HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

iGRADplant
International
Graduate
Program
for Plant Science

Transcriptional adaptation in response to changing environments in plants with specialized photosynthesis types

Inaugural-Dissertation

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

vorgelegt von

Dominik Brilhaus
aus Solingen

Düsseldorf, May 2016

aus dem
Institut für Biochemie der Pflanzen
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. Andreas P. M. Weber
Korreferent: Prof. Dr. Michael Feldbrügge
Tag der mündlichen Prüfung:

Eidesstattliche Versicherung und Selbstständigkeitserklärung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde 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 17.05.2015

_____

Dominik Brilhaus

Table of Contents

Abbreviations

$[CH2O]_n$	carbohydrates
ATP	Adenosine triphosphate
BSC	bundle sheath cell
CAM	Crassulacean acid metabolism
CBC	Calvin-Benson-Cycle
CCM	CO2-concentrating mechanisms
CO_2	carbon dioxide
H_2O	Water
MC	mesophyll cell
NAD(P)-ME	NAD(P)-malic enzyme
NADPH	Nicotinamide adenine dinucleotide phosphate
NUE	nitrogen use efficiency
O_2	oxygen
OAA	oxaloacetate
PEP	phosphoenolpyruvate
PEPC	PEP carboxylase
PEPCK	PEP carboxykinase
PGA	phosphoglycerate
RUBISCO	RuBP carboxylase/oxygenase
RuBP	ribulose-1,5-bisphosphate
WUE	water use efficiency

Summary

Globally rising temperatures and an exploding world population present enormous challenges for agriculture in the 21st century and the requirement for the development of crop plants to cope with these changing environments. As an adaptation to drought, high temperatures and low CO_2 availability, plants have evolved two specialized types of photosynthesis, termed Crassulacean acid metabolism (CAM) and C_4 photosynthesis. These complex traits provide increased efficiencies in the usage of water, nitrogen and light and are hence of major relevance for engineering efficient and drought resistant crops. Key to both CAM and C_4 is the separation of CO_2 prefixation and assimilation managed in two distinct modes, but utilizing a very similar biochemical adaptation. While the physiology, anatomy and biochemistry of both these traits is well studied, there are still gaps in the knowledge of their transcriptional regulation.

In response to drought stress, the tropical herbaceous perennial species *Talinum triangulare* is capable of transitioning, in a facultative, reversible manner, from C_3 photosynthesis to weakly expressed CAM. The transcriptional regulation of this transition was studied combining mRNA-Seq with targeted metabolite measurements. We found highly elevated levels of CAM-cycle enzyme transcripts and their metabolic products in *T. triangulare* leaves upon water deprivation. This large-scale expression dataset of drought-induced CAM demonstrates transcriptional regulation of the C_3–CAM transition. In addition, we identified candidate transcription factors to mediate this photosynthetic plasticity, which may contribute in the future to the design of more drought-tolerant crops via engineered CAM (Manuscript I).

An organism's capability to sync the metabolism to changing source availabilities and sink requirements increases productivity and survival. A large proportion of a plant's metabolism is light/dark-regulated. In addition, the internal circadian clock provides anticipation of the daily and seasonally changing environment. In order to study the specific adaptations of transcriptional control through light and the circadian clock in plants performing highly efficient C_4 photosynthesis, we compared the changes in leaf transcriptomes of two closely related *Flaveria* species employing C_4 (*F. bidentis*) or C_3 (*F. robusta*) photosynthesis. We found largely overlapping diurnal and circadian expression patterns and periodicity between both species and an indication that acquisition of light-regulation of the C_4-cycle predates the evolution of C_4 in the *Flaveriaceae.* Furthermore, C_4-specifically enhanced light-regulation of many photosynthesis related genes was detected. We identified candidate genes to regulate this enhanced light-response in the C_4 plant and discuss possible advantages (Manuscript II).

Zusammenfassung

Globale Klimaerwärmung und die stark wachsende Weltbevölkerung präsentieren enorme Herausforderungen für die Agrarwirtschaft des 21. Jahrhunderts und den Bedarf für die Entwicklung von Nutzpflanzen, die imstande sind in diesen sich ändernden Bedingungen zu gedeihen. Als Anpassung an Dürre, hohe Temperaturen und niedrige CO_2 Verfügbarkeit haben Pflanzen zwei spezialisierte Arten der Photosynthese entwickelt, genannt Crassulacean acid metabolism (CAM, deutsch: Crassulaceensäurestoffwechsel) und C_4 Photosynthese. Pflanzen, die diese komplexen Eigenschaften besitzen, weisen erhöhte Nutzeffizienz von Wasser, Stickstoff und Licht auf und sind daher von besonderer Relevanz für das Erzeugen effizienter, trocken-resistenter Nutzpflanzen. Der Schlüssel zu beiden speziellen Photosynthesearten ist die Isolierung von CO_2- Präfixierung und Assimilierung, die auf zweierlei Arten jedoch durch sich ähnelnde biochemische Anpassungen geschieht. Während die Physiologie, Anatomie und Biochemie beider Photosynthesearten umfassend untersucht sind, gibt es noch Lücken in dem Verständnis ihrer transkriptionellen Regulierung.

Als Antwort auf Trockenstress ist die krautige, einjährige Spezies *Talinum triangulare* in der Lage, auf fakultative und reversible Art zwischen C_3 Photosynthese und schwachem CAM zu wechseln. Die transkriptionelle Regulierung wurde hier mittels Analyse des Transkriptoms und Messungen von Metaboliten untersucht. Nach Trocken-Induktion stellten wir erhöhte Levels der Transkripte von CAM-Enzymen und derer Produkte in Blättern von *T. triangulare* fest. In diesem großangelegten Expressionsdatensatz von trocken-induziertem CAM, konnten wir also die transkriptionelle Regulierung des Wechsels zwischen C_3 und CAM nachweisen. Des Weiteren haben wir Kandidaten für die Regulierung dieser photosynthetischen Plastizität identifiziert, die in Zukunft helfen könnten, trocken-tolerantere Pflanzen mittels Etablierung von CAM zu entwickeln (Manuskript I).

Die Möglichkeit den Stoffwechsel eines Organismus an wechselnde Verfügbarkeit von Ressourcen und dem Bedarf nach deren Nutzung anzupassen erhöht die Produktivität und Überlebensfähigkeit. Ein Großteil des pflanzlichen Stoffwechsels ist durch Licht und Dunkelheit reguliert. Des Weiteren können täglich und saisonal wechselnde Umweltbedingungen mittels innerer biologischer Uhr antizipiert werden. Um die spezifische Anpassung der transkriptionellen Regulierung mittels Licht oder Biorhythmus in Pflanzen mit der hoch-effizienten C_4 Photosynthese zu verstehen, haben wir die Blatttranskriptome zwei nahe verwandter *Flaveria* Spezies untersucht, die C_4 (*F. bidentis*) und C_3 Photosynthese (*F. robusta*)

benutzen. Wir fanden sich zwischen den beiden Spezies überwiegend gleichende diurnale und circadiane Expressionsmuster und –phasen und Anzeichen für eine vor der C_4 Evolution bestehende Licht-Regulierung des C_4-Zykluses in *Flaveriaceae*. Des Weiteren wurde eine C_4-spezifisch verstärkte Licht-Regulation von anderen Photosynthese Genen entdeckt. Wir diskutieren möglichen Vorzüge für diese in C_4 Pflanzen verstärkte Licht-Antwort und präsentieren mögliche Kandidaten für dessen Regulierung (Manuskript II).

Introduction

Improving photosynthetic efficiency

The world population is estimated to grow to 9 billion people by roughly the middle of the current century (Godfray et al., 2010), rapidly increasing the demands for environmentally and socially sustainable production of food, feed and energy resources, on simultaneously shrinking farmlands (Tilman et al., 2011). Plants and other photosynthetic organisms are the basis of the food chain and fossil and non-fossil energy resources, thanks to their ability to photosynthesize, i.e. assimilate carbon dioxide (CO_2) into organic compounds using water and the energy of sun light. The Green Revolution of the late 20th century resulted in a global food production growth of *c.* 160% during the years 1960 to 2005 (Burney et al., 2010). It was primarily enabled through advances in technology (e.g. irrigation infrastructure) and organic chemistry (e.g. production of pesticides and synthetic nitrogen fertilizers) and the selection for plants with increased biomass partitioning into grains (Monteith and Moss, 1977; Ort et al., 2015). While improvement of yield potential by targeting this trait is theoretically exhausted, there is still large potential to improve the plants' use of solar energy conversion, i.e. photosynthetic efficiency (Ort et al., 2015).

One part of photosynthesis is the capture of CO_2 and its conversion into biomass. In this dissertation I will highlight two special types of photosynthetic carbon assimilation, the comprehensive understanding of which may in the future yield plant based solutions to (a) increase the efficiency of plant biomass production, by circumventing the metabolic inefficiency of an ancient enzyme and (b) allow the exploitation of new arable lands by increasing drought resistance or water use efficiency.

Plant photosynthetic pathways: C_3, C_4 and CAM

The enzyme that catalyzes the fixation of CO_2 in all photosynthetic organisms, RUBISCO (ribulose-1,5-*bis*phosphate (RuBP) carboxylase/oxygenase), is the key enzyme and CO_2 entry point of the Calvin Benson Cycle (CBC), a metabolic pathway for the stepwise reduction of CO_2 into three-carbon compound phosphoglycerate (PGA) (Benson and Calvin, 1950). PGA is further converted into sugar-phosphates building the basis of all biomass. RUBISCO evolved in a low-CO_2 world approximately 2.4 million years ago and exhibits a limited capability to distinguish between CO_2 and oxygen (O_2) (Morell et al., 1992; Farquhar et

al., 2011). At ambient air, approximately 25% of RUBISCO's activity catalyze the oxygenation rather than carboxylation of RuBP, which leads to the production of one molecule each PGA and 2-phosphoglycolate, a toxic compound that needs to be recycled via photorespiration. Since previously assimilated CO_2 is lost through photorespiration, this process can have detrimental effect on the plant's biomass productivity (Ogren, 1984). Reduced solubility of CO_2 at high temperature and hence a higher O_2:CO_2 ratio, intensify the oxygenation reaction and result in a higher photorespiration:photosynthesis ratio, further decreasing productivity (Jordan and Ogren, 1984; Long, 1991). In response to the selective pressure of rising O_2 levels from oxygenic photosynthesis, plants evolved two CO_2-concentrating mechanisms (CCM), termed C_4 photosynthesis and Crassulacean acid metabolism (CAM). C_4 and CAM species inhabit primarily hot, arid climates and despite being still essential, photorespiration is markedly reduced in these species (Sage, 2004; Zelitch et al., 2009; Lüttge, 2011)

Shared and distinct biochemistry

Key to both C_4 and CAM is the prefixation of CO_2 into a four-carbon acid by phospho*enol*pyruvate (PEP) carboxylase (PEPC), hence the name C_4 photosynthesis to abstract from C_3 plants in which a C_3 compound is the primary product of carbon assimilation. PEPC has a higher affinity for CO_2 than RUBISCO and does not react with oxygen (Hatch, 1987). In a second step, the C_4 acid is decarboxylated to release CO_2 at the site of RUBISCO. CAM and C_4 differ mainly by their way of separating the routes of carboxylation and decarboxylation (West-Eberhard et al., 2011). In CAM plants, PEPC is active at night and high levels of organic acids (mainly malate) accumulate in the vacuole to be released and decarboxylated in the following light period, separating the enzyme activity temporally (Fig. 1). The resulting diurnal fluctuations in organic acid levels have first been described in the early 19th century in the *Crassulacea* plant *Cotyledon calycina* (Heyne, 1816), giving rise to its name *Crassulacean acid metabolism.* The metabolic processes of CAM are divided into four phases (Osmond, 1978; Winter and Smith, 1996b). The nocturnal opening of stomata during phase I enable the uptake of CO_2 when evapotranspiration is low. Carbonic anhydrase converts external and respiratory CO_2 to bicarbonate, which is fixed into PEP by PEPC, finally resulting in the accumulation of malic acid. Phase II describes a short burst of CO_2 uptake at dawn, in which CO_2 is assimilated via both PEPC and RUBISCO. In phase III, which spans the majority of the warmer day, malic acid is decarboxylated and CO_2 is refixed by RUBISCO, thus allowing stomatal closure. During the

final phase IV in the cooler evening, stomata are re-opened and CO_2 is directly fixed by RUBISCO, the activity of which is coupled to light reactions (Winter and Smith, 1996b).

In C_4 plants the enzyme activities are separated spatially in two different cell types (Fig. 1). PEPC is active in the mesophyll cell (MC) and a resulting C_4-acid, malate or aspartate, is shuttled to the bundle sheath cell (BSC), where RUBISCO is active, for decarboxylation (Johnson and Hatch, 1968; Hatch, 1987). Plants in the family *Chenopodiaceae* and *Hydrocharitaceae* performing single-cell C_4 photosynthesis, in which enzyme activities are light-dependently compartmentalized across a single chlorenchyma cell and dimorphic chloroplasts, are currently considered rare exceptions to this type of photosynthesis (Freitag and Stichler, 2000; Edwards et al., 2004; Caemmerer et al., 2014; Bowes et al., 2002).

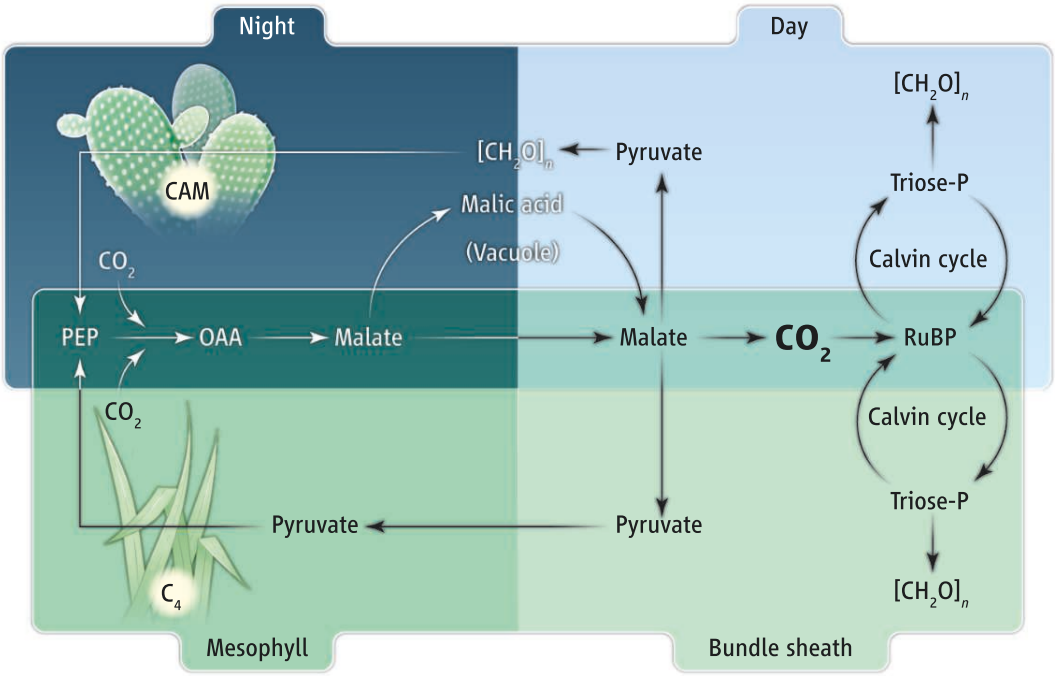


Fig. 1: Similar separation of carboxylation and decarboxylation. The CCMs are separated temporally in CAM (top) and spatially in C_4 (bottom). $[CH_2O]_n$, carbohydrates; OAA, oxaloacetate. Taken from (West-Eberhard et al., 2011).

In C_4 and CAM plants, different biochemical subtypes are traditionally distinguished based on the main decarboxylating enzyme (Hatch, 1987; Holtum et al., 2005). Plastidial NADP-malic enzyme (NADP-ME) and mitochondrial NAD-ME decarboxylate malate to pyruvate, which is subsequently converted to PEP, while cytosolic PEP carboxykinase (PEPCK) converts oxaloacetate to PEP thereby releasing CO_2. More recent evidence of a wide range of C_4 species suggests a less strict distinction in these three subtypes and mixed decarboxylase activities as a means to respond to fluctuating environmental conditions (Furbank, 2011; Wang et al., 2014b).

This is currently not evidenced in CAM plants. In C_4, PEP is shuttled back to the MC to enter another cycle of CO_2 assimilation. In CAM plants PEP is fed forward to synthesis of storage carbohydrates during the day, which are required to replenish the PEP pool in the following night (Holtum and Winter, 1982). This temporary carbohydrate repository can be either chloroplastic starch or extrachloroplastic hexoses, thus in combination with either decarboxylation type (ME vs. PEPCK) four metabolic variants are recognized for CAM plants (Holtum et al., 2005).

In addition to the soluble enzymes, both pathways highly depend on the activity of transport proteins, which are also largely shared between the two pathways. Depending on the biochemical subtype and decarboxylative route, the compartmentalized high metabolic fluxes of the C_4-cycle require many more transport steps compared to C_3 photosynthesis. For instance, NADP-ME plants depend on at least 30 transport steps to produce and export one molecule of triosephosphate, whereas only one step is required in C_3 plants (Weber and Caemmerer, 2010). These include transport of (i) the C_4 acids malate, oxaloacetate and aspartate across plastidial membranes (all subtypes) and mitochondrial membranes (NAD-ME, or PEP-CK subtypes), (ii) their decarboxylation products pyruvate and possibly alanine, (iii) PEP, (iv) triosephosphates as shuttle of ATP and/or NADPH and (v) additional co-transported substrates like protons (Weber and Caemmerer, 2010; Schlüter et al., 2016; for details see Fig. 3 of Manuscript II). Similarly, the CAM-cycle depends on (i) plastidial transport of triose-phosphates, PEP and sugar-phosphates, (ii) transport of hexoses as carbohydrate repositories and co-transport of acids resulting from CO_2 fixation and protons across the tonoplast and (iii) transport of C_4 and C_3 acids across mitochondrial and plastidial membranes (Holtum et al., 2005) (for details see Fig. 3 of Manuscript I).

Anatomical adaptations

In addition to biochemical adaptations, the evolution of the CCMs required typical and markedly distinct anatomical adaptations in each of the photosynthesis types.

Kranz anatomy. The spatial separation of the two CO_2 fixation steps of the C_4-cycle is enabled by separation of the required enzyme activities across MC and BSC. These two cell types are typically organized in a wreath-like anatomy around closely spaced veins (Kennedy et al., 1977; Brown, 1975). The outer layer of this so-called Kranz anatomy is derived from MC, which contain PEPC and take up CO_2 from the leaf intercellular air space. The layer proximal to the vein is derived from BSCs, which are enlarged and harbor RUBISCO activity. The BSC:MC

ratio is reduced and and a dense network of plasmodesmata allows metabolite shuttling between the two compartments. Often the cell walls at the interface of MC and BSC are impregnated by suberin to prevent CO_2 leakage (Dengler and Nelson, 1999).

Succulence. Although also observed in non-CAM plants, leaf succulence (leaf thickness) is often directly connected to CAM. Succulence is the result of large, undifferentiated mesophyll cells with large vacuoles and tight cell packaging (Gibson, 1982; Winter and Smith, 1996a). Since vacuolar storage capacity for acids, sugar and water is a key determinant for the magnitude of CAM (Nelson et al., 2005), succulence broadly correlates with CAM proportion to CO_2 assimilation (Nelson and Sage, 2008).

Ecophysiology and agriculture

C_4 plants typically have lower CO_2 compensation points than C_3 plants, i.e. they achieve positive net photosynthetic productivity at lower internal levels of CO_2 (Sage, 2004). Photosynthesis rates of CAM plants are generally lower than those of C_3 and C_4 plants (Winter et al., 2015). As a result of reduced stomatal conductance, water use efficiency (WUE), i.e. the unit CO_2 taken up per unit H_2O lost through transpiration, is markedly higher in C_4 and CAM plants compared to C_3 plants (Nobel, 1991; Borland et al., 2009; Taylor et al., 2009; Ehleringer and Monson, 1993). With stomata predominantly open at the cooler night when leaf-atmosphere water vapour pressure deficit is low, CAM plants achieve an even higher WUE than C_4 plants (Davis et al., 2014; Winter et al., 2005; Nobel, 1991). Enrichment of CO_2 concentration in the vicinity of RUBISCO allows for reduced levels of this enzyme which results in increased nitrogen use efficiency (NUE), i.e. CO_2 assimilated per molecule leaf nitrogen in C_4 and CAM plants compared to C_3 plants (Long, 1999; Ghannoum et al., 2010; Griffiths, 1989; Raven and Spicer, 1996).

The concentration of CO_2 at the site of RUBISCO comes at a higher energetic cost in both CCMs compared to C_3. In a typical C_3 plant the assimilation of one molecule of CO_2 requires three molecules of ATP and two molecules of NADPH. Depending on the subtype, assimilation of one molecule of CO_2 in C_4 plants requires at least two additional ATP for the phosphorylation of pyruvate to PEP (Zhu et al., 2008). If aspartate is the shuttled C_4 acid, one additional ATP is required to balance the NADPH consumption by conversion of oxaloacetate (OAA) to malate (Meister et al., 1996). Likewise, the diel cycling of malic acid accumulation and degradation in CAM plants poses an estimated 10% additional cost relative to C_3 plants (Winter and Smith, 1996b). However, the suppression of photorespiration results in greater or

comparable theoretical energy conversion efficiency in C_4 and CAM plants compared to C_3 plants (Zhu et al., 2008; Winter and Smith, 1996b). In addition, solar energy is not limiting in the typical (semi-)arid and open environments of C_4 and CAM plants.

The lack of systems-level analyses of CAM plants currently hinders a direct comparison of productivity between C_3, C_4 and CAM plants on an agricultural scale (Davis et al., 2015). However, when comparing the most highly productive species of each type the maximally achievable annual above-ground productivity is observed as $C_3 < CAM < C_4$ (Nobel, 1991). Sage and Zhu (2011) have estimated, that C_4 plants account for approximately a quarter of the earth's primary productivity, despite the fact that they are represented by probably two orders of magnitude less species than C_3 plants. Agronomically important C_4 crop and bioenergy plants are maize, sugar cane and sorghum. Despite their apparent lower annual productivity compared to C_4 plants, CAM plants, particularly Agave species, are also recognized as valuable bioenergy plants (Borland et al., 2009; Holtum et al., 2011). Approximately one third of the global land is semi-arid or arid (Mabbutt, 2009) and productivity in these environments is primarily limited by precipitation. Here, CAM plants are beneficial over C_4 plants due to their higher WUE and would not compete with high irrigation requiring C_3 crops (Davis et al., 2014).

Evolution and plasticity

C_4 and CAM are noticeable examples of convergent evolution. To date, it is estimated that C_4 photosynthesis evolved independently at least 61 times in 19 separate plant families (Sage, 2016) and CAM evolved independently in at least 343 genera in 35 plant families (Silvera et al., 2010). Central to the evolution of both traits is the recruitment and neofunctionalization of enzymes from C_3 ancestors functioning in different contexts (e.g. tricarboxylic acid cycle or gluconeogenesis) (Monson, 2003; Winter et al., 2015). This neofunctionalization may include changes in kinetics, regulation and tissue specificity (Sage, 2004) and usually does not affect the ancestral enzymatic functionality, but occurs on products of duplicated genes (Monson, 2003). As a result, enzymes employed in evolutionarily distant plants of the same CCM are often recruited from different ancestral isoforms (Ludwig, 2011; Christin et al., 2011; Sage et al., 2012). By contrast, there is recent evidence, that the same ancestral gene variant may be recruited for both C_4 and CAM in plant families with multiple recurrent origins of both CCMs (Christin et al., 2015), disclosing their shared evolutionary trajectories (Edwards and Ogburn, 2012; Christin et al., 2014). Despite the striking similarities -the utilization of an equal set of enzymes and adaptation to similar selective pressures- the evolution of C_4 and CAM in the same

plant tissue is mutually exclusive, primarily due to the different solutions to separate PEPC and RUBISCO activity and the anatomical specializations (Sage, 2002). In *Portulaca grandiflora*, currently the only described species to employ both types of photosynthesis, C_4 and CAM activities are separated amongst different tissues (Guralnick et al., 2002).

Although convergent and frequently recurrent, the evolution of the complex trait C_4 is a stepwise process (Sage et al., 2012). Anatomical (e.g. close veins) and genomic (duplicated genes encoding C_4-cycle enzymes) preconditioning as well as formation of a Proto-Kranz arrangement with enlarged BSCs containing more organelles are the first steps along the C_4 evolutionary path (Sage et al., 2012). Once these enablers are given, a photorespiratory CO_2 pump can be established by restriction of glycine decarboxylase (GDC) activity to the BSC, as is still present in extent C_3-C_4 species (Sage et al., 2012; Mallmann et al., 2014; Schlüter and Weber, 2016). Computational modeling approaches suggested that C_4 evolution follows a smooth "Mount Fuji" like path, where the order of individual steps is interchangeable within sequential modules (Heckmann et al., 2013). Intermediate C_3-C_4 plants found in several plant lineages are now widely accepted as *de facto* transitory states between C_3 and C_4, indicating that each incremental step of C_4 evolution grants a fitness advantage (Heckmann et al., 2013; Schlüter and Weber, 2016). This is supported by the family *Moricandia*, which comprises no C_4 but only C_3-C_4 species (Sage, 2004; Schlüter and Weber, 2016). In addition, C_3-C_4 metabolism as a stable feature supports the idea that CO_2 limitation rather than high temperatures or low water availability was the major selective pressure and C_4 evolution merely a way of lowering photorespiration (Sage, 2004). Once the photorespiratory pump is in place, high MC-specific activity of PEPC and the establishment and optimization of other C_4-cycle enzymes lead towards efficient C_4 photosynthesis.

In contrast to the discrete variations in C_4-cycle proportion, i.e. from C_3-C_4 over C_4-like to actual C_4 plants (Schlüter and Weber, 2016), there is a continuous variation in the extent of CAM proportion to CO_2 assimilation relative to C_3, which can be anywhere between 0 to 100% (Winter et al., 2015; Silvera et al., 2010). Three general types of CAM are distinguished based on their CO_2 acquisition strategy. In "CAM-cycling" plants, C_3 photosynthesis behind open stomata contributes primarily to CO_2 assimilation and scavenging of respiratory CO_2 into malate via PEPC only diminishes the nocturnal loss of CO_2 (Borland et al., 2011). The four phases described for CAM earlier are indicative for CAM plants *sensu stricto*, where CO_2 is primarily taken up at night and fixed into organic acids, which cycle reciprocally with carbohydrates over the course of the day. "CAM-idling" plants show no net CO_2 uptake and merely recycle respired CO_2 behind constantly closed stomata (Silvera et al., 2010).

The extent of the engagement in these three types can be determined both developmentally or environmentally. Even in constitutive or obligate CAM species like desert cacti or *Agavaceae* the establishment of CAM is ontogenetically programmed and C_3 photosynthesis may contribute considerably to CO_2 assimilation at some point in their life cycle (Winter et al., 2008; Winter and Holtum, 2011). Particular plasticity regarding CAM proportion is observed in inducible or facultative CAM species. Facultative CAM is induced upon environmental (salt stress, drought, heat) rather than developmental cues and is often reversible (Winter et al., 2008; Winter and Holtum, 2014). In these species, the possibility to switch to CAM enhances survival in environmentally unfavorable conditions. However, if these conditions do not occur, facultative CAM species may fulfill their entire life cycle performing solely C_3 photosynthesis (Winter and Holtum, 2007). Due to the experimentally controlled inducibility and reversibility, facultative CAM species emerge as valuable models to study the basal genetic requirements and signaling events for CAM (Winter and Holtum, 2014). Compared to C_4, the molecular evolution of CAM plants is much less studied (Silvera et al., 2010). It is currently not known whether facultative or weakly expressing CAM, CAM-cycling or CAM-idling species represent transitory states along the evolutionary path from C_3 to CAM (Hancock and Edwards, 2014; Winter et al., 2015), as is the case for C_3-C_4 plants in C_4 evolution.

The developmental establishment of the C_4-cycle in leaf gradients or the separation of C_3 and C_4 photosynthesis across different tissues is frequently observed in C_4 plants e.g. (Pick et al., 2011; Sekhon et al., 2011; Li et al., 2015; Aubry et al., 2014). However, in contrast to CAM plants, there are only few examples of plants with environmentally inducible C_4 photosynthesis (Reiskind et al., 1997; Bowes et al., 2002). Plastic short-term responses to fluctuating environments, such as light intensity or quality, are enabled biochemically through the utilization of alternative C_4 shuttles (Wang et al., 2014b; Stitt and Zhu, 2014).

Engineering C_4 and CAM into C_3 plants

The benefits of naturally existing C_4 and CAM crops in semi-arid to arid environments, validate the idea of engineering these complex traits into C_3 plants to enhance their photosynthetic efficiency or WUE (Leegood, 2013; Yang et al., 2015). Both traits evolved from diverse C_3 species through convergent evolution and there can be plasticity in the degree of C_4 or CAM within the same plant (Monson, 2003; Pick et al., 2011; Winter and Holtum, 2014). Thus, the pathways are theoretically not incompatible with C_3 organisms. Engineering the single-cell CCM of CAM may be more easily achieved than engineering C_4 as it does not depend on two

specialized cell types with their individual metabolic and anatomical adaptations (Yang et al., 2015). Theoretically, the same would be true for engineering single-cell C_4, but inducing spatially separated dimorphic chloroplasts within a single cell is likely to be even more complex than engineering the classical two-cell based CO_2 pump (Häusler et al., 2002).

While research in the CAM field is somewhat lagging behind in this respect (Borland et al., 2014; Yang et al., 2015), there have already been multiple attempts of engineering C_4 enzymes into different C_3 plants (Leegood, 2013; Schuler et al., 2016). However, at present there have been no examples of increased photosynthetic efficiency or growth (Schuler et al., 2016). Major obstacles include the complexity of the pathway and the large array of genes required to be transgenetically transferred in parallel into a suitable target crop plant. For instance, introduction of a few key enzymes like PEPC and NADP-ME, even if achieved cell-type specifically, would not enhance CO_2 pumping, if the required transport activities are not in place (Bräutigam et al., 2008). The number of genes encoding for the C_4-cycle alone, disregarding specialized Kranz anatomy or regulatory components, is beyond the scope of currently available techniques for transformation (Schuler et al., 2016). In addition, some of the essential core C_4 cycle enzymes, including a few transport proteins and likely a large array of regulatory proteins, are not yet identified on the molecular level.

Regulation of gene expression

C_4 and CAM photosynthesis are regulated on the post-transcriptional, post-translational and epigenetic level in addition to the transcriptional level (Sheen, 1999; Hibberd and Covshoff, 2010; Borland et al., 2009). Still, the transcriptional up-regulation and cell-specific expression of enzymes required for the core C_4 or CAM cycles was key to the evolution of these traits as evidenced in accumulating comparative transcriptomes of C_4 and CAM species (Bräutigam et al., 2011; Gowik et al., 2011; Wang et al., 2014a; Li and Brutnell, 2011; Aubry et al., 2014; Cushman et al., 2008).

Despite the overabundance of these datasets for C_4, the search for a "master switch" transcription factor, which could regulate the increased expression of multiple C_4 cycle genes in parallel, has so far yielded only few promising candidates (Westhoff and Gowik, 2010; Langdale, 2011). The GOLDEN2-LIKE transcription factors are associated with cell-type specific chloroplast dimorphism in C_4 monocot and dicot plants (Wang et al., 2013). Kranz patterning and possibly C_4-cycle integration are regulated by the SHORTROOT and SCARECROW network (Fouracre et al., 2014; Külahoglu et al., 2014). Due to the genetic

transformability of *Flaveria bidentis* and the presence of intermediate species representing the evolutionary path from C_3 to C_4 (Chitty et al., 1994; Mallmann et al., 2014), *Flaveria* has emerged as the model genus to study C_4 evolution and regulation of gene expression in dicots (Gowik et al., 2011). We analyzed transcriptomes of C_3 and C_4 *Flaveria spp.* to study the differential timing and light-response of gene expression and regulation under light-dark and constant light conditions (Manuscript II).

By contrast to C_4, only a few transcriptomics datasets are currently available for CAM (Hartwell et al., 2016) and knowledge of gene regulation is limited. Monocot and dicot facultative CAM species represent good models to study the minimal set of genes and regulatory mechanisms required to establish a CAM-cycle (Yang et al., 2015; Hartwell et al., 2016). We analyzed leaf transcriptomes of the reversibly inducible CAM plant *Talinum triangulare* in response to different levels of drought and elucidated the basal events and candidate regulatory factors leading to the establishment of weak CAM (Manuscript I).

References

Aubry, S., Kelly, S., Kümpers, B.M.C., Smith-Unna, R.D., and Hibberd, J.M. (2014). Deep Evolutionary Comparison of Gene Expression Identifies Parallel Recruitment of Trans - Factors in Two Independent Origins of C4 Photosynthesis. PLoS Genet. **10**: e1004365.

Benson, A.A. and Calvin, M. (1950). Carbon Dioxide Fixation By Green Plants. Annu. Rev. Plant. Physiol. **1**: 25–42.

Borland, A.M., Barrera Zambrano, V.A., Ceusters, J., and Shorrock, K. (2011). The photosynthetic plasticity of crassulacean acid metabolism: an evolutionary innovation for sustainable productivity in a changing world. New Phytol **191**: 619–633.

Borland, A.M., Griffiths, H., Hartwell, J., and Smith, J.A.C. (2009). Exploiting the potential of plants with crassulacean acid metabolism for bioenergy production on marginal lands. J Exp Bot **60**: 2879–2896.

Borland, A.M., Hartwell, J., Weston, D.J., Schlauch, K.A., Tschaplinski, T.J., Tuskan, G.A., Yang, X., and Cushman, J.C. (2014). Engineering crassulacean acidmetabolism to improve water-useefficiency. Trends in Plant Science **19**: 327–338.

Bowes, G., Rao, S.K., Estavillo, G.M., and Reiskind, J.B. (2002). C4 mechanisms in aquatic angiosperms: comparisons with terrestrial C4 systems. Functional Plant Biol. **29**: 379–392.

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., Hoffmann-Benning, S., Hofmann-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.

Brilhaus, D., Bräutigam, A., Mettler-Altmann, T., Winter, K., and Weber, A.P.M. (2016). Reversible Burst of Transcriptional Changes during Induction of Crassulacean Acid Metabolism in *Talinum triangulare*. Plant Physiol **170**: 102–122.

Brown, W.V. (1975). Variations in Anatomy, Associations, and Origins of Kranz Tissue. Am. J. Bot. **62**: 395–402.

Burney, J.A., Davis, S.J., and Lobell, D.B. (2010). Greenhouse gas mitigation by agricultural intensification. Proc. Natl. Acad. Sci. U.S.A. **107**: 12052–12057.

Caemmerer, von, S., Edwards, G.E., Koteyeva, N., and Cousins, A.B. (2014). Single cell C4 photosynthesis in aquatic and terrestrial plants: A gas exchange perspective. Aquatic Botany **118**: 71–80.

Chitty, J.A., Furbank, R.T., Marshall, J.S., Chen, Z., and Taylor, W.C. (1994). Genetic

transformation of the C4 plant, *Flaveria bidentis*. Plant J. **6**: 949–956.

Christin, P.-A., Arakaki, M., Osborne, C.P., and Edwards, E.J. (2015). Genetic enablers underlying the clustered evolutionary origins of C4 photosynthesis in angiosperms. Molecular Biology and Evolution **32**: 846–858.

Christin, P.-A., Arakaki, M., Osborne, C.P., Bräutigam, A., Sage, R.F., Hibberd, J.M., Kelly, S., Covshoff, S., Wong, G.K.-S., Hancock, L., and Edwards, E.J. (2014). Shared origins of a key enzyme during the evolution of C4 and CAM metabolism. J Exp Bot **65**: 3609–3621.

Christin, P.-A., Sage, T.L., Edwards, E.J., Ogburn, R.M., Khoshravesh, R., and Sage, R.F. (2011). Complex evolutionary transitions and the significance of C3-C4 intermediate forms of photosynthesis in *Molluginaceae*. Evolution **65**: 643–660.

Cushman, J.C., Tillett, R.L., Wood, J.A., Branco, J.M., and Schlauch, K.A. (2008). Large-scale mRNA expression profiling in the common ice plant, *Mesembryanthemum crystallinum*, performing C3 photosynthesis and Crassulacean acid metabolism (CAM). J Exp Bot **59**: 1875–1894.

Davis, S.C., LeBauer, D.S., and Long, S.P. (2014). Light to liquid fuel: Theoretical and realized energy conversion efficiency of plants using Crassulacean Acid Metabolism (CAM) in arid conditions. J Exp Bot **65**: 3471–3478.

Davis, S.C., Ming, R., LeBauer, D.S., and Long, S.P. (2015). Toward systems-level analysis of agricultural production from crassulacean acid metabolism (CAM): scaling from cell to commercial production. New Phytol **208**: 66–72.

Dengler, N.G. and Nelson, T. (1999). Leaf structure and development in C4 plants.

Edwards, E.J. and Ogburn, R.M. (2012). Angiosperm Responses to a Low-Co2 World: Cam and C4 Photosynthesis as Parallel Evolutionary Trajectories. Int. J Plant Sci. **173**: 724–733.

Edwards, G.E., Franceschi, V.R., and Voznesenskaya, E.V. (2004). Single-cell C_4 photosynthesis versus the dual-cell (Kranz) paradigm. Annu. Rev. Plant Biol. **55**: 173–196.

Ehleringer, J.R. and Monson, R.K. (1993). Evolutionary and ecological aspects of photosynthetic pathway variation. Annual Review of Ecology and Systematics **24**: 411–439.

Farquhar, J., Zerkle, A.L., and Bekker, A. (2011). Geological constraints on the origin of oxygenic photosynthesis. Photosyn. Res. **107**: 11–36.

Fouracre, J.P., Ando, S., and Langdale, J.A. (2014). Cracking the Kranz enigma with systems biology. J Exp Bot **65**: 3327–3339.

Freitag, H. and Stichler, W. (2000). A remarkable new leaf type with unusual photosynthetic tissue in a central asiatic genus of *Chenopodiaceae*. Plant Biology **2**: 154–160.

Furbank, R.T. (2011). Evolution of the C4 photosynthetic mechanism: are there really three C4 acid decarboxylation types? J Exp Bot **62**: 3103–3108.

Ghannoum, O., Evans, J.R., and Caemmerer, von, S. (2010). Chapter 8 Nitrogen and Water Use Efficiency of C4 Plants. In C4 Photosynthesis and Related CO2 Concentrating Mechanisms, Advances in Photosynthesis and Respiration. (Springer Netherlands: Dordrecht), pp. 129–146.

Gibson, A.C. (1982). The anatomy of succulence. In *Crassulacean acid metabolism* (eds. I.P. Ting and M. Gibbs), 1-17. (Proc. V ann. Symp. Bot)

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., and Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. Science **327**: 812–818.

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.

Griffiths, H. (1989). Carbon Dioxide Concentrating Mechanisms and the Evolution of CAM in Vascular Epiphytes. In Vascular Plants as Epiphytes, Ecological Studies. (Springer Berlin Heidelberg: Berlin, Heidelberg), pp. 42–86.

Guralnick, L.J., Edwards, G., Ku, M.S.B., Hockema, B., and Franceschi, V.R. (2002). Photosynthetic and anatomical characteristics in the C4-crassulacean acid metabolism-cycling plant, Portulaca grandiflora. In, pp. 763–773.

Hancock, L. and Edwards, E.J. (2014). Phylogeny and the inference of evolutionary trajectories. J Exp Bot **65**: 3491–3498.

Hartwell, J., Dever, L.V., and Boxall, S.F. (2016). Emerging model systems for functional genomics analysis of Crassulacean acid metabolism. Current Opinion in Plant Biology **31**: 100–108.

Hatch, M.D. (1987). C4 photosynthesis: a unique blend of modified biochemistry, anatomy and ultrastructure. BBA Reviews On Bioenergetics **895**: 81–106.

Häusler, R.E., Hirsch, H.J., Kreuzaler, F., and Peterhansel, C. (2002). Overexpression of C4-cycle enzymes in transgenic C3 plants: a biotechnological approach to improve C3-photosynthesis. J Exp Bot **53**: 591–607.

Heckmann, D., Schulze, S., Denton, A., Gowik, U., Westhoff, P., Weber, A.P.M., and Lercher, M.J. (2013). Predicting C4 Photosynthesis Evolution: Modular, Individually Adaptive Steps on a Mount Fuji Fitness Landscape. Cell **153**: 1579–1588.

Heyne, B. (1816). XVII. On the Deoxidation of the Leaves of *Cotyledon calycina*; in a Letter to A. B. Lambert, Esq., Vice-President of the Linnean Society. Transactions of the Linnean Society of London **11**: 213–215.

Hibberd, J.M. and Covshoff, S. (2010). The regulation of gene expression required for C4 photosynthesis. Annu. Rev. Plant Biol. **61**: 181–207.

Holtum, J.A.M. and Winter, K. (1982). Activity of enzymes of carbon metabolism during the induction of Crassulacean acid metabolism in *Mesembryanthemum crystallinum* L. Planta **155**: 8–16.

Holtum, J.A.M., Chambers, D., MORGAN, T., and Tan, D.K.Y. (2011). Agave as a biofuel feedstock in Australia. GCB Bioenergy **3**: 58–67.

Holtum, J.A.M., Smith, J.A.C., and Neuhaus, H.E. (2005). Intracellular transport and pathways of carbon flow in plants with crassulacean acid metabolism. Functional Plant Biol. **32**: 429–449.

Johnson, H.S. and Hatch, M.D. (1968). Distribution of the C4-dicarboxylic acid pathway of photosynthesis and its occurrence in dicotyledonous plants. Phytochemistry **7**: 375–380.

Jordan, D.B. and Ogren, W.L. (1984). The CO2/O2 specificity of ribulose 1,5-bisphosphate carboxylase/oxygenase. Planta **161**: 308–313.

Kennedy, R.A., Barnes, J.E., and Laetsch, W.M. (1977). Photosynthesis in C4 Plant Tissue Cultures. Plant Physiol.

Külahoglu, C. et al. (2014). Comparative transcriptome atlases reveal altered gene expression modules between two *Cleomaceae* C3 and C4 plant species. Plant Cell **26**: 3243–3260.

Langdale, J.A. (2011). C4 Cycles: Past, Present, and Future Research on C4 Photosynthesis. Plant Cell **23**: 3879–3892.

Leegood, R.C. (2013). Strategies for engineering C4 photosynthesis. Journal of Plant Physiology **170**: 378–388.

Li, P. and Brutnell, T.P. (2011). *Setaria viridis* and *Setaria italica*, model genetic systems for the Panicoid grasses. J Exp Bot **62**: 3031–3037.

Li, Y., Ma, X., Zhao, J., Xu, J., Shi, J., Zhu, X.-G., Zhao, Y., and Zhang, H. (2015). Developmental genetic mechanisms of C4 syndrome based on transcriptome analysis of C3 cotyledons and C4 assimilating shoots in *Haloxylon ammodendron*. PLoS ONE **10**: e0117175.

Long, S.P. (1999). Environmental Responses. In C4 Plant Biology (Elsevier), pp. 215–249.

Long, S.P. (1991). Modification of the response of photosynthetic productivity to rising temperature by atmospheric CO2 concentrations: Has its importance been underestimated? Plant Cell Environ. **14**: 729–739.

Ludwig, M. (2011). The molecular evolution of carbonic anhydrase in *Flaveria*. J Exp Bot **62**: 3071–3081.

Lüttge, U. (2011). Photorespiration in phase III of crassulacean acid metabolism: evolutionary and ecophysiological implications. Progress in Botany 72 **2**: 371–384.

Mabbutt, J.A. (2009). A New Global Assessment of the Status and Trends of Desertification. Envir. Conserv. **11**: 103–113.

Mallmann, J., Heckmann, D., Bräutigam, A., Lercher, M.J., Weber, A.P.M., Westhoff, P., and Gowik, U. (2014). The role of photorespiration during the evolution of C4 photosynthesis in the genus *Flaveria*. eLife **3**: e02478.

Meister, M., Agostino, A., and Hatch, M. (1996). The roles of malate and aspartate in C4 photosynthetic metabolism of *Flaveria bidentis* (L.). Planta **199**.

Monson, R.K. (2003). Gene duplication, neofunctionalization, and the evolution of C4 photosynthesis. Int. J Plant Sci. **164**: S43–S54.

Monteith, J.L. and Moss, C.J. (1977). Climate and the Efficiency of Crop Production in Britain. Philos. Trans. R. Soc. Lond., B, Biol. Sci. **281**: 277–294.

Morell, M.K., Paul, K., Kane, H.J., and Andrews, T.J. (1992). Rubisco: Maladapted or Misunderstood. Australian Journal of Botany **40**: 431–441.

Nelson, E.A. and Sage, R.F. (2008). Functional constraints of CAM leaf anatomy: tight cell packing is associated with increased CAM function across a gradient of CAM expression. J Exp Bot **59**: 1841–1850.

Nelson, E.A., Sage, T.L., and Sage, R.F. (2005). Functional leaf anatomy of plants with crassulacean acid metabolism. Functional Plant Biol. **32**: 409–419.

Nobel, P.S. (1991). Achievable productivities of certain CAM plants: basis for high values compared with C3 and C4 plants. New Phytol **119**: 183–205.

Ogren, W.L. (1984). Photorespiration: pathways, regulation, and modification. Annu. Rev. Plant. Physiol.

Ort, D.R. et al. (2015). Redesigning photosynthesis to sustainably meet global food and bioenergy demand. Proc. Natl. Acad. Sci. U.S.A. **112**: 8529–8536.

Osmond, C.B. (1978). Crassulacean acid metabolism: a curiosity in context. Annu. Rev. Plant. Physiol. **29**: 379–414.

Pick, T.R., Bräutigam, A., Schlüter, 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.

Raven, J.A. and Spicer, R.A. (1996). The Evolution of Crassulacean Acid Metabolism. In Crassulacean Acid Metabolism, Ecological Studies. (Springer Berlin Heidelberg: Berlin, Heidelberg), pp. 360–385.

Reiskind, J.B., Madsen, T.V., and Ginkel, L.C. (1997). Evidence that inducible C4-type photosynthesis is a chloroplastic CO2-concentrating mechanism in *Hydrilla*, a submersed monocot. Plant **20**: 211–220.

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. J Exp Bot: erw156.

Sage, R.F. (2002). Are crassulacean acid metabolism and C4 photosynthesis incompatible? In, pp. 775–785.

Sage, R.F. (2004). The evolution of C4 photosynthesis. New Phytol **161**: 341–370.

Sage, R.F. and Zhu, X.-G. (2011). Exploiting the engine of C(4) photosynthesis. J Exp Bot **62**: 2989–3000.

Sage, R.F., Sage, T.L., and Kocacinar, F. (2012). Photorespiration and the Evolution of C4 Photosynthesis. Annu. Rev. Plant Biol. **63**: 19–47.

Schlüter, U. and Weber, A.P.M. (2016). The road to C4 photosynthesis: Evolution of a complex trait via intermediary states. Plant Cell Physiol.: pcw009.

Schlüter, U., Denton, A.K., and Bräutigam, A. (2016). Understanding metabolite transport and metabolism in C4 plants through RNA-seq. Current Opinion in Plant Biology **31**: 83–90.

Schuler, M.L., Mantegazza, O., and Weber, A.P.M. (2016). Engineering C4 photosynthesis into C3 chassis in the synthetic biology age. Plant J.

Sekhon, R.S., Lin, H., Childs, K.L., Hansey, C.N., Buell, C.R., de Leon, N., and Kaeppler, S.M. (2011). Genome-wide atlas of transcription during maize development. Plant J. **66**: 553–563.

Sheen, J. (1999). C4 gene expression. Annu. Rev. Plant Physiol. Plant Mol. Biol. **50**: 187–217.

Silvera, K., Neubig, K.M., Whitten, W.M., Williams, N.H., Winter, K., and Cushman, J.C. (2010). Evolution along the crassulacean acid metabolism continuum. Functional Plant Biol. **37**: 995–1010.

Stitt, M. and Zhu, X.-G. (2014). The large pools of metabolites involved in intercellular metabolite shuttles in C4 photosynthesis provide enormous flexibility and robustness in a fluctuating light environment. Plant Cell Environ. **37**: 1985–1988.

Taylor, S.H., Hulme, S.P., Rees, M., Ripley, B.S., Ian Woodward, F., and Osborne, C.P. (2009). Ecophysiological traits in C3 and C4 grasses: a phylogenetically controlled screening experiment. New Phytol **185**: 780–791.

Tilman, D., Balzer, C., Hill, J., and Befort, B.L. (2011). Global food demand and the sustainable intensification of agriculture. Proc. Natl. Acad. Sci. U.S.A. **108**: 20260–20264.

Wang, L. et al. (2014a). Comparative analyses of C_4 and C_3 photosynthesis in developing leaves of maize and rice. Nat Biotechnol **32**: 1158–1165.

Wang, P., Kelly, S., Fouracre, J.P., and Langdale, J.A. (2013). Genome-wide transcript analysis of early maize leaf development reveals gene cohorts associated with the differentiation of C4 Kranz anatomy. Plant J. **75**: 656–670.

Wang, Y., Bräutigam, A., Weber, A.P.M., and Zhu, X.-G. (2014b). Three distinct biochemical subtypes of C_4 photosynthesis? A modelling analysis. J Exp Bot **65**: 3567–3578.

Weber, A.P.M. and Caemmerer, von, S. (2010). Plastid transport and metabolism of C3 and C4 plants - comparative analysis and possible biotechnological exploitation. Current Opinion in Plant Biology **13**: 257–265.

West-Eberhard, M.J., Smith, J.A.C., and Winter, K. (2011). Plant science. Photosynthesis, reorganized. Science **332**: 311–312.

Westhoff, P. and Gowik, U. (2010). Evolution of C4 photosynthesis - looking for the master switch. Plant Physiol **154**: 598–601.

Winter, K. and Holtum, J.A.M. (2007). Environment or development? Lifetime net CO2 exchange and control of the expression of Crassulacean acid metabolism in *Mesembryanthemum crystallinum*. Plant Physiol **143**: 98–107.

Winter, K. and Holtum, J.A.M. (2014). Facultative crassulacean acid metabolism (CAM) plants: powerful tools for unravelling the functional elements of CAM photosynthesis. J Exp Bot **65**: 3425–3441.

Winter, K. and Holtum, J.A.M. (2011). Induction and reversal of crassulacean acid metabolism in *Calandrinia polyandra*: effects of soil moisture and nutrients. Functional Plant Biol. **38**: 576.

Winter, K. and Smith, J. (1996a). Crassulacean acid metabolism: current status and perspectives.: 389–426.

Winter, K. and Smith, J.A.C. (1996b). An Introduction to Crassulacean Acid Metabolism. Biochemical Principles and Ecological Diversity. In Crassulacean Acid Metabolism, Ecological Studies. (Springer Berlin Heidelberg: Berlin, Heidelberg), pp. 1–13.

Winter, K., Aranda, J., and Holtum, J.A.M. (2005). Carbon isotope composition and water-use efficiency in plants with crassulacean acid metabolism. Functional Plant Biol. **32**: 381–388.

Winter, K., Garcia, M., and Holtum, J.A.M. (2008). On the nature of facultative and constitutive CAM: environmental and developmental control of CAM expression during early growth of *Clusia*, *Kalanchöe*, and *Opuntia*. J Exp Bot **59**: 1829–1840.

Winter, K., Holtum, J.A.M., and Smith, J.A.C. (2015). Crassulacean acid metabolism: a continuous or discrete trait? New Phytol **208**: 73–78.

Yang, X. et al. (2015). A roadmap for research on crassulacean acid metabolism (CAM) to enhance sustainable food and bioenergy production in a hotter, drier world. New Phytol **207**: 491–504.

Zelitch, I., Schultes, N.P., Peterson, R.B., Brown, P., and Brutnell, T.P. (2009). High glycolate oxidase activity is required for survival of maize in normal air. Plant Physiol **149**: 195–204.

Zhu, X.-G., Long, S.P., and Ort, D.R. (2008). What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? Current Opinion in Biotechnology **19**: 153–159.

