RESEARCH PAPER

An rbcL mRNA-binding protein is associated with C₃ to C₄ evolution and light-induced production of Rubisco in Flaveria

Pradeep Yerramsetty¹,², Erin M. Agar¹,#, Won C. Yim², John C. Cushman² and James O. Berry¹,*

¹ Department of Biological Sciences, State University of New York, Buffalo, NY 14260, USA
² Department of Biochemistry and Molecular Biology, University of Nevada, Reno, NV 89557, USA
# Present address: Cargill, Inc., Fort Collins, CO 80525, USA.
* Correspondence: camjob@buffalo.edu

Received 7 April 2017; Editorial decision 5 July 2017; Accepted 7 July 2017

Editor: Robert Sharwood, Australian National University

Abstract

Nuclear-encoded RLSB protein binds chloroplastic rbcL mRNA encoding the Rubisco large subunit. RLSB is highly conserved across all groups of land plants and is associated with positive post-transcriptional regulation of rbcL expression. In C₃ leaves, RLSB and Rubisco occur in all chlorenchyma cell chloroplasts, while in C₄ leaves these accumulate only within bundle sheath (BS) chloroplasts. RLSB's role in rbcL expression makes modification of its localization a likely prerequisite for the evolutionary restriction of Rubisco to BS cells. Taking advantage of evolutionarily conserved RLSB orthologs in several C₃, C₃–C₄, C₄-like, and C₄ photosynthetic types within the genus Flaveria, we show that low level RLSB sequence divergence and modification to BS specificity coincided with ontogeny of Rubisco specificity and Kranz anatomy during C₃ to C₄ evolution. In both C₃ and C₄ species, Rubisco production reflected RLSB production in all cell types, tissues, and conditions examined. Co-localization occurred only in photosynthetic tissues, and both proteins were co-ordinately induced by light at post-transcriptional levels. RLSB is currently the only mRNA-binding protein to be associated with rbcL gene regulation in any plant, with variations in sequence and acquisition of cell type specificity reflecting the progression of C₄ evolution within the genus Flaveria.

Key words: Cell type specificity, light regulation, nuclear-encoded mRNA-binding protein, phylogeny, plastid-encoded rbcL gene, post-transcriptional control, protein synthesis, RLSB and C₄ evolution, Rubisco regulation, tissue specificity.

Introduction

C₄ photosynthesis is used by only 5% of all terrestrial plants, yet these species yield up to a quarter of the Earth’s primary productivity (Jones, 2011; Sage and Zhu, 2011; Sage et al., 2012). The enhanced photosynthetic productivity of C₄ plants depends on specialized leaf anatomy that compartmentalizes biochemical modifications of the more basic C₃ pathway (Hatch 1987; Berry et al., 2011, 2013; Garner et al., 2016). The leaves of most C₄ plants separate two sets of photosynthetic reactions into different leaf cell types, called Kranz anatomy, that consist of an outer mesophyll (M) cell layer surrounding an internal ring of bundle sheath (BS) cells, which in turn surround the leaf veins. An exception occurs in ‘single-cell C₄’ species of the family Chenopodiaceae that compartmentalize these same reaction sets into two regions of leaf chlorenchyma cells (Edwards and Voznesenskaya, 2011; Koteyeva et al., 2016). In Kranz species, the initial carboxylation of
phosphoenolpyruvate (PEP) occurs only in the M cells, where phosphoenolpyruvate carboxylase (PEPCase) is specifically localized. CO₂ incorporation into the Calvin–Benson cycle occurs within chloroplasts of internal BS cells, where Rubisco is specifically located. These C₄ specializations essentially eliminate metabolically wasteful photorespiration and increase photosynthetic efficiency (Ghannoum et al., 2011; Gowik et al., 2011; Sage et al., 2012; Berry et al., 2013, 2016; Garner et al., 2016).

C₄ photosynthesis originated ~35 million years ago, evolving independently >60 times across many higher plant families, including both monocots and dicots (Osborne and Freckleton, 2009; Westhoff and Gowik, 2010; Sage et al., 2011, 2012, 2013; Khosh ravesh et al., 2016). The occurrence of C₃–C₄ intermediates in several present-day genera provides evidence for distinct stages during the evolutionary transition from C₃ to full C₄ photosynthesis (Westhoff and Gowik, 2010; Sage et al., 2012; Heckmann et al., 2013; Aubry et al. 2014; Khosh ravesh et al., 2016; Schulze et al., 2016). An initial step is thought to have been the confinement of mitochondrial glycine decarboxylase (GDC) to internal BS cells in leaves of an ancient C₁ plant (Sage et al., 2013; Schulze et al., 2013; Mallman et al., 2014; Khosh ravesh et al., 2016; Schulze et al., 2016). Confining the GDC activity to BS cells would establish a photorespiratory pump and increase CO₂ concentrations around Rubisco, a process also referred to as C₂ photosynthesis. It is hypothesized that several additional transitional steps followed establishment of the foundational C₃ state, ultimately leading to present-day full C₄ pathways (Gowik et al., 2011; Schulze et al., 2013; Khosh ravesh et al., 2016; Lin et al., 2016; Schulze et al., 2016). Proposed stages include redistribution of mitochondria within BS cells, the evolution of ‘proto-Kranz’ anatomy, and final optimization/activation of the C₄ cycle. According to this model, the final optimization stage included establishment of BS cell-specific expression of several metabolic genes, some of which already showed BS cell-preferential expression within C₃ leaves (Aubry et al., 2014). Another proposed process was eliminating the expression of genes encoding Rubisco and other Calvin–Benson cycle enzymes from M cells, leading to the cell type specificity pattern characteristic of present-day C₄ species.

Rubisco provides an excellent model system to investigate how BS cell specificity for nuclear and plastid-encoded photosynthetic genes might have occurred during C₄ evolution (Berry et al., 2016). Rubisco is composed of eight large subunits (LSUs) encoded by the chloroplast rbcL gene and eight small subunits (SSUs) encoded by a nuclear RbcS gene family (Andersson, 2008; Andersson and Backlund, 2008). Although synthesized in different cellular compartments, anterograde and retrograde signaling processes ensure proportionate amounts of each subunit for the Lₘₘ holoenzyme (Patel and Berry, 2008; Berry et al., 2013, 2016). In both C₃ and C₄ plants, rbcL and RbcS genes are highly regulated in response to external and internal factors (Patel and Berry, 2008; Berry et al., 2013, 2016). External determinants include light, temperature, disease, water, and nutrient availability (Berry et al., 2011, 2013, 2016). Internal factors include developmental stage, cell type, tissue type, and senescence. While transcriptional control of Rubisco gene expression has been implicated in many of these processes, post-transcriptional mechanisms also represent prominent regulatory steps (Patel and Berry, 2008; Berry et al., 2013, 2016).

Post-transcriptional regulation of gene expression is mediated by cis-acting regulatory sequences on an mRNA, usually within 5’ or 3’ untranslated regions (UTRs) (Raynaud et al., 2007; Barkan, 2011; Brown et al., 2011; Berry et al., 2013). These are recognized by RNA-binding proteins that regulate mRNA translation, processing, or stability (Raynaud et al., 2007; Tillich et al., 2010; Barkan, 2011; Berry et al., 2013; Bowman et al., 2013). Many post-transcriptional regulation studies in plants have focused on plastid-encoded genes, where such processes depend on nuclear-encoded plastid-targeted RNA-binding proteins (Tillich et al., 2010; Barkan, 2011; Berry et al., 2013; Bowman et al., 2013). Several classes of binding proteins specifically regulate ~100 different chloroplast-encoded mRNAs (Raynaud et al., 2007; Tillich et al., 2010; Barkan, 2011; Berry et al., 2013; Bowman et al., 2013).

