



RESEARCH PAPER

# Most photorespiratory genes are preferentially expressed in the bundle sheath cells of the C<sub>4</sub> grass *Sorghum bicolor*

Florian Döring<sup>1</sup>, Monika Streubel<sup>1</sup>, Andrea Bräutigam<sup>2,3,\*</sup> and Udo Gowik<sup>1,†</sup>

<sup>1</sup> Institute of Plant Molecular and Developmental Biology, Universitätsstrasse 1, Heinrich-Heine-University, D-40225 Düsseldorf, Germany

<sup>2</sup> Institute of Plant Biochemistry, Universitätsstrasse 1, Heinrich-Heine-University, D-40225 Düsseldorf, Germany

<sup>3</sup> Cluster of Excellence on Plant Sciences (CEPLAS) ‘From Complex Traits towards Synthetic Modules’, D-40225 Düsseldorf, Germany

\* Present address: Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Corrensstraße 3, D-06466 Stadt Seeland, Germany.

† Correspondence: [gowik@uni-duesseldorf.de](mailto:gowik@uni-duesseldorf.de)

Received 2 December 2015; Accepted 21 January 2016

Editor: Martin Hagemann, University Rostock

## Abstract

One of the hallmarks of C<sub>4</sub> plants is the division of labor between two different photosynthetic cell types, the mesophyll and the bundle sheath cells. C<sub>4</sub> plants are of polyphyletic origin and, during the evolution of C<sub>4</sub> photosynthesis, the expression of thousands of genes was altered and many genes acquired a cell type-specific or preferential expression pattern. Several lines of evidence, including computational modeling and physiological and phylogenetic analyses, indicate that alterations in the expression of a key photorespiration-related gene, encoding the glycine decarboxylase P subunit, was an early and important step during C<sub>4</sub> evolution. Restricting the expression of this gene to the bundle sheath led to the establishment of a photorespiratory CO<sub>2</sub> pump. We were interested in whether the expression of genes related to photorespiration remains bundle sheath specific in a fully optimized C<sub>4</sub> species. Therefore we analyzed the expression of photorespiratory and C<sub>4</sub> cycle genes using RNA *in situ* hybridization and transcriptome analysis of isolated mesophyll and bundle sheath cells in the C<sub>4</sub> grass *Sorghum bicolor*. It turns out that the C<sub>4</sub> metabolism of *Sorghum* is based solely on the NADP-dependent malic enzyme pathway. The majority of photorespiratory gene expression, with some important exceptions, is restricted to the bundle sheath.

**Key words:** C<sub>4</sub> photosynthesis, CO<sub>2</sub> fixation, differential gene expression, evolution, photorespiration, *Sorghum bicolor*.

## Introduction

C<sub>4</sub> plants evolved multiple times from C<sub>3</sub> ancestors. The C<sub>4</sub> photosynthetic pathway leads to concentration of CO<sub>2</sub> around the main carboxylating enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO). This is achieved by a set of anatomical and biochemical modifications to the original C<sub>3</sub> pathway (Hatch, 1987). In the presence of high CO<sub>2</sub> concentrations, the oxygenase activity of RubisCO, which always competes with the carboxylation reaction, is

effectively suppressed and hence photorespiration is strongly reduced in C<sub>4</sub> plants (Hatch, 1987). Photorespiration occurs when O<sub>2</sub> is used by RubisCO, which leads to the production of 2-phosphoglycolate (2-PG), a compound which is toxic for the plant cell and which needs to be detoxified (Anderson, 1971). Photorespiration takes place in chloroplasts, peroxisomes, and mitochondria. Throughout the regeneration of phosphoglycerate from phosphoglycolate, previously fixed

$\text{CO}_2$  is lost and additional energy and reduction equivalents are needed. Hence photorespiration can reduce the efficiency of photosynthesis in  $\text{C}_3$  species by up to 30% (Ogren, 1984; Bauwe et al., 2010; Raines, 2011; Fernie et al., 2013). Therefore,  $\text{C}_4$  photosynthesis can be of great advantage in conditions that promote photorespiration, such as hot, arid, and saline environments, in which plants have to close their stomata in order to avoid water loss through transpiration but which in consequence hinders the uptake of  $\text{CO}_2$  (Sage, 2004).  $\text{C}_4$  plants can keep their stomata closed for a longer time, because the  $\text{CO}_2$  pump facilitates high rates of photosynthesis even under low  $\text{CO}_2$  concentrations in the intercellular air space of the leaf and therefore minimizes water loss.

Leaves of  $\text{C}_4$  plants show anatomical differences compared with those of  $\text{C}_3$  plants. The vascular bundles are surrounded by organelle-rich bundle sheath cells, which, in turn, are surrounded by mostly one layer of mesophyll cells. This leads to a wreath-like appearance, which is termed Kranz anatomy (Haberlandt, 1904; Laetsch, 1974). In  $\text{C}_4$  leaves, bundle sheath cells are enlarged and the interveinal distance is reduced (Dengler and Nelson, 1999). To allow the efficient interchange of metabolites between mesophyll and bundle sheath cells, both cell types are connected through numerous plasmodesmata (Botha, 1992).

In most species,  $\text{C}_4$  photosynthesis largely depends on the division of labor between mesophyll and bundle sheath cells, in which the  $\text{CO}_2$  assimilatory enzymes are compartmentalized. The  $\text{C}_4$  pathway begins with the conversion of  $\text{CO}_2$  to bicarbonate by carbonic anhydrase (CA) in the cytosol of mesophyll cells and the subsequent fixation into the  $\text{C}_4$  acid oxaloacetate by phosphoenolpyruvate carboxylase (PEPC) with the 3-carbon compound phosphoenolpyruvate (PEP) as  $\text{CO}_2$  acceptor. Afterwards, oxaloacetate is either reduced to malate or transaminated to aspartate, which is transported to the bundle sheath cells. There,  $\text{CO}_2$  is released by decarboxylation of the  $\text{C}_4$  compounds through a decarboxylating enzyme, either an NADP-dependent malic enzyme (NADP-ME), an NAD-dependent malic enzyme (NAD-ME), a PEP-carboxykinase (PEP-CK), or, as shown recently, a combination of these (Furbank, 2011; Y. Wang et al., 2014). The released  $\text{CO}_2$  is immediately refixed by RubisCO and enters the Calvin–Benson cycle. Less RubisCO is needed compared with  $\text{C}_3$  plants as it works more efficiently under these conditions (Long, 1999). This results in a better nitrogen use efficiency of  $\text{C}_4$  plants, since RubisCO is by far the most abundant protein in the leaves of higher plants (Long, 1999). Pyruvate, the other product of the decarboxylation, is transferred to the mesophyll cells where PEP is regenerated by pyruvate phosphate dikinase (PPDK).

$\text{C}_4$  photosynthesis has evolved at least 66 times independently from the original  $\text{C}_3$  pathway (Sage et al., 2011, 2012). To better understand the changes underlying the evolution of  $\text{C}_4$  on the gene level, in recent years several studies aimed at creating transcriptome atlases of total leaf RNA of various pairs of closely related  $\text{C}_4$  and  $\text{C}_3$  species (Bräutigam et al., 2011, 2014; Gowik et al., 2011; Mallmann et al., 2014). The development of  $\text{C}_3$  and  $\text{C}_4$  leaves was studied by analyzing the gene expression in different developmental stages of dicot

leaves and the developmental gradients found in the leaves of  $\text{C}_3$  and  $\text{C}_4$  grasses (Li et al., 2010; Pick et al., 2011; Kulahoglu et al., 2014; L. Wang et al., 2014; Ding et al., 2015). The coordination of the two different cell types was analyzed using mesophyll and bundle sheath transcriptomes of the  $\text{C}_4$  grasses maize and *Setaria viridis* (Li et al., 2010; Chang et al., 2012; John et al., 2014; Tausta et al., 2014). It turned out that  $\text{C}_4$  photosynthesis is a complex trait and its evolution involved changes in the expression of thousands of genes. Genes encoding the enzymes and transporters of the  $\text{C}_4$  pathway had to be up-regulated and acquired tissue-specific expression. In addition, several other metabolic pathways must also have been regulated differentially in mesophyll and bundle sheath cells to enable this efficient type of photosynthesis including high nitrogen and water use efficiency attributed to  $\text{C}_4$  plants.

