Received: 28 July 2023 | Revised: 25 March 2024 | Accepted: 7 April 2024

DOI: 10.1002/ppp3.10517

Plants People Planet PPP Open Access

RESEARCH ARTICLE

Sequencing complex plants on a budget: The development of *Kalanchoë blossfeldiana* as a C_3, CAM comparative tool

Daniel Cowan-Turner[1,2] | Bethan A. Morris[1] | Alexandra Sandéhn[3] | Iwona Bernacka-Wojcik[3] | Eleni Stavrinidou[3] | Robyn F. Powell[4] | Ilia J. Leitch[4] | Jessica Taylor[1] | Max Walker[1] | Osita Nwokeocha[1] | Maxim V. Kapralov[1] | Anne M. Borland[1]

[1]School of Natural and Environmental Sciences, Newcastle University, Newcastle Upon Tyne, UK

[2]Division of Plant Sciences, Research School of Biology, Australian National University, Canberra, ACT, Australia

[3]Laboratory of Organic Electronics, Department of Science and Technology, Linköping University, Norrköping, Sweden

[4]Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, England, UK

Correspondence
Daniel Cowan-Turner, Division of Plant Sciences, Research School of Biology, Australian National University, Canberra, ACT 2601, Australia.
Email: daniel.cowan-turner@anu.edu.au

M. V. Kapralov, School of Natural and Environmental Sciences, Newcastle University, Newcastle Upon Tyne, NE1 7RU, UK.
Email: maxim.kapralov@ncl.ac.uk

Funding information
Biotechnology and Biological Sciences Research Council (BBSRC), Grant/Award Number: 2132439; BAM, Grant/Award Number: 2462247

Societal Impact Statement

Research efforts in plant biology have often been focused on sequenced and well-studied 'model' organisms. Despite the advent of relatively inexpensive genome sequencing, most plant taxonomic groups are underrepresented, with few species that 'represent' the diversity of whole genera. This study describes an economical guide to sequencing a non-model organism, which may be useful in reducing the cost of sequencing more species within genera and across plant life. This method was used to develop *Kalanchoë blossfeldiana* as a resource for comparing C_3 and the water-conserving mode of photosynthesis known as Crassulacean acid metabolism (CAM) within the same plant.

Summary

- Despite the increasing number of well-studied plant species with well-annotated genomes across plant life, there are few densely sampled genera with more than a couple of genome sequences representing the diversity of whole genera. Here, we develop an economic approach to full-genome sequencing that could be used to sequence many species within a genus. We made use of the Nanopore rapid sequencing kit to assist in plant genome assembly, dramatically reducing the cost.
- Here we applied this method to cost-effectively develop genomic resources for *Kalanchoë blossfeldiana*, a commercially important ornamental, in which Crassulacean Acid Metabolism (CAM), a water-conserving mode of photosynthesis can be induced. We present a physiological and biochemical characterisation of *Kalanchoe blossfeldiana* with its nuclear and chloroplastic genome and a comparative C_3, CAM dusk transcriptome.
- We apply this approach to a complex tetraploid genome, making use of a relative species for chromosomal scaffolding to reduce assembly ploidy, we provide a resource for future gene expression studies. We highlight its limitations, e.g. the need for deeper sequencing to accurately resolve genome structure and haplotypes without using a relative species for scaffolding.

- The study demonstrates the merits of *K. blossfeldiana* as a comparative system for studying C_3 and CAM within a plant and has identified substantial changes in the dusk transcriptome between young C_3 and mature CAM *K. blossfeldiana* leaves in response to age-induced CAM, and shows that in the absence of abiotic stress, CAM induction still involves the engagement of drought and abscisic acid (ABA) response pathways.

KEYWORDS

abscisic acid signalling, age inducible Crassulacean acid metabolism, Crassulacean acid metabolism (CAM), drought stress, genomics, Nanopore sequencings, transcriptomics

1 | INTRODUCTION

The availability of a genome for a species of interest provides a versatile resource and a robust foundation for biochemical studies. Genomes help answer biological questions, provide the underpinnings of complex biological processes and guide our understanding of evolution. Recent projects have made great progress in increasing the availability and quality of genomes. The Earth BioGenome Project aims to sequence all eukaryotic life (Lewin et al., 2018). Re-sequencing, as technologies improve, to improve assembly quality and completeness has increased the quality of published genomes; for example long-read sequencing of *Arabidopsis thaliana* (Michael et al., 2018; Pucker et al., 2021). These efforts have been aided by reductions in the costs and complexity of genome sequencing and the advent of long-read sequencing, such as Nanopore long-read sequencing, which is highly accessible in terms of cost, ease of use and its ability to be used in remote off-grid environments (Dumschott et al., 2020; Gowers et al., 2019; Pucker et al., 2022).

However, challenges remain in genome sequencing. Obtaining sequences from non-model plant species can be hindered by their tendency to be large, heterozygous and polyploid (Kyriakidou et al., 2018), leading to gaps in genomic resources between core model organisms and those less well-studied. This is particularly notable for species with the water-conserving mode of photosynthesis known as Crassulacean Acid Metabolism (CAM) with the first CAM genome, *Ananas comosus* (Ming et al., 2015), published some 15 years after the first C_3 plant, *Arabidopsis* (The Arabidopsis Genome Initiative, 2000) and 5 years after the first C_4 plant (Paterson et al., 2009).

Given that many labs still do not have access to sequencing infrastructure, here we used an economical and accessible approach to develop the age-inducible CAM plant *Kalanchoë blossfeldiana* Poelln. as a resource for investigating C_3 and CAM physiologies and metabolism. *Kalanchoe blossfeldiana* is an economically important houseplant, with millions of plants sold yearly (Smith & Shtein, 2022). We present a physiological and biochemical characterisation of *K. blossfeldiana* along with its genome and a comparative C_3/CAM dusk transcriptome.

1.1 | Crassulacean acid metabolism

CAM is an alternate form of photosynthetic metabolism found in ~7% of land plants that can confer a ~ 6-fold greater water usage efficiency, expressed as CO_2 fixed per unit water lost, compared to the C_3 and C_4 modes of photosynthesis (Borland et al., 2009, Gilman et al., 2023). By moving net gas exchange and stomatal opening to the night, CAM plants reduce water losses by curtailing day-time evapotranspiration, making them Ideal candidates for growth on lands that are less suitable for drought-sensitive plants (Borland et al., 2015). CAM plants have the potential to expand the areas in which plants can grow while providing foodstuffs and other bio commodities (Yang et al., 2015). Engineering CAM into C_3 crops could be a potential method of improving crop plant water use efficiency (WUE) (Borland et al., 2014), without substantially reducing productivity (Shameer et al., 2018). CAM nocturnally fixes CO_2 (as HCO_3) by phosphoenolpyruvate carboxylate (PEPC) into a C_4 acid, usually malic acid. The C_4 acid is stored in the vacuole until subsequent daytime decarboxylation and re-assimilation of CO_2 into the C_3 cycle by Ribulose-1,-5-bisphosphate carboxylase (Rubisco) whilst stomata remain closed. CAM plants are therefore able to shift atmospheric net CO_2 uptake to the night, the inverse of C_3 and C_4 plants. Alongside nocturnal carboxylation and daytime decarboxylation, CAM involves a number of other key mechanisms: the inverted pattern of nocturnal stomatal opening/daytime closure; the cycling of carbon in and out of transitionary carbon stores to support phosphoenolpyruvate (PEP) regeneration (Borland & Dodd, 2002); metabolic changes from C_3, such as the use of the phosphorolytic pathway to degrade starch at night in starch storing CAM plants (Ceusters et al., 2021); and increased leaf succulence (Nelson & Sage, 2008; Zambrano et al., 2014).

For CAM to emerge from C_3 photosynthesis, it is believed that substantial rewiring and retiming of core metabolism occurred together with the recruitment of enzymes and pathways that are already present in C_3 plants (Borland et al., 2014; Lim et al., 2019). Efforts to understand CAM, its induction and regulation need to consider genome-scale changes in gene expression. In recent years there has been a flurry of 'CAM-omics' studies, with the genomes of constitutive and facultative CAM species from across plant life sequenced, including *A. comosus* (Ming et al., 2015), *Kalanchoë fedtschenkoi* (Yang

et al., 2017), *Kalanchoë laxiflora* (available on Phytozome, (Goodstein et al., 2012)), *Sedum album* (Wai et al., 2019), *Talinum triangulare* (Brilhaus et al., 2016; Maleckova et al., 2019; Maleckova, 2020), *Talinum fruticosum* (Brilhaus et al., 2023), *Isoetes taiwanensis* (Wickell et al., 2021) and *Cistanthe longiscapa* (Ossa et al., 2022). See Gilman et al. (2023) for a comprehensive list of CAM genomes.

Many CAM genomics studies compare diel transcript expression between CAM and C_3 species to identify genes that were either upregulated in the CAM species or showed a significant diel shift in expression pattern compared to that in the C_3 species (e.g. Moseley et al., 2019). This approach has identified and corroborated the importance of many key CAM genes. These works may, however, be complicated by inter-specific differences in gene expression, especially as many of these comparisons are made between CAM species and *Arabidopsis*, which can differ significantly in their morphology, geography and phylogenetic placing. More recently, efforts have been made to compare the gene expressions of more closely related species, such as comparisons between C_3, CAM and C_3/CAM hybrids in *Yucca* (Heyduk et al., 2019).

Another approach, highlighted by Winter et al. (2008), is to study changes that occur when CAM is induced by abiotic stress such as drought or salinity as in facultative CAM plants. Using the same species to compare CAM and C_3 transcription, physiologies and metabolites (Maleckova et al., 2019) reduces potential conflation from interspecific differences. However, while useful in identifying many key elements of CAM that are upregulated upon CAM induction, such studies are complicated by the potential overlap between CAM induction and abiotic stress response. There have been fewer studies comparing changes in age-induced CAM which appears in several genera which exhibit CAM (for example, Ming et al., 2015 and Niechayev et al., 2023), but it is relatively understudied in comparison to abiotic stress induction.

Here we investigate the age-inducible CAM species *Kalanchoë blossfeldiana* to determine whether it is possible to explore differences between C_3 and CAM physiologies within the same plant without the confounding issue of stress-induced changes in metabolism. CAM in *K. blossfeldiana* can be induced by short days (Lerman & Queiroz, 1974), nitrogen deficiency (Ota, 1988), ABA, drought (Taybi et al., 1995) and leaf ageing (Brulfert et al., 1982). The current *Kalanchoe* genus is currently made of up species from three previous genera, *Kalanchoe* (of which *K. blossfeldiana* was a member), *Kitchingia* and *Bryophyllum* (of which *Kalanchoe fedtschenkoi* and *K. laxiflora* were members Smith & Figueiredo, 2018). *K. blossfeldiana* shows some significant morphological and physiological differences from *K. fedtschenkoi* and *K. laxiflora* - it is predominantly green-leafed, lacks the ability to reproduce from leaf tip pups, pendulous flowers, and performs inducible CAM not constitutive CAM as found in *K. fedtschenkoi* and *K. laxiflora*. The current *Kalanchoe* genus reflects these differences with three subgenera, matching the previously separate genera, with *K. blossfeldiana* in the *Eukalanchoe* subgenus. This is well summarised by Hartwell et al. (2016), who suggested the genus as a whole was ripe for further genomic study to utilise the variation in C_3/CAM phenotypes across the genus, rather than picking individual 'model' species as CAM has evolved independently multiple times.

