


RESEARCH PAPER

The role of alanine and aspartate aminotransferases in C₄ photosynthesis

U. Schlüter¹, A. Bräutigam², J.-M. Droz³, J. Schwender⁴ & A. P. M. Weber¹ 

¹ Institute of Plant Biochemistry, Cluster of Excellence on Plant Science (CEPLAS), Heinrich Heine University Düsseldorf, Düsseldorf, Germany

² Computational Biology, Centre for Biotechnology, University Bielefeld, Bielefeld, Germany

³ Nebion AG, Zürich, Switzerland

⁴ Biology Department, Brookhaven National Laboratory, Upton, New York, USA

Keywords

aminotransferase; C₄ photosynthesis; evolution.

Correspondence

A. P. M. Weber, Institute of Plant Biochemistry, Cluster of Excellence on Plant Science (CEPLAS), Heinrich Heine University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany.
E-mail: andreas.weber@uni-duesseldorf.de

Editor

D. Staiger

Received: 3 April 2018; Accepted: 15 August 2018

doi:10.1111/plb.12904

ABSTRACT

- Alanine and aspartate are essential transfer metabolites for C₄ species of the NAD-malic enzyme and phosphoenolpyruvate carboxykinase subtype. To some degree both amino acids are also part of the metabolite shuttle in NADP-malic enzyme plants. In comparison with C₃ species, the majority of C₄ species are therefore characterised by enhanced expression and activity of alanine and aspartate aminotransferases (AT) in the photosynthetically active tissue. Both enzymes exist in multiple copies and have been found in different subcellular compartments. We tested whether different C₄ species show preferential recruitment of enzymes from specific lineages and subcellular compartments.
- Phylogenetic analysis of alanine and aspartate ATs from a variety of monocot and eudicot C₄ species and their C₃ relatives was combined with subcellular prediction tools and analysis of the subsequent transcript amounts in mature leaves.
- Recruitment of aspartate AT from a specific subcellular compartment was strongly connected to the biochemical subtype. Deviation from the main model was however observed in *Gynandropsis gynandra*. The configuration of alanine AT generally differed in monocot and eudicot species. C₄ monocots recruited an alanine AT from a specific cytosolic branch, but eudicots use alanine AT copies from a mitochondrial branch.
- Generally, plants display high plasticity in the setup of the C₄ pathway. Beside the common models for the different C₄ subtypes, individual solutions were found for plant groups or lineages.

INTRODUCTION

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco), the central enzyme of photosynthetic CO₂ assimilation, evolved in a CO₂-rich and O₂-poor environment (Sage & Stata 2015). The spread of O₂-producing photosynthetic organisms and geological processes caused considerable changes in the content of the atmosphere, which currently contains approx. 21% O₂ and 0.04% CO₂ (Shih 2015). Rubisco can react with O₂ as well as CO₂, and in C₃ species, every third to fourth molecule of RuBP is oxygenated rather than carboxylated in current atmospheric conditions (Sharkey 1988). In the oxygenation reaction, 2-phosphoglycolate (2PG) is formed and has to be metabolised *via* the photorespiratory pathway. This process is energy-demanding, and of the four carbon atoms entering the pathway in the form of two molecules of 2PG only three are recycled into one molecule of 3PGA, whereas 25% of the organic carbon is lost as CO₂ (Hagemann & Bauwe 2016). Under conditions of high temperature and low water availability, when photorespiration is high, a variety of plant species evolved mechanisms to repress photorespiration by concentrating CO₂ around the Rubisco, most successfully *via* C₄ photosynthesis (Sage 2004; Sage *et al.* 2011).

During C₄ photosynthesis, CO₂ is fixed first by phosphoenolpyruvate carboxylase (PEPC), an enzyme with no

oxygenase activity but high affinity for bicarbonate, into a C₄ metabolite. The C₄ metabolite is then transported into the vicinity of Rubisco where its decarboxylation creates a CO₂-rich atmosphere. The C₃ reaction product returns for regeneration to the site of the PEPC. An efficient C₄ shuttle mechanism is achieved by coordinated architectural and biochemical adjustments. The components required for such carbon shuttles are present already in all C₃ plant species (Aubry *et al.* 2011; Reyna-Llorens & Hibberd 2017), they just acquire novel expression and activity patterns. This made it possible that despite its complexity, C₄ photosynthesis could evolve over 60 times independently in the plant kingdom (Sage *et al.* 2011; Sage 2016). The evolution from C₃ to C₄ relies on multiple steps influencing the quantitative, spatial and temporal regulation of selected enzymes and/or their kinetics (Heckmann *et al.* 2013; Mallmann *et al.* 2014).

The majority of the C₄ species separate PEPC and Rubisco reaction sites into different cell types, the mesophyll and bundle sheath cells, organised in a Kranz-like pattern, which allows short paths for efficient metabolite exchange. All C₄ species recruited PEPC as the primary CO₂-fixing enzyme and pyruvate orthophosphate dikinase (PPDK) in coordination with AMP kinase (AMK) and pyrophosphatase (PPase) for regeneration of PEP (Schlüter *et al.* 2016). For the decarboxylation

step, however, three different enzymes have been recruited into the C₄ cycle in the different lineages: (i) phosphoenolpyruvate carboxykinase (PCK), (ii) NAD-dependent malic enzyme (NAD-ME), and (iii) NADP-dependent malic enzyme (NADP-ME). The majority of species actually employ a mixture of NADP-ME and PCK, or PCK and NAD-ME activities (Wingler *et al.* 1999; Furbank 2011; Pick *et al.* 2011; Wang *et al.* 2014b). The specific nature of the decarboxylation enzymes further requires the coordinated activity of additional enzymes and transporters (Schlüter *et al.* 2016). So far research has concentrated mainly on characterisation of the C₄-specific carboxylation and decarboxylation steps (Ku *et al.* 1996; Westhoff & Gowik 2004; Hibberd & Covshoff 2010; Maier *et al.* 2011), but for complete understanding of the whole pathway more detailed knowledge about the additional players is necessary. In this paper, we concentrate on the role of the alanine (Ala) and aspartate (Asp) aminotransferases (AT) recruited into the C₄ shuttle. Ala and Asp ATs are essential for functionality of the NAD-ME and PCK C₄ pathway, but also show enhanced expression in NADP-ME monocot and eudicot species (Gowik *et al.* 2011; Pick *et al.* 2011; Covshoff *et al.* 2016). Both ATs are already involved in primary metabolism in C₃ species and they are present in all plant species in multiple copies.

The Ala ATs catalyse the reversible, pyridoxal phosphate-dependent transfer of an amino group from alanine (Ala) to 2-oxoglutarate (2OG), forming glutamate and pyruvate. Ala AT isozymes have been associated with mitochondrial, cytosolic and peroxisomal fractions of the cell (Igarashi *et al.* 2003; Liepman & Olsen 2003). In *Arabidopsis thaliana*, two isozymes with Ala AT-specific activity exist (Ala AT1 and Ala AT2), and two closely related isozymes with additional glutamate:glyoxylate aminotransferase (GGT1 and GGT2) activity are present in the peroxisomes (Nakamura & Tolbert 1983; Igarashi *et al.* 2003; Liepman & Olsen 2003). GGTs have an essential function in the photorespiratory pathway (Liepman & Olsen 2003). Ala ATs are usually associated with Ala synthesis and catabolism; however, they do not seem to be essential for plant metabolism (Miyashita *et al.* 2007). An Ala AT1 knock-out mutant showed a dramatic reduction in AT activity, but no visible phenotype under control conditions (Miyashita *et al.* 2007), and also the double mutant (*alaat1/alaat2*) was viable (McAllister & Good 2015). The overlapping activity of various aminotransferases (GGT, GABA transaminase, Asp AT) can probably compensate for the loss of Ala AT specific activity (Rocha *et al.* 2010). A specific role for Ala ATs had been related to seed dormancy (Sato *et al.* 2016) and to specific stress responses, especially oxygen deprivation and nitrogen limitation (Shrawat *et al.* 2008; Rocha *et al.* 2010). Under O₂ deprivation, Ala accumulates in the affected plant tissue, and the Ala AT plays an important role in the regulation of plant C, N and energy metabolism (Rocha *et al.* 2010). Ala ATs are also responsible for the fast Ala catabolism and provision of pyruvate during re-aeration of anoxic plant tissue. In the *alaat1* mutant, Ala breakdown was delayed during the recovery phase (Miyashita *et al.* 2007). The influence of Ala AT on N metabolism was shown in rice plants, where overexpression of an Ala AT from barley considerably increased the N use efficiency, plant biomass and grain yield (Shrawat *et al.* 2008).

The Asp ATs transfer the amino group from Asp to 2OG, producing glutamate and oxaloacetate (OAA) in a reversible pyridoxal phosphate-dependent reaction. In *A. thaliana*, Asp

ATs could be localised to various cell compartments, including the mitochondria (ASP1), cytosol (ASP2, ASP4), the peroxisome (ASP3) and the plastid (ASP5) (Schultz & Coruzzi 1995; Wilkie *et al.* 1995; Miesak & Coruzzi 2002). An additional enzyme related to prokaryotic enzymes and having Asp AT as well as prephenate aminotransferase activity is present in the plastids (de la Torre *et al.* 2014). Asp AT activity plays a central role in the regulation of plant C and N metabolism, Asp serves as the substrate for plastidial production of other amino acids and numerous other plant metabolites. Silencing of the prephenate AT gene caused severe growth effects and chlorosis, and synthesis of asparagine, phenylalanine and lignin was reduced (de la Torre *et al.* 2014). Mutants of the plastidial Asp AT on the other hand showed no visible changes in their phenotype (Miesak & Coruzzi 2002; de la Torre *et al.* 2014). The organellar Asp ATs are mainly associated with a role in the shuttling of reducing equivalents, the cytosolic form is involved in Asp synthesis in the light and asparagine synthesis in the dark (Miesak & Coruzzi 2002).

Conceptual and computational models of C₄ evolution indicate that ATs play an important role already during early phases of the transition from C₃ to C₄ photosynthesis. Analysis of a variety of C₃–C₄ intermediate species indicates that the interruption of the photorespiratory glycine decarboxylase reaction in the mesophyll tissue represents a decisive initial step for the evolution of C₄ photosynthesis (Rawsthorne *et al.* 1988; Mallmann *et al.* 2014), causing the establishment of a glycine or C₂ shuttle. Decarboxylation of glycine exclusively in the bundle sheath creates an atmosphere of elevated CO₂ at the site of the bundle sheath Rubisco. The glycine transport step, however; not only translocates C, but also N, and hence additional metabolite transport would be required for N rebalancing. Modelling approaches identified glutamate/2OG, Ala/pyruvate and Asp/malate as the most likely transport metabolites, and transcriptomic as well as biochemical analyses of various *Flaveria* species support such a scenario (Mallmann *et al.* 2014). In coordination with the decarboxylating NADP-ME enzyme, Ala AT was already up-regulated in the least advanced *Flaveria* C₃–C₄ intermediates, allowing the establishment of a low capacity C₄-like cycle. Different Asp AT and malate dehydrogenase (MDH) copies increased at later stages of C₃–C₄ intermediates, and PEPC and PPDK reach maximum activity finally in C₄ (Mallmann *et al.* 2014).

