A systems biology approach and single gene analysis to identify transporters and regulatory proteins in the plant carbon cycle

Kumulative Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Thea Renate Pick

aus Aachen

Düsseldorf, September 2012

Aus dem

Institut Biochemie der Pflanzen

der Heinrich-Heine-Universität Düsseldorf

Eidesstattliche Versicherung und Selbstständigkeitserklärung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf" erstellt worden ist. Diese Dissertation habe ich in dieser oder ähnlicher Form noch bei keiner anderen Fakultät vorgelegt. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 10.10.2012

_____

Thea Pick

Für meine Familie...

Summary

Green plants fix inorganic carbon dioxide (CO_2) by Ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco) in a process called photosynthesis. This carboxylation reaction of Rubisco generates two molecules of 3-phosphoglycerate (3-PGA) and finally leads to biomass production. Hence plants account for a significant share of the global organic carbon pool. While most plants use the original C_3 type of photosynthesis for CO_2 assimilation some of the most important crop plants on earth like maize (*Zea mays*) use a highly efficient type of photosynthesis, namely C_4 photosynthesis.

In C_4 photosynthesis CO_2 is prefixed by an oxygen (O_2) insensitive enzyme into a C_4 acid and concentrated in close vicinity of Rubisco. This leads to a high efficiency of photosynthesis through a reduced rate of photorespiration that is caused by the oxygenation reaction of Rubisco. Although Rubisco strongly favors CO_2, the enzyme can also accept O_2 as a substrate. The product of the oxygenation reaction is one molecule of 3-PGA and one molecule of 2-phosphoglycolate (2-PG). The latter can neither enter the Calvin cycle nor be converted to carbohydrates. This can lead to a dramatic decrease in biomass production since 2-PG has to be recycled at the cost of energy and reduction equivalents in the photorespiratory cycle.

The photorespiratory cycle is a highly compartmentalized pathway involving metabolic reactions in the chloroplasts, peroxisomes, mitochondria, and the cytosol. The mass flow through this pathway is only excelled by photosynthesis and therefore requires a tight and fast connection between the different compartments. This is obtained through highly specific metabolite transporters that account for the shuttle of precursors, intermediates and products across the membranes. It has previously been shown that short-circuiting of the photorespiratory cycle led to enhanced biomass production in C_3 plants. This important finding verifies the significance of research for molecular plant engineering as a significant share of the components of the pathway, such as metabolite transporters and regulatory proteins are still unknown.

Therefore one aim of this thesis was the identification of photorespiratory metabolite transporters since all genes encoding enzymes required for appropriate function of the core cycle are known to date while no transporter has been characterized on the molecular level. In a reverse genetics screen Plastidic glycolate glycerate transporter 1 (PLGG1) was designated as a promising candidate. Molecular characterization of the candidate was used to identify expression pattern and subcellular localization. An *A. thaliana* (*Arabidopsis thaliana*) T-DNA insertional knockout mutant (*plgg1-1*) was isolated and characterized concerning its phenotype and steady state and dynamic metabolite accumulation. Biochemical characterization of the candidate was used to identify transport capacity and to define the role of PLGG1 in plant metabolism. The molecular characterization revealed that PLGG1 is expressed in all green

tissues and that it is located at the chloroplast envelope. Analysis of the knockout mutant revealed that *plgg1-1* plants show a photorespiratory phenotype and the photorespiratory metabolites glycolate and glycerate accumulate in the mutant. Finally the biochemical characterization indicated that glycolate and glycerate transport are impaired in the mutant. These analyses lead to the discovery of PLGG1 as the chloroplastidic glycolate glycerate transporter and thereby to the identification of the first transporter of the core photorespiratory pathway.

Surprisingly PLGG1 was found to be high abundant in maize chloroplast what is somehow counter intuitive as PLGG1 is a photorespiratory transporter and the rate of photorespiration is reduced in the C_4 plant maize. Until today the role of PLGG1 in C_4 photosynthesis is still unresolved and emphasizes the question how the C_4 cycle is organized and how the establishment of C_4 photosynthesis is orchestrated since the C_4 cycle is obviously not as linear and simple at is has been assumed.

Therefore the second aim of this thesis was to set up a systems biology approach to find candidates for unidentified components of the C_4 pathway, such as metabolite transporters and regulatory proteins. The third leaf of the C_4 model plant maize was analyzed along a base-to-tip developmental gradient. Ten continuous leaf slices were analyzed individually and transcriptome analysis, oxygen sensitivity of photosynthesis, photosynthetic rate measurements, and chlorophyll and protein measurements revealed a gradual sink-to-source transition for the leaf without a binary on-off switch. Analysis of transcription factors exhibiting a similar expression pattern to key C_4 enzymes along the leaf gradient designated in a list of putative regulatory proteins orchestrating the establishment of C_4 photosynthesis. Finally, transcriptome and metabolome analysis, enzyme activity measurements, and quantification of selected metabolites showed that the C_4 cycle is not linear but rather a branched cycle. These results led to a revised model of maize C_4 photosynthesis.

Zusammenfassung

Grüne Pflanzen fixieren anorganisches Kohlendioxid (CO_2) mit Hilfe des Enzyms Ribulose 1,5-bisphosphate Carboxylase/Oxygenase (Rubisco) in dem Prozess, der als Photosynthese bezeichnet wird. In der Carboxylierungsreaktion werden zwei Moleküle 3-Phosphoglycerat (3-PGA) gebildet, die letztendlich zur Produktion von Biomasse verwendet werden. Daher tragen Pflanzen zu einem erheblichen Teil des weltweiten organischen Kohlenstoff-Pools bei. Während ein Großteil der Pflanzen dabei den ursprünglichen C_3-Typ der Photosynthese betreiben, gibt es einige Pflanzen, unter denen sich wichtige Nutzpflanzen wie Mais (*Zea mays*) befinden, die einen hoch effizienten Photosynthese-Typ betreiben, nämlich C_4 Photosynthese.

In der C_4 Photosynthese wird CO_2 durch ein Sauerstoff (O_2) unempfindliches Enzym in eine C_4 Säure vorfixiert und in unmittelbarer Nähe zur Rubisco angereichert. Diese Vorfixierung führt zu einer hohen Effizienz der Photosynthese, da die Photorespiration, hervorgerufen durch die Oxygenierungsreaktion der Rubisco, gesenkt wird. Denn obwohl die Rubisco verstärkt CO_2 bevorzugt, kann das Enzym auch O_2 als Substrat verwenden. Die Produkte dieser Oxygenierungsreaktion sind je ein Molekül 3-PGA und ein Molekül 2-Phosphoglycolat (2-PG). Letzteres kann weder in den Calvin-Zyklus eintreten, noch in Kohlenhydrate umgewandelt werden. 2-PG muss daher im photorespiratorischen Zyklus unter Verbrauch von Energie und Reduktionsäquivalenten regeneriert werden, wodurch die Produktion von Biomasse drastisch reduziert werden kann.

Der photorespiratorische Zyklus ist stark kompartimentiert und beinhaltet metabolische Reaktionen in den Chloroplasten, den Peroxisomen, den Mitochondrien und dem Zytosol. Die Durchflussmenge von Metaboliten durch diesen Stoffwechselweg wird nur von der Photosynthese übertroffen und erfordert daher eine dichte und schnelle Verknüpfung der Kompartimente untereinander. Diese Verknüpfung wird durch hochspezifische Metabolit Transporter erzielt, die Vor-, Zwischen und Endprodukte über Membranen transportieren. Es konnte bereits gezeigt werde, dass ein Verkürzen und Kurzschließen der Photosynthese zu einer erhöhten Produktion von Biomasse in C_3 Pflanzen führt. Diese wichtige Erkenntnis bekräftigt die Wichtigkeit der Forschung an molekularer Pflanzentechnik, da weiterhin wichtige Komponenten des Stoffwechselweges, wie etwa Metabolit Transporter und regulatorische Proteine, unbekannt sind.

Daher war ein Ziel dieser Arbeit photorespiratorische Metabolit Transporter zu identifizieren, da alle Gene, die für unabdingbare Enzyme des Hauptzyklus kodieren, bekannt sind, während bisher kein Transporter molekular charakterisiert werden konnte. Mit Hilfe reverser Genetik konnte Plastidic glycolate glycerate transporter 1 (PLGG1) als ein vielversprechender Kandidat isoliert werden. Die molekulare Charakterisierung des Kandidaten wurde zur Untersuchung des Expressionsmusters und der zellulären Lokalisation verwendet.

Eine *A. thaliana* (*Arabidopsis thaliana*) knockout Linie, die eine T-DNA Insertion im Gen trägt (*plgg1-1*), konnte isoliert werden und wurde auf den Phänotypen und stationäre und dynamische Metabolit Anreicherung untersucht. Die biochemische Charakterisierung wurde durchgeführt, um die Transporteigenschaften zu bestimmen und die Rolle von PLGG1 im pflanzlichen Stoffwechsel zu klären. Die molekulare Charakterisierung ergab, dass PLGG1 in allen grünen Geweben exprimiert wird und an der Chloroplasten Hülle lokalisiert ist. Die Analyse der knockout Linie zeigte, dass die *plgg1-1* Pflanzen einen photorespiratorischen Phänotyp aufweisen und die photorespiratorischen Metabolite Glycolat und Glycerat angereichert werden. Schließlich zeigte die biochemische Charakterisierung, dass der Transport von Glycolat und Glycerat in *plgg1-1* Pflanzen beeinträchtigt ist. Diese Untersuchungen führten zu der Entdeckung von PLGG1 als den chloroplastiäre Glycolat Glycerat Transporter und damit zur Identifizierung des ersten Transporters des photorespiratorischen Hauptkreislaufs.

Überraschender Weise wurde zudem festgestellt, dass PLGG1 sehr stark in Mais Chloroplasten vertreten ist, was eigentlich widersprüchlich ist, da PLGG1 ein photorespiratorischer Transporter ist und die Photoresirationsrate in der C$_4$ Pflanze Mais stark reduziert ist. Bisher ist die Rolle von PLGG1 in der C$_4$ Photosynthese noch ungeklärt und verstärkt die Frage wie der C$_4$ Zyklus organisiert ist und wie die Ausbildung des C$_4$ Syndroms gesteuert wird, da der C$_4$ Zyklus offenbar nicht so linear und schlicht ist, wie es bisher angenommen wurde.

Daher war das zweite Ziel dieser Arbeit Kandidaten für unbekannte Komponenten des C$_4$ Kreislaufs, wie etwa Metabolit Transporter und regulatorische Proteine, zu finden, indem ein Systembiologischer Ansatz durchgeführt wurde. Das dritte Blatt der C$_4$ Modellpflanze Mais wurde entlang eines Entwicklunsgradienten von der Basis bis zur Spitze untersucht. Dafür wurden zehn fortlaufende Blattabschnitte individuell analysiert und das Transkriptom, Sauerstoff Empfindlichkeit der Photosynthese, Photosyntheserate, Chlorophyll- und Proteingehalt gemessen. Die Ergebnisse zeigten, dass es im Blatt einen schrittweise Übergang von Sink zu Source Geweben gibt, ohne einen binären an-aus Schalter. Eine Liste möglicher regulatorischer Proteine, die die Ausbildung des C$_4$ Syndroms steuern, konnte erstellt werden indem Transkriptionsfaktoren, die ein ähnliches Expressionsmuster wie wichtige C$_4$ Enzyme aufweisen, analysiert wurden. Die Analyse des Transkriptoms und Metaboloms, sowie die Bestimmung der Enzymaktivität und die Quantifizierung ausgewählter Metabolite ergab, dass der C$_4$ Kreislauf nicht linear, sondern eher verzweigt abläuft. Diese Ergebnisse führten zu einer Änderung des bisherigen Modells der C$_4$ Photosynthese in Mais.

Table of contents

1. Abbreviations

2-OG	2-oxoglutarate
2-PG	2-phosphoglycolate
3-PGA	3-phosphoglycerate
A. thaliana	*Arabidopsis thaliana*
ATP	adenine triphosphate
C_3, C_4	three-, four-carbon molecule
CAT	catalase
CO_2	carbon dioxide
DiT1	2-oxoglutarate/malate translocator
DiT2	glutamate/malate translocator
G2	Golden 2 transcription factor
GDC	glycine decarboxylase
GGT	glutamate:glyoxylate aminotransferase
Glk	G2-like
Glu	glutamate
GLYK	glycerate kinase
GO	glycolateoxidase
GOX	flavin-dependent glycolate oxidase
GS/GOGAT	glutamine synthetase/glutamate synthase
H_2O	water
H_2O_2	hydrogen peroxide
HPR	hydoxypyruvate reductase
mMDH	mitochondrial malate dehydrogenase
MS	malatesynthase
NAD	nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide

NH_4^+	ammonium ion
O_2	oxygen
P_i	orthophosphoric acid
PEPC	Phosphoenolpyruvate carboxylase
PGLP	2-phosphoglycolate phosphatase
PGP	phosphoglycolate phosphatase
PLGG1	Plastidic glycolate glycerate transporter 1
pMDH	peroxisomal malate dehydrogenase
Rubisco	ribulose 1,5-bisphosphate carboxylase/oxygenase
RuBP	ribulose-bisphosphate
SHMT	serine hydroxymethyl transferase
SGT	serine:glyoxylate aminotransferase
THF	tetrahydofolate
Z. mays	*Zea mays*

2. Introduction

Research on photorespiration and C_4 photosynthesis has been of major interest for molecular plant breeding in the last decades. To date especially regulatory proteins and molecular transporters are of special interest as a significant share of components involved in those pathways is still unknown. This thesis is concerned with the construction of a data set to generate candidates for regulatory proteins and molecular transporters of the C_4 photorespiratory cycle as well as single candidate analysis for a molecular transporter of the photorespiratory cycle.

2.1   Motivation

Plants fix inorganic carbon in a process called photosynthesis. Hence plants are responsible for a considerable share of the world's carbon pool and energy supply. However, in future we are facing severe problems concerning the conflict of growth of population and a boosted demand for biofuels and crops (Cassman and Liska, 2007). In this context approaches to generate (a) highly productive plants and (b) plants that are able to grow in harsh habitants and cope with critical environmental factors such as heat, salinity and drought are of major interest and current focus of research (Kajala et al., 2011; von Caemmerer et al., 2012). Approximately 15 to 40 million years ago, long after C_3 photosynthesis was established, a highly efficient type of photosynthesis, namely C_4 photosynthesis, evolved (Sage, 2004). Plants performing C_4 photosynthesis are able to grow in harsher habitants than C_3 plants. Therefore C_4 plants like maize are ideal models to study the adaptation to harsh environments and to use the knowledge to boost photosynthesis in C_3 plants like rice (*Oryza sativa*) (Kajala et al., 2011). A further approach to successfully enhance biomass production is short-circuiting of photorespiration (Kebeish et al., 2007; Maier et al., 2012). Hence identification of regulatory proteins and factors involved in both C_4 photosynthesis and photorespiration are of major interest.

2.2   Photorespiration

Photosynthesis is the light dependent process in which inorganic carbon is incorporated into carbohydrates and is thereby made accessible to heterotrophic organisms. Plants performing oxygenetic photosynthesis are therefore indispensable for the land-based biosphere. Fixation of inorganic carbon is catalyzed by Rubisco that feeds CO_2 into the Calvin cycle finally leading to the production of the C_3 acid 3-PGA. Although Rubisco is the most abundant protein on earth, it is inefficient due to its catalytic activity to both carboxylate and oxygenate the acceptor ribulose-

1,5-bisphosphate. The photosynthetic process evolved 2 billion years ago when the atmosphere was free of O_2 (Kirschvink and Kopp, 2008). Hence, the oxygenation reaction played a minor role during land plant evolution while it became significant with increasing atmospheric oxygen concentration 1 billion years ago. The product of this oxygenation reaction in one molecule 3-PGA and one molecule 2-PG. 2-PG is toxic to plants and can neither enter the Calvin cycle nor be converted to carbohydrates and inhibits enzymes of the Calvin cycle (Husic and Tolbert, 1987). Hence, 2-PG has to be detoxified in the photorespiratory cycle. The first enzymatic steps take place in the chloroplast stroma where 2-PG is dephosphorylated to glycolate by phosphoglycolate phosphatase (PGP). The glycolate is transported out of the chloroplast in counter-exchange with glycerate. Glycolate enters the peroxisomes and is oxidized to glyoxylate by flavin-dependent glycolate oxidase (GOX). Glyoxylate is transaminated to glycine by the two aminotransferases Ser:glyoxylate and Glu:glyoxylate aminotransferase (SGT and GGT, respectively). Glycine is further transported into the mitochondria where it is deaminated and decarboxylated by a multienzyme complex. This complex is called the glycine decarboxylase complex (GDC). One product of this reaction, the 5,10-methylene tetrahydrofolate (Collakova et al., 2008), is further converted by serine hydroxymethyltransferase (SHMT) to yield one molecule of serine from two molecules of Glycine (Somerville and Ogren, 1981, 1982). Serine is transported back into the peroxisomes and converted to glycerate by serine glyoxylate aminotransferase (SGT) and hydroxypyruvate reductase (HPR). The glycerate is transported back into the chloroplasts and finally phosphorylated by glycolate kinase (GLYK) to yield 3-PGA that is fed back to the Calvin cycle. The complete cycle is depicted in Figure 1.

Because previously fixed carbon is lost and energy consumed photorespiration has been considered as a wasteful process. Besides those disadvantageous effects photorespiration can also be beneficial for plants under certain stress conditions (Migge et al., 1999). For example under high light stress when the production of photochemical energy exceeds the consumption photorespiration can work as a valve and protect plants from photoinhibition (Kozaki and Takeba, 1996). In addition photorespiration can serve as a source for metabolites like glycine and serine that are used by the plant for other processes (e.g. (Madore and Grodzinski, 1984). For instance glycine can be further converted to glutathione that is used in the plant antioxidant system (Gillham and Dodge, 1986; Noctor et al., 1999). Furthermore H_2O_2 that is produced in the peroxisomes during photorespiration is discussed to be involved in plant defense mechanisms (Taler et al., 2004).

It has previously been successfully shown in two independent approaches that short-circuiting of the photorespiratory cycle led to enhanced biomass production of C_3 plants (Kebeish et al., 2007; Maier et al., 2012).
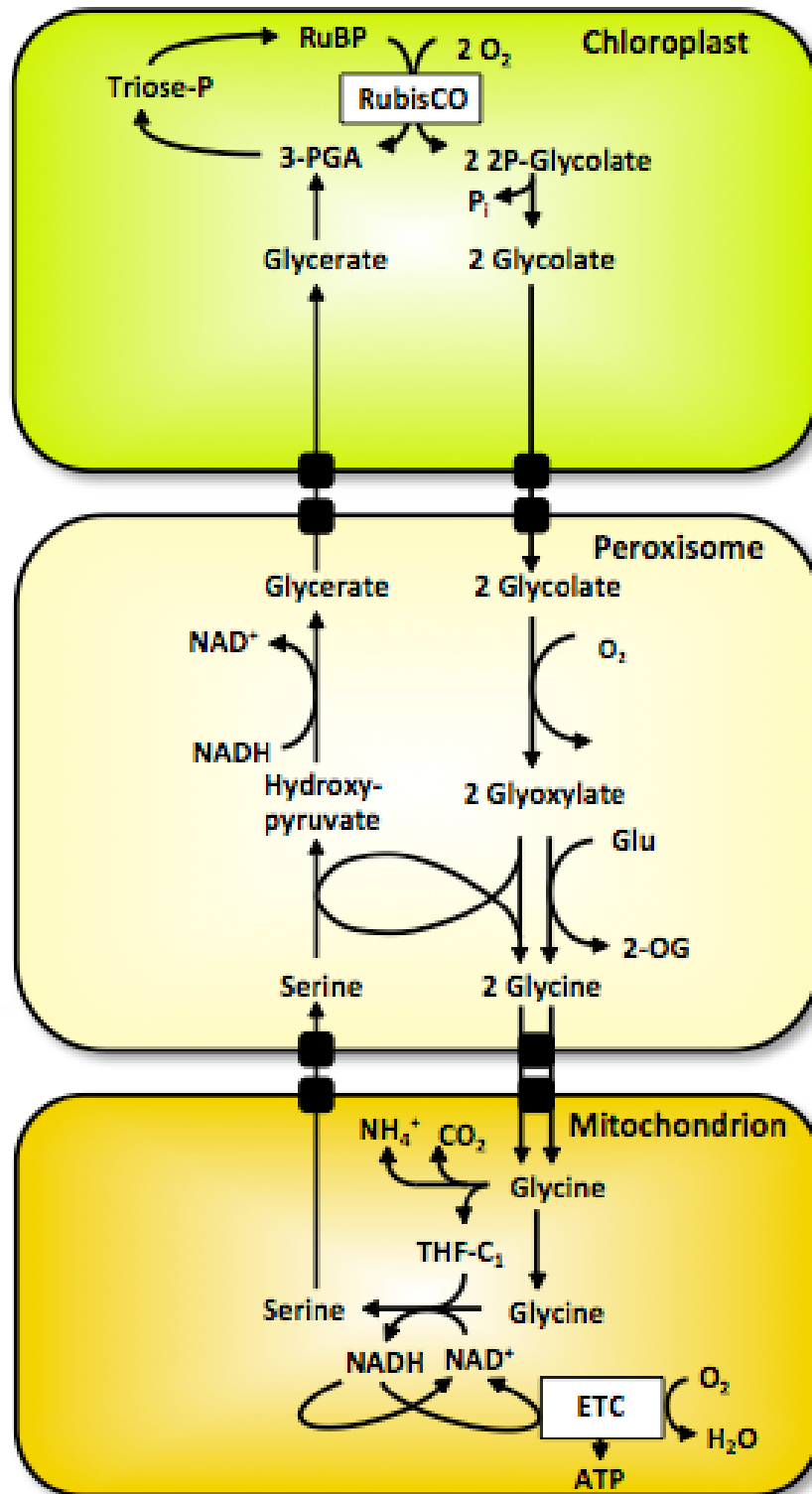
Figure 1: The photorespiratory carbon cycle. Enzymatic reactions take place in the compartments chloroplasts, peroxisomes, and mitochondria; RubisCO, ribulose-1,5-bisphosphate carboxylase/oxygenase; 3-PGA, 3-phosphoglycerate; Glu, glutamate; RuBP, ribulose-bisphosphate; THF, tetrahydofolate. Black boxes symbolize transport proteins. The transport proteins represented in this scheme have not been characterized on the molecular level. The scheme was adapted from (Linka, 2008).

2.2.1 Photorespiratory transporters

As described above the photorespiratory cycle is a highly compartmentalized pathway with metabolic reactions occurring in the chloroplast, the leaf peroxisomes, the mitochondria and the cytosol. Each of the compartments is surrounded by one (peroxisomes) or two membranes (chloroplast and mitochondria) that circumvent unregulated diffusion or leakage of metabolites. However, to allow the metabolic interaction of the different compartments in a controlled manner, transporters account for the shuttle of precursors, intermediates and products across the membranes. To control the selective flux of metabolites most transporters are considered to be highly specific concerning their transport substrate (Weber, 2004; Weber et al., 2004; Weber et al., 2005; Weber and Fischer, 2007). In C_3 plants the ratio of carboxylation to oxygenation of Rubisco is approximately 3:1 when grown in ambient air with a CO_2 concentration of 380 ppm and O_2 concentration of 210,000 ppm (Sharkey, 1988). This results in a high flux rate of metabolites through the photorespiratory cycle. Hence, transporters involved in photorespiration have to cope with an immense metabolite flux and have to have distinct properties. Between 14 and 18 transmembrane transport processes are required for a functional photorespiratory carbon and nitrogen cycle. However, only DiT1 and DiT2 (transport proteins for 2-oxoglutarate and glutamate), both transporters of the nitrogen cycle, are characterized at a molecular level to date (Somerville and Ogren, 1983; Renne et al., 2003). Characterization of transporters of the carbon cycle of photorespiration still remained elusive.

2.2.2 The chloroplastidic glycolate glycerate transporter

In the 1980s in the lab of Howitz and McCarty specific transport activity in extracts of chloroplast envelope fractions was analyzed (Howitz and Mccarty, 1985a, b). Glycolate is produced in the chloroplast and transported out of this compartment. One of the final steps of the photorespiratory cycle is the production of glycerate and its transport into the chloroplast. Howitz and McCarty biochemically analyzed a chloroplastidic protein, which promotes the transport of glycolate and glycerate across the envelope membrane with rates required to cope with the flux through the photorespiratory pathway (Howitz and McCarty, 1991). Additionally, they postulated that the single protein is a transporter for both substrates, glycolate and glycerate, and that the transport mode is not necessarily stoichiometric. Hence, a proton-substrate-symport was measured and it was observed that the presence of the second substrate on the opposite side of the membrane stimulates transport (Howitz and McCarty, 1986, 1991). This is an important finding as for two molecules glycolate exported only one molecule glycerate is imported (Figure 1). Although these findings were completed decades ago the transporter was never characterized on the molecular level.

2.3   *A. thaliana* mutants impaired in the photorespiratory pathway

The majority of molecular components of the photorespiratory pathway have been identified using *A. thaliana* mutants impaired in photorespiration. Prominent examples are mutants impaired in 2-phosphoglycolate phosphatase (PGLP) the enzyme catalyzing the first reaction of the photorespiratory cycle (Somerville and Ogren, 1979), or the mitochondrial SHMT (Somerville and Ogren, 1981; Voll et al., 2006).

A significant similarity of photorespiratory mutants is reduced or no growth under ambient air but normal wild type-like growth under CO_2 enriched air. Only peroxisomal HPR mutants show no typical photorespiratory growth limitation in ambient air due to an alternative cytosolic pathway that suppresses the effect of the mutation (Timm et al., 2008). An additional common observation is the accumulation of specific photorespiratory metabolites in the mutant plants and differences in chlorophyll fluorescence parameters between WT and mutants plants. These shared characteristics of photorespiratory mutants can be used as a tool to identify yet unknown factors involved in the photorespiratory cycle, as it was recently shown for the chlorophyll fluorescence imaging approach (Badger et al., 2009).

2.4 C_4 photosynthesis

As described above high rates of photorespiration can lead to a decrease in biomass production. During evolution several carbon concentrating mechanisms evolved to circumvent the oxygenation reaction of Rubisco and thereby inhibiting the negative effects of photorespiration. One of the most prominent mechanisms is C_4 photosynthesis. This special type of photosynthesis evolved 15 to 40 million years ago (reviewed in (Sage, 2004). In C_4 photosynthesis, CO_2 is prefixed by the O_2 insensitive enzyme phosphoenolpyruvate carboxylase (PEPC) into a C_4 acid and released by decarboxylation of the C_4 acid in close vicinity to Rubisco (Figure 2A). The pathway as it occurs in maize involves enzymatic reactions taking place in the cytosol and chloroplast of two morphological different cell types, namely mesophyll and bundle sheath cells (Figure 2B). This spatial separation results in an increased number of required membrane transport proteins compared to C_3 photosynthesis. Despite detailed understanding of the biochemical reactions in the pathway (Furbank, 2011) we still know comparatively little about the membrane proteins mediating the shuttle of metabolites. As a long term goal of plant researchers is to introduce the highly efficient C_4 photosynthetic cycle in C_3 plants it is also of interest to identify the regulatory proteins that orchestrate the establishment and onset of C_4 photosynthesis.
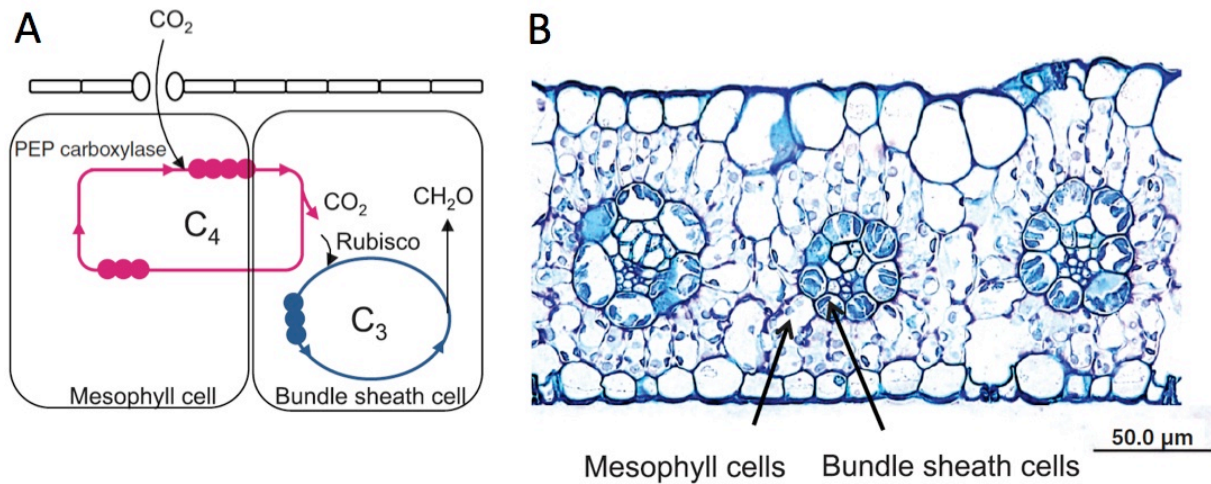
Figure 2: Simplified scheme of the C4 photorespiratory cycle. (A) CO_2 is fixed in the mesophyll cells by Phosphoenolpyruvate (PEP) carboxylase to yield a C4 acid (pink dots and lines). The C4 acid is shuttled to the bundle sheath cells where it is decarboxylated and the CO_2 is concentrated in close vicinity of Rubisco. Rubisco feeds the CO_2 into the Calvin cycle (blue cycle) to produce carbohydrates (CH_2O). (B) Sorghum leaf section showing a typical leaf architecture of a C4 species with only two or three mesophyll cells in between the vascular tissue and chloroplasts in bundle sheath cells. Rubisco, ribulose-1,5-bisphosphate carboxylase/oxygenase. The figure was modified after (von Caemmerer et al., 2012).

2.5 Reverse and forward genetics as a tool to find missing links in plant metabolism and development

As a tool to find regulatory and transport proteins involved in C4 photosynthesis, photorespiration and other metabolic pathways in plants reverse and forward genetics have been of great advantage. For instance the common features of mutants impaired in the photorespiratory pathway (see 2.3) were used for successful forward genetic screens (Somerville and Ogren, 1979, 1980b, 1981; Wallsgrove et al., 1987; Voll et al., 2006).

In the past co-expression analysis as a reverse genetics approaches has been a well studied and successful method to find candidates for missing links in pathways. This analysis is based on the assumption that genes that are co-expressed function in the same pathway and was developed by (Eisen et al., 1998). Hence, unknown genes that are co-regulated with genes of a particular metabolic pathway are probably involved in the same biological process. By this method a wide range of genes have been functional characterized in non-plant organisms like yeast (Wu et al., 2002) or humans (Lee et al., 2004). In *Arabidopsis thaliana* a co-expression analysis in combination with reverse genetics has been a successfully strategy to find genes for example involved in flavonoid metabolism (Yonekura-Sakakibara et al., 2007), cellulose and

(Persson et al., 2005) lignin biosynthesis (Ehlting et al., 2005; Alejandro et al., 2012) and aliphatic glucosinolate biosynthesis (Hansen et al., 2007).