We sought to investigate the utility of age or ontogenetically induced CAM in *K. blossfeldiana*, with newly emerging leaves performing C_3 for a substantial period before transitioning to CAM. Such a system may provide a useful resource for the comparative study of C_3 and CAM between leaf ages without conflation from interspecific differences and between CAM induction and abiotic stress responses. We aimed to develop this resource in an accessible and economical manner with the hope of providing a roadmap for future studies to develop plant species for future genetic and biochemical studies, that could be applied to the sequencing of whole or large portions of genera that are constrained by high cost.

2 | MATERIALS AND METHODS

2.1 | Plant materials and growth conditions

Kalanchoë blossfeldiana plants were propagated from cuttings. Side shoots were cut from the main stem with all but the emerging leaf bud at the apex of the shoot, and the top two leaf pairs removed. Individual cuttings were placed in 5 cm pots with John Innes No.2 compost and grown under clear plastic covers for 2 weeks to maintain humidity. Plants were grown in a Fitotron SGC120 growth chamber (Weiss Technik, Germany) and watered twice weekly. Conditions in the growth cabinet were 190 µmol m^{-2} s^{-1} light at plant height, with 12/12 hrs day/night cycles, 25°C days and 19°C nights, relative humidity was held between 50 and 80%. After 4 weeks, cuttings were re-potted into 12 cm pots with a 3:1 mix of John Ins No.2: Perculite.

2.2 | Diel gas exchange

Diel stomatal conductance (mol H_2O m^{-1} s^{-1}) and net CO_2 uptake (µmol CO_2 m^{-1} s^{-1}) were measured using a LI-COR Portable Photosynthesis System Li-6400XT (LI-COR, USA) over 24 hr periods. All such measurements were run successively in triplicate. Chamber conditions were set to track the conditions of the growth cabinet, with 190 µmol m^{-2} s^{-1} PPFD, unless otherwise noted, with 12/12 hrs day/night cycles and with 25°C days and 19°C nights, relative humidity was held between 35 and 55%. For a standard 24 hr diel run measurements were taken every 15 minutes with an automatic infrared gas analyser (IRGA) matching every 30 minutes, to reduce offsets between IRGAs, and ambient CO_2 set to 400 µmol mol^{-1} CO_2. For all Gas exchange experiments leaves at leaf pairs two (mature leaf) and ten (young leaf, counted from the bottom up) were used with plants 60–90 days after propagation, this was to ensure a balance between optimal leaf size and differences between the photosynthetic metabolism of the leaves.

2.3 | Plant sampling

Plants were sampled 90 days after propagation. *K. blossfeldiana* young leaves (leaf pairs 9 and 10 from the bottom) and mature leaves (leaf

pairs 2 and 3 from the bottom) were sampled. Tissue for biochemical analysis was flash frozen in liquid nitrogen either using a 1 cm leaf disc or the whole leaf. Tissue was ground using a pestle and mortar or a pre-chilled TissueLyser II (Qiagen, USA) for 2 minutes at 30 Hz with a single 2 mm metal bead. Samples were stored at −80°C before analysis.

2.4 | Crassulacean acid metabolism traits – metabolite, morphological and protein analysis

Methanol extracts were used for metabolite analyses, described in detail in Methods S1. The titratable acidity of leaf tissue was determined using 200 µl methanol extracted tissue and titrating against 5 mM NaOH until neutrality, using phenolphthalein as an indicator. Starch was extracted from whole leaves as described previously (Haider et al., 2012). Whole leaf total soluble sugars in methanol extracts were determined by measuring glucose equivalents using a phenol–sulfuric acid colourimetric method (Dubois et al., 1956). Protein extractions and western blots to quantify phosphoenolpyruvate carboxylate (PEPC) and Ribulose-1,5-bisphosphate carboxylase (Rubisco) are described in Methods S2. Methods to quantify CAM morphological traits are described in Methods S3.

2.5 | DNA extraction

Genomic DNA was extracted using an adapted version of Keb-Llanes et al. (2002) followed by a further clean-up using the ChargeSwitch gDNA Mini Bacteria Kit (Invitrogen, USA). Exactly 300 mg of ground leaf tissue was combined with 300 µl buffer A (2% Hexadecyltrimethylammonium bromide (CTAB) (w/v), 100 mM Tris–HCl (pH 8), 20 mM EDTA, 1.4 M NaCl, 4% polyvinylpyrrolidone (PVP-40) (w/v), 0.1% ascorbic acid (w/v), and 10 mM β-mercaptoethanol); and 900 µl buffer B (100 mM Tris–HCl (pH 8), 50 mM EDTA, 100 mM NaCl);100 µl 20% SDS and 10 mM β-mercaptoethanol. Extracts were vortexed and then incubated for ten minutes at 65°C while shaking at 1500 rpm. A total of 410 µl potassium acetate was added, and samples were then spun down at 15,300*g* for 15 minutes at 4°C. 1 ml of supernatant was added to 540 µl ice-cold isopropanol and incubated on ice for 20 minutes. Samples were then washed three times as follows: samples were centrifuged at 9600 *x g* for ten minutes at 4°C; the supernatant was discarded, and the pellet was washed with 500 µl 70% ethanol and immediately air dried; the pellet was then resuspended in 600 µl TE buffer (10 mM Tris, 1 mM EDTA (pH 8)), 60 µl 3 M sodium acetate (pH 5.2) and 360 µl ice-cold isopropanol; the sample was then incubated for 20 minutes on ice. After three washes, the sample was spun down at 9600 g for ten minutes at 4°C then resuspended in 50 µl resuspension buffer (R4, from the ChargeSwitch kit) containing RNase A. The ChargeSwitch protocol followed from this point on, as described in its manual. All samples were accessed for quality by nanodrop 2000 (Thermo Fisher, Massachusetts) and agarose gel with a high molecular weight ladder (N3239S, NEB, Massachusetts).

2.6 | Genome sequencing

The genome size of *K. blossfeldiana* was determined as described in Methods S4. Short-read whole genome library preparation and sequencing of *K. blossfeldiana* was carried out by Novogene (Beijing, China). Sample and library integrity was tested with an Agilent 2,100 Bioanalyzer (Agilent Technologies, USA). Paired end 150 bp libraries were produced using the NEB Next Ultra DNA Library Prep Kit and sequencing was performed in three runs on the same library on an Illumina NovaSeq 6,000. The three runs produced a total of 91.6 Gigabases (Gb), 305 million PE reads, and an estimated sequencing coverage of ~200x for a predicted haploid genome size of 450 Mb. Initial sequence QC and adaptor trimming were carried out by Novogene and then assessed in FastQC 0.11.9 (Andrews, 2010). Long-read sequencing was carried out over three r9.4.1 minion flow cells using either the RAD004 or RBK004 kits, with libraries prepared as per their protocols. Long reads were basecalled and quality filtered by Guppy 6.2.1 (Oxford Nanopore Technologies) and read quality was assessed in LongQC 1.2.0 (Fukasawa et al., 2020). After filtering there was a total of 9.6 Gb from 3.49 million Nanopore long reads. The costs of genome sequencing are broken down in Methods S5.

2.7 | Genome assembly

Genome assembly was performed with a hybrid assembly approach using MaSuRCA 4.0.9 (Zimin et al., 2017), which combines accurate short and error-prone long reads into super reads and is often used in the assembly of large plant genomes (Scott et al., 2020; Wang et al., 2020). Assemblies were also performed using only short reads with abyss 2.3.5 (Jackman et al., 2017) and only long reads with Flye 2.9.1 (Kolmogorov et al., 2019) for comparison. Assembly quality and size statistics were calculated at each step of assembly processing by BUSCO 5.4.3 (Manni et al., 2021) and the abyss-fac script from the abyss assembler.

The MaSuRCA assembly was scaffolded using Ragtag 2.1.0 (Alonge et al., 2019, 2021) with the *Kalanchoë laxiflora* genome (FTBG2000359A v3.1) accessible on Phytozome (Goodstein et al., 2012). Scaffolds shorter than 10 kbp were removed. Multiple methods were used to reduce assembly ploidy to one haplotypic representation of the tetraploid genome. Purge Haplotigs 1.1.2 (Roach et al., 2018) and Purge Dups 1.2.6 (Guan et al., 2020) were run sequentially to reduce assembly ploidy. Scaffolds not contained within chromosome-scale pseudomolecules were filtered using a BUSCO optimisation approach where a combination of scaffolds was retained to give the highest BUSCO score while reducing ploidy. After reducing ploidy, the assembly was polished with the short reads using Pilon 1.24 (Walker et al., 2014). BUSCO was used to assess genome completeness by detecting the presence of key single-copy genes from the embryophyta_odb10 database. The genome assembly was assessed using the long-terminal repeat (LTR) assembly index, which is the proportion of intact long terminal repeats (LAI) within the assembly, from the LTR retriever pipeline 2.9.0 (Ou et al., 2018). A

higher score tends to suggest a more contiguous and complete assembly; and is improved by both short reads increasing accuracy per base and long reads providing resolution over long repetitive regions. A detailed walkthrough of the nuclear genome assembly method is available, see the GitHub link in the data accessibility section.

2.8 | RNA extraction

RNA was extracted from young and mature leaves, at dusk using a method adapted from Liao et al., (2004). The samples were the same as those used for metabolite analysis. Exactly 100 mg of ground tissue was combined with 900 µl of pre-heated (65°C) extraction buffer: 3% Hexadecyltrimethylammonium bromide (CTAB) (w/v), 100 mM Tris–HCl (pH 8), 25 mM EDTA (pH 8), 2 M NaCl, 3% polyvinylpyrrolidone (PVP-40) (w/v), 0.5 g/l spermidine and 4% β-mercaptoethanol added just prior to extraction and incubated at 65°C for ten minutes while shaking. Exactly 900 µl of chloroform: isoamyl alcohol (24:1) was added and mixed by inversion for ten minutes. Samples were centrifuged at max speed for ten minutes at 4°C, with supernatants transferred to fresh tubes with an equal volume of chloroform: isoamyl alcohol (24:1), then incubated at 4°C for ten minutes while inverting. The upper phase was discarded and two times volume of ethanol was added, followed by freezing at −20°C for two hours and centrifuged at 15,000 *x g* for 30 minutes. The supernatant was discarded and 1 ml lithium chloride was added, followed by two hours at −20°C and centrifuged at 15,000 *x g* for 30 minutes. The supernatant was discarded and the RNA pellet was resuspended in 100 µl of DEPC water, with gentle shaking at 65°C for ten minutes. Samples were then DNase treated with an Invitrogen DNase kit, as described by its manual.

2.9 | Dusk RNA-sequencing

Short-read RNA library preparation and sequencing of *K. blossfeldiana* was carried out by Novogene (Beijing, China), with RNA extracted as above. Three biological replicates were collected and sequenced from each leaf age group. A total of 119 million PE reads, and 35.8 Gb were sequenced on an Illumina NovaSeq 6,000, with a range of 15–22 million reads per sample. RNA sequencing QC, alignment and quantification were performed using the nf-core/rnaseq 3.8.1 pipeline (Ewels et al., 2020) with default settings. Briefly, read QC was performed using FastQC, adapters and quality trimming by Trim Galore. STAR was used to align reads to gene models predicted in the *K. blossfeldiana* genome, and Salmon was used to quantify transcript abundance (Patro et al., 2017). DESeq2 (Love et al., 2014) was used to analyse differential expression between samples.

