Characterizing Regulatory and Functional Differentiation between Maize Mesophyll and Bundle Sheath Cells by Transcriptomic Analysis[1][W][OA]

Yao-Ming Chang, Wen-Yu Liu, Arthur Chun-Chieh Shih, Meng-Ni Shen, Chen-Hua Lu, Mei-Yeh Jade Lu, Hui-Wen Yang, Tzi-Yuan Wang, Sean C.-C. Chen, Stella Maris Chen, Wen-Hsiung Li*, and Maurice S.B. Ku

Biodiversity Research Center (Y.-M.C., W.-Y.L., M.-N.S., M.-Y.J.L., T.-Y.W., W.-H.L.), Genomics Research Center (Y.-M.C., W.-Y.L., S.M.C., W.-H.L.), and Institute of Information Science (A.C.-C.S., C.-H.L.), Academia Sinica, Taipei, Taiwan 115; Institute of Bioagricultural Science, National Chiayi University, Chiayi, Taiwan 600 (H.-W.Y., M.S.B.K.); Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637 (S.C.-C.C., W.-H.L.); and School of Biological Sciences, Washington State University, Pullman, Washington 99164–4238 (M.S.B.K.)

To study the regulatory and functional differentiation between the mesophyll (M) and bundle sheath (BS) cells of maize (*Zea mays*), we isolated large quantities of highly homogeneous M and BS cells from newly matured second leaves for transcriptome profiling by RNA sequencing. A total of 52,421 annotated genes with at least one read were found in the two transcriptomes. Defining a gene with more than one read per kilobase per million mapped reads as expressed, we identified 18,482 expressed genes; 14,972 were expressed in M cells, including 53 M-enriched transcription factor (TF) genes, whereas 17,269 were expressed in BS cells, including 214 BS-enriched TF genes. Interestingly, many TF gene families show a conspicuous BS preference in expression. Pathway analyses reveal differentiation between the two cell types in various functional categories, with the M cells playing more important roles in light reaction, protein synthesis and folding, tetrapyrrole synthesis, and RNA binding, while the BS cells specialize in transport, signaling, protein degradation and posttranslational modification, major carbon, hydrogen, and oxygen metabolism, cell division and organization, and development. Genes coding for several transporters involved in the shuttle of C_4 metabolites and BS cell wall development have been identified, to our knowledge, for the first time. This comprehensive data set will be useful for studying M/BS differentiation in regulation and function.

C_4 plants, with few exceptions, require the coordination of the mesophyll (M) and bundle sheath (BS) cells, arranged in a wreath structure called Kranz leaf anatomy (Hatch and Agostino, 1992), to confer high rates of photosynthesis. The initial carboxylation phase of the C_4 pathway takes place in the M cells, while the decarboxylation phase is restricted to the BS cells. The high photosynthetic capacity of C_4 plants implies a massive efflux of C_4-related metabolites between M and BS cells and between the cytosol and organelles in each cell type (Weber and von Caemmerer, 2010).

Although research in the past few decades has greatly increased our understanding of the biochemical reactions and the enzymes involved in the C_4 pathway of photosynthesis, little is known about the specific genes involved in the development of the Kranz leaf anatomy, the C_4 biochemical pathway, and the underlying regulatory mechanisms for the high-level expression of C_4-specific genes in a cell-, organ-, or development-specific manner. With the advancement in genomics, the genomic sequences of several C_3 and C_4 model plants have become available. These advances have allowed in-depth comparative proteomic and transcriptomic analyses of the whole leaves of typical C_3 and C_4 plants and their closely related C_3-C_4 intermediate species of *Cleome* and *Flaveria* (Gowik et al., 2011). These comparative studies allow deduction on how many genes are required to make a C_4 plant and possibly on how they may have been regulated at the genetic level. In addition, attempts have been made recently to characterize the transcriptomic profile of maize (*Zea mays*) leaf in a development-dependent manner (Li et al., 2010). Also, the laser-capture microdissection (LCM) technique was used to isolate both M and BS cells from mature maize leaves for transcriptome analysis. However, LCM has a limitation in the amount of cells that can be isolated and has the potential of cross-contamination.

To gain a comprehensive understanding of the C_4 photosynthesis pathway and related processes, we have modified current methods to isolate large quantities of

[1] This work was supported by Academia Sinica and the National Science Council, Taiwan (grant nos. NSC 99–2321–B–001–041–MY2 and 99B1–114).

* Corresponding author; e-mail whli@sinica.edu.tw.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Wen-Hsiung Li (whli@sinica.edu.tw).

[W] The online version of this article contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.112.203810

highly purified M and BS cells from mature maize leaves for RNA extraction and detailed transcriptome analyses of M and BS cells. Our study reveals the functional differentiation and coordination of the two cell types in various metabolic pathways and provides insights into the possible regulatory network underlying the development of Kranz leaf anatomy and the operation of the photosynthetic and other metabolic pathways in the NADP-malic enzyme (ME) subtype C_4 plants.

RESULTS AND DISCUSSION

Isolation of Highly Pure M Protoplasts and BS Strands

In this study, highly purified mesophyll protoplasts (MP) and bundle sheath strands (BSS) were obtained in large quantities from newly matured second leaves (Supplemental Fig. S1) for transcriptome analyses. Examination by light microscopy showed that the mechanically purified BSS was attached with very little M tissue, except M cell wall and vascular and epidermal tissues. To further assess the purity of the isolated cells, the activities of two key C_4 photosynthesis marker enzymes for M and BS cells were assayed. On a chlorophyll basis, the high activity of NADP-ME, a key C_4 acid decarboxylase located in the BS cells of maize, was almost exclusively associated with the BSS isolated (Supplemental Table S1). On the other hand, the high activity of phospho*enol*pyruvate carboxylase (PEPC) was mainly associated with the MP isolated by enzymatic digestion, although a low activity of PEPC was also detected in the BSS isolated, which might be due to the presence of the C_3-type PEPC in the BS cells. Immunoblot analysis using specific antibodies raised against the maize proteins also demonstrated the high homogeneity of the two cell types isolated (Supplemental Fig. S2).

Raw Read Processing and Definition of Expressed Genes

To clean the sequencing raw reads, we deleted reads that contain any N (unknown nucleotide) or have a poor quality score (Table I). Using Tophat version 1.3.3 (Trapnell et al., 2009; http://tophat.cbcb.umd.edu/), we found 78% and 79% of the cleaned reads can be mapped to the maize B73 genome (Maize Genome AGP version 2, release 5a.59).

We used Cufflinks version 1.2.1 (Trapnell et al., 2010; http://cufflinks.cbcb.umd.edu/) to calculate the expression level of each gene in terms of reads per kilobase per million mapped reads (RPKM; Mortazavi et al., 2008). RPKM is normalized by the length of the exon and the total mappable reads and can be used to quantify the expression level of a gene from the mapped read count. A total of 52,421 annotated genes with at least one read were found in the two transcriptomes (Table II). (All genes with at least one read are listed in Supplemental Data S1 with their RPKM values.) The maximum RPKM were 35,421 and 31,561, and the median RPKM were 0.55 and 0.73 in the M and BS cell transcriptomes, respectively. Only approximately 2% of the genes had a very high expression level (RPKM $\geq$ 100) in either M or BS cells.

The quality of the transcriptomes, the potential contamination of BS cells by epidermal and vascular bundle tissues, and the potential effects of long incubation time during the isolation of M protoplasts are discussed in Supplemental Materials and Methods S1 and Supplemental Fig. S3.

Following the criterion of Li et al. (2010), we defined an "expressed gene" as a gene with RPKM > 1 in the transcriptome. In all analyses below, only expressed genes were considered. If the Filtered Gene Set (FGS) was chosen, we identified 14,972 and 17,269 expressed genes in the M and BS transcriptomes of maize, respectively (Fig. 1A). In total, we identified 18,482 genes expressed in either the M or BS transcriptome or both, which is 1,844 genes more than were identified by Li et al. (2010; Fig. 1A). If the Working Gene Set (WGS) was chosen, then we gained an additional 2,224 expressed genes (Fig. 1A).

Low Degrees of Cross-Contamination between the Two RNA Samples

It is important to know the degree of cross-contamination between the two transcriptomes, because cross-contamination can affect the estimation of the number of expressed genes in a transcriptome (one cell type) and the number of differentially expressed genes between the two transcriptomes (two cell types). To address this issue, we used the expression levels of 10 selected C_4 marker genes to estimate the degrees of cross-contamination in the two RNA samples from the isolated M cells and BS cells (Table III). The proteins encoded by these genes play key roles in C_4 photosynthesis and are assumed to be exclusively or almost exclusively expressed in only one cell type (Supplemental Data S1). Among the 10 selected genes, *NADP-ME* (GRMZM2G085019), *DCT2* (for dicarboxylate transporter; GRMZM2G086258), *PCK*

Table I. *Statistics of the transcriptome data*

Cell Type	Total Reads	Trimmed Reads	Mapped Reads (Mappable Rate)[a]	Unique Hits
M	177,363,768	125,841,884	98,048,470 (78%)	89,522,392
BS	193,418,692	135,111,940	107,090,765 (79%)	100,493,851

[a]The mappable rate is defined as the ratio of the mapped reads to the trimmed reads.

Table II. *Numbers of expressed genes, cell type-enriched genes, and TF genes in the two transcriptomes*

Parameter	M Cells	BS Cells	Total
Detected genes (at least one read)[a]	43,934	48,888	52,421
Expressed genes in the FGS	14,972	17,269	18,482
Expressed TF genes in the FGS	653	819	880
Enriched expressed genes in the FGS	1,545	3,422	4,967
Enriched expressed TF genes in the FGS	53	214	267

[a]The total numbers of annotated genes in the WGS and the FGS of AGPv2 (release 5a.59) were 110,028 and 39,656, respectively.

