Shared characteristics underpinning C₄ leaf maturation derived from analysis of multiple C₃ and C₄ species of Flaveria

Britta M.C. Kümpers*,†, Steven J. Burgess*, Ivan Reyna-Llorens, Richard Smith-Unna, Chris Boursnell and Julian M. Hibberd‡

Department of Plant Sciences, Downing Street, University of Cambridge, Cambridge CB2 3EA, UK

* These authors contributed equally to this work.
† Present address: Centre for Plant Integrative Biology, School of Biosciences, University of Nottingham, Loughborough LE12 5RD, UK.
‡ Correspondence: jmh65@cam.ac.uk

Received 4 November 2016; Editorial decision 7 December 2016; Accepted 13 December 2016

Editor: Susanne von Caemmerer, Australian National University

Abstract

Most terrestrial plants use C₃ photosynthesis to fix carbon. In multiple plant lineages a modified system known as C₄ photosynthesis has evolved. To better understand the molecular patterns associated with induction of C₄ photosynthesis, the genus Flaveria that contains C₃ and C₄ species was used. A base to tip maturation gradient of leaf anatomy was defined, and RNA sequencing was undertaken along this gradient for two C₃ and two C₄ Flaveria species. Key C₄ traits including vein density, mesophyll and bundle sheath cross-sectional area, chloroplast ultrastructure, and abundance of transcripts encoding proteins of C₄ photosynthesis were quantified. Candidate genes underlying each of these C₄ characteristics were identified. Principal components analysis indicated that leaf maturation and the photosynthetic pathway were responsible for the greatest amount of variation in transcript abundance. Photosynthesis genes were over-represented for a prolonged period in the C₄ species. Through comparison with publicly available data sets, we identify a small number of transcriptional regulators that have been up-regulated in diverse C₄ species. The analysis identifies similar patterns of expression in independent C₄ lineages and so indicates that the complex C₄ pathway is associated with parallel as well as convergent evolution.

Key words: C₄ leaf anatomy, C₄ photosynthesis, convergent evolution, Flaveria, gene expression, parallel evolution, RNA-seq.

Introduction

Photosynthesis powers life on earth by harvesting energy from sunlight to fix carbon dioxide. In seed plants, three types of photosynthesis known as C₃, C₄, and Crassulacean acid metabolism (CAM) have been defined. C₃ photosynthesis is considered the ancestral form and is found in the majority of species. All three photosynthetic pathways use Ribulose-1,5-Bisphosphate Carboxylase Oxygenase (RuBisCO) to catalyse CO₂ fixation in the Calvin–Benson–Bassham (CBB) cycle and, in so doing, two molecules of the three-carbon molecule 3-phosphoglycerate are formed. However, RuBisCO is not completely substrate specific, catalysing an oxygenation reaction that produces 2-phosphoglycolate as well as 3-phosphoglycerate (Bowes et al., 1971). As phosphoglycolate is toxic and removes carbon from the CBB cycle, it is rapidly metabolized by the photorespiratory pathway to ensure it does not accumulate, but also to retrieve carbon. Photorespiration
uses energy inputs and is not completely effective in terms of carbon retrieval so some CO₂ is lost (Bauwe et al., 2010). The rate of oxygenation at the active site of RuBisCO increases with temperature, and so in lower latitudes multiple plant lineages have evolved either CAM or C₃ photosynthesis (Sage et al., 2011), which in both cases initially fix CO₂ by an alternative carboxylase, and then subsequently release high concentrations of CO₂ around RuBisCO to limit oxygenation.

Almost all C₄ plants use a two-celled system in which phosphoenolpyruvate carboxylase (PEPC) initially fixes carbon in mesophyll (M) cells, and then, after release of CO₂ in the bundle sheath (BS) cells, RuBisCO re-fixes CO₂ in the CBB cycle. The production of CO₂ within the BS leads to high concentrations of CO₂ around RuBisCO and minimizes oxygenation. PEPC generates oxaloacetate, which is rapidly metabolized to malate and/or aspartate prior to their diffusion to the BS. Three C₄ acid decarboxylase enzymes known as NADP-malic enzyme (NADP-ME), NAD-malic enzyme (NAD-ME), and phosphoenolpyruvate carboxykinase (PEPKC) have been shown to release CO₂ around RuBisCO and, for many years, based on their relative abundance, these decarboxylase enzymes were used to define three C₄ biochemical subtypes. More recently it has been shown that this division into three types is less rigid than previously thought (Furbank, 2011; Bellasio and Griffiths, 2014; Wang et al., 2014). The use of distinct decarboxylases in separate C₄ lineages indicates that C₄ photosynthesis is underpinned by convergent evolution. In addition to the 60 groups of plants known to have evolved the C₄ pathway (Sage et al., 2011), there are some genera containing both C₃ and C₄ species. Probably the most studied of these clades with closely related C₃ and C₄ species is the genus Flaveria found in the Asteraceae (Drinovich et al., 1998; Gowik et al., 2011; Schulze et al., 2013).

Most species of Flaveria are native to Central and North America and grow as annual or perennial herbs or shrubs with decussate leaves (Powell, 1978). C₄ species of Flaveria show high activities of the NADP-ME C₄ acid decarboxylase in chloroplasts of the BS, and leaf anatomy conforms to the Atriplicoid type (McKown and Dengler, 2007). Phylogenetic reconstruction of this group has been conducted using morphological as well as life history and gene sequence data for 21 of the 23 known species (McKown et al., 2005; Lyu et al., 2015). The consensus from this work is that the ancestral condition in Flaveria is C₃ photosynthesis (McKown et al., 2005; Lyu et al., 2015) and that the occurrence of multiple C₃ and C₄ species within the same genus provides an interesting system to study processes associated with the C₄ phenotype (see Supplementary Fig. S1 at JXB online).