2.10 | Genome annotation

RepeatModeler 2.02 (Flynn et al., 2020) and RepeatMasker 4.1.3 (Tarailo-Graovac & Chen, 2009) were used to annotate repetitive regions of the genome assembly. The assembly was soft-masked using Bedtools 2.30 (Quinlan & Hall, 2010). BRAKER2 2.1.6 (Brůna et al., 2021) was used for gene annotation. Two methods were used, as recommended, firstly mRNA based gene predictions were made using the BRAKER pipeline with the Illumina RNA reads as described above. Simultaneously, in a separate BRAKER2 run, gene predictions were made using a set of training proteins from the OrthoDB plant database (Kriventseva et al., 2019) and *K. fedtschenkoi* (Yang et al., 2017). Annotations from both BRAKER runs were combined using the transcript selector TSEBRA (Gabriel et al., 2021). Protein sequences from gene models predicted by the BRAKER pipeline above were functionally characterised. This was primarily performed using OrthoFinder 2.5.4 (Emms & Kelly, 2019) to sort proteins into orthologous groups from *K. fedtschenkoi* (*Kalanchoë fedtschenkoi* v1.1), *K. laxiflora* (FTBG2000359A v3.1), *Vitis vinifera* (Vvinifera 457 v2.1) and *A. thaliana* (TAIR10) from Phytozome (Goodstein et al., 2012). Annotations present from these genomes in each orthogroup were used to fill in *K. blossfeldiana* annotation in the above order.

2.11 | *K. blossfeldiana* chloroplast genome and Phylogenetics

The *K. blossfeldiana* chloroplast genome was assembled with GetOrganelle 1.7.5 (Jin et al., 2020) using the same Illumina PE reads as were used for the nuclear genome. Functional and structural annotation of the chloroplast genome was performed using GeSeq (Tillich et al., 2017) with the *A. thaliana* chloroplast genome used as a reference. The *K. blossfeldiana* chloroplast genome was primarily used to explore the phylogenetic relationship between this variety of *K. blossfeldiana* and other members of the genus *Kalanchoe*, and the wider *Crassulaceae* family. Whole chloroplast genome sequences from 79 species across the order *Saxifragales* were used with a focus on the *Crassulaceae* family and species from the *Haloragaceae* family were used as outgroups. Whole chloroplast genome sequences were obtained from NCBI and aligned using MAFFT 7.505 (Katoh & Standley, 2013). ModelFinder (Kalyaanamoorthy et al., 2017) used the Bayesian Information Criterion (BIC) to select a nucleotide substitution model. IQ-TREE 2.2.0.3 (Minh et al., 2020) was used to construct a phylogenetic tree using maximum likelihood, with 1,000 ultrafast bootstrap replicates (Hoang et al., 2017).

3 | RESULTS AND DISCUSSION

3.1 | *Kalanchoë blossfeldiana* is a useful system for comparing C_3 and CAM within the same plant

Firstly, we determined that *K. blossfeldiana* would be a relevant and useful species in which a comparison of the young and mature leaves would enable an analysis of genetic and mechanistic differences between C_3 and CAM within the same plant. There was a

clear distinction in photosynthetic metabolism between young and mature leaves; the young leaves (10th leaf pair from bottom) showed C_3 photosynthesis, no net nocturnal uptake of CO_2 and no overnight accumulation of titratable acidity (Figure 1). In contrast, mature leaves (2nd leaf pair from bottom) performed CAM, assimilating ~2/3 of their CO_2 overnight, nocturnally accumulating substantial amounts of titratable acidity, and expressing greater levels of PEPC protein than the young leaves (Figure S1). We observed significantly higher rates of net CO_2 assimilation and higher soluble sugar content in the young leaves compared to mature (Figure 1), although it is not clear whether these changes in soluble sugars are due to CAM or if this is a by-product of leaf ageing. *K. blossfeldiana* leaves became more succulent as they aged (Figure 1), in keeping with the often-observed coexistence of succulence and CAM (Nelson & Sage, 2008; Zambrano et al., 2014). Together these results indicate that *K. blossfeldiana* is a useful system for comparing C_3 and CAM physiology and metabolism in response to ageing with a clear distinction noted between the young C_3-performing and mature CAM-performing leaves. Studies have shown the developmental progression of CAM in other species, that tend to involve plants that exhibit weak CAM that progressively become more CAM as they develop (Niechayev et al., 2023; Ping et al., 2018; Winter et al., 2008). A study with *A. comosus* showed a developmental gradient across the leaf, however, the youngest CAM-free section was chlorophyll-free and therefore also unable to complete the light-dependent section of photosynthesis. Together these results indicate that *K. blossfeldiana* is a useful system for comparing C_3 and CAM physiology and metabolism in response to ageing with a clear distinction between the young C_3-performing and mature CAM-performing leaves.

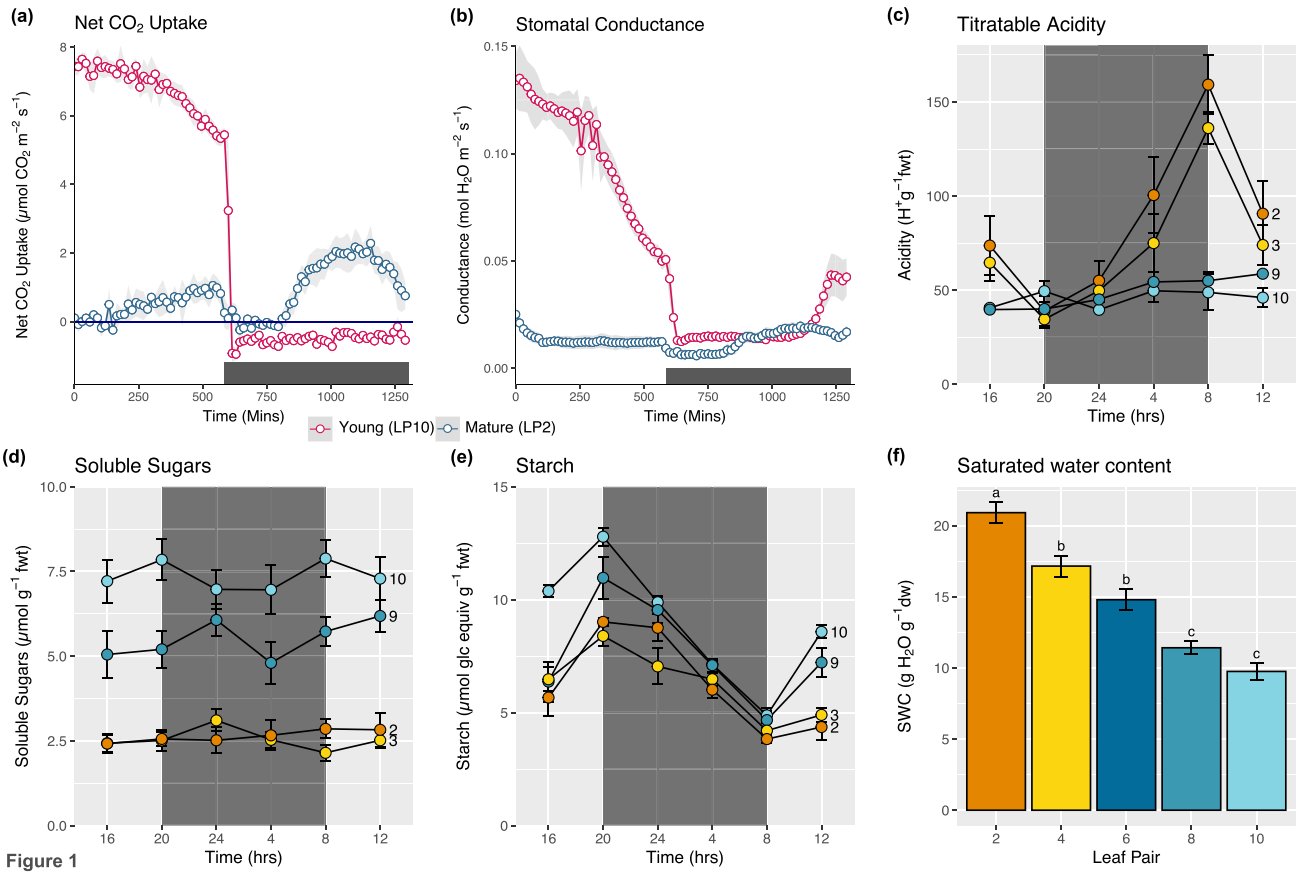

FIGURE 1 *Kalanchoë blossfeldiana* age-specific diel photosynthetic traits, whole leaf titratable acidity, metabolites and Crassulacean acid metabolism (CAM) morphology traits. Diel net CO_2 uptake (panel a, μmol CO_2 m^{-1} s^{-1}) and stomatal conductance (panel B, Mol H_2O m^{-1} s^{-1}) of *Kalanchoe blossfeldiana* young leaves (red points, 10th leaf pair from the bottom) and mature leaves (blue points, 2nd leaf pair from the bottom) between 60 and 90 days after propagation. For panels A and B points represent the mean of three repeats and the shaded interval ± SEM. Horizontal blue line equals zero, panel a. for panels C-E numbers on the left of each graph refer to leaf pair, numbered from the bottom up, mature leaves (LP2: Orange, LP3: yellow) and young leaves (LP9: dark blue, LP10: light blue). Titratable acidity (μmol H^+ g^{-1} fwt) over 24 hrs (panel C), soluble sugars (μmol Glc equiv g^{-1} fwt) over 24 hrs (panel D), starch (μmol Glc equiv g^{-1} fwt) over 24 hrs (panel E). Leaves were sampled every four hours, with the first sampling 4 hrs before dusk. Plants were sampled 90 days after propagation. Values are means of four replicates ± SEM. Leaf-saturated water content (SWC, panel F). Leaf pairs are numbered up the plant, with 12 being the youngest sampled. Numbers above bars on the top left panel indicate sample number, values are means of the noted number of replicates ± SEM. Statistical analysis was performed using one-way ANOVA, significant differences ($P < 0.05$) are indicated by different lowercase letters.

3.2 | *Kalanchoë blossfeldiana* genome assembly and annotation

Many varieties of *K. blossfeldiana* vary in their ploidy with diploid varieties reported with (2n = 34) and tetraploid (2n = 68–72) (van Voorst & Arends, 1982), with 17 or 18 chromosomes expected per haplotype. This variety of *Kalanchoë blossfeldiana* used was estimated to be tetraploid, with a genome size of ~1.94 Gb (2C = 1.99 ± 0.05 pg, Table S1) and with each haplotype 450–490 Mb estimated by flow cytometry and computationally. Initial read counts, and quality statistics for both DNA and RNA sequencing are available in Tables S2-S4.

An initial *K. blossfeldiana* genome assembly of 772 Mb and an N50 of 1 Mb was made using the hybrid assembler MaSuRCA 4.0.9 (Zimin et al., 2017), with 91.6 Gb, 305 million PE Illumina short reads and 10 Gb 3.49 million Nanopore long reads (N50–10 Kb). After determining this variety of Kalanchoë *blossfeldiana* was tetraploid (Figure S2), for downstream transcript abundance analysis we needed to either resolve each haplotype, a process that would require significantly greater long read sequencing depth or attempt to obtain one representative haplotype which could be used to align transcripts to. We decided to obtain one representative haplotype by scaffolding the assembly using Ragtag 2.1.0 (Alonge et al., 2019, 2021) with the *K. laxiflora* genome (FTBG2000359A v3.1, Phytozome); resulting in 18 chromosome-scale pseudomolecules (377 Mb) and 155 smaller scaffolds (268 Mb). Following this we further reduced assembly ploidy by reducing duplication in these remaining 155 smaller scaffolds. This route, while useful in simplifying transcript abundance analysis, does raise issues with the structure of the genome, as this will certainly remove structural features that are specific to *K. blossfeldiana*, and therefore we would not recommend its use for any genome structural analysis.

