

# Natural Variation within a Species for Traits Underpinning $C_4$ Photosynthesis<sup>1[CC-BY]</sup>

Gregory Reeves,<sup>a,2</sup> Pallavi Singh,<sup>a,2</sup> Timo A. Rossberg,<sup>a</sup> E.O. Deedi Sogbohossou,<sup>b</sup> M. Eric Schranz,<sup>b</sup> and Julian M. Hibberd<sup>a,3</sup>

<sup>a</sup>Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, United Kingdom

<sup>b</sup>Biosystematics Group, Department of Plant Sciences, Wageningen University, Wageningen 6700 AA, The Netherlands

ORCID IDs: 0000-0001-9805-4899 (T.A.R.); 0000-0002-9034-0062 (E.O.D.S.); 0000-0001-6777-6565 (M.E.S.); 0000-0003-0662-7958 (J.M.H.).

Engineering  $C_4$  photosynthesis into  $C_3$  crops could substantially increase their yield by alleviating photorespiratory losses. This objective is challenging because the  $C_4$  pathway involves complex modifications to the biochemistry, cell biology, and anatomy of leaves. Forward genetics has provided limited insight into the mechanistic basis of these properties, and there have been no reports of significant quantitative intraspecific variation of  $C_4$  attributes that would allow trait mapping. Here, we show that accessions of the  $C_4$  species *Gynandropsis gynandra* collected from locations across Africa and Asia exhibit natural variation in key characteristics of  $C_4$  photosynthesis. Variable traits include bundle sheath size and vein density, gas-exchange parameters, and carbon isotope discrimination associated with the  $C_4$  state. The abundance of transcripts encoding core enzymes of the  $C_4$  cycle also showed significant variation. Traits relating to water use showed more quantitative variation than those associated with carbon assimilation. We propose that variation in these traits likely adapted the hydraulic system for increased water use efficiency rather than improving carbon fixation, indicating that selection pressure may drive  $C_4$  diversity in *G. gynandra* by modifying water use rather than photosynthesis. The accessions analyzed can be easily crossed and produce fertile offspring. Our findings, therefore, indicate that natural variation within this  $C_4$  species is sufficiently large to allow genetic mapping of key  $C_4$  traits and regulators.

Plants that use  $C_4$  photosynthesis can effectively abolish photorespiratory losses that occur when Rubisco fixes oxygen rather than  $CO_2$  (Bowes et al., 1971; Sharkey, 1988). In  $C_4$  plants, Rubisco is typically sequestered in bundle sheath (BS) cells that are concentrically arranged around the vasculature. The establishment of a molecular  $CO_2$  pump delivers carbon to Rubisco from mesophyll (M) cells via  $C_4$  acid intermediates (Garner et al., 2001).  $C_4$  photosynthesis relies more heavily on the BS for photosynthesis and less on M cells. It also involves more chloroplasts in BS cells, an increased proliferation of plasmodesmata between M and BS cells, and a higher vein density to increase the volume of the leaf occupied by the BS. These morphological alterations to the leaf that facilitate the  $C_4$  cycle are known as Kranz anatomy (Langdale, 2011;

Lundgren et al., 2014). Moreover, photosynthesis gene expression is modified such that genes encoding components of the  $C_4$  and Calvin-Benson-Bassham cycles are strongly and preferentially expressed in either M or BS cells (Garner et al., 2001; Marshall et al., 2007; Hibberd and Covshoff, 2010).

Despite the complex modifications associated with  $C_4$  photosynthesis, current estimates are that the  $C_4$  pathway has independently evolved more than 60 times in angiosperms (Sage, 2016), which suggests that a relatively straightforward route may allow the transition from the ancestral  $C_3$  state to the derived  $C_4$  state. A number of factors appear to have facilitated this apparently complex transition. First, proteins of the  $C_4$  pathway operate in the ancestral  $C_3$  state (Hibberd and Quick, 2002; Brown et al., 2010; Eastmond et al., 2015). Second, the modifications to gene expression associated with  $C_4$  leaves appear to have been built on partly preexisting regulatory architecture found in  $C_3$  species (Brown et al., 2011; Kajala et al., 2012; Burgess et al., 2016; Williams et al., 2016; Reyna-Llorens et al., 2018). Moreover, after a binary categorization of phenotypes into either  $C_3$  or  $C_4$ , mathematical modeling indicates that there is likely more than one route from which to acquire the  $C_4$  state (Williams et al., 2013) but that subsequent modifications, many of which are associated with rebalancing photorespiration, reinforce these improvements in photosynthesis and fitness (Heckmann et al., 2013; Mallmann et al., 2014; Bräutigam and Gowik, 2016).

Genome-wide analysis of transcript abundance in multiple  $C_3$  and  $C_4$  species has provided unbiased in-

<sup>1</sup>G.R. was supported by a Gates Cambridge Trust PhD Fellowship and P.S. by Advanced European Research Council grant 694733 Revolution to J.M.H.

<sup>2</sup>These authors contributed equally to the article.

