

REVIEW ARTICLE

Open Access

# The role of *cis*-elements in the evolution of crassulacean acid metabolism photosynthesis

Li-Yu Chen<sup>1</sup>, Yinghui Xin<sup>2</sup>, Ching Man Wai<sup>2</sup>, Juan Liu<sup>1</sup> and Ray Ming<sup>1,2</sup>

## Abstract

Crassulacean acid metabolism (CAM) photosynthesis is an innovation of carbon concentrating mechanism that is characterized by nocturnal CO<sub>2</sub> fixation. Recent progresses in genomics, transcriptomics, proteomics, and metabolomics of CAM species yielded new knowledge and abundant genomic resources. In this review, we will discuss the pattern of *cis*-elements in stomata movement-related genes and CAM CO<sub>2</sub> fixation genes, and analyze the expression dynamic of CAM related genes in green leaf tissues. We propose that CAM photosynthesis evolved through the re-organization of existing enzymes and associated membrane transporters in central metabolism and stomatal movement-related genes, at least in part by selection of existing circadian clock *cis*-regulatory elements in their promoter regions. Better understanding of CAM evolution will help us to design crops that can thrive in arid or semi-arid regions, which are likely to expand due to global climate change.

## Introduction

Photosynthesis is the process that harvests solar energy to synthesize organic compounds that can ultimately be utilized to drive cellular processes by all forms of life. Photosynthesis is known from cyanobacteria to their descendants including algae and vascular plants<sup>1</sup>. There are three different photosynthetic pathways in terrestrial plants for fixation of carbon dioxide (CO<sub>2</sub>): C<sub>3</sub>, C<sub>4</sub>, and CAM. C<sub>3</sub> photosynthesis is employed by most vascular plants. C<sub>4</sub> plants represent about 3% of vascular plants<sup>2</sup>, while CAM plants represent about 6%<sup>3</sup>. Both C<sub>4</sub> and CAM are add-ons to the C<sub>3</sub> pathway. C<sub>4</sub> and CAM metabolisms are similar in biochemistry but CO<sub>2</sub> concentration steps are spatially separated in C<sub>4</sub> rather than temporally as in CAM. C<sub>4</sub> minimizes photorespiration by concentrating CO<sub>2</sub> in bundle sheath cells, which relies in part on the unique cellular structure (Fig. 1a). Many C<sub>4</sub>

plants are agronomically important species, such as maize and sugarcane<sup>4</sup>. CAM plants have high water-use efficiency (WUE, expressed as mmol CO<sub>2</sub> mol<sup>-1</sup> H<sub>2</sub>O), which is a direct consequence of the fact that they open their stomata at night and keep them closed during the daytime<sup>5</sup>. WUE for carbon assimilation in CAM plants is much higher than in C<sub>3</sub> or C<sub>4</sub> plants. It will be 2–10 times higher than that of C<sub>4</sub> plants and 2.6–20 times higher than that of C<sub>3</sub> plants<sup>6</sup>. One of the major differences between C<sub>4</sub> and CAM photosynthesis centers on the temporal regulation of CO<sub>2</sub> absorption and fixation (Fig. 1).

CAM is found in over 400 genera across 36 families of vascular plants<sup>7</sup> and has evolved multiple times independently from diverse ancestral C<sub>3</sub> plants<sup>3</sup>. While we know ecological factors such as drought condition and CO<sub>2</sub> concentration drive the evolution of CAM<sup>8–10</sup>, far less is known about genetics. Gene family duplication was previously proposed as the driver of CAM metabolism evolution through neofunctionalization of newly duplicated paralogous genes<sup>3</sup>, while others proposed that C<sub>4</sub> and CAM photosynthesis may have arisen through the re-organization of metabolic processes already present in C<sub>3</sub> plants<sup>11–13</sup>.

Correspondence: Ray Ming (rayming@illinois.edu)

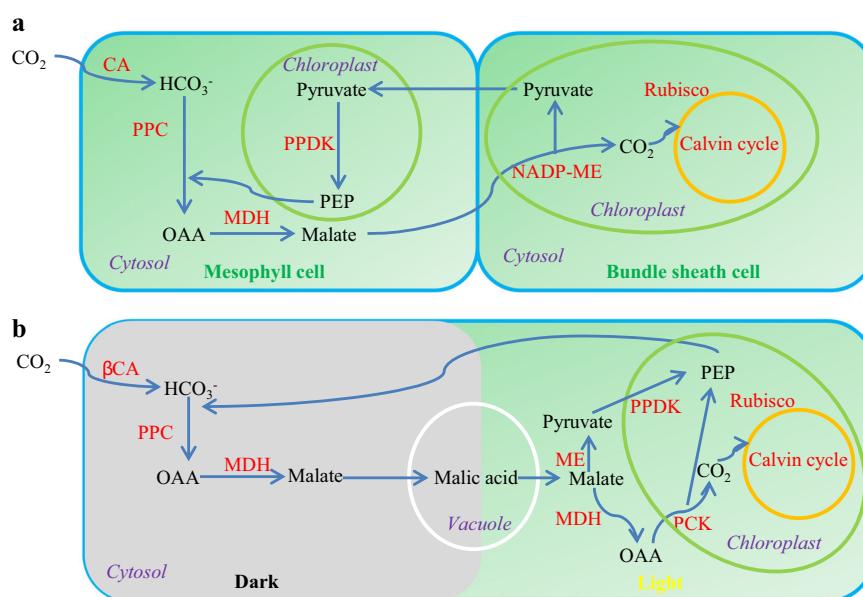
<sup>1</sup>FAFU and UIUC Joint Center for Genomics and Biotechnology, Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, Key Laboratory of Genetics, Breeding and Multiple Utilization of Crops, Ministry of Education, Fujian Agriculture and Forestry University, 350002 Fuzhou, Fujian, China

<sup>2</sup>Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

© The Author(s) 2020



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.



