Investigating intraspecific physiological diversity to understand complex trait evolution

by

Marjorie Ruth Lundgren

Department of Animal and Plant Sciences

University of Sheffield

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

February 2015

ACKNOWLEDGEMENTS

I would like to extend sincere thanks to the countless mentors and supporters who have helped to make this dissertation a reality.

My supervisor, Colin Osborne, always struck a perfect balance in providing me with independence and guidance during my PhD. For this, as well as his seemingly endless patience, support, enthusiasm, and kindness, I am profoundly grateful. An ideal supervisor, Colin has proved worthy of my 5,212-mile move away from the paradise that is sunny California.

Pascal-Antoine Christin has been an invaluable mentor and friend over the last few years. His contributions toward this dissertation are far too numerous to mention and listing them would belittle their significance. I am very appreciative of all the time and resources that he has dedicated to my research. I look forward to many more years of *Alloteropsis*-centric enthusiasm together.

I would not have undertaken a postgraduate degree in Sheffield if it were not for Graham Hymus, my plant physiology rock star. He provided me with a solid eco-physiological foundation on which this dissertation is built. I thank him for his many years of mentorship, friendship, and importantly, his convincing recommendation to work with Colin.

An acknowledgment must also be made to Sonia Sultan, with whom this journey commenced a lifetime ago. As one of my first and most influential mentors, Sonia's academic legacy is evident throughout this dissertation.

In addition to those mentioned above, the research and discussions presented in this dissertation were improved through dialogues with countless people. Although I will inevitably omit far too many of these mentors, I would like to thank the following people for their support, enthusiasm, and academic contributions toward my research; Brad Ripley,

Roger Ellis, Paul Hattersley, Caroline Lehmann, Steve Long, Asaph Cousins, Lawren Sack, Maria Vorontsova, Chandra Bellasio, Christoph Lehmeier, Georg Freck, Guillaume Besnard, Ilia Leitch, Howard Griffiths, Susanne von Caemmerer, Louisa Liebenberg (Saayman), Lyn Fish (Snook), Richard Leegood, Karen Bailey, Ralf Kynast, Mary Namaganda, Jacob Washburn, Joe Hereford, Paula Rudall, Paul Quick, Steve Rolf, and Andrew Fleming. Although we have never met, the work presented in dissertation was facilitated by earlier research by Dough Ibrahim and Christine Long.

The work presented in this dissertation would not have been possible without the assistance of several herbaria. Thus, I am greatly indebted to Maria Vorontsova and Martin Xanthos at the Herbarium at the Royal Botanic Gardens at Kew, Lyn Fish, Catherine Mashau, William Sepheka, and Erich van Wyk of The National Herbarium in Pretoria, South Africa, John Elia of the National Herbarium of Tanzania, Terry Trinder-Smith of the Bolus Herbarium in Cape Town, South Africa, Itambo Malombe of the East African Herbarium in Nairobi, Kenya, Mary Namaganda of the Makerere University Herbarium in Kampala, Uganda, Anna Monro at the Australian National Herbarium in Canberra, Australia, and the National History Museum staff in London.

Sample material was donated by Guillaume Besnard, David Chatelet, Caroline Lehmann, Russell Hall, Hui Lui, Laszlo Csiba, and the Millennium Seed Bank.

I also thank Heather Walker for processing countless Mass Spectrometer samples for me, Brad Ripley for our South African *Alloteropsis* adventure, Thomas Kluvier for R, Python, modeling, and computer help in general, Jen Sloan, Carla Turner, Serina Akhtar, Liam Rasch, and Tom Gregory for assistance with histology methods, Ralf Kynast and Ilia Leitch for assistance with cytology methods and hosting me at the Jodrell Lab, and Adélie Mollimard for experimental assistance. I am also very appreciative of John

Beresford and the APS office staff for their assistance over these past few years.

My sanity was vaguely maintained throughout the PhD with the support of many dear friends. In particular, I would like to express gratitude to Emanuela Samaritani, Jenny Armstrong, Samiya Selim, Isabel Winney, Liz Carabine, Russell Hall, and Carla Turner for their friendship during this trying time.

Finally, I need to voice profound appreciation to Neal Walsh. His patient acceptance of my work addiction and borderline agoraphobia is commendable. I look forward to post-dissertation life together.

This research was funded by The University of Sheffield's Prize Scholarship.

ABSTRACT

The overall aim of this thesis is to identify the small-scale processes that lead to the emergence of novel traits of impressive complexity. Complex traits are composed of multiple individual traits that must all be in place to function as an intricate unit. The evolutionary path from ancestral state to derived complex trait consequently must pass through a series of intermediate stages, and the evolutionary paths and adaptive significance of these intermediary states remain unresolved. The complex trait of C_4 photosynthesis requires anatomical and biochemical modifications to the ancestral C_3 pathway, creating C_3-C_4 intermediate phenotypes. Here, the evolutionary transitions between C_3 and C_4 phenotypes are investigated using the grass *Alloteropsis semialata*, one of the only species known to employ multiple photosynthetic types. An investigation of anatomical and physiological phenotypes of geographically and ecologically diverse *A. semialata* populations discovered the presence of C_3-C_4 intermediates in this species for the first time. The intraspecific history of dispersal across geographic and environmental space was inferred in this species complex to show that non-C_4 *A. semialata* remained geographically and ecologically limited, while in the same length of time, C_4 individuals dispersed across three continents and into an expanded range of environments, encompassing the ancestral one. This first intraspecific investigation of C_4 evolutionary ecology shows that, in otherwise similar plants, C_4 photosynthesis does not directly shift the ecological niche, but broadens it, allowing dispersal into more diverse conditions and over longer distances. Moreover, it is hypothesized that a C_3-C_4 intermediate ancestor gave rise to C_4 photosynthesis in *A. semialata*, while another lineage reverted back to C_3 photosynthesis from this intermediate state. Thus, the small-scale anatomical shifts that accumulated between C_3-C_4 intermediate and C_3 and C_4 accessions are quantified and linked to environmental variability to understand what promoted the emergence of C_4 photosynthesis in this species.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF BOXES

CHAPTER 1

ON THE EVOLUTION OF COMPLEX TRAITS; HOW THE GRASS *ALLOTEROPSIS SEMIALATA* CAN ADVANCE OUR UNDERSTANDING OF PHOTOSYNTHETIC PATHWAY EVOLUTION

Complex trait evolution

Understanding how novel traits form is fundamental to our appreciation of biological diversity, and has been puzzling evolutionary biologists since the early days of the field (Darwin, 1859). Indeed, the processes that underpin the evolution of simple traits, which have direct effects on the phenotype and fitness, are relatively well known. First, random mutations alter the trait quantitatively or qualitatively, and variants that decrease survival or reproductive success are less likely to be passed to subsequent generations, while variants that increase these successes are preferentially transmitted. This process continues over generations, causing new phenotypes to emerge through the gradual accumulation of mutations, each of which confers a slight advantage.

The evolution of complex traits, however, presents a more complicated problem, as they are composed of multiple individual traits that must all be in place to function as a complex unit, such as camera eyes or the ability to fly (Gatesy and Dial, 1996; Fernald, 2006). The evolutionary path from the ancestral state to the derived complex trait consequently must pass through a series of intermediate stages, where only some of the individual traits are present. Each intermediate stage must still be viable and provide an advantage, or at least no disadvantage, over the previous stage to be favoured by natural selection. This paradox was partially solved by showing that intermediate stages can be advantageous, as is the case with photoreceptors, which are precursors in the evolution of camera eyes (Lamb *et al*., 2007). Moreover, individual components of complex traits can be selected for in the absence of others if they have a function in themselves (Wistow and Piatigorsky, 1988; Gatesy and Dial, 1996). These can be diverted from their original function and become part of a complex trait, in a process named 'co-option' (Wistow and Piatigorsky, 1988; True and Carroll, 2002). Feathers, for example, evolved in flightless dinosaurs for purposes linked to thermoregulation and not flight (Xu *et al*., 2012). Despite these few examples, the evolutionary paths to complex traits and the adaptive significance of intermediary states remain unresolved in the vast majority of cases (Monteiro and Podlaha, 2009). This dissertation addresses the evolution of complex traits using the example of photosynthetic pathways.

Photosynthesis is a complex trait

Photosynthesis is the process that most plants use to convert light energy into chemical energy. With carbon dioxide (CO_2), water, and photon inputs, photosynthesis yields carbohydrates and oxygen (O_2). This process occurs in two stages. First, the light reactions harvest photons to store energy in the form of adenosine triphosphate (ATP) and nicotinamide adenine dinucleotide phosphate (NADPH). This step takes place along the thylakoid membrane within chloroplasts. Both ATP and NADPH are needed to fuel the second step, the light-independent reactions, which occur in the liquid stroma of chloroplasts.

The C_3 photosynthetic pathway

The light-independent reactions are known as the Calvin-Benson Cycle (referred to as the Calvin cycle from here onward). In the first step of this cycle, ribulose-1,5 biphosphate carboxylase/oxygenase (Rubisco) catalyses the addition of CO_2 to ribulose-1,5-bisphosphate (RuBP), a five-carbon sugar (Figure 1.1). This produces a six-carbon molecule that immediately splits into two molecules of 3-phosphoglycerate (PGA), the three-carbon organic acid that gives C_3 photosynthesis its name. Next, ATP and NADPH are used to reduce the PGA into phosphoglyceraldehyde (PGAL; Figure 1.1). For every three molecules of CO_2 entering the Calvin cycle, six molecules of PGAL are generated. Of these, one PGAL molecule is released then converted to sucrose or starch - the primary products of photosynthesis - while the remaining five PGAL molecules are used to regenerate RuBP with the addition of ATP. This entire process occurs within the chloroplast stroma of a single cellular compartment, which is usually within the mesophyll in C_3 plants. Because of this, the mesophyll cells of C_3 plants contain high concentrations of chloroplasts in which these light-dependent and independent reactions occur (Stata *et al*., 2014).

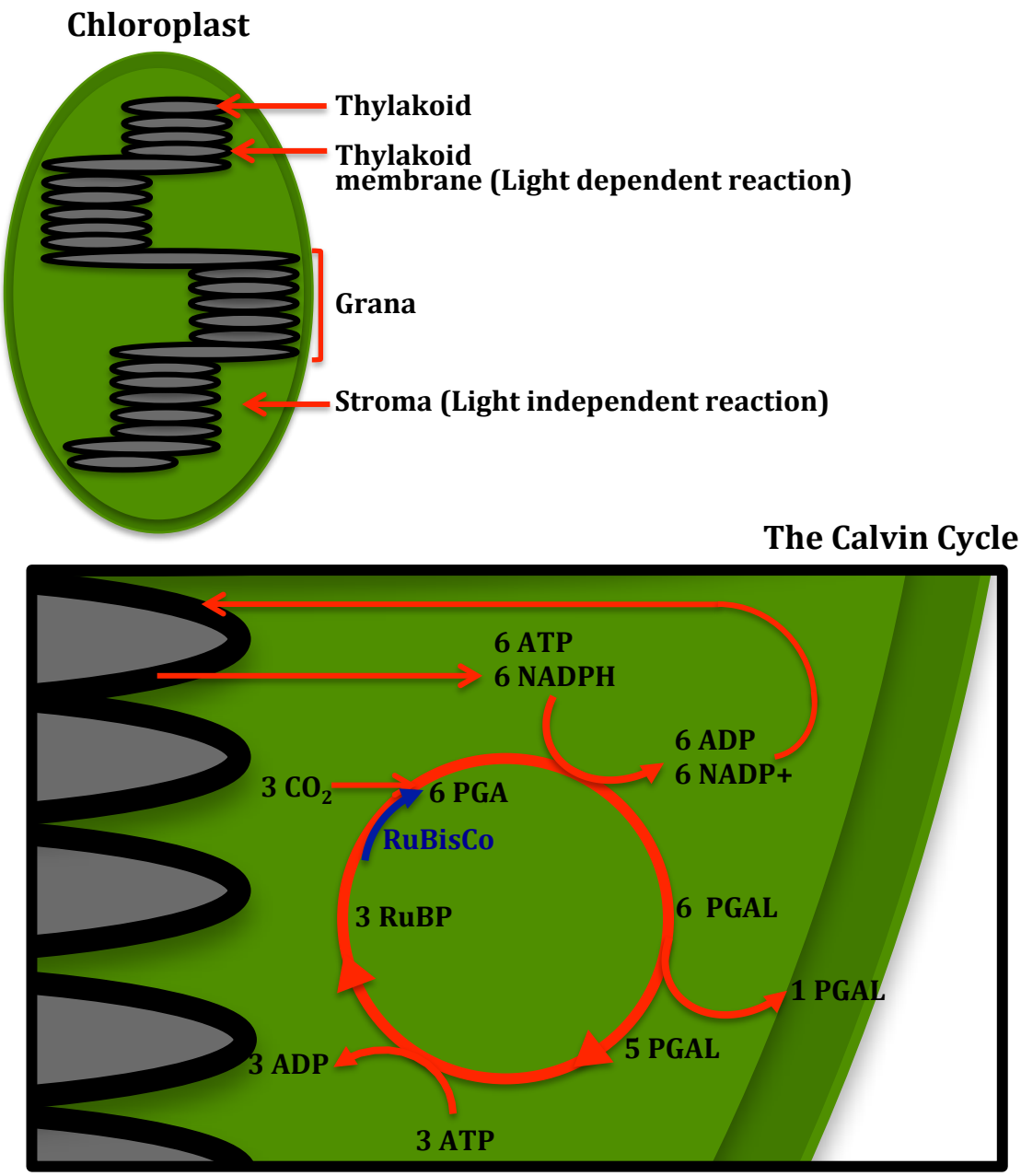


Figure 1.1 The Calvin cycle. At the top, a chloroplast depicts the stroma, or the fluid within the chloroplast where the light independent reaction occurs, and grana, which are stacks of thylakoids. The light independent reaction occurs within the membrane of these thylakoids, contributing energy, in the form of ATP and NADPH, to the light independent reaction. In the Calvin cycle (bottom) Rubisco catalyzes the addition of CO_2 to RuBP to form phosphoglycerate (PGA). Energy from ATP and NADPH reduce PGA into phosphogyceraldehyde (PGAL). Some of the PGAL is release from the cycle to be converted into sucrose or starch, while the rest is used to regenerate RuBP, with energy from ATP.

Inefficiencies of Rubisco

Rubisco can catalyse the addition of both CO_2 and O_2 to RuBP. This confusion arose because O_2 and CO_2 are similarly featureless molecules, and the Rubisco molecule evolved in an anaerobic atmosphere where this property was not detrimental and consequently not counter-selected (Tcherkez *et al.,* 2006). When O_2 is fixed instead of CO_2, the toxic molecule 2-phosphoglycolate (PG) is produced which needs to be recycled via the photorespiratory cycle. This cycle regenerates PG to PGAL, but in the process, it consumes ATP and NADPH, liberates CO_2, and generates ammonia (NH_3) that must be detoxified at additional metabolic cost (Figure 1.2; Sharkey, 1985; Bauwe *et al*., 2010; Fernie *et al*., 2013). Thus, the fixation of O_2 is a competitive inhibitor to CO_2 fixation that reduces photosynthetic efficiency (Kanai and Edwards, 1999; Bauwe *et al*., 2010; Fernie *et al*., 2013). In fact, photorespiration can reduce yield by more than 30 percent in plants using the C_3 pathway (Gerbaud and Andre, 1987; Skillman, 2008; Bauwe *et al*., 2010). The strong detrimental effects of O_2 fixation have sparked photosynthetic novelty, as some lineages have evolved elegant mechanisms to evade these costs.

Photosynthetic diversification

The light-independent reactions are similar across modern plants, however, the source of the CO_2 that fuels these reactions differs, as the ancestral C_3 photosynthetic pathway diversified to create two other primary systems; crassulacean acid metabolism, or CAM photosynthesis, and C_4 photosynthesis. The CAM photosynthetic pathway temporally separates the initial uptake of CO_2 at night from its assimilation during the photoperiod (Osmond, 1978). This temporal separation functions to increase photosynthetic efficiency and minimize water loss, making CAM photosynthesis advantageous in arid and aquatic environments (Nobel, 1996; Drennan and Nobel, 2000). The C_4 pathway also separates CO_2 uptake from the rest of the photosynthetic process, but it does so spatially rather than temporally. In C_4 plants, CO_2 is initially assimilated in a compartment that is in contact with the atmosphere then biochemically transported into a second, isolated compartment where it is reduced (Hatch, 1987). Thus, the C_4 pathway is a complex trait that requires coordination of additional anatomical and biochemical components over the C_3 system. This dissertation focuses on the evolutionary

transitions between C_3 and C_4 photosynthetic types to exemplify complex trait evolution.

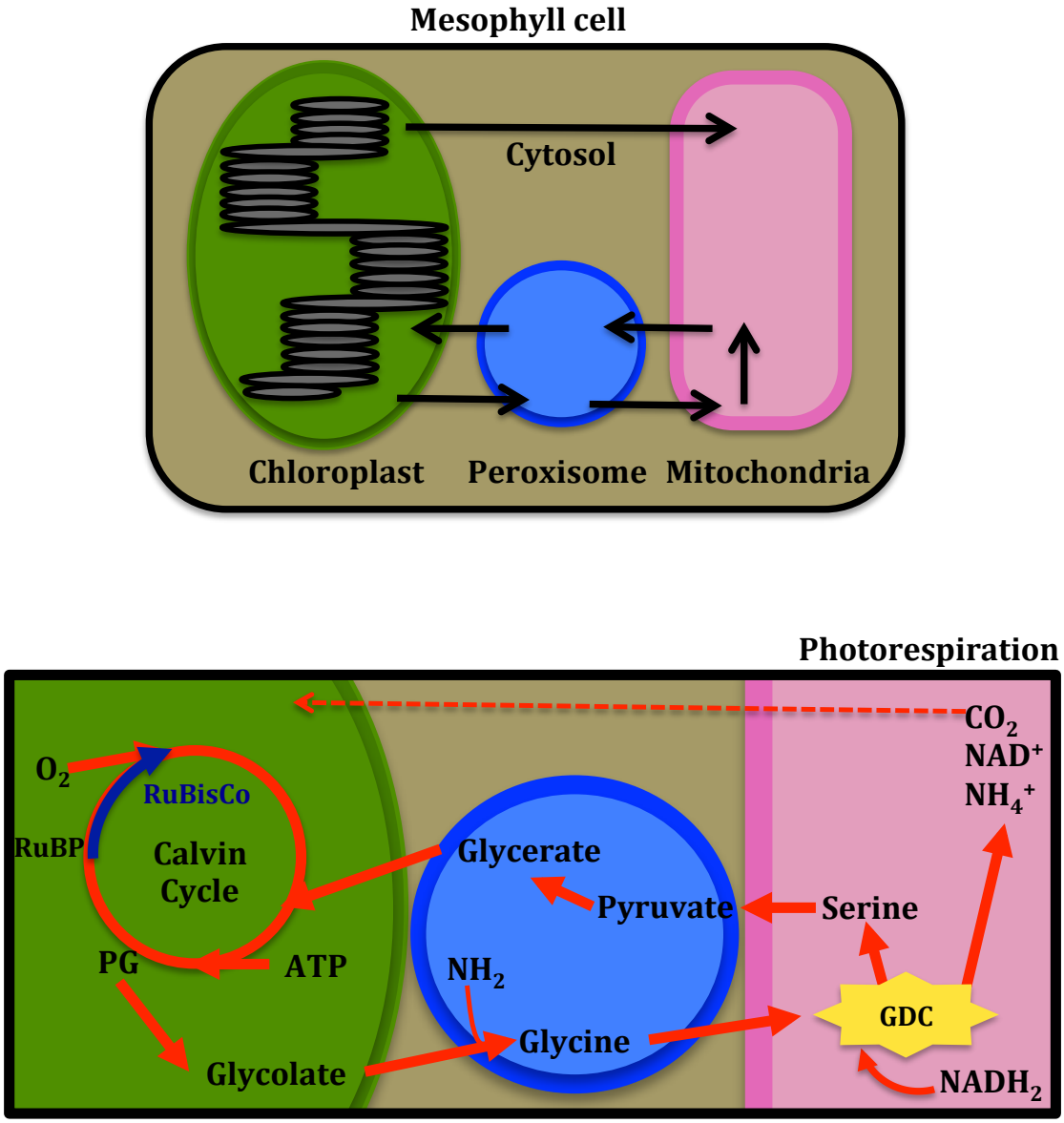


Figure 1.2 The photorespiratory cycle. This process occurs across chloroplasts, peroxisomes, cytosol, and mitochondria. For most plants, this process occurs entirely within the mesophyll. Rubisco catalyzes the addition of O_2 to RuBP to form 2-phosphoglycolate (PG), which combines with NADH to form glycine in the peroxisome. This glycine diffuses into the mitochondria where it is decarboxylated by glycine decarboxylase (GDC) to release CO_2 and produce NH_3 and serine.

The C_4 photosynthetic pathway

The C_4 photosynthetic pathway involves a coordinated system of anatomical and biochemical traits that function to concentrate CO_2 around the active site of Rubisco and substantially reduce O_2 fixation (Chollet and Ogren, 1975). In this pathway, atmospheric CO_2 is initially converted into bicarbonate (HCO_3^-) by the enzyme carbonic anhydrase (CA) within the cytosol of an initial compartment that is in contact with the atmosphere (Figure 1.3; Hatch and Burnell, 1990). The enzyme phosphoenolpyurvate carboxylase (PEPC) catalyses the fixation of this HCO_3^- to phosphoenolpyruvate (PEP) to produce the four-carbon organic acid, oxaloacetate (OAA), which gives C_4 photosynthesis its name. The enzyme malate dehydrogenase (MDH) converts OAA into malate, another four-carbon organic acid. These acids move through plasmodesmata channels to be decarboxylated in a second compartment that is isolated from the atmosphere (Hatch, 1987). The released CO_2 can then enter the Calvin cycle within this second compartment (Brown, 1975). Because this pathway only biochemically shuttles CO_2, and not O_2, into this isolated compartment, CO_2 accumulates in much higher concentrations than O_2 there (von Caemmerer and Furbank, 2003). Indeed, the ratio of CO_2 to O_2 at the active site of Rubisco can be over 25 times greater in C_4 than C_3 plants (Hatch, 1987). Therefore, this CO_2 concentrating mechanism maintains CO_2 at nearly saturating levels around the active site of Rubisco, which nearly eliminates the fixation of O_2 and subsequent photorespiration in plants using the C_4 photosynthetic pathway (Chollet and Ogren, 1975; Hatch and Osmond, 1976).

Functional elements of the C_4 pathway have been implemented in a diversity of ways. One of the most variable components of the C_4 cycle is the decarboxylating enzyme. Some plants use NADP-malic enzyme (NADP-ME), some use NAD-malic enzyme (NAD-ME), and others use phosphoenolpyruvate carboxykinase (PCK) in combination with NADP-ME or NAD-ME (Bräutigam *et al*., 2014). Taxa using these decarboxylating systems do not form discrete groups, but instead form continua across these different enzymes (Bräutigam *et al*., 2014). Furthermore, the primary decarboxylating enzyme may not be rigid within an individual, but plastically determined by its environment and development (Furbank, 2011).