Manuscript 1

Reversible Burst of Transcriptional Changes during Induction of Crassulacean Acid Metabolism in *Talinum triangulare*

Reversible Burst of Transcriptional Changes during Induction of Crassulacean Acid Metabolism in *Talinum triangulare*[1][OPEN]

Dominik Brilhaus, Andrea Bräutigam, Tabea Mettler-Altmann, Klaus Winter, and Andreas P.M. Weber*

Institute of Plant Biochemistry, Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine-University, D–40225 Düsseldorf, Germany (D.B., A.B., T.M.-A., A.P.M.W.); and Smithsonian Tropical Research Institute, Balboa, Ancón, Republic of Panama (K.W.)

ORCID IDS: 0000-0001-9021-3197 (D.B.); 0000-0002-5309-0527 (A.B.); 0000-0002-9161-4889 (T.M.-A.); 0000-0003-0970-4672 (A.P.M.W.).

Drought tolerance is a key factor for agriculture in the 21st century as it is a major determinant of plant survival in natural ecosystems as well as crop productivity. Plants have evolved a range of mechanisms to cope with drought, including a specialized type of photosynthesis termed Crassulacean acid metabolism (CAM). CAM is associated with stomatal closure during the day as atmospheric CO_2 is assimilated primarily during the night, thus reducing transpirational water loss. The tropical herbaceous perennial species *Talinum triangulare* is capable of transitioning, in a facultative, reversible manner, from C_3 photosynthesis to weakly expressed CAM in response to drought stress. The transcriptional regulation of this transition has been studied. Combining mRNA-Seq with targeted metabolite measurements, we found highly elevated levels of CAM-cycle enzyme transcripts and their metabolic products in *T. triangulare* leaves upon water deprivation. The carbohydrate metabolism is rewired to reduce the use of reserves for growth to support the CAM-cycle and the synthesis of compatible solutes. This large-scale expression dataset of drought-induced CAM demonstrates transcriptional regulation of the C_3–CAM transition. We identified candidate transcription factors to mediate this photosynthetic plasticity, which may contribute in the future to the design of more drought-tolerant crops via engineered CAM.

Drought is a major determinant of both plant survival in natural ecosystems and plant productivity in agriculture (Lobell and Gourdji, 2012). Plants have evolved a range of physiological and nonphysiological traits to cope with water deficit stress (Bartels and Sunkar, 2005). Drought adaptation includes leaf shedding in perennials or completing the life cycle while enough water is present in annuals. Few species (e.g. so-called resurrection plants) tolerate extreme dehydration (Ingram and Bartels, 1996) and resume life upon water resupply. Other plants have evolved strategies to cope with drought, such as specialized biochemical pathways, cells, tissues, and organs to survive water scarcity by mitigating the reduction in tissue water loss (Chaves et al., 2003).

Water limitation sensed in the roots or leaves triggers stress signals, which include but are not limited to abscisic acid (ABA; Bray, 1997; Zhu, 2002; Chaves et al., 2003; Bartels and Sunkar, 2005). The immediate consequence of an ABA signal is stomatal closure (Kollist et al., 2014). This can lead to short-term carbon dioxide (CO_2) limitation of photosynthesis, potentially causing oxidative stress, which is mitigated by protective systems (Bartels and Sunkar, 2005; Flexas et al., 2006). The accumulation of compatible solutes is induced to protect the cellular machinery from consequences of leaf water loss and to lower the water potential of the leaf (Hare et al., 1998). These molecules include sugars, such as raffinose, trehalose, and Suc; sugar alcohols like mannitol and inositol; and amino acids and their derivatives, such as Pro or Gly betaine (Hare et al., 1998; Elbein et al., 2003). The signaling cascade has been largely elucidated. ABA is bound by the PYRABACTIN RESISTANCE1 AND PYR1-LIKE REGULATORY COMPONENTS OF ABA RECEPTOR family proteins, which upon binding inhibit the TYPE 2C AND TYPE 2A PROTEIN PHOSPHATASES (PP2C; Cutler et al., 2010). This inhibition, in turn, relieves the inhibition of the protein kinase OPEN STOMATA1 and other kinases, which phosphorylate their targets including the ABA RESPONSIVE ELEMENT BINDING FACTORS. They finally trigger ABA-responsive gene expression (Park et al., 2009). The capability to reduce leaf water loss is variable between species, and leaf water loss exceeding a threshold irreversibly damages a leaf (Lawlor and Cornic, 2002).

[1] This work was supported by the International Graduate Program for Plant Science (iGRAD-Plant; to D.B.) and grants from Deutsche Forschungsgemeinschaft (EXC 1028; IRTG 1525 to A.P.M.W.). K.W. was supported by the Smithsonian Tropical Research Institute.

* Address correspondence to andreas.weber@hhu.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.plantphysiol.org) is: Andreas P.M. Weber (andreas.weber@hhu.de).

D.B., A.B., T.M.-A., K.W., and A.P.M.W. designed the research; D.B., K.W., and T.M.-A. performed the research; D.B., A.B., and T.M.-A. analyzed data; D.B., A.B., T.M.-A., K.W., and A.P.M.W. wrote the article.

[OPEN] This article is available without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.15.01076

The tropical herbaceous dicot *Talinum triangulare* (Jacq). Willd. [according to The Plant List Kew now considered a synonym of the accepted name *Talinum fruticosum* (L.) Juss., but referred to within this work as *T. triangulare* to relate to extensive ecophysiological work on "*T. triangulare*" by Herrera et al. (1991), Taisma and Herrera (1998), Herrera (1999), Taisma and Herrera (2003), and Herrera et al. (2015)], in the family *Talinaceae* (formerly *Portulacaceae*), responds to intermittent drought (Harris and Martin, 1991; Herrera et al., 1991) in two ways. Leaves change from a horizontal to a vertical orientation and exhibit leaf-rolling (Herrera et al., 1991; Taisma and Herrera, 1998; Herrera, 1999). Leaves induce Crassulacean acid metabolism (CAM) in a reversible, facultative manner (Taisma and Herrera, 1998; Herrera, 2009; Winter and Holtum, 2014; Herrera et al., 2015).

CAM is a carbon concentrating mechanism allowing stomatal closure during the day as atmospheric CO_2 is primarily assimilated during the night by PHOSPHO-*ENOL*PYRUVATE CARBOXYLASE (PEPC; Osmond, 1978). The produced organic acids (mainly malic acid) are decarboxylated during the following day to provide CO_2 for the secondary, light-driven carboxylation via Rubisco. The magnitude of CAM varies between and within plant species (Borland et al., 2011). While constitutive CAM plants are ontogenetically determined to engage in this photosynthetic mode at some point in their life cycle, in facultative CAM plants, a transition to CAM occurs in response to water-deficit stress (Winter et al., 2008; Winter and Holtum, 2014). The ability to experimentally control the timing of CAM induction and CAM-to-C_3 reversal makes facultative CAM plants excellent systems to identify key components of CAM.

Extensive physiological, molecular, and mutant-based studies of facultative CAM were focused on, but not restricted to, the halophyte *Mesembryanthemum crystallinum*, and aided to understand central concepts of the CAM-cycle (Winter and von Willert, 1972; Holtum and Winter, 1982; Cushman et al., 2008a). These included a microarray-based large scale gene expression data set (Cushman et al., 2008b), which described major changes in mRNA steady-state levels during the salinity-induced transition from C_3 to CAM. Reversibility of CAM back to C_3 photosynthesis after irrigation was reinitiated has been documented for several species including *T. triangulare* (for review, see Winter and Holtum, 2014). To date, it is not fully understood how this metabolic plasticity is transcriptionally accomplished (Yang et al., 2015). We studied the *T. triangulare* transition from well-watered to water-limited and back to well-watered conditions (1) to test how CAM induction and its reversion are controlled at the level of mRNA abundance and (2) to evaluate CAM- and drought-specific changes in metabolite levels.

RESULTS

Nocturnal Acidification

Following germination, *T. triangulare* was grown well-watered for 28 d. Subsequently, water was withheld for 12 d before rewatering. Fresh weight to dry weight ratio after 4, 9, and 12 d of water deprivation showed slight, continuous (1.1%, 4.8%, and 15.0%, respectively) decreases compared to day 0, which were significant on day 12 (Fig. 1A). Measurements of titratable acidity were used as an indicator of the presence or absence of CAM activity in leaves (Fig. 1B). In well-watered plants, acidity levels were low (10 μmol H^+ g^{-1} fresh weight) and did not change in the course of the night. Withholding water for 4 d did not significantly alter titratable acidity compared to well-watered plants. Significant nocturnal acidification, indicative of CAM, was observed in plants from which water was withheld for 9 and 12 d, consistent with previous findings of Herrera et al. (1991) and Winter and Holtum (2014). Drought stress also resulted in pronounced leaf rolling (Fig. 1C). Following rewatering for 2 d, plants returned to a full C_3 photosynthetic pattern without nocturnal increases in tissue acidity and unrolled leaves (Fig. 1, B and C). However, fresh weight to dry weight ratio was significantly lower (22.7%) compared to day 0 (Fig. 1A).

Changes in Transcript and Metabolite Levels

Samples for mRNA-Seq were taken at midday and midnight prior to the drought treatment (day 0, well-watered), on day 4, day 9, and day 12 of drought, and again 2 d after rewatering (Fig. 2, A and B). Biological triplicates yielded, on average, 40.7 million reads over the time course (Supplemental Table S1). Of those, an average of 52% could be mapped for quantification to the reference genome of Arabidopsis (*Arabidopsis thaliana*; Supplemental Table S1; "Materials and Methods"). The mapped reads matched 16,766 Arabidopsis genes with at least one read per gene. This is comparable to the yield of cross species mapping with equidistant species for both number of genes matched and percentage of mapped reads (Gowik et al., 2011). Within-species mapping on the assembled contigs resulted in 81% mapped reads but indicated that the assembly suffers from the known limitations of transcriptome assembly such as contig fragmentation (Franssen et al., 2011; Supplemental Dataset S2).

Statistical evaluation including multiple hypothesis testing correction (by DESeq2; for details, see "Materials and Methods") identified significantly differential gene expression for 4,628 genes (28% of the whole transcriptome) at midday and 5,191 genes (31% of the transcriptome) at midnight on day 9 as well as 6,143 genes (37% of the transcriptome) at midday and 6,565 genes (39% of the transcriptome) at midnight on day 12 compared to well-watered plants (day 0), i.e. during the 2 d during which pronounced CAM activity was observed (Fig. 2B).

On day 9, 2,117 (12.6%) and 2,213 (13.2%) genes were upregulated, while 2,511 (15%) and 2,978 genes (17.8%) were downregulated at midday and midnight, respectively. On day 12, 2,713 (16.2%) and 2,897 (17.3%) genes were upregulated, while 3,430 (20.5%) and 3,668 genes (21.9%) were downregulated at midday and midnight, respectively. A Venn analysis indicated that 1,634

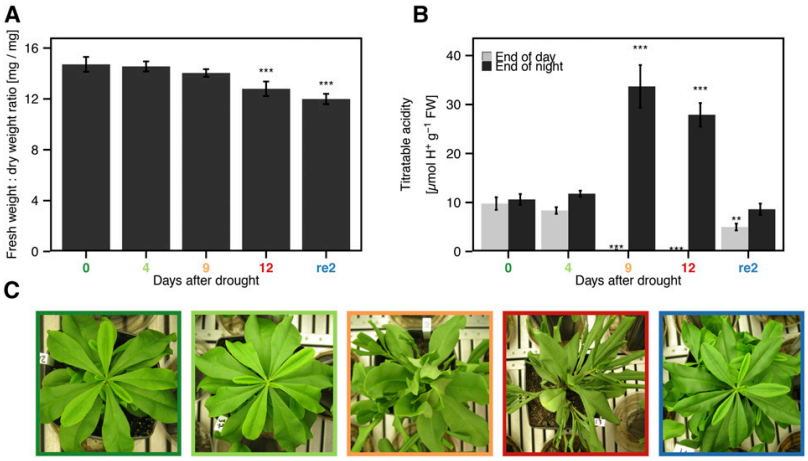


Figure 1. Time course in response to 0 (darkgreen), 4 (light green), 9 (orange), and 12 (red) d of drought and 2 d after rewatering (re2, blue) in *T. triangulare*. A, Fresh weight to dry weight ratio (mean ± SE of leaves harvested at the middle of the day and the middle of the night, $n = 14–16$). B, Levels of titratable acidity of leaves (mean ± SE, $n = 8–16$). C, Representative pictures of plants during the course of the experiment. Asterisks indicate Student's *t* test significance in comparison to day 0 at $***P < 0.001$, $**P < 0.01$, and $*P < 0.05$.

genes and 2,180 genes were shared among the genes up- and downregulated at midday, respectively (Fig. 2B), while at midnight, 1,545 upregulated and 1,865 downregulated genes were shared (Supplemental Fig. S1.). A small percentage (1.9% and 2.1% of upregulated and 1.2% and 1.9% of downregulated genes at midday and midnight, respectively) was exclusively differentially regulated on day 9. On day 4 of water

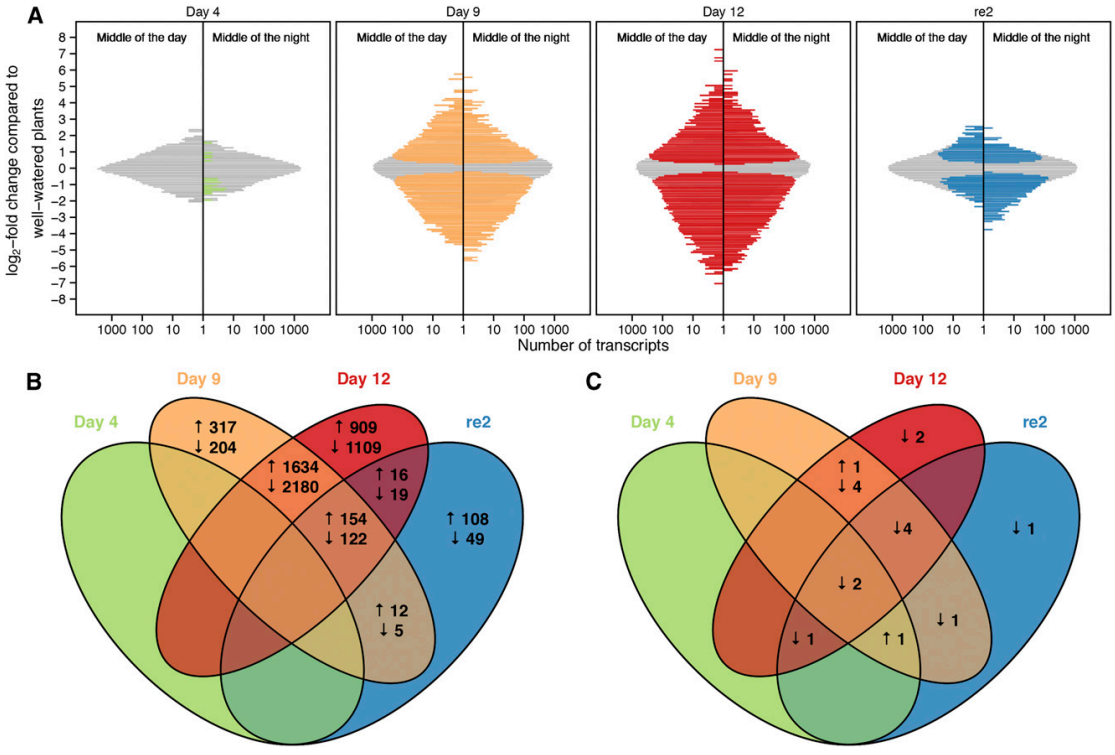


Figure 2. Changes in leaf transcriptomes and metabolomes under varying levels of water availability. A, Histograms of $\log_2$-fold changes in gene expression compared to day 0 ($\log_{10}$-scaled). Colored bars indicate significant changes (DESeq2, $q < 0.01$). B and C, Venn diagrams representing overlapping changes (↑: increased, ↓: depleted) in gene expression (B, DESeq2 $q < 0.01$, $n = 3$, 16,766 genes analyzed in total) or metabolite levels (C, Student's *t* test, $P < 0.05$, $n = 3–4$, 39 metabolites measured in total) at the middle of the day between water-limited stages compared to day 0. See Supplemental Figure S1 for analogous Venn diagrams of changes at the middle of the night.

deprivation, no differentially expressed genes (DEGs) were detectable compared to well-watered conditions at midday, while at midnight 16 genes were up- and 35 were downregulated (Supplemental Fig. S1). Upon rewatering, the number of significantly DEGs was reduced to 485 at midday (3% of the transcriptome) and 1,632 (10% of the transcriptome) at midnight. To compare the changes of steady-state levels of mRNA and metabolites, 39 metabolites were quantified at midday and midnight on days 0, 4, 9, 12, and after rewatering for 2 d (Fig. 2C; Supplemental Fig. S1; Supplemental Table S2). On day 4, nine metabolites were significantly different from well-watered plants at either midday or midnight. In leaves harvested at day 9 and 12, 13 and 14 of the 39 metabolites differed significantly at midday, while 6 and 11 metabolites differed significantly at midnight, respectively. The highest total number of significant changes was observed in rewatered plants, 10 at midday and 16 at midnight compared to day 0. Both the transcriptome analysis and the targeted metabolite profiling showed that transcriptome and metabolome were markedly altered in *T. triangulare* plants, which experience drought. While changes in mRNA abundance were largely reversed upon rewatering (Fig. 2A), the metabolic state remained altered.

CAM-Related Transcriptional Changes

Transcript abundance of known enzymes of the CAM-cycle sensu stricto (i.e. carboxylation and decarboxylation) and sensu lato (i.e. encoding auxiliary steps such as starch turnover and glycolysis for phospho*enol*pyruvate [PEP] generation) were analyzed during C_3-CAM-C_3 transitions. Transcripts encoding four CAM-cycle enzymes sensu stricto, namely PEPC, NADP-MALIC ENZYME (NADP-ME), NAD-MALIC ENZYME (NAD-ME), and PYRUVATE, *ORTHO*PHOSPHATE DIKINASE (PPDK) had higher steady-state levels at both midnight and midday on day 9 and/or day 12 of water-limitation (Fig. 3; Supplemental Dataset S3).

CARBONIC ANHYDRASE (CA) catalyzes the hydration of CO_2 to HCO_3^- at physiological pH and thereby is thought to support providing PEPC with its substrate in CAM plants (Tsuzuki et al., 1982). The gene encoding cytosolic BETA-CARBONIC ANHYDRASE3 (BCA3) was unchanged in expression upon drought but already highly expressed in C_3 conditions (2,498 rpm on day 0). This is in agreement with an earlier study finding no differences in CA activity in *M. crystallinum* in C_3 and CAM mode (Tsuzuki et al., 1982). Two lowly expressed genes encoding CA isoforms (alpha CA3 between 2.7 and 10 rpm and BCA5 between 23.7 and 80.7 rpm) were upregulated at midday on both days. Upregulation of major carbonic anhydrase isoforms as it was found for *M. crystallinum* (Cushman et al., 2008b) could not be detected in *T. triangulare*.

The gene *PPC* encodes PEPC, which catalyzes the CO_2 carboxylation at night. *PPC* was upregulated 25-fold at midnight (to 15,510 rpm, expression rank 4 on day 12) on day 9 and 12, respectively. The coding sequences of *PPC* vary between C_3, C_4, and CAM plants (Bläsing et al., 2000; Paulus et al., 2013). The coding sequences of the four *T. triangulare PPC* contigs with on average at least 100 reads mapped were extracted from the assembly, translated, and aligned with PEPC sequences from various C_3, C_4, and CAM plants (http://www.uniprot.org; Supplemental Fig. S2; Supplemental Table S3). The *T. triangulare* contigs encoding for PEPC showed characteristics of both C_3 and C_4 PEPCs. PEP saturation kinetics are known to be determined by the amino acid at position 780 (counting based on *Zea mays* sequence CAA33317), which in C_4 plants is typically Ser and in C_3 plants Ala (Bläsing et al., 2000). Sensitivity to malate inhibition is determined by the amino acid at position 890, Gly in C_4 plants, and Arg in C_3 plants (Paulus et al., 2013). While at position 890, all contigs encoded for the C_3-typical Arg in *T. triangulare*; at position 780, contig Tt63271 (7,638 rpm at midnight on day 9) and Tt9871_8 (1,198 rpm at midnight on day 12) encoded for the C_3 typical Ala, Tt9871_4 (1,409 rpm at midnight on day 9) and Tt9871_6 (399 rpm at midnight on day 12) encoded for the C_4-typical Ser (Supplemental Fig. S2). Read mapping on the contig level identified all isoforms as upregulated during CAM (Supplemental Table S3); however, it cannot be determined whether Tt63271 and Tt9871_8, Tt9871_4 and Tt9871_6 are alleles or recent duplicates.

Oxaloacetate (OAA) resulting from PEP carboxylation is reduced to malate by MALATE DEHYDROGENASE (MDH). The cytosolic MDH was unchanged but constitutively highly expressed in C_3 conditions (941 rpm at midday on day 0) and during CAM (1,141 rpm, expression rank 110 at midday on day 12). Malate is stored as malic acid in the vacuole during the night (Cheffings et al., 1997). Genes encoding two malate channels of the ALUMINUM-ACTIVATED MALATE TRANSPORTER (ALMT) family were upregulated 6-fold and 28-fold at midnight, but reached only 4 and 19 rpm after upregulation, respectively. While below 1 rpm in abundance at any other day, transcripts encoding the TONOPLAST DICARBOXYLATE TRANSPORTER are detectable on day 12 (MD: 5 rpm, MN: 11 rpm). *VHA-B3,* encoding the subunit B3 of the vacuolar ATPase (V-ATPase), which sustains the electrochemical gradient during import of malic acid (White and Smith, 1989), is the major expressed subunit, being significantly upregulated 2-fold on day 12. Of the 11 other genes encoding for V-ATPase subunits, two were unaltered and eight genes were slightly but significantly downregulated during CAM (Supplemental Table S4). *TYPE I PROTON-TRANSLOCATING PYROPHOSPHATASE* is downregulated 2-fold and 3-fold at midday of days 9 and 12.

During the day, malic acid is released from the vacuole and, depending on the species, is believed to be decarboxylated by NADP-ME, NAD-ME, and/or PEP CARBOXYKINASE (PEPCK; Dittrich, 1976). In *T. triangulare*, the plastidial *NADP-ME4* was upregulated 5-fold at midday to 780 rpm and the apparent cytosolic *NADP-ME1* was upregulated 6-fold to 579 rpm, while

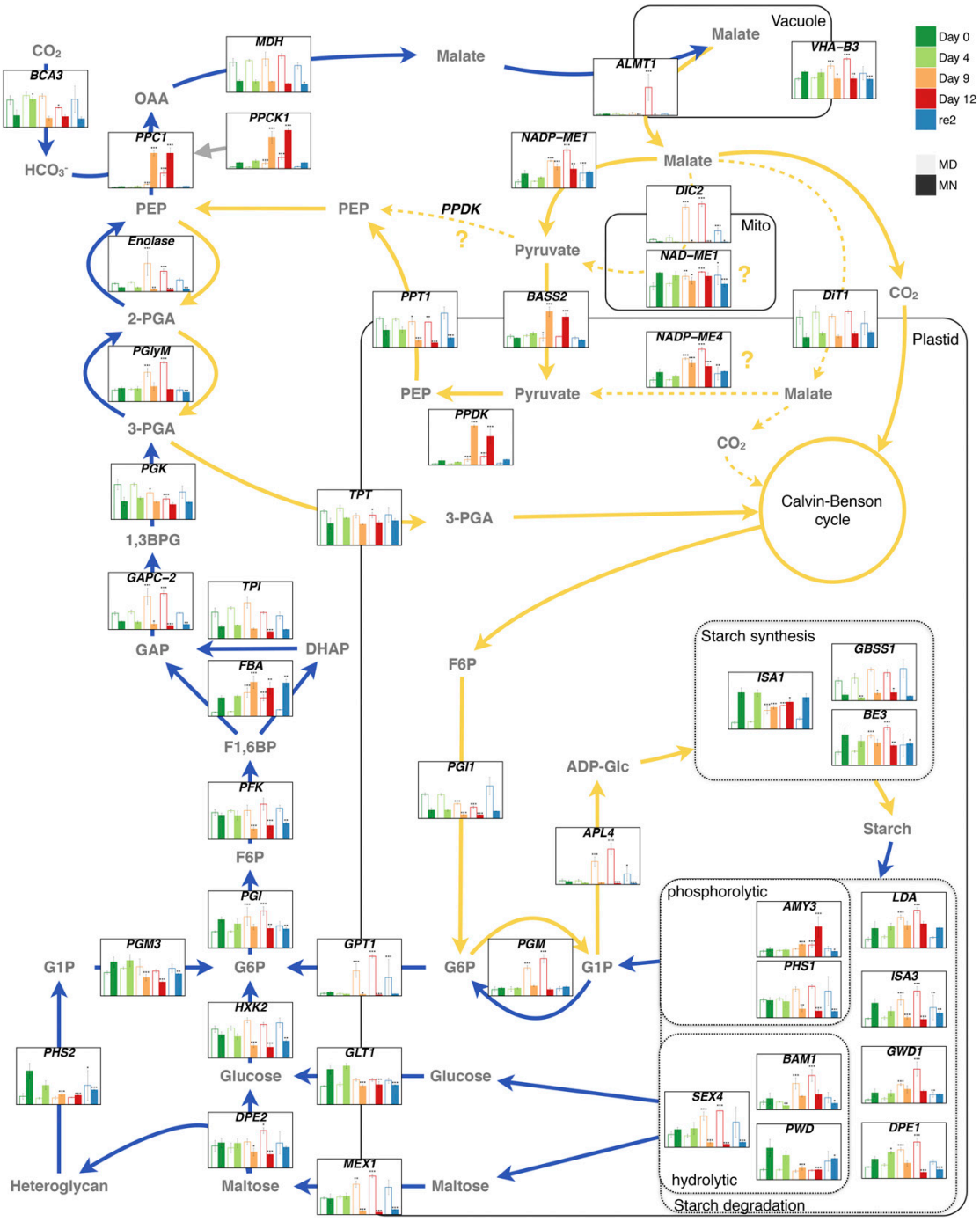


Figure 3. Abundances of CAM genes sensu stricto and sensu lato. Scheme of carbon assimilation via CAM and gene expression of central enzymes and transporters. Metabolites are represented in gray. Transcript levels were measured at the middle of the day and the middle of the night in leaves of *T. triangulare* plants under five different stages of water availability. Scaled to largest expression by gene; mean ± SD, $n = 3$. Asterisks indicate differential gene expression in comparison to day 0 as determined by

NAD-ME1 was 2-fold upregulated to 173 rpm and *PEPCK* was not significantly upregulated. The *T. triangulare* contigs of NADP-ME extracted from the assembly with high-read mappings possess a target peptide and cluster with the plastidial AtNADP-ME4 with high bootstrap support (Supplemental Fig. S3; Supplemental Table S5). In contrast to the downregulation of one isoform of *NADP-ME* at midnight in *M. crystallinum* (Cushman et al., 2008b), both genes encoding NADP-ME isoforms were upregulated at both midday and midnight on both CAM days. Transcript amounts of the plastidial DICARBOXYLATE TRANSPORTERS were unaltered (DiT1, expression between 256 and 381 rpm at midday, DiT2 between 48 and 105 rpm), while all three mitochondrial DICARBOXYLATE CARRIERS were upregulated at midday on both CAM days (DIC1 15-fold to 787 rpm, DIC2 12-fold to 899 rpm, DIC3 14-fold to 310 rpm on day 12). The plastidial pyruvate importer BILE-ACID SODIUM SYMPORTER2 (BASS2) was upregulated at night on both CAM days (4-fold to 411 rpm on day 9). Pyruvate, produced by malate decarboxylation, is phosphorylated to PEP by PPDK encoded by a single gene (9-fold upregulated to 18,371 rpm at midnight on day 9) and fed back to gluconeogenic starch synthesis (Kluge and Osmond, 1971). Extraction, alignment, and quantification of two different contigs encoding for PPDK (Supplemental Table S6) revealed markedly increased transcript amounts during CAM exclusively for one contig, Tt26901, which encodes for a protein with a 77 amino acid shorter N terminus. At midnight on day 12 transcript amounts for Tt26901 were 59-fold higher than for the longer PPDK encoding contig Tt24575 (Supplemental Table S6). Transcript amounts of the PEP/PHOSPHATE TRANSLOCATOR (PPT), catalyzing the export of PEP to the cytosol, were downregulated during CAM (4-fold to 23 rpm at midnight on day 12).

Another important aspect of CAM photosynthesis is starch turnover and its connection to the carboxylation/decarboxylation cycle. During the night, starch is degraded, likely both via the phosphorolytic and hydrolytic pathways, to provide PEP for PEPC via glycolysis (Weise et al., 2011). Four genes encoding enzymes, required both for phosphorolytic and hydrolytic starch degradation according to Weise et al. (2011) and Streb and Zeeman (2012), LIMIT DEXTRINASE (LDA), ISOAMYLASE3 (ISA3), GLUCAN WATER DIKINASE (GWD1), and DISPROPORTIONATING ENZYME1 (DPE1), were upregulated at midday on both CAM days (e.g. on day 12: 4-fold to 178 rpm, 4-fold to 195 rpm, 10-fold to 995 rpm and 3-fold to 197 rpm, respectively). *ISA3* is downregulated 3-fold at midnight on both CAM days and *DPE1* is downregulated at midnight of day 12. Genes encoding enzymes specific for hydrolytic starch degradation, PHOSPHOGLUCAN PHOSPHATASE (abbreviated SEX4 for STARCH EXCESS4) and BETA-AMYLASE1 (BAM1) were upregulated at midday (5-fold to 863 rpm and 3-fold to 150 rpm, respectively) as well, while the gene encoding PHOSPHOGLUCAN, WATER DIKINASE (PWD) was downregulated at midnight on both CAM days. Of the two exporters, MALTOSE EXPORTER1 (MEX1) and PLASTIDIC GLC TRANSLOCATOR1 (GLT1) exporting the hydrolytic breakdown products maltose and Glc (Weber et al., 2000; Niittylä et al., 2004), only the gene encoding MEX1 was significantly upregulated 2-fold on both CAM days at midday. MEX1 and GLT1 were both downregulated at midnight on both CAM days. The cytosolic enzymes catalyzing the conversion of Glc and maltose to Glc-phosphates (HEXOKINASE2 [HXK2], DPE2, STARCH PHOSPHORYLASE2 [PHS2] and PHOSPHOGLUCOMUTASE3 [PGM3]) were all downregulated at midnight upon drought (e.g. on day 12: 2-fold to 78 rpm, 2-fold to 110 rpm and 3-fold to 111 rpm, respectively). Of the enzymes specific for phosphorolytic starch degradation, the gene encoding ALPHA-AMYLASE3 (AMY3) was upregulated (4-fold on day 12 to 702 rpm), while the plastidial PHS1 gene was downregulated (3-fold on day 12 to 252 rpm) at midnight on both CAM days. The Glc-6-phosphate (G6P) exporter (*GPT*) expression was highly induced at midday upon CAM induction (17-fold; 1,095 rpm on day 12) and downregulated at midnight on day 12 (5-fold to 20 rpm).

The genes encoding cytosolic glycolytic enzymes PHOSPHOGLUCOSE ISOMERASE (PGI, upregulated 2-fold to 171 rpm at midday), PHOSPHOFRUCTOKINASE (PFK, constitutive level at midday ranging from 275 to 361 rpm), FRU-BISPHOPHATE ALDOLASE (FBA, upregulated 5-fold at midday on day 12 to 1,247 rpm), TRIOSEPHOSPHATE ISOMERASE (TPI, constitutive level at midday ranging from 980 to 1,541 rpm), SUBUNIT C2 OF GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (GAPC-2, upregulated 2-fold at midday to 2,649 rpm), PHOSPHOGLYCERATE MUTASE (PGlyM, upregulated at midday 4-fold to 675 rpm), and ENOLASE (upregulated 3-fold to 2,273 rpm) were of high abundance or more abundant in CAM at midday. The gene encoding PHOSPHOGLYCERATE MUTASE (PGK) is of lower abundance (2-fold at midday to 1,167 rpm on day 12). The glycolytic enzymes, whose transcripts

Figure 3. (*Continued.*)
DESeq2, * $q < 0.05$, ** $q < 0.01$, and *** $q < 0.001$. Blue and yellow arrows represent reactions occurring at night and day, respectively. Abbreviations are explained in the text and in Supplemental Dataset S3. Question mark and dotted arrows indicate putative activity of plastidial NADP-ME, mitochondrial NAD-ME, and cytosolic PPDK as discussed in the text. Separation into phosphorolytic and hydrolytic starch degradation is based on the models presented by Weise et al. (2011) and Streb and Zeeman (2012). 1,3-BPG, 1,3-Bisphosphoglycerate; 2-PGA, 2-phosphoglycerate; 3-PGA, 3-phosphoglycerate; DHAP, dihydroxyacetone phosphate; F1,6BP, Fru 1,6-bisphosphate; F6P, Fru 6-phosphate; G1P, Glc 1-phosphate; G6P, Glc 6-phosphate; GAP, glyceraldehyde 3-phosphate; Mito, mitochondrion; re2, 2 d after rewatering.

were more abundant, produce 3-phosphoglycerate (3-PGA), which may enter the chloroplast via TRIOSE PHOSPHATE/PHOSPHATE TRANSLOCATOR (TPT, constitutive level ranging from 1,632 to 2,905 rpm). Triose-phosphates (3-PGA and GAP) resulting from Rubisco-based carbon fixation and from recycling the pyruvate out of the decarboxylation reaction can be stored as starch. The genes encoding starch precursor biosynthetic enzymes, PHOSPHOGLUCOMUTASE (PGM) and the LARGE SUBUNIT 4 OF ADP-GLC PYROPHOSPHORYLASE (APL4), were more abundant (6-fold to 1063 rpm and 9-fold to 3790 rpm at midday on day 12, respectively) and transcript levels of the GRANULE-BOUND STARCH SYNTHASES (GBSS) were unchanged while those of ISA1 and STARCH BRANCHING ENZYME3 (BE3) were significantly less abundant (4-fold to 66 rpm and 3-fold to 296 rpm at midday, respectively). The transcript levels of more abundant genes involved in carboxylation, decarboxylation, glycolysis, and gluconeogenesis were reduced to the levels of well-watered plants upon rewatering, except for *GWD* (3-fold upregulated at midday), *GPT1* (4-fold upregulated at midday), and *NADP-ME1* (3-fold upregulated at midday), which remained highly abundant on day 2 after rewatering.

Malate, Citrate, Soluble Sugars, and Starch

In C_3 performing plants, malate levels were somewhat higher at midday than at midnight (Fig. 4, day 0, 4, and 2 d after rewatering). By contrast, in plants exhibiting CAM (Fig. 4, days 9 and 12; Supplemental Table S2), malate levels were up to 8-fold higher at midnight as compared to midday and up to 4-fold higher as compared to midnight during C_3. The increased malate levels at midnight are consistent with the increased acidification measured at the end of the night (Fig. 1B) but did not reflect the full diurnal amplitude during the light/dark cycle. As midnight levels are higher than midday levels, malate decarboxylation may occur quite rapidly during the first 6 h of the light period (Fig. 4). Citrate showed a similar pattern to malate, as observed in other CAM species (Winter and Smith, 1996; Lüttge, 2002), with increased amounts during the night in CAM plants compared to C_3 plants, albeit with a lower amplitude than malate. In agreement with the C_3 plant Arabidopsis (Smith et al., 2004), *T. triangulare* leaves performing C_3 photosynthesis showed high Glc, Fru, and Suc pool sizes at the middle of the day, reflecting high photosynthetic activity (Fig. 4). While Glc and Fru levels dropped at night, Suc levels in the middle of the night were high, probably due to Suc synthesis from starch during the night as reported previously (Chia et al., 2004; Smith et al., 2004). The amounts of Glc, Fru, and Suc during the day and night dropped in CAM conditions (days 9 and 12) and were accompanied by reduced amounts of starch. The total amount of starch showed a continuous decrease from days 4 to 9 to 12. Two days after rewatering, starch levels were increased to the amounts before drought, while Glc, Fru, and Suc pools were not fully restored.

The photorespiratory metabolites glycerate and glycolate were significantly reduced in CAM conditions (Supplemental Fig. S4). During the day, glycerate was depleted 40.2-fold and 238.9-fold on days 9 and 12, and during the night, 3.6-fold and 16.7-fold. Glycolate was depleted during the day on days 9 and 12 (2.2-fold and 3.1-fold).

The contents of the putative compatible solute raffinose increased significantly during day (14.4-fold and 12.3-fold) and night (9.2-fold and 4.7 fold) on days 9 and 12 compared to day 0, and the increased amounts at night were not fully reverted 2 d after rewatering (Fig. 4). Well-known compatible solutes in C_3 plants (Bohnert et al,, 1995; Hare et al., 1998), Pro and the sugar alcohol

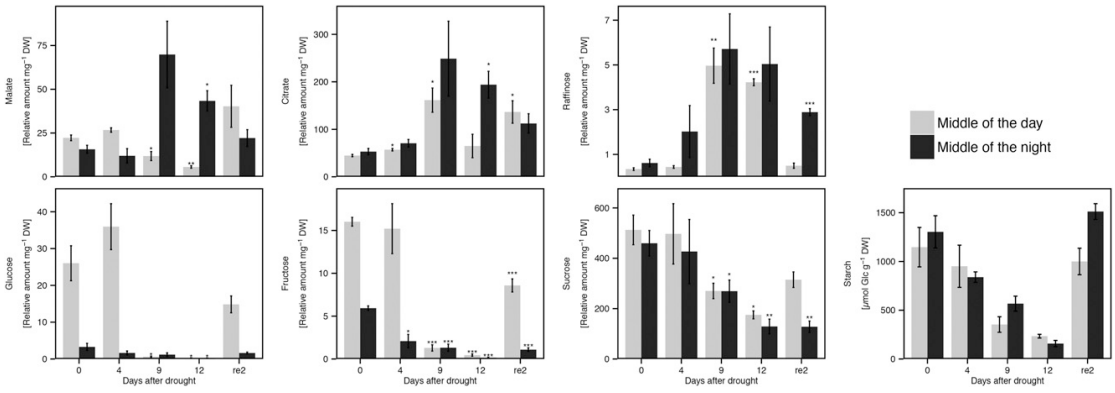


Figure 4. Levels of organic acids, soluble sugars and starch in *T. triangulare* in response to varying levels of water availability. Except for starch, all metabolites were measured by GC-MS and normalized to dry weight (DW) and internal ribitol standard (mean ± SE, $n = 3–4$, asterisks indicate Student's *t* test significance in comparison to day 0 at $^{***}P < 0.001$, $^{**}P < 0.01$, and $^{*}P < 0.05$). Starch was normalized to dry weight (mean ± SE, $n = 2–4$). re2, 2 d after rewatering.

mannitol, were only enriched in individual plants of the biological replicates (Supplemental Fig. S5).

Mapman and *K*-Means Clustering Indicate Multiple Layers of Response and Regulation

In order to understand alterations of mRNA amounts beyond the changes in the CAM-related genes described earlier (Fig. 3), three independent analyses were used to test general changes at the metabolic pathway and single gene level: (1) Mapman-based analysis using all values followed by Wilcoxon Rank Sum test (see "Materials and Methods") for enrichment (Thimm et al., 2004; Supplemental Fig. S6), (2) *k*-means clustering of the significantly changed genes followed by gene ontology (GO) term enrichment analysis (Fig. 5), and (3) manual inspection of the 50 genes with the highest fold changes on day 12 (Tables I and II). In Mapman (Supplemental Fig. S6), the metabolism overview again indicated little to no gene expression changes on day 4 of drought and rewatered plants compared to well-watered samples and massive changes during days 9 and 12 of water-limitation (Supplemental Fig. S6). The changes included downregulation of genes involved in cell wall metabolism (Supplemental Dataset S4). Genes of cell wall proteins of all classes were markedly downregulated ($q < 10^{-6}$ at midday), as were the biosynthesis genes for cell wall polymers (cellulose: $q < 10^{-08}$; modifiers: $q < 10^{-10}$ at midday; pectin esterases: $q < 10^{-11}$; pectate lyase: $q < 10^{-11}$, precursor synthesis: $q < 10^{-7}$; hemicellulose: $q < 10^{-6}$ at midnight). Downregulation of photosynthetic genes of the light reactions was visible on day 9 but more pronounced on day 12 of drought ($q < 10^{-5}$). Consistent with the previous analysis (Fig. 3), starch turnover genes ($q < 10^{-4}$) and genes of glycolysis ($q < 10^{-3}$ on day 12) were strongly upregulated. Additional strong upregulation was limited to the genes in the raffinose synthesis pathway ($q < 10^{-2}$), the committed step of Pro biosynthesis, and to the myoinositol oxidases (Tables I and II).

K-means clustering grouped 6,800 significantly DEGs at midday compared to day 0 by transcriptional pattern into 12 clusters (Fig. 5). Six clusters contained 3,124 genes ascending (clusters 1–6, Fig. 5), and six clusters contained 3,676 genes descending with pronounced water-limitation (7–12, Fig. 5). The clusters with ascending genes included one cluster with gradual increase up to day 12 and only little recovery after 2 d of rewatering (cluster 1). This cluster was enriched in GO terms related to external stimulus, and response. clusters 2 and 3 show mostly stable expression between days 9 and 12 and the expression was fully (cluster 2) and partially (cluster 3) recovered upon 2 d of rewatering. The list of genes with fully restored abundance of cluster 2 is enriched in GO terms related to chloroplast organization, protein localization, and starch. The genes in cluster 3 recovered only partially to well-watered levels after rewatering, and this cluster showed significant enrichment of terms related to

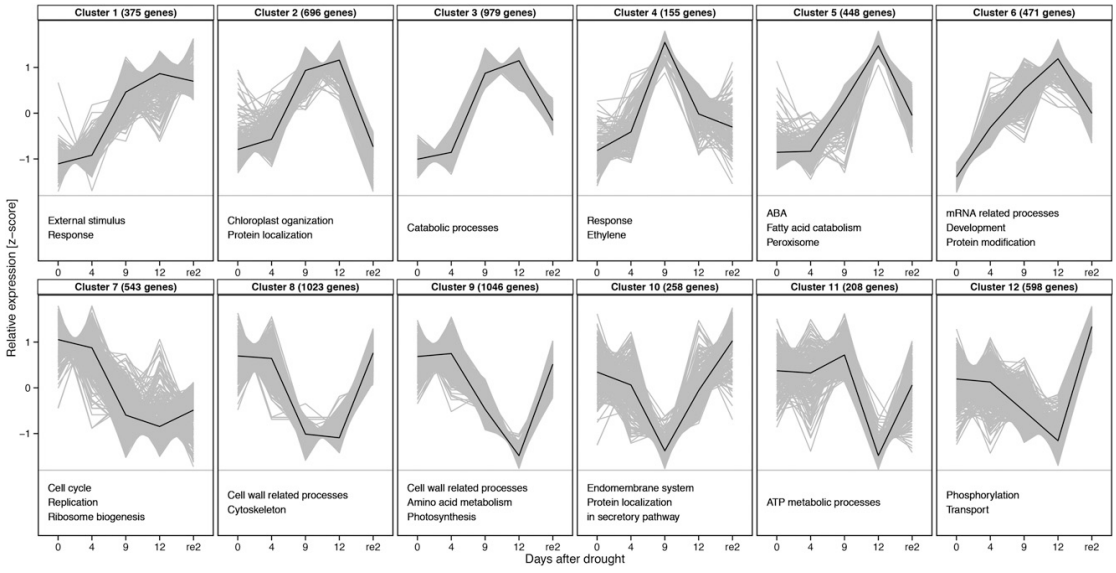


Figure 5. *K*-means clustering of relative gene expression and selected enriched GO terms. Genes that were found to be differentially expressed in the middle of the day between one of the water-limited stages and day 0 (DESeq2, $q < 0.01$) were used for the *k*-means approach (6800 genes in total). For the full list of enriched GO terms, see Supplemental Dataset S4. Gray line, Expression of single genes; black line, average of all genes in cluster; re2, 2 d after rewatering.

Table I. *Fifty most highly upregulated genes on day 12 after drought*

rpm, Reads per million ($n = 3$); MD, middle of the day; MN, middle of the night; $\log_2$-FC, $\log_2$-fold change in expression at given time point on day 12 compared to day 0; ns, not significantly DEG.

Locus	Annotation (TAIR10)	Day 0 MD	Day 0 MN	Day 12 MD	Day 12 MN	$\log_2$-FC (MD ; MN)
		rpm	*rpm*	*rpm*	*rpm*	
AT2G42600	Phospho*enol*pyruvate carboxylase 2	208.67	374.33	4,803.67	9,910.67	4.05 ; 4.7
AT1G53310	Phospho*enol*pyruvate carboxylase 1	312.00	613.33	6,656.33	15,510.33	3.91 ; 4.61
AT4G17030	Expansin-like B1	0.00	0.00	4.67	1.67	4.91 ; 2.55
AT4G26260	Myoinositol oxygenase 4	1.00	0.00	373.33	29.00	7.36 ; 5.02
AT1G14520	Myoinositol oxygenase 1	3.33	1.33	1,290.00	112.67	7.25 ; 5.42
AT2G19800	Myoinositol oxygenase 2	1.67	0.67	451.00	37.00	6.6 ; 4.83
AT1G55740	Seed imbibition 1	9.33	46.67	605.33	2,236.33	5.41 ; 5.38
AT3G08860	PYRIMIDINE 4	3.33	44.33	139.67	23.33	4.68 ; ns
AT2G38400	Ala:glyoxylate aminotransferase 3	2.67	30.00	103.67	17.67	4.54 ; ns
AT3G50980	Dehydrin xero 1	0.00	0.00	7.67	0.00	4.9 ; ns
AT3G22840	Chlorophyll A-B binding family protein	17.67	15.00	4,029.67	17.33	6.74 ; ns
AT5G15250	FTSH protease 6	6.67	2.67	764.33	3.00	6.18 ; ns
AT5G20110	Dynein light chain type 1 family protein	3.67	23.67	368.67	12.33	6.05 ; ns
AT4G27360	Dynein light chain type 1 family protein	19.67	32.33	654.33	54.33	4.62 ; 0.83
AT3G29410	Terpenoid cyclases/Protein prenyltransferases superfamily protein	0.00	0.00	4.67	17.67	4.86 ; 5.96
AT5G16020	Gamete-expressed 3	0.33	1.33	41.00	3.67	5.75 ; 1.59
AT1G80920	Chaperone DnaJ-domain superfamily protein	0.33	12.33	23.33	578.33	4.6 ; 5.46
AT3G22740	Homo-Cys S-methyltransferase 3	2.33	23.33	118.33	145.33	5.08 ; 2.68
AT1G10060	Branched-chain amino acid transaminase 1	1.00	4.67	31.33	11.00	4.63 ; ns
AT1G29900	Carbamoyl phosphate synthetase B	69.67	80.67	309.67	1,955.67	1.75 ; 4.56
AT1G12740	Cytochrome P450, family 87, subfamily A, polypeptide 2	4.00	86.33	290.67	10.67	4.47 ; -2.13
AT1G52240	RHO guanyl-nucleotide exchange factor 11	6.67	113.33	1,652.00	31.00	7.23 ; -1.69
AT1G28480	Thioredoxin superfamily protein	1.00	0.67	2.00	37.00	ns ; 5.76
AT1G18100	PEBP (phosphatidylethanolamine-binding protein) family protein	1.00	1.00	34.67	59.00	4.52 ; 5.57
AT1G03790	SOMNUS	0.00	0.67	19.67	4.33	5.15 ; 2.38
AT1G54070	Dormancy/auxin associated family protein	0.00	1.67	2.33	75.33	2.5 ; 5.01
AT1G80390	Indole-3-acetic acid inducible 15	0.00	0.00	5.67	5.33	4.73 ; 5
AT5G12840	Nuclear factor Y, subunit A1	0.67	2.67	25.67	31.00	4.75 ; 3.56
AT1G15330	Cystathionine beta-synthase (CBS) protein	0.00	0.00	1.33	5.00	3.54 ; 4.74
AT3G21700	Ras-related small GTP-binding family protein	0.67	3.00	17.33	82.33	3.56 ; 4.61
AT1G21000	PLATZ transcription factor family protein	0.33	25.33	18.33	135.00	4.48 ; 2.45
AT3G63060	EID1-like 3	0.00	0.00	22.00	13.33	5.88 ; 5.32
AT3G48530	SNF1-related protein kinase regulatory subunit gamma 1	0.67	31.00	39.67	275.00	5.05 ; 3.1
AT5G56550	Oxidative stress 3	1.33	4.33	18.33	310.33	3.03 ; 5.71
AT5G47560	Tonoplast dicarboxylate transporter	0.00	0.00	5.33	10.67	4.99 ; 5.93
AT1G32450	Nitrate transporter 1.5	4.33	68.00	233.67	34.33	5.04 ; ns
AT4G21680	NITRATE TRANSPORTER 1.8	2.00	20.33	101.33	16.67	5.02 ; ns
AT3G27250	Unknown	0.00	1.00	85.67	58.33	7.08 ; 5.63
AT1G52720	Unknown	0.67	26.67	156.33	80.33	6.72 ; 1.63
AT4G19390	Uncharacterized protein family (UPF0114)	6.33	49.33	1,234.33	276.00	6.58 ; 2.35
AT3G03170	Unknown	0.00	0.67	4.33	66.00	3.56 ; 6.21
AT3G48510	Unknown	2.67	1.33	78.33	106.00	4.16 ; 6.08
AT4G26288	Unknown	0.00	0.33	2.00	32.67	3.02 ; 5.99
AT5G40790	Unknown	1.67	0.33	38.33	48.00	3.96 ; 5.85
AT1G15380	Lactoylglutathione lyase/glyoxalase I family protein	4.67	9.67	337.00	109.33	5.65 ; 3.42
AT5G50360	Unknown	1.00	0.00	3.67	11.67	ns ; 5.3
AT2G44670	Protein of unknown function (DUF581)	0.33	2.33	3.33	65.67	1.92 ; 4.81
AT5G02020	Unknown	1.67	3.33	7.33	96.67	1.74 ; 4.57
AT1G27461	Unknown	0.00	0.00	0.00	3.67	ns ; 4.51
AT2G28780	Unknown	0.33	0.00	16.00	0.00	4.49 ; ns

Table II. *Fifty most highly downregulated genes on day 12 after drought*

rpm, Reads per million ($n = 3$); MD, middle of the day; MN, middle of the night; $\log_2$-FC, $\log_2$-fold change in expression at given time point on day 12 compared to day 0; ns, not significantly DEG.

