Integrative analysis of CAM photosynthesis reveals its impact on primary metabolism in *Yucca*

David Wickell[1*], Richard Field[2,3*], Karl Weitz[4], Rosalie Chu[4], Jesse Trejo[5], Nikola Tolic[4], Nathalie Munoz Munoz[4], Hope Hundley[6], Emily Savage[6], Anna Lipzen[6], Jim Leebens-Mack[2], Karolina Heyduk[1]

[1] Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT, 06269

[2] Department of Plant Biology, University of Georgia, Athens, GA, 30602

[3] USDA-ARS, Athens, GA, 30605

[4] Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, 99354

[5] Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, 99354

[6] US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720

*D.W. and R.C. contributed equally to this work

Corresponding author: David Wickell (david.wickell@uconn.edu)

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 (https://academic.oup.com/plcell/pages/General-Instructions) is: Karolina Heyduk (karolina.heyduk@uconn.edu).

Abstract:

Crassulacean Acid Metabolism (CAM) is an adaptation that temporally separates carbon uptake at night from photosynthesis during the day. CAM has evolved repeatedly across vascular plants, perhaps, in part, because it does not involve any novel gene functions and instead relies on regulatory modification of existing C_3 pathways. The co-option of established processes and corresponding alteration of metabolic flux is likely to impact other primary metabolic pathways such as nitrogen metabolism and photorespiration. However, the majority of CAM research to date has focused on a handful of genes in the core CAM pathway. Understanding the complex downstream effects associated with the emergence of CAM is essential to explaining its convergent evolution as well as efforts to engineer carbon concentrating mechanisms in crops. In this study, we integrate genomic, transcriptomic, proteomic, and metabolomic data to compare primary metabolism between the CAM species *Yucca aloifolia* and closely related C_3 species, *Y. filamentosa*. Despite substantial differences in metabolite abundance across multiple pathways, including citrate and nitrogen metabolism, we observe relatively few gene expression differences outside of the core CAM pathway. Additionally, we observe minimal correlation between protein abundance and mRNA expression, suggesting significant post-transcriptional regulation in CAM species. Our findings provide insights into the metabolic shifts associated with CAM evolution and highlight the complexity of its regulation at multiple biological levels.

Introduction

Crassulacean Acid Metabolism (CAM) is a variation on the more common C_3 metabolism that allows plants to temporally separate carbon uptake and storage from photosynthesis. Briefly, in C_3 plants, uptake of atmospheric CO_2 and carbon fixation by Rubisco co-occur during the day. In CAM plants, however, CO_2 is taken up at night and stored in the vacuole as malate prior to being liberated and fed into the Calvin cycle during the day (Fig. 1, Supplementary Fig. S1). In dry-adapted plants, CAM photosynthesis allows them to keep their stomata closed during the hottest part of the day, greatly reducing transpirational water loss. This, in combination with reduced rates of photorespiration due to the enrichment of CO_2 in photosynthetic tissues during the day, results in dramatic improvement to water use efficiency in CAM lineages relative to C_3 plants (Winter & Smith, 2022).

CAM is also distinct from other carbon concentrating mechanisms in that it does not employ any novel gene functions or anatomical structures as observed in C_4 or pyrenoid-forming plants. Instead, CAM physiology appears to result from largely regulatory modifications to previously existing pathways common to all green photosynthetic organisms. Still, there is some debate as to whether the emergence of CAM requires complex regulatory rewiring across the genome (Wai et al., 2019) or is the product of minimal changes resulting in increased metabolic flux through existing C_3 pathways (Bräutigam et al., 2017). Despite the potential complexity of such fundamental alterations to primary metabolism, CAM has evolved over 60 times across the plant phylogeny (Gilman et al., 2023). It is found in every major lineage of vascular plants, in diverse taxa occupying an equally diverse range of habitats from deserts to lakes to rainforest canopies. While the central components of the CAM pathway are present in highly divergent taxa, studies have found a great deal of lineage specific variation as well. CAM plants can vary in their mechanism of day-time decarboxylation (Ramachandra Reddy & Rama Das, 1978), gene copy recruitment (Heyduk et al., 2022; Wickell et al., 2021), the role of circadian regulation (L.-Y. Chen et al., 2020; Wai et al., 2019; Wickell et al., 2021), and citrate accumulation (Franco et al., 1992; Freschi et al., 2010; U. Lüttge, 1988). This high degree of interspecific variation in the details of CAM makes comparison of closely related CAM and C_3 species essential to our understanding of a complex syndrome that has evolved repeatedly in a diverse array of genomic backgrounds.

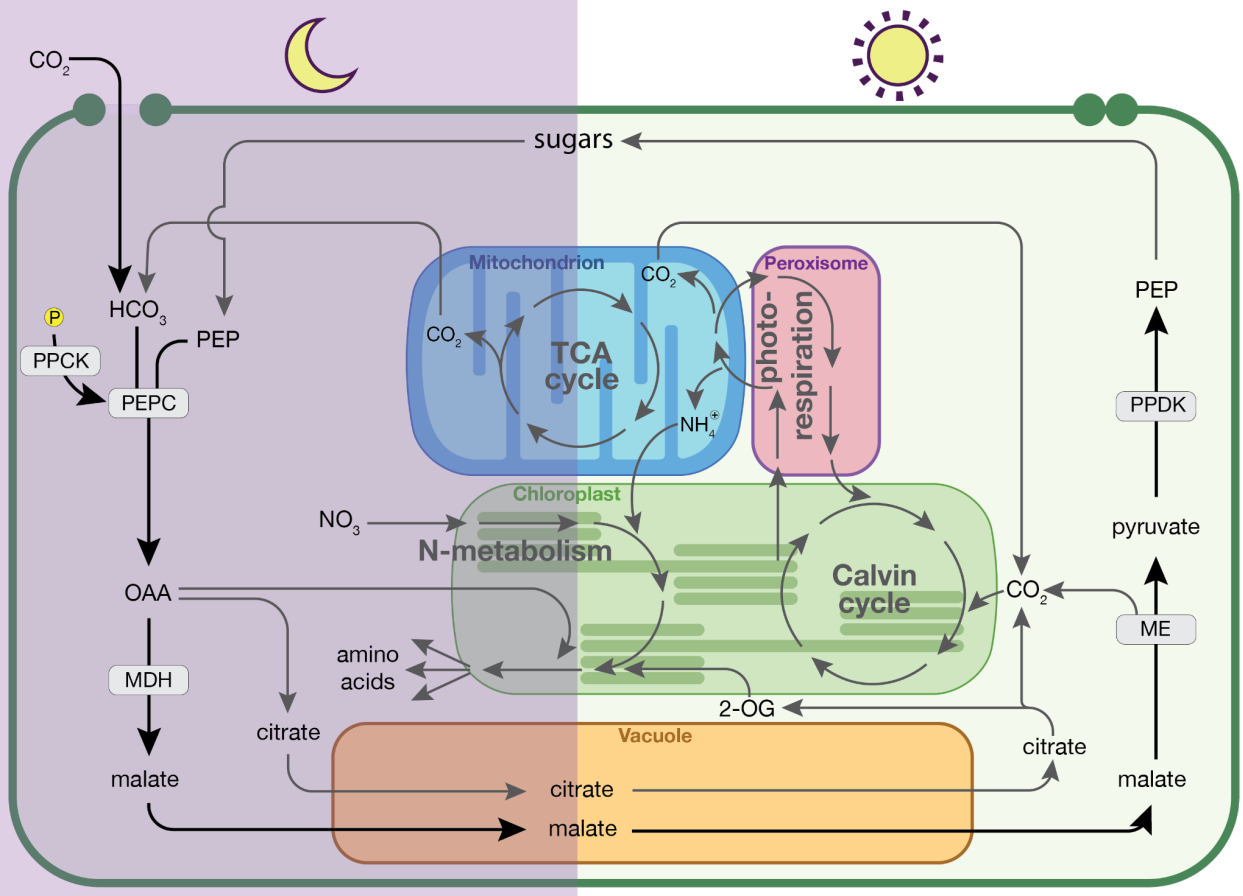

Figure 1: Flow chart of the core CAM pathway (black arrows) and its interactions with other primary metabolic pathways (gray arrows). Abbreviations: 2-OG, 2-oxoglutarate; MDH, malate dehydrogenase; ME, malic enzyme; OAA, oxaloacetate; PEP, phosphoenolpyruvate; PEPC, phosphoenolpyruvate carboxylase; PPDK, pyruvate-phosphate dikinase

To date, the majority of genome-enabled research on CAM species has focused on the central carboxylation mechanisms that fix CO_2 into malate for storage at night and decarboxylation processes that release it during the day (e.g. (Abraham et al., 2016; Ming et al., 2015; Wickell et al., 2021). These core CAM pathways involve genes, proteins, and metabolites that are also central components of primary metabolism in C_3 plants where they replenish intermediates in the citric acid cycle. As such, alterations to their expression and abundance are expected to have complex downstream effects on pathways outside of CAM. A notable example of the effects of CAM outside of the canonical pathway is the night time accumulation of citrate in many CAM plants, often exhibiting a pattern similar to malate. Authors have suggested multiple potential explanations that integrate citrate into the existing CAM framework including

recapture of respired CO_2 (Franco et al., 1992), control of redox potential (Kornas et al., 2009), and mitigation of photoinhibition and photorespiration via efficient concentration of CO_2 around Rubisco (Kornas et al., 2009; Miszalski et al., 2013). However, unlike malate, citrate synthesis is not associated with a net carbon gain (U. Lüttge, 1988). Thus, despite its frequent association with CAM, the precise role of citrate accumulation remains understudied and ultimately unresolved.

Organic acid accumulation, including malate and citrate, has been previously shown to be strongly impacted by both the relative abundance and source of nitrogen and several studies have sought to characterize a mechanistic relationship between CAM and nitrogen metabolism (Gonçalves & Mercier, 2021; Pereira et al., 2018; Pereira & Cushman, 2019; Santos & Salema, 1991). Specifically, researchers have found that some species including *Guzmania monostachia* and several *Kalanchoë* species exhibit increased CAM expression and nocturnal acidification under nitrogen deficiency, further underscoring connections between carbon and nitrogen metabolism (Nobel, 1983; Ota, 1988; Rodrigues et al., 2014; Winter et al., 1982). Additionally, different CAM species seem to show a preference for either nitrate or ammonia as a primary nitrogen source(Ota & Yamamoto, 1991; Pereira et al., 2018). However, few studies have incorporated gene expression data and detailed comparisons of N-metabolism between closely related CAM and C_3 lineages are rare. Similarly, few papers have investigated the photorespiratory pathway in CAM lineages (Duarte & Lüttge, 2007; U. Lüttge, 2011) despite it being widely accepted that increased photorespiration is one of the primary factors driving the convergent evolution of CAM in vascular plants.

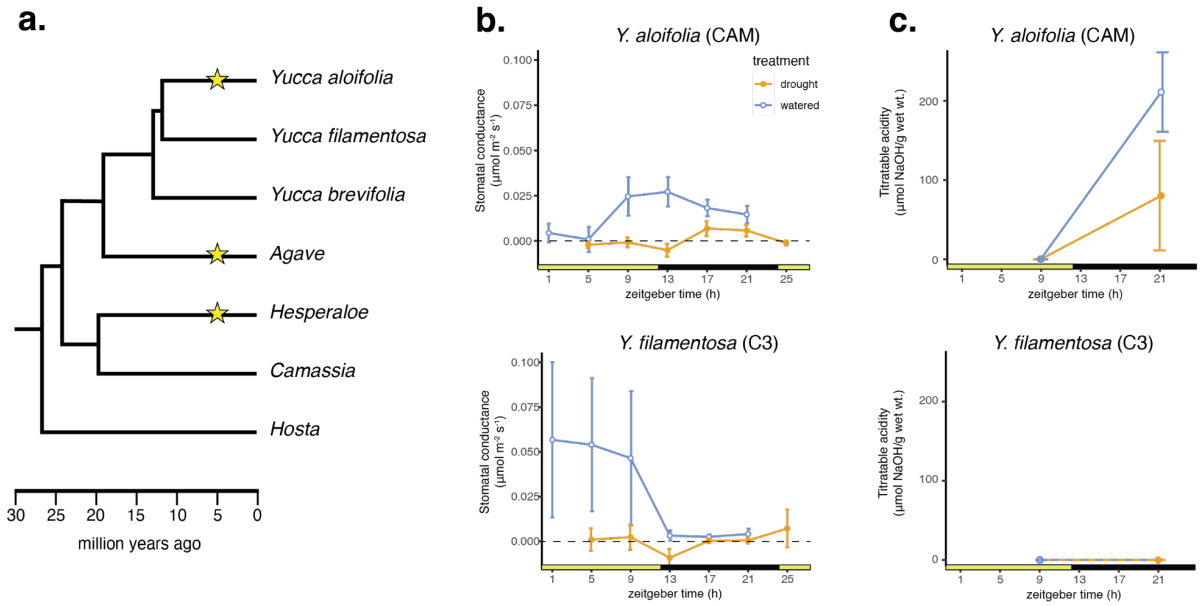

Figure 2: *Yucca aloifolia* and *Y. filamentosa* are closely related species with contrasting photosynthetic pathways. **(a.)** A phylogeny showing placement of *Y. aloifolia* and *Y. filamentosa* within the Agavoideae with hypothesized independent origins of CAM ~5 mya marked with stars, based on Heyduk et al. (2022) with divergence times taken from McKain et al. (2016). **(b.)** Stomatal conductance in *Y. aloifolia* and *Y. filamentosa* over a 24 hr period. Values below zero (black dashed line) are not biologically meaningful. **(c.)** Titratable acidity in *Y. aloifolia* and *Y. filamentosa* between the light period and dark period.