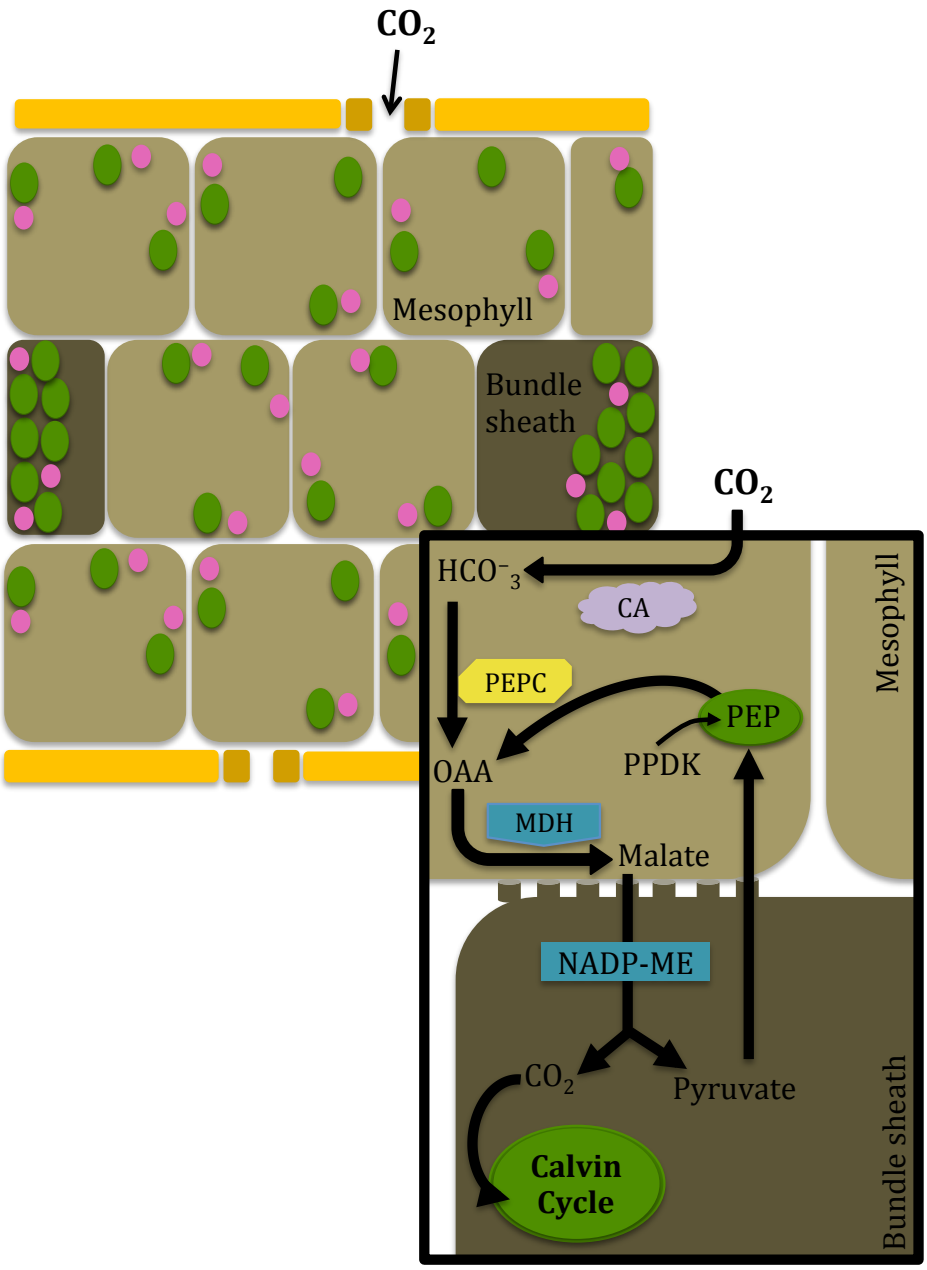


Figure 1.3 The C_4 pathway. The top panel shows arrangement of the two compartments (mesophyll (tan) and bundle sheath (brown) cell types, here) and the localization of chloroplasts (green) and mitochondria (pink) there. Carbonic anhydrase (CA) rapidly converts atmospheric CO_2 and H_2O into bicarbonate (HCO^-_3). Phosphoenolpyurvate carboxylase (PEPC) catalyzes the addition of HCO^-_3 to PEP, forming oxaloacetate (OAA). Malate dehydrogenase (MDH) converts OAA to malate, which diffuses along a concentration gradient into a secondary compartment via plasmodesmata channels. The malate is decarboxylated (by NADP-ME in this example). CO_2 accumulates in high concentrations in this compartment where it enters the Calvin cycle. The released pyruvate returns to the initial compartment where PEP is regenerated in chloroplasts by pyruvate phosphate dikinase (PPDK) and associated enzymes.

Evolution of C_4 photosynthesis

How the C_4 pathway evolves

The evolution of C_4 photosynthesis generally requires genetic and anatomical enablers, as well as anatomical and biochemical modifications over the typical C_3 phenotype. The transition from typical C_3 to C_4 phenotypes is thought to be gradual and smooth, taking shape over time as many small adaptations accumulate (Heckmann *et al*., 2013). Thus, the evolution of C_4 photosynthesis involves numerous intermediary steps, each of which must yield a fitness advantage, or at least no disadvantage, along the way. These enablers, and anatomical and biochemical changes are described below and depicted in Figure 1.4.

1. Enabling phenotypes

Resent work has shown that certain anatomical and genetic phenotypes facilitate the evolution of C_4 photosynthesis (Figure 1.4). First, accessibility of C_4 photosynthesis is dependant on the genomic context of the ancestors as certain gene lineages were repeatedly co-opted for C_4 function in both eudicots and monocots (Christin *et al*., 2013b; John *et al.,* 2014; Christin *et al.,* 2015). Moreover, C_4 photosynthesis likely evolved from C_3 ancestors with higher vein density and larger bundle sheaths than C_3 lineages that did not evolve C_4 photosynthesis (Christin *et al*., 2011a; 2013a; Griffiths *et al*., 2013; reviewed in Sage *et al*., 2014). These phenotypes reduce the distance between mesophyll and bundle sheath cells and increase the relative proportion of bundle sheath to mesophyll area, which also facilitates an increase in the proportion of bundle sheath to mesophyll Rubisco. These modifications move the anatomical phenotype closer to the C_4 requirement (as will be discussed in Chapter 2).

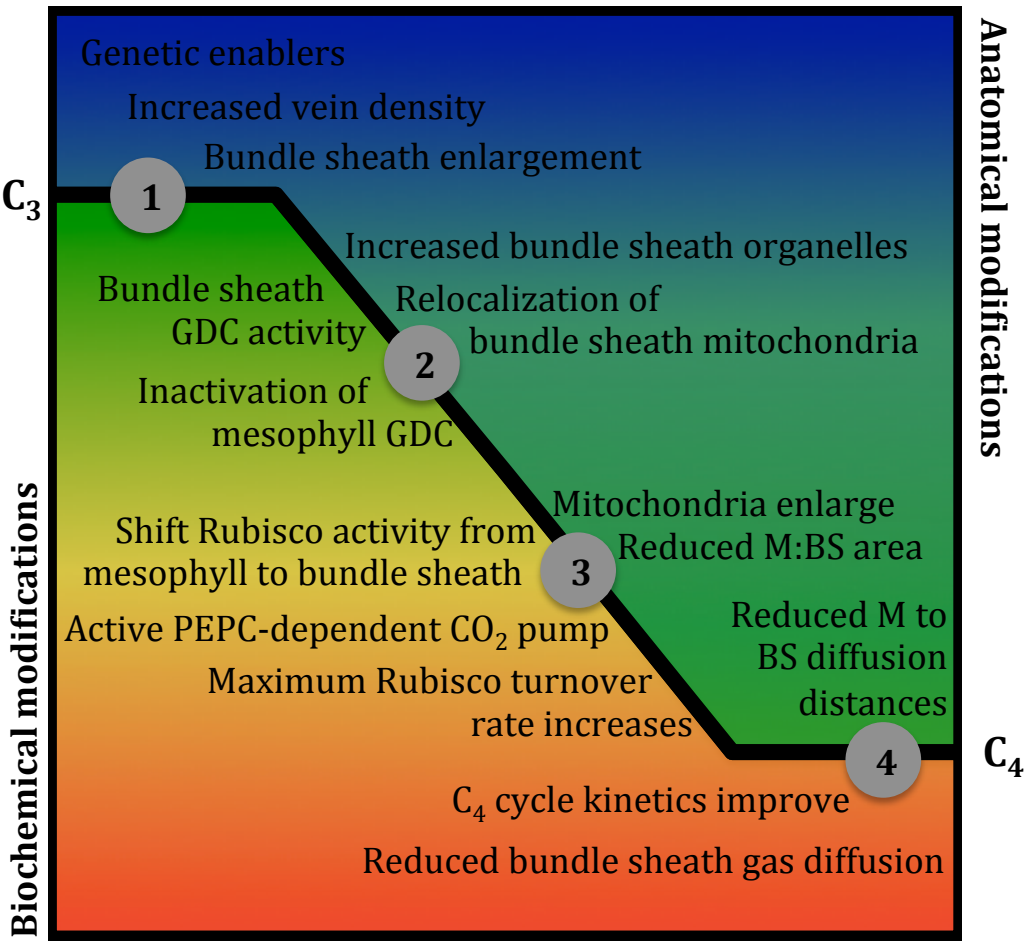


Figure 1.4 Biochemical (left) and anatomical (right) modifications that occur during the transitions between C_3 and C_4 photosynthetic types. The major steps, numbered along this transition, are enabling phenotypes (1), establishment of a C_2 cycle (2), establishment of a C_4 cycle (3), and optimization of a C_4 cycle (4). The minor modifications that facilitate each major step are listed along the sides of the transition landscape. This diagram is primarily based on patterns presented in Sage *et al.*, 2012 and Heckmann *et al.*, 2014. Abbreviations *M* for mesophyll and *BS* for bundle sheath are used.

2. Establishment of the C_2 cycle

The C_2 cycle, also called the glycine shuttle or photorespiratory CO_2 pump (referred to as the C_2 cycle from here onward), rescues and reuses photorespired CO_2. The recycling of photorespiratory CO_2 is not unique to C_2 plants though, as C_3 plants are able to recapture about 50 percent of photorespired CO_2 (Bauwe *et al.*, 1987; Haupt-Herting *et al.*, 2001). However, the C_2 cycle is more efficient at it, as it can reuse over 85 percent of photorespired CO_2 (Bauwe *et al.*, 1987). Under certain

conditions, plants using the C_2 cycle are physiologically more efficient than those that only use a C_3 cycle because of photorespiratory compartmentalization and not photosynthetic metabolism. Indeed, the photosynthetic metabolism of these C_2 plants is similar to C_3 plants.

The C_2 cycle begins with the initial steps of the photorespiration cycle. However, activity of the CO_2-liberating enzyme GDC is shifted away from mesophyll mitochondria and toward the bundle sheath mitochondria (Hylton *et al*., 1988). The glycine that is produced in the photorespiratory pathway must then diffuse into the bundle sheath to be decarboxylated (Figure 1.5). The released CO_2 can then be assimilated within the bundle sheath chloroplasts. The C_2 cycle, therefore, rescues and reuses photorespired CO_2 to create a weak CO_2 concentrating mechanism that elevates CO_2 around bundle sheath Rubisco. A recent study by Keerberg *et al*. (2014) found that the CO_2 concentration at Rubisco active sites was over three times higher in the C_2 *Flaveria pubescens* and nearly six times higher in the C_4 *F. trinervia* than the closely related C_3 species *F. cronquistii*. Thus, the C_2 cycle can concentrate CO_2 around Rubisco at about half the efficiency of C_4 plants. Altogether, the C_2 cycle establishes the bundle sheath Calvin cycle, a weak CO_2 concentrating mechanism, and the two compartment coordination that are characteristic of C_4 photosynthesis (Sage *et al.,* 2013) , making it a clear intermediate step in C_4 evolution as the C_2 cycle shifts the phenotype several steps closer to the C_4 requirement, filling a gap in the selective landscape between C_3 and C_4 phenotypes.

To achieve the C_2 cycle, bundle sheath mitochondria must relocate to the periphery of these cells to efficiency capture diffused glycine, chloroplasts and mitochondria must become more abundant in the bundle sheath but less so in the mesophyll, and GDC must be inactivated in the mesophyll but activated in the bundle sheath (Hylton *et al*., 1988; Devi *et al*., 1995; Sage *et al*., 2013; Figure 1.4). The GDC complex is comprised of four protein subunits; L-, P-, H-, and T- proteins, of which only the P-protein is consistently lacking from mesophyll GDC in C_2 plants (reviewed in Oliver, 1994 and Sage *et al*., 2013). Thus, it seems that the function of the mesophyll GDC complex is blocked in C_2 plants simply by the paucity of this one subunit (Rawshorne, 1992; reviewed in Sage *et al*., 2013).

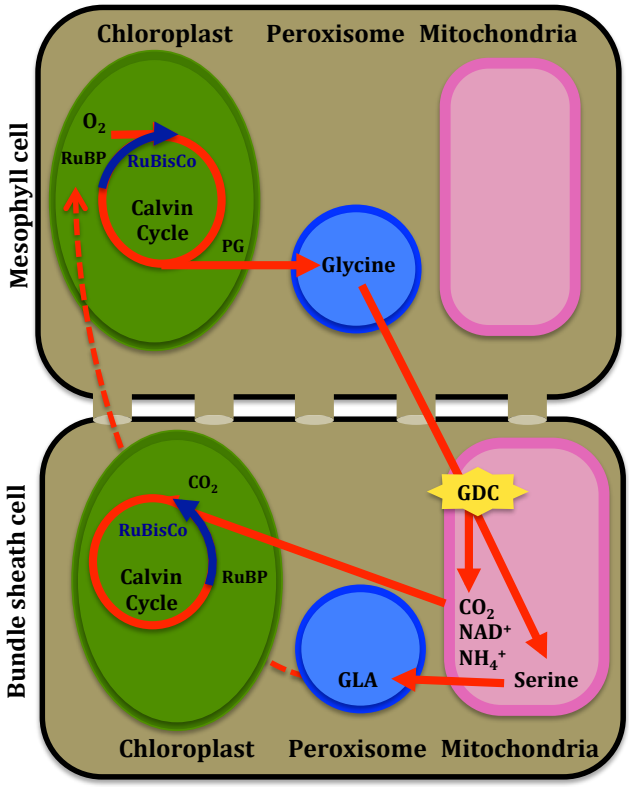


Figure 1.5 The C_2 cycle. When Rubisco catalyzes the addition of O_2 to RuBP in the mesophyll chloroplast Calvin cycle, the resulting phosphoglycolate (PG) is converted to glycine in the peroxysome. This glycine diffuses into the bundle sheath mitochondria, where it is decarboxylated by glycine decarboxylase (GDC) into serine, NAD+, NH_4+, and CO_2. This CO_2 can be refixed in the bundle sheath chloroplast Calvin cycle. Meanwhile, the release serine is converted back to glyceric acid (GLA) in the peroxisomes, which is then converted to RuBP in the mesophyll chloroplasts (reviewed in Sage *et al.*, 2013).

3. Establishment of a C_4 cycle

Once the C_2 cycle is in place, a C_4 cycle may establish in some lineages (Figure 1.4). Anatomically, bundle sheath mitochondria will enlarge and the ratio of mesophyll to bundle sheath tissue area will decrease as lineages transition to a more C_4-like phenotype (reviewed in Sage *et al.*, 2012). These transformations in tissue components are accompanied by shifts in Rubisco activity away from the mesophyll and into bundle sheath cells. This shift in Rubisco activity facilitates the activation of the PEPC-dependent CO_2 pump across these two cell types, which increases the maximum turnover rate of Rubisco (Heckmann *et al.*, 2014). These modifications constitute an active C_4 cycle.

4. Optimization of the C_4 cycle

Once a functional C_4 cycle is in place, additional modifications may occur to optimize the system (Figure 1.4). Anatomically, the diffusion distances between mesophyll and bundle sheath cells can reduce further to facilitate metabolite transfer (reviewed in Sage *et al*., 2012). Biochemically, C_4 cycle kinetics may improve and the maximum turnover rate of Rubisco will continue to increase (Heckmann *et al*., 2014). Bundle sheath gas diffusion may also reduce, which functions to minimize CO_2 leakage and maintain high concentrations of CO_2 in the bundle sheath, increasing the efficiency of the C_4 cycle (Hatch *et al.*, 1995).

Proliferation of the C_4 pathway

Despite its complexity, C_4 photosynthesis is a highly convergent trait that has evolved independently in nearly 70 plant lineages and is distributed across 19 angiosperm families and approximately 7500 species (Sage *et al.*, 1999a; Sage *et al.*, 2011; GPWGII, 2012). This pathway is used by both monocots and eudicots. C_4 monocots are nearly ten times more speciose than C_4 eudicots, yet there were more independent C_4 origins in eudicots (approximately 36 versus 26 origins in eudicots versus monocots; Sage *et al*., 2011), and only three monocot families (Cyperaceae, Hydrocharitaceae, and Poaceae) and as many as fifteen eudicot families have evolved this complex pathway (Sage *et al.,* 1999a).

Despite the intermediary role that the C_2 cycle is thought to play in C_4 evolution, lineages using a C_2 cycle do not necessarily transition to the C_4 pathway and may persist in the C_2 state in perpetuity. The most recent accounts have identified over 40 C_2 species, more broadly known as 'C_3-C_4 intermediate' plants, from around 13 plant families (Osborne *et al*., 2014; Sage *et al.*, 2014). Most of these identified C_2 taxa are eudicots, however, C_2 phenotypes have also been identified in all three monocot families that have evolved C_4 photosynthesis (reviewed in Williams *et al*., 2013 and Sage *et al.*, 2014). Considering this physiological state is an improvement upon the C_3 pathway under high photorespiratory conditions and the important role that it plays in C_3 / C_4 transitions, the true abundance of C_2 species might be much greater than these estimates.

Within the angiosperm phylogeny there are several clusters of C_4 origins, with some large groups completely lacking C_4 plants and others containing a plethora of closely related C_4 origins (Sage, 2001; Sage *et al*., 2011; GPWGII, 2012). In the grasses, for example, the large BEP clade (after the Bambusoideae, Ehrhartoideae, and Pooideae subfamilies that it encompasses) completely lacks C_4 species, while the similarly-sized PACMAD sister clade (after the Panicoideae, Arundinoideae, Chloridoideae, Micrairoideae, Aristidoideae, and Danthonioideae subfamilies) includes over 22 independent C_4 origins (Kellogg, 1999; GPWGII, 2012). This non-random phylogenetic distribution has been attributed to the presence of evolutionary enablers in some groups (Sage, 2001). For example, recent research on the grass family has shown that C_3 members of the PACMAD clade possess C_4-like anatomical characters that are absent from the BEP clade (Christin *et al*., 2013a; Griffiths *et al*., 2013). Thus, these precursors might have been regularly co-opted for C_4 function, facilitating the repeated emergence of C_4 photosynthesis in some lineages.

What drove C_4 evolution and expansion?

C_4 photosynthesis first arose in the grasses and eudicots over 30 million years ago (Fox and Koch, 2003; Kadereit *et al.,* 2003; Christin *et al*., 2008; 2011b; 2014a; Urban *et al*., 2010) when atmospheric CO_2 concentrations were low (Pagani *et al*. 1999; 2005; Beerling and Royer, 2011). Thus, it is believed that C_4 photosynthesis initially evolved in response to a drastic decline in atmospheric CO_2 concentrations. However, C_4 species richness and ecological dominance expanded during the late Miocene when aridity proliferated, rainfall seasonality increased, and forest systems transitioned toward open, higher light environments (Ehleringer *et al.,* 1991; Keeley and Rundel, 2005; Edwards and Still, 2008; Osborne and Freckleton, 2009; Edwards and Smith, 2010; Edwards *et al.,* 2010; Osborne and Sack, 2012; Bouchenak-Khelladi *et al*., 2014; Spriggs *et al*., 2014).

Ecological distribution of photosynthetic types

The current ecological distributions of C_3, C_2, and C_4 photosynthetic types are linked to their physiology. In C_4 plants, the CO_2 concentrating mechanism is costly,

as it requires more ATP than the C_3 system and suffers CO_2 loss through bundle sheath leakage (Pearcy and Ehleringer, 1984). These metabolic and structural costs are imposed in all environments, such that C_4 photosynthesis is only more efficient than the C_3 pathway under conditions promoting high rates of photorespiration, such as low atmospheric CO_2, warm, high light, arid, and saline environments (Ehleringer and Bjorkman, 1977; Osborne and Beerling, 2006; Osborne and Freckleton, 2009; Edwards and Smith, 2010; Bromham and Bennett, 2014).

Habitats with open canopies allow large radiation loads to reach leaves. These high light environments, as well as warm habitats in general, increase leaf temperature and thus, photorespiration because the ratio of CO_2 to O_2 in solution, and the affinity of Rubisco for CO_2 relative to O_2, decrease with leaf temperature (Ehleringer and Bjorkman, 1977; Ogren, 1984; Jordan and Ogren, 1984). Moreover, plants that experience dry or saline environments avoid desiccation by maintaining lower stomatal conductance to reduce water loss (*e.g.*, via fewer stomata or maintaining smaller stomatal apertures). This stomatal restriction also hinders CO_2 exchange, which decreases intercellular CO_2 concentrations and consequently lowers CO_2 to O_2 ratios, leading to more photorespiration (Sage *et al*., 2012).

The distributions of C_4 monocots and eudicots are differently influenced by these photorespiratory stresses. C_4 monocot distribution is often associated with temperature, as it correlates strongly with warmer growth season temperatures and tropical mountain elevation gradients, while C_4 eudicot distribution correlates better with aridity gradients (Teeri and Stowe, 1976; Rundel, 1980; Williams *et al.*, 1995; Ehleringer *et al.,* 1997, Epstein *et al.,* 1997; Edwards and Still, 2008). However these trends simply highlight differences in evolutionary history and cold/dry adaptions, as C_4 lineages tend to inhabit similar environments to their C_3 ancestors. This point is demonstrated well across the monocots, where C_4 grass abundance correlates with temperature, fire frequency, and/or grazing pressures in different lineages (Liu *et al*., 2012; Visser *et al*., 2012). C_4 distribution may also be associated with disturbance from natural and anthropogenic causes, as

photorespiration is an inherent response to these types of abiotic stresses (Voss *et al*., 2013).

Similar to the C_4 pathway, the C_2 cycle is a costly process, as photorespiration uses ATP and NADPH and produces toxic by-products that need to be eliminated. The C_2 cycle, however, does save on photorespiratory costs without incurring C_4 cycle costs. The costs associated with the C_2 cycle are not uniform across all environments, but increase under conditions promoting photorespiration. However, the benefits of running a C_2 cycle over the C_3 pathway alone also increase in these environments, as the C_2 cycle can scavenge more photorespired CO_2. Plants using a C_2 cycle are often fast-growing herbaceous species that grow mainly during the summer months. These C_2 plants often inhabit disturbed habitats and those with high evaporative demands (Monson and Moore, 1989; Sage *et al.,* 2013). Because the C_2 cycle is based on avoiding the costs of photorespiration, it is likely to be common in high photorespiration environments, such as those favouring C_4 plants.

Importance of grasses

Plants using the C_4 photosynthetic pathway are ecologically and economically valuable to humans. They cover nearly twenty percent of the global vegetated land surface and contribute about a quarter of the total terrestrial primary productivity (Still *et al*., 2003), despite accounting for only about three percent of angiosperm species (Sage *et al.*, 1999a). Over half of C_4 species are grasses and include the crops maize, millets, sorghum, sugar cane, miscanthus, and switchgrass, which provide food, biofuel, and livestock feed (Byrt *et al.*, 2011). Furthermore, approximately one-fifth of the global human population depend on grass-dominated savannas for the ecosystem services they provide in the form of pasture land, fuel, food, and medicinal plans (reviewed in Parr *et al*., 2014).

As mentioned above, only the PACMAD clade of grasses evolved C_4 photosynthesis, while the BEP clade continues to use the C_3 pathway and has never evolved C_4 photosynthesis. The world's most widely consumed staple crops, including rice, wheat, and barley, are BEP grasses. International multidisciplinary research teams

have been assembled to focus on improving photosynthesis in these crops. One long-term goal of these projects is to introduce the more efficient C_4 photosynthetic pathway into these C_3 crops. If successful, these projects may have remarkable impacts on improving world hunger and nutrition status (Hibberd *et al*., 2008; von Caemmerer *et al*., 2012). Thus, research that illuminates the transitory steps between C_3 and C_4 phenotypes will inform these crop projects and consequently contribute toward food security.