The final assembly was 461 Mb, consisting of 18 chromosome-scale pseudomolecules and 125 scaffolds with an N50 of 20.6 Mb; after reducing assembly ploidy using Purge Haplotigs 1.1.2 (Roach et al., 2018) and Purge Dups (Guan et al., 2020) and a BUSCO optimisation approach where a combination of scaffolds were retained to give the highest BUSCO single copy score while reducing duplication. Table S5 provides an overview of genome assembly size through these steps. The final assembly had a genomic BUSCO score of 95.6% complete copies, with 70% present in single copies and 25.6% present in duplicated copies using the Embryophyte odb10 database of single-copy genes (Figure 2). A number of methods were used to assess assembly quality and completeness (Table 1). Over half of the genome assembly was identified as repetitive elements (57.3%), the majority marked as LTRs (24.7%), DNA transposons (18.7%) and LINES (7.71%, Figure 2). The BRAKER2 and TSEBRA functional annotation pipeline (Gabriel et al., 2021) identified 38,117 protein-coding genes, with 95.3% of the BUSCO Embryophyte odb10 genes present. The majority (32,096, 84%) of these proteins were functionally annotated.

3.3 | Phylogenetics of *Kalanchoë blossfeldiana*

The *Kalanchoë* genus is estimated to have diverged from the *Sempervivoideae* between 22.78 and 44.29 million years ago (Messerschmid et al., 2020). The current *Kalanchoe* genus is made of up species from three previous genera, *Kalanchoe* (of which *K. blossfeldiana* was a member), *Kitchingia* and *Bryophyllum* (of which *K. fedtschenkoi* and *K. laxiflora* were members and carry out leaf margin embryogenesis (Smith & Figueiredo, 2018). *K. blossfeldiana* shows some morphological and physiological differences from *K. fedtschenkoi* and *K. laxiflora* - it is predominantly green-leafed, lacks the ability to reproduce from leaf tip pups and performs inducible CAM vs constitutive CAM in the latter. Therefore, it was expected that *K. blossfeldiana* would phylogenetically be within the *Kalanchoe* genus but distinct from *K. fedtschenkoi* and *K. laxiflora*. The chloroplast sequence is a useful tool for determining the phylogenetic relationship between *K. blossfeldiana* and other relative species, and recently chloroplasts of a large number of species within the Crassulaceae have been sequenced (Han et al., 2022; Tian et al., 2021). As expected, the *K. blossfeldiana* chloroplast genome was highly conserved between it and other members of the *Kalanchoë* genus with 99.01% sequence similar to *K. fedtschenkoi* chloroplast genome (Figure 3). A *K. blossfeldiana* chloroplast genome sequence recently deposited on NCBI (NCBI accession: OM320795) was identical to *the K. blossfeldiana* chloroplast sequence presented here. Such findings suggest that the chloroplastic genome is highly conserved between *K. blossfeldiana* varieties as would be expected as the varieties are likely to have formed by horticultural breeding programs over the last century. Phylogenies based on whole chloroplast genome sequences (Figure 3) and single copy genes (Figure S3) show *K. blossfeldiana* positioned within the *Kalanchoe* genus but away from *K. fedtschenkoi* and *K. laxiflora*. Its placement and the differences between the three historic genera would be further helped with future genomic sequencing of close relatives of *K. blossfeldiana* such as *Kalanchoë globulifera* and *Kalanchoë darainensis* (Klein et al., 2021).

3.4 | Core CAM genes are upregulated in mature CAM *K. blossfeldiana* leaves

There was a substantial difference in transcript abundances between the young C_3 (9th leaf pair from the bottom) and mature CAM leaves (3rd leaf pair from the bottom), with 685 genes significantly upregulated and 1,491 downregulated in the mature leaves compared to young leaves at dusk. A complete list of up or down-regulated transcript abundances is available in Dataset S1 and key RNA sequencing quality statistics are reported in Figure S4.

Many of the genes upregulated are genes implicated in CAM, with orthologs of the upregulated genes also shown to be upregulated with stress-induced CAM in other inducible CAM species (Maleckova

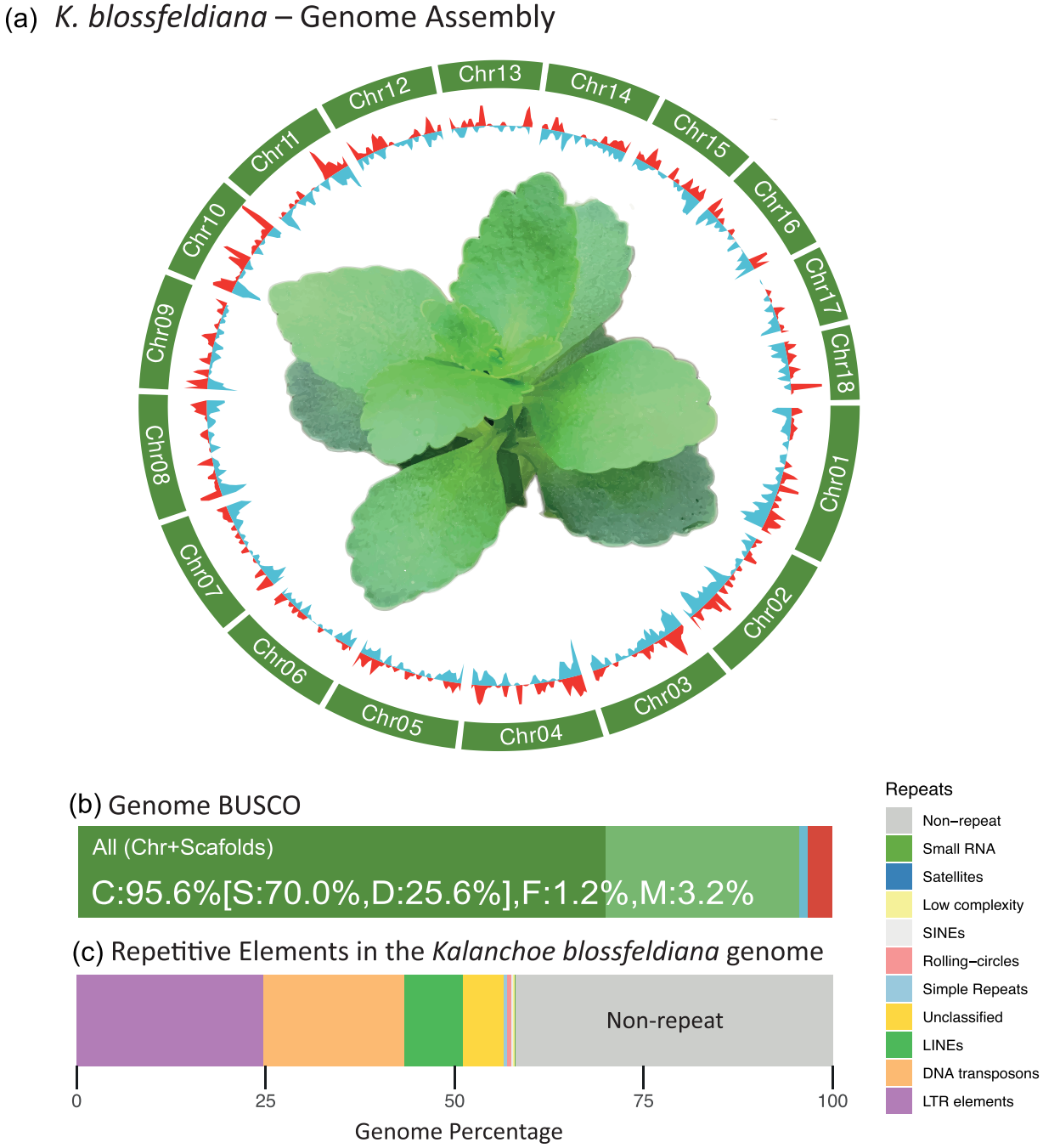

FIGURE 2 *Kalanchoë blossfeldiana* genome assembly and repetitive elements. The genome *of Kalanchoe blossfeldiana* (panel a), scaffolded into 18 chromosome-scale pseudomolecules (377 Mb) and 155 smaller scaffolds (not shown, 268 Mb) using the *Kalanchoë laxiflora* genome (Phytozome). *K. blossfeldiana* is thought to have 17 or 18 chromosomes per haplotype. Density of genes up-regulated (red, outward) or down-regulated (blue, inward) in the mature Crassulacean acid metabolism (CAM) *K. blossfeldiana* leaves (3rd leaf pair from the bottom) compared to the young C_3 *K. blossfeldiana* leaves (9th leaf pair from the bottom). Inner image is of *K. blossfeldiana*. *K. blossfeldiana* Genome completeness (panel B) was assessed using the benchmarking universal single-copy orthologue (BUSCO) tool with the Embryophyte odb10 database of single-copy genes (A), showing the percentage of complete (C) present in single copies (S) or duplicated (D), fragmented (F) or missing (M) genes detected. Proportions of repetitive elements detected within the *K. blossfeldiana* genome (panel C), abbreviations shown include long or short interspersed nuclear elements (LINES or SINES), long terminal repeats (LTRs).

et al., 2019) and/or shown to have CAM-like patterns of expression in constitutive CAM species (Yang et al., 2017). Many of these upregulated 'core-CAM' genes are highlighted in Figure 4. Of these, notably, *PEPC1* appears to be the key PEPC isoform involved in nocturnal malate accumulation in both *K. blossfeldiana* (kblos_g8596) and *K. fedtschenkoi* (Kaladp0095s0055, Yang et al., 2017), with a 3.7-fold

TABLE 1 *Kalanchoë blossfeldiana* genome assembly statistics after scaffolding with the *Kalanchoë laxiflora* genome. The genome was assembled using a hybrid assembly approach using MaSuRCA 4.0.9 (Zimin et al., 2017), followed by scaffolding using the *Kalanchoë laxiflora* genome (FTBG2000359A v3.1) accessible on Phytozome (Goodstein et al., 2012). Multiple methods were used to reduce assembly ploidy to one haplotypic representation of the tetraploid genome. The long-terminal repeat (LTR) assembly index (LAI) refers to the proportion of intact long terminal repeats (LAI) within the assembly, from the LTR retriever pipeline 2.9.0 (Ou et al., 2018). Read mapping rate is the percentage of reads used in genome assembly that map back to the assembly.