The Agavoideae has emerged as an important model group for studying the evolutionary origins of CAM due to its high degree of photosynthetic diversity, containing C_3, CAM, and physiologically intermediate "C_3 + CAM" species (Heyduk et al., 2023). In particular, the genus *Yucca* within the Agavoideae is well suited to comparative analyses of CAM evolution as it comprises both C_3 and CAM species with a single origin of CAM inferred roughly 6 mya in section Yucca (Fig. 2a) (Heyduk et al., 2016b; McKain et al., 2016). The presence of closely related species that differ in photosynthetic pathway allows researchers to investigate CAM through comparative analyses while controlling for the confounding effects of shared ancestry. As a result, considerable research has been done in *Yucca* investigating the genomic (Heyduk et al., 2019, 2022), physiological (Heyduk, et al., 2016a, 2020), and anatomical (Heyduk et al., 2016b, 2020) changes associated with the transition to CAM. A recent study found evidence of

shared expression patterns of core CAM genes between closely related C_3 and CAM species (Heyduk et al., 2019). Similarly, ancestral state reconstruction in Agavoideae indicated that key features of CAM anatomy such as leaf thickness, cell succulence, and 3D venation may have evolved prior to strong CAM (Heyduk et al., 2016b). These studies seem to corroborate work suggesting that the evolution of CAM may merely require an increase in flux through existing carboxylation pathways (Bräutigam et al., 2017). Even so, Heyduk et al. (2019) also found widespread differences in gene expression within and outside of the central CAM pathway alongside differential abundance of several primary metabolites between CAM and C_3 species, indicative of potentially complex downstream effects of CAM across primary metabolism.

Characterizing the wide ranging effects of CAM on primary metabolism requires simultaneous comparative analysis of multiple pathways and molecules. While this would ideally integrate functional genetic analysis, these types of studies are still in their infancy in CAM plants; *Kalanchoë* is the only CAM lineage where knockdown studies have confirmed genetic function (Agarie et al., 2020; Boxall et al., 2020; Dever et al., 2015; Ferrari et al., 2020). As such, comparative multi-omics approaches potentially offer a more accessible way to investigate modification to primary metabolism in a greater diversity of CAM species. In this study we integrated chromosome level genome assemblies with transcriptomic, proteomic, and metabolomic time series data for *Yucca aloifolia* (CAM) and *Y. filamentosa* (C_3) to investigate shifts in primary metabolism associated with the transition from C_3 to CAM. Comparison of transcriptomic and proteomic datasets showed little correlation between fluctuations in transcript and protein abundance over time, suggesting that post-translational modification and activation/inhibition by other proteins and metabolites play a significant role in regulation of CAM following translation of mRNA into proteins. We also found significant alterations to gene expression and metabolite abundance related to citrate metabolism and nitrogen metabolism in *Y. aloifolia* relative to *Y. filamentosa*. However, despite differences in the abundance of photorespiratory metabolites, we observed relatively few alterations to expression of genes in the photorespiratory pathway between the two species. This may suggest a diminished role for photorespiration in the evolution of CAM and support the theory that the evolution of CAM may be driven by increased flux through existing C3 pathways as opposed to extensive regulatory rewiring (Bräutigam et al., 2017). Our findings underscore the utility of comprehensive multi-omics analyses in unraveling the complex metabolic shifts associated with CAM, providing

valuable insights into gene regulation, and generating testable hypotheses that pave the way for further research of CAM evolution and its underlying mechanisms across vascular plants.

Results

Broad comparison of -omics datasets

Based on MaSigPro analysis, 83.4% (34663/41577) of *Y. aloifolia* genes and 50.4% (22195/43996) of *Y. filamentosa* genes were found to cycle across time in at least one treatment. When analysis was restricted to genes that cycle in both treatments, 54% (22523/41577) genes and 28% (12357/43996) genes cycled in *Y. aloifolia* and *Y. filamentosa*, respectively.

Proteomic analysis of leaf tissue recovered abundance data for 6354 proteins in *Y. aloifolia* and 6859 proteins in *Y. filamentosa*, representing 15.2% and 15.6% of annotated genes, respectively. In contrast to analyses of transcriptomic data, relatively few proteins exhibited time-structured abundance patterns (Fig. 3). In proteins, linear regression identified 50.5% (3214/6354) of proteins in *Y. aloifolia* and only 35% (2400/6859) of proteins in *Y. filamentosa* as cycling over time in at least one treatment. When the analysis was restricted to genes that cycle in both treatments, far fewer proteins were identified with 9.8% (622/6354) and 4.5% (310/6859) proteins cycling in *Y. aloifolia* and *Y. filamentosa*, respectively. Transcription factors (TFs) were notably scarce in our proteomic dataset with only 256 of 5636 TFs observed in *Y. aloifolia* and 349 of 5521 TFs observed in *Y. filamentosa*.

Untargeted metabolomic analysis identified 72 metabolites in *Y. aloifolia* and 81 metabolites in *Y. filamentosa* of which 48 were identified in both species. In *Y. aloifolia* 21/72 metabolites showed time-structured expression over the day-night period in at least one treatment compared to 10/81 in *Y. filamentosa*. Of the metabolites identified in both species, just 5/48 metabolites cycled in at least one treatment in both species.

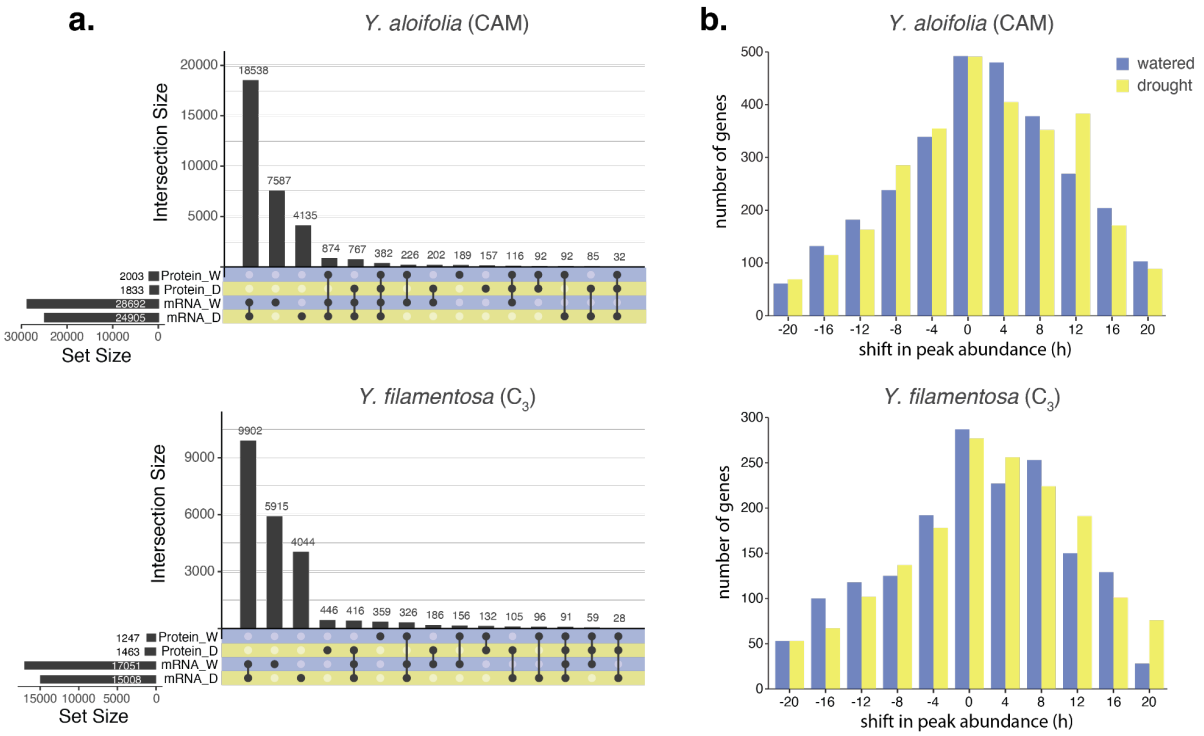

Figure 3: Comparison of cycling transcripts and proteins in *Y. aloifolia* and *Y. filamentosa*. (a.) Upset plot of overlap in cycling genes between treatments and data types and (b.) histograms showing shifts in peak protein abundance relative to peak gene expression in genes that showed cycling dynamics of both transcripts and proteins.

The CAM pathway

CAM associated genes consistently exhibited cycling expression in *Y. aloifolia* relative to *Y. filamentosa*. While at least one copy, and occasionally multiple copies, of CAM genes were found to cycle in *Y. aloifolia*, clustering in MaSigPro (*k*=6; Supplementary Fig. S2) often split genes in the CAM pathway between multiple clusters. In some instances this was expected, as in the partitioning of genes involved in nighttime carboxylation from those involved in decarboxylation during the day or in the case of PEPC and its activator PPCK which are consistently expressed at different times of day in other CAM species (Heyduk et al., 2019; Ming et al., 2015; Wai et al., 2019; Wickell et al., 2021; Yang et al., 2017). Unlike many CAM species, *Y. aloifolia* exhibits cycling expression of two copies of PEPC (Heyduk et al., 2022). In addition to the paralog that is known to cycle in other CAM plants, commonly referred to as

plant-type PEPC (PPPC), *Y. aloifolia* possesses multiple copies of another form, known as bacterial-type PEPC (BPPC) that exhibits cycling expression, identical to that of PPPC (Supplementary Fig. S1). Accordingly, both BPPC and PPPC were assigned to the same expression cluster in *Y. aloifolia*. On the other hand, cytosolic copies of NAD-MDH in *Y. aloifolia* were consistently placed in clusters that differed not only from PEPC and PPCK but also one another despite their common involvement in nighttime carboxylation of PEP. While NAD-MDH is also involved in daytime decarboxylation of malate in some species, *Y. aloifolia* appears to preferentially rely on ME and PPDK for malate decarboxylation based on higher, cycling abundance of these two transcripts relative to MDH and PEPCK. Similar to nighttime carboxylation genes, NADP-ME and PPDK, were split between two MaSigPro clusters based on transcript expression (Supplementary Table 1). In *Y. filamentosa*, there is no correlation in expression between multiple copies of PEPC and all of its paralogs were assigned to different clusters. However, PPCK in *Y. filamentosa* was placed in the same cluster as one copy of PTPC and shows similar expression to PPCK in *Y. aloifolia,* though expression of both genes is roughly 10x lower in *Y. filamentosa*. NAD-MDH expression was ~5-10x lower in *Y. filamentosa* than in *Y. aloifolia* and only a single copy appeared to cycle. Similarly, most copies of NADP-ME and PPDK in *Y. filamentosa* were lowly expressed and tended not to cycle but those that did were divided between multiple clusters.

Protein abundance for CAM genes rarely corresponded to mRNA expression to the extent that it would be difficult to even distinguish between CAM and C_3 on the basis of protein abundance alone (Supplementary Fig. S3). Associated proteins and transcripts did not always cycle in both datasets and when both did cycle, patterns of peak expression were often distinct (Fig. 3b). Of CAM associated proteins, only BPPC abundance appeared to track mRNA expression under drought and watered conditions in *Y. aloifolia*. While two copies of PPPC protein cycled in abundance in *Y. aloifolia* (and a single copy in *Y. filamentosa*), neither corresponded to the gene copies involved in CAM in *Y. aloifolia*. Additionally, while a single copy of PPCK was present in both species, neither protein exhibited cycling abundance and they did not correspond to the most highly expressed cycling transcript in *Y. aloifolia* or *Y. filamentosa*. We found 2 copies of NAD-MDH that cycled in abundance in *Y. filamentosa* but none in *Y. aloifolia*. In the light reactions, both CAM associated copies of NADP-ME and PPDK cycled in protein abundance in *Y. aloifolia*, though peak expression was shifted roughly 8 hours

later in both relative to their transcripts. The same copy of NADP-ME also cycled in *Y. filamentosa* despite the fact that its transcript did not. In addition, a single copy of NADP-ME and NAD-ME that did not show time structured transcript expression cycled in *Y. aloifolia* with peak expression occurring during the dark period.