Alloteropsis semialata as a model system for C_4 evolution

The genus *Alloteropsis* falls within the Paniceae tribe of the Panicoideae PACMAD grasses. Although this genus only includes five species; *A. angusta*, *A. cimicina*, *A. paniculata*, *A. papillosa*, and *A. semialata* (Ibrahim *et al*., 2009), it has remarkably diverse photosynthetic phenotypes. All of these species use C_4 photosynthesis, however, they may differ in how this is realized (Brown, 1977; Hattersley and Watson, 1992). First, *A. cimicina*, *A. paniculata*, and *A. papillosa* use the outer bundle sheath, while *A. semialata* and *A. angusta* use the inner bundle sheath, for the Calvin cycle (Hattersley and Watson, 1976; Brown, 1977; Ellis 1977; 1981; Renvoize, 1987a). Furthermore, analyses of anatomical features suggest that *A. angusta* uses the NADP-ME subtype while *A. cimicina*, *A. paniculata*, and *A. papillosa* use the NAD-ME subtype. However, recent transcriptome data suggest that *A. cimicina* uses NADP-ME (Christin, unpublished). A combination of biochemical and molecular approaches indicate that *A. semialata* uses a combination of the PCK and NADP-ME shuttles (Frean *et al.,* 1983; Prendergast *et al.*, 1987; Hattersley and Watson, 1992; Ueno and Sentoku, 2006; Christin *et al*., 2013b).

The species *A. semialata* also includes populations that use C_3 photosynthesis. The discovery that this species encompasses populations that use either C_3 or C_4 photosynthetic types was made forty years ago, but was based primarily on leaf anatomy (Ellis, 1974a, b; Brown, 1975; Ellis, 1981; Frean *et al*., 1983a, b), which led to the classification of two subspecies supported by morphological disparities; *A. semialata* subsp. *eckloniana* (Nees) Gibbs Russell, stat. nov. and *A. semialata* subsp. *semialata* (Figure 1.6; Gibbs Russell, 1983). A survey of $\delta^{13}C$ measurements

in 50 herbarium specimens found *A. semialata* subsp. *eckloniana* ranged between -25.4 and -28.5‰ and *A. semialata* subsp. *semialata* ranged between -9.3 and -13.8‰ (Box 1.1; Ellis, 1981). These stable isotope results reinforce the morphological divide between the two subspecies and suggest that *A. semialata* subsp. *eckloniana* includes variants that use C_3 photosynthesis while *A. semialata* subsp. *semialata* encompass those individuals that use C_4 photosynthesis.

Figure 1.6 Herbarium specimens exemplifying differences between the gross morphology of *Alloteropsis semialata* subsps. *eckloniana* and *semialata*. The C_3 eckloniana individuals (left) tend to have shorter, wider leaves and very dark, tightly packed inflorescences. The C_4 semialata individuals (right) have longer, narrower leaves with loosely packed inflorescences. All specimens collected in South Africa by G.K. Theron no. 2408 (top left), and R. P. Ellis 3497 (top right), 3812 (bottom left), and 3484 (bottom right). Photographs taken by Marjorie Lundgren.

Box 1.1 Stable isotope discrimination in C_3 and C_4 plants

The two naturally occurring stable isotopes of carbon, ^{12}C and ^{13}C, exist in different relative abundances within plant tissue and the atmosphere (Farquhar *et al*., 1989). This is because the heavier ^{13}C isotope is discriminated against during the incorporation of atmospheric carbon dioxide into plant tissue. Ratios of the relative abundance of ^{13}C to ^{12}C in plant tissue are used to reveal carbon isotope discrimination in plants. Thus, because C_3 and C_4 biochemistry inherently discriminates against the ^{13}C isotope to different degrees, carbon isotope discrimination is a useful tool to distinguish photosynthetic types (Whelan *et al*., 1973; Farquhar, 1983; von Caemmerer *et al*., 2014).

A C_3 plant will discriminate against the heavier ^{13}C isotope at two points, when CO_2 diffuses from the atmosphere into the chloroplast and when Rubisco fixes carbon (Farquhar, 1983). Discrimination in a C_4 plant reflects several steps, including the fractionations of Rubisco and the CA/PEPC system, as well as their interconnectedness, and even bundle sheath leakiness (Reviewed in von Caemmerer *et al*., 2014). Rubisco, however, is by far the strongest driver of ^{13}C discrimination (29-30‰), compared to that of PEPC (5.7‰) or discrimination during diffusion across air (4.4‰) or liquid (1.8‰) (reviewed in von Caemmerer *et al*., 2014). Thus, the ^{13}C isotope signature in plant tissue depends on Rubisco, the carbon-fixing enzyme, which cannot discriminate against ^{13}C when in the C_4 cycle. Thus, measuring the ratio of ^{12}C to ^{13}C in leaf tissue will reflect the degree to which a plant is using a C_4 cycle (Smith and Epstein, 1971; Brown, 1977). Because the discrimination factor of PEPC is so low, $\delta^{13}C$ in C_4 plants is similar to measures of $\delta^{13}C$ in the surrounding air, while values in C_3 plants are much lower. Indeed, $\delta^{13}C$ values in C_3 grasses range between -34 and -19‰ while those of C_4 grasses range between -18 and -9‰ (Table 3.1; Cerling *et al*., 1997).

Intermediacy in Alloteropsis semialata

The designation of two subspecies helped to categorize the abundant variation in *A. semialata*, however, this does not account for all of the photosynthetic variation observed in this species. Several anomalous phenotypes emerged from studies of *A. semialata* herbarium specimens. First, a plant collected in north-western Zambia in 1937 (Milne-Redhead 3021) presented $\delta^{13}C$ values of -20.0‰ and -18.5‰ when measured twice, both of which fall in-between values characteristic of the C_3 and C_4 subspecies of *A. semialata* (Table 1.1; Ellis, 1981). Shortly after this finding, a survey of morphological characteristics in herbarium specimens identified three specimens that fell intermediate between C_3 and C_4 phenotypes in vein width measurements collected on the leaf blade tips and basal sheaths (Table 1.1; Gibbs Russell, 1983). These specimens were also collected in central eastern Africa. Renvoize's (1987a) survey of leaf anatomy in the Paniceae identified four anatomically anomalous *A. semialata* specimens with C_4 biochemistry (Table 1.1; Brown, 1977; Frean *et al*., 1983b). Three of these four specimens averaged more

than four mesophyll cells between veins and lacked quaternary veins, which is abnormal for otherwise C_4 plants. These specimens were collected in Tanzania, Zambia, and South Africa. The fourth specimen, which was collected in Zimbabwe, had quaternary veins and fewer mesophyll cells-which is typical of C_4 plants but had a radiate mesophyll cell arrangement, a C_3-like phenotype. Five years later, a comprehensive review of $\delta^{13}C$ values and leaf anatomy in *A. semialata* herbarium specimens revealed new anomalous specimens, also originating in central eastern Africa (Table 1.1; Long and Hattersley, unpublished data; Hattersley and Watson, 1992). Unfortunately, little work has been done to investigate these anomalous phenotypes since the early 1990s, leaving these reports anecdotal until further analysis.

Biogeography of Alloteropsis semialata

A survey of herbarium specimens suggested the C_3 and C_4 subspecies of *A. semialata* have different geographic and ecological distributions (Ellis, 1981). *A. semialata* subsp. *semialata* is widespread, having been collected in northern and eastern Australia, southeast Asia, and widely distributed throughout western and south eastern Africa. Within Africa, these C_4 specimens were only found sub-Saharan, and never within the Congo Basin, Somaliland, or Kalahari desert (Ellis, 1981). In contrast, *A. semialata* subsp. *eckloniana* had only been collected in southern and eastern South Africa and northward into Zimbabwe and Zambia. The C_3 and C_4 subspecies overlap in these three countries. However, they tend to occupy different habitats. The C_4 subspecies was primarily collected in low altitude areas with slightly acidic soils. The C_3 subspecies was primarily collected in higher altitude areas, often on South facing slopes, with strongly to moderately acidic soils. Both subtypes tolerate relatively infertile soils and seasonal fire (Acocks, 1975; Ellis, 1981). Interestingly, Ellis (1981) noted that the C_3 subspecies, which was found in cool, high altitude areas in South Africa and Zimbabwe, becomes atypical and more C_4-like in the warmer, central African collection locations. Again, these observations were never followed by in-depth investigations and the little research that has been undertaken on this species only focused on specific geographic regions, namely South Africa and Australia, leaving the intriguing story of the central African populations elusive.

Table 1.1 Herbarium specimens of *Alloteropsis semialata* with anomalous $\delta^{13}C$ isotope values (I), morphology (M), and/or leaf blade anatomy (A).

Specimen	Country	Date	$\delta^{13}C$	Anomaly	Study[1]
Milne-Redhead 3021	Zambia	October 1937	-20.0, -18.5, 20.7	I	A, D, E
Eyles 1920	Zimbabwe	November 1919	-10.3	M	B
Greenway 6290	Malawi	-	-	M	B
Milne-Redhead 3371	Zambia	November 1937	-18.6	I, M	B
Milne-Redhead & Taylor 8455	Tanzania	January 1956	-10.7	M	C
Astle 1137	Zambia	-	-	M	C
Norval 106	South Africa	-	-	M	C
Rattray 428	Zimbabwe	-	-11.4	M	C
Emson 340	Tanzania	1932	-21.4	I	D, E
Stowe 495	Zambia	December 1940	-20.8	I	D, E
Brzotowski 26	Tanzania	December 1944	-22.6	M, A	D, E
Robinson 4744	Zambia	December 1961	-22.7	M, A	D, E
Proctor 2165	Tanzania	November 1962	-23.0	M, A	D, E
Mbano DSM812	Tanzania	December 1969	-23.9	M, A	D, E
Proctor 2206	Zambia	December 1962	-23.9	M, A	D, E
Simon 932	Tanzania	October 1966	-25.3	M, A	D, E
Tanner 5076	Tanzania	July 1960	-26.3	M, A	D, E
Bogdon & Williams 238	Kenya	January 1947	-11.8	M, A	D, E

[1]Studies refer to (A) Ellis, 1981; (B) Gibbs Russell, 1983; (C) Renvoize, 1987a; (D) Long and Hattersley, unpublished; and (E) Hattersley and Watson, 1992.

Cytological variation in Alloteropsis semialata

The C_3 and C_4 subspecies of *A. semialata* also differ in their cytology. While only diploid (2n = 2x = 18) populations of *A. semialata* subsp. *eckloniana* have been identified previously (Liebenberg and Fossey, 2001), the C_4 subspecies presents a polyploid series with diploid (2n = 2x = 18), tetraploid (2n = 4x = 36), hexaploid (2n = 6x = 54), octoploid (2n = 8x = 72), and dodecaploid (2n = 12x = 108) individuals (Ellis, 1981; Frean and Marks, 1988; Liebenberg and Fossey, 2001; Ueno and Sentuko, 2006). Of these, the tetraploid and hexaploid forms of the C_4 subspecies are most common in South Africa, while octoploid and dodecaploid individuals are rare, diploids are completely lacking in this region (Liebenberg and Fossey, 2001). These detailed investigations, which included within-population analyses, have been widely reported and led to the belief that C_4 *A. semialata* individuals are polyploid, while individuals of the C_3 subspecies are diploid. However, these studies were conducted only within South Africa, and exploratory work has suggested that C_4 populations from Asia and Australia are diploid, while those inside of Africa present different levels of polyploidy (Ellis, 1981). Again, comprehensive cytological analyses across *A. semialata*'s geographic range are lacking.

Dissertation aims and structure

The overall aim of this thesis is to identify the small-scale processes that lead to the emergence of novel traits of impressive complexity. This problem is addressed here using C_4 photosynthesis as a study system. While different aspects of C_4 evolution have been studied at great length, previous investigations have relied solely on inter-specific comparisons, so that small-scale evolutionary changes remained out of reach. I tackled this problem by studying the photosynthetic variation that exists within a single species, the grass *A. semialata*. I combined approaches from the fields of plant physiology, anatomy, ecology, and phylogenetics to unravel the links between geographic expansion, ecological diversification, and physiological adaptation associated with photosynthetic pathway evolution.

Chapter 2 of this dissertation evaluates the significance of the variation in leaf anatomy within grasses via a literature review and data synthesis aimed at identifying key anatomical differences between C_3 and C_4 plants. Although specific arrangements of leaf anatomy are considered a defining characteristic of C_3 versus C_4 photosynthetic types, the data synthesis presented in Chapter 2 reveals that these anatomies are not discrete, but form continua between photosynthetic types. The implications of this anatomical variation for the repeated evolution of C_4 photosynthesis are discussed.

Taking into account the continua of leaf anatomies expressed across C_3 and C_4 species (Chapter 2), the anatomical variation within *A. semialata* was quantified and coupled with physiological measurements to investigate the relationships between form and function in this taxon in Chapter 3. This chapter elucidates the role that structural reorganization plays in the establishment of photosynthetic types. Stable isotope, leaf component, physiology, and leaf anatomy methods are combined to define the photosynthetic variation that occurs within this species. By including populations from across the species' range, this chapter reveals the simultaneous presence of C_3, C_3-C_4 intermediate, and C_4 photosynthetic types in *A. semialata* for the first time.

Following the characterization of C_3, C_3-C_4 intermediate, and C_4 photosynthetic types in *A. semialata* in Chapter 3, the phylogenetic relationships among these photosynthetic types and their patterns of diversification and dispersal are presented in Chapter 4. A plastid phylogeny of over 60 *A. semialata* accessions is used to determine where, when, and in which environments the transitions between photosynthetic types occurred. Additionally, differences in niche diversification and dispersal patterns between photosynthetic types are highlighted to distinguish the causes from the consequences of C_4 evolution.

Considering that leaf anatomy varies quantitatively among C_3 and C_4 species (Chapter 2), that *A. semialata* includes C_3, C_3-C_4 intermediate, and C_4 photosynthetic types with different anatomies (Chapter 3), and that C_4 photosynthesis likely evolved from C_3-C_4 intermediate ancestors while another lineage reverted back to C_3 photosynthesis in *A. semialata* (Chapter 4), Chapter 5

pieces together how small-scale anatomical shifts in C_3 and C_3-C_4 intermediate accessions accumulated to promote the emergence of C_4 photosynthesis in this species. Here, intraspecific anatomical variation is comprehensively dissected to determine how these traits vary within *A. semialata* across geographic space and environments. The phenotypic distance between C_3 and C_4 anatomies is presented and the importance of the C_3-C_4 intermediate phenotype for bridging that phenotypic landscape is discussed.

The implications of the findings presented in Chapters 2 - 5 on the evolution of photosynthetic types in *A. semialata* are discussed in Chapter 6. These findings are put into a broader context to shed new light on the current understanding of C_4 evolution, and complex trait evolution, in general.

CHAPTER 2

DECONSTRUCTING KRANZ ANATOMY TO UNDERSTAND C_4 EVOLUTION

This chapter is a modified version of the published manuscript:

Lundgren MR, Osborne CP, Christin PA. 2014. Deconstructing Kranz anatomy to understand C_4 evolution. Journal of Experimental Botany 65, 3357-3369.

A reprint of this article is bound to the end of this dissertation.

Introduction

During the diversification of flowering plants, C_4 photosynthesis evolved from C_3 ancestors more than 62 times independently in several distantly related groups (Sage *et al.*, 2011). C_4 photosynthesis is characterised by a biochemical CO_2 pump formed by the coordination of several evolutionary novelties, which increase the relative concentration of CO_2 around Rubisco to nearly eliminate photorespiration (Ludwig and Canvin, 1971; Hatch, 1987; von Caemmerer and Furbank, 2003; Skillman, 2008; Sage *et al.*, 2012). The CO_2-concentrating mechanism relies on the primary fixation of atmospheric carbon by PEPC coupled with CA. These reactions are spatially separated from the secondary refixation of CO_2 by Rubisco (Hatch, 1987; von Caemmerer and Furbank, 2003). An efficient segregation of these C_4 biochemical reactions requires specific leaf properties (Hattersley, 1984; Dengler *et al*., 1994; Muhaidat *et al*., 2007).

As a result of its multiple origins, C_4 photosynthesis does not present a consistent and discrete phenotype, so is better considered a functional trait involving a suite of coordinated leaf anatomical and biochemical characters (Brown and Smith, 1972; Laetsch, 1974). These components can assemble differently during each origin of C_4 photosynthesis, and these divergent evolutionary histories result in high anatomical and biochemical diversity among, and sometimes within C_4 lineages (Hattersley and Watson 1992; Sinha and Kellogg, 1996; Kadereit *et al*., 2003; Muhaidat *et al.,* 2007; Edwards and Voznesenskaya, 2011; Freitag and Kadereit, 2013). An understanding of the evolutionary transitions leading to the recurrent assembly of C_4 photosynthesis requires investigation of the individual characters that together generate C_4 function, not only in C_4 species but also in C_3 species variously related to C_4 taxa (Christin and Osborne, 2013). It is particularly important to differentiate the present function of each component from its identity and developmental origin. In this work, we focus on the variation observed in both C_3 and C_4 plants in each of the anatomical traits that together generate leaf functions compatible with C_4 photosynthesis. We combine a review of the literature with analyses of a quantitative leaf anatomy dataset compiled from 157 C_3 and C_4 grass species (Christin *et al.*, 2013a). The C_4 grasses in this dataset encompass eight of the nine structural C_4 forms described for this family (Edwards and Voznesenskaya, 2011).

What is C_4 leaf anatomy?

Differential arrangements of cells and organelles within the leaves of taxa we now recognise as C_3 and C_4 were first observed and published more than 80 years before the C_4 pathway itself was discovered (Duval-Jouve, 1875; Haberlandt, 1884). The association between specific cell and organelle arrangements and the C_4 pathway was then identified soon after the discovery of C_4 photosynthesis (El-Sharkawy and Hesketh, 1965; Downton and Tregunna, 1968; Berry *et al*., 1970; Welkie and Caldwell, 1970). Since then, C_4 photosynthesis has usually been affiliated closely with a suite of leaf properties referred to as "Kranz" anatomy (after Haberlandt's description in German of a wreath-like arrangement of cells). Kranz anatomy can be described as two distinct concentric layers of chlorenchyma cells, formed by an inner bundle sheath containing most of the chloroplasts, surrounded by an outer layer consisting of a small number of mesophyll cells. The visual identification of such arrangements in transverse section has been used in numerous anatomical surveys of leaves to identify the photosynthetic pathway for hundreds of species (Welkie and Caldwell, 1970; Carolin *et al*., 1973; Carolin *et al*., 1975; Brown, 1977; Carolin *et al.*, 1977; Hattersley *et al.*, 1982; Renvoize, 1987a).

Surveys of numerous C_3 and C_4 species over the past five decades have shown that leaf anatomies cannot be easily and consistently grouped into discrete categories corresponding to the two photosynthetic types, but come in many flavours (Brown, 1975; Edwards and Voznesenskaya, 2011). It is true that the leaf anatomy of a randomly selected C_3 plant is highly likely to deviate significantly from that of a randomly selected C_4 plant. For example, *Viburnum punctatum*, like most C_3 eudicots, has distinct horizontal layers of mesophyll cells in its leaves (Figure 2.1a), arranged such that it does not conform to the general anatomical pattern generally present in C_4 plants, whereby an inner and outer bundle sheath, and sometimes a ring of mesophyll cells, form concentric circles around the vasculature (Figure 2.1c). A similar concentric arrangement of cells can be found in many C_3 grasses though (Figure 2.1b; Figure 2.2; Hattersley *et al.*, 1982; Dengler *et al.*, 1994; Besnard *et al.*, 2013) and, as detailed below, individual leaf characters that are usually associated with a C_4 function can be found in at least some C_3 plants. Furthermore, some plants achieve C_4 photosynthesis without the segregation of photosynthetic reactions into different types of cells (Bowes and Salvucci, 1984;

Bowes and Salvucci, 1989; Freitag and Stichler, 2000; Edwards *et al.*, 2004). Despite this variation, C_4 physiology is still associated with a suite of functional properties (Brown and Smith, 1972; Edwards and Voznesenskaya, 2011), which must first be considered before analysing diversity in the identity and developmental origins of the characters that generate them. Based on the literature, the following functional properties of leaves are considered essential requirements for C_4 photosynthesis (Hattersley *et al.*, 1977; Leegood, 2002; von Caemmerer and Furbank, 2003; Edwards and Voznesenskaya, 2011; Nelson, 2011). Note that these apply equally to all C_4 plants, whether or not they use distinct types of cells.

1. There must be two distinct compartments arranged so that atmospheric gases reach the first compartment more easily than the second. The first compartment houses the PEPC reactions while the second is modified to restrict CO_2 efflux and retain the Calvin cycle.
2. The two compartments must be in close contact to allow the rapid exchange of metabolites.
3. The compartment where the Calvin cycle occurs must occupy a large enough fraction of the leaf to accommodate a significant number of chloroplasts.
4. Chloroplasts must be abundant in the Calvin cycle compartment.

These functional properties are extremely important for C_4 physiology and biochemistry. However, to understand the gradual evolutionary changes leading to the recurrent assembly of C_4 photosynthesis, it is important to account for exact changes in cellular characters, and the genetic determinants of these characters. In the following sections we therefore discuss how each of the four properties listed above are generated from underlying characters. We look at how these characters vary qualitatively and quantitatively among C_3 and C_4 lineages, and show how there is an overlap between the values observed in C_3 and C_4 species.

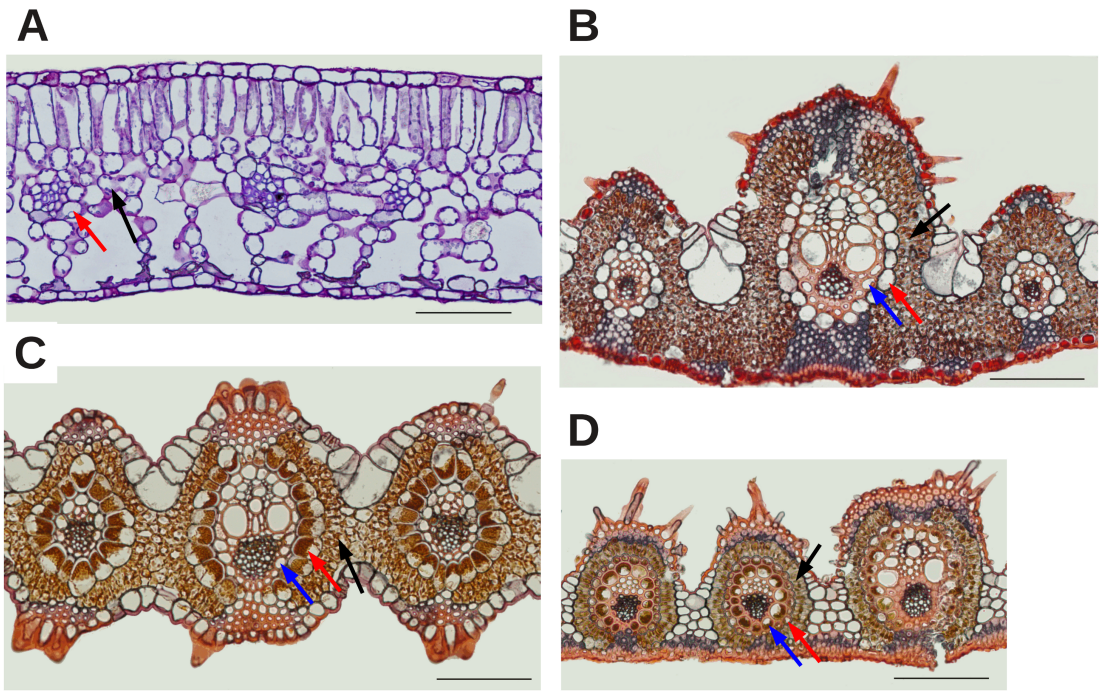


Figure 2.1 Examples of C_3 and C_4 leaf cross-sections. The C_3/C_4 pair on the left (A, C) are unrelated, belonging to different major groups of flowering plants. By contrast, the C_3/C_4 pair on the right (B, D) is composed of closely related species, belonging to the same subfamily of grasses. A. *Viburnum punctatum* (C_3, Adoxaceae), B. *Sartidia angolensis* (C_3, Poaceae), C. *Centropodia mossamedensis* (C_4, Poaceae), and D. *Aristida mollissima* (C_4, Poaceae). The black arrows point to the mesophyll, red arrows to the outer bundle sheath, and blue arrows to the inner sheath of grasses (= mestome sheath). The four cross-sections are shown at the same scale. The black bars represent 100 µm. Picture A. was kindly provided by Dr. David Chatelet from Brown University and pictures B., C. and D. come from collections of Prof J. Travis Columbus from Rancho Santa Ana Botanic Garden.