An additional reverse genetics tool is the analysis of the whole transcriptome of an organism or different tissues of an organism. It has been successfully used by comparing the transcriptome of closely related C_3 and C_4 plants of the genus *Cleome* (Bräutigam et al., 2011). Furthermore, comparative transcriptomics were used in a recent approach to shed light on the question: "Evolution of C_4 photosynthesis in the genus *Flaveria*: How many and which genes does it take to make C_4?" (Gowik et al., 2011). In this approach the transcriptome of five species within the genus *Flaveria*, including C_3 plants, C_3- C_4 intermediates and plants performing C_4 photosynthesis were analyzed to isolate new candidates for yet unknown factors in C_4 photosynthesis, as well as identify C_4 associated pathways.

A different approach is the analysis of the whole proteome of an organism, specific tissues or even specific organelles. For example, comparative proteomics of chloroplasts was used to identify the characteristic composition of proteins in chloroplast of C_4 plants and thereby finding candidates for yet unidentified proteins involved in C_4 photosynthesis. Here two different and independent proteomic approaches were used: on the one hand the set of proteins characteristic for chloroplasts of C_3 and C_4 plants were compared (Bräutigam et al., 2008) and on the other hand chloroplasts of the functionally different cell types bundle sheath and mesophyll cells (Figure 2) in the C_4 plant *Z. mays* (Majeran et al., 2008).

Hence reverse and forward genetics are important tools to find missing links, like regulatory and transport proteins in C_4 photosynthesis and photorespiration.

3. Discussion and Outlook

Plants are able to fix inorganic carbon in photosynthesis and thereby contribute to a significant share of the worldwide carbon cycle what makes them indispensable and important for the land-based biosphere. Oxygenic photosynthesis evolved over 2 billion years ago when the atmosphere was free of O_2 (Kirschvink and Kopp, 2008). When oxygenic photosynthesis emerged the level of O_2 in the atmosphere 'rapidly' increased to relatively high levels about 1 billion years ago. This event resulted in a significant competition of O_2 and CO_2 for the active site of Rubisco, as Rubisco functions as carboxylase as well as oxygenase (Smith, 1976). The oxygenation reaction generates one molecule 2-PG that is toxic to the plant and has to be recycled in the photorespiratory pathway that, in the end, leads to a decrease in biomass production. 15 to 40 million years ago C_4 photosynthesis evolved, what is a highly efficient type of photosynthesis that minimizes the rate of photorespiration. In C_4 photosynthesis CO_2 is concentrated around Rubisco and therefore the oxygenation reaction is reduced.

Both, C_4 photosynthesis and photorespiration are major subjects of plant research that aims to develop highly efficient and robust C_3 plants. The aim of this thesis was to contribute to the knowledge about the unknown factors of these pathways, such as regulatory proteins and metabolite transporters. The long-term goal of studying factors involved in the establishment and the processing of C_4 photosynthesis and photorespiration is a "second green evolution" (von Caemmerer et al., 2012). In this project the C_4 cycle is inserted into the C_3 plant rice to increase the yield. An obstacle on the way to achieve this goal is the observation that the C_4 cycle is not a linear cycle but rather branched and involves several levels of regulation and organization (Bräutigam et al., 2011; Gowik et al., 2011).

Therefore a systems biology approach was used in this thesis to identify candidates for regulatory proteins that orchestrate the establishment of photosynthesis and to finally provide a comprehensive understanding of the dynamic interactions that take place between different cell types and cellular compartments (Pick et al., 2011). As carbon fixation by the oxygen insensitive enzyme PEPC and the fixed carbon entering the Calvin cycle is separated between different cell types a different set of proteins is required in each cell. Hence a highly coordinated expression in the specific cell type is important. The transcription factor Golden 2 (G2) was shown to be a bundle sheath specific transcription factor in maize (Roth et al., 1996). Additionally a whole family of G2-like (Glk) factors was identified in maize offering promising candidates for further research (Rossini et al., 2001). A second attempt to investigate cell specific expression is the analysis of cis-regulatory elements responsible for cell-specific gene expression. The promoter elements responsible for bundle sheath specific expression were identified for the P- subunit of the GDC in the C_4 plant *Flaveria trinervia* (Engelmann et al., 2008; Wiludda et al., 2012) offering an alternative approach. Still, one major gap in the knowledge about the C_4 photorespiratory

cycle are important metabolite transporters that facilitate the transport of metabolites across organellar membranes. Hence, the systems analysis approach presented in this thesis also provides candidates for putative transporters. The next step will be the molecular and biochemical analysis of single candidates to fill the gaps and to complete the knowledge about the establishment and the processing of the C_4 photosynthetic cycle.

In this thesis the single gene analysis of a candidate originating from a reverse genetics screen to identify photorespiratory transporters was shown to be successful and PLGG1 was identified as the photorespiratory glycolate glycerate transporter. As photorespiration leads to a decrease in biomass production it is also an interesting field for plant research.

Therefore a different approach to increase biomass production in C_3 plants deals with photorespiration. A biotechnological approach was shown to successfully short-cut photorespiration by inducing an alternative pathway in *A. thaliana* to recycle 2-PG and increase biomass (Kebeish et al., 2007). In this approach the *Escherichia coli* glycolate catabolic pathway was introduced into the plastid in *A. thaliana.* As the chloroplastidic glycolate transporter is finally identified it needs to be investigated whether biomass production could be further improved by enhancing the flux through the biotechnologically introduced prokaryotic glycolate pathway by trapping the glycolate in the chloroplast as it occurs in *plgg1-1* plants. In this context it is also important to further investigate the transport characteristics and get a detailed view of the structure of the transporter PLGG1. By inserting PLGG1 into the biotechnologically designed plants PLGG1 could be used as a regulated valve to control the amount of glycolate in the chloroplast. A very recent attempt followed an alternative approach by introducing an artificial complete glycolate catabolic cycle in chloroplasts of *A. thaliana* comprising glycolateoxidase (GO), malatesynthase (MS), and catalase (CAT) (Maier et al., 2012).

A surprising and interesting finding was that PLGG1 was found to be highly abundant in maize chloroplasts (Bräutigam et al., 2008). But why is a photorespiratory transporter highly abundant in C_4 photosynthesis that shows very reduced rates of photorespiration? This observation shows that there are still major gaps in our knowledge about the C_4 photosynthetic cycle. The objectives of further research will be to answer the question if PLGG1 took over another function in C_4 photosynthesis or if it is still the chloroplastidic glycolate glycerate transporter. If it is still the glycolate glycerate transporter the functional relevance of its abundance in C_4 needs to be investigated. The observations from these studies will bring further clarity in the C_4 cycle and will help to shed light on yet unknown parts of C_4 photosynthesis.

4. Theses

Despite their importance for economy and plant breeding research many important factors involved in the metabolic processes photorespiration and the highly efficient C_4 photosynthesis are still unknown. Among those factors are regulatory proteins orchestrating the onset of C_4 photosynthesis or transport proteins mediating the shuttling of specific metabolites across organellar membranes in the photorespiratory cycle.

1) The model C_4 plant *Zea mays* displays a developmental gradient along the leaf with young and immature cells at the base of the leaf and old and mature cells at the leaf tip. This developmental gradient is mirrored in the C_4 photosynthetic activity, which is highest at the mature leaf tip and very weak at the young base of the leaf. Against previous hypothesis there is no phase of C_3 photosynthesis at the base of the leaf, where the rate of C_4 photosynthesis is low. Along the gradient specific patterns of transcript and metabolite accumulation appear. These specific patterns can be used to refine the current C_4 model as it occurs in *Z. mays* and to generate candidates for yet unknown regulatory and transport proteins (Manuscript 1: Pick et al., 2011).

2) The photorespiratory cycle is a highly compartmentalized process with enzymatic reactions taking place in the chloroplasts, the peroxisomes, the mitochondria and the cytosol. To ensure the metabolic connection of the different compartments highly selective transporters account for the shuttling of specific metabolites. Despite their importance only two transporters involved in the photorespiratory nitrogen cycle could be identified to date. A reverse genetics approach followed by single gene analysis can be used to identify candidates for yet unknown transport proteins. *Arabidopsis thaliana* T-DNA-insertional k.o. mutants of candidate genes can be used to analyze the role in photorespiration. Metabolite analysis in combination with molecular and biochemical characterization are used to define transport substrates and transport characteristics. With this approach Plastidic glycolate glycerate transporter 1 (PLGG1) could be identified as the chloroplastidic glycolate glycerate transporter (Manuscript 2: Pick et al., 2013).

5. References

Alejandro, S., Lee, Y., Tohge, T., Sudre, D., Osorio, S., Park, J., Bovet, L., Lee, Y., Geldner, N., Fernie, A.R., and Martinoia, E. (2012). AtABCG29 is a monolignol transporter involved in lignin biosynthesis. Curr Biol **22,** 1207-1212.

Badger, M.R., Fallahi, H., Kaines, S., and Takahashi, S. (2009). Chlorophyll fluorescence screening of Arabidopsis thaliana for CO2 sensitive photorespiration and photoinhibition mutants. Funct Plant Biol **36,** 867-873.

Brautigam, A., Hoffmann-Benning, S., and Weber, A.P. (2008). Comparative proteomics of chloroplast envelopes from C3 and C4 plants reveals specific adaptations of the plastid envelope to C4 photosynthesis and candidate proteins required for maintaining C4 metabolite fluxes. Plant Physiol **148,** 568-579.

Brautigam, A., Kajala, K., Wullenweber, J., Sommer, M., Gagneul, D., Weber, K.L., Carr, K.M., Gowik, U., Mass, J., Lercher, M.J., Westhoff, P., Hibberd, J.M., and Weber, A.P. (2011). An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species. Plant Physiol **155,** 142-156.

Cassman, K.G., and Liska, A.J. (2007). Food and fuel for all: realistic or foolish? Biofuel Bioprod Bior **1,** 18-23.

Collakova, E., Goyer, A., Naponelli, V., Krassovskaya, I., Gregory, J.F., 3rd, Hanson, A.D., and Shachar-Hill, Y. (2008). Arabidopsis 10-formyl tetrahydrofolate deformylases are essential for photorespiration. Plant Cell **20,** 1818-1832.

Ehlting, J., Mattheus, N., Aeschliman, D.S., Li, E., Hamberger, B., Cullis, I.F., Zhuang, J., Kaneda, M., Mansfield, S.D., Samuels, L., Ritland, K., Ellis, B.E., Bohlmann, J., and Douglas, C.J. (2005). Global transcript profiling of primary stems from Arabidopsis thaliana identifies candidate genes for missing links in lignin biosynthesis and transcriptional regulators of fiber differentiation. Plant J **42,** 618-640.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc Nat Acad Sci U S A **95,** 14863-14868.

Engelmann, S., Wiludda, C., Burscheidt, J., Gowik, U., Schlue, U., Koczor, M., Streubel, M., Cossu, R., Bauwe, H., and Westhoff, P. (2008). The gene for the P-subunit of glycine decarboxylase from the C4 species Flaveria trinervia: analysis of transcriptional control in transgenic Flaveria bidentis (C4) and Arabidopsis (C3). Plant Physiol **146,** 1773-1785.

Furbank, R.T. (2011). Evolution of the C(4) photosynthetic mechanism: are there really three C(4) acid decarboxylation types? J Exp Bot **62,** 3103-3108.

Gillham, D.J., and Dodge, A.D. (1986). Hydrogen-Peroxide-Scavenging Systems within Pea-Chloroplasts - a Quantitative Study. Planta **167,** 246-251.

Gowik, U., Brautigam, A., Weber, K.L., Weber, A.P., and Westhoff, P. (2011). Evolution of C4 photosynthesis in the genus Flaveria: how many and which genes does it take to make C4? Plant Cell **23,** 2087-2105.

Hansen, B.G., Kliebenstein, D.J., and Halkier, B.A. (2007). Identification of a flavin-monooxygenase as the S-oxygenating enzyme in aliphatic glucosinolate biosynthesis in Arabidopsis. Plant J **50,** 902-910.

Howitz, K.T., and Mccarty, R.E. (1985a). Kinetic Characteristics of the Chloroplast Envelope Glycolate Transporter. Biochemistry **24,** 2645-2652.

Howitz, K.T., and Mccarty, R.E. (1985b). Substrate-Specificity of the Pea Chloroplast Glycolate Transporter. Biochemistry **24,** 3645-3650.

Howitz, K.T., and Mccarty, R.E. (1986). D-Glycerate Transport by the Pea Chloroplast Glycolate Carrier - Studies on [1-C-14] D-Glycerate Uptake and D-Glycerate Dependent O-2 Evolution. Plant Physiol **80,** 390-395.

Howitz, K.T., and McCarty, R.E. (1991). Solubilization, partial purification, and reconstitution of the glycolate/glycerate transporter from chloroplast inner envelope membranes. Plant Physiol **96,** 1060-1069.

Husic, D.W., and Tolbert, N.E. (1987). Inhibition of glycolate and D-lactate metabolism in a Chlamydomonas reinhardtii mutant deficient in mitochondrial respiration. Proc Nat Acad Sci U S A **84,** 1555-1559.

Kajala, K., Covshoff, S., Karki, S., Woodfield, H., Tolley, B.J., Dionora, M.J., Mogul, R.T., Mabilangan, A.E., Danila, F.R., Hibberd, J.M., and Quick, W.P. (2011). Strategies for engineering a two-celled C(4) photosynthetic pathway into rice. J Exp Bot **62,** 3001-3010.

Kebeish, R., Niessen, M., Thiruveedhi, K., Bari, R., Hirsch, H.J., Rosenkranz, R., Stabler, N., Schonfeld, B., Kreuzaler, F., and Peterhansel, C. (2007). Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in Arabidopsis thaliana. Nat Biotechnol **25,** 593-599.

Kirschvink, J.L., and Kopp, R.E. (2008). Palaeoproterozoic ice houses and the evolution of oxygen-mediating enzymes: the case for a late origin of photosystem II. Phil Trans R Soc B Biol Sci **363,** 2755-2765.

Kozaki, A., and Takeba, G. (1996). Photorespiration protects C3 plants from photooxidation. Nature **384,** 557-560.

Lee, H.K., Hsu, A.K., Sajdak, J., Qin, J., and Pavlidis, P. (2004). Coexpression analysis of human genes across many microarray data sets. Genome Res **14,** 1085-1094.

Linka, M. (2008). Understanding the origin and function of organellar metabolite transport proteins in photosynthetic eukaryotes: Galdieria sulphuraria and Arabidopsis thaliana as model systems. PhD thesis.

Madore, M., and Grodzinski, B. (1984). Effect of Oxygen Concentration on C-14-Photoassimilate Transport from Leaves of Salvia-Splendens L. Plant Physiol **76,** 782-786.

Maier, A., Fahnenstich, H., von Caemmerer, S., Engqvist, M.K., Weber, A.P., Flugge, U.I., and Maurino, V.G. (2012). Transgenic Introduction of a Glycolate Oxidative Cycle into A. thaliana Chloroplasts Leads to Growth Improvement. Front Plant Sci **3,** 38.

Majeran, W., Zybailov, B., Ytterberg, A.J., Dunsmore, J., Sun, Q., and van Wijk, K.J. (2008). Consequences of C4 differentiation for chloroplast membrane proteomes in maize mesophyll and bundle sheath cells. Mol Cell Proteomics **7,** 1609-1638.

Migge, A., Kahmann, U., Fock, H.P., and Becker, T.W. (1999). Prolonged exposure of tobacco to a low oxygen atmosphere to suppress photorespiration decreases net photosynthesis and results in changes in plant morphology and chloroplast structure. Photosynthetica **36,** 107-116.

Noctor, G., Arisi, A.C.M., Jouanin, L., and Foyer, C.H. (1999). Photorespiratory glycine enhances glutathione accumulation in both the chloroplastic and cytosolic compartments. J Exp Bot **50,** 1157-1167.

Persson, S., Wei, H., Milne, J., Page, G.P., and Somerville, C.R. (2005). Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. Proc Nat Acad Sci U S A **102,** 8633-8638.

Pick, T.R., Brautigam, A., Schluter, U., Denton, A.K., Colmsee, C., Scholz, U., Fahnenstich, H., Pieruschka, R., Rascher, U., Sonnewald, U., and Weber, A.P.M. (2011). Systems analysis of a maize leaf developmental gradient redefines the current C4 model and provides candidates for regulation. Plant Cell **23,** 4208-4220.

Renne, P., Dressen, U., Hebbeker, U., Hille, D., Flugge, U.I., Westhoff, P., and Weber, A.P.M. (2003). The Arabidopsis mutant dct is deficient in the plastidic glutamate/malate translocator DiT2. Plant J **35,** 316-331.

Rossini, L., Cribb, L., Martin, D.J., and Langdale, J.A. (2001). The maize golden2 gene defines a novel class of transcriptional regulators in plants. Plant Cell **13,** 1231-1244.

Roth, R., Hall, L.N., Brutnell, T.P., and Langdale, J.A. (1996). bundle sheath defective2, a Mutation That Disrupts the Coordinated Development of Bundle Sheath and Mesophyll Cells in the Maize Leaf. Plant Cell **8,** 915-927.

Sage, R.F. (2004). The evolution of C-4 photosynthesis. New Phytol **161,** 341-370.

Sharkey, T.D. (1988). Estimating the Rate of Photorespiration in Leaves. Physiol Plant **73,** 147-152.

Smith, B.N. (1976). Evolution of C-4 Photosynthesis in Response to Changes in Carbon and Oxygen Concentrations in Atmosphere through Time. Biosystems **8,** 24-32.

Somerville, C.R., and Ogren, W.L. (1979). Phosphoglycolate Phosphatase-Deficient Mutant of Arabidopsis. Nature **280,** 833-836.

Somerville, C.R., and Ogren, W.L. (1980). Photorespiration mutants of Arabidopsis thaliana deficient in serine-glyoxylate aminotransferase activity. Proc Nat Acad Sci U S A **77,** 2684-2687.

Somerville, C.R., and Ogren, W.L. (1981). Photorespiration-deficient Mutants of Arabidopsis thaliana Lacking Mitochondrial Serine Transhydroxymethylase Activity. Plant Physiol **67,** 666-671.

Somerville, C.R., and Ogren, W.L. (1982). Mutants of the cruciferous plant Arabidopsis thaliana lacking glycine decarboxylase activity. Biochem J **202,** 373-380.

Somerville, S.C., and Ogren, W.L. (1983). An Arabidopsis thaliana mutant defective in chloroplast dicarboxylate transport. Proc Nat Acad Sci U S A **80,** 1290-1294.

Taler, D., Galperin, M., Benjamin, I., Cohen, Y., and Kenigsbuch, D. (2004). Plant eR genes that encode photorespiratory enzymes confer resistance against disease. Plant Cell **16,** 172-184.

Timm, S., Nunes-Nesi, A., Parnik, T., Morgenthal, K., Wienkoop, S., Keerberg, O., Weckwerth, W., Kleczkowski, L.A., Fernie, A.R., and Bauwe, H. (2008). A cytosolic pathway for the conversion of hydroxypyruvate to glycerate during photorespiration in Arabidopsis. Plant Cell **20,** 2848-2859.

Voll, L.M., Jamai, A., Renne, P., Voll, H., McClung, C.R., and Weber, A.P.M. (2006). The photorespiratory Arabidopsis shm1 mutant is deficient in SHM1. Plant Physiol **140,** 59-66.

von Caemmerer, S., Quick, W.P., and Furbank, R.T. (2012). The development of C(4) rice: current progress and future challenges. Science **336,** 1671-1672.

Wallsgrove, R.M., Turner, J.C., Hall, N.P., Kendall, A.C., and Bright, S.W. (1987). Barley mutants lacking chloroplast glutamine synthetase-biochemical and genetic analysis. Plant Physiol **83,** 155-158.

Weber, A.P.M. (2004). Solute transporters as connecting elements between cytosol and plastid stroma. Curr Opin Plant Biol **7,** 247-253.

Weber, A.P.M., and Fischer, K. (2007). Making the connections - The crucial role of metabolite transporters at the interface between chloroplast and cytosol. FEBS Lett. **581,** 2215-2222.

Weber, A.P.M., Schneidereit, J., and Voll, L.M. (2004). Using mutants to probe the in vivo function of plastid envelope membrane metabolite transporters. J Exp Bot **55,** 1231-1244.

Weber, A.P.M., Schwacke, R., and Flugge, U.I. (2005). Solute transporters of the plastid envelope membrane. Annu Rev Plant Biol **56,** 133-164.

Wiludda, C., Schulze, S., Gowik, U., Engelmann, S., Koczor, M., Streubel, M., Bauwe, H., and Westhoff, P. (2012). Regulation of the photorespiratory GLDPA gene in C(4) flaveria: an intricate interplay of transcriptional and posttranscriptional processes. Plant Cell **24,** 137-151.

Wu, L.F., Hughes, T.R., Davierwala, A.P., Robinson, M.D., Stoughton, R., and Altschuler, S.J. (2002). Large-scale prediction of Saccharomyces cerevisiae gene function using overlapping transcriptional clusters. Nat genet **31,** 255-265.

Yonekura-Sakakibara, K., Tohge, T., Niida, R., and Saito, K. (2007). Identification of a flavonol 7-O-rhamnosyltransferase gene determining flavonoid pattern in Arabidopsis by transcriptome coexpression analysis and reverse genetics. J Biol Chem **282,** 14932-14941.

6. Manuscripts

1. **Thea R. Pick, Andrea Bräutigam, Urte Schlüter, Alisandra K. Denton, Christian Colmsee, Uwe Scholz, Holger Fahnenstich, Roland Pieruschka, Uwe Rascher, Uwe Sonnewald, and Andreas P.M. Weber** (2011). Systems Analysis of a Maize Leaf Developmental Gradient Redefines the Current C_4 Model and Provides Candidates for Regulation. Plant Cell **23**, 4208-4220.

2. **Thea R. Pick, Andrea Bräutigam, Matthias A. Schulz, Toshihiro Obata, Alisdair R. Fernie, and Andreas P.M. Weber** (2013). *PLGG1*, a plastidic glycolate glycerate transporter, is required for photorespiration and defines a new class of metabolite transporters. Submitted for publication to *PNAS*.

Manuscript 1

Systems Analysis of a Maize Leaf Developmental Gradient Redefines
the Current C$_4$ Model and Provides Candidates for Regulation

LARGE-SCALE BIOLOGY ARTICLE

Systems Analysis of a Maize Leaf Developmental Gradient Redefines the Current C$_4$ Model and Provides Candidates for Regulation [W] [OA]

Thea R. Pick,[a,b,1] Andrea Bräutigam,[a,1] Urte Schlüter,[c] Alisandra K. Denton,[a,b] Christian Colmsee,[d] Uwe Scholz,[d] Holger Fahnenstich,[e] Roland Pieruschka,[f] Uwe Rascher,[f] Uwe Sonnewald,[c] and Andreas P.M. Weber[a,2]

[a] Plant Biochemistry, Heinrich Heine University Düsseldorf, 40225 Duesseldorf, Germany
[b] International Graduate Program for Plant Science (iGrad-plant), Heinrich Heine University Düsseldorf, 40225 Duesseldorf, Germany
[c] Department of Biology, Friedrich Alexander University Erlangen-Nürnberg, 91058 Erlangen, Germany
[d] Leibniz Institute of Plant Genetics and Crop Plant Research, 06466 Gatersleben, Germany
[e] Metanomics GmbH, 10589 Berlin, Germany
[f] Forschungszentrum Jülich, Institut für Bio- und Geowissenschaften (Pflanzenwissenschaften), 52425 Juelich, Germany

We systematically analyzed a developmental gradient of the third maize (*Zea mays*) leaf from the point of emergence into the light to the tip in 10 continuous leaf slices to study organ development and physiological and biochemical functions. Transcriptome analysis, oxygen sensitivity of photosynthesis, and photosynthetic rate measurements showed that the maize leaf undergoes a sink-to-source transition without an intermediate phase of C$_3$ photosynthesis or operation of a photorespiratory carbon pump. Metabolome and transcriptome analysis, chlorophyll and protein measurements, as well as dry weight determination, showed continuous gradients for all analyzed items. The absence of binary on–off switches and regulons pointed to a morphogradient along the leaf as the determining factor of developmental stage. Analysis of transcription factors for differential expression along the leaf gradient defined a list of putative regulators orchestrating the sink-to-source transition and establishment of C$_4$ photosynthesis. Finally, transcriptome and metabolome analysis, as well as enzyme activity measurements, and absolute quantification of selected metabolites revised the current model of maize C$_4$ photosynthesis. All data sets are included within the publication to serve as a resource for maize leaf systems biology.

INTRODUCTION

The mechanisms underlying organ development and function are fundamental questions of biology. In plants, grass leaves represent an excellent model in which the establishment of various functions can be followed in a base-to-tip developmental gradient in a single leaf. Cells at the tip of the leaf are the oldest and most mature cells, while cells at the base are the youngest (Nelson and Langdale, 1992). We chose maize (*Zea mays*) to follow the establishment of photosynthetic functions during leaf development. Maize employs the highly efficient C$_4$ type of photosynthesis, which concurrently evolved in multiple seed plant families ∼15 to 40 million years ago, long after C$_3$ photosynthesis had been established (Edwards and Smith, 2010). It

has been previously proposed that the evolutionary progression from C$_3$ to C$_4$ can also be detected in maize leaves along a spatial gradient (Nelson and Langdale, 1992, and references therein), very much like Haeckel suggested that ontogeny recapitulates phylogeny during embryo development in animals (Haeckel, 1866).

C$_4$ photosynthesis has been considered a possible route for spawning a second green revolution in C$_3$ crop plants, such as rice (*Oryza sativa*) (Hibberd et al., 2008). Plants using C$_4$ photosynthesis are capable of producing biomass at faster rates than C$_3$ plants, or, alternatively, these plants can inhabit harsher habitats with limited resources (Sage, 2004, and references therein). The key limitation for more productive photosynthesis is the concentration of carbon dioxide at the site of its assimilation, the reductive pentose phosphate pathway (rPPP) in plant chloroplasts. The enrichment of carbon dioxide around ribulose-1,5-bisphosphate carboxylase/oxygenase minimizes the oxygenation of ribulose-1,5-bisphosphate, which leads to a reduced rate of photorespiration. From an engineering standpoint, C$_4$ photosynthesis, similar to a supercharged combustion engine, enriches the limiting factor, carbon dioxide, via a biochemical cycle operating between the site of initial fixation and final assimilation. Although the C$_4$ cycle as described below appears

---

[1] These authors contributed equally to this work.
[2] Address correspondence to andreas.weber@uni-duesseldorf.de.
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.plantcell.org) is: Andreas P.M. Weber (andreas.weber@uni-duesseldorf.de).
[W] Online version contains Web-only data.
[OA] Open Access articles can be viewed online without a subscription.
www.plantcell.org/cgi/doi/10.1105/tpc.111.090324

deceptively simple, differences between C_3 and C_4 photosynthesis go beyond just the addition of the C_4 cycle on top of the rPPP and include photorespiration, protein translation, cellular and tissue architecture, electron transfer adaptations, cell–cell connections, and likely other still unknown adaptations (Bräutigam et al., 2011; Gowik et al., 2011). In C_4 plants, carbon dioxide is enriched by affixing it to an acceptor, transferring it to the site of final assimilation, liberating it, and returning and recycling the acceptor for a new round. This system is referred to as the C_4 cycle. To avoid a futile cycle, the site of initial fixation, the mesophyll, is spatially separated from the site of assimilation, the bundle sheath. Canonically, maize operates a linear C_4 cycle (Hatch, 1987; Furbank, 2011): In the compartment of initial fixation, the mesophyll, the carbon dioxide acceptor phospho-*enol*pyruvate (PEP) is formed in the chloroplast from pyruvate by pyruvate:phosphate dikinase (PPDK) and then exported to the cytosol. There, carbon dioxide in the form of bicarbonate is fixed by PEP carboxylase (PEPC), creating the dicarboxylic C_4 acid oxaloacetate (OAA) from PEP. OAA is subsequently transferred to the chloroplast and reduced to malate, which is then exported to the cytosol of mesophyll cells. Malate is transported by mass flow to the bundle sheath, the compartment of final assimilation, where it is imported into chloroplasts and decarboxylated by the NADP-dependent malic enzyme (NADP-ME), yielding carbon dioxide, pyruvate, and NADPH. Pyruvate is exported from the chloroplast and returned to the mesophyll for regeneration of the acceptor PEP (Hatch, 1987). Whereas this canonical model of NADP-ME C_4 photosynthesis is depicted in many textbooks, several reports question its simplicity; however, an alternative model has not yet been formulated. For example, bundle sheath strands can efficiently decarboxylate not only malate but also the amino acid Asp (Chapman and Hatch, 1981). Older maize leaves, at least, harbor a second decarboxylation enzyme, PEP carboxykinase (PEP-CK), which releases carbon dioxide from OAA, producing PEP (Wingler et al., 1999). Furthermore, approximately one-quarter of radioactively labeled carbon dioxide that was fed to maize leaves was found to be rapidly incorporated into Asp (Hatch, 1971). Such side routes to the canonical NADP-ME C_4 pathway would require alternative transfer metabolites between mesophyll and bundle sheath cells, such as Asp or Ala, and alternative decarboxylation pathways would alter the demands on the remaining enzymes and the intracellular (Bräutigam and Weber, 2011a) and intercellular (Sowínski et al., 2008) transport systems. Understanding both the intracellular transport system between chloroplasts and cytosol and the intercellular transport between the mesophyll and bundle sheath cells is still in its infancy (Bräutigam et al., 2008; Sowínski et al., 2008; Bräutigam and Weber, 2011a, 2011b; Weber and von Caemmerer, 2010; Weber and Linka, 2011). Finally, understanding the regulatory circuits controlling C_4 photosynthesis is an ongoing quest in plant biology. Although limited information is available, such as the light dependence of C_4 enzyme expression (Chollet et al., 1996), the transcription factors mediating the abundant, cell-specific expression patterns remain unknown.

Recent work demonstrates that the maize leaf displays a gradient with regard to proteins (Majeran et al., 2010) and that large-scale transcriptional changes between four leaf areas can be detected (Li et al., 2010). In this work, we set out to generate a comprehensive systems level picture of the changes in metabolite, enzyme activity, and transcript amounts occurring along a developmental gradient of a growing maize leaf. Using this systems biology data set, we addressed the questions of (1) how photosynthesis is organized along the developmental gradient of the light-exposed leaf with special regard to the presence of C_3 photosynthesis, (2) whether the biochemistry of the C_4 cycle changes along this developmental gradient, and (3) which regulatory modules define the developmental progression in the gradient.

