Journal of Experimental Botany, Vol. 67, No. 10 pp. 3065–3078, 2016
doi:10.1093/jxb/erw150 Advance Access publication 12 April 2016
This paper is available online free of all access charges (see http://jxb.oxfordjournals.org/open_access.html for further details)

RESEARCH PAPER

C_3–C_4 intermediacy in grasses: organelle enrichment and distribution, glycine decarboxylase expression, and the rise of C_2 photosynthesis

Roxana Khoshravesh[1], Corey R. Stinson[1], Matt Stata[1], Florian A. Busch[2], Rowan F. Sage[1], Martha Ludwig[3] and Tammy L. Sage[1,*]

[1] Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks St., Ontario, ON M5S 3B2, Canada
[2] Research School of Biology, Australian National University, Canberra, ACT 2601, Australia
[3] School of Chemistry and Biochemistry, University of Western Australia, Crawley, WA 6009, Australia

* Correspondence: tammy.sage@utoronto.ca

Received 21 February 2016; Accepted 21 March 2016

Editor: Christine Raines, University of Essex

Abstract

Photorespiratory glycine shuttling and decarboxylation in bundle sheath (BS) cells exhibited by C_2 species is proposed to be the evolutionary bridge to C_4 photosynthesis in eudicots. To evaluate this in grasses, we compare anatomy, cellular localization of glycine decarboxylase (GDC), and photosynthetic physiology of a suspected C_2 grass, *Homolepis aturensis*, with these traits in known C_2 grasses, *Neurachne minor* and *Steinchisma hians*, and C_3 *S. laxum* that is sister to *S. hians*. We also use publicly available genome and RNA-sequencing data to examine the evolution of GDC subunits and enhance our understanding of the evolution of BS-specific GDC expression in C_2 and C_4 grasses. Our results confirm the identity of *H. aturensis* as a C_2 species; GDC is confined predominantly to the organelle-enriched BS cells in *H. aturensis* and *S. hians* and to mestome sheath cells of *N. minor*. Phylogenetic analyses and data obtained from immunodetection of the P-subunit of GDC are consistent with the hypothesis that the BS dominant levels of GDC in C_2 and C_4 species are due to changes in expression of a single GLDP gene in M and BS cells. All BS mitochondria and peroxisomes and most chloroplasts in *H. aturensis* and *S. hians* are situated centripetally in a pattern identical to C_2 eudicots. In *S. laxum*, which has C_3-like gas exchange patterns, mitochondria and peroxisomes are positioned centripetally as they are in *S. hians*. This subcellular phenotype, also present in eudicots, is posited to initiate a facilitation cascade leading to C_2 and C_4 photosynthesis.

Key words: Arthropogoninae, bundle sheath, C_2 Kranz anatomy, C_2 photosynthesis, glycine decarboxylase, grasses, mitochondria, *Homolepis*.

Introduction

C_4 photosynthesis has independently evolved >60 times in angiosperms (R.F. Sage *et al.*, 2011; Grass Phylogeny Working Group II, 2012). Within the angiosperms, the Poaceae represents the most prolific family of C_4 origins, with approximately twice as many origins as any other family (Sage, 2016). C_4 grasses also make up the greatest number of C_4 species, comprising ~60% of the 8000 estimated number of C_4 species (Still *et al.*, 2003; Sage, 2016). Roughly a quarter of global net primary productivity on land is due to C_4 photosynthesis (Still *et al.*, 2003), of which the vast majority is contributed by grasses (Sage *et al.*, 1999). C_4 grasses also have great significance for humanity as they dominate the fraction of biomass entering the human food chain as grain (maize, sorghum, and millets), sugar (sugarcane), and fodder for animals, and

© The Author 2016. Published by Oxford University Press on behalf of the Society for Experimental Biology.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

efforts are underway to engineer the C_4 pathway into C_3 grass crops such as rice and wheat to exploit the superior productivity of C_4 photosynthesis (Peterhansel, 2011; von Caemmerer *et al.*, 2012). Given these considerations, there is now a great interest in understanding how C_4 photosynthesis evolved in grasses, to understand both how this complex trait repeatedly arose, and how we might learn from the evolutionary examples to direct C_4 engineering in major crops (Hibberd *et al.*, 2008).

Studies of C_4 evolution are informed by the presence of species exhibiting intermediate stages between fully expressed C_3 and C_4 life forms within a single evolutionary clade (Sage *et al.*, 2014). Ideally, there should be C_3–C_4 relatives from multiple independent lineages of C_4 photosynthesis to facilitate evaluation of evolutionary hypotheses using comparative approaches. Multiple independent clades provide the possibility to assess whether evolutionary trends are replicated, as they should be if C_4 photosynthesis evolved along common trajectories (Heckmann *et al.*, 2013). To date, the majority (85%) of C_3–C_4 species occur in eudicots, with the genus *Flaveria* standing as the major group used in studies of C_4 evolution (R.F. Sage *et al.*, 2011, 2014). *Flaveria* has over twice the number of intermediates (10) as any other evolutionary clade, and these form two distinct clades that each evolved C_3- and C_4-like species (McKown *et al.*, 2005; Lyu *et al.*, 2015). In addition to *Flaveria*, at least 11 other C_4 evolutionary lineages have been identified with C_3–C_4 intermediates branching in sister positions to the C_4 line (Sage *et al.*, 2014). Most of these have only one or two species, although in recent years there is evidence of three clades (*Heliotropium*, *Anticharis*, and *Blepharis*) potentially having more than five intermediates (Muhaidat *et al.*, 2011; Khoshravesh *et al.*, 2012; Fisher *et al.*, 2015). All but three of the C_4 lineages with C_3–C_4 intermediates are eudicots. Among monocots, the genus *Neurachne* contains a C_3–C_4 intermediate that branches in a sister position to a C_4 (Christin *et al.*, 2012). A recent study has reported C_3, C_3–C_4, and C_4 photosynthetic genotypes in *Alloteropsis semialata* (Lundgren *et al.*, 2015). The genus *Steinchisma* also contains C_3–C_4 intermediates (Brown *et al.*, 1983; Hylton *et al.*, 1988), but they lack close C_4 relatives in their subtribe, Otachyrinae (Grass Phylogeny Working Group II, 2012). As a consequence of the discrepancy between C_3–C_4 numbers in eudicots and monocots, our understanding of C_4 evolution is dominated by information from eudicot clades. If there is important variation in eudicot versus monocot patterns of C_4 evolution, as suggested by recent theoretical treatments (Williams *et al.*, 2013), it could be missed because of low monocot representation in the C_3–C_4 intermediate population.

The South American subtribe Arthropogoninae is a hotspot for C_4 evolution within the grasses, with four putative distinct C_4 origins, once in *Mesosetum*/*Arthropogon*, a second time in *Oncorachis*, and twice in *Coleataenia* (Grass Phylogeny Working Group II, 2012). As such, the Arthropogoninae is a strong candidate to contain numerous C_3–C_4 species. This possibility is bolstered by an image from *Homolepis aturensis* in Supplementary fig. S1 of Christin *et al.* (2013) that illustrates enlarged bundle sheath (BS) cells with chloroplasts arranged around the periphery in addition to chloroplast clusters adjacent to the vascular tissue. Although this centripetal chloroplast arrangement led to tentative identification of *H. aturensis* as C_4 (Christin *et al.*, 2013), the presence of centrifugal chloroplasts in the BS cells is a common feature in C_3–C_4 intermediate species (Monson *et al.*, 1984; Sage *et al.*, 2014). Significantly, because the genus *Homolepis* is sister to a clade that contains only C_4 species, it has been identified as a genus that might exhibit precursor traits that could have enabled the evolution of the C_4 phenotype (Grass Phylogeny Working Group II, 2012). To evaluate this possibility, we have collected *H. aturensis* in Costa Rica for study.

A prominent physiological feature of C_3–C_4 intermediates is the transport of photorespiratory glycine from mesophyll (M) to BS cells for decarboxylation by glycine decarboxylase (GDC), with the released CO_2 then being refixed by BS Rubisco (Monson and Rawsthorne, 2000). Photorespiratory glycine shuttling exhibited by C_3–C_4 intermediates has also been termed C_2 photosynthesis in reference to the number of carbons shuttled from M to BS cells (Vogan *et al.*, 2007; Bauwe, 2011). The BS mitochondria of most C_2 species examined to date contain the majority of the GDC within the leaf, with small amounts in M cells potentially for C1 metabolism (Hylton *et al.*, 1988; Morgan *et al.*, 1993; Rawsthorne *et al.*, 1998; Voznesenskaya *et al.*, 2001; Ueno *et al.*, 2003; Marshall *et al.*, 2007; Voznesenskaya *et al.*, 2010; Muhaidat *et al.*, 2011; T.L. Sage *et al.*, 2011). Decarboxylation of glycine in the BS cells establishes a glycine gradient between M and BS cells, and rapid movement to BS cells is facilitated by enhanced vein density in C_2 relative to C_3 species (Monson and Rawsthorne, 2000). Subsequent glycine decarboxylation within BS cells increases CO_2 around BS Rubisco ~3-fold (Keerberg *et al.*, 2014), and the resulting increase in Rubisco efficiency reduces the CO_2 compensation point in C_2 species relative to C_3 by 10–40 μmol mol^{-1} (Holaday *et al.*, 1984; Monson *et al.*, 1984; Vogan *et al.*, 2007; T.L. Sage *et al.*, 2011, 2013).

The functional GDC holoenzyme consists of four subunits encoded by individual genes, GLDH, GLDL, GLDP, and GLDT (Bauwe 2011). Decarboxylase activity of the complex is located in the P-subunit encoded by the GDLP gene. In *Flaveria*, the C_4 photosynthetic mechanism was established through gradual pseudogenization of a ubiquitously expressed GLDP gene, and full activation of a second GLDP gene that shows BS-specific expression in C_3 *Flaveria* species (Schulze *et al.*, 2013). The ancestral duplication of the GLDP gene in *Flaveria* is considered a genetic enabler of C_4 evolution in the genus (Schulze *et al.*, 2013). BS cell-dominant expression of GDC has been reported in the C_2 grass *Steinchisma hians* (=*Panicum milioides*; Hylton *et al.*, 1988). To date, the molecular evolution of this trait in *S. hians* or other C_2 and C_4 grasses has not been assessed. Schulze *et al.* (2013), highlighting the presence of two GLDP genes in rice [a member of the C_3 BEP (Bambusoideae, Ehrhartoideae, Pooideae) clade] and a single copy in maize, sorghum, and *Setaria italica* [of the C_4 PACMAD (Panicoideae, Arundinoideae, Chloridoideae, Micrairoideae, Aristoideae, Danthonioideae) clade], posited that the ubiquitously expressed GLDP gene(s) in C_4 grasses were pseudogenized as in *Flaveria* and subsequently lost from the genomes.