1. Two compartments differentially connected to the atmosphere

In C_3 plants, the Calvin cycle occurs in most of the leaf, while it is restricted to specific locations in C_4 plants. It is well known that the identity of the compartments co-opted for the segregation of the atmospheric CO_2 fixation by PEPC and its refixation by the Calvin cycle differs among C_4 origins (*e.g.,* Brown, 1975; Dengler *et al.*, 1985). For instance, some single-celled C_4 species have evolved separate compartments for the PEPC and Calvin cycle reactions through

the rearrangements of organelles or vacuoles within individual photosynthetic cells (Edwards *et al.*, 2004). In the majority of C_4 plants, however, the PEPC and Calvin cycle reactions are segregated in different types of cells. In C_3 species, the mesophyll and bundle sheath represent two physiologically distinct types of cells, and the central position of bundle sheath cells within the leaf gives the opportunity for minimal contact with the atmosphere (Figure 2.1a, b, Figure 2.3 and Appendix 2.1). The bundle sheaths have consequently been co-opted for Calvin cycle reactions across most C_4 origins, while the mesophyll cells, which are better connected to the atmosphere, are used for the PEPC reactions. Despite this convergence in function, the bundle sheath cells recruited for C_4 photosynthesis are not homologous among all C_4 origins.

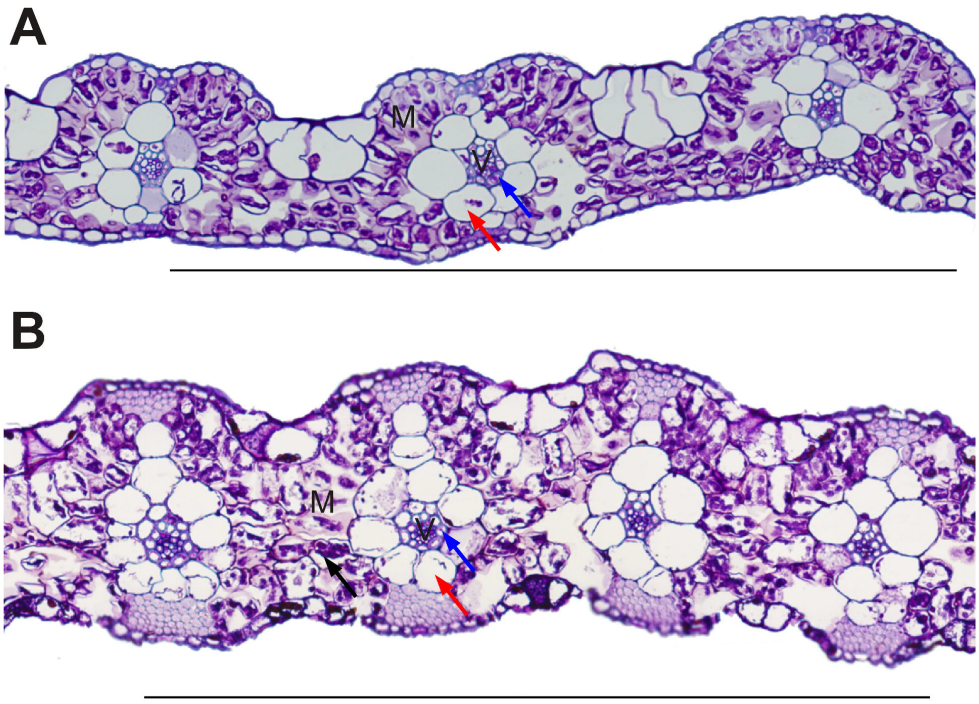


Figure 2.2 Examples of C_3 grasses with leaf anatomy close to the C_4 requirements. A. *Panicum pygmaeum* (C_3), B. *Panicum malacotrichum* (C_3). The mesophyll (M) and vascular tissue (V) are indicated on the sections. The red arrow points to the outer bundle sheath while the blue arrow points to the inner sheath (= mestome sheath). The black bars on the bottom represent 500 µm.

In some C_4 species within the genera *Arundinella*, *Garnotia, Arthropogon*, *Achlaena*, *Dissochondrus*, *Anrthraxon*, and *Microstegium*, the Calvin cycle also occurs in distinctive cells, which are atypical bundle sheath-like cells, differentiated within the mesophyll, but not associated with vascular bundles (Figure 2.3a; Tateoka, 1958; Hattersley and Watson, 1992; Ueno, 1995; Dengler *et al*., 1996; Wakayama *et al.,* 2003). In addition, grasses and sedges possess multiple layers of sheath cells, with inner layers derived from procambium (often referred to as the "mestome sheath") and outer layers from ground meristem (Dengler *et al.,* 1985; Soros and Dengler, 2001; Martins and Scatena, 2011). In studies of C_4 photosynthesis, consideration of the different cells is often based on their function. However, for evolutionary studies, the ontogenic origin of each type of cells needs to be established independently of its function. The C_4 lineages within grasses and sedges have alternatively co-opted one or both of these cell types, while the second cell layer is often lost, for example in the numerous C_4 grasses with a single sheath layer (Figure 2.3a-e; Brown, 1975; Dengler *et al.,* 1996; Soros and Dengler, 2001; Martins and Scatena, 2011). This diversity in the identity of the two compartments co-opted for the segregation of C_4 reactions, together with phylogenetic analyses, has been used previously to argue for multiple independent C_4 origins, rather than fewer origins followed by reversals in closely related C_3 species (Kellogg, 1999; Christin *et al.*, 2010).

The limited connection of the Calvin cycle compartment to the atmosphere is also achieved via different mechanisms in the different C_4 lineages. First, tightly packing mesophyll cells around the bundle sheath reduces the fraction of cells from the latter that are in contact with the atmosphere (Dengler *et al.*, 1994; Muhaidat *et al.*, 2007), although similar packing also occurs in some C_3 grasses (Figure 2.1b; Dengler *et al.*, 1994) and some C_3 eudicots (Muhaidat *et al.*, 2007). In addition, the bundle sheath cell walls can also be covered with a layer of suberin that limits gas diffusion. This is the case in C_4 monocots that have co-opted the inner sheath layer for a C_4 function (Hattersley and Browning, 1981; Ueno *et al.*, 1988a). However, the presence of suberin layers on the inner sheath cell walls can also be found in most C_3 grasses (Hattersley and Browning, 1981). Neither of the characteristics reducing contact of the Calvin cycle with the atmosphere is therefore found exclusively in C_4 plants. However, isolating the Calvin cycle

compartment from the atmosphere may actually be of little functional importance to a C_4 plant as the C_4 system efficiently concentrates CO_2 to the point of saturation in the BS (Jenkins *et al*., 1998), such that any reduction of O_2 achieved by isolating the bundle sheath has little influence on the overall CO_2:O_2 ratio there, and thus, the efficiency of the Calvin cycle in these plants.

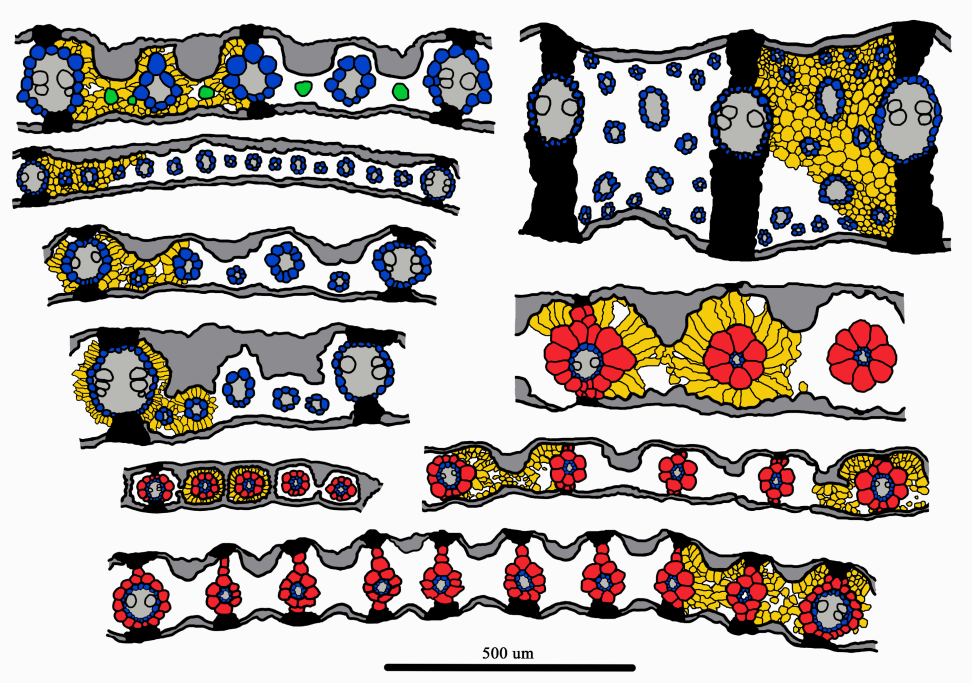


Figure 2.3 Leaf anatomy for selected cross-sections of grasses. A. *Arundinella nepalensis* (C_4), B. *Anthaenantia lanata* (C_4), C. *Axonopus compressus* (C_4), D. *Ischaemum afrum* (C_4), E. *Chrysopogon pallidus* (C_4), F. *Alloteropsis cimicina* (C_4), G. *Panicum pygmaemum* (C_3), H. *Bouteloua stolonifera* (C_4) and I. *Panicum malacotrichum* (C_3). Diagrams highlight the mesophyll cells (yellow), outer bundle sheaths (red), inner bundle sheaths (blue), and distinctive cells (green). Uncolored central areas are composed of mesophyll cells and intercellular airspace. Vein (light grey), epidermis (dark grey), and sclerenchymatous girders (solid black) are also shown. Where only one bundle sheath is present, it is assumed the outer bundle sheath has been lost and the inner bundle sheath remains. All cross-sections are drawn at the same scale, indicated at the bottom. The corresponding pictures can be found in Appendix 2.1.

2. Distance between the two compartments

Close contact between the PEPC and Calvin cycle compartments is guaranteed in plants with a single-celled C_4 system. In plants with a dual-celled C_4 system, the presence of mesophyll cells not directly adjacent to the bundle sheaths will increase the average distance between the compartments containing PEPC and Rubisco. This problem is usually solved in C_4 plants by limiting the number of cells separating consecutive Calvin cycle compartments, and by organizing mesophyll cells into one or two layers around the bundle sheath (Figure 2.1c, d), which produces the classical pattern of Kranz anatomy. In some species, this configuration is achieved through the development of a single bundle sheath layer that encompasses all the vasculature within the leaf and often water storage cells as well, and a single layer of mesophyll surrounds the bundle sheath. Variations on this anatomical theme are common among C_4 eudicots and have been found in the Asteraceae, Amaranthaceae, and Cleomaceae families (Carolin *et al*., 1975; Das and Raghavendra, 1976; Kadereit *et al.*, 2003; Peter and Katinas, 2003; Edwards and Voznesenskaya, 2011; Koteyeva *et al*., 2011). Some C_4 grasses have similar bundle sheaths that extend horizontally from the vascular tissue and join together, such that the mesophyll becomes isolated in small patches (Renvoize, 1983).

For C_4 lineages with multiple photosynthetic units formed by concentric cell layers of mesophyll, bundle sheath and vascular tissue, the presence of fewer mesophyll cells between consecutive veins can be achieved via two different developmental mechanisms. First, the number of cells that develop between consecutive bundle sheaths can be directly reduced during ontogeny. Second, extra Calvin cycle compartments, such as distinctive cells or minor veins, can be added to decrease the average distance between compartments, as has been documented in both monocots (*e.g.*, Poaceae; Figure 2.3a-e; Renvoize, 1987a; Dengler *et al.*, 1994; Ueno *et al*., 2006; Christin *et al.*, 2013a) and eudicots (*e.g.*, Asteraceae; McKown and Dengler, 2007; McKown and Dengler, 2009; and Cleomaceae; Marshall *et al.*, 2007). Interveinal distance (or vein density) is often considered a proxy for the number of mesophyll cells between consecutive bundles, and largely overlaps between C_3 and C_4 grasses (Christin *et al*., 2013a) and eudicots (Muhaidat *et al*., 2007). However, the relationship between interveinal distance and the number of mesophyll cells is only partial. First, because interveinal distance is influenced both by the diameter

of the veins and the size of the bundle sheaths, measuring the actual distance between bundle sheaths is more relevant. This distance is influenced by the size of individual mesophyll cells, their orientation, and finally their number (Figure 2.4c). Some C_4 species, such as *Alloteropsis cimicina*, have relatively large interveinal distances but with only a few large mesophyll cells between consecutive bundles (Figure 2.3f; Figure 2.4c). In addition, the number of mesophyll cells containing PEPC below and above veins can influence the average distance between the PEPC and Calvin cycle reactions independently from the distance between consecutive bundles. Some thick C_4 leaves, such as those of *Anthaenathia lanata* (Figure 2.3b) or some *Portulaca* (Ocampo *et al.*, 2013), consequently require a three-dimensional venation system. Finally, leaf thickness is often reduced between veins so that there are few mesophyll cells in positions most distant from the bundle sheaths, and interveinal distance can greatly exceed the average distance between photosynthetically active mesophyll cells and bundle sheath cells (Figure 2.1; Figure 2.3). For instance, in leaves of the C_3 grass *Panicum pygmaemum*, the average number of mesophyll cells between bundles greatly exceeds four. However, because its leaf thickness decreases between veins, the number of mesophyll cells separated from the bundle sheath by more than one cell is smaller than the number of mesophyll cells separated from the bundle sheath by zero or one cell (38 versus 73 cells between the three veins in Figure 2.2). Finally, the distance between consecutive bundles can be increased by the presence of achlorophyllous cells that do not influence the average path length from PEPC to Calvin cycle cells (*e.g.*, Figure 2.1d).

The number of mesophyll cells between consecutive bundles will distinguish C_3 from C_4 taxa with a high success rate and has consequently been proposed as a criterion to recognize C_4 plants (Hattersley and Watson, 1975; Renvoize, 1987a; Sinha and Kellogg, 1996). However, the C_3 and C_4 distributions for this trait also overlap (Figure 2.4c). For instance, *Panicum malacotrichum* is a C_3 grass with less than four mesophyll cells between veins (Figure 2.2). The variation observed in both C_3 and C_4 taxa is probably due to the importance of vascular architecture for both photosynthetic types. While the distance between consecutive bundles affects the efficiency of C_4 photosynthesis (Ogle, 2003), vein density also influences the transport of metabolites, leaf hydraulics and other physiological characteristics in

C_3 plants (Sack and Scoffoni, 2013; Sack *et al.*, 2013). In summary, both interveinal distance and the number of mesophyll cells between consecutive bundles overlap in C_3 and C_4 taxa, so that C_4 values represent only a subset of those observed among all photosynthetic types (Figure 2.4a-c; Muhaidat *et al*., 2007; Christin *et al.,* 2013a).

The transport of metabolites between the PEPC and Calvin cycle compartments in C_4 plants is also facilitated by a number of plasmodesmata connecting mesophyll and bundle sheath cells that exceeds the number found in C_3 plants (Olesen, 1975; Weiner *et al*., 1988; Botha, 1992). However, plasmodesmata frequency is known in only a few C_3 species, so the overall variation in this trait cannot be established with confidence.

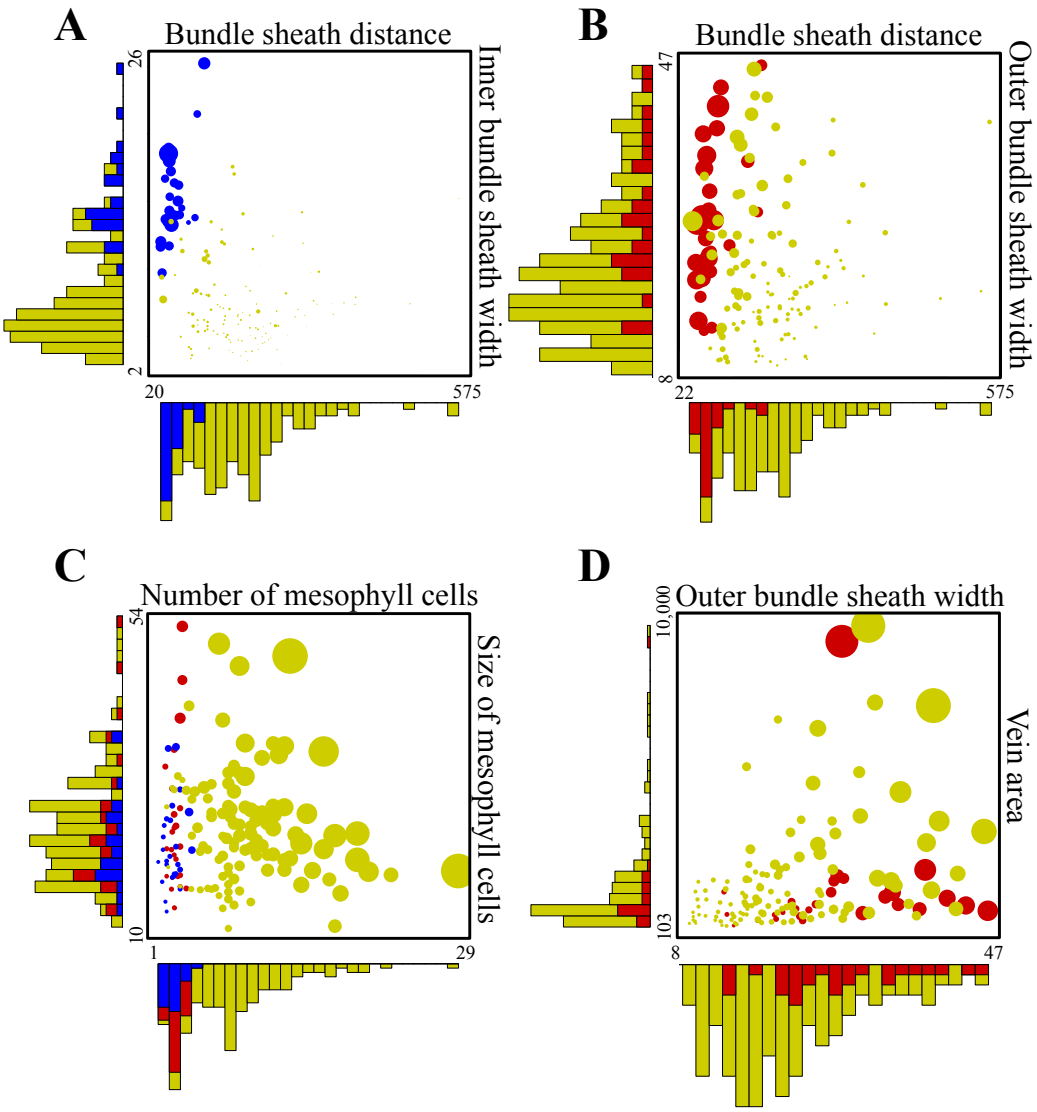


Figure 2.4 Multidimensionality of C_4 anatomy in grasses. Scatter plots for anatomical variables associated with the C_4 syndrome are shown, along with frequency distributions for each trait, arranged along the axes. For each pair of variables, dot size is proportional to a third variable. C_3 grass species are shown in yellow, C_4 grass species using the outer sheath for the Calvin cycle in red, and C_4 grass species using the inner sheath for the Calvin cycle in blue. Relationships between means of: (A) distance between consecutive bundle sheaths (µm) and inner bundle sheath cell width (µm), with dot size proportional to the percentage of inner bundle sheath area (µm^2); (B) distance between consecutive bundle sheaths (µm) and outer bundle sheath cell width (µm), with dot size proportional to the percentage of outer bundle sheath area (µm^2); (C) number of mesophyll cells between consecutive bundles and mesophyll cell length (µm), with dots proportional to the distance between consecutive bundle sheaths (µm); and (D) and outer bundle sheath cell width (µm) and area of vasculature (µm^2), with dot size proportional to the outer bundle sheath area (µm^2) per vein number, are shown. Data for 172 grasses (representing 157 species) come from Christin *et al.* (2013).

3. Large Calvin cycle compartment

The amount of CO_2 that can be re-fixed by Rubisco in the Calvin cycle will depend on the number of chloroplasts within the compartment co-opted for this function. The size of this compartment, not including the volume occupied by the vacuole, will influence the number of chloroplasts that can be accommodated. Thus C_4 plants tend to have enlarged bundle sheath cells able to accommodate numerous chloroplasts. More than the size of individual bundle sheath cells, the cumulative volume of bundle sheath relative to the PEPC compartment (mesophyll) is relevant, and seems to be constrained within a given range in C_4 plants (Hattersley, 1984; Dengler *et al.,* 1994; Muhaidat *et al.*, 2007). This might represent a trade-off between having sufficient chloroplasts in the Calvin cycle compartment and still conserving enough mesophyll volume for PEPC.

Similar bundle sheath : mesophyll ratios can be achieved through different combinations of the numerator (volume of bundle sheath) and denominator (volume of mesophyll). For instance, similar proportions of bundle sheath can be achieved through alternative developmental mechanisms, involving the production of either larger or more numerous bundle sheath cells (the latter is generally achieved through a proliferation of veins; Figure 2.3; Hattersley, 1984; McKown and Dengler, 2009). The cross-sectional area of mesophyll per vein is mainly a function of the distance between veins, the thickness of the leaf (including the thickness between veins in comparison to that at the veins) and the presence of achlorophyllous cells (Christin *et al*., 2013a). On the other hand, when viewed in transverse section, the total area of a given type of bundle sheath per vein is a function of the size of the bundle sheath cells, the diameter of the veins and, in some cases, the completeness of the bundle sheath (Figure 2.4; Christin *et al.,* 2013a). For instance, the external bundle sheath of many grasses is not developed on the abaxial side of the leaf, which reduces the total volume of this tissue (Figure 2.5; *e.g.,* Renvoize, 1985; Renvoize, 1987b). Thus, the relative amount of bundle sheath tissue is a function of at least five distinct traits, which may all vary independently. Functionally similar characteristics can consequently arise through different developmental modifications, as highlighted by the diversity of C_4 leaf anatomy (Figure 2.4).