It is widely accepted that the development of a photorespiratory  $\text{CO}_2$  pump, often termed  $\text{C}_2$  photosynthesis, was an important intermediate step during the evolution of the  $\text{C}_4$  pathway (Bauwe, 2011; Sage et al., 2012; Heckmann et al., 2013; Williams et al., 2013). The photorespiratory pump is based on the restriction of one of the key photorespiratory enzyme complexes, the glycine decarboxylase complex (GDC), to the bundle sheath cells (Rawsthorne et al., 1988a). Photorespiratory glycine has to move to the bundle sheath for decarboxylation, and  $\text{CO}_2$  is released mainly in this compartment, leading to increased  $\text{CO}_2$  concentrations and allowing RubisCO to work more efficiently (Bauwe, 2011; Heckmann et al., 2013). The photorespiratory pump can lead to a 3-fold enrichment of  $\text{CO}_2$  in the bundle sheath cells (Keerberg et al., 2014). The analysis of  $\text{C}_3$ – $\text{C}_4$  intermediate *Flaveria* species implied that the effect of the photorespiratory pump on  $\text{C}_4$  evolution might be quite direct and provided a mechanistic explanation for how the photorespiratory pump and  $\text{C}_4$  photosynthesis interact (Mallmann et al., 2014). The glycine shuttle induces a nitrogen imbalance between mesophyll and bundle sheath cells, and the introduction of important components of the  $\text{C}_4$  pathway, as well as the  $\text{C}_4$  pathway itself, are highly efficient ways to correct this imbalance. This implies that  $\text{C}_4$  evolution is a metabolic exaptation as the  $\text{C}_4$  pathway developed in the first place to transport nitrogen and was not directly related to improving photosynthetic efficiency (Mallmann et al., 2014). Hence, photorespiration and the cell-specific expression of photorespiratory genes in the mesophyll and bundle sheath cells of  $\text{C}_3$ – $\text{C}_4$  intermediates were of key importance for the evolution of  $\text{C}_4$  photosynthesis.

In the present study, we examined how the expression of photorespiratory genes changed after the transition to true  $\text{C}_4$  photosynthesis. Therefore we analyzed the expression of photosynthetic and photorespiratory genes in the  $\text{C}_4$  grass *Sorghum bicolor* by RNA *in situ* hybridization and transcriptome analysis of isolated mesophyll and bundle sheath fractions. *Sorghum bicolor* is a highly optimized plant species with regard to the  $\text{C}_4$  pathway. Methods for the isolation of mesophyll and bundle sheath cells are available (Wyrich et al., 1998) and its genome is fully sequenced (Paterson et al., 2009), allowing transcriptome analysis with plain high-throughput sequencing as well as with a serial analysis of gene expression (SAGE) approach since the short sequence reads could be

directly mapped to the genome or the derived transcriptome sequence (Bräutigam and Gowik, 2010). We determined transcript abundances within our mesophyll and bundle sheath RNA preparations by Illumina sequencing and additionally by SuperSAGE (Matsumura *et al.*, 2003), a combination of SAGE with next-generation sequencing methods.

We hypothesized that the distribution of photorespiratory gene expression is similar to the enzyme distributions determined previously (Ohnishi and Kanai, 1983; Gardeström *et al.*, 1985; Ohnishi *et al.*, 1985) and that it is comparable in specificity with the distribution of genes related to the C<sub>4</sub> pathway.

## Materials and methods

### Plant material, RNA isolation, and cDNA synthesis

*Sorghum bicolor* L. Tx430 (Pioneer Hi-Bred, Plainview, TX, USA) was grown on soil (Floraton 1, Floragard, Oldenburg, Germany) in the greenhouse of the Heinrich-Heine University (Düsseldorf, Germany) with supplementary light for 14 h per day ( $\sim 300 \mu\text{mol m}^{-2} \text{s}^{-1}$ ). For the *in situ* analysis, we harvested the middle thirds of the second leaf from 3-week-old plants and took 2 × 5 mm sections from it. For isolation of mesophyll and bundle sheath RNA, we harvested the upper two-thirds of the second leaf from 10-day-old seedlings. For generation of the cell-specific mRNAs, we separated the bundle sheath and vascular bundles enzymatically from the mesophyll and epidermal cells as described in Wyrich *et al.* (1998). We isolated 15 independent mesophyll and 19 independent bundle sheath samples. Cross-contaminations of the RNA preparations were controlled by dot blot analysis following standard procedures. Five independent mesophyll and bundle sheath preparations were pooled for the SuperSAGE analysis. For cDNA synthesis and Illumina sequencing, we pooled five other preparations for each tissue. Total RNA from intact *Sorghum* leaves was isolated according to Westhoff *et al.* (1991). Poly(A)<sup>+</sup> RNA was enriched by two consecutive rounds of oligo(dT) purification with the Oligotex mRNA Midi Kit (Qiagen, Hilden, Germany). cDNA libraries for Illumina sequencing were prepared with the SMARTer PCR cDNA Synthesis Kit (Clontech-Takara Bio Company, Otsu, Japan), with 300 ng of poly(A)<sup>+</sup> RNA as starting material. The purity and integrity of total RNA, poly(A)<sup>+</sup> RNA, and cDNA were verified spectroscopically with a NanoDrop ND-1000, with the Agilent 2100 Bioanalyzer and by agarose gel electrophoresis.

### SuperSage/Illumina sequencing

The SuperSAGE analysis was performed by GenXPro Inc. (Frankfurt, Germany) (Matsumura *et al.*, 2003). The mesophyll, bundle sheath, and total cDNA libraries were sequenced each in one lane of an Illumina flow cell with an Illumina Genome Analyser II by GATC Biotech AG (Konstanz, Germany) following standard protocols. The read length was 40 bp. The cDNAs were prepared from pooled total RNAs.

### Mapping/statistics

The SuperSAGE tags as well as the Illumina reads were mapped on the *S. bicolor* transcriptome [version 1.4 (Sbicolor\_79\_transcript\_primaryTranscriptOnly.fa) in the case of the SuperSAGE tags, and version 3.1 (Sbicolor\_313\_v3.1.transcript\_primaryTranscriptOnly.fa) in the case of the Illumina reads (<http://phytozome.jgi.doe.gov>)]. The SuperSAGE tags were mapped with BLAST (Altschul *et al.*, 1990) by GenXPro Inc. Two mismatches were allowed and only tags that were found at least twice were counted. Tag counts were transformed to tags per million (tpm). For the mapping of the Illumina

reads, we used BOWTIE (Langmead *et al.*, 2009). The best hit for each Illumina read was retained, and hit counts were then transformed to reads per kilobase and million (RPKM) to normalize for the number of reads available for each cDNA library.

Log<sub>2</sub> ratios were calculated and differentially expressed transcripts were called using the R package DEGseq (Wang *et al.*, 2010) on the non-normalized read counts followed by a Bonferroni correction to account for the accumulation of alpha-type errors when conducting multiple pairwise comparisons.

### qRT-PCR

Quantitative real-time PCR (qRT-PCR) followed standard procedures and was performed with an ABI7500 fast Real Time PCR system. The primers were designed to target photorespiratory genes of *S. bicolor* and to generate amplicons of 170 bp. The specificity of PCRs was verified by melting curve analysis and agarose gel electrophoresis. To estimate the efficiency of the PCRs, four consecutive 5-fold dilutions of the cDNAs were tested with each primer pair. Only reactions with efficiencies >90% were considered for further analysis. As template we used total RNAs pooled from five independent mesophyll and bundle sheath preparations each, not used for SuperSAGE or Illumina sequencing.

### RNA *in situ* hybridization

The tissue was fixed for 16 h in a mixture of 3.7% formaldehyde, 50% ethanol, and 5% acetic acid at 4 °C. Dehydration and embedding was done in the Tissue Processor Leica ASP300S using the following program: 1 h in 50% ethanol, 1 h in 70% ethanol, 1 h in 95% ethanol, 3 × 1 h in 100% ethanol, 2 × 1 h in 100% xylene, 1 h in 100% xylene (37 °C), 2 × 10 min in histowax (62 °C), and 20 min in histowax (62 °C). Subsequently the samples were embedded in paraffin and cut into 12 µm sections with a microtome.

Probe labeling: for the generation of hybridization probes, the respective cDNAs were amplified by PCR and cloned into pJET1.2/blunt plasmid (Thermo Scientific, St. Leon-Rot, Germany). After linearization of the vector with appropriate restriction enzymes, T7 RNA polymerase was used to generate both sense and antisense probes, which were labeled with digoxigenin (DIG)-labeled UTP using the DIG RNA Labeling kit (Roche, Mannheim, Germany). Subsequently the probes were hydrolyzed to a size of ~150–200 bases.

Pre-hybridization, hybridization, and post-hybridization steps were based on the protocol described by Simon (2002). Only deviations from this protocol are mentioned below. First the sections were dewaxed in Roti®-Histol for 10 min and rehydrated in a decreasing ethanol concentration series: 2 × 1 min in 100% ethanol, 1 min in 95% ethanol, 1 min in 85% ethanol, 1 min in 50% ethanol, 1 min in 30% ethanol, and 1 min in ddH<sub>2</sub>O. Afterwards the sections were treated with 10 µg ml<sup>-1</sup> proteinase K for 30 min at 37 °C, post-fixed and acetylated as described by Simon (2002), and finally dehydrated in a reverse order of the ethanol concentration series used before. For the hybridization, 150 ng of probe was used for each slide. The sections were incubated for 16 h at 50 °C in a humid chamber.

After hybridization, the sections were washed three times in washing buffer (2 × SSC, 50% formamide) for 30 min at 50 °C and twice in NTE buffer (500 mM NaCl, 10 mM Tris, 1 mM EDTA, pH 8.0) for 5 min at 37 °C. After RNase A treatment, the sections were washed again twice in NTE at room temperature for 5 min and in washing buffer for 1 h at 50 °C.