The purpose of this study was to determine whether *H. aturensis* exhibits C_2 photosynthesis or is a C_4 species as previously suggested (Christin *et al.*, 2013). We compared anatomy, localization of GLDP, and photosynthetic physiology of *H. aturensis* with patterns previously identified in the C_2 grasses *Steinchisma hians* and *Neurachne minor* (Morgan and Brown, 1979; Hylton *et al.*, 1988). Current phylogenies place the subtribe Otachyrinae, containing *Steinchisma*, as sister to the Arthropogoninae (Grass Phylogeny Working Group II, 2012). In addition, we examined *S. laxum* which has also been identified as a species that might provide information on the early stages of C_2 and C_4 evolution (Grass Phylogeny Working Group II, 2012; Sage *et al.*, 2013). Finally, we use genome sequence data from publicly available databases (the Phytozome and NCBI), as well as assembled RNA-sequencing (RNA-seq) from 16 additional grass species, to examine evolution of GLDP and genes encoding the other GDC subunits and provide a broader understanding of the evolution of BS-specific GDC expression in C_2 and C_4 grasses.

Materials and methods

Plant material

Plants of *Homolepis aturensis* Chase., *Steinchisma hians* Raf., and *S. laxum* (Sw.) Zuloaga obtained from sources described in Supplementary Table S1 at *JXB* online were grown at the University of Toronto in a greenhouse in 10–20 liter pots of a sandy-loam soil and were watered daily to avoid water stress. Fertilizer was supplied weekly as a 50:50 mixture of Miracle-Grow 24-10-10 All Purpose Plant Food and Miracle Grow Evergreen Food (30-10-20) at the recommended dosage (22 ml of fertilizer salt per 6 liters; Scotts Miracle-Gro; www.scotts.ca). Plants of *Neurachne minor* from localities previously described (Christin *et al.*, 2012) were grown in a naturally illuminated glasshouse with mean temperatures of 25 °C/13 °C (day/night) at the Plant Growth Facility (PGF) of the University of Western Australia, Perth, Western Australia (latitude 33°89'S). To provide C_3 and C_4 grass species for comparison, we also examined leaves of PACMAD species *Dichanthelium oligosanthes* (Schult.) Gould (C_3), *Panicum bisulcatum* Thunb. (C_3), and two C_4 species (*Panicum virgatum* L., NAD-ME subtype and *Setaria viridis* P. Beauv., NADP-ME subtype). Seed of these plants, obtained from sources described in Supplementary Table S1, were also grown at the University of Toronto.

Leaf anatomy, ultrastructure, and immunolocalizations

The internal anatomy of leaves was assessed on sections sampled from the middle of the most recent, fully expanded leaves (one leaf per plant; three plants per species). Plants were sampled from 09:00 h to 11:00 h between April and August when day length was >11.5 h and light intensity in the greenhouse regularly exceeded 1400 µmol photons m^{-2} s^{-1}. The youngest cohort of fully expanded leaves was sampled in full sun for all procedures. Samples were prepared for light and transmission electron microscopy (TEM) to assess anatomy as previously described (T.L. Sage *et al.*, 2011, 2013; Stata *et al.*, 2014). For immunolocalization, tissue from the same region of the leaf was fixed overnight in 1% (v/v) paraformaldehyde and 1% (v/v) glutaraldehyde in 0.05 M sodium cacodylate buffer. Tissue was then dehydrated and embedded in LR White (Voznesenskaya *et al.*, 2013). Immunolocalization of GLDP was conducted as outlined by Khosravesh *et al.* (2012). Primary and secondary antibody (18 nm anti-rabbit IgG gold conjugate; Jackson Immunoresearch) dilutions were 1:50 and 1:20, respectively. Immunodetection of the Rubisco large subunit was modified from Ueno (1992). Sections were blocked in 0.5% BSA prior to incubation in primary antibody (1:100) for 3 h. Incubation in secondary antibody (1:40; 18 nm anti-rabbit IgG gold conjugate; Jackson Immunoresearch) was for 1 h. To quantify all BS and M cellular features, TEM images from BS and M cells of the same grids used for immunogold labeling were analyzed using Image J software (Schneider *et al.*, 2012) as previously described (Sage *et al.*, 2013; Stata *et al.*, 2014). The anti-GLDP antiserum was commercially produced (GL Biochem) against a 17 amino acid peptide showing high conservation in both monocots and dicots. Antisera recognizing the Rubisco large subunit (RBCL) were obtained from AgriSera. Three replicate immunolocalizations were conducted on different days with each replicate including sectioned tissue from all species.

Leaf gas exchange analysis

Gas exchange of intact, attached leaves was determined using a LiCor 6400 gas exchange system as previously described (Sage *et al.*, 2013). Measurement conditions were 31 ± 1 °C and a vapor pressure difference between leaf and air of 2 ± 0.2 kPa. For measurement of the response of net CO_2 assimilation rate (A) to intercellular CO_2 concentration (C_i) at light saturation, leaves were first equilibrated to 1200–1500 µmol photons m^{-1} s^{-1} at an ambient CO_2 concentration of 400 µmol m^{-2} s^{-1}. Ambient CO_2 levels were then reduced in steps to 30–50 µmol mol^{-1} (lower end of this range for C_2 and C_4 species, upper end of this range for C_3 species), with measurements at each step after rate equilibration. The ambient CO_2 was then returned to 400 µmol photons m^{-1} s^{-1}, and A re-measured. If A was within 10% of the original rate, the CO_2 concentration around the leaf was increased in steps to near 1600 µmol mol^{-1}, with measurements made at each step. The linear initial slope of the A/C_i response was used as an estimate of carboxylation efficiency (CE).

For estimation of the apparent CO_2 compensation point in the absence of day respiration (C_*), the Laisk method was used as modified by Sage *et al.* (2013). Values of C_* were not estimated in C_4 plants. Leaves were first equilibrated at 400 µmol mol^{-1} and a light intensity near saturation (900–1500 µmol photons m^{-2} s^{-1}). The CO_2 was then reduced to provide a C_i near 100 µmol mol^{-1}, and, after stabilization, the rate was reduced in a series of steps to near the CO_2 compensation point (Γ), with measurements at each step after signal stabilization. The C_i was then returned to near 100 µmol mol^{-1}, and the procedure was repeated at a lower light intensity. This cycle was repeated such that when measurements were completed, there were 4–5 A versus C_i response curves with over five measurements at C_i values <100 µmol mol^{-1}. Each A versus C_i response at a given light intensity was then fitted with a linear regression using the lowest 4–6 measurement points that fell on the regression. Points that fell below the regression line above ~80 µmol mol^{-1} were not included in the regression, as A/C_i responses below light saturation may be prone to non-linearity above 60–100 µmol CO_2 mol^{-1} air. The estimate of C_* was taken as the C_i value where the 4–5 curves at different light intensities converged. Rarely, however, did all curves converge at the exact same C_i; instead, they intersected with each other over a 5–10 ppm range. In such cases, the middle of the intersects was taken as the C_* estimate. However, if there was evidence of a shift to lower photosynthetic photon flux density (PPFD) of the high light response, which often occurs in C_2 species, the high light response was not included in the analysis (Sage *et al.*, 2013).

Phylogenetic analysis of GDC subunit genes

Genome sequences for *Zea mays*, *Sorghum bicolor*, *Panicum hallii*, *Se. viridis*, *S. italica*, *Oryza sativa*, *Brachypodium distachyon*, and *B. stacei* were downloaded from the Phytozome v10.3 (https://phytozome.jgi.doe.gov/pz/portal.html). Polyploid species *Triticum aestivum* and *P. virgatum* were omitted. *Amborella trichopoda* was used as the outgroup, and we included the *Brassicaceae* species *Arabidopsis thaliana*, *Capsella grandiflora*, and *Boechera stricta* to show which duplications in the model plant *A. thaliana* are conserved across

the land plants and which are lineage specific. RNA-seq data for 14 BEP clade species as well as the C_3 PACMAD *Dicanthelium clandestinum* and its close C_4 relative *Megathyrsus maximus* were downloaded from the NCBI short read archive (Supplementary Table S2). Reads mapping to each GDC subunit gene were identified using BLASTN with orthologs in *Z. mays* and *O. sativa* as queries (*S. bicolor* was used in lieu of *Z. mays* for one of the two GLDL paralogs as *Z. mays* appears to have lost this copy). For each gene, alignments were generated based on the Phytozome genomes using Muscle (Edgar, 2004), a highly conserved region was selected, and a consensus sequence was generated. These consensus sequences were used as reference for assembling sequences based on the retrieved reads using Geneious 8 (http://www.geneious.com).

The assemblies and sequences from Phytozome genomes were aligned using Muscle, trimmed using Trimal (Capella-Gutierrez *et al.*, 2009) with the 'strict' heuristic option, and used to generate Bayesian phylogenies using MrBayes (Ronquist and Huelsenbeck, 2003) as follows: four runs, four chains, GTR substitution model, 2 million generations for all trees except GLDT, which was run for 10 million generations to allow better convergence; ≥10 000 trees were sampled from portions of the end of each run where the average SD of split frequencies remained below 1%. Tree figures were generated using Fig Tree (http://tree.bio.ed.ac.uk/software/figtree).

Data analysis

Results were analyzed with Sigmaplot version 12.5 (Systat Software., San Jose, CA, USA) using one-way ANOVA followed by a Tukey's means comparison test. For characterization of leaf anatomy and ultrastructure, leaf samples were collected from three plants. For all traits measured, the values per plant were averaged to give one value for a plant. These individual plant values were the unit of replication for statistical analysis. For characterization of anatomical features, data from 3–5 sections per plant were averaged. For quantitative assessment of organelles in M and BS cells and GLDP immunodetection, the data from 10 imaged cells per cell type per plant were averaged. For leaf gas exchange, 4–15 measurements were conducted on 4–7 plants per species. In *Neurachne* species, vascular tissue is surrounded by two layers of cells, an outermost BS and innermost mestome sheath that functions in the C_4 species as the site of CO_2 refixation (Hattersley *et al.*, 1986). Comparing data collected from the mestome sheath of *N. minor* with data from the BS of the other grasses makes statistical comparisons invalid except for comparisons between either the presence or absence of GDLP in M versus the BS or mestome sheath cells. Hence, *N. minor* was not included in the statistical tests involving these other grasses.

Results

Homolepis aturensis *possesses structural features common to* C_2 *species*

Leaves of *H. aturensis* are anatomically similar to those of the C_2 species *S. hians* and *S. laxum* with respect to M cell structure (Fig. 1). One layer of M cells extends from BS cells to the adaxial and abaxial epidermis (Fig. 1). Approximately six M cells separate the BS of adjacent veins in *S. laxum* whereas four cells separate the BS of adjacent veins in *H. aturensis* and *S. hians* (Fig. 1; Supplementary Table S3). The M:BS tissue ratio for *S. laxum*, *H. aturensis*, and *S. hians* is ~2 (Fig. 1; Supplementary Table S3), which is similar to that of the C_4 species *P. virgatum* but >2.5 times less than that of the C_3 species *D. oligosanthes* and *P. bisulcatum*. A large M and small BS volume contributes to an M:BS of almost 5 in *Se. viridis* even though there are only two cells between each vein (Supplementary Fig. S1C; Supplementary Table S3). M cells in *D. oligosanthes*, *P. bisulcatum*, and *Se. viridis* are more loosely packed and elongate than those of *P. virgatum* (Supplementary Fig. S1). The cellular features of *N. minor* are similar to those previously reported (Hattersley *et al.*, 1986), with 2–3 M cells between veins (Supplementary Table S3).