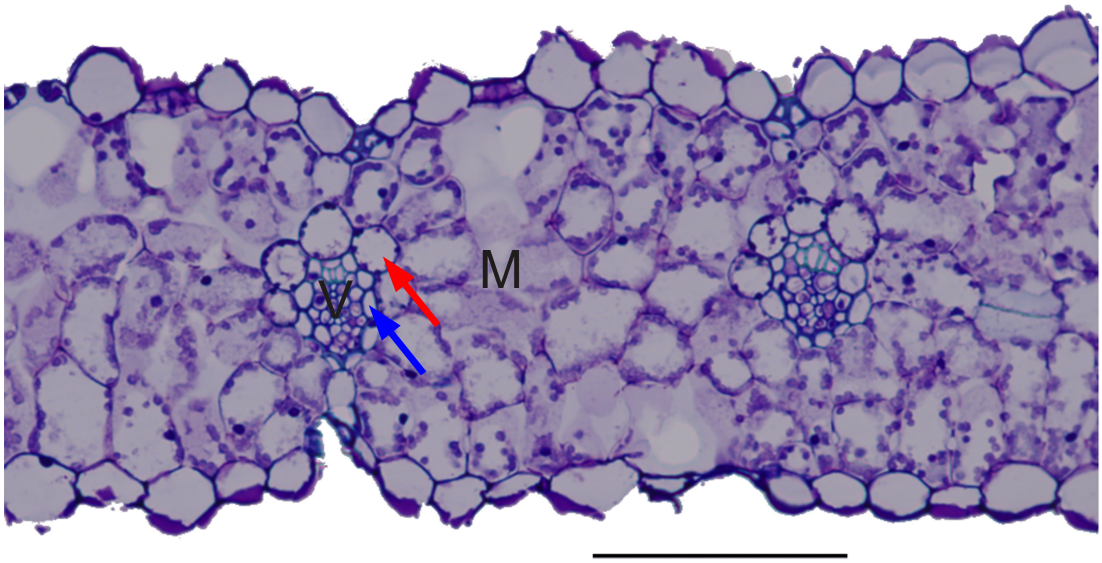


Figure 2.5 Detail of a cross-section for *Dactylis glomerata*. The mesophyll (M) and vascular tissue (V) are indicated on the section of this C_3 species. The red arrow points to the outer bundle sheath while the blue arrow points to the inner sheath (= mestome sheath). The black bar represents 100 µm. Note the incomplete outer sheath.

The five components that dictate the relative amount of bundle sheath tissue are important determinants of the gross leaf anatomy associated with C_4 photosynthesis. However, each component shows an essentially continuous distribution across C_3 and C_4 values, such that C_4-compatible ranges merely represent a subset of the distribution found in C_3 taxa (Figure 2.4; Marshall *et al.*, 2007; McKown and Dengler, 2007). The C_4-suitability of one parameter depends on the values of the other parameters. For instance, large volumes of bundle sheath tissue can arise in the presence of significant distances between consecutive bundles if the bundle sheath cells are enlarged (Figure 2.4a, b). This is highlighted by a comparison of *Alloteropsis cimicina* and *Axonopus compressus* (Figure 2.3f and c, respectively), which achieved similar ratios of bundle sheath per mesophyll area [BS/(BS+M) of 0.26 and 0.21, respectively] through different means. *A. cimicina* has very large outer bundle sheaths that are separated by long distances of mesophyll, while *A. compressus* has small inner sheaths that are separated by very short mesophyll distances in particularly thin leaves (Figure 2.3f and c, respectively).

During the course of evolution, numerous alterations in the characters that generate each leaf function occur either stochastically or in response to selective pressures. For instance, leaf thickness often represents an adaptation to the amount of light received by plants (Boardman, 1977; Terashima *et al*., 2001). The number and size of veins alters the hydraulics of a plant which, in turn, affects the sorting of plants across environments (McKown *et al*., 2010; Sack *et al*., 2012). Finally, the bundle sheath controls water flux between the mesophyll and vascular tissue such that an increase in bundle sheath size might provide better protection against cavitation in arid environments (Sage, 2001; Leegood, 2008; Griffiths *et al*., 2013). Recurrent and independent changes in different leaf properties repeatedly led to the emergence of tissues suitable for C_4 photosynthesis, which characterize numerous extant C_3 plants (Muhaidat *et al*., 2007; Edwards and Voznesenskaya, 2011; Muhaidat *et al*., 2011; Kadereit *et al*., 2012; Christin *et al*., 2013a; Griffiths *et al*., 2013).

4. Distribution of organelles

One of the most important requirements for C_4 photosynthesis probably lies in the distribution of chloroplasts. Although they are present in all photosynthetic cells of C_3 plants, chloroplasts are especially abundant in mesophyll cells and can vary from equally abundant to completely absent in bundle sheath cells (Figure 2.1; Figure 2.2, and Figure 2.5; Crookston and Moss, 1970). In C_4 plants, the light-dependent and light-independent functions of chloroplasts are often decoupled, and chloroplasts of the PEPC and Calvin cycle compartments can become morphologically and functionally differentiated (Woo *et al.*, 1970; Laetsch, 1974; Hattersley *et al.,* 1977; Bowman *et al.*, 2013). Although the characteristics and distribution of organelles vary among C_4 lineages (Ueno *et al*., 1988a; Voznesenskaya *et al*., 2006; Edwards and Voznesenskaya, 2011), the Calvin cycle compartment of C_4 plants consistently has a high concentration of chloroplasts, where the enzymes of the Calvin cycle are preferentially expressed.

No quantitative census of chloroplast distribution is available for randomly selected plants, however the organelle distribution has been investigated in species closely related to C_4 lineages, which shows that some plants maintain

significant numbers of chloroplasts in bundle sheath cells, despite lacking a functional C_4 pathway (Hattersely *et al.*, 1982; Ueno and Sentoku, 2006; Christin *et al*., 2013a). This is particularly common in plants using C_2 photosynthesis, a weak CO_2-concentrating mechanism based on a glycine shuttle from mesophyll to bundle sheath cells (Edwards and Ku, 1987; Sage *et al*., 2012). When chloroplast abundance in bundle sheath cells is compared among taxa, there is a gradient from closely related C_3 to C_2, and then from C_2 to C_4 species (Muhaidait *et al*., 2011; Sage *et al*., 2013). The C_2 trait is consequently often considered an evolutionary intermediate between C_3 and C_4 types (Hylton *et al*., 1988; Sage *et al*., 2012; Williams *et al*., 2013). Therefore, as for other anatomical traits, the number of chloroplasts in bundle sheath cells varies and may form a continuum between C_3 and C_4 species. Despite this, a high concentration of chloroplasts in bundle sheath cells might be the only trait that occurs systematically within dual-celled C_4 photosynthesis that is never present in non-C_4 plants. The tight association between C_4 physiology and chloroplast distribution is explained by the fact that C_4 physiology results from a differential distribution of the Calvin cycle (among other biochemical reactions), which is usually linked to the distribution of chloroplasts.

Other ultrastructural properties associated with some C_4 plants include the distribution of mitochondria and peroxisomes among compartments, the distribution of organelles within compartments and the ultrastructure and photochemical properties of the chloroplasts (Bruhl and Perry, 1995; Edwards and Voznesenskaya, 2011). Some of these properties are also observed in non-C_4 species closely related to C_2 and C_4 taxa (Sage *et al.*, 2012)

Plasticity for C_4-suitable anatomy

Phenotypic plasticity to environmental cues creates an additional layer of variation and further blurs the dichotomy between C_4 and non-C_4 anatomy. Specifically, plasticity for the anatomical traits relevant to photosynthesis (*e.g.*, compartmentalization, interveinal distance, mesophyll cell size and number, bundle sheath cell size, and organelle distribution) could partially explain the variation found in these anatomical characters or, more importantly, the shift of C_3 plants into the C_4-suitable space. Plasticity for these traits has been documented in

the literature. For example, the C_3 grass *Phragmites communis* acquires C_4-like traits when it grows at low soil water potentials (Gong *et al*., 2011). Specifically, interveinal distance decreases, chlorophyll content within bundle sheath cells increases, and the activity of C_4-related enzymes increases as soil water potential becomes more negative across a natural precipitation gradient (Gong *et al*., 2011). The C_4-like *Flaveria brownii* lacks the complete suite of anatomical characteristics required for a fully functioning C_4 system (Araus *et al*., 1990). However, this species can plastically increase its degree of C_4 photosynthesis by nearly doubling its investment in bundle sheath tissue relative to mesophyll in response to high irradiance (1200 μmol m^{-2} s^{-1} PPFD), compared to when it is grown at low irradiance (80 μmol m^{-2} s^{-1} PPFD; Araus *et al*., 1991). Furthermore, interveinal distances decreased in the C_3 grasses, *Festuca arundinaceae* (43% decrease) and the C_3-C_4 intermediate grass, *Panicum miliodies* (34% decrease), when grown in high versus low nitrogen levels (Bolton and Brown, 1980).

In addition to the plasticity of individual anatomical components, two different modes of environmentally induced C_4 photosynthesis exist. First, several aquatic species of Hydrocharitaceae, and possibly some Alismataceae and Cyperaceae, are able to switch from C_3 to single-cell C_4 photosynthesis (Bowes *et al.*, 2002). The environmental cue for this plasticity may be exposure to low CO_2 conditions, as they become submerged underwater, or seasonal variation in temperature (Bowes *et al*., 1979; Bowes, 2011). In contrast, some aquatic *Eleocharis* species use C_3 or C_3-C_4 intermediate photosynthesis when submerged but induce C_3-C_4 or C_4 photosynthesis by developing C_4-compatible leaf anatomy and expressing C_4 enzymes in the emergent leaves (Ueno *et al.*, 1988b; Ueno, 2001; Murphy *et al.*, 2007). Finally, some amphibious C_4 grasses seem to switch from a C_4 system that functions without C_4-associated leaf anatomy in aquatic leaves to a classical dual-cell C_4 cycle in aerial leaves (Keeley, 1998; Boykin *et al*., 2008).

Phenotypic plasticity for C_4-associated traits might have important implications for the evolution of C_4 photosynthesis (Sultan, 1987; West-Eberhard *et al.*, 2011). First, the direction and degree of phenotypic change in response to an environmental gradient is heritable (Schlichting and Levin, 1986; Schlichting and Pigliucci, 1993), and the reaction norm for a trait is genetically distinct from the

trait itself. Selection can therefore act independently on both a trait and on the plasticity for that trait. Plasticity may thus deter the evolutionary transition from C_3 to C_4 photosynthesis by diluting the effects of natural selection. However, adaptive phenotypic plasticity may promote C_4 evolution if the plastic expression of C_4-suitable anatomical traits in C_3 plants allows the colonization of new niches, leading to selective pressures for the gradual acquisition of C_4 biochemistry (Heckmann *et al.,* 2013). Indeed, Sage and McKown (2006) reviewed the literature to find that C_3 plants seem to be inherently more plastic than C_4 plants overall. Thus, this capacity for phenotypic plasticity might affect the probability of evolving C_4 photosynthesis. For instance, differential capacity in the phenotypic plasticity for important C_4 anatomical traits among plant lineages may explain differential propensity for C_4 evolution. However, the plasticity of C_4-associated anatomy remains mostly unknown in C_3 species, and more comparative work is required.

Consequences for the evolution of C_4-associated anatomy

When comparing the anatomy of a randomly selected C_3 taxon with that of a highly efficient C_4 species, the evolutionary transition from C_3 to C_4 anatomy can seem extraordinary (Figure 2.1a, c). However, it is important to note that C_4 photosynthesis did not emerge from the average C_3 taxon, but from C_3 ancestors with leaf anatomical properties much closer to the C_4 requirements (Figure 2.1b; Figure 2.5; Muhaidat *et al*., 2011; Christin *et al.,* 2013a; Sage *et al*., 2013). In the Poaceae, some species apparently using the C_3 photosynthetic type have gross leaf anatomies that closely resemble those of C_4 plants. For instance, *Panicum malacotrichum* and *Panicum pygmaeum* (Figure 2.2) are two C_3 grasses ($\delta^{13}C$ values of -27.4 and -29.7, respectively), which are closely related to several C_4 lineages (namely *Alloteropsis* and *Echinochloa*; GPWGII, 2012). These species possess large proportions of bundle sheath tissue that are firmly in the C_4 range [BS/(BS+M) of 0.26 and 0.23, respectively; Christin *et al.*, 2013a], and most mesophyll cells are directly adjacent to the bundle sheath or separated by only one mesophyll cell (Figure 2.2). Chloroplasts are still almost completely restricted to the mesophyll in these species. However, because the gross leaf anatomy is in place, fewer anatomical changes are necessary for the evolution of C_2 or C_4 photosynthesis. In other cases, such as the grass tribe Neurachninae, C_3 species

that are closely related to C_4 species have both C_4-like gross anatomy [BS/(BS+M) of 0.14-0.16; Christin *et al.*, 2013a] and the presence of conspicuous chloroplasts in the inner sheath, which was co-opted for C_4 photosynthesis in this group (Hattersley *et al*., 1982). These examples show that the evolution of C_4-suitable anatomy might not always require drastic modifications, since C_3 lineages may possess C_4-like values for individual traits that can generate C_4 leaf functions.

Each component of C_4-compatible leaf anatomy may vary independently within C_3 ancestors, such that any combination of mesophyll cell size, bundle sheath cell size, leaf thickness and interveinal distance could theoretically occur. However, the observed range is obviously more limited (Figure 2.4), for a number of reasons. First, multiple traits might be influenced by the same gene (pleiotropy). For instance, genome size theoretically affects the size of all cells (Grime and Mowforth, 1982; Masterson, 1994; Beaulieu *et al.,* 2008; Šímová and Herben, 2012), so that an increase in bundle sheath cell size might co-occur with increases in the sizes of mesophyll cells. Plants often escape this constraint via cell-specific endoreduplication, which allows an increase of one type of cells relative to others (Sugimoto-Shirasu and Roberts, 2003), and comparative analyses show that variation in the cell sizes of different components of C_4 anatomy is only partially correlated (Table 2.1). However, endoreduplication is not involved in the increase of bundle sheath cell size, at least in the C_4 *Cleome gynandra* (Aubry *et al.*, 2013). It is also likely that some combinations of traits are not viable, as the whole leaf structure influences plant fitness (Noblin *et al*., 2008), not its individual components.

The multidimensionality of leaf characters associated with C_4 photosynthesis, as highlighted for the grass family, means that different combinations of underlying traits will generate C_4-compatible leaf anatomies (Figure 2.4). For instance, both a proliferation of veins with small bundle sheath cells and an increase of bundle sheath cell size without additional veins would increase the relative amount of bundle sheath cells (Figure 2.6). This potential for alternative anatomical combinations to achieve the same functional outcome means that C_3 ancestors will repeatedly reach C_4-compatible areas of the multidimensional trait space (Figure 2.6), and increases the likelihood of C_4 anatomy evolving (Williams *et al*., 2013). A

sample of evolutionary trajectories in the Poaceae shows lineages for which repeated and independent alterations of the distance between bundle sheaths and bundle sheath size led several taxa into different C_4-compatible regions of the anatomical space (Figure 2.6). Obviously, not all C_3 lineages that acquired C_4-suitable leaf anatomical characters have evolved C_4 biochemistry. For example, *Panicum malacotrichum* and *Oryza coarctata* have C_4-suitable mesophyll distances between consecutive bundle sheaths and proportions of bundle sheath tissue, but have not developed the C_4 syndrome (Figure 2.2; Figure 2.6; Christin *et al*., 2013a). Furthermore, *Cleome violacea*, *C. africana*, and *C. paradoxa* have small interveinal distances and *C. africana* and *C. paradoxa* also display enlarged bundle sheath cells similar to their C_4 congener *C. gynandra*, yet these three species do not employ the C_4 photosynthetic system (Marshall *et al.*, 2007). However, the presence of these characters likely enables C_4 evolution (pre-adaptation or exaptation *sensu* Gould and Vrba, 1982; Christin *et al*., 2013a; Griffiths *et al*., 2013; Sage *et al*., 2013). Once a C_4-compatible anatomy is in place, the C_4 biochemical pathway can evolve from a C_3 background in a stepwise sequence, where each step incrementally increases the efficiency of photosynthesis (Heckmann *et al*., 2013). However, the multiple anatomical requirements for C_4 photosynthesis do not usually co-occur in C_3 plants. Interesting exceptions include plants with a C_2 physiology, which were likely co-opted for the evolution of C_4 photosynthesis (Christin *et al*., 2011a; Muhaidat *et al*., 2011; Sage *et al*., 2012).

Table 2.1 Degrees of co-variation among anatomical variables.

BS distance				
0.02	OBS cell width			
ns	0.23	IBS cell width		
0.58	ns	0.05	No. M cells	
ns	0.28	0.43	0.12	Leaf thickness

Co-variation in grasses between the mean distance between consecutive bundle sheaths (μm), outer bundle sheath (OBS) cell width (μm), inner bundle sheath (IBS) cell width (μm), number of mesophyll (M) cells between consecutive bundles, and leaf thickness (μm). R^2 values are provided for pairs of variables with significant correlations. Regressions with *p*-values less than 0.05 are considered significant, while those with *p*-values greater than 0.05 are indicated by "ns". Phylogenetically controlled analyses were performed with the pgls function of the caper R package (Orme *et al.*, 2012), using the data for 155 grass species from Christin *et al.* (2013).

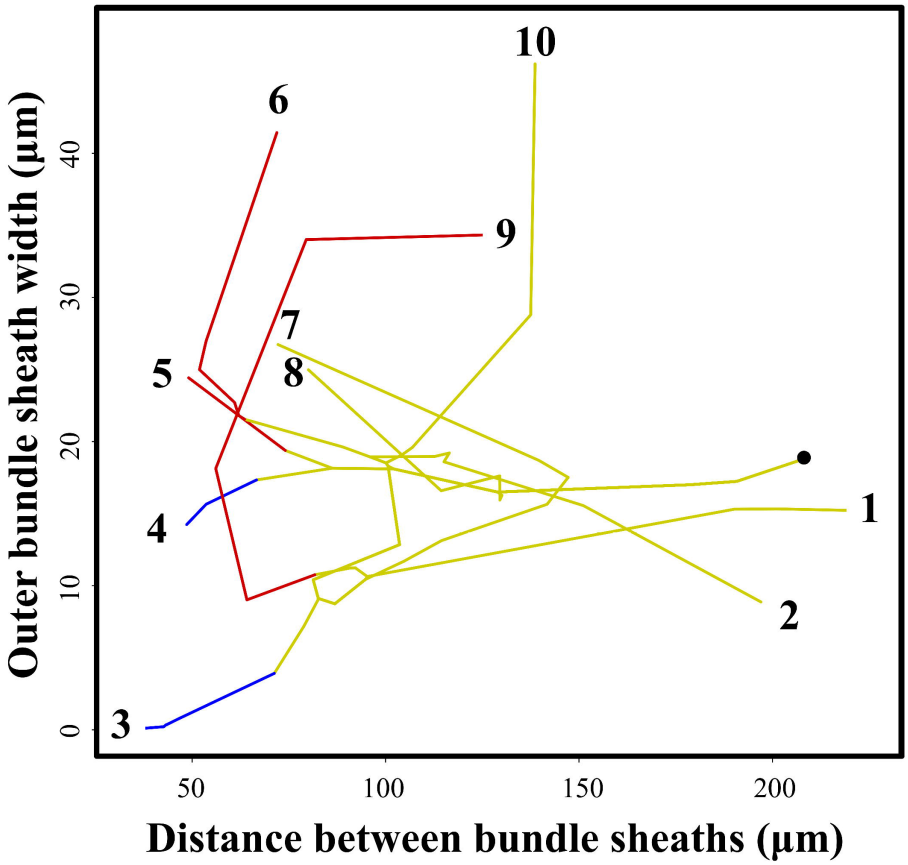


Figure 2.6 Evolutionary trajectories toward C_4-compatible anatomical traits. Phylogenetic relationships are plotted in anatomical space for grass species selected to represent a diversity of anatomical traits. Values are the distance between consecutive bundle sheaths, and the width of outer bundle sheath cells, which are observed for the tips and inferred for the internal nodes. The black point represents the root of the tree (see Christin *et al.*, 2013a for details). Yellow branches indicate a C_3 state, red branches a C_4 state using the outer sheath for the Calvin cycle, and blue branches a C_4 state using the inner sheath for the Calvin cycle. Numbers refer to the extant species: 1 = *Dichanthelium acuminatum* (C_3), 2 = *Danthonia spicata* (C_3), 3 = *Heteropogon contortus* (C_4), 4 = *Aristida congesta* (C_4), 5 = *Stipagrostis obtusa* (C_4), 6 = *Eleusine indica* (C_4), 7 = *Panicum malacotrichum* (C_3), 8 = *Oryza coarctata* (C_3), 9 = *Panicum miliaceum* (C_4), 10 = *Arundo donax* (C_3). Data come from Christin *et al.* (2013).

Functional C_4 diversity as a consequence of evolutionary diversity

Because C_4-compatible leaf anatomy engages multiple components, each C_4 origin may involve different modifications and co-opt different compartments for the Calvin cycle (Brown, 1975; Dengler *et al.*, 1994; Edwards and Voznesenskaya, 2011; Christin *et al.*, 2013a). The anatomy present in the C_3 ancestor might affect

which C_4 phenotypes are possible. For instance, C_3 ancestors with enhanced water storage tissue are likely to give rise to C_4 leaves that maintain the same capacity to store water, with the PEPC and Calvin cycle compartments occupying other parts of the leaves (Voznesenskaya *et al*., 1999; Kadereit *et al*., 2003; Freitag and Kadereit, 2013). Similarly, C_4 species that use the inner bundle sheath for the Calvin cycle must evolve from C_3 ancestors that possessed two differentiated sheaths, as is the case with grasses and sedges (Dengler *et al*., 1994; Soros and Dengler, 2001). Furthermore, similar C_4 phenotypes can be achieved through different modifications, even when starting with similar C_3 ancestors.

Different modifications to fulfil the same C_4 requirements might have functional consequences. Indeed, the adaptation of C_4 photosynthesis through the evolution of thick leaves with large bundle sheath cells (Figure 2.3f) is likely to have different consequences than the evolution of thin leaves with small cells but very short interveinal distance (Figure 2.3c). Furthermore, an increase in vein density will affect not only the hydraulics but also the distribution of stomata, which tend to be located in-between veins (Taylor *et al*., 2012). Leaf thickness will have consequences for light capture efficiency as well as ecologically meaningful traits such as specific leaf area (Wilson *et al.*, 2002). Similarly, light capture will also be affected by the different distribution of chloroplasts in mesophyll and bundle sheath cells, and the relative abundance of each cell type, together with the orientation of mesophyll cells (Vogelmann *et al.*, 1996). The path length from stomata to the photosynthetically active cells will also be influenced by leaf thickness, interveinal distance and amount of intercellular airspace (Noblin *et al*., 2008). Finally, co-opting some areas of the leaf for C_4 photosynthesis while maintaining water storage cells will likely allow the C_4 descendants to thrive in more arid conditions (Voznesenskaya *et al*., 1999; Kadereit *et al*., 2012). All of these characters, which can be directly affected by the evolutionary path a species took to achieve C_4 function, will determine the physiology of a plant and thus its ecological preferences. Therefore, the diversity of evolutionary trajectories toward C_4-compatible leaf anatomy might partially explain the ecological diversity associated with distinct C_4 lineages (*e.g.*, Taub, 2000; Kadereit *et al*., 2012; Liu *et al*., 2012).

Consequences for putative genetic determinism

A detailed discussion of genetic determinants is beyond the scope of this chapter. However, it is worth pointing out that, despite recent important developments (*e.g.,* Slewinski *et al*., 2013; Wang *et al*., 2013; Lundquist *et al*., 2014), the genetic mechanisms necessary to introduce C_4-compatible anatomy into C_3 species remain largely unknown. This has particular implications for the bioengineering of C_4 photosynthesis into major C_3 crops, such as rice and wheat, which has the potential to greatly enhance yield (Covshoff and Hibberd, 2012; von Caemmerer *et al.*, 2012). Firstly, the multiplicity of traits means that there are likely multiple genes involved. For instance, a phylogenetic analysis shows that the distance between consecutive bundle sheaths and the size of these bundle sheaths vary independently in grasses (Table 2.1), suggesting different underlying genetic changes. Secondly, since the variation in most traits presents a continuum from C_3 to C_4 plants, the determinism is likely to involve multiple genes with small effects and no master switch. Thirdly, the diversity of strategies used to achieve C_4-compatible anatomy means that genetic determinism is likely to differ among C_4 lineages. Finally, the genetic changes that occur during the evolution of C_4 photosynthesis are likely to vary as a function of the condition in the C_3 ancestor.

Interestingly, similar variation in some of the underlying traits exists in C_3 and C_4 species, which suggests that useful genetic variants may be identified from the analysis of C_3 taxa that vary in only some of the traits, even if these C_3 taxa do not present C_4-like anatomies. For instance, a C_3 taxon with variation in the number of mesophyll cells between consecutive veins would be a good study system, even if the bundle sheath and distribution of chloroplasts were not C_4-compatible. Considering variation within C_3 taxa that are unrelated to C_4 lineages might therefore expose new ways to identify the adaptive significance of individual C_4 components, as well as their genetic determinism.