The nuclear-encoded rbcL RNA S1 binding domain protein (RLSB) is highly conserved across all groups of land plants (Bowman et al., 2013; Yerramsetty et al., 2016). Biochemical, genetic, and evolutionary studies implicate RLSB as a positive post-transcriptional determinant that binds rbcL mRNA, thereby affecting its stability and/or translation (Bowman et al., 2013; Berry et al., 2016). In the C₃ plant Arabidopsis, RLSB accumulates within the chloroplasts of all leaf chlorenchyma cells. In Kranz-type C₄ species, RLSB accumulates only within Rubisco-containing chloroplasts of internal BS cells (Bowman et al., 2013), and in the single-cell C₄ plant Bienertia sinuspersici only within internal Rubisco-containing central compartment chloroplasts (Rosnow et al., 2014).

As a ubiquitous and highly conserved mRNA-binding protein associated with post-transcriptional rbcL regulation, RLSB could play a role in many processes affecting Rubisco production and localization across a wide range of species. This current study extends our previous findings of RLSB localization and function by revealing the stepwise evolutionary progression to full C₄ in the genus Flaveria (Asteraceae), which contains species possessing a range of photosynthetic types (McKown et al., 2005; Sage et al., 2013; Mallman et al., 2014; Lyu et al., 2015). Findings presented here support a model in which evolutionary modification of RLSB production from an ancestral ‘default’ state in C₃ plants to full BS cell specificity in C₄ plants contributed to the subsequent cell-specific expression of chloroplast-encoded rbcL expression and Rubisco localization, a process most probably initiated at the C₃–C₄/C₃ intermediate stage and completed during the final ‘activation/optimization’ stage of C₄ evolution. We also show the tight relationship between RLSB and Rubisco localization in photosynthetic tissues, and post-transcriptional control of both RLSB and Rubisco in both C₃ and C₄ species. We conclude that the evolutionary acquisition of specialized C₄ patterning did not modify the most basic ‘default’ aspects of RLSB/Rubisco localization or production, such as accumulation only in green tissues or light regulation, that are probably shared among all plants.
Materials and methods

Comparison and phylogeny of RLSB sequences within the genus Flaveria

Translated RLSB ortholog sequences from multiple Flaveria species (used with permission from Dr. Julian Hibberd, Department of Plant Sciences, University at Cambridge) were aligned using the MUSCLE multiple sequence algorithm (Edgar, 2004) implemented using CLC Main Workbench 7.7.2. and CLC Genomics Workbench 8.0.3 (https://www.qiagenbioinformatics.com/). Translated RLSB sequences from Arabidopsis thaliana, Zea mays, and Amborella trichopoda (Yerramsetty et al., 2016) were included in the alignment as references (Supplementary Fig. S1 at JXB online).

For the phylogeny, data from a total of 63 RNA-Seq libraries of 16 Flaveria species were obtained from the National Center for Biotechnology Information (NCBI) sequence read archive (SRA) database (Supplementary Table S1). Initial removal of low quality reads and adapter trimming was performed with Trimmomatic (Bolger et al., 2014). Filtered reads were assembled with SOAPdenovo-Trans (version 1.04) with eight different K-mers (25, 35, 45, 55, 65, 75, 85, and 95) (Xie et al., 2014). The assembled contigs were merged and redundancy removed using Evidential gene (http://arthropods.eugenes.org/EvidentialGene/evigene/).

The number of contigs, total length of transcripts, N50 length, and BUSCO quality assessment (Simão et al., 2015) were reported for each species (Supplementary Table S2). The Arabidopsis RLSB gene (AT1G71720) was mapped against the individually assembled contigs using TBLASTN from BLAST (Camacho et al., 2009), and only full-length RLSB sequence were retained. For fragmented RLSB sequence from low quality assembled transcripts species, filtered RNA-Seq reads were mapped to fragmented RLSB sequence by Bowtie2 with the local alignment option (Langmead and Salzberg, 2012). The mapped reads were converted to fasta and quality file, and assembled with fragmented RLSB sequence using Phrapt (http://phrapt.org/). The contigs from this strategy were examined for full-length coverage by ORFfinder (Wheeler et al., 2003) and SmartBLAST (https://blast.ncbi.nlm.nih.gov/smartblast/).

A set of 16 Flaveria RLSB genes and three strictly defined outgroup RLSB genes from Asteraceae were used to reconstruct phylogeny. The RLSB transcript sequences were converted to coding sequences (CDS) and peptide sequences using TransDecoder (Haas et al., 2013). The peptide sequence from this result were aligned using MUSCLE (Edgar, 2004), and corresponding codon sequence were aligned on peptide alignment by PAL2NAL (Suyama et al., 2006). The alignments were filtered if the position has >50% of gap or the length of the alignment block is smaller than 5 bp by Gblocks (Castresana, 2000). The RLSB tree was constructed using the maximum likelihood (ML) approach with the general time-reversible (GTR) substitution+Γ (gamma) model and 1000 bootstrap replicates by RAxML (Stamatakis, 2014). The proper substitution model was selected by PartitionFinder 2 (Lanfear et al., 2017) among the GTR, GTR+Γ, and GTR+Γ+I (inverse). The tree was rooted to outgroup species (Supplementary Table S3) RLSB genes which were downloaded from BLAST4OneKP (Matasci et al., 2014).

Protein extraction and analysis

Total protein extracts were prepared from Flaveria leaves, hypocotyls, flowers, roots, and stems as described (Bowman et al., 2013). Equal amounts of total protein were loaded into lanes of an SDS–polyacrylamide gel for analysis by immunoblotting. Gels were electrophoresed to either polyvinylidene fluoride (PVDF) (Bio-Rad) or nitrocellulose membranes (GE Healthcare), and then reacted with antisera for RLSB, Rubisco LSU, or PEPCase (Bowman et al., 2013). Images were acquired and analyzed using a Bio-Rad Gel Doc™ XR + System with Image Lab™ Image Capture and Analysis Software.

In vivo protein synthesis in light-grown, etiolated, or 48 h light-transferred (greening) F. pringlei (C₄) and F. bidentis (C₃) was performed as described previously (Berry et al., 1985, 1990). Briefly, freshly cut hypocotyls (immersed in water and cut with a scalpel midway between the root and cotyledons) were placed into 500 μl of labeling solution consisting of 100 μCi of [35S]methionine/cysteine Express Labeling Mix (PerkinElmer NEN Radiochemicals) and 400 μl of water. After 1 h incubation either in complete darkness or under standard lighting conditions, labeled protein extracts were prepared from hypocotyls with both cotyledons (with the stem cut off just below the cotyledons after labeling) using equal wet weight of material, cleared in a microfuge to remove insolubles, and stored at –20 °C. Equal amounts of labeled total protein extracts from both species and all three illumination conditions were immunoprecipitated by incubating with RLSB, LSU, and PEPCase antisera overnight with rotational mixing at 4 °C, and antigen–antibody complexes were precipitated using Pansorbin Staph A cells (Millipore) as described (Berry et al., 1985, 1990; Bowman et al., 2013). The immunoprecipitates were separated by SDS–PAGE and the labeled proteins were visualized using a Storm™ phosphorimager with ImageQuant software version 4 (GE Healthcare).

Plant material, growth conditions, and tissue sampling

Seed for F. pringlei (C₄), F. robusta, F. linearis (C₄–C₃), F. palmeri (C₃-like), and F. bidentis (C₃) were obtained from Dr Rowan Sage, University of Toronto. For standard growth conditions, seeds were germinated and plants were grown in a greenhouse using artificial soil with a 14 h d⁻¹ cycle under 170–280 μmol photons m⁻² s⁻¹. Leaf immunolocalizations used regions midway between the apex and base of young fully expanded 4 cm long leaves, collected from the third node below the apical meristem. Leaf, stem, flowers, and root tissues for immunoblot analysis were collected from 10-week-old plants. Stem samples were taken 10 cm down from the apical meristem. Whole flowers (all four whorls) were used since the small size and compactness of Flaveria flowers made it difficult to separate individual flower parts. Root samples consisted of both primary and lateral clippings.