The aim of this study was a detailed investigation of Ala AT and Asp AT recruitment into the C₄ pathway. We set out to examine: (i) which isoforms of the ATs were recruited into the C₄ pathways in the different lineages; and (ii) if the choice of C₄ ATs can be connected to the biochemical model of the pathway. Ala and Asp AT transcript patterns were recovered from published datasets and supplemented with new data from additional C₄ species representing the different C₄ subtypes from monocot and eudicot species. The isozymes recruited to C₄ photosynthesis were identified using phylogenetic analyses of the respective gene families.

MATERIAL AND METHODS

Phylogenetic analyses

The coding sequences for the Ala and Asp ATs were retrieved from TAIR database for *A. thaliana*, from the MaizeGDB

version 4 (maizegdb.org) for *Zea mays*, from the Amborella Genome Database (<http://amborella.huck.psu.edu/project>) for *Amborella trichopoda*, and from the Phytozome database version 12 (<https://phytozome.jgi.doe.gov/pz/portal.html>) for *Glycine max*, *Vitis vinifera*, *Aquilegia coerulea*, *Oryza sativa*, *Sorghum bicolor*, *Setaria italica*, *Musa acuminata*, *Selaginella moellendorffii*, *Marchantia polymorpha*, *Volvox carteri* and *Chlamydomonas reinhardtii*. The sequences for *Dicranthelium clandestinum*, *Megathyrus maximus*, *Atriplex prostrata*, *Atriplex rosea*, *Amaranthus edulis*, *Panicum miliaceum*, *Alloteropsis semi-alata* and *Alloteropsis eckloniana* were taken from the contigs assembled by Trinity (Haas *et al.* 2013) from transcriptome reads. The sequences for ATs of *Tarenaya hassleriana* and *Gynandropsis gynandra* were assembled from the transcriptomes presented in Kulahoglu *et al.* (2014). The phylogenetic trees were constructed using the SeaView4 software (Gouy *et al.* 2010). The workflow consisted of sequence alignment by Muscle, maximum-likelihood tree generation by PhyML with bootstrap replicate values of 100 and visualisation by FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>). The sequences selected for the phylogenetic trees were also used for prediction of the subcellular localisation using the webtools TargetP (<http://www.cbs.dtu.dk/services/TargetP/>) and MultiLoc2 (abi.inf.uni-tuebingen.de/Services/MultiLoc2). All Ala and Asp AT sequences deriving from new assemblies are presented in Data S1.

Alanine and aspartate aminotransferase transcript abundances

The Ala and Asp ATs and MDH patterns were determined from transcriptome analysis published in Eisenhut *et al.* (2017) for *A. thaliana* wild type ambient CO₂; in Kulahoglu *et al.* (2014) for *Tarenaya hassleriana* and *Gynandropsis gynandra*, mature leaf; in Wang *et al.* (2014a) for *Oryza sativa*, leaf sections 7, 8, 9 and 10; in Chang *et al.* (2012) for *Zea mays* mesophyll and bundle sheath tissue; in Denton *et al.* (2017) for *Z. mays* mesophyll and bundle sheath tissue in a leaf gradient; in Rao *et al.* (2016) for *Panicum virgatum* mesophyll and bundle sheath tissue; in John *et al.* (2014) for *Setaria viridis* mesophyll and bundle sheath tissue; in Aubry *et al.* (2014) for *G. gynandra* mesophyll and bundle sheath tissue; in Doring *et al.* (2016) for *Sorghum bicolor* total leaf tissue. The graphs show the transcript abundances in transcript per million (tpm) as presented in the Supplements of Denton *et al.* (2017) for *Z. mays*, *P. virgatum* and *S. viridis*. Further Ala and Asp ATs transcript abundances were obtained from new transcriptome studies in *Atriplex rosea*, *A. prostrata*, *Amaranthus edulis*, *Dicranthelium clandestinum*, *Megathyrus maximus*, *Setaria italica*, *Alloteropsis eckloniana*, *A. semi-alata*, *Flaveria robusta*, *F. bidentis*, *Panicum miliaceum* (Tables S1, S2). The plants were grown during the summer months in the greenhouses of the Heinrich Heine University in Düsseldorf, Germany; the rooftop greenhouse at the University of Toronto, Canada; and controlled climate chambers at the University of Sheffield, UK. Mature plant tissue was harvested at midday in three replicates. RNA was extracted using the Plant RNeasy kit (Qiagen, Hilden, Germany) and quality controlled using the Agilent 2100 Bioanalyzer (Agilent Genomics, Santa Clara, CA, USA). Paired-end sequencing reads were assembled with Trinity using default parameters. Expression was quantified by back mapping reads using the software RSEM with bowtie1 and normalisation to tpm. The AT specific sequences were identified by BLAST with the corresponding

sequences from *A. thaliana* for eudicots and *Setaria italica* for monocots. For contigs with very high sequence homology and neighbouring position in the tree only the contig with highest read number was included into the final phylogenetic tree. The *Flaveria* samples were sequenced on an Illumina HiSeq 2000 platform in the single end mode with read length of 100 bp. The reads were then aligned to the *Arabidopsis* coding sequences from TAIR 10 release using BLAT as described by Gowik *et al.* (2011). Enzyme activities for the mesophyll and bundle sheath specific tissue along the leaf gradient were taken from Denton *et al.* (2017).

RESULTS

Phylogeny of the plant aspartate aminotransferases

Sequences for the Asp ATs from all cellular compartments and the non-related plastidial prephenate AT genes from various C₃ and C₄ monocot and eudicot species were taken from public databases and new transcriptome experiments. In the phylogenetic tree, the prephenate ATs formed a clear outgroup (Figs. 1 and S1). The plant Asp ATs could be further separated into three main branches, a plastidial and mitochondrial and a branch containing the cytoplasmic and peroxisomal forms, all of them containing members from monocot and eudicot species, as well as C₃ and C₄ plants. Asp ATs from algae and mosses also clustered into the mitochondrial and plastid branch of the tree, indicating that the separation into these organellar forms occurred early in evolution. Within the branches, the investigated species had different copy numbers. The eudicot *A. thaliana* possessed one plastidial and one mitochondrial isoform. Three members of the family could be found in the third Asp AT branch, one of them could be localised to the peroxisome, and the two other very closely related copies were cytoplasmic (Schultz & Coruzzi 1995). The monocot *O. sativa* on the other hand had two mitochondrial copies, one plastidial and also one copy in the cytoplasmic/peroxisomal branch. The C₄ species *Z. mays* showed only one representative in each branch. Generally, vascular plants seemed to have at least one enzyme in each main branch.

The prediction programs mainly supported subcellular localisation of the sequences within the plastidial and mitochondrial subgroups (Fig. 1, Tables S3, S4). As a caveat, for some of the newly assembled contigs used in the tree, the predicted protein sequences might be incomplete, and hence important information for the sorting into the organelles might be missing.

The isoforms, which had been recruited into the C₄ shuttle, were marked in blue and allowed the identification of a C₄ subtype-specific pattern (Fig. 1). The plastidial Asp ATs were only up-regulated in the NADP-ME subtypes, the cytoplasmic Asp AT form in the PCK and NAD-ME subtypes. Additionally, the mitochondrial Asp AT was enhanced in the NAD-ME subtype species. The only exception to this pattern was found in the NAD-ME species *Gynandropsis gynandra*, where only one Asp AT copy from the mitochondrial branch showed considerably high abundance.

Phylogeny of the plant alanine aminotransferases

Plant Ala ATs were very closely related to and showed overlapping specificity with the peroxisomal glutamate:glyoxylate

