20	0.0063
				BN19238	At3g63380, calcium-transporting ATPase	2.01	0.0057
			Others				
BN23926[c]	At1g33700, expressed (unknown) protein (**UKN**)	4.84	0.0062	BN12527	At3g07850, exopolygalacturonase	3.32	0.0080
BN16517	At5g53170, FtsH protease	3.18	0.0014	BN21386	At5g13440, ubiquinol-cytochrome *c* reductase iron−sulfur subunit	3.11	0.0039
BN10702	At3g16000, matrix-localized MAR DNA-binding protein related	2.96	0.0016	BN20003	At1g54000, myrosinase-associated protein	2.93	0.0031
BN27241	At1g72800, nuM1 related	2.69	0.0053	BN16173	At3g06100, major intrinsic family protein	2.67	0.0013
BN19698	At5g02880, HECT-domain-containing protein	2.64	0.0084	BN14088	At2g29310, tropinone reductase	2.55	0.0154
BN19101	At3g23820, NAD-dependent epimerase/dehydratase family protein	2.52	0.0335	BN27473	At1g75330, ornithine carbamoyltransferase	2.39	0.0247

Table 3. Continued

oligo ID	*Arabidopsis* homologue, putative function	fold change[a]	*p* value[b]	oligo ID	*Arabidopsis* homologue, putative functions	fold change[a]	*p* value[b]
BN27445	At5g47210, nuclear RNA-binding protein	2.38	0.0111	BN27149	At4g29090, reverse transcriptase	2.38	0.0025
BN26580	At1g03230, extracellular dermal glycoprotein	2.37	0.0498	BN17184	At3g60570, β-expansin	2.36	0.0015
BN22231	At5g40890, chloride channel protein (CLC-a)	2.35	0.0001	BN27586	At5g49210, expressed protein	2.34	0.0154
BN23654	At3g15310, expressed protein	2.30	0.0104	BN17538	At3g03090, sugar transporter family protein	2.27	0.0018
BN12555	At3g22840, chlorophyll A−B binding family protein	2.26	0.0027	BN12163	At2g42480, meprin and TRAF homology domain-containing protein	2.26	0.0200
BN23479	At1g10240, far-red impaired responsive protein	2.26	0.0116	BN25169	At2g31840, expressed protein	2.25	0.0070
BN26536	At5g39150, germin-like protein	2.24	0.0005	BN23550	At4g28760, expressed protein	2.23	0.0014
BN27592	At2g40290, eukaryotic translation initiation factor 2 subunit 1	2.22	0.0001	BN26050	At5g07730, expressed protein	2.21	0.0019
BN23665	At1g54380, spliceosome protein related	2.18	0.0080	BN27483	At3g23260, F-box family protein	2.14	0.0126
BN25475	At3g56220, expressed protein	2.17	0.0079	BN23689	At5g25500, expressed protein	2.14	
BN11066	At1g05270, TraB family protein	2.13	0.0130	BN15356	At3g48890, cytochrome b_5 domain-containing protein	2.12	0.0032
BN24019	At1g05320, myosin-related	2.11	0.0141	BN22129	At4g24750, expressed protein	2.11	0.0057
BN16213	At5g59750, riboflavin biosynthesis protein	2.10	0.0160	BN26241	At3g05210, nucleotide repair protein	2.10	0.0214
BN13511	At3g61760, dynamin-like protein B (DL1B)	2.09	0.0071	BN23874	At1g71020, armadillo/β-catenin repeat family protein	2.09	0.0034
BN17898	At2g25110, MIR domain-containing protein	2.09	0.0139	BN20600	At3g57990, expressed protein	2.09	0.0079
BN23782	At3g55620, eukaryotic translation initiation factor 6	2.09	0.0013	BN27582	At4g01980, expressed protein	2.08	0.0003
BN18439	At5g43570, serine protease inhibitor	2.08	0.0023	BN19925	At2g16050, DC1 domain-containing protein	2.07	0.0210
BN10518	At1g68620, expressed protein	2.07	0.0044	BN24508	At2g04940, scramblase-related	2.07	0.0034
BN10691	At4g19210, RNase L inhibitor protein	2.06	0.0028	BN23870	At2g27900, expressed protein	2.04	0.0125
BN23750	At5g10630, elongation factor 1-α	2.04	0.0009	BN10809	At2g36530, enolase	2.02	0.0028
BN18111	At3g45630, RNA recognition motif (RRM)-containing protein	2.02	0.0077	BN23586	At1g17530, mitochondrial import inner membrane translocase subunit	2.02	0.0066
BN27262	At5g66470, expressed protein	2.02	0.0138	BN11085	At5g22950, SNF7 family protein	2.00	0.0007