Malate was the only CAM metabolite present in our metabolomic dataset. Our analysis confirmed that malate cycled strongly in *Y. aloifolia* relative to *Y. filamentosa*, accumulating throughout the night and decreasing in abundance during the day as expected. Despite some similarity between species in expression of CAM genes such as PPCK and at least one copy of PPPC, no nighttime accumulation of malate was observed in *Y. filamentosa*. This result was consistent with measurements of titratable acidity that likewise showed a drastic increase in acidity in *Y. aloifolia* during the night but no detectable increase in *Y. filamentosa*.

Citrate accumulation

Citrate exhibited similar expression to malate in *Y. aloifolia*, as previously reported in several other CAM taxa (Borland & Griffiths, 1989; Franco et al., 1992; Freschi et al., 2010) leading us to investigate genes associated with citrate metabolism in both species. Abundance of citrate increased throughout the night and decreased during the day, compared to *Y. filamentosa* where citrate levels were relatively stable between day and night (Fig. 4). We also observed differences between species in genes related to citrate metabolism. In plants, citrate is typically produced by citrate synthase (CS) in the peroxisome as part of the glyoxylate cycle and in the mitochondria, as part of the tricarboxylic acid cycle (TCA) cycle. In the TCA cycle, citric acid is subsequently converted to isocitrate and isocitrate to 2-oxoglutarate (2-OG) by aconitase (ACO) and NAD-dependent isocitrate dehydrogenase (NAD-IDH), respectively. The most highly expressed copy of mitochondrial citrate synthase (CS) cycled in *Y. aloifolia* but not in *Y. filamentosa*, exhibiting peak expression at dawn. Notably, Aconitase (ACO) exhibited strong cycling expression in *Y. aloifolia*, with expression decreasing during the morning before increasing and remaining high throughout the night. Expression of ACO in *Y. filamentosa* did not have a clear temporal pattern and was 3-4 times lower than in *Y. filamentosa*. Differences in expression were further observed in both NAD-dependent (mitochondrial) and NADP-dependent (cytosolic) copies of isocitrate dehydrogenase (IDH). In NAD-IDH expression was relatively stable over time in both species and treatments but roughly twice as high in *Y. aloifolia*.

Cytosolic NADP-IDH exhibited peak expression in both species near dusk but remained high throughout the night in *Y. aloifolia*, strongly resembling the expression of ACO. Many paralogs associated with the TCA cycle did not cycle in *Y.filamentosa*. However in almost every case where TCA cycle transcripts did cycle in *Y. filamentosa* they were assigned to a single cluster by MaSigPro with relatively flat expression profiles (cluster 1, Supplementary Fig. S2). Conversely, 35 out of 40 TCA cycle paralogs cycled in *Y. aloifolia* and were assigned to multiple MaSigPro clusters. Notably, one copy of CS, two copies of ACO, one copy of NAD-IDH, and one copy of NADP-IDH were assigned to a single cluster by MaSigPro with peak expression at dawn (cluster 6, Supplementary Fig. S2), strongly resembling temporal patterns of citrate abundance.

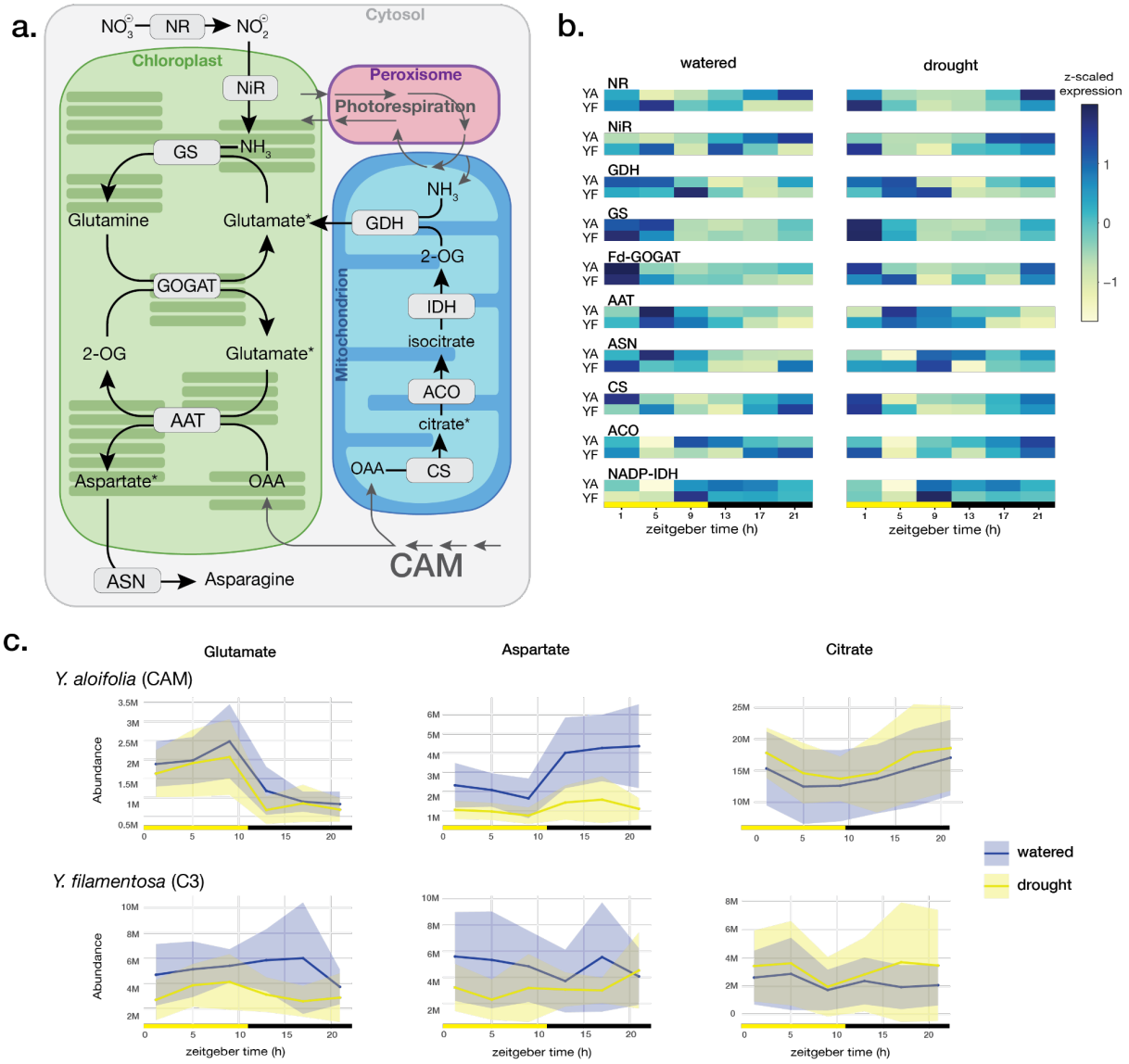

Figure 4: Comparison of nitrogen metabolism in *Y. aloifolia* and *Y. filamentosa.* (a.) Flow-chart of core reactions of the nitrogen metabolism pathway and its interactions with CAM, photorespiration, and the citric acid cycle. **(b.)** Heatmap showing z-scaled gene expression data for *Y. aloifolia* (YA) and *Y. filamentosa* (YF) of most highly expressed genes involved in nitrogen metabolism and the citric acid cycle. **(c.)** Plots of mean abundance over time for key metabolites associated with nitrogen metabolism and the citric acid cycle (indicated in flow-chart by an '*'), shaded areas represent the standard deviation. Abbreviations: 2-OG, 2-oxoglutarate; AAT, aspartate aminotransferase; ACO, aconitase; ASN, asparagine synthetase; CS, citrate synthase; GDH, glutamate dehydrogenase; GOGAT, glutamine oxoglutarate aminotransferase; GS, glutamine synthetase; IDH, isocitrate dehydrogenase; NiR, nitrite reductase; NR, nitrate reductase; OAA, oxaloacetate.

As in CAM genes, cycling expression of proteins did not necessarily follow that of associated transcripts. A single copy of mitochondrial CS protein cycled in *Y. aloifolia*, but did not correspond to the most highly expressed cycling transcript in that species. One mitochondrial and one peroxisomal copy of CS protein appeared to cycle in *Y. filamentosa*. However both exhibited shifts in their abundance profiles relative to transcript expression, particularly under drought conditions. ACO and IDH followed a similar trend, in that the most highly expressed cycling transcripts were not associated with cycling protein abundance, though there were several proteins that cycled under at least one treatment (Supplementary Table 1). In particular, while only 2 out 5 transcripts associated with cluster 6 in *Y. aloifolia* were themselves associated with proteins that cycled, at least one copy of each gene exhibited elevated protein abundance during the night under well-watered conditions.

Nitrogen metabolism

Outside of the CAM pathway, transcriptomic analysis allowed us to explore differences in gene expression associated with nitrogen metabolism. Similar to CAM genes, highly expressed copies of core nitrogen pathway genes were not consistently found in a single cluster in either species. In addition, several genes showed altered expression between species and treatments (Fig. 4). Specifically, both nitrate reductase (NR) and nitrite reductase (NiR),

enzymes catalyzing the initial reduction of nitrate to ammonia, increased throughout the night exhibiting peak expression near dawn in *Y. aloifolia* under both drought and well-watered conditions. While similar expression was observed in *Y. filamentosa* under drought stress, expression levels were relatively stable throughout the day and night in well-watered plants with peak expression shifted to midday in NR and just after dusk in NiR. Glutamate dehydrogenase (GDH), a mitochondrial enzyme which catalyzes the reversible conversion of glutamate to 2-OG and NH_3, showed distinct expression patterns between species with peak expression occurring near dawn in *Y. aloifolia* and around midday in *Y. filamentosa* in both treatments. Conversely, other core nitrogen metabolism genes in the chloroplast including glutamine synthase (GS), glutamine oxoglutarate aminotransferase (Fd-GOGAT), and aspartate aminotransferase (AAT) differed very little between species. Overall expression of both GS and Fd-GOGAT was lower in *Y. aloifolia*, especially under well-watered conditions, but almost four times higher for AAT in both treatments.

Though many genes putatively involved in the nitrogen metabolism pathway cycled in both expression and protein abundance, there were multiple exceptions. GS, mitochondrial-AAT, and asparagine synthetase (ASN) each had a single paralog that exhibited cycling protein abundance without cycling expression of their associated transcript in *Y. aloifolia* (GS, AAT) or *Y. filamentosa* (ASN). In addition there were many transcripts that cycled while their associated protein did not. In *Y. filamentosa* no copies of Fd-GOGAT or AAT cycled in protein abundance and in *Y. aloifolia*, no copies of NR cycled, despite multiple copies of each cycling in transcript expression in both species. In cases where both transcripts and proteins cycled in expression/abundance, peak expression was often shifted either earlier or later in the day/night period. The only instance where the most highly expressed gene cycled in both transcript expression and corresponding protein abundance was GDH, where peak abundance was shifted 8-12 hours later than corresponding transcript expression in both species (Supplementary Fig. S4).

Two primary products of nitrogen metabolism, aspartate and glutamate, were identified by metabolomic analysis. The production of aspartate and glutamate by AAT and Fd-GOGAT, respectively, plays a central role in primary nitrogen metabolism and utilizes OAA, a primary product of CAM dark reactions and TCA cycle intermediate, as a substrate. Despite their similarity in gene expression between CAM and C_3 species, the immediate products of AAT and

Fd-GOGAT appeared to cycle in *Y. aloifolia* but not in *Y. filamentosa*. Though, GDH, which catalyzes the reversible formation of glutamate from 2-OG and ammonia in mitochondria, is differentially expressed between the two species. In *Y. aloifolia*, aspartate and glutamate exhibited inverse abundance with one another, with aspartate peaking during the dark period and glutamate peaking during the light period.