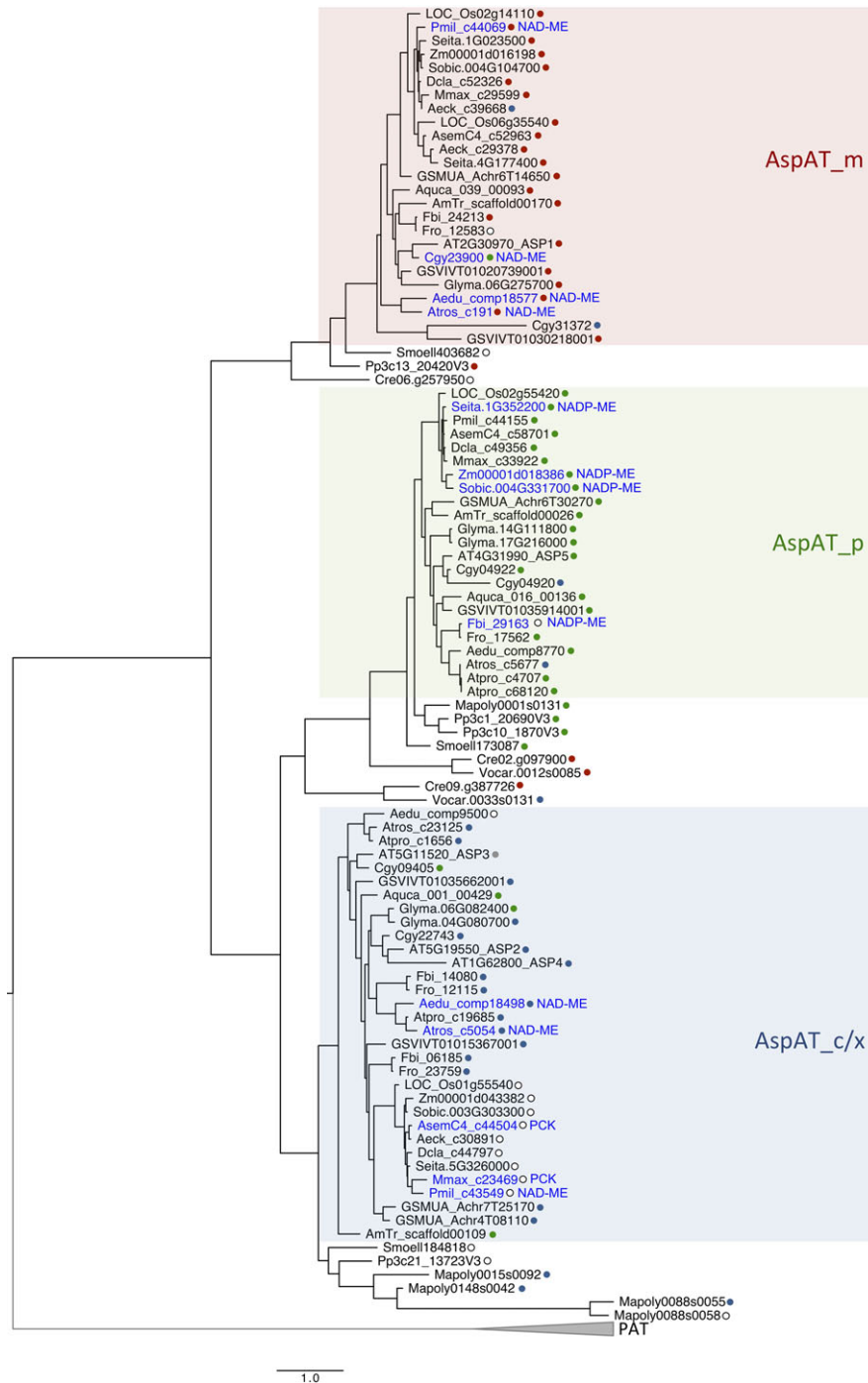


Fig. 1. Phylogenetic tree of the Asp aminotransferase coding sequences in selected C₃ and C₄ species. The copies up-regulated in C₄ species are highlighted in blue. The coloured dots next to the sample names indicate their subcellular prediction: red – mitochondria, blue – cytosol, grey – peroxisome, green – plastid, white – unclear. Bootstrap support values with 100 replicates are shown in Fig. S2. Sequences represent the following species: *Chlamydomonas reinhardtii* (Cre), *Physcomitrella patens* (Pp), *Volvox carteri* (Vocar), *Selaginella moellendorffii* (Smoell), *Marchantia polymorpha* (Mapoly), *Amborella trichopoda* (AmTr), *Aquilegia coerulea* (Aqua), *Arabidopsis thaliana* (AT), *Atriplex rosea* (Atros), *Atriplex prostrata* (Atpro), *Amaranthus edulis* (Aedu), *Glycine max* (Glyma), *Flaveria bidentis* (Fbi), *Flaveria robusta* (Fro), *Gynandropsis gynandra* (Cgy), *Vitis vinifera* (GSVIVT), *Musa acuminata* (GSMUA), *Dichantheium clandestinum* (Dcla), *Megathyrus maximus* (Mmax), *Alloterospis semialata* (AsemC4), *Alloterospis eckloniana* (Aeck), *Panicum miliaceum* (Pmil), *Setaria italica* (Seita), *Sorghum bicolor* (Sobic), *Oryza sativa* (Os), *Zea mays* (Zm). The prephenate aminotransferase branch is abbreviated as PAT.

aminotransferases (GGT), which are involved in the photorespiratory pathway. The separation into GGT and Ala AT specific forms seems to have occurred early in evolution and

branch-specific sequences could already be found in mosses (Figs. 2 and S2). For the Ala ATs, the phylogenetic pattern was very different in monocot and eudicot species. In the

monocots, the *Poaceae* species formed at least four main branches (P1 to P4). The genes from banana and *Amborella trichopoda* localised outside of these main branches (Fig. 2). All tested eudicot Ala AT sequences clustered into just one separate eudicot-specific branch. Multiple copies within this branch, such as the two *A. thaliana* or two *G. gynandra* Ala ATs, could be found next to each other in the tree, indicating a very recent duplication. The subcellular localisation predicted cytoplasmic isoforms for the majority of sequences on the *Poaceae* Ala AT branches P2, P3 and P4; only the sequences in the P1 branch seem to be mitochondrial (Fig. 2, Table S3). The Ala AT branch containing all eudicot Ala ATs clustered close to the P1 branch from the *Poaceae* species; and the majority of eudicot Ala ATs also had predicted mitochondrial localisation, including the C₄ specific forms from *G. gynandra* and *Flaveria bidentis*. Cytosol-specific groups like in the *Poaceae*, on the other hand, were absent from the eudicots. Few eudicot Ala AT sequences were predicted to be cytosolic (GSVIVT010166669001, Aedu_copm18481, Atpro_c25837), it is however possible that the gene sequences available for these Ala ATs are still incomplete.

Transcript abundances of Ala ATs and Asp ATs in photosynthesising leaf tissue

The AT isoforms recruited into the C₄ pathway could be identified by their strong up-regulation in the photosynthetic organs, such as mature leaf tissue. Although transcriptome patterns cannot be directly translated into enzyme activities, earlier studies clearly showed that the majority of C₄-specific enzymes are transcriptionally regulated.

In mature leaf tissue of all tested species, one Ala AT copy dominated; in the monocot species these were Ala ATs from the *Poaceae* P2 branch, and also in the eudicot branch one isoform was more abundant (Fig. 3, Table S1). In *A. thaliana* leaves for instance Ala AT1 expression was higher than that of Ala AT2 (Miyashita *et al.* 2007). In comparison with C₄ species, all C₃ species showed low Ala AT transcript abundance (Fig. 3A). The GGTs, which are involved in photorespiration, were usually more abundant in C₃ than C₄ leaves. The C₄ species with NADP-ME decarboxylation were characterised by only small increases in Ala AT transcript abundance, which was expected since Asp and Ala are responsible only for a minor part of the C₄ shuttle reactions.

Separation of *Z. mays* leaves into mesophyll and bundle sheath along the leaf gradient revealed changes in Ala AT preferences along the leaf. At the base of the maize leaf, where photosynthetic activity of the tissue is low, a predicted mitochondrial isoform dominated in the mesophyll and bundle sheath (Fig. 4A). The cytosolic Ala AT from the P2 branch then increased in transcript abundance along the leaf in both mesophyll and bundle sheath tissues; the transcriptional increase in Ala AT was mirrored in increased enzyme activity. Towards the tip of the leaf, the Ala AT transcripts were higher in the mesophyll than in bundle sheath tissue. Such differences between Ala AT transcripts in mesophyll and bundle sheath tissue were also pronounced in the *Z. mays* experiment of Chang *et al.* (2012; Fig. 4B). In *S. italica* the bundle sheath fraction also had lower Ala AT transcript abundance than the mesophyll fraction (Fig. 4C). Both species belong to the NADP-ME subgroup. In the experiment with the NADP-ME species *Panicum virgatum*

(Fig. 4D), the Ala AT transcript amounts were generally very low, and it is possible that this was due to so far incomplete gene models for Ala ATs in this species. For dicot species the separation of mesophyll and bundle sheath tissue is more difficult, and mesophyll and bundle sheath specific data are only available for the NAD-ME species *G. gynandra*, where the Ala AT transcript amounts were comparable in the mesophyll and bundle sheath fractions of the leaves (Fig. 4E).

The transcript pattern of Asp AT in the mature leaf tissue of the tested plants was generally less uniform than that for Ala ATs, with different subcellularly localised isoforms dominating in the different photosynthetic subtypes. The transcript amount in C₃ species was lower than in the C₄ species (Fig. 3B).

In the NADP-ME C₄ species such as the monocot species *Z. mays*, *Setaria italica*, *Sorghum bicolor* and the eudicot *Flaveria bidentis*, the transcripts of the plastidial Asp ATs were enhanced. In the NAD-ME C₄ species such as the grass *Panicum miliaceum* and the eudicots *Atriplex rosea* and *A. edulis* increases in both the mitochondrial and a cytoplasmic form were observed, but the NAD-ME eudicot *G. gynandra* showed only increased transcript abundance in the mitochondrial Asp AT form. The PCK C₄ monocot species *Alloteropsis semialata* and *Megathyrus maximus* displayed high abundance only for the cytoplasmic Asp AT isoform (Fig. 3B).

Similar to the Ala AT transcripts, also plastidial Asp AT transcript abundance increased in the mesophyll fraction of the maize leaf along the gradient and this was accompanied by increases in Asp AT activity (Fig. 4F). In the bundle sheath fraction on the other hand Asp AT transcript amounts remained stable along the leaf gradient and were lower than in the mesophyll fraction. The Asp AT activity actually decreased along the leaf and was almost undetectable in the photosynthetically highly active part towards the tip (Fig. 4F). The same difference between Asp AT transcript abundance in mesophyll and bundle sheath fractions was also observed in the study of mature maize tissue (Fig. 4G) and in *S. italica* (Fig. 4H). In contrast to these NADP-ME species, the Asp AT transcript abundance was high in both the mesophyll and the bundle sheath fraction of the NAD-ME species *P. virgatum* and *G. gynandra* (Fig. 4I, J). In *P. virgatum* the isoform from the cytosolic Asp AT branch dominated while the mitochondrial isoform was most prominent in the bundle sheath fraction. In *G. gynandra*, the isoform from the mitochondrial Asp AT branch showed by far the highest abundance in mesophyll and bundle sheath fractions of the leaf (cf Sommer *et al.* 2012).

DISCUSSION

The C₄ monocots recruited an Ala AT from a specific cytosolic branch, but eudicots use Ala AT copies from a mitochondrial branch

The use of Asp as transport metabolite in the C₄ shuttle from the mesophyll to the bundle sheath also requires the back transport of an N-containing C₃ compound to balance the N metabolism. The Asp shuttle is therefore accompanied by transport of Ala from the bundle sheath to the mesophyll cells and the activity of Ala AT in both cell types. In comparison with C₃ species, Ala AT transcript abundance was enhanced in C₄ leaves and usually more abundant than the transcript amount of the