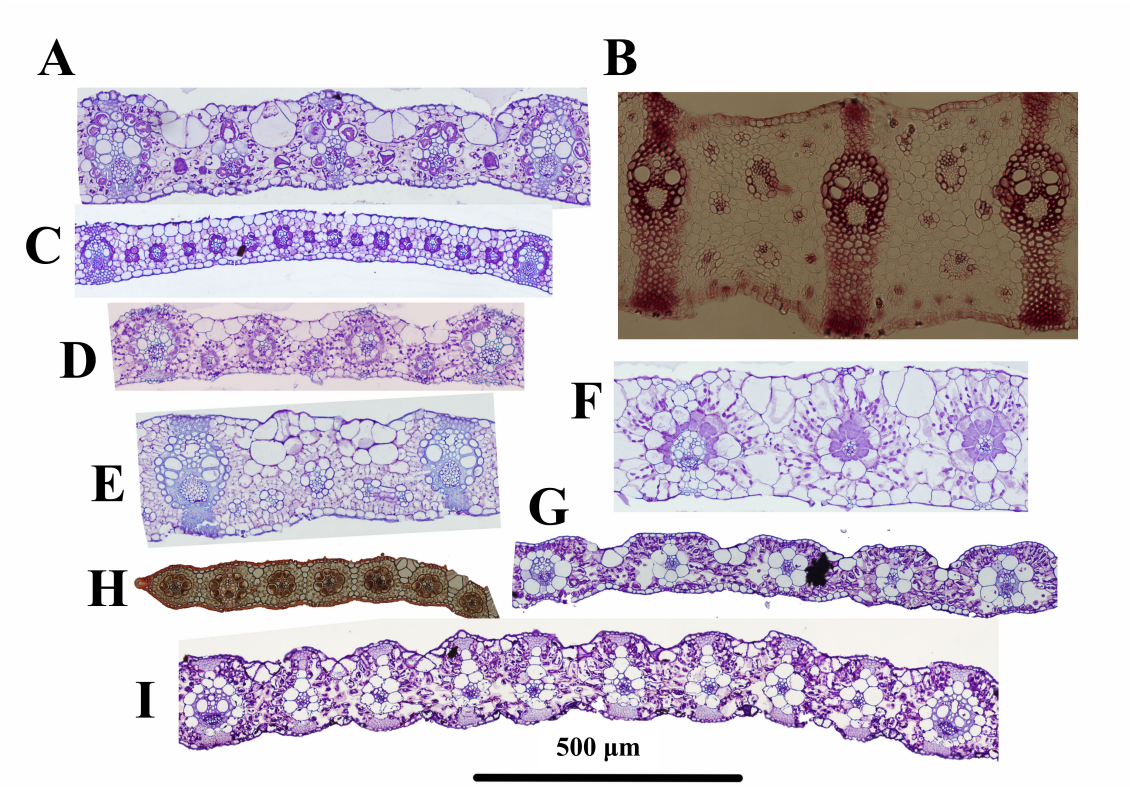
Conclusions

Overall, C_4 leaves can be defined by a set of important properties that characterize all C_4 plants. However, the underlying developmental characters that generate these functional properties are extremely variable, as a consequence of the

taxonomic diversity of C_4 plants. The same functionally important traits are not homologous among all C_4 plants, and this has important implications for the evolution and underlying genetics of C_4-specific leaf anatomy. In addition, the developmental modifications that generate each of the essential requirements of C_4 leaf anatomy can happen independently. Thus, distantly related C_4 groups might arrive at the same phenotype for one of these requirements (*e.g.*, both groups co-opt the same compartment for the Calvin cycle), but not another (*e.g.*, they achieve small distances between the two compartments through either reduction of the number of cells between veins or the development of additional veins).

Most of the anatomical characters that can generate functional properties of C_4 leaves exist in at least some C_3 plants. The only well characterised exception is chloroplast concentration in the compartment co-opted for the segregation of the Calvin cycle, which seems to be specific to C_4 plants, and to some extent C_2 plants. Without considering the distribution of chloroplasts and hence, C_4 physiology, leaves of C_3 and C_4 plants cannot be placed into mutually exclusive categories (see Figure 2.3, for example), and there is continuous variation of the underlying traits among C_4 and C_3 species (Figure 2.4). Hard categorization is meaningful from a functional perspective, but it wrongly suggests that the recurrent emergence of C_4 photosynthesis represents the same number of drastic transitions between two distinct and homogeneous character states. Acknowledging the diversity present within both C_3 and C_4 taxa, and the continuum that exists between these two physiological states, is paramount to understanding the evolutionary processes that led to C_4 plants, as well as the genetic mechanisms responsible for C_4-compatible leaf anatomy.

Appendix 2.1 Cross-sections corresponding to the diagrams shown in Figure 2.3.



A. *Arundinella nepalensis* (C_4), B. *Anthaenantia lanata* (C_4), C. *Axonopus compressus* (C_4), D. *Ischaemum afrum* (C_4), E. *Chrysopogon pallidus* (C_4), F. *Alloteropsis cimicina* (C_4), G. *Panicum pygmaemum* (C_3), H. *Bouteloua stolonifera* (C_4) and I. *Panicum malacotrichum* (C_3). All cross-sections are drawn at the same scale, indicated at the bottom. Cross-section images were obtained from Pascal-Antoine Christin; slides were prepared by Steve Renvoize (B), Travis Columbus (H), and Pascal-Antoine Christin (A, C-G, I). All cross-sections were previously published in Christin *et al.*, 2013a.

CHAPTER 3

DISCOVERY AND EVOLUTIONARY IMPLICATIONS OF C_3-C_4 INTERMEDIATES IN *ALLOTEROPSIS SEMIALATA*

Introduction

As the primary carboxylating enzyme in oxygenic photosynthetic organisms, Rubisco is the most abundant protein on Earth (Ellis, 1979; Raven, 2013), but also one ridden with inefficiencies. Specifically, it has a slow catalytic rate, a low affinity for CO_2, and an inclination toward fixing O_2, which causes O_2 and CO_2 to compete for the active site of Rubisco (Sharkey, 1985; Ehleringer *et al*., 1991; Tcherkez *et al*., 2006; reviewed in Chapter 1). The oxygenation of RuBP produces toxic compounds that have to be recycled in the photorespiratory cycle, which releases CO_2 and consumes additional energy, making this costly (Bauwe *et al*., 2010; Fernie *et al*., 2013). To get around this problem, the derived C_4 pathway relies on a CO_2 concentrating mechanism that increases the concentration of CO_2 at the active site of Rubisco, causing CO_2 to be fixed much more often than O_2 and, thus, nearly eliminating photorespiration (reviewed in Chapter 1; Chollet and Ogren, 1975; Hatch and Osmond, 1976). This is achieved through a separation of CO_2 assimilation and reduction into two compartments, which often constitute mesophyll and bundle sheath cells (Figure 3.1). In short, the enzyme carbonic anhydrase converts atmospheric CO_2 to bicarbonate, which can be fixed by PEPC in the mesophyll of C_4 plants. The resulting four-carbon organic acids are biochemically shuttled into the bundle sheath where they are decarboxylated and Rubisco can fix the released CO_2 in the Calvin cycle (Figure 3.1; reviewed in Chapter 1; Brown, 1975). This system concentrates CO_2 around Rubisco and nearly eliminates O_2 in bundle sheath cells (reviewed in Langdale, 2011). This greatly reduces the opportunity for O_2 to be fixed, causing net CO_2 assimilation to increase, and making C_4 photosynthesis more efficient than the C_3 pathway under conditions where photorespiration would be limiting (reviewed in Chapter 1).

The evolutionary path between C_3 and C_4 photosynthesis includes several anatomical and biochemical modifications that must accumulate and coordinate over time to construct the complex C_4 syndrome (Sage *et al.*, 2014). Because these modifications do not evolve simultaneously, phenotypes intermediate between C_3 and C_4 photosynthetic types exist. These intermediary states fill a gap in the selective landscape between C_3 and C_4 and, as such, are key to our understanding of the transitions that take place during C_4 evolution. Over the last 30 years, greater understanding has emerged of the C_2 cycle as the principal intermediary

step between C_3 and C_4 photosynthesis (Kennedy and Laetsch, 1974; Monson *et al.*, 1984; Monson and Rawsthorne, 2000; Heckman *et al.*, 2013; Sage *et al.*, 2013; Williams *et al.*, 2013; Mallmann *et al.*, 2014; reviewed in Chapter 1).

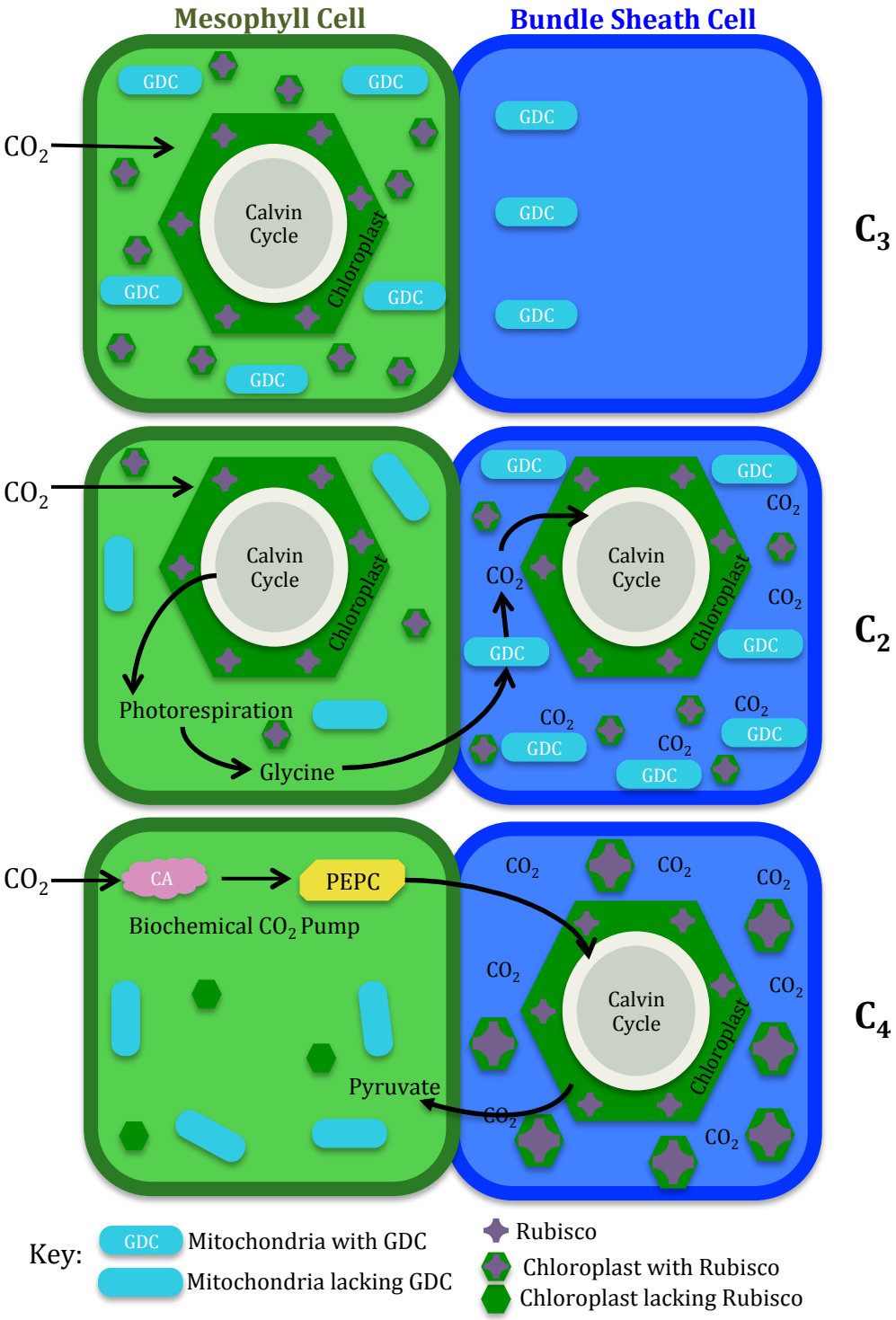


Figure 3.1 Comparison of typical C_3, C_2, and C_4 photosynthetic cycles.

Plants using a C_2 cycle, and C_3-C_4 intermediates in general, share characteristics with both C_3 and C_4 plants (Figure 3.1). C_2 plants will engage the Calvin cycle in mesophyll cells (von Caemmerer, 1989; Monson and Rawsthorne, 2000), similar to typical C_3 plants. However, C_3 plants house the majority of GDC in their mesophyll cell mitochondria while in C_2 plants, GDC is deactivated in mesophyll and activated in bundle sheath mitochondria (Figure 3.1; reviewed in Sage *et al*., 2013). Thus, when photorespiration occurs in the mesophyll cells of a C_2 plant, the resulting glycine must diffuse into the bundle sheath to be decarboxylated by GDC (Rawsthorne, 1992). This releases CO_2 in the bundle sheath, where it can be re-assimilated in the Calvin cycle there (Figure 3.1). The C_2 cycle essentially creates a CO_2 scavenging- and weak CO_2 concentrating- mechanism by rescuing and reusing the photorespired CO_2 that is shuttled into the bundle sheath (Keerberg *et al*., 2014). The C_2 cycle establishes the bundle sheath Calvin cycle and dual-compartment coordination that are characteristic of C_4 photosynthesis (Sage *et al.,* 2013).

Does <u>Alloteropsis semialata</u> have C_3-C_4 intermediates?

The occurrence of C_3 and C_4 photosynthetic types within the single grass species *Alloteropsis semialata* has been known for 40 years (Ellis, 1974a, b; Ellis, 1981; Frean *et al*., 1983a, b; Gibbs Russell, 1983). Interestingly, discoveries of anomalous $\delta^{13}C$ and leaf anatomy phenotypes emerged from these early studies (Ellis, 1981; Hattersley and Watson, 1992; reviewed in Chapter 1), that pointed to intermediate states between C_3 and C_4 photosynthesis. However, the physiological state of this intermediacy was never confirmed. Specifically, four specimens from Tanzania and Zambia presented $\delta^{13}C$ values (*i.e.,* between -21.4 and -18.5‰) intermediate between those typical of C_3 and C_4 plants (Ellis, 1981; Hattersley and Watson, 1992). Moreover, different specimens from Zambia, Tanzania, and Kenya expressed anomalous leaf anatomy that was not consistent with typical C_3 and C_4 specimens of *A. semialata* (Long and Hattersley, unpublished data; Hattersley and Watson, 1992). Yet anomalous physiology measurements have not been recorded in this species. However, previous research on this species almost exclusively compared single populations of C_3 and C_4 *A. semialata* from South Africa and/or Australia and avoided the intriguing central African zone where these anomalous

phenotypes have been found. In this chapter, multiple populations from across the species' range, including those from central and western Africa and Madagascar, as well as South Africa and Australia, are compared to characterize the photosynthetic variation that exists within *A. semialata*.

The extraordinary variation in photosynthetic phenotypes makes C_3, C_3-C_4 intermediates, and C_4 plants difficult to distinguish via any one method. Instead, plants using these different pathways are best distinguished through a combination of stable isotope, anatomy, and physiology methods (Table 3.1).

Table 3.1 Differences in stable isotope, physiology, and anatomy between photosynthetic types in the grasses, as presented in the literature.

	C_3	C_2	C_4	Taxa	Reference
Stable isotopes					
$\delta^{13}C$ range (‰)	-34 to -19; -32 to -23	-27, -25[§]	-18 to -9; -16 to -11	Survey of grasses; *Panicum milliodes*, *Neurachne minor*[§]	Cerling *et al*., 1997 Smith and Brown, 1973 Holaday and Chollet, 1984
Physiology					
CCP (µmol mol^{-1})	>50	7 to 15, 23	<4	*Triticum, Flavaria, Panicum*	Monson *et al*., 1984 Morgan and Brown, 1979 Ku and Edwards, 1978
O_2 inhibition (µmol mol^{-1})	37-63	1 to 3; 25	≅ 0	*Triticum, Flavaria, Panicum*	Monson *et al*., 1984 Ku and Edwards, 1978 Morgan and Brown, 1979
Leaf anatomy[1]					
Mesophyll cells	2 to 22	5 to 8	1 to 8	Survey of grasses	Christin *et al*., 2013a Morgan and Brown, 1979
Bundle sheath size[2]	0.15 to 0.92	0.15 to 0.24	0.69 to 1.5[3] 0.13 to 0.60[4]	Survey of grasses	Christin *et al*., 2013a

[1]Leaf anatomical traits for C_2 individuals are based on measurements of *Steinchisma decipiens*, *Steinchisma hians*, and *Homolepis aturensis*.
[2]Primary CO_2 reducing bundle sheath
[3]These species, similar to *A. semialata*, use the inner bundle sheath for photosynthetic function, so the IBS:OBS ratio is presented.
[4]These species use the outer bundle sheath for photosynthetic function, so the OBS:IBS ratio is presented.

Stable isotope discrimination as a tool to identify photosynthetic type

As described in Chapter 1, measurements of carbon isotope discrimination are a useful tool to distinguish photosynthetic types, as C_3 and C_4 biochemistry yield very different $\delta^{13}C$ signatures (Box 1.1; Smith and Epstein, 1971; Whelan *et al*., 1973; Farquhar, 1983; von Caemmerer *et al*., 2014). Some of the earliest work on the subject found a definitive gap between $\delta^{13}C$ values measured in grasses with C_4 (-16 to -11) and non-C_4 (-32 and -23) anatomy (Smith and Brown, 1973). Consistent with these findings, a study of 50 *A. semialata* specimens found that $\delta^{13}C$ signatures in specimens expressing C_4 anatomy ranged between -14 and -9‰, while these values in non-C_4 specimens ranged between -28.5 and -25‰ (Ellis, 1981). However, a more recent comprehensive review of $\delta^{13}C$ signatures in the grass family shows a continuous distribution of $\delta^{13}C$ values, with C_4 grasses between -18 and -9‰ and C_3 grasses between -34 and -19‰ (Table 3.1; Cerling *et al*., 1997), making it difficult to assign a photosynthetic pathway confidently to plants with $\delta^{13}C$ signatures at the intersection of these ranges (*e.g.,* -20 to -17).

Identification of C_3-C_4 intermediate phenotypes through stable isotope methods may also be difficult. As both C_3 and C_2 plants use Rubisco to fix atmospheric CO_2, $\delta^{13}C$ values are usually similar in these plants (Table 3.1; Holaday and Chollet, 1984; Hattersley *et al*., 1986). In fact, the photorespired CO_2 that is refixed in the C_2 cycle will inherently be enriched in ^{12}C, often causing more negative $\delta^{13}C$ values in C_2 plants than typical C_3 plants. Moreover, some plants with a C_2 cycle may also engage a partial C_4 cycle. When this happens, $\delta^{13}C$ values will be shifted higher than typical C_3 plants and will appear isotopically intermediate between C_3 and C_4 (Hattersley and Watson, 1992). Therefore, C_3-C_4 intermediates have wide-ranging $\delta^{13}C$ values, making them difficult to identify through stable isotope methods alone (Monson *et al*., 1988; von Caemmerer, 1992; Sage *et al*., 2007).

Leaf anatomy as a tool to identify photosynthetic type

As described in detail in Chapter 2, the transition from C_3 to C_4 photosynthesis often requires a rearrangement of cells and organelles to achieve the compartmentalization and biochemical coordination characteristic of the C_4 pathway (Figure 3.1; Brown, 1975). As C_3 plants house the Calvin cycle in the

mesophyll, these cells tend to have dense concentrations of chloroplasts and mitochondria along the cell periphery while there are few, if any, of these organelles in bundle sheath cells (Brown *et al.,* 1983; Stata *et al*., 2014). In contrast, to run the Calvin cycle in the bundle sheath, C_4 plants require larger and denser packing of organelles there than C_3 plants (Brown *et al*., 1983). However, many species have two bundle sheath layers and may use either layer for localization the Calvin cycle, thus there is not a consistent trend across all taxa and instead, bundle sheath phenotypes can differ with each C_4 origin (reviewed in Chapter 2). For example, the C_4 *Panicum prionitis* uses the inner bundle sheath for the Calvin cycle and, as such, has a larger inner bundle sheath and smaller outer bundle sheath then its C_3 congeners (Brown *et al.,* 1983). Similarly, *A. semialata* uses the inner bundle sheath for the Calvin cycle, so it is expected that these cells would be larger in C_4 than C_3 *A. semialata* plants.

C_3 and C_4 plants tend to differ in their mesophyll to bundle sheath area ratios (Hattersley, 1984). One way to raise the ratio of bundle sheath to mesophyll area is to increase the number of bundle sheath units by increasing vein density. Higher vein density also functions to minimize the distance that atmospheric CO_2 travels from its assimilation in the mesophyll to its reduction in the bundle sheath (Muhaidat *et al*., 2011; Christin *et al*., 2013a; Griffiths *et al*., 2013). This is often achieved by decreasing the distance between veins, either through increasing vein density, reducing the number and/or size of mesophyll cells separating bundle sheath units, or some combination of these factors (Chapter 2). Therefore, C_4 plants tend to have larger bundle sheaths, higher vein frequency, and fewer mesophylls separating veins than C_3 plants (Table 3.1; Christin *et al*., 2013a). However, as discussed in Chapter 2, there is continuous variation in nearly all C_4 component traits and most C_4-compatible trait values can occur in C_3 backgrounds, making leaf anatomy a difficult tool to definitively identify photosynthetic type.

As C_3-C_4 plants often represent an intermediary state between C_3 and C_4 photosynthetic types, they do not have a consistent and predictable gross anatomical phenotype. Most characteristics are intermediate between typical C_3 and C_4 phenotypes and depending on where each plant falls along the C_3 to C_4 spectrum, some anatomical traits may be C_4-compatable while others may be more

C_3-like (Table 3.1; Monson *et al.*, 1984). Despite this, certain leaf anatomical characteristics should help to distinguish C_3-C_4 intermediates from C_3 or C_4 plants (Bouton *et al*., 1986; Brown and Hattersley, 1989).

Plants using a C_2 cycle house the Calvin cycle in both mesophyll and bundle sheath cell chloroplasts, and bundle sheath mitochondria are required for the biochemical diffusion of CO_2 from mesophyll to bundle sheath cells. Therefore, unlike most C_3 species, these plants characteristically have a profusion of mesophyll chloroplasts as well as an abundance of chloroplasts, peroxisomes, and enlarged mitochondria in bundle sheath cells (Brown *et al.,* 1983; von Caemmerer, 1989; Monson and Rawsthorne, 2000; Stata *et al*., 2014). The C_2 grasses *Steinchisma milioides*, *S. schenckii*, and *S. decipiens*, for example, express about two, five, and seven times more chloroplasts, peroxisomes, and mitochondria in their bundle sheath, than mesophyll, cells (Brown *et al*., 1983). However, the exact distribution of chloroplasts in C_2 plants may depend on where the particular lineage falls in the C_4 evolutionary process (assuming it is progressing toward a C_4-endpoint, which is a dangerous assumption), as the loss of mesophyll chloroplasts seems to occur late in this process (Stata *et al*., 2014). Furthermore, bundle sheath cells may be enlarged in C_2 plants to accommodate this enhanced organelle content (McKown and Dengler, 2007; Muhaidat *et al*., 2011; Sage *et al*., 2011, 2013). Thus, the bundle sheath size and organelle phenotypes of C_2 plants can appear similar to closely related C_4 plants (Sage *et al*., 2014). C_2 species often have higher vein density than C_3 species and, in combination with their larger bundle sheaths, this leads to lower mesophyll to bundle sheath ratios in C_2 than C_3 plants (Ueno *et al*., 2006; McKown and Dengler, 2007; Muhaidat *et al*., 2011; Sage *et al.,* 2014).