RESULTS AND DISCUSSION

Organization of the Light-Exposed Third Maize Leaf

The transcript and metabolite amounts, as well as protein and chlorophyll contents, displayed characteristic and continuous changes along a tip-to-base gradient of the light-exposed part of the third leaf of maize (Figures 1A and 1B). The relative expression or metabolite contents were most distinct at the distal parts of the leaf compared with relatively minor changes in the center of the leaf. A principal components analysis of transcript and metabolite amounts along the leaf gradient demonstrated clear separation of the leaf slices. The principal components determining this pattern were the distance from the leaf base (component 1) and the distance from the leaf center (component 2). The complete data set is available in readable form as Supplemental Data Set 1 online.

If the leaf was divided from top to bottom into slices, with slice 1 being the tip, gene expression patterns reflecting biochemical pathways could be followed through the development of the leaf (Figure 2A). Since previous work has demonstrated good correlation between transcript and protein abundance in maize (Li et al., 2010), we took transcript amounts as proxies for the corresponding protein amounts. Steady state amounts of transcripts encoding the classical NADP-ME C_4 proteins PEPC, PPDK, and NADP-ME were low toward the leaf base and increased until they reached a maximum around slice 2 or 3 for PEPC and slice 10 for the decarboxylation enzymes (Figure 2A). Transcripts representing subunits of photosystems I and II and of the rPPP had a similar pattern, but their increase was much less pronounced than that of the C_4 transcripts. The pattern of the photorespiratory transcripts mirrored that of the photosystems and of the rPPP (Figure 2A). No peak of photorespiratory transcripts was observed where expression of the C_4 transcripts was low. Photosynthesis, measured as carbon fixation per leaf area, steadily increased between the bottom and the top of the leaf (Figure 2B). Finally, the oxygen sensitivity of photosynthesis was measured to determine whether C_3 photosynthesis or inefficient C_4 photosynthesis would occur in the light-exposed leaf, which should be reflected by a major increase in the apparent photosynthetic rate at low oxygen partial pressure. However, the ratio of photosynthetic rates measured at high and low oxygen partial pressures did not change along the leaf gradient (Figure 2C). Maize leaves were previously hypothesized to undergo a C_3-to-C_4 transition. That is, the program initiating C_4 photosynthesis was proposed being switched on in a particular region of the leaf
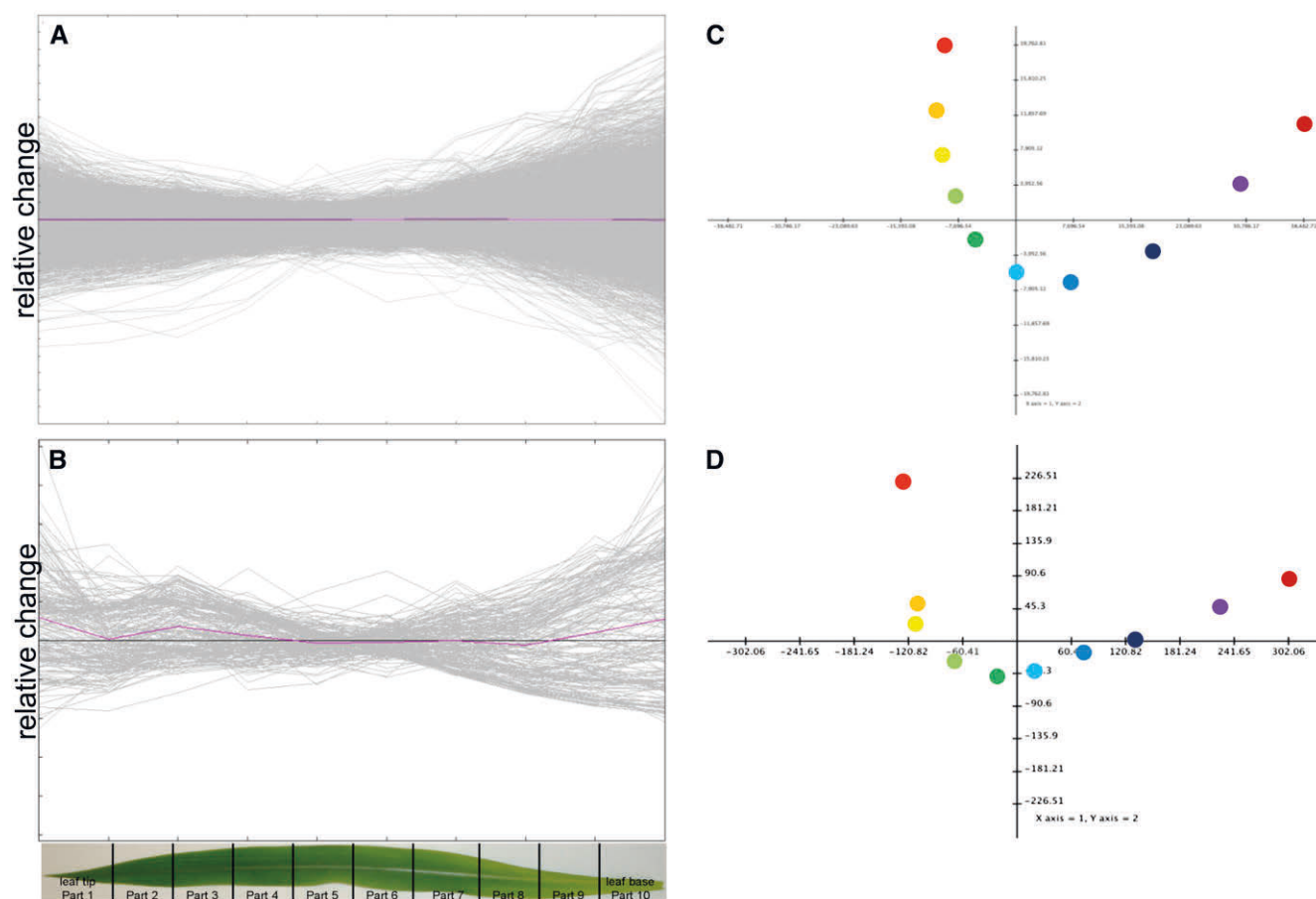
Figure 1. Relative Transcript and Metabolite Levels Are Organized along the Developmental Gradient of the Leaf.

Relative transcript abundance **(A)**, relative metabolite abundance **(B)**, and principal component analysis **(C)** of transcript levels. The first two components explain 83.5% of the variation **(D)** principal component of the metabolite levels. The first two components explain 85.5% of the variation.

(summarized in Nelson and Langdale, 1992). This switch, if existent, was an important target for understanding C_4 genesis and thus replicating it in making C_4 rice. Our systems-level analysis does not support this hypothesis: Photorespiratory transcripts do not peak in the presumed area of C_3-ness (Figure 2A,) and there is no evidence for C_3 photosynthesis or leaky C_4 photosynthesis, as oxygen sensitivity of photosynthesis did not change along the leaf gradient (Figure 2C). We thus conclude that the maize leaf undergoes a gradual sink-to-source transition without a distinct intermediary C_3 phase.

Metabolite Clusters

Although the leaf did not contain a zone of C_3 photosynthesis, it clearly displayed a gradient along its length (Figure 1; Majeran et al., 2010). To investigate the nature of the gradient in detail, extractable metabolites were analyzed by clustering algorithms. For K-means clustering, a figure of merit analysis determined five clusters as the best compromise between cluster formation with limited information loss (Friedman and Stuetzle, 1981) (see Supplemental Figure 1 online). Metabolites in the pyruvate clus-

ter with 23 members, cluster 1, were low at the bottom of the leaf and increased until slice 3 where the increase leveled off (Figure 3A). Cluster 2, with 15 members, contained metabolites that were high at the very bottom and at the tip, while cluster 3, with 10 members, contained metabolites that were level until the middle of the leaf and then increased toward the tip. The building block cluster, cluster 4, was the largest cluster with 60 members. The metabolites in this cluster started high at the bottom and decreased toward the middle of leaf from where the level stabilized. Cluster 5 was the malate cluster whose metabolites had the highest level between slice 3 and slice 7 and lower levels at the tip and the bottom (Figure 3A; condensed list of metabolites in Table 1).

Apart from pyruvate, cluster 1 contained Ala, glycerate, Glu, and citrulline. Five carotenoids, α-tocopherol, glycerol, and Gal of the lipid fraction as well as digalactosylglycerol and 3-O-galactoglycerolipids, four fatty acids, and five other metabolites were members of cluster 1 (Table 1). Surprisingly, the C_4 acids formed a distinct cluster, the malate cluster 5, with Asp, fumarate, citrate, glyoxylate, γ-tocopherol, 3-O-galactosylglycerol, and three other metabolites. Since the C_3 and C_4 acids clearly separated into
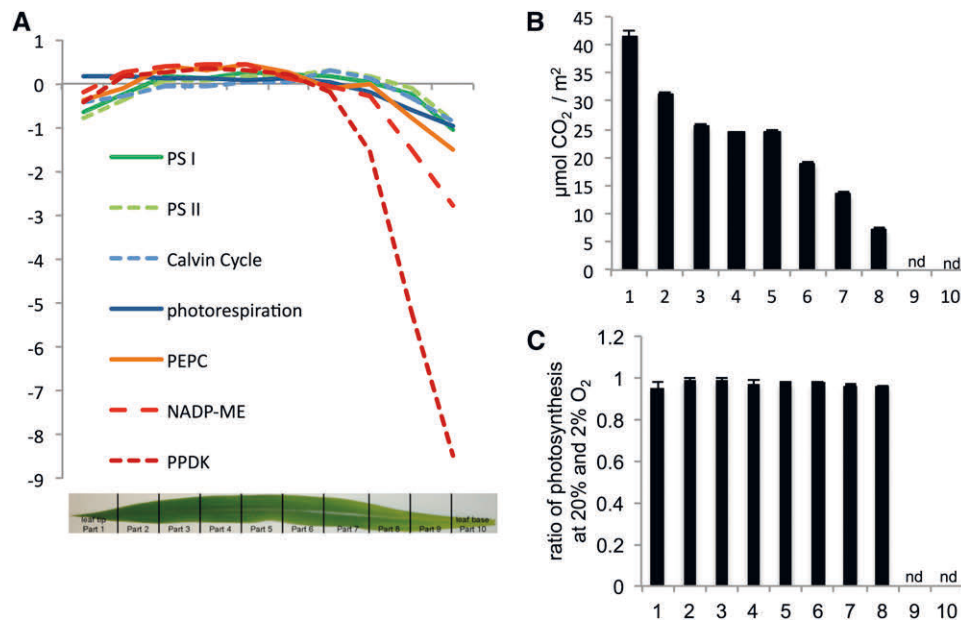
Figure 2. Photosynthetic Transitions in the Maize Leaf.

(A) Average relative expression levels for the transcripts encoding photosystem I, photosystem II, the RPPP and three key C_4 proteins.

(B) Photosynthetic rate along the light-exposed leaf. Error bars depict SD of three biological replicates. nd, not determined since not exposed to light.

(C) Ratio of photosynthesis at 20% and 2% O_2 concentration along the leaf gradient. Error bars depict SD; nd, not determined since not exposed to light, two technical replicates.

distinct clusters, there must be a major shift in the C_4 cycle. One likely explanation was the observed gradient in PEPC activity. At the point where PEPC activity started to decrease in slice 3 (Figure 2A), the C_4 acid pool sizes also sloped downwards. This may indicate that the balance between C_3 and C_4 acid pools shifted toward C_3 acids because carboxylation activity decreased while the sum of decarboxylation activities increased relative to each other (Figure 2A). Some of the pool sizes of tricarboxylic acid (TCA) cycle acids connected to malate also shifted with alterations in PEPC activity. The malate pattern extended to fumarate, citrate, and isocitrate but not to succinate and 2-oxoglutarate. The pool sizes of TCA cycle intermediates were thus only partially isolated from the C_4 cycle. The remaining metabolites shadow the buildup of the chloroplasts and thylakoids, including their pigments. The galactolipids, which dominate the chloroplast membranes (Dörmann and Benning, 2002), were mostly members of the pyruvate cluster, as were the accessory pigments carotenoids that can diffuse excess light energy via the xanthophyll cycle (Bilger and Bjorkman, 1990).

Cluster 2 contained Man, galactitol, diethylene glycol, salicylic acid, five fatty acids, and six other metabolites, while the raffinose cluster 3 contains raffinose, stachyose, galactinol, and seven other metabolites (Table 1). Metabolites from both clusters elevated toward the tip, although metabolites in cluster 2 also elevated at the very bottom (Figure 3A). The abundance of metabolites of the raffinose family from cluster 2 and cluster 3 (myo-inositol-2-P, raffinose, stachyose, and galactinol) pointed toward a drought response (Seki et al., 2007) at the tip, but, strikingly, cluster 2 and cluster 3 did not include amino acids

such as Pro. The dry weight–to–fresh weight ratio increased toward the tip, with only half the water content at the tip (Figure 3B). However, the third maize leaf analyzed in this study did not show any apparent signs of cell death at the tip (see Supplemental Figures 2A and 2D online). We thus hypothesize that despite the low water content at the tip (Figure 3B), the accumulation of compatible solutes at the tip allows photosynthesis to operate efficiently (Figure 2B). On mature field-grown maize plants, the majority of leaf tips are completely dry with only dead cells remaining. We hypothesize that maize leaves undergo a constitutive innate drought response toward the tip of each leaf to continue photosynthesizing (Figure 2B) until water content gets too low to maintain metabolism and cells undergo cell death. Considering the parallel venation pattern of grasses, any drought stress will likely initially manifest in the leaf tip. In leaf tips of the third maize leaf, chlorophyll content was already reduced (Figure 3C); however, the tissue was likely not senescent since the protein content was high (Figure 3D), photosynthesis was highly efficient (Figure 2B), and senescence markers were not highly expressed (see Supplemental Data Set 1 online).

Cluster 4 was termed the building block cluster. It contained 15 proteinogenic amino acids but not Asp, Ala, and Glu, which were part of the pyruvate and malate clusters. In addition to the amino acids, four precursors (shikimate, quinate, homoserine, and S-adenosylhomoserine) were part of this cluster. The major sugars Glc and Fru as well as the minor sugars Rib and Fru had elevated amounts at the leaf base. Ten sphingolipids, four sterols, and six fatty acids were part of cluster 4. Finally, coumaric and ferulic acid, isopentenylpyrophosphate, glucosephosphates, free
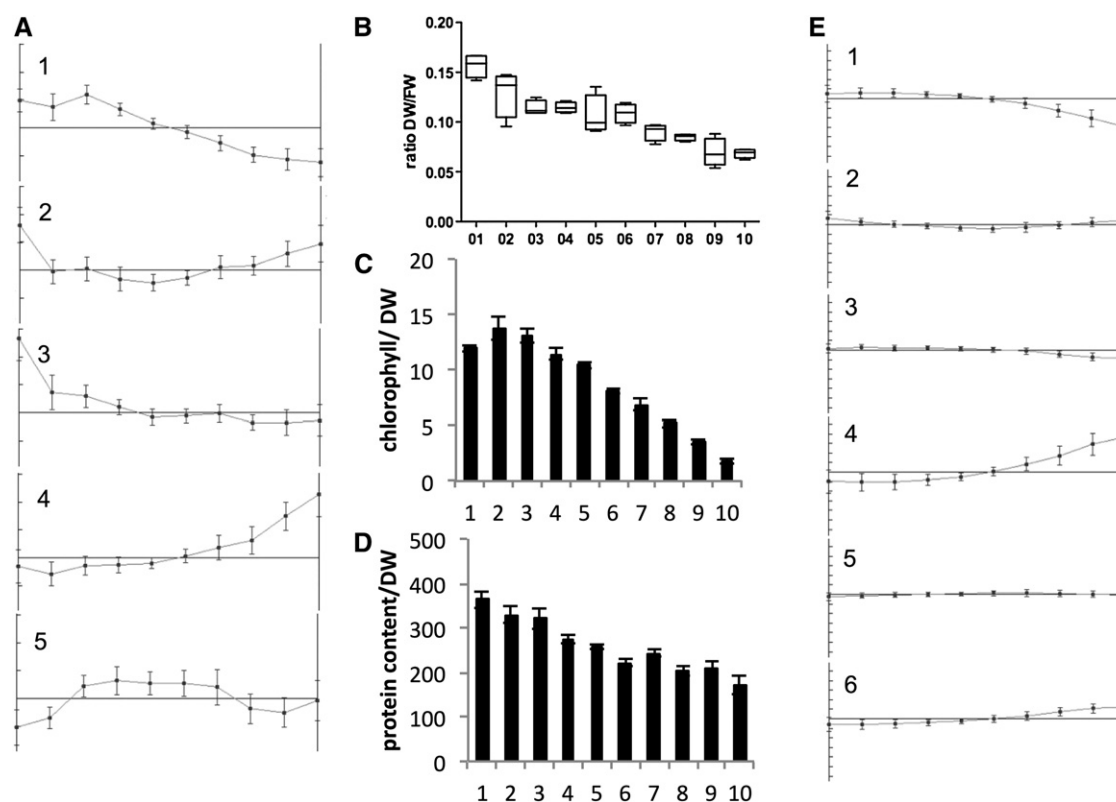
Figure 3. Changes along the Leaf Gradient.

(A) *K*-means clusters of metabolites. Cluster 1 is the pyruvate cluster with 23 members; cluster 2 contains 15 metabolites; cluster 3 is the raffinose cluster with 10 members; cluster 4 is the building block with 60 members; cluster 5 is the malate cluster with 10 members.
(B) The fresh weight (FW)–to–DW ratio indicating a low water content at the leaf tip.
(C) The chlorophyll content.
(D) and **(E)** The protein content **(D)** and *K*-means clustering of transcripts **(E)**. Error bars indicate SD of four biological replicates.

phosphate, and 12 other metabolites finished the cluster (Table 1). The leaf base represents a sink tissue (Evert et al., 1996) with minimal photosynthetic activity (Figure 2B). Chloroplasts began to develop at the leaf base (Evert et al., 1996) and chlorophyll content increased (Figure 3C). Genes encoding components of the rPPP and the electron transfer chain were highly expressed. Consequently, proteinogenic amino acids were in high demand and thus present in large amounts. The major sugars likely reflected transferred carbon, while the minor sugars and lignin precursors pointed to active cell wall synthesis. Membrane buildup was in process. Transcript analysis of four distinct zones of the maize leaf found increased transcript amounts for cell wall, lipids, secondary metabolism, and chloroplast targeting for the area at the bottom of the gradient (Li et al., 2010), thereby corroborating the analysis of metabolites. Proteins involved in lipid synthesis also peak toward the base of the gradient (Majeran et al., 2010).

In summary, four clusters defined the leaf gradient, the building block cluster defined by elevated metabolites at the bottom end of the gradient, which was followed by the C_4 malate cluster with increased C_4 acids and TCA cycle intermediates, which in turn was followed by the C_4 pyruvate cluster with high C_3 acids, carotenoids, and galactolipids. The tip of the leaf contained

elevated amounts of drought-related metabolites of the raffinose family, which are included in clusters 2 and 3 (Figure 3E).

Modules of the Leaf: Transcripts

The changes along the maize leaf were recently investigated using four leaf segments sampled from different parts of the leaf (Li et al., 2010). Li et al. identified 938 transcription factors that showed a differential expression pattern between at least two of the segments. In our study, we followed three different strategies to assess the dynamics and extent of the reprogramming of the transcriptome along the maize leaf developmental gradient: (1) *K*-means clustering to identify patterns of expression along the leaf, (2) hierarchical clustering to identify transcripts with similar patterns as the C_4 transcripts, and (3) a comparison of this leaf gradient with previously published data from four distinct leaf segments (Li et al., 2010).

The *K*-means clustering was prefaced by a figure of merit analysis, which prompted us to choose six clusters as a good solution (see Supplemental Figure 3 online). Four distinct patterns were evident in the clusters: Clusters 1 and 3 contained transcripts that are either very low (1450) or low (8521) toward the

Table 1. Condensed List of Metabolites within Each Cluster

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Pyruvate	Man	Raffinose	Coumaric acid	Asp
Ala	Galactitol	Galactinol	Ferulic acid	Malate
Glycerate	Diethylene glycol	Stachyose	Phosphate	Fumarate
Glu	Salicylic acid	Tryptamine	Glucosephosphates	Citrate (additional: isocitrate)
Citrulline	Ribonic acid	Trp	Four sugars	Glyoxylate
Galactose, lipid fraction	Cys	α-Ketoglutarate	15 proteinogenic amino acids	γ-Tocopherol
Glycerol, lipid fraction	NAD	Nicotinamide	10 sphingolipids	Threonic acid
Digalactosylglycerol	Five fatty acids	Gluconic acid	Four sterols	3-O-galactosylglycerol
3-O-galactoglycerolipids	Three other metabolites	Myristic acid	Six fatty acids	Two other metabolites
Five carotenoids		UDP-glucose	Four precursors	
Four fatty acids			12 other metabolites	
Five other metabolites				

Metabolites were K-means clustered. A figure of merit analysis determined five clusters as the best compromise between cluster formation with limited information loss (Friedman and Stuetzle, 1981).

bottom of the gradient, clusters 4 and 6 contained transcripts that are either very high (1067) or high (4287) at the bottom, cluster 2 transcripts (4459) were high at the bottom and the top, while the largest cluster 5 (10,935) displayed little change (Figure 3E). The rate of change in transcript abundance (Figures 1A and 3E) agrees with those published earlier for differences between noncontinuous leaf segments (Li et al., 2010) and is comparatively modest, especially for the average changes of each cluster (Figure 3E). Our continuous gradient revealed that virtually all changes in transcript abundance changes were gradual along the gradient. That is, no binary switches, which would manifest in extreme changes (Figure 1A), and consequentially no regulons of genes with sudden onset (Figure 3E) could be detected. Transcripts with changes in the transition zone (Li et al., 2010) corresponding to slice 10 in this study showed steady declines or increases throughout the remainder of the gradient rather than an on–off behavior. These patterns indicated that it is highly likely that the transcriptional changes and functional changes were set up by a morphogradient along the leaf, which may be defined either by one or several metabolites (such as those in the pyruvate cluster) or one or more transcripts (Figure 3E). Although light is necessary for transcription of C_4 genes (Langdale et al., 1988), neither the emergence of the leaf into the light at the border between slices 9 and 10 (see Supplemental Figure 2A online) nor the beginning of a planar leaf surface at slices 7 and 8 (see Supplemental Figure 2D online) lead to marked changes in gene expression in the leaf (Figures 1A and 3E). Thus, light was a necessary (Langdale et al., 1988) but not a sufficient cue to alter the gene expression program abruptly, since its availability to the leaf did not cause marked changes. The C_4-related transcripts were members of clusters 1 and 3. In comparison to the metabolite clusters, no cluster resembling the malate cluster, which is elevated in the middle of the gradient, was detected. Thus, the factor determining the metabolite accumulation pattern was likely of posttranscriptional nature.

A large number of transcripts of both transcription factors and other functions increased or decreased consistently along the gradient. To narrow down consistent changes, the continuous gradient was compared with the noncontinuous segmental results from Li et al. (2010). Three patterns were originally defined, high at tip and low at the bottom, a group high in the transition zone and low at the tip and very bottom, and a group high at the bottom and low at the tip. Since the bottom of the continuous gradient corresponded to the transition zone in the earlier experiment, the second and third groups merged into one for the purpose of the comparison.

Of the group that was low at the tip, 632 of 725 (87%) were detected reliably on the microarray and 529 (73%) were significantly changed along the gradient based on analysis of variance. Of the significantly changed ones, 276 (52%) decreased toward the tip. For 11%, various expression patterns were detected, while 37% showed a pattern opposite to expectations (Figure 4A).

Of the group that was high at the tip, 203 of 213 (95%) were detected reliably on the microarray and 186 (92%) were significantly changed along the gradient based on analysis of variance. Of those significantly changed transcripts, 68 (37%) continuously increased in expression toward the leaf tip. The plurality, 52%, did increase in expression but dipped slightly at the tip similar to chlorophyll content and PEPC activity (Figure 4B). This change in pattern was visible only with a continuous gradient and cannot be detected with segmental analysis. Only 10% showed patterns that were not congruous with earlier data. If the morphogradient was set up by transcripts that reflected positional information, only 68 transcription factors increasing in expression would be on the short list of candidate transcription factors at or near the core of the morphogradient. By contrast, 276 transcription factors decrease more or less continuously. Adding this second analysis reduced the list of potential transcription factor from 938 in the earlier study down to 344 in our work. Additional analyses have the potential to reduce the list to the point where single-gene functional analyses become feasible. Three important pieces of information are missing: (1) Which of these factors, if any, display a similar gradient in older and bigger leaves, (2) does this gradual behavior extend throughout the leaf to the point of emergence from the apical meristem, and (3) which factors have a similar gradient in other grass species? In older maize leaves of 40-cm length, enzyme activity measurements clearly show gradients for the C_4 marker enzymes (see Supplemental Figure 4 online), which are similar to those in the
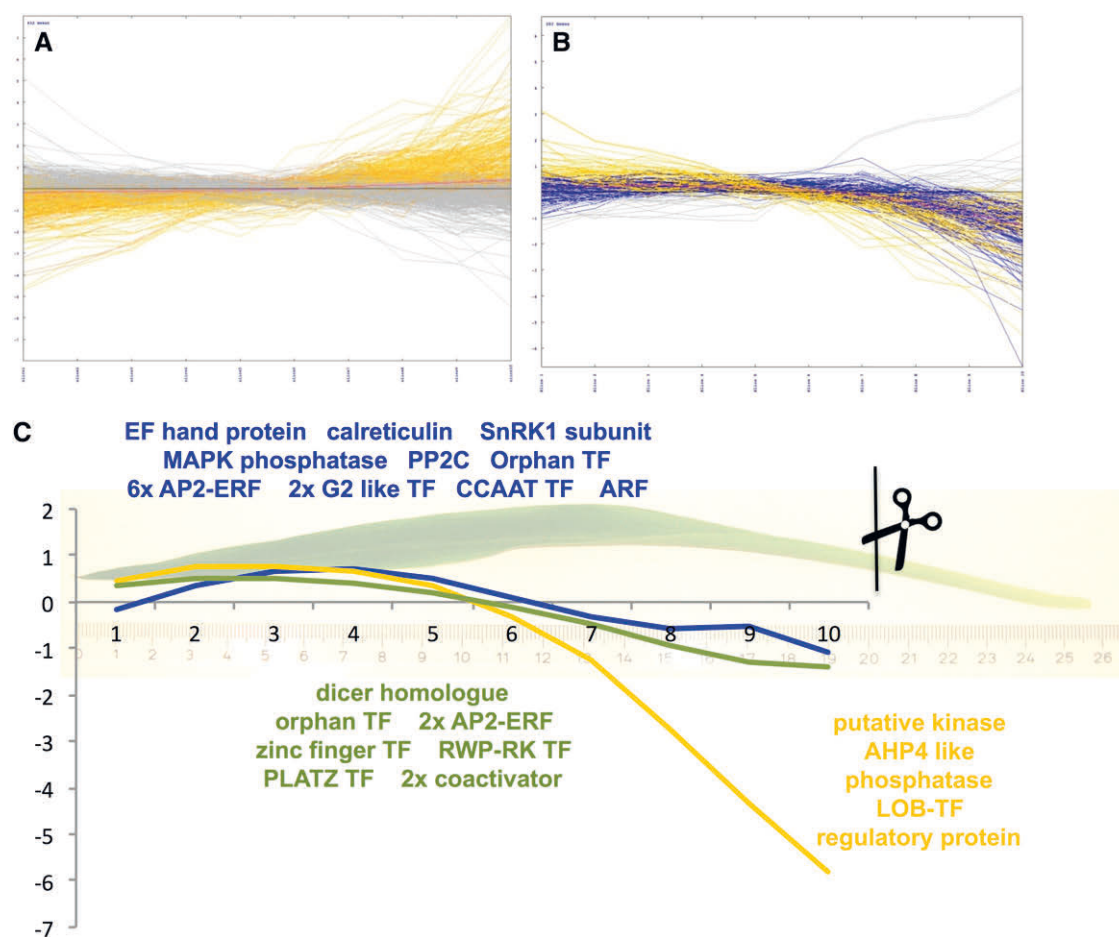
Figure 4. Targeted Expression Analysis of Regulatory Functions.

(A) and **(B)** Expression pattern of transcripts detected as low at the tip **(A)** and high at the tip **(B)** in a previous analysis. Patterns in orange confirm the expectation based on Li et al. (2010), patterns in blue partially confirm, and patterns in gray have different patterns.
(C) Transcripts coexpressed with PEPC (blue), NADP-ME (green), and PPDK (yellow) major isoforms.

young leaves. This at least indicated that older, more mature leaves still display a gradient.

Systematic analyses of dicot C_4 species showed that all C_4 enzyme activities except for malate dehydrogenase are, at least to some degree, regulated at the transcriptional level (Bräutigam et al., 2011; Gowik et al., 2011). Even if the C_4-related transcripts are piggybacking on the developmental gradient, the direct regulators of their transcription would be expected to be coexpressed with or just predating their targets. We identified transcripts encoding putative regulators that were tightly coexpressed with the major C_4 transcripts. The transcript for the major isoform of PEPC accumulated slowly throughout the gradient, reached a plateau between slices 3 and 5, and dipped toward the tip (Figure 4C). In the hierarchical clustering (see Supplemental Data Set 2 online, readable with MeV, www.tm4.org/mev), 16 transcripts representing regulatory functions were identified (see Supplemental Table 1 online). Two of these transcripts related to calcium signaling, one EF hand protein, and a calreticulin. An SnRK1 subunit implicated in sugar and nitrogen signaling (Rolland et al.,

2006) had the same pattern as *PEPC*. In addition, a mitogen-activated protein kinase phosphatase and *PP2C*, which is involved in abscisic acid (ABA) and drought signaling in *Arabidopsis thaliana* (Kuhn et al., 2006), were coregulated with *PEPC*. Finally, 13 transcription factors, one orphan, four APETALA2 (AP2)-ETHYLENE RESPONSE FACTORS (ERFs), two G2-like myb transcription factors, one Auxin Response Factor (ARF), and one CCAAT-type transcription factor tightly correlate with *PEPC* throughout the gradient. The *Arabidopsis* homologs of these transcription factors are involved in ABA signal transduction and ethylene signal transduction (see Supplemental Table 1 online). Possibly, an ABA and/or ethylene-driven regulon was used in evolution of C_4 photosynthesis. Neither DOF1 nor DOF2, which are known to bind the *PEPC* promoter region (Yanagisawa and Sheen, 1998), are tightly coexpressed with *PEPC*. For maize nuclear factor and PEP-I, no sequences were deposited at the National Center for Biotechnology Information (NCBI); hence, they could not be compared with the current data. Notably, none of the tightly correlated transcripts are known to be involved in light signaling, underscoring that light

is necessary but not sufficient to drive expression (see above). The two major *NADP-ME* isoforms showed the same pattern as *PEPC* up to slice 3 but lacked the dip at the tip. Only nine transcripts encoding regulatory functions tightly correlate (see Supplemental Table 1 online): a dicer homolog, one orphan transcription factor, two AP2-ERFs, a zinc-finger transcription factor, a PLATZ transcription factor, and two coactivators. The *PPDK* transcript behaved quite differently; it accumulated from very low levels toward slice 5 and then mirrored *NADP-ME*. Comparatively few regulatory transcripts mirror this more extreme pattern (see Supplemental Table 1 online): a kinase, one phosphatase, a phosphorelay transmitter similar to AHP4 of *Arabidopsis*, a LOB-type transcription factor, and a regulatory protein similar to a flowering regulator from *Arabidopsis*. Selected transcript abundance patterns were confirmed by quantitative RT-PCR (see Supplemental Figure 5 online). The differences in pattern between the three key C_4 transcripts pointed to the fact that a simple generic C_4 regulon may not exist. Rather, additional data sets taken during leaf development of maize and other grasses will increase the resolution of covariation analyses and lead to the identification of the leaf morphogradients and ultimately of regulons that induce expression of the separate C_4 genes.