The BS cells of *H. aturensis* contain chloroplasts arranged around the periphery, with a significantly greater number clustered centripetally (Figs 1A, B, 2A; Supplementary Table S4). Mitochondria and peroxisomes localize almost exclusively to the centripetal BS pole (Fig. 2A; Supplementary Table S4). These spatial arrays of chloroplasts, mitochondria, and peroxisomes are similar in BS cells of *S. hians* (Figs 1E, F, 2C; Supplementary Table S4). As observed for *H. aturensis* and *S. hians*, a significantly greater number of mitochondria and peroxisomes are positioned at the centripetal BS pole in *S. laxum*, although chloroplasts are arranged equally around BS cells (Fig 1C, D; Supplementary Table S4). BS mitochondria are commonly surrounded by chloroplasts in *S. hians* (Fig 3E), *H. aturensis*, and *S. laxum* in a pattern similar to previous reports for *Steinchisma* species (Brown *et al.*, 1983). In contrast to centripetal chloroplasts displayed with their long axis parallel to the BS wall in *S. laxum*, the long axis of these chloroplasts in *H. aturensis* and *S. hians* are perpendicular to the BS wall (Figs 1, 2). BS chloroplasts are situated primarily in a centrifugal position in C_3 grasses *D. oligosanthes* and *P. bisulcatum* (Supplementary Table S4; Supplementary Figs S1, S2), as noted in C_3 eudicots (Muhaidat *et al.*, 2011; Sage *et al.*, 2013). Approximately 65–69% of mitochondria and 19% (*D. oligosanthes*) and 41% (*P. bisulctum*) of peroxisomes, respectively, are located in the centripetal position of the C_3 grasses (Supplementary Table S4).

Quantitative parameters of organelle traits of BS and M cells are summarized in Supplementary Table S5. BS organelle parameters of *H. aturensis* and *S. hians* are not statistically different from each other except for a significantly greater number of mitochondria per planar cell area in *S. hians*. Mitochondria planar area per planar BS cell area of *H. aturensis* and *S. hians* is significantly greater than for *S. laxum*. In addition, mitochondria number per planar BS cell area, chloroplast and peroxisome number, and planar area per planar BS cell area are significantly greater in *S. hians* than in *S. laxum*. Results from a one-way ANOVA on ranks comparing peroxisome planar area per planar BS cell area between *S. hians*, *S. laxum*, and *H. aturensis* only indicated that this trait was significantly higher in *S. hians* and *H. aturensis* than in *S. laxum* ($P \leq 0.001$). The C_3 grasses have the lowest mitochondria planar area per planar BS cell area; however, BS mitochondria parameters of C_4 grasses are not different from those of the C_2 grasses *H. aturensis* and *S. hians*. Mitochondria, peroxisomes, and chloroplasts are abundant in mestome sheath cells of *N. minor* (Hattersley *et al.*, 1986; Sage *et al.*, 2014). Unlike *H. aturensis*, *S. hians*, and *S. laxum*, organelles are not polarized in their distribution (Supplementary Fig. S3), but are instead positioned

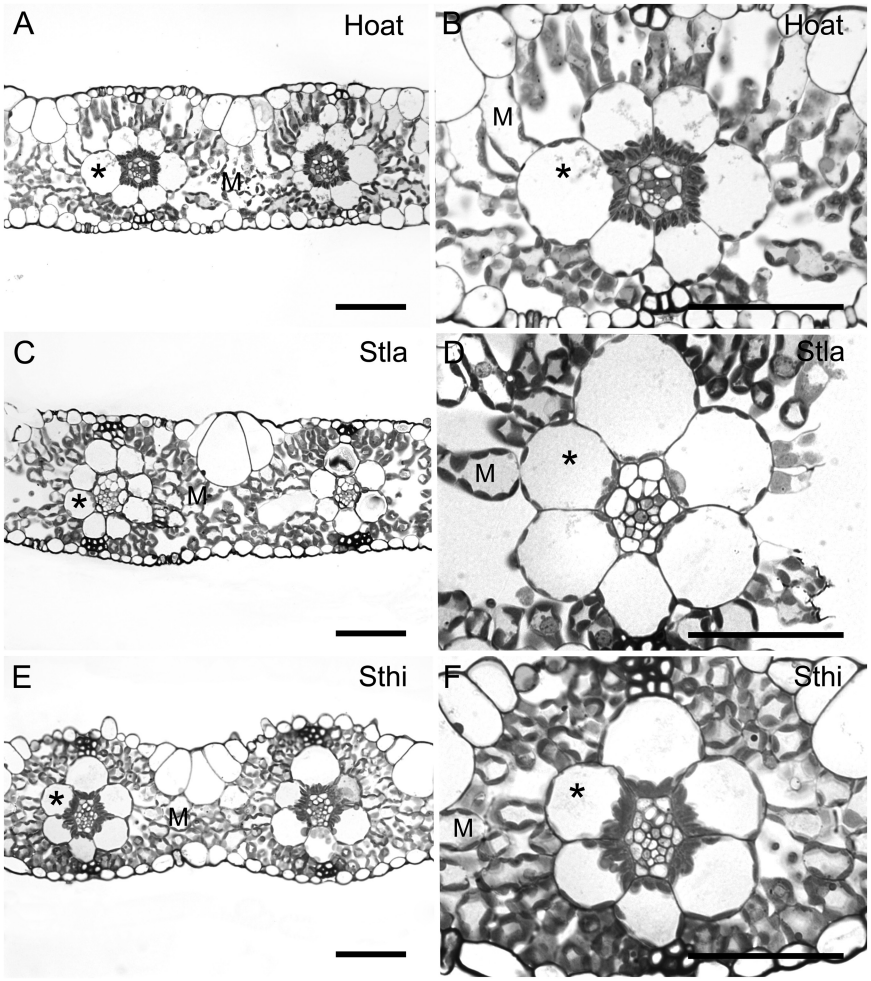

Fig. 1. Leaf cross-sections from (A, B) *Homolepis aturensis* (Hoat), (C, D) *Steinchisma laxum* (Stla), and (E, F) *S. hians* (Sthi). M, mesophyll; asterisk, bundle sheath. Scale bars=50 µm.

around the mestome sheath cell periphery (Hattersley *et al.*, 1986; Sage *et al.*, 2014).

When considering M cell organelle features, significantly fewer mitochondria and peroxisomes are observed in C_4 species relative to the C_3 species and the C_2 species *H. aturensis*, *S. hians*, and *S. laxum* (Supplementary Table S5). The C_4 species as well as *H. aturensis* have significantly lower chloroplast planar area per M planar cell area relative to *S. hians*, *S. laxum*, and the C_3 grass species, and these changes result from either smaller (*H. aturensis*, *P. virgatum*) or fewer (*Se. viridis*) M cell chloroplasts (Supplementary Table S5). M cells of C_4-like and C_4 *Flaveria* and other C_4 species have recently been reported to have significant reductions in chloroplast volume (Stata *et al.*, 2014, 2016).

Homolepis aturensis *exhibits C_2 levels of GLDP in bundle sheath and mesophyll cells*

Results from quantification of gold particles conjugated to secondary antibodies that bind to anti-GLDP are summarized in Supplementary Table S5. GLDP is almost exclusively located in BS cells of *H. aturensis* and *S. hians*, and mestome sheath cells of *N. minor* (Fig. 3A, B, E, F; Supplementary Fig. S3). Both BS and M cells contained high levels of GLDP labeling in *S. laxum* (Fig. 3C, D). In comparison with all other C_3 species examined, *S. laxum* had the highest GLDP labeling in BS mitochondria (Supplementary Table S5). These patterns in GLDP distribution in M and BS mitochondria of *S. laxum* and *S. hians* are similar to earlier reports on these species (Hylton *et al.*, 1988). The gold density in C_3 species *D. oligosanthes* and *P. bisulcatum* is higher in mitochondria of M than BS cells and is significantly greater on a planar M cell area basis than in those of *H. aturensis*, *S. hians*, and the C_4 species (Supplementary Table S5; Supplementary Fig. S4).

Homolepis aturensis *exhibits C_2 levels of Rubisco in bundle sheath and mesophyll cells*

Rubisco occurs in both M and BS cell chloroplasts of *H. aturensis*, and this pattern was similar to that observed in *S. hians*, *S. laxum*, and M and mestome sheath cells of *N. minor* (Supplementary Fig. S5). As expected, Rubisco labeling is present only in the BS chloroplasts of the C_4 species (Supplementary Fig. S6). Although Rubisco is also present in the M and BS cells of the C_3 species *D. oligosanthes* and *P. bisulcatum*, there is qualitatively less labeling in the BS cells (Supplementary Fig. S6).

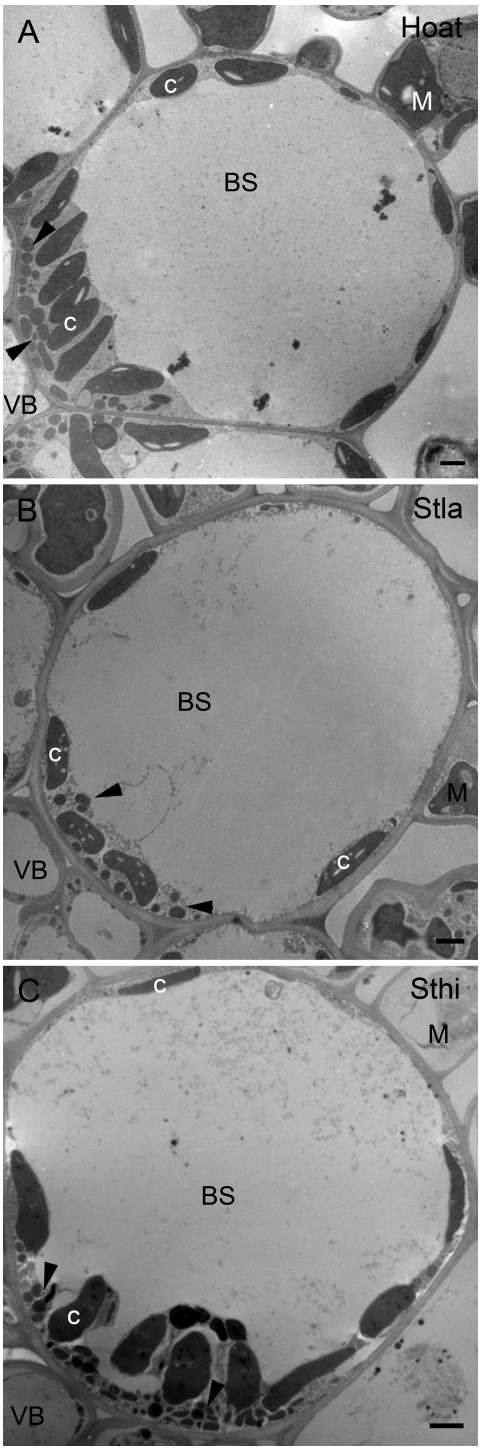

Fig. 2. Bundle sheath ultrastructure of (A) *Homolepis aturensis* (Hato), (B) *Steinchisma laxum* (Stla), and (C) *S. hians* (Sthi). BS, bundle sheath; C, chloroplast; M, mesophyll; VB, vascular tissue; arrowheads, mitochondria. Scale bars=2 µm.