For light regulation studies, etiolated ‘dark-grown’ F. pringlei (C₄) and F. bidentis (C₃) were germinated and grown for 10–14 d in light-proof containers within a dark room (Berry et al., 1988, 1990). Hypocotyls with both cotyledons (stem cut with a scalpel just below the cotyledons) were harvested from etiolated plants under a dim green safelight and immediately placed into a ground glass tissue homogenizer on ice with appropriate buffers for protein or RNA extraction (Bowman et al., 2013). In parallel, hypocotyls/cotyledons were harvested from light-grown Flaveria germinated and grown for 10–14 d in a growth chamber with standard lighting. For plants transferred from dark to light (‘greening’), etiolated Flaveria seedlings grown in darkness as described above for 8–12 d were transferred into the illuminated growth chamber for 48 h, after which the green hypocotyls/cotyledons were harvested and frozen.

Immunolocalization

Serial sectioning of leaf mid-regions from Flaveria species described above were prepared for immunolocalization as described (Bowman et al., 2013). The sections were reacted with RLSB, Rubisco LSU, or PEPCase primary antisera at 4 °C overnight (Berry et al., 1988; Bowman et al., 2013), and then reacted with secondary goat anti-rabbit antisera conjugated to Alexa Fluor® 546 (Life Technologies) for 1 h at room temperature. Visualization and image analysis was performed with an LSM710-InTune Confocal Microscope System using the ×20 objective. A 529 nm laser was used for excitation of Alexa Fluor® 546, and emission was collected at 564–577 nm. Images were processed and analyzed using Zen Imaging software (Carl Zeiss).
RNA isolation and real-time quantitative PCR

RNA was isolated from hypocotyls of light- and dark-grown (etiolated) *F. pringlei* and *F. bidentis* using an RNeasy Plant Mini Kit (Qiagen) according to the manufacturer’s protocol. cDNA synthesis was performed using an iScript cDNA synthesis kit (Bio-Rad) with oligo(dT) and random primers included with the kit. Quantitative reverse transcription–PCR (qRT–PCR) was performed with SYBR Green Supermix (Bio-Rad) on a MyiQ™2 Two Color Real-Time PCR Detection System (Bio-Rad), using primers specific for each transcript analyzed (Supplementary Table S4), as indicated in the figures. For each sample, reactions were normalized to actin mRNA using \(2^{-\Delta\Delta C_t}\) calculations, and statistical significance was calculated using Student’s \(t\)-test. For each bar shown in the graphs, \(P\)-values were <0.05. All data shown represent at least two technical repeats of three independent experiments.

**Results**

**RLSB orthologs in *Flaveria* species**

RLSB orthologs were identified within available transcriptomes from 16 species of the dicot genus *Flaveria* showing \(C_3\), \(C_3\)-\(C_4\) intermediate, \(C_4\)-like, or fully \(C_4\) characteristics. The predicted RLSB protein sequences used for alignment (Supplementary Fig. S1) were all highly conserved along their entire lengths (>70% overall identity), with the most conserved regions (>90%) occurring within the S1 binding domain (red highlight in Supplementary Fig. S1). Gaps at the N-terminal portion of the alignments occur primarily within the plastid transit sequence, so that the size of the proteins showed some variation between species (ranging from 475 to 502 amino acids). Their strong overall similarity with homologs characterized in Arabidopsis and maize suggests they may be associated with the same *rbcL* regulatory activity described for those plant species (Bowman et al., 2013). These findings are consistent with a previous study that identified strongly conserved RLSB homologs within such diverse plant species as the basal angiosperm *A. trichopoda*, the \(C_3\) dicot Arabidopsis, and the \(C_4\) monocot *Z. mays*, even extending as far as Charophyte algae that are considered to be most closely related to the common ancestor of all land plants (Supplementary Fig. S1; Yerramsetty et al., 2016). Previous findings have shown that as with most eudicots, RLSB occurs as a single-copy gene in *Flaveria* species (Yerramsetty et al., 2016).

A phylogenetic tree based on available RLSB sequences from 16 *Flaveria* species shows that divergence between homologs, while very low, does show some correlation with \(C_3\) to \(C_4\) evolution within this genus (Fig. 1). Based on this phylogenetic tree, RLSB sequences of basal species *F. robusta* and *F. cronquistii*, *F. pringlei* representing the ancestral \(C_3\) state of this genus (McKown et al., 2005; Kümpers et al., 2017), and the \(C_3\)-\(C_4\) *F. angustifolia* and *F. sonorensis*, representing the first step toward \(C_4\), are more similar to each other than they are to RLSB sequences of the \(C_3\)-\(C_4\) species (*F. chlorifolia*, *F. floridana*, *F. pubescens*, *F. anomala*, and *F. ramosissima*).

---

**Fig. 1.** Phylogeny of RLSB orthologs of 16 *Flaveria* (\(C_3\), \(C_3\)-\(C_4\)/\(C_2\), \(C_4\)-like, and \(C_4\)) and three outgroup species in Asteraceae. A phylogenetic tree was constructed from protein and coding sequences using maximum likelihood (RAxML) to determine the distance. The tree was rooted to outgroup species belonging to Asteraceae. Numbers on internal nodes indicate support values with 1000 bootstrap samples, and bootstrap values ≥70 are shown. Branch lengths indicate the number of nucleotide substitutions, and the scale bar represents 0.01 substitutions. The leaf labels on the left along with the species name show photosynthesis status. The main groups are defined on the right according to their clade. (This figure is available in colour at *JXB* online.)
Similar to other studies, the C_4-like species *F. brownii* is closer to the C_3–C_4 like group, and does not group with other C_4-like or C_4 species. This finding is consistent with previous predictions that there were at least two independent evolutionary events towards C_4-like and C_4 photosynthesis within the genus *Flaveria* (Powell, 1978; McKown et al., 2005), and suggests the possibility that RLSB might have played a role in these events. RLSB homologs of the later stage C_4-like *F. palmeri* or *F. vaginata*, and the fully C_4 *F. bidentis*, *F. kochiana*, and *F. trinervia*, show increasingly more divergence from the earlier species. The fact that C_4-like *F. vaginata* and *F. palmeri* are closer to completely C_4 *F. bidentis* and *F. kochiana*, respectively, suggests the presence of more than one evolutionary event that led to the establishment of complete ‘C_4-ness’ in this genus. The short branch lengths of C_4-like species *F. vaginata* and *F. palmeri* may also suggest reversal events which led to establishment of a C_4-like state from the advanced C_3 state, although this cannot be said with certainty based on just this level of analysis. While this phylogeny based solely on RLSB sequences shows some variation from other *Flaveria* phylogenies (Kopriva et al., 1996; McKown et al., 2005; Mallman et al., 2014; Lyu et al., 2015), the overall trend indicates that small alterations in RLSB sequence accompanied the progression from ancestral C_3 towards fully C_4 photosynthesis within the genus *Flaveria*. It is possible that such variations in sequence, although not pronounced, could reflect functional or accumulation differences of RLSB orthologs among the different photosynthetic states.

**RLSB co-localizes with Rubisco LSU across the C_3 to C_4 species evolutionary gradient**

As a contributing factor in the post-transcriptional regulation of rbcL expression, RLSB could have been important for establishing BS cell-specific Rubisco localization during C_4 evolution. The genus *Flaveria* provides an excellent model system to test this hypothesis, due to the presence of multiple photosynthetic states associated with the evolutionary transition (McKown and Dengler, 2007; Mallman et al., 2014; Lyu et al., 2015). For this study, immunolocalization was used to characterize RLSB and Rubisco LSU accumulation patterns in the following representative species of *Flaveria*: *F. robusta* and *F. pringlei*, representing C_3 type, *F. linearis* representing C_3–C_4 intermediate, *F. palmeri* with C_4-like traits, and *F. bidentis* representing a fully C_4 species. These species were chosen for this study based on their photosynthetic type, distinctive leaf morphologies, their availability, and being highly amenable to leaf sectioning and antibody labeling.

In the C_3 species *F. robusta* and *F. pringlei*, RLSB and LSU were co-localized and equally distributed within all of the leaf chlorenchyma cells, occurring within the large number of chloroplasts within these cells (Fig. 2). Reminiscent of RLSB and LSU accumulation in the C_3 dicot Arabidopsis (Bowman et al., 2013), there was no cell type specificity observed among the photosynthetic leaf cells of either plant. The leaves of these species showed typical C_3 anatomy with no Kranz-like features.