**Fig. 1** Photosynthetic reactions in C<sub>4</sub> and CAM plants. **a** NADP-malic enzyme type of C<sub>4</sub> pathway. **b** Carbon fixation in CAM plants.

The past few years have seen the rapid progresses in genomics, transcriptomics, proteomics, and metabolomics for an increasing number of plant species including CAM species. The orchid *Phalaenopsis equestris* was the first CAM species for which the genome was assembled in 2015<sup>14</sup>. In the same year, the genome of fruit pineapple (*Ananas comosus* var. *comosus*) cultivar 'F153', which has been cultivated by Del Monte for 50 years, was sequenced, and the evolution of CAM photosynthesis was investigated<sup>15</sup>. In a following work, temporal and spatial transcriptomic profiles of CAM-performing mature leaves have also been studied in fruit pineapple<sup>16</sup>. Recently, we sequenced the *bracteatus* pineapple (*Ananas comosus* var. *bracteatus*) accession CB5 genome, and assembled to chromosomal level<sup>17</sup>. The genome of *Kalanchoë fedtschenkoi*, a eudicot CAM species was also available in 2017<sup>18</sup>. Furthermore, multi-dimensional omics data were available for CAM species, such as *Agave americana*<sup>19,20</sup>, orchids<sup>21,22</sup>, and *Talinum triangulare*<sup>23</sup>. These progresses for species that evolved CAM independently provide an excellent resource for comparative analyses, which will help us have a better understanding on the evolution of CAM photosynthesis.

### Cis-elements of stomatal movement-related genes

Stomata were first described as 'pore-like' structures on the surface of leaves over three centuries ago<sup>24</sup>. Since the earliest examples of stomata were discovered in the leaf fossil record, plants have been evolving in terms of size and density of stomata to maintain the maximum leaf conductance as the atmosphere CO<sub>2</sub> changed<sup>25</sup>. Stomata

play an essential role in controlling of transpiration rate and water homeostasis in plants<sup>26</sup>. Stomatal movement can be stimulated by different environmental factors, such as light, abscisic acid (ABA), pathogens, CO<sub>2</sub> and air humidity<sup>27</sup>. Among them, air humidity and ABA are directly related to water status in plants<sup>28</sup>. In CAM plants, the diel rhythms of stomatal conductance and transpiration are closely linked to the net CO<sub>2</sub>-uptake rhythm<sup>5</sup>.

CAM plants present a reverse stomatal conductance pattern by assimilating CO<sub>2</sub> during the night when the temperature is low resulting in lower evapotranspiration rate compared to C<sub>3</sub> and C<sub>4</sub> plants<sup>29,30</sup>. This unique pattern of stomatal movement leads to the higher WUE in CAM plants<sup>31</sup>. The reverse stomatal rhythm has aroused curiosity and investigation for centuries<sup>32</sup>. Understanding the regulation of stomatal movement-related genes in CAM species may provide promising opportunities for engineering crops with higher WUE<sup>32</sup>.

We identified 118 stomatal movement-related genes in *A. comosus* var. *comosus*, 95 in *A. comosus* var. *bracteatus*, 121 in *P. equestris*, 140 in *Arabidopsis*, 123 in rice, and 121 in sorghum (Supplementary Table S1). Based on the GO annotation, the stomatal movement-related genes were divided into three categories, including genes involved in stomatal opening, stomatal closure, and regulation of stomatal movement. For genes involved in stomatal movement, the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)-binding site (CBS; AAAAATCT) and G-box binding site (CACGTG) showed more than 10% or higher frequency than the expected frequencies based on random chance in *A. comosus* var. *comosus* (Table 1). The

G-box element was enriched in genes involved in all three categories in *A. comosus* var. *comosus* (Table 1). The evening element (EE; AAAATATC) and CBS were enriched in 123 stomatal movement-related genes in rice, whereas the morning element (MOE; CCACAC) was only enriched in stomatal opening category in rice (Table 1). When comparing with non-CAM species, Motif ERF73, ERF7, and ABR1 were enriched in CAM species (Supplementary Table S2). Based on these in silico findings, we propose the hypothesis that these different sets of *cis* elements regulate stomatal opening during the day and closure during the night for these three  $C_3$  and  $C_4$  species. Interestingly, stomatal movement-related genes in *A. comosus* var. *comosus* have higher frequencies of circadian clock *cis*-regulatory elements than *A. comosus* var. *bracteatus* (Table 1).

In *A. americana*, the temporal re-programming of particular genes, including  $CO_2$  and ABA signaling and turgor pressure regulating genes are essential to regulate

stomatal movement<sup>19</sup>. Comparative transcriptomic analyses between the  $C_3$  and CAM *Erycina* species also showed that genes involved in light and ABA signaling are altered<sup>22</sup>. The numbers of genes that contain *cis*-elements involved in several key stomatal movement pathways, such as light, ABA, and stress, are summarized (Table 2). ABA responsiveness-related motif (ABRE) appeared most frequently compared to other signaling pathways in the three species (*Arabidopsis*, *P. equestris*, and sorghum) with different photosynthetic pathways. From previous studies, exogenous ABA can induce stomatal closure and the expression and activity of CAM<sup>33,34</sup>. Moreover, stress-related motif (STRE) was the most frequent in the stomatal related genes of rice and pineapple (Table 2). The stress-induced stomatal movement signaling pathway is closely related to the water status of the plant<sup>35</sup>. Further genomic and molecular analysis of potential stomatal movement genes will enable us to have a comprehensive understanding of stomatal biology of CAM plants, and

**Table 1** Frequency of circadian clock-associated motifs (per kb) in 2 kb promoter regions of genes involved in stomatal movement<sup>a</sup>.

