**Emerging model systems for functional genomics analysis of Crassulacean acid metabolism**

Author names:

James Hartwell\*, Louisa V. Dever and Susanna F. Boxall

Author e-mail addresses:

James Hartwell: james.hartwell@liverpool.ac.uk

Louisa V. Dever: l.dever@liverpool.ac.uk

Susanna F. Boxall: boxall@liverpool.ac.uk

Affiliations/ address:

Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK.

\*Corresponding author:

James Hartwell: \*james.hartwell@liverpool.ac.uk

Abstract (100 – 120 words):

Crassulacean acid metabolism (CAM) is one of three main pathways of photosynthetic carbon dioxide fixation found in higher plants. It stands out for its ability to underpin dramatic improvements in plant water use efficiency, which in turn has led to a recent renaissance in CAM research. The current ease with which candidate CAM-associated genes and proteins can be identified through high-throughput sequencing has opened up a new horizon for the development of diverse model CAM species that are amenable to genetic manipulations. The adoption of these model CAM species is underpinning rapid advances in our understanding of the complete gene set for CAM. We highlight recent breakthroughs in the functional characterization of CAM genes that have been achieved through transgenic approaches.

Highlights:

* Efforts are now underway to decipher the genetic blueprint for CAM.
* Several CAM genomes and transcriptomes have recently been published.
* *Kalanchoë* has been established as readily transformable model genus for CAM.
* *Kalanchoë* transgenics demonstrate the importance of key metabolic steps in CAM.
* A diverse range of CAM species are being developed as amenable models.

**Introduction**

Although the core biochemical pathway of Crassulacean acid metabolism (CAM) was established many decades ago [1], only a very small number of the biochemical steps of CAM have been established as essential through the use of molecular-genetic approaches. Precise details of the co-factor dependency and sub-cellular compartmentalisation of several key CAM enzymes remain largely the subject of unconfirmed speculation. As such, a metabolic pathway that might appear to be completely understood when viewed as a diagram in an undergraduate plant biochemistry textbook is in reality only understood in small fragments with many remaining questions [2]. Furthermore, CAM is well-characterised as an output of the plant circadian clock, but little is known about the underlying signal transduction pathways that allow the central circadian oscillator to coordinate and optimise the daily cycle of CAM CO2 fixation and associated metabolic fluxes relative to the 24 h light/ dark cycle [3,4]. These and many other remaining fundamental questions about CAM are good examples of the types of research challenges that will benefit greatly from the application of functional genomics approaches in ‘model’ CAM species \*\*[5].

The widespread adoption of *Arabidopsis thaliana* as a model flowering plant, and the subsequent delivery of the first complete genome sequence of an Angiosperm for this species in 2000 [6], has driven breakthrough after breakthrough in our understanding of the ways in which genes and their encoded proteins function to deliver phenotypes in higher plants \*\*[7]. However, Arabidopsis is not a panacea for plant biology research, as it doesn’t encompass all aspects of the evolved spectrum of biological functions known from the plant kingdom. Good examples of plant adaptations that cannot be investigated directly in Arabidopsis include the photosynthetic adaptations CAM and C4. *A. thaliana* performs the ancestral C3 form of photosynthetic CO2 fixation. Photosynthesis in Arabidopsis and other C3 species such as rice and wheat becomes inefficient at elevated temperatures as the oxygenase activity of ribulose 1,5-bisphosphate carboxylase oxygenase (Rubisco) is favoured, and thus photorespiration decreases overall photosynthetic efficiency.

CAM and C4 species have evolved CO2 concentrating mechanisms (CCMs) that limit the oxygenase activity of Rubisco by concentrating CO2 around the enzyme. Both adaptations also increase plant water use efficiency (WUE). CAM stands out as the WUE champion because stomatal opening and primary atmospheric CO2 fixation occurs at night via phosphoenolpyruvate carboxylase (PPC) and leads to the vacuolar storage of malate during the cooler more humid dark period, and secondary refixation of CO2 from malate decarboxylation via Rubisco occurs behind closed stomata during the hot, dry light period \*\*[8]. CAM plants thus minimise the transpirational loss of water through their stomatal pores by performing most of their primary atmospheric CO2 fixation in the dark [4].

The increased WUE of CAM plants has led to many researchers recognizing the huge potential and value of this adaptation as a means to facilitate the productive growth of food crops and biomass feedstocks in seasonally dry regions of the world [9,10]. CAM plants are seen as a key option for adapting 21st Century agriculture to the predicted changes in the Earth’s climate, especially in terms of producing biomass for bioenergy using land that is unsuitable for major food crop species [11-14](\*12). Achieving a detailed understanding of how CAM is established and functions efficiently in a diverse range of species that have evolved the adaptation independently will facilitate not only the targeted improvement of existing CAM crop species such as agaves, opuntias, pineapple and vanilla, but will also underpin the engineering of CAM and its improved WUE into C3 crop species using the latest plant genome engineering approaches [5,\*\*8,15].

Despite this huge potential of CAM for climate resilient agriculture, the molecular and biochemical basis for CAM has remained relatively under-studied for many decades. However, recent advances in high-throughput DNA sequencing and the latest accelerated functional genomics approaches are now beginning to deliver significant advances in our understanding of the genetic blueprint for CAM. For example, the first two draft genome sequences for monocot CAM species, namely the moth orchid (*Phalaenopsis equestris*), and the sub-tropical fruit crop, pineapple (*Ananas comosus*), have been published recently \*\*[16-18]. Furthermore, high-throughput transcriptome sequencing (RNA-seq) studies have also been published for the monocot, obligate CAM species *Agave tequilana* and *A. deserti \*\**[19], and for the dicot, inducible weak CAM species *Talinum triangulare* \*\*[20]. More limited expressed sequence tag and RNA-seq transcriptome datasets have also been published for the dicot cacti *Opuntia ficus-indica*, *Lophophora williamsii* and *Hylocereus polyrhizus,* although only limited or no analysis and interpretation was provided concerning the CAM genes in these transcriptome datasets[21-23]. A diverse selection of CAM species were also included in the 1000 plants (1KP) transcriptome sequencing project [24].

