Evolution along the crassulacean acid metabolism continuum

Katia Silvera\textsuperscript{A}, Kurt M. Neubig\textsuperscript{B}, W. Mark Whitten\textsuperscript{B}, Norris H. Williams\textsuperscript{B}, Klaus Winter\textsuperscript{C} and John C. Cushman\textsuperscript{A,D}

\textsuperscript{A}Department of Biochemistry and Molecular Biology, MS200, University of Nevada, Reno, NV 89557-0200, USA.  
\textsuperscript{B}Florida Museum of Natural History, University of Florida, Gainesville, FL 32611-7800, USA.  
\textsuperscript{C}Smithsonian Tropical Research Institute, PO Box 0843-03092, Balboa, Ancón, Republic of Panama.  
\textsuperscript{D}Corresponding author. Email: jcushman@unr.edu

This paper is part of an ongoing series: ‘The Evolution of Plant Functions’.

Abstract.  Crassulacean acid metabolism (CAM) is a specialised mode of photosynthesis that improves atmospheric CO\textsubscript{2} assimilation in water-limited terrestrial and epiphytic habitats and in CO\textsubscript{2}-limited aquatic environments. In contrast with C\textsubscript{3} and C\textsubscript{4} plants, CAM plants take up CO\textsubscript{2} from the atmosphere partially or predominantly at night. CAM is taxonomically widespread among vascular plants and is present in many succulent species that occupy semiarid regions, as well as in tropical epiphytes and in some aquatic macrophytes. This water-conserving photosynthetic pathway has evolved multiple times and is found in close to 6\% of vascular plant species from at least 35 families. Although many aspects of CAM molecular biology, biochemistry and ecophysiology are well understood, relatively little is known about the evolutionary origins of CAM. This review focuses on five main topics: (1) the permutations and plasticity of CAM, (2) the requirements for CAM evolution, (3) the drivers of CAM evolution, (4) the prevalence and taxonomic distribution of CAM among vascular plants with emphasis on the Orchidaceae and (5) the molecular underpinnings of CAM evolution including circadian clock regulation of gene expression.

Additional keywords: phosphoenolpyruvate carboxylase, photosynthesis, \(\delta^{13}\)C.

Introduction

Crassulacean acid metabolism (CAM) is one of three modes of photosynthetic assimilation of atmospheric CO\textsubscript{2}, along with C\textsubscript{3} and C\textsubscript{4} photosynthesis. The net result of CAM is an improvement in water use efficiency (WUE; CO\textsubscript{2} fixed per unit water lost) generally 6-fold higher than for C\textsubscript{3} plants and 3-fold higher than for C\textsubscript{4} plants under comparable conditions (Nobel\textsuperscript{1996}). Thus, CAM is an important ecophysiological metabolic adaptation that permits plants to occupy semiarid habitats and habitats with intermittent or seasonal water availability (Winter and Smith\textsuperscript{1996}; Cushman\textsuperscript{2001}). In this review, we examine the permutations and plasticity of CAM in the context of evolution, discuss the metabolic and genetic requirements for CAM – including leaf succulence – and consider the likely drivers for the evolution of CAM. We next review the current surveys of the taxonomic distribution of CAM species and the several survey methods used to estimate the prevalence of CAM. We then discuss the molecular evolution of CAM, including its origins, describe molecular markers used to study the evolutionary progression of gene family changes and to analyse circadian clock control. Finally, we address future directions for research.

Phases of CAM

The physiological and biochemical temporal sequence of events that constitute CAM have been described in detail as being separable into four discrete phases (Osmond\textsuperscript{1978}; Winter\textsuperscript{1985}; Lüttge\textsuperscript{1987}; Griffiths\textsuperscript{1988}). Phase I is typically characterised by nocturnal stomatal opening, CO\textsubscript{2} uptake and fixation by phosphoenolpyruvate carboxylase (PEPC) in the cytosol and the formation of C\textsubscript{4} organic acids (usually malic acid), which are stored in the vacuole (Fig. 1). The rate of nocturnal CO\textsubscript{2} assimilation is governed by mesophyll processes, such as regulation of carbohydrate storage reserves (Cushman et al.\textsuperscript{2008a}) or vascular storage capacity, rather than stomatal conductance (Winter\textsuperscript{1985}; Winter et al.\textsuperscript{1985}). Depending on the CAM species, a variety of storage carbohydrates (e.g. starch, glucans, soluble hexoses) might be catabolised to produce phosphoenolpyruvate (PEP), the substrate for carboxylation (Christopher and Holturn\textsuperscript{1996, 1998}; Holturn et al.\textsuperscript{2005}). Phase I reflects the fundamental adaptation of CAM that results in reduced transpiration and improved water economy due to lower night-time evaporative demand and associated water losses (Griffiths\textsuperscript{1988}). Phase II describes the transition from PEPC to ribulose-1,5-bisphosphate carboxylase/oxygenase (RUBISCO)-mediated carboxylation during the early light period leading to carbohydrate production. During this phase, CO\textsubscript{2} is derived from both organic acid decarboxylation and direct uptake from the atmosphere. Phase III encompasses the period of major efflux of organic acids from the vacuole and their subsequent decarboxylation (Fig. 1). This decarboxylation can lead to
generation of internal leaf CO2 partial pressures greater than 100 times atmospheric levels (Cockburn et al. 1979; Spalding et al. 1979), reduction in stomatal opening and transpiration and sometimes even release of CO2 from the leaf despite low stomatal conductance (Frimert et al. 1986). Decarboxylation is catalysed by either cytosolic PEP carboxykinase (PEPCK) or cytosolic NADP+- and/or mitochondrial NAD+-malic enzymes (ME) (Smith and Bryce 1992; Christopher and Holtum 1996; Holtum et al. 2005). This CO2-concentrating mechanism or ‘CO2 pump’ effectively suppresses photorespiration during this phase.