Circadian Motif	<i>A. comosus</i> var. <i>comosus</i>	<i>A. comosus</i> var. <i>bracteatus</i>	<i>P. equestris</i>	<i>Arabidopsis</i>	rice	sorghum
<i>Genes involved in stomatal movement<sup>b</sup></i>						
MOE	0.212 (0.250)	0.170 (0.269)	0.132 (0.152)	0.164 (0.215)	0.333 (0.367)	0.281 (0.327)
EE	0.110 (0.105)	0.101 (0.096)	0.107 (0.168)	0.125 (0.124)	<b>0.106</b> (0.059)	0.074 (0.072)
CBS	<b>0.148</b> (0.105)	<b>0.160</b> (0.096)	0.087 (0.168)	0.118 (0.124)	<b>0.069</b> (0.059)	0.058 (0.072)
G-box	<b>0.288</b> (0.250)	0.250 (0.269)	0.120 (0.152)	0.232 (0.215)	0.317 (0.367)	0.248 (0.327)
<i>Regulation of stomatal movement</i>						
MOE	0.209	0.180	0.156	0.174	0.276	0.256
EE	0.115	0.090	0.065	0.107	<b>0.141</b>	<b>0.103</b>
CBS	<b>0.203</b>	<b>0.172</b>	0.097	0.112	<b>0.090</b>	0.064
G-box	<b>0.284</b>	0.270	0.130	<b>0.247</b>	0.359	0.231
<i>Stomatal closure</i>						
MOE	0.185	0.188	0.089	0.121	0.397	0.296
EE	0.093	0.094	<b>0.214</b>	0.136	<b>0.069</b>	0.037
CBS	0.074	<b>0.188</b>	0.071	<b>0.152</b>	0.034	0.074
G-box	<b>0.296</b>	0.000	0.089	0.152	0.259	0.352
<i>Stomatal opening</i>						
MOE	0.250	0.206	0.088	0.184	<b>0.471</b>	0.353
EE	0.111	<b>0.176</b>	0.118	<b>0.184</b>	0.000	0.000
CBS	0.028	0.088	0.059	0.079	0.029	0.000
G-box	<b>0.278</b>	0.235	0.118	<b>0.289</b>	0.206	0.147

<sup>a</sup>Clock-associated motifs with  $\geq 10\%$  frequency higher than expected frequency appearance in stomatal movement-related genes than expected genome wide frequency is highlighted in bold

<sup>b</sup>Species-specific expected frequencies are indicated in brackets. For example, in *A. comosus* var. *comosus*, since the genome GC content is 38%, G/C and A/T occurrence probabilities are 0.19 and 0.31. Both forward and reversed strands were included to calculate the expected frequency of the motif occurrence per kb (For MOE (CCACAC):  $0.19 \times 0.19 \times 0.31 \times 0.19 \times 0.31 \times 0.19 \times 2 \times 1,000 = 0.250$ )

**Table 2** The number of genes and their percentages to the total genes of the genomes that contain *cis*-elements involved in partial key stomatal movement pathways annotated at promoter regions of orthologs in *A. comosus* var. *comosus*, *A. comosus* var. *bracteatus*, *P. equestris*, *Arabidopsis*, rice, and sorghum.

Pathways		<i>A. comosus</i> var. <i>comosus</i>	<i>A. comosus</i> var. <i>bracteatus</i>	<i>P. equestris</i>	<i>Arabidopsis</i>	rice	sorghum
Light responsiveness	Box 4 <sup>a</sup>	107 (0.40%)	69 (0.23%)	30 (0.10%)	116 (0.24%)	89 (0.17%)	76 (0.22%)
	GT1-motif <sup>b</sup>	73 (0.27%)	42 (0.14%)	38 (0.13%)	101 (0.21%)	80 (0.15%)	61 (0.18%)
	GA-motif <sup>c</sup>	25 (0.09%)	11 (0.04%)	19 (0.06%)	29 (0.06%)	17 (0.03%)	33 (0.10%)
	G-Box <sup>d</sup>	40 (0.15%)	18 (0.06%)	36 (0.12%)	40 (0.08%)	50 (0.10%)	105 (0.31%)
	TCT-motif <sup>e</sup>	75 (0.28%)	46 (0.16%)	48 (0.16%)	112 (0.23%)	55 (0.10%)	56 (0.16)
ABA responsiveness	ABRE <sup>f</sup>	101 (0.37%)	65 (0.22%)	58 (0.20%)	123 (0.25%)	115 (0.22%)	112 (0.33%)
Stress	WUN-motif <sup>g</sup>	29 (0.11%)	27 (0.09%)	42 (0.14%)	62 (0.13%)	47 (0.09%)	43 (0.13%)
	STRE <sup>h</sup>	116 (0.43%)	89 (0.30%)	35 (0.12%)	82 (0.17%)	116 (0.22%)	104 (0.30%)

<sup>a</sup> Box 4-motif (ATTAAT)<sup>b</sup> GT1-motif (GGTTAA)<sup>c</sup> GA-motif (ATAGATAA)<sup>d</sup> G-Box (CACGTG)<sup>e</sup> TCT-motif (TCTTAC)<sup>f</sup> ABRE-motif (ACGTG)<sup>g</sup> WUN-motif (AAATTAAT)<sup>h</sup> STRE motif (AGGGG)

might provide candidate genes for engineering crop plants with higher sustainable production<sup>32,36</sup>.

### Diurnal transcript abundance patterns of CAM pathway genes: pineapple as an example

The pineapple genome assembly also allowed the identification of full- and partial-length predicted amino-acid sequences of the key metabolic enzymes comprising the core carboxylation module of CAM responsible for nocturnal fixation of CO<sub>2</sub><sup>15,31,37</sup>. Carbonic anhydrase (CA), catalyzing the conversion of CO<sub>2</sub> into HCO<sub>3</sub><sup>-</sup>, is responsible for the first step in CO<sub>2</sub> assimilation both in C<sub>4</sub> and CAM plants. All three CA subfamily (α, β, and γ) enzymes were identified in pineapple genome (Supplementary Table S3). Only βCA genes (*AccβCA2-1* and *AccβCA2-2*) implicated in CAM-specific roles due to their mRNA abundance in green leaf tissue<sup>15</sup>, indicating that βCA may acts as the enzyme in the initiation of CO<sub>2</sub> fixation.