Further rapid advances in our understanding of the CAM genetic blueprint will be most readily achieved through the development and widespread adoption of appropriate model species that are underpinned by high quality genome sequences and associated CAM-focused RNA-seq transcriptome datasets. Previous studies have championed a range of potential model species for C4 research, including both monocots and dicots, and many C3/ C4 comparative transcriptome sequencing studies have now been published [25-29]. However, model systems for CAM research have only received the briefest of mentions in recent reviews [4,5]. Here we explore the current best model species for CAM research and describe progress with testing the function of candidate CAM genes through molecular-genetic and transgenic approaches.

**The genus *Kalanchoë*: an amenable transgenic system for CAM research**

Members of the eudicot genus *Kalanchoë* (family Crassulaceae, order Saxifragales) have long been viewed as important model species for the study of CAM [3,30-32]. *Kalanchoë* is a relatively large genus of approximately 125 succulent species distributed across Africa, Madagascar and Asia [33]. The greatest species diversity occurs in Madagascar, where approximately 60 species have been recognised [34]. As with the wider family Crassulaceae, *Kalanchoë* are all believed to be capable of some degree of CAM, although several species have been reported to have C3-type 13C ratios when sampled from their native habitats in Madagascar [30]. *Kalanchoës* are mainly perennial, but some species are considered biennial or annual [35]. Their growth habit ranges from shrublets to small trees, and whilst most species are terrestrial, there are several species that grow as epiphytes and/ or lithophytes, and others are lianas [35]. In Madagascar, different *Kalanchoë* species occupy different ecophysiological niches ranging from the arid and semi-arid regions in the South West of the island through to the seasonally dry forests and inselbergs and humid tropical rainforests found in the Central and Northern parts of the island [30].

*Kalanchoë* species have been arranged into taxonomic groupings using molecular phylogenetic approaches and these groups correlate strongly with their physiotypes (Figs. 1 and 2) [36]. The C3-physiotype belong to the basal Kitchingia section of the genus that is endemic to Madagascar. They are mostly thin-leaved epiphytes found in humid mountain forests that experience high annual rainfall. The Kitchingia section species that have been studied in sufficient detail, *K. porphyrocalyx* and *K. miniata,* are however capable of induction into CAM in response to drought stress [37]. The Bryophyllum section of the genus encompasses the flexible or facultative CAM-C3 physiotype, comprising of species that prefer dry habitats with pronounced wet and dry seasons. The *Kalanchoë* species that have historically seen most widespread use for CAM research belong to this section of the genus (e.g. *K. fedtschenkoi* and *K. daigremontiana*). All members of the Bryophyllum section are also capable of clonal reproduction via adventitious plantlets that form in notches on their leaf margins [32,38]. Finally, the Eukalanchoe section of the genus contains strong CAM species that are the most derived in evolutionary terms [39]. This section is comprised of thick-leaved succulent species that inhabit the arid South Western region of Madagascar, and arid parts of Eastern Africa.

*Kalanchoë* species have underpinned a number of key breakthroughs in the understanding of CAM and these have been covered in detail elsewhere [3,4,40]. *K. fedtschenkoi* and *K. daigremontiana* have in particular led the way in terms of understanding circadian rhythms of CAM-associated CO2 fixation [41,42]. Indeed, the first clock-controlled CAM gene, *PPC kinase* (*PPCK*), was cloned and characterised from *K. fedtschenkoi* [43]. The existing wealth of background, CAM-focused literature on *Kalanchoë,* combined with their ease of growth and amenability to stable transformation, has led to the adoption of several species as models not only for CAM research, but also for the developmental embryogenesis and organogenesis processes associated with the growth of adventitious plantlets on the margin of leaves within section Bryophyllum [32,38,44,45].

Draft genome sequences are being decoded, assembled and annotated for *K. fedtschenkoi* and *K. laxiflora* (Fig. 2). The *K. laxiflora* genome is now available in pre-publication form via the Joint Genome Institute’s Phytozome website: https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org\_Kmarnieriana, and *K. fedtschenkoi* and *K. daigremontiana* have been used successfully for transgenic experiments [5,45,46]. Extensive transcriptome sequence datasets are also available for *K. fedtschenkoi* and *K. laxiflora* [4,5]. Leaves of these species display a clear developmental progression from C3 to CAM, providing a powerful system for identifying genes, proteins and metabolites that are associated with increased CAM in older leaves [43,46](\*\*46). *K. laxiflora* sets thousands of dust-like seeds per plant and is thus well-suited to genetic crossing and mutagenesis approaches. However, *Kalanchoës* are relatively large plants compared to Arabidopsis and they can take up to a year to go from seed-to-seed. Small primary transformants can be generated with high efficiency through *Agrobacterium tumefaciens*-mediated transformation and tissue culture regeneration in around 3 to 4 months, and raising large clonal populations of transgenic lines of a suitable size for detailed phenotypic analysis can take a further 4 to 6 months \*\*[46]. Whilst these timescales cannot compete with turnaround times familiar to Arabidopsis researchers, they are comparable to many other model plant species, including transformable C4 model species such as *Setaria* and *Flaveria*.