(GRMZM2G001696), and two *rbcS*s (GRMZM2G098520 and GRMZM2G113033) are supposed to be BS specific, while *PEPC* (GRMZM2G083841), *CA* (GRMZM2G121878), *NADP-MDH* (GRMZM2G129513), *OMT* (for 2-oxoglutarate/malate transporter; GRMZM2G383088), and *PPDK-RP* (GRMZM2G131286) are supposed to be M specific. Using Equations 2 and 4 as shown in "Materials and Methods," we obtained an upper estimate of c_1 = 1.2% for the contamination rate of M cells by BS cells and an upper estimate of c_2 = 4.5% for the contamination rate of BS cells by M cells (Table III). These are considered the upper estimates because the assumption of exclusive expression of a key C_4 gene in a cell-specific manner (e.g. PEPC in M cells) may not be absolutely true. In any case, these calculations suggest that the degrees of cross-contamination in our two RNA samples were reasonably low, which is crucial for us to accurately determine the expression profile of a given gene in a cell-specific manner.

We checked the above conclusion on the purity of our isolated M and BS cells by quantitative reverse transcription-PCR (Supplemental Table S2 and S3). For the three BS-specific genes studied (*rbcS-2*, *rbcS-4*, and *NADP-ME*), the respective M/BS expression ratios were 0.001, 0.049, and 0.004, which are indeed similar to the ratios obtained from the transcriptome data (0.022, 0.022, and 0.007). These data support the conclusion of very low contamination of our M cell sample by BS cells. For the three M-specific genes studied (*PEPC*, *NADP-MDH*, and *OMT*), the respective BS/M ratios were 0.055, 0.050, and 0.037, which are also very close to the BS/M ratios of 0.048, 0.039, and 0.045 from our transcriptome data (Supplemental Table S2), supporting an estimate of approximately 4.5% contamination of our BS sample by M cells.

Li et al. (2010) had reported two sets of deep-sequencing transcriptomic data from the maize M and BS cells isolated by LCM. In their data, nine out of the above 10 C_4 marker genes, except *PPDK-RP* (GRMZM2G131286), were listed as cell type-enriched genes. Applying the same equations to analyze their data, we estimated the upper contamination rates of c_1 = 9.2% and c_2 = 21.6% for their M and BS cell preparations. These results indicate that our isolation methods caused a much lower level of cross-contamination than the LCM method they used. Consequently, our data would give a more accurate estimation of genes differentially expressed between the M and BS cells.

Genes Differentially Expressed between the Two Cell Types

To estimate the differential expression of genes between the two cell types, we defined the degree of cell specificity of gene i as $R_i = |m_i - b_i| / \max(m_i, b_i)$, where m_i and b_i represent the RPKM of gene i in the M and BS RNA samples, respectively. By this definition, $R_i = 1$ if gene i is exclusively expressed in only one cell type, and $R_i = 0$ if gene i is equally expressed in the two cell types. To define "differentially expressed genes," we used the criterion of $R_i = 0.8$, which corresponds to a 5-fold difference in RPKM value between the two cell types. This criterion is more stringent than the criterion used by Li et al. (2010); 72% of their differentially expressed genes had only 4-fold or smaller differences. In the differentially expressed gene set, there were 1,545 and 3,422 genes enriched in the M and BS transcriptomes, respectively (Fig. 1B). That is, 4,697 (1,545 + 3,422) genes, or approximately 25% of all expressed genes in the FGS (18,482), were differentially expressed between the two cell types. In contrast, Li et al. (2010) identified only 3,427 (2,028 + 1,399) differentially expressed genes, or approximately 21% of all expressed genes (16,638), even though they used a looser criterion (Fig. 1B). The lower percentage could be due to the high cross-contamination rates in the study of Li et al. (2010; see above). In our data, there were many more BS-enriched genes than M-enriched genes (Table II). This observation is opposite to the Li et al. (2010) observation of more M- than BS-enriched genes. Only 30% and 16% of M- and BS-enriched genes in this study overlapped with the Li et al. (2010) data (Fig. 1B). Thus, many more cell-specific genes have been identified from our study. For genes with $R_i >$ 0.99, which can be taken as truly cell-specific genes, there were only 65 genes in M cells but 688 genes in BS cells (Supplemental Data S2). This result suggests that maize BS cells have assumed a more important metabolic role in C_4 photosynthesis and other metabolic processes.

The major difference in photosynthetic biochemistry between C_3 and C_4 plants is that in C_4 plants the CO_2-concentrating mechanism and the Calvin cycle are divided between M and BS cells. To achieve an effective CO_2-concentrating mechanism, genes that are involved in carbon fixation, such as *PEPC* (GRMZM2G083841), *CA* (GRMZM2G121878), *NADP-MDH* (GRMZM2G129513), and *PPDK-RP* (GRMZM2G131286), and plastid membrane

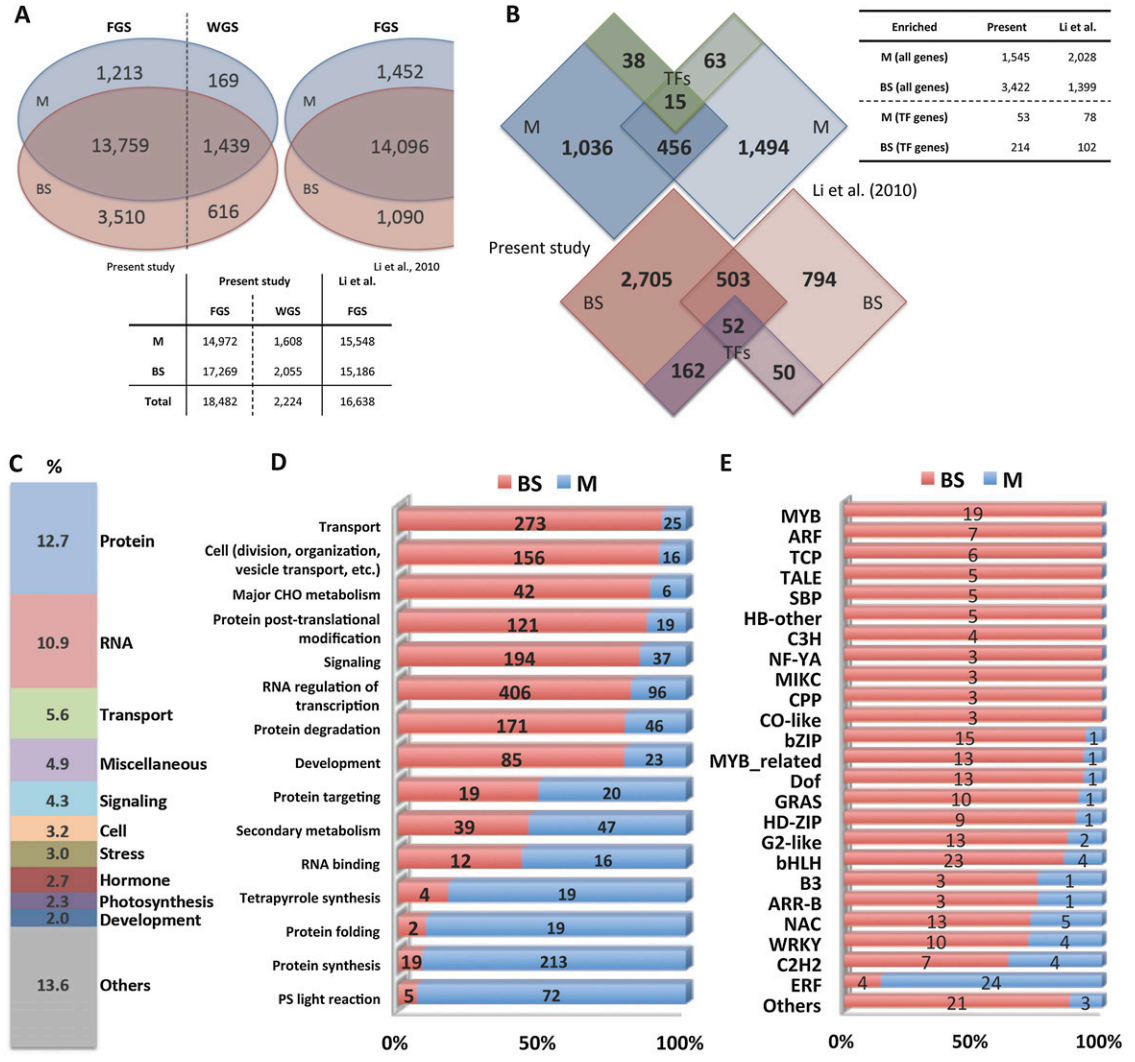

Figure 1. M- and BS-expressed genes and TF genes. A, Numbers of expressed genes in M and BS cells found in the FGS and the additional numbers of expressed genes found in the WGS, excluding transposable elements, gene fragments, and pseudogenes. In Li et al. (2010), no data were available for expressed genes found in the WGS. B, Venn diagrams of M- and BS-enriched genes and TFs compared between this study and Li et al. (2010). C, Functional distribution of all cell type-enriched genes. Genes of unknown function were not included. D, Distribution of functional preferences of cell type-enriched genes. In each row, the red and blue bars represent, respectively, the percentages of BS- and M-enriched genes within the functional category. The numbers in the red and blue bars represent the numbers of M- and BS-enriched genes. E, Distribution of cell type-enriched TF families. The red and blue bars represent the percentages of BS- and M-enriched TFs within each TF family, respectively. The numbers in the red and blue bars represent the numbers of M- and BS-enriched TFs.