Three genes encoding the key enzyme PPC responsible for nocturnal CO<sub>2</sub> fixation were identified in the genome assembly, all of which are predicted to be localized to the cytosol as expected<sup>15</sup>. Three PPC genes were identified in *comosus* pineapple genome (Supplementary Table S3, Supplementary Fig. S1). Among these three PPC genes, *AccPPC1* is the most abundant transcript (>3000 FPKM, fragments per kilobase of exon per million fragments mapped) and displayed highest abundance at 6 pm (>5500 FPKM) in CAM-performing leaf tissues. In *T. triangulare*, a facultative CAM species, PPC was upregulated 25-fold (to 15,510 rpm, reads per million) at midnight on day 9 and 12 of water limitation when indicative of CAM was

observed<sup>23</sup>. Comparative transcriptomic analyses between the C<sub>3</sub> and CAM *Erycina* species also demonstrated that PPC gene in CAM *Erycina* displayed higher abundance than in C<sub>3</sub> *Erycina*<sup>22</sup>. These results suggest that high levels of PPC transcripts are important for CAM.

PPC undergoes reversible N-terminal phosphorylation by a circadian clock-controlled PPC kinase (PPCK), which reduces the sensitivity of the enzyme to allosteric inhibition by L-malate and increases its affinity for its substrate phosphoenolpyruvate (PEP)<sup>38,39</sup>. In *A. americana*, which is an obligate CAM plant, *PPCK1* gene displayed diel transcripts abundance pattern, suggesting its important role in temporal re-programming of CAM<sup>20</sup>. In *K. fedtschenkoi*, PPCK1 is also essential for nocturnal CO<sub>2</sub> fixation; moreover, knock-down of oscillations in the transcript abundance of *PPCK1* will lead to the altered accumulation and periodicity of core circadian clock-related transcripts<sup>40</sup>. In pineapple, *AccPPCK2* was found to exhibit greater mRNA abundance than *AccPPCK1*, and *AccPPCK2* also displayed diel mRNA abundance with high levels at night, suggesting that it functions in CAM<sup>15</sup>.

In the final metabolic step of phase I, the OAA formed as a result of PEP carboxylation is reduced to malate by NAD(P)-dependent malate dehydrogenase (MDH). Fourteen genes in pineapple encode MDH: three genes (*AccMDH4*, *AccMDH5*, and *AccMDH8*) are predicted to be cytosolic-localized and strongly expressed in leaves, suggesting their potential to perform functional roles in CAM; four genes (*AccMDH10*, *AccMDH11*, *AccMDH12*, and *AccMDH13*) are tandemly duplicated and lowly expressed except *AccMDH13*<sup>15</sup>.

In *Arabidopsis*, the malate is transported into the vacuole by an inward-rectifying anion-selective ion channel belonging to the aluminium-activated malate transporter (ALMT) family<sup>41</sup>. In *K. fedtschenkoi*, a putative *ALMT6* gene (Kaladp0062s0038) displays diel mRNA abundance in leaves<sup>18</sup>. There are eight candidate *ALMT* family genes in pineapple, including three *ALMT9* genes (*AccALMT9-1-3*) and five *ALMT1* genes (*AccALMT1-1-5*). Only two *ALMT9* genes (*AccALMT9-1* and *AccALMT9-3*) showed high abundant transcript levels in photosynthetic leaf tissues. *ALMT1* only has higher steady-state transcript levels at the midday on day 9 of water limitation in *T. triangulare*<sup>23</sup>. The malate then undergoes protonation, with protons supplied by the tonoplast H<sup>+</sup>-ATPase and H<sup>+</sup>-PPiase, and is stored as malic acid. In the daytime, malic acid is effluxed out of the vacuole possibly through a putative tonoplast dicarboxylate transporter (tDT)<sup>42</sup>. There are five *DT* genes (*AccDT1-5*) in the pineapple genome, and *AccDT2* and *AccDT3* display specifically high abundant transcripts in daytime in photosynthetic leaf tissues, indicating that they may play a role in malic acid efflux in CAM. Decarboxylation of the malate during phase III of the CAM cycle occurs in pineapple primarily via PEP carboxykinase (PCK)<sup>30,43</sup>, which, following oxidation of malate to OAA by NAD(P)-dependent MDH, decarboxylates OAA to PEP. A single *PCK* gene (*AccPCK1*) is present in the pineapple genome and it is predicted to encode a cytosolic enzyme<sup>15</sup>. It is an ortholog of *AtPCK1* (AT4G37870.1), one of two *PCK* genes in *Arabidopsis*, which is expressed in guard cells and is implicated in stomatal closure<sup>44</sup>. Despite the fact that extractable PCK activity from pineapple leaves is over 15 times higher than that of the malic enzymes (MEs)<sup>45</sup>, and it remains possible that malate may also be decarboxylated, in part, by ME in pineapple<sup>46</sup>. The *comosus* pineapple genome contains five *ME* genes encoding both NAD- and NADP-ME (Supplementary Table S3): two *NADP-ME* genes (*AccNADP-ME1* and *AccNADP-ME3*) exhibit higher mRNA levels during the daytime in photosynthetic leaf tissues and one additional *NADP-ME* gene (*AccNADP-ME2*) shows none mRNA transcript in leaves; two *NAD-ME* genes (*AccNAD-ME1* and *AccNAD-ME2*) encoding isoforms predicted to be localized to the mitochondria exhibit moderate abundant mRNA expression and *AccNAD-ME2* also displayed higher mRNA level during the daytime<sup>15</sup>.