Fig. 1. Postulated major CAM metabolite fluxes and transport requirements during the night and day. During night-time (CAM phase I, black arrows), while stomata are open, atmospheric CO2 enters the cell and is fixed along with respired CO2 as bicarbonate (HCO3-) by PEPC leading to the formation of oxaloacetate and then malate, which undergoes protonation and is stored in the vacuole as malic acid. During day-time (CAM phase III, yellow arrows), while stomata are closed, malic acid exits the vacuole and is decarboxylated, CO2 released in the cytosol or mitochondria is then refixed by RUBISCO through the PCR cycle within the chloroplast. Pyruvate produced in the cytosol or mitochondria moves into the plastid to serve as the substrate for the gluconeogenic pathway where carbohydrates are regenerated and stored as starch. For simplicity, only key reactions of metabolic flux are shown with key enzymes indicated in red font. Cofactors, charge state and allosteric interactions of metabolites are not shown for simplicity. Model shown is for a starch-forming CAM plant such as Mesembryanthemum crystallinum, which uses NADP-ME and NAD-ME for decarboxylation of malate. Some CAM species utilise PEP carboxykinase (PEPCK), which decarboxylates OAA resulting in the formation of PEP (not shown). Hexose-storing CAM species, such as Ananas comosus store hexose in the vacuole that is transported to the cytosol (see Holtum et al. 2005). G6P, glucose-6-phosphate; TP, triose phosphate; PGA, 3-phosphoglycerate; PEP, phosphoenolpyruvate; OAA, oxaloacetate; MAL, malate; FUM, fumarate; PYR, pyruvate; TCA cycle, tricarboxylic acid cycle; PCR cycle, photosynthetic carbon reduction cycle. Dashed arrows indicate multiple metabolic steps or diffusion. Key enzymes include CA, carbonic anhydrase; PEPC, cytosolic phosphoenolpyruvate carboxylase; PPCK, PEPC kinase; NADP-ME, cytosolic NADP-dependent malic enzyme; NAD-ME, mitochondrial NAD-dependent malic enzyme; NAD-MDH, cytosolic and mitochondrial NAD-dependent malate dehydrogenase; PPDK, plastidic pyruvate orthophosphate dikinase; RUBISCO, plastidic ribulose-1,5-bisphosphate carboxylase/oxygenase. Transports indicated by numbered squares include: (1) pGlcT, plastidic glucose transporter; (2) GPT, glucose-6-phosphate/phosphate transporter; (3) TPT, triose-phosphate/phosphate transporter; (4) PYT, plastidic pyruvate transporter; (5) PPT, plastidic phosphoenolpyruvate/phosphate transporter; (6) mDCT, putative mitochondrial dicarboxylate carriers; (7) putative mitochondrial pyruvate transporter; (8) inward-rectifying anion-selective malate channel; (9) V-ATPase, tonoplast vacuolar H+-ATPase complex; (10) V-PPase, tonoplast vacuolar PPase complex; (?) unknown protein proposed to mediate malic acid efflux, possibly a tonoplast dicarboxylate transporter. This model is derived, in part, from information summarised by Holtum et al. (2005).
Phase IV is a second transitional phase marked by the depletion of organic acid stores, slower rates of decarboxylation, reduction in internal partial pressure of CO₂ and increases in stomatal conductance, depending on the prevailing environmental conditions. CO₂ fixation during early phase IV is a mixture of CO₂ assimilation derived mainly from organic acid decarboxylation and direct atmospheric uptake and assimilation via RUBISCO; however, carboxylation into C₄ acids by PEPC may commence as the dark period approaches (Ritz et al. 1986). Because CAM plants perform both CO₂ fixation steps within the same cell, futile cycling of CO₂ is minimised by temporal control of the kinetic properties of PEPC in response to malic acid (Winter 1982) and of RUBISCO in response to light (Maxwell et al. 1999; Griffiths et al. 2002). The diel changes in kinetic properties of PEPC (Winter 1982) are triggered by reversible phosphorylation events catalysed by a dedicated protein kinase (Hartwell et al. 1996, 1999; Taybi et al. 2000), whose expression is controlled by the circadian clock (Hartwell 2005a, 2005b).

**Permutations of CAM**

Although the four CAM phase definitions appear adequate to describe all observed acid metabolism phenomena (Lüttge 1987), additional terminologies have been suggested to describe CAM in some astomatal aquatic species and in the astomatal green aerial roots of epiphytic orchids (Cockburn 1985). Furthermore, environmental conditions can modulate the extent to which each phase is manifested (see also next section) (Cushman 2001; Cushman and Borland 2002). For example, water deficit stress can reduce or eliminate phase IV and light and temperature can regulate the appearance or onset of phases II and III (Griffiths 1988). Under severe water deficit stress, phase I net nocturnal CO₂ uptake can be eliminated completely along with virtually all stomatal conductance across the four phases. This phenomenon, termed ‘CAM idling’, results in small, sustained diel fluctuations in organic acids with essentially all of the CO₂ fixed into malate being derived from internally recycled respiratory CO₂ (Szarek et al. 1973; Ting 1985). CAM idling might play an important role in the prevention of photoinhibition by maintaining photosystem stability (Osmond 1982). The phenomena of ‘CAM cycling’ or ‘weak CAM’ have also been described, wherein organic acid fluctuations are observed, but with little or no net nocturnal CO₂ fixation by PEPC (Sipes and Ting 1985; Ting 1985). In the context of evolution, CAM cycling has been interpreted to be a basal form of CAM, while increasing reliance on nocturnal CO₂ fixation has been associated with an increasingly advanced state among the Crassulaceae (Teeri 1982a, 1982b) and the Bromeliaceae (Smith et al. 1986). The ecophysiological significance of CAM cycling might be to keep plants poised to engage in full CAM once drought conditions end by maintaining the capacity for organic acid fluctuation (Ting 1985). Similarly, the evolutionary importance of weak CAM might be that it serves as a genetic reservoir for CAM radiations in the context of changing environmental conditions or habitat exploitation, such as epiphytism (Silvera et al. 2005, 2009). Finally, the term ‘latent CAM’ has been used to describe an intermediate form of CAM wherein organic acid concentrations remain high but constant throughout the diel cycle (Schuber and Kluge 1981). As with CAM cycling or weak CAM, latent CAM might be regarded as a step along the progression from C₃ to CAM (Lee and Griffiths 1987).

**CAM plasticity**

The degree to which CAM operates can vary greatly depending on the evolutionary history of a given species and its environmental context, resulting in a continuum of differences in the degree to which nocturnal net uptake of CO₂ occurs in relation to day-time net CO₂ uptake (Cushman and Bohnert 1999; Cushman 2001; Cushman and Borland 2002; Dodd et al. 2002). For example, many CAM species engage in ‘obligate’ or ‘constitutive’ CAM in fully mature photosynthetic organs (i.e. leaves and stems), although the extent of gas exchange and nocturnal acidification might be modulated by prevailing environmental conditions (Griffiths 1988). Many members of the Cactaceae and Crassulaceae provide excellent examples of this type of CAM. In contrast, ‘facultative’, ‘inducible’, or ‘optional’ CAM or C₄-CAM intermediate species engage in CAM in response to environmental stimuli such as drought stress (Winter 1985; Griffiths 1988; Winter et al. 2008). The expression of CAM in such C₃-CAM species varies dynamically with experimentally manipulated conditions, such as photoperiod (Brulé and Queiroz 1982), water status, light, temperature, nutritional status, salinity, anoxia, or atmospheric CO₂ concentration (Winter 1985; Lüttge 1987; Griffiths 1988; Roberts et al. 1997). The common ice plant, Mesembryanthemum crystallinum L., a member of the Aizoaceae, is a well studied example of inducible CAM under strict environmental control (Winter 1985; Winter and Holtum 2007; Cushman et al. 2008a, 2008b).

**Requirements for CAM**

Because the basic enzymatic machinery essential for CAM operation is assumed to be present in the chloroplast-containing cells of all green plant species, what are the evolutionary changes that must occur in order for CAM to function? The first and foremost diagnostic indicator of CAM is nocturnal CO₂ uptake (Fig. 2). Second, diel fluctuations in organic acids and reciprocal fluctuations of storage carbohydrates such as starch, glucans, or soluble hexoses typically occur as features of the CAM cycle (Ting 1985). Third, associated transport activities across the tonoplast (e.g. vacuolar H⁺-ATPase), mitochondrial (Holtum et al. 2005; Cushman et al. 2008b), and chloroplast envelope membranes (Häusler et al. 2000; Kore-eda et al. 2005) are needed to support these fluctuations (Fig. 1). Fourth, enhanced expression of PEPC and decarboxylating (e.g. PEPC or NADP⁺/−-NAD⁺-ME) enzymes is necessary. Fifth, enhanced expression of enzymes of both the glycolytic and gluconeogenic pathways is required to support the synthesis of large (typically 40–60% of available reserves) reciprocating pools of carbohydrates (Paul et al. 1993; Borland and Dodd 2002; Dodd et al. 2003). Discrete isoenzymes appear to be recruited selectively in order to conduct the activities necessary for CAM function (Kore-eda et al. 2005; Cushman et al. 2008b).