Homolepis aturensis *exhibits photosynthetic characteristics of a C_2 species*

Values of A at 400 µmol CO_2 mol^{-1} air are statistically similar, being 21 ± 3 (mean±range) µmol m^{-2} s^{-1} for all species in the study except for the C_4 plant *Se. viridis*, which has higher A (Table 1). Thus, any differences in carboxylation efficiency (CE), Γ, or intrinsic water use efficiency (estimated as A/g_s at 400 µmol CO_2 mol^{-1} air) should reflect photosynthetic pathway effects and not variation in photosynthetic capacity. Comparison of the A versus C_i responses at saturating light intensities show three equivalent sets of curves that corresponded to the photosynthetic pathway (Fig. 4). C_3 and C_2 species have similar responses, with the exception that Γ was reduced ≥30 µmol mol^{-1} in the C_2 species; consequently, their A/C_i responses were shifted to lower C_i values. *Homolepis aturensis* has A/C_i responses identical to those of *S. hians* and *N. minor*, leading us to classify *H. aturensis* as a C_2 species. Carboxylation efficiencies and A/g_s of all the C_2 and C_3 species are statistically identical and less than those of the C_4 species (Table 1).

In the C_3 species *D. oligosanthes* and *P. bisulcatum*, we observed C_* values near 50 µmol mol^{-1} (Table 1) which is typical for C_3 species at 31 °C (Busch *et al.*, 2013). *Steinchisma laxum* had a similar C_* value (53 µmol mol^{-1}; Table 1; Fig. 5 A) indicating it is functionally C_3 despite the potential function of the BS organelles. The C_* values of the C_2 species were 10.8−20 µmol mol^{-1} (Table 1; Fig. 5B–D); the 10.8 value is on the lower end of C_* for species with this physiology (Edwards and Ku, 1987). In the grasses studied here, lower C_* values correspond to higher values of BS mitochondria per planar cell area (Fig. 6A) and higher BS GLDP density per planar cell area, except for *Se. viridis* (Fig. 6C). C_* was lower in species where BS chloroplast area per planar cell area was greater and M chloroplast area per planar cell area was lower (Fig. 6E, F).

There is little observed shift to lower C_i in the high light response of A versus C_i in *H. aturensis*, such that all A versus C_i curves converge near a common intersection point (Fig. 5). In *S. hians* and *N. minor* there is a slight reduction by a C_i of ~5-7 µmol mol^{-1} in the high light response. There is no change in Γ in the C_3 species with variation in light intensity (Fig. 7), while in each of the C_2 species, Γ increases at the lower light intensity (Fig. 7). *Neurachne minor* exhibits the greatest increase in Γ as light declines, while the light response of Γ is negligible in *H. aturensis* above 300 µmol m^{-2} s^{-1}. Notably, Γ of *N. minor* is twice that of *H. aturensis* across the range of light intensities.

GDC BS specificity in C_2 and C_4 grasses probably results from changes in expression of a single GLDP gene

Phylogenetic analyses reveal that within the Poaceae, *O. sativa* is the only species examined with two gene copies encoding GLDP, and these sequences are most closely related to each other, even with inclusion of 14 additional BEP clade species (Fig. 8). We find no evidence of broader GLDP gene duplication or loss of gene copies in C_4 grass species. Two paralogs encoding GLDH are present in all angiosperms, owing to an ancient duplication event (Supplementary Figs S7, S8). The two paralogs were treated independently as GLDH1 and GLDH2. Preliminary analysis indicates that there is a Poaceae-wide duplication only of GLDL, and targeted assemblies were made independently for each paralog, which we label GLDL1 and GLDL2 (Supplementary Fig. S9). During each assembly, mapped reads were scrutinized

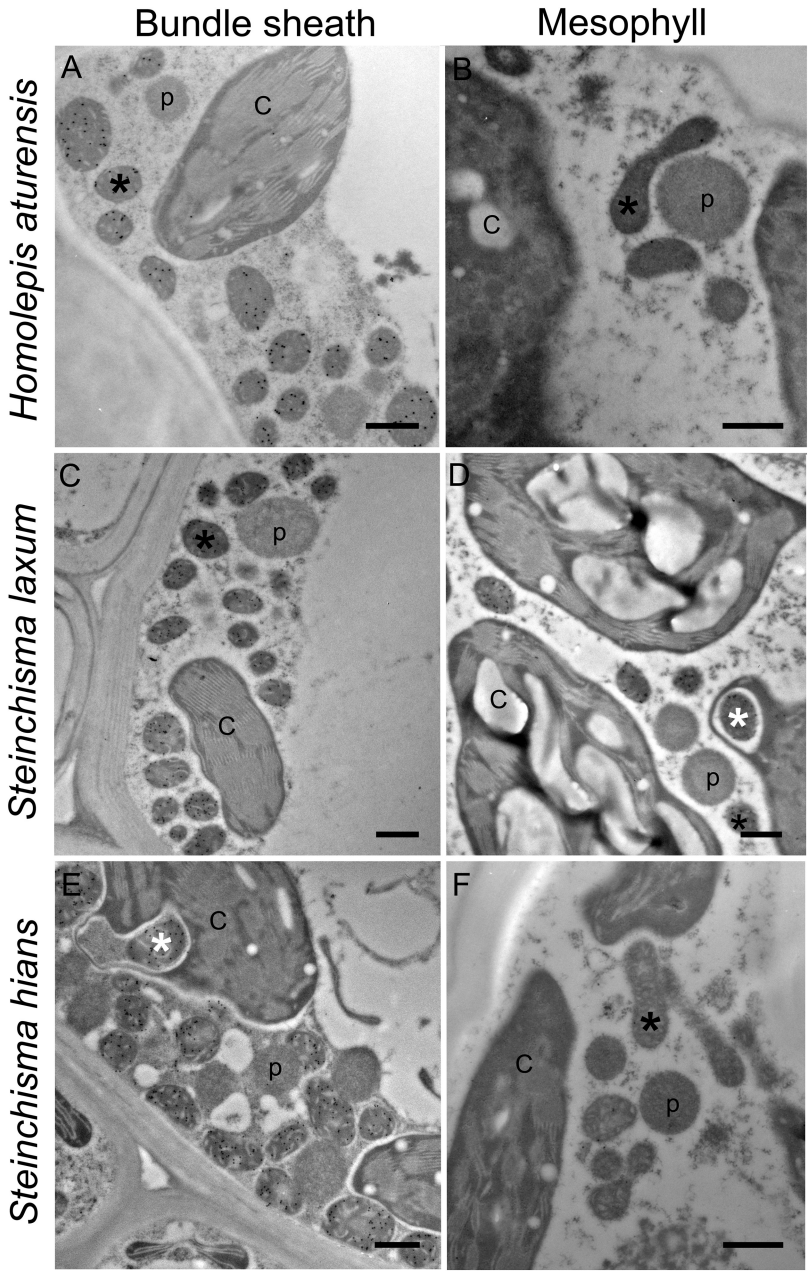

Fig. 3. Immunolocalization of GLDP in bundle sheath and mesophyll cells of (A, B) *Homolepis aturensis*, (C, D) *Steinchisma laxum*, and (E, F) *S. hians*. C, chloroplasts; p, peroxisomes; black asterisk, mitochondria; white asterisk, mitochondria surrounded by chloroplast. Scale bars=500 nm.

manually for evidence of additional paralogs, and none was detected. No reads were found for GLDL2 in the BEP grass *Dendrocalamus sinicus*, suggesting that it may have been lost in this species. Most other grass species possess both GLDL gene copies, with the exception of *Z. mays*, which also lacks GLDL2; *S. bicolor*, which shares a common C_4 origin with *Z. mays*, has both paralogs. For all grass species, full GLDL1 and GLDL2 sequences from genomes and all assemblies which were completed to the start codon were strongly predicted to be mitochondrial-localized by TargetP (Emanuelsson *et al.*, 2000; data not shown). This is consistent with targeting of the two copies in *Arabidopsis* (AT3G17240, AT1G48030; Fig. S9), which are both mitochondrial (Lutziger and Oliver, 2001; Rajinikanth *et al.*, 2007). A plastidial dihydrolipoyl dehydrogenase exists in land plants as well, but is more distantly related and likely dates to a much earlier duplication (Lutziger and Oliver, 2000; Rajinikanth *et al.*, 2007). Finally, GLDT is present as a single-copy gene in all species examined (Supplementary Fig. S10). While we find evidence of local duplication (GLDP, GLDH2) and conserved paralogs (GLDL), we find no evidence of C_4 lineage loss of any GDC subunit genes.

Discussion

Photorespiratory glycine shuttling exhibited by C_2 species is considered to be the evolutionary bridge from C_3 photosynthesis to C_4 photosynthesis, based largely on studies from eudicot species, particularly *Flaveria* (Bauwe, 2011; Sage

Table 1. *Summaries of gas exchange values for species included in this study. Values are means ±SE.*

	Species	*n*	C_* μmol mol^{-1}	CE mol m^{-2} s^{-1}	$A_{\text{at 400}}$ μmol m^{-2} s^{-1}	$A/g_{s\text{ at 400}}$ μmol mol^{-1}
C_3 species						
	Dicanthelium oligosanthes	3, 6	48±1 a	0.13±0.01 b	24.5±1.5 b	56±2
	Panicum bisulcatum	6, 6	50±2 a	0.11±0.01 b	18.7±1.7 b	62±8
C_3-Protokranz						
	Steinchisma laxum	5, 7	53±1 a	0.12±0.0 b	21.7±2.4 b	62±11
C_2 species						
	Homolepis aturensis	4, 7	10.8±1.4 c	0.09±0.01 b	19.8±1.4 b	57±10
	Neurachne minor	3, 3	20.0±2.5 b	0.11±0.02 b	17.7±1.9 b	43±4
	Steinchisma hians	5, 8	11.8±1.4 c	0.10±0.01 b	21.6±1.1 b	58±7
C_4 species						
	Panicum virgatum	0, 2	NA	0.36±0.01 a	24.5±2.5 b	152±5
	Setaria viridis	0, 3	NA	0.54±0.05 a	31.9±0.8 a	76±12

Measurement temperature was 31 ± 1 °C.
Sample sizes given are *n* = 3-5 for C_* estimates, and *n* = 3–8 for all other data, with the exception of *P. virgatum* where *n*= 2).
Letters indicate statistical groupings at *P*<0.5 via one-way ANOVA followed by a Student's–Neumann–Kuehls test.
NA, not applicable.
400 refers to ambient CO_2 concentration in μmol mol^{-1}.