In this study, we used two pairs of C₃ and C₄ Flaveria species, and set out to link the gradual maturation of C₃ and C₄ characteristics in leaves to underlying alterations in transcript abundance. By linking the development of the C₄ phenotype to changes in gene expression in two C₄ species, and comparing these findings with equivalent data from two C₃ species in the same genus, we aimed to identify common traits associated with C₄ photosynthesis, and to remove species-specific characteristics from our data sets. Using the maturing leaf as a dynamic system, we show that in both C₄ species studied, the induction of Kranz anatomy occurs along a base to tip developmental gradient in leaves of >2 cm length. We sampled this maturation gradient and undertook RNA sequencing to correlate the underlying patterns of gene expression with anatomical development.

Materials and methods

Plant growth

Flaveria bidentis (L.) Kuntze, Flaveria pringlei Gandoger, Flaveria robusta Rose, and Flaveria trinervia (Spreng.) C. Mohr were grown in a glasshouse on the roof of the Department of Plant Sciences, Cambridge. Temperature was maintained above 20 °C and supplemental lighting was provided to ensure at least 250–350 µmol m⁻² s⁻¹ photon flux density for 16 h d⁻¹. Seeds were sown directly onto soil (Levington’s M3 potting compost; Scotts Miracle-Gro Company, Godalming, UK) in covered pots as seeds need high humidity for germination. Covers were removed when seedlings were 1 cm high.

Analysis of leaf anatomy

Samples of 4 mm² to 1 cm² were fixed in 4% (w/v) formaldehyde at 4 °C overnight and then placed on ice and dehydrated prior to being placed in 100% (v/v) ethanol, followed by 1:1 ethanol/Technovit mix and then 100% Technovit 7100 (Heraeus Kulzer, Germany). Samples were subsequently left in Technovit solution plus hardener I (1 g of hardener per 100 ml) for at least 1 h. Disposable plastic resin embedding moulds (Agar Scientific, UK) were filled with a mixture of Technovit plus hardener I and II (15 ml of Technovit plus hardener I were mixed with 1 ml of hardener II). Samples were arranged within the embedding moulds which were then covered with unstretched Parafilm® M to seal them from air, and left to harden overnight. Samples were removed, heated to 80 °C, and trimmed for sectioning. Sections of 2 µm thickness were produced using a Thermo Scientific Microm HM340E microtome. Ribbons were mounted onto SuperFrost® white microscope slides (VWR, Leuven, The Netherlands), left to dry, and then stained with 0.1% (w/v) toluidine blue. All sections were analysed with a BX41 light microscope (Olympus, Center Valley, PA, USA), usually using the bright-field setting.

To clear leaves, they were placed into 70% (v/v) ethanol and heated to 60 °C. The next day samples were placed in 5% (w/v) NaOH for ~15 min to clear leaves further, and then mounted in water and analysed by light microscopy.

To quantify leaf anatomical characteristics, Photoshop CS5 was used. The program was calibrated with scale bars and the lasso tool was used to measure cell area for both BS and M cells. Measurements of M cells were taken between BS cells, with 30 BS cells arranged within the embedding moulds which were then covered with SuperFrost® white microscope slides (VWR, Leuven, The Netherlands) and dehydrated prior to being embedded in a mixture of Technovit plus hardener I and II (15 ml of Technovit plus hardener I were mixed with 1 ml of hardener II). Samples were arranged within the embedding moulds which were then covered with unstretched Parafilm® M to seal them from air, and left to harden overnight. Samples were removed, heated to 80 °C, and trimmed for sectioning. Sections of 2 µm thickness were produced using a Thermo Scientific Microm HM340E microtome. Ribbons were mounted onto SuperFrost® white microscope slides (VWR, Leuven, The Netherlands), left to dry, and then stained with 0.1% (w/v) toluidine blue. All sections were analysed with a BX41 light microscope (Olympus, Center Valley, PA, USA), usually using the bright-field setting.

To clear leaves, they were placed into 70% (v/v) ethanol and heated to 60 °C. The next day samples were placed in 5% (w/v) NaOH for ~15 min to clear leaves further, and then mounted in water and analysed by light microscopy.

To quantify leaf anatomical characteristics, Photoshop CS5 was used. The program was calibrated with scale bars and the lasso tool was used to measure cell area for both BS and M cells. Measurements of M cells were taken between BS cells, with 30 BS and 30 M cells being measured per leaf section. This was done for all six leaf stages and for three biological replicates for each species. Images derived from leaf sections were assembled using Photoshop. The background of sections was averaged and in some cases the contrast was increased to improve visibility of cell borders. Vein density was determined using images of cleared leaves. In mature sections of the leaf there was no ambiguity in collecting these data, whereas in the most basal sections of some leaves, the estimates may be underestimated. This analysis was used to calculate vein density. The length of the veins was measured using Q-Capture Pro®, and the length of cells was measured using Q-Capture Pro®.