In the C_3–C_4 intermediate *F. linearis*, leaf cross-sections revealed rudimentary ‘proto’-Kranz-like anatomical features, with discernible BS cells surrounding the veins (Fig. 3). Within these BS cells (Fig. 3), the centripetal localization of RLSB- and LSU-containing chloroplasts was evident (Fig. 3). There was some preferential accumulation of these proteins within the BS cells, relative to the surrounding M (chlorenchyma) cells, based on the stronger fluorescence signal detected within the BS cells. The cell-type-preferential compartmentalization of the RLSB and LSU proteins in the leaves of this intermediate species convergent with proto-C_4-like anatomy appears to represent one step towards the evolution of C_4 from C_3 Rubisco production. Thus in leaves with partial C_4 photosynthesis, there is evidence for the evolutionary beginnings of preferential RLSB and LSU expression within the internal BS cells in leaves of this C_3–C_4 intermediate plant.

Fig. 2. Immunolocalization of RLSB and Rubisco LSU in leaves of C_3 *Flaveria* species. Confocal images of *F. pringlei* (left column) and *F. robusta* (right column) adjacent serial cross-sections taken from a region midway between the leaf apex and base. Sections were incubated with the indicated primary antiserum, and then reacted with Alexafluor 546-conjugated secondary antibody. Images were captured using the x20 objective of a LSM 710 ‘in tune’ confocal microscope. Images show localization of RLSB (top panels) and Rubisco LSU (bottom panels). Note that these images show typical C_3 leaf anatomy, with both photosynthetic proteins distributed throughout all of the leaf chlorenchyma cells. Significant anatomical features identified in the leaf cross-sections are labeled as follows: V, vascular bundles; ue, upper epidermis; le, lower epidermis; C, chlorenchyma cells that harbor chloroplasts. Scale bar=150 µm. (This figure is available in colour at JXB online.)
Rubisco has become compartmentalized to the inner ring of morphologically distinct BS cells. RLSB localization also followed this pattern; similar to LSU, most of the RLSB protein accumulation was observed in BS cells, with low but still easily detectable levels occurring in the M cells (Fig. 4, left middle panel). Consistent with previous findings (Bowman et al., 2013), leaves of the fully C₄ F. bidentis showed both LSU and RLSB proteins exclusively localized within chloroplasts of the fully differentiated leaf BS cells, with little or no accumulation within M cells of this species (Fig. 4, right top and middle panels).

For a cell specificity comparison, in both C₄-like F. palmeri and the fully C₄ F. bidentis, PEPCase protein accumulation was highly specific to the cytoplasm of M cells, and not BS cells, in the characteristic C₄ pattern. Thus, cell specificity for PEPCase was observed in corresponding leaf sections of both photosynthetic types, while RLSB and Rubisco specificity was observed only in the fully C₄ F. bidentis. It is likely that processes responsible for the establishment of complete M cell specificity occur independently and at an earlier stage than those responsible for BS cell specificity along the C₃ to C₄ evolutionary progression.

**RLSB and LSU accumulation in photosynthetic and non-photosynthetic tissues of C₃ and C₄ Flaveria species**

As a potential positive regulator of *rbcL* expression, RLSB would be expected to show the same tissue-specific patterns of accumulation as the protein it regulates, the Rubisco LSU, throughout different photosynthetic and non-photosynthetic plant tissues. To determine if RLSB and Rubisco accumulation correlate in tissues other than leaves, and if tissue specificity patterns are conserved among plant species using the different photosynthetic pathways, immunoblot analysis was performed using protein extracts from seeds, flowers, leaves, stems, and roots of C₃ F. pringlei and C₄ F. bidentis (Fig. 5). PEPCase, used as a loading control for the *F. pringlei* tissues, is a low abundance non-photosynthetic protein in leaves and other tissues of C₃ species (Berry et al., 2011, 2013). To better visualize this protein, the gel used for the PEPCase immunoblot shown in Fig. 5 was loaded with a higher concentration (10-fold higher) of the equalized protein extracts than used for the RLSB and LSU gels. It should be noted that while at least some C₃ and all C₄ plants contain different forms of PEPCase (Berry et al., 2011; Ruckle et al., 2007; Burgess et al., 2016), the polyclonal antiserum used here was prepared against full-length protein purified from amaranth leaves (Wang et al., 1992) and would not distinguish between different forms in western analysis. For C₄ F. bidentis, the NAD-dependent malic enzyme (NAD-ME), a non-photosynthetic protein in this C₄ species, was used as a loading control.

The tissue-specific accumulation patterns for RLSB and LSU did not vary between C₃ and C₄ species. Both proteins were most abundant in leaves and in green stems, both of which are photosynthetic tissues (Fig. 5). These proteins were also detected in flowers, most probably due to the presence in these extracts of green tissue from the calyx. No RLSB
protein was detected in the non-photosynthetic seeds and roots, where LSU was also absent. In contrast, the non-photosynthetic PEPCase and NAD-ME proteins were found at approximately equal levels in each tissue type of *F. pringlei* and *F. bidentis*, respectively.

The close correlation between RLSB and LSU accumulation in these different plant tissues provides further support for RLSB as a key determinant of LSU production and localization. The modified accumulation patterns for the regulatory protein RLSB between C\textsubscript{3} and C\textsubscript{4} plants, like Rubisco, occurred only in leaves (from Figs 2–4). Thus, the co-ordinated C\textsubscript{3} to C\textsubscript{4} evolutionary progression of RLSB and LSU cell type expression patterns in the direction of enhanced ‘C\textsubscript{4}-ness’ was an occurrence specific to leaf cells that did not affect the accumulation of these proteins in any tissues other than leaves.

**Light regulation of RLSB in C\textsubscript{3} and C\textsubscript{4} Flaveria species**

The post-transcriptional regulation of photosynthetic gene expression by light has been well documented and is especially prominent in the regulation of *rbcL* and other plastid-encoded genes (Patel and Berry, 2008; Berry *et al.*, 2011, 2013, 2016). To determine if the nuclear-encoded RLSB *rbcL*...
mRNA-binding protein itself is regulated by light, and if this regulation has been modified during the transition from C3 to C4 photosynthesis, the accumulation and synthesis of RBCL and LSU proteins was examined in hypocotyls from light-grown, dark-grown (etiolated), and 48 h ‘greening’ F. pringlei (C3) and F. bidentis (C4) plants. Hypocotyls were used for the light regulation experiments because the etiolated Flaveria seedlings did not produce leaves. Previous studies from our laboratory demonstrated that these tissues provide an excellent system to study light-mediated gene expression (reviewed in Patel and Berry, 2008; Berry et al., 2011, 2013). For the experiments shown in Fig. 6, seeds of each species were germinated and grown under normal illumination (Light), in complete darkness (Dark), or in darkness and then transferred to light for 48 h (Greening), as described in the Materials and methods. For both Flaveria species, we found that 10–14 d growth of seedlings under light and dark conditions was optimal for experimental viability, light responsiveness, and cotyledon development.

For immunoblot analysis (Fig. 6A), equal amounts of total protein from hypocotyls of each species grown under the conditions indicated were loaded and separated by SDS–PAGE, and subjected to immunoblot analysis as described for Fig. 5 (including using higher concentration of extracts to detect PEPCase in C3 F. pringlei). For both the C3 and C4 Flaveria species, there was substantially more LSU and RLSB in light-grown seedlings, relative to plants of the same age grown in total darkness. When the etiolated plants were transferred to light for 48 h, levels of both proteins increased to levels observed in seedlings grown under normal illumination conditions. The reduction in RLSB in response to darkness (etiolation) occurred in both species, although to a lesser degree in the C4 species relative to the C3 species (Fig. 6A). LSU accumulation in these species reflected RLSB accumulation, with significant reductions in the dark-grown seedlings relative to light-grown, and an increase in dark-grown seedlings after 48 h transfer to light (Fig. 6A).