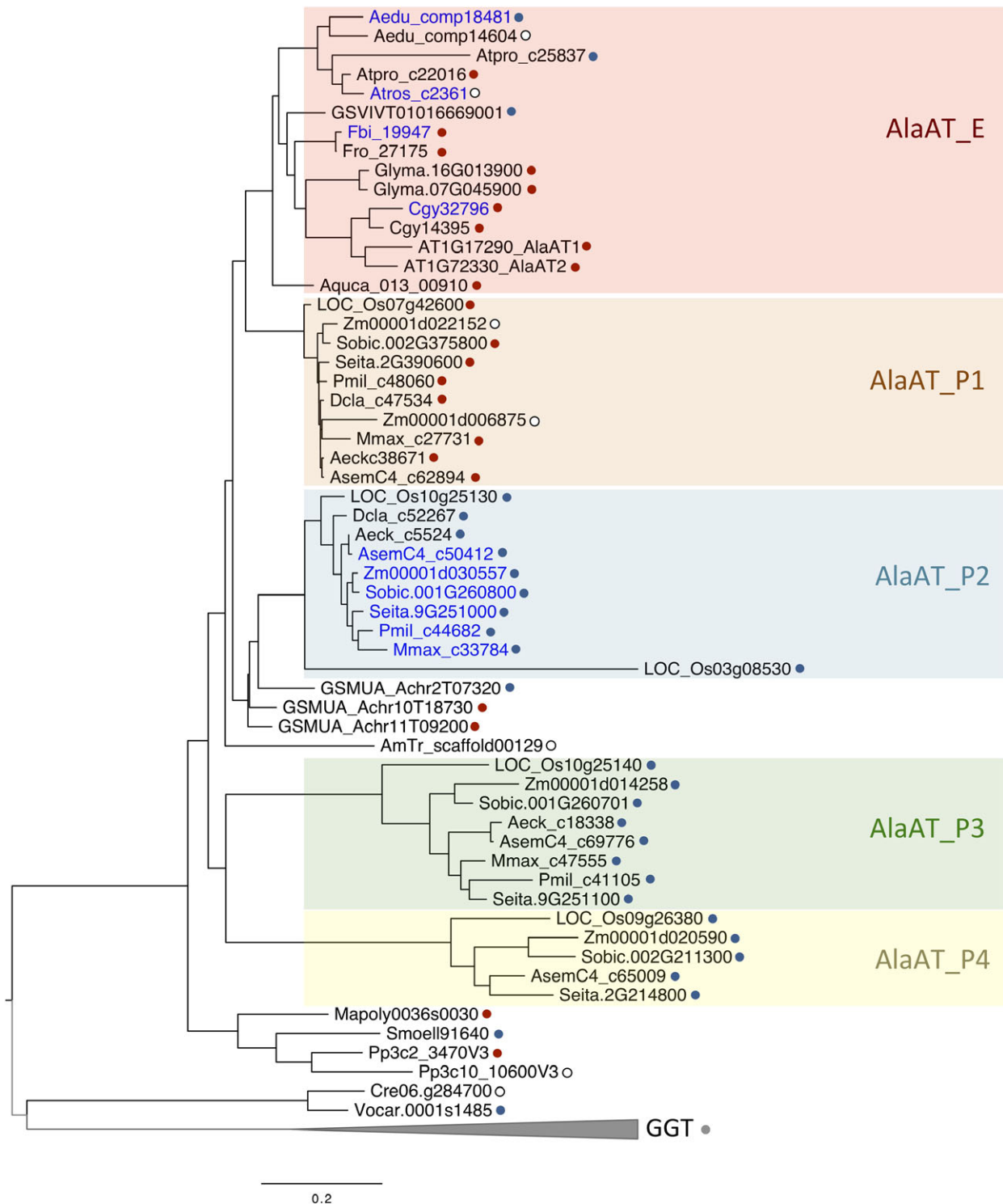


Fig. 2. Phylogenetic tree of the Ala aminotransferase coding sequences in selected C₃ and C₄ species. The copies up-regulated in C₄ species are highlighted in blue. The coloured dots next to the sample names indicate their subcellular prediction: red – mitochondria, blue – cytosol, grey – peroxisome, white – unclear. Bootstrap support values with 100 replicates are shown in Fig. S3. Species abbreviations are the same as in Fig. 1. The glutamate:glyoxylate aminotransferase branch is abbreviated with GGT.

closely related GGT genes. GGT is localised in the peroxisomes, it plays an important role in photorespiration (Igarashi *et al.* 2003; Liepman & Olsen 2003; Niessen *et al.* 2012), and is

therefore in higher demand in C₃ than C₄ species. The ratio in transcript amount of GGT and Ala AT was therefore a good indicator of the photosynthesis type in the tested tissue.

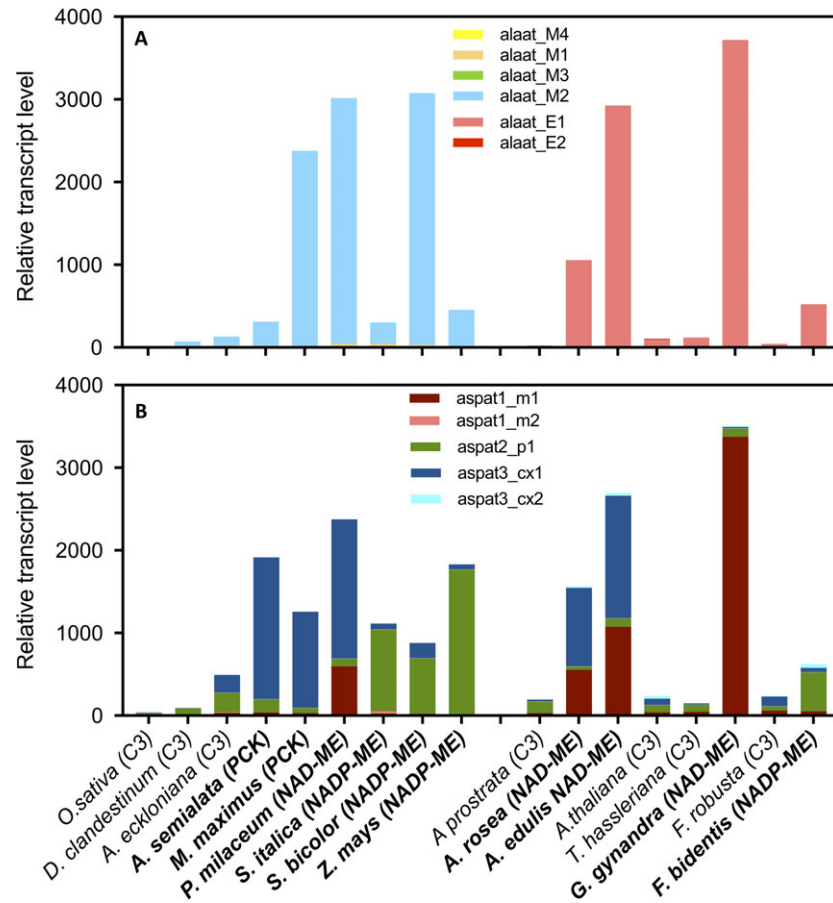


Fig. 3. Ala and Asp aminotransferase transcript amounts in mature leaves of different C₃ and C₄ species. A. Transcript abundance of different Ala aminotransferase copies; B. transcript abundance of different Asp aminotransferase copies. The colouring follows the same schema as in the phylogenetic tree.

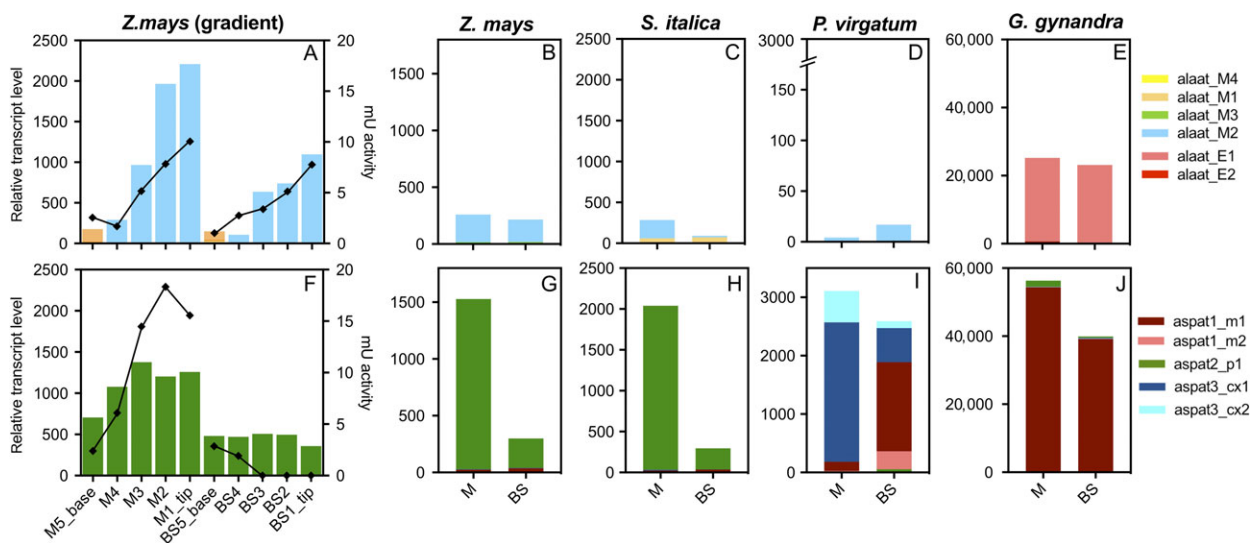


Fig. 4. Mesophyll and bundle sheath specific transcript amounts in mature leaves. Transcript abundance of aminotransferase copies (AT) in mesophyll and bundle sheath of *Zea mays* leaf tissue with NADP-ME photosynthesis along a leaf gradient (A: Ala AT, F: Asp AT), the black dots indicate the associated enzyme activity measured in the same sample (Denton *et al.* 2017); in mature *Z. mays* leaf tissue with NADP-ME photosynthesis (B: Ala AT, G: Asp AT; Chang *et al.* 2012); in mature leaf tissue of *Setaria italica* with NADP-ME photosynthesis (C: Ala AT, H: Asp AT; John *et al.* 2014); in mature *Panicum virgatum* leaf tissue with NAD-ME photosynthesis (D: Ala AT, I: Asp AT; Rao *et al.* 2016); in mature *Gynandropsis gynandra* leaf tissue with NAD-ME photosynthesis (E: Ala AT, J: Asp AT; Aubry *et al.* 2014).

The phylogenetic tree showed that the copy number of Ala ATs differed considerably between monocot and eudicot species. All tested eudicot species possess closely related Ala ATs, probably originating from comparably recent duplication events. In the grasses, on the other hand, multiple branches of Ala ATs were present, and these different Ala AT copies were retained during evolution in the tested grass species. Maize and rice, for instance, possess Ala AT members in all four *Poaceae* branches. Three out of these four Ala AT groups were predicted to localise to the cytosol and only one to the mitochondria. In our study, all tested species recruited Ala AT copies from the same phylogenetic group, with expected cytosolic localisation into the C₄ pathway. Such cytosolic activity of the C₄ Ala AT is in agreement with the current models for all C₄ subtypes (Aubry *et al.* 2011; Maier *et al.* 2011; Ludwig 2016; Covshoff *et al.* 2016; Fig. 5). It would be interesting to know why all tested C₄ grasses recruited Ala ATs from the same branch. Preferences for recruitment of specific copies to the C₄ shuttle could be related to suitable existing spatial and temporal expression patterns and/or kinetic properties (Christin *et al.* 2013). Since all tested C₄ grasses possessed more than one cytosolic Ala AT, the subcellular localisation did not seem to have played a decisive role. However, comparison with the Ala AT pattern in related species showed that the Ala AT from the P2 branch also represented the most abundant copy in C₃ grasses, and recruitment of the most abundant C₃ form into the C₄ cycle was observed for various C₄ enzymes (Christin *et al.* 2013).