transporters such as *OMT* (GRMZM2G383088), are all enriched or specifically expressed only in the M cells. In our M transcriptome, the RPKM values for *PEPC*, *CA*, *NADP-MDH*, *PPDK-RP*, and *OMT* were 9,485, 6,692, 2,621, 476, and 988, respectively, much higher than the corresponding values 431, 281, 95, 22, and 74 in the BS transcriptome (Table III). Therefore, the cell-specificity R_i values for these five genes were 0.95, 0.96, 0.96, 0.95, and 0.93, respectively. (These R_i values were probably underestimated because of the potential contamination of our BS cell sample by M cells.) On the other hand, the genes involved in releasing CO_2 from C_4 acid decarboxylation, such as *NADP-ME* (GRMZM2G085019) and *PCK* (GRMZM2G001696), plastid membrane transporter, such as *DCT2* (GRMZM2G086258), and the Calvin cycle, such as *rbcS*s (GRMZM2G098520 and GRMZM2G113033), were all enriched or specifically expressed in BS cells, each with $R_i > 0.98$ (Table III). Taken together, these results indicate that most of the known C_4 genes are very tightly regulated and specifically expressed in M or BS cells in mature maize leaves, allowing an effective CO_2-concentrating mechanism through effective carboxylation and decarboxylation in the two cell types. An exception to this rule is *PPDK* (GRMZM2G306345): it is preferentially expressed in M cells with a R_i of only 0.78. This is consistent with the

Table III. *Differential expression of C_4 photosynthesis-specific genes between M and BS cells, calculated from the transcriptome data*

Gene Identifier	Symbol	Our Data			Li et al. (2010)		
		M Cells	BS Cells	R_i	M Cells	BS Cells	R_i
GRMZM2G083841	PEPC	9,485	431	0.95	14,963	3,289	0.78
GRMZM2G121878	CA	6,692	281	0.96	10,615	2,261	0.79
GRMZM2G129513	NADP-MDH	2,621	95	0.96	2,523	531	0.79
GRMZM2G383088	OMT	476	22	0.95	13	5	0.65
GRMZM2G131286	PPDK-RP	988	74	0.93	0	0	0
GRMZM2G085019	NADP-ME	87	12,192	0.99	2,016	18,850	0.89
GRMZM2G086258	DCT2	7	545	0.99	1	7	0.93
GRMZM2G001696	PCK	21	7,801	1.00	1,431	12,987	0.89
GRMZM2G098520	rbsS-4	87	4,256	0.98	1,754	23,683	0.93
GRMZM2G113033	rbsS-2	109	5,280	0.98	1,323	15,420	0.91

biochemical study by Aoyagi and Nakamoto (1985), which found substantial activity and protein of PPDK present in the maize BS cells. The conversion of phospho*enol*pyruvate (PEP) to pyruvate and ATP catalyzed by the reversible PPDK is another metabolic way for cells to produce ATP.

There is strong biochemical evidence that the PCK subtype C_4 photosynthesis plays a substantial role in some NADP-ME subtype C_4 plants, such as maize (Wingler et al., 1999; Leegood and Walker, 2003). Consistent with this notion, *PCK* was found to be expressed at an extremely high level and exclusively in BS cells (i.e. R_i = 1.00; Table III). Two other enzymes are also required to support the operation of the PCK decarboxylation system in both cell types, namely aspartate aminotransferase (AspAT) and alanine aminotransferase (AlaAT). There are two major paralogs of *AspAT* found in maize with a high level of expression: one (GRMZM5G836910) is M specific (R_i = 0.97), which encodes a chloroplast enzyme, confirming the biochemical results of Gutierrez et al. (1974), while the other (GRMZM2G094712) is BS enriched (R_i = 0.79), which encodes a mitochondrial enzyme (Supplemental Data S3). In contrast, the two AlaAT paralogous genes showed moderate expression levels, both with an M cell preference. The presence of a large amount of *AspAT* transcript in M cells is necessary for the operation of the PCK decarboxylation system in maize leaves, as OAA is very labile and once produced should be taken up quickly by the M chloroplast for reduction to malate by NADP-MDH or transaminated to Asp by AspAT.

Functional Differentiation between M and BS Cells and Identification of Genes Involved in Kranz Structure Formation

For the 1,545 M-enriched and 3,422 BS-enriched genes identified, their biological functions were further elucidated using MapMan (http://mapman.gabipd.org/): 71% of the enriched genes (3,358 genes: 1,015 M enriched and 2,343 BS enriched) have been assigned to known biological processes or metabolic pathways. Except for those that have been classified as "others" (14%), the three most abundant functional groups are protein metabolism (13%), RNA metabolism (11%), and transport (6%; Fig. 1C). The other abundant functional groups are miscellaneous (5%), signaling (4%), cell (3%), stress (3%), hormone metabolism (3%), photosynthesis (2%), and development (2%; Fig. 1C). When considering the functions of these genes in M or BS cells, we identified 15 cell type-enriched functional groups using the Fisher exact test ($P < 0.05$ for each group; Fig. 1D). The difference in the number of cell type-enriched genes between M and BS cells within each function is quite large, except for the protein-targeting category. Overall, the number of BS-enriched genes (3,422) is much higher than that of M-enriched genes (1,545). As expected, the transcripts of genes coding for photosynthesis light reaction components were more abundant in M than in BS cells. Interestingly, over 90% of the genes related to protein synthesis (for ribosomal proteins) and protein folding were preferentially expressed in M cells, while those involved in protein degradation were predominantly expressed in BS cells. These observations indicate that many proteins were synthesized and folded in M cells but eventually degraded in BS cells, which surround the vasculature. This arrangement facilitates the transport of amino acids released from protein degradation to other plant parts. Genes involved in RNA regulation of transcription, transport, signaling, cell division, organization and vesicle transport, development, and major carbon, hydrogen, and oxygen metabolism were predominantly expressed in BS cells, but genes involved in secondary metabolism, tetrapyrrole synthesis, and RNA binding showed an M cell expression preference. Most of the secondary metabolites accumulate in the epidermis cells, adjacent to M cells.

For the 85 BS-enriched genes involved in development, we found that 22 genes encode the nodulin MtN3 and MtN21 family proteins, which were originally identified as root nodule proteins of the legume *Medicago truncatula* (Gamas et al., 1996). MtN3 and MtN21 also have recently been shown to be crucial for exine pattern formation and cell integrity of

microspores (Guan et al., 2008) and secondary wall formation in xylary and cotton fibers (Ranocha et al., 2010), respectively. These genes code for transmembrane proteins involved in the cell wall development of specialized plant cells that require stringent impermeability to specific substances (e.g. oxygen). Therefore, the counterparts in maize leaves might be involved in the development of the BS cell wall, which plays a key role in preventing the leakage of CO_2 released from C_4 acid decarboxylation in order to maintain a high CO_2 concentration in BS cells (Hatch et al., 1995). At present, little is known about the biochemical composition of the C_4 BS cell wall. For the first time, to our knowledge, our data suggest that C_4 plants may have adopted the exine composition of microspore and pollen for its impermeability to a number of substances, such as CO_2. Further analysis on the expression profile of these genes along the development of the Kranz anatomy in C_4 leaves should provide insights into their functional roles.

Cell Type-Enriched Transcription Factors

Among the 2,056 annotated transcription factor (TF) genes in maize, 654 and 820 were expressed in M and BS cells, respectively (Table II). The median RPKM for all expressed TFs in the two cell types were 5.7 and 7.1, respectively (Supplemental Data S4). We identified 53 TF genes as M enriched and 214 TF genes as BS enriched (Table II), which, according to the TF family database PlantTFDB version 2 (Zhang et al., 2011), belong to 16 and 38 TF families, respectively. Six of the BS-enriched TFs were assigned to two different families simultaneously, because of two different types of DNA-binding domains in a protein. Three TFs, *MYB59* (GRMZM2G130149), *MYB111* (GRMZM2G305856), and another MYB family protein (GRMZM2G002128), belong to both MYB and MYB-related families. The other three TFs are *AGL12* (GRMZM2G105387), which belongs to both M-type and MIKC families, *RPL* (GRMZM2G154641), which belongs to both HB-other and TALE families, and a zinc finger (B-box type) family protein (GRMZM2G075562), which belongs to both CO-like and DBB families. Interestingly, in 11 of the families with cell type-enriched TF genes, all members showed BS enrichment, whereas none has all members showing M enrichment (Fig. 1E). This observation suggests that many TF families have a cell type preference, indicating a differentiation in regulatory network for the functional differentiation of the two photosynthetic cell types in maize leaves. Comparison of the distributions of cell type-enriched TF families between this study and Li et al. (2010) shows some similarities but many differences. For example, Li et al. (2010) identified only two ARF genes enriched in M cells and none enriched in BS cells (Fig. 5d in Li et al., 2010), while we identified seven ARF genes enriched in BS cells but none enriched in M cells (Fig. 1E). Cross-contamination of isolated M and BS cells will lead to underestimation of cell-specific genes.

We searched the literature to identify the functions or pathways where a TF might be involved (Table IV). Most M-enriched TFs are involved in the pathways of responses to environmental stimuli. For example, the three paralogs of *DREB1A* (GRMZM2G380377, GRMZM2G069146, and GRMZM2G069126) are related to the function of cold acclimation (Novillo et al., 2007). The two ERF TFs, *CBF4* (GRMZM2G124037) and *ERF7* (GRMZM2G089995), are involved in the pathways of responses to drought and abscisic acid (Haake et al., 2002; Song et al., 2005), respectively. In addition, three TFs in the NAC family and two in the WRKY family are induced by wounding and pathogens (Collinge and Boller, 2001; Xu et al., 2006; Zheng et al., 2006), respectively. Two other M-enriched TFs were *GLK1* (GRMZM2G026833), which is involved in pathways of chloroplast development specific to M cells (Nakamura et al., 2009), and *ARR10* (GRMZM2G013612), which is related to protoxylem differentiation (Yokoyama et al., 2007; Table IV).