C_4 Photosynthesis along the Developmental Gradient of the Leaf

It was recently proposed that C_4 plants undergo changes in their mode of C_4 photosynthesis based on developmental stage and in response to environmental cues (Furbank, 2011). The well-defined maize leaf developmental gradient analyzed in our study represented a unique opportunity to test this hypothesis.

The classical C_4 genes *PEPC*, *PPDK*, and *NADP-ME* were identified from the literature, and their abundance and expression pattern was used to identify transcripts with similar abundance and pattern. The C_4 genes were among the transcripts that occupy more than one per thousand of the total (see Supplemental Data Set 3 online). Aside from *PEPC*, *PPDK*, and *NADP-ME*, the list of abundant transcripts contained mainly transcripts that encode the chloroplast electron transfer chain and the rPPP (see Supplemental Data Set 3 online). Surprisingly, genes for a plastid-localized Asp aminotransferase (AspAT) and PEP-CK were also members of this group of 147 transcripts. Coexpressed transcripts frequently act in the same or in connected pathways (Eisen et al., 1998; Reumann and Weber, 2006). Hence, all transcripts were clustered to identify transcripts that are coexpressed with known C_4 transcripts. The known C_4 transcripts are low toward the bottom of the leaf and increase toward the tip. The very tip portion is slightly lower in expression compared with middle of the leaf blade (Figure 2A). If transcripts, which might be active in any of the C_4 types, were plotted, one Ala aminotransferase (AlaAT) and a plastidic AspAT as well as a PEP-CK would display a comparable pattern (see Supplemental Figure 6 online). Taken together with the observation that in maize, 25% of the carbon label initially was located in Asp (Hatch, 1971) and the observation that Asp is a carbon donor to the bundle sheath (Chapman and Hatch, 1981), we decided to

investigate the seemingly simple C_4 cycle of maize at the levels of transcripts, metabolites, and enzyme activity.

We initiated the analysis by testing whether the leaf had reached C_4 configuration at the point where the analysis commenced. In slice 10, the bottom of the gradient, the leaf was already differentiated into a vein, bundle sheath, mesophyll, mesophyll, bundle sheath, vein configuration (see Supplemental Figure 2 online). PEPC activity increased from the bottom, reached a maximum at slice 3, and decreased only slightly toward the tip of the leaf. Maximal activity was 18 milli units/mg dry weight (DW). The major decarboxylation enzyme NADP-ME increased from the bottom toward the top and reached its maximal activity at the leaf tip with close to 15 mU/mg DW. Both AspAT and AlaAT had similar patterns compared with NADP-ME and reached activities of 25 mU/mg and 15 mU/mg DW. Although PEP-CK activity was only a quarter of NADP-ME, it had a comparable pattern and reached up to 4 mU/mg DW (Figure 5A). The pattern for all enzymes except PEPC was similar; that is, the activity was low at the bottom of the leaf and increased toward the tip.

On the basis of transcript abundance, the major enzyme isoform of PEPC mirrored the pattern of PEPC activity in the leaf and also peaked around slice 3 (Figure 5B). Transcripts of the major isoforms of NADP-ME, PEP-CK, AspAT, AlaAT, and PPDK, displayed a pattern comparable to that of *PEPC* but different than the extractable activities of the enzymes. Hence, the total activity was likely composed of multiple isoforms of NADP-ME, PEP-CK, AspAT, and AlaAT and/or subject to posttranscriptional regulation. Indeed, there were other isoforms that were of appreciable transcript abundance (see Supplemental Data Set 3 online) and patterns unlike that displayed by the major isoform (Figure 5B). The enzyme activity of AspAT was sufficient to support the carboxylation and decarboxylation activity, while AlaAT fell short for the majority of the leaf. PEP-CK activity was appreciable. While certainly not the major decarboxylation activity, its activity was high enough to catalyze at least one-fifth of the decarboxylation reactions. This was almost certainly an underestimation since PEP-CK was assayed in the unfavorable reverse reaction (Ashton et al., 1990). If amino acids carried part of the carbon flow in the C_4 cycle, their abundance should mirror that of the canonical C_4 cycle acids malate and pyruvate. Ala and Asp mirrored the accumulation pattern of pyruvate and malate, respectively (Figure 5C). In addition, not only their pattern but also their absolute abundance should be comparable to that of malate. The absolute abundance of Asp and Ala were about one-fourth of the abundance of malate (see Supplemental Figure 7 online).

Taken together, these results suggested a revised model of the C_4 cycle in maize (Figure 5D): After PEP is carboxylated to OAA, it is moved to the chloroplast, either in exchange with malate through DiT1 (Kinoshita et al., 2011) or in exchange with Asp through DiT2 (Renne et al., 2003), which are produced by malate dehydrogenase and AspAT, respectively, in the chloroplast. The major AspAT in maize is predicted to be chloroplast localized. Labeling experiments by Hatch (1971) indicated that as much as 25% of the carbon initially labels Asp, not malate. Both C_4 acids diffuse to the bundle sheath, reducing the necessary mass flow compared with either C_4 acid carrying the full load. In the bundle
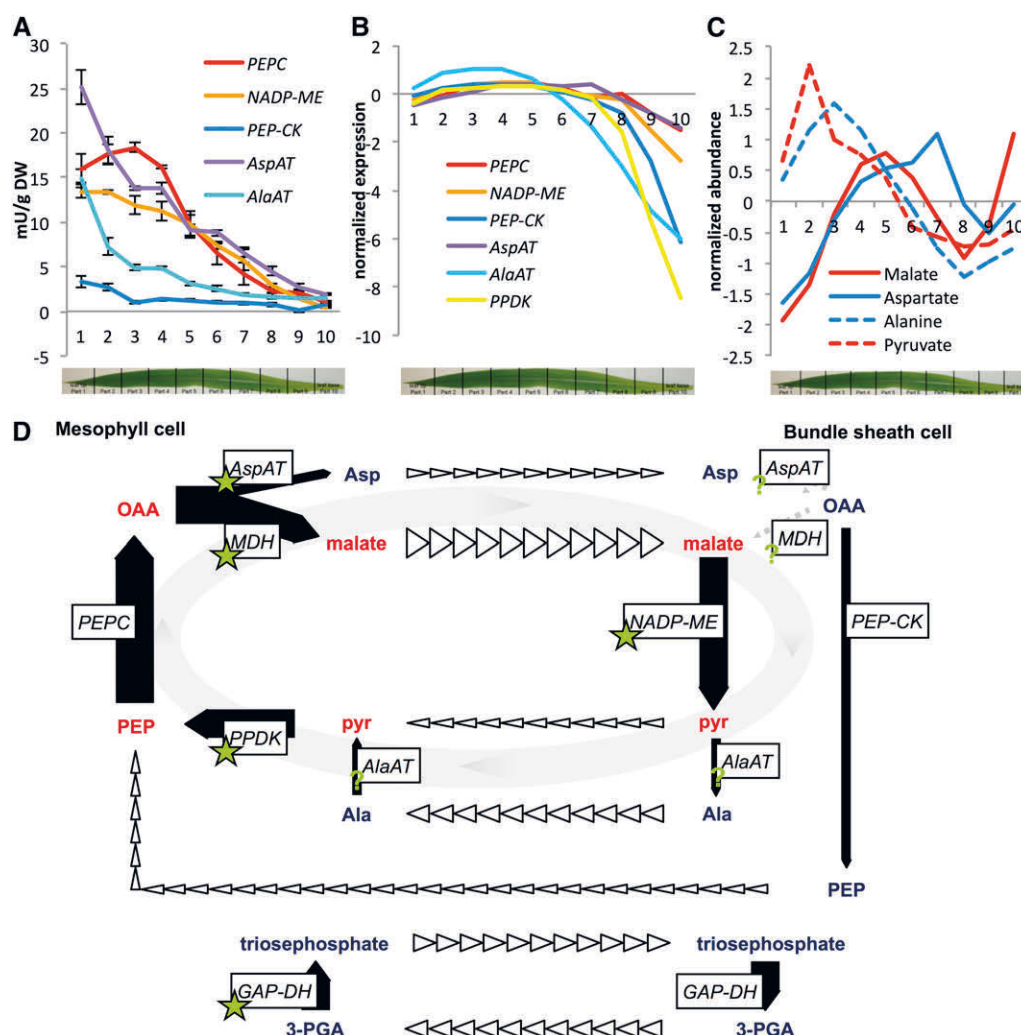
Figure 5. Selected Enzyme Activities and Metabolite Contents along a Maize Leaf.

(A) Enzyme activities. Error bars indicate SD of four biological replicates.
(B) Expression pattern of the major isoform for each enzyme.
(C) Normalized abundance of four C_4 cycle metabolites, with absolute values in slice three: malate, 13.5 mmol/mg DW; Asp,4 mmol/mg DW; Ala, 4 mmol/mg DW.
(D) Model of the C_4 cycle in maize. Arrow widths equal approximate flows, bold arrows represent enzyme activities, open arrowheads indicate transport between the cells, and green stars denote plastid-localized steps. pyr, pyruvate; 3-PGA, 3-phosphoglycerate; MDH, malate dehydrogenase; GAP-DH, glycerinaldehyde-3-phosphate dehydrogenase.

sheath, malate is taken up by a currently unknown mechanism into the chloroplast and decarboxylated. Asp may have two fates: It may be transaminated to OAA and decarboxylated by PEP-CK to PEP, or it may enter the chloroplast by an as yet unknown mechanism and be transformed via OAA to malate to serve as the substrate for NADP-ME. It has long been assumed that malate transfer is preferable to Asp transfer since malate carries a reducing equivalent while Asp does not. However, it has been shown that triosephosphate translocator is one of the most abundant chloroplast envelope proteins (Bräutigam et al., 2008) and that the reduction of 3-phosphoglycerate is almost entirely located in the mesophyll (Majeran et al., 2005), making the

generation of reducing equivalents unnecessary in the bundle sheath. The regeneration of the carbon acceptor and its transfer to the mesophyll may be dissected by analysis of metabolite compartmentation, flux, and gradients and may occur as PEP, pyruvate, or Ala (Figure 5D).

We propose that independent of environmental or developmental cues, the core C_4 cycle in maize is set up already as a branched rather than a linear cycle. In addition to the scheme presented (Figure 5D), the branched core C_4 cycle is also connected to basal metabolism (for example, see Leegood and von Caemmerer, 1988). A distribution of carbon between two C_4 acids and three C_3 acids reduces the diffusion requirements for

any one molecule between mesophyll and bundle sheath. This distribution becomes especially important considering that distribution by diffusion is by no means proven (Sowínski et al., 2008; Bräutigam and Weber, 2011a). It remains to be investigated whether the distribution of carbon to malate and Asp is fixed at 3:1 as reported (Hatch, 1971) or whether this ratio is adjusted by the plant during its life cycle (Furbank, 2011). Within the age gradient in a single leaf, there is no evidence in the enzyme activities, transcript abundance, or metabolite accumulation pattern to suggest that operation of the cycle switches from one transfer acid to another (Figure 5). The presence of higher PEP-CK activity in older maize plants with older leaves (Wingler et al., 1999), however, points to a developmental regulation between leaves rather than within a leaf, similar to what has been recently observed in the dicotyledonous C_4 plant *Cleome gynandra* (Sommer et al., 2012). Environmental adaptation of Asp metabolism based on N availability in maize leaves with regard to pool size and turnover has also been demonstrated (Khamis et al., 1992). Hence, the C_4 cycle is apparently quite flexible.

Conclusion

On the basis of a comprehensive systems biology data set, we conclude that C_4 photosynthesis is established from sink tissue without an intermediate phase of C_3 or C_2 photosynthesis. That is, the likely evolutionary events are not recapitulated during ontogeny. No binary on–off switches were detected within the leaf gradient, pointing to gradual onset of features and, therefore, morphogradients as the determinants for leaf development. Finally, the biochemistry of C_4 photosynthesis is more complex than anticipated but stays constant throughout the leaf.

METHODS

Plant Growth and Harvest

Maize (*Zea mays*) plants of the ecotype B73 were grown in the greenhouse for 14 to 15 d in clay pots in Floraton soil. Natural light, a shading system, and artificial light were used to extend the daylight period to 16 h at a photon flux density of $\sim$500 μmol m^{-2} s^{-1}. The humidity in the greenhouse was between 75 and 90%. The greenhouse's ventilation system kept the temperature at 24°C.

The third leaf was harvested at 18-cm length measured from tip to emergence from the stem. Leaves were harvested by placing them atop a custom-made leaf guillotine where they were snap frozen (see Supplemental Figure 8 online). By closing the lid, the leaf is cut into 10 pieces of 2-cm width each, the last of which had not yet emerged (see Supplemental Figure 2 online). Twenty plants were pooled for each biological replicate. The sections were ground to a fine powder in a porcelain mortar cooled with liquid nitrogen. Frozen powder was used for enzyme assays and metabolomics and transcriptomics analysis. For the DW–to–fresh weight ratio, 30 mg of ground and frozen plant material was dried in a vacuum dryer overnight, and the weight was recorded before and after drying. Δ-^{13}C values were determined according to Coplen et al. (2006). Chlorophyll was determined according to Porra et al. (1989), and protein content was measured with the BCA method (Thermo Fisher Scientific). Enzymatic activities were determined as summarized by Ashton et al. (1990). Photosynthetic rate was measured with a LI-6400XT portable photosynthesis analyzer (LI-COR Environmental) under greenhouse con-

ditions at the time of sampling for the invasive experiments. Oxygen sensitivity of photosynthesis was measured according to Dai et al. (1996). Oxygen partial pressure was controlled by a custom-built gas exchange system. Four biological replicates were measured in all analyses except where otherwise noted.

Metabolite Profiling

Lyophilized tissue equivalent to 200 mg of fresh weight was used for metabolite profiling. Metabolites were extracted with the use of accelerated solvent extraction with polar (methanol + water, 80 + 20 by volume) and nonpolar (methanol + dichlormethan, 40 + 60 by volume) solvents. Subsequent analyses of metabolites by gas chromatography–mass spectrometry (GC-MS) were performed as described elsewhere (Roessner et al., 2000; Walk et al., 2007). In addition, liquid chromatography–tandem mass spectrometry (Niessen, 2003) analyses were performed with the use of an Agilent 1100 capillary LC system (Agilent Technologies) coupled with an Applied Biosystems/MDS SCIEX API 4000 triple quadrupole mass spectrometer (AB Sciex). After reverse-phase HPLC separation, detection and quantification of metabolites were performed in the multiple reaction monitoring and full scan mode (Gergov et al., 2003). Absolute Ala, Asp, and malate contents were estimated by GC-MS (Fiehn et al., 2000), which included an external complex standard and were quantified by coupled enzymatic assays (Bergmeyer, 1974).

Transcript Profiling

The mRNA was isolated after the method of Logemann et al. (1987) and from the same plant material in which the enzyme activities and metabolites were measured. The isolated RNA was purified with the RNeasy purification kit according to the manufacturer's instructions (Qiagen). The quality was checked with the Agilent 2100 Bioanalyzer using the RNA 6000 Nano kit. The cDNA and following antisense cRNA synthesis was performed according the one-color microarray-based gene expression analysis protocol (Agilent Technologies). An aliquot of 1.65 μg of this RNA was loaded on one-color microarrays with custom-designed oligonucleotide probes (Agilent 025271). Transcripts were normalized to the 75th percentile within each array using the Agilent Gene spring program. Arrays can be accessed under submission number GSE33861 in the NCBI Gene Expression Omnibus database. Quantitative RT-PCR was performed with three biological replicates using the SYBR-green technique (MESA GREEN qPCR MasterMix Plus; Eurogentec) and gene-specific primers (see Supplemental Table 2 online) as described by Schmittgen and Livak (2008). Relative expression values were calculated with the $2^{-\Delta\Delta ct}$ (cycle threshold) method after Pfaffl (2001) with threshold values normalized to expression of 18S rRNA.

Data Analysis

For data analysis, the maize transcript list was downloaded from www.maizesequence.org. For each transcript, a best BLAST hit was produced with *Sorghum bicolor* and *Arabidopsis thaliana* as databases (Altschul et al., 1997). Gene Ontology terms were added based on the *S. bicolor* annotation. Information about putative and known transcription factors (Pérez-Rodríguez et al., 2010) and transport proteins (http://membranetransport.org/) were added based on the *Arabidopsis* annotation. A Mapman annotation was downloaded from (http://mapman.gabipd.org; Thimm et al., 2004). Protein localization was predicted based on amino acid sequence (Emanuelsson et al., 2000). For each maize transcript, an annotation was created based on the *Arabidopsis* TAIR10 description (Swarbreck et al., 2008) and, if not available, manually added based on the *Sorghum* data. Transcripts without known or predicted functions were labeled POUF (for protein of unknown function). Based on all information, transcripts were grouped into classes in a

hierarchical manner. *Arabidopsis* information was given precedence over other information given that *Arabidopsis* annotations are currently the best within the plant genomes. Group and functional assignments throughout the publication are based on this annotation table. The complete annotation table, including all raw data, can be accessed as Supplemental Data Set 1 online. The major isoform of C$_4$ enzymes were determined by read mapping of raw data from Li et al. (2010) on the maize transcriptome since analyzed data were not included in the original publication. Read mappings were normalized to reads per million without any further correction factors applied. The data are included in Supplemental Data Set 1 online.

All large-scale data analyses were performed with the MultiExperiment Viewer (http://www.tm4.org/mev/; Saeed et al., 2003). Average metabolite contents were expressed as z-scores (the number of standard deviations the value is different from the mean of all values), resulting in mean centered values. Only metabolites detectable in all biological replicates of eight or more slices were analyzed. Transcripts were normalized to the 75th percentile within each array; the mean of replicates was calculated for each slice, followed by mean centering along each row. For *K*-means cluster analysis, the ideal number of clusters was determined by figure of merit analysis as implemented in MeV (Saeed et al., 2003). Metabolites and transcripts were clustered by Euclidian average linkage clustering and visualized in MeV. For comparison with the Li et al. (2010) data set, transcription factors of different groups were extracted from Li et al. (2010) supplemental data and visualized in MeV (see Supplemental Data Sets 4 and 5 online). Transcripts coexpressed with major C$_4$ enzymes were determined by hierarchical clustering followed by list extraction from MeV. All raw data, including the MeV readable files, are provided as supplemental material accompanying the publication (see Supplemental Data Sets 1 and 2 online).

Accession Numbers

Microarray data from this article can be found in the NCBI Gene Expression Omnibus database under accession number GSE33861.

Supplemental Data

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

Supplemental Figure 1. Figure of Merit Analysis of Metabolite Clustering.

Supplemental Figure 2. Configuration of the Leaf Anatomy along the Stem.

Supplemental Figure 3. Figure of Merit Analysis of Transcript Clustering.

Supplemental Figure 4. Enzyme Activity for Three C$_4$ Marker Enzymes in Maize Leaves of 40-cm Length.

Supplemental Figure 5. qRT-PCR Results for Selected Regulatory Transcripts.

Supplemental Figure 6. Expression Pattern for Enzymes Likely Involved in C$_4$ Photosynthesis.

Supplemental Figure 7. Absolute Concentrations of Malate, Aspartate, and Alanine along the Leaf Gradient.

Supplemental Figure 8. The Guillotine Used for Sampling the Gradient.

Supplemental Table 1. List of Maize Identifiers of Regulatory Transcripts Coregulated with PEPC, PPDK, or NADP-ME.

Supplemental Table 2. Primers Used for qRT-PCR.

Supplemental Data Set 1. The Complete Data Set in Human-Readable Form.

Supplemental Data Set 2. MeV Readable Hierarchical Clustering of All Data to Identify Transcripts Coregulated with Major C$_4$ Enzymes.

Supplemental Data Set 3. Maize Transcripts with the Highest Number of Read Mappings Based on SRR039507 and SRR039508 Originally Published by Li et al. (2010).

Supplemental Data Set 4. MeV Readable Microarray Data for G1 and G2 Upregulated Transcription Factors Published as a Supplemental File by Li et al. (2010).

Supplemental Data Set 5. MeV Readable Microarray Data for G3 Upregulated Transcription Factors Published as a Supplemental File by Li et al. (2010)

AUTHOR CONTRIBUTIONS

T.R.P. sampled the gradient, measured all parameters except transcripts and relative metabolite content, analyzed the data, and cowrote the article. A.B. designed the research, produced the custom annotation, analyzed data, and wrote the article. U. Schlüter measured transcripts. A.K.D. measured absolute metabolite contents and photosynthesis rates. C.C. and U. Scholz provided tables for the custom annotation. H.F. measured relative metabolite contents. R.P. and U.R. supported the oxygen sensitivity measurements. U. Sonnewald designed the research and analyzed data. A.P.M.W. designed the research, analyzed data, and cowrote the article.

REFERENCES

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J.H., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. **25:** 3389–3402.

Ashton, A.R., Burnell, J.N., Furbank, R.T., Jenkins, C.L.D., and Hatch, M.D. (1990). The enzymes in C4 photosynthesis. In Enzymes of Primary Metabolism. Methods in Plant Biochemistry, P.M. Dey and J.B. Harborne, eds (London: Academic Press), pp. 39–72.

Bergmeyer, H.U. (1974). Methoden der Enzymatischen Analyse, 3rd ed. (Weinheim/Bergstrasse, Germany: Verlag Chemie).

Bilger, W., and Bjorkman, O. (1990). Role of the xanthophyll cycle in photoprotection elucidated by measurements of light-induced absorbancy changes, fluorescence and photosynthesis in leaves of *Hedera canariensis*. Photosynth. Res. **25:** 173–185.

Bräutigam, A., Hoffmann-Benning, S., and Weber, A.P.M. (2008). Comparative proteomics of chloroplast envelopes from C3 and C4 plants reveals specific adaptations of the plastid envelope to C4 photosynthesis and candidate proteins required for maintaining C4 metabolite fluxes. Plant Physiol. **148:** 568–579. Erratum. Plant Physiol. **148:** 1734.

Bräutigam, A., et al. (2011). An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species. Plant Physiol. **155:** 142–156.

Bräutigam, A., and Weber, A.P.M. (2011a). Transport processes – Connecting the reactions of C4 photosynthesis. In Advances in Photosynthesis and Respiration, 1st ed, Vol. 32, A.S. Raghavendra and R.F. Sage, eds (Dordrecht, The Netherlands: Springer), pp. 199–219.

Bräutigam, A., and Weber, A.P.M. (2011b). Do metabolite transport processes limit photosynthesis? Plant Physiol. **155:** 43–48.

Chapman, K.S.R., and Hatch, M.D. (1981). Aspartate decarboylation in bundle sheath cells of *Zea mays* and its possible contribution to C4 photosynthesis. Aust. J. Plant Physiol. **8:** 237–248.

Chollet, R., Vidal, J., and O'Leary, M.H. (1996). Phosphoenolpyruvate carboxylase: A ubiquitous, highly regulated enzyme in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. **47:** 273–298.

Coplen, T.B., Brand, W.A., Gehre, M., Gröning, M., Meijer, H.A.J., Toman, B., and Verkouteren, R.M. (2006). New guidelines for delta13C measurements. Anal. Chem. **78:** 2439–2441.

Dai, Z., Ku, M.S.B., and Edwards, G.E. (1996). Oxygen sensitivity of photosynthesis and photorespiration in different photosynthetic types in the genus Flaveria. Planta **198:** 563–571.

Dörmann, P., and Benning, C. (2002). Galactolipids rule in seed plants. Trends Plant Sci. **7:** 112–118.

Edwards, E.J., and Smith, S.A. (2010). Phylogenetic analyses reveal the shady history of C4 grasses. Proc. Natl. Acad. Sci. USA **107:** 2532–2537.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA **95:** 14863–14868.

Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J. Mol. Biol. **300:** 1005–1016.

Evert, R.F., Russin, W.A., and Bosabalidis, A.M. (1996). Anatomical and ultrastructural changes associated with sink-to-source transition in developing maize leaves. Int. J. Plant Sci. **157:** 247–261.

Fiehn, O., Kopka, J., Dörmann, P., Altmann, T., Trethewey, R.N., and Willmitzer, L. (2000). Metabolite profiling for plant functional genomics. Nat. Biotechnol. **18:** 1157–1161.

Friedman, J.H., and Stuetzle, W. (1981). Projection pursuit regression. J. Am. Stat. Assoc. **76:** 817–823.

Furbank, R.T. (2011). Evolution of the C(4) photosynthetic mechanism: Are there really three C(4) acid decarboxylation types? J. Exp. Bot. **62:** 3103–3108.

Gergov, M., Ojanperä, I., and Vuori, E. (2003). Simultaneous screening for 238 drugs in blood by liquid chromatography-ion spray tandem mass spectrometry with multiple-reaction monitoring. J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. **795:** 41–53.

Gowik, U., Bräutigam, A., Weber, K.L., Weber, A.P.M., and Westhoff, P. (2011). Evolution of C4 photosynthesis in the genus flaveria: How many and which genes does it take to make C4? Plant Cell **23:** 2087–2105.

Haeckel, E. (1866). Generelle Morphologie. I: Allgemeine Anatomie der Organismen. II: Allgemeine Entwicklungsgeschichte der Organismen. (Berlin: Verlag Georg Reimer).

Hatch, M.D. (1971). The C4 pathway of photosynthesis. Evidence for an intermediate pool of carbon dioxide and the identity of the donor C4 dicarboxylic acid. Biochem. J. **125:** 425–432.

Hatch, M.D. (1987). C-4 photosynthesis - A unique blend of modified biochemistry, anatomy and ultrastructure. Biochim. Biophys. Acta **895:** 81–106.

Hibberd, J.M., Sheehy, J.E., and Langdale, J.A. (2008). Using C4 photosynthesis to increase the yield of rice-rationale and feasibility. Curr. Opin. Plant Biol. **11:** 228–231.

Khamis, S., Lamaze, T., and Farineau, J. (1992). Effect of nitrate limitation on the photosynthetically active pools of aspartate and malate in maize, a NADP malic enzyme C4 plant. Physiol. Plant. **85:** 223–229.

Kinoshita, H., Nagasaki, J., Yoshikawa, N., Yamamoto, A., Takito, S., Kawasaki, M., Sugiyama, T., Miyake, H., Weber, A.P.M., and Taniguchi, M. (2011). The chloroplastic 2-oxoglutarate/malate transporter has dual function as the malate valve and in carbon/nitrogen metabolism. Plant J. **65:** 15–26.

Kuhn, J.M., Boisson-Dernier, A., Dizon, M.B., Maktabi, M.H., and Schroeder, J.I. (2006). The protein phosphatase AtPP2CA negatively regulates abscisic acid signal transduction in Arabidopsis, and effects of abh1 on AtPP2CA mRNA. Plant Physiol. **140:** 127–139.

Langdale, J.A., Zelitch, I., Miller, E., and Nelson, T. (1988). Cell position and light influence C4 versus C3 patterns of photosynthetic gene expression in maize. EMBO J. **7:** 3643–3651.

Leegood, R.C., and von Caemmerer, S. (1988). The relationship between contents of photosynthetic metabolites and the rate of photosynthetic carbon assimilation in leaves of *Amaranthus edulis L*. Planta **174:** 253–262.

Li, P.H., et al. (2010). The developmental dynamics of the maize leaf transcriptome. Nat. Genet. **42:** 1060–1067.

Logemann, J., Schell, J., and Willmitzer, L. (1987). Improved method for the isolation of RNA from plant tissues. Anal. Biochem. **163:** 16–20.

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

Majeran, W., Friso, G., Ponnala, L., Connolly, B., Huang, M.S., Reidel, E., Zhang, C.K., Asakura, Y., Bhuiyan, N.H., Sun, Q., Turgeon, R., and van Wijk, K.J. (2010). Structural and metabolic transitions of C4 leaf development and differentiation defined by microscopy and quantitative proteomics in maize. Plant Cell **22:** 3509–3542.

Nelson, T., and Langdale, J.A. (1992). Developmental genetics of C-4 photosynthesis. Annu. Rev. Plant Physiol. Plant Mol. Biol. **43:** 25–47.

Niessen, W.M. (2003). Progress in liquid chromatography-mass spectrometry instrumentation and its impact on high-throughput screening. J. Chromatogr. A **1000:** 413–436.

Pérez-Rodríguez, P., Riaño-Pachón, D.M., Corrêa, L.G.G., Rensing, S.A., Kersten, B., and Mueller-Roeber, B. (2010). PlnTFDB: Updated content and new features of the plant transcription factor database. Nucleic Acids Res. **38**(Database issue): D822–D827.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. **29:** e45.

Porra, R.J., Thompson, W.A., and Kriedemann, P.E. (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophyll a and chlorophyll b extracted with 4 different solvents: Verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. Biochim. Biophys. Acta **975:** 384–394.

Renné, P., Dreßen, U., Hebbeker, U., Hille, D., Flügge, U.I., Westhoff, P., and Weber, A.P.M. (2003). The Arabidopsis mutant dct is deficient in the plastidic glutamate/malate translocator DiT2. Plant J. **35:** 316–331.

Reumann, S., and Weber, A.P.M. (2006). Plant peroxisomes respire in the light: some gaps of the photorespiratory C2 cycle have become filled—others remain. Biochim. Biophys. Acta **1763:** 1496–1510.

Roessner, U., Wagner, C., Kopka, J., Trethewey, R.N., and Willmitzer, L. (2000). Technical advance: Simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. Plant J. **23:** 131–142.

Rolland, F., Baena-Gonzalez, E., and Sheen, J. (2006). Sugar sensing and signaling in plants: Conserved and novel mechanisms. Annu. Rev. Plant Biol. **57:** 675–709.

Saeed, A.I., et al. (2003). TM4: A free, open-source system for microarray data management and analysis. Biotechniques **34:** 374–378.

Sage, R.F. (2004). The evolution of C-4 photosynthesis. New Phytol. **161:** 341–370.

Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. Nat. Protoc. **3:** 1101–1108.

Seki, M., Umezawa, T., Urano, K., and Shinozaki, K. (2007). Regulatory metabolic networks in drought stress responses. Curr. Opin. Plant Biol. **10:** 296–302.

Sommer, M., Bräutigam, A., and Weber, A.P.M. (2012). The dicotyledonous NAD-malic enzyme C4 plant *Cleome gynandra* displays age-dependent plasticity of C4 decarboxylation biochemistry. Plant Biol. (vol.) **14:** in press.

Sowiński, P., Szczepanik, J., and Minchin, P.E.H. (2008). On the mechanism of C4 photosynthesis intermediate exchange between Kranz mesophyll and bundle sheath cells in grasses. J. Exp. Bot. **59:** 1137–1147.

Swarbreck, D., et al. (2008). The Arabidopsis Information Resource (TAIR): Gene structure and function annotation. Nucleic Acids Res. **36** (Database issue): D1009–D1014.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P.,

Selbig, J., Müller, L.A., Rhee, S.Y., and 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.

Weber, A.P.M., and Linka, N. (2011). Connecting the plastid: Transporters of the plastid envelope and their role in linking plastidial with cytosolic metabolism. Annu. Rev. Plant Biol. **62:** 53–77.

Weber, A.P.M., and von Caemmerer, S. (2010). Plastid transport and metabolism of C3 and C4 plants—Comparative analysis and possible biotechnological exploitation. Curr. Opin. Plant Biol. **13:** 257–265.

Walk, T., et al. (2007). Means and methods for analyzing a sample by means of chromatography–mass spectrometry. International Patent WO 2007/012643.

Wingler, A., Walker, R.P., Chen, Z.H., and Leegood, R.C. (1999). Phosphoenolpyruvate carboxykinase is involved in the decarboxylation of aspartate in the bundle sheath of maize. Plant Physiol. **120:** 539–546.

Yanagisawa, S., and Sheen, J. (1998). Involvement of maize Dof zinc finger proteins in tissue-specific and light-regulated gene expression. Plant Cell **10:** 75–89.
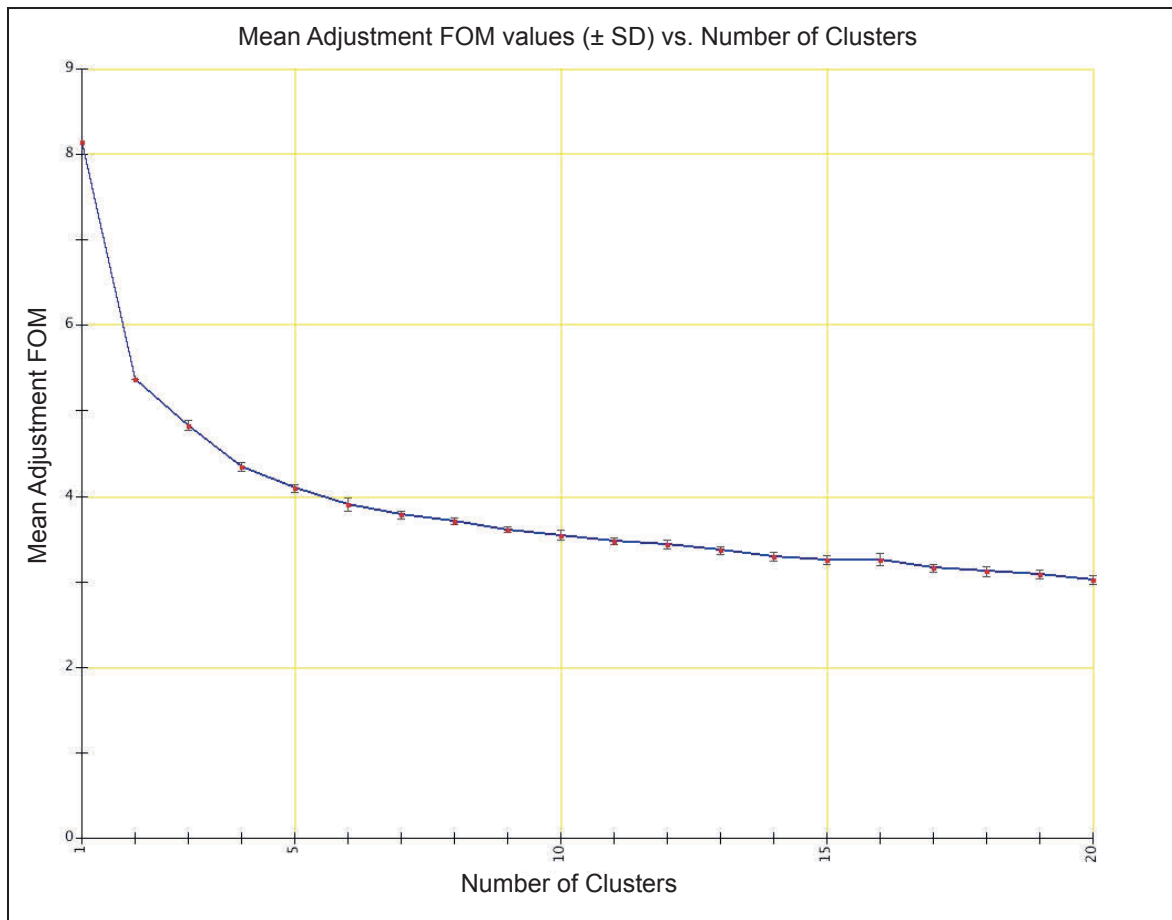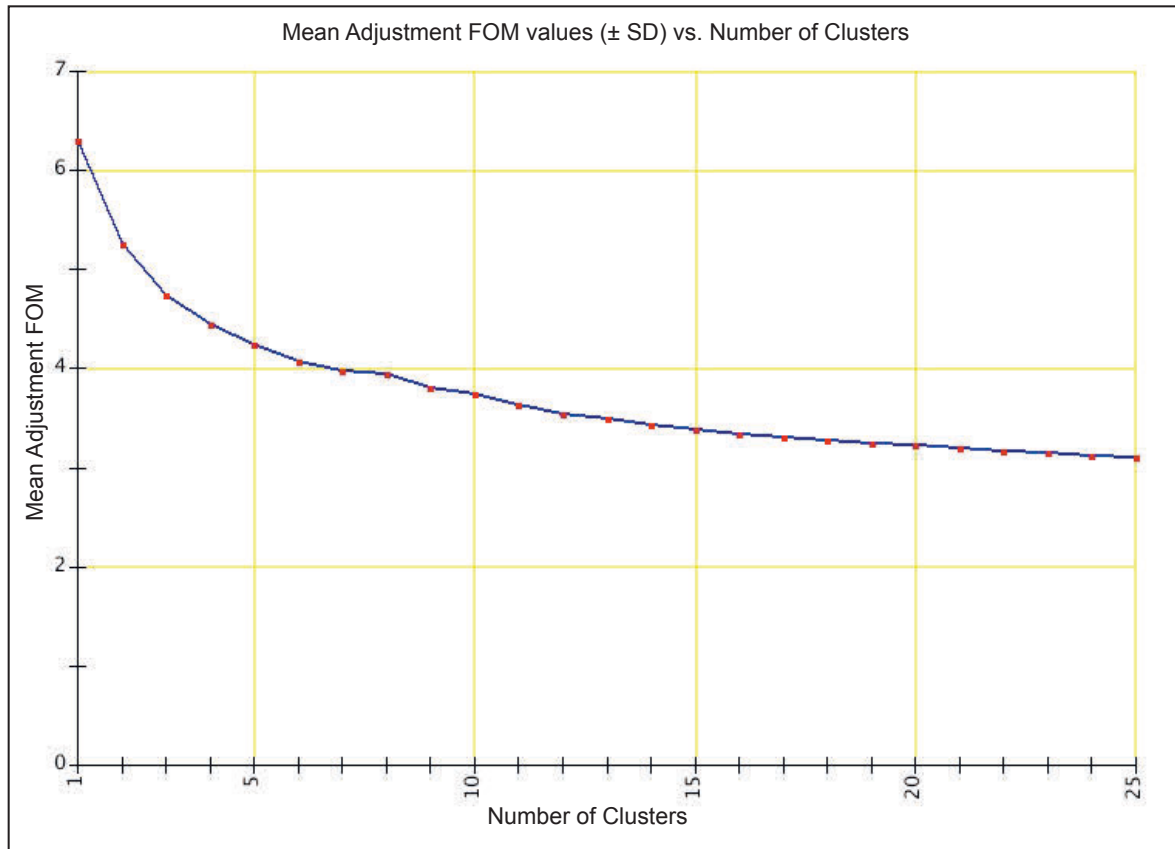
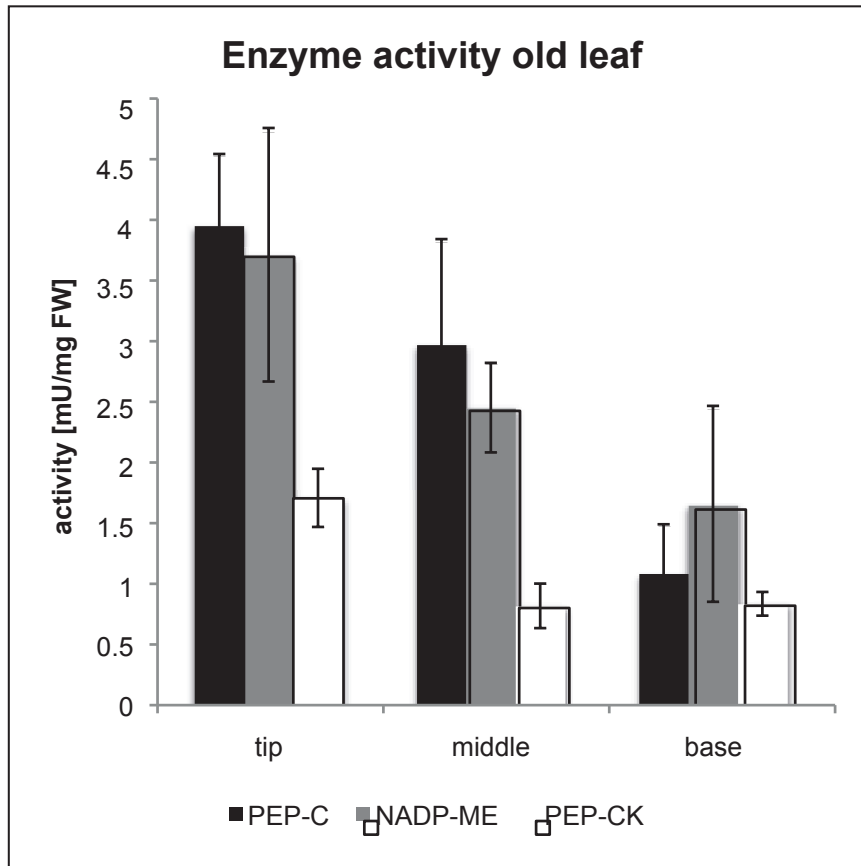**Supplemental Figure 1. Figure of merit analysis of metabolite clustering.**
The algorithm tries 20 times to cluster the data and calculates the information loss based on the cluster pattern compared with the patterns of the items in the cluster. Ideally, at some point, additional clusters no longer lead to more information in the clusters. In practice, the information gain for an additional cluster will reach a constant level. This point was chosen for k-means clustering.
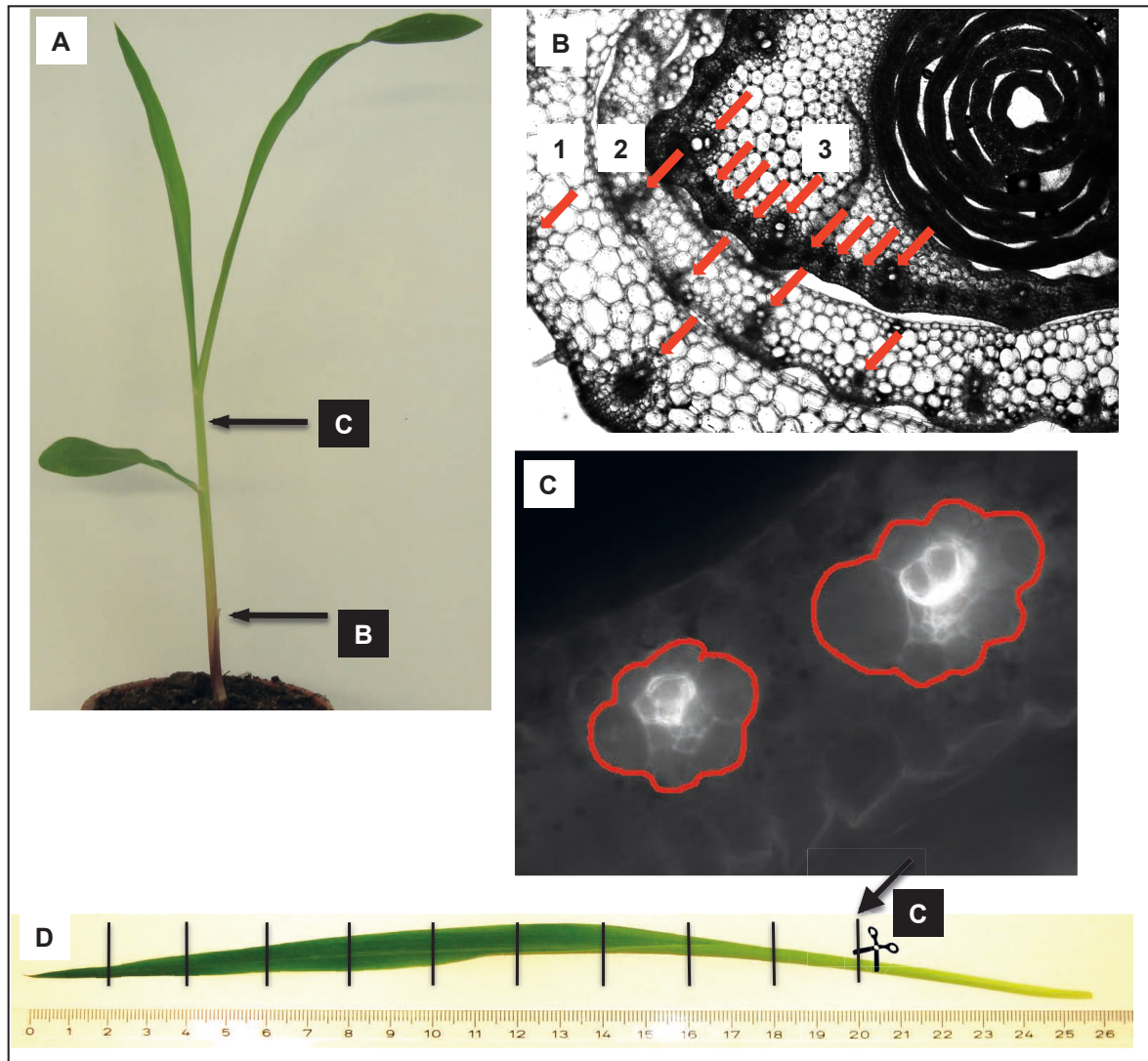
Supplemental Figure 2. Configuration of the leaf anatomy along the stem.
(A) position of cut sites marked by arrows **(B)** cross section of the maize stem; red arrows mark veins; in contrast to the first and second leaf, the third leaf is already in C_4 configuration **(C)** UV illuminated cross section of the third leaf; two bundles are surrounded in red, the distance between them allows not more than two intervening mesophyll cells; **(D)** position of the leaf slices along the third maize leaf.

Mean Adjustment FOM values (± SD) vs. Number of Clusters

Supplemental Figure 3. Figure of merit analysis of transcript clustering. The algorithm tries 20 times to cluster the data and calculates the information loss based on the cluster pattern compared with the patterns of the items in the cluster. Ideally, at some point, additional clusters no longer lead to more information in the clusters. In practice, the information gain for an additional cluster will reach a constant level. This point was chosen for k-means clustering.

Enzyme activity old leaf

Supplemental Figure 4. Enzyme activity for three C$_4$ marker enzymes in maize leaves of 40 cm length. All C$_4$ marker enzymes are more active in the tip region of the leaf compared with the point of emergence into the light. Error bars depict standard deviation; n=3.
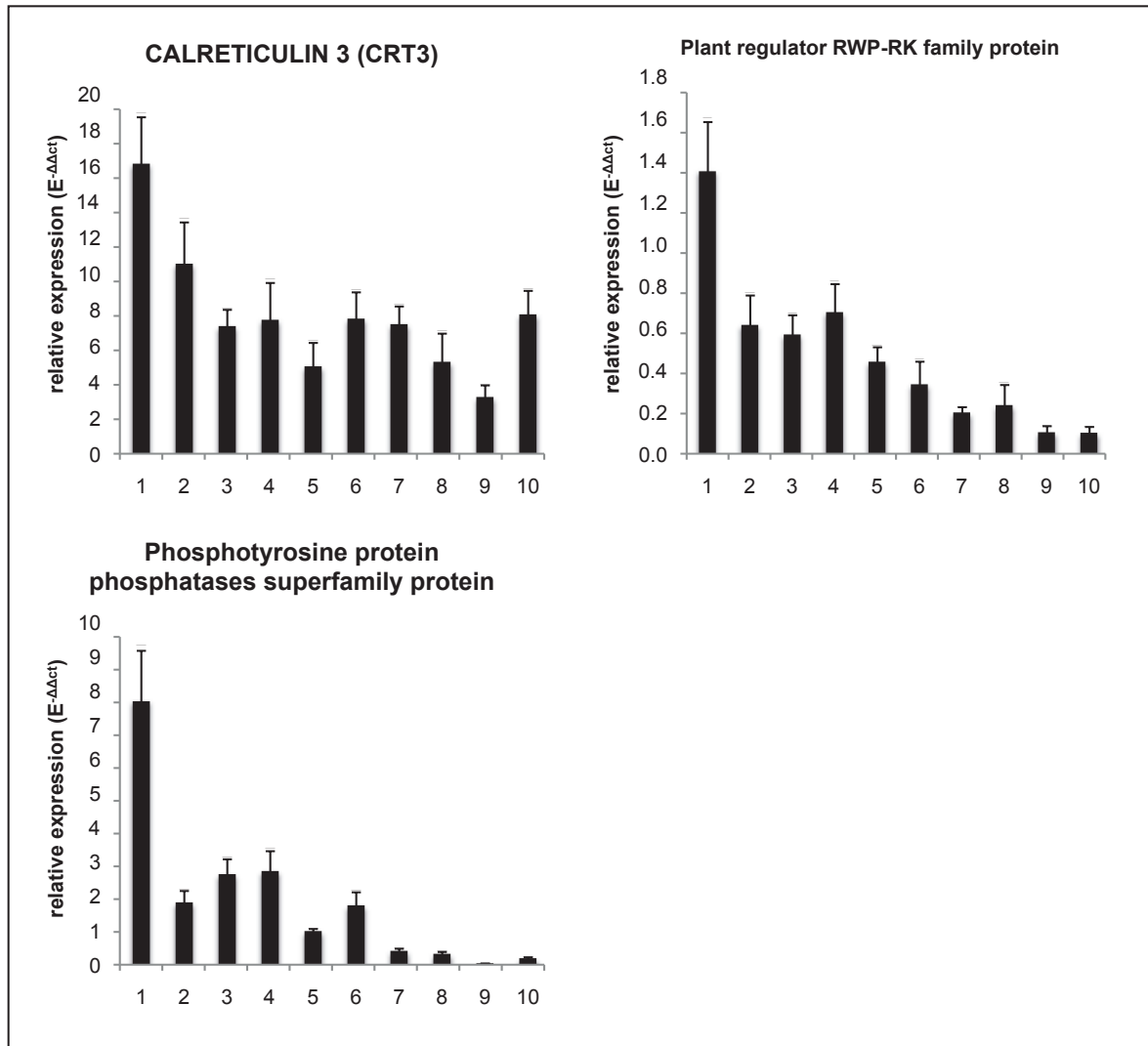
Supplemental Figure 5. qRT-PCR results for selected regulatory transcripts; 1 = leaf slice at the tip; 10 = leaf slice at the leaf base; expression values were normalized to expression levels of 18srRNA. Error bars indicate standard error of three technical replicates of three individual experiments.
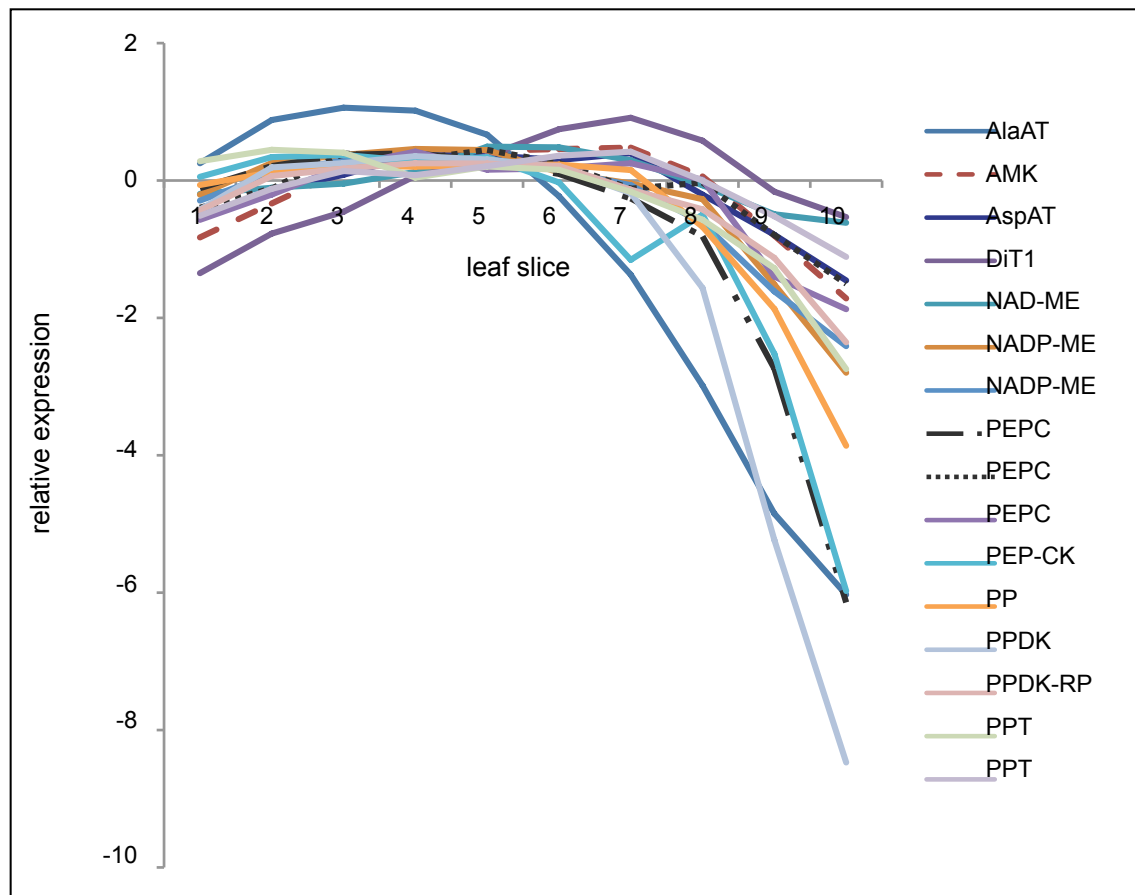
**Supplemental Figure 6. Expression pattern for enzymes likely involved in C$_4$ photosynthesis.** Candidate C$_4$ transcripts follow a 'PEPC' pattern, low at the base, peaking in the middle of the leaf, equal of dipping at the very tip. Abbreviations: AlaAT alanine aminotransferase, AMK adenosin monophosphate kinase, AspAT aspartate aminotransferase; Dit1 dicarboxylate transporter 1, NAD-ME NAD dependent malic enzyme, NADP-ME NADP dependent malic enzyme, PEPC phosphoenolpyruvate carboxylase, PEP-CK phosphoenolpyruvate carboxykinase, PP pyrophosphorylase, PPDK pyruvate phosphate dikinase, PPDK-RP regulatory protein of PPDK, PPT phosphoenolpyruvate phosphate translocator.
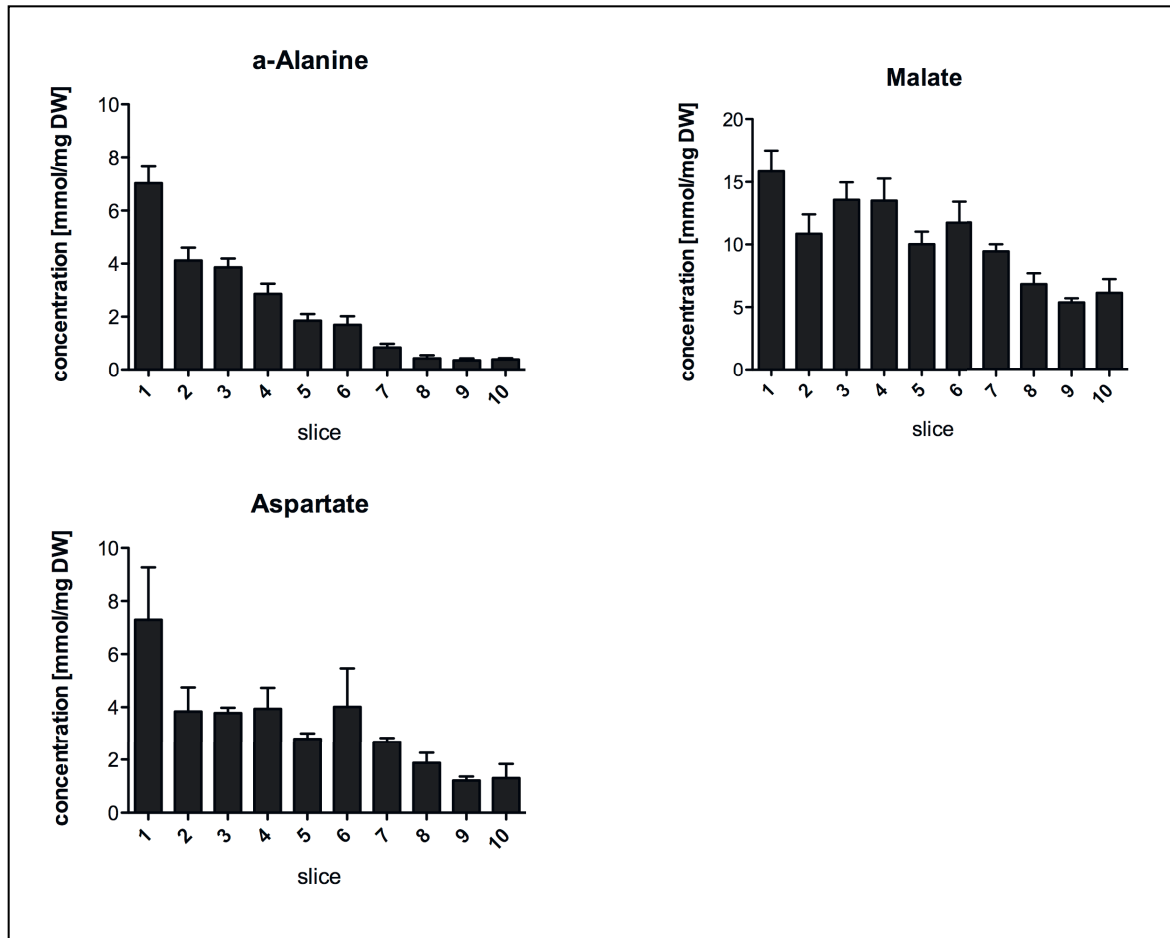
Supplemental Figure 7. Absolute concentrations of malate, aspartate and alanine along the leaf gradient (determined in a separate experiment). The absolute concentration of alanine and aspartate are in the same order of magnitude compared to that of malate. Error bars indicate standard error; n=3.
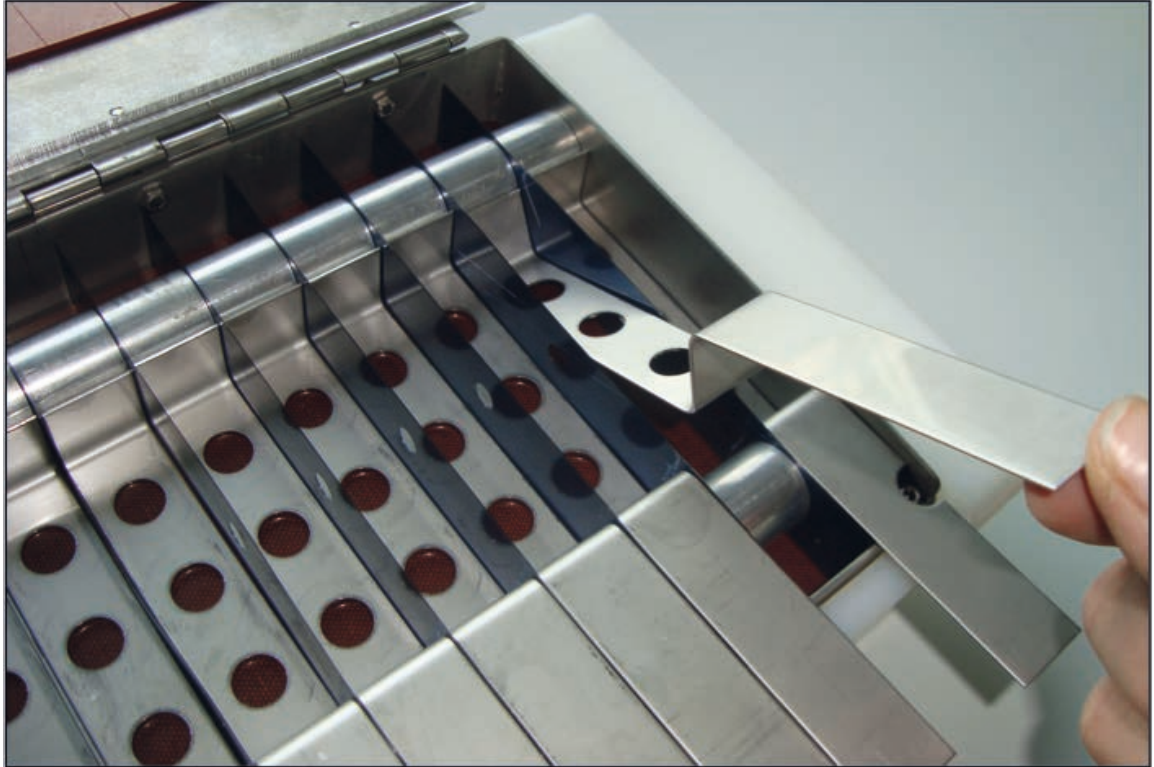
Supplemental Figure 8. The guillotine used for sampling the gradient. The guillotine is filled with liquid nitrogen. A leaf is placed on top of the open guillotine, snap frozen and cut by closing the lid. Each leaf slice falls into its own compartment and can be retrieved on its ledger.