In the eudicots only one main Ala AT branch existed, and the majority of tested eudicot Ala ATs had predicted mitochondrial localisation. In the phylogenetic tree the dicot Ala ATs also clustered close to the mitochondrial branch of the grasses. The presented phylogenetic tree and prediction tools did not allow clear conclusions about the subcellular ancestry of this gene. More detailed experimental data on the subcellular localisation is only available from *A. thaliana*. The two copies of the genes AtAlaAT1 and AtAlaAT2 are present in roots as well as shoots, but in both organs the AtAlaAT1 copy shows much higher expression (Miyashita *et al.* 2007). Both gene copies possess N-terminal extensions predicting mitochondrial localisation (Table S3, Fig. S3). Both Ala AT proteins could also be detected in the mitochondrial fraction from *A. thaliana* mature leaf samples (Lee *et al.* 2013). Fusion of the AtAlaAT1 coding sequences to an RFP (red fluorescent protein) reporter gene at the C-terminus further supported their mitochondrial localisation (Niessen *et al.* 2012). Earlier studies also reported a cytosolic localisation of the AtAlaAT1. However, in the study of Igarashi *et al.* (2003) the gene was fused to GFP at the N-terminus and the mitochondrial signal was probably masked in their assay. Liepman & Olsen (2003) predicted the localisation of the protein based on a now superseded gene model that codes for a shortened N-terminal, therefore missing the mitochondrial targeting signal. Considering all these details, the studies in *Arabidopsis* actually indicate mitochondrial localisation of both Ala AT copies. The complete absence of Ala AT activity in the cytosol of eudicots such as *A. thaliana* would be surprising, since cytosolic Ala AT activity had been examined in various plant processes. These processes include metabolic adjustment to hypoxia and following re-aeration (Miyashita *et al.* 2007) or adjustment to low N conditions (McAllister & Good 2015). Under hypoxic conditions, the *alaat1* mutants

showed considerably reduced Ala AT activity and delayed Ala catabolism during re-aeration (Miyashita *et al.* 2007). Overexpression of a cytosolic Ala AT from barley in *A. thaliana* induced accelerated Ala catabolism during re-aeration (Miyashita *et al.* 2007), indicating that there might also be a cytosolic function for Ala ATs in eudicots.

Evolution of C₄ photosynthesis is thought to have occurred in a very similar way in monocots and eudicots; and eudicot C₄ species of the NAD-ME and NADP-ME subtype have been identified. Especially for the NAD-ME species, the contribution of the Ala AT would be essential, and for fast rates of pyruvate–Ala conversion cytosolic localisation had been assumed. It could of course be possible that very recent changes in the C₄ eudicot Ala AT genes redirected them to the cytosol. However, in our study also the C₄-specific Ala ATs in *Atriplex rosea*, and *Gynandropsis gynandra* were placed into the mitochondria by the prediction tools. If Ala ATs were really exclusively localised to the mitochondria in the C₄ eudicot species, this would require further involvement of additional transporters between the cytosol and the mitochondria in the mesophyll cells of the NAD-ME species (Fig. 5B). The contribution of an Asp–Ala shuttle in the NADP-ME dicot *Flaveria bidentis* had been predicted through transcript analysis and labelling experiments (Meister *et al.* 1996). Ala AT activity exclusively in the mitochondria would in this case depend on additional transport between the cytosol and the mitochondria in mesophyll and bundle sheath cells.

The existence of Ala ATs in the cytosol of eudicot species could, however, still be possible. Prediction programs and protein fusion to marker proteins often aim at the identification of a single target compartment, but dual targeting mechanisms have been shown for a variety of plant proteins (Carrie & Whelan 2013). These allow the localisation of proteins from one gene to different compartments *via* the use of alternative start codons or alternative promoters, *via* the use of two different targeting signals in the same protein, or *via* ambiguous targeting signals (Parsley & Hibberd 2006; Carrie & Whelan 2013). The C₄ enzyme PPK for instance exists in two versions: the longer transcript is generated from a promoter upstream of the first exon and its product is delivered to the plastid, the shorter transcript is derived from a promoter within the first intron of the longer form and targeted to the cytosol (Rosche & Westhoff 1995; Parsley & Hibberd 2006). Delivery of the same protein to an organelle and the cytosol can also be realized through the existence of two starting sites for the translation of the gene, one with and one without the organellar targeting signal, as shown for the valyl-tRNA synthetase (Souciet *et al.* 1999). A closer inspection of the N-termini of the selected dicot Ala ATs and the monocot *O. sativa* revealed that a second methionine could be found in all sequences containing the mitochondrial extension, and this methionine is always close to the start of the mature protein in the cytosolic forms from the monocots (Fig. S3). Regulation of the alternative translation initiation could be realized by elements in the UTRs (untranslated regions) or the intron, which had not been included in the reporter constructs. A different scenario resulted from analysis of the N-terminus in the C₄ Ala AT from the eudicot *Amaranthus edulis*. In this species the Ala AT still possessed an N-terminal extension but was predicted to localise to the cytosol. It would be interesting to test experimentally whether changes in the N-terminus extension indeed cause cytosolic targeting of the Ala AT protein.

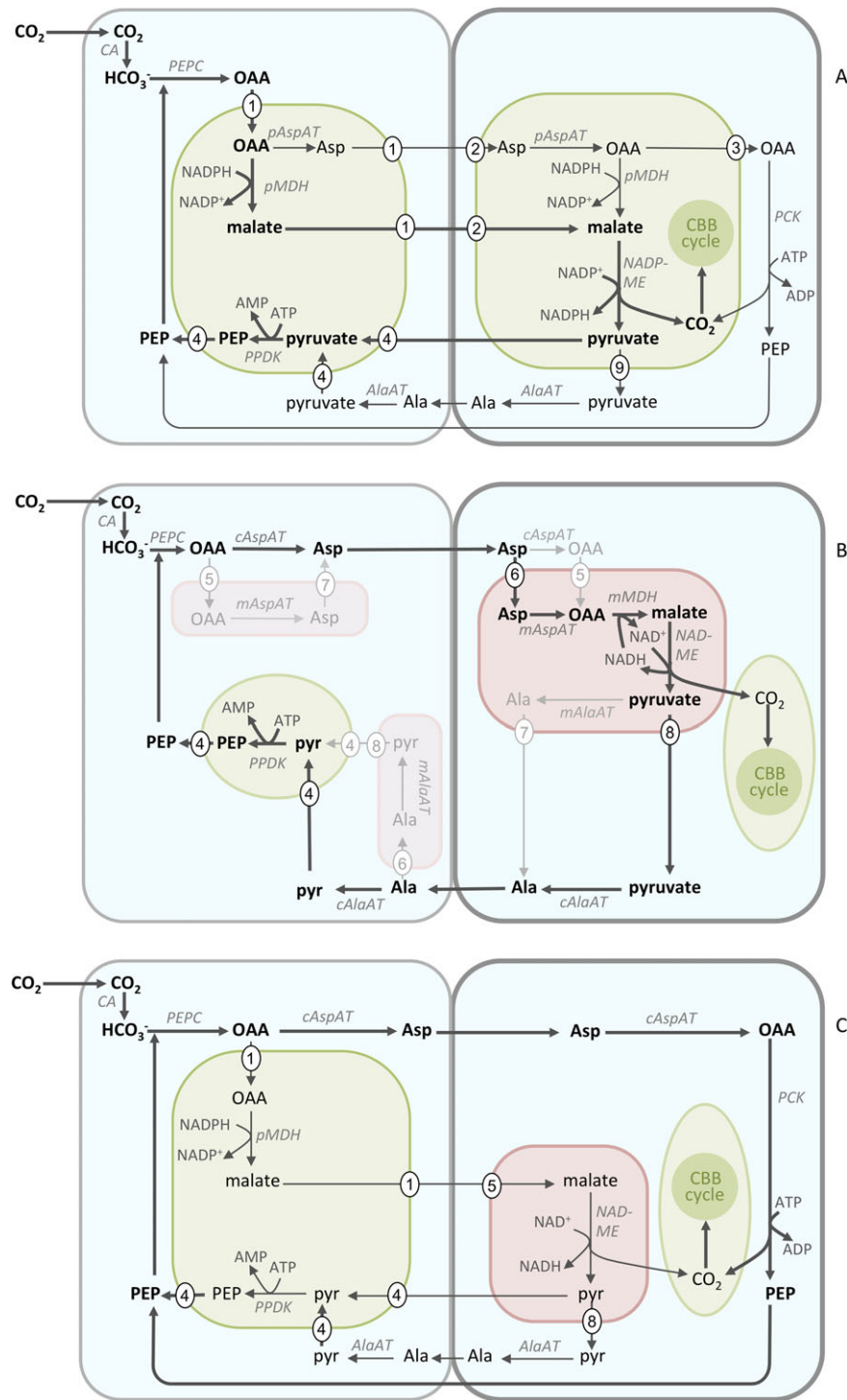


Fig. 5. Overview of C₄ pathways. (A) NADP-malic enzyme photosynthetic pathway including the major malate/pyruvate as well as the minor aspartate (Asp)/alanine (Ala) shuttle; (B) NAD-malic enzyme photosynthetic pathway, the light grey reactions indicate possible pathways in *Gynandropsis gynandra* assuming that all C₄-specific Ala ATs and Asp ATs are indeed localised to the mitochondria of mesophyll and bundles sheath cells; C. PEP carboxykinase photosynthetic pathway. Participating enzymes are carbonic anhydrase (CA); phosphoenolpyruvate carboxylase (PEPC), plastidial NADP-dependent malate dehydrogenase (pMDH), mitochondrial NAD-dependent malate dehydrogenase (mMDH), plastidial Asp aminotransferase (pAsp AT), cytosolic Asp aminotransferase (cAsp AT), mitochondrial Asp aminotransferase (mAsp AT), Ala aminotransferase (Ala AT), phosphoenolpyruvate carboxykinase (PCK), NADP-dependent malic enzyme (NADP-ME), NAD-dependent malic enzyme (NAD-ME), pyruvate Pi dikinase (PPDK) and participating abbreviated metabolites are bicarbonate (HCO₃⁻), oxaloacetate (OAA), aspartate (Asp), alanine (Ala), phosphoenolpyruvate (PEP). CBB = Calvin-Benson-Bassham cycle. Metabolite transport is realised as (1) plastidial exchange malate/Asp versus OAA (DIT1/DIT2); (2) plastidial malate/Asp exchange (DCT2), (3) an unknown plastidial OAA exporter; (4) plastidial pyruvate/PEP exchanger (BASS2/NHD/PPT or an unknown transporter in the *Andropogonae*); (5) mitochondrial dicarboxylate exchanger; (6) unknown mitochondrial amino acid importer; (7) unknown mitochondrial exporter; (8) unknown mitochondrial pyruvate exporter; (9) unknown plastidial pyruvate exporter.