Samples for TEM were processed by the Multi-Imaging Centre in the Department of Physiology, Development and Neuroscience (Cambridge). Sections of ~1 mm thickness of all six stages along a developing leaf were used for embedding. Tissues were fixed in 4% (w/v) glutaraldehyde in 0.1 M HEPES buffer with a pH of 7.4 for 12 h at 4 °C. Subsequently they were rinsed in 0.1 M HEPES buffer five times, and then treated with 1% osmium ferricyanide at room
temperature for 2 h and rinsed in deionized water five times before being treated with 2% (w/v) uranyl acetate in 0.05 M maleate buffer with a pH of 5.5 for 2 h at room temperature. They were rinsed again in deionized water and dehydrated in an ascending series of ethanol solutions from 70% to 100% (v/v). This was followed by treatment with two changes of dry acetonitrile and infiltration with Quetol 651 epoxy resin. Sections 50–70 nm thick were cut with a Leica Ultracut UCT and stained with lead citrate and uranyl acetate. Images were taken with a Tecnai G2 transmission electron microscope (FEI, Hillsboro, USA) and operated at 120 kV using an AMT XR60B digital camera (Advanced Microscopy Techniques, Woburn, USA) running Deben software (Deben UK Limited, Bury St. Edmunds, UK). Images were taken at ×1700, ×3500, and ×5000 magnification.

RNA extraction and sequencing

Deep sequencing was carried out to analyse transcriptomes associated with the leaf developmental gradient defined by the anatomical analysis. Opposite leaves from the third leaf pair after the cotyledons were harvested in the morning between 10:00 h and 11:00 h (4 h after the start of the photoperiod in the glasshouse). The leaves were harvested when they had reached 1.8–2 cm length (from base to tip). Leaves were measured on an ice-cold glass plate and cut into six portions of equal length (3–3.3 mm). Each sample was then immediately placed into liquid nitrogen. Sections of 12 leaves were collected before extraction of RNA using the Qiagen Plant RNeasy Mini Kit. The optional DNA digestion step with RNase-free DNase was performed. Three biological replicates were analysed for each of the four Flaveria species. Total leaf RNA was sent to the Genomics & Transcriptomics Laboratory (GTL) in Düsseldorf (Germany) for paired-end sequencing. RNA samples were prepared following the TrueSeq RNA sample preparation v2 guide (Revision F) and sequenced using an Illumina/Solexa HiSeq 2000 machine. Prior to assembly, reads were trimmed using Trimmomatic v0.32 (Bolger et al., 2014), corrected using BayesHammer v2.5.1 (Nikolenko et al., 2013), and then digitally normalized using the khmer package (Crusoe et al., 2015). SOAPdenovo-trans (Xie et al., 2014), IDBA- tran v1.0.13 (Peng et al., 2013), and SGA v0.10.12 (Simpson and Durbin, 2012) were used for de novo assembly. For each species, the individual assemblies were combined and clustered using CD-HIT-EST (Simpson and Durbin, 2012; Li and Godzik, 2006). Annotation of contigs used Conditional Reciprocal Best Blast (github.com/eboursnell/crb-blast) to the Arabidopsis thaliana peptide sequences as previously described (Aubry et al., 2014). For principal components analysis, Z-scores were calculated from transcripts per million (TPM) reads from for each gene across the four species. Expression patterns of genes of interest were analysed using a custom-made R script to extract the data of interest and to visualize expression patterns. Gene Ontology (GO) term analysis was performed according to Burgess et al. (2016).

Results and Discussion

Leaf maturation in C3 and C4 Flaveria species

We first confirmed that fully expanded mature leaves of C3 F. pringlei and F. robusta as well as C4 F. bidentis and F. trinervia showed characteristic C3 and C4 anatomy under our growth conditions. In all cases, the third leaf pair was chosen as the first and second leaf pairs have previously been described as juvenile (McKown and Dengler, 2009). Mature third leaves of C3 and C4 Flaveria ranged from 6 cm to 10 cm in length, but the C3 species had narrower leaves with an entire margin whereas the C4 leaves were wider and the margin was slightly dentate (Fig. 1A). Both C3 species had closer veins than the C3 species (Fig. 1B; Supplementary Fig. S2). In addition, analysis of transverse leaf sections indicated that C3 leaves were thicker because of larger cells and more cell layers (Fig. 1C). Clear differences in BS and M cell arrangement were visible between the C3 and C4 pairs, with the BS in both C3 species being more uniform in shape than in the C3 species (Fig. 1C). In addition, compared with the C3 species, M cells in the C4 leaves were smaller (Supplementary Fig. S2) and appeared to show increased contact with the BS. In agreement with previous analysis (McKown and Dengler, 2007), leaves of the two C3 and two C4 species typically contained five and eight ground cell layers between the adaxial and abaxial epidermis, respectively.