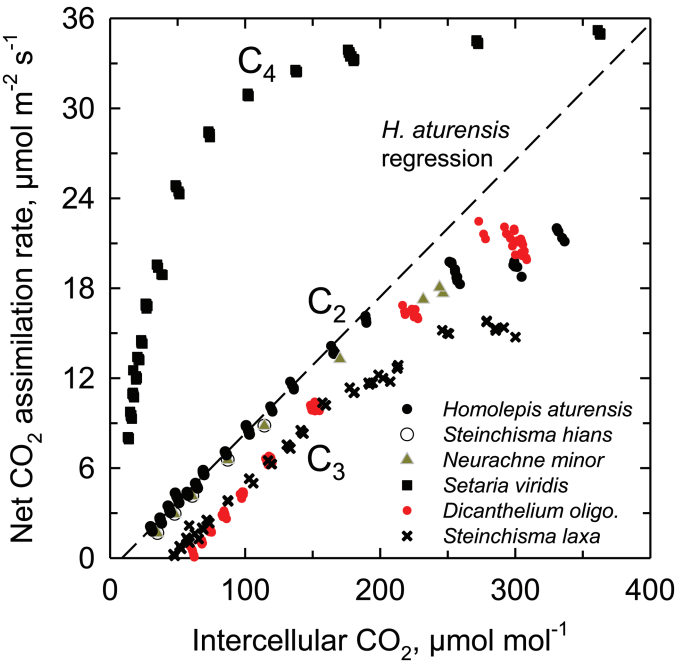

Fig. 4. Response of net CO_2 assimilation rate to intercellular CO_2 concentration in *Homolepis aturensis*, two C_2 species (*Neurachne minor* and *Steinchisma hians*), a C_4 species (*Setaria viridis*), a C_3 species (*Dicanthelium oligosanthes*), and a C_3 species with proto-Kranz anatomy (*Steinchisma laxum*). Measurement conditions were 31 ± 1 °C and saturating light intensities (1200–1500 μmol m^{-2} s^{-1}). The curves shown are representative of 2–3 individual A/C_i responses per species. The linear regression for points below 100 μmol mol^{-1} is shown for *H. aturensis* (dashed line).

et al., 2013; Schulze *et al.*, 2013; Mallmann *et al.*, 2014). It has been posited that C_2 photosynthesis may serve a bridging role to C_4 photosynthesis in grasses as well (Schulze *et al.*, 2013), although the timing of trait acquisition, to include GDC BS cell specificity and abundance, has been proposed to differ between eudicots and monocots (Williams *et al.*, 2013). Here, we evaluated photosynthetic pathway characteristics and cellular features of BS and M cells in *H. aturensis*, a candidate C_2 species which branches in a position sister to a C_4 clade in the subtribe Arthropogoninae. We used comparative studies with confirmed C_2 grasses, *S. hians* and *N. minor*, and the C_3 grass *S. laxum* to facilitate our classification of the photosynthetic type of *H. aturensis*. In addition, we conducted a phylogenetic study of genes for the GDC subunits to test the Schulze *et al.* (2013) inference that the evolution of BS-specific GDC expression in C_4 grasses was similar to that in *Flaveria*. From our physiological and structural results, we conclude that *H. aturensis* is indeed a C_2 species, supporting a hypothesis that the photorespiratory glycine shuttle is a bridge to C_4 photosynthesis in grasses in the subtribe Arthropogoninae. The characterization of *S. laxum* and *S. hians* also allows us to conclude that activation of the BS cells during transition from C_3 to C_2 in grasses in the subtribe sister to the Arthropogoninae is similar to what has been reported for eudicots with respect to BS organelle positioning and organelle and GDC enrichment (Muhaidat *et al.*, 2011; Sage *et al.*, 2013). Finally, our phylogenetic and immunohistochemical data are consistent with the notion that BS GDC in C_2 and C_4 grasses results from changes in expression levels of a single GLDP gene in M and BS cells.

Homolepis aturensis exhibits characteristic features common to species that concentrate photorespired CO_2 in BS cells using the C_2 metabolic cycle. One of the key proteins essential for GDC activity, GLDP, localizes almost exclusively in BS mitochondria in *H. aturensis* and *S. hians*. This pattern is ubiquitous in C_2 eudicots (Hylton *et al.*, 1988; Voznesenskaya *et al.*, 2001; Ueno and Sentoku, 2006; Voznesenskaya *et al.*, 2010; Muhaidat *et al.*, 2011; T.L. Sage *et al.*, 2011). A second characteristic of C_2 species is an abundance of Rubisco in M and BS cells (Monson *et al.*, 1984). Rubisco is abundant in M and BS cells in *H. aturensis* and *S. hians*, in contrast to the typical C_3 pattern (high M Rubisco; low BS Rubisco) and C_4 pattern (Rubisco only in BS cells). Lastly, the number of

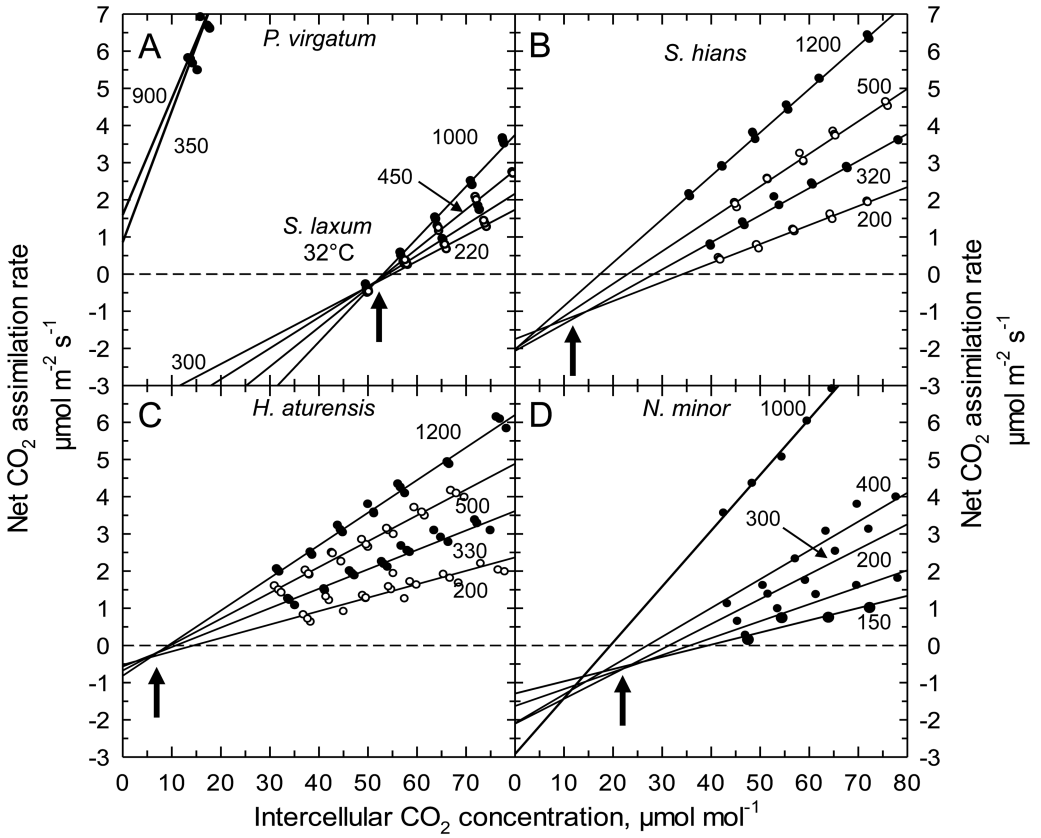

Fig. 5. Representative A/C_i responses below 80 µmol mol^{-1} determined on single leaves at four distinct light intensities of (A) *Panicum virgatum* (C_4) and *Steinchisma laxum* (C_3 proto-Kranz), (B) the C_2 species *Steinchisma hians*, (C) the C_2 species *Homolepis aturensis*, and (D) the C_2 species *Neurachne minor*. Measurement light intensities are indicated beside each curve in µmol photons m^{-1} s^{-1}. The C_* estimate is indicated by arrows. Curves shown are representative of 3–6 measurement sets per species, except for *P. virgatum* where two sets of measurements were obtained. Measurement temperature was 31 ± 1 °C.

M cells between BS cells in *H. aturensis* and *S. hians* results in a reduced M:BS ratio and increased vein density common to C_2 species (reviewed in Sage *et al.*, 2012). These features promote rapid flux of photorespiratory metabolites between M and BS compartments (Monson and Rawsthorne, 2000), improve water relations under high photorespiratory conditions (Osborne and Sack, 2012), and facilitate an increase in the volume of leaf tissue where photorespired CO_2 is concentrated around Rubisco in C_2 species.

As the refixed fraction of photorespiratory CO_2 increases, C_* (and Γ) declines (von Caemmerer, 2000). The effectiveness of the C_2 process in *H. aturensis* as well as *S. hians* is reflected in the Γ and C_* values that are at the lower range of values reported for C_2 species (Holaday *et al.*, 1984; Monson *et al.*, 1984; Ueno *et al.*, 2003; Vogan *et al.*, 2007; Voznesenskaya *et al.*, 2010; T.L. Sage *et al.*, 2011, 2013). Values of Γ and C_* below 15 ppm indicate either that a C_4 cycle is active to complement the C_2 cycle, or that the CO_2 trap in the inner BS is particularly effective at recapturing the photorespired CO_2 (von Caemmerer, 2000). *Steinchisma hians* has weak to negligible C_4 cycle activity (Edwards *et al.*, 1982), and cellular characteristics of organelle orientation in BS cells may explain the low C_* in *H. aturensis* and *S. hians*. The organelle-enriched BS cells of these two species exhibit polarity in organelle positioning such that almost all peroxisomes and GDC-containing mitochondria, and over half of the chloroplasts are situated adjacent to the vascular tissue. This arrangement of BS organelles is posited to be particularly effective in enhancing recapture efficiency of photorespired CO_2 before it can escape the BS cell in C_2 species (Rawsthorne, 1992). One notable feature we observed is that the parallel orientation of the centripetal BS chloroplasts to the inner wall in *S. laxum* shifts to a perpendicular orientation in *S. hians*. A similar perpendicular chloroplast orientation is present in *H. aturensis*. This pattern allows for packaging of the more numerous, larger chloroplasts in the centripetal position, which could be important for increasing the surface area for refixing photorespired CO_2 from adjacent mitochondria. Moreover, BS mitochondria are physically surrounded by chloroplasts in *S. hians* and *H. aturensis*, and these close physical associations have been proposed to enhance refixation of photorespired CO_2 (Brown *et al.*, 1983).