For immunological detection, all steps were performed on a shaking platform. First the sections were washed in buffer 1 (100 mM Tris-HCl pH 7.5, 150 mM NaCl) for 5 min, before they were incubated in buffer 2 (buffer 1 containing 0.5% blocking reagent; Roche) for 40 min. Subsequently they were incubated in buffer 3 (buffer 1 containing 0.3% Triton X-100, 1% normal sheep serum, and sheep anti-DIG-alkaline phosphatase at a dilution of 1:2000) for 2 h, after which they were washed four times

in buffer 1 containing 0.3% Triton X-100 for 15 min. Then the sections were washed in buffer 1 for 5 min, incubated in buffer 4 (0.1 M Tris-HCl pH 9.5, 0.1 M NaCl, and 50 mM MgCl<sub>2</sub>) for 5 min, and finally stained in buffer 5 [buffer 4 containing 10% polyvinyl alcohol, 0.16 mM nitroblue tetrazolium (NBT), and 0.15 mM BZIP] in a humid chamber for 12–16 h. The reaction was stopped by washing the sections twice in distilled water, after which they were mounted with Entellan® (Merck Millipore, Darmstadt, Germany).

## Results

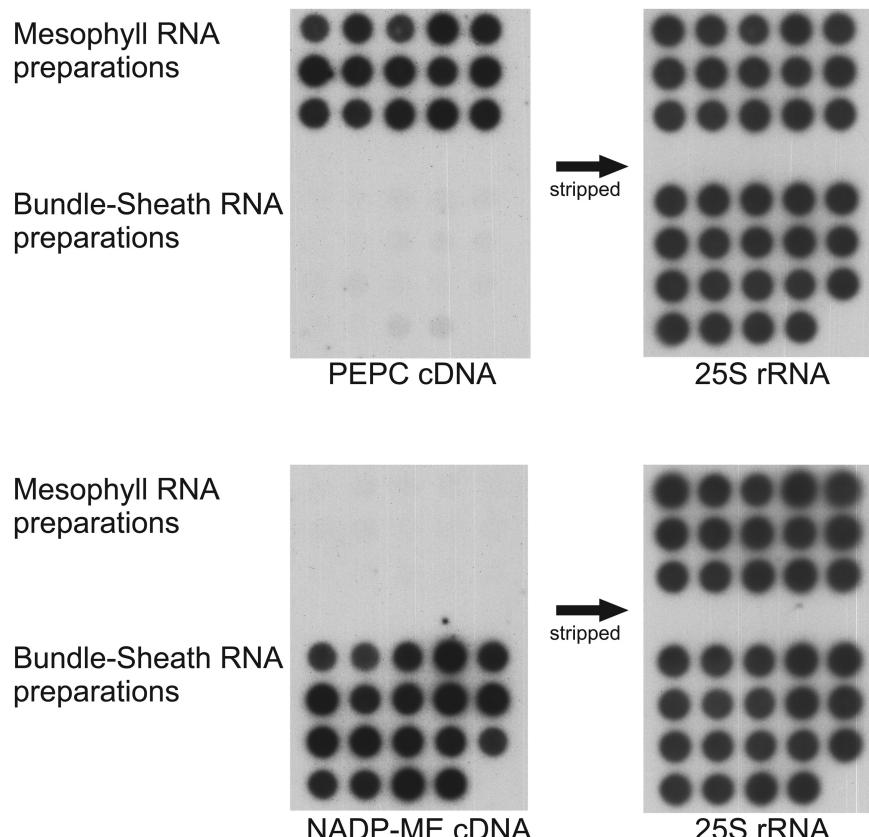
### Mesophyll and bundle sheath RNAs

Mesophyll and bundle sheath cells of *S. bicolor* for RNA preparations were separated by enzymatic digestion of leaf cell walls as described in Wyrich *et al.* (1998). It has to be considered that the mesophyll fraction also contains epidermis cells whereas the bundle sheath fraction contains all vascular tissues. The cross-contamination of mesophyll and bundle sheath preparations was analyzed by dot blot analysis using a PEPC and an NADP-ME cDNA as hybridization probes (Fig. 1). PEPC is thought to be mesophyll specific in *Sorghum* whereas NADP-ME was shown to be exclusively expressed in bundle sheath cells (Wyrich *et al.*, 1998). Since no signals indicating cross-contamination were visible, it can be assumed that the RNA preparations are pure and that the cross-contamination of mesophyll and bundle sheath RNAs is <5% (Fig. 1).

### Transcriptome analysis via SuperSage and RNA-Seq

To create transcriptome atlases of *Sorghum* bundle sheath and mesophyll tissue, we performed plain Illumina sequencing and a SuperSage analysis. With the SuperSage method, we obtained  $>6.8 \times 10^6$  tags (total leaf, 1 098 800; mesophyll, 3 349 814; bundle sheath, 2 421 27) that could be assigned to >12 000 (12 937) of the 34 211 predicted *Sorghum* genes, whereas 2327 genes exhibited a significantly different expression between mesophyll and bundle sheath cells ( $P < 0.01$ ) (Table 1). With plain Illumina sequencing we produced  $>36 \times 10^6$  reads (total leaf, 17 704 772; mesophyll, 10 420 446; bundle sheath, 8 695 328) which could be mapped to 23 244 *Sorghum* genes. With this method, we identified 1705 genes as being expressed significantly differentially between mesophyll and bundle sheath cells ( $P < 0.01$ ) (Table 1; Supplementary Table S1 at JXB online). With the SuperSage approach, we detected far fewer transcripts compared with the plain Illumina sequencing. This is most probably due to limitations of the SuperSage method. A transcript will not be recognized if the cleavage site of the anchoring enzyme, which is needed to produce the DNA fragments used as tags, is not present in the transcript (Matsumura *et al.*, 2003).

In total, we were able to detect 12 154 transcripts expressed within the *Sorghum* leaf with at least one read in both experiments, which corresponds to 35% of the total number of transcripts predicted from the *Sorghum* genome sequence (Paterson *et al.*, 2009). A total of 455 (3.7%) of them were more abundant in mesophyll cells and 401 (3.2%) in the bundle sheath in both experiments.



**Fig. 1.** Dot-blot analysis of independent mesophyll and bundle sheath RNA preparations. *Sorghum bicolor* PEPC cDNA, NADP-ME cDNA, and 25S rRNA were used as probes.

*The enzymatic separation of mesophyll and bundle sheath cells influences gene expression*

During the separation of mesophyll and bundle sheath cells by enzymatic digest, the tissue is incubated for up to 2.5 h at 25 °C. It is known that this treatment stresses the plant cells and leads to the expression of stress-related genes (Sawers *et al.*, 2007). To account for this problem, we isolated RNA from complete, unstressed *Sorghum* leaves. We assumed that mesophyll and bundle sheath RNA accounts for a comparable fraction of the whole leaf RNA. Based on this premise, we identified 3697 genes within the SuperSage experiment and 3724 genes within the RNA-Seq experiment that were up-regulated >3-fold apparently due to the enzymatic treatment. To test this assumption, we analyzed the representation of Gene Ontology (GO) terms for the up-regulated genes. Indeed, we found an over-representation of GO terms related to stress response among these 3-fold up-regulated genes in the SuperSage as well as in the RNA-Seq experiment (Tables 2, 3). The genes found to be >3-fold up- or down-regulated after enzyme treatment were tagged.

*The photorespiratory cycle mainly takes place in the bundle sheath in S. bicolor*

It was assumed earlier that in C<sub>4</sub> plants the photorespiratory pathway is mainly located in the bundle sheath cells since in C<sub>4</sub> plants, RubisCO, the entry enzyme of photorespiration, is restricted to

**Table 1.** Overview of the SuperSage and RNA-Seq results

	SuperSage	RNA-Seq
Total reads:	6 870 541	36 820 546
Genes detected ( <i>S. bicolor</i> 34 211 genes):	12 937	23 244
Percentage:	37	67
Differentially expressed:	2327	1705
Percentage:	6.8	4.9

**Table 2.** GO term over-representation analysis of genes up-regulated >3-fold in mesophyll or bundle sheath RNAs compared with total leaf RNA within the Illumina RNA-Seq experiment

The 10 most strongly over-represented GO terms are shown. Analysis was performed using the Gene Ontology Consortium database (<http://geneontology.org>).

GO term	GO name	P-value
GO:0050896	Response to stimulus	9.40E-17
GO:1901701	Response to oxygen-containing compound	1.57E-13
GO:0042221	Response to chemical	5.65E-12
GO:0001101	Response to acid chemical	5.65E-12
GO:0006950	Response to stress	2.83E-11
GO:0044699	Single-organism process	4.35E-11
GO:0071704	Single-organism cellular process	5.77E-11
GO:0009719	Response to endogenous stimulus	1.20E-10
GO:0071229	Cellular response to acid chemical	2.57E-10
GO:0010033	Response to organic substance	2.79E-10

P-values are corrected by the Bonferroni method.

this cell type (Bauwe, 2011). One exception is glycerate kinase (GLYK), which catalyzes the regeneration of 3-phosphoglycerate (3-PG) and was found to be restricted to the mesophyll cells (Usuda and Edwards, 1980). The present transcriptome analysis largely supports these expectations (Fig. 2; Supplementary Table S2), as do the *in situ* hybridizations (Fig. 2; Supplementary Fig. S1). We detected a strong signal in the bundle sheath for most transcripts of the core photorespiratory pathway with genes that show virtually no expression in the mesophyll and can be seen as bundle sheath specific, such as phosphoglycolate phosphatase (PGLP), glycolate oxidase (GOX), serine hydroxymethyl transferase (SHM), and the H, P, and T subunit of the GDC (Fig. 2; Supplementary Fig. S1). However, there are also genes such as glycine 2-oxoglutarate aminotransferase (GGT) and the GDC L subunit that, although preferentially expressed in the bundle sheath, still seem to be expressed to a certain extent in the mesophyll (Fig. 2; Supplementary Fig. S1). Taken together, this implies that all genes of the core photorespiratory pathway are at least preferentially if not specifically expressed in the bundle sheath, except for GLYK that is expressed to a much higher level in the mesophyll than in the bundle sheath (Fig. 2; Supplementary Table S2). We did not obtain any *in situ* hybridization signal for GLYK. This may be caused by the low absolute expression of the gene observed even in the mesophyll (Supplementary Table S2).