In accordance with metabolic C₄ models, a cytosolic form of the Ala AT was recruited into the C₄ shuttle in monocot species. In eudicots, on the other hand, no exclusive cytosolic form of Ala AT seemed to be encoded by the genome, and Ala ATs were integrated into the C₄ pathways from mitochondrial forms. Further experiments are necessary in order to find out whether these eudicot Ala ATs exclusively reside in the mitochondria, which would also require the activity of suitable transporters for realisation of a fast flux through the C₄ system, or whether cytosolic localisation of the Ala AT proteins is realized in the eudicots by a so far unknown mechanism.

Recruitment of specific Asp ATs into C₄ is strongly influenced by the predominant decarboxylating enzyme

In the PCK and NAD-ME C₄ subtypes, Asp is used as the main shuttle metabolite for delivery of CO₂ to the site of Rubisco in the bundle sheath. Also, in the NADP-ME subtype Asp can function as an additional C₄ transport metabolite. C₄ species of all subtypes are therefore characterised by an increase in Asp AT abundance when compared to related C₃ species. In contrast to the situation for Ala ATs, plants possess specific Asp AT copies in different cell compartments, the cytosol, mitochondria, plastid and peroxisome (Schultz & Coruzzi 1995; Wilkie *et al.* 1995; Miesak & Coruzzi 2002). During C₄ photosynthesis, the Asp needs to be transaminated and decarboxylated in the bundle sheath. In all current C₄ models, the transamination of Asp is catalysed by an Asp AT forming OAA. This step is followed either by direct decarboxylation *via* PCK in the cytosol or *via* conversion of OAA to malate by an MDH and decarboxylation of malate by an NAD-ME in the mitochondria or NADP-ME in the plastid. The transcript data indicated that the C₄-specific Asp AT in the bundle sheath localises to the same compartment as the decarboxylating enzyme.

In the PCK species *Alloteropsis semialata* and *Megathyrus maximus*, considerable increases were only observed for the cytosolic Asp AT transcript. Since the product of the Asp AT reaction, OAA, can be directly decarboxylated by the PCK in the cytosol, further transport processes are avoided. CO₂ is released in the cytosol and can diffuse to the site of Rubisco in the bundle sheath plastids. PEP can then be returned directly to the mesophyll cells. For re-balancing of the N metabolism between the cells, additional cycles including the exchange of malate and alanine are thought to operate in the PCK plants (Furbank 2011).

In the NAD-ME species, transcripts of two Asp ATs, a cytosolic and a mitochondrial form, were enhanced in the grasses *Panicum miliaceum*, and *P. virgatum* as well as in the eudicots *A. edulis* and *A. rosea*. Separation of the mesophyll and bundle sheath fraction in *P. virgatum* (Fig. 4, Rao *et al.* 2016) showed that the mitochondrial Asp AT was specific for the bundle sheath and the cytosolic one for the mesophyll cells. The results supported the early studies in NAD-ME species *Panicum miliaceum* and *A. spongiosa* using cell and organelle specific fractionation and activity assays (Hatch & Mau 1973; Moore *et al.* 1984; Taniguchi *et al.* 1995). In these experiments, different isoenzymes of Asp AT were enhanced in the mesophyll and bundle sheath fraction. In the mesophyll cells, the Asp ATs could be associated with the cytosol, and in the bundle sheath with the mitochondria (Hatch & Mau 1973; Moore *et al.* 1984). The transcript pattern of *P. virgatum* showed that the

mitochondrial form of NAD-MDH was also strongly enhanced in the bundle sheath fraction (Fig. S4). In Asp AT enzyme assays from *P. miliaceum* leaves, the Asp consumption was only 50–60% compared to Asp production (Taniguchi *et al.* 1995), but immediate removal of the Asp AT reaction product OAA by the MDH in the same cellular compartment would promote the continued transamination of the Asp. The NADH consumed for the reduction of OAA to malate could be supplied by the following NAD-ME decarboxylation reaction (Fig. 5). The localisation of the bundle sheath Asp AT in mitochondria in NAD-ME species implied that Asp rather than OAA or malate is imported into the mitochondria. In the mesophyll cells of NAD-ME plants, on the other hand, a cytosolic placement of the Asp AT reaction would avoid further transport processes within the shuttle. In contrast to the Asp AT pattern in the majority of NAD-ME species, only the Asp AT transcript from the mitochondrial branch increased in abundance in mesophyll and bundle sheath cells of *G. gynandra*. The prediction programs actually indicated plastidic localisation for this specific protein, but GFP-fusion experiments with the C₄-specific Asp AT form proved its targeting to the mitochondria (Sommer *et al.* 2012). The possible pathway in *G. gynandra* including a mitochondrial placement of the Asp ATs as well the Ala AT, as predicted from the phylogenetic tree, is shown in light grey in Fig. 5B. The exception to the common model, however, underlines the general flexibility of the pathway in the individual lines.

In the NADP-ME species, malate represents the main C₄ shuttle metabolite, but to varying degrees also Asp contributes to the C₄ cycle (Chapman & Hatch 1981; Pick *et al.* 2011; Arrivault *et al.* 2017). Such a contribution was calculated to amount to about one-third in *Flaveria bidentis* (Meister *et al.* 1996) or between 10% (Arrivault *et al.* 2017) and 25% (Chapman & Hatch 1981) in maize. The total leaf transcript pattern also showed up-regulation of the plastidic Asp AT in monocot (*Z. mays*, *S. italica*, and to a lesser degree also *S. bicolor*) as well as eudicot (*F. bidentis*) NADP-ME species. Abundant Asp AT isoforms associated with the plastidial fraction could also be identified in maize leaves using proteome analysis (Majeran *et al.* 2005). In the mesophyll cells of NADP-ME species, OAA is reduced to malate by NADP-MDH within the plastid, where the NADPH is provided by the photosynthetic electron transport chain (Fig. 5A). Placement of the Asp AT into the same organelle would provide an alternative reaction using OAA without the consumption of NADPH. Both reactions could therefore contribute to the adjustment of reductant availability within the mesophyll plastid.

After Asp transport to the bundle sheath cells, it could be transaminated and decarboxylated either directly by PCK in the cytosol or by conversion to malate *via* plastid-localised MDH by the NADP-ME in the plastid. Both pathways would require the activity of an Asp AT. The transcript abundance of Asp ATs as well as MDH was, however, much lower in the bundle sheath than in the mesophyll cells of *Z. mays* (Fig. S4). The same was true for the total Asp AT activity in the bundle sheath fraction (Fig. 4; Denton *et al.* 2017) and Asp AT protein abundance in the maize leaf (Majeran *et al.* 2005). The fate of Asp in the bundle sheath is therefore not immediately clear from the currently available data (Rao & Dixon 2016). Asp AT activity could contribute to general amino acid production and protein biosynthesis, which is preferentially located in the mesophyll

cells (Majeran *et al.* 2005; Brautigam *et al.* 2011), and not all Asp produced in the mesophyll cells therefore participates in the C₄ shuttle. Experiments by Chapman & Hatch (1981) suggested that Asp might not be immediately metabolised in the bundle sheath but plays a beneficial role in the transport of malate into the plastid. It is also possible that Asp in the bundle sheath feeds into multiple reactions. Labelling of Ala in maize leaves also occurred more slowly than that of pyruvate (Arrivault *et al.* 2017). Transcript studies indicated that although lower than in the mesophyll cells, the plastidial form of Asp AT was still most abundant in the bundle sheath of NADP-ME plants (Fig. 4F, G, H), so that Asp involved in the C₄ shuttle was preferentially transaminated in the plastids followed by conversion to malate and decarboxylation by NADP-ME. In contrast to these results from grasses, the NADP-ME eudicot species *F. bidensis* showed considerable Asp and Ala AT activity also in the bundle sheath fraction of the leaf (Meister *et al.* 1996), thus providing complete support for the mixed NADP-ME – PCK decarboxylation model (Fig. 5A).

Hence, in PCK and NAD-ME species, the localisation of the decarboxylating enzyme determines the targeting of the contributing Asp AT in the bundle sheath cells. In the mesophyll cells the cytosolic form is preferred, but divergent solutions such as in *G. gynandra* are possible. In the NADP-ME species, the Asp shuttle generally contributes to the robustness of the system under changing environmental conditions, supplying flexibility in the N and energy balancing of the cells (Khamis *et al.* 1992; Arrivault *et al.* 2017).

Plants display plasticity in the setup of C₄ pathways in different lineages

All C₄ species rely on the same enzymes for carboxylation (PEPC) and PEP regeneration (PPDK in coordination with AMK, PPase) in the mesophyll cells, but at least three different enzymes (NADP-ME, NAD-ME, PCK) can act as the main decarboxylating enzyme in the bundle sheath cells. In many C₄ species, the main decarboxylating enzyme is also supplemented by the contribution of a second decarboxylating enzyme (Furbank 2011; Wang *et al.* 2014b; Brautigam & Gowik 2016). On an evolutionary scale, the decarboxylation unit likely developed before the strong enhancement of the PEPC carboxylation in the mesophyll cells (Mallmann *et al.* 2014). The reasons for the preference for a specific C₄ type are not yet resolved. It could be related to the genetic components of the plants resulting in lineage-specific characteristics, or could represent specific ecophysiological adaptations resulting in subtype-specific characteristics (Pinto *et al.* 2016). The carbon concentrating efficiency was similar for grasses from all subtypes (Sonawane *et al.* 2017). Lineage specificity was observed for water use-related features, while high N use efficiency could be associated with the C₄ subtype (Pinto *et al.* 2016). Fine-tuning of the C₄ cycle to varying conditions may be easier if relevant enzymes occur in proximity in the same compartment. Alternatively, stability of cycle intermediates, such as OAA, limits enzyme placement.