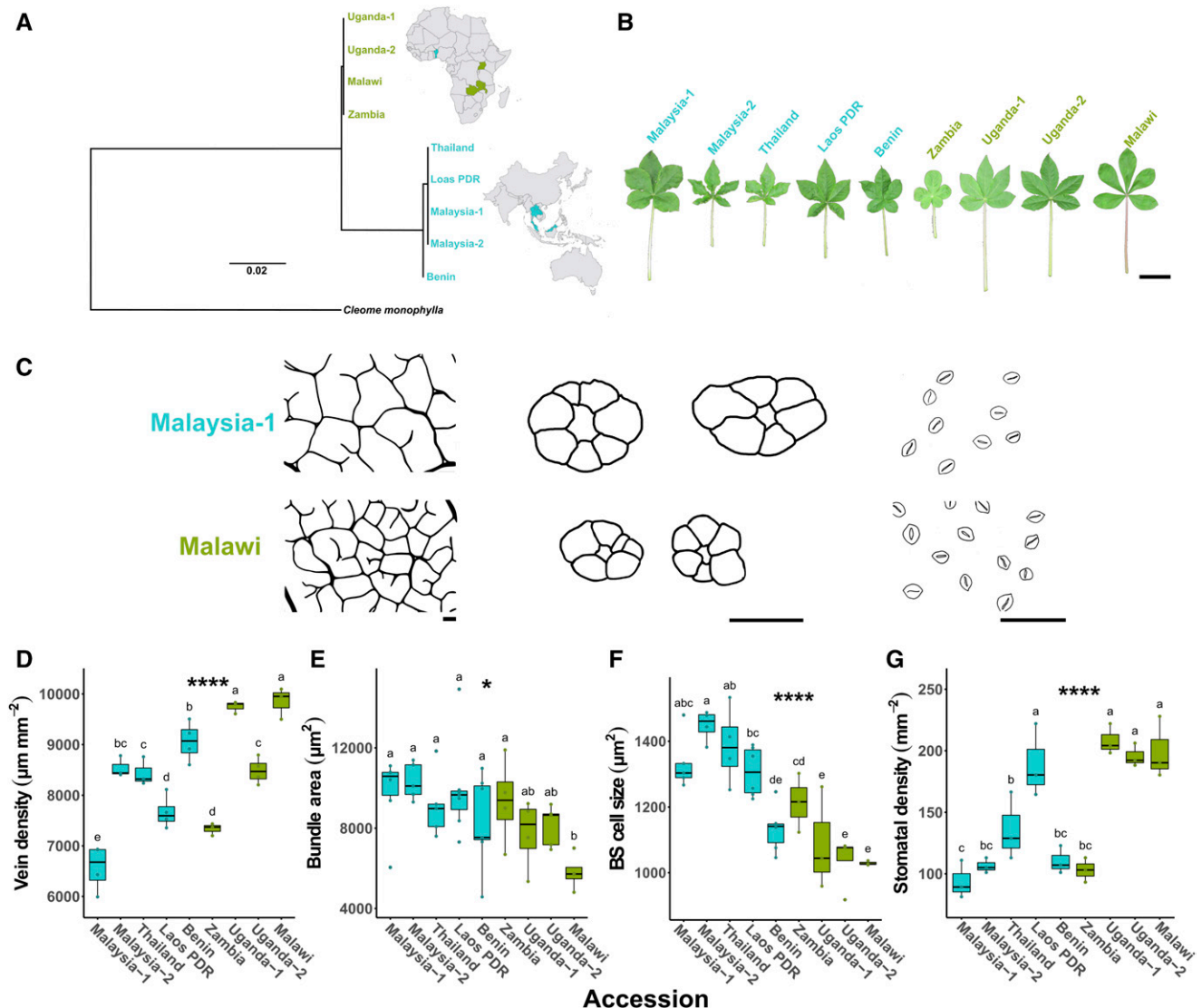
<sup>3</sup>Address correspondence to [jmh65@cam.ac.uk](mailto:jmh65@cam.ac.uk).

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors ([www.plantphysiol.org](http://www.plantphysiol.org)) is: Julian M. Hibberd ([jmh65@cam.ac.uk](mailto:jmh65@cam.ac.uk)).

G.R., P.S., and J.M.H. designed the study; G.R., P.S., T.A.R., and E.O.D.S. carried out experimental work; G.R., P.S., T.A.R., M.E.S., and J.M.H. wrote the article.

[CC-BY] Article free via Creative Commons CC-BY 4.0 license.

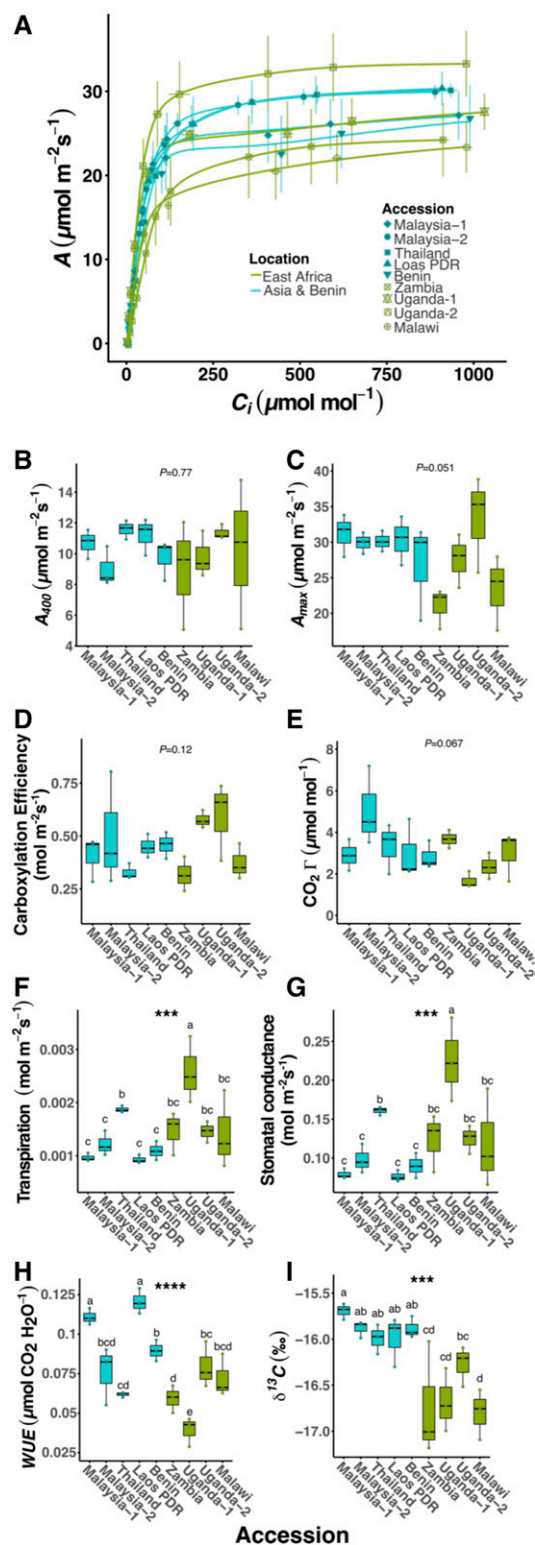
[www.plantphysiol.org/cgi/doi/10.1104/pp.18.00168](http://www.plantphysiol.org/cgi/doi/10.1104/pp.18.00168)



**Figure 1.** Natural variation in Kranz anatomy among a diverse panel of *G. gynandra* accessions. A, Geographic and phylogenetic relationships for nine accessions from seven countries across Africa and Asia. B, Variation in fully mature whole leaves of 6-week-old plants. C, Traces indicate variation in venation, bundle area, BS cell size, and stomatal density of fully mature leaves for two extreme accessions. D to G, Vein density (D), bundle area (E), BS cell size (F), and stomatal density (G) for all accessions. Asterisks indicate significant differences between accessions (one-way ANOVA: \*,  $P < 0.05$ ; and \*\*\*\*,  $P < 0.0001$ ). Letters above individual box-scatterplots indicate significant groupings according to Duncan's multiple range test ( $\alpha = 0.05$ ). The number of replicates for each accession ranged from three to seven. Bars = 5 cm (B) and 100  $\mu\text{m}$  (C).

sight into processes that likely change in C<sub>4</sub> compared with C<sub>3</sub> leaves (Aubry et al., 2014; Bräutigam et al., 2014; Lauterbach et al., 2017). Furthermore, cis-elements that control the expression of genes encoding components of the C<sub>4</sub> cycle have been documented (Akyildiz et al., 2007; Williams et al., 2016; Reyna-Llorens et al., 2018). However, the regulators that recognize these motifs have not been isolated (Reeves et al., 2017). Thus, despite the progress in our understanding of C<sub>4</sub> photosynthesis, it is currently not possible to rationally design a C<sub>4</sub> pathway in a C<sub>3</sub> leaf.