*The transcriptome analysis reveals detailed insight into the C<sub>4</sub> pathway of S. bicolor*

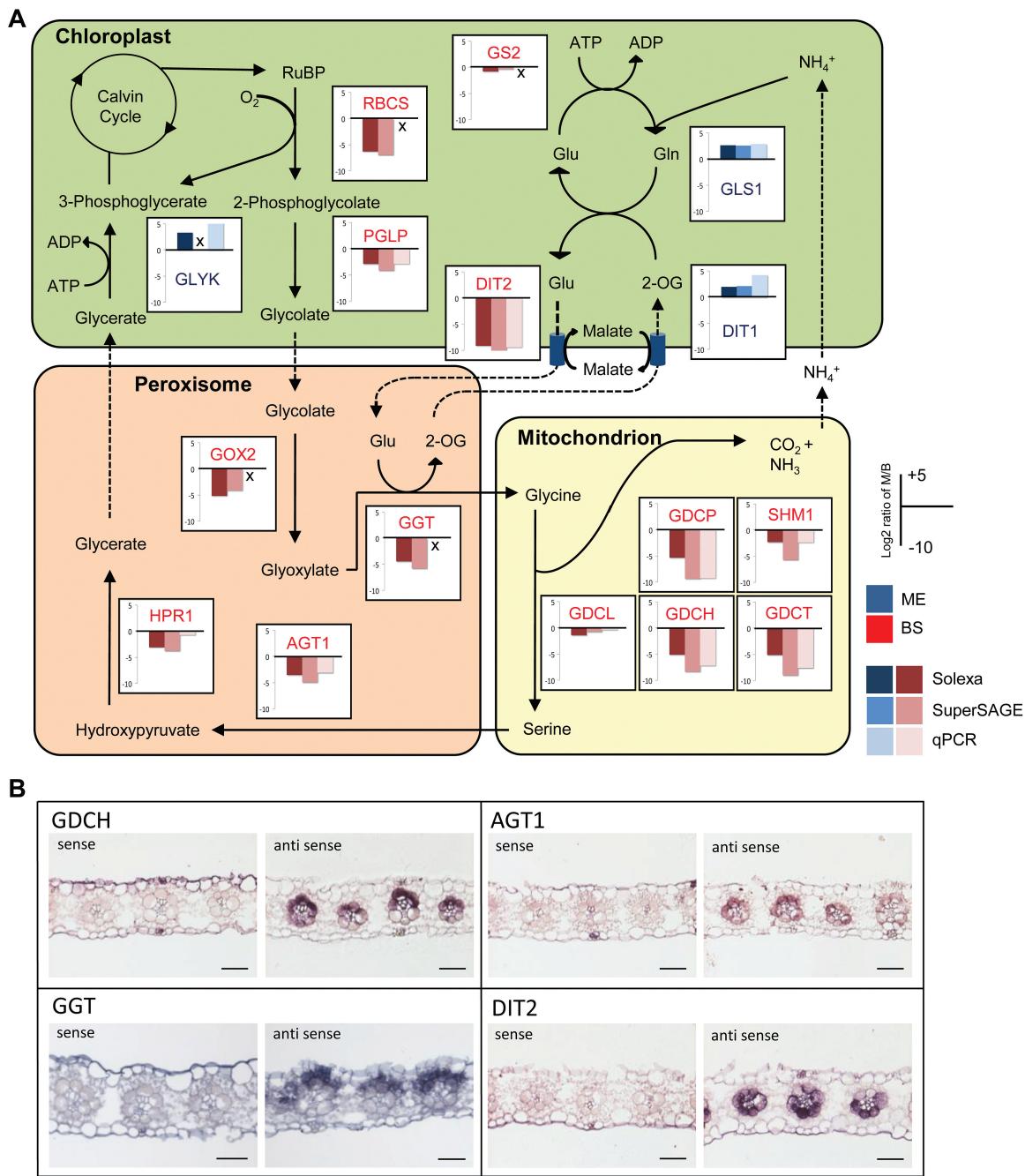
*Sorghum bicolor* belongs to the NADP-ME type of C<sub>4</sub> plants. The genes encoding PEPC, malate dehydrogenase (MDH), or PPDK are expected to be expressed specifically or at least strongly preferentially in the mesophyll in these plants, whereas the genes encoding NADP-ME or RubisCO are bundle sheath specific. The results of our transcriptome analyses are essentially in line with these expectations (Fig. 3; Supplementary Table S3). Although PEPC

**Table 3.** GO term over-representation analysis of genes up-regulated >3-fold in mesophyll or bundle sheath RNAs compared with total leaf RNA within the SuperSage experiment

The 10 most strongly over-represented GO terms are shown. Analysis was performed using the Gene Ontology Consortium database (<http://geneontology.org>).

GO term	GO name	P-value
GO:0042221	Response to chemical	3.88E-17
GO:1901700	Response to oxygen-containing compound	1.24E-16
GO:0050896	Response to stimulus	2.02E-16
GO:0009987	Cellular process	8.90E-16
GO:0044237	Cellular metabolic process	1.98E-15
GO:0044699	Single-organism process	1.73E-14
GO:0009628	Response to abiotic stimulus	1.83E-14
GO:0044710	Single-organism metabolic process	1.92E-14
GO:0006950	Response to stress	4.25E-14
GO:0010033	Response to organic substance	4.74E-14

P-values are corrected by the Bonferroni method.