Very different from the M-enriched TFs, most BS-enriched TFs are related to development or the biosynthesis of cells and organelles. These functions include chlorophyll biosynthesis, secondary cell wall biosynthesis, cell division, cell differentiation, or development of chloroplast and leaves. The observations also match some functional preferences of BS-enriched genes: categories of development and cell (division, organization, vesicle transport, etc.; Fig. 1D). Three TFs, *MYB20* (GRMZM2G055158), *MYB52* (GRMZM2G455869), and *MYB54* (GRMZM2G077147), in the MYB family, and *ATAF1* (GRMZM2G179049) in the NAC family, are related to secondary cell wall biosynthesis (Zhong et al., 2006, 2008). Two TFs, *TCP4* (AC205574.3_FG006) and *TCP5* (GRMZM2G120151), in the TCP family are related to cell division (Cubas et al., 1999; Aggarwal et al., 2011). The GRAS TFs, *RGA1* (GRMZM2G023872) and *SHR* (GRMZM2G132794), are related to asymmetric cell division (Truong et al., 1997; Heidstra et al., 2004). Two bHLH TFs, *PIL5* (GRMZM2G165042) and *FMA* (GRMZM2G162450), are related to chlorophyll biosynthesis and cell differentiation (Huq et al., 2004; Gray, 2007), respectively. Three ARF TFs, *ARF3* (GRMZM2G056120), *ARF6* (GRMZM2G317900), and *ARF8* (GRMZM2G078274), are related to leaf or floral organ development (Finet et al., 2010; Tabata et al., 2010; Table IV).

Interestingly, the ERF family was the only one that showed a larger number of M-enriched TFs than BS-enriched TFs. Many of these M-enriched ERF TFs have been shown to play a role in response to environment stimuli, but none of the BS-enriched ERF TFs are known to play such a role. Our data also showed that *GOLDEN2* (*GLK2*; GRMZM2G087804) was preferentially expressed in BS cells of mature maize leaves at a moderate level (RPKM = 29) with R_i = 0.78 (Supplemental Data S1). *GLK2* has been shown to be involved in maize chloroplast development in the BS cells independent of light, but it is not involved in *rbcS* accumulation (Hall et al., 1998; Cribb et al., 2001). On

Table IV. *Selected cell type-enriched TFs and related functions or pathways*

Cell	TF Family[a]	Gene Identifier	M	BS	R_i	Gene Name	Function or Pathway
M	G2-like	GRMZM2G026833	125.41	15.16	0.88	GLK1	Chloroplast development
M	ARR-B	GRMZM2G013612	55.22	10.27	0.81	ARR10	Protoxylem differentiation
M	WRKY	GRMZM2G449681	22.82	4.42	0.81	WRKY33	Pathogen induced
M	WRKY	AC209050.3_FG003	2.21	0.21	0.91	WRKY18	Pathogen induced
M	ERF	GRMZM2G380377	4.81	0.32	0.93	DREB1A	Cold acclimation
M	ERF	GRMZM2G069146	69.81	5.53	0.92	DREB1A	Cold acclimation
M	ERF	GRMZM2G069126	39.81	2.78	0.93	DREB1A	Cold acclimation
M	ERF	GRMZM2G124037	7.19	0.03	1.00	CBF4	Drought adaptation
M	NAC	GRMZM2G336533	2.16	0.13	0.94	ATAF1	Induced by wounding
M	NAC	GRMZM2G060116	11.36	0.86	0.92	ATAF2	Induced by wounding
M	NAC	GRMZM2G126936	19.29	1.73	0.91	ATAF2	Induced by wounding
BS	bHLH	GRMZM2G165042	0.80	41.36	0.98	PIL5	Chlorophyll biosynthesis
BS	MYB	GRMZM2G455869	0.04	13.31	1.00	MYB52	Secondary wall
BS	MYB	GRMZM2G077147	0.01	4.13	1.00	MYB54	Secondary wall
BS	MYB	GRMZM2G055158	0.08	1.07	0.93	MYB20	Secondary wall
BS	NAC	GRMZM2G179049	0.12	2.82	0.96	ATAF1	Secondary wall
BS	TCP	GRMZM2G120151	1.34	42.49	0.97	TCP5	Cell division, proliferation
BS	TCP	AC205574.3_FG006	0.00	1.53	1.00	TCP4	Cell division, proliferation
BS	bHLH	GRMZM2G162450	0.06	3.13	0.98	FMA	Cell differentiation
BS	ARR-B	GRMZM2G479110	1.05	14.69	0.93	ARR12	Protoxylem differentiation
BS	ARF	GRMZM2G056120	0.63	11.66	0.95	ARF3	Development of floral organs and leaves
BS	ARF	GRMZM2G078274	0.44	7.86	0.94	ARF8	Development of floral organs
BS	ARF	GRMZM2G317900	0.17	12.34	0.99	ARF6	Development of floral organs
BS	GRAS	GRMZM2G023872	0.00	1.20	1.00	RGA1	Asymmetric cell division
BS	GRAS	GRMZM2G132794	0.02	2.29	0.99	SHR	Asymmetric cell division

[a]TF family names are as follows: G2-like, GOLDEN2-like; ARR-B, type B phospho-Accepting Response Regulator; WRKY, conserved WRKY domains, about 60 amino acid residues with the WRKYGQK sequence followed by a C2H2 or C2HC zinc finger motif; ERF, Ethylene-Responsive Element Binding Factor; NAC, NAM (no apical meristem), ATAF, and CUC (cup-shaped cotyledon) TFs; bHLH, basic helix-loop-helix proteins; MYB, myeloblastosis virus; TCP, TEOSINTE BRANCHED1 from maize, CYCLOIDEA from snapdragon (*Antirrhinum majus*), and PROLIFERATING CELL FACTOR1 and -2 from rice (*Oryza sativa*); ARF, Auxin Response Factor; GRAS, GAI, RGA, and SCR genes.

the other hand, *GLK1* was highly and preferentially expressed in M cells ($R_i = 0.89$). These data suggest that GLK1 may regulate plastid development in M cells, while GLK2 may regulate plastid development in BS cells of C_4 leaves.

More Genes Are Expressed in M and BS Cells Than Previously Known

Recently, Li et al. (2010) examined the transcriptomes of M and BS cells isolated from fully expanded third maize leaves by LCM. About 41 million 32-bp sequencing reads from the two cell transcriptomes were obtained. In this study, we generated approximately 370 million 120-bp reads from M and BS RNA samples (approximately 177 and approximately 193 million reads, respectively). Our data allowed the identification of 18,482 expressed genes in total, while Li et al. (2010) detected only 16,638 expressed genes. Another possible reason for the above difference between the two studies is that Li et al. (2010) used an earlier annotation version (APGv1) and only included genes with complementary DNA (cDNA) support, while we used APGv2. As a result, the total number of expressed genes obtained from our data is higher than that obtained by Li et al. (2010).

In addition to the expressed genes in the FGS, we detected an additional 2,224 expressed genes in the WGS (Fig. 1A), with the exclusion of transposable elements, gene fragments, and pseudogenes. Many genes in this additional gene set were highly expressed, with RPKM $>$ 100 for 85 M-enriched genes and 86 BS-enriched genes. The highest RPKM values in this set were 4,900 (GRMZM5G809350; a paralog of LHCII type 1 CAB2) and 31,561 (GRMZM2G142891; a paralog of β-glucosidase aggregating factor) in M and BS cells, respectively. Furthermore, we detected 178 M-enriched and 610 BS-enriched genes in the WGS. Although there is no MapMan annotation for the genes in the WGS, 263 of the 2,224 expressed genes in the WGS are found to have one or more Gene Ontology terms (http://www.geneontology.org/); the five most abundant terms in molecular functions are ATP binding (23%), ice binding (13%), zinc ion binding (12%), DNA binding (10%), and protein binding (10%). Three M-enriched genes (GRMZM2G411457, GRMZM2G030616, and GRMZM2G152576) are related to light reaction in photosynthesis (Supplemental Data S1).

Notably, our data reveal more genes expressed in BS than in M cells in mature maize leaves. Expressed genes from vascular tissues (e.g. sieve element) in the BS cell preparation might have contributed to the category of BS-expressed genes, as the BSS isolated in

this study contained the vascular tissue (Supplemental Table S4) and sieve elements are active cells for transport function. Thus, as discussed above, a small portion of the RNA from vascular tissues might have contributed to our BS transcriptome. However, it is most likely that more genes are expressed in BS cells than in M cells because of functional necessity, in agreement with the observation that the BS cells in C_4 plants play much more important roles than the M cells in various physiological processes (Majeran et al., 2005; Leegood, 2008). In addition to genes related to C_4 acid decarboxylation, expression of specific genes related to the metabolism of photorespiration, nitrate, sulfate, ammonium, starch, ion transport, secondary metabolites, and cell wall structure is largely confined to BS cells.

Light Reaction, Carbon Fixation Pathway, and Related Transporters

Many of the maize genes coding for the PSI components are highly expressed in both cell types, but with higher expression levels in M cells (Supplemental Data S3). In contrast, many of the maize genes coding for the PSII components are highly and predominantly expressed in M cells, with $R_i > 0.9$ for most of the genes. Transcripts for the two genes coding for ATP synthase, however, are found in both cell types. Most of the NDF (for NDH-dependent cyclic electron flow) genes coding for the iron-sulfur cluster-binding protein involved in cyclic electron transport exhibited a preferential expression in BS cells, with R_i ranging from 0.67 to 0.86, whereas two NDF genes (GRMZM2G865543 and GRMZM2G162233) showed a moderate preference in expression in BS cells, with R_i of 0.38 and 0.36, respectively. These data are in agreement with the observations that maize BS cells possess a higher capacity for cyclic electron transport (Ivanov et al., 2007).