The fine structure of C_2 BS cells is defined as C_2 Kranz, reflecting a view that this photosynthetic carbon-concentrating mechanism is associated with its own enabling Kranz-like structure (Sage *et al.*, 2014). Multiple convergence of C_2 Kranz in eudicots and grasses is strong evidence that this particular BS anatomy is specifically adapted for the C_2 pathway. The earliest recognizable subcellular events that 'increase the accessibility' (Grass Phylogeny Working Group II, 2012) of C_2 Kranz from C_3 in eudicots are an enhancement in numbers and size of mitochondria per BS cell, and positioning of

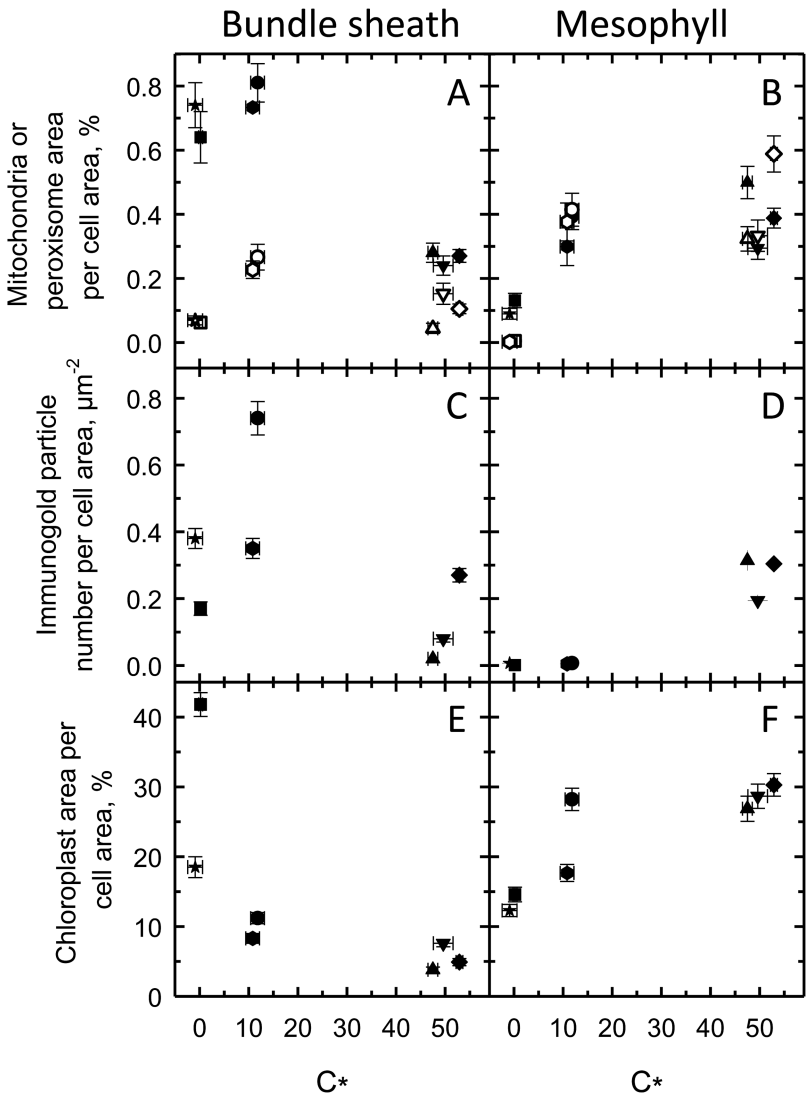

Fig. 6. C_* (Γ, for C_4 species), organelle traits, and GLDP labeling in bundle sheath (A, C, E) and mesophyll (B, D, F) cells of *Dichanthelium oligosanthes* (triangle, C_3), *Panicum bisulcatum* (inverted triangle, C_3), *Steinchisma laxum* (diamond, C_3), *S. hians* (circle, C_2), *Homolepis aturensis* (hexagon, C_2), *Se. viridis* (square, C_4), and *P. virgatum* (star, C_4). Filled and open symbols represent mitochondria and peroxisomes, respectively. Mean ±SE.

these organelles from the centrifugal C_3 position to the centripetal BS pole. The BS chloroplast numbers also increase in tandem with alterations in mitochondria placement, along with a rearrangement of many, but not all, BS chloroplasts from the centrifugal to centripetal pole (Muhaidat *et al.*, 2011; Sage *et al.*, 2013; Voznesenskaya *et al.*, 2013). The anatomy associated with these earliest subcellular events has been termed proto-Kranz (Muhaidat *et al.*, 2011; Sage *et al.*, 2013). The transition to full C_2 BS patterns from proto-Kranz in eudicots results from further amplification in centripetal mitochondria volume (size and numbers) and relocation of a greater fraction of enlarged chloroplasts to the centripetal pole (Muhaidat *et al.*, 2011; Sage *et al.*, 2013). Proto-Kranz and the shift to C_2 Kranz occurs with increasing vein density in *Flaveria* and *Heliotropium* (Muhaidat *et al.*, 2011; Sage *et al.*, 2013).

The subcellular framework of C_3 *S. laxum* BS cells and subsequent changes to that configuration from *S. laxum* to C_2 *S. hians* are similar to those observed in eudicots, supporting a hypothesis that proto-Kranz facilitates the C_3 to C_2 transition (Sage *et al.*, 2014). In comparison with the C_3 species *D. oligosanthes* and *P. bisulcatum*, mitochondria and peroxisomes are situated along the centripetal poles of BS cells in *S. laxum*, classifying this species as proto-Kranz. Previous characterizations of proto-Kranz species have not presented data on peroxisomes, and this additional focus in the present study provides critical information on the positioning of the other organelle involved in C_2 photosynthesis in the BS. A 3-fold increase in mitochondrial volume and corresponding increase in GDC, as well as a 2.5-fold increase in peroxisome volume, and 2-fold increase in chloroplast volume accompany the transition to the C_2 BS pattern in *S. hians*. Also, as observed in eudicots, the increase in BS chloroplast volume in *S. hians* from proto-Kranz is associated with more of these organelles in the centripetal location. Notably, although the number and size of mitochondria per BS cell area are lower in *S. laxum* than in *S. hians*, the GLDP label intensity is similar per mitochondrion and significantly higher than that observed in the

C_3 grasses. These results indicate that increased BS GDC per mitochondrion is also a functionally important development early in C_2 evolution in grasses. The enhanced BS GDC density per mitochondrion in *S. laxum* is present in tandem with C_3-like values of vein density. The C_2 levels of BS GDC in *S. hians* are present in high vein density leaves. The high levels of BS GDC in *S. laxum* and *S hians* contrasts with the theoretical predictions of Williams *et al.* (2013) who modeled C_4 evolution in eudicots and monocots based on observed patterns of trait acquisition in C_3–C_4 intermediates. For monocots, they predicted that an increase in vein density preceded enhanced GDC specificity and abundance in BS cells. However, their model relied on a relatively small data set with significant gaps. The results here indicate that C_2 evolution in grasses follows a pattern more typical of eudicots, which the model of Williams *et al.* (2013) may support when reparameterized with a richer data set.

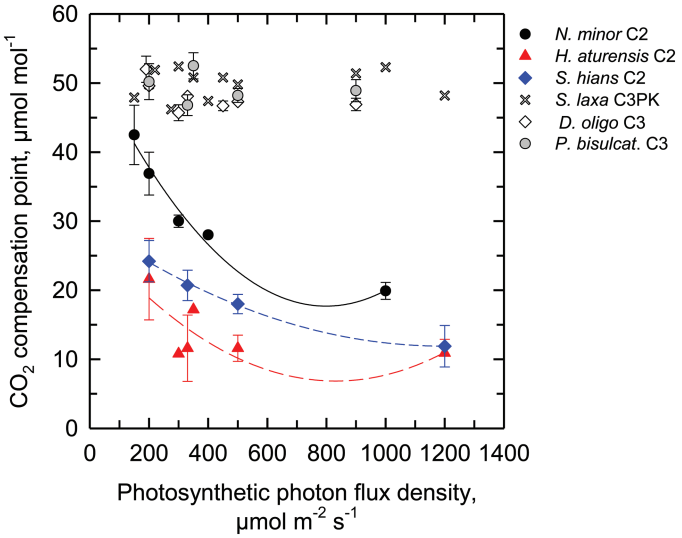

Fig. 7. The response of the CO_2 compensation point of *A* (Γ) as a function of measurement light intensity in the C_2 species *Homolepis aturensis*, *Neurachne minor*, *Steinchisma hians*, the proto-Kranz species *Steinchisma laxum*, and the C_3 species *Dicanthelium oligosanthes* and *Panicum bisulcatum*. Values of Γ were determined from the A/C_i curves used in the sequence of measurements to determine C_c. Symbols are means ±SE (n=2–6), except where no error bar is shown, in which case n=1.

In *Neurachne*, as in many other grasses, the mestome sheath cell is the site of the Calvin–Benson cycle in C_4 species (Hattersley and Browning, 1981; Hattersley *et al.*, 1986; Dengler and Nelson, 1999; Edwards and Voznesenskaya, 2011). We demonstrate that the mestome sheath cells in the C_2 species *N. minor* are functionally similar to the C_2 BS cells of *H. aturensis* and *S. hians* because GDC is almost exclusively located in organelle-enriched mestome sheath cells in the high vein density leaves. However, unlike *H. aturensis* and *S. hians*, there is no polarized orientation of organelles in the GDC-enriched mestome sheath cells of *N. minor* (Hattersley *et al.*, 1986; this study), indicating that *N. minor* utilizes a different strategy from *H. aturensis* and *S. hians* to trap photorespired CO_2. In *Neurachne*, as in many other grasses, the thick mestome sheath cell wall with a suberized lamella becomes the

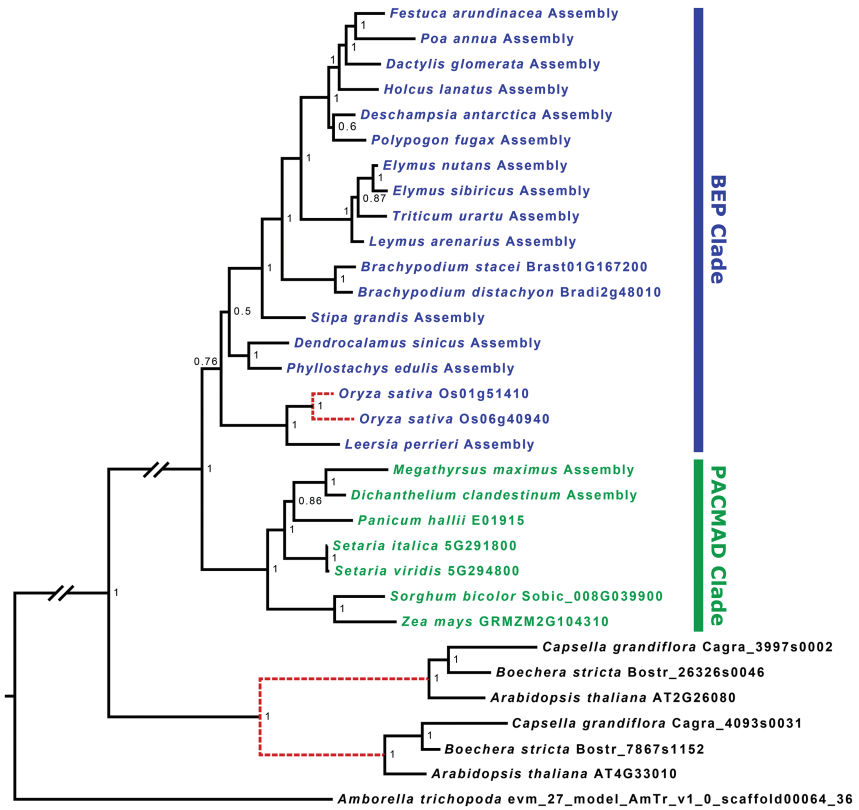

Fig. 8. A Bayesian phylogenetic tree of GLDP nucleotide sequences from 24 grasses, three Brassicaceae species, and *Amborella*. Long branches between distantly related groups are condensed for visibility, denoted with a gap. Dashed red lines denote inferred gene duplication events. Numbers at nodes indicate posterior probability.

trap (Hattersley and Browning, 1981; Hattersley *et al.*, 1986; Dengler and Nelson, 1999; Edwards and Voznesenskaya, 2011). The C_2 species *Alloteropsis semialata* ssp *semialata* has a similar C_2 Kranz anatomy to *N. minor*; GDC levels are highest in mestome sheath cells with a suberized lamella and the abundant organelles are equally partitioned within those cells (Hattersley and Browning, 1981; Ueno and Sentoku, 2006). Intriguingly, although organelle orientation in mestome sheath cells is not important in the evolution of C_2 photosynthesis in *N. minor*, chloroplasts do have a centrifugal orientation in C_4 *Neurachne* species (Hattersley *et al.*, 1986). Qualitative observations on C_3 *Neurachne* species indicate that some of the C_3 species have enhanced numbers of chloroplasts and mitochondria in mestome sheath cells (Hattersley *et al.*, 1986), leading us to posit that organelle and GDC enrichment may have been important during the early stages of C_2 evolution in the genus.