Sixth, some degree of leaf succulence, characterised by increased mesophyll cell size due to large storage vacuoles and increased mesophyll tissue and leaf thickness are often characteristic of CAM species. Such large cell volumes per unit leaf or stem area
Increased gene expression

Increased leaf succulence

Circadian clock control

Fig. 2. Major requirements hypothesised for the evolution of CAM along the evolutionary progression from C3 photosynthesis to weak CAM and strong CAM indicated by colour transitions from green to orange and red, respectively. The relative contribution of nocturnal CO2 uptake increases as indicated by increased dark shading and height of triangle. Fluctuations in titratable acidity increase due to diel fluctuations in organic acid accumulation and reciprocal fluctuations in storage carbohydrates such as starch, glucans, or soluble hexoses as indicated by white (day-time) and black (night-time) arrows demonstrating increasing diel flux dynamics. These fluctuations are accompanied by associated increases in transport activities across the tonoplast (e.g. vacuolar H+ ATPase), mitochondrial and chloroplast envelope membranes (not shown). Putative expansion in the complexity of gene families due to gene duplication events accompanied by selective recruitment of CAM-specific isoforms is indicated by orange and red lines. Diagnostic indicators of gene recruitment include a progressive increase in leaf-specific or leaf-preferential mRNA and protein expression and overall increased expression of PEPC, decarboxylating (e.g. PEPC kinase or NADP+/- NAD+ ME) enzymes and enzymes of both the glycolytic and gluconeogenic pathways indicated by orange and red coloured leaves and bars, respectively. CAM evolution is well correlated with increased leaf succulence characterised by mesophyll cells with increased size, increased mesophyll tissue and leaf thickness due to large storage vacuoles as indicated by the increased thickness of leaf models drawn in side view. Finally, circadian clock control is thought to progress by both shifts in the phase of circadian clock output and increased magnitude of circadian expression patterns, particularly at the level of mRNA expression. White shaded box, day-time, black shaded box, night-time. ensure a high capacity for nocturnal organic acid storage and water storage (Gibson 1982; Nelson et al. 2005). Lastly, circadian clock control of CO2 fixation, and mRNA and post-translational regulatory events, such as the reversible phosphorylation of PEPC by PEPC kinase (Fig. 1), are required to ensure that reciprocating organic acid and carbohydrate pools are properly synchronised along the diel CAM cycle (Borland et al. 1999; Hartwell et al. 1999; Taybi et al. 2000; Dodd et al. 2003; Cushman et al. 2008a, 2008b).

The differential enzyme-mediated discrimination against 13CO2 during photosynthetic carbon assimilation between C3 photosynthesis and CAM results in different whole-tissue carbon isotope ratios (δ13C) (Ehleringer and Osmond 1989). Species exhibiting pronounced CAM typically show δ13C values less negative than −20‰; whereas, for C3 plants, δ13C values may range from about −33 to −22‰ (Ehleringer and Osmond 1989; Griffiths 1992) depending, for example, on interspecific variation in the chemical and diffusional processes that contribute to the carbon isotopic signature, plant water status, altitude and plant position within a forest canopy. Thus, δ13C values have become widely used as a rapid and relatively inexpensive screening method for determining the presence of strong CAM. However, low level CAM activity or ‘weak CAM’ species have δ13C values that overlap with those of C3 species because the majority of CO2 is being fixed by the C3 pathway (Table 1) (Winter and Holtum 2002). Small changes in integrated tissue δ13C values caused by small amounts of dark CO2 fixation typical of weak CAM species are generally not significantly different from the wide range of δ13C values exhibited by C3 species in large species surveys due to variations in plant biochemistry, plant-environment interactions, and the 13C/12C composition of the source air under field conditions (Griffiths 1992; Winter and Holtum 2002).

Table 1. δ13C values, nocturnal fluctuation in titratable acidity, and leaf traits from 173 orchid species

<table>
<thead>
<tr>
<th>Trait</th>
<th>C3</th>
<th>Weak CAM</th>
<th>Strong CAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>δ13C (%) mean ± s.d.</td>
<td>−27.7 ± 1.8</td>
<td>−26.7 ± 2.3</td>
<td>−16.1 ± 2.6</td>
</tr>
<tr>
<td>Titratable acidity (μmol H+ g⁻¹ FW)</td>
<td>2.0 ± 3.2</td>
<td>12.2 ± 9.5 (+6.1)</td>
<td>75.8 ± 60.2 (+6.2)</td>
</tr>
<tr>
<td>Leaf thickness (mm)</td>
<td>0.5 ± 0.4</td>
<td>0.5 ± 0.3 (0)</td>
<td>2.0 ± 2.01 (+4.1)</td>
</tr>
<tr>
<td>SLA (cm² g⁻¹)</td>
<td>167 ± 78</td>
<td>145 ± 59 (−0.8)</td>
<td>72 ± 31 (−2.0)</td>
</tr>
<tr>
<td>FM/DM</td>
<td>6.1 ± 2.1</td>
<td>6.1 ± 1.8 (0)</td>
<td>10.5 ± 4.1 (+1.8)</td>
</tr>
</tbody>
</table>

Convergence of leaf succulence in CAM species

A general anatomical feature of CAM plants and apparent evolutionary co-requisite for CAM is leaf succulence (Fig. 2) with vacuoles occupying 90–95% of the volume of cells with
dimensions of greater than 100 μm (Gibson 1982; Smith 1984). A tight correlation between greater tissue succulence and increased magnitude of CAM has been observed within the Crassulaceae (Teeri et al. 1981; Kluge et al. 1991, 1993), in the Orchidaceae (Winter et al. 1983; Silvera et al. 2005), as well as in many other diverse CAM families (Nelson et al. 2005; Nelson and Sage 2008). Large cell size leads to a tightly packed chlorenchyma with reduced intercellular air spaces (IAS) and reduced surface area exposure of mesophyll cells to IAS ($L_{\text{mes/area}}$), which likely results in low internal conductance of CO$_2$ ($g_t$) and restriction of CO$_2$ efflux (particularly internal CO$_2$ leakage during phase III), thereby enhancing CAM carbon economy. Moreover, in CAM plants, C$_3$ photosynthetic CO$_2$ uptake during phases II and IV is believed to be limited more strongly by low $g_t$ than PEPC mediated nocturnal CO$_2$ uptake during phase I (Maxwell et al. 1997; Maxwell 2002; Nelson and Sage 2008), increasing the reliance on CAM in highly succulent CAM species. 

If the above anatomical traits associated with leaf succulence enhance the degree of CAM photosynthesis and limit the degree of C$_3$ photosynthesis, then all CAM species, regardless of their evolutionary lineage, would be expected to converge to a common succulent leaf anatomy. Supporting evidence for this hypothesis has been obtained by comparing the degree of leaf succulence, indicated by leaf thickness, to leaf δ$^{13}$C values (Winter et al. 1983; Zotz and Ziegler 1997). For example, a survey of leaf thickness and leaf δ$^{13}$C values in 173 tropical orchid species (Table 1) revealed that in species with leaf δ$^{13}$C values commonly observed for C$_3$ plants (~32 to ~22‰), leaf thickness averaged 0.5 ± 0.4 mm, whereas in species with δ$^{13}$C values typical of strong CAM (~21.9 to ~12‰), leaf thickness averaged 2.03 ± 2.01 mm. In weak CAM species with δ$^{13}$C values between ~32 and ~22‰, leaf thickness averaged 0.5 mm ± 0.3 mm (Silvera et al. 2005).