Locus	Annotation (TAIR10)	Day 0 MD	Day 0 MN	Day 12 MD	Day 12 MN	$\log_2$-FC (MD ; MN)
		rpm	*rpm*	*rpm*	*rpm*	
AT5G20630	Germin 3	2,449.67	648.33	2.00	1.67	−9.13 ; −7.24
AT1G72610	Germin-like protein 1	932.67	274.67	1.00	1.00	−8.55 ; −6.45
AT1G64390	Glycosyl hydrolase 9C2	177.00	80.00	0.67	1.67	−7.45 ; −5.27
AT4G21960	Peroxidase superfamily protein	1,529.33	8,040.00	10.67	349.67	−7.09 ; −4.22
AT4G11050	Glycosyl hydrolase 9C3	70.67	31.67	0.00	0.33	−7.06 ; −5.32
AT5G22740	Cellulose synthase-like A02	272.67	207.00	3.33	4.67	−6.42 ; −5.27
AT2G32990	Glycosyl hydrolase 9B8	42.00	34.33	0.33	0.00	−6.36 ; −6.19
AT2G04780	FASCICLIN-like arabinoogalactan 7	15.00	24.33	0.00	0.00	−6.12 ; −6.27
AT4G37450	Arabinogalactan protein 18	10.33	22.67	0.00	0.00	−5.64 ; −6.24
AT2G35860	FASCICLIN-like arabinogalactan protein 16 precursor	94.33	125.67	2.67	1.33	−5.11 ; −6.12
AT5G03760	Nucleotide-diphospho-sugar transferases superfamily protein	61.00	64.33	1.00	0.67	−5.75 ; −6.11
AT2G37130	Peroxidase superfamily protein	21.00	7.67	0.00	0.00	−6.03 ; −4.26
AT3G11700	FASCICLIN-like arabinogalactan protein 18 precursor	129.33	195.67	4.67	2.67	−5.04 ; −5.92
AT1G02335	germin-like protein subfamily 2 member 2 precursor	9.33	16.33	0.00	0.00	−5.47 ; −5.91
AT1G67750	Pectate lyase family protein	15.00	53.33	0.00	0.67	−5.82 ; −5.5
AT4G13410	Nucleotide-diphospho-sugar transferases superfamily protein	27.67	23.00	0.67	0.00	−5.65 ; −5.81
AT3G28150	TRICHOME BIREFRINGENCE-LIKE 22	35.67	12.00	0.33	0.00	−5.79 ; −5.33
AT3G53190	Pectin lyase-like superfamily protein	14.67	23.33	0.00	1.00	−5.79 ; −4.07
AT3G62020	Germin-like protein 10	8.33	14.33	0.00	0.00	−5.66 ; −5.65
AT1G12090	Extensin-like protein	151.33	29.33	3.33	1.33	−5.65 ; −4.22
AT1G41830	SKU5-similar 6	19.00	47.33	0.33	0.33	−5.03 ; −5.65
AT5G46890	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	28.67	5.67	0.00	0.00	−6.42 ; −3.84
AT3G04290	Li-tolerant lipase 1	235.00	817.67	1.00	13.00	−6.39 ; −4.91
AT1G76160	SKU5 similar 5	37.33	89.33	0.33	1.00	−6.13 ; −6.34
AT5G33370	GDSL-like Lipase/Acylhydrolase superfamily protein	162.00	675.00	1.00	6.33	−6.17 ; −5.57
AT5G23940	HXXXD-type acyl-transferase family protein	95.00	148.33	1.33	11.67	−6.13 ; −3.47
AT3G16370	GDSL-like Lipase/Acylhydrolase superfamily protein	298.00	307.67	5.33	4.67	−5.98 ; −5.78
AT3G15850	Fatty acid desaturase 5	20.00	9.00	0.00	2.00	−5.91 ; −1.93
AT1G21850	SKU5 similar 8	17.33	35.33	0.33	0.67	−5.17 ; −5.73
AT4G28780	GDSL-like Lipase/Acylhydrolase superfamily protein	111.00	569.33	0.33	4.33	−5.63 ; −5.47
AT1G63710	Cytochrome P450, family 86, subfamily A, polypeptide 7	45.00	193.33	1.00	1.00	−5.27 ; −6.59
AT3G10185	Gibberellin-regulated family protein	177.33	292.33	1.00	4.33	−6.46 ; −4.94
AT1G12570	Glc-methanol-choline (GMC) oxidoreductase family protein	88.67	59.33	1.00	2.67	−6.09 ; −4.2
AT2G45970	Cytochrome P450, family 86, subfamily A, polypeptide 8	34.33	140.33	1.00	1.33	−5.3 ; −6.07
AT1G74670	Gibberellin-regulated family protein	19.67	28.00	0.00	0.33	−6.03 ; −5.03
AT1G61720	NAD(P)-binding Rossmann-fold superfamily protein	6.00	89.67	0.00	0.67	−3.32 ; −5.82
AT3G52500	Eukaryotic aspartyl protease family protein	43.33	35.67	1.33	0.33	−5.01 ; −5.75
AT5G44635	Minichromosome maintenance (MCM2/3/5) family protein	37.67	6.67	0.00	0.00	−6.6 ; −4.12
AT5G67100	DNA-directed DNA polymerases	21.67	9.67	0.33	1.00	−5.93 ; −3.47
AT5G46280	Minichromosome maintenance (MCM2/3/5) family protein	26.67	5.33	0.00	0.00	−5.91 ; −3.97
AT1G44900	minichromosome maintenance (MCM2/3/5) family protein	66.00	12.00	0.67	0.33	−5.87 ; −3.62
AT2G07690	Minichromosome maintenance (MCM2/3/5) family protein	73.67	22.67	1.33	5.67	−5.74 ; −1.74
AT4G02060	Minichromosome maintenance (MCM2/3/5) family protein	38.00	6.33	0.67	0.00	−5.73 ; −3.52
AT1G27040	Major facilitator superfamily protein	44.33	2.67	0.00	2.00	−6.87 ; ns
AT5G62730	Major facilitator superfamily protein	40.00	3.00	0.00	3.00	−6.49 ; ns
AT3G02500	Unknown	28.33	4.00	0.33	0.00	−5.89 ; −3.27
AT1G27930	Protein of unknown function (DUF579)	8.67	19.00	0.00	0.00	−4.65 ; −5.79
AT2G21100	Disease resistance-responsive (dirigent-like protein) family protein	16.33	49.67	0.67	0.33	−4.04 ; −5.66
AT4G17340	Tonoplast intrinsic protein 2;2	321.33	125.00	3.67	31.00	−5.74 ; ns
AT1G13260	Related to ABI3/VP1 1	18.33	5.33	0.00	0.33	−5.86 ; −3.45

catabolic processes. Cluster 4 was only transiently upregulated on day 9 of water withholding but dropped back to background levels on day 12. It was enriched with GO terms relating to responses to signals, which include ethylene and mechanical stimulus. The expression of genes in clusters 5 and 6 peaked on day 12. Genes of cluster 5 only responded after 9 d of drought, increased sharply toward day 12, and mostly recovered after 2 d of rewatering and were enriched in GO terms related to ABA, fatty acid catabolism, and peroxisomes. The expression of genes in cluster 6 responded already after 4 d of drought, was only partially reverted and was enriched for mRNA related processes, development, and protein modification.

Similar to cluster 1, where genes were upregulated and stayed mostly up despite rewatering, cluster 7 contained genes that dropped in expression during days 9 and 12 of drought and barely recovered. Cluster 7 was enriched in genes related to growth, i.e. in GO terms referring to DNA replication, cell cycle, and ribosome genesis. Genes of cluster 8 were similarly downregulated during CAM conditions but fully recovered upon rewatering. This cluster is enriched in GO terms related to cell wall processes and cytoskeleton. Genes in cluster 9 only responded on day 9, further increased on day 12 of drought, and fully recovered after rewatering. This cluster was enriched in GO terms related to cell wall related processes, amino acid metabolism, and photosynthesis. Cluster 10 was the mirror cluster of cluster 4, which peaked on day 9, and included genes with peak downregulation on day 9. It was enriched in genes related to protein localization in the secretory pathway and the endomembrane system, including the protein population of the plasma membrane. Downregulation of genes in clusters 11 and 12 peaked on day 12. Genes in cluster 11 responded only after 12 d of drought and fully recovered after rewatering. This cluster was enriched in processes related to nucleotide metabolic processes. Expression of genes in cluster 12 began decreasing on day 4, continued to decrease until day 12, and after rewatering, were increased above well-watered levels. This overshoot cluster was enriched in genes related to posttranslational modification in particular phosphorylation and transport. Analogous *k*-means clustering of 7,563 significant DEGs at midnight revealed large overlap of major changes in comparison to day 0 at midday (Supplemental Fig. S7; Supplemental Dataset S5). Taken together, the clusters with elevated expression during drought mostly contained genes enriched in GO terms related to catabolism, RNA metabolism and also related to signaling. The clusters with reduced gene expression contained genes related to growth from photosynthesis over amino acid synthesis and cell wall processes to DNA replication and cell cycle.

To highlight the results of the analysis at the pathway level, the top 50 of significantly induced (Table I) and repressed (Table II) genes on day 12 of water limitation were analyzed. Of the fifty genes with the highest upregulation on day 12 in *T. triangulare*, 22 were upregulated by ABA in Arabidopsis (Table I). The CAM gene *PPC* was among the top fifty upregulated genes. Six genes were related to production of compatible solutes or their precursors including raffinose, beta-Ala, and sugar alcohols. One gene encoding a LATE EMBRYOGENESIS ABUNDANT (LEA) protein was among the top 50 induced genes as well as a heat inducible chaperone and the oxidative stress responsive OXIDATIVE STRESS3 (OXS3). Light protection was seemingly strengthened as evidenced by EARLY LIGHT INDUCIBLE1 (ELIP1), FTSH PROTEASE6 (FTSH6), and two dyneins, which may be involved in organelle repositioning (Heddad and Adamska, 2000; Hutin et al., 2003; Zelisko et al., 2005). Proteins encoded by 11 of the 50 most induced genes were involved in regulation: the SNF1 RELATED KINASE REGULATORY SUBUNIT GAMMA1 (KING1), the F-box protein EID1-LIKE3 (EDL3), RHO GUANYL-NUCLEOTIDE EXCHANGE FACTOR11, and four transcription factors including SOMNUS and a PLATZ transcription factor. There were also three transport proteins affected.

The fifty genes most reduced in expression on day 12 included 21 genes involved in cell wall synthesis, nine genes related to lipid metabolism and the vacuolar aquaporin TONOPLAST INTRINSIC PROTEIN2;2 (Table II). Downregulation of DNA replication via downregulation of the genes (*MINICHROMOSOME MAINTENANCE2-3* and *5-7*) encoding five of six subunits of the helicase and downregulation of the gene encoding DNA-directed DNA polymerase INCURVATA2 was also evident among the fifty most downregulated genes, indicating cell-cycle arrest.

Mediators of the Transcriptional Changes

K-means clustering identified a multilayer response with genes only transiently regulated on day 9 of drought and with genes with sustained change on both day 9 and day 12 of drought. To identify the candidate transcriptional regulators mediating the responses, genes encoding for putative transcription factors were extracted from PlnTFDB (Pérez-Rodríguez et al., 2010) and tested for differential expression on either day 9 or day 12 of water limitation relative to the well-watered state (day 0) and for cluster membership. Of 1,449 identified transcription factor genes, 582 (40%) were differentially regulated on either day 9 or day 12 or both compared to well-watered plants (Supplemental Dataset S6). Of the 582, a subset of 19 transcription factor genes belonged to cluster 4, the genes, which were only transiently upregulated on day 9 of drought (Table III). These included eight TFs of the APETALA2 AND ETHYLENE-RESPONSIVE ELEMENT BINDING PROTEINS (AP2-EREBP) class including C-REPEAT/DRE BINDING FACTORS (CBF2 and CBF3), which are known to be stress induced (Zou et al., 2011), REDOX RESPONSIVE TRANSCRIPTION FACTOR1 (RRTF1) and ETHYLENE RESPONSE FACTOR8 (ERF8). Both CBFs were shown to bind the drought responsive

Table III. *Transcription factors of k-means cluster midday 4*

TF Family, Transcription factor family based on Pérez-Rodríguez et al. (2010).

Locus	Annotation (TAIR10)	TF Family
AT1G12610	Integrase-type DNA-binding superfamily protein	AP2-EREBP
AT1G19210	Integrase-type DNA-binding superfamily protein	AP2-EREBP
AT1G33760	Integrase-type DNA-binding superfamily protein	AP2-EREBP
AT1G53170	Ethylene response factor 8	AP2-EREBP
AT4G25470	C-repeat/DRE binding factor 2	AP2-EREBP
AT4G25480	Dehydration response element B1A	AP2-EREBP
AT4G34410	Redox responsive transcription factor 1	AP2-EREBP
AT5G21960	Integrase-type DNA-binding superfamily protein	AP2-EREBP
AT3G47640	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	bHLH
AT2G21230	Basic-Leu zipper (bZIP) transcription factor family protein	bZIP
AT1G27730	Salt tolerance zinc finger	C2H2
AT2G40140	SALT-INDUCIBLE ZINC FINGER 2	C3H
AT4G17230	SCARECROW-like 13	GRAS
AT1G18710	Myb domain protein 47	MYB
AT3G13540	Myb domain protein 5	MYB
AT1G01380	Homeodomain-like superfamily protein	MYB-related
AT1G33060	NAC 014	NAC
AT3G49530	NAC domain containing protein 62	NAC
AT4G35580	NAC transcription factor-like 9	NAC

element DNA sequence (Liu et al., 1998). Two zinc finger transcription factors, SALT TOLERANCE ZINC FINGER (STZ) and SALT-INDUCIBLE ZINC FINGER2 (SFZ2), belonged to cluster 4. STZ was shown to function as a transcriptional repressor in response to drought and ABA (Sakamoto et al., 2000, 2004), while SFZ2 is induced by salt stress (Sun et al., 2007). In addition, three genes encoding NO APICAL MERISTEM (NAC) domain proteins, one GRAS, one bZIP, and three MYB-domain proteins made up the transcriptional part of cluster 4. The known functions of transcription factors in the transient group indicated that *T. triangulare* underwent a transient general stress response commonly observed in plants mediated primarily by the drought responsive element binding transcription factors.

Of the remaining 563 genes encoding transcription factor candidates, the top 25 upregulated were analyzed for known functions (Table IV). The ABA-responsive proliferation inhibitor SOMNUS (SOM) was the most highly upregulated transcription factor on day 12. The top 25 also included four genes encoding transcription factors known to be involved in ABA signaling, NUCLEAR FACTOR Y, SUBUNIT A1 (NF-YA1), NF-YA9, and HOMEOBOX7 (HB7) and one NAC-like transcription factor, NAC-LIKE, ACTIVATED BY AP3/PI (NAP), which acts upstream of ABA biosynthesis and promotes chlorophyll degradation (Yang et al., 2014). In addition, two transcription factors of the Orphans family, ETHYLENE RESPONSE2 (ETR2) and ETHYLENE INSENSITIVE4 (EIN4), which were shown to antagonistically control seed germination under salt stress (Wilson et al., 2014) as well as three heat shock factor family proteins were among the 25 most upregulated transcription factors. The sustained response of *T. triangulare* included ABA responsive transcription factors and growth associated regulators, while the transient response included the DRE responsive TFs of the CBF family.

To test if phytohormones other than ABA are also involved in the drought response, the overlap between genes significantly changed on day 12 at midday and genes specifically altered by the application of different hormones in Arabidopsis (Goda et al., 2008) was determined and statistically evaluated. Among the genes shown to be upregulated during hormone treatment of Arabidopsis with ABA, auxin, brassinosteroids, cytokinine, and ethylene, only the ABA-upregulated genes were enriched among the *T. triangulare* drought-inducible genes (Supplemental Fig. S8). A total of 434 ABA-regulated genes of Arabidopsis were drought-induced genes in *T. triangulare* and, accordingly, 541 ABA-regulated genes of Arabidopsis were repressed by drought in *T. triangulare.* The ABA-, auxin-, brassinosteroid-, and ethylene-downregulated genes were enriched among the *T. triangulare* drought-repressed genes. However, only few genes were shown to be repressed by auxin, brassinosteroids, and ethylene (90, 11, and 23, respectively). Thus, among the phytohormones, ABA made the major contribution to the control of drought-controlled genes and likely controls at least one-fifth of the genes differentially regulated under drought conditions in *T. triangulare*.

To visualize the ABA contribution to changes in general metabolism, the Mapman map of ABA induced changes in Arabidopsis was compared with those occurring in *T. triangulare* (Supplemental Fig. S9). The map of ABA-responsive genes replicates the induction of raffinose synthesis genes, and a mild reduction of photosynthesis genes but failed to result in the major changes in cell wall synthesis and changed photorespiration, sulfur metabolism, and secondary metabolism. The ABA signal detected in Arabidopsis at the level of metabolic gene expression is not fully congruent with the signals detected in *T. triangulare*, indicating additional layers of regulation in *T. triangulare*.

DISCUSSION

Facultative CAM

During CAM induction in *T. triangulare*, several key components of the CAM-cycle sensu stricto were transcriptionally upregulated such as genes encoding

Table IV. *Twenty-five most highly upregulated transcription factors on day 12 after drought*

TF Family, Transcription factor family based on Pérez-Rodríguez et al. (2010); rpm, reads per million ($n = 3$); MD, middle of the day; MN, middle of the night; log_2-FC, log_2-fold change in expression at given time point on day 12 compared to day 0; ns, not significantly DEG.

Locus	Annotation (TAIR10)	TF Family	Day 0 MD	Day 0 MN	Day 12 MD	Day 12 MN	Log_2-FC (MD ; MN)
			rpm	*rpm*	*rpm*	*rpm*	
AT1G03790	SOMNUS	C3H	0.00	0.67	19.67	4.33	5.15 ; 2.38
AT5G12840	Nuclear factor Y, subunit A1	CCAAT	0.67	2.67	25.67	31.00	4.75 ; 3.56
AT1G21000	PLATZ transcription factor family protein	PLATZ	0.33	25.33	18.33	135.00	4.48 ; 2.45
AT3G54320	Integrase-type DNA-binding superfamily protein	AP2-EREBP	1.00	1.33	3.33	27.00	1.4 ; 4.41
AT5G67480	BTB and TAZ domain protein 4	TAZ	19.67	71.00	461.33	129.00	4.15 ; 0.95
AT2G46680	Homeobox 7	HB	7.67	38.00	165.33	445.00	3.92 ; 3.56
AT1G69490	NAC-like, activated by AP3/PI	NAC	8.67	76.00	206.33	366.00	3.73 ; 2.09
AT1G52880	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein	NAC	0.00	2.67	5.00	10.67	3.51 ; 1.8
AT3G02550	LOB domain-containing protein 41	LOB	0.00	0.00	0.00	1.33	ns ; 3.3
AT1G17590	Nuclear factor Y, subunit A8	CCAAT	3.67	9.33	41.67	63.00	3.19 ; 2.75
AT3G23150	ETHYLENE RESPONSE 2	Orphans	12.33	10.67	86.00	94.00	2.47 ; 3.13
AT3G04580	ETHYLENE INSENSITIVE 4	Orphans	14.33	9.67	94.00	82.33	2.38 ; 3.13
AT2G26150	Heat shock transcription factor A2	HSF	4.67	6.00	9.00	47.00	ns ; 3.06
AT1G80840	WRKY DNA-binding protein 40	WRKY	2.33	0.67	0.33	6.67	−2.38 ; 3.04
AT4G17900	PLATZ transcription factor family protein	PLATZ	8.33	38.67	91.67	55.33	3.01 ; ns
AT3G11580	AP2/B3-like transcriptional factor family protein	ABI3VP1	0.33	1.33	4.33	1.67	3 ; ns
AT4G13980	Winged-helix DNA-binding transcription factor family protein	HSF	2.67	9.00	4.33	68.67	ns ; 2.98
AT3G24520	Heat shock transcription factor C1	HSF	2.67	11.33	5.00	88.00	ns ; 2.96
AT3G20910	Nuclear factor Y, subunit A9	CCAAT	11.00	25.67	72.67	181.00	2.32 ; 2.87
AT4G02640	bZIP transcription factor family protein	bZIP	1.67	6.67	0.67	48.33	ns ; 2.83
AT2G25900	Zinc finger C-x8-C-x5-C-x3-H type family protein	C3H	12.67	12.67	119.00	45.33	2.83 ; 1.91
AT3G15500	NAC domain containing protein 3	NAC	0.67	8.00	8.33	21.00	2.82 ; 1.39
AT1G32700	PLATZ transcription factor family protein	PLATZ	16.00	56.00	141.33	75.67	2.78 ; ns
AT4G39070	B-box zinc finger family protein	Orphans	1.00	3.00	1.67	22.00	ns ; 2.77
AT5G28770	bZIP transcription factor family protein	bZIP	2.67	11.00	2.00	74.00	ns ; 2.77

PEPC, NADP-ME, NAD-ME, and PPDK (Fig. 3). Other components, including transcripts of MDH and BCA3, remained unchanged but were constitutively highly expressed already prior to CAM induction. This is consistent with enzyme activity data for MDH (Holtum and Winter, 1982) and CA (Tsuzuki et al., 1982) in *M. crystallinum* operating in the C_3 and CAM modes, that were already very high in the C_3 state. Nonetheless, Cushman et al. (2008b) found additional upregulation of these enzymes during salt-induced CAM in *M. crystallinum*. In C_4 photosynthesis, activities of the corresponding carboxylation and decarboxylation enzymes required for CAM are not temporally but spatially separated into different cell types (Sage, 2003). In C_4 species of the genera *Flaveria* and *Cleome*, NAD-MDH is also not transcriptionally elevated relative to C_3 sister species (Bräutigam et al., 2011; Gowik et al., 2011) and CA only in some, but not in all, cases (Bräutigam et al., 2014).

In contrast to the canonical scheme of the CAM-cycle with NADP-ME localized in the cytosol (Holtum et al. 2005 and references therein), we found two isoforms of NADP-ME to be upregulated at both midday and midnight during CAM in the cross species mapping (Fig. 3; Supplemental Dataset S3). The contig analysis indicates that indeed two contigs are highly expressed; however, both possess a target peptide (Emanuelsson et al., 2000) and cluster with the plastidial AtNADP-ME4 with high bootstrap support. (Supplemental Fig. S3). The putative plastidic malate importers of the DiT family are not upregulated (Supplemental Dataset S3). NAD-ME was also upregulated, albeit at lower absolute levels but concomitant with the required malate importer DIC (Supplemental Dataset S3). Taken together, the data suggest that multiple routes for decarboxylation are possible. Unlike in the transcriptomic analyses of C_4 species (Bräutigam et al., 2011; Gowik et al., 2011), the transcriptomic evidence favors neither pathway in *T. triangulare*. It cannot be excluded that the low levels of CAM photosynthesis detected in *T. triangulare* (Winter and Holtum, 2014) are supported by multiple decarboxylation routes. Ideally our findings on the transcript levels of NAD(P)-ME would be confirmed by enzyme activity data, but so far all biochemical assays to measure activity in leaf material failed due to a suppression effect derived from an unknown constitute of the *T. triangulare* plant extract (Supplemental Fig. S10). Experimental differentiation

of the biochemical pathway of malate decarboxylation in the light during CAM in *T. triangulare* will be part of future work.

PPDK was strongly upregulated during CAM-induction in *T. triangulare* (Fig. 3). In *M. crystallinum*, the plastidic location of PPDK was demonstrated by measurements of enzyme acvitity combined with studies of pyruvate metabolism of intact isolated chloroplasts (Winter et al., 1982; Holtum et al., 2005). However, immunogold labeling of PPDK in the CAM species *Kalanchoë blossfeldiana* (Kondo et al., 2001) also raises the possibility of cytosolic PEP regeneration. In support of this hypothesis, the contig Tt26901, which encodes PPDK with a truncated N terminus, showed markedly higher transcript amounts on days 9 and 12 compared to day 0 and compared to the longer contig, which encodes for a putative target peptide. The contigs possibly represent splicing variants of the same gene and encode for PPDKs dually targeted to the cytosol (Tt26901, truncated N terminus) or the chloroplast (Tt24575) as found in Arabidopsis (Parsley and Hibberd, 2006). In C_4 species (NADP-ME and NAD-ME type), where the regeneration of PEP is invariably localized in the plastids, the transporters for pyruvate import (i.e. BASS2 and NHD1) and PEP export (i.e. PPT) are increased in expression compared to C_3 sister plants (Bräutigam et al., 2008, 2014; Furumoto et al., 2011). Interestingly, while BASS2 was upregulated but lowly expressed at midnight, transcripts encoding transporters of the PPT family were found to be downregulated during the C_3-CAM switch in *T. triangulare* (Supplemental Dataset S3). This contrasts with *M. crystallinum* where transcriptional induction of PPT was detected (Kore-eda et al., 2005). Transcripts matching the NHD1 gene were not detected. The absence of a target peptide in the major contig encoding for PPDK suggests that concomitant with the missing NHD1 and downregulated PPT, *T. triangulare* employs cytosolic PPDK. In addition to PPDK, both isoforms of PPDK regulatory protein, which promote circadian regulation of PPDK activity by dark-phased phosphorolytic inactivation (Dever et al., 2015), were slightly but significantly induced during CAM in *T. triangulare* (Supplemental Dataset S3).

During the evolution of C_4, the change of a single amino acid residue in the N-terminal region of PEPC, namely from Ala to Ser, was shown to increase the enzyme's kinetic efficiency compared to C_3 sister plants (Svensson et al., 1997; Bläsing et al., 2000; Svensson et al., 2003; Gowik et al., 2006). It was recently shown that in C_4 and CAM species within the *Caryophyllales*, to which the *Talinaceae* belong, the isogenes of a recurrently recruited gene lineage encoding PEPC1, namely *ppc1-E1c*, almost exclusively encodes the C_4-like Ser-780 (Christin et al., 2014). In the same study, in response to reduced irrigation, the expression of *ppc1-E1c* was upregulated at night in *Portulaca oleracea*, a C_4 species with facultative CAM (Christin et al., 2014). The two *PPC* contigs, Tt9871_4 and Tt9871_6, that exhibit the C_4-like Ser-780 (Supplemental Fig. S2A; Supplemental Table S3) indeed cluster together with *ppc1-E1c* isoforms of *Talinaceae* and *P. oleracea* (Supplemental Fig. S2B). However, highest transcript levels and the most drastic elevation in abundance during CAM induction was found for contig Tt63271, which encodes for the C_3-like Ala-780 and clusters in the *ppc1-E1b* clade. Low-level CAM photosynthesis apparently employs different PEPC genes compared to C_4 photosynthesis or fully expressed CAM photosynthesis. At night, PEPC is activated through phosphorylation via PEPC kinase (Carter et al., 1991), the activity of which is regulated transcriptionally through a circadian oscillator (Hartwell et al., 1996, 1999; Borland et al., 1999; Taybi et al., 2000). Phosphorylated PEPC is less sensitive to feedback inhibition by malate (Hartwell et al., 1996; Taybi et al., 2000, 2004). Consistent with the circadian regulation and strong upregulation of PEPC kinase at night during CAM induction in *M. crystallinum* (Hartwell et al., 1999; Cushman et al., 2008b; Dever et al., 2015), PEPC kinase was most markedly upregulated at midnight during CAM in *T. triangulare* (expression rank 300 on day 12, Supplemental Dataset S3), enabling increased PEPC activity at night. In contrast to C_4 plants, PEPC of CAM plants needs to be efficiently switched off during the light through feedback inhibition by malate exported from the vacuole to prevent futile cycling of CO_2 (Borland et al., 1999). All of the PEPC isoforms identified as upregulated in *Talinum* carry an Arg residue at position 890 (Supplemental Fig. S2; Supplemental Table S3; Paulus et al., 2013) and are thus likely feedback inhibited both by malate export during the day and once vacuolar storage capacity is exceeded during the night.

The upregulation of CAM-cycle genes was accompanied by elevated malate and citrate levels during the night and reduced malate levels during the day (Fig. 4) in agreement with increase in nocturnal acidity (Fig. 1). Photorespiration is not inactivated during the decarboxylation phase of CAM (Niewiadomska and Borland, 2007; Lüttge, 2011), and we observed a mostly stable expression of photorespiratory genes during the C_3-CAM transition (Supplemental Table S7). Reduced steady-state metabolite pools of glycerate and glycolate (Supplemental Fig. S4), as well as the reduced sugar levels, were likely a consequence of reduced overall photosynthetic rate. Nocturnal carbon assimilation rates during CAM were estimated to be reduced to approx. 5% of CO_2 fixation in the light in the C_3 mode (Winter and Holtum, 2014). Hence, the associated metabolite fluxes during CAM are expected to be low.

The carboxylation reaction during the night depends on degrading starch for PEP production. The degradation can be either phosphorolytic or hydrolytic (Häusler et al., 2000; Weise et al., 2011). Some, but not all, transcripts encoding these two alternate starch breakdown pathways were upregulated (Fig. 3). Phosphorolytic starch breakdown preserves some of the energy released during breakage of the glycosidic bonds in starch and therefore consumes less ATP for activation of the released Glc residues. ATP, ADP, and/

or free phosphate may balance hydrolytic and phosphorolytic starch breakdown depending on the ATP demand of the cytosol. Using starch and malate contents at the end of the night and the end of the day, the amount of starch turnover was calculated as sufficient to support the level of photosynthesis during CAM (Supplemental Table S8). Efficient starch turnover apparently required elevated expression of both starch synthesis and starch degradation genes, as well as sufficient branching of the starch to allow rapid degradation (Fig. 3). Based on strong upregulation of *GPT*, which was among the top 50 upregulated genes at midday, and the upregulation of *MEX1* (Fig. 3), glycolysis likely operated in the cytosol, which was confirmed by the transcriptional upregulation of all but one of the cytosolic glycolytic enzymes (Fig. 3). Taken together, inducible CAM required upregulation of key CAM-cycle genes, elevation of starch turnover by changes in selected genes and increased expression of cytosolic glycolytic genes. While starch turnover capacity is increased (Fig. 3), the amount of starch is reduced (Fig. 4), which may be explained by lowered photosynthetic capacity during CAM (Winter and Holtum, 2014). In contrast to *T. triangulare*, CAM can substantially contribute to carbon gain in *M. crystallinum* (Winter and Holtum, 2007), and following salinity-induced CAM, starch accumulation is increased and directly correlates with acidification (Haider et al., 2012).

The rapid induction of CAM-related gene expression and the coexpression of ABA-related signaling networks upon drought stress points to a direct connection to the ABA signaling network as shown for *Kalanchoë blossfeldiana* (Taybi et al., 1995) and *M. crystallinum* (Taybi and Cushman, 2002). Control of CAM induction may be exerted either directly by ABA-mediated phosphorylation and activation of transcription factors or by downstream effectors (Table IV).

Drought Response and Carbon Resource Utilization

Beyond CAM photosynthesis, additional processes contribute to the adaptation of *T. triangulare* to drought, such as a classical drought response and management of carbon resources. Water limitation leads to stomatal closure during the day (Herrera et al., 1991; Winter and Holtum, 2014), which reduces the photosynthetic carbon assimilation and hence pool sizes of soluble sugars such as Suc, Glc, and Fru during the day and the night (Fig. 4). Under severe drought stress, part of these remaining soluble sugars may act as compatible solutes, retain water in the cells and protect cellular structures against damage. In contrast to the overall loss of these sugars, substantial accumulation of raffinose was observed (Fig. 4) likely triggered by transcriptional upregulation of raffinose biosynthesis (Supplemental Fig. S6; Table I). During cold acclimation in Arabidopsis (Nägele and Heyer, 2013) and *Ajuga reptans* (Findling et al., 2015), nonaqueous fractionations localized raffinose to variable extents within the vacuole, cytosol and chloroplast stroma. At this point we do not know whether raffinose accumulating in drought-stressed *T. triangulare* functions as a compatible solute in the cytosol, supports osmotic pressure maintenance in the vacuole, or has protective functions for membranes and/or proteins within the plastid or cytosol (Nishizawa et al., 2008). Although the gene encoding PYRROLINE-5-CARBOXYLATE SYNTHASE (P5CS), which catalyzes the committed step for Pro synthesis, was upregulated during CAM (expression rank 57 on day 12 at midday, Supplemental Dataset S1), Pro accumulated only in some, but not all, plants tested (Supplemental Fig. S5). Similarly, levels of the sugar alcohols glycerol, mannitol, myoinositol, and sorbitol, which could potentially act as compatible solutes, were not significantly altered during drought, although mannitol accumulated in individual plants (Supplemental Figs. S5 and S11). The discrepancy between transcript and metabolite measurements is likely based on the fact that the transcriptome precedes the metabolome. Although genes were upregulated in all replicates (Supplemental Dataset S1), the metabolic status may be delayed.

Stomatal closure reduces the consumption of absorbed light energy in photosynthesis, potentially leading to short-term redox stress. Among the top 50 upregulated genes, four were involved in light stress, e.g. the gene of ELIP1 (Hutin et al., 2003), which is ABA induced in Arabidopsis (Table I). Genes encoding components of light stress-induced oxidation of fatty acids in the peroxisomes were enriched in cluster 3, representing induced genes that do not quite reach well-watered levels after rewatering (Fig. 5). Thus, despite leaf-angle changes and leaf rolling (Taisma and Herrera, 2003; Fig. 1) that reduce light exposure, a transcriptional upregulation of light-protection mechanisms is seemingly required to aid leaf functioning. Additionally, the increase of raffinose, shown to act as a scavenger of hydroxyl radicals in Arabidopsis (Nishizawa et al., 2008), might play a role in supporting leaf functioning.

Under drought, increased use of resources such as fatty acids and proteins was induced (Fig. 5), which might explain the slight increase in free amino acids (Noctor and Foyer, 2000; Supplemental Fig. S11). Catabolism of fatty acids through peroxisomal processes was induced by day 9 already, maintained on day 12, and only partially reversed upon rewatering. Genes involved in processes related to mRNA turnover were also increased during drought (Fig. 5, cluster 6). These induced catabolic processes were, however, not associated with permanent damage as the transcriptome returned to prestress status upon rewatering (Fig. 2). Most genes encoding anabolic processes such as photosynthesis and amino acid biosynthesis and transport were transcriptionally downregulated (Fig. 5; Supplemental Fig. S6), but returned to prestress levels after rewatering. For example, cell wall biosynthesis was downregulated and completely reversed upon rewatering (Fig. 5, clusters 8 and 9; Supplemental Fig. S6). Finally, genes involved in replication and the cell cycle were among the most downregulated genes overall on day 12 (Table II), and processes related to

replication and ribosome biogenesis were among those, which were downregulated upon drought but did not return to well-watered levels after rewatering (Fig. 5, cluster 7). Apparently, after 2 d, the expression levels of genes involved in primary production are increased however transcript levels of growth-related genes (e.g. genes involved in replication, translation, and cell division) have not yet completely recovered. This might be due to the fact that the full resource availability has not been restored within two days reflected by the still altered starch and metabolite pools (Fig. 2).

Regulation of the Coordinated Drought Response

Clustering differential gene expression indicated that the response of *T. triangulare* is multilayered since clusters respond and recover at different time points to different degrees (Fig. 3). Initially, drought stress induced a general stress response (clusters 4 and 9, Fig. 5) and a specific ABA-mediated response (cluster 5, Fig. 5; Supplemental Fig. S8). A conceptual model integrating the transcriptomic data is depicted in Figure 6. The general stress response included the CBFs, which are known to bind the drought responsive element (Zhu, 2002). However, their upregulation was transient and no longer detectable on day 12 of water limitation (Supplemental Dataset S6). During this initial stress response, our data indicate changes of the endomembrane system (cluster 9, Fig. 5). In contrast, the ABA-mediated response was further increased on day 12 (cluster 5, Fig. 5). The transcriptomic response also indicated a feed-forward loop with higher expression of

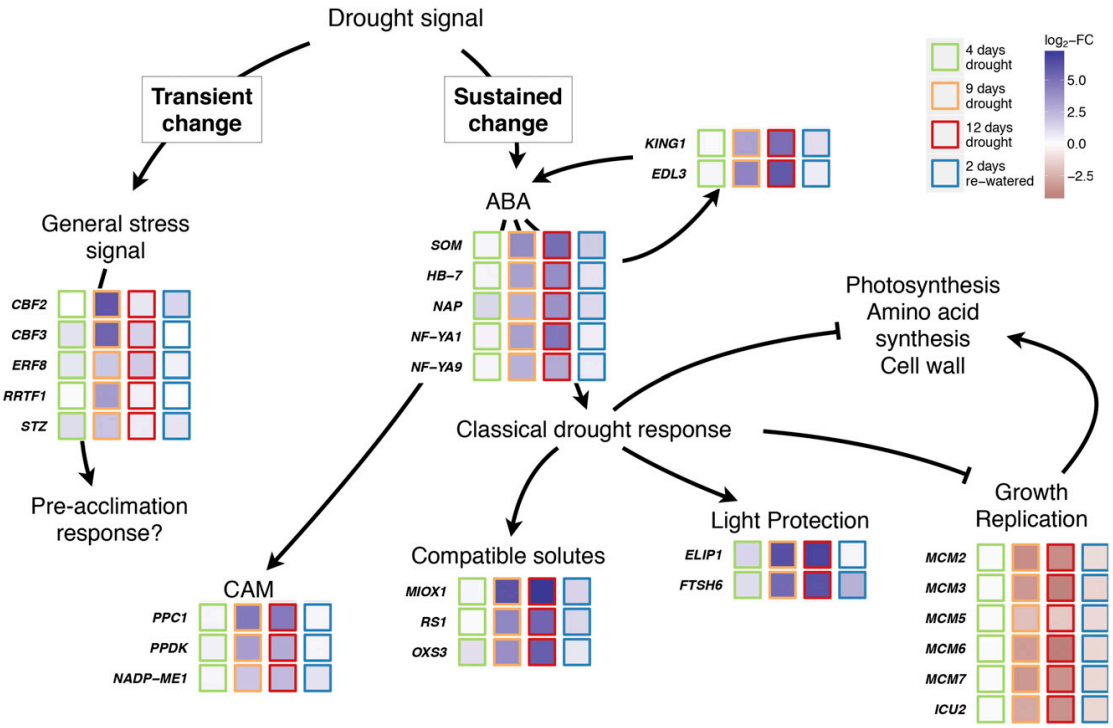


Figure 6. Proposed regulation of the coordinated drought response in *T. triangulare*. Shown are genes discussed in the text. A general stress signal was mediated by transcription factors, which were transiently upregulated on day 9 (*CBF2, C-repeat/DRE binding factor 2*; *CBF3, C-repeat/DRE binding factor 3*; *ERF8, ethylene response factor 8*; *RRTF1, redox responsive transcription factor 1*; *STZ, salt tolerance zinc finger*) and triggered a yet unclear response. The sustained response to persistent drought was primarily mediated via ABA, the activity of which was maintained through a feed-forward loop (*KING1, SNF1-related protein kinase regulatory subunit gamma 1*; *EDL3, EID1-like 3*), through ABA-responsive transcription factors (*SOM, SOMNUS*; *HB-7, homeobox 7*; *NAP, NAC-like, activated by AP3/PI*; *NF-YA1, nuclear factor Y subunit A1*; *NF-YA9, nuclear factor Y subunit A9*) and besides the induction of CAM (see Fig. 3 for details; *PPC1, phosphoenolpyruvate carboxylase 1*; *PPDK, Pyruvate phosphate dikinase, PEP/pyruvate binding domain*; *NADP-ME1, NADP-malic enzyme 1*) included a classical drought response, i.e. synthesis of compatible solutes (*RS1, Raffinose synthase 1*; *MIOX1, myoinositol oxygenase 1*; *OXS3, oxidative stress 3*) and light protection (*ELIP1, Early light inducible protein 1*; *FTSH6, FTSH protease 6*). At the same time, growth was repressed through downregulation of the replication cassette (*MCMs, minichromosome maintenance*; *ICU2, incurvata 2*) and processes related to photosynthesis, amino acid synthesis, and cell wall synthesis (see Fig. 5 and Supplemental Fig. S4) allowing *T. triangulare* to engage in a state of quiescence during drought. Log_2-FC, Absolute maximum log_2-fold changes in gene expression compared to day 0 at midday or midnight.

KING1, which activates ABA signaling, and EDL3, which mediates ABA signaling (Table I). At the same time, the expression of a PP2C was increased (expression rank 424 on day 12 at midnight, Supplemental Dataset S1), which is required to attenuate the signaling pathway once the stress signal is removed (Umezawa et al., 2009). Hence, *T. triangulare* plants are transcriptionally enabled to both sustain the ABA signal to the transcriptome and also to switch it off once the stress is removed upon rewatering. The ABA signal is likely overlaid with a carbon starvation signal. This additional signal may explain why several clusters did not recover to well-watered expression levels although soil moisture was restored. Catabolism-related processes only recovered partially (Fig. 5, clusters 3 and 5), and replication and ribosome biogenesis related genes recovered only marginally (Fig. 5, cluster 7).

The candidate transcription factors downstream of the initial signal, which is probably at least in part ABA mediated, are those that are significantly upregulated with a large magnitude. They include highly induced ABA-dependent transcription factors, such as HB7, NF-YA1 and 9, or NAP (Table IV), as well as the phosphorylation-controlled ABA-dependent transcription factors not detectable in the transcriptome analysis. Stalled growth, which is part of the response, may also at least in part be dependent on ABA as SOM is highly induced upon drought (Table I; Lim et al., 2013). SOM is required to control germination cessation in seeds downstream of the signal perception cascade (Kim et al., 2008) and thus likely controls effector genes required for growth. The trehalose-phosphate and SnRK1 mediated signaling (Baena-Gonzalez et al., 2007) does not significantly overlap with the signals observed in *T. triangulare.*

CONCLUSION

This study showed that the response of *T. triangulare* to drought stress was associated with profound and reversible changes in the transcriptome. Known key enzymes of the CAM-cycle and some of those connected to carbohydrate metabolism were upregulated, as were genes encoding functions associated with a classical drought response, including light protection and the synthesis of compatible solutes. Genes related to growth, photosynthesis, the synthesis of cell walls, and amino acids were downregulated in response to drought while catabolic processes were upregulated. A set of transcription factors likely to be involved in mediating these responses was identified. *T. triangulare* is easy to grow, flowers and produces seeds within few weeks after germination, and rapidly induces CAM in response to drought stress, thereby exhibiting important attributes of a potential model system to study key components of a drought response that includes CAM and its signaling components under stress. Research on *T. triangulare* may in the future contribute to engineering drought coping strategies such as CAM via synthetic biology into C_3 crop plants with the goal to improve their performance in water-limited environments (Yang et al., 2015).

MATERIALS AND METHODS

Plant Material and Growth Conditions

Talinum triangulare plants were grown in Miracle-Gro Potting Mix (Miracle-Gro) in "Short-One" treepots, 1.6 l (Stuewe and Sons). The experiment was initiated with 28-d-old plants in a controlled environment chamber (Environmental Growth Chambers) maintained under 12 h light (30°C, 37% relative humidity)/12 h dark (22°C) cycles. Photon flux density at leaf level was 425 μmol m^{-2} s^{-1}. Irrigation was withheld on day 1 and recommenced on day 14. Leaves were harvested when plants were well-watered as well as after 4, 9, and 12 d of water deprivation and watered for two days following the drought period.

Titratable Acidity

Mature leaves of the same plants used for RNASeq and metabolite profiling were harvested at the end of the light and dark periods, respectively, and frozen in liquid nitrogen after measurement of fresh weight. Organic acids were extracted by sequentially boiling leaves in 50% methanol and water. Titratable acidity was determined by measuring the volume of 5 mM KOH required to neutralize the aqueous extract to pH 6.5.

Metabolite Profiling

Lyophilized leaf material was extracted for metabolite analysis by gas chromatography-mass spectrometry (GC-MS) according to Fiehn et al. (2000) using a 7200 GC-QTOF (Agilent). Data analysis was conducted with the Mass Hunter Software (Agilent). For relative quantification, all metabolite peak areas were normalized to the peak area of the internal standard ribitol added prior to extraction.

RNA Extraction, Preparation, and Sequencing of Illumina Libraries

The topmost mature unshaded leaves (of approximately 3–4.5 cm length) of *T. triangulare* were harvested in the middle of the light or the middle of the dark period and immediately frozen in liquid nitrogen. RNA was isolated from ground tissue using the GeneMatrix Universal RNA Purification Kit (EURx Ltd.). Residues of DNA were removed with DNase (New England Biolabs). RNA integrity, sequencing library, and fragment size were analyzed on a 2100 Bioanalyzer (Agilent). Libraries were prepared using the TruSeq RNA Sample Prep Kit v2 (Illumina) and quantified with a Qubit 2.0 (Invitrogen). Samples were multiplexed with 12 libraries per lane and sequenced in single-end mode (Rapid Run, 150 bp read length) on an Illumina HiSEquation 2000 platform, yielding ~14 million reads per library.

Contig Assembly and Sequence Alignment

To substantiate cross-species mapping against the sequenced reference genome of Arabidopsis, Illumina libraries from representative samples (day 0, day 9, and 2 d after rewatering) were pooled and paired-end sequenced (Rapid Run, 100 bp read length) on an Illumina HiSEquation 2000 platform for assembly of contigs. Trinity assembly of the reads using Trinity (Grabherr et al., 2011) in default mode yielded 105,520 contigs, which were collapsed to 39,781 putative open reading frames using Transdecoder (transdecoder.github.io) and manual removal of duplicated contigs. Remapping placed 81% of the original 150bp reads on the ORFs. Contigs were annotated via BlastX implemented in BLAST+ (Camacho et al., 2009) against peptide sequences of the *Beta vulgaris* reference genome RefBeet-1.1 (Dohm et al., 2014) and peptide sequences of a minimal set of the TAIR10 release of the Arabidopsis genome (http://www.arabidopsis.org/). Protein sequences were aligned using Clustal Omega at www.ebi.ac.uk (Sievers et al., 2011). Phylogenetic analyses were performed with the PhyML tool (Guindon et al., 2010) implemented in SeaView 4 (Gouy et al., 2010) in default mode after Gblocks creation with the least stringent parameters.

Read Mapping and Gene Expression Profiling

Illumina reads were aligned to a minimal set of coding sequences of the TAIR 10 release of the Arabidopsis genome (http://www.arabidopsis.org/) using BLAT (Kent, 2002) in protein space. The minimal transcriptome was obtained as described in Gowik et al. (2011) and contains 21,869 nuclear encoded protein-coding genes. The best BLAT hit for each read was determined by (1) lowest e-value and (2) highest bit score. Raw read counts were transformed to reads per million (rpm) to normalize for the number of reads available at each sampling stage. Cross-species mapping takes advantage of the completeness and annotation of the Arabidopsis genome and overcomes the limitations of transcriptome assembly (Franssen et al., 2011; Schliesky et al., 2012). To quantify abundance of contigs, the 150 bp single-end reads were mapped onto the assembled contigs using the RNA-Seq analysis tool implemented in the CLC genomics workbench version 8.5 (www.clcbio.com) in default mode.

Data Analysis

Data analysis was performed using R statistics software (R version 3.2.1 provided by the CRAN project, http://www.R-project.org). Differential gene expression was analyzed with the DESeq2 package (Love et al., 2014) in default mode. Automatic independent filtering retained 16,766 genes with nonzero total read count. A significance threshold of 0.01 was applied after p-value adjustment with false discovery rate via Benjamini-Hochberg correction (Benjamini and Hochberg, 1995). Log_2 expression ratios calculated by DESeq2 were used for downstream analysis. For *k*-means clustering, transcript levels were log_2-transformed and scaled via calculation of z-scores by gene. One *k*-means clustering was performed for each daytime of sampling (MD and MN) by R statistics software with 6,800 (MD) and 7,563 (MN) genes that were determined as differentially expressed genes (DEG) compared to well-watered (day 0), respectively. Enrichment for GO terms of biological processes within *k*-means clusters was performed via Fisher's Exact Test implemented in the R-package topGO (Alexa et al., 2006; including DEG as the background set, significance level of alpha = 0.01). All log_2-fold changes of water-limited days compared to day 0 were used for pathway analysis in Mapman (Thimm et al., 2004). Wilcoxon Rank Sum test implemented in Mapman was used to test for enrichment and corrected for multiple hypothesis testing via Benjamini-Hochberg correction (Benjamini and Hochberg, 1995). Arabidopsis response to treatments with the phytohormones ABA, 1-aminocyclopropane-1-carboxylic acid (a precursor of ethylene), brassinolide (a brassinosteroid), indole-3-acetic acid (i.e. auxin), and zeatin (a cytokinine) were extracted from AtGenExpress and are based on microarrays by Goda et al. (2008). Response to hormone treatment ($n = 3$) compared to mock treatment ($n = 3$) was called significant at $q < 0.01$ (corrected after Benjamini and Hochberg [1995]). Overlap of *T. triangulare* reponse to drought with Arabidopsis response to hormone treatments was calculated using Fisher's Exact Test ($p < 0.001$). All log_2-fold changes (treatment vs. mock) of Arabidopsis response to ABA treatment (Goda et al., 2008) were used for pathway analysis in Mapman. Data were annotated with publicly available information for Arabidopsis (TAIR10, http://www.arabidopsis.org/). Full quantitative data with annotation is available as Supplemental Dataset S1.

Accession Number

The read data have been submitted to the National Center for Biotechnology Information Gene Expression Omnibus under accession number GSE70601 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=wryvugcifxexnml&acc=GSE70601).

Supplemental Data

The following supplemental materials are available.

Supplemental Data Set S1. Quantitative information and annotation for all reads mapped onto the reference genome of Arabidopsis.

Supplemental Data Set S2. Quantitative information and annotation for all reads mapped onto the *Talinum* contigs.

Supplemental Data Set S3. Levels of all putative CAM genes sensu stricto and sensu lato.

Supplemental Data Set S4. Enriched Mapman categories for all plants harvested 4, 9, and 12 d after water deprivation and 2 d after rewatering (re2) at the middle of the day (MD) or night (MN).

Supplemental Data Set S5. Full list of enriched GO terms in *k*-means clusters (Fig. 5; Supplemental Fig. S7).

Supplemental Data Set S6. Transcription factors differentially expressed on day 9 and/or day 12.

Supplemental Figure S1. Changes in leaf transcriptomes and metabolomes under varying levels of water availability in the middle of the night (analogous to Fig. 1, B and C).

Supplemental Figure S2. *T. triangulare* PEPCs show characteristics of both C_3 and C_4 plants.

Supplemental Figure S3. *Talinum* NADP-ME with highest transcript are predicted to be localized in the plastid.

Supplemental Figure S4. Levels of photorespiratory metabolites.

Supplemental Figure S5. Levels of compatible solutes.

Supplemental Figure S6. Mapman overview of metabolism for all drought samples (day 4, day 9, and day 12 of water deprivation and 2 d after rewatering [re2] compared to day 0).

Supplemental Figure S7. *K*-means clustering of relative gene expression at the middle of the night.

Supplemental Figure S8. Overlap of differentially expressed genes during CAM in *T. triangulare* with hormone treated leaves of Arabidopsis (Goda et al., 2008).

Supplemental Figure S9. Mapman overview of metabolism for ABA response in Arabidopsis.

Supplemental Figure S10. Experiments to assay activity of NADP-ME (panels 1–4) or MDH (panel 5) with leaf extracts of *T. triangulare* ("*Talinum* extract") or recombinant NADP-ME from *Z. mays* (ZmNADP-ME, provided by Anastasiia Bovdilova, group of Veronica G. Maurino).

Supplemental Figure S11. Levels of additional metabolites.

Supplemental Table S1. mRNA-Seq statistics.

Supplemental Table S2. Metabolite profiling statistics.

Supplemental Table S3. Expression levels of *T. triangulare* contigs encoding PEPC.

Supplemental Table S4. Expression levels of all *T. triangulare* contigs encoding for subunits of vacuolar ATP synthases.

Supplemental Table S5. Expression levels of *T. triangulare* contigs encoding NADP-ME.

Supplemental Table S6. Expression levels of two *T. triangulare* contigs encoding PPDK.

Supplemental Table S7. Expression levels of genes encoding for photorespiratory enzymes.

Supplemental Table S8. Estimated partitioning of carbon in organic acids and starch on day 12.

ACKNOWLEDGMENTS

We acknowledge the excellent technical assistance of M. Graf, K. Weber, and E. Klemp for GC-MS measurements, S. Kurz for help with Illumina library preparation, and A. Virgo for titration of organic acids.

Received July 22, 2015; accepted November 3, 2015; published November 3, 2015.