The abundant transcript level for *ME* genes in pineapple suggests that malate decarboxylation also results in the formation of pyruvate, which must then be phosphorylated to PEP by pyruvate phosphate dikinase (PPDK). Consistent with this supposition, a single candidate *PPDK1* gene (*AccPPDK1*) was identified in the pineapple genome<sup>15</sup>, providing the metabolic flexibility to allow gluconeogenesis via both the PCK and ME/PPDK

routes<sup>47</sup>. *AccPPDK1* displayed higher transcript abundance during the daytime. The *AtPPDK1* gene encodes an enzyme predicted to be localized to the cytosol, but this enzyme might be localized to either the chloroplast or the cytosol depending upon the production of alternative transcripts arising from two different promoters<sup>48</sup>. More detailed examination of this locus in pineapple is needed to verify this possibility. Overall, the enzymes making up the carboxylation and decarboxylation pathways in the CAM cycle in pineapple are encoded by gene families that are generally smaller than those encoded by the *A. thaliana* genome, because pineapple has one fewer whole-genome duplications than that have been reported for *Arabidopsis* and the grass family<sup>49</sup>.

### Circadian clock-associated *cis*-elements in CAM genes

In most living organisms, internally synchronized circadian clocks make it possible for them to coordinate behavior and physiology corresponding with the 24 h light-dark cycle. CCA1 and LATE ELONGATED HYPOCOTYL (LHY), two single-MYB domain transcription factors, are central to the circadian oscillator of angiosperms<sup>50,51</sup>. *CCA1* and *LHY* are morning expressed genes. They act to suppress the expression of the DNA sequence they bind to. *CCA1* and *LHY* are partially redundant, and they can directly bind to the *TIMING OF CAB EXPRESSION 1 (TOC1)* also known as *PRR1 (PSEUDO-RESPONSE REGULATOR 1)* promoter to negatively regulate its expression<sup>52</sup>.

Circadian control of CAM has been implicated as a core component in diel re-programming of metabolism in CAM plants<sup>20,53</sup>. A comprehensive spatial and temporal survey of gene co-expression clusters in pineapple leaf tissues reveals CAM pathway genes are enriched with clock-associated *cis*-elements, suggesting circadian regulation of CAM<sup>15,16</sup>. At dawn, *CCA1* and *LHY* repress evening-phased genes by binding to CBS and EE<sup>49</sup>. In addition to CBS and EE, the G-box is also enriched in the *CCA1* binding regions<sup>54,55</sup>. *TOC1* can bind to MOE as a negative regulator<sup>56</sup>. In pineapple, all of the three  $\beta$ *CA* genes contain CBS in their promoter regions (Table 3), suggesting they may have function in  $\beta$ *CA* genes' nighttime and early-morning transcripts abundance pattern in photosynthetic leaf tissues. All three copies of *PPC* genes also contain CBS in their promoter regions, along with MOE or G-box (Table 3). Interestingly, CAM pathway genes in *A. comosus* var. *comosus*, contain more circadian clock *cis*-regulatory elements than *A. comosus* var. *bracteatus* (Table 3). Besides the core CAM genes, more than 40% of transcription factors and transcription co-regulators displayed diel rhythmic expression in pineapple, suggesting it is a global adaptation<sup>57</sup>. In a recent work by Heyduk and colleagues (2018), they



**Table 3 Cis-elements annotated at promoter regions of selected CAM photosynthetic genes in pineapple.**

CAM enzyme	<i>A. comosus</i> var. <i>comosus</i>		<i>A. comosus</i> var. <i>bracteatus</i>	
	Gene ID	TF binding motif	Gene ID	TF binding motif
βCA	Aco002732	CBS (2)	CB5.v30091940	CBS (2), G-Box (3)
	Aco005402	CBS (1)	CB5.v30297520	CBS (2)
	Aco006181	CBS (2)	CB5.v30069370	–
PPC	Aco010025	EE (1), MOE (2), CBS (2)	CB5.v30072230	–
	Aco018093	MOE (2), CBS (1)	CB5.v30160110	EE (1), MOE (1)
	Aco016429	G-box (1), CBS (1)	CB5.v30185780	–
PPCK	Aco010095	G-box (1)	CB5.v30098990	G-box (1)
	Aco013938	G-box (1)	CB5.v30291670	–
	Aco013938	G-box (1)	CB5.v30308180	G-box (3)
MDH	Aco006122	CBS (1)	CB5.v30030450	–
	Aco007734	CBS (2)	CB5.v30063990	–
	Aco013935	MOE (2)	CB5.v30119070	–
	Aco002885	MOE (2), G-box (1)	CB5.v30217860	–
	Aco004349	MOE (2), G-box (1)	CB5.v30283140	CBS (2), G-box (2)
	Aco014690	G-box (2)	CB5.v30016950	CBS (1)
	Aco017525	CBS (3)	CB5.v30081100	MOE (1)
	Aco017526	–	CB5.v30308160	MOE (2)
	Aco017527	CBS (1)	CB5.v30167990	–
	Aco017528	CBS (2)	CB5.v30057750	EE (1), CBS (1)
	Aco019631	CBS (1)	CB5.v30057770	–
	Aco010232	MOE (1), G-box (1)	CB5.v30175150	CBS (1)
	Aco004996	MOE (3)	CB5.v30175160	CBS (2)
	Aco008626	–	CB5.v30057740	CBS (1)
	NAD_ME	Aco016569	–	CB5.v30057760
Aco016569		–	CB5.v30106390	–
Aco007622		CBS (1), G-box (1)	CB5.v30132860	CBS (1), G-box (2)
NADP_ME	Aco009967	G-box (4), MOE (1)	CB5.v30038850	G-box (2)
	Aco005631	CBS (1)	CB5.v30285200	CBS (1)
	Aco005989	G-box (2)	CB5.v30300470	G-box (3), ME (1)
PCK	Aco017762	MOE (2), G-box (1), CBS (1)	CB5.v30124740	MOE (1), G-box (1), CBS (1)
PPDK	Aco024818	EE (2), G-box (1), CBS (1)	CB5.v30137360	EE (1), G-box (1), CBS (1)

demonstrated that some canonical CAM genes were unaltered by comparative transcriptomic analyses between the  $C_3$  and CAM *Erycina* species. However, 149 gene families, including genes involved in light and ABA signaling, had significant differences in network connectivity, indicating that transcriptional cascades changes are critical for the transition from  $C_3$  to CAM in *Erycina*<sup>22</sup>.