Thus, relative leaf thickness might serve as a useful, surrogate indicator for the presence of CAM activity provided that hydrrenchyma (without chloroplast containing cells) is not the largest contributor of leaf thickness (Fig. 2). Just as increased leaf thickness and concomitant increase in storage capacity for malic acid confer a selective advantage for committing to CAM, if, in CAM plants, increased leaf succulence and associated decreases in $g_t$ indeed affect CO$_2$ uptake in the light more than in the dark (Nelson and Sage 2008), then, evolutionary progression from the C$_3$ to CAM state would appear to favour either retention of C$_3$ photosynthesis or full conversion to CAM, but not the intermediate state. Indeed, such a pattern is reflected in differences in high and low IAS and $L_{\text{mes/area}}$ values between weak and strong CAM species (Nelson and Sage 2008), as well as the bimodal distribution of δ$^{13}$C values observed in large surveys of plant families with mixtures of C$_3$/weak CAM and CAM species (see text below) (Zotz and Ziegler 1997; Pierce et al. 2002; Crayn et al. 2004; Holtum et al. 2004; Silvera et al. 2005, 2009, 2010).

Drivers of CAM evolution
Numerous reports have postulated that C$_3$ photosynthesis is the evolutionary ancestral or progenitor state for CAM, with a progression towards strong CAM taking place in several incremental steps (Teeri 1982a, 1982b; Pilon-Smits et al. 1996; Crayn et al. 2004; Silvera et al. 2009). Reversion from CAM to the C$_3$ state is also possible (Teeri 1982a, 1982b) and evidence for likely reversal events, associated with radiations into less xeric habitats, has come from large-scale isotopic surveys within the Bromeliaceae (Crayn et al. 2004) and the Orchidaceae (Silvera et al. 2009). The reversion of CAM to C$_3$ photosynthesis points to the complex evolutionary histories within these taxa. Although the main driver for CAM evolution remains unclear, several hypotheses have been put forward. Water limitation and the resulting limitation of CO$_2$ brought about by stomatal closure and reductions in atmospheric CO$_2$ concentrations during the late Tertiary (Pearson and Palmer 2000) might have provided the selective pressures for the evolution of CAM over the last 40–100 million years (Monson 1989; Ehleringer and Monson 1993; Raven and Spicer 1996). It is difficult, however, to determine the first origin of CAM in plants, especially because the majority of families in which CAM is present originated recently and fossil evidence of CAM has not been discovered (Raven and Spicer 1996). Indeed, Dendrobiurn and Earina (Epidendroideae) macrofossils of orchid specimens from the early Miocene (23–20 MYA) have been described, however, these were not investigated for the presence of CAM-related characters (Conran et al. 2009). Based on the broad diversity of taxa showing CAM compared with species exhibiting C$_4$ CAM likely evolved first, and because of the presence of CAM in ancient groups such as the isoeitids and cycads, CAM might have appeared as early as the Triassic (Griffiths 1992; Ehleringer and Monson 1993). In any event, CAM likely evolved in response to selection for increased carbon gain and increased water use efficiency (Ehleringer and Monson 1993) after the global reduction in atmospheric CO$_2$ concentration during the Miocene and early Pleistocene or perhaps even earlier during the Oligocene (Edwards et al. 2010). Notably, a large CAM radiation event in the most species-rich epiphytic clade in orchids (Fig. 3), the Epidendroideae, was predicted to have originated ~65 MYA and linked to the decline of atmospheric CO$_2$ during the Tertiary (Silvera et al. 2009). CAM has contributed to the exploitation of wider epiphytic habitat ranges, from low elevation sites where CAM orchids are mostly present, to mid-elevation tropical forest sites of around 1000 m, where moist suitable microenvironments exist for epiphytic orchid colonisation (Silvera et al. 2009).

Taxonomic distribution of CAM
Numerous past studies have estimated the phylogenetic distribution of CAM (Moore 1982; Winter 1985; Lüttge 1987; Griffiths 1988; Ehleringer and Monson 1993). CAM is widespread within the plant kingdom across at least 343 genera in 35 plant families comprising ~6% of flowering plant species (Table 2; JAC Smith, unpubl. data) (Griffiths 1989; Smith and Winter 1996; Holtum et al. 2007). In contrast, C$_4$ photosynthesis occurs in only 19 families and accounts for ~3% of plant species comprising mainly grasses and sedges and some dicots (Sage 2001, 2004). The oldest lineage with CAM described to date is represented by Isoetes, a mostly aquatic or semi-aquatic group distributed in oligotrophic lakes or mesotrophic shallow seasonal pools (Keeley 1998). The retention of CAM in this group is hypothesised to be due to the chronically low or day-time
decline in levels of dissolved CO2 in these aquatic environments (Keeley 1996). CAM has also been documented within the Gnetales, Welwitschia mirabilis Hook. F. (Welwitschiaceae) (von Willert et al. 2005), the Cycadales, Dioon edule Lindl. (Zamiaceae) (Vovides et al. 2002) and in several epiphytic families of ferns within the Polypodiaceae (Holtum and Winter 1999) and the Vittariaceae (Martin et al. 2005).

The widespread taxonomic distribution of both C4 and CAM plants indicates that C4 and CAM plants must have evolved independently multiple times, even within a single genus (Monson 1989, 1999; Ehleringer and Monson 1993; Kellogg 1999; Silvera et al. 2009). Studies with limited taxon sampling by δ13C analysis have been reported for the Crassulaceae (Kalanchoë) (Kluge et al. 1991), Sedum and Aeonium (Pilon-Smits et al. 1996), Clusiaceae (Gehrig et al. 2003; Gustafson et al. 2007), and Orchidaceae (Cymbidium) (Motonura et al. 2008). More extensive combined taxon and isotopic sampling has been completed within the Bromeliaceae (Crayn et al. 2004) and the Orchidaceae (Silvera et al. 2009, 2010).

Estimating the prevalence of CAM

CAM species are widely distributed throughout semiarid tropical and subtropical environments, including epiphytes in the humid tropics that must endure frequent reductions in water availability. Excluding the Orchidaceae, ~9000 species are estimated to perform CAM (Winter and Smith 1996). However, the Orchidaceae alone could contribute an additional 7800 species assuming that of the estimated 26 000 orchid species (Pfahl et al. 2008), 75% are epiphytic, and that ~10% of these engage in strong CAM, and 30% engage in weak CAM (Silvera et al. 2005). Thus, ~16 800 species or close to 6% of an estimated 300 000 vascular plant species (Kreft and Jetz 2007) might engage in CAM to varying degrees. The extent of CAM expression generally correlates with the degree of adaptation to more xeric ecological niches (Kluge et al. 2001; Pierce et al. 2002; Zotz 2004). A recent survey of 1022 orchid species from Panama and Costa Rica using stable isotopic measurements documented that the number of CAM species increases with decreasing precipitation with the majority of CAM species occurring at