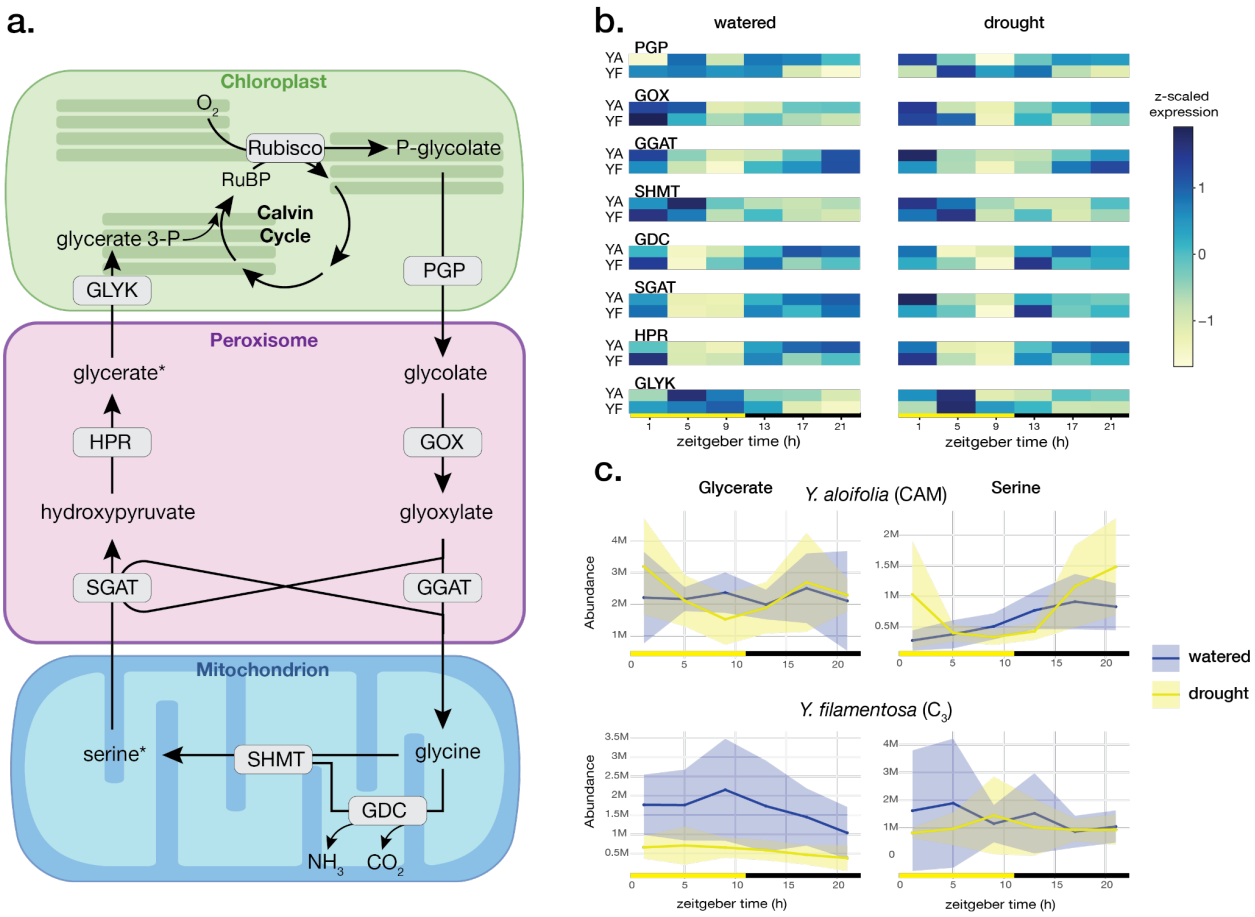

Figure 5: Comparison of photorespiration in *Y. aloifolia* and *Y. filamentosa*. (a.) Flow-chart of core reactions of the photorespiratory pathway. **(b.)** Heatmap showing z-scaled gene expression data for *Y. aloifolia* (YA) and *Y. filamentosa* (YF) of most highly expressed photorespiratory genes. **(c.)** Plots of abundance over time for key metabolites associated with photorespiration (indicated in flow-chart by an '*') shaded areas represent the standard deviation. Abbreviations: GDC, glycine decarboxylase complex; GGAT, glutamate:glyoxylate aminotransferase; GLYK, glycerate kinase; GOX, glycolate oxidase; HPR, hydroxypyruvate

reductase; PGP, phosphoglycolate phosphatase; Rubisco, ribulose bisphosphate carboxylase; SGAT, serine:glyoxylate aminotransferase; SHMT, serine hydroxymethyl transferase.

Photorespiration

Despite the putative involvement of photorespiration in the convergent evolution of CAM, overall expression of genes in the photorespiratory pathway was relatively similar between *Y. aloifolia* and *Y. filamentosa* (Fig. 5). A notable exception is phosphoglycolate phosphatase (PGP), the enzyme that catalyzes the first step of the photorespiratory pathway, converting 2-phosphoglycolate to glycolate in the chloroplast. Both species possess two copies of PGP, and both copies exhibit similar expression under drought and watered treatments in *Y. filamentosa*, with elevated expression throughout the daylight hours and reduced expression at night. This pattern is consistent with a presumed response to elevated photorespiratory stress during the day. Conversely, in *Y. aloifolia*, a single copy of PGP is expressed. Under well watered conditions, PGP exhibits peak expression at midday but increases again to similar levels during the dark period. Under drought, expression peaks at dawn in *Y. aloifolia*, steadily decreasing throughout the day before increasing again during the dark period (Fig. 5b). In addition to the relatively similar expression between species, cycling genes in the photorespiratory pathway were consistently assigned to the same MaSigPro cluster in each species. 15 out of 24 cycling transcripts including 6 out of 8 of the most highly expressed copies in *Y.aloifolia* were assigned to cluster 1 with peak expression at the end of the night period. In *Y. filamentosa* a similar pattern was observed with 10 out of 20 cycling transcripts assigned to cluster 3 with peak expression during the dark period. However, 5 of 20 were assigned to cluster 6 with peak expression at midday (Supplementary Fig. S2, Supplementary Data). There were no genes assigned to a corresponding midday cluster in *Y. aloifolia*. Nearly every transcript that was not assigned to cluster 1 in *Y. aloifolia* or clusters 3 and 6 in *Y. filamentosa* was a paralog of SHMT (6 out of 8 in *Y. aloifolia*, 4 out of 5 in *Y. filamentosa*).

Once again, we found little correlation between transcript expression and protein abundance in genes associated with the photorespiratory pathway (Fig. 3, Supplementary Fig. S5). In particular, very few genes displayed cycling protein abundance in either species. Out of 8 core genes involved in photorespiration only SHMT and GDC-H cycled in protein abundance in both species as well as two copies of GOX in *Y. filamentosa*. SHMT and the GDC enzyme

complex facilitate the conversion of glycine to serine in the mitochondria as part of the photorespiratory pathway. Specifically, GDC-H plays a essential role in the decarboxylation of glycine and subsequent release of photorespiratory CO_2 and NH_3, by interacting with the other three enzymes in the complex (GDC-L, GDC-P, and GDC-T; (Timm et al., 2012)). In *Y. aloifolia*, 5 out of 9 copies of SHMT exhibited cycling abundance, all of which also cycled in transcript expression. In contrast, only a single copy of SHMT cycled in protein abundance in *Y. filamentosa*, correlated with a transcript that did not cycle in expression.

In our metabolomic dataset, serine and glycerate showed differential abundance patterns between *Y. aloifolia* and *Y. filamentosa*. Serine and glycerate are produced in plants as intermediates in the photorespiratory pathway (Fig. 5a). Both exhibited elevated abundance during the dark period in drought and well watered conditions in *Y. aloifolia* though the fluctuation was much stronger under drought (Fig. 5c). However, as photorespiration is a light dependent reaction its metabolites are typically expected to exhibit peak abundance during daylight hours, as observed in *Y. filamentosa*. It may be of note that while production of serine and glycerate do not use OAA as a substrate, flux of OAA into the mitochondrion is essential to replenishing stores of NAD^+ used for the conversion of glycine to serine. In this way, increased nocturnal abundance of OAA could potentially drive serine production and thus affect downstream abundance of glycerate by providing unreduced NAD^+ to the GDC enzyme complex.

Discussion

Protein abundance shows reduced temporal expression relative to mRNA

Metabolites exhibiting time structured expression were frequently associated with transcripts that also fluctuated over the day-night period. That being said, metabolites and mRNA typically do not have a one to one relationship as a single metabolite may be involved in multiple pathways. Conversely, protein abundance rarely cycled along with their associated transcripts, or in cases where they did, they frequently exhibited shifts in peak abundance relative to transcript expression (Fig. 3). Our results align with previous studies that find relatively low correlation between mRNA expression and protein abundance (Baerenfaller et al., 2012; de Sousa Abreu et al., 2009; Kundrátová et al., 2021). The lack of correlation between transcript

expression and protein abundance is likely driven by multiple factors. First, far fewer genes were present in our protein abundance dataset. Many genes with average expression below 100 TPM did not have associated protein data. As a result, it is possible that some correlations were not detected. Specifically, proteins that may be present in low abundances including transcription factors (e.g. PRR1, CCA1/LHY,) and protein kinases (e.g. PPCK) were absent from our proteomic dataset regardless of their expression at the transcript level. Many of these proteins are expected to exhibit strong cycling abundance and are associated with transcripts that, likewise, cycle in expression throughout the day. Second, there is evidence that shifts in protein abundance are driven by factors other than mRNA expression including post-transcriptional regulation, translation, and degradation of protein products (Vogel & Marcotte, 2012; Wu et al., 2024). Finally, it is important to keep in mind that enzymatic activity is not simply regulated by protein abundance. Processes such as protein-protein interaction (Rugen et al., 2021; Zhang et al., 2018), allosteric regulation by metabolites (Hartman et al., 2023), and post-translational modification (PTM) have significant effects on primary metabolism (Balparda et al., 2023). For instance, several proteins in the CAM pathway, including PEPC (Giglioli-Guivarc'h et al., 1996), MDH (Carr et al., 1999), and PPDK (Y.-B. Chen et al., 2014), are subject to post-translational interactions and modification that augment their activity. Likewise, multiple enzymes involved in nitrogen metabolism (Costa-Broseta et al., 2020; Kaiser & Huber, 2001; Seabra et al., 2013) and photorespiration (Jossier et al., 2019; Liu et al., 2020; Palmieri et al., 2010) are regulated by PTMs.

Modification of primary metabolism in Y. aloifolia

The evolution of Crassulacean Acid Metabolism (CAM) involves substantial regulatory rewiring of gene expression, accompanied by shifts in primary metabolite fluxes. These broad scale modifications to core processes of carbon fixation and circadian regulation likely necessitate simultaneous adjustments across other pathways that share metabolic substrates and overlap in regulation. In particular, PEPC activity is strongly associated with N-metabolism in C_3 plants as it is responsible for replenishing carbon skeletons diverted from the TCA cycle for amino acid synthesis (O'Leary & Plaxton, 2015). Our investigation into downstream effects of CAM in *Yucca* reveals significant differences in N-metabolism between *Y. aloifolia* and *Y. filamentosa*, in agreement with previous studies highlighting links between CAM photosynthesis

and nitrogen assimilation in various CAM and C_3 + CAM lineages (Freschi et al., 2010; Gonçalves & Mercier, 2021; Pereira et al., 2018; Santos & Salema, 1991). Under well-watered conditions, we observed distinct changes in gene expression related to nitrate reduction and amino acid synthesis. Specifically, there were clear differences in accumulation of glutamate and aspartate, immediate downstream products of primary nitrogen metabolism. Aspartate and glutamate have been previously shown to be powerful allosteric regulators of the interaction between respiration and N-metabolism in plants (O'Leary & Plaxton, 2015). Accumulation of glutamate during the light period in *Y. aloifolia* was associated with a shift in expression of GDH from midday to early morning. Conversely, the sharp increase in aspartate abundance coupled with the drop in glutamate abundance over night was not accompanied by any changes in expression of AAT, which converts OAA and glutamate to 2-OG and aspartate. Thus, it is possible that the increased flux of carbon through OAA may be sufficient to explain nocturnal synthesis of aspartate via AAT without any alteration to transcript expression or protein abundance. In addition to GDH, we observed a similar shift in genes associated with nitrate reduction with expression in *Y. aloifolia* occurring near dawn under well-watered conditions. Interestingly, interspecific differences in GDH, NR, and NiR were less pronounced under drought conditions. Accordingly, *Y. filamentosa's* similarity to *Y. aloifolia* in gene expression under drought stress could imply that alterations to gene regulation in this pathway may have also played a role in the early evolution of CAM under increasingly arid conditions.