### Evolution of CAM photosynthesis

$C_4$  and CAM photosynthesis are innovations that evolved in response to decreasing atmospheric levels of  $CO_2$  and water-limiting environments<sup>2,9</sup>. CAM has a higher incidence<sup>3</sup>, and mutation of CAM genes in CAM species is not lethal<sup>40,58</sup>. Both  $C_4$  and CAM have evolved independently multiple times, even within individual families, or even genera during angiosperm evolution<sup>59–61</sup>.

For example, in the Neotropical family Bromeliaceae, to which pineapple belongs, CAM photosynthesis evolved independently at least four, and probably five times<sup>59</sup>.

Recruitment of pre-existing mechanisms underlying C<sub>3</sub> photosynthesis is adopted in *Gynandropsis gynandra* (referred to previously as *Cleome gynandra*), a C<sub>4</sub> plant which is relatively closely related to *Arabidopsis*<sup>62</sup>. Furthermore, gene duplication also plays a profound role in the evolution of C<sub>4</sub>. For example,  $\beta$ CA genes are tandemly duplicated in sorghum<sup>63</sup>. After duplication, some C<sub>4</sub> genes, such as C<sub>4</sub> PPC genes, NADP-MDH genes, and PPK genes, underwent adaptive evolution<sup>63</sup>.

Comparative analyses demonstrated signatures of convergence in protein sequence and re-scheduling of diel transcript abundance of genes involved in nocturnal CO<sub>2</sub> fixation, stomatal movement, heat tolerance, the circadian clock, and carbohydrate metabolism<sup>14,18,21</sup>. Firstly, convergent evolution has been detected in terms of diel cycles of gene transcript abundance<sup>18</sup>. PPK is a key regulator of PPC, which can activate PPC by phosphorylating it. Both *AccPPK2* and *KfPPK2* showed diel expression patterns<sup>18</sup>. Secondly, a convergent amino-acid change in PPC2 was discovered to be shared by *K. fedtschenkoi* and *P. equestris* and the *PPC2* gene in *K. fedtschenkoi* is a much lower abundance transcript relative to the CAM-associated *PPC1* gene, so the function of PPC2 has yet to be linked to CAM directly in either *K. fedtschenkoi* or *P. equestris*<sup>18</sup>.

These findings are consistent with the hypothesis that the CAM photosynthesis evolved as a result of a re-organization of pre-existing metabolic pathways<sup>11,15</sup>. These different features were later coordinated to form the functional CAM photosynthesis.

### Concluding remarks

Genomic studies have led to a renaissance in CAM research. Recent genomic and transcriptomic information from CAM species has improved our understanding of the evolution of CAM photosynthesis<sup>14–21</sup>. The identified candidate genes provide initial targets for detailed functional studies of how the CAM genes have evolved through regulation of gene expression to gain the observed spatial and temporal expression patterns, and loss of repressors is certainly involved. It may be possible for us to apply genome editing to verify functions of candidate CAM genes. CRISPR/Cas9 technology will be a powerful tool to get higher order mutants of tandemly duplicated genes in the same chromosome, which is impossible to generate by traditional mutagenesis methods.

Water loss from stomata for C<sub>3</sub> plants can be very substantial under hot and dry condition. Adjusting the temporal pattern of stomatal movement genes may be a key evolutionary step for switching stomatal opening from the light period to dark<sup>32</sup>. Enrichment of different sets of circadian clock regulatory *cis* elements may have played a

role in this dramatic shift in gene regulation in pineapple and *P. equestris*. CAM photosynthesis and its associated high WUE are key evolutionary innovations that adapted to arid environments and/or low CO<sub>2</sub> environment and this valuable trait is a direct consequence of stomatal closure throughout hottest and driest part for the 24 h cycle, and leaf succulence.

### Acknowledgements

This work was supported by the grant 2016NZ0001-1 from the Department of Science and Technology of Fujian Province, National Natural Science Foundation of China (31701052), Distinguished Young Scholars Fund in Fujian Agriculture and Forestry University (xjq201609), Natural Science Foundation of Fujian Province (Kd18032A), and Program for Excellent Youth Talents in Fujian Province University.

### Conflict of interest

The authors declare that they have no conflict of interest.