LITERATURE CITED

Alexa A, Rahnenführer J, Lengauer T (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics **22:** 1600–1607

Ashton AR, Burnell JN, Furbank RT, Jenkins CLD, Hatch MD (1990) 3 enzymes of C4 photosynthesis. Enzymes of Primary Metabolism **3:** 39

Baena-González E, Rolland F, Thevelein JM, Sheen J (2007) A central integrator of transcription networks in plant stress and energy signalling. Nature **448:** 938–942

Bartels D, Sunkar R (2005) Drought and salt tolerance in plants. Crit Rev Plant Sci **24:** 23–58

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. J R Stat Soc, B **57:** 289–300

Bläsing OE, Westhoff P, Svensson P (2000) Evolution of C4 phospho*enol*pyruvate carboxylase in *Flaveria*, a conserved serine residue in the carboxyl-terminal part of the enzyme is a major determinant for C4-specific characteristics. J Biol Chem **275:** 27917–27923

Bohnert HJ, Nelson DE, Jensen RG (1995) Adaptations to environmental stresses. Plant Cell **7:** 1099–1111

Borland AM, Barrera Zambrano VA, Ceusters J, Shorrock K (2011) The photosynthetic plasticity of crassulacean acid metabolism: an evolutionary innovation for sustainable productivity in a changing world. New Phytol **191:** 619–633

Borland AM, Hartwell J, Jenkins GI, Wilkins MB, Nimmo HG (1999) Metabolite control overrides circadian regulation of phospho*enol*pyruvate carboxylase kinase and CO2 fixation in crassulacean acid metabolism. Plant Physiol **121:** 889–896

Bray EA (1997) Plant responses to water deficit. Trends Plant Sci **2:** 48–54

Bräutigam A, Hoffmann-Benning S, Weber AP, 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:** 568–579

Bräutigam A, Kajala K, Wullenweber J, Sommer M, Gagneul D, Weber KL, Carr KM, Gowik U, Mass J, Lercher MJ, 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, Schliesky S, Külahoglu C, Osborne CP, Weber APM (2014) Towards an integrative model of C4 photosynthetic subtypes: insights from comparative transcriptome analysis of NAD-ME, NADP-ME, and PEP-CK C4 species. J Exp Bot **65:** 3579–3593

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics **10:** 421

Carter PJ, Nimmo HG, Fewson CA, Wilkins MB (1991) Circadian rhythms in the activity of a plant protein kinase. EMBO J **10:** 2063–2068

Chaves MM, Maroco J, Pereira JS (2003) Understanding plant responses to drought - from genes to the whole plant. Funct Plant Biol **30:** 239–264

Cheffings CM, Pantoja O, Ashcroft FM, Smith JA (1997) Malate transport and vacuolar ion channels in CAM plants. J Exp Bot **48:** 623–631

Chia T, Thorneycroft D, Chapple A, Messerli G, Chen J, Zeeman SC, Smith SM, Smith AM (2004) A cytosolic glucosyltransferase is required for conversion of starch to sucrose in Arabidopsis leaves at night. Plant J **37:** 853–863

Christin P-A, Arakaki M, Osborne CP, Bräutigam A, Sage RF, Hibberd JM, Kelly S, Covshoff S, Wong GK-S, Hancock L, et al (2014) Shared origins of a key enzyme during the evolution of C4 and CAM metabolism. J Exp Bot **65:** 3609–3621

Cushman JC, Agarie S, Albion RL, Elliot SM, Taybi T, Borland AM (2008a) Isolation and characterization of mutants of common ice plant deficient in crassulacean acid metabolism. Plant Physiol **147:** 228–238

Cushman JC, Tillett RL, Wood JA, Branco JM, Schlauch KA (2008b) Large-scale mRNA expression profiling in the common ice plant, *Mesembryanthemum crystallinum*, performing C3 photosynthesis and Crassulacean acid metabolism (CAM). J Exp Bot **59:** 1875–1894

Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR (2010) Abscisic acid: emergence of a core signaling network. Annu Rev Plant Biol **61:** 651–679

Dever LV, Boxall SF, Kneřová J, Hartwell J (2015) Transgenic perturbation of the decarboxylation phase of Crassulacean acid metabolism alters physiology and metabolism but has only a small effect on growth. Plant Physiol **167:** 44–59

Dittrich P (1976) Nicotinamide adenine dinucleotide-specific "malic" enzyme in *Kalanchoë daigremontiana* and other plants exhibiting crassulacean acid metabolism. Plant Physiol **57:** 310–314

Dohm JC, Minoche AE, Holtgräwe D, Capella-Gutiérrez S, Zakrzewski F, Tafer H, Rupp O, Sörensen TR, Stracke R, Reinhardt R, et al (2014) The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*). Nature **505:** 546–549

Elbein AD, Pan YT, Pastuszak I, Carroll D (2003) New insights on trehalose: a multifunctional molecule. Glycobiology **13:** 17R–27R

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

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

Findling S, Zanger K, Krueger S, Lohaus G (2015) Subcellular distribution of raffinose oligosaccharides and other metabolites in summer and winter leaves of *Ajuga reptans* (*Lamiaceae*). Planta **241:** 229–241

Flexas J, Bota J, Galmés J, Medrano H, Ribas-Carbó M (2006) Keeping a positive carbon balance under adverse conditions: responses of photosynthesis and respiration to water stress. Physiol Plant **127:** 343–352

Franssen SU, Shrestha RP, Bräutigam A, Bornberg-Bauer E, Weber APM (2011) Comprehensive transcriptome analysis of the highly complex *Pisum sativum* genome using next generation sequencing. BMC Genomics **12:** 227–242

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

Goda H, Sasaki E, Akiyama K, Maruyama-Nakashita A, Nakabayashi K, Li W, Ogawa M, Yamauchi Y, Preston J, Aoki K, et al (2008) The AtGenExpress hormone and chemical treatment data set: experimental design, data evaluation, model data analysis and data access. Plant J **55:** 526–542

Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol **27:** 221–224

Gowik U, Bräutigam A, Weber KL, Weber APM, 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

Gowik U, Engelmann S, Bläsing OE, Raghavendra AS, Westhoff P (2006) Evolution of C(4) phospho*enol*pyruvate carboxylase in the genus *Alternanthera*: gene families and the enzymatic characteristics of the C(4) isozyme and its orthologues in C(3) and C(3)/C(4) *Alternantheras*. Planta **223:** 359–368

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol **29:** 644–652

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol **59:** 307–321

Haider MS, Barnes JD, Cushman JC, Borland AM (2012) A CAM- and starch-deficient mutant of the facultative CAM species *Mesembryanthemum crystallinum* reconciles sink demands by repartitioning carbon during acclimation to salinity. J Exp Bot **63:** 1985–1996

Hare PD, Cress WA, Van Staden J (1998) Dissecting the roles of osmolyte accumulation during stress. Plant Cell Environ **21:** 535–553

Harris D, Martin C (1991) Plasticity in the degree of CAM-cycling and its relationship to drought stress in 5 species of *Talinum* (*Portulacaceae*). Oecologia **86:** 575–584

Hartwell J, Gill A, Nimmo GA, Wilkins MB, Jenkins GI, Nimmo HG (1999) Phospho*enol*pyruvate carboxylase kinase is a novel protein kinase regulated at the level of expression. Plant J **20:** 333–342

Hartwell J, Smith LH, Wilkins MB, Jenkins GI, Nimmo HG (1996) Higher plant phospho*enol*pyruvate carboxylase kinase is regulated at the level of translatable mRNA in response to light or a circadian rhythm. Plant J **10:** 1071–1078

Häusler RE, Baur B, Scharte J, Teichmann T, Eicks M, Fischer KL, Flügge U-I, Schubert S, Weber A, Fischer K (2000) Plastidic metabolite transporters and their physiological functions in the inducible crassulacean acid metabolism plant *Mesembryanthemum crystallinum*. Plant J **24:** 285–296

Heddad M, Adamska I (2000) Light stress-regulated two-helix proteins in *Arabidopsis thaliana* related to the chlorophyll a/b-binding gene family. Proc Natl Acad Sci USA **97:** 3741–3746

Herrera A (1999) Effects of photoperiod and drought on the induction of *Crassulacean* acid metabolism and the reproduction of plants of *Talinum triangulare*. Can J Bot **77:** 404–409

Herrera A (2009) Crassulacean acid metabolism and fitness under water deficit stress: if not for carbon gain, what is facultative CAM good for? Ann Bot (Lond) **103:** 645–653

Herrera A, Ballestrini C, Montes E (2015) What is the potential for dark CO2 fixation in the facultative crassulacean acid metabolism species *Talinum triangulare?* J Plant Physiol **174:** 55–61

Herrera A, Delgado J, Paraguatey I (1991) Occurrence of inducible crassulacean acid metabolism in leaves of *Talinum triangulare* (*Portulacaceae*). J Exp Bot **42:** 493–499

Holtum JAM, Smith JAC, Neuhaus HE (2005) Intracellular transport and pathways of carbon flow in plants with crassulacean acid metabolism. Funct Plant Biol **32:** 429–449

Holtum JAM, Winter K (1982) Activity of enzymes of carbon metabolism during the induction of Crassulacean acid metabolism in *Mesembryanthemum crystallinum* L. Planta **155:** 8–16

Hutin C, Nussaume L, Moise N, Moya I, Kloppstech K, Havaux M (2003) Early light-induced proteins protect *Arabidopsis* from photooxidative stress. Proc Natl Acad Sci USA **100:** 4921–4926

Ingram J, Bartels D (1996) The molecular basis of dehydration tolerance in plants. Annu Rev Plant Physiol Plant Mol Biol **47:** 377–403

Kent WJ (2002) BLAT - The BLAST-like alignment tool. Genome Res **12:** 656–664

Kim DH, Yamaguchi S, Lim S, Oh E, Park J, Hanada A, Kamiya Y, Choi G (2008) SOMNUS, a CCCH-type zinc finger protein in *Arabidopsis*, negatively regulates light-dependent seed germination downstream of PIL5. Plant Cell **20:** 1260–1277

Kluge M, Osmond CB (1971) Pyruvate Pi dikinase in crassulacean acid metabolism. Naturwissenschaften **58:** 414–415

Kollist H, Nuhkat M, Roelfsema MRG (2014) Closing gaps: linking elements that control stomatal movement. New Phytol **203:** 44–62

Kondo A, Nose A, Ueno O (2001) Coordinated accumulation of the chloroplastic and cytosolic pyruvate, Pi dikinases with enhanced expression of CAM in *Kalanchoë blossfeldiana*. Physiol Plant **111:** 116–122

Kore-eda S, Noake C, Ohishi M, Ohnishi J, Cushman JC (2005) Transcriptional profiles of organellar metabolite transporters during induction of Crassulacean acid metabolism in *Mesembryanthemum crystallinum*. Funct Plant Biol **32:** 451–466

Lawlor DW, Cornic G (2002) Photosynthetic carbon assimilation and associated metabolism in relation to water deficits in higher plants. Plant Cell Environ **25:** 275–294

Lim S, Park J, Lee N, Jeong J, Toh S, Watanabe A, Kim J, Kang H, Kim DH, Kawakami N, et al (2013) ABA-insensitive3, ABA-insensitive5, and DELLAs interact to activate the expression of SOMNUS and other high-temperature-inducible genes in imbibed seeds in Arabidopsis. Plant Cell **25:** 4863–4878

Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. Plant Cell **10:** 1391–1406

Lobell DB, Gourdji SM (2012) The influence of climate change on global crop productivity. Plant Physiol **160:** 1686–1697

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **15:** 550–570

Lüttge U (2002) CO_2-concentrating: consequences in crassulacean acid metabolism. J Exp Bot **53:** 2131–2142

Lüttge U (2011) Photorespiration in phase III of crassulacean acid metabolism: evolutionary and ecophysiological implications. Prog Bot **72:** 371–384

Nägele T, Heyer AG (2013) Approximating subcellular organisation of carbohydrate metabolism during cold acclimation in different natural accessions of *Arabidopsis thaliana*. New Phytol **198:** 777–787

Niewiadomska E, Borland AM (2007) Crassulacean acid metabolism: a cause or consequence of oxidative stress in planta? Prog Bot **69:** 247–266

Niittylä T, Messerli G, Trevisan M, Chen J, Smith AM, Zeeman SC (2004) A previously unknown maltose transporter essential for starch degradation in leaves. Science **303:** 87–89

Nishizawa A, Yabuta Y, Shigeoka S (2008) Galactinol and raffinose constitute a novel function to protect plants from oxidative damage. Plant Physiol **147:** 1251–1263

Noctor G, Foyer CH (2000) Homeostasis of adenylate status during photosynthesis in a fluctuating environment. J Exp Bot **51:** 347–356

Osmond CB (1978) Crassulacean acid metabolism: a curiosity in context. Annu Rev Plant Physiol **29:** 379–414

Park S-Y, Fung P, Nishimura N, Jensen DR, Fujii H, Zhao Y, Lumba S, Santiago J, Rodrigues A, Chow T-FF, et al (2009) Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. Science **324:** 1068–1071

Parsley K, Hibberd JM (2006) The *Arabidopsis* PPDK gene is transcribed from two promoters to produce differentially expressed transcripts responsible for cytosolic and plastidic proteins. Plant Mol Biol **62:** 339–349

Paulus JK, Schlieper D, Groth G (2013) Greater efficiency of photosynthetic carbon fixation due to single amino-acid substitution. Nat Commun **4:** 1518–1524

Pérez-Rodríguez P, Riaño-Pachón DM, Corrêa LGG, Rensing SA, Kersten B, Mueller-Roeber B (2010) PlnTFDB: updated content and new features of the plant transcription factor database. Nucleic Acids Res **38:** D822–D827

Sage RF (2003) The evolution of C4 photosynthesis. New Phytol **161:** 341–370

Sakamoto H, Araki T, Meshi T, Iwabuchi M (2000) Expression of a subset of the Arabidopsis Cys(2)/His(2)-type zinc-finger protein gene family under water stress. Gene **248:** 23–32

Sakamoto H, Maruyama K, Sakuma Y, Meshi T, Iwabuchi M, Shinozaki K, Yamaguchi-Shinozaki K (2004) Arabidopsis Cys2/His2-type zinc-finger proteins function as transcription repressors under drought, cold, and high-salinity stress conditions. Plant Physiol **136:** 2734–2746

Schliesky S, Gowik U, Weber APM, Bräutigam A (2012) RNA-Seq assembly - Are we there yet? Front Plant Sci **3:** 220

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, et al (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol **7:** 539–544

Smith SM, Fulton DC, Chia T, Thorneycroft D, Chapple A, Dunstan H, Hylton C, Zeeman SC, Smith AM (2004) Diurnal changes in the transcriptome encoding enzymes of starch metabolism provide evidence for both transcriptional and posttranscriptional regulation of starch metabolism in *Arabidopsis* leaves. Plant Physiol **136:** 2687–2699

Streb S, Zeeman SC (2012) Starch metabolism in Arabidopsis. The Arabidopsis Book **10:** doi/10.199/tab.e0160

Sun J, Jiang H, Xu Y, Li H, Wu X, Xie Q, Li C (2007) The CCCH-type zinc finger proteins AtSZF1 and AtSZF2 regulate salt stress responses in *Arabidopsis*. Plant Cell Physiol **48:** 1148–1158

Svensson P, Bläsing OE, Westhoff P (1997) Evolution of the enzymatic characteristics of C4 phospho*enol*pyruvate carboxylase–a comparison of the orthologous PPCA phospho*enol*pyruvate carboxylases of *Flaveria trinervia* (C4) and *Flaveria pringlei* (C3). Eur J Biochem **246:** 452–460

Svensson P, Bläsing OE, Westhoff P (2003) Evolution of C4 phospho*enol*pyruvate carboxylase. Arch Biochem Biophys **414:** 180–188

Taisma MA, Herrera A (1998) A relationship between fecundity, survival, and the operation of crassulacean acid metabolism in *Talinum triangulare*. Can J Bot **76:** 1908–1915

Taisma MA, Herrera A (2003) Drought under natural conditions affects leaf properties, induces CAM and promotes reproduction in plants of *Talinum triangulare*. Interciencia **28:** 292–297

Taybi T, Cushman JC (2002) Abscisic acid signaling and protein synthesis requirements for phospho*enol*pyruvate carboxylase transcript induction in the common ice plant. J Plant Physiol **159:** 1235–1243

Taybi T, Nimmo HG, Borland AM (2004) Expression of phospho*enol*pyruvate carboxylase and phospho*enol*pyruvate carboxylase kinase genes. Implications for genotypic capacity and phenotypic plasticity in the expression of crassulacean acid metabolism. Plant Physiol **135:** 587–598

Taybi T, Patil S, Chollet R, Cushman JC (2000) A minimal serine/threonine protein kinase circadianly regulates phospho*enol*pyruvate carboxylase activity in crassulacean acid metabolism-induced leaves of the common ice plant. Plant Physiol **123:** 1471–1482

Taybi T, Sotta B, Gehrig H, Guclu S, Kluge M, Brulfert J (1995) Differential effects of abscisic acid on phospho*enol*pyruvate carboxylase and CAM operation in *Kalanchoë blossfeldiana*. Bot Acta **108:** 240–246

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

Tsuzuki M, Miyachi S, Winter K, Edwards GE (1982) Localization of carbonic anhydrase in crassulacean acid metabolism plants. Plant Sci Lett **24:** 211–218

Umezawa T, Sugiyama N, Mizoguchi M, Hayashi S, Myouga F, Yamaguchi-Shinozaki K, Ishihama Y, Hirayama T, Shinozaki K (2009) Type 2C protein phosphatases directly regulate abscisic acid-activated protein kinases in *Arabidopsis*. Proc Natl Acad Sci USA **106:** 17588–17593

Weber A, Servaites JC, Geiger DR, Kofler H, Hille D, Gröner F, Hebbeker U, Flügge U-I (2000) Identification, purification, and molecular cloning of a putative plastidic glucose translocator. Plant Cell **12:** 787–802

Weise SE, van Wijk KJ, Sharkey TD (2011) The role of transitory starch in C(3), CAM, and C(4) metabolism and opportunities for engineering leaf starch accumulation. J Exp Bot **62:** 3109–3118

White PJ, Smith JA (1989) Proton and anion transport at the tonoplast in crassulacean-acid-metabolism plants: specificity of the malate-influx system in *Kalanchoë daigremontiana.* Planta **179:** 265–274

Wilson RL, Kim H, Bakshi A, Binder BM (2014) The ethylene receptors ETHYLENE RESPONSE1 and ETHYLENE RESPONSE2 have contrasting roles in seed germination of Arabidopsis during salt stress. Plant Physiol **165:** 1353–1366

Winter K, Foster JG, Edwards GE, Holtum JAM (1982) Intracellular localization of enzymes of carbon metabolism in *Mesembryanthemum crystallinum* exhibiting C_3 photosynthesic characteristics or performing Crassulacean acid metabolism. Plant Physiol **69:** 300–307

Winter K, Garcia M, Holtum JAM (2008) On the nature of facultative and constitutive CAM: environmental and developmental control of CAM expression during early growth of *Clusia, Kalanchöe*, and *Opuntia.* J Exp Bot **59:** 1829–1840

Winter K, Holtum JAM (2007) Environment or development? Lifetime net CO_2 exchange and control of the expression of Crassulacean acid metabolism in *Mesembryanthemum crystallinum.* Plant Physiol **143:** 98–107

Winter K, Holtum JAM (2014) Facultative crassulacean acid metabolism (CAM) plants: powerful tools for unravelling the functional elements of CAM photosynthesis. J Exp Bot **65:** 3425–3441

Winter K, Smith JAC (1996) An introduction to crassulacean acid metabolism. Biochemical principles and ecological diversity. *In* Crassulacean Acid Metabolism. Springer, Berlin Germany, pp 1–13

Winter K, von Willert DJ (1972) NaCl-induzierter Crassulaceansäurestoffwechsel bei *Mesembryanthemum crystallinum.* Z Pflanzenphysiol **67:** 166–170

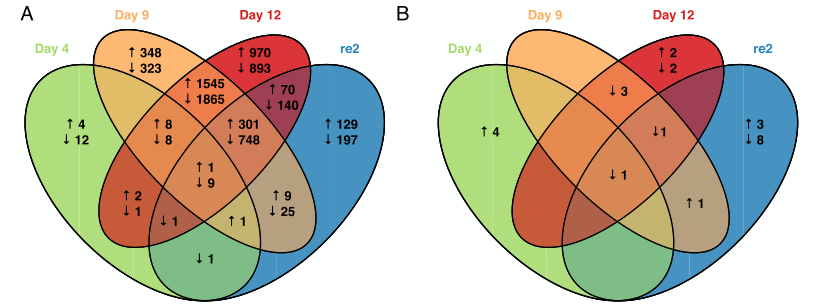
Yang J, Worley E, Udvardi M (2014) A NAP-AAO3 regulatory module promotes chlorophyll degradation via ABA biosynthesis in Arabidopsis leaves. Plant Cell **26:** 4862–4874

Yang X, Cushman JC, Borland AM, Edwards EJ, Wullschleger SD, Tuskan GA, Owen NA, Griffiths H, Smith JAC, De Paoli HC, et al (2015) A roadmap for research on crassulacean acid metabolism (CAM) to enhance sustainable food and bioenergy production in a hotter, drier world. New Phytol **207:** 491–504

Zelisko A, García-Lorenzo M, Jackowski G, Jansson S, Funk C (2005) AtFtsH6 is involved in the degradation of the light-harvesting complex II during high-light acclimation and senescence. Proc Natl Acad Sci USA **102:** 13699–13704

Zhu J-K (2002) Salt and drought stress signal transduction in plants. Annu Rev Plant Biol **53:** 247–273

Zou C, Sun K, Mackaluso JD, Seddon AE, Jin R, Thomashow MF, Shiu S-H (2011) Cis-regulatory code of stress-responsive transcription in *Arabidopsis thaliana.* Proc Natl Acad Sci USA **108:** 14992–14997

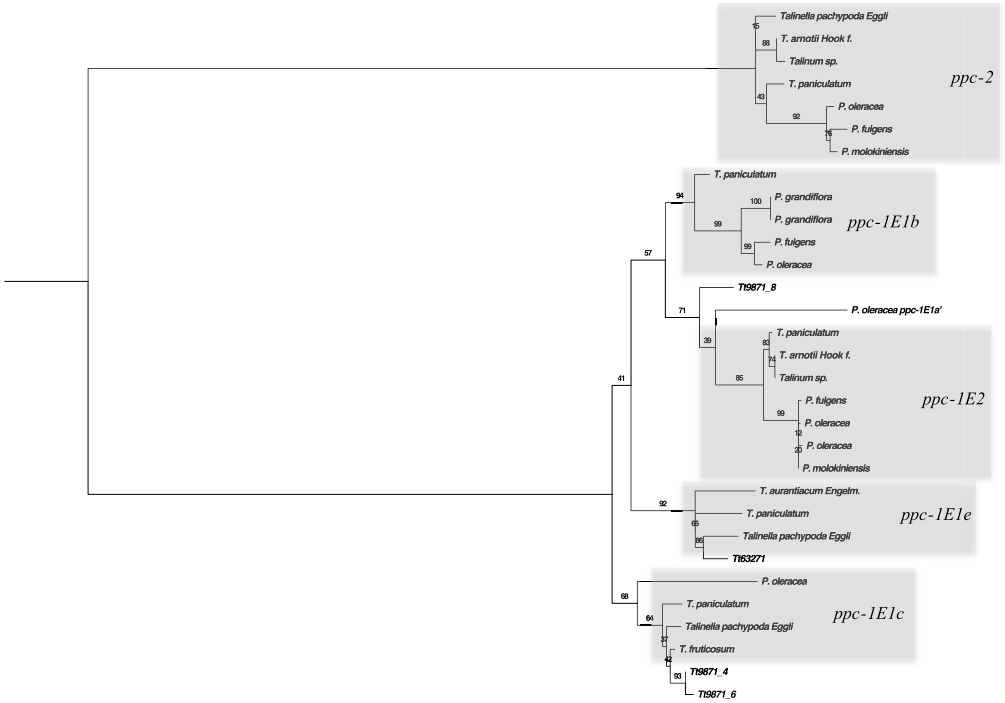


Supplemental Figure S1. Changes in leaf transcriptomes and metabolomes under varying levels of water availability in the middle of the night (Analogous to Figure 1. B and C). Venn diagrams representing overlapping changes (↑: increased, ↓: depleted) in gene expression (A, DESeq2 $q < 0.01$, $n = 3$, 16,766 genes analyzed in total), or metabolite levels (B, Student's *t* test, $p < 0.05$, $n = 3$-4, 39 metabolites measured in total) at the middle of the night between water-limited stages compared to day 0.

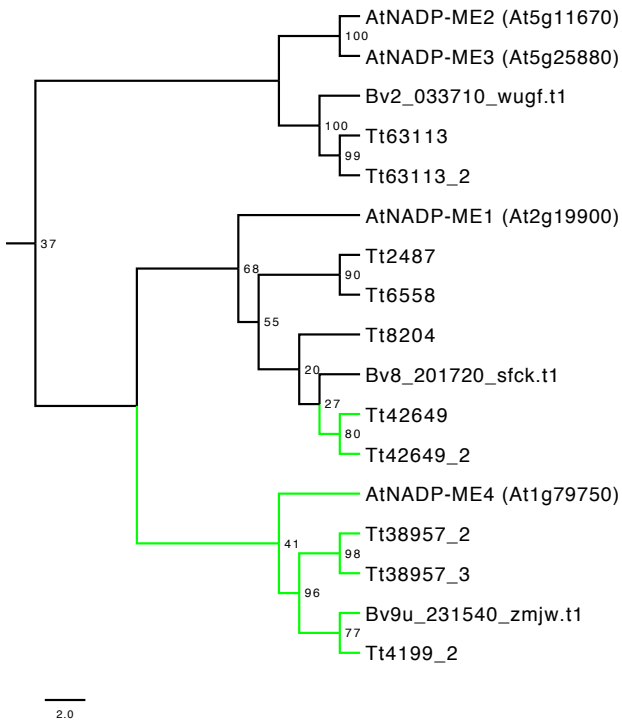
A

```
Zea mays (C4)                                779 FSWTQTRFHLPV(…)ILEGDPFLKQGLVLRNPY     897
Flaveria trinervia (C4)                      773 FSWTQTRFHLPV(…)LLEGDPYLKQGIRLRDPY     891
Sorghum bicolor (C4)                         770 FSWTQTRFHLPV(…)ILEGDPYLKQGLRLRNPY     888
Flaveria pringlei (C3)                       774 FAWTQTRFHLPV(…)LLEGDPYLKQRIRLRDSY     892
Arabidopsis thaliana (C3)                    775 FAWTQTRFHLPV(…)LLEGDPYLKQRLRLRDSY     893
Mesembryanthemum crystallinum (CAM)          774 FAWTQTRFHLPV(…)LLEGDPYLKQRLRLRDPY     892
Portulaca fulgens ppc-1E1b (C4)              341 FAWTQTRFHLPV(…)VLEGDPYLKQRLRLRDPY     459
Portulaca grandiflora ppc-1E1b (C4-CAM)      341 FAWTQTRFHLPV(…)VLEGDPYLKQRLRLRDPY     459
Portulaca oleracea ppc-1E1a' (C4-CAM)        334 FAWTQTRFHLPV(…)VLEGDPYLKQRLRLRDPY     452
Portulaca oleracea ppc-1E1b (C4-CAM)         673 FSWTQTRFHLPV(…)LLERNPSLKQRLHLRGPY     791
Portulaca oleracea ppc-1E1c (C4-CAM)         762 FSWTQTRLHLPV(…)LLENDPYLKQMLRLRDPY     881
T. triangulare ppc1E1c                       19  FSWTQTRFHLPV(…)LLEGNPSLKQRLHLRGPY     137
Tt63271                                      192 FAWTQTRFHLPV(…)LLERDPYLKHRLHLRDPY     310
Tt9871_4                                     750 FSWTQTRFHLPV(…)LLEGNPSLKQRLHLRGPY     868
Tt9871_6                                     750 FSWTQTRFHLPV(…)LLEGNPSLKQRLHLRGPY     868
Tt9871_8                                     246 FAWTQTRFHLPV(…)LLEGDPYLKQRLRLRDPY     364
```

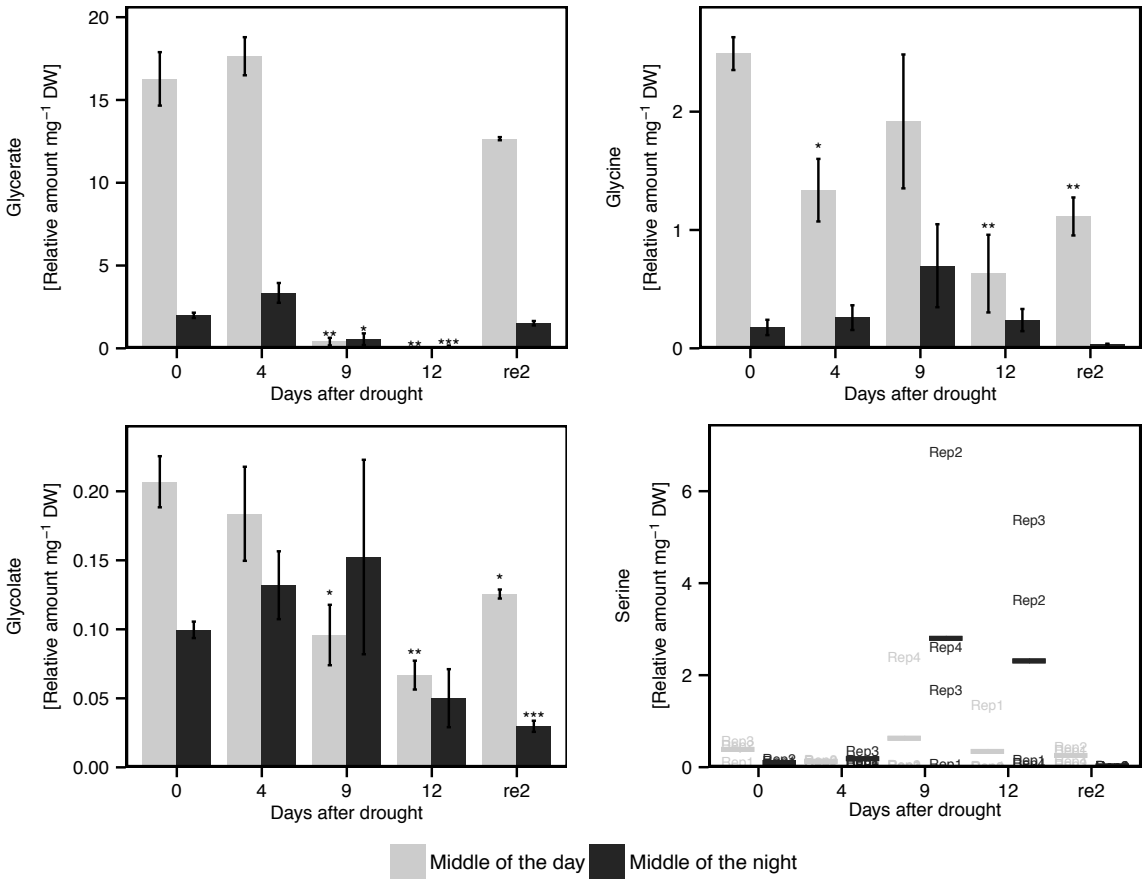
B



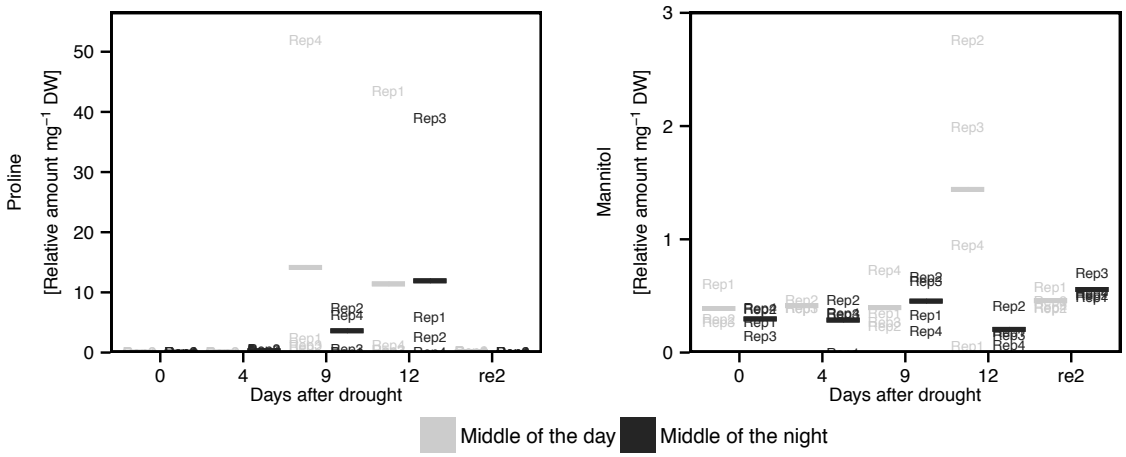
Supplemental Figure S2. *T. triangulare* PEPCs show characteristics of both C_3 and C_4 plants. A, Alignment of translated *T. triangulare* contigs encoding PEPC with protein sequences of C_4 plants (*Zea mays* PPC1, UniProt No.: CAA33317; *Flaveria trinervia* PPCA, UniProt No.: P30694; *Sorghum bicolor* PEPC3, UniProt No.: P15804; *Portulaca fulgens* ppc-1E1b, UniProt No.: KJ161577), C_3 plants (*Flaveria pringlei* PPCA1, UniProt No.: Q01647; *Arabidopsis thaliana* PEPC1, UniProt No.: Q9MAH0), facultative CAM plants (*Mesembryanthemum crystallinum* PEPC1, UniProt No.: P10490; *T. triangulare* ppc1E1c, UniProt No.: A0A075J1W2) and C_4-CAM plants (*Portulaca grandiflora* ppc-1E1b, UniProt No.: KJ161578; *Portulaca oleracea* ppc-1E1b, UniProt No.: KJ161580; *P. oleracea* ppc1E1a' and ppc1E1c are translated contigs from Christin et al. (2014)). Sequences of *M. crystallinum*, *Portulaceae* and *T. triangulare* ppc1E1c were retrieved from Christin et al. (2014). Shaded areas highlight the amino residues determining PEP saturation kinetics (Ser780, counting based on *Zea mays*, Bläsing et al. (2000)) and sensitivity towards malate inhibition (Gly890, Paulus et al. (2013)). B, Maximum likelihood estimate of phylogenetic relationships of *T. triangulare* contigs encoding PEPC and protein sequences of the *ppc* gene families of *Portulacaceae* and *Talinaceae, ppc-1E1b* (*T. paniculatum*, UniProt No.: KJ161572; *P. fulgens*, UniProt No.: KJ161577; *P. grandiflora*, UniProt No.: KJ161578; *P. grandiflora*, UniProt No.: KJ161579; *P. oleracea*, UniProt No.: KJ161580), *ppc-1E1c* (*T. paniculatum*, UniProt No.: KJ161605; *T. fruticosum*, UniProt No.: KJ161606; *Talinella pachypoda* Eggli, UniProt No.: KJ161607), *ppc-1E1e* (*T. aurantiacum* Engelm., UniProt No.: KJ161647; *T. paniculatum*, UniProt No.: KJ161648; *Talinella pachypoda* Eggli, UniProt No.: KJ161649), *ppc-1E2* (*T. arnotii* Hook f, UniProt No.: KJ161736; *T. paniculatum*, UniProt No.: KJ161737; *Talinum* sp., UniProt No.: KJ161738; *P. fulgens*, UniProt No.: KJ161742; *P. molokiniensis*, UniProt No.: KJ161743; *P. oleracea*, UniProt No.: KJ161744; *P. oleracea*, UniProt No.: KJ161745), *ppc-2* (*P. fulgens*, UniProt No.: KJ161807; *P. molokiniensis*, UniProt No.: KJ161808; *P. oleracea*, UniProt No.: KJ161809; *T. arnotii* Hook, UniProt No.: KJ161799; *T. paniculatum*, UniProt No.: KJ161800; *Talinum sp.*, UniProt No.: KJ161802; *Talinella pachypoda* Eggli, UniProt No.: KJ161801) extracted from Christin et al. (2014). Numbers above nodes represent maximum likelihood bootstrap values (n = 100). For details see Material and Methods. See Supplemental Table S3 for quantification of the *T. triangulare* contigs. Only contigs with an average expression of over 100 reads per million across all 30 samples were included.



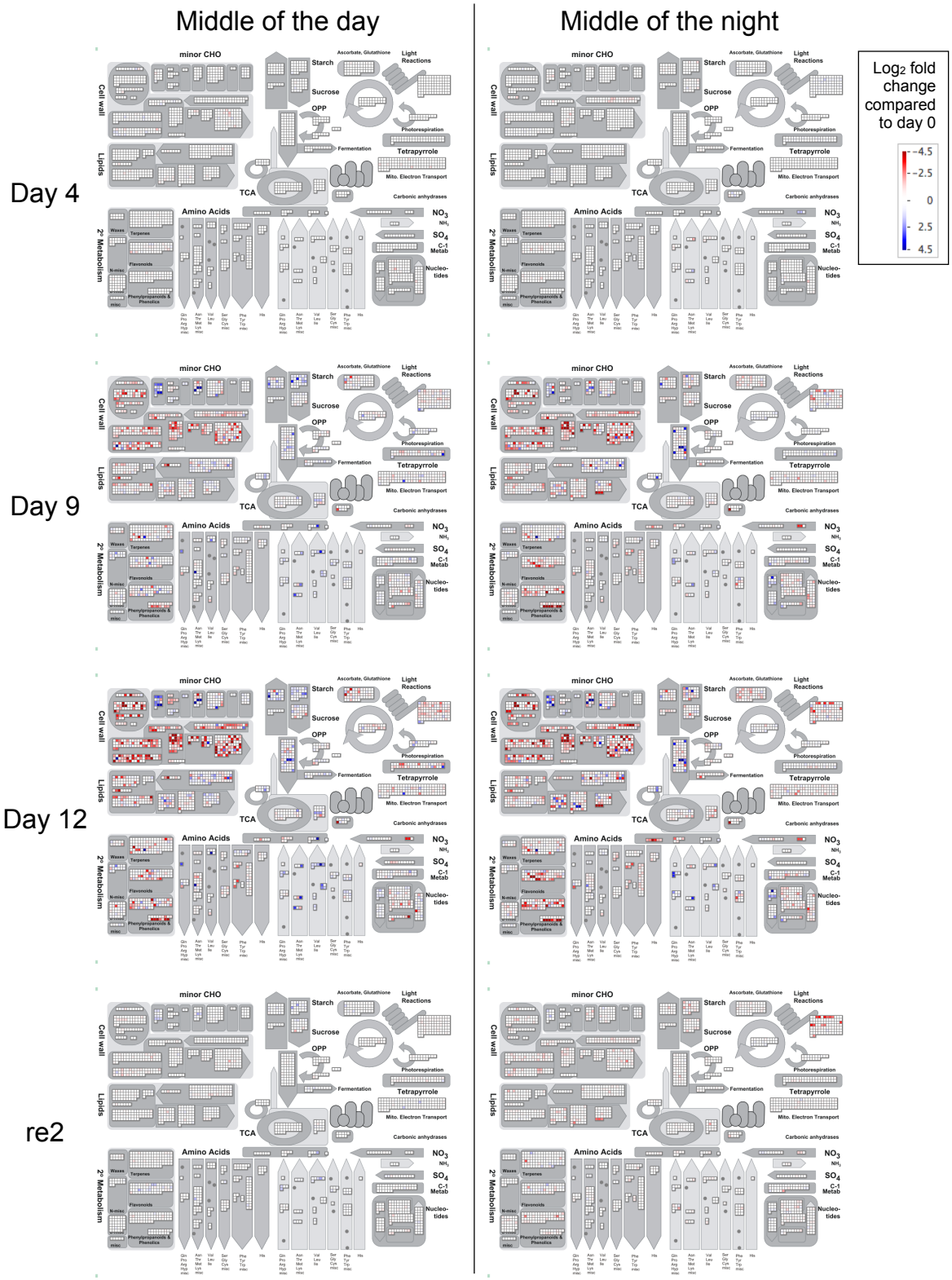
Supplemental Figure S3. *Talinum* NADP-ME with highest transcript are predicted to be localized in the plastid. Maximum likelihood estimate of phylogenetic relationships of *T. triangulare* contigs and *B. vulgaris* and *A. thaliana* genes encoding isoforms of NADP-ME. PhyML analysis was performed on protein level. Numbers above nodes represent maximum likelihood bootstrap values (n = 100). Green nodes indicate predicted chloroplast targeting based on TargetP (Emanuelsson et al., 2000). For details see Material and Methods. See Supplemental Table S4 for quantification of the *T. triangulare* contigs.



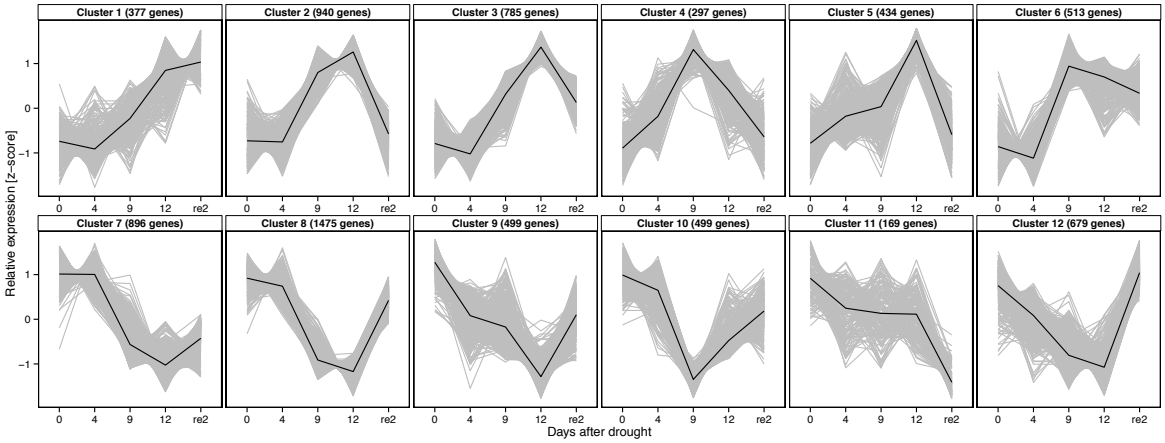
Supplemental Figure S4. Levels of photorespiratory metabolites. Metabolites were measured by GC-MS and normalized to dry weight (DW) and internal ribitol standard (Mean ± SE, n = 3-4, Asterisks indicate Student's *t* test significance in comparison to day 0 at ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$). re2, two days after re-watering. Bar in serine plot represents the mean of biological replicates (Rep).



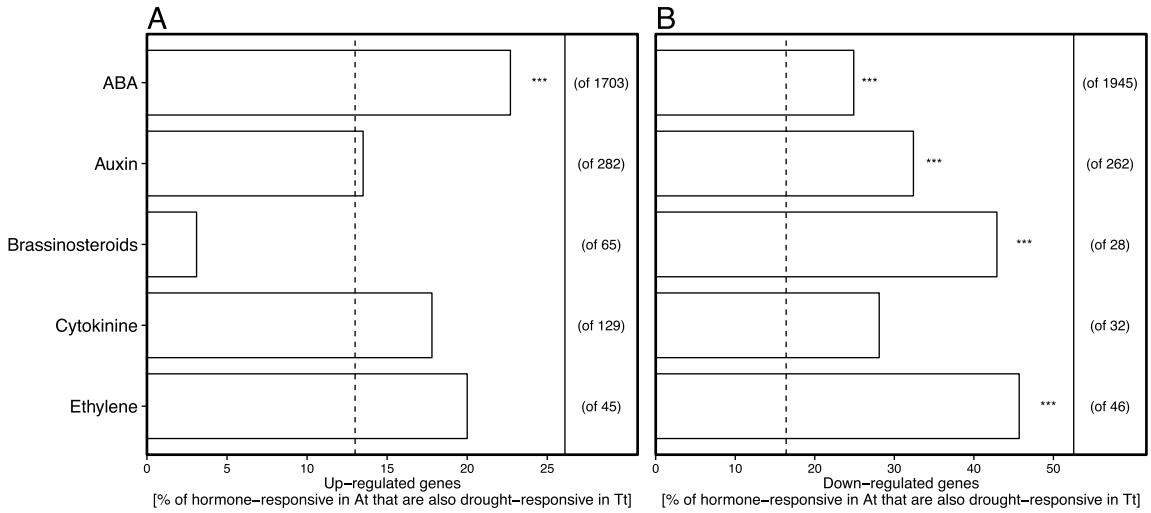
Supplemental Figure S5. Levels of compatible solutes. Metabolites were measured by GC-MS and normalized to dry weight (DW) and internal ribitol standard. Bar represents the mean of biological replicates (Rep).



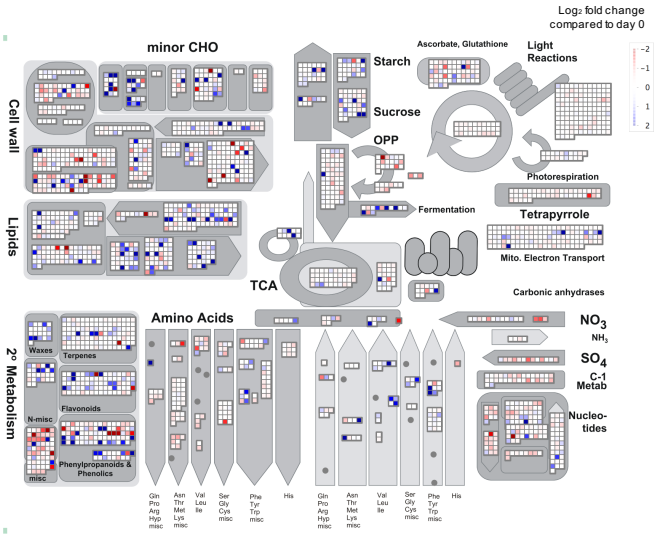
Supplemental Figure S6. Mapman overview of metabolism for all drought samples (day 4, day 9, day 12 of water-deprivation and 2 days after re-watering (re2) compared to day 0. Details see Materials and Methods.



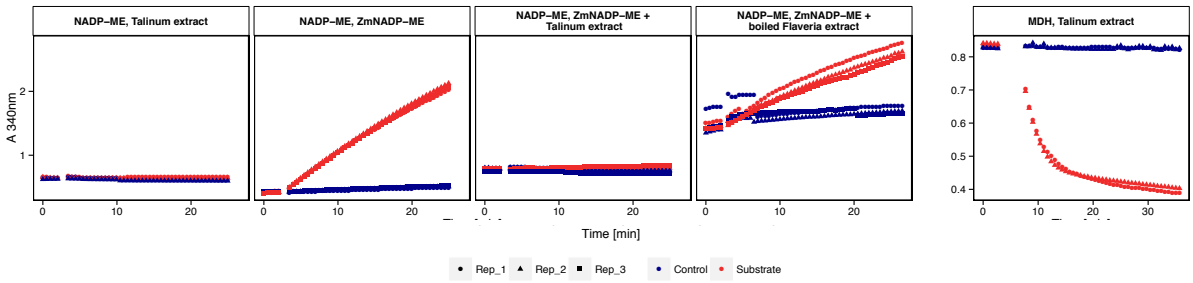
Supplemental Figure S7. *K*-means clustering of relative gene expression at the middle of the night. Genes that were found to be differentially expressed in the middle of the night between one of the water-limited stages and day 0 (DESeq2, $q < 0.01$) were used for the *k*-means approach (7563 genes in total). For a list of enriched GO Terms see Supplemental Dataset S4. Grey line, expression of single genes; black line, average of all genes in cluster; re2, two days after re-watering.



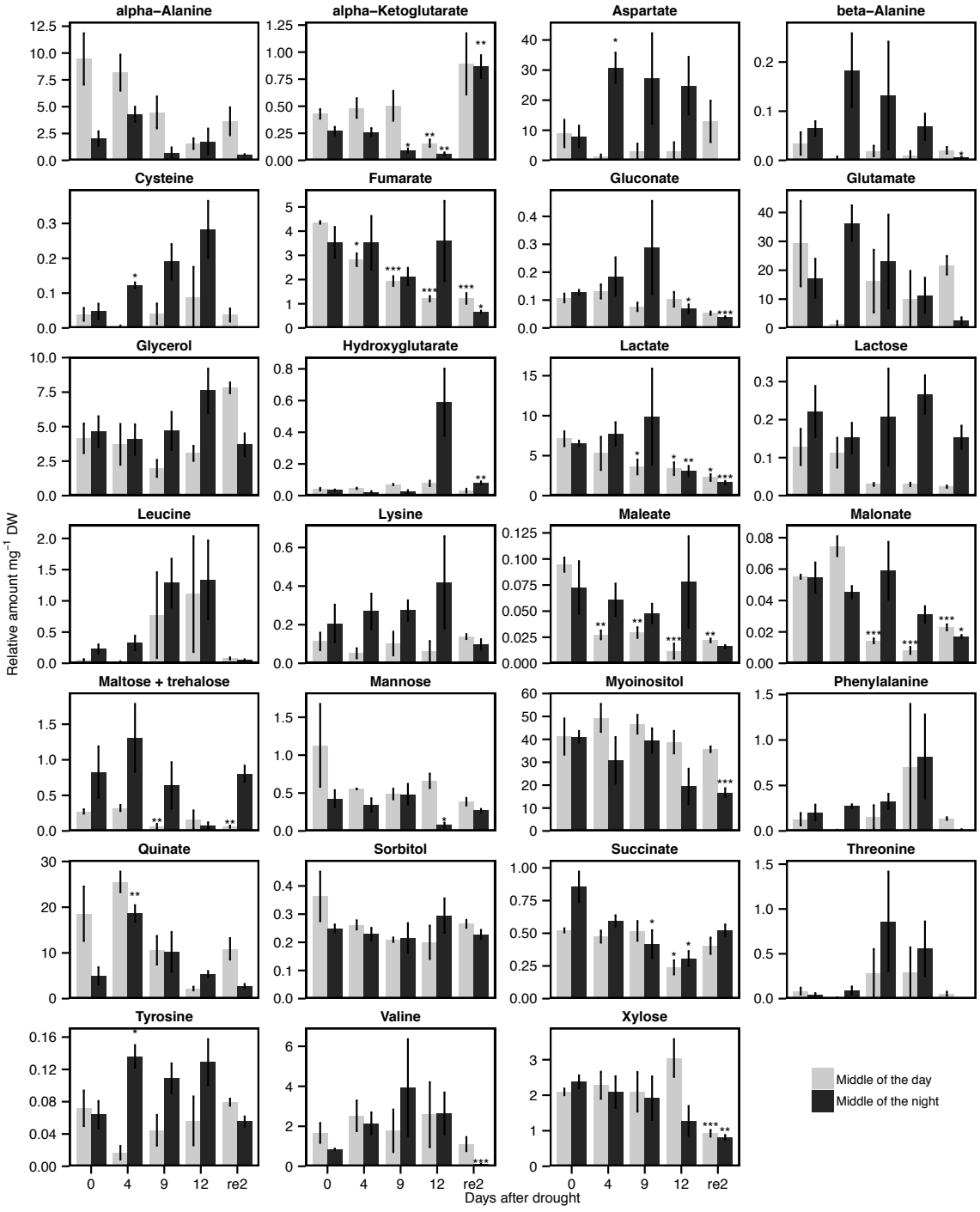
Supplemental Figure S8. Overlap of differentially expressed genes during CAM in *T. triangulare* with hormone treated leaves of *A. thaliana* (Goda et al., 2008). Bars represent the percentage of *T. triangulare* genes that are upregulated (A) or downregulated (B) on day 12 compared to day 0 (DESeq2, $q < 0.01$) in response to the respective hormone treatment in *A. thaliana*. Dashed line indicates the expected value (i.e. average of upregulated and downregulated genes on day 12, respectively). Asterisks indicate significant enrichment compared to the average response (Fisher's exact test, $***P < 0.001$). Numbers to the right represent the total number of responsive genes in *A. thaliana*. Details see Materials and Methods and Supplemental Dataset S1.



Supplemental Figure S9. Mapman overview of metabolism for ABA-response in *A. thailana.* Details see Materials and Methods.



Supplemental Figure S10. Experiments to assay activity of NADP-malic enzyme (panels 1-4) or malate dehydrogenase (panel 5) with leaf extracts of *T. triangulare* ("*Talinum* extract") or recombinant NADP-malic enzyme from *Zea mays* (ZmNADP-ME, provided by Anastasiia Bovdilova, group of Veronica G. Maurino). Leave extraction and enzyme assays were performed as described in Ashton et al. (1990). NADP-ME activity is not detectable in *Talinum* leaf extracts (panel 1). Strong activity of recombinant ZmNADP-ME (panel 2) is fully inhibited by addition of *Talinum* leaf extract (panel 3), while only slightly inhibited by boiled leaf extract of the NADP-ME C4 plant *Flaveria bidentis* ("Flaveria extract", panel 4). This indicates the formation of an inhibitor of unknown kind during the leaf extraction of *T. triangulare* leaves, which affects NADP-ME activity, but not MDH activity (panel 5).



Supplemental Figure S11. Levels of additional metabolites. Metabolites were measured by GC-MS and normalized to dry weight (DW) and internal ribitol standard (Mean ± SE, n = 3-4, Asterisks indicate Student's *t* test significance in comparison to day 0 at ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$). re2, two days after re-watering.

Supplemental Table S1. mRNA-Seq statistics

[a]Leaves were harvested from mature *T. triangulare* plants (n= 3) after 0 (Day 0), 4 (Day 4), 9 (Day 9) and 12 days (Day 12) of drought as well as after 2 d of re-watering (re2). [b]Total reads are the total number of 150 bp Illumina reads of three biological replicates per sample. [c]Mapped reads are the number of reads that could be matched to the TAIR10 reference genome of *A. thaliana*. [d]Mapping efficiency is the percentage of c/b. [e]Indicates how many genes (out of 21,869 genes of the minimal reference genome) matched at least one read. [f]Representation of TAIR10 gives the percentage of 21,869 genes of the minimal reference genome, that were hit by at least one read.