Alongside well established differences in expression in genes associated with the central CAM carboxylation and decarboxylation pathways, we observed altered flux through carbon compounds other than malate including citrate, malonic acid, and maleic acid in the CAM species *Y. aloifolia* relative to the C_3 species *Y. filamentosa* (Fig. 4, Supplementary Fig. S6). Nocturnal accumulation of citrate has been documented in multiple CAM species (Borland & Griffiths, 1989; Franco et al., 1992; Freschi et al., 2010; Kornas et al., 2009; U. Lüttge, 1988, 1990). However, its role in CAM is less clear than that of malate, as fixation of CO_2 into citrate does not result in a net carbon gain (U. Lüttge, 1988). A number of theories have been put forth regarding putative benefits of citrate accumulation in CAM plants. In the CAM pathway, the production of malate by MDH is dependent on NADH, as opposed to citrate synthesis which results in a net gain of NADH. As such, nocturnal citrate production may provide additional reducing power for reduction of malate while also serving as a secondary storage mechanism for

excess CO_2 as internal stores of NADH are exhausted (Trípodi & Podestá, 2003). Additionally, in *Yucca*, we observed shifts in expression of NR and NiR from midday in *Y. filamentosa* to the end of the night in *Y. aloifolia*. This may suggest that the additional reducing power generated by the production of citrate is also being used for reduction of nitrate during the dark period in *Y. aloifolia* as opposed to C_3 plants and algae which primarily generate reducing power for N-assimilation via photosynthesis during the day. It is also possible that accumulation of citrate is merely an atavistic trait retained in CAM plants from their C_3 ancestors (Bräutigam et al., 2017). However, this scenario seems unlikely as we do not observe any nocturnal accumulation of malate in *Y. filamentosa*. Even so, in C_3 species, the decarboxylation of citrate during the day is thought to supply the carbon skeletons for amino acid synthesis during the day (Scheible et al., 2000). Citrate's role in linking carbon and nitrogen metabolism in C_3 plants may offer an explanation for the inverse cycling of aspartate and glutamate throughout the day and night. In C_3 plants, 2-OG is supplied for nitrogen assimilation from the decarboxylation of citrate in the mitochondria or the cytosol (Szal & Podgórska, 2012). Additionally, there is evidence that 2-OG utilized for day time nitrogen assimilation originates, at least in part, from citrate accumulated during the night (Tcherkez et al., 2009). As such, the high accumulation of glutamate during the day in *Y. aloifolia* relative to *Y. filamentosa* could be driven by the increase in citrate production during the dark period and its subsequent degradation during the day (Fig. 4).

Additionally, in both CAM and C_3 plants, there seems to be a clear association between citrate production and various types of abiotic stress (U. Lüttge, 2006; Tahjib-Ul-Arif et al., 2021). Notably, citrate:malate ratios are known to increase in CAM plants in response to high light stress in *Tillandsia (Borland & Griffiths, 1989)* and high light stress and drought in *Clusia* (Franco et al., 1992). Citrate does not offer an effective means for assimilation of atmospheric CO_2, due to the release of CO_2 during generation of its substrate Acetyl-CoA. However, citrate production is more energetically favorable than that of malate, as the conversion of hexose to citrate results in a net gain of ATP and NAD(P)H relative to fixation of CO_2 into malate (U. Lüttge, 1988). Additionally, the breakdown of citric acid releases three molecules of CO_2 whereas the conversion of malate to OAA during the light reactions of CAM releases only a single molecule of CO_2. As such, production of citrate may represent a way to recycle respiratory CO_2 during periods of drought where stomatal conductance is severely reduced during the day and night limiting access to atmospheric CO_2 (U. Lüttge, 1988). Likewise, under high light,

accumulation of citrate at night and the subsequent release of CO_2 upon its breakdown during the day is thought to minimize photorespiration and photoinhibition by concentrating CO_2 around rubisco as supplies of malate are exhausted. Finally, it has been suggested in *Clusia* that accumulating citrate at night could increase the buffering capacity of the vacuole, leading to an increase in storage capacity for CO_2 (Franco et al., 1992). Our observations in *Y. aloifolia* were consistent with those in other species that accumulate citrate with the ratio of citrate to malate increasing under drought stress. This may be due to citrate's role in minimizing oxidative stress or alternatively may reflect the lower energetic requirements of citrate production relative to malate (U. Lüttge, 1988). Regardless, our findings corroborate a potentially important link between citrate metabolism and CAM in *Yucca*.

Comparison of transcript and protein abundance over time between *Y. aloifolia* and *Y. filamentosa* demonstrates that the metabolic changes underpinning the evolution of CAM occur at many levels with shifts in protein kinetics and post-translational modification potentially having a significant impact on CAM and other downstream pathways. Furthermore, analysis of gene expression alongside fluctuations in metabolite abundance demonstrate a clear effect on overlapping pathways of nitrogen assimilation and to a lesser extent, photorespiration.The CAM metabolite OAA seemed to play a key role in driving shifts in primary metabolism and may act as a link between regulation of carbon and nitrogen metabolism. Many metabolites, including those in the nitrogen cycle, that exhibited day-night fluctuations in abundance either use OAA as a substrate or rely on it to supply NAD^+ in the case of photorespiration. The cycling abundance of these primary metabolites is notable as OAA is produced in great abundance by carboxylation of PEP by the enzyme PPC in CAM plants. The subsequent reduction of OAA to malate by MDH is dependent on the presence of NADH. Under conditions of high light or drought, it's possible that OAA synthesis by PPC outstrips the ability of MDH to convert OAA to malate. This could potentially result in OAA being funneled into other processes that either generate reducing power, such as citrate synthesis, or at the very least do not require NADH, such as aspartate synthesis. As such, it's possible that cycling of multiple metabolites outside of the central CAM pathway may simply require greater flux through OAA rather than extensive regulatory rewiring outside of the core CAM pathway.

Photorespiratory pathway largely conserved between CAM and C_3 plants

Photorespiration is thought to play a major role in CAM evolution, and it is assumed to decrease alongside the evolution of carbon concentrating mechanisms like CAM and C_4. However, the photorespiratory pathway itself has only rarely been the focus of CAM research. Our results show that while there are subtle differences in gene expression and metabolite abundance related to photorespiratory pathways between CAM and C_3 species of *Yucca*, gene expression in the photorespiratory pathway as a whole remains largely conserved between species in drought and well-watered conditions. A notable exception to this similarity in expression is found in PGP, which has daytime expression in *Y. filamentosa* but is highly expressed at night in *Y. aloifolia* under both treatments. The difference between species is particularly pronounced during drought where expression of PGP is nearly inverse. Additionally, expression of PGP in both species and conditions roughly tracks metabolite abundance of glycerate and serine, key products of the photorespiratory pathway (Fig. 5). This trend may reflect the regulatory role of PGP as it has been previously shown to play a major role in controlling flux through the photorespiratory pathway, along with having downstream effects on the TCA cycle and stomatal control (Eisenhut et al., 2017; Timm & Hagemann, 2020). In this case, altered expression of PGP may be an important component of CAM evolution. Alternatively, glycerate and serine are involved in multiple pathways. Specifically, both are produced by the poorly characterized, non-phosphorylated “glycerate pathway” of serine biosynthesis (Igamberdiev & Kleczkowski, 2018). Thus, it is possible that their nocturnal accumulation under drought stress in *Y. aloifolia* does not reflect flux through the photorespiratory pathway, but production of serine by an alternative pathway in the absence of photorespiration.

It is also of interest that when *Y. filamentosa* experiences drought, the levels of some metabolites involved in primary nitrogen metabolism have similar abundance patterns as those in *Y. aloifolia*. However, this similarity does not appear to extend to photorespiratory metabolites found in our data. This may suggest that the evolution of CAM was not primarily driven by photorespiratory stress, as drought induces "CAM-like" behavior in some pathways but not others. The highly similar expression of photorespiratory genes indicates that significant changes in gene regulation outside the CAM pathway may be less critical to establishment of CAM. Instead, CAM evolution may only require minimal changes in gene expression, which primarily enhance metabolic flux through existing C_3 pathways rather than requiring complex, global

regulatory shifts. Additional, detailed comparisons of molecular time series and functional analysis of photorespiratory genes in related CAM and C_3 species are needed to fully understand the dynamics we observe in *Y. aloifolia* and *Y. filamentosa*.

Considerations for future studies in CAM evolution

Our study shows the importance of conducting genomics research in closely related C_3 and CAM taxa and future studies can continue to expand the scope of our knowledge of CAM evolution by conducting similar studies in other lineages. In addition to continuing to develop molecular resources in the Agavoideae and other model CAM taxa, researchers should endeavor to incorporate an ever greater diversity of taxa including aquatic CAM species and physiologically diverse C_3 + CAM species. While C_3 + CAM should not necessarily be treated as a transitional state, its study is likely to provide valuable insights into initial molecular changes that selection might act upon to drive the emergence of CAM in a C_3 background. Furthermore, the overall lack of concordance between transcript expression, protein abundance, and metabolite abundance seen in *Yucca* suggests that while analyses based on transcriptomic data are highly accessible and capable of providing valuable insight into CAM evolution, they are but one piece of a much larger puzzle. Future studies would do well to keep this in mind and supplement traditional, bulk-transcriptomic analyses with other molecular data types and integrating emerging methods such as single cell and spatial transcriptomics/proteomics.

Similarly, while genomic resources continue to advance in many CAM lineages, there has been little to no functional characterization of the distribution and binding affinity of cis-regulatory elements in CAM species. Instead, current methods rely almost entirely on *in-silico* analyses based on homology to motifs and transcription factors from distantly related taxa (Gilman et al., 2022; Wickell et al., 2021). Accurate identification of regulatory elements is especially pertinent given that CAM is fundamentally viewed as a regulatory syndrome that is frequently evaluated through the lens of gene expression. Thus, functional exploration of promoter regions through techniques such as ChIP-seq, DAP-seq, and/or ATAC-seq are essential to deepening our understanding of the mechanisms underpinning the repeated evolution of CAM.

Indeed, much of our current understanding of CAM genetics is based almost entirely on a combination of transcriptomic data combined with homology to C_3 plants. While this can potentially provide valuable information and identify key differences between CAM and C_3 taxa,

there are many cases where homology is vague, unhelpful, or downright misleading. This is particularly true in genes from large, diverse families such as transport proteins and transcription factors or those which are primarily regulated by post transcriptional/translational mechanisms. While core functions of genes involved in primary metabolism appear to be conserved across deep evolutionary time, other aspects including regulation, interaction, localization and kinetics are likely to vary between lineages or even between paralogs within lineages. In many cases we found that multiple paralogs were highly expressed with distinct diel profiles, as exemplified by our clustering analyses where cycling paralogs were frequently split between multiple clusters. While these results are intriguing, they suggest that comparative methods alone are insufficient to elucidate the complex interactions between environment, genetics, and physiology. Targeted functional research is necessary to tease apart these relationships and clarify the roles and multi-level regulation of individual genes within and alongside the CAM pathway. This will necessitate the development of plant transformation resources in CAM and closely related C_3 lineages.

Our research advances the study of CAM evolution by identifying genes that are promising targets for focused investigations into metabolism beyond the core CAM pathway. However, the substantial discordance between our datasets also highlights some of the shortcomings of an approach based solely on comparative genomics. While continued expansion of genomic resources into new lineages is essential, advancing our understanding of CAM physiology will ultimately require a shift from broad comparative analyses to targeted studies that functionally validate the molecular mechanisms and regulatory controls that promote the convergent evolution of CAM.

Methods

Physiology

Six mature *Y. aloifolia* plants, each a different genotype, were selected from a collection of wild-collected plants at the University of Georgia (UGA) Plant Biology greenhouse. Six mature *Y. filamentosa* plants, also each a different genotype, were obtained either from the UGA collection or were purchased from landscape nurseries. Plants obtained from nurseries were re-potted in 5-gallon containers with a 50/50 mixture of standard potting soil and sand (the same mixture in which the *Y. aloifolia* plants were grown and maintained) and allowed to acclimate to

the greenhouse for at least 4 weeks. Plants were introduced into a growth chamber at the UGA Plant Biology greenhouse in randomized positions one week before the onset of sampling to allow entrainment to the light regime and were watered daily. Growth chamber conditions were programmed with the following conditions: the light period commenced at 7:00 am with relative humidity of 40%, ambient temperature at 30° C, and photosynthetically active radiation (PAR) approximately 270 µmol m^{-2} s^{-1} measured at leaf level. The dark period commenced at 7:00 pm with relative humidity of 40%, ambient temperature at 17° C, and photosynthetically active radiation (PAR) 0 µmol m^{-2} s^{-1}.

Gas exchange and leaf tissue sampling for the well-watered treatment commenced on December 03, 2021. Beginning one hour after the onset of the light period ("Zeitgeber time 1", ZT1), instantaneous gas exchange measurements were obtained with a LiCOR-6400XT portable photosynthesis system on mid-leaf sections of mature leaves. Immediately following gas exchange measurements, a randomly selected, mature leaf was excised and snap frozen in liquid N_2 and stored at -80° C. Gas exchange measurements were repeated on the same leaf, and a random leaf was sampled for each plant, in the same order every four hours for the next five time points (ZT5, ZT9, ZT13, ZT17, ZT21). Soil moisture was measured at the end of the 24-hour sampling period with a digital soil moisture probe inserted into the top ~3 inches of soil for each plant. Watering was then withheld for the remainder of the experiment to induce the drought response. Soil moisture levels were periodically checked over the next 18 days. At day 13, it was observed that the *Y. aloifolia* species' soil was drying down faster than that of *Y. filamentosa*. To maintain comparable soil conditions, each fast-drying plant was top-watered with 500 mL water, bringing those plants back to similar soil water content as the others. By day 18 (December 20th, 2021), the average soil moisture for each species was below 5% and gas exchange and leaf tissue sampling were repeated in the same manner as the well-watered samples.