[a] Fold change of gene expression in yellow-seeded developing siliques versus brown-seeded developing siliques when tested on an Agriculture Canada 15K *B. napus* microarray. Data were expressed as the mean of three individually replicated experiments from independent RNA extractions. [b] *p* values indicate probability of differential expression between yellow- and brown-seeded samples at a significance level of 0.05. [c] Bold font indicate genes also tested by Q-PCR and shown in **Figure 5**.

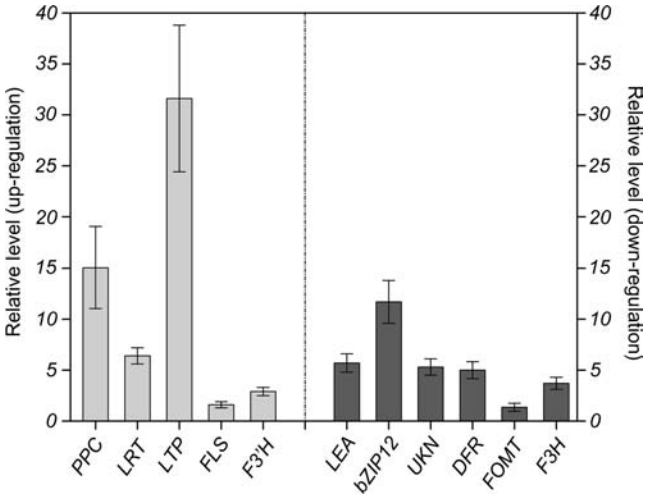

Figure 5. qRT-PCR of 11 genes confirming the microarray expression profile. Full names for the confirmed genes are found in **Tables 2** and **3**.

noteworthy (i.e., downregulated 2.9-fold in yellow-seeded siliques compared to brown-seeded siliques).

Because kaempferols were accumulated preferentially in seed coats of the yellow-seeded line, we examined the microarray data for changes in individual flavonoid and cytochrome P450 genes in developing siliques. A flavonol synthase (*FLS*), a flavonoid 3′ hydroxylase (*F3′H*) closely related to the *Arabidopsis* Transparent Testa 7 (*TT7*) *F3′H* gene At5g07990, and three cytochrome P450 genes were upregulated more than 2-fold in the yellow-seeded siliques relative to the brown-seeded siliques (**Table 2**). Q-PCR analysis of FLS and F3′H correlated well with the upregulated microarray data (**Figure 5**). A substantial number of flavonoid genes were also downregulated more than 2-fold in the yellow-seeded siliques. These included a chalcone synthase (*CHS*) closely related to At5g13930 (*TT4*), two distinct flavanone 3-hydroxylases (*F3H*) (including one closely related to At3g51240, *TT6*), a dihydroflavonol-4-reductase (*DFR*) related to At1g25460 (and also annotated as a potential cinnamoyl CoA reductase), Transparent Testa Glabra 1 (*TTG1*), glutathione *S*-transferase (*GST*) closely related to At5g17220 (*TT19*), and a flavonol-3-*O*-methyl transferase (*FOMT*) (**Table 3**). F3H, DFR, and FOMT expression as measured by Q-PCR correlated well with the downregulated microarray expression pattern for these three genes (**Figure 5**). Expression of *PAP1* and chalcone isomerase (*CHI*, *TT5*) was not significantly different between the lines (data not shown). Expression patterns for other *TT* genes [*TT1*, *TT2*, *TT8*, *TTG2*, *TT3* (*DFR*), *ANS* (*LDOX*), *BAN*, *AHA10*, and *TT10*] could not be measured, because oligos for them were not represented on the microarray.

Seven genes in the transcription factor category were more than 3-fold higher in expression in the yellow-seeded developing siliques compared to brown-seeded developing siliques (**Table 2**). These included a B3 transcription factor family protein (hybridizing to BN26979, homologous to At2g24650, ~5-fold up), a WRKY factor (hybridizing to BN24410, homologous to At1g69810, 3.3-fold up), a bZIP protein (hybridizing to BN18693, homologous to At5g10030, 3-fold up), and four zinc ring finger genes (ranging from 2.8- to 3-fold higher). WD40 repeat proteins and several other factors were also elevated in the yellow-seeded siliques. Genes in the transcription factor category that were downregulated in the yellow-seeded line included bZIP12 (hybridizing to BN21104, homologous to At2g41070, 6.3-fold down) (**Table 3**).