When natural variation is present, the development of a mapping population enables molecular marker-trait associations by methods such as quantitative trait locus mapping and genome-wide association studies. Molecular marker-trait association studies help delineate genomic regions associated with trait variation. Fine genetic mapping of these genomic regions with major effects allows the identification of causal genes underpinning the variation. Such approaches have been used extensively to map loci responsible for numerous complex traits in plants (Mauricio, 2001).



**Figure 2.** Physiological variation for photosynthetic gas-exchange parameters among a diverse panel of *G. gynandra* accessions. A,  $A$  versus  $C_i$  response curve. B,  $A_{400}$  ( $400 \mu\text{L L}^{-1}$  atmospheric  $\text{CO}_2$  concentration [ $C_i$ ] and photosynthetic photon flux density [PPFD] of  $350 \mu\text{mol m}^{-2} \text{s}^{-1}$ ). C,  $A_{\text{max}}$  ( $1,200 \mu\text{L L}^{-1}$   $C_a$  and PPFD of  $2,000 \mu\text{mol m}^{-2} \text{s}^{-1}$ ). D, F, E, Carboxylation efficiency. F, Transpiration. G, Stomatal conductance. H, WUE. I,  $\delta^{13}\text{C}$ . Asterisks indicate significant differences between acces-

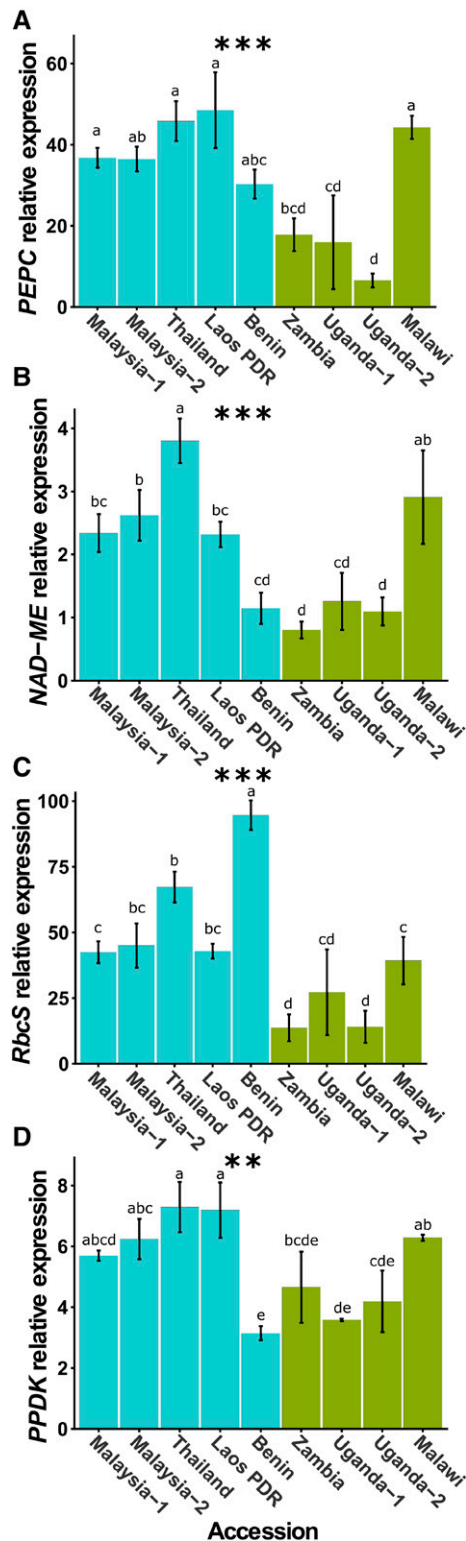
sions (one-way ANOVA,  $n = 3$ ; \*\*\*,  $P < 0.001$ ; and \*\*\*\*,  $P < 0.0001$ ). Letters above individual box-scatterplots indicate significant groupings according to Duncan's multiple range test ( $\alpha = 0.05$ ).

The application of molecular marker-trait associations to study  $C_4$  photosynthesis would expedite the discovery of key regulators to engineer increased photosynthetic efficiency in  $C_3$  plants. However, to our knowledge, there are currently no examples showing that variation in  $C_4$  traits within a single species is sufficient to allow breeding and molecular trait mapping. Interspecific hybrids have been generated between  $C_3$  and  $C_4$  species of the dicotyledon genus *Atriplex* (Oakley et al., 2014). Although the progeny possessed variation in  $C_4$  phenotypes, specific traits showed limited penetrance, and there were high rates of sterility (Brown and Bouton, 1993). In the grass family, *Allotheropsis semialata* shows natural variation in  $C_4$  parameters and has been classified into  $C_3$  or  $C_4$  subspecies (Ueno and Sentoku, 2006; Lundgren et al., 2016), but there are currently no reports that these populations have been bred. Therefore, we investigated the extent to which key  $C_4$  traits varied in the  $C_4$  dicotyledon *Gynandropsis gynandra*, which is a leafy green vegetable (Sogbohossou et al., 2018) in a clade sister to the Brassicaceae that contains both  $C_3$  and  $C_4$  species (Brown et al., 2005; Marshall et al., 2007; Feodorova et al., 2010). Here, we show that accessions of *G. gynandra* show significant variation in both anatomical and physiological traits associated with  $C_4$  photosynthesis. These accessions have short generation spans, are sexually compatible, and produce fertile offspring. Our findings indicate that there is sufficient natural variation to allow the use of classical genetics to identify loci controlling the complex  $C_4$  phenotype.