In all four Flaveria species, leaves of 1.8–2 cm length showed a basipetal maturation gradient with differentiating tissues at the base and fully differentiated tissues at the tip (Fig. 2A–C, Supplementary Fig. S3). This base-to-tip maturation programme is typical of dicotyledons, but it is notable that in Flaveria maturation occurred in larger leaves than in

Fig. 1. Characteristics of mature leaves from C3 and C4 species of Flaveria. (A) Representative outlines of mature leaves showing the variation between species and the dentate nature of the two C3 species. (B) Vein traces taken in the middle of the leaf to the right of the mid-vein indicating denser venation in mature leaves of both C3 species. (C) Transverse sections showing the reduced mesophyll cross-sectional area and the closer veins with more symmetric bundle sheaths in the C4 compared with the C3 species. Each cell type is false colour-coded: veins (red), bundle sheath (dark green), spongy mesophyll (light green), palisade mesophyll (turquoise), parenchyma layer with no or few chloroplasts (blue), and epidermis (ilac). Species abbreviations are as follows: Flaveria pringlei (Fp), Flaveria robusta (Fr), Flaveria bidentis (Fb), Flaveria trinervia (Ft). Scale bars represent 1 cm (A) or 100 μm (B and C).
Gynandropsis gynandra, where the maturation gradient was detected in leaflets of 0.4 cm length (Aubry et al., 2014). This prolonged development of leaf maturation in Flaveria thus provides an excellent system to analyse the induction of C_{4} photosynthesis as the larger leaf allows these leaves to be divided into more stages for functional analysis.

To better understand differences in leaf maturation in C_{3} and C_{4} Flaveria, leaves from each species were divided into six portions, and RNA was extracted, quantified, integrity determined (Supplementary Fig. S4) and then subjected to deep sequencing. On average, 20 million reads were recovered from each replicate (Supplementary Table S1). After de novo assembly of these reads, the average number of annotated transcripts per species was 12 475 (Supplementary Table S1).

Read mapping was used to quantify transcript abundance in TPM (Supplementary Data S1). By grouping the two base, mid, and tip stages, it was possible to analyse transcript behaviour quickly across the maturation gradient. This revealed that on average 40% of annotated transcripts showed descending behaviours, meaning that they were expressed more strongly at the base of the leaf than at the tip (Supplementary Table S1).

Using the 8000 annotations found in all four species, correlation analysis showed that patterns of gene expression first clustered by species and then by photosynthetic type, meaning that the correlation between neighbouring developmental stages was highest within species and that the correlation was higher between species of the same photosynthetic type than between species of different photosynthetic types (Supplementary Fig. S5). A gradient from base to tip was clear in all four species, but was more pronounced in the two C_{4} species (Supplementary Fig. S5).

The six portions from the base to the tip of the leaf showed a clear induction of genes known to encode components of the C_{4} cycle (Fig. 3). An important step in the evolution of C_{4} photosynthesis is thought to be a co-ordinate alteration in expression of genes encoding the photorespiratory cycle and nitrogen metabolism (Mallman et al., 2014). Consistent with a reduction in photorespiration in the C_{4} species, genes of photorespiration increased along the developmental gradient in C_{3} species but to a lesser extent in the C_{4} Flaveria species. The main exception was DIT1, a transporter gene which is involved in the C_{4} cycle and is more abundant in both C_{4}
species (Supplementary Data S1). Around 200 genes are consistently up-regulated in the upper tip tissue using \( \text{C}_4 \) compared with \( \text{C}_3 \) photosynthesis (Fig. 4A; Supplementary Data S2) and, in young tissue that is still undergoing differentiation, the number of genes up-regulated in the \( \text{C}_4 \) species is around double this number. These estimates are significantly lower than those reported from analysis of whole leaves at various developmental stages from \( \text{C}_3 \) Tarenaya hassleriana and \( \text{C}_4 \) G. gynandra in the Cleomaceae (Külahoglu et al., 2014), where between 2500 and 3500 genes were estimated to be up-regulated in the \( \text{C}_4 \) species. The reduction in differential gene expression in the \( \text{C}_4 \) compared with \( \text{C}_3 \) Flaveria species in this study compared with the estimates from the Cleomaceae may be due to the use of two \( \text{C}_3 \) and two \( \text{C}_4 \) species, removing species-specific differences, but also to the reduced phylogenetic distance between the Flaveria species (Sage et al., 2011). Principal components analysis indicated that the first two dimensions contributed to 37% and 19% of the variation, respectively, and that whereas the first dimension was associated with the leaf maturation signal from all species, the second dimension was associated with the photosynthetic pathway being used (Fig. 4B). This finding implies that \( \text{C}_4 \) metabolism has a large impact on the differences in gene expression associated with \( \text{C}_3 \) and \( \text{C}_4 \) leaves within one genus. When MapMan categories associated with the top decile of differentially expressed genes were assessed it was notable that photosynthesis genes were over-represented for a prolonged period along the gradient in the \( \text{C}_4 \) species (Fig. 4C). This was also the case for genes related to DNA and transport. In contrast, the MapMan category associated with RNA was over-represented for longer in the \( \text{C}_3 \) species (Fig. 4C). Equivalent analysis for GO terms indicated that cell wall loosening was over-represented for longer in both \( \text{C}_4 \) species compared with the \( \text{C}_3 \) species (Supplementary Fig. S6). Overall, these data indicate that the maturation of Flaveria leaves captures changes in transcript abundance associated with induction of \( \text{C}_3 \) or \( \text{C}_4 \) photosynthesis, and provide insight into the broader alterations in gene expression. We next aimed to relate the underlying anatomical changes associated with leaf maturation in these \( \text{C}_4 \) and \( \text{C}_3 \) species of Flaveria to changes in gene expression determined from the deep sequencing.