In vivo protein synthesis was analyzed by using seedlings of F. pringlei (C3) and F. bidentis (C4) that were labeled with [35S]methionine/cysteine while growing under the different illumination conditions. Labeled proteins were extracted from the hypocotyls, and LSU, RLSB, or PEPCase were immunoprecipitated from labeled total protein. In both the C3 and C4 species, the light-associated changes in RLSB and LSU accumulation were mirrored by corresponding changes in in vivo synthesis for each protein. For.
both species, in vivo synthesis of LSU and RLSB was easily detectable in light-grown seedlings (Fig. 6B), while synthesis of both proteins was significantly reduced in the dark-grown seedlings. After transfer of etiolated seedlings to light, levels of LSU and RLSB synthesis increased, reaching normal light-grown levels by 48 h following transfer.

Unlike RLSB and LSU, the amounts and synthesis of non-photosynthetic PEPCase in the C₃ F. pringlei hypocots was not affected by changes in illumination (Fig. 6A, B, bottom left panels). In contrast, in C₄ F. bidentis, accumulation and synthesis of photosynthetic PEPCase were reduced in darkness, relative to light, and increased in response to 48 h of illumination (Fig. 6A, B, bottom right panels). This is in agreement with previous studies showing that the photosynthetic form of this enzyme acquired light regulation during its modification from metabolic to photosynthetic function during C₄ evolution (Berry et al., 2011, 2013). However, it is important to note that 48 h after transfer to light increased PEPCase and its synthesis had not reached the more abundant levels observed in the hypocots of light-grown seedlings. Therefore, while light regulation is a characteristic of all three photosynthetic proteins in C₄ F. bidentis, for these tissues and conditions, light-induced synthesis of photosynthetic PEPCase appears to lag behind RLSB and LSU, probably requiring longer period of growth under illumination to achieve normal levels of synthesis and accumulation.

Taken together, it is apparent that at the levels of protein accumulation and synthesis, nuclear-encoded RLSB is similar to the chloroplast-encoded LSU, with production of both proteins being light regulated in C₃ and C₄ plants. The shared light regulation of RLSB at opposite ends of the C₃ to C₄ evolutionary spectrum differs from PEPCase, a nuclear-encoded protein recruited to a photosynthetic function that shows light-regulated production only in the C₄ species.

Analysis of RLSB and rbcL mRNA using qRT–PCR showed that in both Flaveria species, as in the C₃ dicot Arabidopsis and the C₄ monocot maize (Bowman et al., 2013), mRNAs encoding RLSB are much less abundant than those encoding rbcL. (note the difference in y-axis scales for RLSB and rbcL in Fig. 7). In both C₃ F. pringlei and C₄ F. bidentis, the abundance of rbcL and RLSB mRNAs was not affected by illumination, with approximately equal levels of each transcript present in both light-grown and dark-grown plants (Fig. 7). This is in clear contrast to the RLSB and

![Fig. 7. RLSB and rbcL mRNA in F. pringlei (C₃) and F. bidentis (C₄) in response to light. For each panel, RNA was extracted from hypocots of F. pringlei (C₃) and F. bidentis (C₄) seedlings grown in light or complete darkness (etiolated). Transcript levels of both genes were analyzed by qRT–PCR using primers specific for each sequence. Quantification of transcript levels was standardized to actin mRNA. (A) Relative levels of RLSB and rbcL mRNA accumulation in F. pringlei seedlings grown under the conditions indicated. (B) Relative levels of RLSB and rbcL mRNA in F. bidentis seedlings grown under the conditions indicated. Data in both panels represent at least two technical repeat reactions of three independent experiments. Note differences in scale for panels showing RLSB and rbcL mRNAs, indicating that these two transcripts accumulate to substantially different levels in both plants under each condition. (This figure is available in colour at JXB online.)](https://academic.oup.com/jxb/article-abstract/68/16/4635/4079623)
LSU proteins, both of which showed significant reductions in accumulation and synthesis in dark-grown seedlings relative to those grown in light (Fig. 6A, B). The lack of correlation between the accumulation of transcripts and their encoded proteins is indicative of regulation at the level of translation, or possibly protein stability. Thus, like many proteins associated with C$_3$ and C$_4$ photosynthetic processes, including the Rubisco LSU and SSU subunits (Patel and Berry, 2008; Berry et al., 2013, 2016), RLSB expression/accumulation appears to be post-transcriptionally regulated by light.

**Discussion**

**RLSB and C$_4$ evolution**

Plastid-encoded genes are regulated at multiple steps, from transcription, RNA processing, transcript stabilization, and translation (Raynaud et al., 2007; Tillich et al., 2010; Barkan, 2011; Berry et al., 2013). Each of these steps involves multi-component complexes of interacting proteins, many of which are encoded in the nucleus as participants in anterograde signaling and gene regulation. Our previous findings demonstrated that reduced RLSB leads to corresponding reductions in rbcL mRNA (Bowman et al., 2013), providing evidence that this protein in itself is a determinant of transcript stability. RLSB was purified based on its ability to bind the 5’ portion of this mRNA (Bowman et al., 2013), where sequences required for stability/degradation of rbcL and other plastid transcripts are located (Salvador et al., 2011; Berry et al., 2013). However, RLSB probably does not function on its own. We suspect that other proteins, some cell or tissue specific and others more general, might interact with RLSB as part of a complex to mediate final translation of LSU protein from stabilized rbcL mRNA. Reductions in RLSB or any other single component would lead to inactivation/destabilization of the entire complex, causing the observed decreases in rbcL mRNA and LSU protein observed in our previous studies. This phenomenon occurs for many interacting proteins, including Rubisco and other chloroplast complexes (Choquet et al., 2003; Cohen et al., 2005, 2006; Duncan and Mata, 2011). During C$_4$ evolution, only one key component of such a regulatory complex, such as RLSB, would need to become down-regulated in M cells to achieve BS cell-specific activity. As a conceptual example, the BS cell-specific localization of only one of four GDC subunits, the P subunit, in C$_3$–C$_4$ intermediates of Flaveria and other genera demonstrates how a single component might become the cell specificity determinant for an entire complex (Sage et al., 2012; Schulze et al., 2013; Khoshravesh et al., 2016; Schulze et al., 2016). Similarly, reducing the expression of only the H subunit in M cells of transgenic rice correspondingly reduces GDC activity within those cells (Lin et al., 2016). On the other hand, any potential overproduction of RLSB alone in any tissues or conditions might not have an effect on overall Rubisco production, at least not a complete effect. Increased Rubisco holoenzyme might not occur without co-overexpression with other interacting components, or even the SSU to interact with and stabilize any excess LSU produced.

As a post-transcriptional regulatory component of rbcL mRNA metabolism, RLSB would probably be involved with many cell- and tissue-specific aspects of Rubisco gene expression in C$_3$ and C$_4$ plant species (Bowman et al., 2013; Rosnow et al., 2014; Yerramsetty et al., 2016). This is especially significant for the evolution of C$_4$ photosynthesis where nuclear-encoded RLSB, through anterograde regulation, could assist in localizing the expression of the chloroplast-encoded rbcL gene and therefore overall Rubisco production to leaf BS cells. According to this model, cell type-specific regulation of plastid genes would be anchored to cell-specific expression of nuclear-encoded regulatory genes, thereby co-ordinating cell specificity between genes encoded within the different cell compartments. Other mechanisms are also likely to be involved in co-ordinating the two compartments, including retrograde signaling and light regulation (Berry et al., 2013; Burgess et al., 2016). This hypothesis is supported by findings presented in the current study, in which RLSB and Rubisco co-localization was observed in several members of the genus Flaveria that display different patterns of Rubisco localization depending on the photosynthetic pathway utilized. Co-localization of LSU with its associated transcript-binding protein RLSB occurred within mature leaves across a range of Flaveria species representing C$_3$, C$_3$–C$_4$, C$_4$-like, and C$_4$ photosynthetic types. This tight association across an evolutionary continuum of ‘C$_3$ to C$_4$-ness’ provides evidence that RLSB localization has laid down the pattern of Rubisco localization in leaves of the different photosynthetic types, gradually leading to the confinement of Rubisco to BS cells in Kranz-type species with full C$_4$ development. Under this scenario, the rbcL regulatory protein RLSB is proposed to have played a role in the evolutionary transition from C$_3$ to C$_4$, with its localization essential for the downstream BS cell-specific localization patterns exhibited by Rubisco.