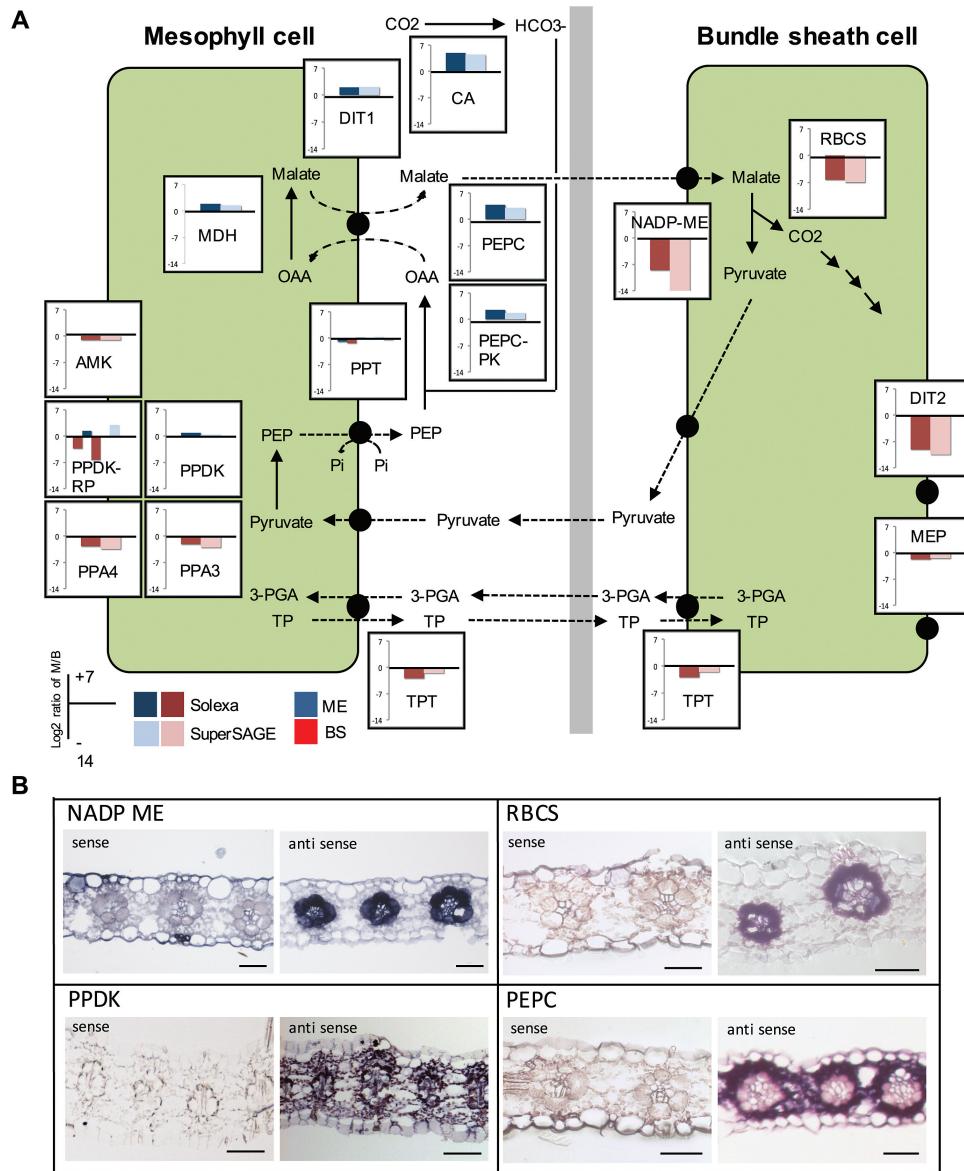


**Fig. 2.** (A) Distribution of photorespiratory genes between mesophyll and bundle sheath cells. Preferential gene expression in the mesophyll and bundle sheath is indicated by blue or red color, respectively. AGT, serine glyoxylate aminotransferase; DIT1+2, dicarboxylate transporter 1+2; GDCH/GDCL, glycine decarboxylase H, L, P, and T subunit; GGT, glutamate glyoxylate aminotransferase; GLS, glutamate synthase; GLYK, glycerate kinase; GOX2, glycolate oxidase 2; GS, glutamine synthetase; HPR, hydroxypyruvate reductases; PGLP, phosphoglycolate phosphatase; SHM, serine hydroxymethyltransferase; RBCS, ribulose bisphosphate carboxylase/oxygenase small subunit. (B) RNA *in situ* hybridization of *Sorghum bicolor* leaves with probes for transcripts related to photorespiration. Scale bars=50 µm.

was found to be expressed preferentially in the mesophyll, as expected, the absolute transcript levels as estimated by the Illumina sequencing appear to be quite low compared with NADP-ME or PPDK. In contrast, PEPC transcript levels turned out to be quite high when determined by the SuperSage method (Supplementary Table S3). If and how we selected against detecting high levels of the PEPC during the Illumina analysis is unclear. We detected virtually no expression of bundle sheath genes such as NADP-ME

or RubisCO in the mesophyll, indicating that our mesophyll RNA preparations were not cross-contaminated with bundle sheath RNA (Supplementary Tables S2, S3). The fact that we detected some expression of typical mesophyll genes such as PEPC in the bundle sheath indicates some contamination of our bundle sheath RNA preparation with mesophyll RNA in the range of ~5% (Supplementary Table S3).

Recent results indicate that the classification of the different types of the C<sub>4</sub> pathway is not as clear-cut as previously



**Fig. 3.** (A) Distribution of C<sub>4</sub> cycle genes between mesophyll and bundle sheath cells. Preferential gene expression in the mesophyll and bundle sheath is indicated by blue or red color, respectively. AMK, AMP kinase; CA, carbonic anhydrase; DIT1+2, dicarboxylate transporter 1+2; MDH, NADP-dependent malate dehydrogenase; MEP, mesophyll envelope protein; NADP-ME, NADP-dependent malic enzyme; PEPC, phosphoenolpyruvate carboxylase; PEPC-PK, phosphoenolpyruvate carboxylase protein kinase; PPA3+4, pyrophosphorylase 3+4; PPDK, pyruvate phosphate dikinase; PPT, phosphoenolpyruvate phosphate translocator; PPDK-RP, PPDK regulatory protein; RubisCO, ribulose bisphosphate carboxylase/oxygenase; TPT, triosephosphate phosphate translocator. (B) RNA *in situ* hybridization of *Sorghum bicolor* leaves with probes for transcripts related to the C<sub>4</sub> pathway. Scale bars=50 µm.

thought (Furbank, 2011; Pick *et al.*, 2011; Y. Wang *et al.*, 2014). Maize, which was assumed to be an archetypal NADP-ME-type C<sub>4</sub> plant, uses in parallel the PEP-CK type pathway to a considerable extent (Wingler *et al.*, 1999; Pick *et al.*, 2011). Interestingly this does not apply for *Sorghum*, although maize and *Sorghum* possess a common C<sub>4</sub> origin. We did not find a highly expressed PEP-CK gene in bundle sheath cells and no abundantly expressed NAD-ME genes could be detected (*Supplementary Table S3*). It follows that *Sorghum* instead of maize should be considered as the NADP-ME C<sub>4</sub> archetype. It was shown earlier that, in contrast to the common textbook models of this pathway, some NADP-ME species use alanine and aspartate as transport metabolites in parallel to malate and pyruvate (Meister *et al.*, 1996; Gowik *et al.*, 2011).

We were interested in whether the same is true for *Sorghum*, but the results are inconclusive. While we could identify a highly expressed aspartate aminotransferase (AspAT) gene in mesophyll as well as in bundle sheath cells, we have not found an alanine aminotransferase (AlaAT) that is highly expressed in both cell types. The most highly abundant AlaAT transcript, which belongs to the most abundant transcripts identified in this study, is clearly mesophyll specific. The function of this highly abundant AlaAT in the mesophyll remains unknown. We found another AlaAT gene which was significantly more highly expressed in the bundle sheath compared with the mesophyll (*Supplementary Table S3*) but, since its overall abundance is much lower, it is unclear if the overall AlaAT transcript abundance in the bundle sheath

allows the considerable usage of alanine and aspartate as transport metabolites. The up-regulated AspAT, on the other hand, is predicted to be localized in the chloroplast (TargetP score: 0.968). This is in line with other NADP-ME species that synthesize and decarboxylate aspartate in the chloroplasts of mesophyll and bundle sheath cells (Meister *et al.*, 1996; Gowik *et al.*, 2011).

Transcripts related to most of the known transporters thought to be directly involved in the NADP-ME C<sub>4</sub> pathway, such as the triosephosphate phosphate translocator (TPT), the PEP phosphate translocator (PPT), the dicarboxylate transporter (DIT/DCT/OMT), or the inner chloroplast envelope transporter MEP (Weber and von Caemmerer, 2010) could be identified, and most of them showed high abundance in agreement with their probable role in the C<sub>4</sub> pathway (Fig. 3; Supplementary Table S3). However, it has to be considered that they did not always show the expected distribution in the two cell types (e.g. the PPT was expected to be mesophyll specific but we also found high amounts of PPT transcripts in the bundle sheath). We could not detect high expression for the BASS2 and the NHD transporter that were shown to catalyze pyruvate transport across the chloroplast membrane in the C<sub>4</sub> *Flaveria* species (Furumoto *et al.*, 2011). This is in line with earlier results indicating that *Sorghum* uses a proton-dependent pyruvate transporter (Aoki *et al.*, 1992) instead of BASS, which was shown to be a pyruvate–sodium symporter (Furumoto *et al.*, 2011).

While most of the core C<sub>4</sub> genes are expressed either mesophyll or bundle sheath specifically, as expected, we found that PPDK transcripts are not only highly abundant in the mesophyll, but were also present in respectable amounts in the bundle sheath, with a mesophyll to bundle sheath ratio of only ~1 to 2 (Fig. 3). Along with that, we also found that transcripts related to the PPDK reaction such as pyrophosphatases, AMP kinase, or the PPT exhibit high levels in the bundle sheath cells and are partly even preferentially expressed in the bundle sheath (Supplementary Table S3).

To verify the tissue distribution of selected transcripts, we performed *in situ* hybridizations for typical C<sub>4</sub> genes such as PEPC, NADP-ME, PPDK, and RBCS (RubisCO small subunit). The obtained results largely support the outcome of the transcriptome analysis using SuperSage or RNA-Seq (compare Fig. 3A and B). *In situ* hybridization confirmed bundle sheath-specific expression for RBCS and NADP-ME, mesophyll-specific expression for the PEPC gene, and the preferential expression in the mesophyll cells of PPDK, with high PPDK transcript levels also in the bundle sheath.

#### *Expression patterns of genes associated with photorespiration are variable*

During photorespiration not only CO<sub>2</sub>, but also nitrogen is released in the mitochondria in the form of NH<sub>3</sub> that becomes reassimilated in the chloroplasts. In contrast to the core photorespiratory pathway, the genes for nitrogen assimilation and the dedicated transporters do not show a tissue-specific expression pattern. Glutamine synthetase as well as glutamate synthase genes are expressed in mesophyll and bundle sheath cells, but glutamine synthetase is more highly

expressed in the bundle sheath, and a ferredoxin-dependent glutamine oxoglutarate aminotransferase (Fd-GOGAT) shows higher transcript abundance in the mesophyll (Fig. 2; Supplementary Table D2).