The evolutionary transition from C_3 to C_2 has been proposed first to involve a change in cell type-specific expression of GDC from M to BS in tandem with a loss of M GDC (Bauwe, 2011). A comparison of the cellular site of GDC expression in proto-Kranz *S. laxum* with that of the sister species *S. hians* indicates that the severe reduction in M GDC in the C_2 species is preceded by increased expression in BS cells. This is consistent with patterns observed in the eudicots *Flaveria* and *Heliotropium*, and supports a model of gradual GDC loss in M cells following a physiological activation of the BS (Muhaidat *et al.*, 2011; Sage *et al.*, 2013; Schulze *et al.*, 2013). C_3 species of *Flaveria* contain two copies of the gene encoding GLDP resulting from gene duplication (Schulze *et al.*, 2013). One of these is BS dominant in expression and the second is expressed ubiquitously throughout the leaf in C_3 *Flaveria* (Schulze *et al.*, 2013). During the evolution of C_4 photosynthesis, the loss of M GDC function in *Flaveria* resulted from pseudogenization of the gene coding for the ubiquitously expressed GLDP (Schulze *et al.*, 2013). Schulze *et al.* (2013) speculated that BS-dominant GLDP expression arose in a similar manner in C_4 grasses, because *O. sativa* (C_3 BEP clade) has two GLDP genes, but *Z. mays* and other C_4 species in the C_4 PACMAD clade have only one. To provide an understanding of the evolution of BS-specific GDC expression in C_2 and C_4 grasses, we conducted phylogenetic analyses of genes encoding GLDP using 17 BEP and seven PACMAD grass species, three Brassicaceae species, and *Amborella* as outgroup. Our analyses demonstrate that, with the exception of *O. sativa*, all grasses have one copy of GLDP. The two copies of the genes encoding GLDP in rice are more closely related to each other than to any other GLDP gene included in the analysis and therefore represent a local duplication. Combined, the phylogenetic and immunohistochemical observations on C_3 and C_4 PACMAD species are consistent with a hypothesis that BS-dominant quantities of GDC in C_2 and C_4 grasses resulted from modifications in regulatory mechanisms controlling the levels of expression of a single GLDP gene present in M and BS cells; C_3 species have high levels of M GDC and low levels of BS GDC, and the opposite pattern is present in the C_2 and C_4 species. Mesophyll tissue specificity of phosphoenolpyruvate carboxylase has evolved through modification of *cis*-regulatory elements in C_4 *Flaveria* (Gowik *et al.*, 2004). Studies examining promoter regions of the GLDP subunits in closely related C_3, C_2, and C_4 species should provide insights into the evolution of the regulatory mechanisms that confer the requisite expression patterns for C_2 and subsequently C_4 photosynthesis in grasses.

Since it is conceivable that BS specificity of the GDC complex arose via duplication and pseudogenization of one of the other subunits, we also included analyses for GLDL, GLDH1 and GLDH2, and GLDT. In land plants, GLDH1 functions in photorespiration and GLDH2 is associated with C1 metabolism (Rajinikanth *et al.*, 2007). We found no evidence of broad duplication for either of the ancient GLDH copies or GLDT in the grasses; however, GLDL is encoded by two conserved paralogs in most grass species, and may have arisen from the whole-genome duplication in the ancestor of the grass family. GLDL is the only GDC subunit gene for which we find evidence of subfunctionalized and conserved paralogs analogous to the two GLDP copies in *Flaveria*. There is, however, no evidence that either of these copies has been lost or pseudogenized (e.g. via nonsense or frameshift mutation) in any C_4 species except *Z. mays*, which has a local duplication of GLDL1 and lacks a gene encoding GLDL2. Yet it is unlikely that this loss was involved in the evolution of C_4 photosynthesis, as GLDL2 is present in *S. bicolor*, which shares a common C_4 origin with maize (Grass Phylogeny Working Group II, 2012). The two GLDL paralogs in grasses may be partially reduntant with one another (Rajinikanth *et al.*, 2007). The nature of the subfunctionalization of GLDL which resulted in the evolutionary retention of two paralogs in Poaceae is not known, but our results indicate that these paralogs did not play a role in the evolution of C_4 photosynthesis analogous to the GLDP paralogs in *Flaveria*. It is similarly unlikely that the two ancestral copies of GLDH played a role in C_4 evolution analogous to that of GLDP in *Flaveria* as both are present in all grass species examined here, and only GLDH1 plays a role in photorespiration (Rajinikanth *et al.*, 2007).

Conclusion

Evolution of C_2 photosynthesis in the grasses and eudicots is conferred by organelle enrichment, BS- or mestome sheath-dominant GDC accumulation, and centripetal positioning of organelles when the BS is the carbon-concentrating tissue. In *Steinchisma*, the earliest recognized subcellular event that facilitates C_2 photosynthesis is the placement of BS mitochondria and peroxisomes exclusively to the centripetal pole. This feature, also present in eudicots, is posited to set in motion a feed-forward facilitation cascade that leads to C_2 and subsequently C_4 photosynthesis (Sage *et al.*, 2013). How and why changes in BS or mestome sheath GDC levels and organelle volume, and BS organelle positioning were initiated during the early stages of C_2 evolution in grasses remains a mystery. Identification of mechanisms controlling these processes should be a primary focus of research. The Arthropogoninae, the subtribe containing *Steinchisma* (Otachyrinae), and Neurachninae will be key to these future studies.

Supplementary data

Supplementary data are available at *JXB* online.

Table S1. List of species studied and source of species.

Table S2. Species for which RNA-seq data were downloaded and assembled for phylogenetic analysis of GDC subunits.

Table S3. Anatomical parameters of C_3, C_2, and C_4 leaves.

Table S4. Organelle distribution in bundle sheath cells of C_3 and C_2 species.

Table S5. Quantification of organelle numbers, size, and density of gold labeling (GLDP) in mesophyll and bundle sheath cells of C_3, C_2, and C_4 leaves.

Figure S1. Light micrographs of C_3 and C_4 species.

Figure S2. Bundle sheath cell ultrastructure of C_3 and C_4 species.

Figure S3. Leaf structure and anatomy and immunolocalization of GLDP in *N. minor*.

Figure S4. Immunolocalization of GLDP in M and BS cells of C_3 and C_4 species.

Figure S5. Immunolocalization of Rubisco large subunit in M and BS cells of C_3 and C_2 species.

Figure S6. Immunolocalization of Rubisco large subunit in M and BS cells of C_3 and C_4 species.

Figure S7. A Bayesian phylogenetic tree of GLDH1.

Figure S8. A Bayesian phylogenetic tree of GLDH2.

Figure S9. A Bayesian phylogenetic tree of GLDL.

Figure S10. A Bayesian phylogenetic tree of GLDT.

Acknowledgements

We thank E. Kellog, T. Brutnell, and M. Tanaiguchi for providing seed, and N. Dakin for assistance with preparation of *N. minor* leaves. This research was supported by funding from the Natural Science and Engineering Research Council of Canada (grant nos 2015-04878 to TLS and 154273-2012 to RFS) and an Australian Research Council Discovery Project Grant to ML, TLS, and RFS (DP130102243).

References

Bauwe H. 2011. Photorespiration: the bridge to C_4 photosynthesis. In: Raghavendra AS, Sage RF, eds. C_4 photosynthesis and related CO_2 concentrating mechanisms. Advances in Photosynthesis and Respiration, Vol. **32**. Dordrecht, The Netherlands: Springer, 81–108.

Brown H, Bouton JH, Rigsby L, Rigler M. 1983. Photosynthesis of grass species differing in carbon dioxide fixation pathways. VIII. Ultrastructural characteristics of *Panicum* species in the LAXA group. Plant Physiology **71,** 425–431.

Busch FA, Sage TL, Cousins AB, Sage RF. 2013. C_3 plants enhance rates of photosynthesis by reassimilating photorespired and respired CO_2. Plant, Cell and Environment **36,** 200–212.

Capella-Gutierrez S, Silla-Martinez J, Gabaldon T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics **25,** 1972–1973.

Christin P-A, Osborne CP, Chatelet DS, Columbus JT, Besnard G, Hodkinson TR, Garrison LM, Vorontsova MS, Edwards EJ. 2013. Anatomical enablers and the evolution of C_4 photosynthesis in grasses. Proceedings of the National Academy of Sciences, USA **110,** 1381–1386.

Christin PA, Wallace MJ, Clayton H, Edwards EJ, Furbank RT, Hattersley PW, Sage RF, Macfarlane TD, Ludwig M. 2012. Multiple photosynthetic transitions, polyploidy, and lateral gene transfer in the grass subtribe Neurachninae. Journal of Experimental Botany **63,** 6297–6308.

Dengler NG, Nelson T. 1999. Leaf structure and development in C_4 plants In: Sage RF, Monson RK, eds. C_4 plant biology . New York: Academic Press, 133–172.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research **32,** 1792–1797.

Edwards GE, Ku MSB. 1987. Biochemistry of C_3–C_4 intermediates. In: Hatch MD, Boardman HK, eds. The biochemistry of plants, Vol. 10. London: Academic Press, 275–325.

Edwards GE, Ku MSB, Hatch MD. 1982. Photosynthesis in *Panicum milioides*, a species with reduced photorespiration. Plant and Cell Physiology **23,** 1185–1195.

Edwards GE, Voznesenskaya EV. 2011. C_4 photosynthesis: Kranz forms and single-cell C_4 in terrestrial plants. In: Raghavendra AS, Sage RF, eds. C_4 photosynthesis and related CO_2 concentrating mechanisms. Advances in Photosynthesis and Respiration, Vol. **32**. Dordrecht, The Netherlands: Springer, 29–61.

Emanuelsson O, Nielsen H, Brunak S, von Heijne G. 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. Journal of Molecular Biology **300,** 1005–1016.

Fisher AE, McDade LA, Kiel CA, Khoshravesh R, Johnson MA, Stata M, Sage TL, Sage RF. 2015. Evolutionary history of *Blepharis* (Acanthaceae) and the origin of C_4 photosynthesis in section Acanthodium. International Journal of Plant Sciences **176,** 770–790.

Gowik U, Burscheidt J, Akyildiz M, Schlue U, Koczor M, Streubel M, Wessthoff P. 2004. cis-Regulatory elements for mesophyll-specific gene expression in the C_4 plant *Flaveria trinervia*, the promoter of the C_4 phosphoenolpyruvate carboxylase gene. The Plant Cell **16,** 1077–1090.

Grass Phylogeny Working Group II. 2012. New grass phylogeny resolves deep evolutionary relationships and discovers C_4 origins. New Phytologist **193,** 304–312.