<table>
<thead>
<tr>
<th>Family</th>
<th>Order</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>LYCOPODIOPHYTA</td>
<td>Isoetales</td>
<td>Isoetes (quillworts)</td>
</tr>
<tr>
<td>PTERIDOPHYTA</td>
<td>Polyptiales</td>
<td>Pyrosis</td>
</tr>
<tr>
<td>Magnoliophyta</td>
<td>Polypiales</td>
<td>Vittaria</td>
</tr>
<tr>
<td>Agavaceae = Asparagaceae</td>
<td>Asparagales</td>
<td>Agave</td>
</tr>
<tr>
<td>Aizoaceae</td>
<td>Caryophyllales</td>
<td>Mesembryanthemum</td>
</tr>
<tr>
<td>Alismataceae</td>
<td>Alismatales</td>
<td>Sagittaria</td>
</tr>
<tr>
<td>Anacampserotaceae</td>
<td>Caryophyllales</td>
<td>Grahamia</td>
</tr>
<tr>
<td>Araceae</td>
<td>Alismatales</td>
<td>Zamioculcas</td>
</tr>
<tr>
<td>Asphodelaceae = Xanthorrhoeaceae</td>
<td>Asparagales</td>
<td>Aloe, Haworthia</td>
</tr>
<tr>
<td>Apiales</td>
<td>Gentianales</td>
<td>Pachypodium</td>
</tr>
<tr>
<td>Asteraceae</td>
<td>Asterales</td>
<td>Senecio</td>
</tr>
<tr>
<td>Bromeliaceae</td>
<td>Poales</td>
<td>Aechmea, Tillandsia</td>
</tr>
<tr>
<td>Commelinaceae</td>
<td>Commelinales</td>
<td>Callisia</td>
</tr>
<tr>
<td>Cactaceae</td>
<td>Caryophyllales</td>
<td>Opuntia</td>
</tr>
<tr>
<td>Clusiaceae</td>
<td>Malpighiales</td>
<td>Clusia</td>
</tr>
<tr>
<td>Crassulaceae</td>
<td>Saxifragales</td>
<td>Crassula, Kalanchoë</td>
</tr>
<tr>
<td>Cucurbitaceae</td>
<td>Cucurbites</td>
<td>Xerosicyos</td>
</tr>
<tr>
<td>Didiereaceae</td>
<td>Caryophyllales</td>
<td>Didiera</td>
</tr>
<tr>
<td>Euphorbiaceae</td>
<td>Malpighiales</td>
<td>Euphorbia, Monadenium</td>
</tr>
<tr>
<td>Geraniaceae</td>
<td>Geraniales</td>
<td>Sarcoaulon</td>
</tr>
<tr>
<td>Gesneriaceae</td>
<td>Lamiaces</td>
<td>Codonanthe</td>
</tr>
<tr>
<td>Hydrocharitaceae</td>
<td>Alismatales</td>
<td>Vallisneria</td>
</tr>
<tr>
<td>Lamiales</td>
<td>Lamiaces</td>
<td>Plecranthus</td>
</tr>
<tr>
<td>Montiaceae</td>
<td>Caryophyllales</td>
<td>Calandrinia</td>
</tr>
<tr>
<td>Orchidaceae</td>
<td>Asparagales</td>
<td>Oncidium, Phalaenopsis</td>
</tr>
<tr>
<td>Orchidaceae</td>
<td>Oncidales</td>
<td>Oxalis</td>
</tr>
<tr>
<td>Passifloraceae</td>
<td>Malpighiales</td>
<td>Adenia</td>
</tr>
<tr>
<td>Piperaceae</td>
<td>Piperales</td>
<td>Peperomia</td>
</tr>
<tr>
<td>Plantaginaceae</td>
<td>Lamiaces</td>
<td>Litorrella</td>
</tr>
<tr>
<td>Portulacaceae</td>
<td>Caryophyllales</td>
<td>Portulaca, Portulacaria</td>
</tr>
<tr>
<td>Rubiaceae</td>
<td>Gentianales</td>
<td>Myrmecodia</td>
</tr>
<tr>
<td>Ruscaceae = Asparagaceae</td>
<td>Asparagales</td>
<td>Sansevieria, Dracaena</td>
</tr>
<tr>
<td>Vitaceae</td>
<td>Vitales</td>
<td>Cissus</td>
</tr>
<tr>
<td>Welwitschiaceae</td>
<td>Gnetales</td>
<td>Welwitschia</td>
</tr>
<tr>
<td>Zamiaceae</td>
<td>Cycadaceae</td>
<td>Dioon</td>
</tr>
</tbody>
</table>
sites between sea level and 500 m and no CAM species occurring above 2400 m (Silvera et al. 2009). At two lowland forest sites in Panama, 36% and 42% of epiphytic orchid species displayed CAM isotopic values (Zotz and Ziegler 1997; Zotz 2004), and these percentages were 26% and over 60% among tropical epiphytic orchids collected in Papua New Guinea (Earnshaw et al. 1987) and Australia (Winter et al. 1983), respectively. Also, within a single site, the percentage of CAM epiphytes tended to increase from shaded understory sites to exposed canopy sites. For example, in a moist tropical forest in Panama, CAM was more prevalent in emergent layers and exposed tree canopies than in understory sites (Zotz and Ziegler 1997).

A recent study of how closely $\delta^{13}$C values reflect the proportion of CO$_2$ fixed during day and night revealed that ‘the typical CAM plant’ gains ~71–77% of its carbon through nocturnal fixation (Winter and Holtum 2002). However, surveys using only $\delta^{13}$C values to determine the number of CAM-equipped species do not take into account CAM species obtaining less than one-third of their carbon in the dark (Winter and Holtum 2002). Recent surveys that include measurements of nocturnal tissue acidification have identified a greater number of CAM species than surveys using isotopic composition measurements alone (Pierce et al. 2002; Silvera et al. 2005). Furthermore, surveys conducted during the rainy season might not reveal the presence of facultative CAM species that exhibit CAM only under water deficit stress conditions. Thus, estimates of the taxonomic distribution of CAM using only stable isotopes measurements are likely to underestimate the prevalence of CAM.

Integrative studies that attempt to map carbon-isotopic ratio surveys with molecular phylogenies remain limited. Although a well-resolved and comprehensively sampled molecular phylogeny of the Aizoaceae exists (Klak et al. 2003, 2004), the occurrence of CAM has not been mapped onto the available phylogenetic tree. Similarly, an incomplete molecular phylogeny has been established for about two-thirds of species within the Agavaceae (Good-Avila et al. 2006); however, most of those species are expected to perform CAM. A detailed molecular phylogenetic reconstruction of the Vanilloideae with emphasis on the genus Vanilla, which surveyed 47 of the 110 different species, has also been constructed using four plastid genes (Bougard et al. 2010). However, no attempt was made in this study to map the occurrence of CAM. DNA-based molecular phylogenies are well established for the Bromeliaceae (Crayn et al. 2004; Barfuss et al. 2005; Jabaily and Sytsma 2010). Carbon-isotopic ratios collected from 1873 of 2885 bromeliad species revealed that CAM photosynthesis and the epiphytic habit evolved a minimum of three times in this family (Crayn et al. 2004). Molecular phylogenies have also been established for 87 of the estimated 400 species within the Clusiaceae (Gustafson et al. 2007), revealing CAM arose independently within two of the three major groups of Clusia species with multiple reversal events as determined by carbon isotope ratio analysis (Vaasen et al. 2002; Gehrig et al. 2003; Gustafson et al. 2007).

**Orchids as a model for the study of CAM evolution**

The Orchidaceae is the largest family of flowering plants with >800 genera and ~26 000 species worldwide, of which ~75% are estimated to be epiphytic (Atwood 1986; Dressler 1993; Gravendeel et al. 2004). Orchids exhibit a large number of morphological, anatomical, ecological and physiological characteristics that allow them to exist within diverse ecosystems and ecological niches with the greatest diversity in mountainous regions of the tropics (Cribb and Govaerts 2005). One such characteristic is the expression of CAM. Orchids are a large and taxonomically well studied family of CAM plants. For example, from 1994–2004 DNA sequences from 4262 orchids had been deposited in GenBank (Cameron 2005). Chase et al. (2003) proposed an updated classification for the family based upon many recent and ongoing molecular phylogenetic studies. Orchids contain a mixed distribution of C$_3$ and CAM species, a feature that is useful for tracing the occurrence of CAM within discrete lineages (Silvera et al. 2010). In contrast, nearly all species within some other families, such as the Agavaceae, Cactaceae and the Didieraceae, display CAM and, thus, do not permit the evaluation of CAM evolutionary progression. Of families that display a mixture of both C$_3$ and CAM species (e.g. Aizoaceae, Bromeliaceae, Clusiaceae, Crassulaceae), the Orchidaceae has an advanced, well-resolved molecular phylogeny as summarised in the five volumes of Genera Orchidacearum (Pridgeon et al. 1999, 2009).