A number of known genes in categories other than secondary metabolism and transcription factors were highly upregulated in the yellow-seeded siliques (**Table 2**). These included a lipid transfer protein with protease inhibitor function (hybridizing to BN24247, homologous to At5g46900, 17.8-fold up), a phosphoenol pyruvate carboxylase (hybridizing to BN24435, homologous to At5g65690, 17-fold up), a leucine-rich repeat transmembrane kinase (hybridizing to BN25290, homologous to At5g45780, 9.4-fold up), a glycerophosphoryl diester phosphodiesterase (hybridizing to BN17169, homologous to At1g74210, 6-fold up), an auxin responsive protein in the AUX/IAA family (hybridizing to BN25488, homologous to At4g32280, 5.2-fold up), and a ubiquitin-conjugating enzyme COP10 (hybridizing to BN20290, homologous to At3g13550, 4-fold up). These six upregulated genes had the strongest differences in expression between the yellow- and brown-seeded siliques. Overall, downregulated genes had much less extreme expression differences.

A substantial portion of the transcriptome changes were genes for proteins with unknown function. A total of 33 of these unknown genes were upregulated ≥2-fold in yellow-seeded siliques (**Table 2**). Fewer genes of unknown function were downregulated in the yellow-seeded siliques compared to those that were upregulated. Noteworthy genes in this category included an expressed protein hybridizing to BN24171 (homologous to At1g80240) and 5.2-fold higher in the yellow-seeded line and an expressed protein hybridizing to BN23926 (homologous to At1g33700) and 4.8-fold lower in the yellow-seeded line.

To obtain a sense for whether "alerted" *B. carinata* genes were clustered into specific regions of the genome, *Arabidopsis* gene loci homologous to *B. napus* genes and highlighted by the changes in the *B. carinata* transcriptome were layered onto a physical representation of the *Arabidopsis* genome. Their relative positions were evaluated visually for clustering in *Arabidopsis* and *B. napus* based on segmental similarities between these two genomes (*29*). Although one area on the upper arm of At5g contained a cluster of 13 genes highlighted on the array, including *TT7* (*F3′H*), the remainder of the highlighted genes were scattered relatively evenly across the five *Arabidopsis* chromosomes (data not shown).

DISCUSSION

This study elucidated differences in seed coat phenolics and developing silique gene expression profiles for two genetically related lines of yellow- and brown-seeded *B. carinata*. We extended previous studies on phenolics in *Brassica* species seed coats (*4*, *5*, *26*, *27*) by quantifying phenylpropanoids, lignans, flavonol glycosides, and two types of PA (extractable and nonextractable) to give a more detailed "picture" of seed-coat-specific phenolic composition than previously known in this species. Five kaempferol glycosides recovered from the yellow-seeded seed coats have never been detected before in *B. carinata*. Although two of these studies showed that yellow-seeded seed coats of *B. carinata* have reduced lignin content (*4*, *5*), the present study shows that structurally related methanol-soluble compounds (sinapoyl glycosides and identifiable lignans) (*28*) did not change in either line. This reduction in lignin occurred without a change in phenylpropanoids (i.e., sinapoyl glycosides or lignans) and was distinct from the change in flavonoid composition, which accompanied the reduction of PA in the *B. carinata* yellow-seeded seed coats. Here, there was a shift in accumulation of new kaempferol glycosides in the yellow-seeded seed coats (compounds **7**, **13**, **15**, and **18**). Our phytochemical results are consistent with the rise in dihydroflavonols and flavonols observed earlier in yellow-seeded seed coats (*5*), although the rise in kaempferols in yellow-seeded seed coats was not proportional to the reduction in PAs and lignin.