RNA-sequencing and analysis

Whole leaf tissue for each sample was ground to a powder in a mortar and pestle under liquid N_2, transferred to a 15 mL Falcon tube and stored at -80° C freezer. RNA extraction for each sample from an aliquot of approximately 100 mg of powdered leaf tissue was carried out with the Qiagen Plant RNeasy kit with DNase I on-column digestion according to the Qiagen standard protocol. RNA quantity was measured with a Qubit fluorometer, and RNA integrity was

measured on an Agilent 2100 Bioanalyzer with an RNA Nano 6000 kit. RNA samples were shipped on dry ice to the DOE's Joint Genome Institute (JGI) labs in Berkeley, CA for cDNA library construction and paired-end 150 bp Illumina sequencing on an Illumina NovaSeq sequencer. All but five libraries produced deep sequence data (total PE150 libraries = 139, median fragment count = 28,066,771, mean Q30 = 0.95).

RNA-seq read quantification

RNA sequence data was downloaded from the JGI server onto the Georgia Advanced Computing Research Center computing cluster. Reads for each species were pseudo-aligned to their respective primary transcript nucleotide fasta sequences and quantification was calculated with *kallisto* (Bray et al., 2016) with default parameters. Raw read counts were imported into expression matrices in R with *tximport* (Soneson et al., 2016).

Time-structured expression profiling of transcriptomic data:

To assess differential gene expression patterns due to watering regimes over the 24-hour sampling period, differential expression analysis taking time and treatments as variables was carried out in the R software package *MaSigPro* (Nueda et al., 2014). For each species, trimmed mean of *M* values (TMM) were calculated and used to normalize read counts within species, and counts per-million (CPM) were computed from the normalized libraries in the R software package *EdgeR* (Robinson et al. 2009). Genes with less than 0.5 CPM across all time points were removed. Dispersion was estimated using "classic mode," which uses conditional maximum-likelihood to estimate dispersion and an empirical Bayes procedure to minimize the dispersion around a consensus value within replicates (here, species replicates). Outlier libraries were checked with multidimensional scaling plots. Two and six libraries were removed from *Y. aloifolia* and *Y. filamentosa*, respectively. Genes were tested for significant time structure with functions *p.vector (counts=TRUE)*, and the *T.fit(step method = "backward", alpha = 0.05)*. Genes with high variability between replicates were classified as overly influential by *MaSigPro*, removed, and the analyses were rerun on the filtered data set. Genes with significant treatment effects across time were called with *sig.siggenes*(*rsq* = 0.6). To sort genes into clusters based on time-structured expression similarity, *MaSigPro* requires the user to input the expected numbers of clusters (*k*). To estimate *k,* within-group sums of squares were calculated and plotted so the *k*

value at which marginal reductions in the sums of squares can be determined. *k*=6 was chosen for both species and genes were sorted into clusters using *see.genes(cluster.method="hclust").*

Metabolite extraction and protein isolation

2.0 ml tubes (Safe-Lock, Eppendorf) were filled with leaf material to approximately ¼ capacity, covered with tube membranes, and samples were lyophilized overnight. Lyophilized leaves were then homogenized in a Tissue Lyser II (Qiagen) using tungsten beads, specifically two-3mm beads and one-5mm in each tube. Samples were lysed for 2 min at a frequency of 23 Hz until ground to a powder. Metabolites from 20 mg of the ground leaves were extracted with 800 µl of a 4:1 methanol: water mix, and each sample spiked with 50 µl of a 330 µg/ml solution of glycine d5 (Cambridge Isotopes). Samples were then placed in a thermomixer and shaken at 1200 rpm and 21°C for 1 hr. Following centrifugation at 5000 x g, supernatant containing the metabolite fraction was removed and 10% of the total metabolite extract was transferred to a glass vial to be dried under vacuum for subsequent metabolomic analysis (described below). After removing the metabolite fraction, protein pellets were washed twice by adding 500 µl 100% MeOH, vortexing for 15 seconds, and centrifuging at 10000 x g and 21°C for 10 min. The resulting pellet was dried in a fume hood until remaining MeOH had completely evaporated and stored at -80°C until protein digestion.

Protein pellets were resuspended in 120 µl of an aqueous solution of 8 M urea and 50 mM HEPES by vortexing and sonication for 30 sec. Following resuspension, 6.3 µL 1 M aqueous solution of DTT was added to each sample and samples were placed in a thermomixer and shaken at 500 rpm and 37°C for 30 min for reduction. Next, 40 µL 0.5 M aqueous solution of iodoacetamide was added to each sample and samples were placed in a thermomixer and shaken for another hour at 500 rpm and 37°C for alkylation. Samples were then diluted with 800 µL 50 mM aqueous solution of HEPES prior to digestion with trypsin (1:50 w/w trypsin:protein) for 3 h at 500 rpm and 37°C. Sample clean-up was performed using C18 SPE media and samples were concentrated to ~100 µL and peptide concentration was determined via BCA assay. Finally, 30 µg of each sample were transferred to a new 1.5 mL tube, dried down in a speedvac and stored at -80°C for TMT labeling.

TMT isobaric tag labeling

Samples were designed into 12 TMT sets (6 per plant) including two pools, one consisting of a mix of the well-watered plants and one of all the drought plants to ensure correlation across sets. Each set included 1 replicate of each condition and blanks to establish a base background noise signal and then each set completely randomized. Each sample was diluted in 40 µl 500 mM HEPES, pH 8.5 and labeled using amine-reactive Tandem Mass Tag Isobaric Mass Tagging Kits (Thermo Scientific, Rockford, IL) according to the manufacturer's instructions. Briefly, 250 µL of anhydrous acetonitrile was added to each 5 mg reagent, vortexed, and allowed to dissolve for 5 mins with occasional vortexing. Reagents (5 µl) were then added to each sample and incubated for 1 hour at room temperature with shaking at 400 rpm. Each sample was then diluted with 30 µl 20% acetonitrile. A portion from each sample was collected as a premix to run on a mass spectrometer to ensure complete labeling. The samples were frozen at -80°C until the results showed efficient labeling. At that point the frozen samples were thawed and the reaction was quenched by adding 8 µL of 5 % hydroxylamine to the sample with incubation for 15 mins at room temperature with shaking at 400 rpm. The samples within each set were then combined and completely dried in the speedvac. Each of the samples were then resuspended in 0.1% trifluoroacetic acid solution and cleaned using Discovery C18 50 mg/1 mL solid phase extraction tubes as described above and once again assayed with bicinchoninic acid to determine the final peptide concentration and vialed for microfractionation.

Microfractionation

Peptide mixtures were separated by high resolution reversed phase UPLC using a nanoACQUITY UPLC® system (Waters Corporation, Milford, MA) equipped with an autosampler. Capillary columns, 200 µm i.d. × 65 cm long, were packed with 300Å 3.0 µm p.s. Jupiter C18 bonded particles (Phenomenex, Torrence, CA). Separations were performed at a flow rate of 2.2 µL/min on binary pump systems, using 10 mM ammonium formate (pH 7.5) as mobile phase A and 100% acetonitrile as mobile phase B. 20 µL of TMT labeled peptide mixtures (0.5 µg/µL) were loaded onto the column and separated using a binary gradient of 1 % B for 35 min, 1-10 % B in 2 min, 10-15 % B in 5 min, and 15-25 % B in 35 min, 25-32 % B in 25min, 32-40 % B in 13 min, 40-90 % B in 43 min, held at 90 % B for 2 min (column washed and equilibrated from 90-50 % B in 2 min, 50-95% B in 2 min, held at 95 % B for 2 min and 95-1 % in 4 min). The capillary column eluent was automatically deposited every minute into 12x32

mm polypropylene vials (Waters Corporation, Milford, MA) starting at minute 60 and ending at minute 170, over the course of the 180 min LC run. Fractions were concatenated by collecting fractions from vial 1 to vial 24 and then returning to vial 1 back to vial 24 (etc.). Prior to peptide fraction collection, 100 µL of water was added to each vial to aid the addition of small droplets into the vial. Each vial was completely dried in a vacuum concentrator (Labconco, Kansas City, MO), reconstituted in 15 µL 2 % acetonitrile, 0.1 % formic acid and stored at -20°C until LC-MS/MS analysis.

TMT quantitation of isolated peptides

A Waters nano-Acquity dual pumping UPLC system (Milford, MA) was configured for on-line trapping of a 5 µL injection at 5 µL/min for 5 min with reverse-flow elution onto the analytical column at 200 nL/min using mobile phase A. The analytical column was slurry packed in-house using Waters BEH C18 1.7 um particles (Waters Chromatography, Drinagh, Ireland) into a 25 cm PicoFrit 360um od x 75um id column (New Objective, Littleton, MA) and held at 45C during use by a 15 cm AgileSleeve flexible capillary heater (Analytical Sales and Services, Inc., Flanders, NJ). The trapping column was slurry packed in-house using Jupiter 5um C18 particles (Phenomenex, Torrence, CA) into a 4cm long piece of 360um od x 150um id fused silica (Polymicro Technologies, Phoenix, AZ) and solgel fritted at both ends. Mobile phases consisted of (A) 0.1% formic acid in water and (B) 0.1% formic acid in acetonitrile with the following gradient profile (min, %B): 0, 1; 10, 8; 105, 25; 115, 35; 120, 75; 123, 95; 129, 95; 130, 50; 132, 95; 138, 95; 140, 1.

MS analysis of 24 fractions per plex set was performed using a Orbitrap Fusion Lumos Tribrid MS (ThermoFisher Scientific) outfitted with a homemade nano-electrospray ionization interface. Electrospray emitters were homemade using 150 µm o.d. × 20 µm i.d. chemically etched fused silica. The ion transfer tube temperature and spray voltage were 250°C and 2.2 kV, respectively. Peptides were separated on an in-house prepared 40 cm x 75 um 1.7-um Waters BEH C18 column, mobile phases consisted of 0.1% formic acid in water (MP-A) and 0.1% formic acid in acetonitrile (MP-B). Data were collected for 120 min following a 10 min delay after completion of sample trapping and start of LC gradient.

FT-MS spectra were acquired from 350 to 1800 m/z at a resolution of 60 k (AGC target 4e5) and the FT-HCD-MS/MS spectra were acquired in data-dependent mode for ions with

detected charge states 2-6 and minimum intensity of 2.5e3 using a normalized collision energy of 30. Other MS/MS parameters were as follows - isolation window was set to 0.7 m/z at a resolution of 50 k, AGC target 1e5, and dynamic exclusion time was 45 s.

Acquired raw spectra were processed using automated MSGF+ software (S. Kim et al., 2008) in target/decoy mode with 20 ppm parent ion mass tolerance, partial tryptic rule with oxidation of methionine (+15.9949) as dynamic modification and static modifications carbamidomethylation of cysteine (+57.0215) and TMTpro modification of lysine on peptide N termini (+304.2071). Each set of samples was matched with corresponding protein collections for *Y. aloifolia* (43,466 protein sequences) and *Y. filamentosa* (45,266 protein sequences). Common contaminant proteins from trypsin and keratins (16 protein sequences) were considered in each search.

Top peptide-spectral matches (PSM) passing 1% FDR cut-off established for the whole datasets of *Y. aloifolia* and *Y. filamentosa* were used for subsequent TMT quantification. MASIC (Monroe et al., 2008) software was used to pull abundances of the sample reporter ions for all PSMs. Measured reporter ion abundances were log transformed to remove skewness of abundance distribution and normalized using mean central tendency normalization. Peptide abundances from all fractions were summed prior to normalization, normalization was performed using InfernoRDN (Polpitiya et al., 2008) software. Protein level roll-up of normalized peptide abundances was normalized again using the same strategy and software.

Proteomic abundance analysis

Normalized protein abundance data was filtered to remove proteins supported by less than two peptide matches as well as proteins that were missing abundance data for more than ⅓ of samples. After filtering, abundance values were converted to z-scores in R to facilitate removal of outlier data points, identification of proteins that fluctuate in abundance throughout the day (a.k.a. “cycling” proteins), and high level comparisons between species and timepoints. For each gene and treatment, individual abundance values more than two standard deviations above or below the next closest abundance value at a given time point were considered outliers and removed prior to identifying cycling proteins.

Cycling proteins were identified and peak mean expression was calculated using a custom R script. To identify cycling proteins, filtered, z-scaled abundance data was divided into drought and well watered conditions prior to fitting polynomial functions for linear, quadratic, cubic, and quartic models. The best fitting model was determined using an ANOVA test. Proteins were considered cycling if any model provided a significantly better fit than the linear model ($p < 0.05$). R scripts used for z-scaling, outlier removal, and identification of cycling proteins are available on GitHub at https://github.com/dawickell/Yucca-omics.