Sample[a]	Time of sampling	Replicate	Total reads[b]	Sum total reads of triplicate group	Mapped reads[c]	Sum mapped reads of triplicate group	Mapping efficiency against TAIR10[d]	Average mapping efficiency of triplicate group	Genes with at least one read[e]	Genes with at least one read by triplicate group	Represantation of TAIR10[f]
Day 0	MD	Replicate 1	15,213,727	41,908,047	8,756,533	23,688,643	58%	57%	15,192	15,746	72%
		Replicate 2	13,142,342		7,160,926		54%		15,178		
		Replicate 3	13,551,978		7,771,184		57%		15,111		
	MN	Replicate 1	14,666,449	40,863,248	6,731,139	20,025,434	46%	49%	15,423	15,903	73%
		Replicate 2	11,862,902		5,829,721		49%		15,306		
		Replicate 3	14,333,897		7,464,574		52%		15,397		
Day 4	MD	Replicate 1	13,276,380	38,821,817	7,140,769	21,441,452	54%	55%	15,219	15,728	72%
		Replicate 2	13,120,679		7,373,641		56%		15,106		
		Replicate 3	12,424,758		6,927,042		56%		15,145		
	MN	Replicate 1	9,661,691	30,562,513	4,266,915	14,318,644	44%	47%	15,049	15,660	72%
		Replicate 2	10,922,104		5,177,651		47%		15,189		
		Replicate 3	9,978,718		4,874,078		49%		15,123		
Day 9	MD	Replicate 1	14,467,089	45,931,558	6,418,919	21,819,528	44%	48%	15,004	15,645	72%
		Replicate 2	16,888,163		8,110,932		48%		15,183		
		Replicate 3	14,576,306		7,289,677		50%		14,965		
	MN	Replicate 1	14,280,800	44,283,472	7,382,494	22,471,234	52%	51%	15,111	15,677	72%
		Replicate 2	13,930,814		7,247,181		52%		15,017		
		Replicate 3	16,071,858		7,841,559		49%		15,199		
Day 12	MD	Replicate 1	11,060,708	38,449,744	5,712,636	20,015,280	52%	52%	14,620	15,431	71%
		Replicate 2	14,318,061		7,579,969		53%		14,783		
		Replicate 3	13,070,975		6,722,675		51%		14,849		
	MN	Replicate 1	14,541,469	46,449,946	6,962,404	23,231,393	48%	50%	14,951	15,638	72%
		Replicate 2	14,440,636		7,690,020		53%		14,989		
		Replicate 3	17,467,841		8,578,969		49%		15,143		
re2	MD	Replicate 1	12,772,394	40,739,535	7,433,710	22,686,620	58%	56%	15,111	15,719	72%
		Replicate 2	14,298,024		8,137,123		57%		15,151		
		Replicate 3	13,669,117		7,115,787		52%		15,202		
	MN	Replicate 1	12,994,986	38,667,014	6,714,748	20,015,370	52%	52%	15,175	15,757	72%
		Replicate 2	12,849,492		6,673,096		52%		15,185		
		Replicate 3	12,822,536		6,627,526		52%		15,246		

Supplemental Table S2. Metabolite profiling statistics

[a]Leaves were harvested from mature *T. triangulare* plants (n= 3-4) after 0 (Day 0), 4 (Day 4), 9 (Day 9) and 12 days (Day 12) of drought as well as after 2 d of re-watering (re2). Numbers indicate the fold change of metabolites with significantly different accumulation (negative values indicate decrease) compared to day 0 (Student's *t* test, $p < 0.05$). ns, not significant.

	Middle of the day				Middle of the night			
	Day 4	**Day 9**	**Day 12**	**re2**	**Day 4**	**Day 9**	**Day 12**	**re2**
Aspartate	ns	ns	ns	ns	3.9	ns	ns	ns
Citrate	1.3	3.6	ns	3.1	ns	ns	3.7	ns
Cysteine	ns	ns	ns	ns	2.6	ns	ns	ns
Fructose	ns	-12.4	-35.0	-1.9	-2.8	-4.5	-40.2	-5.4
Fumarate	-1.5	-2.2	-3.6	-3.6	ns	ns	ns	-5.3
Gluconate	ns	ns	ns	ns	ns	ns	-1.9	-3.4
Glucose	ns	-42.5	-119.5	ns	ns	ns	-16.0	ns
Glutamate	ns	ns	ns	ns	ns	ns	ns	ns
Glycerate	ns	-40.2	-238.9	ns	ns	-3.6	-16.7	ns
Glycerol	ns	ns	ns	ns	ns	ns	ns	ns
Glycine	-1.9	ns	-3.9	-2.2	ns	ns	ns	ns
Glycolate	ns	-2.2	-3.1	-1.6	ns	ns	ns	-3.3
Hydroxyglutarate	ns	ns	ns	ns	ns	ns	ns	2.4
Lactate	ns	-2.0	-2.1	-3.1	ns	ns	-2.1	-4.0
Lactose	ns	ns	ns	ns	ns	ns	ns	ns
Leucine	ns	ns	ns	ns	ns	ns	ns	ns
Lysine	ns	ns	ns	ns	ns	ns	ns	ns
Maleate	-3.5	-3.2	-8.1	-4.3	ns	ns	ns	ns
Malate	ns	-1.9	-4.0	ns	ns	ns	2.8	ns
Malonate	ns	-3.8	-6.7	-2.4	ns	ns	ns	-3.2
Maltose	ns	-4.1	ns	-4.5	ns	ns	ns	ns
Mannitol	ns	ns	ns	ns	ns	ns	ns	1.9
Mannose	ns	ns	ns	ns	ns	ns	-5.5	ns
Myoinositol	ns	ns	ns	ns	ns	ns	ns	-2.5
Phenylalanine	ns	ns	ns	ns	ns	ns	ns	ns
Proline	ns	ns	ns	ns	ns	ns	ns	-2.6
Quinate	ns	ns	ns	ns	3.8	ns	ns	ns
Raffinose	ns	14.4	12.3	ns	ns	9.2	ns	4.7
Serine	ns	ns	ns	ns	ns	ns	ns	ns
Sorbitol	ns	ns	ns	ns	ns	ns	ns	ns
Succinate	ns	ns	-2.2	ns	ns	-2.1	-2.8	ns
Sucrose	ns	-1.9	-2.9	ns	ns	-1.7	-3.6	-3.6
Threonine	ns	ns	ns	ns	ns	ns	ns	ns
Tyrosine	ns	ns	ns	ns	2.1	ns	ns	ns
Valine	ns	ns	ns	ns	ns	ns	ns	-8.4
Xylose	ns	ns	ns	-2.3	ns	ns	ns	-2.9
α-Alanine	ns	ns	ns	ns	ns	ns	ns	ns
α-Ketoglutarate	ns	ns	-2.7	ns	ns	-3.0	-4.5	3.2
β-Alanine	ns	ns	ns	ns	ns	ns	ns	-12.2

Supplemental Table S3. Expression levels of *T. triangulare* contigs encoding PEPC

Contigs of *T. triangulare* plants were assembled from Illumina paired-end reads (100 bp length) as described in Materials and Methods and matched to the reference genome RefBeet-1.1 of *Beta vulgaris* (Dohm et al., 2014) and the minimal reference genome of *Arabidopsis thaliana* (TAIR10) via BlastX for annotation. The single-end Illumina reads (150 bp) were mapped to the assembled contigs. Numbers indicate mean expression (rpm, reads per million) of all contigs orthologous to *Beta vulgaris* or *Arabidopsis thaliana* PEPC isoforms at five different stages of water-availability (n = 3) in the middle of the day (MD) and the middle of the night (MN) in *T. triangulare* (bold numbers indicate significantly DEG compared to day 0, q < 0.01, DESeq2). re2, 2 days after re-watering.

Talinum triangulare Contig ID	Arabidopsis thaliana gene ID	Annotation (TAIR10)	Beta vulgaris gene ID	Annotation (RefBeet-1.1)	Day 0 MD [rpm]	Day 0 MN [rpm]	Day 4 MD [rpm]	Day 4 MN [rpm]	Day 9 MD [rpm]	Day 9 MN [rpm]	Day 12 MD [rpm]	Day 12 MN [rpm]	re2 MD [rpm]	re2 MN [rpm]
Tt14682	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.00	1.33	0.00	1.33	0.00	**11.67**	**1.00**	**54.33**	0.00	2.67
Tt14682_2	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv4_084890_ihek.t1	Phosphoenolpyruvate carboxylase 2 Short=PEPCase 2; Short=PEPC 2; EC=4.1.1.31;	27.00	20.67	22.67	19.67	**13.33**	**6.33**	**19.00**	**6.33**	35.00	**7.00**
Tt16423	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv4_084890_ihek.t1	Phosphoenolpyruvate carboxylase 2 Short=PEPCase 2; Short=PEPC 2; EC=4.1.1.31;	12.00	12.67	15.67	10.33	9.67	**3.33**	13.00	**4.33**	20.33	**5.67**
Tt22933	AT3G42628	phosphoenolpyruvate carboxylase-related / PEP carboxylase-related	Bv_55700_psaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	1.00	1.67	0.67	1.33	2.00	**7.00**	2.67	**9.33**	1.00	1.67
Tt32711	AT2G42600	phosphoenolpyruvate carboxylase 2	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.00	1.00	0.00	0.67	0.00	**10.00**	0.33	**37.00**	0.00	1.67
Tt32711_2	AT2G42600	phosphoenolpyruvate carboxylase 2	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tt33875	AT2G42600	phosphoenolpyruvate carboxylase 2	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.67	6.33	0.33	3.33	2.00	**36.33**	**4.33**	**23.67**	3.00	16.00
Tt37959	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.00	1.33	0.00	1.33	0.00	**14.33**	**1.00**	**62.33**	0.00	3.00
Tt37959_2	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.00	0.00	0.00	0.00	0.00	0.33	0.00	1.00	0.00	0.00
Tt41156	AT1G68750	phosphoenolpyruvate carboxylase 4	Bv6_138910_yout.t1	Phosphoenolpyruvate carboxylase 4 Short=PEPCase 4; Short=PEPC 4; Short=AtPPC4; EC=4.1.1.31;	1.67	3.00	1.67	2.00	1.00	**0.00**	**0.67**	**0.33**	1.33	1.00
Tt43681	AT2G42600	phosphoenolpyruvate carboxylase 2	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.00	1.00	0.00	0.33	0.00	**9.67**	**1.00**	**39.00**	0.00	1.67
Tt47449	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.33	12.33	0.67	5.33	**4.67**	**64.00**	**8.00**	**47.00**	**5.33**	31.67
Tt53912	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.00	1.00	0.00	0.33	0.00	**9.67**	0.33	**41.00**	0.00	2.00
Tt55601	AT1G68750	phosphoenolpyruvate carboxylase 4	Bv6_138910_yout.t1	Phosphoenolpyruvate carboxylase 4 Short=PEPCase 4; Short=PEPC 4; Short=AtPPC4; EC=4.1.1.31;	1.67	2.00	1.33	2.67	1.00	**0.33**	**0.67**	**0.33**	1.00	1.00
Tt63271	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	4.00	66.33	1.00	**8.67**	**65.67**	**7638.00**	**63.33**	**6462.33**	2.33	168.00
Tt9871	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.00	0.67	0.00	0.00	0.67	**5.67**	**1.33**	**5.00**	0.33	2.33
Tt9871_2	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.00	0.33	0.00	0.00	0.00	1.00	**1.00**	1.00	0.00	1.00
Tt9871_3	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tt9871_4	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.33	10.67	0.33	**1.33**	**10.00**	**1408.67**	**12.00**	**1196.33**	0.33	**39.67**
Tt9871_5	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.33	5.33	0.00	**0.33**	**4.33**	**476.67**	**4.33**	**481.33**	0.00	**15.33**
Tt9871_6	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	0.33	4.33	0.33	2.33	**47.33**	**325.00**	**334.67**	**399.00**	0.33	4.33
Tt9871_7	AT1G53310	phosphoenolpyruvate carboxylase 1	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	16.00	18.67	19.67	30.00	27.33	**44.33**	**50.00**	**72.67**	18.33	19.67
Tt9871_8	AT2G42600	phosphoenolpyruvate carboxylase 2	Bv9_215080_xeaz.t1	Phosphoenolpyruvate carboxylase 1 Short=PEPCase 1; Short=PEPC 1; EC=4.1.1.31;	300.33	357.67	356.33	591.33	491.67	641.00	**856.33**	**1198.33**	350.00	281.00

Supplemental Table S4. Expression levels of all *T. triangulare* contigs encoding for subunits of vacuolar ATP synthases

Contigs of *T. triangulare* plants were assembled from Illumina paired-end reads (100 bp length) as described in Materials and Methods. Contigs encoding V-ATPases were extracted via BlastP in protein space using protein sequences of *A. thaliana* V-ATPases identified with SUBA3 (http://suba3.plantenergy.uwa.edu.au) and retrieved from TAIR (http://arabidopsis.org). The single-end Illumina reads (150 bp) were mapped to the assembled contigs. Numbers indicate mean expression (rpm, reads per million) at five different stages of water-availability (n = 3) in the middle of the day (MD) and the middle of the night (MN) in *T. triangulare* (bold numbers indicate significantly DEG compared to day 0, q < 0.01, DESeq2). re2, 2 days after re-watering.

AGI	Description (SUBA)	**Talinum triangulare Contig ID**	**Day 0 MD [rpm]**	**Day 0 MN [rpm]**	**Day 4 MD [rpm]**	**Day 4 MN [rpm]**	**Day 9 MD [rpm]**	**Day 9 MN [rpm]**	**Day 12 MD [rpm]**	**Day 12 MN [rpm]**	**re2 MD [rpm]**	**re2 MN [rpm]**
AT1G20260.1	ATPase, V1 complex, subunit B protein;	Tt13863	81.67	151.67	87.67	142.67	67.67	139.67	73.33	155.00	87.67	139.00
AT1G76030.1	ATPase, V1 complex, subunit B protein;Encodes the vacuolar ATP synthase subunit B1. This subunit was shown to interact with the gene product of hexokinase1 (ATHXK1). This interaction, however, is solely restricted to the nucleus.	Tt13863	81.67	151.67	87.67	142.67	67.67	139.67	73.33	155.00	87.67	139.00
AT4G38510.1	ATPase, V1 complex, subunit B protein;	Tt13863	81.67	151.67	87.67	142.67	67.67	139.67	73.33	155.00	87.67	139.00
AT4G11150.1	vacuolar ATP synthase subunit E1;Encodes a vacuolar H+-ATPase subunit E isoform 1 which is required for Golgi organization and vacuole function in embryogenesis.	Tt14339	277.67	361.67	308.33	406.33	279.33	295.33	261.00	**270.67**	299.67	**257.67**
AT1G16820.1	vacuolar ATP synthase catalytic subunit-related / V-ATPase-related / vacuolar proton pump-related;	Tt23235	43.67	65.33	47.67	71.33	**29.33**	**40.00**	**30.67**	**34.00**	48.67	54.00
AT1G12840.1	vacuolar ATP synthase subunit C (VATC) / V-ATPase C subunit / vacuolar proton pump C subunit (DET3);Encodes subunit C of the vacuolar H(+)-ATPase (V-ATPase). Bound and phosphorylated by AtWNK8.	Tt33042	21.00	25.33	19.67	31.00	24.00	25.33	26.33	22.67	20.00	22.67
AT3G58730.1	vacuolar ATP synthase subunit D (VATD) / V-ATPase D subunit / vacuolar proton pump D subunit (VATPD);	Tt41117	114.00	160.33	110.00	155.33	**92.33**	**106.33**	**97.67**	**80.00**	108.33	**89.67**
AT3G42050.1	vacuolar ATP synthase subunit H family protein;	Tt43428_4	47.67	60.67	53.00	74.33	**28.67**	**30.00**	**28.67**	**30.67**	46.33	41.67
AT1G78900.1	vacuolar ATP synthase subunit A;Encodes catalytic subunit A of the vacuolar ATP synthase. Mutants are devoid of vacuolar ATPase activity as subunit A is encoded only by this gene and show strong defects in male gametophyte development and in Golgi stack morphology.	Tt49354	110.00	205.00	131.33	188.67	**81.33**	**126.00**	101.00	**127.00**	139.67	191.67
AT4G02620.1	vacuolar ATPase subunit F family protein;	Tt56663	63.67	71.00	59.67	67.00	53.33	51.00	**50.33**	**34.67**	73.33	**45.00**
AT4G23710.1	vacuolar ATP synthase subunit G2;	Tt8658	122.67	169.67	109.00	152.67	134.00	126.33	115.33	**74.67**	117.00	**95.67**
AT4G25950.1	vacuolar ATP synthase G3;V-ATPase G-subunit like protein	Tt8658	122.67	169.67	109.00	152.67	134.00	126.33	115.33	**74.67**	117.00	**95.67**

Supplemental Table S5. Expression levels of *T. triangulare* contigs encoding NADP-ME

Contigs of *T. triangulare* plants were assembled from Illumina paired-end reads (100 bp length) as described in Materials and Methods and matched to the reference genome RefBeet-1.1 of *Beta vulgaris* (Dohm et al., 2014) and the minimal reference genome of *Arabidopsis thaliana* (TAIR10) via BlastX for annotation. The single-end Illumina reads (150 bp) were mapped to the assembled contigs. Numbers indicate mean expression (rpm, reads per million) of all contigs orthologous to *Beta vulgaris* or *Arabidopsis thaliana* NADP-ME isoforms at five different stages of water-availability (n = 3) in the middle of the day (MD) and the middle of the night (MN) in *T. triangulare* (bold numbers indicate significantly DEG compared to day 0, q < 0.01, DESeq2). re2, 2 days after re-watering.

Talinum triangulare Contig ID	Arabidopsis thaliana gene ID	Annotation (TAIR10)	Beta vulgaris gene ID	Annotation (RefBeet-1.1)	Day 0 MD [rpm]	Day 0 MN [rpm]	Day 4 MD [rpm]	Day 4 MN [rpm]	Day 9 MD [rpm]	Day 9 MN [rpm]	Day 12 MD [rpm]	Day 12 MN [rpm]	re2 MD [rpm]	re2 MN [rpm]
Tt2487	AT2G19900	NADP-malic enzyme 1	Bv8_201720_sfck.t1	NADP-dependent malic enzyme Short=NADP-ME; EC=1.1.1.40;	1.00	6.00	1.67	4.00	2.67	6.33	**7.67**	3.67	4.33	7.33
Tt38957_2	AT1G79750	NADP-malic enzyme 4	Bv9u_231540_zmjw.t1	NADP-dependent malic enzyme Short=NADP-ME; EC=1.1.1.40;	107.33	115.33	100.33	50.33	**302.00**	86.33	**369.33**	**52.67**	231.33	**53.67**
Tt38957_3	AT1G79750	NADP-malic enzyme 4	Bv9u_231540_zmjw.t1	NADP-dependent malic enzyme Short=NADP-ME; EC=1.1.1.40;	105.00	313.00	110.67	271.33	**460.00**	**611.33**	**727.33**	**580.67**	149.33	417.67
Tt4199_2	AT1G79750	NADP-malic enzyme 4	Bv9u_231540_zmjw.t1	NADP-dependent malic enzyme Short=NADP-ME; EC=1.1.1.40;	30.00	40.33	36.00	17.00	**101.00**	30.33	**138.67**	**20.33**	**78.33**	21.67
Tt42649	AT2G19900	NADP-malic enzyme 1	Bv8_201720_sfck.t1	NADP-dependent malic enzyme Short=NADP-ME; EC=1.1.1.40;	1.67	5.00	1.67	3.67	**20.00**	8.33	**36.67**	7.33	**21.33**	**12.00**
Tt42649_2	AT2G19900	NADP-malic enzyme 1	Bv8_201720_sfck.t1	NADP-dependent malic enzyme Short=NADP-ME; EC=1.1.1.40;	1.00	4.67	1.67	3.67	**21.33**	7.33	**36.33**	6.67	**21.00**	**10.67**
Tt63113	AT1G79750	NADP-malic enzyme 4	Bv8_201720_sfck.t1	NADP-dependent malic enzyme Short=NADP-ME; EC=1.1.1.40;	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tt63113_2	AT1G79750	NADP-malic enzyme 4	Bv8_201720_sfck.t1	NADP-dependent malic enzyme Short=NADP-ME; EC=1.1.1.40;	1.67	2.33	2.00	1.33	1.00	1.00	**0.00**	1.33	1.33	1.00
Tt6558	AT2G19900	NADP-malic enzyme 1	Bv8_201720_sfck.t1	NADP-dependent malic enzyme Short=NADP-ME; EC=1.1.1.40;	0.67	3.33	1.00	2.67	**1.67**	2.67	**3.33**	2.00	**2.33**	3.33
Tt8204	AT5G11670	NADP-malic enzyme 2	Bv8_201720_sfck.t1	NADP-dependent malic enzyme Short=NADP-ME; EC=1.1.1.40;	0.33	1.00	0.00	1.67	0.33	2.00	**1.67**	1.00	1.33	2.33

Supplemental Table S6. Expression levels of two *T. triangulare* contigs encoding PPDK

Contigs of *T. triangulare* plants were assembled from Illumina paired-end reads (100 bp length) as described in Materials and Methods and matched to the reference genome RefBeet-1.1 of *Beta vulgaris* (Dohm et al., 2014) and the minimal reference genome of *Arabidopsis thaliana* (TAIR10) via BlastX for annotation. The single-end Illumina reads (150 bp) were mapped to the assembled contigs. Numbers indicate mean expression (rpm, reads per million) of all contigs orthologous to *Beta vulgaris* or *Arabidopsis thaliana* PPDK isoforms at five different stages of water-availability (n = 3) in the middle of the day (MD) and the middle of the night (MN) in *T. triangulare* (bold numbers indicate significantly DEG compared to day 0, q < 0.01, DESeq2). re2, 2 days after re-watering.

Talinum triangulare Contig ID	Length of encoded peptide (aa)	Arabidopsis thaliana gene ID	Annotation (TAIR10)	Beta vulgaris gene ID	Annotation (RefBeet-1.1)	Day 0 MD [rpm]	Day 0 MN [rpm]	Day 4 MD [rpm]	Day 4 MN [rpm]	Day 9 MD [rpm]	Day 9 MN [rpm]	Day 12 MD [rpm]	Day 12 MN [rpm]	re2 MD [rpm]	re2 MN [rpm]
Tt24575	964	AT4G15530	pyruvate orthophosphate dikinase	Bv1_015500_fjqs.t1	Pyruvate, phosphate dikinase, chloroplastic EC=2.7.9.1; AltName: Full=Pyruvate, orthophosphate dikinase; Flags: Precursor;	399.00	358.00	328.33	198.67	498.67	267.00	589.67	220.00	541.67	367.67
Tt26901	887	AT4G15530	pyruvate orthophosphate dikinase	Bv1_015500_fjqs.t1	Pyruvate, phosphate dikinase, chloroplastic EC=2.7.9.1; AltName: Full=Pyruvate, orthophosphate dikinase; Flags: Precursor;	160.33	1558.33	195.33	1117.67	**1759.33**	**17488.33**	**3285.33**	**12924.33**	262.33	2049.67

Supplemental Table S7. Expression levels of genes encoding for photorespiratory enzymes

Expression levels were extracted from the cross-species mapping (Supplemental Dataset S1). Numbers indicate mean expression (rpm, reads per million) at five different stages of water-availability (n = 3) in the middle of the day (MD) and the middle of the night (MN) in *T. triangulare* (bold numbers indicate significantly DEG compared to day 0, q < 0.01, DESeq2). re2, 2 days after re-watering.

Locus[a]	Annotation (TAIR10)	Day 0 MD [rpm]	Day 0 MN [rpm]	Day 4 MD [rpm]	Day 4 MN [rpm]	Day 9 MD [rpm]	Day 9 MN [rpm]	Day 12 MD [rpm]	Day 12 MN [rpm]	re2 MD [rpm]	re2 MN [rpm]
AT4G32520	serine hydroxymethyltransferase 3	34.0	43.0	38.3	34.3	**170.3**	48.3	**202.7**	**23.3**	41.7	45.3
AT3G14130	Aldolase-type TIM barrel family protein	3.3	7.0	4.7	6.7	**12.0**	**11.0**	**17.3**	**14.7**	6.0	11.0
AT2G26080	glycine decarboxylase P-protein 2	2331.3	5268.0	2261.7	4127.7	**7350.7**	4898.0	**6474.0**	3832.7	3548.3	7660.0
AT2G13360	alanine:glyoxylate aminotransferase	2080.7	3847.0	2037.3	3514.3	3563.7	**6019.3**	3226.3	**6226.7**	2304.7	5590.0
AT1G68010	hydroxypyruvate reductase	1115.3	2214.0	1146.0	2229.7	**2262.3**	**3560.3**	1772.7	**3317.3**	1070.7	2810.0
AT1G12550	D-isomer specific 2-hydroxyacid dehydrogenase family protein	27.0	78.7	25.7	91.7	**44.7**	**106.0**	**50.7**	**109.0**	35.7	105.0
AT1G11860	Glycine cleavage T-protein family	1956.7	3114.7	1903.3	3233.0	1872.0	3304.3	**1362.0**	2605.7	1711.3	3027.3
AT1G32470	Single hybrid motif superfamily protein	1755.0	2474.3	1721.3	2773.7	2480.0	2054.7	1681.0	**1386.7**	1520.3	1710.7
AT1G79870	D-isomer specific 2-hydroxyacid dehydrogenase family protein	51.3	102.7	53.7	107.7	58.0	82.3	51.7	**63.0**	65.0	81.0
AT1G80380	P-loop containing nucleoside triphosphate hydrolases superfamily protein	262.3	179.3	244.7	186.7	276.7	178.0	326.3	167.7	241.0	140.3
AT2G35120	Single hybrid motif superfamily protein	107.3	146.3	98.7	178.7	135.7	96.3	97.7	**73.3**	92.0	101.0
AT2G45630	D-isomer specific 2-hydroxyacid dehydrogenase family protein	29.3	55.0	30.0	67.3	29.7	54.3	31.0	60.0	31.7	59.3
AT3G14415	Glycolate oxidase 2	1186.7	1608.0	1292.0	1806.3	1735.3	1946.3	1385.3	1824.0	1236.0	1899.0
AT3G14420	Glycolate oxidase 1	2684.3	3855.7	2877.0	4180.3	3995.0	4692.7	3329.3	4365.3	2900.3	4609.3
AT4G17360	Formyl transferase	33.7	56.3	32.0	55.0	44.7	53.0	40.0	44.3	30.7	44.7
AT4G18360	Glycolate oxidase 3	212.0	321.3	243.7	378.0	349.0	397.7	279.7	364.7	236.0	395.0
AT5G36700	2-phosphoglycolate phosphatase 1	1340.3	366.7	1245.3	355.7	**821.0**	427.3	**624.3**	440.7	1006.7	391.0

Supplemental Table S8. Estimated partitioning of carbon in organic acids and starch on day 12

Concentrations of starch and organic acids were measured from the same plants (n = 4) but independent mature leaves at the end of the night and the end of the day, respectively. Starch was measured with an assay. Organic acids were measured via acid titration (See Fig. 1). Amount of carbon in starch was estimated as six multiplied with the concentration of glucose (accounting for six C atoms). For approximation of carbon in organic acids, acid concentrations were multiplied with three under the assumption that malate (C4) accounts for the major day-night fluctuations during CAM (Fig. 4) and subtracting one molecule of CO_2 which is released to the Calvin-Benson cycle after decarboxylation during the day (Fig. 3).

Time	Compound	Concentration	Carbon
End of the day	Starch	20.8 µmol Glc / g FW	125
	Acid	0 µmol H+/ g FW	0
End of the night	Starch	6.4 µmol Glc / g FW	38
	Acid	31.5 µmol H+/ g FW	95

Manuscript 1: Reversible Burst of Transcriptional Changes during Induction of Crassulacean Acid Metabolism in *Talinum triangulare*

Status: Published (Januar 2016)

Authors: Brilhaus, D., Bräutigam, A., Mettler-Altmann, T., Winter, K., and Weber, A.P.M.

Journal: "Plant Physiology"

Author contributions

D.B., A.B., T.M.-A., K.W., and A.P.M.W. designed the research.

D.B. and K.W. conducted the drought experiment.

D.B. and T.M.-A. performed the metabolite analysis.

D.B. sampled and prepared leaves for all experiments, extracted RNA, metabolites and enzymes, prepared Illumina libraries, conducted enzyme assays and measurement of starch content, graphically presented all data and drafted the manuscript.

D.B., A.B., and T.M.-A. analyzed data

D.B., A.B., T.M.-A., K.W., and A.P.M.W. wrote the final article.

Manuscript 2

Diurnal and circadian leaf transcriptomes reveal enhanced light-responsive gene expression during the evolution of C_4

TITLE

Diurnal and circadian leaf transcriptomes reveal enhanced light-responsive gene expression during the evolution of C_4

Dominik Brilhaus, Thea R. Pick, Tabea Mettler-Altmann, Andreas P. M. Weber

Institute of Plant Biochemistry, Center of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine-University, Universitätsstraße 1, D-40225 Düsseldorf, Germany (D.B., T. M.-A., A. P. M. W.)
Joint BioEnergy Institute, Emeryville, California 94608, USA (T. R. P.)

dominik.brilhaus@hhu.de
tabea.mettler@hhu.de
trpick@lbl.gov
andreas.weber@hhu.de

Corresponding author:
Andreas P.M. Weber,
Tel.: +49 211 81 12347
Fax: +49 211 81 13706
E-mail address: andreas.weber@hhu.de

Running title: Diurnal and circadian transcriptome of C_3 and C_4 *Flaveriaceae*

ABSTRACT

C_4 plants are of special interest to overcome global challenges of the current century such as growing demands for food and energy due to their high efficiency to use light, water, and nitrogen. Approaches to transgenically engineer the complex trait C_4 photosynthesis into C_3 plants depend on a comprehensive knowledge of all proteins involved. Extensive studies elucidated the molecular identity of most core C_4 enzymes and transporters, as well as their regulation of spatial, temporal, and (often) light-driven expression. However, our understanding of the adaptation of diurnal and circadian clock regulated expression of genes directly or indirectly involved in the core C_4 network is currently incomplete. How is gene expression C_4-specifically controlled by light or the circadian clock?

We compared the changes in leaf transcriptomes of two closely related *Flaveria* species employing C_4 (*F. bidentis*) or C_3 (*F. robusta*) photosynthesis under diurnal (light-dark) and circadian (constant light) conditions. In both species approximately two thirds of the transcriptome were periodically expressed under diurnal conditions, while circadian regulation appeared to be limited to less than 6% of the genes.

Transcript abundances of the C_4 cycle genes were markedly increased in *F. bidentis* compared to *F. robusta*. However, diurnal and circadian expression patterns as well as periodicity of most C_4 cycle genes largely overlapped between both species, raising the hypothesis that acquisition of transcriptional light-regulation predates the evolution of C_4 in the *Flaveriaceae.* Many photosynthesis related genes were under stronger circadian control in *F. robusta*, while light-regulated in *F. bidentis*. We identified candidate genes for the regulation of an enhanced light-response in the C_4 plant and discuss possible advantages. These genes represent good candidates for the fine tuning of gene expression in C_4 plants and may support efforts to engineer C_4 photosynthesis into C_3 plants.

INTRODUCTION

C_4 photosynthesis is a CO_2 concentrating mechanism that evolved as an adaptation to, hot, arid environments (Sage et al., 2012). The metabolic reactions of C_4 are typically partitioned across two types of cells, the mesophyll cell (MC) and the bundle sheath cell (BSC), which are positioned around the leaf veins in a wreath-like anatomical arrangement termed Kranz anatomy. Phospho*enol*pyruvate (PEP) carboxylase (PEPC) fixes HCO_3^- into oxaloacetate (OAA) in the MC, which is then reduced to one of the C_4 acids malate or aspartate. The C_4 acid is shuttled to the BSC, where it is decarboxylated by either of three decarboxylases. The released CO_2 is then finally fixed via the Calvin-Benson-Cycle (CBC) by RUBISCO (ribulose-1,5-*bis*phosphate (RuBP) carboxylase/oxygenase) (Hatch, 1987). Based on the kind of decarboxylase (NADP-ME, NAD-ME or PEP-CK) employed, three biochemical subtypes of C_4 were traditionally distinguished (Furbank, 2011). Due to CO_2 enrichment in the vicinity of RUBISCO, photorespiration is suppressed and the synthesis of CBC and photorespiratory enzymes is reduced in C_4 plants, resulting in an enhanced nitrogen use efficiency compared to C_3 plants (Long, 1999). Stomatal conductance is also reduced, since PEPC has a higher affinity for CO_2 than RUBISCO and does not react with oxygen (Hatch, 1987), thus resulting in an enhanced water use efficiency (Ehleringer and Monson, 1993).

The suite of enzymes and transporters required to run the C_4 cycle were already present in ancestral C_3 plants and were recruited for this new context via gene duplication, neofunctionalization of isogenes and increased or changed cell-type specific expression (Monson, 2003; Bräutigam et al., 2011; Gowik et al., 2011). The allocation of key enzymes of photorespiration to the BSC evidenced in extant C_3-C_4 species within several plant lineages was suggested as an intermediate step on the evolutionary path from C_3 to C_4 photosynthesis (Sage et al., 2012; Mallmann et al., 2014; Schulze et al., 2013; 2016; Sage, 2004). In addition to photorespiration, the RUBISCO and a large part of the CBC are BSC-specifically expressed, and the activity of photosystem II is reduced in this compartment (Meierhoff and Westhoff, 1993; Höfer et al., 1992).

Although C_4 photosynthesis is regulated on the post-transcriptional, post-translational and epigenetic level in addition to the transcriptional level (Sheen, 1999; Hibberd and Covshoff, 2010), the transcriptional up-regulation and cell-specific expression of C_4 cycle enzymes was key to the evolution of the C_4 syndrome as evidenced in accumulating comparative transcriptomes of dicotyledonous and monocotyledonous C_4 species (Bräutigam et al., 2011; Gowik et al., 2011; Wang et al., 2014a; Li and Brutnell, 2011; Aubry et al., 2014). However, the

search for a "master switch" transcription factor, the increased expression of which would increase expression of multiple C_4 cycle genes in parallel, has so far yielded only few promising candidates (Westhoff and Gowik, 2010; Langdale, 2011; Wang et al., 2013).

The *Asteraceae* genus *Flaveria* comprises closely-related species performing C_3, C_3-C_4, or C_4 photosynthesis and has in the past decades been explored for studies of several aspects of C_4 gene expression and regulation. *Flaveria* C_4 plants employ the most frequently found NADP-ME subtype (Drincovich et al., 1998) and the transformability of the C_4 species *F. bidentis* facilitated the establishment of *Flaveria* as a model C_4 genus (Chitty et al., 1994). Several studies with *Flaveria* have elucidated the cell-type specific and light-dependent accumulation of C_4 cycle proteins and transcripts. Alteration of cis-acting elements in the *Ppc* promoter induces MC-specific expression of PEPC in the C_4 plant *F. trinervia* (Stockhaus et al., 1997; Gowik et al., 2004). Accumulation of NADP-ME1 is light-dependent and correlates with C_4 photosynthesis in *F. bidentis* (Marshall et al., 1996). In *F. trinervia*, BSC-specific and light-dependent mRNA accumulation of *NADP-ME1* correlating with photosynthetic activity are both encoded in the promoter (Lai et al., 2002). In this study, light was shown to both enhance expression in BSC and suppress expression in the MC. In addition to the promoter region, sequence parts of its 3' untranslated region and downstream of it are required for high levels of *NADP-ME* in the BSC (Lai et al., 2002). The accumulation of *PPDK* and *PPCK* mRNAs was also shown to be under light control in *F. trinervia* (Rosche and Westhoff, 1995; Tsuchida et al., 2001; Aldous et al., 2014). Besides the light-dependent mRNA accumulation, differential diurnal phases of mRNA abundances between two PPCK isoforms in both C_3 and C_4 *Flaveria spp.* were shown (Aldous et al., 2014). This study raised the hypothesis for differential light-regulated expression and involvement of the circadian clock in the transcriptional regulation of C_4 photosynthesis. Circadian clock regulation of *PPCK* expression has so far only been evidenced in CAM and C_3 plants (Taybi et al., 2000; Sullivan et al., 2005). Beside core C_4 cycle enzymes, thorough studies on the promoter region of the key photorespiratory protein GDC-P in several *Flaveria spp.* gained insights on BSC-specific gene regulation (Wiludda et al., 2012; Schulze et al., 2013).

Light-regulation of gene expression is dependent on the perception of light via photoreceptor molecules. Four families of photoreceptors are distinguished in plants (Chen et al., 2004). Two blue-light sensing phototropins, which sense light via an FMN (flavin mononucleotide) molecule tightly bound to the so-called LOV (Light, Oxygen, Voltage) domain, were found to regulate photosynthesis, phototropism, chloroplast movements and stomatal opening in *Arabidopsis thaliana* (Christie and Briggs, 2001), however direct involvement in

transcriptional regulation has not been shown in plants. Another family of blue-light sensing, LOV-domain containing photoreceptors, comprise of ZEITLUPE (ZTL), FLAVIN-BINDING, KELCH REPEAT, F-BOX 1 (FKF1), and LOV KELCH PROTEIN 2 (LKP2), and are hence termed the ZTL/FKF1/LKP2 group (Zoltowski and Imaizumi, 2013). These proteins are involved in the control of flowering time as well as sensing and transduction of photoperiod to the circadian clock (Imaizumi et al., 2003; Zoltowski and Imaizumi, 2013), their gene regulatory function is thus secondary through the action of the clock. Cryptochromes (CRYs), a third class of UV-A/blue light photoreceptors, and Phytochromes (PHYs), red light/far red light receptors, undergo a conformational change upon photoexcitation and directly regulate gene transcription via binding of specific transcription factors. CRYs guide the transcriptional reprogramming during seedling de-etiolation (Zoltowski and Imaizumi, 2013), while PHYs regulate genes encoding for proteins involved in photosynthesis, flavonoid biosynthesis, photorespiration, and other parts of metabolisms (Quail, 2002). In addition, CRYs and PHYs work in concert and also transduce photoperiod to the circadian clock. The best studied transcription factors to interact with PHYs are the PHY interacting factors (PIFs), which bind to specific *cis*-regulatory elements comprising a G-box motif. More specifically related to C_4, (Cockburn et al., 1996) described, that PHY is involved in light regulation of PEPC transcription in the CAM plant *Mesembryanthemum crystallinum*. However, evidence for PHY-specific induction of C_4 cycle genes expression is currently unavailable. The interplay of photoreceptors specialized to sense different light qualities (blue, red, far-red) and quantities allow the plant to perceive both daytime and season.

In addition to the directly light/dark dependent diurnal changes of gene expression, a complex anticipatory signaling network, termed the circadian clock is, described in all domains of life with the common properties of self-sustaining periodicity, entrainment and temperature compensation (McClung, 2006). The plant circadian clock is based on interlocked feedback loops operating on the levels of transcription, post-transcription, post-translational modification and protein turnover (Harmer, 2009). These loops maintain the proper timing of expression of downstream target genes over the course of the day (Greenham and McClung, 2015 and references therein). Targets of daily and seasonal transcriptional regulation include photosynthesis, growth, flowering time, hormone metabolism, abiotic and biotic stress response (Gehan et al., 2015; Covington et al., 2008; Michael et al., 2008). Their anticipatory regulation grants a fitness advantage (Dodd et al., 2005).

To gain insight on the C_4-specific regulation of gene expression by light or the circadian clock, we compared the changes in leaf transcriptomes of two closely related *Flaveria* species

employing C_4 (*F. bidentis*) or C_3 (*F. robusta*) photosynthesis under light / dark cycle ("diurnal") followed by two days under constant light ("circadian").

MATERIALS AND METHODS

Plant Material and Growth Conditions

Flaveria spp. plants were grown for three weeks in the green house in soil (C-400 with Cocopor [Stender Erden] fertilized with 3 g/L Osmocote exact standard 3 to 4 M [Scotts]). In order to synchronize gene expression between all plants of both species, the plants were transferred to controlled growth chambers (Percival Scientific) maintained under 10 h light (22 °C, 50% relative humidity) / 14 h dark (19 °C) cycles for seven days prior to the experiment. In the light, photon flux density at leaf level was 120 µmol m^{-2} s^{-1}. These conditions were maintained during the first two days of the experiment (i.e. "diurnal"). During the subsequent two days the, the light was constantly on (i.e. "circadian").

RNA Extraction, Preparation and Sequencing of Illumina Libraries

Every four hours, the four uppermost leaves of two biological replicates were harvested, pooled and immediately frozen in liquid nitrogen, providing representative leaves over several developmental stages. RNA was isolated from ground tissue using the RNeasy mini plant kit (Qiagen). Residues of DNA were removed with DNase (New England Biolabs). RNA integrity, sequencing library and fragment size were analyzed on a 2100 Bioanalyzer (Agilent). Libraries were prepared using the TruSeq RNA Sample Prep Kit v2 (Illumina), and quantified with a Qubit 2.0 (Invitrogen). Samples were multiplexed with 12 libraries per lane and sequenced in single-end mode (Rapid Run, 100 bp read length) on an Illumina HiSeq 2000 platform, yielding ~17 million (*F. bidentis*) and ~16 million (*F. robusta*) reads per library.

Read Mapping and Gene Expression Profiling

Illumina reads were aligned to a minimal set of coding sequences of the TAIR 10 release of the *A. thaliana* genome (http://www.arabidopsis.org/) using BLAT (Kent, 2002) in protein space. The minimal transcriptome was obtained as described in Gowik et al. (2011) and contained 21,869 nuclear encoded protein-coding genes. The best BLAT hit for each read was determined by (i) lowest e-value and (ii) highest bit score. Raw read counts were transformed to reads per million (rpm) to normalize for the number of reads available at each sampling stage. Cross-species mapping takes advantage of the completeness and annotation of the *A. thaliana*

genome and overcomes the limitations of transcriptome assembly (Franssen et al., 2011; Schliesky et al., 2012).

Data Analysis

Data analysis was performed using R statistics software (R version 3.2.4 provided by the CRAN project, http://www.R-project.org). For statistical analysis, the complete dataset was subset into "diurnal" and "circadian" as depicted in Figure 1. Note that the last time point of the diurnal and the first time point of circadian are the same, in order to have two full cycles with 13 time points for each subset. Principle component analysis was performed using the function prcomp implemented in R on normalized transcript abundances (reads per million), $\log_2$-transformed (after adding a pseudo-count of 1) and z-score scaled by species. For differential gene expression, two time points of peak expression were extracted for each species in each subset (diurnal and circadian) based on averages of normalized reads (rpm), resulting in four independent samples. Differential gene expression was analyzed with the DESeq2 package (Love et al., 2014) in default mode, contrasting *F. bidentis* against *F. robusta*. Automatic independent filtering retained 15,655 (diurnal) and 15,548 (circadian) genes with nonzero total read count. An alpha of 0.01 was applied after p-value adjustment with false discovery rate via Benjamini-Hochberg correction (Benjamini and Hochberg, 1995). Analysis of periodicity was performed using the R implementation of JTK_cycle (Hughes et al., 2010) on normalized reads (rpm) of biological replicates with the default period input, seeking for genes with 20, 24 or 28 hours periods. Note, applying a denser period input of only 24 hours resulted in a marked reduction of genes that were identified as periodically cycling exclusively in the circadian subset of *F. robusta* (data not shown), which was the reason to choose the broader range of periods, albeit a 24-hour period would probably fit better to the 24-hour entrainment set by the experiment. We found, that the vast majority of genes periodically cycling in the circadian *F. robusta* subset were assigned a 20-hour period, which might explain the high accumulation of genes with a phase of ZT10 in this subset (Supplemental Fig. S1). If not specifically indicated, data was annotated and categorized with publicly available information for *A. thaliana* (TAIR10, http://www.arabidopsis.org/). Full quantitative data with annotation is available as Supplemental Dataset 1.

Accession number

The read data have been submitted to the National Center for Biotechnology Information Gene Expression Omnibus under accession number GSE74768. The following link has been created to allow review of record GSE74768, while it remains in private status:
http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE74768

RESULTS

Large and shared proportions of the leaf transcriptomes are diurnally cycling

Leaf samples for mRNA-Seq analysis were collected every four hours in four consecutive days from *Flaveria* plants grown under light-dark for two days ("diurnal") followed by two days of constant light ("circadian") (Fig. 1A). Over the complete time course, biological duplicates yielded on average of 34.1 million and 32.9 million reads for *F. bidentis* and *F. robusta*, respectively (Supplemental Table S1). These reads mapped to 15,344 (*F. bidentis*) and 15,449 (*F. robusta*) genes of the reduced *A. thaliana* TAIR10 reference genome with at least one read, resulting in an average mapping efficiency of 53% and a representation of approximately 66% of the TAIR10 genome. This is comparable to the previously published cross-species mapping of *Flaveria* reads against *Arabidopsis* TAIR9 via the same pipeline (Gowik et al., 2011; Mallmann et al., 2014). For statistical analyses the dataset was divided into four parts (*F. bidentis* diurnal and circadian, *F. robusta* diurnal and circadian).

Principle component analysis of the separated diurnal and circadian subsets revealed the species as the major source of variance (principle component 1, Supplemental Fig. S2). In the diurnal subset, individual samples of the two cycles clustered closely together by the time of sampling for each species and in the same manner between the two species resulting in a divergent clustering. By contrast, the clustering of samples yielded by the consecutive cycles under circadian conditions was less divergent.

Under diurnal conditions, 9,835 *F. bidentis* genes (64.5% of the transcriptome) and 9,467 *F. robusta* genes (61.8% of the transcriptome) were identified as periodically cycling (JTK_cycle V 3.1, $q < 0.05$, Hughes et al., 2010), of which 7,555 were shared amongst the two species (Fig. 1B and Supplemental Dataset S2). In both species, the expression of a large proportion (59.9% in *F. bidentis* and 65.5% in *F. robusta*) of significantly periodic genes peaked (i.e. "phase") during the middle of the night (Fig. 1C, zeitgeber time (ZT) 14 to 18). Comparison of the phases across the two species revealed largely overlapping timing of expression (Fig. 1D). Strikingly, only 532 (*F. bidentis*) and 882 (*F. robusta*) genes were identified as periodic in constant light (Supplemental Fig. S2 and Supplemental Dataset S3). The distribution of their phases was more divergent than under diurnal conditions. The fact that more than 50% of the periodic genes in *F. robusta* were phased to ZT10 may be an artifact of the JTK_cycle analysis (as discussed in Materials and Methods). Still, the phases of the 116 genes that are periodic in both species overlapped largely between the species.

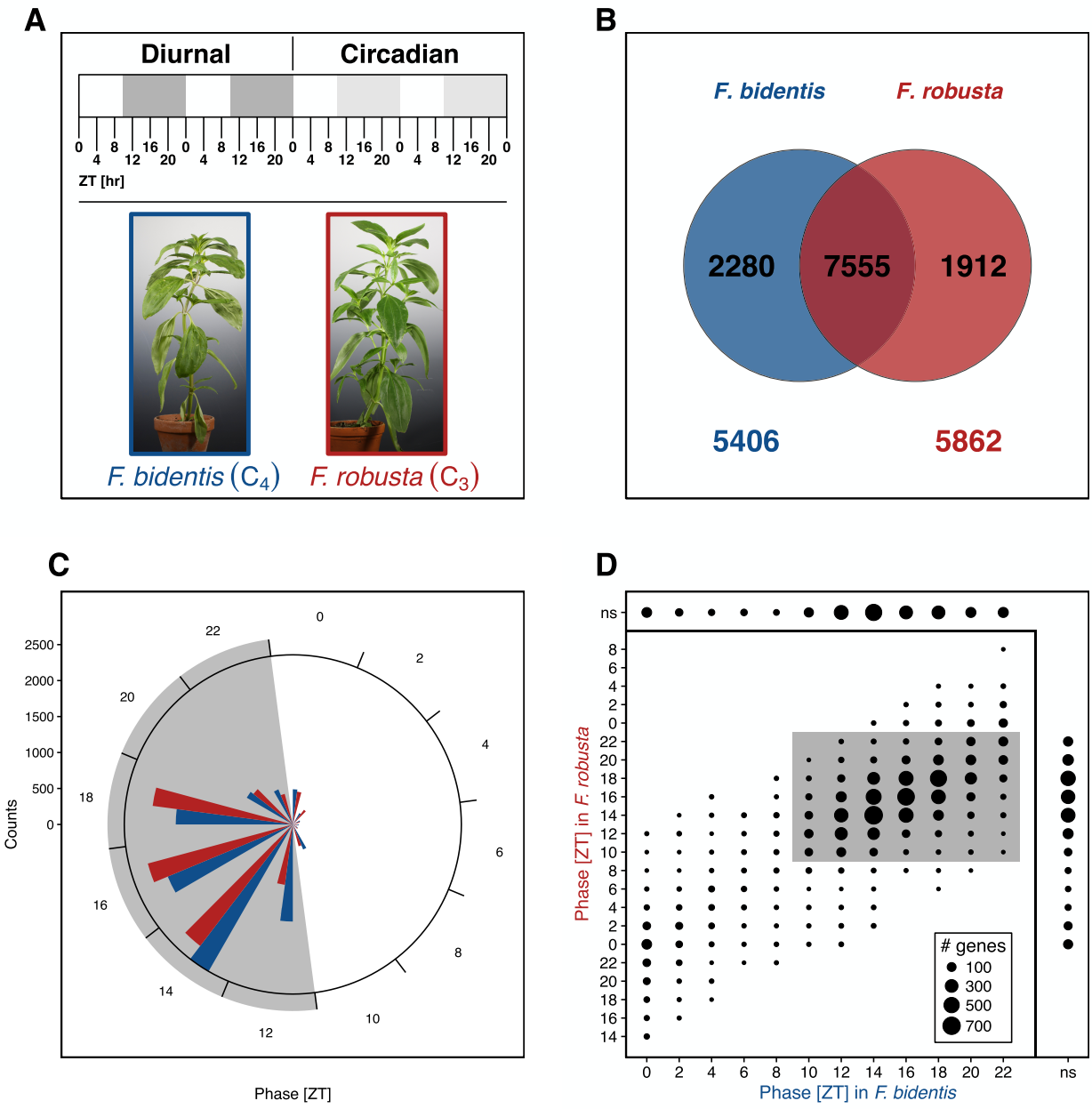


Fig. 1. Experimental design and periodicity analysis of the diurnal dataset. A, Time course and representative pictures of plants during the experiment. Leaves were collected at the time points indicated in zeitgeber time (ZT) during two consecutive cycles each under light-dark ("diurnal") and constant light ("circadian") conditions. B-D, results of JTK_cycle analysis of the diurnal subset (analogous results of the circadian subset are presented in Supplemental Fig. S2). B, Venn diagram of genes identified as periodic (JTK_cycle, $q < 0.05$) in the two *Flaveria* species under diurnal conditions. Numbers outside the venn diagram represent non-periodic genes. C, Phase of expression of significantly periodic genes (JTK_cycle, $q < 0.05$). D, Comparison of phases of significantly periodic genes between the two *Flaveria* species. ns, no significant periodicity. Shaded areas in A, C and D indicate night or subjective night (panel A, "circadian").

Phase adjusted analysis of differential gene expression

Although, the phases of 7,555 diurnally periodic genes in both species were largely overlapping (Fig. 1D), the phases of 20.8% of these genes were shifted by four or more hours between the two species, raising the possibility of under- and overestimating the proportion of differentially expressed genes (DEG) in single time point experiments (Fig. 2A and Hsu and Harmer, 2012). For each gene, species and subset the transcript abundances at two peak time points (ergo biological quadruplicates) were extracted and used for statistical analysis of differential gene expression (DESeq2, $q < 0.01$, Love et al., 2014) between the two *Flaveria* species. This approach resulted in the identification of 4913 significantly DEG in the diurnal subset, 2,412 genes with higher transcript abundance in *F. bidentis* and 2,501 genes with higher transcript abundance in *F. robusta* (Supplemental Dataset S4). Gowik et al. (2011) identified 200 and 410 significantly more abundant as well as 263 and 585 less abundant transcripts in their comparison of *Flaveria* C_4 species (*F. bidentis* and *F. trinervia)* against their closely related C_3 species (*F. pringlei* and *F. robusta*), respectively (Fig. 2B). After removing 19 genes with trends contrasting (ambiguous C_3 or C_4 specificity) between the two-species comparison, the remaining DEG 1,207 (489 C_4 and 718 C_3 specific) were compared to results of the current study (Fig. 2C). While a large proportion of previously identified DEG, 43% of C_4 specific and 39% of C_3 specific, were confirmed with this approach, we found conflicting results for 80 genes, i.e. genes that were shown to be more abundant in the diurnal data in *F. bidentis,* while previously identified as less abundant in C_4 and vice versa. In addition to the previously identified DEGs, we identified 2,152 and 2,190 genes with higher transcript abundances in *F. bidentis* and *F. robusta*, respectively. The same approach was used to analyze differential expression under circadian conditions yielding 2,817 DEGs, 1,378 enriched in *F. bidentis* and 1,439 enriched in *F. robusta*. The results of the circadian DEGs overlapped largely with those of the diurnal DEGs (Supplemental Fig. S3), 1,071 (*F. bidentis*) and 1,158 (*F. robusta*) genes with higher expression under diurnal conditions were still more highly expressed under circadian conditions. Additionally, 306 and 270 genes in *F. bidentis* and *F. robusta* gained significantly higher transcript abundances under circadian conditions and 12 genes showed contrasting results.