The decrease in yellow-seeded silique transcripts for Transparent Testa genes *CHS* (*TT4*), *F3H* (*TT6*), *GST* (*TT19*), and the *WD40* regulatory gene, *TTG1*, and the increase in transcripts for FLS were consistent with the strong decrease in both types of PAs and the increase in kaempferol glycosides in yellow-seeded seed coats. In addition, they are consistent with the reduction in DFR expression noted in the earlier study of these lines (*5*). *B. carinata* genes homologous to *TT* genes could not be assessed, because they were not represented on the array. Genes specifying phenolic glycosides, lignans, or lignins on the microarray, i.e., phenylalanine ammonia lyase, cinnamoyl alcohol dehydrogenase, cinnamoyl CoA reductase, and dirigent proteins, were mainly not affected by the yellow-seeded trait, except potentially for one homologue of At1g25460 (annotated as both a DFR and a CCR) and a flavonol-*O*-methyl transferase (FOMT, homologous to At5g54160, also annotated as a phenolic acid OMT). These developing silique expression data, together with the fact that both seed coat types accumulate equivalent levels of lignans and phenolic glycosides, suggest that non-structural phenolics may accumulate through a different regulatory mechanism compared to seed coat structural phenolics, such as lignin. Regulatory gene(s), which specify PA and the dominant yellow-seeded color trait in *B. carinata*, may affect less known steps in lignin biosynthesis.

The increase in F3′H transcripts in the yellow-seeded developing siliques is consistent with the accumulation of quercetin in embryos of the near-isogenic yellow line in an earlier study (*5*) but does not correlate with the appearance of new kaempferol derivatives and the reduction in higher molecular-weight quercetin derivatives in HPLC peak 2. Increased kaempferol glycosides and reduced quercetin derivatives in the pale seed coats suggests that mono- and dihydroxylated flavonols (and therefore F3′H) may be unequally represented in developing embryos and developing seed coats in *B. carinata*. Additionally, lignans could be "tied up" in the larger molecular-weight lignan-type structures in HPLC peak 2. Correlations between flavonoid gene expression and phenolics will only be fully understood when the complete composition and proportion of quercetin and lignan derivatives in peak 2 are known for both genetically related lines and when the tissue-specific nature and copy number of F3′H in *B. carinata* is known.

Transcriptome analysis indicated elevated transcript profiles for a disproportionately large number of regulatory genes and proteins of unknown function in yellow-seeded developing siliques compared to brown-seeded developing siliques. These genes are inherited along with the yellow-seeded, reduced seed PA/lignin traits as a single dominant recombinant unit in *B. carinata*. The yellow-seeded trait in most other *Brassica* species in the triangle of U is specified by at least two and sometimes three recessive genes (*7*−*9*), although at least two semi-dominant yellow-seeded phenotypes are now known for *B. napus* (*14*). The dominant nature of the yellow-seeded trait and the disproportionate number of regulatory and unknown developing silique genes with increased expression suggest that a single global regulatory gene may control these upregulated genes. Because *B. carinata* has a duplicated genome, a duplicated regulatory gene may also be involved. The chalcone synthase gene in soybean has been shown to have a complex structure composed of duplicated copies in a sense orientation and one additional copy in an antisense orientation reducing their impact (*30*).

Determining the contribution of alerted *B. carinata* transcription factors and unknown genes to the dominant yellow-seeded phenotype could be accelerated by testing the phenotypes of *Arabidopsis* mutations made available to the public over the past 20 years, including knockdown/out and activation lines (*28*). The utility of *Arabidopsis* mutations will depend upon whether a dominant regulatory gene for the absence of PA exists in *Arabidopsis*. Global regulatory factors could also be determined using the "alerted" regulatory genes and "unknown" genes as bait in yeast two-hybrid screens and RNAi gene "knockout" studies or gene overexpression studies in *B. carinata*. It is possible that only small sequence differences exist between a global regulatory gene specifying the dominant dark seed coat color in other *Brassica* species and the semi-dominant yellow-seeded color found in *B. carinata*. For example, a small sequence change to the activation site of the *C1* anthocyanin regulatory gene changed it from an activator into an inhibitor (*31*). *B. napus* physical gene maps and bacterial artificial chromosome libraries also are available to dissect out the dominant yellow-seeded trait in *B. napus*. Physical maps for *B. carinata* will become available once the sequencing of the *Brassica* B genome is complete. Amplified fragment length polymorphism (AFLP) markers were recently developed to define genetic diversity in *B. carinata* (*32*,*33*). Expression changes for additional genes (*TT* and others) also could be determined using a new 90K *B. napus* Canadian oligonucleotide array (http://www.dotm.ca/) or a 105K *B. napus* oligoarray (Agilent Technologies) and with new arrays that include B genome sequences as they become available.