**Kranz anatomy during leaf maturation**

Using cleared leaves, veins were traced to visualize their development. In the most basal part of the leaf, major veins were present but, in the two \( \text{C}_4 \) species, the highest order veins were still being laid down. Some of these developing higher order veins could be detected by the specific patterns of periclinal cell divisions associated with their production (Supplementary Fig. S7). Vein density was determined in each of the six sections along the leaf developmental gradient, and this showed markedly different developmental patterns in the \( \text{C}_3 \) and \( \text{C}_4 \) pairs of Flaveria (Fig. 5A). Vein density was lower at the base in \( \text{C}_4 \) compared with \( \text{C}_3 \) leaves of Flaveria, but, during the transition from the base to the middle of the leaf, density increased dramatically in the \( \text{C}_4 \) species (Fig. 5B). In both \( \text{C}_3 \) and \( \text{C}_4 \) pairs, vein density then decreased from the middle to the tip of the leaf. The reduction in vein density as leaves matured is probably associated with expansion of M and BS cells. In the two most basal sections of the leaf, cell division was still ongoing and procambial strands were initiated first with a periclinal and then an anticlinal division, to derive BS cells (Supplementary Fig. S7). These basal portions of the leaf therefore appear most interesting to interrogate for candidate genes associated with these processes. The \textit{de novo}

![Fig. 3. Transcripts encoding the C4 cycle increase dramatically during leaf maturation of the C4 species. The schematic of major components of the C4 pathway located in mesophyll and bundle sheath cells of C4 leaves is surrounded by plots of transcript abundance for key genes of the NADP-ME subtype in transcripts per million (TPM) for the four species. RBCS1A transcripts are, as expected, higher in the C4 species. All C4 transcripts other than CARBONIC ANHYDRASE1 (CA1) are barely detectable in the C3 species. PEPC1, PHOSPHOENOLPYRUVATE CARBOXYLASE 1; NADP-MDH, NADP-DEPENDENT MALATE DEHYDROGENASE; NADP-ME1, NADP-DEPENDENT MALIC ENZYME1; PPDK, PYRUVATE, ORTHOPHOSPHATE D kinase; PPDK-RP1, PYRUVATE, ORTHOPHOSPHATE D kinase REGULATORY PROTEIN; NHD1, SODIUM-HYDROGEN ANTIPORTER; BASS2, BILE ACID SODIUM SYMPORTER FAMILY2. The two C4 species are shown with solid lines, and the two C3 species with dashed lines.](https://academic.oup.com/jxb/article-abstract/68/2/177/2849803)
assembled transcriptomes from the four *Flaveria* species were therefore analysed for homologues of genes known to impact on vein formation in other species.

Six genes that affect vein formation in *A. thaliana* showed different behaviours in the C₄ *Flaveria* species compared with the C₃ species (Fig. 6). In all cases, absolute transcript abundance tended to be higher at the base of the C₄ leaves compared with the base of the C₃ species. However, the most consistent differences in expression were found for *A. thaliana* HOMEBOX-GENE-8 (ATHB8). SHR and SCR have been implicated in the development of C₄ Kranz anatomy (Slewinski et al., 2012; Slewinski, 2013). SHR was detected in three of the species with a descending pattern, but with no clear
difference between C₃ and C₄. SCR also showed a descending pattern, but again with no clear difference between photosynthetic types. Of Scarecrow-like genes, only SCARECROW-LIKE 14 was clearly higher in C₄ and showed a parabolic expression pattern, being most highly expressed at the base and the tip (Supplementary Data S3). SCARECROW-LIKE 14 was recently identified from cross-species selection scans in the grasses (Huang et al., 2016). Transcripts encoding the auxin response factors ARF3, ARF8, and also IAA7 showed differences in the timing of expression in C₄ compared with C₃ species (Supplementary Data S3).

The same six regions from leaf base to tip used to define vein density were next used to quantify maturation of BS and M cells (Fig. 7). Leaf thickness was higher in the C₃ compared with the C₄ species, and cell expansion continued for longer in the C₃ leaves (Fig. 7A). The BS in the C₃ species was less regular than in the C₄ species, and individual cell size within the BS was more variable (Fig. 7A). However, it was notable that the cross-sectional area of BS cells in the C₃ species was larger than that of C₄ species, particularly towards the tip.

It has been proposed that a large BS cell size is a key early event associated with the evolution of C₄ photosynthesis (Christin et al., 2013; Griffiths et al., 2013; Williams et al., 2013), and so it appears that within Flaveria this is a pre-existing trait. It was noticeable that the cross-sectional area of M cells from the two C₃ species was around five times greater than that of M cells of the C₄ species (Fig. 7B). This strongly implies that compared with C₃ species, the reduced leaf depth of C₄ species is associated with an inhibition of M cell expansion as well as a reduced number of cell layers. A number of genes that have been annotated as having roles in cell proliferation or expansion showed behaviours that may explain the reduced cell expansion in leaves of both C₄ species. Most notable was DWF4, a 22α-hydroxylase that catalyses the rate-limiting step of brassinosteroid synthesis and so controls cell expansion. In both C₃ species, DWF4 transcript abundance increased along the leaf gradient but in the C₄ species its transcripts were barely detectable (Supplementary Data S4). These data imply that DWF4 is a strong candidate for the reduced expansion associated with maturation of the C₄ leaf.