The nature of the decarboxylating enzyme then requires the up-regulation of specific enzymes, especially AT and MDH at specific cellular and subcellular localisations, plus the relevant transporters for the establishment of an efficient C₄ shuttle. In the tested monocot and eudicot C₄ species a subtype-dependent pattern of ATs was enhanced. The absence of a cytosol-

specific Ala AT branch in the eudicots did not prevent the evolution of multiple C₄ lineages. Individual lineages, however, also show a certain plasticity of the C₄ setup. In contrast to the majority of NAD-ME species using mitochondrial Asp AT in the bundle sheath and cytosolic Asp AT in the mesophyll, *G. gynandra* relies on operation of Asp AT exclusively from the mitochondrial branch in both cell types. Divergent solutions are also found for the plastidial pyruvate transporter. The majority of C₄ plants use a sodium-dependent transport system (BASS2/NHD/PPT), but this is not the case in the *Andropogonae* (Furumoto *et al.* 2011); C₄ species from this family seem to rely on a direct proton-pyruvate exchange system that has not been identified at the molecular level to date (Schlüter *et al.* 2016). Hence, beside the small core set of enzymes and transporters, more than one solution seems to be possible for several steps of the C₄ cycle in the individual lineages. Plasticity was also observed in the architecture of C₄ species. The majority and most efficient C₄ species possess the classical Kranz anatomy, but bundle sheath arrangements around a central water body or the setup of a C₄ shuttle within a single cell have also been successful (Chuong *et al.* 2006; Bohley *et al.* 2015).

The plasticity of the C₄ cycle, with diverging solutions for individual steps, might also explain why multiple C₃ versus C₄ transcript comparisons yielded only a limited number of genes and did not allow the identification of the still missing transporters (Brautigam *et al.* 2014). For engineering approaches, such as the introduction of a C₄ shuttle into C₃ rice plants, knowledge of the transfer acid-generating enzymes is required for rational choices and – in the absence thereof – choice of enzyme close to a naturally occurring C₄ constellation is suggested.

ACKNOWLEDGEMENTS

We would like to thank all members of the 3 to 4 EU project for help with data analysis and extensive discussions. For the transcriptome analysis, RNA was kindly provided by the groups of Richard Leegood and Pascal-Antoine Christin in Sheffield for *Amaranthus edulis* and the *Alloteropsis* species, and by Rowan and Tammy Sage in Toronto for the *Atriplex* species. Dominik Brillhaus kindly provided transcript and sequence data for the *Flaveria* species. We would like to acknowledge funding of the work by the EU (289582) and CEPLAS (EXC 1028). JS acknowledges support from the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under contract number DE-SC0012704 – specifically through the Physical Biosciences program of the Chemical Sciences, Geosciences and Biosciences Division.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Data S1. Ala and Asp AT protein sequences for phylogenetic tree.

Table S1. Transcript amounts of Ala and Asp aminotransferases in mature leaves of different C₃ and C₄ species.

Table S2. Transcript amounts of Ala and Asp aminotransferases in mesophyll (M) and bundle sheath (BS) from mature leaf samples.

Table S3. Prediction of subcellular localisation for Ala aminotransferases.

Table S4. Prediction of subcellular localisation for Asp aminotransferases.

Figure S1. Asp AT phylogenetic tree with bootstrap support values (100 iterations).

Figure S2. Ala AT phylogenetic tree with bootstrap support values (100 iterations).

REFERENCES

- Arrivault S., Obata T., Szcwowska M., Mengin V., Guenther M., Hoehne M., Fernie A.R., Stitt M. (2017) Metabolite pools and carbon flow during C₄ photosynthesis in maize: ¹³CO₂ labeling kinetics and cell type fractionation. *Journal of Experimental Botany*, **68**, 283–298.
- Aubry S., Brown N.J., Hibberd J.M. (2011) The role of proteins in C₃ plants prior to their recruitment into the C₄ pathway. *Journal of Experimental Botany*, **62**, 3049–3059.
- Aubry S., Smith-Unna R.D., Bournsnell C.M., Kopriva S., Hibberd J.M. (2014) Transcript residency on ribosomes reveals a key role for the *Arabidopsis thaliana* bundle sheath in sulfur and glucosinolate metabolism. *The Plant Journal*, **78**, 659–673.
- Bohley K., Joos O., Hartmann H., Sage R., Liede-Schumann S., Kadereit G. (2015) Phylogeny of Sesuvioideae (Aizoaceae) – Biogeography, leaf anatomy and the evolution of C₄ photosynthesis. *Perspectives in Plant Ecology, Evolution and Systematics*, **17**, 116–130.
- Brautigam A., Gowik U. (2016) Photorespiration connects C₃ and C₄ photosynthesis. *Journal of Experimental Botany*, **67**, 2953–2962.
- 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., Weber A.P.M. (2011) An mRNA blueprint for C₄ photosynthesis derived from comparative transcriptomics of closely related C₃ and C₄ species. *Plant Physiology*, **155**, 142–156.
- Brautigam A., Schliesky S., Kulahoglu C., Osborne C.P., Weber A.P.M. (2014) Towards an integrative model of C₄ photosynthetic subtypes: insights from comparative transcriptome analysis of NAD-ME, NADP-ME, and PEP-CK C₄ species. *Journal of Experimental Botany*, **65**, 3579–3593.
- Carrie C., Whelan J. (2013) Widespread dual targeting of proteins in land plants. When, where, how and why. *Plant Signaling & Behaviour*, **8**, e25034.
- Chang Y.M., Liu W.Y., Shih A.C., Shen M.N., Lu C.H., Lu M.Y., Yang H.W., Wang T.Y., Chen S.C.C., Chen S.M., Li W.H., Ku M.S.B. (2012) Characterizing regulatory and functional differentiation between maize mesophyll and bundle sheath cells by transcriptomic analysis. *Plant Physiology*, **160**, 165–177.
- Chapman K.S.R., Hatch M.D. (1981) Aspartate decarboxylation in bundle sheath cells of *Zea mays* and its possible contribution to C₄ photosynthesis. *Australian Journal of Plant Biology*, **8**, 237–248.
- Christin P.A., Boxall S.F., Gregory R., Edwards E.J., Hartwell J., Osborne C.P. (2013) Parallel recruitment of multiple genes into C₄ photosynthesis. *Genome Biology and Evolution*, **5**, 2174–2187.
- Chuang S.D., Franceschi V.R., Edwards G.E. (2006) The cytoskeleton maintains organelle partitioning required for single-cell C₄ photosynthesis in Chenopodiaceae species. *The Plant Cell*, **18**, 2207–2223.
- Covshoff S., Szcwowska M., Hughes T.E., Smith-Unna R., Kelly S., Bailey K.J., Sage T.L., Pachebat J.A., Leegood R., Hibberd J.M. (2016) C₄ photosynthesis in the rice paddy: insights from the noxious weed *Echinochloa glabrescens*. *Plant Physiology*, **170**, 57–73.
- Denton A.K., Mass J., Kulahoglu C., Lercher M.J., Brautigam A., Weber A.P.M. (2017) Freeze-quenched maize mesophyll and bundle sheath separation uncovers bias in previous tissue-specific RNA-seq data. *Journal of Experimental Botany*, **68**, 147–160.
- Doring F., Streubel M., Brautigam A., Gowik U. (2016) Most photorespiratory genes are preferentially expressed in the bundle sheath cells of the C₄ grass *Sorghum bicolor*. *Journal of Experimental Botany*, **67**, 3053–3064.
- Eisenhut M., Brautigam A., Timm S., Florian A., Tohge T., Fernie A.R., Bauwe H., Weber A.P.M. (2017) Photorespiration is crucial for dynamic response of photosynthetic metabolism and stomatal movement to altered CO₂ availability. *Molecular Plant*, **10**, 47–61.
- Furbank R.T. (2011) Evolution of the C₄ photosynthetic mechanism: are there really three C₄ acid decarboxylation types? *Journal of Experimental Botany*, **62**, 3103–3108.
- Furumoto T., Yamaguchi T., Ohshima-Ichikawa Y., Nakamura M., Tsuchida-Iwata Y., Shimamura M., Ohnishi J., Hata S., Gowik U., Westhoff P., Brautigam A., Weber A.P.M. (2011) A plastidial sodium-dependent pyruvate transporter. *Nature*, **476**, 472–475.
- Gouy M., Guindon S., Gascuel O. (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, **27**, 221–224.
- Gowik U., Brautigam A., Weber K.L., Weber A.P., Westhoff P. (2011) Evolution of C₄ photosynthesis in the genus *Flaveria*: how many and which genes does it take to make C₄? *The Plant Cell*, **23**, 2087–2105.
- Haas B.J., Papanicolaou A., Yassour M., Grabherr M., Blood P.D., Bowden J., Couger M.B., Eccles D., Li B., Lieber M., MacManes M.D., Ott M., Orvis J., Pochet N., Strozzi F., Weeks N., Westerman R., William T., Dewey C.N., Henschel R., LeDuc R.D., Friedman N., Regev A. (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, **8**, 1494–1512.
- Hagemann M., Bauwe H. (2016) Photorespiration and the potential to improve photosynthesis. *Current Opinion in Chemical Biology*, **35**, 109–116.
- Hatch M.D., Mau S.-L. (1973) Activity, location, and role of aspartate aminotransferase and alanine aminotransferase isoenzymes in leaves with C₄ pathway photosynthesis. *Archives of Biochemistry and Biophysics*, **156**, 195–206.
- Heckmann D., Schulze S., Denton A., Gowik U., Westhoff P., Weber A.P., Lercher M.J. (2013) Predicting C₄ photosynthesis evolution: modular, individually adaptive steps on a Mount Fuji fitness landscape. *Cell*, **153**, 1579–1588.
- Hibberd J.M., Covshoff S. (2010) The regulation of gene expression required for C₄ photosynthesis. *Annual Review of Plant Biology*, **61**, 181–207.
- Igarashi D., Miwa T., Seki M., Kobayashi M., Kato T., Tabata S., Shinozaki K., Ohsumi C. (2003) Identification of photorespiratory glutamate:glyoxylate aminotransferase (GGAT) gene in *Arabidopsis*. *The Plant Journal*, **33**, 975–987.
- John C.R., Smith-Unna R.D., Woodfield H., Covshoff S., Hibberd J.M. (2014) Evolutionary convergence of cell-specific gene expression in independent lineages of C₄ grasses. *Plant Physiology*, **165**, 62–75.
- Khamis S., Lamaze T., Farineau J. (1992) Effect of nitrate limitation on the photosynthetically active pool of aspartate and malate in maize, a NADP malic enzyme C₄ plant. *Physiologia Plantarum*, **85**, 223–229.
- Ku M.S.B., Kano-Murakami Y., Matsuoka M. (1996) Evolution and expression of C₄ photosynthesis genes. *Plant Physiology*, **111**, 949–957.
- Kulahoglu C., Denton A.K., Sommer M., Mass J., Schliesky S., Wrobel T.J., Berckmans B., Gongora-Castillo E., Buell R., Simon R., De Veylder L., Brautigam A., Weber A.P.M. (2014) Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C₃ and C₄ plant species. *The Plant Cell*, **26**, 3243–3260.
- Lee C.P., Taylor N.L., Millar A.H. (2013) Recent advances in the composition and heterogeneity of the *Arabidopsis* mitochondrial proteome. *Frontiers in Plant Science*, **4**, 4.
- Liepmann A.H., Olsen L.J. (2003) Alanine aminotransferase homologs catalyze the glutamate:glyoxylate aminotransferase reaction in peroxisomes of *Arabidopsis*. *Plant Physiology*, **131**, 215–227.
- Ludwig M. (2016) The roles of organic acids in C₄ photosynthesis. *Frontiers in Plant Science*, **7**, 647.
- Maier A., Zell M.B., Maurino V.G. (2011) Malate decarboxylases: evolution and roles of NAD(P)-ME isoforms in species performing C₄ and C₃ photosynthesis. *Journal of Experimental Botany*, **62**, 3061–3069.
- Majeran W., Cai Y., Sun Q., van Wijk K.J. (2005) Functional differentiation of bundle sheath and mesophyll maize chloroplasts determined by comparative proteomics. *The Plant Cell*, **17**, 3111–3140.
- Mallmann J., Heckmann D., Brautigam A., Lercher M.J., Weber A.P.M., Westhoff P., Gowik U. (2014) The role of photorespiration during the evolution of C₄ photosynthesis in the genus *Flaveria*. *ELife*, **3**, e02478.
- McAllister C.H., Good A.G. (2015) Alanine aminotransferase variants conferring diverse NUE