## RESULTS AND DISCUSSION

Accessions of *G. gynandra* were collected from Africa and Asia (Supplemental Table S1). DNA sequencing and phylogenetic reconstruction generated a taxonomy that was consistent with geographical origin but also indicated that the accession from Benin (West Africa) was more similar to Asian accessions than to those from East Africa (Fig. 1A). These accessions displayed considerable variation in macroscopic characteristics associated with leaf appearance (Fig. 1B; Supplemental Fig. S1A). For example, fully expanded leaflets varied in size and shape, and there was variation in petiole length, the presence of trichomes, and anthocyanin pigmentation. As there was considerable macroscopic variation in leaf characteristics, we then evaluated these accessions for variation in features of Kranz anatomy. Interestingly, there were statistically significant differences in vein density (Fig. 1, C and D; Supplemental Fig. S1B), cross-sectional area of BS strands (Fig. 1, C and E; Supplemental Fig. S1C), size of individual BS cells (Fig. 1F), and stomatal density (Fig. 1, C



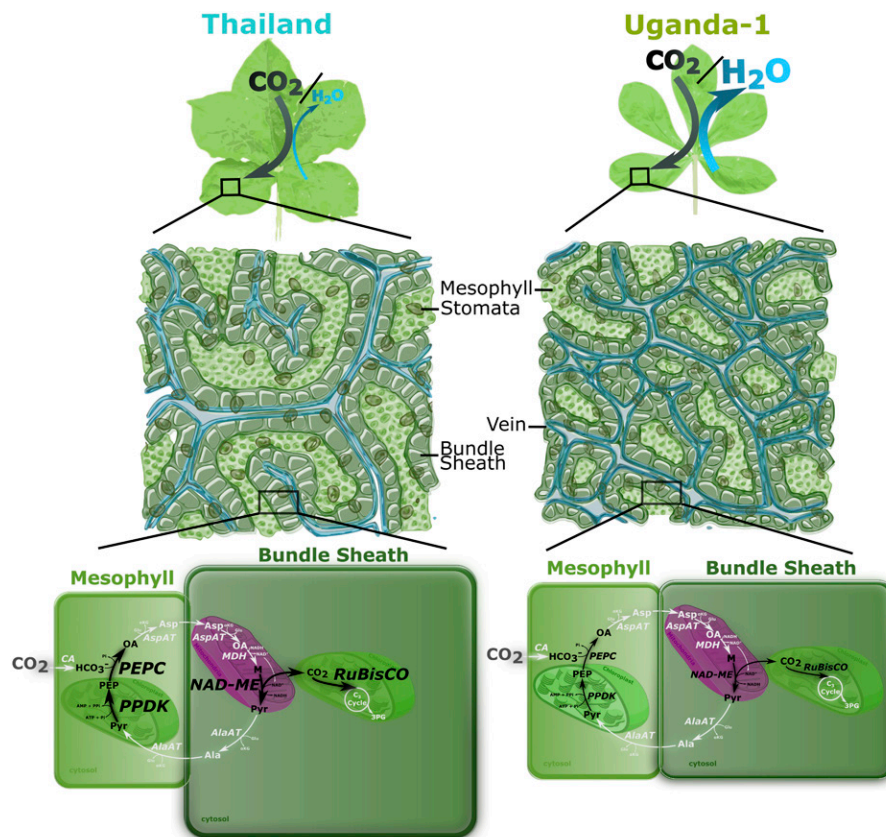


**Figure 3.** Transcript abundance for key enzymes in the C<sub>4</sub> cycle among diverse *G. gynandra* accessions. Differences in transcript abundance were determined by reverse transcription-quantitative PCR (RT-qPCR) relative to the *ACTIN7* gene. Asterisks indicate significant differences between accessions (one-way ANOVA,  $n = 3$ : \*\*,  $P < 0.01$ ; and \*\*\*,  $P < 0.001$ ). Letters above individual bar charts indicate significant groupings among accessions according to Duncan's multiple range test ( $\alpha = 0.05$ ).

and G). There was no discernible variation in intervein distance or number of BS cells surrounding each vein (Supplemental Fig. S2). However, on average, East African accessions showed higher vein density, reduced distance between veins, and greater stomatal density than Asian accessions (Supplemental Fig. S3, A–C). Asian accessions typically had larger BS areas and cell sizes than those from the African continent (Supplemental Fig. S3, D and E). Vein density was inversely correlated with BS area and BS cell size but positively correlated with stomatal density (Supplemental Table S2). The average number of BS cells around each vein showed no statistically significant differences between lines (Supplemental Fig. S3F), but the cross-sectional area of the BS and the size of individual BS cells were positively correlated ( $\rho = 0.8$ ,  $P < 0.0001$ ). Therefore, we conclude that the area of individual BS cells, rather than the number of these cells per vein bundle, drives the increased BS strand area observed in some accessions. This suggests that genetic determinants of cell size rather than cell proliferation are involved in the variation in BS tissue in *G. gynandra*. Despite the lower phenotypic variation associated with C<sub>4</sub> compared with C<sub>3</sub> leaves (Sage and McKown, 2006), our findings demonstrate that flexibility is still possible within individual species that are fully C<sub>4</sub>.

We next investigated whether differences observed in Kranz anatomy affected photosynthetic performance. For all accessions, their CO<sub>2</sub> response curves (assimilation [*A*] and response to the concentration of CO<sub>2</sub> inside the leaf [*C<sub>i</sub>*]) were typical of C<sub>4</sub> plants, with high carboxylation efficiencies and low CO<sub>2</sub> compensation points ( $\Gamma$ ; Fig. 2A; Supplemental Fig. S4A). Parameters associated with instantaneous gas exchange, such as maximum rate of photosynthesis ( $A_{max}$ ), rate of photosynthesis under the conditions of growth ( $A_{400}$ ), CO<sub>2</sub> carboxylation efficiencies, and  $\Gamma$ , showed little variation between accessions (Fig. 2, B–E; Supplemental Fig. S4, B–E). As C<sub>4</sub> species typically grow under relatively high light intensities, variation in these traits may become apparent at higher light. In fact, modifications to stomatal density in *Arabidopsis* impact assimilation rates at high light but not at lower light intensities (Schlüter et al., 2003). Thus, it is possible that, when grown under different conditions, these *G. gynandra* accessions would demonstrate additional variation in photosynthetic characteristics.

There were statistically significant differences in transpiration (Fig. 2F), stomatal conductance (Fig. 2G), and water use efficiency (WUE; Fig. 2H). Furthermore, there was also significant variation in the carbon isotope discrimination against <sup>13</sup>C ( $\delta^{13}$ C) in leaf dry matter (Fig. 2I), which is a measure of the efficiency of the C<sub>4</sub> carbon pump over the lifetime of the leaf. Accessions from Asia-Benin showed reduced discrimination against <sup>13</sup>C compared with the East African lines (Supplemental Fig. S4I). Therefore, these data indicate that accessions of *G. gynandra* possess significant variation in parameters linked to the balance between water use



**Figure 4.** Summary of differences in anatomy, physiology, and  $\text{C}_4$ -related transcript abundance exhibited by extreme *G. gynandra* accessions. The accessions Thailand and Uganda-1 had larger BS cells, lower vein and stomatal density, and higher  $\text{WUE}$ ,  $\delta^{13}\text{C}$ , and transcript abundance of core  $\text{C}_4$  cycle enzymes. Differential transcript abundance among accessions is indicated by black arrows, and larger font represents higher relative transcript abundance. Enzymatic steps in white were not investigated.

and photosynthesis that influenced the efficiency of the  $\text{C}_4$  cycle over their lifetimes.