					Read mapping rate (%)		
Scaffolds:	N50 (Mb)	Largest scaffold (Mb)	Assembly size (Mb)	Gaps (%)	Short reads	Long reads	Long-terminal repeat (LTR) assembly index (LAI)
143	20.06	29.6	461	0.138	89.03	94.38	11.12

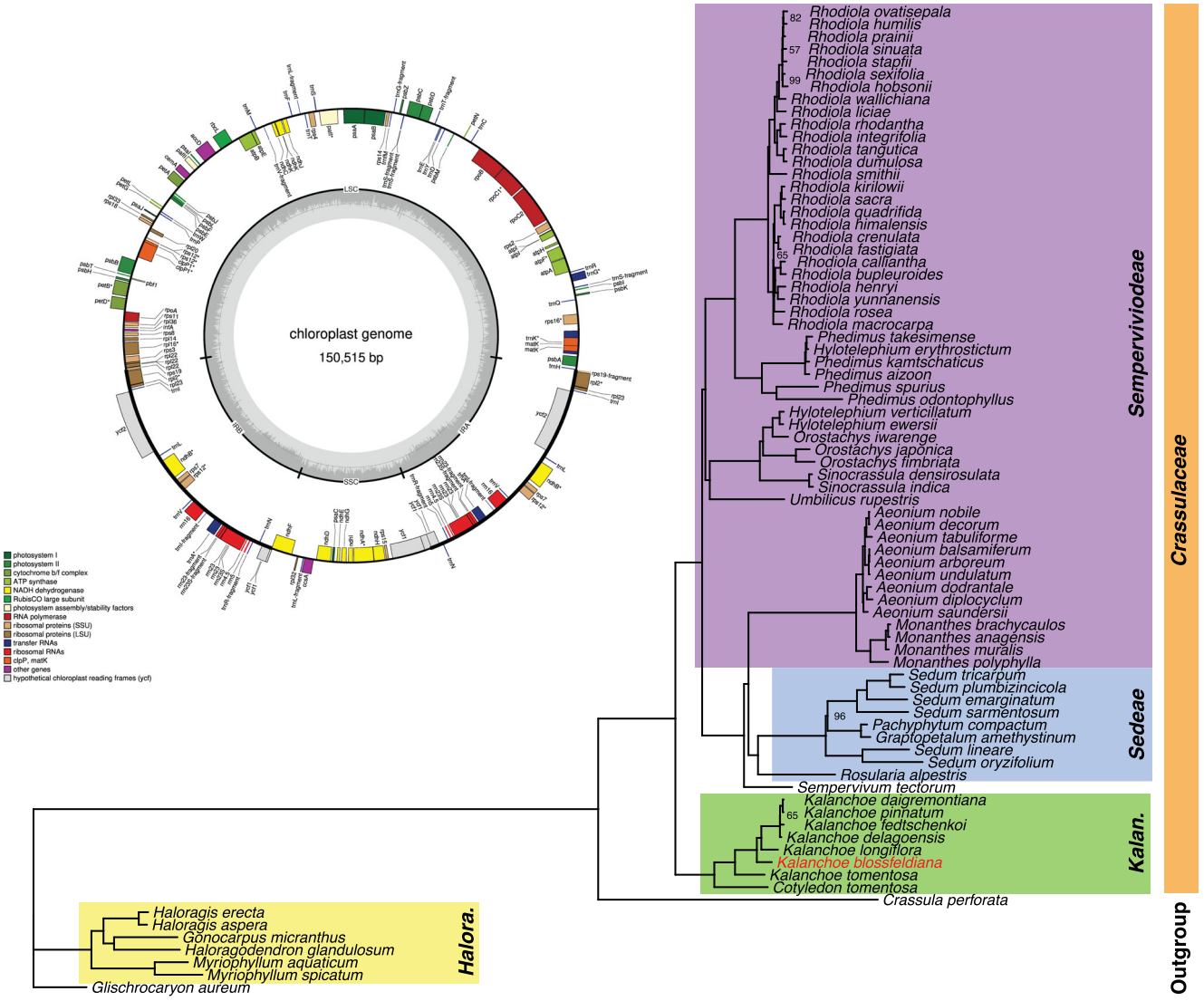

FIGURE 3 Annotated features of the *Kalanchoë blossfeldiana* chloroplast genome (left). Genes transcribed are displayed on the outer ring, on the inside if transcribed clockwise or outside if counter-clockwise. The chloroplastic genome is divided into a large single copy (LSC) region, a small single copy (SSC) region and a pair of inverted repeat regions (IRA and IRB). The intertrack conveys GC content. The annotation and diagram were produced using GeSeq (Tillich et al., 2017). The phylogenetic position of *Kalanchoë blossfeldiana* (right). Whole chloroplast genome sequences from 79 species across the order *Saxifragales* were aligned and reconstructed into a phylogenetic tree with a focus on the *Crassulaceae* family, species from the *Haloragaceae* family were used as outgroups. Whole chloroplast genome sequences were aligned using MAFFT 7.505 (Katoh & Standley, 2013). IQ-TREE 2.2.0.3 (Minh et al., 2020) was used to construct a phylogenetic tree using maximum likelihood, with 1,000 ultrafast bootstrap replicates (Hoang et al., 2017), with nodes labelled if below 100% bootstrap support.

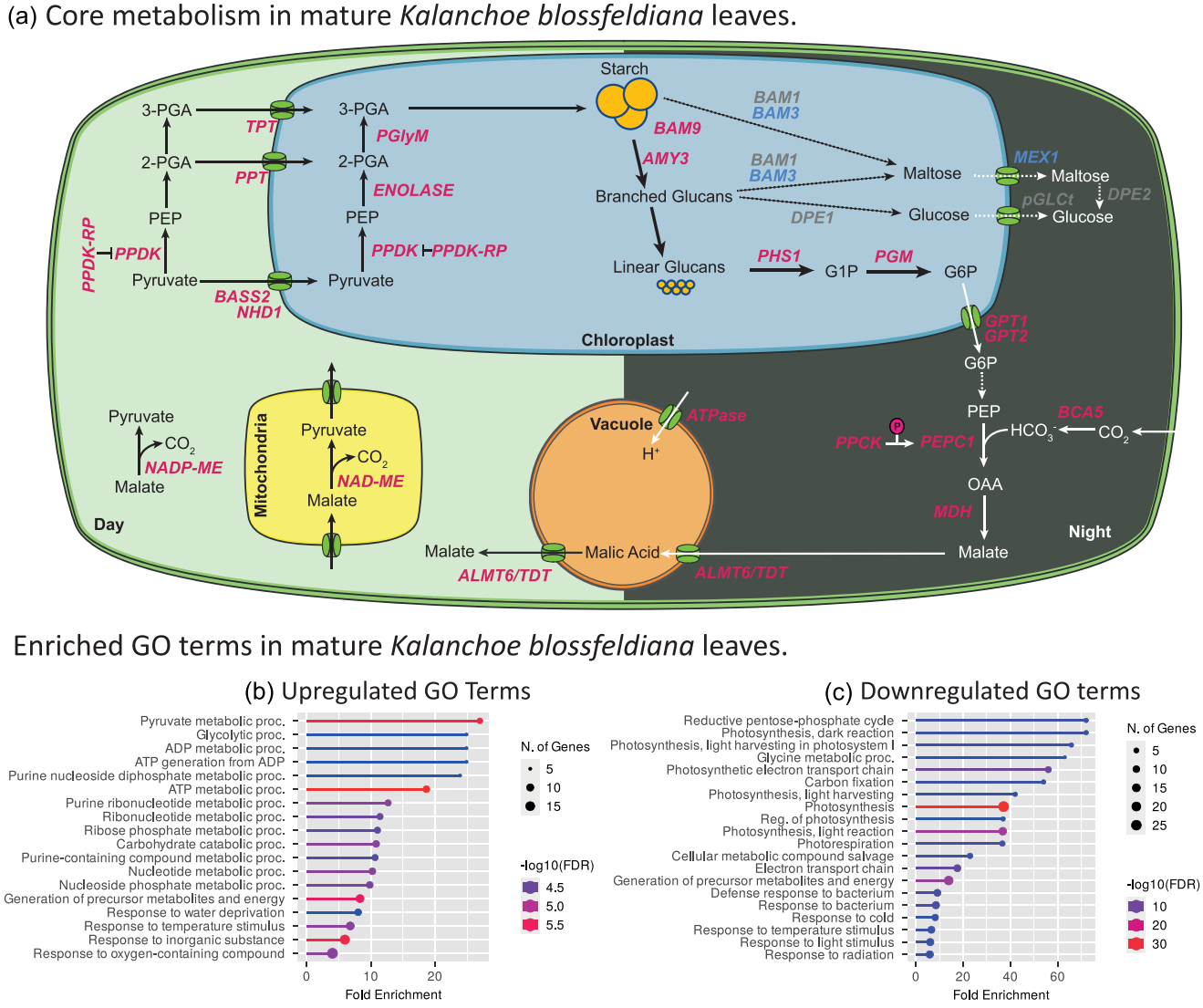

FIGURE 4 Changes in transcript abundance of key genes involved in Crassulacean acid metabolism (CAM) in the mature CAM *Kalanchoë blossfeldiana* leaves compared to young C_3 leaves at dusk (Panel A). Upregulated genes (red), downregulated genes (blue) and genes with no change in expression (grey) are shown in italics. Bold lines indicate the upregulation of the phosphorolytic starch degradation pathway in the mature leaves. GO-term enrichment for top 100 genes (Panel B) or (Panel C) in the mature (3rd leaf pair from the bottom) compared to the young *Kalanchoë blossfeldiana* leaves (9th leaf pair from the bottom) sampled at dusk. Abbreviations shown include: *Phosphoglyceromutase (PGlyM); tonoplast dicarboxylate transporter (TDT); aluminium activated malate transporter (ALMT);bile acid sodium symporter 2 (BASS2); debranching enzyme (DBE); glucose-1-phosphate (G1P);glucose-6-phosphate (G6P); malate dehydrogenase (MDH); NAD(P)-dependent malic enzyme (NAD(P)-ME); sodium hydrogen antiporter 1 (NHD1); oxaloacetate (OAA); phosphoenolpyruvate (PEP); phosphoenolpyruvate carboxylase (PEPC); phosphoglycerate (PGA); pyruvate phosphate dikinase (PPDK); PPDK regulatory protein (PPDK-RP); PPT, phosphoenolpyruvate phosphate translocator; triosephosphate/phosphate translocator (TPT); AMY3 (α-amylase 3); BAM1 (β-amylase 1); BAM3 (β-amylase 3); Phosphoglucomutase (PGM);BCA5 (β-carbonic anhydrase 1); PPCK (phosphoenolpyruvate carboxylase kinase); DPE1 (chloroplastic disproportionating enzyme); PHS1 (chloroplastic α-glucan phosphorylase); cytosolic disproportionating enzyme (DPE2); MEX1 (maltose transporter); pGLCT (plastidic glucose transporter), GPT1 (glucose phosphate translocator 1);* and *GPT2 (glucose phosphate translocator 2).*

increase in expression in the CAM mature *K. blossfeldiana* leaves. Alongside this, transcript abundance of the PEPC-activating *PHOSPHOENOLPYRUVATE CARBOXYLASE KINASE 1/2* (*PPCK*, kblos_g22485) increased in the mature CAM *K. blossfeldiana* leaves. Together these results highlight the clear distinction between the young C_3 and CAM performing mature *K. blossfeldiana* leaves in terms of transcript abundance of key genes implicated in CAM.

3.5 | Switch to phosphorolytic starch degradation in CAM *K. blossfeldiana* leaves

Starch-storing CAM plants have been shown to re-route starch degradation via the phosphorolytic pathway instead of the hydrolytic pathway used in C_3 *Arabidopsis*, with starch broken down to form glucose-6-phosphate (G6P) that is exported from the chloroplasts to

ultimately regenerate PEP for the CAM cycle (Borland & Dodd, 2002; Neuhaus & Schulte, 1996). CAM was impaired in *K. fedtschenkoi PHS1* mutants unable to degrade starch with the phosphorolytic pathway (Ceusters et al., 2021). Following CAM induction in the facultative species *Mesembryanthemum crystallinum* increased G6P export from the chloroplasts has been reported along with increased transcript abundance of *GLUCOSE PHOSPHATE TRANSLOCATOR* (*GPT*), thought to export G6P from chloroplasts (Hausler et al., 2000; Kore-eda et al., 2013; Kore-eda & Kanai, 1997; Neuhaus & Schulte, 1996). During CAM induction, *K. blossfeldiana* also appears to upregulate the phosphorolytic starch degradation pathway with upregulation of *PHS1* and two *GPT* isoforms along with the down-regulation of the chloroplastic maltose exporter *MEX1* involved in the hydrolytic starch degradation pathway. Increases in transcript abundances of genes involved in the phosphorolytic starch degradation pathway were also observed in ABA-induced CAM induction in *T. triangulare* (Maleckova et al., 2019). This move to phosphorolytic starch degradation could be an effort to energetically balance the CAM cycle, as it allows for the provision of cytosolic ATP during the conversion of G6P to PEP (Ceusters et al., 2021; Shameer et al., 2018). The switch to phosphorolytic starch degradation across *M. crystallinum*, *K. blossfeldiana* and *K. fedtschenkoi* implies its importance to starch-storing CAM plants.