Recently, the targeted silencing of the CAM decarboxylation pathway enzymes NAD-malic enzyme (NAD-ME) and pyruvate orthophosphate dikinase (PPDK) using a hairpin RNA transgene RNAi approach in *K. fedtschenkoi* revealed that these enzymes were essential for CAM \*\*[46]. Importantly, the NAD-ME RNAi lines revealed that *K. fedtschenkoi* performs the majority of its malate decarboxylation via NAD-ME in the mitochondrion, suggesting that NADP-ME in the cytosol or chloroplast plays at most a minor role in the light-period turnover of malate and liberation of internal CO2 in *Kalanchoë* leaf mesophyll cells (Fig. 3). This evidence that NAD-ME is the major malic enzyme for malate decarboxylation in the light invokes a requirement for malate import into the mitochondrion, most probably via the dicarboxylate carrier (DiC), and pyruvate export from the mitochondrion via an as yet unidentified mitochondrial pyruvate export protein (Fig. 3). A mitochondrial pyruvate carrier (MPC) from the mitochondrial inner envelope membrane has recently been characterised from yeast, Drosophila and human mitochondria, but this protein has only been reported to import pyruvate into mitochondria [47,48] (\*47, \*\*48).

A further fascinating insight from silencing NAD-ME in *K. fedtschenkoi* was that the loss of NAD-ME led to arrhythmia of both CO2 fixation and gene transcript oscillations for the central circadian clock gene, *TIMING OF CHLOROPHYLL A/B-BINDING PROTEIN1* (*KfTOC1*) \*\*[46]. This latter result suggested for the first time that metabolic perturbations associated with a lack of CAM, for example a dampened daily oscillation in malate levels, could feed back to influence rhythmicity within the core circadian oscillator.

This first report of the targeted down-regulation of CAM-associated genes using transgenic approaches sets the stage for the systematic transgenic perturbation of CAM gene candidates in order to identify the complete minimal gene set for efficient CAM [5]. This work is now ongoing and beginning to bear fruit, highlighting the importance of developing amenable transgenic model systems for CAM research.

**Box1**

**Alternative CAM model species**

**Iceplant**

The Common or Crystalline Ice Plant (*Mesembryanthemum crystallinum L.*; Fig. 4A) has been employed as a key model for the study of facultative or inducible CAM ever since the original report of salt stress-induced CAM in this species over four decades ago [49]. Again, key breakthroughs in CAM research have been achieved using this species, and these have also been covered in detail elsewhere [31]. *M. crystallinum* belongs to the Aizoaceae, which in turn resides within the order Caryophyllales and thus represents a distinct evolutionary origin of CAM. Caryophyllales includes other key families that have evolved CAM, including the iconic Cactaceae plus specific members of the Portulacaceae that are the only species known to be capable of switching from C4 to CAM.

Iceplant was the first CAM species for which a substantial transcriptome sequence database was established in the form of traditional expressed sequence tags (ESTs) [50,51], and further transcriptome data has been generated more recently using Illumina sequencing \*[52]. Iceplant remains the only CAM species for which a mutant population has been generated, using fast neutrons, and screened [53]. This led to the identification of a starch-less phosphoglucomutase mutant as the first CAM mutant to be isolated [54]. A microarray study of the light/ dark and circadian regulation of thousands of Iceplant genes in leaves in the C3 and CAM state has revealed widespread circadian clock control of CAM-associated transcripts [51].

Iceplant continues to be an attractive and valuable model species for CAM research, but is currently held back due to the lack of a simple and efficient stable transformation system, which precludes the production of targeted transgenic gene knockout and/ or over-expression lines, and/ or confirmation of gene mutations through genetic complementation. Iceplant does however have a relatively small, diploid genome, a rapid, annual life cycle, and produces thousands of small seed that make it amenable to further mutagenesis studies. In particular, Iceplant could be an excellent species in which to develop a TILLING population [55], especially once a high quality genome sequence is available [5].

**Other emerging models**

Further, alternative facultative CAM model species are currently under development, including *Talinum triangulare* (Fig. 4B), *Calandrinia polyandra* (Fig. 4C), and *Portulaca oleracea* (Fig. 4D) in the Caryophyllales, and *Clusia pratensis* (Fig. 4E) in the Malpighiales, which all display reversible induction of CAM in response to drought stress \*\*[56]. *P. oleracea* is particularly noteworthy as it is a C4 species that has been found to induce weak CAM in response to drought stress \*\*[57].

In addition, assembled and annotated genome and transcriptome resources are available or under development for various CAM crop species including pineapple (Fig. 4F), *Agave tequilana* (Fig. 4G), *Agave sisalana* (Fig. 4H) and *Opuntia ficus-indica* (Fig. 4I) \*\*[5,16,19]. Whilst these species are not as readily manipulated in the lab environment as the other mentioned CAM models, their high-productivity CAM and proven use in agriculture will provide important insights.

**Conclusions and perspectives**

The pressing need for diverse CAM molecular-genetic model species is exemplified by the insights into CAM that have been made possible by leveraging various *Kalanchoë* speciesand *M. crystallinum* as model systems. Developing models for CAM research is a more complex challenge than that which faced the wider plant biology community at the dawn of the Arabidopsis era. Back then having a single model plant that would allow a broad spectrum of fundamental plant biology questions to be addressed through molecular-genetic approaches was a massive advantage. CAM has however evolved independently many times, and thus it would be very blinkered to focus all effort on a single CAM model species in the longer term. Other candidate CAM model species have been proposed (Box 1., Fig. 4) \*\*[56], and some of these have genome and transcriptome resources available that provide preliminary insights into their CAM genetic blueprint [20,57] (\*\*57). The development of simple and efficient stable transformation systems for these species, and the subsequent targeted transgenic manipulation of candidate CAM genes identified from the transcriptome studies, will contribute greatly to understanding the convergent, independent evolution of CAM in multiple lineages.