The diversity expressed by orchids is crucial in linking CAM expression to vegetative morphology such as leaf thickness, to habitat specialisation such as epiphytism, and to adaptive radiation spanning moisture gradients (Dressler 1993; Williams et al. 2001a). Silvera et al. (2005) used a combination of $\delta^{13}$C isotopic ratios and titratable acidity measurements to survey for the presence of CAM in 200 native Panamanian orchid species. The survey produced a bimodal distribution of $\delta^{13}$C values with peaks around –15‰ (signifying strong CAM) and –28‰ (signifying C$_3$ photosynthesis), comparable to other broad surveys employing $\delta^{13}$C value measurements. Within the peak of C$_3$ photosynthesis $\delta^{13}$C values, titratable acidity measurements revealed a second CAM cluster indicative of species with low capacities for nocturnal CO$_2$ fixation (weak CAM). Taking into account both $\delta^{13}$C values and titratable acidity measurements, CAM appears to be widespread among tropical epiphytic orchids. However, fewer than 4% of all known orchid species have been sampled for isotope analysis to date (Silvera et al. 2009, 2010).

**Mapping the occurrence of CAM within the Orchidaceae**

A key prerequisite for the phylogenetic reconstruction of the evolutionary origins of CAM is a sufficiently robust and densely sampled phylogeny for the family based on molecular and morphological characters. The subtribe Oncidiinae is one of the most highly derived clades of orchids of the New World, with great variation in chromosome number, vegetative features and floral characteristics (Chase et al. 2005). Oncidiinae is the second largest orchid subtribe and comprises ~69 genera and ~1600 species, most of which are epiphytic (Williams et al. 2001a, 2001b). This subtribe is one of the most intensively sampled clades within the Orchidaceae. Approximately 600 of 1600 (37%) species have been sampled, employing data from both nuclear and plastid DNA sequences as well as morphological characters (Williams et al. 2001a, 2001b) to provide an excellent
basis from which to study CAM evolution. The monophyly of the Oncidiinae and phylogenetic relationships of related genera have been evaluated by combined data from the internal transcribed spacer of nuclear rDNA (nrrITS) and three plastid regions (matK, trnL intron, and the trnL-F intergenic spacer) producing highly resolved cladograms (Williams et al. 2001a, 2001b). Members of the Oncidiinae occupy a wide variety of epiphytic sites, from large limbs that are exposed in the canopy of tropical forests, to densely shaded sites in the understory (Chase 1988; Chase et al. 2005). Leaf morphology is also highly variable, with species exhibiting a gradient from thick-succulent terete or conduplicate leaves to species showing thin conduplicate leaves. Species within Oncidiinae also show a gradient of CAM expression, from C3 photosynthesis to weak and strong CAM. Ancestral state reconstruction of the occurrence of CAM onto a phylogeny of orchids shows multiple independent origins of CAM with several reversal events (Fig. 3) and a positive cross-genera relationship between epiphytism and photosynthetic pathways, indicating that divergence of photosynthetic pathways has been correlated through evolutionary time (Silvera et al. 2009). CAM is prevalent in low-elevation epiphytes, especially in those from habitats with a strong dry season, and less prevalent in those from cooler habitats with a more even moisture regime that includes both rainfall and fog (Silvera et al. 2009). Ancestral state reconstruction of CAM onto a phylogeny of Oncidiinae species indicates at least eight independent origins of CAM within the clade (Fig. 4).

**Molecular evolution of CAM**

The progression of photosynthetic pathways has been shown consistently to be from C3 ancestors to CAM photosynthesis. However, the genetic changes required for this progression (and reversion) remain unclear. The multiple independent evolutionary origins of CAM and the observation that presumably all of the enzymatic requirements to perform CAM already exist in most plant cells, particularly stomatal guard cells, might suggest that CAM evolution involves relatively few genetic changes. The available molecular data from C4 cycle enzymes support this view in that none of the relatively few genetic changes. The available molecular data guard cells, might suggest that CAM evolution involves CAM already exist in most plant cells, particularly stomatal

**Molecular markers for studying CAM evolution**

The cytosolic enzyme phosphoenolpyruvate carboxylase (EC 4.1.1.31; PEPC) catalyses the β-carboxylation of phosphoenolpyruvate, with oxaloacetate and inorganic phosphate as products, and serves various functions in plants (Chollet et al. 1996; Nimmo 2000). In addition to its anaplerotic roles in leaves and nonphotosynthetic tissues, PEPC catalyses the initial fixation of atmospheric CO2 into C4-dicarboxylic acids in CAM and C4 photosynthesis. For the PEPC gene family, both non-photosynthetic and photosynthetic isoforms are present in C3, C4 and CAM species. These non-photosynthetic, ‘C3 isoforms’ might have served as the starting point for the evolution of the C4 and CAM isogenes. In C4 plants, key determinants for the evolution of the C4 cycle include duplication of ancestral non-photosynthetic or C3 isogenes, followed by the acquisition of increased mRNA and protein expression, with organ- and cell-type-specific expression patterns of the C4 photosynthetic isogenes largely due to transcriptional changes in gene expression (Furumoto et al. 2000; Westhoff and Gowik 2004; Hibberd and Covshoff 2010). As in C4 plants, CAM-specific isofoms of PEPC are distinguished by their elevated mRNA and protein expression in leaf tissues. Evidence from comparative analysis of C3, C3-C4 intermediates, and C4 Flaveria species suggests that C4 photosynthetic PEPC isoforms have evolved from non-photosynthetic or C3 isoforms by gene duplication, and have acquired distinct kinetic and regulatory properties mediated by discrete amino acid changes (Bläsing et al. 2000, 2002; Engelmann et al. 2002, 2003; Westhoff and Gowik 2004).

In CAM plants, an evolutionary progression of gene family changes similar to that described for C4 plants is thought to have occurred in PEPC gene families from C3, weak CAM and strong CAM orchid species (Fig. 5; K Silvera, unpubl. data). This assumption is based on the premise that the molecular mechanisms that drive evolutionary changes in gene family structure are conserved throughout the plant kingdom regardless of the type of photosynthetic pathway. Among CAM species, multiple PEPC isoforms have been described in various species, with CAM-specific PEPC isoforms exhibiting enhanced mRNA expression relative to the expression of C3 PEPC isoforms (Cushman & al. 1989; Gehrig et al. 1995, 2001). For example, comparison of four Clusia species (one C3, two C3-CAM intermediates with water-deficit stress-inducible CAM and one strong, constitutive CAM species), revealed that the ability to conduct nocturnal CO2 fixation was well correlated with PEPC quantity and activity (Borland et al. 1998) and that PEPC mRNA and protein expression was a major factor underpinning the genotypic capacity for CAM (Taybi et al. 2004).