![Fig. 6. Patterns of transcript abundance in the four Flaveria species for genes known to be involved in vein formation. Transcript abundance is shown in transcripts per million (TPM), and the parts of the leaf are annotated on the x-axis. The two C₄ species are shown with solid lines, and the two C₃ species with dashed lines. Fp, Flaveria pringlei; Fr, Flaveria robusta; Fb, Flaveria bidentis; Ft, Flaveria trinervia.](https://academic.oup.com/jxb/article-abstract/68/2/177/2849803)
mutation of CYP450D2, a Cyt P450 involved in brassinosteroid synthesis, leads to perturbed vein spacing (Rizal et al., 2015). Although the orthologue to CYP450D2 is not present in A. thaliana, transcripts encoding CYP90D1 and CYP90C1 appear to be slightly higher in the base of the C₄ compared with the C₃ Flaveria species (Supplementary Data S4).

A prolonged rate of cell division in the C₄ Flaveria species is also supported by transcript profiles. A number of genes implicated in cell division are found at higher levels in basal sections of C₄ than C₃ leaves including PLE (Müller et al., 2004), MORI (Whittington et al., 2001), ATK1 (Marcus et al., 2003), and ATK3 (Mitsui et al., 1994) which are involved in microtubule organization, and SCC3 (Chelysheva et al., 2005) and SYNI (Cai et al., 2003) which are implicated in sister chromatid pairing (Supplementary Data S4). Transcripts derived from genes associated with DNA replication and repair including SOG1 (Yoshiyama et al., 2009), TOP6B (Gilkerson and Callis, 2014), MutS homologues MSH (MSH3, MSH4, and MSH7) (Culligan and Hays, 2000), and multiple components of the mini-chromosome maintenance (MCM) complex (MCM2, MCM3, MCM5, and MCM6) (Tuteja et al., 2011) were also more abundant in the C₄ species at the base of the leaf (Supplementary Data S4). Good candidate regulators of cell division include MYB3R4 which forms a complex with E2FB and RETINOBLASTOMA-RELATED PROTEIN (RBR1) to activate cell cycle progression (Haga et al., 2011), and DEL1 which maintains cell division by repressing entry into the endocycle (Vlieghe et al., 2005), both of which are higher in the base of C₄ leaves and maintain expression longer into the leaf gradient.

A reduction along the leaf gradient in the abundance of transcripts encoding CPP SYNTHASE 1 (CPS1), which catalyses the first step of gibberellin synthesis, and a spike in transcripts of GA2ox2 (Supplementary Data S4), which degrades gibberellin, between the upper base and lower mid in C₄ leaves is in keeping with findings that a decrease in gibberellin levels controls the transition between cell division and expansion in maize (Nelissen et al., 2012). A corresponding pattern cannot be seen in C₃ leaves, which may mean that the transition between cell division had already occurred at the base of C₃ lineages, or is mediated by another mechanism. We did not undertake analysis of sink-to-source transitions within the Flaveria leaves. However, as sugar signalling impacts on the transition from cell division to expansion, it is possible that the altered maturation of C₄ compared with C₃ leaves is related to sink-to-source transition as well as the C₄ paradigm per se (Li et al., 2010).

Chloroplast maturation within the leaf gradient

Species that primarily use NADP-ME to decarboxylate malate in the C₄ BS commonly develop dimorphic chloroplasts with
M cells containing stacked thylakoids (grana) and the BS cells containing fewer grana (Edwards and Walker, 1983; Edwards and Voznesenskaya, 2011). To provide insight into the dynamics of chloroplast development and maturation in C₃ and C₄ Flaveria species, TEM was used to investigate chloroplast structure in each of the six stages along the leaf gradient. Dimorphic chloroplasts were observed in the C₄ species but not in the C₃ species (Fig. 8A–D). Towards the base of the C₄ leaves, both BS and M chloroplasts contained 2–3 granal lamellae per stack (Supplementary Fig. S8). In more mature parts of the leaf, levels of stacking were reduced in BS cells but increased in M cells (Fig. 8A–D). In both C₃ species, a developmental gradient in granal stacking was also visible from base to tip, with more mature parts of the leaf showing increased levels of granal stacking in both cell types (Supplementary Fig. S8). Thus, a key difference between the C₃ and C₄ leaves was the reduction in granal stacking in C₄ BS cells from the middle of the leaf.

The control of photosynthesis gene expression and chloroplast development has previously been linked to a pair of transcription factors known as GOLDEN-LIKE 1 and 2 (GLK1 and GLK2) (Supplementary Data S5). Unfortunately, de novo assembly did not generate a complete set of contigs for either GLK1 or GLK2 from all four species, so we were unable to determine if their abundance correlated with the observed changes in chloroplast development. However, other candidate genes that can be associated with chloroplast maturation were identified. For example, cytokinin is known to play a role in chloroplast maturation, and increased abundance of transcripts encoding the cytokinin receptors AHK2 and AHK3, along with the downstream effector ARR1 in C₄ species suggests that enhanced cytokinin signalling may play a role in these alterations to the C₄ chloroplast. Additionally, it was also notable that the behaviour of two nuclear-encoded chloroplast RNA polymerases (RpoTp and RpoTmp) (Swiatecka-Hagenbruch et al., 2008) showed clear differences between the C₃ and C₄ Flaveria species. Transcripts encoding RpoTp, which controls ycfI that is involved in plastid protein import, were very low in both C₄ species. Further indications to changes in chloroplast protein import related to components of the TIC/TOC complex (Supplementary Data S5). Transcripts encoding the outer envelope receptor proteins TOC34 and TOC159 were higher in the base and peaked later in the C₄ than the C₃ species. Additionally, TOC75, the major channel constituent for protein import into the chloroplast (Li and Chiu, 2010), was lower at the base of C₄ leaves. These data would be consistent with protein import playing a role in the establishment of dimorphic chloroplasts. Chloroplast maturation also correlated with increased abundance of nuclear-encoded chloroplast sigma factors, which direct the nuclear-encoded chloroplast RNA polymerase (NEP) to promoters.