However, there is, for the time being, a strong argument in favour of focusing global research efforts on *Kalanchoë* species as models for CAM research. The key criterion underpinning this proposal relates to the ease with which all *Kalanchoë* species that have been tested to date can be transformed stably using *A. tumefaciens* [46,58-60] (\*\*46). Their small genome size (250 – 500 Mbp) and dust-like seed also aid their use as molecular-genetic models. Many species are also diploid [61], simplifying genetic approaches. This coupled with the broad spectrum of CAM physiotypes found within the genus*,* from C3/ inducible weak CAM to strong, constitutive CAM (Fig. 1), sets them apart as a model genus that will allow many fundamental aspects of CAM evolution, development, induction, 24 h light/ dark and circadian clock regulation, and metabolic flexibility to be understood.

*Kalanchoës* should prove to be a particularly fruitful study system in terms of comparative genomic analyses performed using representative C3 physiotype (e.g. *K. gracilipes*), CAM-C3 physiotype (e.g. *K. fedtschenkoi*, *K. laxiflora*), and strong CAM physiotype (e.g. *K. millotii, K. beharensis, K. rhombopilosa*) species (Figs. 1 and 2). If the same ancestral copies of CAM-associated genes have been recruited to CAM-associated functions across the entire genus, then comparative analysis of these genes between C3 (Kitchingia), CAM-C3 (Bryophyllum) and strong CAM (Eukalanchoe) species, leveraging both genome and epigenome sequence analysis techniques, will help to reveal the sequence-level changes that underpinned the evolution of strong CAM in section Eukalanchoe. This in turn should provide valuable insights into the genetic changes that will be required to re-engineer a C3 species to perform CAM \*\*[8].

**Acknowledgements**

We thank Prof. Luciano Freschi, University of São Paulo, Brazil for very kindly providing the photographs of *Mesembryanthemum crystallinum*, *Talinum triangulare*, *Calandrinia polyandra*, *Clusia pratensis* and *Agave sisalana* shown in Figure 4. We would like to thank the many members of the CAM and *Kalanchoë* research communities who have contributed through discussions and collaborations to the ideas expressed in this article, and apologise to those whose work we were unable to cite due to space limitations. The work described in this article that was performed in the Hartwell lab was funded in part by the Biotechnology and Biological Sciences Research Council, UK (grant no. BB/F009313/1 awarded to J.H.), in part by the Royal Society Newton Advanced Fellowship (grant no. NA140007), and in part by the US Department of Energy (DOE) Office of Science, Genomic Science Program under Award Number DE-SC0008834. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the DOE.

**References**

1. Osmond CB: **Crassulacean acid metabolism: a curiosity in context**. *Annu Rev Plant Physiol Plant Mol Biol* 1978, **29**:379-414.

2. Ting IP: **Crassulacean acid metabolism**. *Annu Rev Plant Physiol Plant Mol Biol* 1985, **36**:595-622.

3. Hartwell J: **The circadian clock in CAM plants**. In *Endogenous Plant Rhythms*. Edited by Hall AJW, McWatters HG: Blackwell Publishing Ltd; 2006:211 - 236. [Roberts J (Series Editor): Annual Plant Reviews, vol 21.]

4. Borland AM, Griffiths H, Hartwell J, Smith JAC: **Exploiting the potential of plants with crassulacean acid metabolism for bioenergy production on marginal lands**. *J Exp Bot* 2009, **60**:2879-2896.

\*\*5. Yang XH, Cushman JC, Borland AM, Edwards EJ, Wullschleger SD, Tuskan GA, Owen NA, Griffiths H, Smith JAC, De Paoli HC, et al.: **A roadmap for research on crassulacean acid metabolism (CAM) to enhance sustainable food and bioenergy production in a hotter, drier world**. *New Phytol* 2015, **207**:491-504.

The authors describe a research roadmap for the near future of CAM research, including coverage of aspects of CAM evolution and ecology, ecophysiology and physiology, and genomics and functional genomics. In particular, a brief overview is provided of current and ongoing progress with the development of genomic and transcriptomic resources for potential CAM model species.

6. Kaul S, Koo HL, Jenkins J, Rizzo M, Rooney T, Tallon LJ, Feldblyum T, Nierman W, Benito MI, Lin XY, et al.: **Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana***. *Nature* 2000, **408**:796-815.

\*\*7. Provart NJ, Alonso J, Assmann SM, Bergmann D, Brady SM, Brkljacic J, Browse J, Chapple C, Colot V, Cutler S, et al.: **50 years of Arabidopsis research: highlights and future directions**. *New Phytol* 2016, **209**:921 - 924.

This article summarises some key highlights from 50 years of research using *Arabidopsis thaliana* as a model species. It exemplifies the power of using model species to accelerate progress in research across a broad spectrum of modern plant biology sub-disciplines.

\*\*8. Borland AM, Hartwell J, Weston DJ, Schlauch KA, Tschaplinski TJ, Tuskan GA, Yang XH, Cushman JC: **Engineering crassulacean acid metabolism to improve water-use efficiency**. *Trends Plant Sci* 2014, **19**:327-338.

 The value of understanding CAM in sufficient detail to permit the forward engineering of CAM into existing C3 crop species is discussed in detail. Initial ideas are presented for 'modules' of genes for the CAM carboxylation and decarboxylation pathways that may allow engineering of CAM into C3 species.

9. Borland AM, Wullschleger SD, Weston DJ, Hartwell J, Tuskan GA, Yang XH, Cushman JC: **Climate-resilient agroforestry: physiological responses to climate change and engineering of crassulacean acid metabolism (CAM) as a mitigation strategy**. *Plant Cell Environ* 2015, **38**:1833-1849.