Only a few transporters involved in the intracellular transport of photorespiratory metabolites are known to date. We could identify two transcripts corresponding to the plastid glycolate/glycerate transporter (Pick *et al.*, 2013). Whereas one of the genes appears not to be expressed at all in the *Sorghum* leaf, the other one exhibits high amounts of transcripts in both cell types, but the expression in the bundle sheath is higher than in the mesophyll (Fig. 2; Supplementary Table S2). The mitochondrial transporter BOU, known to be needed for functional photorespiration in *Arabidopsis thaliana* (Eisenhut *et al.*, 2013), appears to be expressed only at a low level in the leaves of the C<sub>4</sub> plant *Sorghum* and does not show a strong tissue preference (Supplementary Table S2). *Sorghum* contains five genes encoding dicarboxylate transporters (DITs); four of these transporters are classified as DIT2 and one is classified as a DIT1 gene. The DIT1 gene is expressed to moderate levels and clearly is expressed preferentially in the mesophyll. One of the DIT2 genes is highly expressed in the bundle sheath (Fig. 2; Supplementary Table S2). The two transporters are thought to interact in the glutamate–oxoglutarate exchange across the chloroplast membrane during NH<sub>3</sub> reassimilation (Renne *et al.*, 2003; Bauwe *et al.*, 2010). Additionally the DIT proteins might be involved in the C<sub>4</sub> cycle of NADP-ME C<sub>4</sub> species and facilitate the exchange of malate and/or aspartate across the chloroplast membrane (Gowik *et al.*, 2011; Kinoshita *et al.*, 2011), which may explain the highly tissue-preferential expression of these genes in *Sorghum*.

## Discussion

C<sub>4</sub> photosynthesis mainly evolved to enhance photosynthetic efficiency by avoiding photorespiration. It is widely accepted that an important initial step towards the evolution of C<sub>4</sub> was the establishment of a photorespiratory CO<sub>2</sub> pump (Bauwe, 2011; Sage *et al.*, 2012). This was achieved by restricting the activity of a central photorespiratory protein complex, the GDC, to the bundle sheath cells, allowing the release of photorespiratory CO<sub>2</sub> exclusively in this cell type (Hylton *et al.*, 1988; Rawsthorne *et al.*, 1988b). Finally that was realized by restricting the expression of either single GDC subunit genes or all GDC and SHM genes to the bundle sheath (Morgan *et al.*, 1993). Nevertheless, photorespiration is still essential in C<sub>4</sub> plants (Zelitch *et al.*, 2008) and we were interested in the tissue-specific expression of photorespiratory genes in the mesophyll and bundle sheath cells of a widely optimized C<sub>4</sub> species. Therefore we analyzed gene expression in leaves of *S. bicolor* using RNA-Seq on isolated mesophyll and bundle sheath transcripts and RNA *in situ* hybridization.

#### *Photorespiration is largely confined to the bundle sheath cells in Sorghum*

In C<sub>4</sub> plants, photorespiration is reduced to low levels compared with C<sub>3</sub> plants as a result of concentrating CO<sub>2</sub> around RubisCO (Hatch, 1987). Using RNA-Seq and SuperSage,

we were able to detect the transcripts of all core photorespiratory genes as well as of the genes encoding transporters known to be involved in photorespiration. The vast majority of the core photorespiratory genes are expressed preferentially in the bundle sheath. The only noticeable exceptions are GLYK, which is expressed preferentially in the mesophyll, and the two genes encoding the L subunit of the GDC complex (GDCL), which are nearly equally expressed in both cell types. This largely reflects earlier results from the analysis of mesophyll and bundle sheath transcriptomes and proteomes of the C<sub>4</sub> grass maize (Li *et al.*, 2010; Majeran *et al.*, 2010; Chang *et al.*, 2012) and studies on the enzyme activities in different C<sub>4</sub> species (Usuda and Edwards, 1980; Ohnishi and Kanai, 1983; Ohnishi *et al.*, 1985). Since in C<sub>4</sub> plants RubisCO is missing from the mesophyll cells, no 2-PG can be produced there and 2-PG detoxification in this cell type is no longer necessary. Consequently, the expression of photorespiratory genes was switched off in the mesophyll during C<sub>4</sub> evolution. The photorespiratory enzymes belong to the most highly abundant proteins in the leaves of C<sub>3</sub> species (Osborne and Freckleton, 2009; Bauwe, 2011). Accordingly, the decrease in these proteins adds to the reduction of RubisCO in C<sub>4</sub> plants and contributes to the better nitrogen use efficiency found for C<sub>4</sub> species (Oaks, 1994; Osborne and Freckleton, 2009).

GDCL is not only part of the GDC but is also connected to other multienzyme complexes such as the pyruvate dehydrogenase complex, the 2-oxoglutarate dehydrogenase complex, and the branched-chained 2-oxoacid dehydrogenase that are not involved in photorespiration and have important functions in general cell metabolism (Millar *et al.*, 1999; Marrott *et al.*, 2014). This explains why the genes encoding GDCL have to stay active in the mesophyll of C<sub>4</sub> plants. An explanation for the preferential expression of GLYK in the mesophyll is less obvious. In advanced C<sub>4</sub> species using the NADP-ME pathway, such as maize or *Sorghum*, the activity of photosystem II is greatly reduced in the bundle sheath (Woo *et al.*, 1970; Oswald *et al.*, 1990). This requires the reductive phase of the Calvin–Benson cycle to take place in the mesophyll cells, due to a lack of reducing equivalents in the bundle sheath, and is achieved by a phosphoglycerate–triose phosphate shuttle (Weber and von Caemmerer, 2010). It appears to be more efficient to transfer the photorespiratory glycerate directly to the mesophyll chloroplasts to regenerate 3-PG instead of importing it into the bundle sheath chloroplast for regeneration.

The genes involved in photorespiratory ammonia refixation, glutamine synthetase and glutamate synthase, show different expression patterns in mesophyll and bundle sheath cells. While two glutamine synthetase genes are expressed in both cell types with a bundle sheath preference, Fd-GOGAT is preferentially expressed in the mesophyll. This makes sense in the light of lacking reducing equivalents in the bundle sheath and one can assume that the released ammonia is fixed by glutamine synthetase and the resulting glutamine is partially transferred to the mesophyll to generate glutamate.

The plastidic glycolate transporter PLGG1 (Pick *et al.*, 2013) is expressed in both cell types. This might be due to the fact that glycolate has to be exported from bundle

sheath chloroplasts and glycerate must be imported into the chloroplasts in the mesophyll. It is known that the mitochondrial transporter BOU is essential for photorespiration in *A. thaliana* (Eisenhut *et al.*, 2013). Like PLGG, BOU is expressed in both cell types, but the overall transcript abundance is much lower. Since the specific substrate for the BOU transporter is not known (Eisenhut *et al.*, 2013), one can only speculate about possible functions beside photorespiration.

#### *Specificity of photorespiratory genes is as variable as that of C<sub>4</sub> genes*

With the transcriptome analysis, we confirmed that *S. bicolor* belongs to the NADP-ME type of C<sub>4</sub> plants since all participating C<sub>4</sub> genes (Wang *et al.*, 2009) are expressed in a tissue-preferential manner as expected for the NADP-ME archetype. Recent studies in maize revealed that not only the NADP-ME pathway is operating, but a respectable level of PEP-CK activity, up to 25% of the NADP-ME activity, was also found (Pick *et al.*, 2011). In the leaf transcriptome of *S. bicolor* we could find neither any highly expressed PEP-CK gene nor any significantly expressed NAD-ME gene in the bundle sheath. Taken together, these results indicate that *Sorghum* relies solely on the NADP-ME pathway.

As expected, we found PPDK to be one of the most highly expressed genes in the *Sorghum* leaf. Surprisingly, the transcriptome analysis indicated that PPDK transcripts are not restricted to the mesophyll but are also found in high amounts in the bundle sheath, with a mesophyll to bundle sheath ratio of only ~1 to 2 (Fig. 3). We confirmed that the analysis detects only the gene encoding the chloroplast-targeted PPDK isoform and indeed the gene encoding the cytosolic isoform showed quite low expression in *Sorghum* leaves. Also the RNA *in situ* analysis indicates high amounts of PPDK transcripts in the bundle sheath cells (Fig. 3B). Since this analysis is strictly independent of the transcriptome analysis, it must be considered that *Sorghum* contains considerable amounts of PPDK in its bundle sheath cells. This is in contrast to the analysis of mesophyll and bundle sheath cells of maize or *S. viridis* where PPDK transcripts were found to be five and 20 times more abundant in the mesophyll than in the bundle sheath, respectively (Chang *et al.*, 2012; John *et al.*, 2014). Very similar patterns were also found for the transcripts of genes that functionally interact with PPDK such as the PPDK regulatory proteins, plastid-localized pyrophosphatases, an AMP kinase, and the plastid PEP translocator PPT (Fig. 3; Supplementary Table S3). For all these genes, we found considerable amounts of transcripts in the bundle sheath preparations that were often even higher than in the mesophyll. The most parsimonious explanation is that *Sorghum* is capable of regenerating substantial amounts of PEP in the bundle sheath cells. The existence of plants using extensively the PEP-CK type of the C<sub>4</sub> pathway shows that PEP can serve as a transport metabolite in the C<sub>4</sub> cycle. Due to up-regulation of photosystem I and cyclic electron transport in the bundle sheath chloroplasts (Supplementary Table S1; Kubicki *et al.*, 1994, 1996), *Sorghum* potentially produces high amounts of ATP in this compartment that can be used

for PEP regeneration. By regenerating PEP in the bundle sheath chloroplasts, the number of transport processes would be reduced since PEP can be exported by PPT and diffuse into the mesophyll where it could be carboxylated by PEPC in the cytosol.