Previous studies have demonstrated that C₄ plants alter photosystem activity to adjust ADP/ATP and NADPH ratios to ensure proper functioning of the carbon pump. Consistent with this was the C₄-specific up-regulation of SIG3, which initiates transcription of psbN (Zghidi et al., 2007) that is involved in repair of PSII complexes (Torabi et al., 2014). It was also notable that transcripts encoding components of the

---

**Fig. 8.** Chloroplast ultrastructure of mesophyll and bundle sheath cells from C₃ F. pringlei and C₄ F. bidentis. Transmission electron microscopy was used to investigate chloroplast ultrastructure. (A, B) Bundle sheath chloroplasts from C₃ Flaveria pringlei and C₄ Flaveria bidentis. (C, D) Mesophyll chloroplasts from C₃ Flaveria pringlei and C₄ Flaveria bidentis. Little granal stacking was seen in bundle sheath chloroplasts from F. bidentis (B) whereas it could be observed in C₃ bundle sheath and in all mesophyll chloroplasts. Insets show a close-up of thylakoids, with white arrowheads indicating thylakoid stacks (A, C, D) or the lack of extensive stacking (B). Scale bars represent 500 nm.
NDH complex (NDF1, NDF6, and PIFI), as well as genes involved in cyclic electron transport (PGR5 and PGR/LIA), were more abundant in both C₄ species, and these differences became more apparent from base to tip (Supplementary Data S5). Chloroplast dimorphism also coincided with an increase in transcript abundance of CRR1, which is proposed to be involved in NDH complex assembly.

Analysis of electron micrographs also revealed that chloroplast length in the BS increased from base to tip in the C₄ species more than in the C₃ species (Fig. 9). Two members of the REDUCED CHLOROPLAST COVERAGE (REC) gene family, REC2 and REC3, knockout of which causes a reduction in chloroplast size (Larkin et al., 2016), were up-regulated in the C₄ species. Further, FtsZ1-1, which is involved in chloroplast division, was significantly down-regulated in the C₄ compared with the C₃ species (Supplementary Data S5). As reduced division of chloroplasts leads to increased size (Pyke and Leech, 1994; Pyke, 1997), the reduced expression of chloroplast division genes in C₄ plants could lead to their increased size in BS cells.

Convergence in transcript abundance between independent C₄ lineages

By combining our data from the four Flaveria species with publicly available data sets, we next sought to investigate whether separate lineages of C₄ plants shared common changes in transcript abundance compared with their C₃ congeners. A recent study compared transcriptomes derived from leaf developmental gradients of T. hassleriana (C₃) and G. gynandra (C₄) from the Cleomaceae (Külahoglu et al., 2014). In contrast to our work, whole leaves were staged by age (Külahoglu et al., 2014) and changes in transcript abundance compared with non-photosynthetic tissues such as roots. This analysis identified a set of 33 genes that could be annotated with unique A. thaliana identifiers. This set of genes showed root expression in the C₃ species T. hassleriana but leaf expression in the C₄ species G. gynandra (Külahoglu et al., 2014), and so it was proposed that these were key genes that neofunctionalize to become involved in C₄ photosynthesis. Of these 33 genes, homologues to 15 were detected in all four Flaveria species, but only three showed up-regulation in the C₄ compared with C₃ species in our study and three showed higher expression in the two C₃ species (Supplementary Data S6).

Previously, transcript abundance in developing leaves of the C₄ dicot G. gynandra (Aubry et al., 2014) has been compared with that in the C₄ monocot maize (Li et al., 2010; Pick et al., 2011). Despite the very wide phylogenetic distance between these species, 18 transcription factors that showed the same patterns of behaviour in both C₄ G. gynandra and C₄ maize were identified (Aubry et al., 2014). Ten of these
18 candidates were also detected in all four *Flaveria* species. Five of those showed ascending behaviour and higher transcript abundance in the two C₄ species. These were RAD-LIKE6 (RL6, AT1G75250); AT2G05160, a CCCH-type zing finger family protein with an RNA-binding domain; AT2G21530, a SMAD/FHA domain-containing protein; RNA Polymerase Sigma Subunit C (SIGC, AT3G35920), and BEL1-Like Homeodomain1 (BLH1, AT5G67030) (Supplementary Data S7). The five genes have therefore been shown to exhibit the same behaviour along leaf developmental gradients in C₄ species in the monocots and dicots [both rosids (Clemaceae) and asterids (Asteraceae)]. This strongly supports the notion that independent lineages of C₄ plants have recruited homologous trans-factors to induce the C₄ system.