10. Mason PM, Glover K, Smith JAC, Willis KJ, Woods J, Thompson IP: **The potential of CAM crops as a globally significant bioenergy resource: moving from 'fuel or food' to 'fuel and more food'**. *Energ Environ Sci* 2015, **8**:2320-2329.

11. Owen NA, Griffiths H: **Marginal land bioethanol yield potential of four crassulacean acid metabolism candidates (*Agave fourcroydes*, *Agave salmiana*, *Agave tequilana* and *Opuntia ficus-indica*) in Australia**. *GCB Bioenergy* 2014, **6**:687-703.

\*12. Corbin KR, Byrt CS, Bauer S, DeBolt S, Chambers D, Holtum JAM, Karem G, Henderson M, Lahnstein J, Beahan CT, et al.: **Prospecting for Energy-Rich Renewable Raw Materials: Agave Leaf Case Study**. *PLOS One* 2015, **10**:23.

Results are presented from one of the first field trials of *Agave tequilana* grown in Australia specifically as a source of renewable biomass for energy and/ or industrial chemicals.

13. Cushman JC, Davis SC, Yang XH, Borland AM: **Development and use of bioenergy feedstocks for semi-arid and arid lands**. *J Exp Bot* 2015, **66**:4177-4193.

14. Davis SC, Ming R, LeBauer DS, Long SP: **Toward systems-level analysis of agricultural production from crassulacean acid metabolism (CAM): scaling from cell to commercial production**. *New Phytol* 2015, **208**:66-72.

15. DePaoli HC, Borland AM, Tuskan GA, Cushman JC, Yang XH: **Synthetic biology as it relates to CAM photosynthesis: challenges and opportunities**. *J Exp Bot* 2014, **65**:3381-3393.

\*\*16. Ming R, VanBuren R, Wai CM, Tang HB, Schatz MC, Bowers JE, Lyons E, Wang ML, Chen J, Biggers E, et al.: **The pineapple genome and the evolution of CAM photosynthesis**. *Nat Genet* 2015, **47**:1435-1442.

 The second published draft whole genome sequence for a monocot CAM species. RNA-seq data for the tip and base of leaves of field grown pineapple sampled over a light/ dark cycle was used to support the identification of putative CAM associated genes.

\*\*17. Cai J, Liu X, Vanneste K, Proost S, Tsai WC, Liu KW, Chen LJ, He Y, Xu Q, Bian C, et al.: **The genome sequence of the orchid *Phalaenopsis equestris***. *Nat Genet* 2015, **47**:65-72.

 The first published draft genome sequence for a CAM species. Preliminary analysis was presented highlighting the evolution of CAM-associated gene families, although no experimental evidence was presented to aid the identification of the members of the CAM-associated gene families that have been recruited to function in CAM in this orchid.

\*\*18. Albert VA, Carretero-Paulet L: **A genome to unveil the mysteries of orchids**. *Nat Genet* 2015, **47**:3-4.

 An overview of the significance of the publication of the draft *Phalaenopsis equestris* genome as the first genome for a member of the Orchidaceae. Of particular interest here due to mentioning the CAM orchid *Erycina pusilla* as a potential model for the study of CAM evolution in orchids due to its relatively short life-cycle and the availability of a stable genetic transformation procedure.

\*\*19. Gross SM, Martin JA, Simpson J, Abraham-Juarez MJ, Wang Z, Visel A: **De novo transcriptome assembly of drought tolerant CAM plants, *Agave deserti* and *Agave tequilana***. *BMC Genomics* 2013, **14**:14.

 The first published *de novo* sequenced and assembled transcriptome for a monocot CAM species. High quality assemblies of the transcriptomes were presented and annotated. Quantitative RNA-seq analysis was performed on the C3 base and CAM-performing tip of *A. deserti* leaves, supporting the identification of CAM-associated genes.

\*\*20. Brilhaus D, Brautigam A, Mettler-Altmann T, Winter K, Weber APM: **Reversible Burst of Transcriptional Changes during Induction of Crassulacean Acid Metabolism in *Talinum triangulare***. *Plant Physiol* 2016, **170**:102-122.

 Illumina RNA-seq data was used to study gene transcript abundance changes associated with drought-induced weak CAM in *T. triangulare* leaves sampled in both the light and the dark. Complementary metabolite analysis was also presented. This article is particularly noteworthy for providing the most comprehensive analysis of CAM-associated genes out of all of the currently available published genome and transcriptome studies on CAM species.

21. Ibarra-Laclette E, Zamudio-Hernandez F, Perez-Torres CA, Albert VA, Ramirez-Chavez E, Molina-Torres J, Fernandez-Cortes A, Calderon-Vazquez C, Olivares-Romero JL, Herrera-Estrella A, et al.: ***De novo* sequencing and analysis of *Lophophora williamsii* transcriptome, and searching for putative genes involved in mescaline biosynthesis**. *BMC Genomics* 2015, **16**:14.

22. Mallona I, Egea-Cortines M, Weiss J: **Conserved and Divergent Rhythms of Crassulacean Acid Metabolism-Related and Core Clock Gene Expression in the Cactus *Opuntia ficus-indica***. *Plant Physiol* 2011, **156**:1978-1989.

23. Qingzhu H, Chengjie C, Pengkun C, Yuewen M, Jingyu W, Jian Z, Guibing H, Jietang Z, Yonghua Q: **Transcriptomic analysis reveals key genes related to betalain biosynthesis in pulp coloration of *Hylocereus polyrhizus***. *Front Plant Sci* 2016, **6**.

24. Matasci N, Hung LH, Yan ZX, Carpenter EJ, Wickett NJ, Mirarab S, Nguyen N, Warnow T, Ayyampalayam S, Barker M, et al.: **Data access for the 1,000 Plants (1KP) project**. *GigaScience* 2014, **3**:10.