Leaf physiology as a tool to identify photosynthetic type

Measures of leaf physiology distinguish C_3, C_3-C_4 intermediate, and C_4 photosynthetic types effectively (Table 3.1; Monson and Rawsthorne, 2000). This is mostly due to the differing effects of photorespiration on these plants and is reflected best by their CO_2 compensation point (CCP), carboxylation efficiency (CE), and O_2 inhibition (OI; Monson *et al*., 1984). The photosynthetic CCP is the CO_2 concentration at which assimilated CO_2 equals respired CO_2, causing the net CO_2

flux from the leaf to be zero. Thus, plants with higher rates of photorespiration (*e.g.,* those using C_3 photosynthesis) will have a greater CCP than those with lower rates of photorespiration (*e.g.,* those using C_4 photosynthesis; Monson *et al*., 1984).

The CE quantifies the amount of CO_2 that is assimilated per concentration of available intercellular CO_2 (C_i). In other words, CE is a product of the capacity for CO_2-fixation and the competitive interaction between CO_2 and O_2. The enzyme system of C_4 plants is more efficient at converting C_i into assimilated CO_2, particularly because of reduced photorespiration rates. Thus, C_4 plants tend to have higher CE than C_3 and C_2 plants (Ku and Edwards, 1978). However, C_2 plants with a partially engaged C_4 cycle should express intermediate, more C_4-like, CE.

Plants using a C_2 cycle may still have high rates of photorespiration like C_3 plants. However, photorespired CO_2 is more often captured and reassimilated in C_2 than C_3 plants. This reduces the overall negative effects of photorespiration on net CO_2 assimilation in C_2 plants, yielding lower CCP and higher CE in C_2 than C_3 plants (Table 3.1). Moreover, net photosynthesis rates are approximately between 30 - 35 and 20 – 25 per cent lower at ambient (21%), than low (2%), O_2 in C_3 and C_2 plants, respectively (Ku and Edwards, 1978). Thus, CCP and CE will generally be highest in C_3 plants, intermediate in C_2 plants, and lowest in C_4 plants (Table 3.1; Ku and Edwards, 1978; Morgan *et al*., 1980). Furthermore, measurements of CCP and CE at different O_2 concentrations will reflect the degree to which photorespiration is lowering net CO_2 assimilation rates (Monson *et al*., 1984). Photosynthesis in C_3 plants is inhibited by high O_2 concentrations, where rates of photorespiration are elevated. Meanwhile, photosynthesis in C_4 plants is minimally inhibited by O_2, regardless of concentration (Table 3.1; Morgan and Brown, 1979). Thus, C_3 plants are characterized by larger OI than C_4 plants and, because C_2 plants refix photorespired CO_2, the negatives effects of photorespiration are reduced, and they experience less OI than C_3 plants, but still more than C_4 plants (Ku and Edwards, 1978; Morgan *et al.,* 1980). The same pattern is seen in CE_{delta}, or the difference in CE measured at ambient and low O_2 concentrations (Monson *et al.,* 1984).

Aims of Chapter 3

Significant progress has been made over the last 30 years to understand the evolutionary transitions to C_4 photosynthesis (Monson *et al.,* 1984; Heckmann *et al*., 2013; Sage *et al*., 2014). However, studies designed to elucidate the nature of these transitions often rely on distantly related taxa, which differ in many traits unrelated to photosynthesis. These interspecific comparisons can only evaluate a limited number of steps in the C_3 to C_4 transition. Improving upon this approach, this chapter presents the photosynthetic variation that exists within a single species, the grass *Alloteropsis semialata*, to untangle the intricacies of C_4 evolution. Moreover, this chapter aims to decipher how the small anatomical and biochemical variations that drive these transitions can yield physiological benefits that explain the selective advantage of intermediate photosynthetic states. This photosynthetic variation is then used to characterize how structural reorganization corresponds to the physiological variation seen in this species. This information is used to extrapolate the possible evolutionary trajectory linking C_3 and C_4 photosynthetic types in this species.

Methods

Plant material

Twenty-two accessions of *A. semialata* from 11 different populations across *A. semialata's* geographic range were chosen for this study (Figure 3.2; Table 3.2). Previous work suggested populations of *A. semialata* in western Africa, Asia, and Australia would use C_4 photosynthesis, while those from Madagascar and south-eastern Africa may use either C_3 or C_4 photosynthesis (Ellis, 1981). Therefore, accessions from across these different regions were chosen to comprehensively characterize the photosynthetic variation in this species.

Accessions from Australia ("AUS"; seedbank, AusTRCF 322458), Burkina Faso ("BF"; seedbank, MSB 0506063), and Madagascar ("MAJUNGA", collected by G. Besnard, CNRS, Université Paul Sabatier, France) were germinated from seed (Table 3.2). These seeds were sterilized in EtOH (60%; 2 minutes) and bleach (30%; 5 minutes), distributed onto sterile agar plates, and germinated at 30°C,

high light (4LS), and 65% humidity in a controlled environment plant chamber (Sanyo, Bensenville, IL, USA).

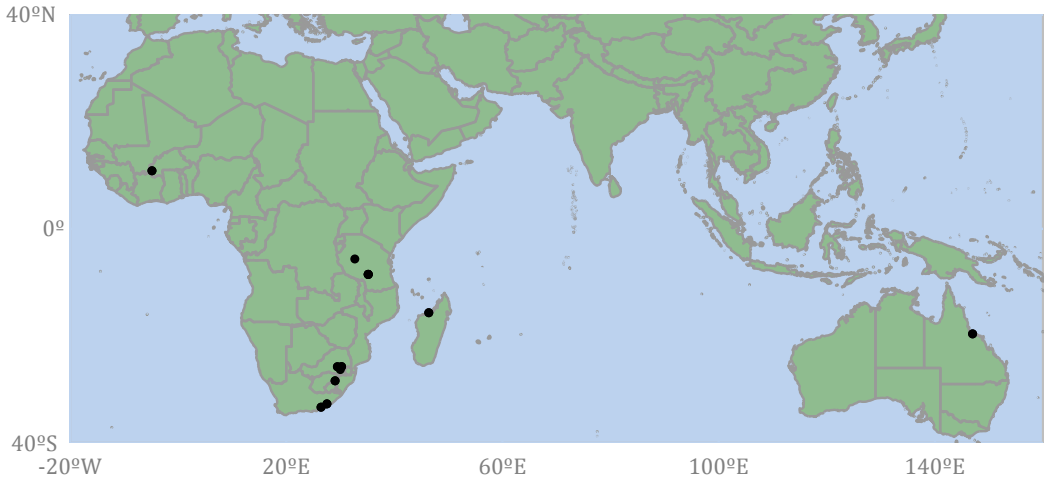


Figure 3.2 Geographic locations from which populations of *Alloteropsis semialata* presented in this study were originally collected in the wild.

Table 3.2 Details of *Alloteropsis semialata* accessions included in this study.

Accession	Country	Latitude	Longitude	Source
AUS-3	Australia	-19.62	146.96	Germplasm
AUS-4	Australia	-19.62	146.96	Germplasm
MAJUNGA-1	Madagascar	-15.67	46.37	Germplasm
MAJUNGA-3	Madagascar	-15.67	46.37	Germplasm
BF-2	Burkina Faso	10.85	-4.83	Germplasm
BF-3	Burkina Faso	10.85	-4.83	Germplasm
MDG-1	South Africa	-25.76	29.47	Field cutting
MDG-2	South Africa	-25.76	29.47	Field cutting
SFD-1	South Africa	-28.39	29.04	Field cutting
SFD-3	South Africa	-28.39	29.04	Field cutting
CRL-2	South Africa	-25.74	30.24	Field cutting
CRL-4-5	South Africa	-25.74	30.24	Field cutting
EML-11-3	South Africa	-26.29	30.00	Field cutting
JMS-201	South Africa	-33.32	26.44	Field cutting
JMS-202	South Africa	-33.32	26.44	Field cutting
KWT-3-5	South Africa	-32.70	27.53	Field cutting
KWT-5-4	South Africa	-32.70	27.53	Field cutting
LO1-A	Tanzania	-5.63	32.69	Field cutting
LO4-A	Tanzania	-8.51	35.17	Field cutting
LO4-B	Tanzania	-8.51	35.17	Field cutting
LO4-D	Tanzania	-8.51	35.17	Field cutting
LO4-E	Tanzania	-8.51	35.17	Field cutting

Latitude and longitude are in decimal degrees.

For populations that were difficult to grow from seed, live cuttings were collected in the field. A comprehensive review of *A. semialata* herbarium specimens indicated specific locations where this species had been found in South Africa in the past. These localities were targeted during a collection trip across eastern South Africa in March 2012 (Figure 3.3a) that provided all the MDG, SFD, CRL, EML, JMS, and KWT accessions used here (Table 3.2). Herbarium records were also used to locate potentially anomalous populations. A review of $\delta^{13}C$ values measured on herbarium specimens showed that isotopically intermediate *A. semialata* plants had been collected in Tanzania and Zambia in the past (Ellis, 1981; Gibbs Russell, 1983; Long and Hattersley, unpublished data; Hattersley and Watson, 1992). A collecting trip to Tanzania was organized in January 2014 (Figure 3.3b), that yielded the L01 and L04 accessions included in this study (Table 3.2). Live tillers collected from South Africa and Tanzania were transported to the University of Sheffield's Sir David Read Controlled Environment Facility for re-establishment.

All adult plants were grown in controlled environment growth chambers (CEGC; Conviron Ltd, Winnipeg, Canada) at the University of Sheffield until stable isotope, physiology, and anatomy data were collected in May 2014. Growth conditions were controlled at 25°C day/20°C night temperatures, 500 μmol m^{-2} s^{-1} photosynthetic photon flux density (PPFD), 60% relative humidity, and ambient CO_2 concentrations over 14 hour day/10 hour night cycles. Plants were grown in John Innes No. 2 or 3 compost (John Innes Manufacturers Association, Reading, England), maintained at well-watered soil water conditions, and fertilized biweekly with Scotts Evergreen Lawn Food (The Scotts Company, Surrey, England).

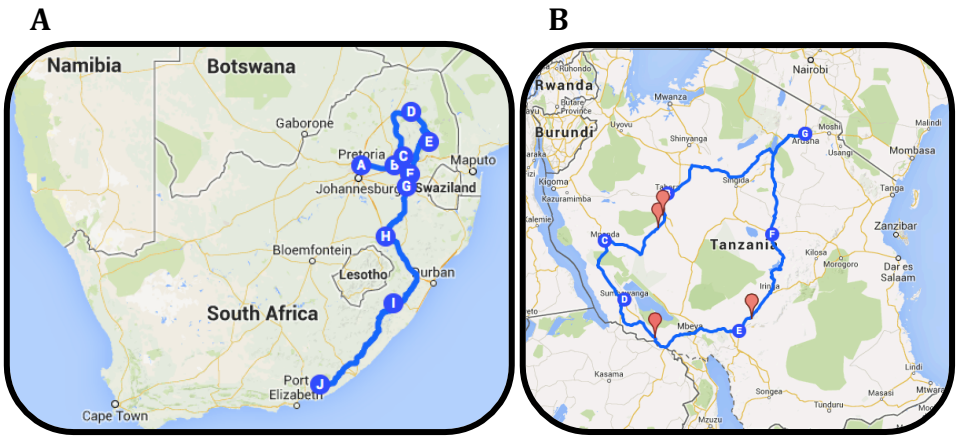


Figure 3.3 Routes taken during *Alloteropsis semialata* collecting trips to South Africa (A) and Tanzania (B). South African map data from 2015 AfriGIS (Pty) Lty, Google (www.google.co.uk/maps/@28.0341155,26.7381432,6z?hl=en) and Tanzanian data from Google 2015 (www.google.co.uk/maps/@6.2854704,35.507717,6z?hl=en).

Ploidy and genome size variation

To determine ploidy level for each population, chromosome number was determined from root squashes. Healthy root tips were harvested from well-watered plants four hours into the photoperiod to maximize the number of cells entering metaphase. After cleaning off soil debris, roots were pre-treated in 8-hydroxyquinoline for three hours to arrest cells in mitotic metaphase and to contract chromosomes for easier counting. Roots were then fixed in O'mara's fixative, a solution of absolute methanol, chloroform, and propionic acid in 6:3:2 ratio for 48 hours. After this time, the fixative solution was washed off thoroughly in distilled water, and roots were hydrolysed for 10 minutes in 1M hydrochloric acid at 60°C for 10 minutes to remove purine bases from the DNA and macerate the root tissue. Hydrolysed roots were then stained in Schiff's reagent for at least 30 minutes before being dissected and squashed in the counterstain propiono-orcein to enhance contrast. Squashed roots were viewed and imaged at 40X or 100X magnification using microscopy imaging software and a camera mounted on a microscope (CellA; Olympus DP71; BX51, respectively. Olympus, Hamburg, Germany). Genome size measurements were made by Dr. Ilia Leitch and Dr. Oriane Hidalgo at the Jodrell Laboratory at Kew Royal Botanical Gardens.

Leaf composition and stable isotope dataset

The youngest fully expanded leaf from one tiller on each plant was used to quantify leaf composition and stable isotope traits. A chlorophyll meter (SPAD 502, Spectrum Technologies, Inc. Illinois, USA) was used to measure the SPAD value, a rapid and reasonably accurate estimate of leaf chlorophyll content (Markwell *et al*., 1995), along three spaced points of this fresh and turgid leaf. Leaves were then detached and scanned on a desktop scanner and the leaf area was measured using ImageJ software (Schneider *et al*., 2012). Then leaves were dried in an oven and the dry masses weighed. Specific leaf area was calculated as leaf area (cm^2) per mass (g).

Dried leaves were milled into a homogeneous powder and 1-2 mg subsamples of this product were analyzed using an ANCA GSL preparation module coupled to a 20–20 stable isotope analyzer (PDZ Europa, Cheshire, UK) by the University of Sheffield, Faculty of Science biOMICS facility. Analyses by this mass spectrometer yielded the percent carbon and nitrogen content of the leaf, the carbon to nitrogen ratio, and the $\delta^{15}N$ and $\delta^{13}C$ isotopic signatures. $\delta^{13}C$ and $\delta^{15}N$ data are presented as isotopic ratios (per mil, ‰), relative to the isotopic standard Pee Dee Belemnite (PDB) for carbon and atmospheric N_2 for nitrogen (Equations 3.1 and 3.2).

Equation 3.1

$$\delta^{13}C = ([(^{13}C/^{12}C)_{sample} / (^{13}C/^{12}C)_{standard}] -1) \times 1000$$

Equation 3.2

$$\delta^{15}N = ([(^{15}N/^{14}N)_{sample} - (^{15}N/^{14}N)_{standard}] /(^{15}N/^{14}N)_{standard}) \times 1000$$

Because variations in $\delta^{13}C$ result from both source isotopic compositions and carbon isotope discriminations, measure on plants grown in CEGC may be different from natural populations (Farquhar *et al*., 1989). Ambient air surrounding the controlled environment growth facilities at the University of Sheffield, and in most cities, is enriched with fossil CO_2 released from anthropogenic sources. Thus, plants grow in an isotopically light atmosphere in these facilities, with lower proportions of ^{13}C than wild populations would experience. This causes $\delta^{13}C$ values to be shifted in these controlled environment grown plants, such that they appear

to discriminate more against ^{13}C than they would in the wild. This chapter presents $\delta^{13}C$ data for both CEGC and field grown plants. For CEGC measurements, leaves were collected for $\delta^{13}C$ analysis at the same time and on the same plant as physiology and anatomy data collections. Thus, for all comparisons between $\delta^{13}C$, physiology, and anatomy in this chapter, the CEGC $\delta^{13}C$ values are used. Field values are only presented to show the range of natural $\delta^{13}C$ values observed in this species.

For most accessions, field values of $\delta^{13}C$ were measured on silica-dried leaves that were collected from the same population, and often the same plant, growing in the wild. These equivalent field values were available for the CRL, EML, JMS, KWT, L01, and L04 accessions. For accessions without equivalent field material available, $\delta^{13}C$ values were adjusted based on a known difference between field and CEGC plants measured in the same population. CEGC $\delta^{13}C$ values for C_4 plants MDG-1 and MDG-2 were -13.91‰ and -13.89‰, respectively. $\delta^{13}C$ values from ten field grown replicate plants from the same MDG population averaged -11.10‰ (Ibrahim, 2007). Thus, the average shift in $\delta^{13}C$ between CEGC and field grown plants was 2.66‰. CEGC $\delta^{13}C$ values in C_4 accessions lacking field data were consequently increased by 2.66‰, to yield an estimate of the field $\delta^{13}C$ value for comparison.

Leaf anatomy

Tissue was collected from the centre portion of the leaf blade on a youngest fully expanded leaf of each plant. Leaf pieces 3-5 mm in length were fixed in Carnoy's fixative (4:1 EtOH:acetic acid) and then embedded in Methacrylate embedding resin (Technovit 7100, Heraeus Kulzer GmbH, Wehrhein, Germany). Leaves were sectioned between 6-8 µm thick on a manual rotary microtome (Leica Biosystems, Newcastle, UK) and stained with Toluidine Blue O (Sigma-Aldrich, St. Louis, MO, USA). Stained leaf sections were photographed using microscopy imaging software and a camera mounted on a microscope (CellA; Olympus DP71; BX51, respectively. Olympus, Hamburg, Germany). Images were stitched together (DoubleTake, Echo One, Frederikssund, Denmark) to recreate the continuous width of the whole leaf blade in cross-section.

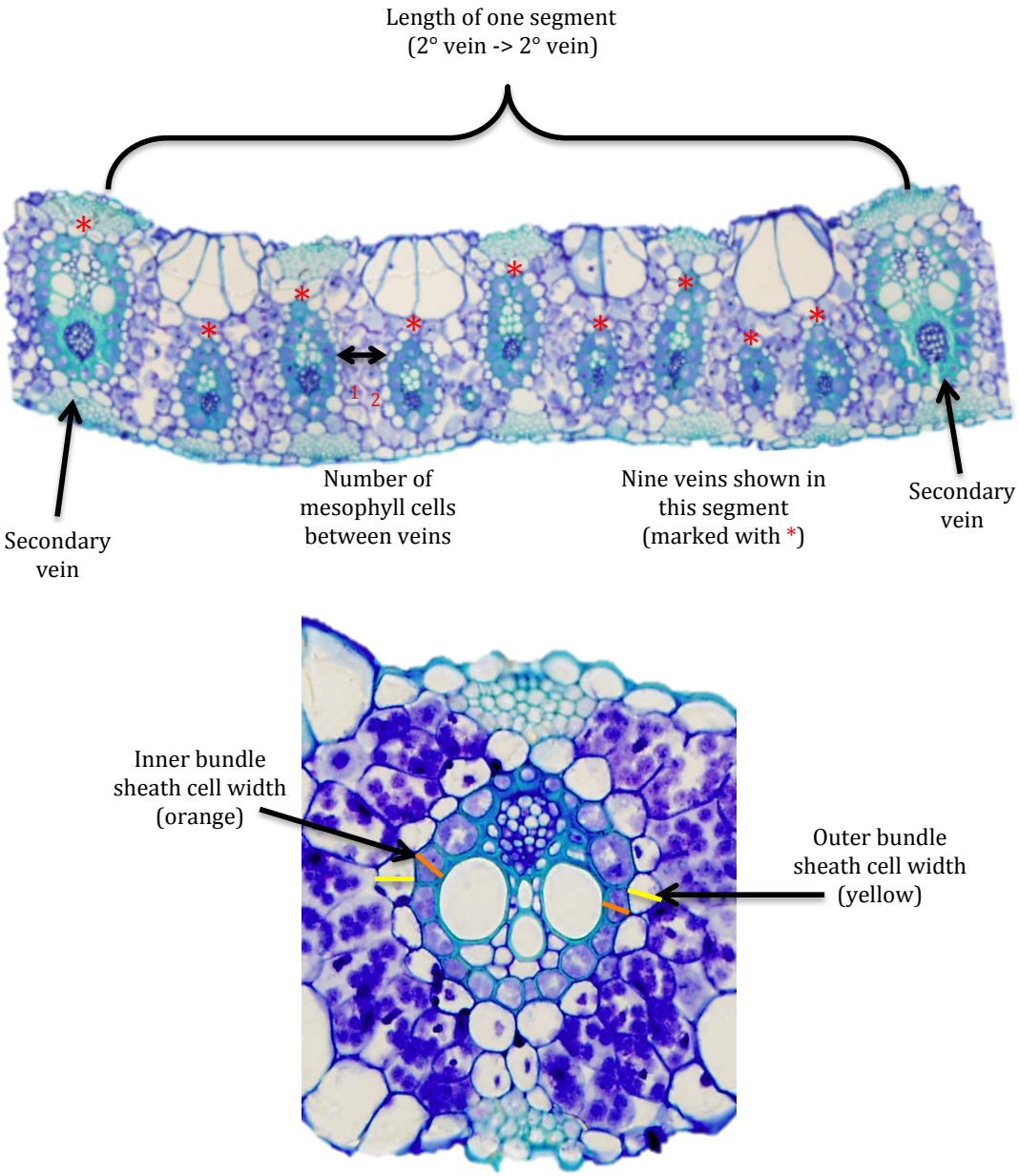


Figure 3.4 Demonstration of leaf anatomy measurements collected on *Alloteropsis semialata* cross sections. The top image shows a single segment, defined as the leaf portion between two secondary veins. The number of veins per segment was counted for three segments per leaf. The minimum number of adjacent mesophyll cells separating vein pairs was counted for ten vein pairs per leaf. The bottom image shows bundle sheath measurements. The width of one representative OBS (yellow) and one IBS (orange) on both sides of three primary veins were measured and the IBS:OBS cell width ratio was calculated.

Anatomical traits were measured using ImageJ software as shown in Figure 3.4. Only segments in the middle of the leaf blade were used and those immediately adjacent to the mid-rib and lateral tips of each cross-section were avoided. A single "segment" was considered the leaf portion between two 2nd order veins. These secondary veins are large and contain metaxylem while minor veins (*e.g.,* tertiary, quaternary, and quinary veins) lack metaxylem. Of the veins that lacked metaxylem, those with bundle sheath extensions were categorized as tertiary veins and those lacking these extensions were either quanternary or quinary veins, distinguished by distinctive size classes. The number of veins per segment was counted and averaged along 3 segments per leaf. The minimum number of adjacent mesophyll cells separating vein pairs was counted for ten pairs of veins per leaf and averaged. To maintain consistency, the representative bundle sheath cells used for measurements were those that fell parallel to the leaf surface, as often as possible (Figure 3.4). The width of one of these outer bundle sheath cells (OBS) and one inner bundle sheath cell (IBS) on both sides of a secondary vein were measured and the ratio of IBS to OBS cell widths was calculated and averaged for three secondary veins per leaf. Images were also captured of fresh hand cut cross sections of *A. semialata* accessions to show chloroplast localization.

Physiology

All leaf physiology measurements were collected between two and seven hours into the photoperiod on each plant's youngest fully expanded leaf using an open system, incorporating an infra-red gas analyser (IRGA) with leaf chamber fluorometer attachment (LI-6400XT and LI-6400-40, respectively; LICOR, Lincoln, Nebraska, USA).

Rapid light response curves were collected at 27°C leaf temperature, 400 μmol CO_2 mol air^{-1} reference CO_2 concentration, ambient relative humidity (approximately 65%), and 250 μmol s^{-1} flow rate. First, leaves were acclimated to high light, then a series of measurements were collected with decreasing light intensities of 2000, 1750, 1500, 1250, 1000, 750, 500, 250, 100, 50, 25, 0 PPFD. Measurements were made when gas exchange values stabilized, after 2-3 minutes of acclimation to each light level. Light use efficiency (LUE), light compensation point (LCP), and dark

respiration (R_{dark}) were derived from these curves. LUE is calculated as the slope of the light-limited linear portion of the curve and LCP is the PPFD where net photosynthetic assimilation equals zero (Figure 3.5a). R_{dark} is the absolute value of the CO_2 assimilation rate where PPFD equals zero (Figure 3.5a).