Detailed comparisons of PEPC isoforms from C3 photosynthesis, weak CAM and strong CAM species within the Orchidaceae are underway. As in C4 plants, distinct kinetic and regulatory properties might be expected to be conferred by discrete amino acid changes in CAM-specific isoforms of PEPC, which do show distinct amino acid differences from housekeeping or C3 photosynthesis isoforms (Fig. 5; K Silvera, unpubl. data). In addition, genes of CAM isoforms for PEPC and PEPC kinase appear to have evolved unique expression patterns that are under circadian clock control with expression patterns that are distinct from those in C3 plants (Taybi et al. 2004; Boxall et al. 2005; Cushman et al. 2008b). However, the evolutionary recruitment of gene family members must extend beyond those involved in C4 acid metabolism and include those genes that control the large reciprocating pools of storage carbohydrates, which can account for up to 20% of the total leaf dry weight (Winter and Smith 1996; Dodd et al. 2002, 2003). Some CAM plants can accumulate soluble sugars (e.g. sucrose, glucose, fructose) and polysaccharides (e.g. fructan,
galactomannan) in extra-chloroplastic compartments, while other species store both plastidic starch and cytoplasmic glucose (Christopher and Holtum 1996). Detailed studies have revealed at least eight distinct combinations of malate decarboxylation and carbohydrate storage strategies in CAM plants (Christopher and Holtum 1996). These various
carbohydrate accumulation patterns likely reflect the evolutionary history of the species, rather than the carbon flow constraints of the pathway (Winter and Smith 1996; Winter and Holtum 2002).

Studies of developmental and spatial expression patterns using partial nucleotide sequences of PEPC have provided valuable molecular tools for understanding the evolution of metabolic pathways in which PEPC is involved. PEPC sequences are useful, not only because the gene is ubiquitous in prokaryotes and plants, but also because the marker can provide information about the tissue-specific expression patterns and metabolic roles of specific gene family members (Gehrig et al. 2001). For example, Gehrig et al. (2001) used expression changes during leaf development to infer potential CAM-related isogenes of PEPC relative to non-CAM isoforms expressed predominantly in non-photosynthetic roots. However, tissue-specific expression alone is inadequate to infer a CAM-related function. The relative abundance of each isoform in CAM-performing tissue must be confirmed in order to designate the most abundantly expressed isoform as CAM-specific (Cushman et al. 1989; Gehrig et al. 1995, 2005; Taybi et al. 2004). However, this does not necessarily preclude the possibility that multiple isoforms contribute to CAM-specific function. Interestingly, three leafless orchid species with chloroplast-containing, CAM-performing aerial roots (Winter et al. 1985) also expressed PEPC isoforms that clustered with PEPC isoforms recovered from CAM-performing leaves of other species, but not with PEPC isoforms from nonphotosynthetic aerial roots (Gehrig et al. 2001). This observation suggests that such ‘shootless’ species do not make use of the root-inherent isoform for photosynthetic carbon assimilation, but express either an ancestral leaf, CAM-derived or an additional PEPC isoform that conducts the initial fixation of CO₂ needed for CAM.

Characterisation of PEPC isogenes from Kalanchoë pinnata (Lam.) Pers. revealed seven distinct PEPC isogenes: four in leaves and three in roots. Sequence similarity comparisons and distance neighbour-joining calculations separate the seven PEPC isoforms into two clades, one of which contains the three PEPCs found in roots (Gehrig et al. 2005). The second clade contains the four isoforms found in leaves and is divided into two branches, one of which contains two PEPCs most similar to previously described CAM isoforms. Of these two isoforms, however, only one exhibited abundant expression in CAM-performing leaves, but not in very young leaves, which do not exhibit CAM, suggesting that this isoform encodes a CAM-specific PEPC. Protein sequence comparison and phylogenetic analysis using the neighbour joining method suggest that all isogenes are likely derived from a common ancestor gene, presumably by serial gene duplication events (Sánchez and Cejudo 2003). In addition to plant-type PEPCs, higher plant genomes also encode a bacterial-type, non-phosphorylatable form of PEPC that phylogenetic analysis suggests diverged early during the evolution of plants from a common ancestral PEPC gene probably from γ-proteobacteria (Sánchez and Cejudo 2003).

Comprehensive examination of PEPC gene families together with phylogenetic tree construction from C₃, weak CAM and strong CAM orchids, suggests an emerging model for the evolutionary appearance of paralogous PEPC genes (Fig. 5). Abundantly expressed CAM-PPC genes from CAM species cluster together and belong to a sister group of weak CAM and C₃-most-abundant PPC-genes. Less abundantly expressed isoforms from C₃, weak CAM, and CAM species cluster separately and presumably belong with PPC genes involved in

Fig. 5. Proposed model for the evolutionary progression of paralogous PEPC genes (Ppc) in the Orchidaceae. CAM-Ppc genes from CAM species cluster together (red shading), and belong to a sister group of weak CAM (orange shading) and C₃ (green shading) most abundant Ppc-genes. The least abundant isoforms from C₃, weak CAM and CAM species cluster separately and presumably represent Ppc genes having anaplerotic, as opposed to photosynthetic functions (K. Silvera, unpublished data). The E. coli-like Ppc4 gene from Arabidopsis was used as outgroup.
anaplerotic function (Fig. 5). Although this model requires additional experimental support to confirm these proposed evolutionary events, it should be possible to discern CAM isoforms not only from developmental and tissue-specific expression pattern data, but also by comparing nucleotide or amino acid sequences.

\textit{Circadian clock-regulated markers}

In addition to PEPC, two other genes have been characterised recently that are likely to be excellent markers for tracing CAM evolution within the context of circadian clock biology: PEPC kinase and the glucose-6-phosphate/Pi translocator. In CAM plants, PEPC is activated at night via phosphorylation of a Ser residue near the N-terminus, which renders the enzyme more sensitive to PEP and the positive effectors glucose-6-phosphate (G6P) and triose-phosphate (TP) and less sensitive to the allosteric inhibitor, malate (Chollet et al. 1996; Nimmo 2000). This phosphorylation is carried out by PEPC kinase (PPCK), a dedicated, calcium-independent Ser/Thr protein kinase, the steady-state transcript abundance and activity of which is controlled by the circadian clock (Hartwell et al. 1999; Taybi et al. 2000). However, the circadian regulation of PPCK mRNA abundance can also be regulated by metabolic signals, such as malate accumulation (Borland et al. 1999). Thus, both circadian and metabolic signals appear to modulate PPCK transcript abundance, which in turn regulates PPCK activity, the phosphorylation/activation state of PEPC, and the degree of nocturnal carbon assimilation (Nimmo 2000). In a comparison of four \textit{Clusia} species including a C3 species, two C3-CAM intermediates and a strong, constitutive CAM species, the circadian modulation of PPCK mRNA abundance correlated with the performance of CAM and with day/night changes in malate and soluble sugar content. However, circadian fluctuations in PPCK mRNA abundance were not evident in the C3 species and one of the C3-CAM intermediates (Taybi et al. 2004).

In a more recent study in the C3-CAM plant, \textit{M. crystallinum}, the expression of the glucose-6-phosphate/Pi translocator gene (\textit{Gpt2}) was undetectable in plants performing C3, but was preferentially enhanced in leaves of CAM-induced plants (Kore-edo et al. 2005) and was under circadian clock control (Cushman et al. 2008b). In summary, three major changes appear to have occurred during the evolution of progenitor genes in order to function in CAM (Fig. 2): (1) high expression in plants performing CAM, (2) leaf-specific or leaf-preferential expression patterns and (3) expression patterns coming under circadian clock control. We know very little about the \textit{cis}-regulatory elements and cognate trans-factors involved in controlling the CAM-specific expression of PEPC, PEPC kinase and glucose-6-phosphate/Pi translocator and other CAM-related genes. We also do not know whether there are common regulatory genes responsible for directing the expression of coordinately regulated sets of genes within these regulatory networks.