25. Brutnell TP: **Model grasses hold key to crop improvement**. *Nat Plants* 2015, **1**: 15026.

26. Brutnell TP, Bennetzen JL, Vogel JP: ***Brachypodium distachyon* and *Setaria viridis*: Model Genetic Systems for the Grasses**. *Annu Rev Plant Biol* 2015, **66**:465-485.

27. Brown NJ, Parsley K, Hibberd JM: **The future of C4 research: Maize, *Flaveria* or *Cleome*?** *Trends Plant Sci* 2005, **10**:215-221.

28. Kulahoglu C, Denton AK, Sommer M, Mass J, Schliesky S, Wrobel TJ, Berckmans B, Gongora-Castillo E, Buell CR, Simon R, et al.: **Comparative Transcriptome Atlases Reveal Altered Gene Expression Modules between Two Cleomaceae C3 and C4 Plant Species**. *Plant Cell* 2014, **26**:3243-3260.

29. Gowik U, Brautigam A, Weber KL, Weber APM, Westhoff P: **Evolution of C4 Photosynthesis in the Genus *Flaveria*: How Many and Which Genes Does It Take to Make C4?** *Plant Cell* 2011, **23**:2087-2105.

30. Kluge M, Brulfert J, Ravelomanana D, Lipp J, Ziegler H: **Crassulacean acid metabolism in *Kalanchoë* species collected in various climatic zones of Madagascar: a survey by d13C analysis**. *Oecologia* 1991, **88**:407-414.

31. Cushman JC, Bohnert HJ: **Crassulacean acid metabolism: Molecular genetics**. *Annu Rev Plant Physiol Plant Mol Biol* 1999, **50**:305-332.

32. Garcês H, Sinha N: **The ‘Mother of Thousands’ (*Kalanchoë daigremontiana*): A Plant Model for Asexual Reproduction and CAM Studies**. *Cold Spring Harb Protoc* 2009, **2009**:pdb.emo133.

33. Boiteau P, Allorgue-Boiteau L: *Kalanchoë (Crassulacées) de Madagascar: systématique, écophysiologie et phytochimie*. Editions Karthala, Paris; 1995.

34. Allorge-Boiteau L: **Madagascar as the center of speciation and origin of the genus *Kalanchoë* (Crassulaceae)**. In *Biogeography of Madagascar*. Edited by Lourenço WR: Editions de l'ORSTOM, Paris; 1996:137-145.

35. Descoings B: ***Kalanchoë***. In *Illustrated Handbook of Succulent Plants: Crassulaceae*. Edited by Eggli U: Springer, Berlin Heidelberg; 2003:143-181.

36. Gehrig H, Gaussmann O, Marx H, Schwarzott D, Kluge M: **Molecular phylogeny of the genus *Kalanchoë* (Crassulaceae) inferred from nucleotide sequences of the ITS-1 and ITS-2 regions**. *Plant Sci.* 2001, **160**:827-835.

37. Brulfert J, Ravelomanana D, Guclu S, Kluge M: **Ecophysiological studies in *Kalanchoë porphyrocalyx* (Baker) and *K. miniata* (Hils et Bojer), two species performing highly flexible CAM**. *Photosynthesis Res* 1996, **49**:29-36.

38. Garces HMP, Champagne CEM, Townsley BT, Park S, Malho R, Pedroso MC, Harada JJ, Sinha NR: **Evolution of asexual reproduction in leaves of the genus *Kalanchoë***. *P Natl Acad Sci USA* 2007, **104**:15578-15583.

39. Kluge M, Brulfert J, Lipp J, Ravelomanana D, Ziegler H: **A comparative study by d13C analysis of crassulacean acid metabolism (CAM) in *Kalanchoë* (Crassulaceae) species of Africa and Madagascar**. *Bot Acta* 1993, **106**:320-324.

40. Hartwell J, Nimmo GA, Wilkins MB, Jenkins GI, Nimmo HG: **Probing the circadian control of phosphoenolpyruvate carboxylase kinase expression in *Kalanchoë fedtschenkoi***. *Funct Plant Biol* 2002, **29**:663-668.

41. Wilkins MB: **Circadian rhythms: their origin and control.** *New Phytol* 1992, **121**:347-375.

42. Carter PJ, Nimmo HG, Fewson CA, Wilkins MB: **Circadian Rhythms in the Activity of a Plant Protein Kinase**. *EMBO J* 1991, **10**:2063-2068.

43. Hartwell J, Gill A, Nimmo GA, Wilkins MB, Jenkins GI, Nimmo HG: **Phosphoenolpyruvate carboxylase kinase is a novel protein kinase regulated at the level of expression**. *Plant J* 1999, **20**:333-342.

44. Kulka RG: **Cytokinins inhibit epiphyllous plantlet development on leaves of *Bryophyllum* (*Kalanchoë*) *marnierianum***. *J Exp Bot* 2006, **57**:4089-4098.

45. Garces HMP, Koenig D, Townsley BT, Kim M, Sinha NR: **Truncation of LEAFY COTYLEDON1 Protein Is Required for Asexual Reproduction in *Kalanchoë daigremontiana***. *Plant Physiol* 2014, **165**:196-206.

\*\*46. Dever LV, Boxall SF, 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* 2015, **167**:44-59.

 The first report of targeted transgenic disruption of individual steps in the CAM pathway using the model CAM species *K. fedtschenkoi*. Reduction of either NAD-ME or PPDK to <10 % of wild type activity resulted in a loss of malate turnover in the light, and an associated failure to undertake primary atmospheric CO2 fixation in the dark,. Mitochondrial NAD-ME was demonstrated to be the major route for malate decarboxylation in the light period. Loss of CAM in the NAD-ME RNAi lines also resulted in arrhythmia of transcript oscillations in the central circadain oscillator.