This hypothesis is also consistent with the fact that, like RLSB (Yerramsetty et al., 2016), rbcL regulatory and coding sequences are for the most part very conserved throughout all higher plants (Manen et al. 1994; Salvolainen et al., 2000; Kapralov et al., 2010, 2011; Sharwood et al., 2016). rbcL genes in different plant species do show variations that occur primarily within their coding sequences, with amino acid changes affecting holoenzyme assembly, interactions with Rubisco activase, and enzyme kinetics (Kapralov et al., 2010, 2011; Sage et al., 2012; Sharwood et al., 2016). Some of these have been linked with functional adaptations during the evolutionary transition from C$_3$ to C$_4$ photosynthesis. However, changes to the rbcL coding sequence itself would probably occur independently from modifications in C$_3$ to C$_4$ gene expression patterns. In fact, non-coding regulatory regions of this gene appear to have been highly conserved, at least among dicots (Manen et al., 1994).

A recent study indicated that Arabidopsis RLSB (designated as PRB1 in that study) shows in vitro interactions with biotinylated ycf1 mRNA (Yang et al., 2016). In our previous study, we suggested that RLSB could interact with and regulate one or more plastid mRNAs, in addition to rbcL, that were not included in our in vitro or in vivo binding analysis (Bowman et al., 2013). Plastid-encoded ycf1 is an essential
cell viability gene in many plant species (Asakura and Barkan, 2006; Bölt er and Soll, 2017), but it does not occur in all plants. For example, the ycf1 gene appears to have been evolutionarily deleted from chloroplast genomes of most grasses, including C4 maize (Maier et al., 1995; Drescher et al., 2000; Asakura and Barkan, 2006) which has two RLSB paralogs (Yerramsetty et al., 2016). The function of ycf1 has not been clearly established in any plant, and a potential role for this protein in C4 capability or evolution is not known.

In the two C3 species, F. robusta and F. pringlei, both RLSB and LSU were found within chloroplasts that were distributed throughout all of the leaf chlorenchyma cells, and were not specific to any one photosynthetic cell type (Fig. 2). In the C3-C4 intermediate F. linearis, some C4-like anatomical features, such as the presence of morphologically distinguishable M cells and BS cells, were clearly apparent (Fig. 3). In these leaves, proto-Kranz BS cells were located immediately surrounding the vascular bundles. These cells were larger and contained more chloroplasts than the adjacent M cells. The reduction in M cell chloroplast number in this species relative to BS cells (clearly observable in Fig. 3B) is characteristic of this early stage towards increased C4-ness (Stata et al., 2014; Khosh ravesh et al., 2016; Lin et al., 2016). In addition to the more numerous chloroplasts, BS cells of F. linearis leaves showed increased fluorescence signal representing increased amounts of chloroplast-localized RLSB and LSU proteins relative to the M cells, indicating co-ordinated BS-preferential accumulation of both proteins at this intermediate stage of C4 evolutionary development (Fig. 3). As the C4-ness increases, the M cell-associated reduction in chloroplast density becomes magnified, as observed in the C4-like species F. palmeri and the fully C4 species F. bidentis (Fig. 4). In leaves of the C4-like F. palmeri (Fig. 4, left panels) the formation of well-defined Kranz anatomy was clearly observable, with the majority of RLSB and LSU being localized to the BS chloroplasts. In these cells, cell specificity for LSU and the rbcL regulator RLSB was not complete, with low levels of both proteins still found within the M cell chloroplasts. In F. bidentis (Fig. 4, right panels), the species showing the most advanced full C4 stage of evolutionary development, Rubisco and RLSB were both highly specific to leaf BS cells, with little if any of either protein observed within the fully differentiated M cells. At this final stage along the C3-C4 species gradient, specific localization of RLSB and LSU to leaf BS cells is complete. These observed evolutionary changes exhibited by both RLSB and LSU across these different Flaveria photosynthetic types provides strong evidence that C4 evolution has incorporated cell-specific modifications to genes encoding both proteins, the nuclear-encoded regulatory protein RLSB and the chloroplast gene it regulates, rbcL. These modifications may have occurred in co-ordination with modifications to leaf anatomy, since the small anatomical changes in the C3-C4 intermediate F. linearis were accompanied by changes in the localization of these two proteins.

The progression towards C4-ness in Flaveria species is associated with changes in the CO2 compensation point, with values of C3-C4 species approximately midway between those of C3 and fully C4 species (Holaday et al., 1984; Edwards and Voznesenskaya, 2011; Sage et al., 2012; Mallmann et al., 2014; Khosh ravesh et al., 2016). Reduced photorespiration is another factor associated with the degree of C4-ness, and the photorespiration avoidance efficiencies of C3-C4 intermediates of Flaveria also lie in between the true C3 and fully C4 species (Sage et al., 2012; Schulze et al., 2013; Mallmann et al., 2014; Khosh ravesh et al., 2016; Schulze et al., 2016). Enhanced photosynthetic efficiency based on a lowered CO2 compensation point and reduced photorespiration are directly related to the internalization of Rubisco within C4 leaves. In Kranz species, the only way for Rubisco to become BS cell specific is by modification of default Rubisco gene expression in C3 plants to BS cell specificity in C4 plants. Post-transcriptional regulation of Rubisco gene expression is likely to play a major role in this process, probably mediated by the rbcL mRNA-binding protein RLSB (Patel and Berry, 2008; Hibberd and Covshoff, 2010; Berry et al., 2013, 2016; Bowman et al., 2013). It is notable that recent translatome data have shown that RLSB is not preferentially expressed in C4 Arabidopsis BS cells (Aubry et al., 2014). This is consistent with our findings of a progressive C3 to C4 evolutionary transition towards BS specificity for this mRNA-binding protein in Flaveria species.

While our previous studies (Bowman et al., 2013; Rosnow et al., 2014) demonstrated the co-localization for RLSB and Rubisco in mature leaves of several C3 and C4 plants, they did not address if there was any correlation in how the localization patterns developed. If these had in fact occurred independently (i.e. with no progression or co-ordination), this would suggest that RLSB was not associated with the progressive evolution of C4 Rubisco localization, and that their co-localization in mature C3, as well as Kranz-type and single-cell C4 leaves might be more circumstantial and possibly unrelated. In fact, the progressive correlation for RLSB and Rubisco localization across the C3 to C4 spectrum was striking. Changes in localization for both proteins occurred together, with the incremental morphological development of Kranz anatomy and correspond to other progressive changes known to be associated with C4 evolution (Holaday et al., 1984; Edwards and Voznesenskaya, 2011; Sage et al., 2012; Mallmann et al., 2014; Khosh ravesh et al., 2016). Considered together, these findings suggest that change in localization of RLSB-binding protein was one prominent factor, integrated and working in unison with other molecular, physiological, and morphological processes, during the evolutionary progression leading from C3 to full C4 capability.

RLSB is associated with tissue specificity and light regulation of rbcL expression in C3 and C4 Flaveria

Tissue-specific RLSB and Rubisco LSU accumulation patterns in the C3 and C4 Flaveria species mirror patterns of mRNA accumulation observed for the C3 dicot Arabidopsis (Bowman et al., 2013). Rubisco LSU and RLSB were found to accumulate only in photosynthetic tissues, which included leaves, green stems, and flowers with the green calyx. The finding that RLSB, like LSU, was only in these same green tissues, and not in non-green tissues such as roots and seed,
provides further evidence for its role as a regulator of photosynthetic activity (Berry et al., 2013; Bowman et al., 2013; Yerramsetty et al., 2016). Although RLSB accumulation is BS cell specific in the fully C₄ species, it has retained the same conserved pattern of accumulation in the stems, leaves, and flower sepals as the C₃ species. Whatever regulatory modification was responsible for restricting RLSB/rbcL expression to BS cells during C₃ to C₄ evolution appears to have occurred only in leaves, without affecting overall tissue-specific accumulation patterns shared with other plant species. Consistent with our hypothesis, this finding suggests tissue-specific accumulation patterns of Rubisco accumulation may be defined by patterns of RLSB accumulation in the different plant tissues. Furthermore, regulatory processes responsible for limiting RLSB accumulation to C₄ BS cells are probably separable from those that limit its expression to photosynthetic tissues in both C₃ and C₄ species, and were not affected by the evolutionary transition to the more evolutionarily derived photosynthetic pathway.