All in all, it appears that the degree of cell specificity is quite comparable for photorespiratory and C<sub>4</sub> cycle genes. While most of the genes encoding core pathway enzymes are expressed in a highly cell type-specific manner, exceptions are the PPDK in the case of the C<sub>4</sub> cycle and GDCL in the case of photorespiration. This is notable since tissue specificity for C<sub>4</sub> enzymes such as PEPC or NADP-ME is necessary to avoid futile cycles and ensure the efficiency of the pathway, whereas tissue-specific expression of most photorespiratory genes has to be seen as optimization that saves nitrogen. The expression of auxiliary genes of both pathways was found to be not very tissue specific. This might be due to additional roles of the encoded protein in other important pathways as can be envisaged for the genes involved in primary nitrogen and amino acid metabolism.

#### *Evolutionary aspects of restricting photorespiration to the bundle sheath*

As discussed above, photorespiration was important for the evolution of C<sub>4</sub> photosynthesis in different ways. The avoidance of photorespiration was one of the driving forces towards C<sub>4</sub> photosynthesis, and the establishment of a photorespiratory pump was an important intermediate step during C<sub>4</sub> evolution (Bauwe, 2011; Sage et al., 2012). The reduction and exclusion of the majority of photorespiratory reactions from the mesophyll represents an optimization and enhances the nitrogen use efficiency. This optimization could only happen after the implementation of a fully functional C<sub>4</sub> pathway and the complete down-regulation of RubisCO in the mesophyll since the oxygenase reaction of RubisCO would be fatal without PGLP and GOX activity present in the same compartment. This has a further implication for C<sub>4</sub> evolution: once PGLP and GOX are switched off in the mesophyll, the reintroduction of RubisCO into this compartment would be detrimental. Once these photorespiratory reactions are gone from the mesophyll due to optimization, a reversal from C<sub>4</sub> to C<sub>3</sub> photosynthesis becomes impossible.

#### Supplementary data

Supplementary data are available at *JXB* online.

**Table S1.** Excel worksheet providing quantitative information for all reads and all SuperSage tags mapped onto the reference transcriptome from *Sorghum bicolor*.

**Table S2.** Transcript abundance of genes related to photorespiration

**Table S3.** Transcript abundance of C<sub>4</sub> cycle genes and C<sub>4</sub>-related transporters.

**Table S4.** Gene-specific primers used for qPCR and RNA *in situ* analysis.

**Figure S1,** RNA *in situ* hybridization of *Sorghum bicolor* leaves with probes for transcripts related to photorespiration

#### Acknowledgements

This work was supported by the Bill and Melinda Gates Foundation through the C4 Rice Project, the Deutsche Forschungsgemeinschaft through the Research Group FOR1186, and the Excellence Cluster EXC 1028 (From Complex Traits towards Synthetic Modules).

#### References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403–410.
- Anderson LE.** 1971. Chloroplast and cytoplasmic enzymes. II. Pea leaf triose phosphate isomerases. *Biochimica et Biophysica Acta* **235**, 237–244.
- Aoki N, Ohnishi J, Kanai R.** 1992. Two different mechanisms for transport of pyruvate into mesophyll chloroplasts of C<sub>4</sub> plants—a comparative study. *Plant and Cell Physiology* **33**, 805–809.
- Bauwe H, Hagemann M, Fernie AR.** 2010. Photorespiration: players, partners and origin. *Trends in Plant Science* **15**, 330–336.
- Bauwe H.** 2011. Photorespiration: the bridge to C<sub>4</sub> photosynthesis. In: Raghavendra AS, Sage RF, eds. C<sub>4</sub> photosynthesis and related CO<sub>2</sub> concentrating mechanisms . Dordrecht: Springer, 81–108.
- Botha CE.** 1992. Plasmodesmatal distribution, structure and frequency in relation to assimilation in C<sub>3</sub> and C<sub>4</sub> grasses in southern Africa. *Planta* **187**, 348–358.
- Bräutigam A, Gowik U.** 2010. What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. *Plant Biology* **12**, 831–841.
- Bräutigam A, Kajala K, Wullenweber J, et al.** 2011. An mRNA blueprint for C<sub>4</sub> photosynthesis derived from comparative transcriptomics of closely related C<sub>3</sub> and C<sub>4</sub> species. *Plant Physiology* **155**, 142–156.
- Bräutigam A, Schliesky S, Külahoglu C, Osborne CP, Weber APM.** 2014. Towards an integrative model of C<sub>4</sub> photosynthetic subtypes: insights from comparative transcriptome analysis of NAD-ME, NADP-ME, and PEP-CK C<sub>4</sub> species. *Journal of Experimental Botany* **65**, 3579–3593.
- Chang YM, Liu WY, Shih AC, et al.** 2012. Characterizing regulatory and functional differentiation between maize mesophyll and bundle sheath cells by transcriptomic analysis. *Plant Physiology* **160**, 165–177.
- Dengler NG, Nelson T.** 1999. Leaf structure and development in C<sub>4</sub> plants. In: Rowan FS, Russell KM, eds. C<sub>4</sub> plant biology . San Diego: Academic Press, 133–172.
- Ding Z, Weissmann S, Wang M, et al.** 2015. Identification of photosynthesis-associated C<sub>4</sub> candidate genes through comparative leaf gradient transcriptome in multiple lineages of C<sub>3</sub> and C<sub>4</sub> species. *PLoS One* **10**, e0140629.
- Eisenhut M, Planchais S, Cabassa C, et al.** 2013. Arabidopsis A BOUT DE SOUFFLE is a putative mitochondrial transporter involved in photorespiratory metabolism and is required for meristem growth at ambient CO<sub>2</sub> levels. *The Plant Journal* **73**, 836–849.
- Fernie AR, Bauwe H, Eisenhut M, et al.** 2013. Perspectives on plant photorespiratory metabolism. *Plant Biology* **15**, 748–753.
- Furbank RT.** 2011. Evolution of the C<sub>4</sub> photosynthetic mechanism: are there really three C<sub>4</sub> acid decarboxylation types? *Journal of Experimental Botany* **62**, 3103–3108.
- Furumoto T, Yamaguchi T, Ohshima-Ichie Y, et al.** 2011. A plastidial sodium-dependent pyruvate transporter. *Nature* **476**, 472–475.
- Gardeström P, Edwards GE, Henricson D, Ericson I.** 1985. The localization of serine hydroxymethyltransferase in leaves of C<sub>3</sub> and C<sub>4</sub> species. *Physiologia Plantarum* **64**, 29–33.
- Gowik U, Bräutigam A, Weber KL, Weber APM, Westhoff P.** 2011. Evolution of C<sub>4</sub> photosynthesis in the genus Flaveria: how many and which genes does it take to make C<sub>4</sub>? *The Plant Cell* **23**, 2087–2105.
- Haberlandt G.** 1904. Physiologische Pflanzenanatomie . Leipzig: Verlag von Wilhelm Engelmann.