ACKNOWLEDGMENT

The authors thanks Dr. Isobel Parkin for editorial comments on an earlier version of the manuscript.

Supporting Information Available: ^{1}H NMR spectra of a 70% methanolic extract of (A) brown-seeded and (B) yellow-seeded *B. carinata* seed coats (Supplementary Figure 1) and compiled NMR data of flavonoids, lignans, and phenylpropanoids found in a yellow-seeded *B. carinata* seed coat (Supplementary Data Set 1). This material is available free of charge via the Internet at http://pubs.acs.org.

LITERATURE CITED

(1) Scarth, R.; Tang, J. Modification of *Brassica* oil using conventional and transgenic approaches. *Crop Sci*. **2006**, *46*, 1225−1236.
(2) Mnzava, N. A.; Msikita, W. W. Leaf yield response of Ethiopian mustard (*Brassica carinata* A Br.) selections to defoliation regimes. *Acta Hortic*. **1988**, *218*, 77−82.
(3) Bouaid, A.; Martinez, M.; Aracil, J. Production of biodiesel from bioethanol and *Brassica carinata* oil: Oxidation stability study. *Bioresour. Technol*. **2009**, *100*, 2234−2239.
(4) Marles, M. A. S.; Gruber, M. Histochemical characterisation of unextractable seed coat pigments and quantification of extractable lignin in the Brassicaceae. *J. Sci. Food Agric*. **2004**, *84*, 251−262.
(5) Marles, M. A. S.; Gruber, M. Y.; Scoles, G. J.; Muir, A. D. Pigmentation in the developing seed coat and seedling leaves of *Brassica carinata* is controlled at the dihydroflavonol reductase locus. *Phytochemistry* **2003**, *62*, 663−672.
(6) Getinet, A.; Rakow, G.; Downey, R. K. Seed colour inheritance in *B. carinata* A. Braun, cultivar S-67. *Plant Breed*. **1987**, *99*, 80−82.
(7) Nagaharu, U. Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jpn. J. Bot*. **1935**, *7*, 389−452.
(8) Somers, D. J.; Rakow, G.; Prabhu, V. K.; Friesen, K. R. Identification of a major gene and RAPD markers for yellow seed coat colour in *Brassica napus*. *Genome* **2001**, *44*, 1077−1082.
(9) Mohmood, T.; Rahman, M. H.; Stringham, G. R.; Raney, J. P.; Good, A. G. Molecular markers for seed colour in *Brassica juncea*. *Genome* **2005**, *48*, 755−760.
(10) Yu, O.; Jez, J. M. Nature's assembly line: Biosynthesis of simple phenylpropanoids and polyketides. *Plant J*. **2008**, *54*, 750−762.