Untargeted metabolomic quantitation and analysis

Dried metabolite extracts were chemically derivatized as previously reported (Y.-M. Kim et al., 2015). Briefly, the extracts were derivatized by methoxyamination and trimethylsilyation (TMS), then the samples were analyzed by GC-MS. Samples were acquired in an Agilent GC 7890A using a HP-5MS column (30 m × 0.25 mm × 0.25 µm; Agilent Technologies, Santa Clara, CA) coupled with a single quadrupole MSD 5975C (Agilent Technologies). One microliter of sample was injected into a splitless port at constant temperature of 250°C. The GC temperature gradient started at 60 °C, with a hold of temperature for 1 minute after injection, followed by increase to 325 °C at a rate of 10 °C/minute and a 10-minute hold at this temperature. A fatty acid methyl ester standard mix (C8-28) (Sigma-Aldrich) was analyzed in parallel as standard for retention time calibration.

GC-MS raw data files were processed using the Metabolite Detector software (Hiller et al., 2009). Retention indices (RI) of detected metabolites were calculated based on the analysis of a FAMEs mixture, followed by their chromatographic alignment across all analyses after deconvolution. Metabolites were initially identified by matching experimental spectra to a PNNL augmented version of Agilent GC-MS metabolomics Library, containing spectra and validated retention indices for over 1200 metabolites. Then, the unknown peaks were additionally matched with the NIST20/Wiley11 GC-MS library. All metabolite identification and quantification ions were validated and confirmed to reduce deconvolution errors during automated data-processing and to eliminate false identifications. Peak areas for identified metabolites and unknowns were normalized to the peak area of the internal standard glycine d5, reported and used for further statistical analysis.

To assess shifts in metabolite abundance over time in *Y. aloifolia* and *Y. filamentosa* we used the MaSigPro package in R. Spectral counts of identified metabolite abundance were tested for significant time structure with the *p.vector (counts=TRUE)* function.

Data Availability

Transcript expression, protein abundance, and metabolite abundance data and associated R and python scripts are available at https://github.com/dawickell/Yucca-omics.

Accession Numbers

Sequence data from this article can be found in NCBI's Sequencing Read Archive; SRA BioProject numbers: 'PRJNA1162856', 'PRJNA1162850', 'PRJNA1162828', 'PRJNA1162846', 'PRJNA1162829', 'PRJNA1162847', 'PRJNA1162862', 'PRJNA1162861', 'PRJNA1162852', 'PRJNA1162845', 'PRJNA1162853', 'PRJNA1162857'.

Supplementary Data

Supplemental Fig 1: Comparison of CAM pathway gene expression in *Y. aloifolia* and *Y. filamentosa*.
Supplemental Fig 2: Mean profiles of mRNA transcripts clustered by similar expression.
Supplemental Figs 3: Comparisons of core CAM genes show a lack of correlation between protein abundance and mRNA expression over time.
Supplemental Figs 4: Comparisons of nitrogen metabolism genes show a lack of correlation between protein abundance and mRNA expression over time.
Supplemental Figs 5: Comparisons of photorespiratory genes show a lack of correlation between protein abundance and mRNA expression over time.
Supplemental Fig 6: Metabolomic data shows nocturnal co-accumulation of organic acids in *Y. aloifolia* relative to *Y. filamentosa*.
Supplementary Table 1: Summary statistics of proteomic and transcriptomic data for genes of interest referenced in this article

Acknowledgements

The authors would like to thank Jen Liddle and Jeremy Balsbaugh for consultation on proteomic analysis, Kerrie Berry for project management at JGI, and Amir Ahkami and Sarah Leichty for project management at EMSL. We would also like to thank the Department of Energy Joint Genome Institute and collaborators for prepublication use of the genome and annotations for *Y.*

aloifolia and *Y. filamentosa* for transcriptome analysis. A portion of this research was performed under the Facilities Integrating Collaborations for User Science (FICUS) program (proposal:10.46936/fics.proj.2020.51552/60000215) and used resources at the DOE Joint Genome Institute (https://ror.org/04xm1d337) and the Environmental Molecular Sciences Laboratory (https://ror.org/04rc0xn13), which are DOE Office of Science User Facilities operated under Contract Nos. DE-AC02-05CH11231 (JGI) and DE-AC05-76RL01830 (EMSL).

Author Contributions

R.F., J.L-M, and K.H designed the research; R.F., K.W., R.C., J.T., N.T., N.M.M., H.H., E.S., and A.L. contributed to data collection, including physiology, transcriptomic, metabolomic, and proteomic; D.W. led analyses with assistance from R.F. and K.H.; D.W. wrote the paper with assistance from K.H.

References:

Abraham, P. E., Yin, H., Borland, A. M., Weighill, D., Lim, S. D., De Paoli, H. C., Engle, N., Jones, P. C., Agh, R., Weston, D. J., Wullschleger, S. D., Tschaplinski, T., Jacobson, D., Cushman, J. C., Hettich, R. L., Tuskan, G. A., & Yang, X. (2016). Transcript, protein and metabolite temporal dynamics in the CAM plant Agave. *Nature Plants*, *2*, 16178.

Agarie, S., Umemoto, M., Sunagawa, H., Anai, T., & Cushman, J. C. (2020). An Agrobacterium-mediated transformation via organogenesis regeneration of a facultative CAM plant, the common ice plant *Mesembryanthemum crystallinum* L. *Plant Production Science*, *23*(3), 343–349.

Baerenfaller, K., Massonnet, C., Walsh, S., Baginsky, S., Bühlmann, P., Hennig, L., Hirsch-Hoffmann, M., Howell, K. A., Kahlau, S., Radziejwoski, A., Russenberger, D., Rutishauser, D., Small, I., Stekhoven, D., Sulpice, R., Svozil, J., Wuyts, N., Stitt, M., Hilson, P., … Gruissem, W. (2012). Systems-based analysis of Arabidopsis leaf growth reveals adaptation to water deficit. *Molecular Systems Biology*, *8*(1), 606.

Balparda, M., Bouzid, M., Martinez, M. D. P., Zheng, K., Schwarzländer, M., & Maurino, V. G. (2023). Regulation of plant carbon assimilation metabolism by post-translational modifications. *The Plant Journal: For Cell and Molecular Biology*, *114*(5), 1059–1079.

Borland, A. M., & Griffiths, H. (1989). The Regulation of Citric Acid Accumulation and Carbon Recycling During CAM in *Ananas comosus*. *Journal of Experimental Botany*, *40*(1), 53–60.

Boxall, S. F., Kadu, N., Dever, L. V., Kneřová, J., Waller, J. L., Gould, P. J., & Hartwell, J. (2020). *Kalanchoë* PPC1 is essential for crassulacean acid metabolism and the regulation of core circadian clock and guard cell signaling genes. *The Plant Cell*, *32*(4), 1136–1160.

Bräutigam, A., Schlüter, U., Eisenhut, M., & Gowik, U. (2017). On the Evolutionary Origin of

CAM Photosynthesis. *Plant Physiology*, *174*(2), 473–477.

Carr, P. D., Verger, D., Ashton, A. R., & Ollis, D. L. (1999). Chloroplast NADP-malate dehydrogenase: structural basis of light-dependent regulation of activity by thiol oxidation and reduction. *Structure* , *7*(4), 461–475.

Chen, L.-Y., Xin, Y., Wai, C. M., Liu, J., & Ming, R. (2020). The role of cis-elements in the evolution of crassulacean acid metabolism photosynthesis. *Horticulture Research*, *7*, 5.

Chen, Y.-B., Lu, T.-C., Wang, H.-X., Shen, J., Bu, T.-T., Chao, Q., Gao, Z.-F., Zhu, X.-G., Wang, Y.-F., & Wang, B.-C. (2014). Posttranslational Modification of Maize Chloroplast Pyruvate Orthophosphate Dikinase Reveals the Precise Regulatory Mechanism of Its Enzymatic Activity. *Plant Physiology*, *165*(2), 534–549.

Costa-Broseta, Á., Castillo, M., & León, J. (2020). Nitrite Reductase 1 Is a Target of Nitric Oxide-Mediated Post-Translational Modifications and Controls Nitrogen Flux and Growth in Arabidopsis. *International Journal of Molecular Sciences*, *21*(19), 7270.

de Sousa Abreu, R., Penalva, L. O., Marcotte, E. M., & Vogel, C. (2009). Global signatures of protein and mRNA expression levels. *Molecular bioSystems*, *5*(12), 1512–1526.

Dever, L. V., Boxall, S. F., 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 Physiology*, *167*(1), 44–59.

Duarte, H. M., & Lüttge, U. (2007). Correlation between photorespiration, CO2-assimilation and spatiotemporal dynamics of photosynthesis in leaves of the C3-photosynthesis/crassulacean acid metabolism-intermediate species *Clusia minor* L. (Clusiaceae). *Trees*, *21*(5), 531–540.

Eisenhut, M., Bräutigam, A., Timm, S., Florian, A., Tohge, T., Fernie, A. R., Bauwe, H., & Weber, A. P. M. (2017). Photorespiration Is Crucial for Dynamic Response of

Photosynthetic Metabolism and Stomatal Movement to Altered CO2 Availability. *Molecular Plant*, *10*(1), 47–61.

Ferrari, R. C., Bittencourt, P. P., Nagumo, P. Y., Oliveira, W. S., Rodrigues, M. A., Hartwell, J., & Freschi, L. (2020). Developing Portulaca oleracea as a model system for functional genomics analysis of C4/CAM photosynthesis. *Functional Plant Biology: FPB*, *48*(7), 666–682.

Franco, A. C., Ball, E., & Lüttge, U. (1992). Differential effects of drought and light levels on accumulation of citric and malic acids during CAM in *Clusia*. *Plant, Cell & Environment*, *15*(7), 821–829.

Freschi, L., Rodrigues, M. A., Tiné, M. A. S., & Mercier, H. (2010). Correlation between citric acid and nitrate metabolisms during CAM cycle in the atmospheric bromeliad *Tillandsia pohliana*. *Journal of Plant Physiology*, *167*(18), 1577–1583.

Giglioli-Guivarc'h, N., Pierre, J. N., Brown, S., Chollet, R., Vidal, J., & Gadal, P. (1996). The Light-Dependent Transduction Pathway Controlling the Regulatory Phosphorylation of C4 Phosphoenolpyruvate Carboxylase in Protoplasts from *Digitaria sanguinalis*. *The Plant Cell*, *8*(4), 573–586.

Gilman, I. S., Moreno-Villena, J. J., Lewis, Z. R., Goolsby, E. W., & Edwards, E. J. (2022). Gene co-expression reveals the modularity and integration of C4 and CAM in *Portulaca*. *Plant Physiology*, *189*(2), 735–753.

Gilman, I. S., Smith, J. A. C., Holtum, J. A. M., Sage, R. F., Silvera, K., Winter, K., & Edwards, E. J. (2023). The CAM lineages of planet Earth. *Annals of Botany*, *132*(4), 627–654.

Gonçalves, A. Z., & Mercier, H. (2021). Transcriptomic and Biochemical Analysis Reveal Integrative Pathways Between Carbon and Nitrogen Metabolism in *Guzmania monostachia*

(Bromeliaceae) Under Drought. *Frontiers in Plant Science*, *12*, 715289.

Hartman, M. D., Rojas, B. E., Iglesias, A. A., & Figueroa, C. M. (2023). The involvement of allosteric effectors and post-translational modifications in the control of plant central carbon metabolism. *The Plant Journal: For Cell and Molecular Biology*, *114*(5), 1037–1058.

Heyduk, K., Burrell, N., Lalani, F., & Leebens-Mack, J. (2016a). Gas exchange and leaf anatomy of a C3-CAM hybrid, *Yucca gloriosa* (Asparagaceae). *Journal of Experimental Botany*, *67*(5), 1369–1379.

Heyduk, K., McAssey, E. V., Field, R., & Leebens-Mack, J. (2023). The Agavoideae: an emergent model clade for CAM evolutionary biology. *Annals of Botany*, *132*(4), 727–737.

Heyduk, K., McAssey, E. V., & Leebens-Mack, J. (2022). Differential timing of gene expression and recruitment in independent origins of CAM in the Agavoideae (Asparagaceae). *The New Phytologist*, *235*(5), 2111–2126.

Heyduk, K., McKain, M. R., Lalani, F., & Leebens-Mack, J. (2016b). Evolution of a CAM anatomy predates the origins of Crassulacean acid metabolism in the Agavoideae (Asparagaceae). *Molecular Phylogenetics and Evolution*, *105*, 102–113.

Heyduk, K., Ray, J. N., Ayyampalayam, S., Moledina, N., Borland, A., Harding, S. A., Tsai, C.-J., & Leebens-Mack, J. (2019). Shared expression of crassulacean acid metabolism (CAM) genes pre-dates the origin of CAM in the genus *Yucca*. *Journal of Experimental Botany*, *70*(22), 6597–6609.

Heyduk, K., Ray, J. N., & Leebens-Mack, J. (2020). Leaf anatomy is not correlated to CAM function in a C3+CAM hybrid species, *Yucca gloriosa*. *Annals of Botany*, *127*(4), 437–449.