**Summary and conclusions**

Most previous analysis derived from comparisons of RNA-seq of C₃ and C₄ species have relied on either sampling one congeneric C₃ and C₄ species pair, or analysis of mature leaf tissues. The number of genes that are differentially expressed between any two species can be very high simply due to species-specific differences, which do not relate to their photosynthetic pathway. In this work, we identify genes that are consistently differentially expressed in multiple C₃ and C₄ species from the same genus along a leaf maturation gradient. In the *Flaveria* genus, the data set therefore indicates that ~200 genes are consistently up-regulated in mature tissue using C₄ compared with C₃ photosynthesis. In young tissue that is still undergoing differentiation, the number of genes up-regulated in the C₄ species is around double this number. Leaf maturation and the photosynthetic pathway were responsible for the greatest amount of variation in transcript abundance. In *Flaveria*, the evolution of C₄ photosynthesis is not associated with an expansion of BS cell size, but rather a reduction in the cross-sectional area of M cells in C₄ compared with C₃ species. Overall, the analysis therefore provides quantitation of changes in gene expression associated with C₃ and C₄ photosynthesis, candidate genes underlying the alterations in leaf characteristics associated with these pathways, and, through comparison with independent C₄ lineages outside of the Asteraceae, insight into the extent to which parallel evolution underlies the convergent and complex C₄ trait.

**Supplementary data**

Supplementary data are available at *JXB* online.

- Fig. S1. Phylogeny of the *Flaveria* genus.
- Fig. S2. Mature leaf vein density and cell size.
- Fig. S3. Leaf anatomy maturation gradient in C₃ and C₄ *Flaveria*.
- Fig. S4. RNA quality from sections of each species.
- Fig. S5. TPM correlation matrix.
- Fig. S6. GO term analysis presented as heatmaps for each *Flaveria* species.
- Fig. S7. Cell division and vein development in the base samples of *Flaveria*.

Data S1. TPM values for all four species for all stages and all replicates.

Data S2. Summary of transcripts that were up-regulated in both C₃ species or both C₄ species at each stage.

Data S3. Transcript abundance of SHR, SCR, and SCR-like genes as well as auxin response genes.

Data S4. Transcript abundance of genes associated with cell division.

Data S5. Transcript abundance of genes associated with chloroplastic maturation and development.

Data S6. Comparison of data derived from this study with genes identified by Külahoglu et al. (2014).

Data S7. Comparison of data derived from this study with genes identified by Aubry et al. (2014).

Table S1. Number of reads and transcripts, per stage and species.

**Author contributions**

BMCK and JMH designed the study. BMCK grew the plants, undertook the analytical analysis, and isolated the RNA; RS-U and CB undertook the *de novo* transcriptome assembly and annotation of contigs; BMCK, SJB, and IR-L analysed the RNA-seq data; BMCK, SJB, and JMH wrote the paper.

**Acknowledgements**

Seeds for all *Flaveria* species used were kindly provided by Udo Gowik (Universität Düsseldorf). The European Union *3to4* project and the Biotechnology and Biological Sciences Research Council (BBSRC) grant BB/J011754/1 funded the project.

**References**

Aubry S, Kelly S, Kümpers BM, Smith-Unna RD, Hibberd JM. 2014. Deep evolutionary comparison of gene expression identifies parallel recruitment of trans-factors in two independent origins of C₄ photosynthesis. *PLoS Genetics* 10, e1004365.

Bauwe H, Hageman M, Fernie AR. 2010. Photorespiration: players, partners and origin. Trends in Plant Science 15, 300–308.

Bellasio C, Griffiths H. 2014. The operation of two decarboxylases, transamination, and partitioning of C₄ metabolic processes between mesophyll and bundle sheath cells allows light capture to be balanced for the maize C₄ pathway. Plant Physiology 164, 466–480.

Bowes G, Ogren WL, Hageman RH. 1971. Phosphoglycolate production catalyzed by ribulose diphosphate carboxylase. Biochemical and Biophysical Research Communications 45, 716–722.

Burgess SJ, Granero-Moya I, Grangé-Guermente MJ, Boursnell C, Terry MJ, Hibberd JM. 2016. Ancestral light and chloroplast regulation form the foundations for C₄ gene expression. Nature Plants 2, 16161.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.

Cai X, Dong F, Edelmann RE, Makaroff CA. 2003. The Arabidopsis SYN1 cohesin protein is required for sister chromatid arm cohesion and homologous chromosome pairing. *Journal of Cell Science* 116, 2999–3007.

Chelysheva L, Diallo S, Vezon D, et al. 2005. AtREC8 and AtSCC3 are essential to the monopolar orientation of the kinetochores during meiosis. *Journal of Cell Science* 118, 4621–4632.

Christin P-A, Osborne CP, Chatelot DS, Columbus JT, Besnard G, Hodkinson TR, Garrison LM, Vorontsova MS, Edwards EJ. 2013. Anatomical enablers and the evolution of C₃ photosynthesis in grasses. Proceedings of the National Academy of Sciences, USA 110, 1381–1386.

Fig. S8. Chloroplast maturation in mesophyll and bundle sheath cells for each species.

The number of genes that are consistently differentially expressed in multiple C₃ and C₄ species in the monocots and dicots [both rosids (Clemaceae) and asterids (Asteraceae)]. This strongly supports the notion that independent lineages of C₄ plants have recruited homologous trans-factors to induce the C₄ system.
Xie Y, Wu G, Tang J, et al. 2014. SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. Bioinformatics 30, 1660–1666.

Yoshiyama K, Conklin PA, Huefner ND, Britt AB. 2009. Suppressor of gamma response 1 (SOG1) encodes a putative transcription factor governing multiple responses to DNA damage. Proceedings of the National Academy of Sciences, USA 106, 12843–8.

Zghidi W, Merendino L, Cottet A, Mache R, Lerbs-Mache S. 2007. Nucleus-encoded plastid sigma factor SIG3 tranecribes specifically the psbN gene in plastids. Nucleic Acids Research 35, 455–464.