Based on the expression profiles of C_4-specific genes (Table III; Supplemental Data S1 and S3), we propose the partition of the key biochemical reactions of the C_4 photosynthesis pathway between M and BS cells of maize, an NADP-ME subtype C_4 plant (Fig. 2). Hydration of CO_2 to bicarbonate, generation of PEP from pyruvate, carboxylation of bicarbonate to OAA, and its subsequent transamination to Asp or reduction to malate in the chloroplast all take place in the M cells. In contrast, decarboxylation of OAA to PEP and CO_2 and decarboxylation of malate to pyruvate and CO_2 occur in the cytosol and chloroplast of BS cells, respectively. The high expression levels of *NADP-ME* and *PCK* suggest that the PCK decarboxylation system plays a substantial role in maize C_4 photosynthesis, consistent with previous biochemical studies (Furumoto et al., 1999, 2000; Wingler et al., 1999; Furbank, 2011). The dual decarboxylation system in some NADP-ME subtype C_4 plants, such as maize, probably makes them the most efficient photosynthetic plants.

High rates of photosynthesis by C_4 plants imply high fluxes of metabolites intercellularly and intracellularly. In C_4 plants, most of the photosynthetic metabolites are known to be transported symplastically between M and BS cells through plasmodesmata. However, within each cell type, transporters on the chloroplast envelope must be active to substantiate the heavy traffic of related metabolites between cytosol and chloroplast. A maize gene coding for a 2-oxoglutarate/malate translocator (*ZmpOMT1*; GRMZM2G383088) was highly and predominantly expressed in M cells, and a dicarboxylate transporter gene (*ZmpDCT1*; GRMZM2G040933) was also preferentially expressed in M cells at a moderate level. Thus, these two transporters may be responsible for the uptake of OAA and the export of malate by M chloroplasts, respectively (Taniguchi et al., 2004). In contrast, the transcript for another maize gene coding for DCT (*ZmpDCT2*; GRMZM2G086258) was highly and almost exclusively expressed in BS cells, which may be responsible for the import of malate by BS chloroplasts, as suggested previously by Taniguchi et al. (2004).

In C_3 plants the chloroplast phosphate translocator triose phosphate/phosphate translocator (TPT) is thought to have a central role in the partitioning of fixed carbon between starch in the chloroplast and Suc in the cytosol. Inorganic phosphate is taken up by the translocator in exchange for triose phosphate, which is then metabolized to Suc in the cytosol. In NADP-ME subtype C_4 plants, BS chloroplasts are deficient in PSII, so the triose phosphate produced must be transported to M chloroplasts for reduction. Consistent with this role, the transcript of a maize TPT gene (*APE2*; GRMZM2G070605) was expressed with very high RPKM values of 2,069 and 1,551 in M and BS cells, respectively (Supplemental Data S3).

In C_4 plants, pyruvate produced from malate decarboxylation in the BS chloroplasts must be returned to M chloroplasts for regeneration of PEP via PPDK with the input of ATP, and PEP is subsequently exported to the cytosol for carboxylation by PEPC. Earlier biochemical and molecular studies suggested that to accommodate these new transport functions in C_4 plants, TPT has been adopted and modified for the exchange of phosphate and PEP (PPT; Flugge, 1999), and two types of light-dependent pyruvate transporter have been employed by the M chloroplasts for the uptake of pyruvate in a wide range of C_4 plants, one sodium dependent (Ohnishi et al., 1990; Ohnishi and Kanai, 1990; Aoki et al., 1992) and one proton dependent (Aoki et al., 1992). The NADP-ME subtype C_4 *Flaveria* spp. use the sodium/pyruvate symporter (Gowik et al., 2011), but maize, a NADP-ME subtype monocot, uses the proton/pyruvate symporter (Ohnishi et al., 1990; Ohnishi and Kanai, 1990; Aoki et al., 1992). The gene coding for the sodium-dependent pyruvate transporter BASS2, a bile acid:sodium symporter family protein 2, has been cloned recently from C_4 *Flaveria* spp. (Furumoto et al., 2011), but the gene(s) coding for the proton-dependent one remains elusive. Transcripts of several maize *PPT*s were detected in this study. One of them (GRMZM2G174107) was highly enriched in

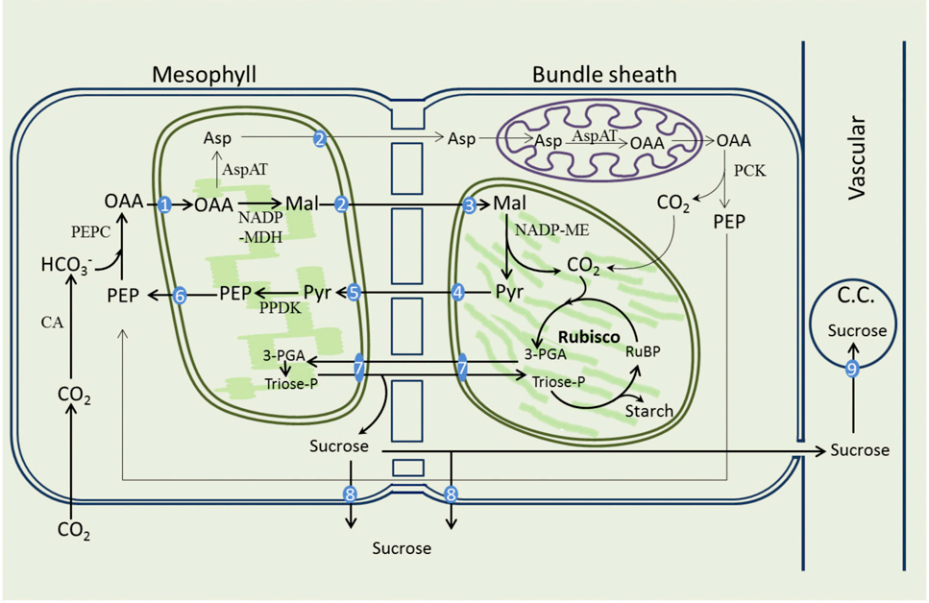

Figure 2. Cellular compartmentation of the key C_4 biochemical reactions in the NADP-ME subtype maize leaves. Enzymes presented are CA (carbonic anhydrase; GRMZM2G121878) and PEPC (phospho*enol*pyruvate carboxylase; GRMZM2G083841) in the cytosol and AspAT (aspartate aminotransferase; GRMZM5G836910), NADP-MDH (NADP-malate dehydrogenase; GRMZM2G129513), and PPDK (pyruvate, Pi dikinase; GRMZM2G306345) in the chloroplast of M cells, as well as PCK (phospho*enol*pyruvate carboxykinase; GRMZM2G001696) and AspAT (aspartate aminotransferase; GRMZM2G094712) in the cytosol and mitochondria, respectively, and NADP-ME (NADP-malic enzyme; GRMZM2G085019) and Rubisco (large subunit, GRMZM5G815453; small subunit, GRMZM2G098520 and GRMZM2G113033) in the chloroplast in BS cells. The major metabolite transporters involved are as follows: (1) OMT1 (2-oxoglutarate/malate transporter; GRMZM2G383088), (2) DIT1 (dicarboxylate transporter; GRMZM2G040933), (3) DCT2 (dicarboxylate transporter; GRMZM2G086258), (4) MEP3 (BS; putative proton/pyruvate symporter; GRMZM2G138258), (5) MEP3 (M; putative proton/pyruvate symporter; GRMZM2G305851), (6) PPT (phospho*enol*pyruvate/phosphate translocator; GRMZM2G174107), (7) TPT (triose phosphate/phosphate antiporter; GRMZM2G070605), (8) SUT1/2 (Suc transporters; GRMZM2G034302 and GRMZM2G087901) in companion cells (C.C.), and (9) SUT4 (Suc transporter; GRMZM2G307561).

M cells, whereas a second one (GRMZM2G066413) was moderately enriched in BS cells. Thus, the GRMZM2G174107 gene may specifically code for the exchanger of triose phosphate and PEP in the M chloroplasts of maize leaves. In addition, our transcriptome data showed that for the three *MEP* (for methyl erythritol phosphate) genes coding for the inner envelope transporter proteins, *MEP3* (GRMZM2G138258) was expressed exclusively in BS cells (R_i = 0.99) at an extremely high level (RPKM = 2,734), GRMZM2G305851 was moderately expressed with a preference in M cells, whereas *MEP2* (GRMZM2G077222) was equally expressed in both cell types at relatively low levels. The exact physiological roles of MEP transporters in maize have not been established, but all three MEP transporters have chloroplast-targeting sequences (by TargetP; http://www.cbs.dtu.dk/services/TargetP/) and transmembrane domains (by DAS; http://mendel.imp.ac.at/sat/DAS/DAS.html), which suggest that they are plastidial transporters. Since pyruvate is one of the two crucial substrates of the MEP pathway in plastids, we propose that the two *MEP3* paralogs may code for the inner chloroplast envelope proteins responsible for pyruvate export from BS chloroplasts and import of pyruvate by M chloroplasts, respectively, whereas *MEP2* may mainly function to supply pyruvate to the MEP pathway, as in C_3 plants. In agreement with our data, Friso et al. (2010) suggested earlier that MEP2,4 and MEP3/4 may function as pyruvate transporters in M and BS chloroplasts, respectively, based on proteomic studies of isolated maize M and BS chloroplasts.