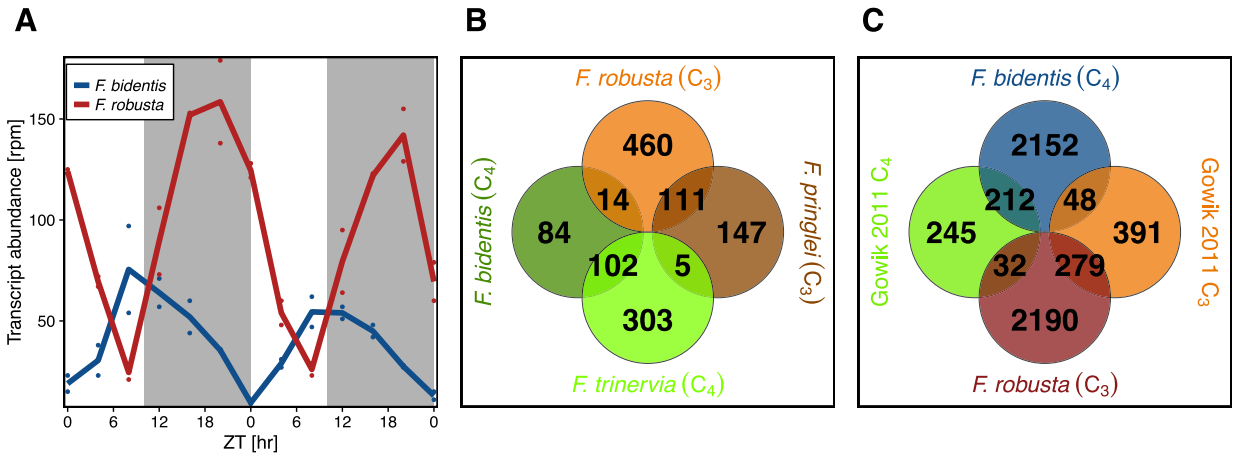


Fig. 2. Phase-adjusted analysis of differential gene expression. A, Transcript abundances under diurnal conditions of an exemplary gene (At2g42690) with differential phases between the two species. Differential expression is only detected when regarding the expression over the whole diurnal course rather than at midday only (e.g. ZT 4 to 8). B, Venn diagram of differential gene expression between *Flaveria* sister pairs (reproduced from Gowik et al., 2011). In Gowik et al. (2011), differential gene expression was determined by comparison of *F. bidentis* against *F. pringlei* and *F. trinervia* against *F. robusta*. The numbers represent genes with significantly higher expression in the indicated species compared to its sister species. C, Venn diagram comparing the results of phase-adjusted differential expression of the diurnal subset with the results of Gowik et al. (2011). The results presented in panel B, where summarized as C_3 and C_4 specific expression, after removal of contrasting results. See text for explanation.

Diurnal and circadian expression of the C_4-cycle

In good agreement with previous *Flaveria* transcriptome studies (Gowik et al., 2011; Mallmann et al., 2014), genes encoding for core C_4-cycle enzymes and transporters of the NADP-ME type were amongst the most highly significantly upregulated genes in *F. bidentis* (Fig. 3 and Supplemental Dataset S4). The genes encoding for enzymes involved in the initial carboxylation reactions in the BSC, BETA CARBONIC ANHYDRASE 3 (BCA3), PHOSPHO*ENOL*PYRUVATE CARBOXYLASE 1 (PPC1) and its regulating kinase PEP CARBOXYLASE KINASE 1 (PPCK1) showed two-fold, 34-fold and seven-fold higher transcript abundances in *F. bidentis* under diurnal conditions (Supplemental Dataset S5). Transcripts encoding for the plastidial decarboxylase NADP-MALIC ENZYME 4 (NADP-ME4) were enriched ten-fold in *F. bidentis*, while the alternative decarboxylases, PHOSPHO*ENOL*PYRUVATE CARBOXYKINASE 1 (PEP-CK1) and NAD-MALIC ENZYME 1 (NAD-ME1) were two-fold and three-fold enriched in the C_3 species *F. robusta*. Transcripts of NADP-DEPENDENT MALATE DEHYDROGENASE (NADP-MDH) were eight-fold more abundant in *F. bidentis*. Transcripts encoding for transaminases, which allow for alternating shuttling of aspartate and alanine in addition to malate and pyruvate between the BSC and MC,

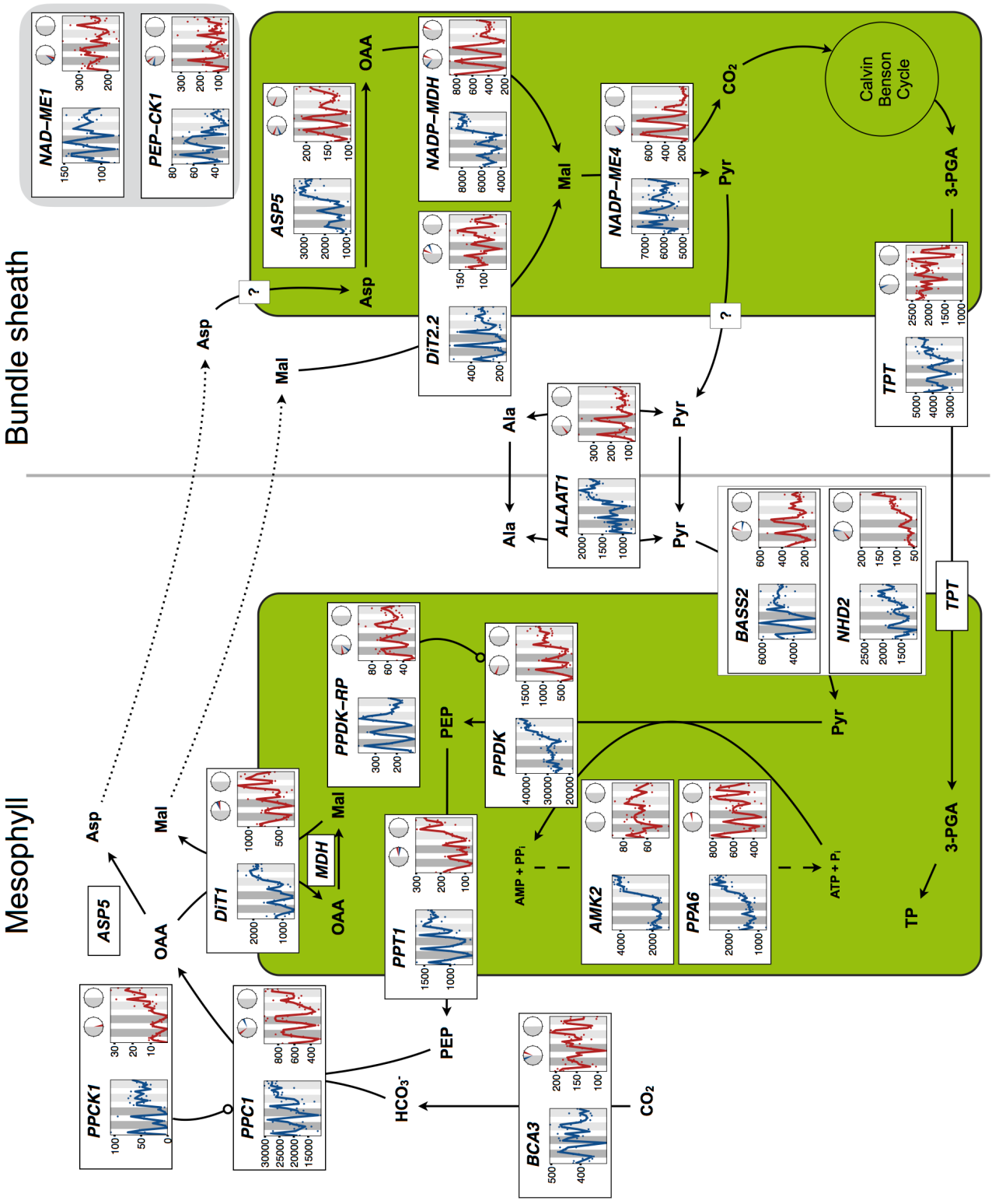
ALANINE AMINOTRANSFERASE (ALAAT1) and ASPARTATE AMINOTRANSFERASE 5 (ASP5), were five-fold and ten-fold more abundant in *F. bidentis*. The C_4-cycle is replenished with PEP in the BSC via PYRUVATE, *ORTHO*PHOSPHATE DIKINASE (PPDK), which shows 27-fold higher transcript abundances in *F. bidentis*, and is overall the second most highly expressed gene in *F. bidentis* ranging from 23,477 to 43,061 rpm (reached after two days of constant light). The PPDK REGULATORY PROTEIN (PPDK-RP) was five-fold enriched. In addition, transcripts encoding for enzymes catalyzing the reconversion of pyrophosphate and AMP accumulating during the conversion of pyruvate to PEP, ADENYLATE KINASE 2 (AMK2) and PYROPHOSPHORYLASE 6 (PPA6) were 29-fold and 2-fold more abundant in the C_4 species. Besides the soluble proteins, all known transport proteins involved in the C_4-cycle, PEP/PHOSPHATE TRANSLOCATOR 1 (PPT1, eight-fold), DICARBOXYLATE TRANSPORTER 1 (DiT1, two-fold), DICARBOXYLATE TRANSPORTER 2.2 (DiT2.2, three-fold), SODIUM-DEPENDENT PYRUVATE TRANSPORTER 2 (BASS2, twelve-fold) and TRIOSE-PHOSPHATE/PHOSPHATE TRANSLOCATOR (TPT, two-fold) had higher transcript abundances in *F. bidentis*. The gene encoding for SODIUM:HYDROGEN ANTIPORTER 1 (NHD1), which works in concert with BASS2, was not included in the reference genome. Consequently, transcripts encoding for NHD1 likely mapped to NHD2, which was 21-fold enriched in *F. bidentis.* Under circadian conditions, transcripts encoding for all of the abovementioned 17 C_4 genes (excl. NAD-ME1 and PEP-CK1) were still significantly more abundant in *F. bidentis* and their ranges of expression and fold-changes compared to *F. robusta* closely resembled those under diurnal conditions (Fig. 3 and Supplemental Dataset S4).

Transcripts encoding for eight C_4-cycle enzymes, NADP-ME4, NAD-ME1, PEP-CK1, BCA3, DiT1, PPT1, ASP5 and PPDK-RP were identified as significantly periodic and expressed at similar phases in both species under diurnal conditions. The expression of PPDK, ALAAT1 and PPA6 was exclusively periodic in *F. robusta*, while their expression only moderately changed over the diurnal course in *F. bidentis*. PPCK1 was also only identified as significantly periodic in *F. robusta* with peak expression at ZT10. However, manual inspection of the expression pattern in *F. bidentis* suggested periodicity similar to that in *F. robusta* with peak expression at ZT8. Transcripts mapping against another PPCK isogene, namely PPCK2, were likewise only detected as periodic in *F. robusta*, but with peak expression in the night at ZT14, while repeatedly peaking at ZT8 in *F. bidentis* (Supplemental Fig. S4). While the phase of expression of DiT2.2 and BASS2 was delayed compared to *F. robusta* by approximately four and six hours with peak expression post-dawn in *F. bidentis*, the phase of NADP-MDH was advanced by four hours to peak post-dusk. The most drastic differences in phase between the

species were identified for NHD2 (10 hours later in *F. bidentis*, peaking pre-dawn) and PPC1 (8 hours later in *F. bidentis*, peaking post-dawn at ZT4). The same phases were identified for transcripts mapping against another PPC isoform, PPC2, which was also highly upregulated in C_4 (Supplemental Fig. S4). The expression of TPT was only significantly periodic in *F. bidentis.* No significant periodic expression of AMK2 was detected in either species, but transcript abundances repeatedly peaked at approximately ZT10 in *F. bidentis*. In summary, a large proportion of the C_4-cycle enzymes was expressed in a diurnally periodic pattern and the majority indicated equal phase in both species.

After the transition to constant light, the average expression levels over all time points of all C_4-cycle genes was either sustained or increased compared to the diurnal conditions in *F. bidentis* (Fig. 3 and Supplemental Dataset S4). Largest increases of transcript abundances were observed for *PPCK1* (2-fold), *AMK2* (1.9-fold), *DiT1* and *ASP5* (1.7-fold), *PPA6* (1.5-fold), *NADP-MDH* and *ALAAT1* (1.4-fold). In constant light, constant expression of these genes markedly surpassed their peak levels under diurnal conditions. By contrast, with the exception of *PPCK1* (2.6-fold increase in constant light), *NHD2* (1.6-fold increase), *PPDK* and *PPT1* (both 1.5-fold increase), expression of C_4-cycle genes was either maintained or slightly depleted in *F. robusta.* In addition, although statistically only captured for two C_4-cycle genes, *NADP-MDH* and *ASP5*, patterns of *PPDK-RP*, *BCA3*, *CA1*, *PPA6*, *AMK2*, *TPT*, *DiT1* and *DiT2* strongly suggest periodic (i.e. circadian) expression in constant light in *F. robusta*. In *F. bidentis,* transcript abundances of PPDK-RP and BCA3 followed the diurnal phase in the first cycle of the circadian subset. This apparent circadian regulation was completely depleted in the second cycle. Circadian periodicity was not detected for any other of the C_4-cycle genes in *F. bidentis* (Fig. 3).

Fig. 3. Expression profiles of core C_4-cycle genes. Scheme of the NADP-ME C_4 pathway with expression profiles (in reads per million) of genes encoding for core C_4 enzymes and transporters over the diurnal and circadian time course. Blue, *F. bidentis*; Red, *F. robusta*; dark grey bars, night; light grey bars, subjective night. For detailed explanation of the time course, see Fig. 1A. Abbreviations are explained in the text and in Supplemental Dataset S3. Top right: phase of gene expression for significantly periodic genes (JTK_cycle, $q < 0.05$) und diurnal (left clock) and circadian (right clock) conditions. Notes: Key enzymes of other C_4 biochemical subtypes (NAD-ME1 and PEP-CK1) are displayed for comparison. NHD2 levels likely represent NHD1 (as explained in the text). "Circle-head" indicates post-translational regulation. Enzymes with activity in both cell types are only presented once for simplicity. ALAAT1 is likely active in the cytosols of both MC and BSC. Unknown transport proteins are marked with a question mark. 3-PGA, 3-phosphoglycerate; Ala, alanine; Asp, aspartate; Mal, malate; OAA, oxaloacetate, $(P)P_i$, (pyro)phosphate; Pyr, pyruvate, TP, triose-phosphate


Bundle sheath
Mesophyll
NAD–ME1
PEP–CK1
ASP5
OAA
Asp
NADP–MDH
Mal
DiT2.2
NADP–ME4
CO2
Pyr
Calvin Benson Cycle
3-PGA
TPT
?
Ala
ALAAT1
BASS2
NHD2
PPDK–RP
PPDK
PEP
DiT1
MDH
PPT1
AMP + PPi
AMK2
PPA6
ATP + Pi
TP
PPCK1
PPC1
HCO3-
BCA3

Diurnal and circadian expression of light reaction, CBC and photorespiration genes

Photosystem II (PSII). Abundances of 10 out of 25 transcripts encoding for subunits of PSII and the associated LIGHT-HARVESTING COMPLEX (LHCB) proteins were significantly lower in *F. bidentis* compared to *F. robusta* under diurnal conditions (Supplemental Dataset S6). The expression patterns of all PSII-related genes showed clear diurnal rhythmicity in both species, although not detected by JTK_cycle for all genes (Fig. 4A). With onset of the first subjective night, the expression of all genes followed the same path, albeit with a reduced amplitude, as in the two night parts of the diurnal subset in both species. This maintained increase or decrease with beginning of the night is indicative for circadian control of gene expression. While the expression of almost all genes had a trend of increase in constant light in *F. bidentis*, their abundances tended to deplete towards the end of the second cycle of constant light in *F. robusta*. This differential pattern was particularly prominent for genes encoding for LHCB proteins. Expression of the PSII subunits encoded by *Psb* genes (Fig. 4A), appeared to be under stronger circadian control in *F. robusta* than in *F. bidentis* judged by the largely differential amplitudes between the two species in constant light.

Cyclic electron flow (CEF). The ferredoxin:plastoquinone oxidoreductase pathway of CEF is composed of two proteins in *Flaveriaceae* (Nakajima Munekage, 2016), PROTON GRADIENT REGULATION 5 (PGR5) and PGR-like 1B (PGRL1B). Periodicity of PGR5 transcripts overlapped between both species in both subsets (with the same two-hour phase advance after transition to constant light). The phases of PGRL1B were maximally different with a midday peak (ZT04) in *F. bidentis* and a midnight peak (ZT16) in *F. robusta* (Fig. 4B).

Out of 21 genes encoding for subunits of the NADH dehydrogenase-like (NDH) complex (Ifuku et al., 2011), 15 were significantly more highly expressed in *F. bidentis* than in *F. robusta* under diurnal conditions (Fig. 4C). Significant diurnal periodicity with between-species-phase-differences smaller than four hours were assigned to all genes, except for *NdhN*, *NdhS, PnsL1* (not significant but detectable in *F. bidentis*) and *NdhU*, *PnsL4* and *PnsL5* (not significant but detectable in *F. robusta*). Although only in the first of the circadian cycles in *F. bidentis*, the circadian response detected in response to constant light for all NDH encoding genes in both species was more pronounced in *F. robusta*. In particular, levels of all five genes encoding for subunits of the NDH luminal subcomplex (*PnsL1-5*), four subcomplex B subunits (*PnsB1-3* and *PnsB5*) as well as *NdhS* were highly increased and largely maintained in constant light in *F. bidentis*.

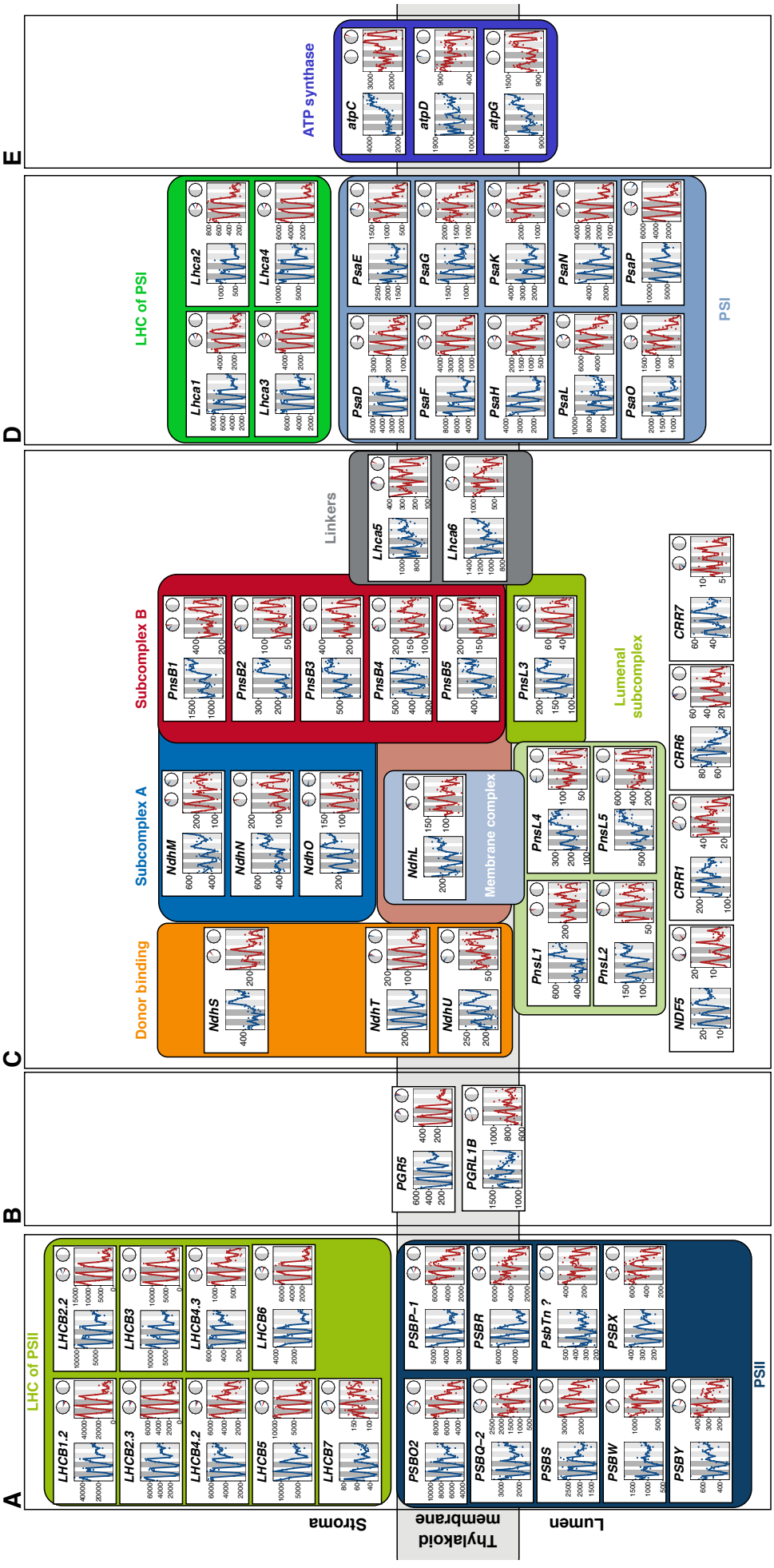

A
B
C
D
E
LHC of PSII
PSII
PSBO2
PSBQ–2
PSBS
PSBW
PSBY
PSBP–1
PSBR
PsbTn ?
PSBX
LHCB1.2
LHCB2.3
LHCB4.2
LHCB5
LHCB7
LHCB2.2
LHCB3
LHCB4.3
LHCB6
PGR5
PGRL1B
Donor binding
NdhS
NdhT
NdhU
Subcomplex A
NdhM
NdhN
NdhO
Membrane complex
NdhL
Subcomplex B
PnsB1
PnsB2
PnsB3
PnsB4
PnsB5
Linkers
Lhca5
Lhca6
Lumenal subcomplex
PnsL3
PnsL4
PnsL5
PnsL1
PnsL2
CRR7
CRR6
CRR1
NDF5
LHC of PSI
Lhca1
Lhca3
Lhca2
Lhca4
PSI
PsaD
PsaF
PsaH
PsaL
PsaO
PsaE
PsaG
PsaK
PsaN
PsaP
ATP synthase
atpC
atpD
atpG
Stroma
Thylakoid membrane
Lumen

Fig. 4. Transcript profiles of genes encoding for proteins involved in light reactions.
A, Photosystem II (PSII) and light harvesting complex (LHC) of PSII; B, ferredoxin:plastoquinone oxidoreductase pathway of CEF; C, Subunits of the chloroplast NADH dehydrogenase-like (NDH) complex. D, PSI and LHC of PSI; E, ATP synthase subunits. Except for panel C (Nomenclature and categorisation based on Ifuku et al. (2011), organisation with respect to thylakoid membrane an, storm and lumen is only schematic for simplicity. Genes in panel D based on Jensen et al. (2007). Only major isoforms and subunits present in the minimal reference shown (in particular, plastome encoded genes absent). Legend, see Fig. 1A and Fig. 3. Full gene names see text and Supplemental Dataset S6.

PSI. Transcripts encoding for seven out of 20 PSI-related proteins, four PSI subunits (PsaF/H/N/P), LHCA1 and 5, and PSBP-DOMAIN PROTEIN1 (PPD1) (Jensen et al., 2007), were significantly more abundant in *F. bidentis* than in *F. robusta* under diurnal conditions (Supplemental Dataset S6). All PSI-related genes exhibited significantly diurnal periodicity in both species with phases differing no more than two hours (Fig. 4D), except for *PsaE* (*F. bidentis*: ZT00, *F. robusta*: ZT06 and *PsaG* (not significant in *F. robusta*, but similar pattern as in *F. bidentis*). A circadian response comparable to the PSII-related genes, with diminished amplitudes was also visible for all PSI-related genes. Likewise, most genes trended upwards in *F. bidentis* and downwards in *F. robusta* after two days of constant light. Patterns of *LHCA5* and *LHCA6* (Fig. 4C), which function as linker proteins between the NDH complex and PSI (Ifuku et al., 2011), were markedly different from those of the other *LHCA* genes with (i) sharper peaks under diurnal conditions and a pronounced upward trend in constant light in *F. bidentis* and (ii) stronger circadian amplitudes in *F. robusta*.

ATP synthase. Transcript abundances of three nuclear encoded subunits of the ATP synthase complex (Friso et al., 2004) increased in constant light in *F. bidentis*, while they gained circadian periodicity in *F. robusta* with almost 2-fold higher amplitudes compared to diurnal expression (Fig. 4E and Supplemental Dataset S6).

Calvin-Benson-Cycle (CBC). Transcripts encoding for nine CBC enzymes were identified as significantly less abundant in *F. bidentis* than *F. robusta* under diurnal conditions, with RUBISCO SMALL CHAIN 1A and 3B (rbcS1A and rbsC3B, 4-fold and 3-fold lower) and two RUBISCO activases (RCA, 4-fold and 3-fold) showing the largest differences (Supplemental Dataset S6). Transcripts encoding for the CP12 DOMAIN-CONTAINING PROTEIN 2 (CP12-2) and a putative RUBISCO large subunit N-methyltransferase (RbcMT) were 1.3-fold and 2.1-fold more abundant in *F. bidentis*.

Diurnal periodicity and phases with no more than four hours difference between the two species was assigned to RCA (*F. bidentis*: ZT00 and *F. robusta*: ZT02), fructose-bisphosphate

aldolase (FBA) 1 and 3 (both ZT00 and ZT02), FBA5 (ZT18 and ZT14), ribulose-phosphate 3-epimerase (RPE, ZT22 and ZT20), sedoheptulose-1,7-bisphosphatase (SBPASE, ZT00 and ZT22), glyceraldehyde 3-phosphate (GAP) dehydrogenase A subunit 2 (GapA-2, ZT00 and ZT04), phosphoribulokinase (PRK, ZT00 and ZT02), CP12-1 and CP12-2 (both ZT10 and ZT12) (Fig. 5A). Under circadian conditions, *RCA*, *FBA1* and *FBA3* were the only genes to exhibit periodicity in both species. Their phases of expression remained the same as under diurnal conditions in the respective species, except for *RCA* in *F. bidentis*, with a phase delay of two hours (Fig. 5A). Except for *RPE* and *SBPASE*, no additional gene showed circadian periodicity in *F. bidentis*, while in *F. robusta* all of the CBC genes show circadian expression. Noteworthy were the transcriptional patterns of constant increase with prolonged light (rbcS1A, rbcS3B, ribose-5-phosphate isomerase (RPI3), and phosphoglycerate kinase (PGK)) or increased and maintained expression surpassing the diurnal level (GapB; FBPase (fructose 1,6-bisphosphatase) and transketolase 2 (TKL2)) in *F. bidentis*, while in *F. robusta* transfer to constant light resulted in a gain of amplitude and periodicity in these genes. In addition, expression of *GapB* was phase-advanced by six hours in *F. bidentis*.

Another, CP12 protein, namely CP12-3, which is supposedly not involved in the CBC (Groben et al., 2010), was 13.3-fold more abundant in *F. bidentis* under diurnal conditions and also showed the pattern of diurnal periodicity in both species, while constant, high expression in constant light in *F. bidentis* was contrasted by circadian expression in *F. robusta* (Fig. 5B).

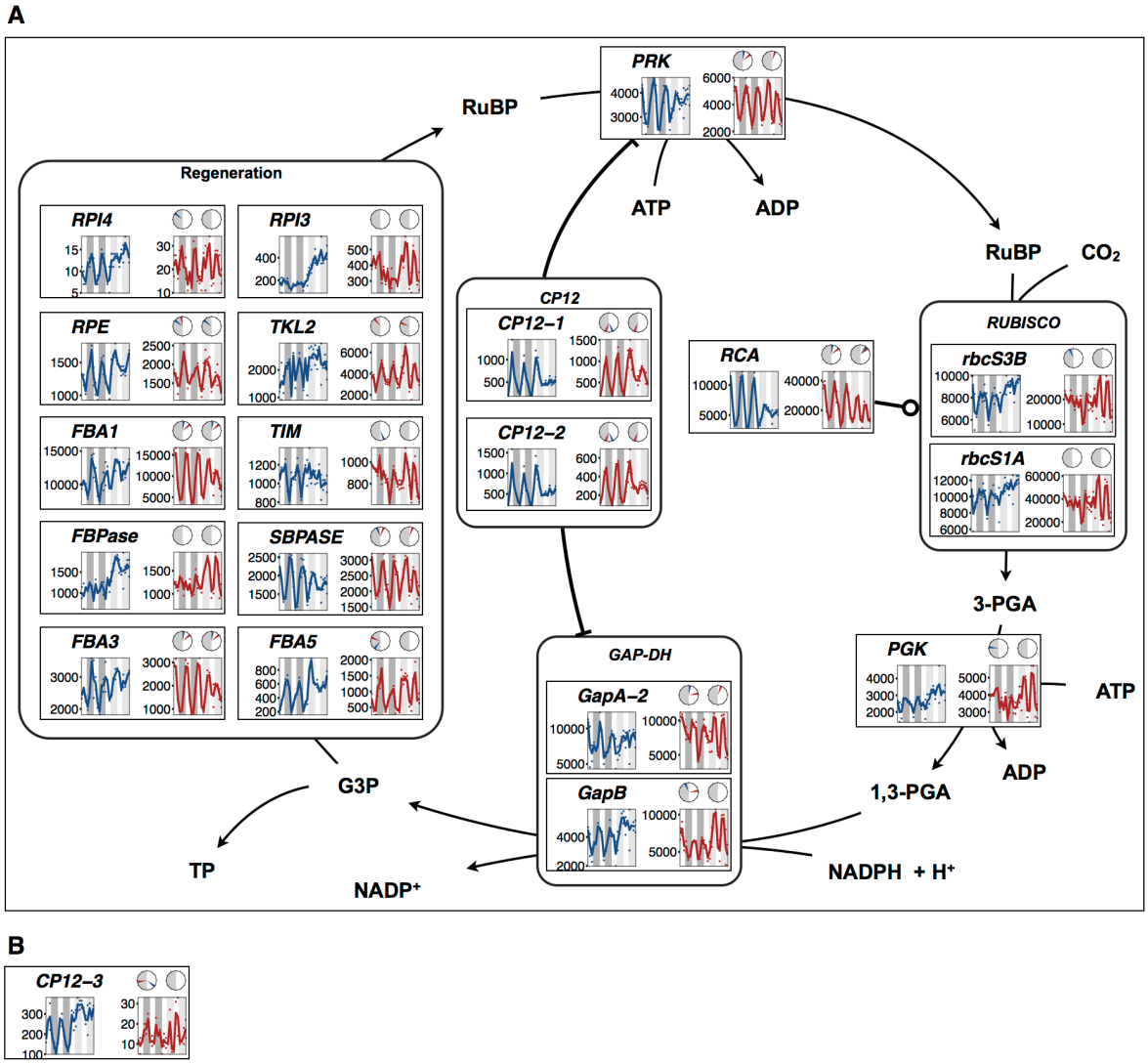


Fig. 5. Transcript profiles of CBC genes and CP12-3.
A, Scheme of CBC. Only major isoforms presented. B, CP12-3. Legend, see Fig. 1A and Fig. 3. Full gene names see text and Supplemental Dataset S6.

Photorespiration (PR). Transcripts encoding for 19 (diurnal) or 14 (circadian) out of 25 enzymes and transporters of PR were significantly less abundant in *F. bidentis* compared to *F. robusta* (Supplemental Dataset S6). Abundances of transcripts encoding CATALASE 3 (CAT3), HYDROXYPYRUVATE REDUCTASE 1 (HPR1), GLYCOLATE OXIDASE 3 (GOX3), SERINE HYDROXYMETHYLTRANSFERASE 3 (SHMT3) were not significantly different between the two species. Only the two DiT proteins, which are also part of the C_4-cycle, were upregulated in *F. bidentis*. After transition to constant light, average expression over all time points was as high or higher than the average expression under diurnal conditions for 13 genes (not counting DiT proteins) in *F. bidentis* (Fig. 6 and Supplemental Fig. S5). These included all but one genes encoding for subunits of GLYCINE DECARBOXYLASE (GDC), SERINE-GLUTAMATE AMINOTRANSFERASE (SGAT), PHOSPHOGLYCOLATE PHOSPHATASE (PGLP), GLUTAMATE:GLYOXYLATE AMINOTRANSFERASE (GGAT), SHMT3, HPR1, GLYCOLATE OXIDASES 1-3 (GOX 1-3) and the PLASTIDIC GLYCOLATE GLYCERATE TRANSPORTER (PLGG1). Transcript abundances of GDC subunit P (GDC), SHMT1, GLYCERATE KINASE (GLYK), GLUTAMATE SYNTHASE 1 (Fd-GOGAT), the four GLUTAMINE SYNTHETASES, GS2 and GLN1;2/4/5 and CAT2 and 3. By contrast, in *F. robusta* except for four slightly depleted genes (*GLYK*, *GLN1;4*, *GLN1;5* and *SHMT3*), all genes encoding for PR proteins were expressed in the same range as under diurnal conditions or increased after transition to constant light (Fig. 6 and Supplemental Fig. S5).

Under diurnal conditions, significantly periodic expression with equal phases in both species was assigned to *GGAT*, *GOX1-3*, *HPR1* and *FD-GOGAT*. Expression of the GDC subunits H, L and T, CAT2 and 3, SGAT, SHMT3, and GS2 was phase advanced by approximately two to four hours in *F. bidentis* compared to *F. robusta*. Exclusively periodic expression in *F. bidentis* was detected for transcripts encoding SHMT1 (phase ZT04), GDC-P (phase ZT20), PGLP and PLGG1 (both phase ZT 18). Analogously to the previously highlighted C_4-cycle genes, periodic expression was diminished after transition to constant light in *F. bidentis* for a large proportion of PR genes, while following a circadian expression in *F. robusta* (although likewise not detected by the JTK_cycle algorithm). This difference between the two species was particularly noticeable for PGLP, SGAT, GDC subunits L and T, GS2, HPR1 and PLGG1.

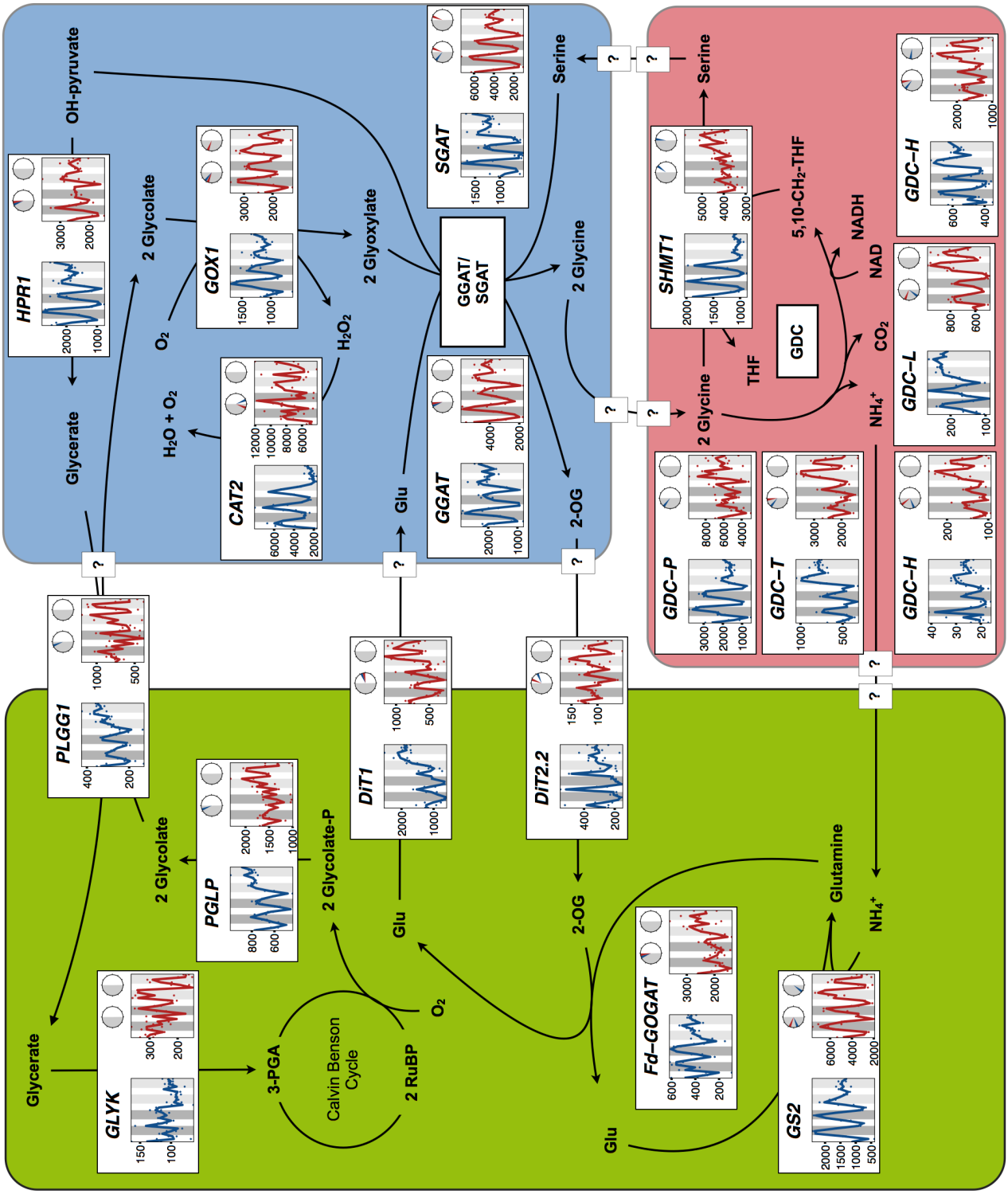


Fig. 6. Transcript profiles of photorespiration genes.
Scheme of photorespiration across chloroplast (green), peroxisome (blue) and mitochondrium (red). Only major isoforms presented. Legend, see Fig. 1A and Fig. 3. Full gene names see text and Supplemental Dataset S6.

Light perception and signaling

Photoreceptors. Transcripts encoding for the far-red light photoreceptor phytochrome A (PHYA) and a blue light receptor of the ZEITLUPE (ZTL) protein family, flavin-binding, kelch repeat, f box 1 (FKF1) were 3.6-fold and 2.4-fold more abundant in *F. bidentis* compared to *F. robusta* under diurnal conditions (Supplemental Dataset S6). Diurnal periodicity with phases differing no more than two hours between species was assigned to all genes encoding for photoreceptor proteins, except for LOV KELCH protein 2 (LKP2), which was not significant but detectable and equally phased in both species (Fig. 7). Transition to constant light indicated circadian response, albeit with diminished amplitude expression of all photoreceptor encoding genes in both species. While transcript patterns of PHYB to PHYE and PHOTOTROPINS largely overlapped between the species, transcripts encoding for both cryptochromes and all members of the ZTL family showed larger amplitudes in *F. robusta* in constant light.

Light signaling. Based on the observation that PHYA was more abundantly expressed in *F. bidentis*, the known and partially overlapping signaling network downstream of PHYA and PHYB was analyzed regarding differential expression profiles between the species (Supplemental Fig. S6 Quail, 2002). Under diurnal conditions, three genes were more abundantly expressed in *F. robusta*, SHORT UNDER BLUE LIGHT 1 (SUB1) as well as RED ELONGATED 1 (RED1) and FUSCA 9 (FUS9). Transcripts encoding for CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1) were more abundant in *F. bidentis* and transcript abundances of the PHYTOCHROME-INTERACTING FACTOR 4 (PIF4) were negligible in *F. robusta*, while low but significantly periodic levels were detected in *F. bidentis*. Although not detected as significant DEG, transcripts encoding for GIGANTEA (GI) were more abundant under diurnal conditions in *F. bidentis*, while following the same pattern of expression as in *F. robusta*.

The most drastically differential patterns between the species were detected for two genes, NUCLEOSIDE-DIPHOSPHATE KINASE 2 (*NDPK2*) and *PIF3*, which are induced by both PHYA and B (Supplemental Fig. S6). NDPK2 transcripts were diurnally periodic and increased in constant light in *F. bidentis*, while in *F. robusta* expression was more distorted under diurnal conditions and gained periodicity with a high amplitude in constant light. Expression of PIF3 decreased in *F. bidentis* and increased with a gain of periodicity and amplitude in *F. robusta* in constant light. In addition, LONG HYPOCOTYL 5 (HY5) expression

was phase advanced in *F. bidentis* under diurnal conditions and the circadian response was delayed.

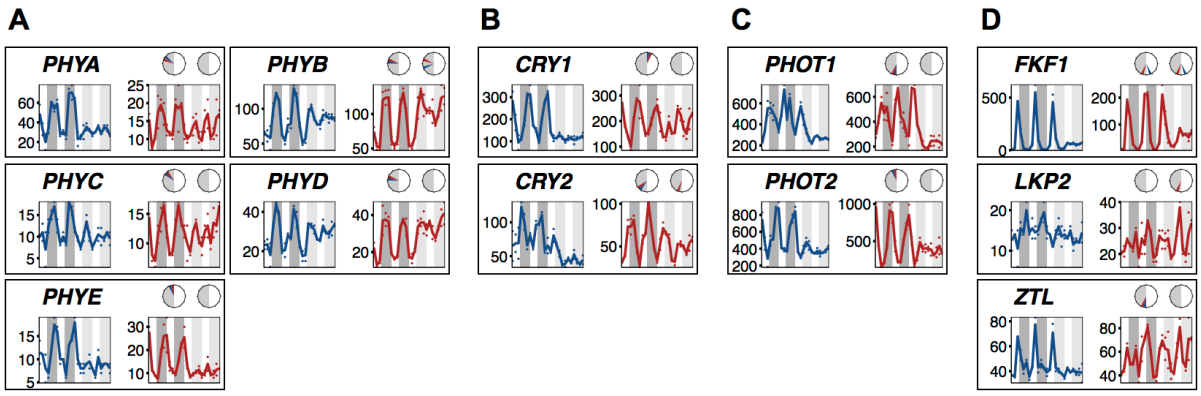


Fig. 7. Transcript profiles of photoreceptor genes.
A, Phytochromes; B, Cryptochromes; C, Phototropins; D, FKF/LKP/ZTL family proteins.
Legend, see Fig. 1A and Fig. 3. Full gene names see text and Supplemental Dataset S6.

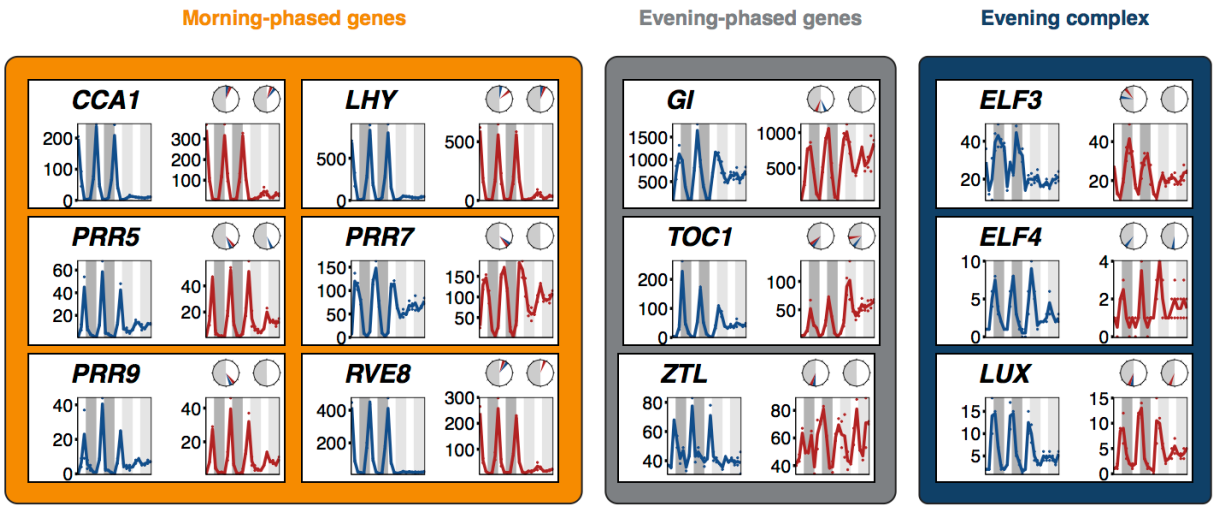


Fig. 8. Transcript profiles of core circadian clock genes.
Legend, see Fig. 1A and Fig. 3. Full gene names see text and Supplemental Dataset S6. See Greenham and McClung (2015) for details.

The circadian clock

Expression profiles of genes encoding for the core clock proteins CIRCADIAN CLOCK ASSOCIATED 1 (CCA1), LATE ELONGATED HYPOCOTYL (LHY), the PSEUDO-RESPONSE REGULATORs (PRR) 5, 7, and 9, EARLY FLOWERING 3 (ELF3), LUX ARRHYTHMO (LUX) and REVEILLE 8 (RVE8) strongly correlated between the two species over the complete time course with significantly periodic, equal diurnal phases and similar amplitudes (Fig. 8 based on Greenham and McClung, 2015). Higher transcript abundances were detected for TIMING OF CAB EXPRESSION 1 (TOC1, 3.2-fold), ELF4 (2.6-fold, albeit lowly expressed in both species) and RVE8 (1.7-fold). After transition to constant light, TOC1 expression dampened in *F. bidentis*, while increasing in *F. robusta*.

Amongst the genes encoding for additional proteins related to the circadian clock (reviewed in Nakamichi, 2011; Hsu and Harmer, 2014; McClung, 2006), the expression profiles of six genes showed the most drastic differences between the two species (Supplemental Fig S7). *CASEIN KINASE II BETA SUBUNIT 4* (*CKB4*) was similarly diurnally periodic. In constant light, expression markedly dampened in *F. bidentis*, while increasing with large amplitude in *F. robusta*. The expression of *CHLOROPLAST STEM–LOOP-BINDING PROTEIN OF 41 kDA B* (*CSP41B*) was two-fold decreased and six-hour phase-advanced to ZT20 in *F. bidentis* under diurnal conditions. After transition to constant light, *CSP41B* periodicity was diminished with abundances surpassing the diurnal peaks in *F. bidentis*, while strong circadian response and gain in amplitude was observed in *F. robusta*. Circadian response was pronounced in *F. robusta* for *LIGHT-REGULATED WD 1* (*LWD1*), which was equally phased under diurnal conditions. Similarly, diurnal expression of *TIME FOR COFFEE* (*TIC*) and *XAP5 CIRCADIAN TIMEKEEPER* (*XCT*) was equally periodic between the species and periodicity, albeit with a shorter period, was maintained under circadian in *F. robusta*, while dampened in *F. bidentis*. Expression profiles of *SENSITIVE TO FREEZING 6* (*SFR6*) were distorted under both conditions in both species, but seemed to gain abundance and periodicity under circadian conditions in *F. robusta*, while dampening in *F. bidentis*. Expression of *RVE6* was similarly depleted upon constant light in both species, but phase-delayed by six hours to ZT06 under diurnal conditions in *F. bidentis*. In addition, transcripts of COLD, CIRCADIAN RHYTHM, AND RNA BINDING 2 (CCR2), which showed otherwise alike expression patterns, was two-fold more abundant in *F. bidentis*.

DISCUSSION

Large and shared proportions of the leaf transcriptomes are diurnally periodic

Comparative analyses of the timing of gene expression raise opportunities to understand the fine tuning of processes leading to phenotypes and metabolic adaptations differing between plant species. In order to identify differential timing, gene expression profiles are analyzed regarding their timing of peak expression (phase), their period and amplitudes (Harmer, 2009). With accumulating RNASeq scale diurnal and circadian datasets, there is a growing choice of statistical analyses and tools to detect periodicity in such datasets, which are constantly refined and adapted to the specific requirements (Deckard et al., 2013). Using the JTK_cycle (Hughes et al., 2010) algorithm, we showed largely overlapping expression phases between the two closely related *Flaveria* species of different photosynthesis types (Fig. 1). Two thirds of the leaf transcriptomes phased to the night in both species under diurnal conditions. After transition to constant light, approximately 4% (*F. bidentis*) and 6% (*F. robusta*) of the transcriptome was detected as periodically expressed, and thus under control of the circadian clock (Supplemental Fig. S1). These numbers are lower than typically reported for the model plant *A. thaliana* (e.g. between 10% in Covington and Harmer, 2007 and 16% in Edwards et al., 2006). Mathematical detection of periodicity gave a hint to whether genes were maintained cycling under constant light conditions. However, the currently available algorithms failed to detect a weak circadian response, which may be of biological relevance, as explored in the detailed analysis of categorical gene subsets (Figures 3 - 8 and Supplemental Figures S4 – S7). In order to call gene expression "periodic", JTK_cycle fits the expression levels over the time course against sinusoidal curves of user-specified periods and a respective range of phases. This worked well for genes with a sinus like and very regular expression over at least two repetitive cycles (i.e. equal light-dark conditions as in the diurnal subset). Transition from light-dark to constant light often induces prolonged period of gene expression (Harmer, 2009). The resulting more irregular gene expression in constant light did not seem to be captured as cycling, especially when considering that (i) gene expression was likely adapted gradually and (ii) the first half of the first of the circadian cycles was not perceived by the plants as an environmental irregularity, ergo the cue "constant light" was only perceived with the first missing darkness (i.e. subjective night). However, the quickly dampened amplitudes detected for the vast majority of the genes in both species after more than one day of constant light argue against an alternative experimental setup

with one day of acclimatization to constant light before collection of the circadian samples. In addition, sampling uninterrupted by a gap for light acclimatization allowed for a valuable, direct comparison of the gene expression response to constant light versus light-dark conditions. Many publicly available datasets, especially for the model plant *A. thaliana*, neglect this feature, hindering the distinction of light- versus circadian clock regulated processes.
The parameters q-value, amplitude, phase and period retrieved from the periodicity analysis alone do not serve as valid statistical metrics to identify significantly species-specific diurnal or circadian gene expression, but need to be combined with additional analyses.

Phase adjusted analysis of differential gene expression

The analysis of periodic gene expression revealed that more than 60% of the leaf transcriptomes of both *Flaveria spp.* were periodically cycling over the diurnal course and the major proportion of these genes was phased to peak in the middle of the dark period (Fig. 1). However, the phases of at least one fifth of these genes were shifted by four or more hours between the two species. Due to the high sequencing cost of next generation sequencing, comparative transcriptomics approaches are frequently restricted to a few samples collected at a single time point (e.g. Gowik et al., 2011; Mallmann et al. 2014). Genes with differential phases between two comparison groups may thus be falsely identified or rejected as DEG (regarding their absolute transcript abundance), or their degree of differential expression may be over- or underestimated depending on the time point of sampling (Fig. 2A). This is especially emphasized when a time point in the middle of the light period is chosen as the single time point of sampling (i.e. the time of lowest expression for many genes), which is commonly the case. With two consecutive, independent cycles of sampling at a four-hour resolution, the dataset presented here has the strength to overcome this potential burden.