We next sought to investigate the extent to which the transcript abundance of core genes in the  $\text{C}_4$  cycle differed between the accessions. Interestingly, there were statistically significant differences in the abundance of transcripts encoding phosphoenolpyruvate carboxylase ( $\text{PEPC}$ ), which catalyzes the first committed step of the  $\text{C}_4$  cycle, the BS-specific decarboxylase NAD-dependent malic enzyme ( $\text{NAD-ME}$ ), the small subunit of Rubisco ( $\text{RbcS}$ ), and pyruvate, orthophosphate dikinase ( $\text{PPDK}$ ) that regenerates phosphoenolpyruvate, the primary acceptor of  $\text{HCO}_3^-$  (Fig. 3). These differences in  $\text{C}_4$ -related transcript abundance were associated with the geographical location and phylogeny of the accessions, with the accessions from Asia-Benin accumulating greater levels of  $\text{C}_4$ -related transcripts than those from East Africa (Supplemental Fig. S5, A–D). Understanding how the genes encoding photosynthetic enzymes become strongly expressed and patterned to either M or BS cells of  $\text{C}_4$  species is a longstanding area of research. Although progress has been made in identifying the cis-elements responsible

for this phenomenon, little is known about the transcription factors involved. Therefore, the intraspecific variation in expression of genes encoding enzymes of the  $\text{C}_4$  cycle in *G. gynandra* provides an opportunity to identify trans-factors important for  $\text{C}_4$  photosynthesis. Despite accessions functioning with similar photosynthetic efficiencies under ambient  $\text{CO}_2$  and light conditions, when assessed by phylogenetic grouping, those with more pronounced Kranz traits (e.g. larger BS tissues and lower vein densities) exhibited increased  $A_{\text{max}}$ ,  $\text{WUE}$ , and  $\delta^{13}\text{C}$  (Supplemental Fig. S4, C, H, and I) and stronger expression of the core genes of the  $\text{C}_4$  cycle (Supplemental Fig. S5). To summarize, there were striking differences between accessions, such that some possessed higher  $\text{WUE}$ , lower density of stomata and veins, larger BS areas and cell sizes, higher  $\delta^{13}\text{C}$  (indicative of a stronger  $\text{C}_4$  cycle), and increased expression of genes encoding  $\text{C}_4$  enzymes (Fig. 4).

The considerable variation reported in this study offers a valuable germplasm resource to identify regulators of the  $\text{C}_4$  pathway and Kranz anatomy through genetic mapping. All accessions in this study

hybridize easily. Emasculation and pollination need only take 15 to 30 s per flower. For example, the most divergent accessions regarding anatomy, Malaysia-1 × Malawi, Malaysia-2 × Malawi, and their reciprocal crosses, produce an average  $52 \pm 11$  seeds per silique ( $n = 6$ ), whose offspring are fully fertile. The F<sub>1</sub> hybrids displayed intermediate characteristics, differing in leaflet number, presence of trichomes on the petiole, and leaf size and shape (Supplemental Fig. S6). Therefore, these hybrid populations provide a breeding foundation to delineate regulatory mechanisms. More broadly, the discovery of intraspecific variation in a C<sub>4</sub> grass would be particularly useful in mapping traits relevant to improving photosynthesis in crops and, thus, introduce C<sub>4</sub> photosynthesis into C<sub>3</sub> cereals.

While our understanding of the regulatory mechanisms underlying C<sub>4</sub> metabolism is growing, there is still a significant gap in tools to expand our understanding of the regulation behind Kranz anatomy and the C<sub>4</sub> biochemical cycle. Methods such as quantitative trait locus mapping or genome-wide association studies in *G. gynandra* or an equally diverse C<sub>4</sub> species may provide insights into the regulation of Kranz development. In this study, most trait variation in *G. gynandra* was associated with characteristics relating to water use that impact carbon capture. It is noteworthy that modifications to C<sub>3</sub> leaves considered to represent early steps on the path toward the C<sub>4</sub> phenotype also are associated with water use rather than CO<sub>2</sub> fixation (Sage, 2004; Williams et al., 2013). As natural vegetation is not considered to be under strong selection pressure to optimize photosynthesis (Long et al., 2015; Ort et al., 2015), it seems likely that C<sub>4</sub> trait variation continues to be driven at least partially through the optimization of water use rather than photosynthesis per se.

## MATERIALS AND METHODS

### Plant Accessions and Growth Conditions

A selection of nine diverse accessions of *Gynandropsis gynandra* was made from a larger germplasm collection based on initial phenotypic and genetic screening. Five accessions were from Africa and four from Asia (Supplemental Table S1; materials available on request from M.E.S.). Plants from all *G. gynandra* accessions were grown under identical conditions prior to sampling. After germination, all seeds were planted in 5:1 F2 compost (Levington Advance):fine vermiculite premixed with 0.17 g L<sup>-1</sup> insecticide (Imidasect 5GR; Fargro) in 5-cm<sup>3</sup> cells. Seedlings were kept in a growth chamber at 350 μmol photons m<sup>-2</sup> s<sup>-1</sup> PPFD, with a 16-h photoperiod, at 25°C, 60% relative humidity (RH), and ambient [CO<sub>2</sub>]. A single dose of 3 mL L<sup>-1</sup> slow-release 17N-9P-11K fertilizer (All Purpose Continuous Release Plant Food; Miracle-Gro) was applied after 1.5 weeks. Plants for physiological measurements were grown under identical conditions to those for Kranz measurements for the first 3 weeks, after which the plants for Kranz measurements were assayed and the plants for physiological measurements were replanted in 13-cm<sup>3</sup> pots with a 5:1 M3 soil (Levington Advance):medium vermiculite soil mixture and moved to a growth room set to 23°C, 60% RH, ambient [CO<sub>2</sub>], and 350 μmol photons m<sup>-2</sup> s<sup>-1</sup> PPFD with a 16-h photoperiod. All plants were watered by an automated system whereby the bottom of the trays containing the 13-cm<sup>3</sup> pots was flooded to a depth of 4 cm every 48 h for 10 min, following which the irrigation water was drained.

### DNA Sequencing and Phylogenetic Analyses

Genomic DNA was extracted using a cetyltrimethylammonium bromide extraction method. Briefly, leaf samples were ground uniformly in 300 μL of cetyltrimethylammonium bromide extraction buffer containing 2% (w/v) polyvinylpyrrolidone-40 and 1.7% (v/v) β-mercaptoethanol. After vortexing, samples were incubated at 65°C for 30 min, and 300 μL of 24:1 chloroform:isoamyl alcohol was then added. After vortexing and centrifugation, 300 μL of supernatant was removed and placed in 300 μL of isopropanol. DNA was precipitated overnight at -20°C and then centrifuged at full speed for 15 min. The supernatant was removed, the pellet was washed with 200 μL of 96% (v/v) ethanol prior to drying, and the DNA was resuspended in nuclease-free water.