Hattersley PW, Browning AJ. 1981. Occurrence of the suberized lamella in leaves of grasses of different photosynthetic types. I. In parenchymatous bundle sheaths and PCR ('Kranz') sheaths. Protoplasma **109,** 371–401.

Hattersley PW, Wong S-C, Perry S, Roksandic Z. 1986. Comparative ultrastructure and gas exchange characteristics of the C_3–C_4 intermediate *Neurachne minor* S. T. Blake (Poaceae). Plant, Cell and Environment **9,** 217–233.

Heckmann D, Schulze S, Denton A, Gowik U, Westhoff P, Weber APM, Lercher MJ. 2013. Predicting C4 photosynthesis evolution: modular, individually adaptive steps on a Mount Fuji fitness landscape. Cell **153,** 1579–1588.

Hibberd JM, Sheehy JE, Langdale JA. 2008. Using C_4 photosynthesis to increase the yield of rice—rationale and feasibility. Current Opinion in Plant Biology **11,** 228–231.

Holaday AS, Lee KW, Chollet R. 1984. C_3–C_4 intermediate species in the genus *Flaveria*: leaf anatomy, ultrastructure, and the effect of O_2 on the CO_2 compensation concentration. Planta **160,** 25–32.

Hylton CM, Rawsthorne S, Smith AM, Jones DA, Woolhouse HW. 1988. Glycine decarboxylase is confined to the bundle-sheath cells of leaves of C_3–C_4 intermediate species. Planta **175,** 452–459.

Keerberg O, Pärnik T, Ivanova H, Bassüner B, Bauwe H. 2014. C_2 photosynthesis generates about 3-fold elevated leaf CO_2 levels in the C_3–C_4 intermediate species *Flaveria pubescens*. Journal of Experimental Botany **65,** 3649–3656.

Khoshravesh R, Akhani H, Sage TL, Nordenstam B, Sage RF. 2012. Phylogeny and photosynthetic pathway distribution in *Anticharis* Endl. (Scrophulariaceae). Journal of Experimental Botany **63,** 5645–5658.

Lundgren MR, Christin PA, Gonzalez Escobar E, Ripley BS, Besnard G, Long CM, Hattersley PW, Ellis RP, Leegood RC, Osborne CP. 2015. Evolutionary implications of C_3–C_4 intermediates in the grass *Alloteropsis semialata*. Plant, Cell and Environment (in press).

Lutziger I, Oliver DJ. 2000. Molecular evidence of a unique lipoamide dehydrogenase in plastids: analysis of plastidic lipoamide dehydrogenase from *Arabidopsis thaliana*. FEBS Letters **484**, 12–16.

Lutziger I, Oliver DJ. 2001. Characterization of two cDNAs encoding mitochondrial lipoamide dehydrogenase from Arabidopsis. Plant Physiology **127**, 615–623.

Lyu MJ, Gowik U, Kelly S, *et al*. 2015. RNA-Seq based phylogeny recapitulates previous phylogeny of the genus *Flaveria* (Asteraceae) with some modifications. BMC Evolutionary Biology **15,** 116.

Mallmann J, Heckmann D, Brautigam A, Lercher MJ, Weber AP, Westhoff P, Gowik U. 2014. The role of photorespiration during the evolution of C_4 photosynthesis in the genus *Flaveria*. Elife **3,** e02478.

Marshall DM, Muhaidat R, Brown NJ, Liu Z, Stanley S, Griffiths H, Sage RF, Hibberd JM. 2007. *Cleome*, a genus closely related to Arabidopsis, contains species spanning a developmental progression from C_3 to C_4 photosynthesis. The Plant Journal **51,** 886–896.

McKown AD, Moncalvo JM, Dengler NG. 2005. Phylogeny of *Flaveria* (Asteraceae) and inference of C_4 photosynthesis evolution. American Journal of Botany **92,** 1911–1928.

Monson R, Rawsthorne S. 2000. CO_2 assimilation in C_3–C_4 intermediate plants. In: Leegood R, Sharkey T, von Caemmerer S, eds. Photosynthesis, Vol. **9**. Dordrecht, The Netherlands: Springer, 533–550.

Monson RK, Edwards GE, Ku MSB. 1984. C_3–C_4 intermediate photosynthesis in plants. BioScience **34,** 563–574.

Morgan JA, Brown RH. 1979. Photosynthesis in grass species differing in carbon dioxide fixation pathways: II. A search for species with intermediate gas exchange and anatomical characteristics. Plant Physiology **64,** 257–262.

Morgan CL, Turner SR, Rawsthorne S. 1993. Coordination of the cell-specific distribution of the 4 subunits of glycine decarboxylase and of serine hydroxymethyltransferase in leaves of C_3–C_4 intermediate species from different genera. Planta **190,** 468–473.

Muhaidat R, Sage TL, Frohlich MW, Dengler NG, Sage RF. 2011. Characterization of C_3–C_4 intermediate species in the genus *Heliotropium* L. (Boraginaceae): anatomy, ultrastructure and enzyme activity. Plant, Cell and Environment **34,** 1723–1736.

Osborne CP, Sack L. 2012. Evolution of C_4 plants: a new hypothesis for an interaction of CO_2 and water relations mediated by plant hydraulics. Philosophical Transactions of the Royal Society B: Biological Sciences **367,** 583–600.

Peterhansel C. 2011. Best practice procedures for the establishment of a C_4 cycle in transgenic C_3 plants. Journal of Experimental Botany **62,** 3011–3019.

Rawsthorne S. 1992. C_3–C_4 intermediate photosynthesis—linking physiology to gene-expression. The Plant Journal **2,** 267–274.

Rawsthorne S, Morgan CL, O'Neill CM, Hylton CM, Jones DA, Frean ML. 1998. Cellular expression pattern of the glycine decarboxylase P protein in leaves of an intergeneric hybrid between the C_3–C_4 intermediate species *Moricandia nitens* and the C_3 species *Brassica napus*. Theoretical and Applied Genetics **96,** 922–927

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19,** 1572–1574.

Rajinikanth M, Harding SA, Tsai CJ. 2007. The glycine decarboxylase complex multienzyme family in *Populus*. Journal of Experimental Botany **58,** 1761–1770.

Sage R. 2016. A portrait of the C_4 photosynthetic family on the 50th anniversary of its discovery: species number, evolutionary lineages, and hall of fame. Journal of Experimental Botany **67,** 2919–2922.

Sage RF, Christin PA, Edwards EJ. 2011. The C_4 plant lineages of planet Earth. Journal of Experimental Botany **62,** 3155–3169.

Sage RF, Khoshravesh R, Sage TL. 2014. From proto-Kranz to C_4 Kranz: building the bridge to C_4 photosynthesis. Journal of Experimental Botany **65,** 3341–3356.

Sage RF, Sage TL, Kocacinar F. 2012. Photorespiration and the evolution of C_4 photosynthesis. Annual Review of Plant Biology **63,** 19–47.

Sage RF, Wedin DA, Li M. 1999. The biogeography of C_4 photosynthesis: patterns and controlling factors. In: Sage RF, Monson RK, eds. C_4 plant biology . San Diego: Academic Press: 313–374.

Sage TL, Busch FA, Johnson DC, Friesen PC, Stinson CR, Stata M, Sultmanis S, Rahman BA, Rawsthorne S, Sage RF. 2013. Initial events during the evolution of C_4 photosynthesis in C_3 species of *Flaveria*. Plant Physiology **163,** 1266–1276.

Sage TL, Sage RF,Vogan PJ, Rahman B, Johnson DC, Oakley JC, Heckel MA. 2011. The occurrence of C_2 photosynthesis in *Euphorbia* subgenus Chamaesyce (Euphorbiaceae). Journal of Experimental Botany **62,** 3183–3195.

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. Nature Methods **9,** 671–675.

Schulze S, Mallmann J, Burscheidt J, Koczor M, Streubel M, Bauwe H, Gowik U, Westhoff P. 2013. Evolution of C_4 photosynthesis in the genus *Flaveria*: establishment of a photorespiratory CO_2 pump. The Plant Cell **25,** 2522–2535.

Stata M, Sage TL, Hoffman N, Covshoff S, Ka-Shu Wong G, Sage RF. 2016. Mesophyll chloroplast investment in C_3, C_4 and C_2 species of the genus *Flaveria*. Plant and Cell Physiology **57** (in press).

Stata M, Sage TL, Rennie TD, Khoshravesh R, Sultmanis S, Khaikin Y, Ludwig M, Sage RF. 2014. Mesophyll cells of C_4 plants have fewer chloroplasts than those of closely related C_3 plants. Plant, Cell and Environment **37,** 2587–2600.

Still CJ, Berry JA, Collatz GJ, DeFries RS. 2003. Global distribution of C_3 and C_4 vegetation: carbon cycle implications. Global Biogeochemical Cycles **17**.

Ueno O. 1992. Immunogold localization of photosynthetic enzymes in leaves of *Aristida latifolia*, a unique C_4 grass with a double chlorenchymatous bundle sheath. Physiologia Plantarum **85,** 189–196.

Ueno O, Bang SW, Wada Y, Kondo A, Ishihara K, Kaneko Y, Matsuzawa Y. 2003. Structural and biochemical dissection of photorespiration in hybrids differing in genome constitution between *Diplotaxis tenuifolia* (C_3–C_4) and radish (C_3). Plant Physiology **132,** 1550–1559.

Ueno O, Sentoku N. 2006. Comparison of leaf structure and photosynthetic characteristics of C_3 and C_4 *Alloteropsis semialata* subspecies. Plant, Cell and Environment **29,** 257–268.

Vogan PJ, Frohlich MW, Sage RF. 2007. The functional significance of C_3–C_4 intermediate traits in *Heliotropium* L. (Boraginaceae): gas exchange perspectives. Plant, Cell and Environment **30,** 1337–1345.

von Caemmerer S. 2000. Biochemical models of leaf photosynthesis . Australia: CSIRO publishing.

von Caemmerer S, Quick WP, Furbank RT. 2012. The development of C_4 rice: current progress and future challenges. Science **336,** 1671–1672.

Voznesenskaya EV, Artyusheva EG, Franceschi VR, Pyankov VI, Kiirats O, Ku MSB, Edwards GE. 2001. *Salsola arbusculiformis*, a C_3–C_4 intermediate in Salsoleae (Chenopodiaceae). Annals of Botany **88,** 337–348.

Voznesenskaya EV, Koteyeva NK, Akhani H, Roalson EH, Edwards GE. 2013. Structural and physiological analyses in Salsoleae (Chenopodiaceae) indicate multiple transitions among C_3, intermediate, and C_4 photosynthesis. Journal of Experimental Botany **64,** 3583–3604.

Voznesenskaya EV, Koteyeva NK, Edwards GE, Ocampo G. 2010. Revealing diversity in structural and biochemical forms of C_4 photosynthesis and a C_3–C_4 intermediate in genus *Portulaca* L. (Portulacaceae). Journal of Experimental Botany **61,** 3647–3662.

Williams BP, Johnston IG, Covshoff S, Hibberd JM. 2013. Phenotypic landscape inference reveals multiple evolutionary paths to C_4 photosynthesis. Elife **2,** e00961.