Gene expression levels extracted at two peak time points of each subset and species were used for statistical analysis of DEG with DESeq2 (Love et al., 2014), resulting in the identification of 4,913 DEG under diurnal and 2,817 DEG under circadian conditions. Since this approach yielded an order of magnitude more DEG than the five-species comparison with *Flaveria spp.* by Gowik et al. (2011), this may be an over-estimation of DEG and the detected DEG are certainly not all connected to evolution of the C_4 syndrome. Also, the Gowik et al. (2011) study (i) was restricted to single samples without replication and (ii) compared different species pairs as in the given study.

The approach failed to capture genes, if variance of peak expression between two cycles was too high as was exemplified with the circadian clock and light signaling related gene *GI* (Fig. 8). Abundances in *F. bidentis* of periodically expressed *GI* were constantly above those in *F. robusta*.

Alternative approaches for analysis of absolute differential expression in mRNA-Seq scale time series data are required. Analysis of variance (ANOVA) and the ANOVA-like implementation in the DESeq2 package so far yielded unexpected or less powerful results (data not shown). Nonetheless, the DEG analysis presented here provides a reasonable approximation for detection of highly different transcript abundances between the two species.

Diurnal expression and light-regulation of the C_4-cycle is already partially present in the C_3 species

In order to identify differential gene regulation by light or the circadian clock between C_3 and C_4 *Flaveria* species, we analyzed the expression profiles of genes encoding the core C_4-cycle as well as genes related to photosynthesis, photorespiration, light signaling and the circadian clock.

The phase-adjusted DEG analysis confirmed previous *Flaveria* transcriptomic studies showing the transcriptional enrichment of all NADP-ME C_4-cycle genes in the C_4 species *F. bidentis* (Gowik et al., 2011; Mallmann et al., 2014 and Supplemental Dataset S5). While transcript abundances, i.e. diurnal amplitudes, were largely enhanced in *F. bidentis* compared to *F. robusta*, the diurnal expression phases of only a few genes was found to be substantially different between the two species (Fig. 3). The phase of *PPC1* was shifted from a late night expression in *F. robusta* towards midday expression in *F. bidentis*, where it preceded the gene expression of its kinase *PPCK1*. Aldous et al. (2014) showed differential diurnal phases of mRNA abundances between two *PPCK* isoforms in *F. trinervia* (C_4) and *F. pringlei* (C_3). While abundances of *PPCKA* peaked during the middle of the light, abundances of *PPCKB* peaked during the middle of the dark. PPCK regulates the activity of PPC via (de–)phosphorylation (Nimmo, 2003). The light-phased isoform, *PPCKA*, was (i) significantly more abundant in the C_4 species than in the C_3 species and (ii) its diurnal expression profile correlated closely with the diurnal phosphorylation state of the PPC protein in *F. trinervia*, while the phosphorylation state of PPC in *F. pringlei* did not show diurnal oscillation (Aldous et al., 2014). Here, differential diurnal phases between *PPCK* isogene transcripts were only detected in *F. robusta* (*PPCK1*: ZT10, *PPCK2*: ZT14) and werw not as drastic as in their study. The expression profiles of both

PPCK isogenes largely overlapped in *F. bidentis*, which may be specific for this species (Fig. 3 and Supplemental Fig. S4). In addition, it is important to note, that isogene distinction is possibly hindered by the cross-species mapping against *A. thaliana*. Re-mapping of the reads against *Flaveria* genomes will hopefully help in this respect in the future. Differential phases of *NHD*, *MDH*, *DiT2* and *BASS2* did not suggest co-regulation with proximal enzymes in the C_4-cycle.

The presence of diurnal periodicity in expression of C_4 genes with minor changes in phases between the two species allows to draw two major hypotheses. First, diurnally periodic, and thus light- or circadian clock regulated expression of C_4-cycle genes was already present in the C_3 ancestor. Second, during the evolutionary recruitment of C_4-cycle enzymes from C_3 ancestral functions, the expression of their encoding genes was not set under a common regulatory control mechanism providing equally phased expression that overrides or alters this diurnal control. Similarly, a circadian transcriptome study of the C_4 monocot maize detected periodic but not equally phased C_4 gene expression (Khan et al., 2010).

Major differences between the two species emerged after transition to constant light. Light-dependent mRNA accumulation has previously been shown for *NADP-ME*, *PPDK*, and *PPCK* in *F. trinervia* (Lai et al., 2002; Rosche and Westhoff, 1995; Tsuchida et al., 2001; Aldous et al., 2014). Surprisingly, light-dependent increases in transcript abundances of *NADP-ME*s were not detected (Fig. 3). *PPDK* expression was not diminished in *F. robusta* in constant light, but it did not resemble the high and constant increase seen in *F. bidentis* (Fig. 3). High *PPDK* levels in the C_3 plant may be attributed to its original light-dependent function in gluconeogenesis. Interestingly, although more prominently circadian regulated, abundances of *PPCK* transcripts had an increasing trend in *F. robusta*, while rather constant in *F. bidentis*. Furthermore, markedly increasing transcript abundances of *DiT1*, *AMK2*, *PPA6, ASP5, ALAAT1, BASS2, CA1* and *NADP-MDH* in *F. bidentis* were contrasted by a circadian response or diminished abundances in *F. robusta* (Fig. 3 and Supplemental Fig. S4).

The fact that *ASP5* and *NADP-MDH*, but not *DiT2*, were highly upregulated in constant light in *F. bidentis* emphasizes the role of the aspartate route for *NADP-ME* C_4 *Flaveria* under high light intensities (Meister et al., 1996). Nitrogen transfer between MC and BSC via shuttling of aspartate is likely compensated by shuttling back alanine after transamination of pyruvate by *ALAAT1* (Wang et al., 2014b). *ALAAT1* transcripts also showed a light-responsive increase in *F. bidentis*. However, the minor light-response of *NADP-ME4* does not fit into this picture.

Transcripts encoding for all three *PPC* isoforms increased to high levels in constant light in both species (Fig. 3 and Supplemental Fig. S4). Interestingly, this trend was pronounced for the non-photosynthetic *PPC4* in *F. bidentis*, which showed strong circadian periodicity in

F. robusta. Similarly, expression of *PPT1* seemed to have a stronger light-response in *F. robusta* than in *F. bidentis*, and in both species compared to *PPT2*. Further support of the hypothesis of the recruitment of readily light-regulated isogenes for C_4 (Steve Kelly, unpublished), require *Flaveria* genome sequences. Minor differential periodic expression between the two species and absence of the *F. bidentis* specific light-dependent accumulation of genes encoding for the alternative decarboxylases *NAD-ME1* and *PEP-CK1* argue against their putative involvement in CO_2 assimilation in *F. bidentis* at least under the conditions tested (Wang et al., 2014b).

In summary, while most C_4-cycle genes where maintained periodic in the C_3 species *F. robusta* indicating regulation by the circadian clock, prominent circadian responses were limited to a few genes in the C_4 species. By contrast, in the C_4 species their expression was maintained or upregulated to high levels surpassing the peak expressions reached under diurnal conditions (Fig. 3). This suggests acquisition of direct or indirect light-regulation that overshadows the circadian control or a complete liberation from the circadian clock. In order to substantiate the hypothesis that C_4-cycle genes were liberated from circadian control during C_4 evolution, their expression in the evolutionarily more distant C_3 species *A. thaliana* was analyzed (Supplemental Fig. S8, Edwards et al., 2006; Covington and Harmer, 2007). This indicated circadian regulation of *ALAAT1, ASP5, BASS2, DiT1, DiT2, PPCK1, PPDK, PPDK-RP and TPT* in another C_3 species.

Further proof for C_4-specific acquisition of cis-regulatory elements, e.g. G-boxes for light-regulation, will likely be substantiated with the availability of *Flaveria* whole genome sequences. In addition, liberation from circadian control could unequivocally be disclosed in plants with mutations in one of the core clock genes or by analyzing the response to constant darkness. However, as a side effect in such experiments, the circadian or light/dark-responses would likely be overridden by a starvation response.

Light-regulated and circadian clock liberated gene expression of photosynthetic processes in C_4

In addition to the core C_4-cycle genes, gene expression profiles related to other photosynthetic processes in C_3 and C_4 plants were analyzed. A large proportion of the genes encoding for proteins functioning in PSII, CBC or PR were markedly downregulated in *F. bidentis* compared to *F. robusta* as has previously been shown for *Flaveria* (Gowik et al., 2011; Mallmann et al., 2014) and other C_4 species (e.g. Bräutigam et al., 2011; Külahoglu et al.,

2014). By contrast, the majority of genes encoding for PSI subunits and CEF was more abundantly expressed in *F. bidentis* (Supplemental Dataset S6).

The expression profiles of most genes encoding for subunits of PSII and PSI were diurnally periodic in both species (Fig. 4). Their circadian regulation was pronounced in the C_3 species, while increasing accumulation was observed in constant light in the C_4 species. However, the ranges of gene expression in constant light of only a few of these genes surpassed the peaks under diurnal conditions, which abstracts their profiles from the C_4, CBC and PR genes. On a side note, in contrast to many other observed sharp peaks, the high and constant transcript abundance of LHC proteins during the light with a sharp drop after onset of darkness represented beautiful examples for direct but rather rare correlation between transcription and the demand for the gene product and is likewise seen in *A. thaliana* (Harmer et al., 2000).

In all C_4 plants, two ATP per fixed CO_2 are required for phosphorylation of pyruvate in the MC. Hence, the operation of the C_4-cycle in addition to the CBC results in a higher energy requirement compared to C_3 plants (Munekage and Taniguchi, 2016). In NADP-ME type *Flaveria spp.* approximately 40% of the fixed CO_2 pass through aspartate (Moore and Edwards, 1986), additionally increasing the ATP consumption due to high MDH activity in the BSC (Meister et al., 1996). During the evolution of C_4, the additional ATP requirement prompted an increase in CEF activity around PSI (Munekage and Taniguchi, 2016). Decreased abundance of stacked grana in the BSC during the evolution from C_3 to C_3-C_4 to C_4 *Flaveria* species is a visible indicator of the promotion of CEF, which occurs at unstacked grana (Nakajima Munekage, 2016). Two partially redundant pathways of CEF are distinguished, the PGR5-dependent pathway and the NDH-dependent (Shikanai, 2007). In the C_3 plant *A. thaliana*, the PGR-dependent is the major pathway of CEF (Munekage et al., 2004), while the NDH complex appears only essential under stress conditions (Shikanai, 2007). By contrast, in the C_4 plant *F. bidentis* NDH is the major CEF pathway, indicated by lowered rates of CO_2 assimilation in NDH knock-out mutants (Yuri Munekage, unpublished) as well as four-fold and 14-fold elevation of NDH-H abundance in MC and BSC chloroplast compared to MC of C_3 species (Nakamura et al., 2013). Three-fold increases of PGR5 and PGRL1 in C_4 compared with C_3 species were not cell-specific (Nakamura et al. 2013). In good agreement, we observed significantly higher abundances of the majority of NDH subunit encoding genes in *F. bidentis* compared to *F. robusta* (Supplemental Dataset S6). In addition, constant light prompted increased expression of the *NDH* genes surpassing the diurnal levels in *F. bidentis*, while clear circadian regulation was observed in *F. robusta* (Fig. 4). Transcripts of PGR proteins were only slightly more abundant in the C_4 species and did not show a clear differential response to

constant light. Interestingly, the diurnal phases of *PGRL1* were maximally different between the two species (Fig. 4). In *F. bidentis*, dawn-phased expression of *PGRL1B* was more similar to that of *PGR5*, raising the possibility of a C_4-specific importance of yet unknown kind.

A large array of CBC genes, most prominently *rbcS1A*, *rbcS3B*, *RPI3*, *GapB*, *FBPase*, and *TKL2* as well as *TPT*, which encodes for the TP transporter and is required for TP shuttling in C_4 plants and chloroplast export of TPs synthesized by the CBC showed a light-responsive increase in transcript abundance in *F. bidentis* and an obvious circadian regulation in *F. robusta* (Fig. 5A). The same pattern has been detected for many genes encoding for PR proteins, in particular PGLP, SGAT, GDC subunits L and T, GS2, HPR1 and PLGG, (Fig. 6).

Additionally, differential gene expression profiles were detected for *CP12-3*, which was 13.3-fold more abundant in *F. bidentis* under diurnal conditions and diurnally periodic in both species, but phase-shifted by ten hours (Fig. 5B). CP12-3 proteins form a clade distinct from other CP12 proteins, which are involved in post-translational CBC regulation via complexation of GAPDH and PRK under oxidative stress (Groben et al., 2010; Marri et al., 2014). *Flaveria* CP12-3 knock out plants exhibit decreased photosynthetic activity under high light. This effect is emphasized under low CO_2 conditions (Tsuyoshi Furumoto, unpublished). Different energy requirements increase the contribution of aspartate to CO_2 assimilation under high light conditions (Meister et al., 1996; Nakajima Munekage, 2016). Recent evidence of *Flaveria* CP12-3 binding MDH instead of GAPDH led to the hypothesis, that MDH is set free from CP12-3 under high light (i.e. oxidative stress) to catalyze the malate to OAA conversion when aspartate is the primary C_4 acid (Tsuyoshi Furumoto, unpublished). The light-responsive transcript accumulation in *F. bidentis* versus circadian expression in *F. robutsa* support the C_4 specific of function CP12-3.

Candidate regulators of CC-liberated or light-pronounced gene expression in C_4

Diurnally changing gene expression is likely a consequence of regulation by the circadian clock, by light or metabolic state, e.g. sugar signaling through trehalose-6-phosphate, which itself is a consequence of light-regulated photosynthetic processes (Rolland et al., 2006).

In order to determine, whether sugar-signaling plays a role in the light-dependent accumulation of transcripts encoding for photosynthetic functions in *F. bidentis*, we analyzed the expression profiles of genes encoding for known steps of sugar-signaling (Supplemental Fig. S9). Transcript profiles of trehalose-6-synthases (TPS), trehalos-6-phosphatases (TPP) and SNF1-RELATED PROTEIN KINASEs (SnRK1) were not markedly different regarding

periodicity and absolute abundances between the two species. While a maintained circadian periodicity and increasing abundance were observed for *HEXOKINASE 2* (*HXK2*) in *F. robusta*, periodicity was maintained only for the first circadian cycle in *F. bidentis* and expression depleted thereafter, a pattern that does not resemble the pattern seen for photosynthesis related genes.

Amongst known genes encoding for photoreceptors or the downstream light-signaling network, two photoreceptor genes, *PHYA* and *FKF1*, were found to be more abundantly expressed in *F. bidentis*. Analysis of the levels of these genes across multiple *Flaveria* species in the data of Mallman et al. (2014) suggests that the observed high abundances appeared to be specific to *F. bidentis* rather than C_4 *Flaveria* in general. Still, as soon as a *Flaveria* (*bidentis*) genome becomes available, promoter sequences of possible downstream genes can be analyzed for C_4 specific acquisition of G-boxes indicating PIF binding and initiation of light-responsive transcription. Only negligible differential patterns were observed for known PIF transcripts. Transcriptomic responses to experiments with monochromatic (e.g. far-red, red or blue light) could underline photoreceptor specific regulation in C_4 *Flaveria*. Differential abundances of transcripts encoding RED1, SUB1, COP10 and GI, which is also part of the clock, appeared likewise species-specific (Mallmann et al., 2014).

Intriguingly, the expression profile of *NDPK2* resembled the pattern of the highlighted photosynthesis-related genes (Fig. 9A). NDPK2 is known to mediate phytochrome signaling and its activity is increased upon interaction with the Pfr form (Choi et al., 1999). *A. thaliana ndpk2* mutants show higher levels of reactive oxygen species (ROS) and partially defected red/far-red light responses, including cotyledon opening and greening (Choi et al., 1999; Moon et al., 2003). Heterologous expression of *A. thaliana* NDPK2 in potato, sweet potato, poplar and alfalfa (Kim et al., 2009; 2011; Wang et al., 2014c) and its overexpression in *A. thaliana* (Moon et al., 2003) confirmed its role in oxidative stress signaling and was associated with increased growth and tolerance to abiotic stresses, higher relative water and chlorophyll contents, increased activities of antioxidative enzymes and lowered ROS levels. The NDPK2 signaling pathway is not yet fully elucidated. Moon et al. (2003) showed ROS-induced gene expression and autophosphorylation of NDPK2 and suggested H_2O_2-mediated mitogen-activated protein kinase (MPK) activity. Increased levels of auxin-related genes (Kim et al., 2005; 2009; 2011) suggested a possible role for auxin-mediated responses, including auxin transport. However, correlation of *NDPK2* transcript profiles to those of genes related to MPKs or auxin were not detected in the *F. bidentis* transcriptome (Data not shown). Further analyses will be required to determine, whether the observed differential *NDPK2* profiles indicate their involvement in enhancing

transcription or activity of photosynthesis-related genes or are simply the response to increased levels of ROS production due to higher photosynthetic and photoassimilatory activities during C_3 photosynthesis. In this context it is to be noted, that the overall levels of *NDPK2* were lower in C_4 plants compared to C_3 relatives (Fig. 9C). Peak expression of *NDPK2* was detected in the middle leaf of *F. bidentis* in a comparative transcriptomics study of *Flaveria* leaf segments (Billakurthi, Wrobel, unpublished and Fig. 9B) thus not significantly correlated with C_4-cycle genes.

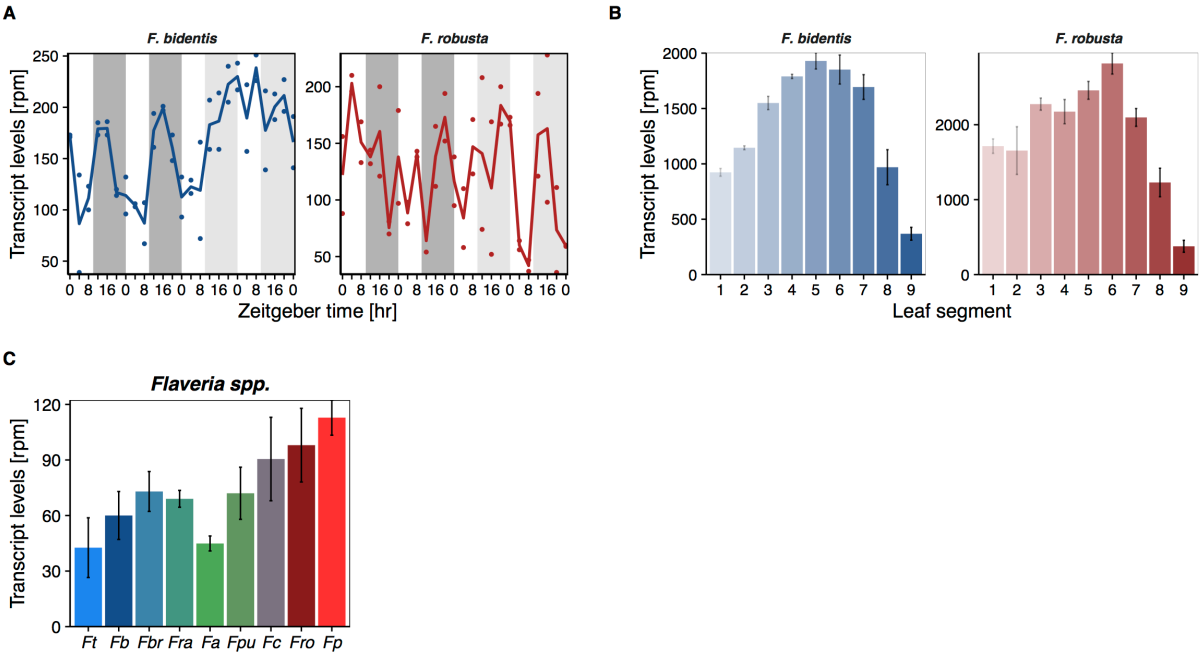


Fig. 9. Transcript levels of NDPK2 in different C_4 transcriptome studies.
A, Transcript profile this study (see Fig. 1A for legend); B, Comparative transcriptome of *Flaveria* leaf segments (Billakurthi, Wrobel, unpublished) ranging from bottom (1) to tip (9); C, Study of Mallmann et al. (2014) comparing the levels in C_4 (*Ft*, *F. trinervia*; *Fb, F. bidentis*), C_4-like (Fbr, *F. brownii*), C_3-C_4 (*Fra*, *F. ramosissima*; *Fa*, *F. anomala*; *Fpu*, *F. pubescens*; *Fc*, *F. chloraefolia*) and C_3 (*Fr*, *F. robusta*; *Fp*, *F. pubescens*) species of *Flaveria*.

Most circadian clock genes did not exhibit markedly altered expression profiles between the two species. Phase and range of expression were highly correlated between the two *Flaveria* species (Fig. 8 and Supplemental Fig. S7), emphasizing a likely evolutionary conservation of the circadian clock in the plant kingdom (Harmer, 2009). However, the transcriptomes of the leaf circadian clock genes appeared uncoupled in both species (Endo et al., 2014), which was revealed by quickly diminished amplitudes of the vast majority of these genes and the detection of comparably small proportions of circadian gene expression within both transcriptomes (Supplemental Fig. S1). As previously mentioned, cross-species mapping hinders confident distinction of isogenes. Hence, if a single isogene of a large gene family is periodically expressed, while the sister genes are not, the periodicity of this isogene may be diluted by the indistinct mapping. The forthcoming genome sequences might help in this aspect (Udo Gowik, unpublished). Analysis of *TOC1*, *ELF4* and *CCR2* in the Mallmann et al. (2014) transcriptomes disclosed their upregulation as likely species-specific.

During plant evolution, the genetic information of the majority of chloroplast functions, including most photosynthesis-related genes, was transferred to the nucleus (Schnarrenberger and Martin, 2002; Richly and Leister, 2004). However, the regulation of their expression is largely signaled by the chloroplast, a process known as retrograde signaling, in the form of tetrapyrrole intermediates of chlorophyll biosynthesis, plastid proteins, sugar status or ROS (Leister, 2005). Hassidim et al. (2007) showed alterations in the timing of mRNA accumulation of core clock genes *CCA1* and *LHY* in plants with mutated *csp41b* (in their study called *crb* for chloroplast RNA binding), which is involved in retrograde signaling, highlighting its possible role in connecting the circadian system to the status of the chloroplast. *A. thaliana csp41b* mutants are smaller and paler than wild-type plants, have lower fresh weight, reduced levels of chlorophylls, impaired photosynthesis and altered thylakoid structure, including increased granal stacking (Hassidim et al., 2007). Furthermore, lack of CSP41b results in decreased chloroplast translation rate, reduced levels of thylakoid proteins and decreased photosynthetic activity (Bollenbach et al., 2009; Qi et al., 2012). There is also evidence, that CSP41B enhances chloroplast RNA-stability of photosynthetic core proteins and ribosomal RNAs and that its RNA-binding properties are regulated through the redox state of the stroma (Qi et al., 2012), which is directly correlated to light/dark environment.

In our study, transcript abundances of *CSP41B* in response to constant light were markedly increased and constantly high in *F. bidentis*, while under circadian regulation in *F. robusta* (Fig. 10A), thus resembling the expression profiles of many C_4, CBC and PR genes. Although transcriptionally down-regulated in C_4 *Flaveria* and *Cleome* plants (Supplemental

Dataset S6 and (Mallmann et al., 2014; Külahoglu et al., 2014; Gowik et al., 2011 and Fig. 10), *CSP41B* represents a promising candidate as part of a signaling network guiding differential regulation of photosynthesis genes. This is supported by highest expression of *CSP41B* the upper third of the *F. bidentis* leaf, where C_4 activity is maximal (Billakurthi, Wrobel, unpublished and Fig. 10B). Furthermore, the expression of *CSP41B* is under circadian control in the C_3 plant *Arabidopsis* (Fig. 10E). It could be speculated, that the transcription of *CSP41B* C_4-specifically acquired light-regulation via respective *cis*-elements and concomitantly the downstream stabilized photosynthetic and other chloroplastic mRNAs were set under light-control.

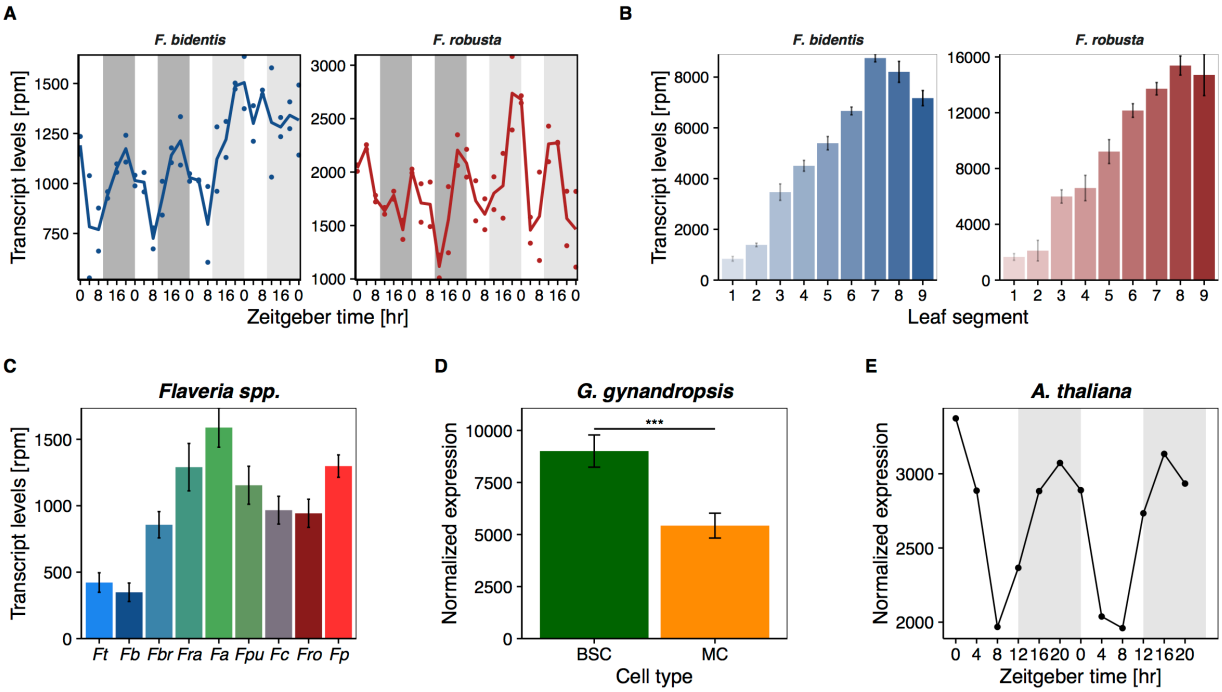


Fig. 10. Transcript levels of CSP41B in different C_4 and *Arabidopsis* transcriptome studies.
A-C, see Fig. 9. D, Comparative transcriptome of BSC and MC levels (Aubry et al., 2014); E, Microarray of *Arabidopsis* circadian response (Covington and Harmer, 2007). See publications for details.

Does circadian clock liberation of photosynthesis increase fitness in C_4?

The benefit of increasing abundances in response to constant light observed in genes related to C_4, CBC, PR (at least in C_3-C_4 species) and CEF in the C_4 species may be obvious: as long as light of sub-stressful levels is available, there should be high activity and synthesis of enzymes to support photosynthesis and increase biomass. However, it has often been shown that the well-timed alignment of photosynthetic processes with the environment through a discrete circadian control grant a fitness advantage (Dodd et al., 2005; Graf et al., 2010). In other words, liberation from said control as seen in the C_4 species would decrease its fitness.

In the C_3 species *F. robusta*, overall transcript abundances of photosynthetic genes did not drastically alter after transition to constant light, i.e. neither a clear up- or downregulation, while the levels of most of these genes seemed to constantly increase with prolonged light in *F. bidentis*. Perhaps the photosynthetic capacity is simply saturated in C_3 and a further transcriptional increase alone would not increase photosynthesis and hence be wasteful. If the circadian response, which from a C_4 point of view resembles the repeated downregulation of these genes, would prevent or is the result of light stress or over-energization, we would possibly rather see downregulation of the PS genes and detect a typical light stress response in *F. robusta*, which was not the case.

The patterns described here in detail are not enough to speculate, that clock-liberated gene expression enhances fitness in C_4 plants. In the future, *F. bidentis* mutant lines with knocked-down circadian clock genes or the candidates *CSP41B* and *NDPK2*, especially when realized BSC-specifically, combined with assays of growth, carbohydrate and protein content, and seed set could shed light on the advantage of clock-liberation. In addition, it would be interesting to compare the transcriptomes of C_3 and C_4 species in response to prolonged (i.e. longer than two days) constant light or high light intensities. Is high transcription of photosynthetic genes maintained over many days?

BSC-specific circadian clock liberation

Genes encoding for C_4 BSC-specific processes, e.g. CBC and photorespiration, appeared to be enriched amongst those genes exhibiting the pattern of high expression in high light in *F. bidentis* contrasted by circadian *F. robusta* expression. Mathematical approaches to describe and quantify this pattern or identify co-regulated candidate genes, for instance via *k*-means

clustering, have been challenging so far, but may improve with the design of algorithms targeted at analysis of time-series data at RNA-Seq scale (Bar-Joseph et al., 2012; Omranian et al., 2015).

The observation, that photorespiration appears to be largely light-regulated although less essential and thus lowly expressed in the C_4 plant may be an evolutionary relic. In C_3-C_4 plants, where up to 43% of the CO_2 may be fixed through the BSC-specific activity of GDC (Vogan and Sage, 2011). BSC-specific light-regulation of photorespiration might be similarly beneficial as is light-regulation of C_4-cycle enzymes for C_4 plants. To our knowledge, diurnal or circadian accumulation of GDC-P or other photorespiration enzymes has so far not been tested in either C_3 or C_4 plants. In favor of the hypothesis of BSC-specific CC-liberation, (i) the BSC-specific NDH subunits but not PGR genes show the pattern, while (ii) genes that are not restricted to the BSC, e.g. PSII and PSI show a similar but distinct pattern. However, this cell-specific distinction does not hold true for some C_4 genes, which show the pattern despite their MC-specificity (e.g. *PPC*, *AMK*).

The pattern could simply be guided by cell-type specific transcription factor or regulator set under light-control. Significantly enriched levels of *CSP41B* in *Cleome* BSC compared to MC could support this candidate's role in the C_4 specific light-regulation (Fig. 10D). Again, *Flaveria* genome and cell type specific transcriptome data will shed light on this aspect and help to disentangle cell specific and isogene-specific regulation.

Conclusion and perspective

This study showed largely overlapping diurnal timing of expression in two *Flaveria* leaf transcriptomes. While the amplitudes of core C_4-cycle gene expression were markedly increased in the C_4 species *F. bidentis* compared to the C_3 species *F. robusta*, the phases were mainly identical, raising the possibility for a diurnal regulation mechanism of the C_4-cycle that existed already before the evolution of C_4 in *Flaveria*. In response to constant light many photosynthesis related genes, including C_4-cycle, CBC, photorespiration and photosystems, showed highly increasing transcript levels and diminished periodicity in *F. bidentis*, contrasted by circadian regulation in *F. robusta*. This suggests liberation of photosynthetic processes from control of the circadian clock in C_4.

Further experiments are required to determine, whether the changes in transcript levels actually translate to the protein level. The advent of a *Flaveria* genome will facilitate to analyze acquisition or loss of *cis*-regulatory elements. This will help to disentangle the contribution of light regulation through phytochromes versus circadian clock regulation. In addition, promoter-

luciferase assays will support the distinction between transcriptional or post-transcriptional regulation of these genes. If C_4-specific liberation from circadian clock regulation can be confirmed on the protein (and metabolic and physiological) level, it will be interesting to analyze whether re-setting photosynthetic genes under clock control in *F. bidentis* will result in a decrease in photosynthetic efficiency. Consequently, efforts to engineer C_4 into C_3 plants could be substantially improved by integrating clock liberated gene expression.

SUPPLEMENTAL DATA

Supplemental Dataset S1. Quantitative information and annotation for all reads mapped onto the reference genome of *Arabidopsis*.
Supplemental Dataset S2. Significantly diurnal genes in either species (JTK_cycle, $q < 0.05$)
Supplemental Dataset S3. Significantly circadian genes in either species (JTK_cycle, $q < 0.05$)
Supplemental Dataset S4. Genes with differential expression at species peak expression (DESeq2, $q < 0.01$)
Supplemental Dataset S5. C4-cycle genes
Supplemental Dataset S6. Gene categories analyzed in detail

Supplemental Figure S1. Periodicity analysis of the circadian subset analogous to Fig. 1 B-D.
Supplemental Figure S2. Principle component analysis of transcript abundances in the diurnal and circadian subsets.
Supplemental Figure S3. Venn diagram of differential gene expression under diurnal and circadian conditions
Supplemental Figure S4. Additional isogenes of the C_4 cycle.
Supplemental Figure S5. Additional isogenes encoding for photorespiratory enzymes.
Supplemental Figure S6. Light signalling.
Supplemental Figure S7. Genes encoding for additional proteins associated to the circadian clock.
Supplemental Figure S8. Circadian response of selected C_4-cycle genes in *A. thaliana*.
Supplemental Figure S9. Genes encoding for proteins involved in sugar signaling.

Supplemental Table S1. mRNA-Seq statistics.

ACKNOWLEDGEMENTS

We acknowledge the excellent technical assistance of U. Gowik with plant growth, M. Sommer with collection of samples for mRNA-Seq and S. Kurz with cDNA library preparations. This work was supported by the International Graduate Program for Plant Science (iGRAD-Plant; to D. B.), grants from Deutsche Forschungsgemeinschaft (DFG) (EXC 1028; IRTG 1525 to A. P. M. W.).

REFERENCES

Aldous, S.H., Weise, S.E., Sharkey, T.D., Waldera-Lupa, D.M., Stühler, K., Mallmann, J., Groth, G., Gowik, U., Westhoff, P., and Arsova, B. (2014). Evolution of the Phospho*enol*pyruvate Carboxylase Protein Kinase Family in C3 and C4 *Flaveria* spp. Plant Physiol **165**: 1076–1091.

Aubry, S., Kelly, S., Kümpers, B.M.C., Smith-Unna, R.D., and Hibberd, J.M. (2014). Deep Evolutionary Comparison of Gene Expression Identifies Parallel Recruitment of Trans - Factors in Two Independent Origins of C 4 Photosynthesis. PLoS Genet. **10**: e1004365.

Bar-Joseph, Z., Gitter, A., and Simon, I. (2012). Studying and modelling dynamic biological processes using time-series gene expression data. Nat Rev Genet **13**: 552–564.

Bollenbach, T.J., Sharwood, R.E., Gutierrez, R., Lerbs-Mache, S., and Stern, D.B. (2009). The RNA-binding proteins CSP41a and CSP41b may regulate transcription and translation of chloroplast-encoded RNAs in Arabidopsis. Plant Mol Biol **69**: 541–552.

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.

Chen, M., Chory, J., and Fankhauser, C. (2004). Light Signal Transduction in Higher Plants. Annu. Rev. Genet. **38**: 87–117.

Chitty, J.A., Furbank, R.T., Marshall, J.S., Chen, Z., and Taylor, W.C. (1994). Genetic transformation of the C4 plant, *Flaveria bidentis*. Plant J. **6**: 949–956.

Choi, G., Yi, H., Lee, J., Kwon, Y.K., Soh, M.S., Shin, B.C., Luka, Z., Hahn, T.R., and Song, P.S. (1999). Phytochrome signalling is mediated through nucleoside diphosphate kinase 2. Nature **401**: 610–613.

Christie, J.M. and Briggs, W.R. (2001). Blue light sensing in higher plants. J. Biol. Chem. **276**: 11457–11460.

Cockburn, W., Whitelam, G.C., Broad, A., and Smith, J. (1996). The participation of phytochrome in the signal transduction pathway of salt stress responses in *Mesembryanthemum crystallinum* L. J Exp Bot **47**: 647–653.

Covington, M.F. and Harmer, S.L. (2007). The circadian clock regulates auxin signaling and responses in *Arabidopsis*. PLoS Biol. **5**: e222.

Covington, M.F., Maloof, J.N., Straume, M., Kay, S.A., and Harmer, S.L. (2007). Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development. Genome Biol **9**: R130.1–R130.18.

Deckard, A., Anafi, R.C., Hogenesch, J.B., Haase, S.B., and Harer, J. (2013). Design and analysis of large-scale biological rhythm studies: a comparison of algorithms for detecting periodic signals in biological data. Bioinformatics **29**: 3174–3180.

Dodd, A.N., Salathia, N., Hall, A., Kévei, E., Tóth, R., Nagy, F., Hibberd, J.M., Millar, A.J.,

and Webb, A.A.R. (2005). Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. Science **309**: 630–633.

Drincovich, M., Casati, P., Andreo, C., Chessin, S., Franceschi, V., Edwards, G., and Ku, M. (1998). Evolution of C4 photosynthesis in *Flaveria* species. Isoforms of NADP-malic enzyme. Plant Physiol **117**: 733–744.

Edwards, K.D., Anderson, P.E., Hall, A., Salathia, N.S., Locke, J., Lynn, J.R., Straume, M., Smith, J.Q., and Millar, A.J. (2006). FLOWERING LOCUS C mediates natural variation in the high-temperature response of the Arabidopsis circadian clock. Plant Cell **18**: 639–650.

Ehleringer, J.R. and Monson, R.K. (1993). Evolutionary and ecological aspects of photosynthetic pathway variation. Annual Review of Ecology and Systematics **24**: 411–439.

Endo, M., Shimizu, H., Nohales, M.A., Araki, T., and Kay, S.A. (2014). Tissue-specific clocks in Arabidopsis show asymmetric coupling. Nature: 1–16.

Friso, G., Giacomelli, L., Ytterberg, A.J., Peltier, J.-B., Rudella, A., Sun, Q., and Wijk, K.J.V. (2004). In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: new proteins, new functions, and a plastid proteome database. Plant Cell **16**: 478–499.

Furbank, R.T. (2011). Evolution of the C4 photosynthetic mechanism: are there really three C4 acid decarboxylation types? J Exp Bot **62**: 3103–3108.

Gehan, M.A., Greenham, K., Mockler, T.C., and McClung, C.R. (2015). Transcriptional networks - crops, clocks, and abiotic stress. Current Opinion in Plant Biology **24**: 39–46.

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.

Gowik, U., Burscheidt, J., Akyildiz, M., Schlue, U., Koczor, M., Streubel, M., and Westhoff, P. (2004). *cis*-Regulatory elements for mesophyll-specific gene expression in the C4 plant *Flaveria trinervia*, the promoter of the C4 phosphoenolpyruvate carboxylase gene. Plant Cell **16**: 1077–1090.

Graf, A., Schlereth, A., Stitt, M., and Smith, A.M. (2010). Circadian control of carbohydrate availability for growth in *Arabidopsis* plants at night. Proc. Natl. Acad. Sci. U.S.A. **107**: 9458–9463.

Greenham, K. and McClung, C.R. (2015). Integrating circadian dynamics with physiological processes in plants. Nat Rev Genet **16**: 598–610.

Groben, R., Kaloudas, D., Raines, C.A., Offmann, B., Maberly, S.C., and Gontero, B. (2010). Comparative sequence analysis of CP12, a small protein involved in the formation of a Calvin cycle complex in photosynthetic organisms. Photosyn. Res. **103**: 183–194.

Harmer, S.L. (2009). The circadian system in higher plants. Annu. Rev. Plant Biol. **60**: 357–377.

Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.S., Han, B., Zhu, T., Wang, X.,

Kreps, J.A., and Kay, S.A. (2000). Orchestrated Transcription of Key Pathways in *Arabidopsis* by the Circadian Clock. Science **290**: 2110–2113.

Hassidim, M., Yakir, E., Fradkin, D., Hilman, D., Kron, I., Keren, N., Harir, Y., Yerushalmi, S., and Green, R.M. (2007). Mutations in CHLOROPLAST RNA BINDING provide evidence for the involvement of the chloroplast in the regulation of the circadian clock in *Arabidopsis*. Plant J. **51**: 551–562.

Hatch, M.D. (1987). C4 photosynthesis: a unique blend of modified biochemistry, anatomy and ultrastructure. BBA Reviews On Bioenergetics **895**: 81–106.

Hibberd, J.M. and Covshoff, S. (2010). The regulation of gene expression required for C4 photosynthesis. Annu. Rev. Plant Biol. **61**: 181–207.

Höfer, M.U., Santore, U.J., and Westhoff, P. (1992). Differential accumulation of the 10-, 16- and 23-kDa peripheral components of the water-splitting complex of photosystem II in mesophyll and bundle-sheath chloroplasts of the dicotyledonous C4 plant *Flaveria trinervia* (Spreng.) C. Mohr. Planta **186**: 304–312.

Hsu, P.Y. and Harmer, S.L. (2012). Circadian phase has profound effects on differential expression analysis. PLoS ONE **7**: e49853.

Hsu, P.Y. and Harmer, S.L. (2014). Global profiling of the circadian transcriptome using microarrays. Methods Mol. Biol. **1158**: 45–56.

Hughes, M.E., Hogenesch, J.B., and Kornacker, K. (2010). JTK_CYCLE: An Efficient Nonparametric Algorithm for Detecting Rhythmic Components in Genome-Scale Data Sets. Journal of Biological Rhythms **25**: 372–380.

Ifuku, K., Endo, T., Shikanai, T., and Aro, E.-M. (2011). Structure of the Chloroplast NADH Dehydrogenase-Like Complex: Nomenclature for Nuclear-Encoded Subunits. Plant Cell Physiol. **52**: 1560–1568.

Imaizumi, T., Tran, H.G., Swartz, T.E., Briggs, W.R., and Kay, S.A. (2003). FKF1 is essential for photoperiodic-specific light signalling in *Arabidopsis*. Nature **426**: 302–306.

Jensen, P.E., Bassi, R., Boekema, E.J., Dekker, J.P., Jansson, S., Leister, D., Robinson, C., and Scheller, H.V. (2007). Structure, function and regulation of plant photosystem I. Biochimica Et Biophysica Acta-Bioenergetics **1767**: 335–352.

Khan, S., Rowe, S.C., and Harmon, F.G. (2010). Coordination of the maize transcriptome by a conserved circadian clock. BMC Plant Biology **10**: 126–115.

Kim, J.I., Hong, S.W., Shin, B., Choi, G., Blakeslee, J.J., Murphy, A.S., Seo, Y.W., Kim, K., Koh, E.J., Song, P.S., and Lee, H. (2005). A possible role for NDPK2 in the regulation of auxin-mediated responses for plant growth and development. Plant Cell Physiol. **46**: 1246–1254.

Kim, Y.-H., Kim, M.D., Choi, Y.I., Park, S.-C., Yun, D.-J., Noh, E.W., Lee, H.-S., and Kwak, S.-S. (2011). Transgenic poplar expressing *Arabidopsis* NDPK2 enhances growth as well as oxidative stress tolerance. Plant Biotechnology Journal **9**: 334–347.

Kim, Y.-H., Lim, S., Yang, K.-S., Kim, C.Y., Kwon, S.-Y., Lee, H.-S., Wang, X., Zhou, Z., Ma, D., Yun, D.-J., and Kwak, S.-S. (2009). Expression of *Arabidopsis* NDPK2 increases antioxidant enzyme activities and enhances tolerance to multiple environmental stresses in transgenic sweet potato plants. Molecular Breeding **24**: 233–244.

Külahoglu, C. et al. (2014). Comparative transcriptome atlases reveal altered gene expression modules between two *Cleomaceae* C3 and C4 plant species. Plant Cell **26**: 3243–3260.

Lai, L.B., Wang, L., and Nelson, T.M. (2002). Distinct but conserved functions for two chloroplastic NADP-malic enzyme isoforms in C-3 and C-4 *Flaveria* species. Plant Physiol **128**: 125–139.

Langdale, J.A. (2011). C4 Cycles: Past, Present, and Future Research on C4 Photosynthesis. Plant Cell **23**: 3879–3892.

Leister, D. (2005). Genomics-based dissection of the cross-talk of chloroplasts with the nucleus and mitochondria in *Arabidopsis*. Gene **354**: 110–116.

Li, P. and Brutnell, T.P. (2011). *Setaria viridis* and *Setaria italica*, model genetic systems for the Panicoid grasses. J Exp Bot **62**: 3031–3037.

Liu, H., Liu, B., Zhao, C., Pepper, M., and Lin, C. (2011). The action mechanisms of plant cryptochromes. Trends in Plant Science **16**: 684–691.

Long, S.P. (1999). Environmental Responses. In C4 Plant Biology (Elsevier), pp. 215–249.

Love, M.I., Huber, W., and Anders, S. (2014). Differential analysis of count data – the DESeq2 package. Genome Biol **15**: 550–48.

Mallmann, J., Heckmann, D., Bräutigam, A., Lercher, M.J., Weber, A.P.M., Westhoff, P., and Gowik, U. (2014). The role of photorespiration during the evolution of C4 photosynthesis in the genus *Flaveria*. eLife **3**: e02478.

Marri, L., Thieulin-Pardo, G., Lebrun, R., and Puppo, R. (2014). CP12-mediated protection of Calvin–Benson-Cycle enzymes from oxidative stress. Biochimie **97**: 228–237.

Marshall, J.S., Stubbs, J.D., and Taylor, W.C. (1996). Two genes encode highly similar chloroplastic NADP-malic enzymes in *Flaveria* - Implications for the evolution of C4 photosynthesis. Plant Physiol **111**: 1251–1261.

McClung, C.R. (2006). Plant circadian rhythms. Plant Cell **18**: 792–803.

Meierhoff, K. and Westhoff, P. (1993). Differential biogenesis of photosystem II in mesophyll and bundle-sheath cells of monocotyledonous NADP-malic enzyme-type C4 plants: the non-stoichiometric abundance of the subunits of photosystem II in the bundle-sheath chloroplasts and the translational activity of the plastome-encoded genes. Planta **191**: 23–33.

Meister, M., Agostino, A., and Hatch, M. (1996). The roles of malate and aspartate in C4 photosynthetic metabolism of *Flaveria bidentis* (L.). Planta **199**.

Michael, T.P. et al. (2008). Network discovery pipeline elucidates conserved time-of-day-specific cis-regulatory modules. PLoS Genet. **4**: e14.

Monson, R.K. (2003). Gene duplication, neofunctionalization, and the evolution of C4 photosynthesis. Int. J Plant Sci. **164**: S43–S54.

Moon, H. et al. (2003). NDP kinase 2 interacts with two oxidative stress-activated MAPKs to regulate cellular redox state and enhances multiple stress tolerance in transgenic plants. Proc. Natl. Acad. Sci. U.S.A. **100**: 358–363.

Moore, B.D. and Edwards, G.E. (1986). Photosynthetic Induction in a C4 Dicot, *Flaveria-Trinervia* .1. Initial Products of (Co2)-C-14 Assimilation and Levels of Whole Leaf C-4 Metabolites. Plant Physiol **81**: 663–668.

Munekage, Y., Hashimoto, M., Miyake, C., Tomizawa, K.-I., Endo, T., Tasaka, M., and Shikanai, T. (2004). Cyclic electron flow around photosystem I is essential for photosynthesis. Nature **429**: 579–582.

Munekage, Y.N. and Taniguchi, Y.Y. (2016). Promotion of Cyclic Electron Transport around Photosystem I with the Development of C4 Photosynthesis. Plant Cell Physiol. **57**: 897–903.

Nakajima Munekage, Y. (2016). Light harvesting and chloroplast electron transport in NADP-malic enzyme type C4 plants. Current Opinion in Plant Biology **31**: 9–15.

Nakamichi, N. (2011). Molecular mechanisms underlying the *Arabidopsis* circadian clock. Plant Cell Physiol. **52**: 1709–1718.

Nakamura, N., Iwano, M., Havaux, M., Yokota, A., and Munekage, Y.N. (2013). Promotion of cyclic electron transport around photosystem I during the evolution of NADP–malic enzyme-type C4 photosynthesis in the genus *Flaveria*. New Phytol **199**: 832–842.

Nickelsen, J. and Rengstl, B. (2013). Photosystem II Assembly: From Cyanobacteria to Plants. Annu. Rev. Plant Biol. **64**: 609–635.

Nimmo, H.G. (2003). Control of the phosphorylation of phosphoenolpyruvate carboxylase in higher plants. Arch. Biochem. Biophys. **414**: 189–196.

Omranian, N., Mueller-Roeber, B., and Nikoloski, Z. (2015). Segmentation of biological multivariate time-series data. Sci. Rep. **5**: 8937–6.

Qi, Y., Armbruster, U., Schmitz-Linneweber, C., Delannoy, E., de Longevialle, A.F., Rühle, T., Small, I., Jahns, P., and Leister, D. (2012). Arabidopsis CSP41 proteins form multimeric complexes that bind and stabilize distinct plastid transcripts. J Exp Bot **63**: 1251–1270.

Quail, P.H. (2002). Phytochrome photosensory signalling networks. Nat Rev Mol Cell Biol **3**: 85–93.

Richly, E. and Leister, D. (2004). An improved prediction of chloroplast proteins reveals diversities and commonalities in the chloroplast proteomes of *Arabidopsis* and rice. Gene **329**: 11–16.

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.

Rosche, E. and Westhoff, P. (1995). Genomic structure and expression of the pyruvate, *ortho*phosphate dikinase gene of the dicotyledonous C4 plant *Flaveria trinervia* (*Asteraceae*). Plant Mol Biol **29**: 663–678.

Sage, R.F. (2004). The evolution of C4 photosynthesis. New Phytol **161**: 341–370.

Sage, R.F., Sage, T.L., and Kocacinar, F. (2012). Photorespiration and the Evolution of C4 Photosynthesis. Annu. Rev. Plant Biol. **63**: 19–47.

Schnarrenberger, C. and Martin, W. (2002). Evolution of the enzymes of the citric acid cycle and the glyoxylate cycle of higher plants. A case study of endosymbiotic gene transfer. Eur. J. Biochem. **269**: 868–883.

Schulze, S., Mallmann, J., Burscheidt, J., Koczor, M., Streubel, M., Bauwe, H., Gowik, U., and Westhoff, P. (2013). Evolution of C4 photosynthesis in the genus *Flaveria*: establishment of a photorespiratory CO2 pump. Plant Cell **25**: 2522–2535.

Schulze, S., Westhoff, P., and Gowik, U. (2016). Glycine decarboxylase in C3, C4 and C3-C4 intermediate species. Current Opinion in Plant Biology **31**: 29–35.

Sharrock, R.A. and Clack, T. (2002). Patterns of expression and normalized levels of the five *Arabidopsis* phytochromes. Plant Physiol **130**: 442–456.

Sheen, J. (1999). C4 gene expression. Annu. Rev. Plant Physiol. Plant Mol. Biol. **50**: 187–217.

Shikanai, T. (2007). Cyclic electron transport around photosystem I: genetic approaches. Annu. Rev. Plant Biol. **58**: 199–217.

Stockhaus, J., Schlue, U., Koczor, M., Chitty, J.A., Taylor, W.C., and Westhoff, P. (1997). The promoter of the gene encoding the C4 form of phosphoenolpyruvate carboxylase directs mesophyll-specific expression in transgenic C4 *Flaveria* spp. Plant Cell **9**: 479–489.

Sullivan, S., Shenton, M., and Nimmo, H.G. (2005). Organ specificity in the circadian control of plant gene expression. Biochem. Soc. Trans. **33**: 943–944.

Taybi, T., Patil, S., Chollet, R., and Cushman, J.C. (2000). A minimal serine/threonine protein kinase circadianly regulates phosphoenolpyruvate carboxylase activity in crassulacean acid metabolism-induced leaves of the common ice plant. Plant Physiol **123**: 1471–1482.

Tsuchida, Y., Furumoto, T., Izumida, A., Hata, S., and Izui, K. (2001). Phospho*enol*pyruvate carboxylase kinase involved in C 4photosynthesis in *Flaveria trinervia*: cDNA cloning and characterization 1. FEBS Letters **507**: 318–322.

Vogan, P.J. and Sage, R.F. (2011). Water-use efficiency and nitrogen-use efficiency of C3-C4 intermediate species of *Flaveria Juss.* (*Asteraceae*). Plant Cell Environ. **34**: 1415–1430.

Wang, L. et al. (2014a). Comparative analyses of C_4 and C_3 photosynthesis in developing leaves of maize and rice. Nat Biotechnol **32**: 1158–1165.

Wang, P., Fouracre, J., Kelly, S., Karki, S., Gowik, U., Aubry, S., Shaw, M.K., Westhoff, P., Slamet-Loedin, I.H., Quick, W.P., Hibberd, J.M., and Langdale, J.A. (2013).

Evolution of GOLDEN2-LIKE gene function in C3 and C4 plants. Planta **237**: 481–495.