- Hatch MD.** 1987. C4 photosynthesis: a unique blend of modified biochemistry, anatomy and ultrastructure. *Biochimica et Biophysica Acta* **895**, 81–106.
- Heckmann D, Schulze S, Denton A, Gowik U, Westhoff P, Weber AP, Lercher MJ.** 2013. Predicting C4 photosynthesis evolution: modular, individually adaptive steps on a Mount Fuji fitness landscape. *Cell* **153**, 1579–1588.
- Hilton C, Rawsthorne S, Smith A, Jones DA, Woolhouse H.** 1988. Glycine decarboxylase is confined to the bundle-sheath cells of leaves of C3–C4 intermediate species. *Planta* **175**, 452–459.
- John CR, Smith-Unna RD, Woodfield H, Covshoff S, Hibberd JM.** 2014. Evolutionary convergence of cell-specific gene expression in independent lineages of C4 grasses. *Plant Physiology* **165**, 62–75.
- Keerberg O, Parnik T, Ivanova H, Bassuner B, Bauwe H.** 2014. C2 photosynthesis generates about 3-fold elevated leaf CO<sub>2</sub> levels in the C3–C4 intermediate species *Flaveria pubescens*. *Journal of Experimental Botany* **65**, 3649–3656.
- Kinoshita H, Nagasaki J, Yoshikawa N, Yamamoto A, Takito S, Kawasaki M, Sugiyama T, Miyake H, Weber AP, Taniguchi M.** 2011. The chloroplastic 2-oxoglutarate/malate transporter has dual function as the malate valve and in carbon/nitrogen metabolism. *The Plant Journal* **65**, 15–26.
- Kubicki A, Steinmuller K, Westhoff P.** 1994. Differential transcription of plastome-encoded genes in the mesophyll and bundle-sheath chloroplasts of the monocotyledonous NADP-malic enzyme-type C4 plants maize and Sorghum. *Plant Molecular Biology* **25**, 669–679.
- Kubicki A, Funk E, Westhoff P, Steinmüller K.** 1996. Differential expression of plastome-encoded ndh genes in mesophyll and bundle-sheath chloroplasts of the C4 plant *Sorghum bicolor* indicates that the complex I-homologous NAD(P)H-plastoquinone oxidoreductase is involved in cyclic electron transport. *Planta* **199**, 276–281.
- Kulahoglu C, Denton AK, Sommer M, et al.** 2014. Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. *The Plant Cell* **26**, 3243–3260.
- Laetsch WM.** 1974. C4 syndrome—structural-analysis. Annual Review of Plant Physiology and Plant Molecular Biology **25**, 27–52.
- Langmead B, Trapnell C, Pop M, Salzberg SL.** 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**, R25.
- Li P, Ponnala L, Gandotra N, et al.** 2010. The developmental dynamics of the maize leaf transcriptome. *Nature Genetics* **42**, 1060–1067.
- Long SP.** 1999. Environmental responses. In: Sage RF, Monson RK, eds. C4 plant biology. San Diego: Academic Press, 215–249.
- Majeran W, Friso G, Ponnala L, et al.** 2010. Structural and metabolic transitions of C4 leaf development and differentiation defined by microscopy and quantitative proteomics in maize. *The Plant Cell* **22**, 3509–3542.
- Mallmann J, Heckmann D, Bräutigam A, Lercher MJ, Weber AP, Westhoff P, Gowik U.** 2014. The role of photorespiration during the evolution of C4 photosynthesis in the genus *Flaveria*. *Elife* **3**, e02478.
- Marrott NL, Marshall JJ, Svergun DI, Crennell SJ, Hough DW, van den Elsen JM, Danson MJ.** 2014. Why are the 2-oxoacid dehydrogenase complexes so large? Generation of an active trimeric complex. *Biochemical Journal* **463**, 405–412.
- Matsumura H, Reich S, Ito A, Saitoh H, Kamoun S, Winter P, Kahl G, Reuter M, Kruger DH, Terauchi R.** 2003. Gene expression analysis of plant host-pathogen interactions by SuperSAGE. *Proceedings of the National Academy of Sciences, USA* **100**, 15718–15723.
- Meister M, Agostino A, Hatch MD.** 1996. The roles of malate and aspartate in C4 photosynthetic metabolism of *Flaveria bidentis* (L.). *Planta* **199**, 262–269.
- Millar A, Hill S, Leaver C.** 1999. Plant mitochondrial 2-oxoglutarate dehydrogenase complex: purification and characterization in potato. *Biochemical Journal* **343**, 327–334.
- Morgan CL, Turner SR, Rawsthorne S.** 1993. Coordination of the cell-specific distribution of the four subunits of glycine decarboxylase and of serine hydroxymethyltransferase in leaves of C3–C4 intermediate species from different genera. *Planta* **190**, 468–473.
- Oaks A.** 1994. Efficiency of nitrogen utilization in C3 and C4 cereals. *Plant Physiology* **106**, 407–414.
- Ogren WL.** 1984. Photorespiration: pathways, regulation, and modification. *Annual Review of Plant Physiology* **35**, 415–442.
- Ohnishi J-i, Kanai R.** 1983. Differentiation of photorespiratory activity between mesophyll and bundle sheath cells of C4 plants I. Glycine oxidation by mitochondria. *Plant and Cell Physiology* **24**, 1411–1420.
- Ohnishi J-i, Yamazaki M, Kanai R.** 1985. Differentiation of photorespiratory activity between mesophyll and bundle sheath cells of C4 plants II. Peroxisomes of *Panicum miliaceum* L. *Plant and Cell Physiology* **26**, 797–803.
- Osborne CP, Freckleton RP.** 2009. Ecological selection pressures for C4 photosynthesis in the grasses. *Proceedings of the Royal Society B: Biological Sciences* **276**, 1753–1760.
- Oswald A, Streubel M, Ljungberg U, Hermans J, Eskins K, Westhoff P.** 1990. Differential biogenesis of photosystem-II in mesophyll and bundle-sheath cells of ‘malic’ enzyme NADP(+)-type C4 plants. A comparative protein and RNA analysis. *European Journal of Biochemistry* **190**, 185–194.
- Paterson AH, Bowers JE, Bruggmann R, et al.** 2009. The *Sorghum bicolor* genome and the diversification of grasses. *Nature* **457**, 551–556.
- Pick TR, Bräutigam A, Schlüter U, et al.** 2011. Systems analysis of a maize leaf developmental gradient redefines the current C4 model and provides candidates for regulation. *The Plant Cell* **23**, 4208–4220.
- Pick TR, Bräutigam A, Schulz MA, Obata T, Fernie AR, Weber AP.** 2013. PLGG1, a plastidic glycolate transporter, is required for photorespiration and defines a unique class of metabolite transporters. *Proceedings of the National Academy of Sciences, USA* **110**, 3185–3190.
- Raines CA.** 2011. Increasing photosynthetic carbon assimilation in C3 plants to improve crop yield: current and future strategies. *Plant Physiology* **155**, 36–42.
- Rawsthorne S, Hilton CM, Smith AM, Woolhouse HW.** 1988a. Photorespiratory metabolism and immunogold localization of photorespiratory enzymes in leaves of C3 and C3–C4 intermediate species of Moricandia. *Planta* **173**, 298–308.
- Rawsthorne S, Hilton CM, Smith AM, Woolhouse HW.** 1988b. Distribution of photorespiratory enzymes between bundle-sheath and mesophyll cells in leaves of the C3–C4 intermediate species Moricandia arvensis (L.) DC. *Planta* **176**, 527–532.
- Renne P, Dressen U, Hebbeker U, Hille D, Flugge UI, Westhoff P, Weber AP.** 2003. The *Arabidopsis* mutant *dct* is deficient in the plastidic glutamate/malate translocator *DIT2*. *The Plant Journal* **35**, 316–331.
- Sage RF.** 2004. The evolution of C4 photosynthesis. *New Phytologist* **161**, 341–370.
- Sage RF, Christin P-A, Edwards EJ.** 2011. The C4 plant lineages of planet Earth. *Journal of Experimental Botany* **62**, 3155–3169.
- Sage RF, Sage TL, Kocacinar F.** 2012. Photorespiration and the evolution of C4 photosynthesis. *Annual Review of Plant Biology* **63**, 19–47.
- Sawers RJ, Liu P, Anufrieva K, Hwang JT, Brutnell TP.** 2007. A multi-treatment experimental system to examine photosynthetic differentiation in the maize leaf. *BMC Genomics* **8**, 12.
- Simon R.** 2002. Molecular and biochemical analysis of *Arabidopsis*. DIG application manual for nonradioactive *in situ* hybridization. Mannheim, Germany: Roche Applied Science, 197–207.
- Tausta SL, Li P, Si Y, Gandotra N, Liu P, Sun Q, Brutnell TP, Nelson T.** 2014. Developmental dynamics of Kranz cell transcriptional specificity in maize leaf reveals early onset of C4-related processes. *Journal of Experimental Botany* **65**, 3543–3555.
- Usuda H, Edwards GE.** 1980. Localization of glycerate kinase and some enzymes for sucrose synthesis in C3 and C4 plants. *Plant Physiology* **65**, 1017–1022.
- Wang L, Feng Z, Wang X, Wang X, Zhang X.** 2010. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **26**, 136–138.
- Wang L, Czedik-Eysenberg A, Mertz RA, et al.** 2014. Comparative analyses of C4 and C3 photosynthesis in developing leaves of maize and rice. *Nature Biotechnology* **32**, 1158–1165.
- Wang X, Gowik U, Tang H, Bowers JE, Westhoff P, Paterson AH.** 2009. Comparative genomic analysis of C4 photosynthetic pathway evolution in grasses. *Genome Biology* **10**, R68.

- Wang Y, Bräutigam A, Weber AP, Zhu XG.** 2014. Three distinct biochemical subtypes of C4 photosynthesis? A modelling analysis. *Journal of Experimental Botany* **65**, 3567–3578.
- Weber AP, von Caemmerer S.** 2010. Plastid transport and metabolism of C3 and C4 plants—comparative analysis and possible biotechnological exploitation. *Current Opinion Plant Biology* **13**, 257–265.
- Westhoff P, Offermann-Steinhard K, Höfer M, Eskins K, Oswald A, Streubel M.** 1991. Differential accumulation of plastid transcripts encoding photosystem II components in the mesophyll and bundle-sheath cells of monocotyledonous NADP-malic enzyme-type C4 plants. *Planta* **184**, 377–388.
- Williams BP, Johnston IG, Covshoff S, Hibberd JM.** 2013. Phenotypic landscape inference reveals multiple evolutionary paths to C4 photosynthesis. *Elife* **2**, e00961.
- Wingler A, Walker RP, Chen ZH, Leegood RC.** 1999. Phosphoenolpyruvate carboxykinase is involved in the decarboxylation of aspartate in the bundle sheath of maize. *Plant Physiology* **120**, 539–546.
- Woo KC, Anderson JM, Boardman NK, Downton WJ, Osmond CB, Thorne SW.** 1970. Deficient photosystem II in agranal bundle sheath chloroplasts of C4 plants. *Proceedings of the National Academy of Sciences, USA* **67**, 18–25.
- Wyrich R, Dressen U, Brockmann S, Streubel M, Chang C, Qiang D, Paterson AH, Westhoff P.** 1998. The molecular basis of C4 photosynthesis in sorghum: isolation, characterization and RFLP mapping of mesophyll- and bundle-sheath-specific cDNAs obtained by differential screening. *Plant Molecular Biology* **37**, 319–335.
- Zelitch I, Schultes NP, Peterson RB, Brown P, Brutnell TP.** 2008. High glycolate oxidase activity is required for survival of maize in normal air. *Plant Physiology* **149**, 195–204.