The phosphorolytic route of starch degradation in C_3 plants has been suggested to act in response to stress conditions when photorespiration is elevated (Weise et al., 2006; Zeeman et al., 2007, 2004). There is the intriguing possibility that the upregulation of phosphorolytic starch degradation in CAM species suggests that the evolution of CAM is founded on the 'hijacking' of stress-responsive aspects of metabolism that are present in C_3 plants.

3.6 | Age-induced CAM shows substantial overlap with abiotic stress-induced CAM

We hypothesized that age-induced CAM may be a 'cleaner' system for detecting CAM-related genes with fewer genes differentially expressed between young C_3 and mature CAM *K. blossfeldiana* leaves compared to species where CAM is induced in response to stressors such as drought (Wai et al., 2019) or following treatment with ABA (Maleckova et al., 2019). However, we observed a substantial overlap between age-induced CAM and the drought and ABA-CAM induction pathways. Genes involved in the GO-term 'Response to Water Deprivation' were significantly enriched in the top 100 upregulated genes in mature *K. blossfeldiana* leaves, including several genes involved in drought and ABA response. In *T. triangulare*, drought-induced an increase in both CAM and ABA genes, which subsequently returned to non-stressed expression levels after re-watering (Brilhaus et al., 2016). Exogenous ABA application has been shown to induce CAM in *T. triangulare* (Maleckova et al., 2019), *M. crystallinum* (Holtum & Winter, 1982) and *K. blossfeldiana* (Taybi et al., 1995). In the present work, there was an overlap between age and ABA-induced CAM, with the transcript upregulation of *9-CIS-EPOXYCAROTENOID DIOXYGENASE (NCED3)*, a key ABA biosynthesis gene, in the mature CAM *K. blossfeldiana* leaves. Genes reported to be upregulated in response to ABA and drought induction of CAM in *T. triangulare* (Brilhaus et al., 2016; Maleckova et al., 2019) and drought-induced CAM in *S. album* (Wai et al., 2019) were also upregulated in the mature *K. blossfeldiana* leaves. Five ABA-responsive protein phosphatases were upregulated, homologous to *Highly ABA-Induced PP2C Gene 2* and *3*, and were also observed in ABA-induced CAM induction in *T. triangulare* (Maleckova et al., 2019). Downstream of this, several ABA-responsive transcription factors were upregulated in the mature *K. blossfeldiana* leaves, such as *NUCLEAR FACTOR Y (NF-Y) SUBUNITS A9* and *B3*, which were also upregulated in drought (Brilhaus et al., 2016) and ABA (Maleckova et al., 2019) induced CAM in *T. triangulare* and may go onto regulate gene expression by binding to ABA-responsive elements. Interestingly, ABA signalling has also been shown to play a central role in the induction of C_4 photosynthesis in inducible aquatic C_4 species (Chen et al., 2014; Ueno, 1998; Wang et al., 2016). Furthermore, genes involved in auxin and ethylene signalling were also upregulated in both ABA-induced CAM in *T. triangulare* and age-induced CAM in mature *K. blossfeldiana* leaves.

Together this suggests there is overlap between the different CAM induction pathways, and they may convergently make use of drought and ABA response pathways, already present in C_3 species, to induce CAM rather than using an entirely novel 'CAM induction pathway'. Future work could utilize the ability to induce CAM in *K. blossfeldiana* by ageing, drought, ABA, or short days to further elucidate CAM induction and its evolution. The data also suggests ABA may mediate age-induced-CAM, but whether it is required and triggers CAM induction by accumulating as the leaves age remains to be confirmed. If so, it raises some interesting questions; did inducible CAM evolve by hijacking ABA and drought signalling pathways?; did the use of ABA signalling in drought-induced CAM open the door to CAM induction via ABA by other means such as ageing? Key questions remain, including how ABA signalling is linked to ageing in *K. blossfeldiana* and what role ABA plays in constitutive CAM.

4 | CONCLUSIONS

Here we have taken a plant species of interest, and determined its relevance and usefulness in tackling our biological questions, in the context of CAM research. We economically and accessibly sequenced and assembled the nuclear and chloroplast genomes of *K. blossfeldiana* and provided a young vs mature leaf comparative transcriptome sampled at dusk. By utilizing low-cost Illumina short reads (~200x) and Nanopore long reads (~20x) we have shown that it is possible to assemble a medium-sized polyploid plant genome economically and quickly to enable quantification of transcript abundance. A key cost saving was the use of the Nanopore Rapid Barcoding/Sequencing Kit, removing the requirement for expensive additional consumables and shortening sequencing procedures. All of this was carried out rapidly and for ~£3,300 in the absence of any core genomics facilities. (see a full breakdown of costs in Methods S1). This is significantly cheaper than methods including a combination of Nanopore and PacBio

sequencing and/or those using Hi-C chromosomal contact mapping for further scaffolding. Further improvements in cost and throughput of plant genome assembly may come from multiplexed sequencing with the Nanopore Rapid Barcoding Kit.

While this assembly has shown it is possible to reduce the cost of plant genome sequencing, we believe there are areas for improvement. We would not classify this genome as a reference genome, as we had to align it with another more contiguous genome to reduce its ploidy for downstream transcript analysis. Improvements here will come with increased sequencing throughput vs cost, which has already occurred within the last few years. We do, however, think this method would be useful when trying to 'fill in' missing genomes and sequencing whole clades, where a key pressure would be to sequence cheaply. In this example, we also used a complex plant, with a tetraploid genome, had the genome been highly homozygous and diploid it is likely that the output would have been closer to a reference genome without the need to use a relative species for scaffolding. Together we hope this work provides a guide to repeating plant sequencing projects with many other plant species that are of interest to other researchers.

This study has also demonstrated the merits of *K. blossfeldiana* as a comparative system for studying C_3 and CAM physiologies and metabolism within a plant and has identified substantial changes in the dusk transcriptome between young C_3 and mature CAM *K. blossfeldiana* leaves in response to age-induced CAM. Finally, we have shown that in the absence of abiotic stress, CAM induction still involves the engagement of drought and ABA response pathways; and we have identified an overlap between age-induced CAM in *K. blossfeldiana* and the drought and ABA CAM induction pathways of *T. triangulare* and *S. album*. Together this supports suggestions that CAM may have evolved from the 'hi-jacking' of stress-related metabolism and signalling, making use of genes and signalling pathways already present in C_3 plants.

AUTHOR CONTRIBUTIONS

DCT, BAM, AS, IBW, ES, ON, AMB and MK discussed and planned the work. DCT, BAM, AS, IBW, RP, IL, JT, MW conducted the experiments. DCT carried out the data analysis, created the figures and drafted the initial manuscript. DCT, BAM, AMB and MK redrafted the final manuscript.

ACKNOWLEDGEMENTS

This work was supported by studentships awarded by the Biotechnology and Biological Sciences Research Council (BBSRC) to DCT (2132439) and BAM (2462247).

CONFLICT OF INTEREST STATEMENT

The authors report no conflict of interest.

DATA AVAILABILITY STATEMENT

All sequencing data used for the *K. blossfeldiana* genome assembly and dusk young versus mature *K. blossfeldiana* RNA-seq were deposited at NCBI under Bio Project number PRJNA915757. The *K. blossfeldiana* genome assembly has been deposited to https://zenodo.org/record/8188997. A detailed walkthrough of the nuclear genome assembly method is available at Github. (github.com/dcowanturner/Kb_genome).

ORCID

Daniel Cowan-Turner https://orcid.org/0000-0002-5277-3342
Maxim V. Kapralov https://orcid.org/0000-0001-7966-0295
Anne M. Borland https://orcid.org/0000-0003-4560-9998

REFERENCES

Alonge, M., Lebeigle, L., Kirsche, M., Aganezov, S., Wang, X., Lippman, Z. B., Schatz, M. C., & Soyk, S. (2021). Automated assembly scaffolding elevates a new tomato system for high-throughput genome editing. *bioRxiv* 2021.11.18.469135. https://doi.org/10.1101/2021.11.18.469135

Alonge, M., Soyk, S., Ramakrishnan, S., Wang, X., Goodwin, S., Sedlazeck, F. J., Lippman, Z. B., & Schatz, M. C. (2019). RaGOO: Fast and accurate reference-guided scaffolding of draft genomes. *Genome Biology*, *20*, 1–17. https://doi.org/10.1186/s13059-019-1829-6

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Borland, A. M., & Dodd, A. N. (2002). Carbohydrate partitioning in crassulacean acid metabolism plants: Reconciling potential conflicts of interest. *Functional Plant Biology*, *29*, 707–716. https://doi.org/10.1071/PP01221

Borland, A. M., Griffiths, H., Hartwell, J., & Smith, J. A. C. (2009). Exploiting the potential of plants with crassulacean acid metabolism for bioenergy production on marginal lands. *Journal of Experimental Botany*, *60*, 2879–2896. https://doi.org/10.1093/jxb/erp118

Borland, A. M., Hartwell, J., Weston, D. J., Schlauch, K. A., Tschaplinski, T. J., Tuskan, G. A., Yang, X., & Cushman, J. C. (2014). Engineering crassulacean acid metabolism to improve water-use efficiency. *Trends in Plant Science*, *19*, 327–338. https://doi.org/10.1016/j.tplants.2014.01.006

Borland, A. M., Wullschleger, S. D., Weston, D. J., Hartwell, J., Tuskan, G. A., Yang, X. H., & Cushman, J. C. (2015). Climate-resilient agroforestry: Physiological responses to climate change and engineering of crassulacean acid metabolism (CAM) as a mitigation strategy. *Plant, Cell & Environment*, *38*, 1833–1849. https://doi.org/10.1111/pce.12479

Brilhaus, D., Bräutigam, A., Mettler-Altmann, T., Winter, K., & Weber, A. P. M. (2016). Reversible burst of transcriptional changes during induction of crassulacean acid metabolism in Talinum triangulare1[OPEN]. *Plant Physiology*, *170*, 102–122. https://doi.org/10.1104/pp.15.01076

Brilhaus, D., Denton, A. K., Maleckova, E., Reichel-Deland, V., & Weber, A. P. M. (2023). The Genome of *Talinum fruticosum*. *bioRxiv*, 2023.04.20.537669. https://doi.org/10.1101/2023.04.20.537669

Brulfert, J., Guerrier, D., & Queiroz, O. (1982). Photoperiodism and Crassulacean acid metabolism. *Planta*, *154*, 332–338. https://doi.org/10.1007/BF00393911

Brůna, T., Hoff, K. J., Lomsadze, A., Stanke, M., & Borodovsky, M. (2021). BRAKER2: Automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genomics and Bioinformatics*, *3*, lqaa108. https://doi.org/10.1093/nargab/lqaa108

Ceusters, N., Ceusters, J., Hurtado-Castano, N., Dever, L. V., Boxall, S. F., Kneřová, J., Waller, J. L., Rodick, R., van den Ende, W., Hartwell, J., & Borland, A. M. (2021). Phosphorolytic degradation of leaf starch via plastidic α-glucan phosphorylase leads to optimized plant growth and water use efficiency over the diel phases of Crassulacean acid

metabolism. *Journal of Experimental Botany*, *72*, 4419–4434. https://doi.org/10.1093/jxb/erab132