Although photorespiration is not apparent in C_4 plants under ambient atmospheric conditions, measurable photorespiratory activity can be detected at low CO_2 conditions (de Veau and Burris, 1989; Dai et al., 1993) and in young leaves (Dai et al., 1993, 1995). Consistent with the localization of Rubisco in the BS chloroplasts of C_4 plants, the expression of genes coding for most of the key photorespiratory enzymes exhibited a BS specificity (Supplemental Data S3). High levels of transcript were detected in the BS cells for glycolate oxidase (GRMZM2G129246), Gly decarboxylase (GRMZM2G399183 and GRMZM2G104310), and Ser hydroxymethyltransferase (GRMZM2G135283), with R_i values ranging from 0.94 to 1.00. However, expression of two of the photorespiratory pathway enzymes, chloroplast phosphoglycolate phosphatase and peroxisomal NAD-hydroxypyruvate reductase,

did not show strong BS preference, presumably due to their housekeeping roles in the organelles. Also, chloroplast is known to be capable of assimilating ammonium, a toxic compound released from Ser decarboxylation in mitochondria, through Glu synthetase and Gln synthase. In accordance, genes for Glu synthetase (GRMZM2G098290) and Gln synthase (GRMZM2G036609) were highly expressed in both cells. It is also notable that several paralogs of Ser hydroxymethyltransferase were expressed in maize, but only one of them (GRMZM2G135283) was exclusively expressed in BS cells, which could have been acquired during the evolution of C_4 plants.

Development, Cell Wall, and Plasmodesmata Formation

Kranz leaf anatomy is essential for most C_4 plants to concentrate CO_2 around Rubisco in the chloroplasts of BS cells to suppress its oxygenase and the associated photorespiratory activity. Consequently, the BS cell wall in many C_4 plants is thickened and specialized with a suberin layer to prevent CO_2 leakage (Evans and Von Caemmerer, 1996; Leegood, 2002). Our data depict the expression of many cell wall-associated genes specifically expressed in BS cells, which include genes coding for cell wall-associated kinase (GRMZM2G135291 and GRMZM2G062471) and several highly expressed nodulin-type proteins (MtN3 and MtN21). Therefore, these genes may be related to the differentiation of C_4 BS cell walls. In contrast, only one cell wall-related gene coding for expansin (GRMZM2G094990) is specifically expressed in M cells.

Intercellular transport of metabolites between M and BS cells (Evert et al., 1977; Robards and Lucas, 1990) and between the BS and the adjacent vascular parenchyma cells (Botha et al., 2000) in C_4 leaves is mainly by diffusion through interconnecting plasmodesmata, which are concentrated along the cell wall between these cell types. Interestingly, our data reveal that the genes coding for plasmodesmata-located proteins (signaling receptor kinases), plasmodesmata callose-binding proteins (β-1,3-glucan hydrolases), and callose synthase are almost exclusively expressed in BS cells (Supplemental Data S3), suggesting that genes related to plasmodesmata development may be controlled by the regulatory network of BS cells. Callose, which differs from cellulose in consisting of a β-1,3-glucan chain, is known to be made by a few cell types at a specific stage during cell wall development, such as microspore and pollen. It is also ectopically deposited over the cell wall between the BS and vascular parenchyma cells in maize leaf minor veins, thereby occluding the plasmodesmata (Botha et al., 2000). Therefore, we speculate that a similar cell wall composition may be shared by nodule, microspore, pollen, xylary fiber, and C_4 BS cells. C_4 plants regulate the expression of these cell wall-related genes in an organ- and cell-specific manner to function in the CO_2-concentrating mechanism.

CONCLUSION

The C_4 physiology is an integrated syndrome of developmental, anatomical, cellular, and biochemical traits that must rely on regulatory networks that control the functional differentiation of M and BS cells. Our transcriptome profiling of the two photosynthetic cell types of mature maize leaves reveals the cell differentiation of gene expression at both regulatory and functional levels. To our knowledge for the first time, genes coding for the transporters of several C_4 metabolites and the development of the BS cell wall have been identified that express in a cell-specific manner. Clearly, both cell types are specialized in different metabolic functions but are also coordinated to provide a high efficiency in carbon, mineral, and water use. In addition, many TF genes have been shown to express in a cell-specific manner. Thus, the comprehensive data sets from this study have laid a solid foundation for further study of the regulation of the expression and evolution of C_4-specific and -related genes.

Although C_4 plants represent only approximately 3% of plant species, they contribute to about 30% of terrestrial carbon fixation, and 60% of C_4 plants are grasses, which include some of our most important food crops, such as maize, sugarcane (*Saccharum officinarum*), millet (*Pennisetum americanum*), and sorghum (*Sorghum bicolor*) of the NADP-ME subtype. Despite this scarcity, some of the most productive crops and noxious weeds on earth belong to the C_4 plants, which make this photosynthetic property a desirable trait for introduction into the less productive C_3 crops (Hibberd et al., 2008). However, at present, very little is known about the molecular regulatory networks underlying the development of Kranz leaf anatomy and the functional differentiation of its photosynthetic cells in C_4 photosynthesis and other metabolism. A comprehensive understanding of the C_4 syndrome is required before we can effectively transfer the C_4 traits into C_3 crops, such as rice (*Oryza sativa*), wheat (*Triticum aestivum*), and soybean (*Glycine max*). A thorough and systematic examination of the regulatory networks involved in C_4 syndrome from comparative C_3, C_3-C_4 intermediate, and C_4 transcriptome study (Gowik and Westhoff, 2011) and from developmental and cellular transcriptome profiling of C_4 leaves (Li et al., 2010; this study) should streamline this process.

MATERIALS AND METHODS

Plant Growth Conditions and Leaf Sample Collection

Seeds of maize (*Zea mays* 'White Crystal'), a glutinous cultivar, were germinated and cultivated in the greenhouse in June 2011 under natural sunlight conditions (13–14 h) with a maximum photosynthetic photon flux density of 1,600 μmol m^{-2} s^{-1}, day/night temperatures of 28°C to 32°C/25°C to 28°C, and 60% to 70% relative humidity. The second newly expanded leaves of 9-d-old plants were harvested around 9 AM for isolation of M and BS cells in the laboratory and extraction of total RNA.

Enzymatic Isolation of M Protoplasts

MP were isolated from the mid section (5 cm) of newly matured second leaves by digesting leaf segments in a cell wall digestive medium, according to

the procedures of Kanai and Edwards (1973), Sheen (1995), Markelz et al. (2003), Sawers et al. (2007), and Sharpe et al. (2011) with minor modifications. All solutions and glassware were sterilized either by filtration or by treatment with diethyl pyrocarbonate water followed by autoclave. The digestive medium, containing 1.5% (w/v) cellulose (Onozuka RS), 0.1% (w/v) macerozyme (Onozuka R-10), 20 mM MES (pH 5.8), 0.6 M sorbitol, 5 mM β-mercaptoethanol, 1 mM $CaCl_2$, 0.1% bovine serum albumin, and 10 mM dithiothreitol (DTT), was prepared freshly from filter-sterilized stock solutions. MES buffer solution was preheated to 70°C for 5 min and cooled to room temperature before the addition of sorbitol and digestive enzymes. The medium was then heated to 55°C for 10 min to solubilize the enzymes as well as to inactivate DNase, and other components were added. Several leaf blades were stacked and cut perpendicularly to the long axis into 0.5- to 1-mm slices with sterilized razor blades and then quickly transferred into a 50-mL sterile beaker containing 20 mL of digestive medium. The medium with leaf slices was vacuum infiltrated twice and incubated at 23°C under low light (100 μmol m^{-2} s^{-1}) for 3 h without shaking.

At the end of digestion, the mixture was first filtered through a 500-μm mesh, and the materials retained on the mesh were washed gently with 100 mL of sterile sorbitol wash medium (0.6 M sorbitol, 5 mM HEPES-KOH, pH 7.0, 100 mM β-mercaptoethanol, and 1 mM $CaCl_2$) to release more MP. The filtrate was then filtered through an 80-μm nylon net to retain undigested tissue and BS, while MP (15–35 μm diameter) in the filtrate was collected by centrifugation at 300*g* for 5 min. The crude MP pellet was gently resuspended in 4 to 6 mL of 13% (w/v) dextran-Suc solution (13% dextran, 0.6 M Suc, 5 mM HEPES-KOH, pH 7.0, 100 mM β-mercaptoethanol, 1 mM $CaCl_2$, and 10 mM DTT). For purification, the MP crude suspension was transferred to a 13-mL test tube and gently overlaid sequentially with 2 mL of 10.7% (w/v) dextran-Suc solution and 2 mL of sorbitol wash medium. The gradient was centrifuged at 300*g* in a swing-bucket rotor for 20 min, and pure MP was collected from the interphase between the sorbitol wash medium and 10.7% dextran-Suc solution and diluted in a small volume of sorbitol wash medium before isolation of RNA. The purity of isolated MP was examined with a light microscope.

Mechanical Isolation of BSS

To avoid damage to BS cells during enzymatic digestion, a mechanical method was adopted for rapid isolation of BS cells. The isolation medium contained 0.6 M sorbitol, 50 mM Tris-HCl (pH 8.0), 5 mM EDTA, 0.5% polyvinylpyrrolidone-10, 10 mM DTT, and 100 mM β-mercaptoethanol. Leaf slices, prepared as described above, were quickly added to 50 mL of cold BS cell isolation medium and homogenized twice (10 s each) at high speed in a Waring blender. After blending, the mixture was first filtered through a 500-μm mesh, the crude tissue retained on the 500-μm mesh was resuspended in another 50-mL isolation medium, and the blending and filtration processes were repeated. BS cells that passed through the 500-μm mesh were filtered and retained on an 80-μm nylon net. The purity of BS cells was monitored with a light microscope, and if deemed necessary, the collected BS cells were further washed with the cold isolation medium to remove attached M tissue and the epidermal tissue floating on top of the isolation medium. The whole isolation procedure was completed within 10 min.