To generate a phylogeny for the accessions, internal transcribed spacer (ITS) regions were sequenced using universal primers ITS 1 and ITS 4, following a protocol published previously (White et al., 1990). In short, ITS regions I and II and the 5.8 S polycistronic rRNA precursor transcribed region were amplified by PCR. These PCR products were purified, and fragments were sequenced. Sequences were then aligned using the MAFFT online server (MAFFT version 7; <https://mafft.cbrc.jp/alignment/server/>), and a best scoring maximum likelihood tree was generated with RAxML version 8.0.0 on the CIPRES science gateway (Miller et al., 2010).

### Preparation of Leaf Tissue Sections for Kranz Measurements

Three weeks after germination, tissue was harvested from healthy plants from the center trifoliate leaves of the second pair of fully expanded true leaves. A 3-mm<sup>2</sup> rectangle was cut from the leaf adjacent to the midvein with a razor blade for transverse sections. Two slightly larger rectangles were cut from identical regions for paradermal sectioning and RT-qPCR analysis.

For transverse sections, leaf tissue in plastic cuvettes was submerged in a 4% (v/v) paraformaldehyde in PBS solution (Sigma-Aldrich), placed in a vacuum chamber for 1 h, and incubated at 4°C overnight for fixing. Cuvettes then underwent an ethanol dehydration series from 30% to 90% (v/v) ethanol solutions (Thermo Fisher Scientific) in 10% (v/v) increments for 45 min each at 4°C with a final overnight treatment at 4°C in 95% ethanol with 0.1% (w/v) eosin dye solution (Sigma-Aldrich). The dye solution was washed three times with 100% (v/v) ethanol at room temperature. The samples were embedded in resin according to the Technovit 7100 (Kulzer) manufacturer's protocol. Hardened resin blocks were cut with a manual rotary microtome (Thermo Fisher Scientific). Sections were placed on microscope slides and stained with 0.1% (w/v) Toluidine Blue solution (Sigma-Aldrich) prior to imaging on a light microscope.

For paradermal sections, fresh tissue samples were placed in plastic cuvettes and incubated in 3:1 100% (v/v) ethanol:acetic acid solution before treatment with 70% (v/v) ethanol solution (refreshed once) at 37°C overnight. To clear the samples, cuvettes were submerged in 5% NaOH solution for 3 h at 37°C, rinsed twice with 70% (v/v) ethanol, and stored at 4°C. After storage, the samples were stained with a 0.1% (v/v) eosin dye in 95% (v/v) ethanol solution. Samples were stored overnight at 4°C and washed with 70% (v/v) ethanol three times before transfer to slides for imaging. To determine stomatal density, impressions of the abaxial epidermis of each central leaflet were generated by applying a thin coat of transparent nail varnish (Boots). After drying, the varnish was peeled off and mounted onto a glass slide for imaging.

### Measurement of Kranz Anatomy Traits

Slides of all leaf sections were imaged with an Olympus BX41 light microscope with a mounted Micropublisher 3.3 RTV camera (Q Imaging). Images were captured with Q-Capture Pro 7 software, and measurements were analyzed with the software ImageJ (Schneider et al., 2012). To maximize comparability, strict criteria were applied for all image analyses. Microscopy of transverse leaf sections was used to quantify the BS in terms of average BS tissue area (the total cross-sectional area of all BS cells immediately surrounding a vein) and BS cell size (the average cross-sectional area of individual BS cells around the vein). To quantify BS tissue area, the freehand selection tool was used to subtract the integrated area of each vein from the integrated area of all BS cells in direct contact with the vein on images with 200× total magnification. This value was divided by the number of BS cells in each vein bundle to



obtain the average BS cell size. For intervein distance (the distance between the centers of adjacent veins in transverse sections), only vein bundles were measured for which the following criteria did not apply: wide (indicates branching) or extremely large veins, veins with distorted BS cells due to contact with adjacent BS tissues (indicates merging), and veins with damaged BS cells. The line selection tool was used to measure the linear distance between the centers of adjacent veins on images with 40× total magnification. Vein density (vein length per unit area of leaf) was quantified on paradermal sections on images with 100× total magnification. Slides were imaged with the same microscopy equipment as transverse sections but set to Ph3 (phase contrast). Three images (from three different leaves per plant) were randomly selected for measurement. The freehand line tool was used to trace all veins (both major and minor) along their center. As it was not possible to trace all veins in an image simultaneously, individual vein sections were measured progressively without overlap and the individual lengths summed. The total vein length was divided by the image area to obtain the density. Stomatal density (the number of stomata per unit area of leaf) was measured on three subsampled leaves from three random plants on images with 200× total magnification. The total number of stomata was divided by the image area to obtain the density.

## Photosynthetic Performance

An LI-6800 portable photosynthesis infrared gas analyzer system (LI-COR) equipped with a multiphase flash fluorimeter was used to assess physiological differences for photosynthetic parameters between *G. gynandra* accessions. All physiological measurements were performed on the central leaflet of three independent 5-week-old plants, grown in separate pots. For stomatal conductance, transpiration ( $E$ ), and  $A_{400}$  measurements were taken during ambient conditions of growth (400  $\mu\text{L L}^{-1} \text{C}_a$  and PPFD of 350  $\mu\text{mol m}^{-2} \text{s}^{-1}$ ).  $WUE$  was defined as  $A_{400}/E$ . A combination chlorophyll fluorescence and  $A/C_i$  curve was measured for three plants from each accession.  $C_a$  reference values were 400, 400, 300, 200, 100, 50, 25, 400, 400, 400, 600, 800, 1,000, 1,200, and 400  $\mu\text{L L}^{-1}$ , with a saturating rectangular pulse of 12,000  $\mu\text{mol m}^{-2} \text{s}^{-1}$  at each reference point. Otherwise, measurements were made at a PPFD of 2,000  $\mu\text{mol m}^{-2} \text{s}^{-1}$ , 23°C, and 60% RH at each reference point. All leaves covered the full area of the cuvette. Measurements were carried out on consecutive days between 1 and 8 h postdawn, measuring one plant selected at random from each accession per day. Rates of gas exchange including  $A_{\text{max}}$  at 1,200  $\mu\text{L L}^{-1} \text{C}_a$  represent the measurements made, but a line of best fit also was plotted to fit the  $A/C_i$  response data to a logarithmic trend using the generalized additive model feature of the tidyverse package in R.  $I$  was calculated using linear regression of  $A$  and  $C_i$  measurements made between  $C_a$  of 25 to 200  $\mu\text{L L}^{-1}$  (Sharkey, 1988). Adjusted  $R^2$  values for all regression lines ranged between 0.993 and 0.997. Carboxylation efficiency was calculated as the partial derivative  $\frac{\partial A}{\partial C_i}$  at  $A = 0$ .  $\delta^{13}\text{C}$  analysis was performed according to methods described previously (Royles et al., 2016) on three biological replicates per accession with 500  $\mu\text{g}$  of dried leaf tissue.