Supplemental Table 1. List of maize identifiers of regulatory transcripts co-regulated with PEPC, PPDK and NADP-ME, respectively.

coregulated with PEPC

Parent Gene	best hit AGI	gene name of corresponding AGI (where available)	transcription factor class
GRMZM2G467184			
GRMZM2G305115	AT1G08450	CALRETICULIN 3 (CRT3)	
GRMZM2G025459	AT5G21170	(AKINBETA1)	
GRMZM2G465287	AT2G30020		
GRMZM2G177693	AT3G11410	ARABIDOPSIS THALIANA PROTEIN PHOSPHATASE 2CA (PP2CA)	
GRMZM2G062218	AT5G41380		Orphans
GRMZM2G363908	AT4G25480	DEHYDRATION RESPONSE ELEMENT B1A (DREB1A)	AP2-EREBP
GRMZM2G307152	AT1G71520		AP2-EREBP
GRMZM2G369912	AT1G53170	ETHYLENE RESPONSE FACTOR 8 (ERF8)	AP2-EREBP
GRMZM2G125460	AT2G47520	HYPOXIA RESPONSIVE ERF (ETHYLENE RESPONSE FACTOR) 2 (HRE2)	AP2-EREBP
GRMZM2G384386	AT5G19790	RELATED TO AP2 11 (RAP2.11)	AP2-EREBP
GRMZM2G382379	AT3G48590	NUCLEAR FACTOR Y, SUBUNIT C1 (NF-YC1)	CCAAT
GRMZM2G386337	AT5G64750	ABA REPRESSOR1 (ABR1)	AP2-EREBP
GRMZM2G390641	AT4G30080	AUXIN RESPONSE FACTOR 16 (ARF16)	ARF
GRMZM2G159119	AT1G25550		G2-like
GRMZM2G016370	AT1G25550		G2-like
GRMZM2G114389			
GRMZM2G009598			

coregulated with PPDK

Parent Gene	best hit AGI	gene name of corresponding AGI (where available)	transcription factor class
GRMZM2G330751	AT4G21390	(B120)	
GRMZM2G039246	AT3G16360		
GRMZM2G070315	AT1G05000		
GRMZM2G079768	AT3G13850	LOB DOMAIN-CONTAINING PROTEIN 22 (LBD22)	LOB
GRMZM2G450273	AT5G24860	FLOWERING PROMOTING FACTOR 1 (FPF1)	

coregulated with NADP-ME

Parent Gene	best hit AGI	gene name of corresponding AGI (where available)	transcription factor class
GRMZM2G040762	AT1G01040	DICER-LIKE 1 (DCL1)	
GRMZM2G346466	AT2G21320		Orphans
GRMZM2G052667	AT1G53910	RELATED TO AP2 12 (RAP2.12)	AP2-EREBP
GRMZM2G110333	AT3G14230	RELATED TO AP2 2 (RAP2.2)	AP2-EREBP
GRMZM2G092483	AT1G66140	ZINC FINGER PROTEIN 4 (ZFP4)	C2H2
GRMZM2G427618	AT4G35270		RWP-RK
GRMZM2G086403	AT4G17900		PLATZ
GRMZM2G074821	AT1G29810	CELLULOSE SYNTHASE LIKE B5	
GRMZM2G074821	AT1G29810	CELLULOSE SYNTHASE LIKE B5	

Supplemental Table 2: Primer used for qRT-PCR. All sequences in 5'-3' orientation.

	Forward Primer	Reverse Primer
18SrRNA	CTGTCGGCCAAGGCTATATACT	CAGAACATCTAAGGGCATCACAG
CALRETICULIN 3 (CRT3)	GGCACGGGAGGAAGGG	GGTAGTCGTAGTGCCTATGACGC
Phosphotyrosine protein phosphatases superfamily protein	CCTCAAGTCCCTCAACCTCC	GAGCCTGATCCCATTCTTTTC
Plant regulator RWP-RK family protein	CGCCTTCTTCGTCATCGTC	CAGGCTGTTGGCCACCA

6.1 First publication: Systems Analysis of a Maize Leaf Developmental Gradient Redefines the Current C$_4$ Model and Provides Candidates for Regulation

Status: **Published** (December 2011)

Thea R. Pick, Andrea Bräutigam, Urte Schlüter, Alisandra K. Denton, Christian Colmsee, Uwe Scholz, Holger Fahnenstich, Roland Pieruschka, Uwe Rascher, Uwe Sonnewald, and Andreas P.M. Weber

Journal: "Plant Cell"

Impact factor: 10.224

1. Co-Author

Own contribution: 75%

- Preliminary experiments
- Sampling of the gradient
- Measured all parameters except transcripts and relative metabolite content
- Absolute metabolite content
- Data analysis
- Co-writing the manuscript

Manuscript 2

PLGG1, a plastidic glycolate glycerate transporter, is required for photorespiration and defines a new class of metabolite transporters.

BIOLOGICAL SCIENCES: Plant Biology

PLGG1, a plastidic glycolate glycerate transporter, is required for photorespiration and defines a new class of metabolite transporters

Thea R. Pick*[a], Andrea Bräutigam*[a],Matthias Schulz[a], Toshihiro Obata[b], Alisdair R. Fernie[b] and Andreas P.M. Weber[a]

*both authors contributed equally to this work
a Institute of Plant Biochemistry, Heinrich Heine University Düsseldorf, Center of Excellence on Plant Sciences (CEPLAS), 40225 Duesseldorf, Germany
b Max-Planck Institute for Molecular Plant Physiology, 14476, Potsdam-Golm, Germany

Photorespiratory carbon flux reaches up to a third of photosynthetic flux and thus contributes massively to the global carbon cycle. The pathway recycles glycolate-2-phosphate, the most abundant byproduct of RubisCO reactions. This oxygenation reaction of RubisCO and subsequent photorespiration limit significantly the biomass gains of many crop plants. Although photorespiration is a compartmentalized process with enzymatic reactions in the chloroplast, the peroxisomes, the mitochondria and the cytosol, no transporter required for the core photorespiratory cycle has been identified at the molecular level to date. Using transcript co-expression analyses, we identified Plastidal glycolate glycerate translocator 1 (PLGG1) as a candidate core photorespiratory transporter. Related genes are encoded in the genomes of archaea, bacteria, fungi, and all Archaeplastida and have previously been associated with a function in programmed cell-death. A mutant deficient in PLGG1 shows WT-like growth only in an elevated carbon dioxide atmosphere. The mutant accumulates glycolate and glycerate, leading to the hypothesis that PLGG1 is a glycolate/glycerate transporter. This hypothesis was tested and supported by *in vivo* and *in vitro* transport assays and $^{18}O_2$-metabolic flux profiling. Our results indicate that PLGG1 is the chloroplastidic glycolate/glycerate transporter, which is required for the function of the photorespiratory cycle. Identification of the PLGG1 transport function will facilitate unraveling the role of similar proteins in bacteria, archaea, and fungi in the future.

Introduction

Carbon flux through photorespiration is second in magnitude only to photosynthesis and thus this metabolic pathway constitutes a major component of the global carbon cycle. Photorespiration is essential since the enzyme RubisCO, which assimilates CO_2 from the atmosphere into biomass, also catalyzes a futile reaction, the oxygenation of the CO_2 acceptor, ribulose 1,5-bisphosphate (RuBP). This latter reaction leads to the formation of the toxic metabolite 2-phosphoglycolate, which is detoxified and recycled to RuBP by a complex metabolic pathway called photorespiration. Photorespiration became an essential requirement for photosynthesis after the carbon composition of the atmosphere changed to an oxygen rich atmosphere, as a consequence of oxygenic photosynthesis by cyanobacteria, algae, and plants. Large gains in photosynthetic efficiency can be achieved if photorespiration is suppressed by enriching CO_2 in the vicinity of RubisCO. For example, plants carrying a metabolic bypass for photorespiration indeed produce more biomass (1, 2) providing a promising approach for increasing the productivity of some the most important crop plants, such as rice (*Oryza sativa*) or wheat (*Triticum aestivum*) that have to cope with high rates of photorespiration.

In plants, the photorespiratory cycle is a highly compartmentalized process with enzymatic reactions in chloroplasts, peroxisomes, and mitochondria as well as in the cytosol. In the chloroplast stroma 2-PG resulting from the RubisCO oxygenation reaction is dephosphorylated to glycolate by 2-phosphoglycolate phosphatase (PGLP). Glycolate is exported from the chloroplasts to the peroxisomes where it is oxidized to glyoxylate by glycolate oxidase (GOX) and transaminated to glycine by Ser:glyoxylate and Glu:glyoxylate aminotransferase (SGT and GGT, respectively). Glycine leaves the peroxisomes and enters the mitochondria where two molecules of glycine are deaminated and decarboxylated by the glycine decarboxylase complex (GDC) and serine hydroxymethyltransferase (SHMT) to form one molecule of each, serine, ammonia, and carbon dioxide. Serine is exported from the mitochondria to the peroxisomes where it is predominantly converted to glycerate by SGT and hydroxypyruvate reductase (HPR). Glycerate leaves the peroxisomes and is taken up into the chloroplast where it is phosphorylated by glycerate kinase (GLYK) to yield 3-PGA. In essence, one of the four carbon atoms contained in two molecules of 2-PG entering the pathway is lost as CO_2, whereas three are shuttled back into the Calvin-Benson cycle. Thus, photorespiration constitutes a metabolic repair cycle that is required to detoxify 2-PG, at the expense of energy and loss of carbon dioxide and ammonia. Most of the components of the photorespiratory cycle have been identified by forward genetic analyses starting in the 1980s. For example the first enzyme of the photorespiratory cycle PGLP (3), or the mitochondrial SHMT (4, 5) have been discovered by this approach. The hallmark of photorespiratory mutants is reduced or no growth under ambient air but normal WT-like growth under CO_2 enriched air. Only peroxisomal HPR mutants show no typical photorespiratory growth limitation in ambient air due to an alternative cytosolic pathway that

suppresses the effect of the mutation (6). Photorespiratory mutants additionally contain elevated pools of photorespiratory metabolites.

The photorespiratory pathway intermediates have to be transported across multiple cellular membranes at high flux rates, which is facilitated by metabolite transporters residing in the organellar membranes. In total two molecules of glycolate and one molecule of glycerate cross the chloroplast envelope and the peroxisomal membrane in one turn of the photorespiratory cycle. Two molecules of glycine and one molecule of serine cross the peroxisomal membrane and the mitochondrial envelope. These six transport steps in the carbon cycle of photorespiration together with the transport steps required for the nitrogen cycle result in 18 postulated transport processes (see (7), Eisenhut et al, 2012 for review). While all soluble enzymes of the photorespiratory cycle are identified on the molecular level, this is only true for the chloroplastidic transporters DiT1 and DiT2 that are involved in nitrogen recycling during photorespiration (8-11). Transporters required for the photorespiratory carbon cycle are still unknown, which represents a major gap in the knowledge about this important pathway.

While most enzymes required for photorespiration have been identified by forward genetics, only one transporter associated with photorespiration has been found by this approach to date (8, 10). An alternative means for the identification of candidate genes is transcript co-expression analysis, which was developed by Eisen et al. (12) using microarray data for yeast. Co-expression analysis is based on the assumption that genes that function in the same pathway tend to display similar expression patterns. Hence, unknown genes that are co-regulated with genes of a particular metabolic pathway are hypothesized to be involved in the same biological process. By this method a wide range of genes were functionally characterized in yeast (13) or humans (14). In *Arabidopsis thaliana* a co-expression analysis in combination with reverse genetics has been a successfully strategy to find genes involved in, e.g., flavonoid metabolism (15), cellulose and (16) lignin biosynthesis (17, 18), and aliphatic glucosinolate biosynthesis (19).

In this work we have used co-expression analysis in combination with reverse genetics to identify a transporter involved in the photorespiratory cycle. Co-expression analysis revealed that the genes encoding photorespiratory enzymes are co-expressed with each other and with the candidate transporter Plastidal glycolate glycerate translocator 1 (PLGG1), which was previously identified in proteomics of chloroplasts (20), and was assumed to be involved in programmed cell death (21, 22). By analyzing an *A. thaliana* T-DNA knockout mutant deficient in PLGG1 (*plgg1-1*), we could demonstrate that the mutant has a strong photorespiratory phenotype and is no longer able to transport glycolate and glycerate across the chloroplast envelope. The chloroplastidic glycolate/glycerate transporter PLGG1 thus defines a new class of metabolite transporters that is present in Archaeplastida, fungi, bacteria, and archaea.

Results

PLGG1 is co-expressed with known photorespiratory enzymes. To identify transporters involved in the photorespiratory pathway a co-expression analysis with eleven known photorespiratory pathway enzymes was conducted using publicly available co-expression databases (AtGenExpress & NASC Array, CSB.DB (http://csbdb.mpimp-golm.mpg.de/). Photorespiratory genes were significantly co-expressed with approx. 100 to 150 genes, depending on the co-expression matrix used. These included genes involved in photorespiration, photosynthesis, and chloroplast function (23). The Spearman and Pearson co-expression coefficients for ten photorespiratory genes exceeded 0.9 for 73 out of 98 tested cases (Table 1). PGLP is co-expressed with all genes involved in the pathway with the exception of GLYK (Table 1). For some of the enzymes several isoforms of the enzymes exist. In these cases, only one distinct isoform was found to be co-expressed with other photorespiratory genes. For example, of the five isoforms encoding enzymes for glycolate oxidation, only AtGOX1 is strongly correlated (AtGOX2 is indistinguishable due to the probe on the ATH1 chip) (Table 1). Likewise only GGT1 but not GGT2, only one isoform each of the subunits of GDC, and only the SHM1 isoform is co-expressed in the context of photorespiration (Table 1). Only GLYK is not co-expressed with the other genes in the pathway (Table1).

To identify unknown transporters in the photorespiratory cycle, proteins of unknown function with a co-expression coefficient of at least 0.9 with the majority of photorespiratory genes were tested for the presence of predicted membrane spanning helices. One protein of unknown function previously identified by proteomics and named PLGG1 (20) was both co-expressed with genes involved in photorespiration and had twelve predicted membrane spanning helices (ARAMEMNON, (24)). We hypothesized that PLGG1 is a transporter involved in photorespiration.

The *plgg1-1* insertion mutant exhibits a photorespiratory phenotype

To test the hypothesis that PLGG1 is involved in photorespiration, we isolated an *A. thaliana* T-DNA insertional mutant. The *plgg1-1* mutant carries the T-DNA insertion in the first intron of the respective gene At1g32080 (SI1). PLGG1 cDNA can only be detected in WT but not in *plgg1-1* mutant plants (SI1). Under ambient CO_2 conditions (380 ppm) the *plgg1-1* mutant developed yellow and bleached lesions on the leaf lamina but not along the veins (Fig. 1A). This phenotype was suppressed under high CO_2 (3000 ppm) conditions (Fig. 1B). The phenotype was complemented when PLGG1 was expressed under its own promoter in *plgg1-1* plants (Fig. 1C). We conclude that *plgg1-1* displays a bleached leaf phenotype in ambient CO_2 that can be suppressed by elevated CO_2, which is consistent with a function of PLGG1 in photorespiration.

Photorespiratory metabolites accumulate in *plgg1-1* plants

To determine the position of PLGG1 in the photorespiratory pathway, we analyzed metabolites in WT and *plgg1-1* plants under high and ambient CO_2 conditions. The chosen time points represent metabolite levels under: (i) high CO_2 (0 days), conditions under which the rate of photosynthesis was nearly identical in WT and mutant (14.56 ± 0.42 and 13.87 ± 0.23 μmol CO_2 m^{-2} s^{-1}, respectively, SI2); (ii) after shift to ambient CO_2 conditions of plants with no or very weak visible phenotype but already decreased photosynthetic capacity in the mutant (2 days, 6.7 ± 0.76 μmol CO_2 m^{-2} s^{-1}); and (iii) after a pronounced visible phenotype and low rates of photosynthesis were observed in the mutant (5 days, 4.62 ± 0.47 μmol CO_2 m^{-2} s^{-1}). The largest and most significant differences between *plgg1-1* mutant plants and WT were found for glycolate, glycine, serine, hydroxypyruvate, and glycerate, which are all photorespiratory intermediates. Already in CO_2 enriched air (0 days) glycolate and glycerate were significantly elevated in the mutant compared to WT (Fig. 2A and F). This accumulation escalated when plants were shifted from elevated CO_2 to ambient air. Glycine, serine, and hydroxypyruvate did not accumulate in mutant plants under high CO_2 but when shifted to ambient air (Fig. 2C-E). Steady state glyoxylate levels in mutant plants did not differ significantly from those in WT. After shift from high CO_2 to ambient CO_2 conditions, only serine levels changed in WT plants (Fig. 2C).

Since photorespiration occurs only during the day when RubisCO is active, we expected day-time dependent metabolite accumulation patterns in the mutant. During the night in elevated CO_2, in WT and *plgg1-1* plants, photorespiratory metabolite levels with the exception of glycerate did not differ strongly (Fig. 3). During the day in elevated CO_2, only glycerate and glycolate accumulated in a time dependent manner (Fig. 3A and E). After shift to ambient air, all five metabolites showed a light dependent accumulation that became more pronounced throughout the light period. During the night in ambient air glycolate and glycine levels in *plgg1-1* plants dropped to WT levels. Serine and hydroxypyruvate levels did not mirror the values during the night in CO_2 enriched air, but stayed elevated (Fig. 3C-D). The metabolite profile of the *plgg1-1* mutant is consistent with a role of PLGG1 in photorespiration. The pronounced accumulation of glycolate and glycerate and the localization of this transporter at the chloroplast envelope membrane ((25), SI2) point to PLGG1 as a glycolate/glycerate transporter.

Glycerate- and light-dependent O_2 evolution in intact chloroplasts

Transport experiments with isolated chloroplasts showed that a single transporter transports glycolate and glycerate across the envelope (26, 27). Therefore we expressed PLGG1 heterologously and measured active ^{14}C-glycerate uptake in an *in vitro* uptake system. A Michaelis-Menten-type saturation kinetics of glycerate uptake was observed when liposomes were preloaded with glycerate or glycolate. However, a high background signal was detected in the liposome system due to high rates of unspecific diffusion of glycerate in the liposome system (SI3). Therefore to corroborate the result of the liposome uptake assays, we employed an *in vivo*

system that was successfully used before (26) to test whether the *plgg1-1* mutant was affected in the transport of glycerate. Chloroplasts show light-dependent oxygen evolution in the presence of glycerate (26) because imported glycerate is converted by glycerate kinase to 3-PGA, which is further reduced to triosephophate. The reduction consumes NADPH and the photosynthetic regeneration of NADPH produces oxygen. Hence, feeding of glycerate to isolated chloroplasts drives oxygen evolution (for a detailed scheme see SI4).

To test whether *plgg1-1* chloroplasts were physiologically intact and transport competent, 3-PGA-dependent oxygen evolution was measured. The rates of 3-PGA-dependent oxygen evolution observed for WT and *comp* chloroplasts (20.32 µmol/mg Chl*h) did not differ significantly from the rates observed for *plgg1-1* chloroplasts (Table 2). With WT and *comp* chloroplasts O_2 evolution rates of 22.54 and 18.77 µmol/mg Chl*h, respectively, were observed with 1mM glycerate. The rates observed with *plgg1-1* chloroplasts did not exceed the background signal and differed significantly (P < 0.0001) from the WT rates. Hence chloroplasts of the *plgg1-1* mutant do not display glycerate-dependent oxygen-evolution, which supports PLGG1's function as a glycerate transporter.

Glycolate and glycerate flux are impaired in *plgg1-1* plants

PLGG1's function as a glycolate transporter was further tested by [18O]oxygen flux analysis. Transfer of label from glycolate to downstream metabolites would be impaired in *plgg1-1* if PLGG1 was involved in catalyzing the efflux of glycolate from the chloroplast. To follow photorespiratory flux in *plgg1-1* in relation to WT plants we incubated plants in an [18O]oxygen-atmosphere and followed the incorporation of the label into metabolites, a method that was previously used to identify metabolites and the kinetics of the photorespiratory cycle (28). The label was incorporated into glycolate within 5 sec and into glycerate after 3 min (28). Therefore we chose timepoints 30 sec, 1 min, 2 min and 5 min for a kinetic analysis. All values were expressed relative to the initial label (30 sec). For glycolate, the values observed with WT and *comp* plants were similar to those obtained by Berry et al. (Fig. 4). After 30 sec exposure to 18O, the incorporation reached a plateau and no further increase in label incorporation was detectable after 1 (0.93 fold for WT and 0.86 fold for *comp*), 2 (0.88 fold for WT and 0.81 fold for *comp*), and 5 min (0.88 fold for WT and 0.88 fold for *comp*), respectively, indicating that a steady state was achieved. For *plgg1-1* plants, no increase in label incorporation into glycolate was obtained after 1 min as compared to the value measured after 30 sec. In contrast to WT, the label in glycolate continued to accumulate after 2 min (1.11 fold) and 5 min (1.86 fold) in the mutant, which indicates a slower removal of label from the glycolate pool. This was consistent with a reduced export of glycolate from the chloroplast. In WT and *comp* plants, the 18O label incorporation in glycerate increased slowly. This was not true for the mutant plants. Here no increase in label incorporation into glycerate could be observed with continuous low values for

all time points. Again, this data indicates that the transfer of label from the glycolate to the glycerate pool is impaired, which can be explained by reduced export of glycolate from the chloroplast. Thus, *plgg1-1* is impaired in the transport of both glycolate and glycerate.

Discussion

Photorespiration is a highly compartmentalized pathway that requires multiple transmembrane transport steps across organellar membranes to enable the high flux of metabolites in this process. Despite their importance, no transporter involved in the carbon cycle of photorespiration has been identified to date. Here we report the identification of the chloroplastidic glycolate/glycerate transporter PLGG1 as the first transporter of the photorespiratory carbon cycle.

Co-expression analysis showed that PLGG1 is co-regulated with photorespiratory enzymes (Table 1) and an *A. thaliana* T-DNA insertional mutant deficient in *plgg1-1* (SI1) is only able to grow WT-like and exhibit WT-like rates of photosynthesis in elevated CO_2 (Fig. 1A-B, SI2). In previous studies this CO_2-dependent phenotype was also observed for other photorespiratory mutants (4, 29-31). Those plants grow normally in CO_2 enriched air but have reduced growth or are even inviable under ambient CO_2. Compared to other photorespiratory mutants, the phenotype observed for *plgg1-1* plants is relatively mild because plants can still grow, albeit more slowly and with visible symptoms (Fig. 1), in ambient air, which is likely the reason it was not identified in previous photorespiratory mutant screens. Glycolate, a small organic acid, is probably able to diffuse out of chloroplasts, as do other small organic acids (32) once it accumulates to high levels (Fig. 2). In the absence of high photorespiratory flux during nighttime, the glycerate and glycolate pools approach wild type levels (Fig. 3), which likely alleviated the photorespiratory symptoms in the *plgg1-1* mutant. A shift in night time metabolism at least partially metabolized extra-plastidial glycerate, likely towards serine synthesis which explains why, during the night in ambient air, serine and hydroxypyruvate levels continue to rise while glycerate levels fall but do not reach WT levels. Accumulation of the photorespiratory metabolites glycine, serine and hydroxypyruvate is likely driven by feedback through cytoplasmatic and peroxisomal glycerate accumulation and was previously observed in other photorespiratory mutants with elevated glycerate levels (6, 30).

Glycerate and glycolate are transported across the chloroplast envelope in the photorespiratory cycle by the same transport protein (26). The localization of PLGG1 in the chloroplast envelope (SI 3) together with the metabolite accumulation patterns (Fig. 2 and 3) identify PLGG1 as the chloroplastidic glycolate/glycerate transporter, that was characterized biochemically the 1980s and 1990s in McCarthy's laboratory (26, 27, 33-35). It was shown that the transport rates observed with isolated chloroplasts are sufficient to cope with the high

photorespiratory carbon flux, that the transporter exhibits a proton/substrate symport activity, and that glycolate and glycerate are transported through the same transport protein.

To verify that PLGG1 is the glycolate/glycerate transporter, we expressed PLGG1 heterologously and measured ^{14}C-glycerate/glycolate counter-exchange activity in a reconstituted liposome assay. While a saturable and preloading-dependent uptake kinetics could be observed, also high rates of unspecific diffusion were detected (SI3). This high rate of diffusion is typical for small organic acids in artificial membrane systems and was detected before (27). We therefore next employed an *in vivo* approach that is more suitable to verify glycolate/glycerate transport. To this end, we used an *in vivo* approach with isolated chloroplasts that was developed by Howitz and McCarty (26) and ^{18}O$_2$ flux analysis in *plgg1-1* plants. WT chloroplasts evolved oxygen when provided with 3-PGA or glycerate, proving that they were transport competent and biochemically active. *PLGG1-1* chloroplasts were also transport competent and biochemically active as they were capable of evolving oxygen when supplied with 3-PGA (Table 2). They were only unable to transport glycerate as no glycerate-dependent oxygen evolution was observed with *plgg1-1* chloroplasts (Table 2). Flux analysis further supported PLGG1's role as the glycolate/glycerate transporter since the ^{18}O$_2$ label accumulation in glycolate increased in the *plgg1-1* plants over time indicating that glycolate is trapped in the chloroplast and is not accessible to enzymes that process glycolate outside of the chloroplast (Fig. 3). In contrast, the label accumulation in glycolate in WT plants already reached a plateau after 30 sec and did not increase, indicating that glycolate was processed outside of the chloroplast and can thus proceed in the photorespiratory cycle. In WT plants label incorporation into glycerate increased over time as it takes a few minutes until turnover of the majority of the glycerate pool is achieved (28). In *plgg1-1* plants the transfer of the label from glycolate to other metabolites is mostly blocked and thus no increase in label in the glycerate pool is detectable.

Two recent publications hypothesized that PLGG1 is involved in chloroplast development or functions against cell death (21, 22) but in neither publication was the molecular function of PLGG1 identified. The conclusions were drawn from the visible observations that true leaves develop chlorotic regions in which chloroplasts are destroyed. We demonstrate that that elevated CO$_2$ alleviates the symptoms in all developmental stages (Fig. 1, SI5), that the accumulation of photorespiratory metabolites predates the occurrence of visible symptoms (Fig. 2), and that PLGG1 is the chloroplast glycerate/glycolate carrier (Fig. 4, Table 2). Since the phenotype can be suppressed with high CO$_2$ (Fig. 1) and biochemically active chloroplasts can be isolated from *plgg1-1* mutant plants grown in high CO$_2$ (Table 2) we posit that the visible symptoms are due to the accumulation of toxic concentrations of glycolate and glycerate. Indeed photorespiratory intermediates including glycerate can be toxic to chloroplasts at high concentrations (36) explaining chloroplast and cell disruption in the *plgg1-1*

mutant. Thus, the chloroplastidic glycolate/glycerate transporter Plgg1 defines a new class of metabolite transporters, which is present in Archaeplastida, fungi, bacteria, and archaea (SI6).

Material and Methods

Plant growth and conditions

A. thaliana ecotype Columbia (Col-0) was used as wild-type reference (WT). The SALK line SALK_053469 (*plgg1-1*) was obtained from the Nottingham Arabidopsis Stock Centre (37). Complemented *plgg1-1* mutant plants (*comp*) were used as control for the SALK line. Unless stated otherwise, plants were grown in normal air (380 ppm CO_2) and in air with elevated CO_2 (3000 ppm; high CO_2) at a 12-h-light/12-h-dark cycle (22/18°C) in growth chambers (150 µmol m^{-2} sec^{-1} light intensity) on soil (mixture of 1/4 Floraton and ¾ Arabidopsis root substrate).

Isolation of the T-DNA insertion line

PCR-based screening was used to isolate a homozygous T-DNA insertion line for *Plgg1*. Primers P1, P2 and P3 were used for the genomic DNA screening (for primer sequences see SI Table 1). P1 and P2 for amplification of the WT gene and P1 and P3 for the T-DNA/gene junction. The effect of the T-DNA insertion on the amount of *PLGG1* transcript amounts was tested using qPCR with cDNA of WT and *plgg1-1* plants as template and primers P4 and P5 for amplification. As a positive control *ACTIN7* (At*ACT7*, *At5g09810*) was amplified using P6 and P7.

Statistical analysis

Curve fits and Student's t-tests were performed with PRISM5.0a (GraphPad, http://www.graphpad.com/prism/prism.htm). Results were called as 'extremely significant' if the *P* value was < 0.01% ($P<0.0001$) and 'very significant' if the *P* value was < 0.1% ($P<0.001$), and is indicated by three and two asterisks, respectively.

beta-glucuronidase (GUS) expression and establishment of a complementation line

To assess the expression profile of the *PLGG1* gene, a 3.5 kb gDNA fragment upstream of the ATG- start site including the first nine bases of the *PLGG1* gene was amplified using P8/P9 and cloned into vector pCAMBIA3301 for a C-terminal GUS fusion. For complementation analysis a 9 kb gDNA fragment, including 3.5 kb upstream of the ATG-start site, the full genomic *PLGG1* sequence und 0.5 kb 3'-UTR was amplified using P10/P11 and cloned into vector pCAMBIA3301. *A. thaliana* plants were stably transformed using the floral dip method (38). GUS staining of two week old seedlings was performed using the method described in (39).

Metabolite Extraction, Gas Chromatography–Time of Flight–Mass Spectrometry Analysis

Methanolic extraction of leaf material was performed according to the method described by

Fiehn et al. (40) and GC/MS was performed according to Lee & Fiehn (41). Analysis of metabolites was performed by GC/MS (Agilent Technologies 5973 (Santa Clara, CA, USA)) . Results were analyzed using the MassLynx software package supplied with the instrument (Waters). As an internal standard ribitol was added and relative metabolite levels were determined from the ratio of the area of each metabolite and the corresponding ribitol area. A detailed description of the methods is available in the supplement.

Transient expression of GFP fusions in tobacco protoplast

For localization studies a C-terminal GFP fusion construct was cloned using Gateway vectors to insert PLGG1 into pMDC83 (42) via pDONR207 (Invitrogen, Gateway®). For amplification primer P12/P13 were used. Vectors were transformed into *Nicotiana benthamiana* leaves via infection with *Agrobacterium tumefaciens* strain *GV3101* (43, 44), protoplasts were isolated after three days and localization was visualized using a Zeiss Laser scanning microscope 510 Meta (Zeiss, http://www.zeiss.com) as described in detail in Breuers et al. (25).

Chloroplast isolation and glycerate dependent oxygen evolution

Chloroplasts were isolated according to the method described by Aronsson and Jarvis (45) using a two-step-Percoll gradient, with the modification that plants were grown on soil for three weeks in high CO_2. Intactness of chloroplasts was determined using the Hill-reaction (46). 3-PGA and glycerate dependent oxygen evolution was measured according to the method described by Howitz and McCarty (26).