Enzyme Assay and Immunoblot Analysis

The activities of NADP-ME and PEPC were assayed at 30°C using a spectrophotometer according to Kanai and Edwards (1973). The procedure of Murphy et al. (2007) was followed for immunoblot analysis using specific antibodies.

RNA Isolation, Library Preparation, and Deep Sequencing

Total RNA from isolated cells was extracted using Trizol reagent (Invitrogen) and purified using acid phenol-chloroform extraction. The quality of extracted RNA was examined by gel electrophoresis and by BioAnalyzer (Agilent).

Using the Illumina mRNA-seq kit, 10 to 15 μg of a total RNA sample was purified for poly(A) RNA and subjected to fragmentation for 3 min at 94°C. The mRNA fragments were then converted into single-stranded cDNA, and the ends were modified and ligated to the paired-end adaptors. The ligation products were size selected on agarose gels (independent slices of 250–350 and 350–450 bp), amplified by 15 cycles of PCR, cleaned up using Ampure beads (Beckman Agencourt), and examined by Qubit (Invitrogen) and BioAnalyzer High Sensitivity DNA chip (Agilent). Paired-end 2- × 120-nucleotide sequencing was performed on Illumina Genome Analyzer IIx, and two lanes of data were obtained for each cell type by sequencing two independently sized selected libraries. Approximately 177.4 million and 193.4 million reads were obtained from the M and BS RNA samples, respectively.

Deep-Sequencing Data Processing, Mapping, and Expression Level Calculation

We processed and mapped raw reads to the maize genome (B73) using the mapping tool Tophat version 1.3.3 (http://tophat.cbcb.umd.edu/) with default parameters (maximum allowed mismatches, two; maximum number of hits for a read, 10). Based on the APGv2 annotation downloaded from http://ftp.maizesequence.org/, we calculated the expression levels of each annotated gene in RPKM (Mortazavi et al., 2008) using the expression calculation tool Cufflinks version 1.2.1 (http://cufflinks.cbcb.umd.edu/).

Estimation of Cell Type Specificity

As described in "Results," to estimate the degree of differential expression of genes between the two cell types, we defined the degree of cell specificity of gene i as $R_i = |m_i - b_i| / \max(m_i, b_i)$, where m_i and b_i represent the RPKM of gene i in the M and BS RNA samples, respectively. Applying this measure to the transcriptome data, one can study the gene expression differences between the M and BS cells.

Estimation of Contamination Rates

We can use the transcriptome data to obtain an approximate estimate of the contamination rate in the M (or BS) cells isolated. Let c_1 be the percentage contamination of the isolated M cells by BS cells. That is, c_1 is the proportion of BS cells in the sample of M cells. For gene i, let x_i and y_i be the expected expression levels in M cells and in BS cells, respectively, under no contamination. Then, the expected expression level of gene i in the isolated M cells, denoted by m_i, should be given by

$$m_i = (1 - c_1)x_i + c_1 y_i \quad (1)$$

Equation 1 cannot be solved because there are two unknowns, c_1 and m_i. However, if gene i is supposed not to be expressed in M cells, we can assume $x_i = 0$ and obtain $c_1 = m_i / y_i$.

Since using one gene to estimate c_1 may not be reliable, we estimate c by

$$c_1 = \sum m_i / \sum y_i \quad (2)$$

where the summation is over the genes that are supposed to be expressed only in BS cells but not in M cells. We can take the estimate as an upper estimate because one or more of the genes used may not be absolutely only expressed in BS cells.

Similarly, let c_2 be the rate of contamination of BS cells by M cells and let b_i be the observed level of expression in the transcriptome data of the BS cells isolated. Then, we have

$$b_i = c_2 x_i + (1 - c_2)\, y_i \quad (3)$$

$$c_2 = \sum b_i / \sum x_i \quad (4)$$

where the summation is over the genes that are supposed to be expressed only in M cells but not in BS cells.

Functional Annotation

The identified cell type-enriched genes were annotated with biological functions and metabolic pathways by MapMan (Thimm et al., 2004; http://mapman.gabipd.org/). The TF information was downloaded from Plant TFDB version 2.0 (Zhang et al., 2011; http://planttfdb.cbi.pku.edu.cn/).

The read data have been submitted to the National Center for Biotechnology Information Short Read Archive under accession number SRP009063.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Micrographs of isolated M and BS cells.

Supplemental Figure S2. Immunoblot analysis in mature leaves and isolated M and BS cells (two preparations).

Supplemental Figure S3. Median depth of read-mapping coverage along the cDNA.

Supplemental Table S1. Activities of PEPC and NADP-ME in the whole leaf and in the isolated M and BS cells of maize.

Supplemental Table S2. Validation of cell type-specific gene expression.

Supplemental Table S3. List of quantitative reverse transcription-PCR primers, Universal Probe Library probe, and PCR efficiency.

Supplemental Table S4. RPKM of selected genes of vascular and epidermal tissues.

Supplemental Data S1. RPKM values of all genes with at least one read in one of the two transcriptomes.

Supplemental Data S2. List of genes with $R_i > 0.99$ in the two cell types.

Supplemental Data S3. Differential expression between M and BS cells of genes.

Supplemental Data S4. List of cell type-enriched TFs.

Supplemental Materials and Methods S1.

ACKNOWLEDGMENTS

We thank Jen Sheen and Shin-Han Shiu for valuable comments and suggestions on the manuscript. The transcriptome sequencing was conducted by the High Throughput Sequencing Core, Academia Sinica, Taiwan.

Received July 17, 2012; accepted July 23, 2012; published July 24, 2012.

LITERATURE CITED

Aggarwal P, Padmanabhan B, Bhat A, Sarvepalli K, Sadhale PP, Nath U (2011) The TCP4 transcription factor of Arabidopsis blocks cell division in yeast at G1→S transition. Biochem Biophys Res Commun **410:** 276–281

Aoki N, Ohnishi J-i, Kanai R (1992) Two different mechanisms for transport of pyruvate into mesophyll chloroplasts of C_4 plants: a comparative study. Plant Cell Physiol **33:** 805–809

Aoyagi K, Nakamoto H (1985) Pyruvate, Pi dikinase in bundle sheath strands as well as in mesophyll cells in maize leaves. Plant Physiol **78:** 661–664

Botha C, Cross R, van Bel A, Peter C (2000) Phloem loading in the sucrose-export-defective (*SXD-1*) mutant maize is limited by callose deposition at plasmodesmata in bundle sheath-vascular parenchyma interface. Protoplasma **214:** 65–72

Collinge M, Boller T (2001) Differential induction of two potato genes, Stprx2 and StNAC, in response to infection by Phytophthora infestans and to wounding. Plant Mol Biol **46:** 521–529

Cribb L, Hall LN, Langdale JA (2001) Four mutant alleles elucidate the role of the G2 protein in the development of $C(_4)$ and $C(_3)$ photosynthesizing maize tissues. Genetics **159:** 787–797

Cubas P, Lauter N, Doebley J, Coen E (1999) The TCP domain: a motif found in proteins regulating plant growth and development. Plant J **18:** 215–222

Dai Z, Ku M, Edwards GE (1993) C_4 photosynthesis (the CO_2-concentrating mechanism and photorespiration). Plant Physiol **103:** 83–90

Dai Z, Ku M, Edwards GE (1995) C_4 photosynthesis (the effects of leaf development on the CO_2-concentrating mechanism and photorespiration in maize). Plant Physiol **107:** 815–825

de Veau EJ, Burris JE (1989) Photorespiratory rates in wheat and maize as determined by o-labeling. Plant Physiol **90:** 500–511

Evans JR, Von Caemmerer S (1996) Carbon dioxide diffusion inside leaves. Plant Physiol **110:** 339–346

Evert RF, Eschrich W, Heyser W (1977) Distribution and structure of the plasmodesmata in mesophyll and bundle-sheath cells of *Zea mays* L. Planta **136:** 77–89

Finet C, Fourquin C, Vinauger M, Berne-Dedieu A, Chambrier P, Paindavoine S, Scutt CP (2010) Parallel structural evolution of auxin response factors in the angiosperms. Plant J **63:** 952–959

Flugge UI (1999) Phosphate translocators in plastids. Annu Rev Plant Physiol Plant Mol Biol **50:** 27–45

Friso G, Majeran W, Huang M, Sun Q, van Wijk KJ (2010) Reconstruction of metabolic pathways, protein expression, and homeostasis machineries across maize bundle sheath and mesophyll chloroplasts: large-scale quantitative proteomics using the first maize genome assembly. Plant Physiol **152:** 1219–1250

Furbank RT (2011) Evolution of the $C(_4)$ photosynthetic mechanism: are there really three $C(_4)$ acid decarboxylation types? J Exp Bot **62:** 3103–3108

Furumoto T, Hata S, Izui K (1999) cDNA cloning and characterization of maize phosphoenolpyruvate carboxykinase, a bundle sheath cell-specific enzyme. Plant Mol Biol **41:** 301–311

Furumoto T, Hata S, Izui K (2000) Isolation and characterization of cDNAs for differentially accumulated transcripts between mesophyll cells and bundle sheath strands of maize leaves. Plant Cell Physiol **41:** 1200–1209