## Statistical Analysis

For all tests, individual plants were considered experimental units in a complete randomized design. Data were analyzed in SAS (University Version; SAS Institute) and in R (version 3.4.2; R Studio). The R packages agricolae, car, cowplot, dplyr, ggplot2, multcompView, and tidyverse were used for data analysis and downloaded via the Install Packages Tool in RStudio (version 1.0.143; R Studio). A one-way ANOVA ( $\alpha = 0.05$ ) with type III sums of squares, which accounts for unequal replication, compared all means from anatomical and physiological measurements among *G. gynandra* accessions, and Student's  $t$  test ( $\alpha = 0.05$ ) was used to compare means of accessions by phylogenetic proximity, using the PROC GLM and PROC TTEST functions in SAS, respectively. Null hypotheses were rejected for specific ANOVAs or Student's  $t$  tests for any population with  $P \leq 0.05$ . Levene's test was used to evaluate homoscedasticity and was centered at the median (Levene, 1960) using the car package in R. All analyses showed equal variance ( $\alpha = 0.05$ ), except for the number of BS cells (Supplemental Table S3). Duncan's multiple range posthoc test was used for mean separation on accessions ( $\alpha = 0.05$ ) with statistically significant ANOVAs in the R package agricolae (Duncan, 1955). To provide conservative lower type I experiment error rates, Tukey's studentized range test (HSD) also was used ( $\alpha = 0.05$ ) with statistically significant ANOVAs in SAS and was plotted in R with the packages dplyr and multcompView (Supplemental Fig. S7). Pearson

product-moment correlation coefficients were calculated to find associations among features of Kranz traits (Pearson, 1895) using the PROC CORR function in SAS. Linear regression was performed using the PROC REG function in SAS to calculate  $I$  and carboxylation efficiency. The total number of replicates used in each experiment is summarized in Supplemental Table S4.

## Analysis of Transcript Abundance

Leaf tissue samples for RNA extraction were harvested simultaneously with samples for Kranz trait measurements after 3 weeks of growth using the opposite paired trifoliate leaf to that harvested for sectioning and imaging. These fresh leaf samples were immediately flash frozen with liquid nitrogen and stored at  $-80^\circ\text{C}$ . Total RNA was extracted from three tissue samples per accession with an RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. An on-column DNase digestion protocol was applied to remove genomic DNA contamination (Qiagen) before cDNA was synthesized with Invitrogen SuperScript II reverse transcriptase enzyme according to the manufacturer's instructions (Thermo Fisher Scientific). All cDNA samples were stored at  $-20^\circ\text{C}$  before RT-qPCR. Primers were designed for quantitative PCR of the  $C_4$  cycle genes *PEPC*, *NAD-ME*, *RbcS*, and *PPDK* (Supplemental Table S5), and reactions were carried out as reported previously (Burgess et al., 2016) on three biological and three technical replicates.

## Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession numbers MH188307 to MH188315.

## Supplemental Data

The following supplemental materials are available.

**Supplemental Figure S1.** Representative images for macroscopic and microscopic variation in leaf anatomy across a panel of *G. gynandra* accessions.

**Supplemental Figure S2.** Nonvariable features of Kranz anatomy among accessions.

**Supplemental Figure S3.** Natural variation in features of Kranz anatomy between *G. gynandra* accessions from Asia-Benin and those from East Africa.

**Supplemental Figure S4.** Physiological variation for photosynthetic gas-exchange parameters between *G. gynandra* accessions from Asia-Benin and those from East Africa.

**Supplemental Figure S5.** Transcript abundance differences for key enzymes in the  $C_4$  cycle between *G. gynandra* phylogenetic clusters.

**Supplemental Figure S6.** Leaf appearance of F1 hybrids between *G. gynandra* lines that exhibited the greatest differences in Kranz features.

**Supplemental Figure S7.** Quantitative natural variation among diverse *G. gynandra* accessions controlling for lower type I error rates.

**Supplemental Table S1.** Accessions of *G. gynandra* investigated and their source regions.

**Supplemental Table S2.** Pearson product-moment correlation coefficients for Kranz anatomy traits.

**Supplemental Table S3.** Summary statistics for Levene's test for homogeneity of variance for individual ANOVAs ( $\alpha = 0.05$ ).

**Supplemental Table S4.** The number of replicates for anatomical measurements.

**Supplemental Table S5.** List of primers for RT-qPCR analyses.

## ACKNOWLEDGMENTS

We thank Frank Becker and Beatrice Landoni (Wageningen University) for providing ITS sequence data and initial information on the selected lines, Blandine Gilbert for help with image processing, and James Rolfe for carbon isotope analysis.

Received February 9, 2018; accepted March 30, 2018; published April 20, 2018.