\*47. Bender T, Pena G, Martinou JC: **Regulation of mitochondrial pyruvate uptake by alternative pyruvate carrier complexes**. *EMBO J* 2015, **34**:911-924.

 Different mitochondrial pyruvate carrier (MPC) genes in yeast were found to form complexes that were more active in pyruvate import into yeast mitochondria during fermentative growth or respiratory growth. MPC1 remained constitutive, but MPC2 formed the active complex in fermentative conditions, whereas MPC3 formed the complex in respiratory conditions.

\*\*48. Bricker DK, Taylor EB, Schell JC, Orsak T, Boutron A, Chen Y-C, Cox JE, Cardon CM, Van Vranken JG, Dephoure N, et al.: **A Mitochondrial Pyruvate Carrier Required for Pyruvate Uptake in Yeast, Drosophila, and Humans**. *Science* 2012, **337**:96-100.

 This study identified the genes encoding the two subunits of the mitochondrial pyruvate carrier responsible for the import of pyruvate into mitochondria; a key step in respiratory metabolism.

49. Winter K, Willert DJV: **NaCl Induced Crassulacean Acid Metabolism in *Mesembryanthemum crystallinum***. *Z Pflanzenphysiol* 1972, **67**:166-&.

50. Kore-Eda S, Cushman MA, Akselrod I, Bufford D, Fredrickson M, Clark E, Cushman JC: **Transcript profiling of salinity stress responses by large-scale expressed sequence tag analysis in *Mesembryanthemum crystallinum***. *Gene* 2004, **341**:83-92.

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

\*52. Oh DH, Barkla BJ, Vera-Estrella R, Pantoja O, Lee SY, Bohnert HJ, Dassanayake M: **Cell type-specific responses to salinity: the epidermal bladder cell transcriptome of *Mesembryanthemum crystallinum***. *New Phytol* 2015, **207**:627-644.

 The first published report of quantitative Illumina RNA-seq data for iceplant. Although the transcriptome of epidermal bladder cells was the main focus, a few CAM genes were highlighted as being induced in RNA-seq data when well-watered leaf tissue was contrasted with salt-stressed samples.

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

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

55. Chen L, Hao LG, Parry MAJ, Phillips AL, Hu YG: **Progress in TILLING as a tool for functional genomics and improvement of crops**. *J Integr Plant Biol* 2014, **56**:425-443.

\*\*56. Winter K, Holtum JAM: **Facultative crassulacean acid metabolism (CAM) plants: powerful tools for unravelling the functional elements of CAM photosynthesis**. *J Exp Bot* 2014, **65**:3425-3441.

 An important study that provides high quality gas exchange data characterising drought-induction and reversal of CAM in several potential model facultative CAM species, including *Clusia pratensis*, *Calandrinia polyandra*, *Mesembryanthemum crystallinum*, *Talinum triangulare*, and the facultative C4-CAM species *Portulaca oleracea*.

\*\*57. Christin P-A, Arakaki M, Osborne CP, Braeutigam A, Sage RF, Hibberd JM, Kelly S, Covshoff S, Wong GK-S, Hancock L, et al.: **Shared origins of a key enzyme during the evolution of C4 and CAM metabolism**. *J Exp Bot* 2014, **65**:3609-3621.

 The most detailed study to date on the evolution of the phosphoenolpyruvate carboxylase (*PPC*) gene family in C3, C4 and CAM dicot species within the order Caryophyllales. RNA-seq data for the facultative C4-CAM species *Portulaca oleracea* helped to define distinct *PPC* gene family members that are used for C4 and CAM. Other candidate CAM-associated genes were indentified by virtue of their increased transcript abundance in *P. oleracea* following drought stress.

58. Garcia-Sogo B, Pineda B, Castelblanque L, Anton T, Medina M, Roque E, Torresi C, Beltran JP, Moreno V, Canas LA: **Efficient transformation of *Kalanchoë blossfeldiana* and production of male-sterile plants by engineered anther ablation**. *Plant Cell Rep* 2010, **29**:61-77.

59. Jia SR, Yang MZ, Ott R, Chua NH: **High-Frequency Transformation of *Kalanchoë laciniata***. *Plant Cell Rep* 1989, **8**:336-340.

60. Jung Y, Rhee Y, Auh CK, Shim H, Choi JJ, Kwon ST, Yang JS, Kim D, Kwon MH, Kim YS, et al.: **Production of recombinant single chain antibodies (scFv) in vegetatively reproductive *Kalanchoë pinnata* by *in planta* transformation**. *Plant Cell Rep* 2009, **28**:1593-1602.

61. Baldwin JT: ***Kalanchoë* the genus and its chromosomes**. *Am J Bot* 1938, **25**:572-579.

**Figure captions**

**Figure 1. Carbon isotope ratios correlate with phylogenetic relationships of *Kalanchoë* species.** Basal, group I *Kalanchoës* (Kitchingia section) have been found to possess C3-like carbon isotope signatures, whereas more derived group II (Bryophyllum) and group III (Eukalanchoë) *Kalanchoës* generally have less negative 13C ratios, although group II do tend to show flexible CAM and thus large variations in 13C values. The most derived *Kalanchoës* in group III have the least negative 13C ratios indicative of strong CAM. These species also have the most succulent leaves and inhabit the driest habitats in the South West of their native Madgascar. This strong correlation between phylogeny and the strength of CAM physiology suggests that comparison of the genome sequences of these species will reveal genetic and epigenetic changes that correlate with strong CAM in group III species. The tree is representative only of the major groups within the genus *Kalanchoë* and was redrawn from the published *Kalanchoë* phylogeny [37].