Chen, T., Zhu, X.-G., & Lin, Y. (2014). Major alterations in transcript profiles between C3–C4 and C4 photosynthesis of an amphibious species Eleocharis baldwinii. *Plant Molecular Biology*, *86*, 93–110. https://doi.org/10.1007/s11103-014-0215-8

DuBois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., & Smith, F. (1956). Colorimetric Method for Determination of Sugars and Related Substances. *Analytical Chemistry*, *28*(3), 350–356. https://doi.org/10.1021/ac60111a017

Dumschott, K., Schmidt, M. H. W., Chawla, H. S., Snowdon, R., & Usadel, B. (2020). Oxford Nanopore sequencing: New opportunities for plant genomics? *Journal of Experimental Botany*, *71*, 5313–5322. https://doi.org/10.1093/jxb/eraa263

Emms, D. M., & Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*, 1–14. https://doi.org/10.1186/s13059-019-1832-y

Ewels, P. A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M. U., Di Tommaso, P., & Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. *Nature Biotechnology*, *38*, 276–278. https://doi.org/10.1038/s41587-020-0439-x

Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., & Smit, A. F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences of the United States of America*, *117*, 9451–9457. https://doi.org/10.1073/pnas.1921046117

Fukasawa, Y., Ermini, L., Wang, H., Carty, K., & Cheung, M. S. (2020). LongQC: A quality control tool for third generation sequencing long read data. G3: Genes. *Genomes*, *Genetics*, *10*, 1193–1196. https://doi.org/10.1534/g3.119.400864

Gabriel, L., Hoff, K. J., Brůna, T., Borodovsky, M., & Stanke, M. (2021). TSEBRA: Transcript selector for BRAKER. *BMC Bioinformatics*, *22*, 566. https://doi.org/10.1186/s12859-021-04482-0

Gilman, I. S., Smith, J. A. C., Holtum, J. A. M., Sage, R. F., Silvera, K., Winter, K., & Edwards, E. J. (2023). The CAM Lineages of Planet Earth. *Annals of Botany*, *132*(4), 627–654. https://doi.org/10.1093/aob/mcad135

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research*, *40*, 1178–1186. https://doi.org/10.1093/nar/gkr944

Gowers, G. O. F., Vince, O., Charles, J. H., Klarenberg, I., Ellis, T., & Edwards, A. (2019). Entirely off-grid and solar-powered DNA sequencing of microbial communities during an ice cap traverse expedition. *Genes (Basel)*, *10*, 902. https://doi.org/10.3390/genes10110902

Guan, D., Guan, D., McCarthy, S. A., Wood, J., Howe, K., Wang, Y., Durbin, R., & Durbin, R. (2020). Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics*, *36*, 2896–2898. https://doi.org/10.1093/bioinformatics/btaa025

Haider, M. S., Barnes, J. D., Cushman, J. C., & Borland, A. M. (2012). A CAM- and starch-deficient mutant of the facultative CAM species Mesembryanthemum crystallinum reconciles sink demands by repartitioning carbon during acclimation to salinity. *Journal of Experimental Botany*, *63*, 1985–1996. https://doi.org/10.1093/jxb/err412

Han, S., Bi, D., Yi, R., Ding, H., Wu, L., & Kan, X. (2022). Plastome Evolution of Aeonium and Monanthes (Crassulaceae): Insights Into the Variation of plastomic tRNAs, and the Patterns of Codon Usage and Aversion. *Planta*, *256*(2), 35. https://doi.org/10.1007/s00425-022-03950-y

Hartwell, J., Dever, L. V., & Boxall, S. F. (2016). Emerging model systems for functional genomics analysis of Crassulacean acid metabolism. *Current Opinion in Plant Biology*, *31*, 100–108. https://doi.org/10.1016/j.pbi.2016.03.019

Hausler, R. E., Baur, B., Scharte, J., Teichmann, T., Eicks, M., Fischer, K. L., Flugge, U. I., Schubert, S., Weber, A., & Fischer, K. (2000). Plastidic metabolite transporters and their physiological functions in the inducible crassulacean acid metabolism plant Mesembryanthemum crystallinum. *Plant Journal*, *24*, 285–296. https://doi.org/10.1046/j.1365-313x.2000.00876.x

Heyduk, K., Ray, J. N., Ayyampalayam, S., Moledina, N., Borland, A., Harding, S. A., Tsai, C. J., & Leebens-Mack, J. (2019). Shared expression of crassulacean acid metabolism (CAM) genes pre-dates the origin of CAM in the genus yucca. *Journal of Experimental Botany*, *70*, 6597–6609. https://doi.org/10.1093/jxb/erz105

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2017). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. https://doi.org/10.1093/molbev/msx281

Holtum, J. A. M., & Winter, K. (1982). Activity of enzymes of carbon metabolism during the induction of Crassulacean acid metabolism in Mesembryanthemum crystallinum L. *Planta*, *155*, 8–16. https://doi.org/10.1007/BF00402925

Jackman, S. D., Vandervalk, B. P., Mohamadi, H., Chu, J., Yeo, S., Hammond, S. A., Jahesh, G., Khan, H., Coombe, L., Warren, R. L., & Birol, I. (2017). ABySS 2. 0: Resource-efficient assembly of large genomes using a bloom filter effect of bloom filter false positive rate. *Genome Research*, *27*, 768–777. https://doi.org/10.1101/gr.214346.116.Freely

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Keb-Llanes, M., González, G., Chi-Manzanero, B., & Infante, D. (2002). A rapid and simple method for small-scale DNA extraction in Agavaceae and other tropical plants. *Plant Molecular Biology Reporter*, *20*, 299. https://doi.org/10.1007/BF02782465

Klein, D.-P., Shtein, R., Nusbaumer, L., & Callmander, M. W. (2021). Kalanchoe darainensis (Crassulaceae), a New Species From Northeastern Madagascar. *Candollea*, *76*(1). https://doi.org/10.15553/c2021v761a12

Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*, *37*, 540–546. https://doi.org/10.1038/s41587-019-0072-8

Kore-eda, S., & Kanai, R. (1997). Induction of glucose 6-phosphate transport activity in chloroplasts of Mesembryanthemum crystallinum by the C-3-CAM transition. *Plant & Cell Physiology*, *38*, 895–901. https://doi.org/10.1093/oxfordjournals.pcp.a029249

Kore-eda, S., Nozawa, A., Okada, Y., Takashi, K., Azad, M. A. K., Ohnishi, J., Nishiyama, Y., & Tozawa, Y. (2013). Characterization of the Plastidic phosphate translocators in the inducible Crassulacean acid metabolism plant Mesembryanthemum crystallinum. *Bioscience, Biotechnology, and Biochemistry*, *77*, 1511–1516. https://doi.org/10.1271/bbb.130174

Kriventseva, E. V., Kuznetsov, D., Tegenfeldt, F., Manni, M., Dias, R., Simão, F. A., & Zdobnov, E. M. (2019). OrthoDB v10: Sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Research*, *47*, D807–D811. https://doi.org/10.1093/nar/gky1053

Kyriakidou, M., Tai, H. H., Anglin, N. L., Ellis, D., & Strömvik, M. V. (2018). Current strategies of polyploid plant genome sequence assembly. *Frontiers in Plant Science*, *9*, 1660. https://doi.org/10.3389/fpls.2018.01660

Lerman, J. C., & Queiroz, O. (1974). Carbon fixation and isotope discrimination by a Crassulacean plant: Dependence on the photoperiod. *Science*, *183*, 1207–1209. https://doi.org/10.1126/science.183.4130.1207

Lewin, H. A., Robinson, G. E., Kress, W. J., Baker, W. J., Coddington, J., Crandall, K. A., Durbin, R., Edwards, S. V., Forest, F., Gilbert, M. T. P., Goldstein, M. M., Grigoriev, I. V., Hackett, K. J., Haussler, D., Jarvis, E. D., Johnson, W. E., Patrinos, A., Richards, S., Castilla-Rubio, J. C., … Zhang, G. (2018). Earth BioGenome project: Sequencing

life for the future of life. *Proceedings of the National Academy of Sciences*, *115*, 4325–4333. https://doi.org/10.1073/pnas.1720115115

Liao, Z., Chen, M., Guo, L., Gong, Y., Tang, F., Sun, X., & Tang, K. (2004). Rapid Isolation of High-Quality Total RNA From Taxus and Ginkgo. *Preparative Biochemistry & Biotechnology*, *34*, 209–214. https://doi.org/10.1081/PB-200026790

Lim, S. D., Lee, S., Choi, W.-G., Yim, W. C., & Cushman, J. C. (2019). Laying the Foundation for Crassulacean Acid Metabolism (CAM) Biodesign: Expression of the C4 Metabolism Cycle Genes of CAM in Arabidopsis. *Frontiers in Plant Science*, *10*, 101. https://doi.org/10.3389/fpls.2019.00101

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*, 550. https://doi.org/10.1186/s13059-014-0550-8

Maleckova, E. (2020). *Regulation of crassulacean acid metabolism (CAM) in the facultative CAM species Talinum triangulare*. Dissertation (Vol. 2020). Heinrich-Heine-Universität.

Maleckova, E., Brilhaus, D., Wrobel, T. J., & Weber, A. P. M. (2019). Transcript and metabolite changes during the early phase of abscisic acid-mediated induction of crassulacean acid metabolism in Talinum triangulare. *Journal of Experimental Botany*, *70*, 6581–6596. https://doi.org/10.1093/jxb/erz189

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., & Zdobnov, E. M. (2021). BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology and Evolution*, *38*, 4647–4654. https://doi.org/10.1093/molbev/msab199

Messerschmid, T. F. E., Klein, J. T., Kadereit, G., & Kadereit, J. W. (2020). Linnaeus's Folly – Phylogeny, Evolution and Classification of Sedum (Crassulaceae) and Crassulaceae Subfamily Sempervivoideae. *Taxon*, *69*(5), 892–926. Portico. https://doi.org/10.1002/tax.12316

Michael, T. P., Jupe, F., Bemm, F., Motley, S. T., Sandoval, J. P., Lanz, C., Loudet, O., Weigel, D., & Ecker, J. R. (2018). High contiguity Arabidopsis thaliana genome assembly with a single nanopore flow cell. *Nature Communications*, *9*, 541. https://doi.org/10.1038/s41467-018-03016-2

Ming, R., VanBuren, R., Wai, C. M., Tang, H., Schatz, M. C., Bowers, J. E., Lyons, E., Wang, M. L., Chen, J., Biggers, E., Zhang, J., Huang, L., Zhang, L., Miao, W., Zhang, J., Ye, Z., Miao, C., Lin, Z., Wang, H., … Yu, Q. (2015). The pineapple genome and the evolution of CAM photosynthesis. *Nature Genetics*, *47*, 1435–1442. https://doi.org/10.1038/ng.3435

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., Lanfear, R., & Teeling, E. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Moseley, R. C., Tuskan, G. A., & Yang, X. (2019). Comparative genomics analysis provides new insight into molecular basis of stomatal movement in Kalanchoë fedtschenkoi. *Frontiers in Plant Science*, *10*, 292. https://doi.org/10.3389/fpls.2019.00292

Nelson, E. A., & Sage, R. F. (2008). Functional constraints of CAM leaf anatomy: Tight cell packing is associated with increased CAM function across a gradient of CAM expression. *Journal of Experimental Botany*, *59*, 1841–1850. https://doi.org/10.1093/jxb/erm346