## LITERATURE CITED

- Akyildiz M, Gowik U, Engelmann S, Koczor M, Streubel M, Westhoff P (2007) Evolution and function of a cis-regulatory module for mesophyll-specific gene expression in the C4 dicot *Flaveria trinervia*. *Plant Cell* 19: 3391–3402
- Aubry S, Kelly S, Kümpers BMC, Smith-Unna RD, Hibberd JM (2014) Deep evolutionary comparison of gene expression identifies parallel recruitment of *trans*-factors in two independent origins of C4 photosynthesis. *PLoS Genet* 10: e1004365
- Bowes G, Ogren WL, Hageman RH (1971) Phosphoglycolate production catalyzed by ribulose diphosphate carboxylase. *Biochem Biophys Res Commun* 45: 716–722
- Bräutigam A, Gowik U (2016) Photorespiration connects C3 and C4 photosynthesis. *J Exp Bot* 67: 2953–2962
- Bräutigam A, Schliesky S, Kūlahoglu C, Osborne CP, Weber APM (2014) Towards an integrative model of C4 photosynthetic subtypes: insights from comparative transcriptome analysis of NAD-ME, NADP-ME, and PEP-CK C4 species. *J Exp Bot* 65: 3579–3593
- Brown HR, Bouton JH (1993) Physiology and genetics of interspecific hybrids between photosynthetic types. *Annu Rev Plant Physiol Plant Mol Biol* 44: 435–456
- 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
- Brown NJ, Palmer BG, Stanley S, Hajaji H, Janacek SH, Astley HM, Parsley K, Kajala K, Quick WP, Trenkamp S, (2010) C4 acid decarboxylases required for C4 photosynthesis are active in the mid-vein of the C3 species *Arabidopsis thaliana*, and are important in sugar and amino acid metabolism. *Plant J* 61: 122–133
- Brown NJ, Parsley K, Hibberd JM (2005) The future of C4 research: maize, *Flaveria* or *Cleome*? *Trends Plant Sci* 10: 215–221
- Burgess SJ, Granero-Moya I, Grangé-Guermente MJ, Bournsnel C, Terry MJ, Hibberd JM (2016) Ancestral light and chloroplast regulation form the foundations for C4 gene expression. *Nat Plants* 2: 16161
- Duncan DB (1955) Multiple range and multiple F tests. *Biometrics* 11: 1–42
- Eastmond PJ, Astley HM, Parsley K, Aubry S, Williams BP, Menard GN, Craddock CP, Nunes-Nesi A, Fernie AR, Hibberd JM (2015) *Arabidopsis* uses two gluconeogenic gateways for organic acids to fuel seedling establishment. *Nat Commun* 6: 6659
- Feodorova TA, Voznesenskaya EV, Edwards GE, Roalson EH (2010) Biogeographic patterns of diversification and the origins of C4 in *Cleome* (Clemnaceae). *Syst Bot* 35: 811–826
- Garner DMG, Mure CM, Yerramsetty P, Berry JO (2001) Kranz anatomy and the C4 pathway. *eLS*
- 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, Covshoff S (2010) The regulation of gene expression required for C4 photosynthesis. *Annu Rev Plant Biol* 61: 181–207
- Hibberd JM, Quick WP (2002) Characteristics of C4 photosynthesis in stems and petioles of C3 flowering plants. *Nature* 415: 451–454
- Kajala K, Brown NJ, Williams BP, Borrill P, Taylor LE, Hibberd JM (2012) Multiple *Arabidopsis* genes primed for recruitment into C4 photosynthesis. *Plant J* 69: 47–56
- Langdale JA (2011) C4 cycles: past, present, and future research on C4 photosynthesis. *Plant Cell* 23: 3879–3892
- Lauterbach M, Schmidt H, Billakurthi K, Hankeln T, Westhoff P, Gowik U, Kadereit G (2017) *De novo* transcriptome assembly and comparison of C3, C3-C4, and C4 species of tribe Salsoleae (Chenopodiaceae). *Front Plant Sci* 8: 1939
- Levene H (1960) Robust tests for equality of variances. In O Ingram, H Harold, eds, *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford University Press, Stanford, CA, pp 278–292
- Long SP, Marshall-Colon A, Zhu XG (2015) Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* 161: 56–66
- Lundgren MR, Christin PA, Escobar EG, Ripley BS, Besnard G, Long CM, Hattersley PW, Ellis RP, Leegood RC, Osborne CP (2016) Evolutionary implications of C3-C4 intermediates in the grass *Alloteropsis semialata*. *Plant Cell Environ* 39: 1874–1885
- Lundgren MR, Osborne CP, Christin PA (2014) Deconstructing Kranz anatomy to understand C4 evolution. *J Exp Bot* 65: 3357–3369
- Mallmann J, Heckmann D, Bräutigam A, Lercher MJ, Weber APM (2014) The role of photorespiration during the evolution of C4 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 C3 to C4 photosynthesis. *Plant J* 51: 886–896
- Mauricio R (2001) Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. *Nat Rev Genet* 2: 370–381
- Miller M, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In 2010 Gateway Computing Environments Workshop. GCE, pp 1–8
- Oakley JC, Sultmanis S, Stinson CR, Sage TL, Sage RF (2014) Comparative studies of C3 and C4 Atriplex hybrids in the genomics era: physiological assessments. *J Exp Bot* 65: 3637–3647
- Ort DR, Merchant SS, Alric J, Barkan A, Blankenship RE, Bock R, Croce R, Hanson MR, Hibberd JM, Long SP, (2015) Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proc Natl Acad Sci USA* 112: 8529–8536
- Pearson K (1895) Notes on regression and inheritance in the case of two parents. *Proc R Soc Lond* 58: 240–242
- Reeves G, Grangé-Guermente MJ, Hibberd JM (2017) Regulatory gateways for cell-specific gene expression in C4 leaves with Kranz anatomy. *J Exp Bot* 68: 107–116
- Reyna-Llorens I, Burgess SJ, Reeves G, Singh P, Stevenson SR, Williams BP, Stanley S, Hibberd JM (2018) Ancient duons may underpin spatial patterning of gene expression in C4 leaves. *Proc Natl Acad Sci USA* 115: 1931–1936
- Royles J, Amesbury MJ, Roland TP, Jones GD, Convey P, Griffiths H, Hodgson DA, Charman DJ (2016) Moss stable isotopes (carbon-13, oxygen-18) and testate amoebae reflect environmental inputs and microclimate along a latitudinal gradient on the Antarctic Peninsula. *Oecologia* 181: 931–945
- Sage RF (2004) Tansley review. The evolution of C4 photosynthesis. *New Phytol* 161: 30
- Sage RF (2016) A portrait of the C4 photosynthetic family on the 50th anniversary of its discovery: species number, evolutionary lineages, and Hall of Fame. *J Exp Bot* 67: 4039–4056
- Sage RF, McKown AD (2006) Is C4 photosynthesis less phenotypically plastic than C3 photosynthesis? *J Exp Bot* 57: 303–317
- Schlüter U, Muschak M, Berger D, Altmann T (2003) Photosynthetic performance of an *Arabidopsis* mutant with elevated stomatal density (*sdd1-1*) under different light regimes. *J Exp Bot* 54: 867–874
- Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9: 671–675
- Sharkey TD (1988) Estimating the rate of photorespiration in leaves. *Physiol Plant* 73: 147–152
- Sogbohossou EOD, Achigan-Dako EG, Maundu P, Solberg S, Deguenon EMS, Mumm RH, Hale J, Van Deynze A, Schranz ME (2018) A roadmap for breeding orphan leafy vegetable species: a case study of *Gynandropsis gynandra* (Clemnaceae). *Hortic Res* 5: 2
- Ueno O, Sentoku N (2006) Comparison of leaf structure and photosynthetic characteristics of C3 and C4 *Alloteropsis semialata* subspecies. *Plant Cell Environ* 29: 257–268
- White T, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protoc* 18: 315–322



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. *Plant Cell* **28**: 454–465

Williams BP, Johnston IG, Covshoff S, Hibberd JM (2013) Phenotypic landscape inference reveals multiple evolutionary paths to C4 photosynthesis. *eLife* **2**: e00961