- phenotypes in *Arabidopsis thaliana*. *PLoS ONE*, **10**, e0121830.
- Meister M., Agostino A., Hatch M.D. (1996) The roles of malate and aspartate in C-4 photosynthetic metabolism of *Flaveria*. *Planta*, **199**, 262–269.
- Miesak B.H., Coruzzi G.M. (2002) Molecular and physiological analysis of *Arabidopsis* mutants defective in cytosolic or chloroplastic aspartate aminotransferase. *Plant Physiology*, **129**, 650–660.
- Miyashita Y., Dolferus R., Ismond K.P., Good A.G. (2007) Alanine aminotransferase catalyses the breakdown of alanine after hypoxia in *Arabidopsis thaliana*. *The Plant Journal*, **49**, 1108–1121.
- Moore B., Ku M.S.B., Edwards G.E. (1984) Isolation of leaf bundle sheath protoplasts from C₄ dicot species and intercellular localization of selected enzymes. *Plant Science Letters*, **35**, 127–138.
- Nakamura Y., Tolbert N.E. (1983) Serine:glyoxylate, alanine:glyoxylate, and glutamate:glyoxylate aminotransferase reactions in peroxisomes from spinach leaves. *Journal of Biological Chemistry*, **258**, 7631–7638.
- Niessen M., Krause K., Horst I., Staebler N., Klaus S., Gaertner S., Kebeish R., Araujo W.L., Fernie A.R., Peterhansel C. (2012) Two alanine aminotransferases link mitochondrial glycolate oxidation to the major photorespiratory pathway in *Arabidopsis* and rice. *Journal of Experimental Botany*, **63**, 2705–2716.
- Parsley K., Hibberd J.M. (2006) The *Arabidopsis* PPKK gene is transcribed from two promoters to produce differentially expressed transcripts responsible for cytosolic and plastidic proteins. *Plant Molecular Biology*, **62**, 339–349.
- Pick T.R., Bräutigam A., Schlüter U., Denton A.K., Colmsee C., Scholz U., Fahnstich H., Pieruschka R., Rascher U., Sonnewald U., Weber A.P.M. (2011) Systems analysis of a maize leaf developmental gradient redefines the current C₄ model and provides candidates for regulation. *The Plant Cell*, **23**, 4208–4220.
- Pinto H., Powell J.R., Sharwood R.E., Tissue D.T., Ghannoum O. (2016) Variations in nitrogen use efficiency reflect the biochemical subtype while variations in water use efficiency reflect the evolutionary lineage of C₄ grasses at inter-glacial CO₂. *Plant, Cell and Environment*, **39**, 514–526.
- Rao X., Dixon R.A. (2016) The differences between NAD-ME and NADP-ME subtypes of C₄ photosynthesis: more than decarboxylating enzymes. *Frontiers in Plant Science*, **7**, 1525.
- Rao X., Lu N., Li G., Nakashima J., Tang Y., Dixon R.A. (2016) Comparative cell-specific transcriptomics reveals differentiation of C₄ photosynthesis pathways in switchgrass and other C₄ lineages. *Journal of Experimental Botany*, **67**, 1649–1662.
- Rawsthorne S., Hylton C.M., Smith A.M., Woolhouse H.W. (1988) Photorespiratory metabolism and immunogold localisation of photorespiratory enzymes in leaves of C₃ and C₃-C₄ intermediate species of *Moricaandia*. *Planta*, **173**, 298–308.
- Reyna-Llorens I., Hibberd J.M. (2017) Recruitment of pre-existing networks during the evolution of C₄ photosynthesis. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, **372**, 20160386.
- Rocha M., Licausi F., Araujo W.L., Nunes-Nesi A., Sodek L., Fernie A.R., van Dongen J.T. (2010) Glycolysis and the tricarboxylic acid cycle are linked by alanine aminotransferase during hypoxia induced by waterlogging of *Lotus japonicus*. *Plant Physiology*, **152**, 1501–1513.
- Rosche E., Westhoff P. (1995) Genomic structure and expression of the pyruvate, orthophosphate dikinase gene of the dicotyledonous C₄ plant *Flaveria trinervia*. *Plant Molecular Biology*, **29**, 663–678.
- Sage R.F. (2004) The evolution of C₄ photosynthesis. *New Phytologist*, **161**, 341–370.
- Sage R.F. (2016) A portrait of the C₄ photosynthetic family on the 50th anniversary of its discovery: species number, evolutionary lineages, and Hall of Fame. *Journal of Experimental Botany*, **67**, 4039–4056.
- Sage R.F., Stata M. (2015) Photosynthetic diversity meets biodiversity: the C₄ plant example. *Journal of Plant Physiology*, **172**, 104–119.
- Sage R.F., Christin P.A., Edwards E.J. (2011) The C₄ plant lineages of planet Earth. *Journal of Experimental Botany*, **62**, 3155–3169.
- Sato K., Yamane M., Yamaji N., Kanamori H., Tagiri A., Schwerdt J.G., Fincher G.B., Matsumoto T., Takeda K., Komatsuda T. (2016) Alanine aminotransferase controls seed dormancy in barley. *Nature Communications*, **7**, 11625.
- Schlüter U., Denton A.K., Bräutigam A. (2016) Understanding metabolite transport and metabolism in C₄ plants through RNA-seq. *Current Opinion in Plant Biology*, **31**, 83–90.
- Schultz C.J., Coruzzi G.M. (1995) The aspartate aminotransferase gene family of *Arabidopsis* encodes isoenzymes localized to three distinct subcellular compartments. *The Plant Journal*, **7**, 61–75.
- Sharkey T.D. (1988) Estimating the rate of photorespiration in leaves. *Physiologia Plantarum*, **73**, 147–152.
- Shih P.M. (2015) Photosynthesis and early Earth. *Current Biology*, **25**, R855.
- Shrawat A.K., Carroll R.T., DePauw M., Taylor G.J., Good A.G. (2008) Genetic engineering of improved nitrogen use efficiency in rice by the tissue-specific expression of alanine aminotransferase. *Plant Biotechnology Journal*, **6**, 722–732.
- Sommer M., Bräutigam A., Weber A.P. (2012) The dicotyledonous NAD malic enzyme C₄ plant *Cleome gynandra* displays age-dependent plasticity of C₄ decarboxylation biochemistry. *Plant Biology*, **14**, 621–629.
- Sonawane B.V., Sharwood R.E., von Caemmerer S., Whitney S.M., Ghannoum O. (2017) Short-term thermal photosynthetic responses of C₄ grasses are independent of the biochemical subtype. *Journal of Experimental Botany*, **68**, 5583–5597.
- Souciet G., Menand B., Ovesna J., Cosset A., Dietrich A., Wintz H. (1999) Characterization of two bifunctional *Arabidopsis thaliana* genes coding for mitochondrial and cytosolic forms of valyl-tRNA synthetase and threonyl-tRNA synthetase by alternative use of two in frame AUGs. *European Journal of Biochemistry*, **266**, 848–854.
- Taniguchi M., Kobe A., Kato M., Sugiyama T. (1995) Aspartate Aminotransferase Isozymes in *Panicum milaceum* L., an NAD-Malic Enzyme-type C₄ Plant: comparison of enzymatic properties, primary structures, and expression patterns. *Archives of Biochemistry and Biophysics*, **318**, 295–306.
- de la Torre F., El-Azaz J., Avila C., Canovas F.M. (2014) Deciphering the role of aspartate and prephenate aminotransferase activities in plastid nitrogen metabolism. *Plant Physiology*, **164**, 92–104.
- Wang L., Czédik-Eysenberg A., Mertz R.A., Si Y., Tohge T., Nunes-Nesi A., Arrivault S., Dedow L.K., Bryant D.W., Zhou W., Xu J., Weissmann S., Studer A., Li P., Zhang C., LaRue T., Shao Y., Ding Z., Sun Q., Patel R.V., Turgeon R., Zhu X., Provart N.J., Mockler T.C., Fernie A.R., Stitt M., Liu P., Brutnell T.P. (2014a) Comparative analyses of C₄ and C₃ photosynthesis in developing leaves of maize and rice. *Nature Biotechnology*, **32**, 1158–1165.
- Wang Y., Bräutigam A., Weber A.P., Zhu X.G. (2014b) Three distinct biochemical subtypes of C₄ photosynthesis? A modelling analysis. *Journal of Experimental Botany*, **65**, 3567–3578.
- Westhoff P., Gowik U. (2004) Evolution of C₄ phosphoenolpyruvate carboxylase. Genes and proteins: a case study with the genus *Flaveria*. *Annals of Botany*, **93**, 13–23.
- Wilkie S.E., Roper J.M., Smith A.G., Warren M. (1995) Isolation, characterisation and expression of a cDNA clone encoding plastid aspartate aminotransferase from *Arabidopsis thaliana*. *Plant Molecular Biology*, **27**, 1227–1233.
- Wingler A., Walker R.P., Chen Z.-H., Leegood R. (1999) Phosphoenolpyruvate carboxylase is involved in the decarboxylation of aspartate in the bundle sheath of maize. *Plant Physiology*, **120**, 539–545.