**Figure 2. Examples of representative current and future model *Kalanchoë* species. A.** *K. gracilipes*. This species belongs to the most basal Kitchingia group within the genus (group I) and performs C3 photosynthesis in the well-watered state. Published 13C values for this species sampled from its native Madagascar indicate predominately C3 metabolism. **B**. *K. fedtschenkoi* and **C.** *K. laxiflora*. These intermediate species from group II perform CAM in their older leaves but C3 in their younger leaves. The two species shown are the main species being developed as models for CAM research, and both have genome sequencing efforts underway [5]. 13C values from wild collected material ranged from strong CAM to weak CAM (Fig. 1). Note the visible plantlets forming in the notches along the leaf margins of *K. laxiflora* at the base of the picture in C. The *K. laxiflora* plant also displays a large number of ripening seed pods held above the flowers in C. **D.** *K. millotii*, **E.** *K. beharensis* and **F.** *K. rhombopilosa.* These strong CAM species have the most succulent leaves of the species shown here. The published 13C values for wild collected material from Madagascar indicated that they rely predominantly on dark CO2 fixation (strong CAM) in their native habitat (see Fig. 1 for 13C values).

**Figure 3. Transgenic experiments in *Kalanchoë fedtschenkoi* revealed that NAD-ME in the mitochondrion was the major malate decarboxylating enzyme functioning during the light phase of CAM.** Silencing of NAD-ME using an RNAi construct in transgenic *K. fedtschenkoi* diminished dark CO2 fixation, and greatly reduced light period malate decarboxylation. This indicated that decarboxylation of malate via NADP-ME in either the cytosol and/ or the chloroplast was not able to rescue the down-regulation of NAD-ME in these transgenic lines. The low flux through NADP-ME is indicated by the pale grey arrows in the diagram. Note that the CO2 liberated from malate in the mitochondrion would be refixed by Rubisco in the Calvin Benson cycle in the chloroplast, but these details are absent from the diagram for clarity. Also note that several of the transporters are hypothesised rather than proven to mediate the indicated metabolite movements during CAM. The discovery that NAD-ME was the main route for malate decarboxylation in the light has focused research attention on several key questions concerning metabolite transport steps during the daily CAM cycle. A priority amongst these questions concerns the identity of the pyruvate transporter that exports pyruvate through the inner mitochondrial membrane following malate decarboxylation by NAD-ME. A mitochondrial pyruvate carrier (MPC) has recently been characterised from yeast, Drosophilla and human mitochondria, but the ability of this protein to export pyruvate from the mitochondrion is in doubt as it has only been demonstrated to function as a pyruvate import protein in yeast and animal mitochondria \*\*[47, 48]. Abbreviations: oxaloacetate (OAA), pyruvate (Pyr), glucose 6-phosphate (G6P), phosphoenolpyruvate (PEP), tonoplast dicarboxylate transporter (tDT), mitochondrial dicarboxylate carrier (DiC), NAD-malic enzyme (NAD-ME), NADP-malic enzyme (NADP-ME), mitochondrial pyruvate carrier (MPC), pyruvate orthophosphate dikinase (PPDK), Na+ pyruvate co-transporter (BASS2), PEP:Pi translocator (PPT), G6P:Pi translocator (GPT), chloroplast dicarboxylate transporter (DiT).

**Figure 4. Examples of representative facultative/ inducible model CAM species and model CAM crop species. A.** *Mesembryanthem crystallinum* (Common Iceplant). This eudicot species belongs to the family Aizoaceae in the order Caryophyllales. Main photograph shows an 8-week-old plant and the inset shows a single flower from a 3-month-old plant. **B.** *Talinum triangulare* (Waterleaf). This eudicot species belongs to the family Talinaceae within the Caryophyllales. The photograph shows an 8-week-old plant. **C.** *Calandrinia polyandra* (Parakeelya). This eudicot species resides in the Montiaceae within the Caryophyllales. Main photograph shows an 8 week-old plant prior to flowering, whereas the inset shows a close-up view of the flower. **D.** *Portulaca oleracea* (Common Purslane). This eudicot species is a member of the Portulacaceae within the Caryophyllales. The plants shown are 6-weeks-old and were about to flower. **E.** *Clusia pratensis*. A eudicot member of the Clusiaceae in the order Malphigiales. The photograph show a 2.5-year-old plant. **F.** *Ananas comosus* (Pineapple). A model, monocot CAM crop species in the Bromeliaceae, order Poales. The photograph shows a plant with developing pineapple fruit growing at the Eden Project in Cornwall, England. **G.** *Agave tequilana* Weber var. *azul* (Blue Agave). A model, monocot CAM crop species in the Asparagaceae, order Asparagales. The plant shown was cultivated in the research greenhouse at the Unversity of Liverpool, UK. This species is cultivated on a large scale in the Jalisco region of Mexico for the production of the alcoholic beverage tequila. **H.** *Agave sisalana* (Sisal). A model, monocot CAM crop in the Asparagaceae, order Asparagales. Sisal is grown on a large scale for fibre production in several seasonally dry, sub-tropical countries including Brazil, Tanzania, Kenya and Madagascar. The photograph shows a plant of harvestable age under commercial cultivation in a field near Monteiro, Paraíba, Brazil. **I.** *Opuntia ficus-indica* (Prickly Pear). A model, eudicot CAM crop species in the family *Cactaceae*, order Caryophyllales. The pads of this species are used for cattle fodder, or as a vegetable by humans, and the fruits are also consumed by humans. The photograph shows a plant growing in Spain where *Opuntia ficus-indica* is a common invasive species.