LSU and RLSB protein accumulation (Fig. 6A) and in vivo synthesis (Fig. 6B) were found to be light dependent in hypocotyls of C₃ and C₄ Flaveria species. This is consistent with previous studies demonstrating light regulation of rbcL expression in many plant species (Patel and Berry, 2008; Berry et al., 2013, 2016). As with cell type and tissue type specificity, light-induced changes in LSU accumulation were mirrored by changes in the accumulation and synthesis of RLSB. These results clearly show that the tight association between RLSB and Rubisco production is maintained in different light conditions for both C₃ and C₄ photosynthetic types.

As expected from other studies (Kausch et al., 2001; Berry et al., 2011), synthesis and accumulation of non-photosynthetic PEPCase protein in hypocotyls of C₃ F. pringlei was not regulated by light, while the PEPCase in C₄ F. bidentis did show light regulation. The accumulation of photosynthetic PEPCase is known to be up-regulated by light in C₄ plants, due primarily to regulation of transcription. However, light-induced expression of photosynthetic PEPCase was delayed in hypocotyls of F. bidentis, so that a longer period of greening may be required for induction of this M cell-specific gene, at least in these tissues. Such findings suggest that the acquisition of light-regulated protein production for the nuclear-encoded RLSB in C₃ and C₄ plants and nuclear-encoded photosynthetic PEPCase in C₄ plants were non-synchronous events that occurred independently during the evolutionary progression from C₃ to C₄ in Flaveria. In contrast to protein production, a recent study indicates that transcription and accumulation of PEPCase mRNA is light regulated in both C₃ Arabidopsis and C₄ Gynandropsis gynandra (Burgess et al., 2016). Different analytical approaches used in this current study versus the previous study, such as chromatin immunoprecipitation and sequencing analysis (ChIP-SEQ) versus western blot and in vivo protein synthesis in different species, suggest that transcriptional and post-transcriptional regulatory mechanisms determining final PEPCase levels may not be the same in all plants. Accumulated findings regarding PEPCase regulation in C₃ and C₄ plants are consistent with the model that genes encoding different C₄ photosynthesis proteins are regulated independently, involving both shared and divergent regulatory processes that vary between species (Hibberd and Covshoff, 2010; Berry et al., 2011, 2013; Burgess et al., 2016; Garner et al., 2016; Williams et al., 2016; Kümpers et al., 2017).

Similar to rbcL and RbcS in the C₄ dicot amaranth and other plants (Patel and Berry, 2008; Berry et al., 2013, 2016), differences in the synthesis and accumulation of Rubisco LSU and RLSB proteins in response to light and dark growth conditions did not correlate with levels of their corresponding transcripts for either C₃ F. pringlei (Fig. 7A) or C₄ F. bidentis (Fig. 7B). Nuclear-encoded RLSB mRNAs and plastid-encoded rbcL mRNAs were present at nearly identical levels in seedlings grown under normal illumination (when accumulation and synthesis of both proteins occurred) and in complete darkness (when their accumulation and synthesis did not occur). This lack of correlation between protein and transcript accumulation is indicative of post-transcriptional regulation at the level of translation, or possibly protein stability. Such regulation is characteristic of nuclear as well as plastid-encoded photosynthesis genes in many C₃ and C₄ plant species (Patel and Berry, 2008; Berry et al., 2011, 2013). This finding provides further evidence that the RLSB mRNA-binding protein is closely correlated with post-transcriptional rbcL expression, possessing regulatory properties characteristic of many other post-transcriptionally regulated photosynthetic genes.

Several studies have shown that Rubisco activity, protein levels, and transcript accumulation are reduced in the leaves of C₃ relative to C₄ plants, which contributes to the increased nitrogen-use efficiency of C₄ species (Patel and Berry, 2008; Gowik and Westhoff, 2011; Sage et al., 2012; Carmona-Silva et al., 2015; Garner et al., 2016; Kümpers et al., 2017). In Figs 6 and 7, similar levels of LSU protein synthesis and accumulation, as well as rbcL mRNA levels, were observed in light-grown and greening hypocotyls of both species. Disparity between this current and previous studies might be related to our use of hypocotyls rather than leaves for analysis of light regulation. For most dicots, true leaves do not develop on etiolated seedlings (Wang et al., 1992; Patel and Berry, 2008). However, the cotyledons from early seedlings have been shown to undergo light regulation for Rubisco and other photosynthetic genes (Berry et al., 1985; 1990; Wang et al., 1992; Patel and Berry, 2008). In amaranth, these also show similar patterns of C₄ development and cell specificity to leaves, but there are differences (Wang et al., 1992). Leaves originate from vegetative meristems and undergo several stages of growth, development, and differentiation. Cotyledons present on early hypocotyls develop from cell divisions that occur during seed development, with no cell division and limited morphological development after germination. Thus, while hypocotyls provide an excellent system for studying light regulation and early C₄ development, levels of Rubisco mRNA and protein production in these embryonically derived tissues may not necessarily correspond to those observed in mature C₃ and C₄ leaves.
RLSB as a unique determinant of post-transcriptional rbcL expression and Rubisco accumulation

Plants that utilize C₄ photosynthesis are critical for many agricultural and industrial applications, including food and biofuel production (Jones, 2011; Sage and Zhu, 2011; Von Caemmerer and Furbank, 2016; Sharwood et al., 2016), as well as lesser known applications such as the bourbon distillation industry (Arnold and Simanek, 2016). Although characteristics of C₄ plants have been known for decades, uncoating the molecular basis of C₄ photosynthesis remains an elusive goal (Hibberd and Cowshoff, 2010; Berry et al., 2011, 2013, 2016; Langdale, 2011; Huang and Brutnell, 2016). Many studies have provided evidence that post-transcriptional regulation plays a role in photosynthetic gene expression (Berry et al., 2011, 2013, 2016; Brown et al., 2011; Garner et al., 2016; Williams et al., 2016), and regulation of C₄ genes at this level may be more significant than previously thought. In support of this hypothesis, a recent transcriptome study found post-transcriptional regulation of mRNA stability to be more prominent in C₄ plants relative to C₃ plants (Fankhauser and Aubry, 2017). The results presented here provide further evidence for RLSB as a unique mRNA-binding protein involved with rbcL gene expression in all plants, with stepwise modifications in its leaf localization correlating with the acquisition of BS cell-specific Rubisco production along the C₃ to C₄ species gradient. These modifications are superimposed on the basic shared characteristics of tissue-specific and light-mediated control, which were not altered during C₄ evolution in this genus. As one of the few post-transcriptional regulatory components implicated in C₄ expression, the ancient highly conserved RLSB protein may serve as a paradigm for the identification, functional characterization, and evolutionary analysis of such regulators as studies into the origins and processes of this essential photosynthetic pathway move forward.

Supplementary data

Supplementary data are available at JXB online.

Fig. S1. Amino acid sequences of RLSB homologs in the genus *Flaveria*.

Fig. S2. Original unedited loading order of Fig. 6B.

Table S1. Transcriptome sequencing data of *Flaveria* in the NCBI database.

Table S2. *Flaveria* transcriptome assembly.

Table S3. Three outgroup species from BLAST4OneK.

Table S4. List of primer sequences used for this study.