Neuhaus, H. E., & Schulte, N. (1996). Starch degradation in chloroplasts isolated from C-3 or CAM (crassulacean acid metabolism)-induced Mesembryanthemum crystallinum L. *Biochemical Journal*, *318*, 945–953. https://doi.org/10.1042/bj3180945

Niechayev, N. A., Mayer, J. A., & Cushman, J. C. (2023). Developmental dynamics of crassulacean acid metabolism (CAM) in Opuntia ficus-indica. *Annals of Botany*, *132*, 869–879. https://doi.org/10.1371/journal.pgen.1008209

Ossa, P. G., Moreno, A. A., Orellana, D., Toro, M., Carrasco-Valenzuela, T., Riveros, A., Meneses, C. C., Nilo-Poyanco, R., & Orellana, A. (2022). Transcriptomic Analysis of the C3-CAM transition in Cistanthe longiscapa, a Drought Tolerant Plant in the Atacama Desert. *bioRxiv*, 2022.03.16.484649. https://doi.org/10.1101/2022.03.16.484649

Ota, K. (1988). Stimulation of CAM photosynthesis in Kalanchoë blossfeldiana by transferring to nitrogen-deficient conditions. *Plant Physiology*, *87*, 454–457. https://doi.org/10.1104/pp.87.2.454

Ou, S., Chen, J., & Jiang, N. (2018). Assessing genome assembly quality using the LTR assembly index (LAI). *Nucleic Acids Research*, *46*, e126. https://doi.org/10.1093/nar/gky730

Paterson, A. H., Bowers, J. E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., Poliakov, A., Schmutz, J., Spannagl, M., Tang, H., Wang, X., Wicker, T., Bharti, A. K., Chapman, J., Feltus, F. A., Gowik, U., … Rokhsar, D. S. (2009). The Sorghum bicolor genome and the diversification of grasses. *Nature*, *457*, 551–556. https://doi.org/10.1038/nature07723

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, *14*, 417–419. https://doi.org/10.1038/nmeth.4197

Ping, C. Y., Chen, F. C., Cheng, T. C., Lin, H. L., Lin, T. S., Yang, W. J., & Lee, Y. I. (2018). Expression profiles of phosphoenolpyruvate carboxylase and phosphoenolpyruvate carboxylase kinase genes in Phalaenopsis, implications for regulating the performance of Crassulacean acid metabolism. *Frontiers in Plant Science*, *9*, 1587. https://doi.org/10.3389/fpls.2018.01587

Pucker, B., Irisarri, I., de Vries, J., & Xu, B. (2022). Plant genome sequence assembly in the era of long reads: Progress, challenges and future directions. *Quantitative Plant Biology*, *3*, e5. https://doi.org/10.1017/qpb.2021.18

Pucker, B., Kleinbölting, N., & Weisshaar, B. (2021). Large scale genomic rearrangements in selected Arabidopsis thaliana T-DNA lines are caused by T-DNA insertion mutagenesis. *BMC Genomics*, *22*, 599. https://doi.org/10.1186/s12864-021-07877-8

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033

Roach, M. J., Schmidt, S., & Borneman, A. R. (2018). Purge Haplotigs: Allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics*, *19*, 460. https://doi.org/10.1186/s12859-018-2485-7

Scott, A. D., Zimin, A. V., Puiu, D., Workman, R., Britton, M., Zaman, S., Caballero, M., Read, A. C., Bogdanove, A. J., Burns, E., Wegrzyn, J., Timp, W., Salzberg, S. L., & Neale, D. B. (2020). A reference genome sequence for giant sequoia. G3: Genes. *Genomes*, *Genetics*, *10*, 3907–3919. https://doi.org/10.1534/g3.120.401612

Shameer, S., Baghalian, K., Cheung, C. Y. M., Ratcliffe, R. G., & Sweetlove, L. J. (2018). Computational analysis of the productivity potential of CAM. *Nature Plants*, *4*, 165–171. https://doi.org/10.1038/s41477-018-0112-2

Smith, G. F., & Figueiredo, E. (2018). The infrageneric classification and nomenclature of Kalanchoe Adans. (Crassulaceae), with special reference to the southern African species. *Bradleya*, *36*, 162–172. https://doi.org/10.25223/brad.n36.2018.a10

Smith, G. F., & Shtein, R. (2022). A review of horticulturally desirable characters in Kalanchoe (Crassulaceae subfam. Kalanchooideae): Variable and deviating vegetative and reproductive morphologies useful in breeding Programmes. *Haseltonia*, *28*, 106–119. https://doi.org/10.2985/026.028.0110

Tarailo-Graovac, M., & Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. *Current Protocols in Bioinformatics*, *25*, 4.10.1-4.10.14. https://doi.org/10.1002/0471250953.bi0410s25

Taybi, T., Sotta, B., Gehrig, H., Güçlü, S., Kluge, M., & Brulfert, J. (1995). Differential effects of abscisic acid on phosphoenolpyruvate carboxylase and CAM operation in Kalanchoë blossfeldiana. *Botanica Acta*:

Journal of the German Botanical Society, *108*, 240–246. https://doi.org/10.1111/j.1438-8677.1995.tb00856.x

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, *408*, 796–815. https://doi.org/10.1038/35048692

Tian, X., Guo, J., Zhou, X., Ma, K., Ma, Y., Shi, T., & Shi, Y. (2021). Comparative and Evolutionary Analyses on the Complete Plastomes of Five Kalanchoe Horticultural Plants. *Frontiers in Plant Science*, *12*, 705874. https://doi.org/10.3389/fpls.2021.705874

Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E. S., Fischer, A., Bock, R., & Greiner, S. (2017). GeSeq - versatile and accurate annotation of organelle genomes. *Nucleic Acids Research*, *45*(W1), W6–W11. https://doi.org/10.1093/nar/gkx391

Ueno, O. (1998). Induction of Kranz anatomy and C4-like biochemical characteristics in a submerged amphibious plant by abscisic acid. *Plant Cell*, *10*, 571–583. https://doi.org/10.1105/tpc.10.4.571

van Voorst, A., & Arends, J. C. (1982). The origin and chromosome numbers of cultivars of Kalanchoe blossfeldiana V on P oelln.: Their history and evolution. *Euphytica*, *31*, 573–584. https://doi.org/10.1007/BF00039195

Wai, C. M., Weise, S. E., Ozersky, P., Mockler, T. C., Michael, T. P., & Vanburen, R. (2019). Time of903 day and network reprogramming during drought induced CAM photosynthesis in904 Sedum album. *PLoS Genet 15*. https://doi.org/10.1371/journal.pgen.1008209

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., & Earl, A. M. (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE*, *9*, e112963. https://doi.org/10.1371/journal.pone.0112963

Wang, W., Das, A., Kainer, D., Schalamun, M., Morales-Suarez, A., Schwessinger, B., & Lanfear, R. (2020). The draft nuclear genome assembly of Eucalyptus pauciflora: A pipeline for comparing de novo assemblies. *GigaScience*, *9*, 1–12. https://doi.org/10.1093/gigascience/giz160

Wang, P., Vlad, D., & Langdale, J. A. (2016). Finding the genes to build C4 rice. *Current Opinion in Plant Biology*, *31*, 44–50. https://doi.org/10.1016/j.pbi.2016.03.012

Weise, S. E., Schrader, S. M., Kleinbeck, K. R., & Sharkey, T. D. (2006). Carbon balance and circadian regulation of hydrolytic and phosphorolytic breakdown of transitory starch. *Plant Physiology*, *141*, 879–886. https://doi.org/10.1104/pp.106.081174

Wickell, D., Kuo, L. Y., Yang, H. P., Dhabalia Ashok, A., Irisarri, I., Dadras, A., de Vries, S., de Vries, J., Huang, Y. M., Li, Z., Barker, M. S., Hartwick, N. T., Michael, T. P., & Li, F. W. (2021). Underwater CAM photosynthesis elucidated by Isoetes genome. *Nature Communications*, *12*, 6348. https://doi.org/10.1038/s41467-021-26644-7

Winter, K., Garcia, M., & Holtum, J. A. M. (2008). On the nature of facultative and constitutive CAM: Environmental and developmental control of CAM expression during early growth of Clusia, Kalanchoë, and Opuntia. *Journal of Experimental Botany*, *59*, 1829–1840. https://doi.org/10.1093/jxb/ern080

Yang, X., Cushman, J. C., Borland, A. M., Edwards, E. J., Wullschleger, S. D., Tuskan, G. A., Owen, N. A., Griffiths, H., Smith, J. A. C., De Paoli, H. C., Weston, D. J., Cottingham, R., Hartwell, J., Davis, S. C., Silvera, K., Ming, R., Schlauch, K., Abraham, P., Stewart, J. R., … Holtum, J. A. M. (2015). A roadmap for research on crassulacean acid metabolism (CAM) to enhance sustainable food and bioenergy production in a hotter, drier world. *The New Phytologist*, *207*, 491–504. https://doi.org/10.1111/nph.13393

Yang, X., Hu, R., Yin, H., Jenkins, J., Shu, S., Tang, H., Liu, D., Weighill, D. A., Cheol Yim, W., Ha, J., Heyduk, K., Goodstein, D. M., Guo, H. B., Moseley, R. C., Fitzek, E., Jawdy, S., Zhang, Z., Xie, M., Hartwell, J., … Tuskan, G. A. (2017). The Kalanchoë genome provides insights into convergent evolution and building blocks of crassulacean acid metabolism. *Nature Communications*, *8*, 1899. https://doi.org/10.1038/s41467-017-01491-7

Zambrano, A. B., Lawson, T., Olmos, E., Fernández-García, N., & Borland, A. M. (2014). Leaf anatomical traits which accommodate the facultative engagement of crassulacean acid metabolism in tropical trees of the genus Clusia. *Journal of Experimental Botany*, *65*, 3513–3523. https://doi.org/10.1093/jxb/eru022

Zeeman, S. C., Delatte, T., Messerli, G., Umhang, M., Stettler, M., Mettler, T., Streb, S., Reinhold, H., & Kotting, O. (2007). Starch breakdown: Recent discoveries suggest distinct pathways and novel mechanisms. *Functional Plant Biology*, *34*, 465–473. https://doi.org/10.1071/FP06313

Zeeman, S. C., Thorneycroft, D., Schupp, N., Chapple, A., Weck, M., Dunstan, H., Haldimann, P., Bechtold, N., Smith, A. M., & Smith, S. M. (2004). Plastidial alpha-glucan phosphorylase is not required for starch degradation in arabidopsis leaves but has a role in the tolerance of abiotic stress. *Plant Physiology*, *135*, 849–858. https://doi.org/10.1104/pp.103.032631

Zimin, A. V., Puiu, D., Luo, M. C., Zhu, T., Koren, S., Marçais, G., Yorke, J. A., Dvořák, J., & Salzberg, S. L. (2017). Hybrid assembly of the large and highly repetitive genome of Aegilops tauschii, a progenitor of bread wheat, with the MaSuRCA mega-reads algorithm. *Genome Research*, *27*, 787–792. https://doi.org/10.1101/gr.213405.116

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Cowan-Turner, D., Morris, B. A., Sandéhn, A., Bernacka-Wojcik, I., Stavrinidou, E., Powell, R. F., Leitch, I. J., Taylor, J., Walker, M., Nwokeocha, O., Kapralov, M. V., & Borland, A. M. (2024). Sequencing complex plants on a budget: The development of *Kalanchoë blossfeldiana* as a C_3, CAM comparative tool. *Plants, People, Planet*, 1–15. https://doi.org/10.1002/ppp3.10517