\textit{Circadian clock specialisation during CAM evolution}

CAM represents an noteworthy example of circadian clock specialisation and is one of the best-characterised physiological rhythms in plants (Wilkins 1992; Lüttge 2003; Wyka et al. 2004). The presence of the CAM enzymatic machinery within a single cell requires strict temporal control of the competing carboxylation reactions by PEPC and RUBISCO. PEPC activity is regulated by reversible protein phosphorylation by PEPC kinase, whose expression is under circadian clock control (Hartwell et al. 1999; Taybi et al. 2000). RUBISCO activity also appears to be modulated, with peak activity apparent during the mid-to-late part of the light period, which would reduce the likelihood that RUBISCO and PEPC would compete for CO\textsubscript{2} during the early morning (Maxwell et al. 1999; Griffiths et al. 2002). Indeed, cytosolic malate (or a related metabolite) concentration appears to exert a negative effect on PEPC kinase gene expression or mRNA stability and override its circadian control in \textit{K. daigremontiana} (Borland and Dodd 1999; Borland and Taybi 2004). This observation has led to the suggestion that circadian control of PEPC kinase expression is a secondary response to malate transport across the tonoplast membrane of the vacuole (Nimmo 2000). However, it should be noted that this metabolite override mechanism has only been reported in this single CAM species. Circadian control of the large, reciprocating pools of carbohydrates (Dodd et al. 2003), changes in mRNA abundance for starch synthesis and degradation enzymes (Dodd et al. 2003; Cushman et al. 2008b), diurnal or circadian expression of genes encoding plastidic triose phosphate/Pi and glucose-6-phosphate/Pi translocators (Häusler et al. 2000; Kore-edo et al. 2005; Cushman et al. 2008b) and the partitioning of isotopically distinct, C\textsubscript{3}- or C\textsubscript{4}-derived classes of carbon pools are also likely to be critical for the optimal performance of CAM (Borland and Dodd 2002; Ceusters et al. 2008).

Comparison of the steady-state mRNA abundance patterns of seven circadian clock components in the facultative CAM plant \textit{M. crystallinum} operating in either C\textsubscript{3} or CAM mode, indicated that its central clock is very similar to that in \textit{Arabidopsis} and is not perturbed by development or salinity stress (Boxall et al. 2005). However, various clock components can be used in different ways to alter clock outputs. Evidence for alternate clock component functions comes from the observation that \textit{ZETTLUPE (ZTL)}, a gene that does not exhibit an oscillating pattern of mRNA expression in \textit{Arabidopsis}, does so in \textit{M. crystallinum}. Furthermore, the circadian abundance profile of \textit{NcZTL} transcripts exhibits a slightly more prolonged period of expression. Additional support for alterations in the circadian clock control outputs in \textit{M. crystallinum} comes from oligonucleotide-based microarray experiments that document that the shift from C\textsubscript{3} to CAM is accompanied by shifts in the phase at which peak expression occurs (Cushman et al. 2008b). A large proportion (70\%) of \textit{Arabidopsis} genes that exhibit circadian fluctuations in transcript abundance also respond to environmental stress (i.e. low temperature, salt, and drought; Kreps et al. 2002). Such rhythmic expression of stress-adaptive genes might prepare the plant to better withstand a stress or exploit a limiting resource (Eriksson and Millar 2003). Given that water deficit stress is likely to be one of the driving forces behind CAM evolution (Raven and Spicer 1996), ancestral C\textsubscript{3} progenitors of CAM plants might have evolved clocks which exerted pervasive control over metabolism as a means of maintaining metabolic homeostasis under stressful environments (Borland and Taybi 2004).
Concluding remarks
The varying degrees to which CAM is expressed reflect a continuum of photosynthetic metabolism from C3 photosynthesis to weakly expressed CAM to fully expressed CAM arising from the unique evolutionary history of a particular species. The plasticity of CAM is governed by the evolutionary disposition of each species, whether under developmental control in a constitutive CAM species or under environmental control in a facultative CAM species. Our understanding of the CAM photosynthetic pathway is advancing, especially at the molecular genetic level. Gene sequence information has proliferated quickly and will provide a solid foundation for future research into CAM evolution. For example, transcriptome sequencing has been performed in a strong CAM orchid species (Riossioglossum ampliatum (Lindl.) M.W.Chase & N.H.Williams) along with the fabrication of a custom oligonucleotide microarray (K Silvera and JC Cushman, unpublished data), which will permit mRNA expression patterns to be compared within closely related C3 photosynthesis and weak CAM species as a way to define large-scale gene expression changes associated with CAM evolution. Future projects aimed at analysing the presence or absence of cis-regulatory elements responsible for circadian clock-controlled expression patterns will determine whether or not CAM-specific expression patterns are regulated by evolutionary changes within 5′ flanking or other control regions. Weak CAM species are of particular interest because a reservoir of duplicated genes that have undergone neo-functionalisation from C3 ancestral genes is expected to be present, in addition to C3 genes. Gene duplication events followed by neo-functionalisation and sub-functionalisation are likely to occur through the differentiation of cis-regulatory elements that control tissue- and clock-specific patterns of expression (Monson 2003). However, the major distinction for CAM evolution is that temporal regulation of gene expression patterns will have greater importance than the cell-specific expression patterns found in C4 plants. In both cases, differentiation within the coding regions can also be expected to produce modified functional domains within proteins.

Undoubtedly, the presence of CAM in evolutionary lineages must be defined at the molecular level, in order to understand the genetic changes responsible for the evolutionary progression from C3 to strong CAM, and possible reversal events linked to a changing environment. Molecular analyses will also provide insights into whether or not weak CAM in certain lineages might have served as a genetic reservoir for adaptive radiations leading to strong CAM. The use of phylogenetic comparative methods will be particularly useful for the testing of correlated evolutionary changes of multiple CAM traits (e.g. molecular, physiological, anatomical and environmental traits). Larger carbon isotope ratio surveys should be attempted and these should be performed in conjunction with titratable acidity measurements from live specimens under defined water status. Future surveys should be conducted under both well watered and water deficit conditions to discover facultative CAM species that might be missed by carbon-isotope ratio surveys alone. High throughput RNA/DNA sequencing strategies should be used to compare gene expression patterns in closely related C3, weak CAM and strong CAM species. Such information will provide important insights into the molecular genetic requirements for CAM evolution within discrete lineages. Such large-scale sequencing strategies can also be applied in a comparative genomics context in order to investigate the convergent evolution of CAM in lineages that evolved CAM independently. The ultimate goal of this approach will be to define the molecular ‘parts list’ required for CAM. Complementary proteomic studies targeting temporal changes in protein abundance or post-translational modifications are also expected to improve our understanding of circadian regulation, especially when coupled with mRNA expression profiling of both coding mRNAs and non-coding micro RNAs (miRNAs) in selected CAM species. Ultimately, integrated approaches that combine molecular genetic strategies, genetic approaches, phylogenetic analysis, ecophysiology, and bioinformatics will aid in our understanding of the molecular evolution of CAM.

Acknowledgements
This work was supported by funding from the USA Environmental Protection Agency under the Greater Research Opportunities Graduate Program (Agreement no. MA 91685201 to KS), National Science Foundation NSF 10B-0543659 (to ICC), and Smithsonian Tropical Research Institute (to KW). We are indebted to Dr J Andrew C Smith (Oxford University) for his contributions to Table 2. We would also like to thank the two anonymous reviewers for their helpful comments and Mary Ann Cushman for her critical reading of the manuscript. EPA has not formally reviewed this publication. The views expressed in this publication are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this publication. This publication was also made possible by NIH Grant Number P20 RR-016464 from the INBRE Program of the National Center for Research Resources through its support of the Nevada Genomics, Proteomics and Bioinformatics Centers.

References


Evolution along the CAM continuum


Manuscript received 15 April 2010, accepted 2 August 2010