$^{18}O_2$ feeding

A. thaliana WT and *plgg1-1* plants were fed with $^{18}O_2$ according to the method described in (28). Plants were grown on soil for four weeks in ambient air in 8-h-light/16-h-dark cycle (20/16°C). A single plant in a pot was placed in a 1 L plastic bag with a plastic seal (Paclan GmbH, Henfenfeld). Air was removed by vacuum and replaced by air-mixture of 0.03% CO_2, 78.97% N_2 and 21% $^{18}O_2$. After 0.5, 1, 2 and 5 min plants were freeze-quenched within less than 2 s in liquid nitrogen and processed for metabolite analysis. Metabolites were extracted by the method optimized for photorespiratory metabolites. A detailed description of the methods is available in the supplement.

References:

1. Kebeish R*, et al.* (2007) Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in Arabidopsis thaliana. *Nat Biotechnol* 25(5):593-599.

2. Maier A*, et al.* (2012) Transgenic Introduction of a Glycolate Oxidative Cycle into A. thaliana Chloroplasts Leads to Growth Improvement. *Front Plant Sci* 3:38.

3. Somerville CR & Ogren WL (1979) Phosphoglycolate Phosphatase-Deficient Mutant of Arabidopsis. *Nature* 280(5725):833-836.

4. Somerville CR & Ogren WL (1981) Photorespiration-deficient Mutants of Arabidopsis thaliana Lacking Mitochondrial Serine Transhydroxymethylase Activity. *Plant Physiol* 67(4):666-671.

5. Voll LM*, et al.* (2006) The photorespiratory Arabidopsis shm1 mutant is deficient in SHM1. *Plant Physiol* 140(1):59-66.

6. Timm S*, et al.* (2008) A cytosolic pathway for the conversion of hydroxypyruvate to glycerate during photorespiration in Arabidopsis. *Plant Cell* 20(10):2848-2859.

7. Reumann S & Weber APM (2006) Plant peroxisomes respire in the light: Some gaps of the photorespiratory C-2 cycle have become filled - Others remain. *Bba-Mol Cell Res* 1763(12):1496-1510.

8. Somerville SC & Ogren WL (1983) An Arabidopsis-Thaliana Mutant Defective in Chloroplast Dicarboxylate Transport. *Proc Nat Acad Sci U S A* 80(5):1290-1294.

9. Somerville SC & Somerville CR (1985) A Mutant of Arabidopsis Deficient in Chloroplast Dicarboxylate Transport Is Missing an Envelope Protein. *Plant Sci Let* 37(3):217-220.

10. Renne P*, et al.* (2003) The Arabidopsis mutant dct is deficient in the plastidic glutamate/malate translocator DiT2. *Plant J* 35(3):316-331.

11. Schneidereit J, Hausler RE, Fiene G, Kaiser WM, & Weber APM (2006) Antisense repression reveals a crucial role of the plastidic 2-oxoglutarate/malate translocator DiT1 at the interface between carbon and nitrogen metabolism. *Plant J* 45(2):206-224.

12. Eisen MB, Spellman PT, Brown PO, & Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Nat Acad Sci U S A* 95(25):14863-14868.

13. Wu LF*, et al.* (2002) Large-scale prediction of Saccharomyces cerevisiae gene function using overlapping transcriptional clusters. *Nat Genet* 31(3):255-265.

14. Lee HK, Hsu AK, Sajdak J, Qin J, & Pavlidis P (2004) Coexpression analysis of human genes across many microarray data sets. *Genome Res.* 14(6):1085-1094.

15. Yonekura-Sakakibara K, Tohge T, Niida R, & Saito K (2007) Identification of a flavonol 7-O-rhamnosyltransferase gene determining flavonoid pattern in Arabidopsis by transcriptome coexpression analysis and reverse genetics. *J Biol Chem* 282(20):14932-14941.

16.   Persson S, Wei HR, Milne J, Page GP, & Somerville CR (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc Nat Acad Sci U S A* 102(24):8633-8638.

17.   Ehlting J, *et al.* (2005) Global transcript profiling of primary stems from Arabidopsis thaliana identifies candidate genes for missing links in lignin biosynthesis and transcriptional regulators of fiber differentiation. *Plant J* 42(5):618-640.

18.   Alejandro S, *et al.* (2012) AtABCG29 Is a Monolignol Transporter Involved in Lignin Biosynthesis. *Curr Biol* 22(13):1207-1212.

19.   Hansen BG, Kliebenstein DJ, & Halkier BA (2007) Identification of a flavin-monooxygenase as the S-oxygenating enzyme in aliphatic glucosinolate biosynthesis in Arabidopsis. *Plant J* 50(5):902-910.

20.   Bräutigam A, Hofmann-Benning S, & Weber APM (2008) Comparative Proteomics of Chloroplast Envelopes from C3 and C4 Plants Reveals Specific Adaptations of the Plastid Envelope to C4 Photosynthesis and Candidate Proteins Required for Maintaining C4 Metabolite Fluxes. *Plant Physiol* 148(1):568-579.

21.   Horan K, *et al.* (2008) Annotating genes of known and unknown function by large-scale coexpression analysis. *Plant Physiol* 147(1):41-57.

22.   Schwacke R, *et al.* (2003) ARAMEMNON, a novel database for Arabidopsis integral membrane proteins. *Plant Physiol* 131(1):16-26.

23.   Breuers FKH, *et al.* (2012) Dynamic Remodeling of the Plastid Envelope Membranes - A Tool for Chloroplast Envelope in vivo Localizations. *Front Plant Sci* 3:7.

24.   Howitz KT & McCarty RE (1986) D-Glycerate Transport By The Pea Chloroplast Glycolate Carrier - Studies On [1-C-14] D-Glycerate Uptake And D-Glycerate Dependent O-2 Evolution. *Plant Physiol* 80(2):390-395.

25.   Howitz KT & McCarty RE (1991) Solubilization, Partial-Purification, And Reconstitution Of The Glycolate Glycerate Transporter From Chloroplast Inner Envelope Membranes. *Plant Physiol* 96(4):1060-1069.

26.   Berry JA, Osmond, C.B., and Lorimer, G.H. (1978) Fixation of 1802 during Photorespiration. *Plant Physiol* 62:954-967.

27.   Somerville CR & Ogren WL (1980) Inhibition of Photosynthesis in Arabidopsis Mutants Lacking Leaf Glutamate Synthase Activity. *Nature* 286(5770):257-259.

28.   Boldt R, *et al.* (2005) D-GLYCERATE 3-KINASE, the last unknown enzyme in the photorespiratory cycle in Arabidopsis, belongs to a novel kinase family. *Plant Cell* 17(8):2413-2420.

29.   Igarashi D, *et al.* (2003) Identification of photorespiratory glutamate : glyoxylate aminotransferase (GGAT) gene in Arabidopsis. *Plant J* 33(6):975-987.

30. Benning C (1986) Evidence Supporting a Model of Voltage-Dependent Uptake of Auxin into Cucurbita Vesicles. *Planta* 169(2):228-237.

31. Howitz KT & Mccarty RE (1985) Kinetic Characteristics of the Chloroplast Envelope Glycolate Transporter. *Biochemistry* 24(11):2645-2652.

32. Howitz KT & Mccarty RE (1985) Substrate-Specificity of the Pea Chloroplast Glycolate Transporter. *Biochemistry* 24(14):3645-3650.

33. Young XK & McCarty RE (1993) Assay Of Proton-Coupled Glycolate And D-Glycerate Transport Into Chloroplast Inner Envelope Membrane-Vesicles By Stopped-Flow Fluorescence. *Plant Physiol* 101(3):793-799.

34. Yang Y, *et al.* (2012) A chloroplast envelope membrane protein containing a putative LrgB domain related to the control of bacterial death and lysis is required for chloroplast development in Arabidopsis thaliana. *New Phytol* 193(1):81-95.

35. Yamaguchi M, *et al.* (2012) Loss of the plastid envelope protein AtLrgB causes spontaneous chlorotic cell death in Arabidopsis thaliana. *Plant Cell Physiol* 53(1):125-134.

36. Enser U & Heber U (1980) Metabolic regulation by pH gradients. Inhibition of photosynthesis by indirect proton transfer across the chloroplast envelope. *Biochim Biophys Acta* 592(3):577-591.

37. Alonso JM, *et al.* (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science* 301(5633):653-657.

38. Clough SJ & Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* 16(6):735-743.

39. Lagarde D, *et al.* (1996) Tissue-specific expression of arabidopsis AKT1 gene is consistent with a role in K+ nutrition. *Plant J* 9(2):195-203.

40. Fiehn O, *et al.* (2001) Metabolite profiling for plant functional genomics *Nature Biotechnol* 19(2):173-173.

41. Lee DY & Fiehn O (2008) High quality metabolomic data for Chlamydomonas reinhardtii. *Plant Methods* 4.

42. Curtis MD & Grossniklaus U (2003) A Gateway Cloning Vector Set for High-Throughput Functional Analysis of Genes in Planta. *Plant Physiol* 133(2):462-469.

43. Koncz C & Schell J (1986) The Promoter of Tl-DNA Gene 5 Controls the Tissue-Specific Expression of Chimeric Genes Carried by a Novel Type of Agrobacterium Binary Vector. *Mol Gen Genet* 204(3):383-396.

44. Bendahmane A, Kanyuka K, & Baulcombe DC (1999) The Rx gene from potato controls separate virus resistance and cell death responses. *Plant Cell* 11(5):781-792.

45. Aronsson H & Jarvis P (2002) A simple method for isolating import-competent Arabidopsis chloroplasts. *FEBS Letters* 529(2-3):215-220.

46.   Bregman A (1990) Laboratory Investigations in Cell and Molecular Biology. *John Wiley & Sons, New York*.

Figure legends:

Fig. 1. Photorespiratory phenotype of *plgg1-1* plants; (A) *plgg1-1*, (C) WT and (D) *comp* plants grown in high CO_2 for four weeks, shifted to ambient CO_2 for one week. (B) *Plgg1-1* plant grown in high CO_2 for five weeks.

Fig. 2. Accumulation of photorespiratory metabolites in WT and *plgg1-1* plants. WT and *plgg1-1* plants were grown in 3000ppm CO_2 conditions for four weeks and then shifted to ambient air. Steady state metabolite levels were measured for plants kept at high CO_2 and after two and five day shift to ambient air; all values are measured in μmol/ g fresh weight except hydroxypyruvate which is shown as arbitrary units; glycolate (A), glyoxylate (B), glycine (C), serine (D), hydroxypyruvate (E) and glycerate (F). Error bars indicate Standard deviation, n = 3.

Fig. 3. Time course of photorespiratory metabolite accumulation in WT and *plgg1-1* plants. WT and *plgg1-1* plants were grown in 3000ppm CO_2 conditions for four weeks and shifted to ambient air (300ppm, at 24 hours, indicated by black arrow); all values are measured in μmol/g fresh weight except hydroxypyruvate which is shown as arbitrary units; glycolate (A), glycine (B), serine (C), hydroxypyruvate (D) and glycerate (E). Error bars indicate Standard deviation, n = 3.

Fig. 4. Revised model of the photorespiratory cycle and $^{18}O_2$ incorporation in glycolate and glycerate; (A) The photorespiratory cycle including the new chloroplastidic glycolate/glycerate transporter PLGG1. $^{18}O_2$ incorporation into glycolate (A) and glycerate (B) was determined in WT, complemented mutants (*comp*) and *plgg1-1* mutant plants. Plants were grown in ambient air for four weeks, incubated in $^{18}O_2$ air for 0.5, 1,2 and 5 min and metabolite levels were analyzed using GC/MS. Metabolite levels were normalized to the relative abundance after 0.5 min in order to calculate the enrichment of $^{18}O_2$ incorporation in glycolate and glycerate. 3-PGA: 3-Phosphoglycerate; ETC: electron transport chain; 2-OG: 2-Oxoglutarate; TP: triose phosphates. N = 3.

Table legends:

Table 1: Co-expression Spearman coefficients for genes involved in photorespiration and for the new candidate gene PLGG1; only the isoform with the best correlation coefficient is shown; correlations with p<0.1 are printed boldface.

Table 2. 3-PGA and glycerate dependent O$_2$ evolution in isolated intact WT, *plgg1-1* and *comp* chloroplasts [µmol/mg Chl*h]. ± indicates Standard deviation, n = 3.

Supporting Information:

SI Fig. 1: T-DNA insertion leads to full k.o. of the gene At1g32080. (A) The T-DNA is inserted in the first intron of the gene At1g32080. The arrow symbolizes the transcription start side. Grey boxes stand for exons, lines for introns, dark gray boxes for 5'-UTR, light gray boxes for 3'-UTR. (B) Expression control of the k.o. gene. WT and mutant gene show actin expression. Only WT shows PLGG1 expression.

SI Fig. 2. PLGG1 is located at the chloroplast envelope. (A) GFP fluorescence; (B) chlorophyll auto-fluorescence and (C) merge of both signals; GFP was fused C-terminal translationally to the full-length PLGG1 protein; the construct was transiently expressed in *Nicotiana benthaminana*; protoplasts were used for microscopic detection of the GFP signal.

SI Fig. 3. PLGG1 shows an active and time- dependent uptake of [14C]Glycerate, that can be inhibited by increasing concentrations of unlabeled glycerate.

(A) Time-dependent uptake of radiolabelled [14C]Glycerate (50 µM external concentration) in the presence (filled circles) or absence of glycolate (open rectangles) as internal substrate (20 mM). In the presence of a suitable counter-exchange substrate inside the liposomes (e.g., glycolate) a Michaelis-Menten-type saturation kinetics was observed. In control liposomes preloaded with equal amounts of buffer or in liposomes that did not contain reconstituted PLGG1, only unspecific diffusion of the labeled substrate was observed.

(B) Concentration-dependent inhibition of active transport in the presence (black bars) or absence of glycolate (white bars) as internal substrate (20 mM). Active transport is inhibited with increasing external glycerate concentrations, while passive diffusion is independent on external glycerate concentration.

Error bars represent SD of three independent uptake experiments.

Recombinant At1g32080 protein was expressed *in vitro* using the 5PRIME wheat germ expression kit (5PRIME GmbH, Hamburg, Germany) as described previously (1) and reconstituted into liposomes using reconstitution and uptake buffers as described in Howitz et al. (2).

SI Fig. 4. Scheme of glycerate dependent oxygen evolution. Glycerate is transported by PLGG1 into the chloroplast in a proton-dependent manner, converted to 3-PGA under ATP consumption. 3-PGA is converted to TP under NADPH consumption. NADP is regenerated to NADPH at the photosynthetic electron transport chain under water consumption and oxygen evolution (red arrow). 3-PGA: 3-Phosphoglycerate; TP: triose phosphates.

SI Fig 6. Phylogenetic Tree of PLGG1. AtPLGG1 was used as the query to Blast Explorer (http://www.phylogeny.fr/version2_cgi/one_task.cgi?task_type=blast) and the sequences suggested by the program were manually curated and fed to an "à la carte" phylogeny analysis (http://www.phylogeny.fr/version2_cgi/phylogeny.cgi). Briefly, sequences were aligned with Muscle and all positions with gaps were removed. The phylogenetic tree was calculated with PhyML and bootstrapped (100 repetitions). For display, all branches with <30% branch support were collapsed.

The tree divides in two large branches, a bacterial and an eukaryotic branch. The sole cyanobacterium *Acaryochloris* spec is nested within the bacteria. In eukaryotes, the tree divides the fungi (more precisely the dikarya) from the archaeplastida. Within the archaeplastida, *Cyanophora paradoxa* separates from the red and green lineages which themselves form separate branches. In the green lineage, the green algae separate from the streptophyta in which the moss branches before the fern. Monocots and dicots form separate branches.

Since the fungi branch with the archaeplastida with a bootstrap support of 62%, the gene likely evolved before the archaeplastida and the fungi split and it was lost from the other branches of life. Had the gene been acquired by a lateral gene transfer, one would expect it to branch within the bacteria. In higher plants PLGG1 is nuclear encoded. For transport proteins of the chloroplasts, at least three different phylogenetic roots have been demonstrated: rooted in the Chlamydiae, rooted in the cyanobacteria and a host origin. PLGG1-like proteins do not appear in the Clamydiae but in Cyanobacteria. The phylogenetic tree indicates that the gene is likely not acquired from the cyanobacteria since the single cyanobacterium PLGG1 is nested within the bacteria. Therefore it is more likely that the cyanobacterium acquired it by lateral gene transfer from a bacterium. Hence PLGG1 likely originated from the host.

(1) Howitz KT & McCarty RE (1991) Solubilization, Partial-Purification, And Reconstitution Of The Glycolate Glycerate Transporter From Chloroplast Inner Envelope Membranes. *Plant Physiology* 96(4):1060-1069.

(2) Nozawa A et al. (2007) A cell-free translation and proteoliposome reconstitution system for functional analysis of plant solute transporters. *Plant Cell Physiol* 48:1815–1820.

(3) Dereeper A., Guignon V., Blanc G., Audic S., Buffet S., Chevenet F., Dufayard J.-F., Guindon S., Lefort V., Lescot M., Claverie J.-M., Gascuel O. *Phylogeny.fr: robust phylogenetic analysis for the non-specialist* Nucleic Acids Research. 2008 Jul 1; 36 (Web Server Issue):W465-9. Epub 2008 Apr 19.

Tables

Table 1: Co-expression Spearman coefficients for genes involved in photorespiration and for the new candidate gene PLGG1; only the isoform with the best correlation coefficient is shown; correlations with p<0.1 are printed boldface.

Enzyme	AGI	Abbre-viation	PGLP1	GOX1	AGT1	GGT1	GLDP1	GLDH1	GLDH3	GLDT1	SHM1	HPR1	GlyK	PLGG1
2-PG phosphatase	At5g36790	PGLP1	1.00	-	-	-	-	-	-	-	-	-	-	-
Glycolate oxidase	At3g14420	GOX1	**0.93**	1.00	-	-	-	-	-	-	-	-	-	-
Ser:glyoxylate aminotransferase	At2g13360	AGT1	**0.91**	**0.95**	1.00	-	-	-	-	-	-	-	-	-
Glu:glyoxylate aminotransferase	At1g23310	GGT1	**0.92**	**0.96**	**0.96**	1.00	-	-	-	-	-	-	-	-
GDC P-protein	At4g33010	GLDP1	**0.93**	**0.95**	**0.93**	**0.94**	1.00	-	-	-	-	-	-	-
GDC H-protein	At2g35370	GLDH1	**0.93**	**0.92**	<0.9	<0.9	<0.9	1.00	-	-	-	-	-	-
GDC H-protein	At1g32470	GLDH3	**0.91**	<0.9	<0.9	<0.9	<0.9	**0.94**	1.00	-	-	-	-	-
Gly decarboxylase T-protein	At1g11860	GLDT1	**0.92**	**0.92**	<0.9	<0.9	<0.9	**0.92**	**0.95**	1.00	-	-	-	-
Serine:hydroxy methyltransferase	At4g37930	SHM1	**0.93**	**0.96**	**0.94**	**0.96**	**0.95**	**0.91**	<0.9	1.00	1.00	-	-	-
Hydroxypyruvate reductases	At1g68010	HPR1	**0.93**	**0.97**	**0.96**	**0.97**	**0.94**	<0.9	<0.9	**0.91**	**0.96**	1.00	-	-
Glycerate kinase	At1g80380	GlyK	<0.9	<0.9	<0.9	<0.9	<0.9	<0.9	<0.9	<0.9	<0.9	<0.9	1.00	-
AtPLGG1	At1g32080	PLGG1	**0.93**	**0.9**	**0.9**	**0.9**	<0.9	**0.93**	**0.93**	**0.93**	**0.91**	**0.93**	<0.9	1.00

Table 2. 3-PGA and glycerate dependent O_2 evolution in isolated intact WT, *plgg1-1* and *comp* chloroplasts [μmol/mg Chl*h]. ± indicates Standard deviation, n = 3.

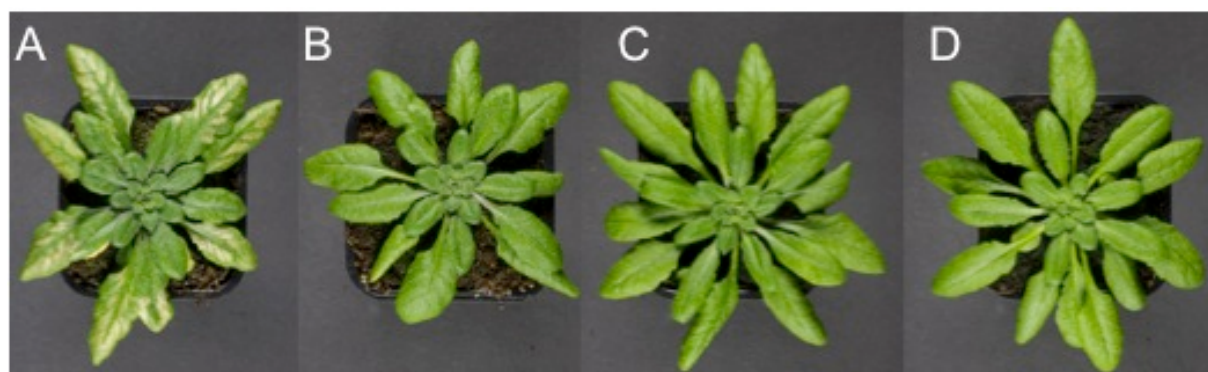|	glycerate	3-PGA
WT	22.54 (± 1.11)	22.12 (± 3.58)
plgg1-1	1.61 (± 1.72)	19.91 (± 1.72)
comp	18.77 (± 0.66)	20.32 (± 2.22)

Figures

Fig. 1. Photorespiratory phenotype of *plgg1-1* plants; (A) *plgg1-1*, (C) WT and (D) *comp* plants grown in high CO_2 for four weeks, shifted to ambient CO_2 for one week. (B) *Plgg1-1* plant grown in high CO_2 for five weeks.
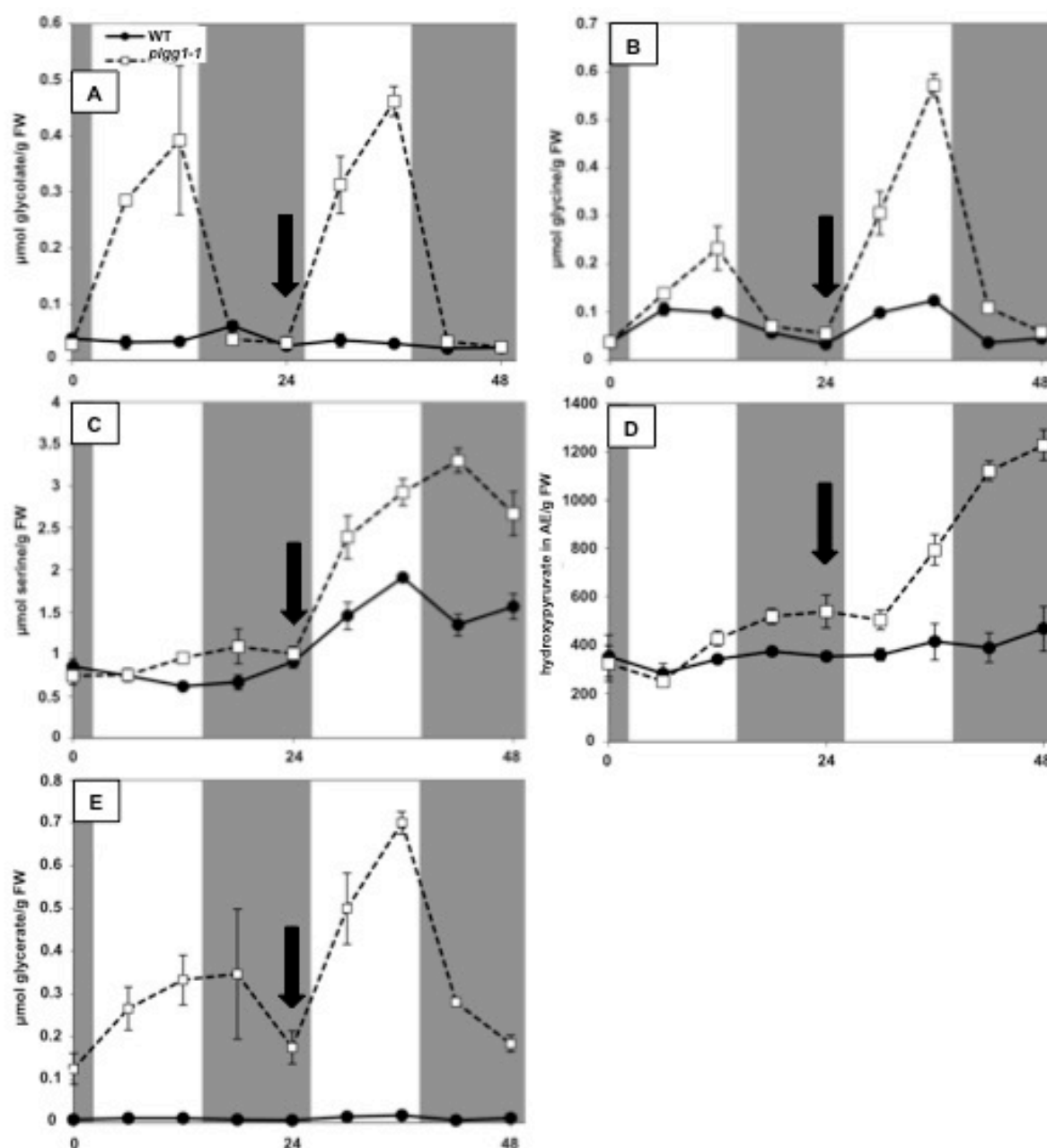
Fig. 2. Accumulation of photorespiratory metabolites in WT and *plgg1-1* plants. WT and *plgg1-1* plants were grown in 3000ppm CO_2 conditions for four weeks and then shifted to ambient air. Steady state metabolite levels were measured for plants kept at high CO_2 and after two and five day shift to ambient air; all values are measured in µmol/ g fresh weight except hydroxypyruvate which is shown as arbitrary units; glycolate (A), glyoxylate (B), glycine (C), serine (D), hydroxypyruvate (E) and glycerate (F). Error bars indicate Standard deviation, n = 3.
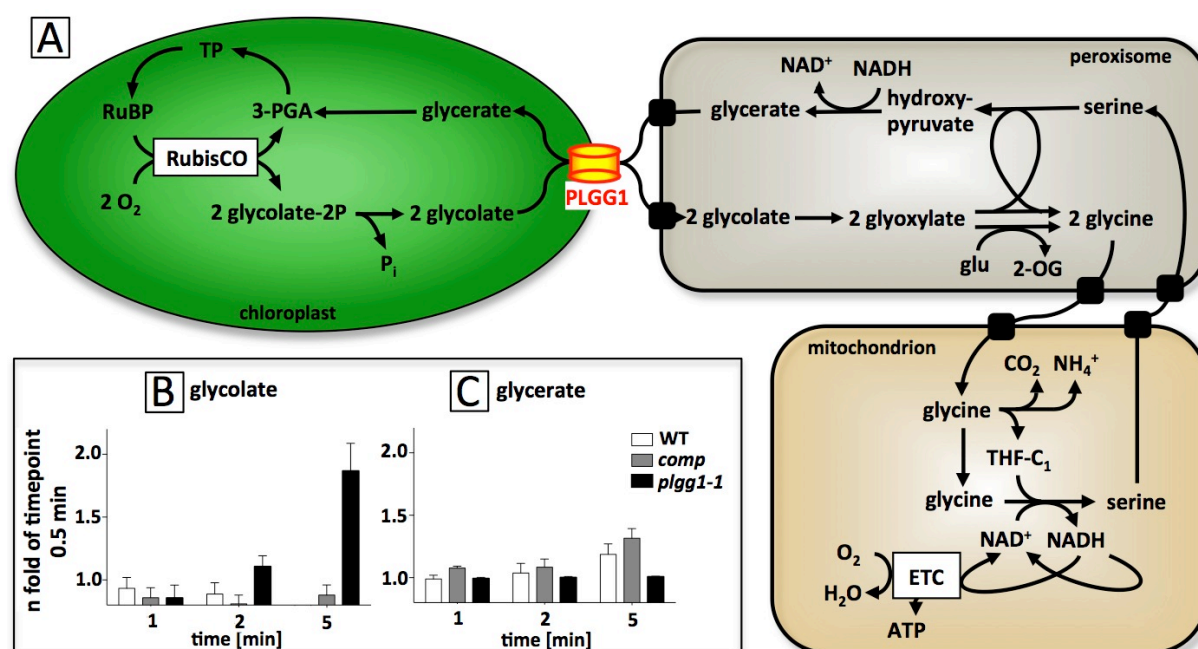
Fig. 3. Time course of photorespiratory metabolite accumulation in WT and *plgg1-1* plants. WT and *plgg1-1* plants were grown in 3000ppm CO_2 conditions for four weeks and shifted to ambient air (300ppm, at 24 hours, indicated by black arrow); all values are measured in µmol/g fresh weight except hydroxypyruvate which is shown as arbitrary units; glycolate (A), glycine (B), serine (C), hydroxypyruvate (D) and glycerate (E). Error bars indicate Standard deviation, n = 3.
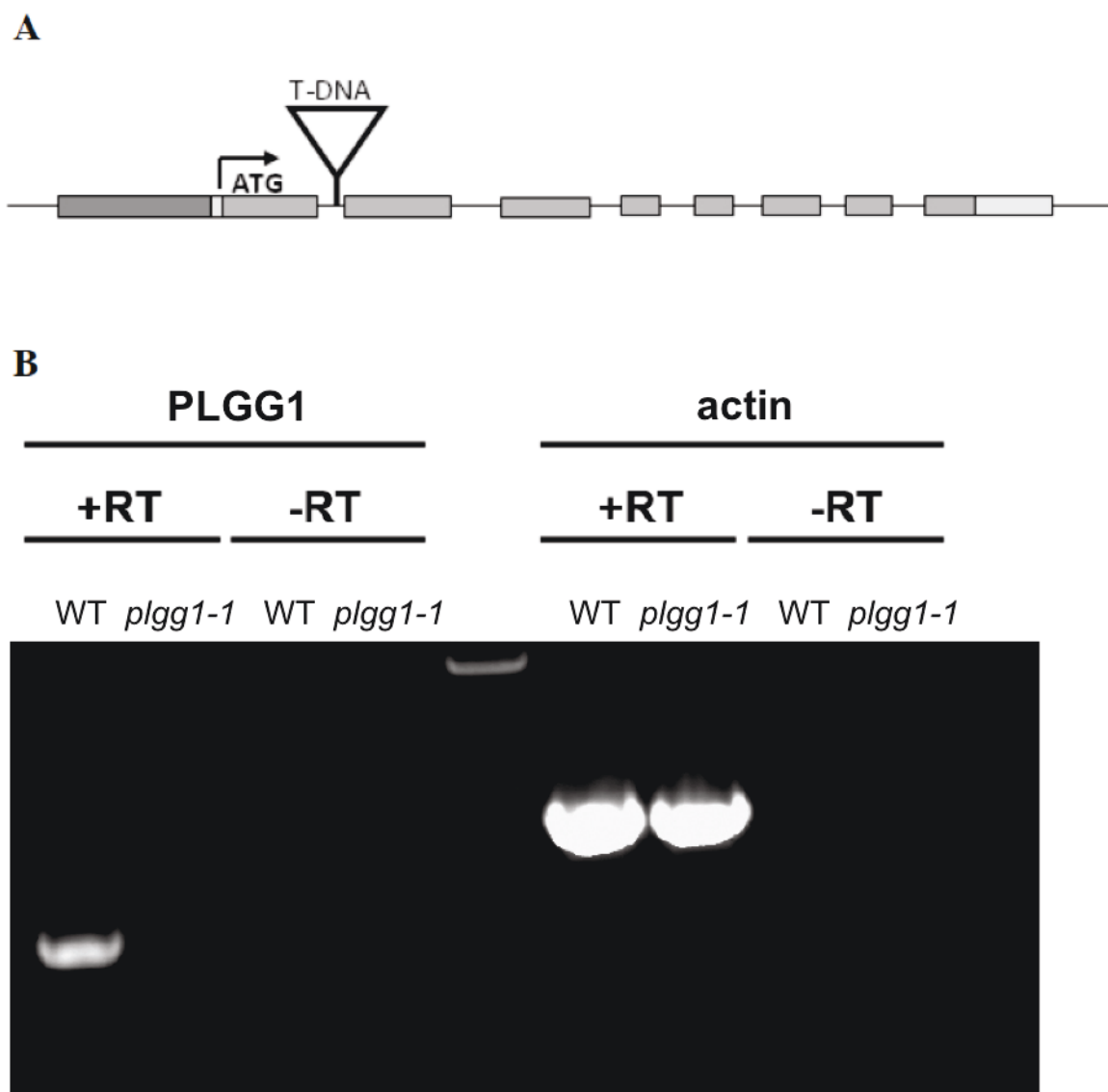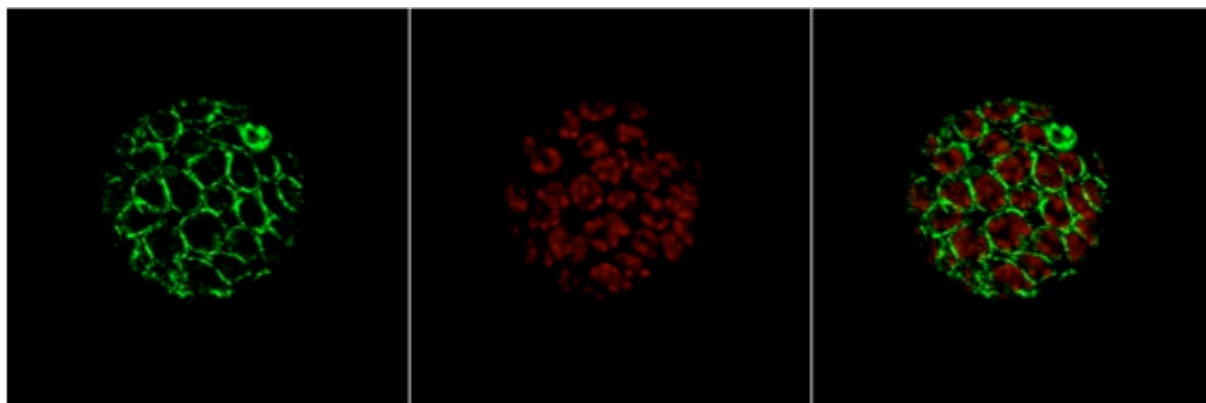
Fig. 4. Revised model of the photorespiratory cycle and $^{18}O_2$ incorporation in glycolate and glycerate; (A) The photorespiratory cycle including the new chloroplastidic glycolate/glycerate transporter PLGG1. $^{18}O_2$ incorporation into glycolate (A) and glycerate (B) was determined in WT, complemented mutants (*comp*) and *plgg1-1* mutant plants. Plants were grown in ambient air for four weeks, incubated in $^{18}O_2$ air for 0.5, 1,2 and 5 min and metabolite levels were analyzed using GC/MS. Metabolite levels were normalized to the relative abundance after 0.5 min in order to calculate the enrichment of $^{18}O_2$ incorporation in glycolate and glycerate. 3-PGA: 3-Phosphoglycerate; ETC: electron transport chain; 2-OG: 2-Oxoglutarate; TP: triose phosphates. N = 3.
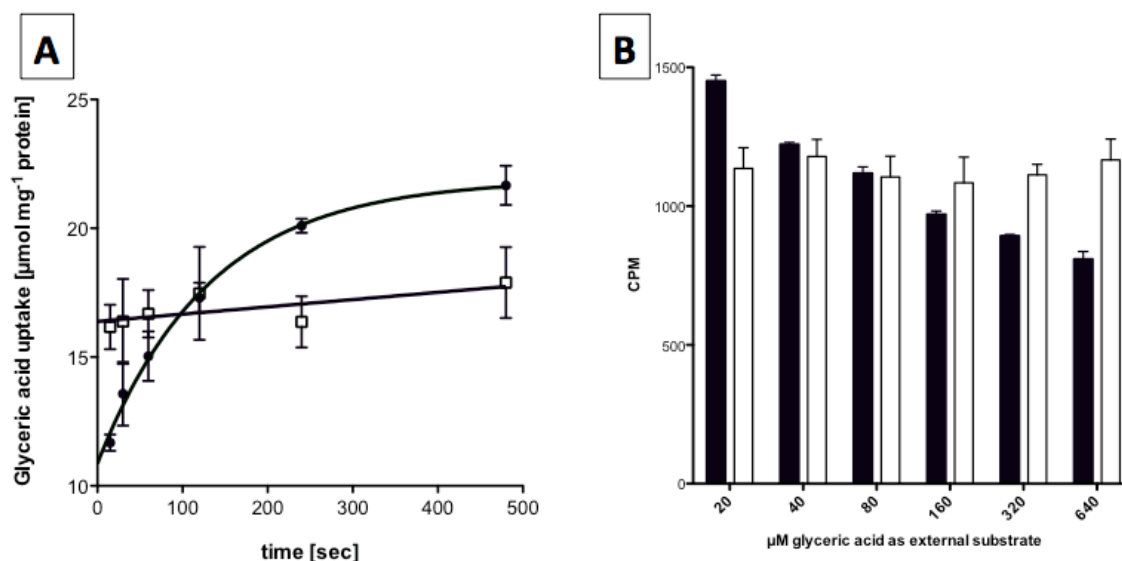
Supporting Information