**Supplementary Information** accompanies this paper at (<https://doi.org/10.1038/s41438-019-0229-0>).

Received: 23 August 2019 Revised: 1 December 2019 Accepted: 4 December 2019

Published online: 01 January 2020

### References

- Margulis, L. *Symbiosis in Cell Evolution*. (WH Freeman, 1992).
- Sage, R. F. The evolution of C<sub>4</sub> photosynthesis. *New Phytol.* **161**, 341–370 (2004).
- Silvera, K. et al. Evolution along the crassulacean acid metabolism continuum. *Funct. Plant Biol.* **37**, 995–1010 (2010).
- Hibberd, J. M., Sheehy, J. E. & Langdale, J. A. Using C<sub>4</sub> photosynthesis to increase the yield of rice—rationale and feasibility. *Curr. Opin. Plant Biol.* **11**, 228–231 (2008).
- Borland, A. M. et al. Engineering crassulacean acid metabolism to improve water-use efficiency. *Trends Plant Sci.* **19**, 327–338 (2014).
- Borland, A. M., Griffiths, H., Hartwell, J. & Smith, J. A. C. Exploiting the potential of plants with crassulacean acid metabolism for bioenergy production on marginal lands. *J. Exp. Bot.* **60**, 2879–2896 (2009).
- Smith, J. A. C. & Winter, K. in *Crassulacean Acid Metabolism: Biochemistry, Ecophysiology and Evolution* (eds K. Winter & J. A. C. Smith) 1–13 (Springer, 1996).
- Borland, A. M. et al. Climate-resilient agroforestry: physiological responses to climate change and engineering of crassulacean acid metabolism (CAM) as a mitigation strategy. *Plant Cell Environ.* **38**, 1833–1849 (2015).
- Cushman, J. C. Crassulacean acid metabolism. A plastic photosynthetic adaptation to arid environments. *Plant Physiol.* **127**, 1439–1448 (2001).
- Griffiths, H., Cousins, A. B., Badger, M. R. & von Caemmerer, S. Discrimination in the dark. Resolving the interplay between metabolic and physical constraints to phosphoenolpyruvate carboxylase activity during the crassulacean acid metabolism cycle. *Plant Physiol.* **143**, 1055–1067 (2007).
- West-Eberhard, M. J., Smith, J. A. C. & Winter, K. Photosynthesis, reorganized. *Science* **332**, 311–312 (2011).
- Brautigam, A., Schluter, U., Eisenhut, M. & Gowik, U. On the evolutionary origin of CAM photosynthesis. *Plant Physiol.* **174**, 473–477 (2017).
- Heyduk, K., Moreno-Villena, J. J., Gilman, I. S., Christin, P. A. & Edwards, E. J. The genetics of convergent evolution: insights from plant photosynthesis. *Nat. Rev. Genet.* **20**, 485–493 (2019).
- Cai, J. et al. The genome sequence of the orchid *Phalaenopsis equestris*. *Nat. Genet.* **47**, 65–72 (2015).
- Ming, R. et al. The pineapple genome and the evolution of CAM photosynthesis. *Nat. Genet.* **47**, 1435–1442 (2015).
- Wai, C. M. et al. Temporal and spatial transcriptomic and microRNA dynamics of CAM photosynthesis in pineapple. *Plant J.* **92**, 19–30 (2017).
- Chen, L. Y., et al. The *bracteatus* pineapple genome and domestication of clonally propagated crops. *Nat. Genet.* **51**, 1549–1558 (2019).