Wang, Y., Bräutigam, A., Weber, A.P.M., and Zhu, X.-G. (2014b). Three distinct biochemical subtypes of C_4 photosynthesis? A modelling analysis. J Exp Bot **65**: 3567–3578.

Wang, Z., Li, H., Ke, Q., Jeong, J.C., Lee, H.-S., Xu, B., Deng, X.-P., Lim, Y.P., and Kwak, S.-S. (2014c). Transgenic alfalfa plants expressing AtNDPK2 exhibit increased growth and tolerance to abiotic stresses. Plant Physiol. Biochem. **84**: 67–77.

Westhoff, P. and Gowik, U. (2010). Evolution of C4 photosynthesis - looking for the master switch. Plant Physiol **154**: 598–601.

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 C4 *Flaveria*: an intricate interplay of transcriptional and posttranscriptional processes. Plant Cell **24**: 137–151.

Zoltowski, B.D. and Imaizumi, T. (2013). Structure and Function of the ZTL/FKF1/LKP2 Group Proteins in *Arabidopsis*. The Enzymes **35**: 213–239.

FIGURE LEGENDS

Fig. 1. Experimental design and periodicity analysis of the diurnal dataset. A, Time course and representative pictures of plants during the experiment. Leaves were collected at the time points indicated in zeitgeber time (ZT) during two consecutive cycles each under light-dark ("diurnal") and constant light ("circadian") conditions. B-D, results of JTK_cycle analysis of the diurnal subset (analogous results of the circadian subset are presented in Supplemental Fig. S2). B, Venn diagram of genes identified as periodic (JTK_cycle, $q < 0.05$) in the two *Flaveria* species under diurnal conditions. Numbers outside the venn diagram represent non-periodic genes. C, Phase of expression of significantly periodic genes (JTK_cycle, $q < 0.05$). D, Comparison of phases of significantly periodic genes between the two *Flaveria* species. ns, no significant periodicity. Shaded areas in A, C and D indicate night or subjective night (panel A, "circadian").

Fig. 2. Phase-adjusted analysis of differential gene expression. A, Transcript abundances under diurnal conditions of an exemplary gene (At2g42690) with differential phases between the two species. Differential expression is only detected when regarding the expression over the whole diurnal course rather than at midday only (e.g. ZT 4 to 8). B, Venn diagram of differential gene expression between *Flaveria* sister pairs (reproduced from Gowik et al., 2011). In Gowik et al. (2011), differential gene expression was determined by comparison of *F. bidentis* against *F. pringlei* and *F. trinervia* against *F. robusta.* The numbers represent genes with significantly higher expression in the indicated species compared to its sister species. C, Venn diagram comparing the results of phase-adjusted differential expression of the diurnal subset with the results of Gowik et al. (2011). The results presented in panel B, where summarised as C_3 and C_4 specific expression, after removal of contrasting results. See text for explanation.

Fig. 3. Transcript profiles of core C_4-cycle genes. Scheme of the NADP-ME C_4 pathway with expression profiles (in reads per million) of genes encoding for core C_4 enzymes and transporters over the diurnal and circadian time course. Blue, *F. bidentis*; Red, *F. robusta*; dark grey bars, night; light grey bars, subjective night. For detailed explanation of the time course, see Fig. 1. Abbreviations are explained in the text and in Supplemental Dataset S3. Top right: phase of gene expression for significantly periodic genes (JTK_cycle, $q < 0.05$) und diurnal (left clock) and circadian (right clock) conditions. Notes: Key enzymes of other C_4 biochemical subtypes (NAD-ME1 and PEP-CK1) are displayed for comparison. NHD2 levels likely represent NHD1 (as explained in the text). "Circle-head" indicates post-translational regulation. Enzymes with activity in both cell types are only presented once for simplicity. ALAAT1 is likely active in the cytosols of both MC and BSC. Unknown transport proteins are marked with a question mark. 3-PGA, 3-phosphoglycerate; Ala, alanine; Asp, aspartate; Mal, malate; OAA, oxaloacetate, (P)P_i, (pyro)phosphate; Pyr, pyruvate, TP, triose-phosphate

Fig. 4. Transcript profiles of genes encoding for proteins involved in light reactions.
A, Photosystem II (PSII) and light harvesting complex (LHC) of PSII; B, ferredoxin:plastoquinone oxidoreductase pathway of CEF; C, Subunits of the chloroplast NADH dehydrogenase-like (NDH) complex. D, PSI and LHC of PSI; E, ATP synthase subunits. Except for panel C (Nomenclature and categorisation based on Ifuku et al. (2011), organisation with respect to thylakoid membrane an, storm and lumen is only schematic for simplicity. Genes in panel D based on Jensen et al. (2007). Only major isoforms and subunits present in the minimal reference shown (in particular, plastome encoded genes absent). Legend, see Fig. 1A and Fig. 3. Full gene names see text and Supplemental Dataset S6.

Fig. 5. Transcript profiles of CBC genes and CP12-3.
A, Scheme of CBC. Only major isoforms presented. B, CP12-3. Legend, see Fig. 1A and Fig. 3. Full gene names see text and Supplemental Dataset S6.

Fig. 6. Transcript profiles of photorespiration genes.
Scheme of photorespiration across chloroplast (green), peroxisome (blue) and mitochondrium (red). Only major isoforms presented. Legend, see Fig. 1A and Fig. 3. Full gene names see text and Supplemental Dataset S6.

Fig. 7. Transcript profiles of photoreceptor genes.
A, Phytochromes; B, Cryptochromes; C, Phototropins; D, FKF/LKP/ZTL family proteins.
Legend, see Fig. 1A and Fig. 3. Full gene names see text and Supplemental Dataset S6.

Fig. 8. Transcript profiles of core circadian clock genes.
Legend, see Fig. 1A and Fig. 3. Full gene names see text and Supplemental Dataset S6. See Greenham and McClung (2015) for details.

Fig. 9. Transcript levels of NDPK2 in different C_4 transcriptome studies.
A, Transcript profile this study (see Fig. 1A for legend); B, Comparative transcriptome of *Flaveria* leaf segments (Billakurthi, Wrobel, unpublished) ranging from bottom (1) to tip (9); C, Study of Mallmann et al. (2014) comparing the levels in C_4 (*Ft*, *F. trinervia*; *Fb, F. bidentis*), C_4-like (Fbr, *F. brownii*), C_3-C_4 (*Fra*, *F. ramosissima*; *Fa*, *F. anomala*; *Fpu*, *F. pubescens*; *Fc*, *F. chloraefolia*) and C_3 (*Fr*, *F. robusta*; *Fp*, *F. pubescens*) species of *Flaveria*.

Fig. 10. Transcript levels of CSP41B in different C_4 and *Arabidopsis* transcriptome studies.
A-C, see Fig. 9. D, Comparative transcriptome of BSC and MC levels (Aubry et al., 2014); E, Microarray of *Arabidopsis* circadian response (Covington and Harmer, 2007). See publications for details.

SUPPLEMENTAL TABLE

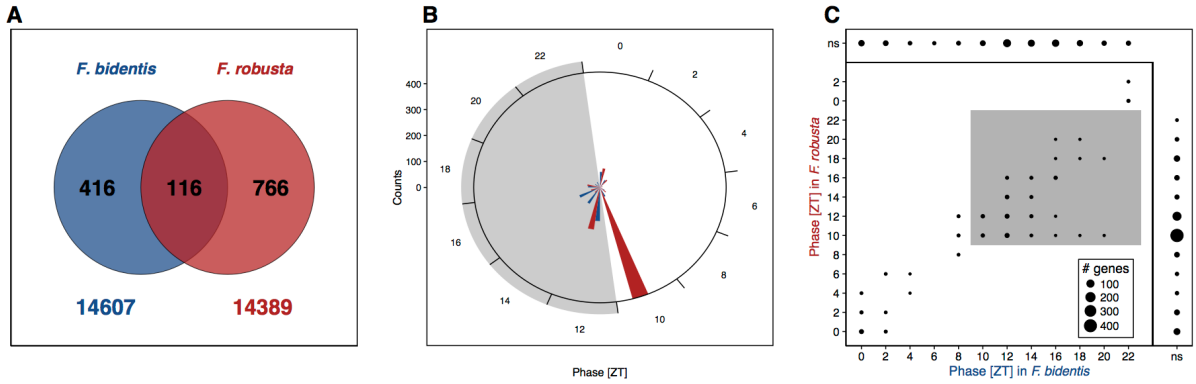
Supplemental Table S1. mRNA-Seq statistics.

Species	TimeAfterSampling [hr]	Zeitgeber Time	Light or Dark	Diurnal or circadian	Replicate	Sample	Total reads	Mapped reads	Mapping efficiency	Genes with at least one read	Sum total reads	Sum mapped reads	Averge mapping efficiency	Genes with at least one read in rep group	Representation of reduc TAIR10 reference (21,8 genes total)
F. bidentis	0	ZT00	Light	diurnal	Replicate1	Fb00_rep1	17,600,749	8,972,813	51.0%	13,970	34,007,922	17,339,111	51.0%	14,286	65.3%
					Replicate 2	Fb00_rep2	16,407,173	8,366,298	51.0%	13,908					
	4	ZT04			Replicate1	Fb04_rep1	18,592,961	7,171,364	38.6%	14,020	37,599,061	17,617,778	46.8%	14,387	65.8%
					Replicate 2	Fb04_rep2	19,006,100	10,446,414	55.0%	13,908					
	8	ZT08			Replicate1	Fb08_rep1	17,111,245	9,370,544	54.8%	13,938	30,463,843	16,425,421	53.8%	14,336	65.6%
					Replicate 2	Fb08_rep2	13,352,598	7,054,877	52.8%	13,956					
	12	ZT12	Dark		Replicate1	Fb12_rep1	18,661,320	9,502,515	50.9%	14,186	40,698,832	20,844,201	51.2%	14,415	65.9%
					Replicate 2	Fb12_rep2	22,037,512	11,341,686	51.5%	14,102					
	16	ZT16			Replicate1	Fb16_rep1	16,563,629	8,240,810	49.8%	14,333	34,345,056	17,204,288	50.1%	14,613	66.8%
					Replicate 2	Fb16_rep2	17,781,427	8,963,478	50.4%	14,387					
	20	ZT20			Replicate1	Fb20_rep1	15,815,061	7,923,480	50.1%	14,143	34,948,362	17,489,231	50.1%	14,466	66.1%
					Replicate 2	Fb20_rep2	19,133,301	9,565,751	50.0%	14,221					
	24	ZT00	Light		Replicate1	Fb24_rep1	17,528,960	9,502,423	54.2%	14,003	32,896,758	17,573,961	53.4%	14,193	64.9%
					Replicate 2	Fb24_rep2	15,367,798	8,071,538	52.5%	13,540					
	28	ZT04			Replicate1	Fb28_rep1	19,491,657	10,719,946	55.0%	13,894	38,408,779	21,246,692	55.3%	14,103	64.5%
					Replicate 2	Fb28_rep2	18,917,122	10,526,746	55.7%	13,645					
	32	ZT08			Replicate1	Fb32_rep1	16,043,731	8,489,664	52.9%	13,889	35,877,271	19,221,115	53.5%	14,323	65.5%
					Replicate 2	Fb32_rep2	19,833,540	10,731,451	54.1%	14,059					
	36	ZT12	Dark		Replicate1	Fb36_rep1	16,211,925	8,345,890	51.5%	14,069	37,157,799	19,133,728	51.5%	14,424	66.0%
					Replicate 2	Fb36_rep2	20,945,874	10,787,838	51.5%	14,206					
	40	ZT16			Replicate1	Fb40_rep1	18,487,680	9,321,684	50.4%	14,307	35,553,624	17,885,913	50.3%	14,502	66.3%
					Replicate 2	Fb40_rep2	17,065,944	8,564,229	50.2%	14,180					
	44	ZT20			Replicate1	Fb44_rep1	16,818,834	8,562,935	50.9%	14,005	35,376,012	17,924,110	50.7%	14,593	66.7%
					Replicate 2	Fb44_rep2	18,557,178	9,361,175	50.5%	14,451					
	48	ZT00		diurnal and circadian	Replicate1	Fb48_rep1	15,919,083	8,349,179	52.5%	13,863	32,188,538	16,914,262	52.6%	14,337	65.6%
					Replicate 2	Fb48_rep2	16,269,455	8,565,083	52.7%	14,085					
	52	ZT04			Replicate1	Fb52_rep1	17,911,779	9,970,906	55.7%	13,983	34,581,989	19,324,583	55.9%	14,227	65.1%
					Replicate 2	Fb52_rep2	16,670,210	9,353,677	56.1%	13,801					
	56	ZT08			Replicate1	Fb56_rep1	15,897,003	8,343,402	52.5%	14,032	34,865,362	18,891,506	54.1%	14,277	65.3%
					Replicate 2	Fb56_rep2	18,968,359	10,548,104	55.6%	13,817					
	60	ZT12			Replicate1	Fb60_rep1	17,382,108	8,872,247	51.0%	14,204	38,215,392	20,000,523	52.2%	14,406	65.9%
					Replicate 2	Fb60_rep2	20,833,284	11,128,276	53.4%	13,989					
	64	ZT16			Replicate1	Fb64_rep1	16,868,974	8,736,624	51.8%	14,077	33,703,635	17,805,220	52.8%	14,417	65.9%
					Replicate 2	Fb64_rep2	16,834,661	9,068,596	53.9%	14,134					
	68	ZT20			Replicate1	Fb68_rep1	13,864,344	7,596,633	54.8%	14,115	29,062,393	15,972,660	55.0%	14,313	65.4%
					Replicate 2	Fb68_rep2	15,198,049	8,376,027	55.1%	13,805					
	72	ZT00	Light		Replicate1	Fb72_rep1	18,070,706	9,938,553	55.0%	14,137	34,038,201	18,755,544	55.1%	14,332	65.5%
					Replicate 2	Fb72_rep2	15,967,495	8,816,991	55.2%	13,907					
	76	ZT04			Replicate1	Fb76_rep1	15,882,527	8,632,910	54.4%	14,027	35,138,226	19,454,830	55.3%	14,232	65.1%
					Replicate 2	Fb76_rep2	19,255,699	10,821,920	56.2%	13,811					
	80	ZT08			Replicate1	Fb80_rep1	8,305,385	4,566,591	55.0%	13,839	27,336,743	15,297,874	55.7%	14,224	65.0%
					Replicate 2	Fb80_rep2	19,031,358	10,731,283	56.4%	13,930					
	84	ZT12			Replicate1	Fb84_rep1	19,385,469	9,804,070	50.6%	14,302	30,482,160	14,857,553	48.1%	14,457	66.1%
					Replicate 2	Fb84_rep2	11,096,691	5,053,483	45.5%	13,707					
	88	ZT16			Replicate1	Fb88_rep1	17,895,751	10,088,941	56.4%	13,826	36,126,177	20,143,781	55.8%	14,249	65.2%
					Replicate 2	Fb88_rep2	18,230,426	10,054,840	55.2%	14,070					
	92	ZT20			Replicate1	Fb92_rep1	15,396,415	8,595,725	55.8%	13,834	32,785,014	18,226,771	55.6%	14,331	65.5%
					Replicate 2	Fb92_rep2	17,388,599	9,631,046	55.4%	14,140					
	96	ZT00			Replicate1	Fb96_rep1	18,113,346	9,800,723	54.1%	14,056	27,631,405	15,095,851	54.9%	14,329	65.5%
					Replicate 2	Fb96_rep2	9,518,059	5,295,128	55.6%	13,828					

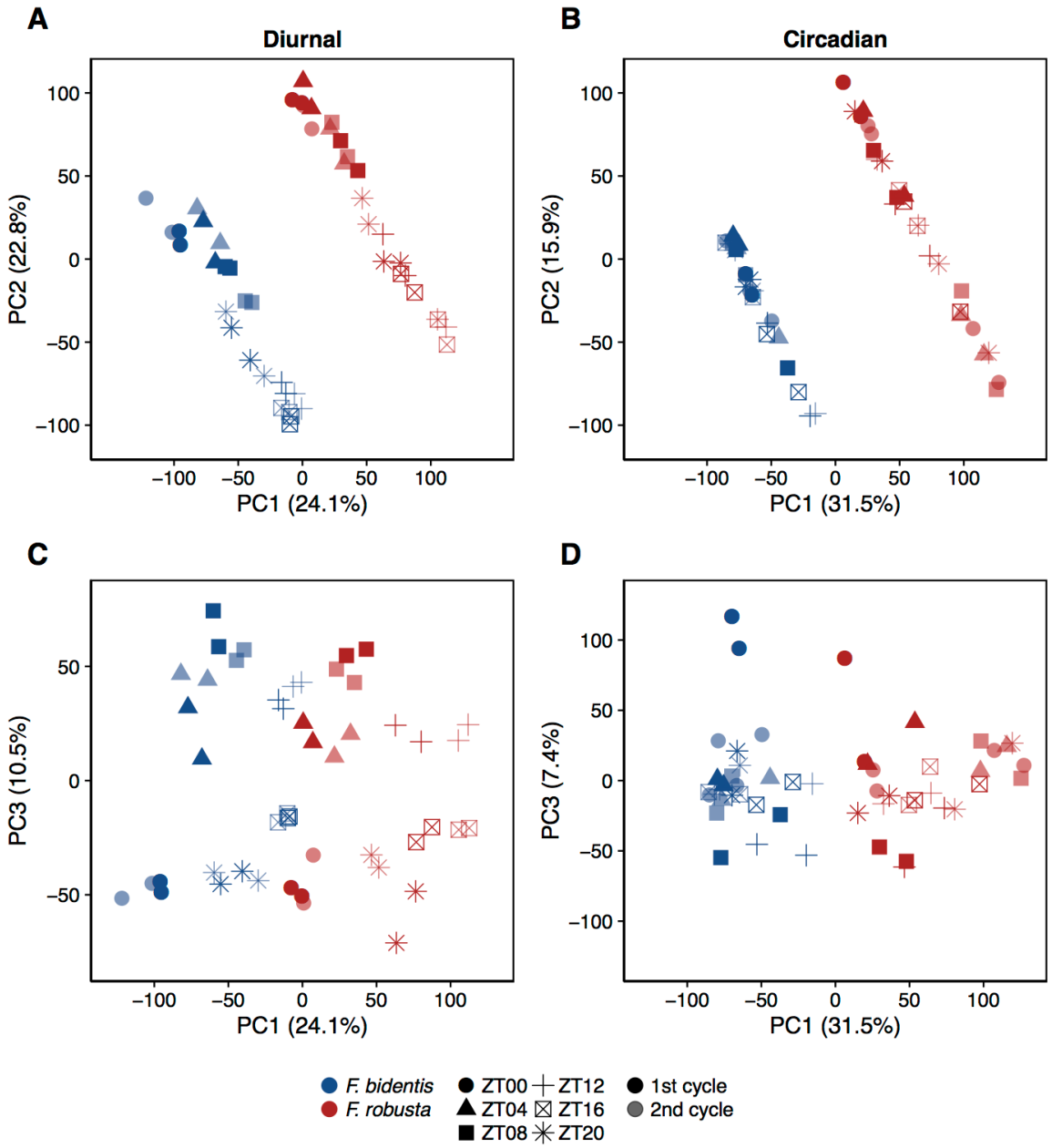
continuation of **Supplemental Table S1.** mRNA-Seq statistics.

Species	TimeAfterSampling [hr]	Zeitgeber Time	Light or Dark	Diurnal or circadian	Replicate	Sample	Total reads	Mapped reads	Mapping efficiency	Genes with at least one read	Sum total reads	Sum mapped reads	Averge mapping efficiency	Genes with at least one read in rep group	Representation of reduc TAIR10 reference (21,86 genes total)
F. bidentis	0	ZT00	Light	diurnal	Replicate1	Fr00_rep1	14,972,796	7,898,003	52.8%	14,038	28,262,077	15,255,095	54.1%	14,336	65.6%
					Replicate 2	Fr00_rep2	13,289,281	7,357,092	55.4%	13,912					
	4	ZT04			Replicate1	Fr04_rep1	7,984,282	4,481,292	56.1%	13,694	26,960,231	15,407,776	56.9%	14,111	64.5%
					Replicate 2	Fr04_rep2	18,975,949	10,926,484	57.6%	13,714					
	8	ZT08			Replicate1	Fr08_rep1	15,949,046	9,076,174	56.9%	13,819	35,134,108	19,792,620	56.4%	14,172	64.8%
					Replicate 2	Fr08_rep2	19,185,062	10,716,446	55.9%	13,888					
	12	ZT12	Dark		Replicate1	Fr12_rep1	9,907,052	5,113,590	51.6%	14,146	31,066,588	16,513,218	52.7%	14,508	66.3%
					Replicate 2	Fr12_rep2	21,159,536	11,399,628	53.9%	14,109					
	16	ZT16			Replicate1	Fr16_rep1	23,389,869	12,077,045	51.6%	14,158	42,314,243	21,848,980	51.6%	14,562	66.6%
					Replicate 2	Fr16_rep2	18,924,374	9,771,935	51.6%	14,386					
	20	ZT20			Replicate1	Fr20_rep1	16,678,258	8,243,012	49.4%	14,230	33,752,842	16,795,996	49.8%	14,579	66.7%
					Replicate 2	Fr20_rep2	17,074,584	8,552,984	50.1%	14,223					
	24	ZT00	Light		Replicate1	Fr24_rep1	16,042,812	9,026,760	56.3%	14,000	28,986,350	16,062,303	55.3%	14,335	65.5%
					Replicate 2	Fr24_rep2	12,943,538	7,035,543	54.4%	13,826					
	28	ZT04			Replicate1	Fr28_rep1	18,338,389	10,122,366	55.2%	13,997	39,701,830	21,833,582	55.0%	14,354	65.6%
					Replicate 2	Fr28_rep2	21,363,441	11,711,216	54.8%	14,003					
	32	ZT08			Replicate1	Fr32_rep1	16,628,825	9,344,536	56.2%	13,801	33,993,819	18,975,505	55.8%	14,231	65.1%
					Replicate 2	Fr32_rep2	17,364,994	9,630,969	55.5%	13,980					
	36	ZT12	Dark		Replicate1	Fr36_rep1	17,384,177	8,761,643	50.4%	14,222	39,467,181	19,863,133	50.3%	14,547	66.5%
					Replicate 2	Fr36_rep2	22,083,004	11,101,490	50.3%	14,320					
	40	ZT16			Replicate1	Fr40_rep1	15,000,302	7,431,858	49.5%	14,368	29,448,296	14,381,047	48.8%	14,726	67.3%
					Replicate 2	Fr40_rep2	14,447,994	6,949,189	48.1%	14,317					
	44	ZT20			Replicate1	Fr44_rep1	16,407,359	8,538,844	52.0%	13,963	36,308,950	18,931,094	52.1%	14,429	66.0%
					Replicate 2	Fr44_rep2	19,901,591	10,392,250	52.2%	14,162					
	48	ZT00		diurnal and circadian	Replicate1	Fr48_rep1	15,531,009	8,283,924	53.3%	13,713	32,205,351	17,315,949	53.8%	14,184	64.9%
					Replicate 2	Fr48_rep2	16,674,342	9,032,025	54.2%	13,821					
	52	ZT04		diurnal	Replicate1	Fr52_rep1	17,731,465	9,378,590	52.9%	14,076	30,941,091	16,611,193	53.8%	14,322	65.5%
					Replicate 2	Fr52_rep2	13,209,626	7,232,603	54.8%	13,786					
	56	ZT08			Replicate1	Fr56_rep1	16,051,018	8,876,572	55.3%	13,865	31,915,021	17,472,086	54.7%	14,286	65.3%
					Replicate 2	Fr56_rep2	15,864,003	8,595,514	54.2%	13,933					
	60	ZT12			Replicate1	Fr60_rep1	18,984,585	9,441,108	49.7%	14,108	40,227,490	20,421,021	50.7%	14,462	66.1%
					Replicate 2	Fr60_rep2	21,242,905	10,979,913	51.7%	14,108					
	64	ZT16			Replicate1	Fr64_rep1	12,414,097	6,436,603	51.9%	14,054	25,854,203	13,663,087	52.8%	14,470	66.2%
					Replicate 2	Fr64_rep2	13,440,106	7,226,484	53.8%	14,066					
	68	ZT20			Replicate1	Fr68_rep1	17,598,070	9,383,612	53.3%	13,942	31,088,753	17,056,412	55.1%	14,310	65.4%
					Replicate 2	Fr68_rep2	13,490,683	7,672,800	56.9%	13,952					
	72	ZT00	Light	diurnal	Replicate1	Fr72_rep1	16,635,103	9,184,970	55.2%	13,889	32,926,678	18,222,954	55.3%	14,243	65.1%
					Replicate 2	Fr72_rep2	16,291,575	9,037,984	55.5%	13,845					
	76	ZT04			Replicate1	Fr76_rep1	18,673,194	9,429,143	50.5%	14,310	35,881,385	18,027,616	50.2%	14,603	66.8%
					Replicate 2	Fr76_rep2	17,208,191	8,598,473	50.0%	14,304					
	80	ZT08			Replicate1	Fr80_rep1	15,591,905	7,680,934	49.3%	14,429	31,741,834	15,980,730	50.3%	14,619	66.8%
					Replicate 2	Fr80_rep2	16,149,929	8,299,796	51.4%	14,080					
	84	ZT12			Replicate1	Fr84_rep1	15,432,584	8,473,558	54.9%	13,944	33,603,711	18,095,939	53.9%	14,420	65.9%
					Replicate 2	Fr84_rep2	18,171,127	9,622,381	53.0%	14,139					
	88	ZT16			Replicate1	Fr88_rep1	17,872,957	9,662,835	54.1%	14,098	33,109,233	17,649,278	53.2%	14,511	66.4%
					Replicate 2	Fr88_rep2	15,236,276	7,986,443	52.4%	14,266					
	92	ZT20			Replicate1	Fr92_rep1	12,680,294	6,143,062	48.5%	14,219	30,011,291	15,124,958	50.1%	14,536	66.5%
					Replicate 2	Fr92_rep2	17,330,997	8,981,896	51.8%	14,128					
	96	ZT00			Replicate1	Fr96_rep1	15,023,831	7,544,208	50.2%	14,386	27,485,641	13,548,535	49.2%	14,685	67.1%
					Replicate 2	Fr96_rep2	12,461,810	6,004,327	48.2%	14,298					

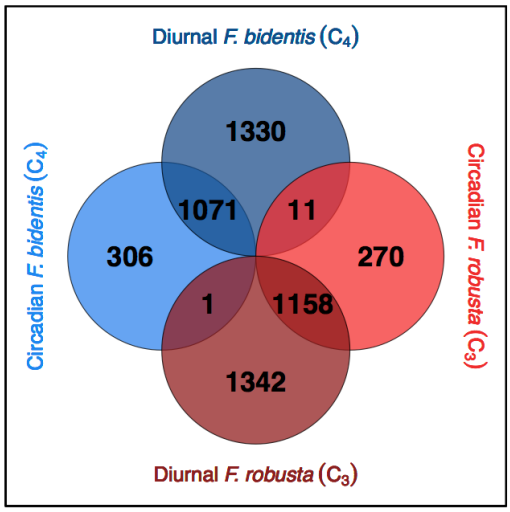
SUPPLEMENTAL FIGURES



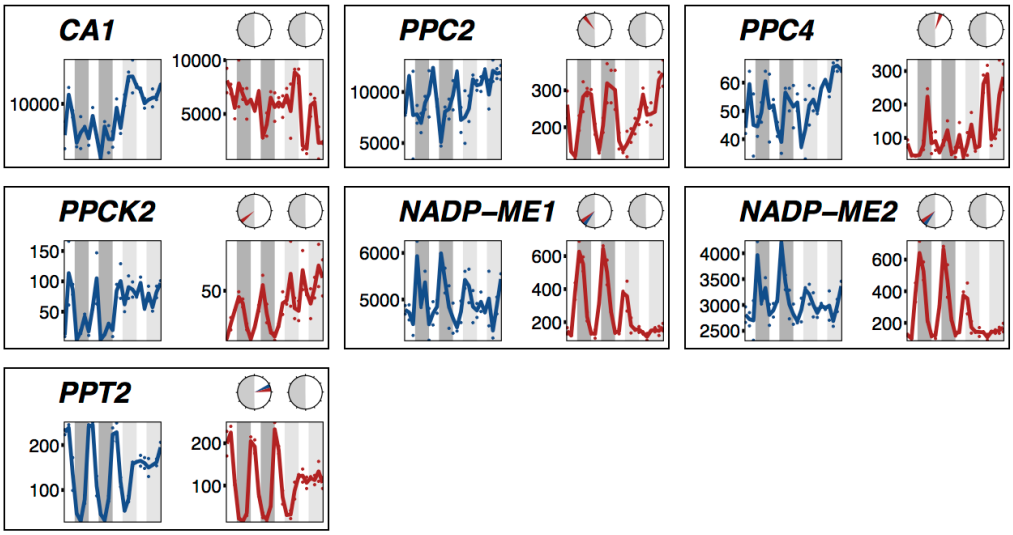
Supplemental Figure S1. Periodicity analysis of the circadian subset analogous to Fig. 1 B-D. For explanation, see Figure 1.



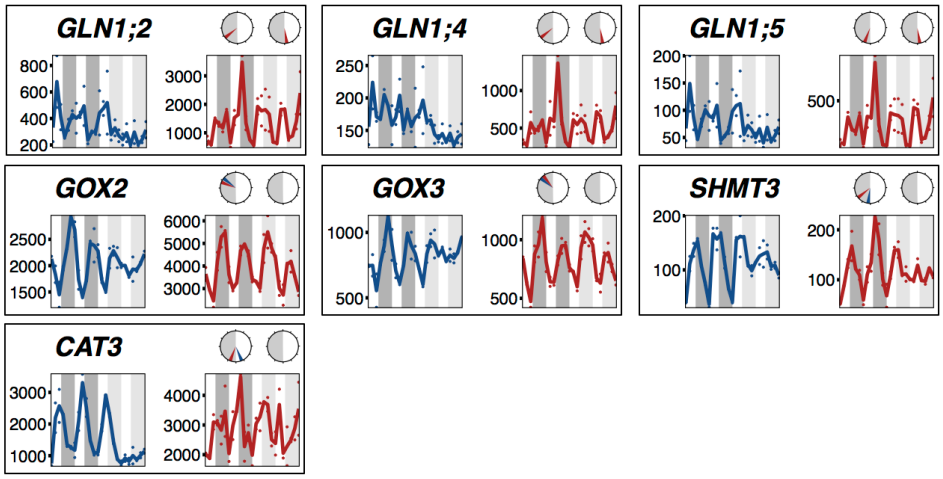
Supplemental Figure S2. Principle component analysis of transcript abundances in the diurnal (A, C) and circadian (B, D) subsets. PC, principle component.



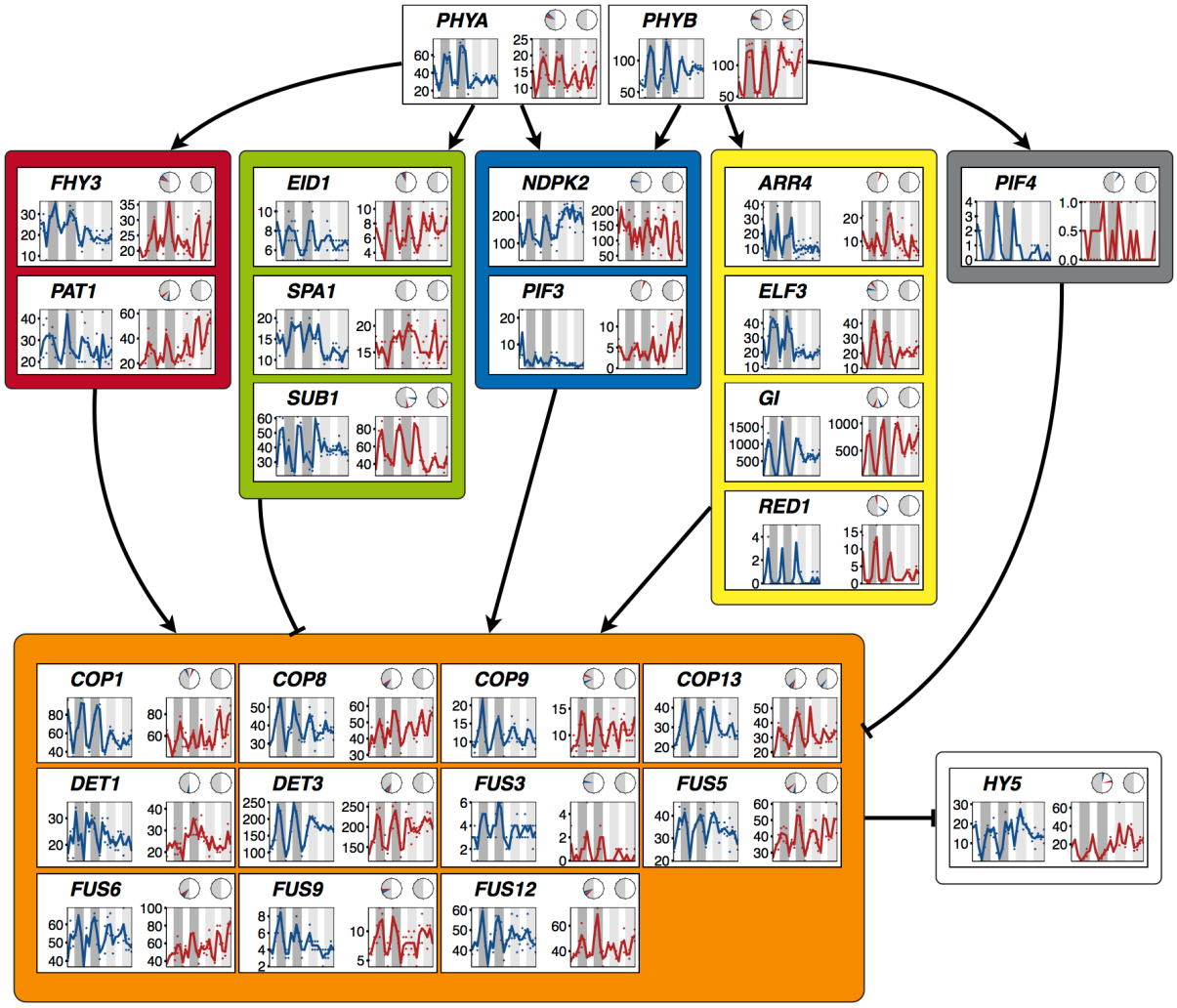
Supplemental Figure S3. Venn diagram of differential gene expression under diurnal and circadian conditions



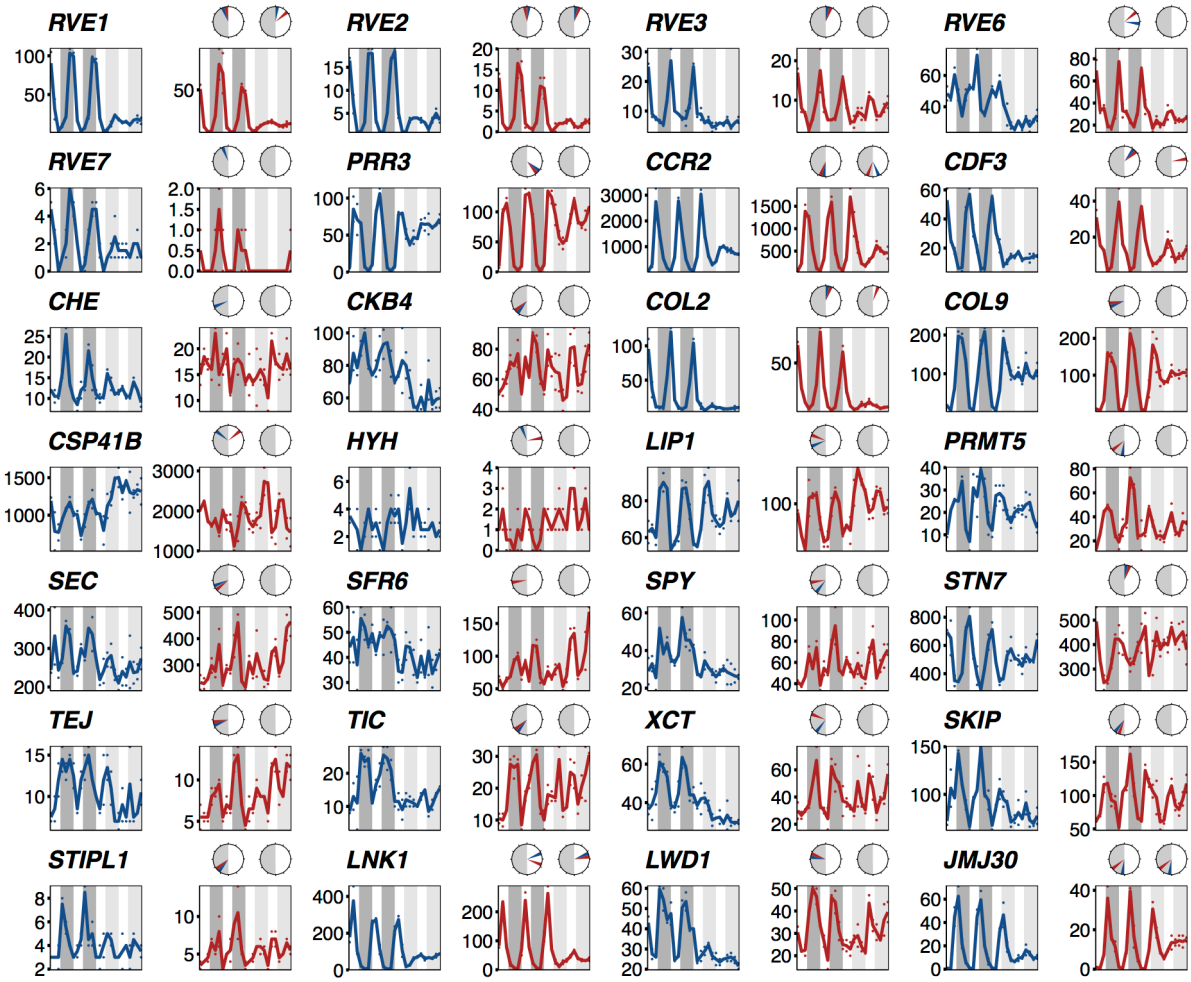
Supplemental Figure S4. Additional isogenes of the C_4 cycle. For explanation, legend and context see Fig. 1 and Fig. 3. *CA1*, *CARBONIC ANHYDRASE 1*



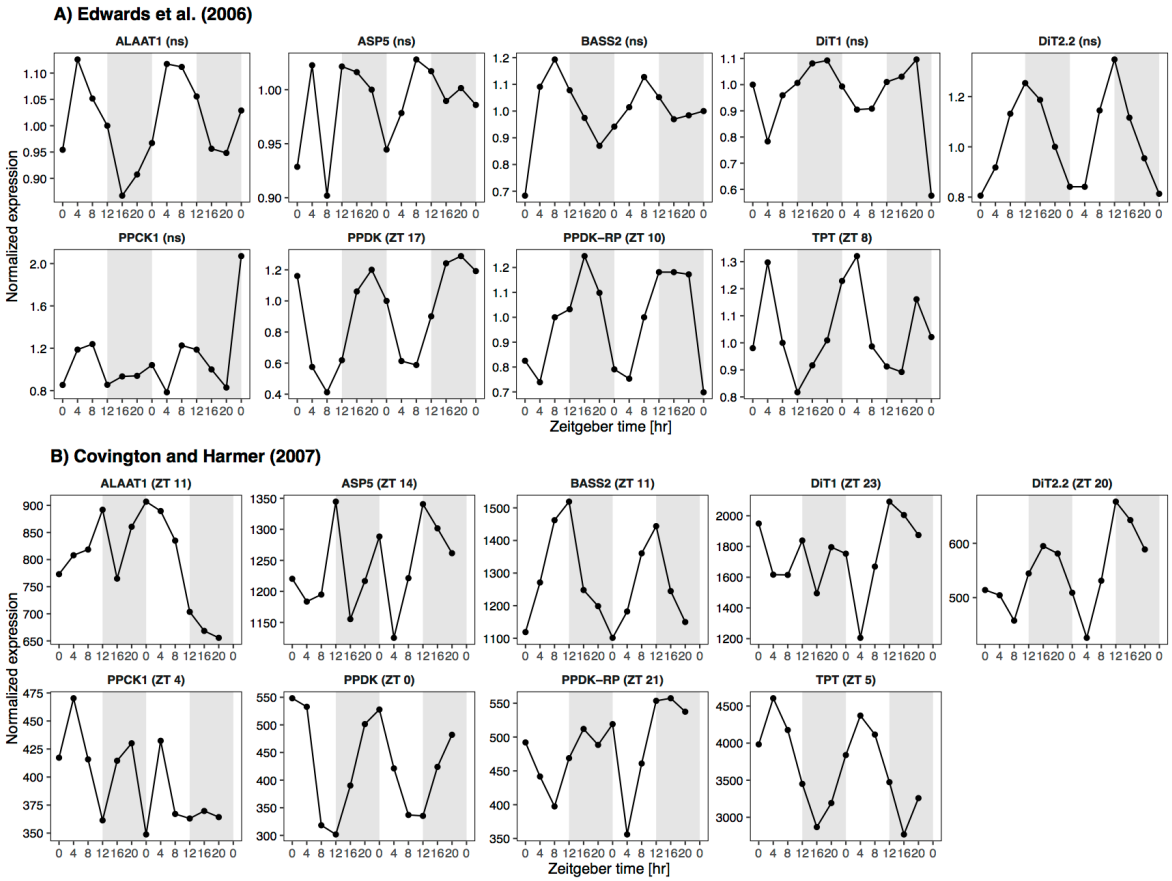
Supplemental Figure S5. Additional isogenes encoding for photorespiratory enzymes. For explanation, legend and context see Fig. 1 and Fig. 7. *GLN, GLUTAMINE SYNTHETASE*



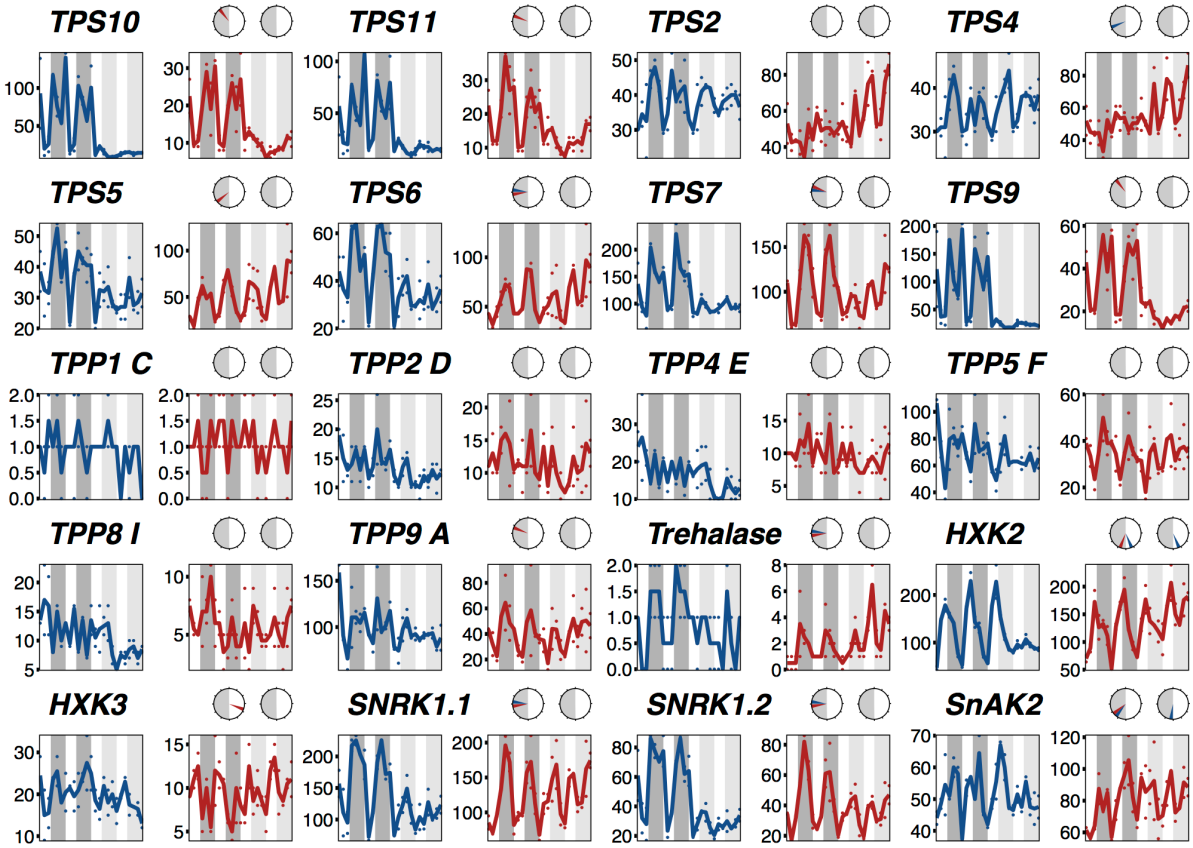
Supplemental Figure S6. Light signalling. For legend see Fig. 1. Genes encoding for regulatory proteins signalling a PhyA-induced (red box; FHY3, far-red elongated hypocotyl 3; PAT1, PHYTOCHROME A SIGNAL TRANSDUCTION 1), PhyA-repressed (green box; EID1, EMPFINDLICHER IM DUNKELROTEN LICHT 1; SPA1, SUPPRESSOR OF PHYA 1; SUB1, SHORT UNDER BLUE LIGHT 1), PhyA/PhyB-induced (blue box; NDPK2, NUCLEOSIDE DIPHOSPHATE KINASE 2; PIF3, PHYTOCHROME-INTERACTING FACTOR 3), PhyB-induced (yellow box; ARR4, ARABIDOPSIS RESPONSE REGULATOR 4; ELF3, EARLY FLOWERING 3; GI, GIGANTEA; RED1, red elongated 1) , PhyB-repressed (grey box; PIF4, PHYTOCHROME-INTERACTING FACTOR 4) response towards the COP/DET/FUS complex (orange box; COP, CONSTITUTIVE PHOTOMORPHOGENIC; DET, DE-ETIOLATED; FUS, FUSCA) and HY5, LONG HYPOCOTYL 5. Based on Quail (2002). Arrowhead, inducing activity. Flathead, repressing activity. For details, see Quail (2002).



Supplemental Figure S7. Genes encoding for additional proteins associated to the circadian clock. For legend see Fig. 1 and Fig. 3. See McClung (2006), Nakamichi (2011) and Hsu and Harmer (2014) for details on the molecular functions. *CCR2, COLD, CIRCADIAN RHYTHM, AND RNA BINDING 2; CDF3, CYCLING DOF FACTOR 3; CHE, CCA1 HIKING EXPEDITION; CKB4, CASEIN KINASE II BETA SUBUNIT 4; COL2, CONSTANS-LIKE 2; COL9, CONSTANS-LIKE 9; CSP41B, CHLOROPLAST RNA BINDING; FT, FLOWERING LOCUS T; HYH, HY5-HOMOLOG; JMJ30, JUMONJI C DOMAIN-CONTAINING PROTEIN 30; LIP1, LIPOIC ACID SYNTHASE 1; LNK1, NIGHT LIGHT-INDUCIBLE AND CLOCK-REGULATED 1; LWD1, ANTHOCYANIN11; PRMT5, SHK1 BINDING PROTEIN 1; PRR3, PSEUDO-RESPONSE REGULATOR 3; RVE1, REVEILLE 1; RVE2, REVEILLE 2; RVE3, REVEILLE 3; RVE6, REVEILLE 6; RVE7, REVEILLE 7; SEC, SECRET AGENT; SFR6, SENSITIVE TO FREEZING 6; 23SPY, SPINDLY; STIPL1, SPLICEOSOMAL TIMEKEEPER LOCUS1; STN7, Stt7 homolog; TEJ, Sanskrit for bright'; XCT, XAP5 CIRCADIAN TIMEKEEPER.*



Supplemental Figure S8. Circadian response of selected C_4-cycle genes in *A. thaliana*. Abbreviations see Fig. 3 and text. Numbers in brackets indicate significant phase as determined in the respective study. Shaded areas, subjective night. Microarray data from Edwards et al. (2006) and Covington and Harmer (2007). See publications for details.



Supplemental Figure S9. Genes encoding for proteins involved in sugar signaling. For legend see Fig. 1 and Fig. 3. *TPS*, *TREHALOSE PHOSPHATE SYNTHASE*; *TPP*, *TREHALOSE PHOSPHATE PHOSPHATASE*; *HXK*, *HEXOKINASE*; *SNRK1*, *SNF1-RELATED PROTEIN KINASE*; *SnAK2*, *SNRK1-ACTIVATING PROTEIN KINASE*

Manuscript 2: Diurnal and circadian leaf transcriptomes reveal enhanced light-responsive gene expression during the evolution of C_4

Status: Draft

Brilhaus, D., Pick, T. R., Mettler-Altmann, T. and Weber, A.P.M.

Journal: Submission to "Journal of Experimental Botany" planned

Author contributions of the current draft

T.R.P. and A.P.M.W. designed the research.

T.R.P. conducted the experiment, sampled leaves, extracted RNA and prepared Illumina libraries.

D.B. analyzed and graphically presented all data and drafted the manuscript.

T.M.-A. proofread the manuscript.

Acknowledgements

I want to thank the following people for their support during my PhD.

Prof. Dr. Andreas Weber for your support ever since my time as an undergrad in your lab, for your faith and for letting me make the most of being part of an **international** graduate program (from Michigan, over Panama to Berkeley and several conferences and retreats in the US, Australia and Portugal). I never took this for granted and my freedom of action always made the time as a PhD much less of a job and much more of a profession and a time well spent!

Prof. Dr. Michael Feldbrügge for being the second member of my PhD committee and corrector of this dissertation.

Dr. Andrea Bräutigam for being my PhD supervisor and for inspiring thoughts on cognitive bias. Your ways of supervision taught and challenged me on plenty levels and I will not forget in the future.

Prof. Dr. Klaus Winter, who welcomed and introduced me to Panama and his lab, grew *Talinum* plants, helped with performing the experiment and writing of the manuscript, which led to the main part of this dissertation.

Dr. Frank Harmon and his PhD students **Carine Marshall** and **Emma Kovak,** for welcoming me in Frank's lab in Berkeley, California, for nice discussions about the circadian clock and a refreshing work atmosphere in the finishing stretch of my PhD

Dr. Tabea Mettler-Altmann for your support and understanding as a scientist and a friend, for help with metabolite profiling, writing the *Talinum* manuscript and comments on the *Flaveria* manuscript. Your perfectionism is challenging and often inspiring.

Dr. Thea Pick for being the closest friend and co-worker in the Weberlab ever since my time as an undergrad. I thank you for your support and for comforting me when in doubt, for fun at parties, BBQs, retreats, field trips, talks and seminars, coffees, sports, dancing and finally for sharing your *Flaveria* project with me and opening the gates to a nice time in Berkeley together.

Dr. Manuel Sommer for endless discussions about science, life, the world, the best ways to brew coffee, your alternative ways of thinking and doing things, for a few Cuba Libres and plenty of beer, here and in California.

PD Dr. Nicole Linka for your encouragement ever since the first practical class in plant biochemistry. If it was not for you and **Marc Linka**, who supervised my Bachelor thesis, I probably would not have specialized in plant biology. I thank you for a lot of fun, plenty of support, good discussions, good cheese, and always new computer challenges.

Dr. Fabio Facchinelli, **Lisa Leson** and **Pia Dahlhoff** for your support and friendship at different stages during my PhD

The RISE-students **Cynthia L. Amstutz** and **Ryan Bower** for investing your summer months to work on my projects

Anja Nöcker for your non-bureaucratic ways to support me in many bureaucratic issues.

iGRAD-plant for funding and the **iGRAD-plant fellows** for nice discussions over PBS pizza lounges and plenty of fun on the retreats. Special thanks to **Dr. Sigrun Wegener-Feldbrügge** for her support during my Bachelor and Master International and during the first half of my PhD.

The **C_4 and CAM research** communities for some great and always refreshing conferences

Various people of the **Weberlab** for making it the place to be for plant sciences in the past few years, for nice coffee and lunch breaks, discussions, field trips, jogging, playing music, scientific and non-scientific support, challenge, motivation and a nice working atmosphere.

My family for the considerate support and interest in my work throughout the years.

Jessica Albers for your love and support that made the last years of my PhD much easier, for believing in me, backing me up, making the best sandwiches, and for the goat.