A

B

SI Fig. 1: T-DNA insertion leads to full k.o. of the gene At1g32080. (A) The T-DNA is inserted in the first intron of the gene At1g32080. The arrow symbolizes the transcription start side. Grey boxes stand for exons, lines for introns, dark gray boxes for 5'-UTR, light gray boxes for 3'-UTR. (B) Expression control of the k.o. gene. WT and mutant gene show actin expression. Only WT shows PLGG1 expression.

SI Fig. 2. PLGG1 is located at the chloroplast envelope. (A) GFP fluorescence; (B) chlorophyll auto-fluorescence and (C) merge of both signals; GFP was fused C-terminal translationally to the full-length PLGG1 protein; the construct was transiently expressed in *Nicotiana benthaminana*; protoplasts were used for microscopic detection of the GFP signal.
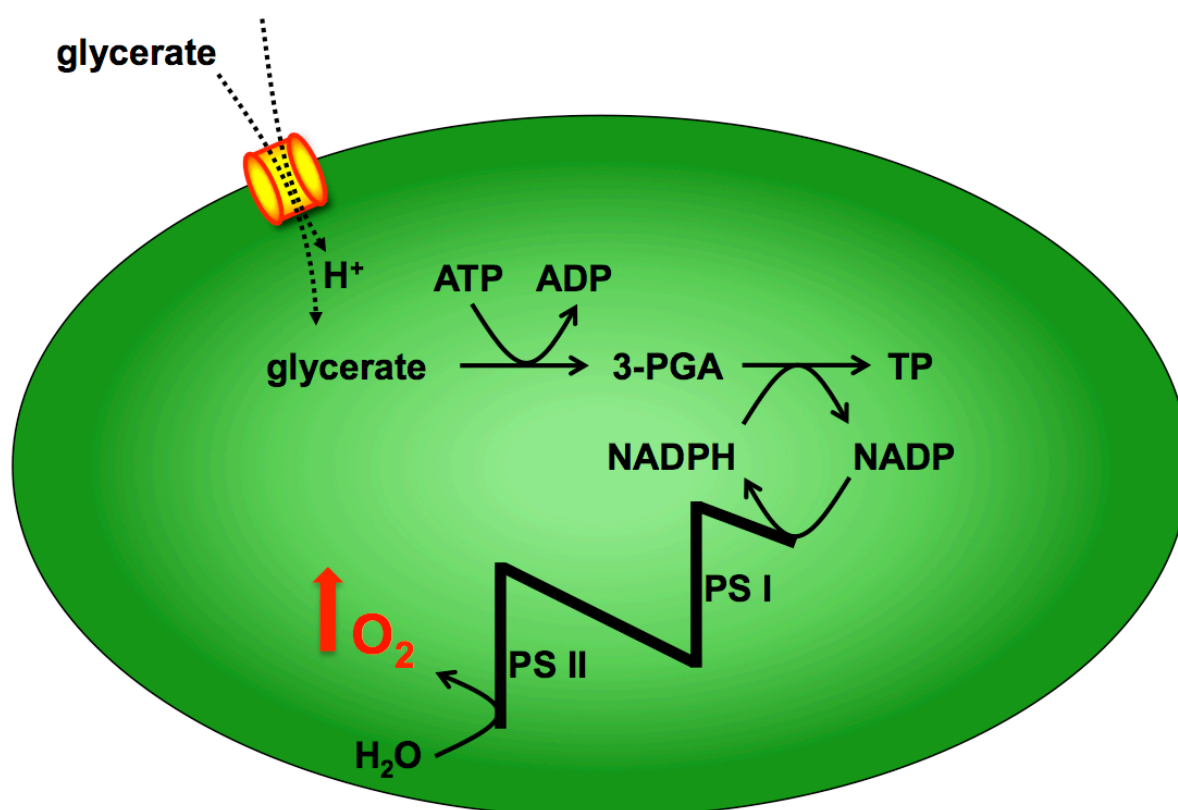
SI Fig. 3. PLGG1 shows an active and time- dependent uptake of [^{14}C]Glycerate, that can be inhibited by increasing concentrations of unlabeled glycerate.

(A) Time-dependent uptake of radiolabelled [^{14}C]Glycerate (50 µM external concentration) in the presence (filled circles) or absence of glycolate (open rectangles) as internal substrate (20 mM). In the presence of a suitable counter-exchange substrate inside the liposomes (e.g., glycolate) a Michaelis-Menten-type saturation kinetics was observed. In control liposomes preloaded with equal amounts of buffer or in liposomes that did not contain reconstituted PLGG1, only unspecific diffusion of the labeled substrate was observed.
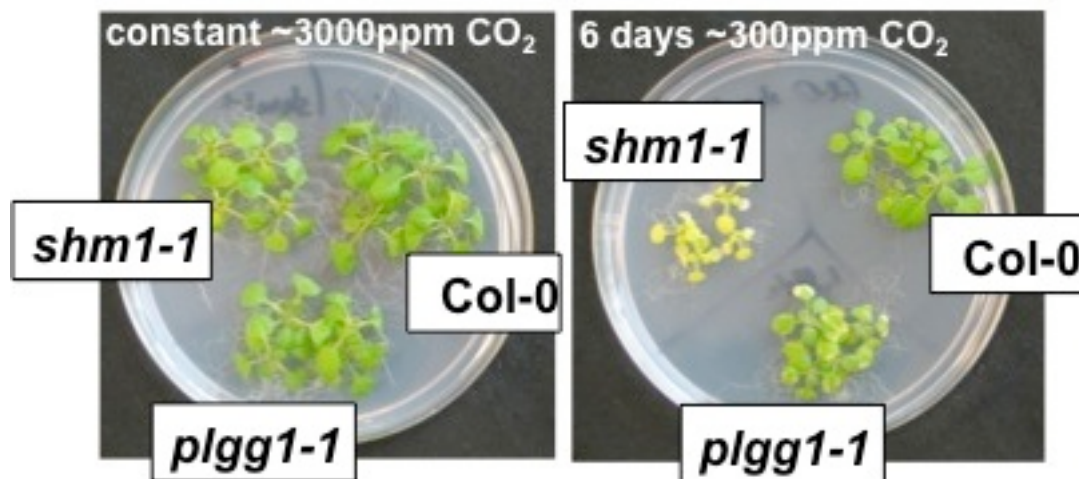
(B) Concentration-dependent inhibition of active transport in the presence (black bars) or absence of glycolate (white bars) as internal substrate (20 mM). Active transport is inhibited with increasing external glycerate concentrations, while passive diffusion is independent on external glycerate concentration.

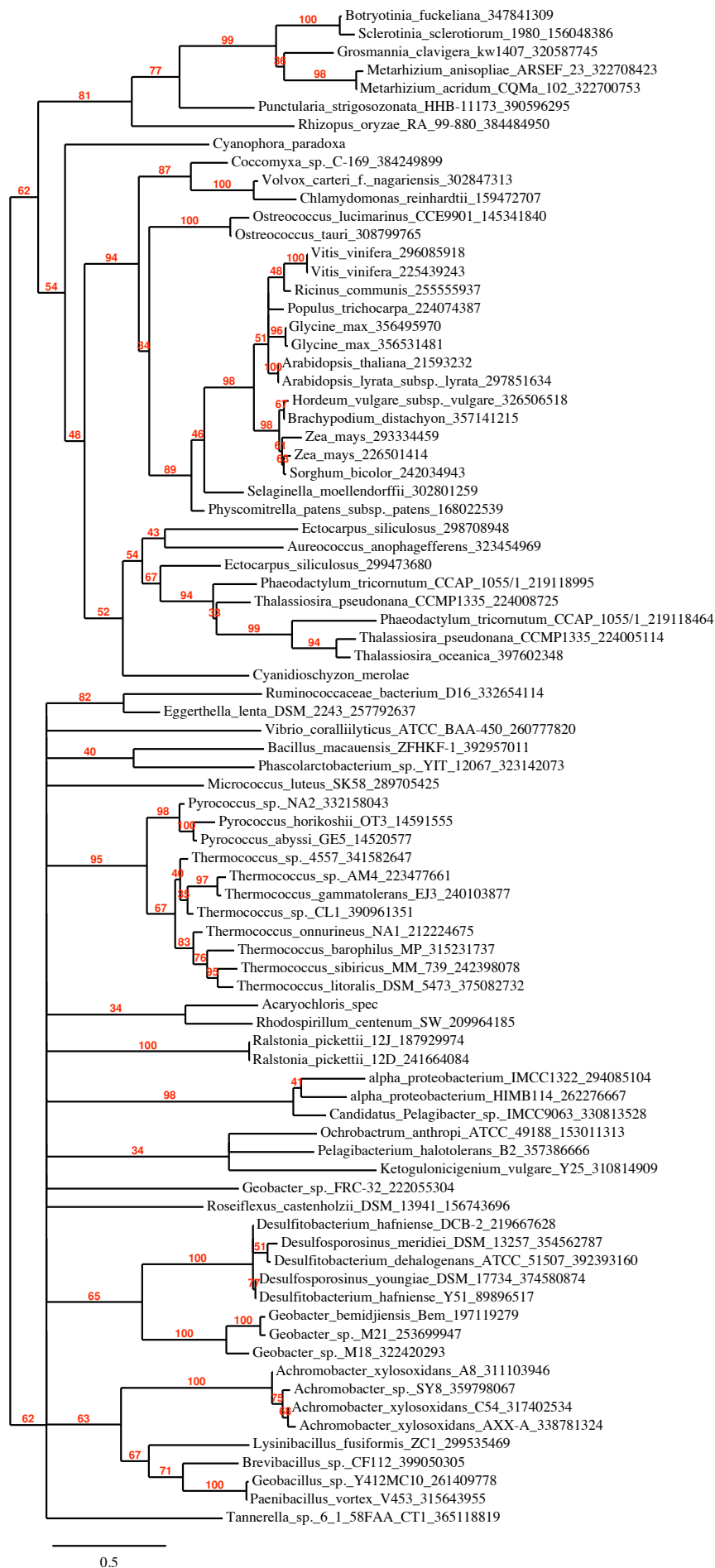Error bars represent SD of three independent uptake experiments.

Recombinant At1g32080 protein was expressed *in vitro* using the 5PRIME wheat germ expression kit (5PRIME GmbH, Hamburg, Germany) as described previously (1) and reconstituted into liposomes using reconstitution and uptake buffers as described in Howitz et al. (2).

SI Fig. 4. Scheme of glycerate dependent oxygen evolution. Glycerate is transported by PLGG1 into the chloroplast in a proton-dependent manner, converted to 3-PGA under ATP consumption. 3-PGA is converted to TP under NADPH consumption. NADP is regenerated to NADPH at the photosynthetic electron transport chain under water consumption and oxygen evolution (red arrow). 3-PGA: 3-Phosphoglycerate; TP: triose phosphates.

SI Fig. 5. Photorespiratory phenotype of *plgg1-1* seedlings. Seedlings from *Arabidopsis thaliana* Col-0 (WT), the photorespiratory mutant *shm1-1* (impaired in mitochondrial SHMT) and the T-DNA insertion line *plgg1-1*. Plants were grown on MS-media under elevated CO2 concentrations (3000 ppm) for ten days (left panel) or for four days and shifted to ambient CO2 (300 ppm) for six days (right panel).

SI Fig 6. Phylogenetic Tree of PLGG1. AtPLGG1 was used as the query to Blast Explorer (http://www.phylogeny.fr/version2_cgi/one_task.cgi?task_type=blast) and the sequences suggested by the program were manually curated and fed to an "à la carte" phylogeny analysis (http://www.phylogeny.fr/version2_cgi/phylogeny.cgi). Briefly, sequences were aligned with Muscle and all positions with gaps were removed. The phylogenetic tree was calculated with PhyML and bootstrapped (100 repetitions). For display, all branches with <30% branch support were collapsed.

The tree divides in two large branches, a bacterial and an eukaryotic branch. The sole cyanobacterium *Acaryochloris* spec is nested within the bacteria. In eukaryotes, the tree divides the fungi (more precisely the dikarya) from the archaeplastida. Within the archaeplastida, *Cyanophora paradoxa* separates from the red and green lineages which themselves form separate branches. In the green lineage, the green algae separate from the streptophyta in which the moss branches before the fern. Monocots and dicots form separate branches.

Since the fungi branch with the archaeplastida with a bootstrap support of 62%, the gene likely evolved before the archaeplastida and the fungi split and it was lost from the other branches of life. Had the gene been acquired by a lateral gene transfer, one would expect it to branch within the bacteria. In higher plants PLGG1 is nuclear encoded. For transport proteins of the chloroplasts, at least three different phylogenetic roots have been demonstrated: rooted in the Chlamydiae, rooted in the cyanobacteria and a host origin. PLGG1-like proteins do not appear in the Clamydiae but in Cyanobacteria. The phylogenetic tree indicates that the gene is likely not acquired from the cyanobacteria since the single cyanobacterium PLGG1 is nested within the bacteria. Therefore it is more likely that the cyanobacterium acquired it by lateral gene transfer from a bacterium. Hence PLGG1 likely originated from the host.

(1) Howitz KT & McCarty RE (1991) Solubilization, Partial-Purification, And Reconstitution Of The Glycolate Glycerate Transporter From Chloroplast Inner Envelope Membranes. *Plant Physiology* 96(4):1060-1069.

(2) Nozawa A et al. (2007) A cell-free translation and proteoliposome reconstitution system for functional analysis of plant solute transporters. *Plant Cell Physiol* 48:1815–1820.

(3) Dereeper A., Guignon V., Blanc G., Audic S., Buffet S., Chevenet F., Dufayard J.-F., Guindon S., Lefort V., Lescot M., Claverie J.-M., Gascuel O. *Phylogeny.fr: robust phylogenetic analysis for the non-specialist* Nucleic Acids Research. 2008 Jul 1; 36 (Web Server Issue):W465-9. Epub 2008 Apr 19.

SI Table 1: Primer used for PCR amplification. All sequences are displayed in 5'-3' orientation.

Primer name	Sequence (5'-3')
P1	CGTCGTCGTCTCCATACCCAT
P2	GTTTTGCCATAGGCTCGGCTT
P3	GCGTGGACCGCTTGCTGCAACT
P4	CACCATGGCTACTCTTTTAGCCACTCC
P5	GACACCTGAAGCAGCCGG
P6	TTCAATGTCCCTGCCATGTA
P7	TGAACAATCGATGGACCTGA
P8	CACCGGAGCATGAGTACCGTTACAATTG
P9	GACACCTGAAGCAGCCGG
P10	CACCGGAGCATGAGTACCGTTACAATTG
P11	ACTAACGGCCTTCGAGGAT
P12	CACCATGGCTACTCTTTTAGCCACTCC
P13	GACGACCGCTAGCAAACTCTG

SI Table 2: Plants were germinated and grown under different light regimens and both the size difference compared to wildtype and the percentage of plants with lesions were scored.

While very young plants up to two weeks of age benefit from extended light periods which alleviate the phenotype, plants from week three on suffer from extended light periods. nd not determined

		8h L, 16h D	16h L, 8 hD	24h L
14 days	size ratio WT/ mutant	1.58	1.48	1.36
	percentage of mutant plants with lesions	50%	10%	nd
21 days	size ratio WT/ mutant	1.94	1.94	1.49
	percentage of mutant plants with lesions	60%	90%	nd

Metabolite extraction.

For metabolite analysis, 100mg (fresh weight) of powdered tissue was extracted with 700uL of methanol for 15 minutes at 70°C. 700uL of water and 375uL of chloroform were added and the samples were incubated in a rotating shaker for 30 minutes at 4°C (1). The phases were separated and the methanol/water phase retained for further experiments. Amino acid contents were determined directly from the extract exactly as described in (2) except that the samples were extracted in 1:1 (v/v) methanol/water. Polar metabolites were analyzed as described by (1). Briefly, 50ul of the extract was dried in a speed-vac and carbonyl moieties were protected by methoximation using a 20 mg/mL solution of methoxyamine hydrochloride in pyridine at 30 °C for 90 min. Acidic protons were modified with N-methyl-N-trimethylsilyltrifluoroacetamide (MSTFA) at 37 °C for 30 min. Samples were injected at 230°C, separated on HP5 columns for 30 minutes and analyzed from $m/z=50$ to $m/z=800$ in total ion scans in an Agilent technologies 5973 Inert mass spectrometer. Metabolites were identified by comparison to the retention time and the fractionation pattern of standards and quantified with external standard curves of complex standards. The complex standard for LC/MS contained all amino acids. The complex standard for GC/MS contained glucose, fructose, glycolate, succinate, glycerate, fumarate, malate, 2-oxoglutarate, citrate, isocitrate, myo-inositol, sucrose, glycine, serine, pyruvate, maltose and ribitol at concentrations of 0.1, 0.5, 1, 5, 10, 50 and 100 μM. The complex standard was injected prior to the randomized samples, after injection of 21 samples and at the end of the measurement. Retention times and peak size for the complex standards remained stable.

Metabolite extraction optimized for photorespiratory metabolites.

The extraction procedure was conducted at 4°C. Fifty-milligram of powdered material was extracted with 600 μl of ice-cold N,N-dimethiylformamide. Following the addition of 400 μl of water, the samples were shaken for 10 min and separated into phases by centrifugation for 8 min at 22,000x*g*. The upper aqueous phase was mixed with 600 μl of xylene for 10 min and centrifuged for 3 min. Upper organic phase was discarded and 300 μl of lower aqueous phase was dried for the analysis. The GC-MS analysis was conducted exactly as described by Lisec et al. (2006) (3). Chromatograms and mass spectra were evaluated by Chroma TOF® 1.6 (Leco, St Joseph, MI) and TagFinder 4.0 (4) for the quantification and annotation of the peaks. The [18]O enrichment was calculated according to Berry et al. (5)

1.   Fiehn O, Kopka J, Trethewey RN, & Willmitzer L (2000) Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. *Analytical Chemistry* 72(15):3573-3580.

2.   Lu Y*, et al.* (2008) New connections across pathways and cellular processes: Industrialized mutant screening reveals novel associations between diverse phenotypes in Arabidopsis. *Plant Physiology* 146(4):1482-1500.

3.   Lisec J, Schauer N, Kopka J, Willmitzer L, & Fernie AR (2006) Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nature Protocols* 1(1):387-396.

4.   Luedemann A, Strassburg K, Erban A, & Kopka J (2008) TagFinder for the quantitative analysis of gas chromatography - mass spectrometry (GC-MS)-based metabolite profiling experiments. *Bioinformatics* 24(5):732-737.

5.   Berry JA, Osmond, C.B., and Lorimer, G.H. (1978) Fixation of 1802 during Photorespiration. *Plant Physiology* 62:954-967.

6.2 Second publication: *PLGG1*, a plastidic glycolate glycerate transporter, is required for photorespiration and defines a new class of metabolite transporters

Status: **Submitted** (August 2012)

Thea R. Pick, Andrea Bräutigam, Matthias A. Schulz, Toshihiro Obata, Alisdair R. Fernie, and Andreas P.M. Weber

Journal:" Proceedings of the National Academy of Sciences of the United States of America"

Impact factor: 9.681

1. Co-Author

Own contribution: 60%

- Visualization of the phenotypes
- GFP Localization
- Data analysis
- Oxygen evolution
- $^{18}O_2$ labeling and metabolite measurement and analysis
- Manuscript writing

Danksagungen

Zu guter Letzt möchte ich den Menschen danken, ohne die ich meine Promotion nicht hätte abschließen können und denen ich zu tiefstem Dank verpflichtet bin.

Vielen Dank...

... **Prof. Dr. Andreas Weber** für Möglichkeit an Ihrem Institut promovieren zu können und ein Teil des Weber-Universums zu sein. Sie haben mir ermöglicht an begeisternden und spannenden Forschungsthemen zu arbeiten, die keine „Bad Projects" waren und im Rahmen der PROMICS und OPTIMAS meetings interessante Forscher zu treffen, an lehrreichen Diskussionen teilzunehmen und Amerika zu entdecken. In Ihrem Institut haben Sie mir ermöglicht mich als Wissenschaftler weiter zu entwickeln und weiterhin so viel Spaß an der Wissenschaft zu haben.

... **Prof. Dr. Peter Westhoff** für die Übernahme des Koreferats, hilfreiche Denkanstöße und Ratschläge als Committee member.

... **Andrea**, dass du mir das interessante Mep1 Thema überlassen hast, so viel Geduld mit meinen holprigen Schreib-Versuchen hattest und meine Doktorarbeit betreut hast.

... allen **iGRAD plant** Mitgliedern und ganz besonders **Dr. Sigrun Wegener-Feldbrügge** für anregende Diskussionen und sehr unterhaltsame Retreats.

... der PROMICS-Crew **Marion, Christian und Daria** für schöne gemeinsame Meetings und Konferenzen.

... **Frau Nöcker** für Ihre unkomplizierte Hilfe bei organisatorischen Angelegenheiten.

... den **Gärtnern** für die Aufzucht und Pflege der Pflanzen.

... **Katja** für die großartigen und unendlich lustigen Geschichten, einmaligen Parodien, unterhaltsamen Progress-Reports, die lustigen Kniffel-Spiele zwischendurch und vor allem die SOS E-Mail Hilfe, die du so oft leisten musstest und wahrscheinlich auch noch weiter musst.

... **Katrin**, meiner Lauf-Partnerin, für die schönen Gespräche. Das Joggen mit dir macht einfach riesigen Spaß und ich bin froh, dass wir uns als Lauf-Team gefunden haben.

... **Daria**, aka Chrobonaut, Chrobofant, Chrobosaurier oder einfach Chrobo, für ihre lustigen, unterhaltsamen Aktionen und Geschichten und die entspannte Zeit als Zimmer-Partnerin in Warnemünde.

... meiner **Kaffee-Familie**. Ihr seid einfach die Besten und die Pausen mit euch sind einfach wunderschön und keeelasse!

... **Jan, Christian, Sarah, Ali, Nadine, Chris, Sam, Alex, Nicole, Fabio** und allen anderen, die die Mittagspausen immer zu einem sehr unterhaltsamen Erlebnis gemacht haben und hoffentlich weiterhin machen werden.

... allen weiteren aktuellen und ehemaligen „**Weber-Knechten**" für die tolle, unvergessliche Zeit, das angenehme Arbeitsklima und produktive Miteinander, die tollen Betriebsausflüge und die schönsten Geburtstags-Ständchen, die man bekommen kann.

... **Lisa** für ihre unkomplizierte, lockere, selbstlose und unendlich liebe Art. Ich hoffe, dass unsere Freundschaft noch lange so schön bleibt wie sie ist, egal wo auf der Welt wir uns grade rumtreiben sollten. Bleib so, wie du bist, denn so bist du einfach super!

... **Dominik**, für so vieles. Deinen Kaffee-Service, die für dich so selbstverständliche Hilfe (die überhaupt nicht selbstverständlich ist), das Rumhampeln beim Sporti, die Spieleabende, das Rumalbern im Labor, die gemeinsamen Skate-Nights, die leckeren Rezepte, die ständigen peinlichen Aktionen, die mich so herzhaft lachen lassen.... Du bist einfach immer so lustig und herzlich und gleichzeitig auch immer für ein gutes Gespräch zu haben. Ich hoffe, dass das noch lange so schön bleibt.

... **Mareike**, für deine sonnige und optimistische Art. Wenn ich an dich denke muss ich irgendwie immer lächeln, weil deine unerschütterlich gute Laune so ansteckend ist. Die Zeit mit dir ist immer so schön – egal auf welcher Welthalbkugel wir grade zusammen sind :-) und du bist meine absolute Postkarten-Königin! Ich hoffe, dass wir weiterhin so viel zusammen erleben und lachen wie bisher.

... **Manuel**, weil du einfach das Beste bist, was mir je passiert ist. Du hilfst mir so sehr in allen Lebenslagen ruhig zu bleiben und die richtigen Prioritäten zu setzen. Mit dir ist das Leben so relaxt und schön und aufregend und spannend zugleich. Und vor allem lustig... :-) mit dir werde ich immer ein Lächeln auf den Lippen haben...

... meiner Mutter **Hanne**, meinem Vater **Rolf** und meiner Schwester **Lena**, weil ihr mich zu dem Menschen gemacht habt, der ich jetzt bin. Mit euch konnte ich mich frei entfalten und wusste immer, dass ich einen Rückzugsort und sicheren Hafen habe in dem ich geliebt werde. Ich freue mich schon auf ganz ganz viele gemeinsame Familienurlaube in Frankreich und alles was wir in der Zukunft zusammen erleben werden.