18. Yang, X. H. et al. The *Kalanchoë* genome provides insights into convergent evolution and building blocks of crassulacean acid metabolism. *Nat. Commun.* **8**, 1899 (2017).
19. Abraham, P. E. et al. Transcript, protein and metabolite temporal dynamics in the CAM plant *Agave*. *Nat. Plants* **2**, 16178 (2016).
20. Yin, H. F. et al. Diel rewiring and positive selection of ancient plant proteins enabled evolution of CAM photosynthesis in *Agave*. *BMC Genom.* **19**, 588 (2018).
21. Zhang, L. S. et al. Origin and mechanism of crassulacean acid metabolism in orchids as implied by comparative transcriptomics and genomics of the carbon fixation pathway. *Plant J.* **86**, 545–545 (2016).
22. Heyduk, K. et al. Altered gene regulatory networks are associated with the transition from C<sub>3</sub> to crassulacean acid metabolism in *Erycina* (Oncidiinae: Orchidaceae). *Front. Plant Sci.* **9**, 2000 (2018).
23. Brihlhaus, D., Brautigam, A., Mettler-Altmann, T., Winter, K. & Weber, A. P. Reversible burst of transcriptional changes during induction of crassulacean acid metabolism in *Talinum triangulare*. *Plant Physiol.* **170**, 102–122 (2016).
24. Woolfenden, H. C. et al. Models and mechanisms of stomatal mechanics. *Trends Plant Sci.* **23**, 822–832 (2018).
25. Beerling, D. J. & Franks, P. J. Evolution of stomatal function in 'lower' land plants. *New Phytol.* **183**, 921–925 (2009).
26. Grondin, A. et al. Aquaporins contribute to ABA-triggered stomatal closure through OST1-mediated phosphorylation. *Plant Cell* **27**, 1945–1954 (2015).
27. Kollist, H., Nuhkat, M. & Roelfsema, M. R. G. Closing gaps: linking elements that control stomatal movement. *New Phytol.* **203**, 44–62 (2014).
28. Taybi, T. & Cushman, J. C. Signaling events leading to crassulacean acid metabolism induction in the common ice plant. *Plant Physiol.* **121**, 545 (1999).
29. And, R. A. F. & Turner, N. C. Plant productivity in the arid and semiarid zones. *Annu. Rev. Plant Physiol.* **29**, 277–317 (1978).
30. Osmond, C. B. Crassulacean acid metabolism: a curiosity in context. *Annu. Rev. Plant Biol.* **29**, 379–414 (1978).
31. Ming, R., Wai, C. M. & Guyot, R. Pineapple genome: a reference for monocots and CAM photosynthesis. *Trends Genet.* **32**, 690–696 (2016).
32. Males, J. & Griffiths, H. Stomatal biology of CAM plants. *Plant Physiol.* **174**, 550 (2017).
33. Taybi, T. et al. Differential effects of abscisic acid on phosphoenolpyruvate carboxylase and CAM operation in *Kalanchoë blossfeldiana*. *Plant Biol.* **108**, 240–246 (2015).
34. Chu, C. & Edwards, G. E. Induction of crassulacean acid metabolism in the facultative halophyte *Mesembryanthemum crystallinum* by abscisic acid. *Plant Physiol.* **93**, 1253–1260 (1990).
35. El-Sharkaway, M. A. & Cock, J. H. Water use efficiency of cassava. I. Effects of air humidity and water stress on stomatal conductance and gas exchange. *Crop Sci.* **24**, 497–502 (1984).
36. Lee, J. S. Stomatal opening mechanism of CAM plants. *J. Plant Biol.* **53**, 19–23 (2010).
37. Xu, H. M. et al. PGD: Pineapple genomics database. *Hortic. Res.* **5**, 66 (2018).
38. Nimmo, G. A., Nimmo, H. G., Fewson, C. A. & Wilkins, M. B. Diurnal changes in the properties of phosphoenolpyruvate carboxylase in *Bryophyllum* leaves: a possible covalent modification. *FEBS Lett.* **178**, 199–203 (1984).
39. Carter, P. J., Nimmo, H. G., Fewson, C. A. & Wilkins, M. B. Circadian rhythms in the activity of a plant protein kinase. *EMBO J.* **10**, 2063–2068 (1991).
40. Boxall, S. F., Dever, L. V., Knerova, J., Gould, P. D. & Hartwell, J. Phosphorylation of phosphoenolpyruvate carboxylase is essential for maximal and sustained dark CO<sub>2</sub> fixation and core circadian clock operation in the obligate crassulacean acid metabolism species *Kalanchoë fedtschenkoi*. *Plant Cell* **29**, 2519–2536 (2017).
41. Kovermann, P. et al. The Arabidopsis vacuolar malate channel is a member of the ALMT family. *Plant J.* **52**, 1169–1180 (2007).
42. Emmerlich, V. et al. The plant homolog to the human sodium/dicarboxylic cotransporter is the vacuolar malate carrier. *Proc. Natl Acad. Sci. USA* **100**, 11122–11126 (2003).
43. Christopher, J. T. & Holtum, J. A. Patterns of carbon partitioning in leaves of crassulacean acid metabolism species during deacidification. *Plant Physiol.* **112**, 393–399 (1996).
44. Penfield, S. et al. Expression and manipulation of PHOSPHOENOLPYRUVATE CARBOXYKINASE 1 identifies a role for malate metabolism in stomatal closure. *Plant J.* **69**, 679–688 (2012).
45. Ditttrich, P., Campbell, W. H. & Black, C. C. Phosphoenolpyruvate carboxykinase in plants exhibiting crassulacean acid metabolism. *Plant Physiol.* **52**, 357–361 (1973).
46. Hong, H. T. K., Nose, A. & Agarie, S. Respiratory properties and malate metabolism in percoll-purified mitochondria isolated from pineapple, *Ananas comosus* (L.) Merr. cv. smooth cayenne. *J. Exp. Bot.* **55**, 2201–2211 (2004).
47. Eastmond, P. J. et al. Arabidopsis uses two gluconeogenic gateways for organic acids to fuel seedling establishment. *Nat. Commun.* **6**, 6659 (2015).
48. Parsley, K. & Hibberd, J. M. The Arabidopsis *PPDK* gene is transcribed from two promoters to produce differentially expressed transcripts responsible for cytosolic and plastidic proteins. *Plant Mol. Biol.* **62**, 339–349 (2006).
49. Harmer, S. L. & Kay, S. A. Positive and negative factors confer phase-specific circadian regulation of transcription in Arabidopsis. *Plant Cell* **17**, 1926–1940 (2005).
50. Wang, Z. Y. & Tobin, E. M. Constitutive expression of the *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) gene disrupts circadian rhythms and suppresses its own expression. *Cell* **93**, 1207–1217 (1998).
51. Schaffer, R. et al. The *late elongated hypocotyl* mutation of *Arabidopsis* disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* **93**, 1219–1229 (1998).
52. Alabadi, D. et al. Reciprocal regulation between *TOC1* and *LHY/CCA1* within the *Arabidopsis* circadian clock. *Science* **293**, 880–883 (2001).
53. Wilkins, M. B. The role of the epidermis in the generation of the circadian rhythm of carbon dioxide fixation in leaves of *Bryophyllum fedtschenkoi*. *Planta* **185**, 425–431 (1991).
54. Nagel, D. H. et al. Genome-wide identification of CCA1 targets uncovers an expanded clock network in Arabidopsis. *Proc. Natl Acad. Sci. USA* **112**, E4802–E4810 (2015).
55. Kamioka, M. et al. Direct repression of evening genes by *CIRCADIAN CLOCK-ASSOCIATED1* in the Arabidopsis circadian clock. *Plant cell* **28**, 696–711 (2016).
56. Gendron, J. M. et al. Arabidopsis circadian clock protein, *TOC1*, is a DNA-binding transcription factor. *Proc. Natl Acad. Sci. USA* **109**, 3167–3172 (2012).
57. Sharma, A., Wai, C. M., Ming, R. & Yu, Q. Y. Diurnal cycling transcription factors of pineapple revealed by genome-wide annotation and global transcriptomic analysis. *Genome Biol. Evol.* **9**, 2170–2190 (2017).
58. Dever, L. V., Boxall, S. F., Knerova, J. & Hartwell, J. Transgenic perturbation of the decarboxylation phase of crassulacean acid metabolism alters physiology and metabolism but has only a small effect on growth. *Plant Physiol.* **167**, 44–59 (2015).
59. Crayn, D. M., Winter, K., Schulte, K. & Smith, J. A. C. Photosynthetic pathways in Bromeliaceae: phylogenetic and ecological significance of CAM and C<sub>3</sub> based on carbon isotope ratios for 1893 species. *Bot. J. Linn. Soc.* **178**, 169–221 (2015).
60. Fisher, A. E. et al. Evolutionary history of *Blepharis* (Acanthaceae) and the origin of C<sub>4</sub> photosynthesis in section *Acanthodium*. *Int. J. Plant Sci.* **176**, 770–790 (2015).
61. Heyduk, K., McKain, M. R., Lalani, F. & Leebens-Mack, J. Evolution of a CAM anatomy predates the origins of crassulacean acid metabolism in the Agavoideae (Asparagaceae). *Mol. Phylogenet. Evol.* **105**, 102–113 (2016).
62. Brown, N. J. et al. Independent and parallel recruitment of preexisting mechanisms underlying C<sub>4</sub> photosynthesis. *Science* **331**, 1436–1439 (2011).
63. Wang, X. Y. et al. Comparative genomic analysis of C<sub>4</sub> photosynthetic pathway evolution in grasses. *Genome Biol.* **10**, R68 (2009).