Hiller, K., Hangebrauk, J., Jäger, C., Spura, J., Schreiber, K., & Schomburg, D. (2009). MetaboliteDetector: comprehensive analysis tool for targeted and nontargeted GC/MS

based metabolome analysis. *Analytical Chemistry*, *81*(9), 3429–3439.

Igamberdiev, A. U., & Kleczkowski, L. A. (2018). The glycerate and phosphorylated pathways of Serine synthesis in plants: The branches of plant glycolysis linking carbon and nitrogen metabolism. *Frontiers in Plant Science*, *9*, 318.

Jossier, M., Liu, Y., Massot, S., & Hodges, M. (2019). Enzymatic Properties of Recombinant Phospho-Mimetic Photorespiratory Glycolate Oxidases from *Arabidopsis thaliana* and *Zea mays*. *Plants*, *9*(1), 27.

Kaiser, W. M., & Huber, S. C. (2001). Post-translational regulation of nitrate reductase: mechanism, physiological relevance and environmental triggers. *Journal of Experimental Botany*, *52*(363), 1981–1989.

Kim, S., Gupta, N., & Pevzner, P. A. (2008). Spectral probabilities and generating functions of tandem mass spectra: a strike against decoy databases. *Journal of Proteome Research*, *7*(8), 3354–3363.

Kim, Y.-M., Nowack, S., Olsen, M. T., Becraft, E. D., Wood, J. M., Thiel, V., Klapper, I., Kühl, M., Fredrickson, J. K., Bryant, D. A., Ward, D. M., & Metz, T. O. (2015). Diel metabolomics analysis of a hot spring chlorophototrophic microbial mat leads to new hypotheses of community member metabolisms. *Frontiers in Microbiology*, *6*, 209.

Kornas, A., Fischer-Schliebs, E., Lüttge, U., & Miszalski, Z. (2009). Adaptation of the obligate CAM plant *Clusia alata* to light stress: Metabolic responses. *Journal of Plant Physiology*, *166*(17), 1914–1922.

Kundrátová, K., Bartas, M., Pečinka, P., Hejna, O., Rychlá, A., Čurn, V., & Červeň, J. (2021). Transcriptomic and Proteomic Analysis of Drought Stress Response in Opium Poppy Plants during the First Week of Germination. *Plants*, *10*(9), 1878.

Liu, Y., Guérard, F., Hodges, M., & Jossier, M. (2020). Phosphomimetic T335D Mutation of Hydroxypyruvate Reductase 1 Modifies Cofactor Specificity and Impacts Arabidopsis Growth in Air. *Plant Physiology*, *183*(1), 194–205.

Lüttge, U. (1988). Day-night changes of citric-acid levels in crassulacean acid metabolism: phenomenon and ecophysiological significance. *Plant, Cell & Environment*, *11*(6), 445–451.

Lüttge, U. (1990). Nocturnal citrate accumulation and its response to environmental stress in the CAM plant *Kalanchoe pinnata* (Lam.) Pers. *Plant, Cell & Environment*, *13*(9), 977–982.

Lüttge, U. (2006). Photosynthetic flexibility and ecophysiological plasticity: questions and lessons from *Clusia*, the only CAM tree, in the neotropics. *The New Phytologist*, *171*(1), 7–25.

Lüttge, U. (2011). Photorespiration in Phase III of Crassulacean Acid Metabolism: Evolutionary and Ecophysiological Implications. In U. E. Lüttge, W. Beyschlag, B. Büdel, & D. Francis (Eds.), *Progress in Botany 72* (pp. 371–384). Springer Berlin Heidelberg.

McKain, M. R., McNeal, J. R., Kellar, P. R., Eguiarte, L. E., Pires, J. C., & Leebens-Mack, J. (2016). Timing of rapid diversification and convergent origins of active pollination within Agavoideae (Asparagaceae). *American Journal of Botany*, *103*(10), 1717–1729.

Ming, R., VanBuren, R., Wai, C. M., Tang, H., Schatz, M. C., Bowers, J. E., Lyons, E., Wang, M.-L., Chen, J., Biggers, E., Zhang, J., Huang, L., Zhang, L., Miao, W., Zhang, J., Ye, Z., Miao, C., Lin, Z., Wang, H., … Yu, Q. (2015). The pineapple genome and the evolution of CAM photosynthesis. *Nature Genetics*, *47*(12), 1435–1442.

Miszalski, Z., Kornas, A., Rozpądek, P., Fischer-Schliebs, E., & Lüttge, U. (2013). Independent fluctuations of malate and citrate in the CAM species *Clusia hilariana* Schltdl. under low

light and high light in relation to photoprotection. *Journal of Plant Physiology*, *170*(5), 453–458.

Monroe, M. E., Shaw, J. L., Daly, D. S., Adkins, J. N., & Smith, R. D. (2008). MASIC: a software program for fast quantitation and flexible visualization of chromatographic profiles from detected LC-MS(/MS) features. *Computational Biology and Chemistry*, *32*(3), 215–217.

Nobel, P. S. (1983). Nutrient levels in cacti-relation to nocturnal acid accumulation and growth. *American Journal of Botany*, *70*(8), 1244–1253.

O'Leary, B., & Plaxton, W. C. (2015). The central role of glutamate and aspartate in the post-translational control of respiration and nitrogen assimilation in plant cells. In J. P. F. D'Mello (Ed.), *Amino acids in higher plants* (1st ed., pp. 277–297). CAB International.

Ota, K. (1988). CAM Photosynthesis under Drought Conditions in *Kalanchoe blossfeldiana* Grown with Nitrate or Ammonium as the Sole Nitrogen Source. *Plant & Cell Physiology*, *29*(5), 801–806.

Ota, K., & Yamamoto, Y. (1991). Effects of Different Nitrogen Sources and Concentrations on CAM Photosynthesis in *Kalanchoë blossfeldiana*. *Journal of Experimental Botany*, *42*(10), 1271–1277.

Palmieri, M. C., Lindermayr, C., Bauwe, H., Steinhauser, C., & Durner, J. (2010). Regulation of plant glycine decarboxylase by s-nitrosylation and glutathionylation. *Plant Physiology*, *152*(3), 1514–1528.

Pereira, P. N., & Cushman, J. C. (2019). Exploring the Relationship between Crassulacean Acid Metabolism (CAM) and Mineral Nutrition with a Special Focus on Nitrogen. *International Journal of Molecular Sciences*, *20*(18), 4363.

Pereira, P. N., Gaspar, M., Smith, J. A. C., & Mercier, H. (2018). Ammonium intensifies CAM photosynthesis and counteracts drought effects by increasing malate transport and antioxidant capacity in *Guzmania monostachia*. *Journal of Experimental Botany*, *69*(8), 1993–2003.

Polpitiya, A. D., Qian, W.-J., Jaitly, N., Petyuk, V. A., Adkins, J. N., Camp, D. G., 2nd, Anderson, G. A., & Smith, R. D. (2008). DAnTE: a statistical tool for quantitative analysis of -omics data. *Bioinformatics* , *24*(13), 1556–1558.

Ramachandra Reddy, A., & Rama Das, V. S. (1978). The Decarboxylating Systems in Fourteen Taxa Exhibiting CAM Pathway. *Zeitschrift Für Pflanzenphysiologie*, *86*(2), 141–146.

Rodrigues, M. A., Freschi, L., Pereira, P. N., & Mercier, H. (2014). Interactions Between Nutrients and Crassulacean Acid Metabolism. In U. Lüttge, W. Beyschlag, & J. Cushman (Eds.), *Progress in Botany: Vol. 75* (pp. 167–186). Springer Berlin Heidelberg.

Rugen, N., Schaarschmidt, F., Eirich, J., Finkemeier, I., Braun, H.-P., & Eubel, H. (2021). Protein interaction patterns in Arabidopsis thaliana leaf mitochondria change in dependence to light. *Biochimica et Biophysica Acta, Bioenergetics*, *1862*(8), 148443.

Santos, I., & Salema, R. (1991). Nitrogen nutrition and the level of Crassulacean acid metabolism in *Kalanchoë lateritia* Engl. *Plant, Cell & Environment*, *14*(3), 311–317.

Scheible, W.-R., Krapp, A., & Stitt, M. (2000). Reciprocal diurnal changes of phosphoenolpyruvate carboxylase expression and cytosolic pyruvate kinase, citrate synthase and NADP-isocitrate dehydrogenase expression regulate organic acid metabolism during nitrate assimilation in tobacco leaves. *Plant, Cell & Environment*, *23*(11), 1155–1167.

Seabra, A. R., Silva, L. S., & Carvalho, H. G. (2013). Novel aspects of glutamine synthetase (GS) regulation revealed by a detailed expression analysis of the entire GS gene family of

Medicago truncatula under different physiological conditions. *BMC Plant Biology*, *13*(1), 137.

Szal, B., & Podgórska, A. (2012). The role of mitochondria in leaf nitrogen metabolism: Mitochondria and N metabolism. *Plant, Cell & Environment*, *35*(10), 1756–1768.

Tahjib-Ul-Arif, M., Zahan, M. I., Karim, M. M., Imran, S., Hunter, C. T., Islam, M. S., Mia, M. A., Hannan, M. A., Rhaman, M. S., Hossain, M. A., Brestic, M., Skalicky, M., & Murata, Y. (2021). Citric Acid-Mediated Abiotic Stress Tolerance in Plants. *International Journal of Molecular Sciences*, *22*(13), 7235.

Tcherkez, G., Mahé, A., Gauthier, P., Mauve, C., Gout, E., Bligny, R., Cornic, G., & Hodges, M. (2009). *In folio* respiratory fluxomics revealed by 13C isotopic labeling and H/D isotope effects highlight the noncyclic nature of the tricarboxylic acid "cycle" in illuminated leaves. *Plant Physiology*, *151*(2), 620–630.

Timm, S., Florian, A., Arrivault, S., Stitt, M., Fernie, A. R., & Bauwe, H. (2012). Glycine decarboxylase controls photosynthesis and plant growth. *FEBS Letters*, *586*(20), 3692–3697.

Timm, S., & Hagemann, M. (2020). Photorespiration—how is it regulated and how does it regulate overall plant metabolism? *Journal of Experimental Botany*, *71*(14), 3955–3965.

Trípodi, K. E. J., & Podestá, F. E. (2003). Purification and characterization of an NAD-dependent malate dehydrogenase from leaves of the crassulacean acid metabolism plant *Aptenia cordifolia*. *Plant Physiology and Biochemistry*, *41*(2), 97–105.

Vogel, C., & Marcotte, E. M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews. Genetics*, *13*(4), 227–232.

Wai, C. M., Weise, S. E., Ozersky, P., Mockler, T. C., Michael, T. P., & VanBuren, R. (2019).

Time of day and network reprogramming during drought induced CAM photosynthesis in *Sedum album*. *PLOS Genetics*, *15*(6), e1008209.

Wickell, D., Kuo, L.-Y., Yang, H.-P., Dhabalia Ashok, A., Irisarri, I., Dadras, A., de Vries, S., de Vries, J., Huang, Y.-M., Li, Z., Barker, M. S., Hartwick, N. T., Michael, T. P., & Li, F.-W. (2021). Underwater CAM photosynthesis elucidated by *Isoetes* genome. *Nature Communications*, *12*(1), 6348.

Winter, K., Foster, J. G., Schmitt, M. R., & Edwards, G. E. (1982). Activity and quantity of ribulose bisphosphate carboxylase-and phosphoenolpyruvate carboxylase-protein in two Crassulacean acid metabolism plants in relation to leaf age, nitrogen nutrition, and point in time during a day/night cycle. *Planta*, *154*(4), 309–317.

Winter, K., & Smith, J. A. C. (2022). CAM photosynthesis: the acid test. *The New Phytologist*, *233*(2), 599–609.

Wu, H.-Y. L., Jen, J., & Hsu, P. Y. (2024). What, where, and how: Regulation of translation and the translational landscape in plants. *The Plant Cell*, *36*(5), 1540–1564.

Yang, X., Hu, R., Yin, H., Jenkins, J., Shu, S., Tang, H., Liu, D., Weighill, D. A., Cheol Yim, W., Ha, J., Heyduk, K., Goodstein, D. M., Guo, H.-B., Moseley, R. C., Fitzek, E., Jawdy, S., Zhang, Z., Xie, M., Hartwell, J., … Tuskan, G. A. (2017). The *Kalanchoë* genome provides insights into convergent evolution and building blocks of crassulacean acid metabolism. *Nature Communications*, *8*(1), 1899.

Zhang, Y., Swart, C., Alseekh, S., Scossa, F., Jiang, L., Obata, T., Graf, A., & Fernie, A. R. (2018). The Extra-Pathway Interactome of the TCA Cycle: Expected and Unexpected Metabolic Interactions. *Plant Physiology*, *177*(3), 966–979.