(11) Gullett, N. P.; Ruhul Amin, A. R.; Bayraktar, S.; Pezzuto, J. M.; Shin, D. M.; Khuri, F. R.; Aggarwal, B. B.; Surh, Y. J.; Kucuk, O. Cancer prevention with natural compounds. *Semin. Oncol.* **2010**, *37*, 258−281.
(12) Li, X.; Gao, P.; Gjetvaj, B.; Westcott, N.; Gruber, M. Analysis of the metabolome and transcriptome of *Brassica carinata* seedlings after lithium chloride exposure. *Plant Sci.* **2009**, *177*, 68−80.
(13) Guttman, M. Localization of PAs using *in situ* hydrolysis with sulfuric acid. *Biotech. Histochem.* **1993**, *68*, 161−165.
(14) Getinet, A.; Rakow, G. Repression of seed coat pigmentation in Ethiopian mustard. *Can. J. Plant Sci.* **1997**, *77*, 501−505.
(15) Rosch, D.; Krumbein, A.; Mugge, C.; Kroh, L. W. Structural investigations of flavonol glycosides from area buckthorn (*Hippophae rhamnoides*) pomace by NMR spectroscopy and HPLC−ESI−MS^n. *J. Agric. Food Chem.* **2004**, *52*, 4039−4046.
(16) Cuca, L.; Martinez, J.; Monache, F. 7,9′-Epoxylignan and other constituents of *Zanthoxylum culantrillo*. *Phytochemistry* **1998**, *47*, 1437−1439.
(17) Wilson, K. E.; Wilson, M. I.; Greenberg, B. M. Identification of the flavonoid glycosides that accumulate in *Brassica napus* L. cv. Tobas specifically in response to ultraviolet B radiation. *Photochem. Photobiol.* **2008**, *67*, 547−553.
(18) Olsson, L. C.; Veit, M.; Weissenbock, G.; Bornman, J. F. Differential flavonoid response to enhanced UV-B radiation in *Brassica napus*. *Phytochemistry* **1998**, *49*, 1021−1028.
(19) Sharaf, M.; El-Ansar, M. A.; Saleh, N. A. M. Flavonoids of four *Cleome* and three *Capparis* species. *Biochem. Syst. Ecol.* **1997**, *25*, 161−166.
(20) Price, K.; Casuscelli, F.; Colquhiuu, I.; Rhodes, M. Hydroxycinnamic acid esters from broccoli florets. *Phytochemistry* **1997**, *45*, 1683−1687.
(21) Bloor, S.; Falshaw, R. Covalently linked anthocyanin−flavonol pigments from blue *Agapanthus* flowers. *Phytochemistry* **2000**, *53*, 575−579.
(22) Fursa, N.; Litvinenko, V. Lepidoside from *Lepidium perfoliatum*. *Chem. Nat. Compd.* **1970**, *6*, 127−128.
(23) Baumert, A.; Milkowski, C.; Schmidt, J.; Nimtz, M.; Wray, V.; Strack, D. Formation of a complex pattern of sinapate esters in *Brassica napus* seeds, catalyzed by enzymes of a serine carboxypeptidase-like acyltransferase family? *Phytochemistry* **2005**, *66*, 1334−1345.
(24) Pedras, S. M.; Zheng, Q.; Gadagi, C.; Rimmer, R. S. Phytoalexins and polar metabolites from the oilseeds canola and rapeseed: Differential metabolic responses to the biotroph *Albugo candida* and to abiotic stress. *Phytochemistry* **2008**, *69*, 894−910.
(25) Yahara, S.; Domoto, H.; Sugimura, C.; Nohara, T.; Niiho, Y.; Nakajima, Y.; Ito, H. An alkaloid and two lignans from *Cuscuta chinensis*. *Phytochemistry* **1994**, *37*, 1755−1757.
(26) Naczk, M.; Nichols, T; Pink, D.; Sosulski, F. Condensed tannins in canola hulls. *J. Agric. Food Chem.* **1994**, *42*, 2196−2200.
(27) Naczk, M.; Amarowicz, R.; Pink, D.; Shahidi, F. Insoluble condensed tannins of canola/rapeseed. *J. Agric. Food Chem.* **2000**, *48*, 1758−1762.
(28) Li, X.; Gao, M. J.; Pan, H. Y.; Cui, D. J.; Gruber, M. Y. Purple canola: *Arabidopsis* PAP1 increases antioxidants and phenolics in *Brassica napus* leaves. *J. Agric. Food Chem.* **2010**, *58*, 1639−1645.
(29) Parkin, I.; Gulden, S. M.; Sharpe, A. G.; Lukens, L.; Trick, M.; Osborn, T. C.; Lydiate, D. J. Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics* **2005**, *171*, 765−781.
(30) Tuteja, J. H.; Vodkin, L. O. Structural features of the endogenous CHS silencing and target loci in the soybean genome. *Crop Sci.* **2008**, S49−S68.
(31) Paz-Arez, J.; Ghosal, D.; Saedler, H. Molecular analysis of the C1-I allele from *Zea mays*: A dominant mutant of the regulatory C1 locus. *EMBO J.* **1990**, *9*, 315−321.
(32) Genet, T.; Viljoen, C. D.; Labuschagne, M. T. Genetic analysis of Ethiopian mustard genotypes using amplified fragment length polymorphism (AFLP) markers. *Afr. J. Biotechnol.* **2005**, *4*, 891−897.
(33) Teklewold, A.; Becker, H. C. Geographic pattern of geneic diversity among 43 Ethiopian mustard (*Brassica carinata* A. Braun) accessions as revealed by RAPD analysis. *Genet. Resour. Crop Evol.* **2006**, *53*, 1173−1185.

Received for review June 8, 2010. Revised manuscript received August 24, 2010. Accepted September 4, 2010. X. Li was a recipient of a Visiting Fellowship to a Government of Canada Laboratory. This work was supported by the Canadian Crop Genomic Initiative, the Canadian Cellulosic Biofuel Network, and the Program for New Century Excellent Talents in University (X.L., NCE09-0423), Ministry of Education of the People's Republic of China.