Furumoto T, Yamaguchi T, Ohshima-Ichie Y, Nakamura M, Tsuchida-Iwata Y, Shimamura M, Ohnishi J, Hata S, Gowik U, Westhoff P, et al (2011) A plastidial sodium-dependent pyruvate transporter. Nature **476:** 472–475

Gamas P, Niebel FdeC, Lescure N, Cullimore J (1996) Use of a subtractive hybridization approach to identify new Medicago truncatula genes induced during root nodule development. Mol Plant Microbe Interact **9:** 233–242

Gowik U, Bräutigam A, Weber KL, Weber AP, Westhoff P (2011) Evolution of C_4 photosynthesis in the genus *Flaveria*: how many and which genes does it take to make C_4? Plant Cell **23:** 2087–2105

Gowik U, Westhoff P (2011) The path from C_3 to C_4 photosynthesis. Plant Physiol **155:** 56–63

Gray JE (2007) Plant development: three steps for stomata. Curr Biol **17:** R213–R215

Guan YF, Huang XY, Zhu J, Gao JF, Zhang HX, Yang ZN (2008) *RUPTURED POLLEN GRAIN1*, a member of the MtN3/saliva gene family, is crucial for exine pattern formation and cell integrity of microspores in Arabidopsis. Plant Physiol **147:** 852–863

Gutierrez M, Huber SC, Ku SB, Kanai R, Edwards GE (1974). Intracellular localization of carbon metabolism in mesophyll cells of C_4 plants. *In* M Avron, ed, III International Congress on Photosynthesis Research. Elsevier Science Publishers, Amsterdam, The Netherlands, pp 1219–1230

Haake V, Cook D, Riechmann JL, Pineda O, Thomashow MF, Zhang JZ (2002) Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis. Plant Physiol **130:** 639–648

Hall LN, Rossini L, Cribb L, Langdale JA (1998) GOLDEN 2: a novel transcriptional regulator of cellular differentiation in the maize leaf. Plant Cell **10:** 925–936

Hatch MD, Agostino A (1992) Bilevel disulfide group reduction in the activation of C_4 leaf nicotinamide adenine dinucleotide phosphate-malate dehydrogenase. Plant Physiol **100:** 360–366

Hatch MD, Agostino A, Jenkins C (1995) Measurement of the leakage of CO_2 from bundle-sheath cells of leaves during C_4 photosynthesis. Plant Physiol **108:** 173–181

Heidstra R, Welch D, Scheres B (2004) Mosaic analyses using marked activation and deletion clones dissect Arabidopsis SCARECROW action in asymmetric cell division. Genes Dev **18:** 1964–1969

Hibberd JM, Sheehy JE, Langdale JA (2008) Using C_4 photosynthesis to increase the yield of rice: rationale and feasibility. Curr Opin Plant Biol **11:** 228–231

Huq E, Al-Sady B, Hudson M, Kim C, Apel K, Quail PH (2004) Phytochrome-interacting factor 1 is a critical bHLH regulator of chlorophyll biosynthesis. Science **305:** 1937–1941

Ivanov B, Asada K, Edwards GE (2007) Analysis of donors of electrons to photosystem I and cyclic electron flow by redox kinetics of P700 in chloroplasts of isolated bundle sheath strands of maize. Photosynth Res **92:** 65–74

Kanai R, Edwards GE (1973) Separation of mesophyll protoplasts and bundle sheath cells from maize leaves for photosynthetic studies. Plant Physiol **51:** 1133–1137

Leegood RC (2002) $C(_4)$ photosynthesis: principles of $CO(_2)$ concentration and prospects for its introduction into $C(_3)$ plants. J Exp Bot **53:** 581–590

Leegood RC (2008) Roles of the bundle sheath cells in leaves of C_3 plants. J Exp Bot **59:** 1663–1673

Leegood RC, Walker RP (2003) Regulation and roles of phosphoenolpyruvate carboxykinase in plants. Arch Biochem Biophys **414:** 204–210

Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, Kebrom TH, Provart N, Patel R, Myers CR, et al (2010) The developmental dynamics of the maize leaf transcriptome. Nat Genet **42:** 1060–1067

Majeran W, Cai Y, Sun Q, van Wijk KJ (2005) Functional differentiation of bundle sheath and mesophyll maize chloroplasts determined by comparative proteomics. Plant Cell **17:** 3111–3140

Markelz NH, Costich DE, Brutnell TP (2003) Photomorphogenic responses in maize seedling development. Plant Physiol **133:** 1578–1591

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods **5:** 621–628

Murphy LR, Barroca J, Franceschi VR, Lee R, Roalson EH, Edwards GE, Ku MSB (2007) Diversity and plasticity of C_4 photosynthesis in *Eleocharis* (Cyperaceae). Funct Plant Biol **34:** 571–580

Nakamura H, Muramatsu M, Hakata M, Ueno O, Nagamura Y, Hirochika H, Takano M, Ichikawa H (2009) Ectopic overexpression of the transcription factor OsGLK1 induces chloroplast development in non-green rice cells. Plant Cell Physiol **50:** 1933–1949

Nakazono M, Qiu F, Borsuk LA, Schnable PS (2003) Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: identification of genes expressed differentially in epidermal cells or vascular tissues of maize. Plant Cell **15:** 583–596

Novillo F, Medina J, Salinas J (2007) Arabidopsis CBF1 and CBF3 have a different function than CBF2 in cold acclimation and define different gene classes in the CBF regulon. Proc Natl Acad Sci USA **104:** 21002–21007

Ohnishi J, Flügge UI, Heldt HW, Kanai R (1990) Involvement of Na^+ in active uptake of pyruvate in mesophyll chloroplasts of some C_4 plants: Na^+/pyruvate cotransport. Plant Physiol **94:** 950–959

Ohnishi J, Kanai R (1990) Pyruvate uptake induced by a pH jump in mesophyll chloroplasts of maize and sorghum, NADP-malic enzyme type C_4 species. FEBS Lett **269:** 122–124

Ranocha P, Denancé N, Vanholme R, Freydier A, Martinez Y, Hoffmann L, Köhler L, Pouzet C, Renou JP, Sundberg B, et al (2010) *Walls are thin 1* (*WAT1*), an Arabidopsis homolog of Medicago truncatula *NODULIN21*, is a tonoplast-localized protein required for secondary wall formation in fibers. Plant J **63:** 469–483

Robards AW, Lucas WJ (1990) Plasmodesmata. Annu Rev Plant Physiol Plant Mol Biol **41:** 369–419

Sawers RJ, Liu P, Anufrikova K, Hwang JT, Brutnell TP (2007) A multitreatment experimental system to examine photosynthetic differentiation in the maize leaf. BMC Genomics **8:** 12

Sharpe RM, Mahajan A, Takacs EM, Stern DB, Cahoon AB (2011) Developmental and cell type characterization of bundle sheath and mesophyll chloroplast transcript abundance in maize. Curr Genet **57:** 89–102

Sheen J (1995) Methods for mesophyll and bundle sheath cell separation. Methods Cell Biol **49:** 305–314

Song CP, Agarwal M, Ohta M, Guo Y, Halfter U, Wang P, Zhu JK (2005) Role of an *Arabidopsis* AP2/EREBP-type transcriptional repressor in abscisic acid and drought stress responses. Plant Cell **17:** 2384–2396

Tabata R, Ikezaki M, Fujibe T, Aida M, Tian CE, Ueno Y, Yamamoto KT, Machida Y, Nakamura K, Ishiguro S (2010) Arabidopsis auxin response factor6 and 8 regulate jasmonic acid biosynthesis and floral organ development via repression of class 1 *KNOX* genes. Plant Cell Physiol **51:** 164–175

Taniguchi Y, Nagasaki J, Kawasaki M, Miyake H, Sugiyama T, Taniguchi M (2004) Differentiation of dicarboxylate transporters in mesophyll and bundle sheath chloroplasts of maize. Plant Cell Physiol **45:** 187–200

Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J **37:** 914–939

Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25:** 1105–1111

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol **28:** 511–515

Truong HN, Caboche M, Daniel-Vedele F (1997) Sequence and characterization of two *Arabidopsis thaliana* cDNAs isolated by functional complementation of a yeast *gln3 gdh1* mutant. FEBS Lett **410:** 213–218

Weber AP, von Caemmerer S (2010) Plastid transport and metabolism of C_3 and C_4 plants: comparative analysis and possible biotechnological exploitation. Curr Opin Plant Biol **13:** 257–265

Wingler A, Walker RP, Chen ZH, Leegood RC (1999) Phospho*enol*pyruvate carboxykinase is involved in the decarboxylation of aspartate in the bundle sheath of maize. Plant Physiol **120:** 539–546

Xu X, Chen C, Fan B, Chen Z (2006) Physical and functional interactions between pathogen-induced *Arabidopsis* WRKY18, WRKY40, and WRKY60 transcription factors. Plant Cell **18:** 1310–1326

Yokoyama A, Yamashino T, Amano Y, Tajima Y, Imamura A, Sakakibara H, Mizuno T (2007) Type-B ARR transcription factors, ARR10 and ARR12, are implicated in cytokinin-mediated regulation of protoxylem differentiation in roots of *Arabidopsis thaliana*. Plant Cell Physiol **48:** 84–96

Zhang H, Jin J, Tang L, Zhao Y, Gu X, Gao G, Luo J (2011) PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database. Nucleic Acids Res **39:** D1114–D1117

Zheng Z, Qamar SA, Chen Z, Mengiste T (2006) Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. Plant J **48:** 592–605

Zhong R, Demura T, Ye ZH (2006) SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of *Arabidopsis*. Plant Cell **18:** 3158–3170

Zhong R, Lee C, Zhou J, McCarthy RL, Ye ZH (2008) A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. Plant Cell **20:** 2763–2782