Acknowledgements

We are grateful to Julian Hibberd for sharing the *Flaveria* transcriptomes. We thank Jim Stamos for help in preparing the figures, Alan Siegel for confocal microscopy support, Victor Albert for use of the PCR detection system, and Chris Mure for valuable discussion and insight. This work was initiated under support from USDA/NRI Grant 2008-01070 to JOB, and sustained with bridge funding from the University at Buffalo College of Arts and Sciences and the Department of Biological Sciences to JOB. PY was supported in part by a University at Buffalo Dissertation Fellowship WCY and JCC acknowledge support from the Department of Energy, Office of Science, Genomic Science Program under Award Number DE-SC0008834. The Zeiss LSM 710 ‘In Tune’ Confocal Microscope was purchased through NSF Major Research Instrumentation grant DBI 0923313.

References

Andersson I. 2008. Catalysis and regulation in Rubisco. Journal of Experimental Botany 59, 1555–1568.

Andersson I, Backlund A. 2008. Structure and function of Rubisco. Plant Physiology and Biochemistry 46, 275–291.

Arnold R, Simanek E. 2016. Rubisco and C₄ plants. In: Shots of knowledge, the science of whiskey. Fort Worth, TX: Texas Christian University Press, 30–31.

Asakura Y, Barkan A. 2006. Arabidopsis orthologs of maize chloroplast splicing factors promote splicing of orthologous and species-specific group II introns. Plant Physiolo 142, 1656–1663.

Aubry S, Smith-Unna RD, Boursnell CM, Koprina S, Hibberd JM. 2014. Transcript residency on ribosomes reveals a key role for the Arabidopsis thaliana bundle sheath in sulfur and glucosinolate metabolism. The Plant Journal 78, 659–673.

Barkan A. 2011. Expression of plastid genes: organelle-specific elaborations on a prokaryotic scaffold. Plant Physiology 155, 1520–1532.

Berry JO, Breiding DE, Klessig DF. 1990. Light-mediated control of translational initiation of ribulose-1, 5-bisphosphate carboxylase in amaranth cotyledons. The Plant Cell 2, 795–803.

Berry JO, Carr JP, Klessig DF. 1988. mRNAs encoding ribulose-1,5-bisphosphate carboxylase remain bound to polysomes but are not translated in amaranth seedlings transferred to darkness. Proceedings of the National Academy of Sciences, USA 85, 4190–4194.

Berry JO, Mure CM, Yerramsetty P. 2016. Regulation of Rubisco gene expression in C₄ plants. Current Opinion in Plant Biology 31, 23–28.

Berry JO, Nikolau BJ, Carr JP, Klessig DF. 1985. Transcriptional and post-transcriptional regulation of ribulose 1,5-bisphosphate carboxylase gene expression in light- and dark-grown amaranth cotyledons. Molecular and Cellular Biology 5, 2238–2246.

Berry JO, Yerramsetty P, Zielinski AM, Mure CM. 2013. Photosynthetic gene expression in higher plants. Photosynthesis Research 117, 91–120.

Berry JO, Zielinski AM, Patel M. 2011. Gene expression in mesophyll and bundle sheath cells of C₃ plants. In: Raghavendra AS, Sage RF, eds. C₃ photosynthesis and related CO₂ concentrating mechanisms. Advances in photosynthesis and respiration, Vol. 32. Dordecht, The Netherlands: Springer, 221–256.

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

Böttler B, Soll J. 2017. Ycf1/Tic214 is not essential for the accumulation of plastid proteins. Molecular Plant 10, 219–221.

Bowman SM, Patel M, Yerramsetty P, Mure CM, Zielinski AM, Bruenn JA, Berry JO. 2013. A novel RNA binding protein affects rbcL gene expression and is specific to bundle sheath chloroplasts in C₄ plants. BMC Plant Biology 13, 138.

Brown NJ, Newell CA, Stanley S, Chen JE, Perrin AJ, Kajala K, Hibberd JM. 2011. Independent and parallel recruitment of preexisting mechanisms underlying C₄ photosynthesis. Science 331, 1436–1439.

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.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421.

Carmo-Silva E, Scales JC, Madwick PJ, Parry MA. 2015. Optimizing Rubisco and its regulation for greater resource use efficiency. Plant, Cell and Environment 38, 1817–1832.

Castrigana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular Biology and Evolution 17, 540–552.

Choquet Y, Zito F, Wostrikoff K, Wollman FA. 2003. Cytochrome f translation in Chlamydomonas chloroplast is autoregulated by its carboxyl-terminal domain. The Plant Cell 15, 1443–1454.
Sage TL, Busch FA, Johnson DC, Friesen PC, Stinson CR, Staia M, Sultmanis S, Rahman BA, Rawsthorne S, Sage RF. 2013. Initial events during the evolution of C₄ photosynthesis in C₃ species of Flaveria. Plant Physiology 163, 1266–1276.

Salvador ML, Suay L, Klein U. 2011. Messenger RNA degradation is initiated at the 5’ end and follows sequence- and condition-dependent modes in chloroplasts. Nucleic Acids Research 39, 6213–6222.

Salvolainen V, Chase MW, Hoot SB, Morton CM, Soltis DE, Bayer C, Fay MF, de Bruijn AY, Sullivan S, Qiu YL. 2000. Phylogenetics of flowering plants based on combined analysis of plastid atpB and rbcL gene sequences. Systematic Biology 49, 306–362.

Schulze S, Mallmann J, Burscheidt J, Koczor M, Streubel M, Bauwe H, Gowik U, Westhoff P. 2013. Evolution of C₄ photosynthesis in the genus Flaveria: establishment of a photorespiratory CO₂ pump. The Plant Cell 25, 2522–2535.

Schulze S, Westhoff P, Gowik U. 2016. Glycine decarboxylase in C₃, C₄ and C₃–C₄ intermediate species. Current Opinion in Plant Biology 31, 29–35.

Sharwood RE, Ghannoum O, Whitney SM. 2016. Prospects for improving CO₂ fixation in C₃-crops through understanding C₄-Rubisco biogenesis and catalytic diversity. Current Opinion in Plant Biology 31, 135–142.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Stata M, Sage TL, Rennie TD, Khoshravesh R, Sultmanis S, Khaikin Y, Ludwig M, Sage RF. 2014. Mesophyll cells of C₄ plants have fewer chloroplasts than those of closely related C₃ plants. Plant, Cell and Environment 37, 2587–2600.

Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Research 34, W609–W612.

Tillich M, Beick S, Schmitz-Linneweber C. 2010. Chloroplast RNA-binding proteins: repair and regulation of chloroplast transcripts. RNA Biology 7, 172–178.

von Caemmerer S, Furbank RT. 2016. Strategies for improving C₄ photosynthesis. Current Opinion in Plant Biology 31, 125–134.

Wang JL, Klessig DF, Berry JO. 1992. Regulation of C₄ gene expression in developing Amaranth leaves. The Plant Cell 4, 173–184.

Westhoff P, Gowik U. 2010. Evolution of C₄ photosynthesis—looking for the master switch. Plant Physiology 154, 598–601.

Wheeler DL, Church DM, Federhen S, et al. 2003. Database resources of the National Center for Biotechnology. Nucleic Acids Research 31, 28–33.

Williams BP, Burgess SJ, Reyna-Llorens I, Knerova J, Aubry S, Stanley S, Hibberd JM. 2016. An untranslated cis-element regulates the accumulation of multiple C₄ enzymes in Gynandropsis gynandra mesophyll cells. The Plant Cell 28, 454–465.

Xie Y, Wu G, Tang J, et al. 2014. SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. Bioinformatics 30, 1660–1666.

Yang XF, Wang YT, Chen ST, Li JK, Shen HT, Guo FQ. 2016. PBR1 selectively controls biogenesis of photosynthetic complexes by modulating translation of the large chloroplast gene Ycf1 in Arabidopsis. Cell Discovery 2, 16003.

Yerramsetty P, Stata M, Siford R, Sage TL, Sage RF, Wong GK, Albert VA, Berry JO. 2016. Evolution of RLSB, a nuclear-encoded S1 domain RNA binding protein associated with post-transcriptional regulation of plastid-encoded rbcL mRNA in vascular plants. BMC Evolutionary Biology 16, 141.