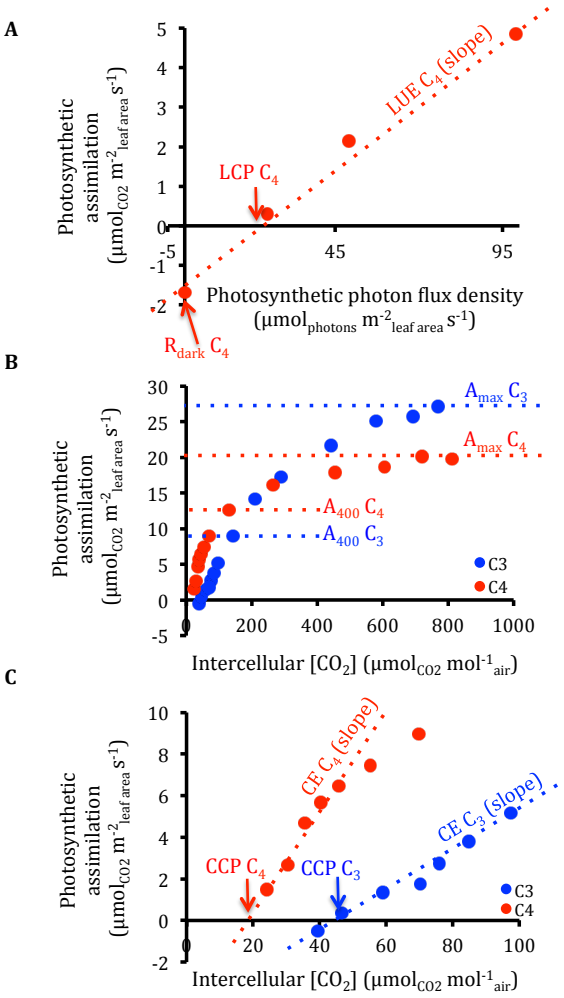


Figure 3.5 Example light response (A) and A/C_i (B and C) curves demonstrating the derivation of physiological parameters. Plot A shows the light-limited portion of the light response curve where the light use efficiency (LUE) is the slope of the linear portion of this curve and the light compensation point (LCP) is the PPFD where net photosynthetic assimilation equals zero. The dark respiration rate (R_{dark}) is calculated as the negative CO_2 assimilation rate where PPFD equals zero. The absolute value of R_{dark} is presented in this chapter. Plot B shows a whole A/C_i curve for C_3 and C_4 individuals showing how A_{max} represents light and CO_2 saturating photosynthetic rates and A_{400} represents ambient CO_2 and light saturating photosynthetic rates. Plot C shows the CO_2 limited portion of an A/C_i curve, from which the carboxylation efficiency (CE) is calculated as the slope of the linear portion of this curve and the CO_2 compensation point (CCP) is calculated as the intercellular CO_2 concentration where net photosynthetic assimilation equals zero.

Curves measuring the photosynthetic assimilation response to intercellular CO_2 concentrations (A/C_i curves) were generated at ambient (21%) and low (2%) oxygen concentrations. All A/C_i measurements were collected at light saturation (1600 μmol m^{-2} s^{-1} PPFD), 27°C leaf temperature, ambient relative humidity (approximately 70%), and 250 μmol s^{-1} flow rate. At ambient O_2, full A/Ci curves were collected at reference CO_2 concentrations of 1200, 1000, 800, 600, 400, 250, 150, 120, 100, 85, 70, 50, and 35 μmol mol^{-1} using an A/C_i auto-program with LICOR recommended stability settings. Measurements started at ambient CO_2 levels (400 μmol mol^{-1}) after sufficient acclimation to chamber conditions, then continued at sequentially lower reference CO_2 concentrations. After a measurement was collected at 35 μmol mol^{-1} (the lowest C_i value that the Licor can maintain stably), a second measurement was collected at 400 μmol mol^{-1} to check that no reduction in rate had been caused by exposure to low C_i (*e.g.,* through Rubisco deactivation), then the curve proceeded to measure at sequentially higher CO_2 concentrations and was stopped after the 1200 μmol mol^{-1} reference CO_2 measurement.

The maximum rates of photosynthesis (A_{max}) presented in this chapter were taken from the top of these A/C_i curves, where light and CO_2 were saturating, but before any triose phosphate utilization (TPU) limitation was evident (Figure 3.5b). This chapter also presents measures of photosynthetic rate at ambient CO_2 concentrations (A_{400}), the ratio of intercellular to ambient CO_2 concentrations (C_i/C_a), values of stomatal conductance to water (g_s), and intrinsic water use efficiency ($_i$WUE; calculated as A_{400}/transpiration). These five parameters were taken from the initial data point of these curves, collected at 400 μmol mol^{-1} and light saturation (Figure 3.5b).

At low O_2 concentrations (2%), A/C_i curves were only collected at reference CO_2 concentrations of 400, 250, 150, 120, 100, 85, 70, 50, and 35 μmol mol^{-1} using an A/C_i auto-program with LICOR recommended stability settings. Low O_2 A/C_i curves were performed after supplying the gas exchange machine with 2% O_2 and recalibrating the IRGA for low O_2 measurements. Measurements started at ambient CO_2 levels (400 μmol mol^{-1}) after sufficient acclimation to light saturation, then

continued at sequentially lower reference CO_2 concentrations, stopping at 35 μmol mol^{-1} reference CO_2.

The CE and CCP were calculated from A/C_i curves collected at ambient and low O_2. CE was derived from the slope of the linear portion of the A/C_i curve, while CCP was calculated as the intercellular CO_2 concentration where net photosynthetic assimilation equals zero (Figure 3.5c). O_2 inhibition was calculated as the difference between CCP measured at ambient and low O_2 concentrations, while CE_{delta} is the difference in CE measured at these two O_2 concentrations.

Statistical analysis

Correlation analyses were performed using the cor.test function in R (version 3.1.0, R Development Core Team, 2014) to test whether anatomical traits were correlated with one another. General linear models were performed in R to characterize the relationship between leaf component, anatomy, and physiology traits and the $\delta^{13}C$ gradient. Parameters showing significant relationships with $\delta^{13}C$ at $p<0.05$ level were included in a principal components and hierarchical clustering analyses to test whether the variation across 22 accessions naturally clustered into groups consistent with expected photosynthetic types. The following traits were included in this principal components analysis (PCA): $\delta^{13}C$, CCP, OI, CE, A_{max}, vein frequency, minimum number of mesophyll cells separating veins, outer bundle sheath cell size, and the IBS:OBS ratio. All PCA and hierarchical clustering analyses were performed using the FACTOMINER package (Lê *et al*., 2008) in R. A second set of linear model analyses were performed to explain how leaf physiology varied with anatomical gradients. A PCA was performed on all measured physiology traits and overlaid with anatomy and leaf component traits to aid interpretation.

Results

Stable isotope ratios

Alloteropsis semialata accessions included in this study had $\delta^{13}C$ values ranging from -28.3 to -11.0‰, spanning both C_3 and C_4 values (-33.4 and -13.7‰, in controlled environment grown plants; Figure 3.6a; Table 3.3). Ten accessions had $\delta^{13}C$ values within the typical C_4 range (referred to as C_4 accessions from this point onward; Table 3.1 and 3.3). These were from Australia, Madagascar, Burkina Faso, and South Africa. The remaining seven accessions from South Africa and all five Tanzanian accessions had $\delta^{13}C$ values in the C_3/C_2 range (referred to as non-C_4 accessions from this point onward; Table 3.1 and 3.3). There was an obvious gap between -23 and -13‰, with C_4 accessions ranging between -12.6 and -11.0‰ while non-C_4 accessions between -28.3 and -23.1‰ (Figure 3.6a; Table 3.3).

While there was a considerable amount of variation in %C, %N, C:N, SPAD, and $\delta^{15}N$, none of these parameters correlated with $\delta^{13}C$ (Figure 3.6b-f), suggesting that these traits did not differ with photosynthetic type. Interestingly, the five Tanzanian accessions expressed higher %N than the other accessions. However, these plants were initially transplanted into John Innes No. 3 soil media, a higher nutrient blend than John Innes No. 2, which may account for the high leaf nitrogen content. Indeed, when %N was measured in these plants after they were transplanted into No. 2 soil media, the %N was consistent with the other populations.

Cytology

Most of the populations included in this study were diploid (2n=2x=18; Figure 3.7; Table 3.3). However, the South African C_4 populations (SFD and MDG) were hexaploid (2n=6x=54; Table 3.3). Genome size was still variable across the diploid populations (Figure 3.7; Table 3.3). The hexaploid genome size was more than twice, but less than three times, the size of the diploid genomes (Table 3.3). This suggests that the polyploids are undergoing diploidization.

Table 3.3 Stable isotope ratio, ploidy, and genome size listed for each accession.

ID	Cluster	Plant	Country	$\delta^{13}C^{C}$	$\delta^{13}C^{F}$	Ploidy	Genome size	Type
1	2	CRL-2	South Africa	-27.8	-27.6	Diploid		C_3
2	2	CRL-4-5	South Africa	-28.3	-27.6			C_3
3	2	EML-11-3	South Africa	-29.1	-27.3	Diploid		C_3
4	2	JMS-201	South Africa	-30.8	-28.3	Diploid	1.84±0.02	C_3
5	2	JMS-202	South Africa	-27.9	-28.3			C_3
6	2	KWT-3-5	South Africa	-28.7	-27.1	Diploid		C_3
7	2	KWT-5-4	South Africa	-33.4	-27.1			C_3
8	3	LO1-A	Tanzania	-26.5	-26.3	Diploid	2.24±0.03	C_3-C_4
9	3	LO4-A	Tanzania	-25.3	-23.1			C_3-C_4
10	3	LO4-B	Tanzania	-25.6	-23.9	Diploid	1.92±0.01	C_3-C_4
11	3	LO4-D	Tanzania	-26.9	-26.4			C_3-C_4
12	3	LO4-E	Tanzania	-28.3	-25.0			C_3-C_4
13	1	AUS-3	Australia	-14.3	-11.6 *	Diploid	2.25±0.00	C_4
14	1	AUS-4	Australia	-15.3	-12.6*			C_4
15	1	BF-2	Burkina Faso	-13.9	-11.2 *	Diploid	1.99±0.02	C_4
16	1	BF-3	Burkina Faso	-14.1	-11.4 *			C_4
17	1	MAJUNGA-1	Madagascar	-14.2	-11.6 *	Diploid	2.10±0.01	C_4
18	1	MAJUNGA-3	Madagascar	-14.7	-12.0 *			C_4
19	1	MDG-1	South Africa	-13.9	-11.1	Hexaploid	5.34±0.02	C_4
20	1	MDG-2	South Africa	-13.9	-11.1			C_4
21	1	SFD-1	South Africa	-14.1	-11.5*	Hexaploid		C_4
22	1	SFD-3	South Africa	-13.7	-11.0 *			C_4

Accession identifiers (ID) also used to label accessions in the PCA analyses associated with Figures 3.14 and 3.15. Clusters assigned to each accession in the hierarchical cluster analysis associated with Figure 3.15. $\delta^{13}C$ measured on plants grown under controlled environment (C) or (F) conditions, (in ‰). For accessions without field $\delta^{13}C$ data available (*), controlled environment values were adjusted as described in the method section.

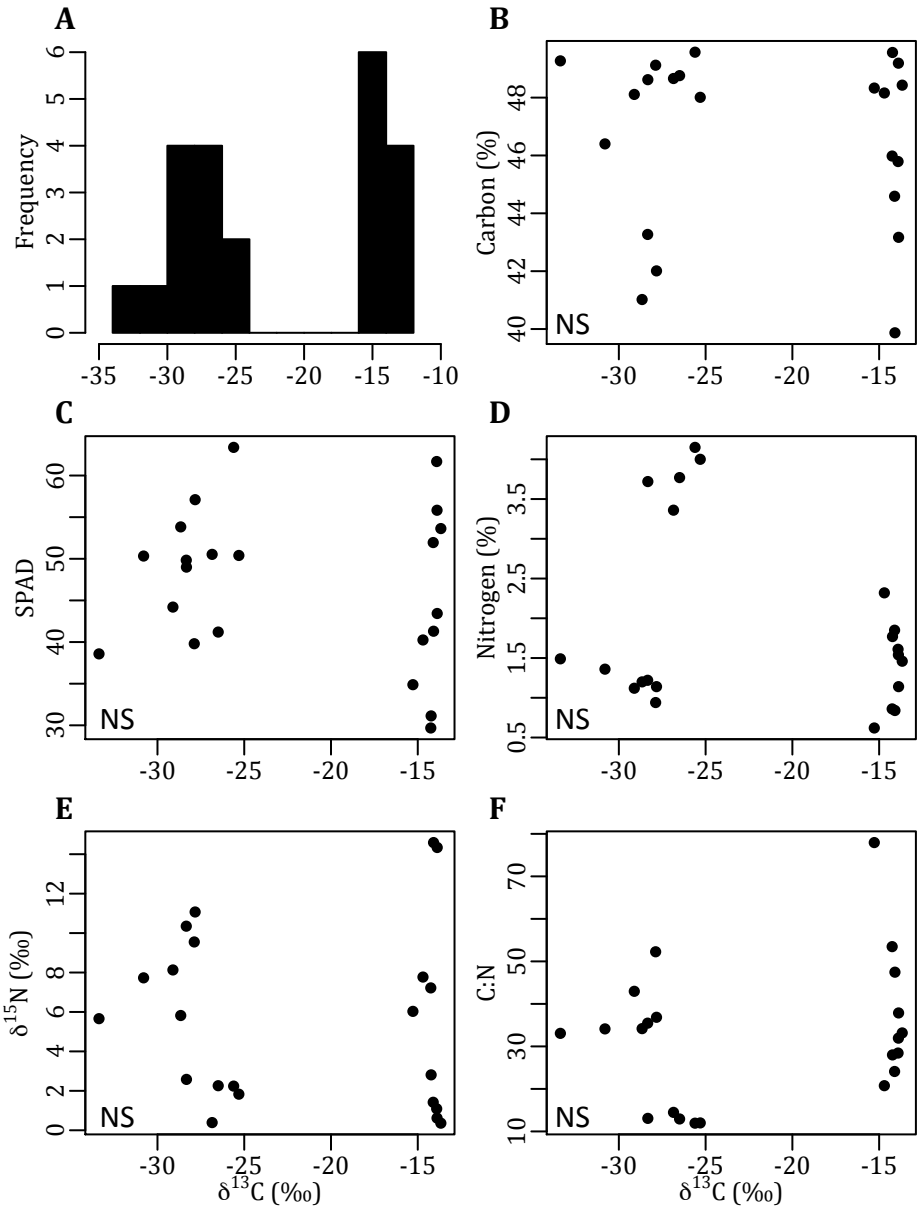


Figure 3.6 Variation in leaf composition traits. Plots show frequency distribution of $\delta^{13}C$ values (A) and how the percentage of carbon (B) and nitrogen (D), the carbon to nitrogen ratio in leaf tissue (F), SPAD (C), and $\delta^{15}N$ (E) varied across this $\delta^{13}C$ range. All relationships were not significant (NS) at the $p<0.05$ level in linear model analyses. N=22.

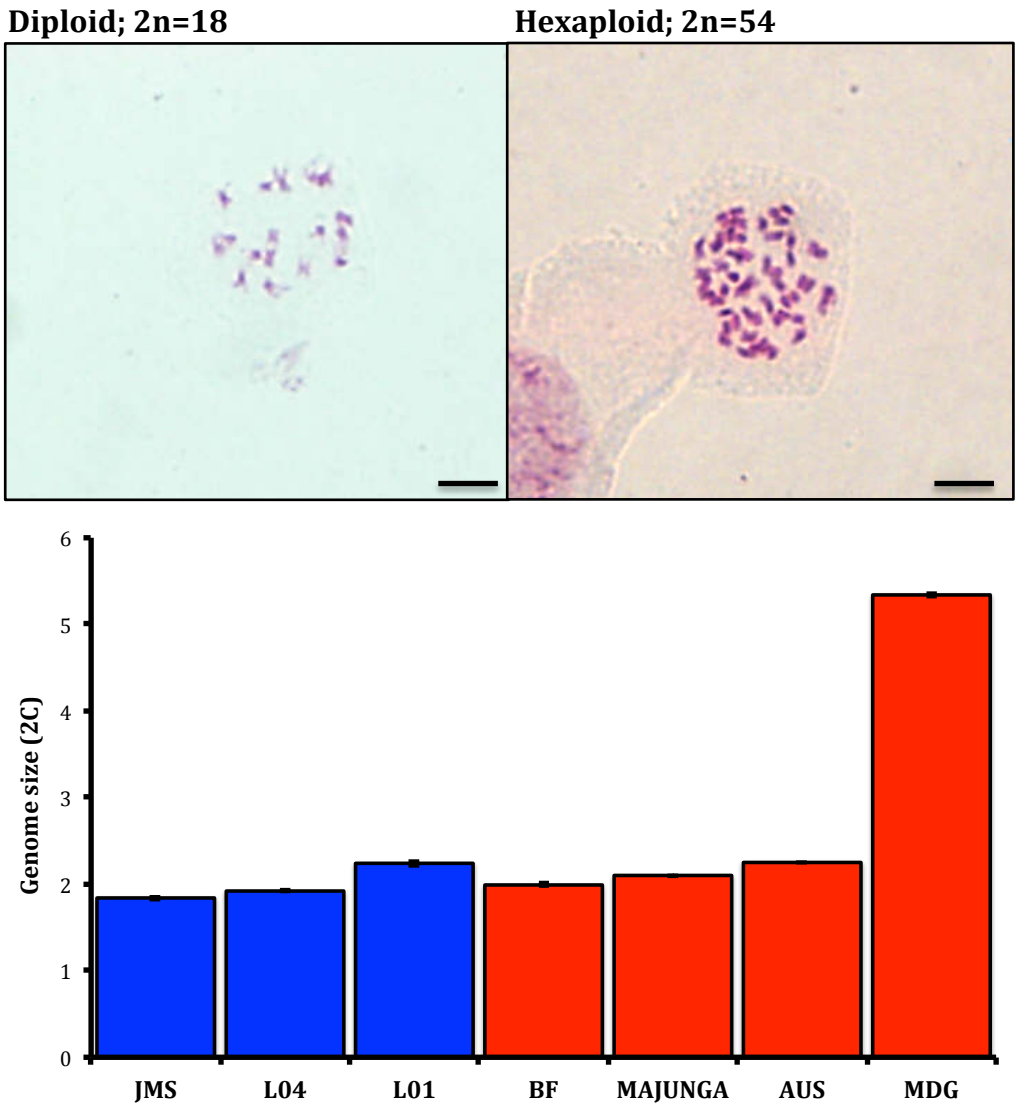


Figure 3.7 Cytology and genome sizes of *Alloteropsis semialata*. Example images of *A. semialata* diploid (top left) and hexaploid (top right) chromosome sets. Genome size (2C) data are presented for a single individual from seven different populations (bottom). C_4 populations are represented in red and non-C_4 populations in blue.

Leaf anatomy

Leaf anatomy varied widely across *A. semialata* accessions (Figure 3.8). There were between two and twelve veins per segment and between one and nine mesophyll cells separating these veins. The IBS cells were between two-thirds smaller than, and one third larger than, OBS cells. Moreover, variation in these anatomical traits correlated well with $\delta^{13}C$ signatures, as C_4 accessions had fewer mesophyll cells, more veins per segment, smaller OBS cell widths and higher IBS:OBS ratios than non-C_4-like accessions (Figure 3.9). Non-C_4 accessions had

quite variable anatomies that showed progression toward a more C_4-like phenotype with increasing $\delta^{13}C$ signatures (Figure 3.9). This trend is particularly evident in the number of mesophyll cells between veins, OBS cell width, number of veins per segment, and the IBS:OBS ratio (Figure 3.9 b, c, d, and e). IBS cell size presented a different trend across the $\delta^{13}C$ gradient as non-C_4 plants expressed two discrete IBS phenotypes (Figure 3.9a). IBS cell widths were larger than 10µm in six accessions but smaller than 8µm in the remaining non-C_4 plants, creating a distinctive gap between 8 and 10µm. These six accessions with large IBS, and consequently high IBS:OBS ratios, were within the C_4 range for these traits. The six data points came from the five Tanzanian accessions and one South African accession (CRL-4-5). SLA was conserved across the $\delta^{13}C$ gradient, however a significant effect is seen because of one particularly high SLA value in a South African C_3 plant (KWT-5-4; Figure 3.9f).

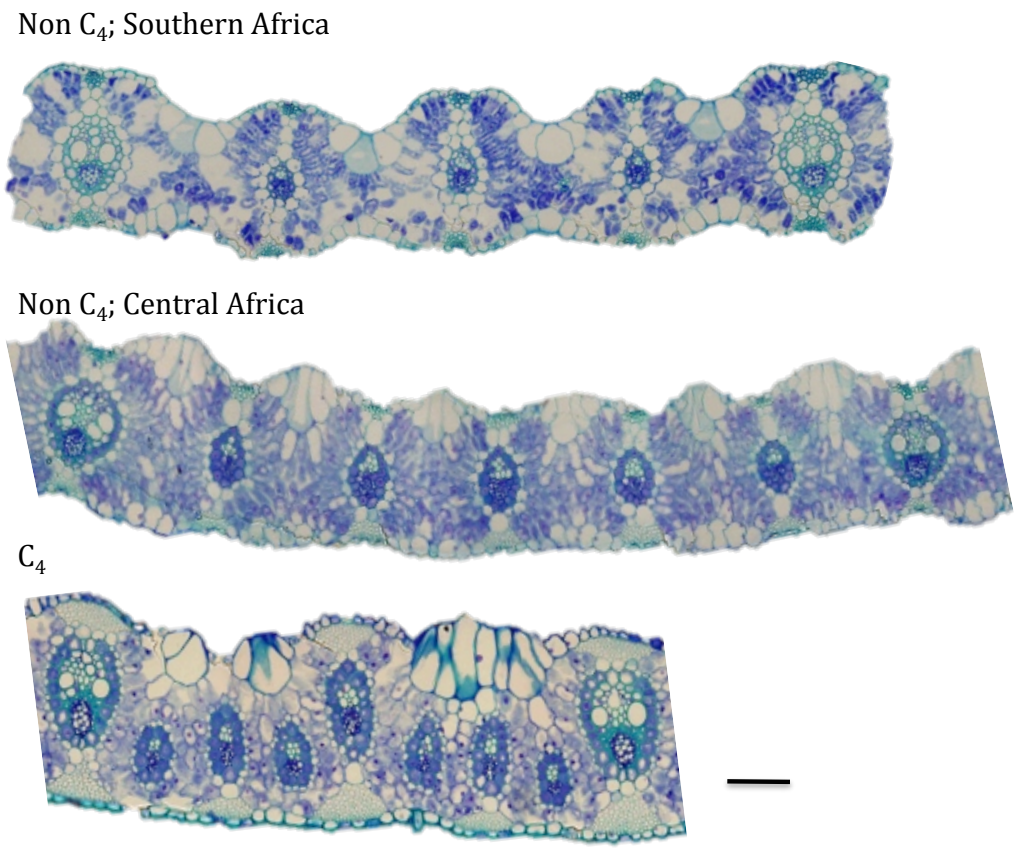


Figure 3.8 Cross-sections showing anatomical differences between typical *Alloteropsis semialata* accessions. The bottom image is C_4 ($\delta^{13}C$ is -13.9; BF-2, Burkina Faso), while the middle ($\delta^{13}C$ is -25.6, L04-B, Tanzania) and top ($\delta^{13}C$ is -29.1; EML-11, South Africa) images are non-C_4. Note differences in the number of veins per segment, the number of mesophyll cells separating these veins, and the relative sizes of inner and outer bundle sheath cells in these three accessions. Scale bar = 100 µm.

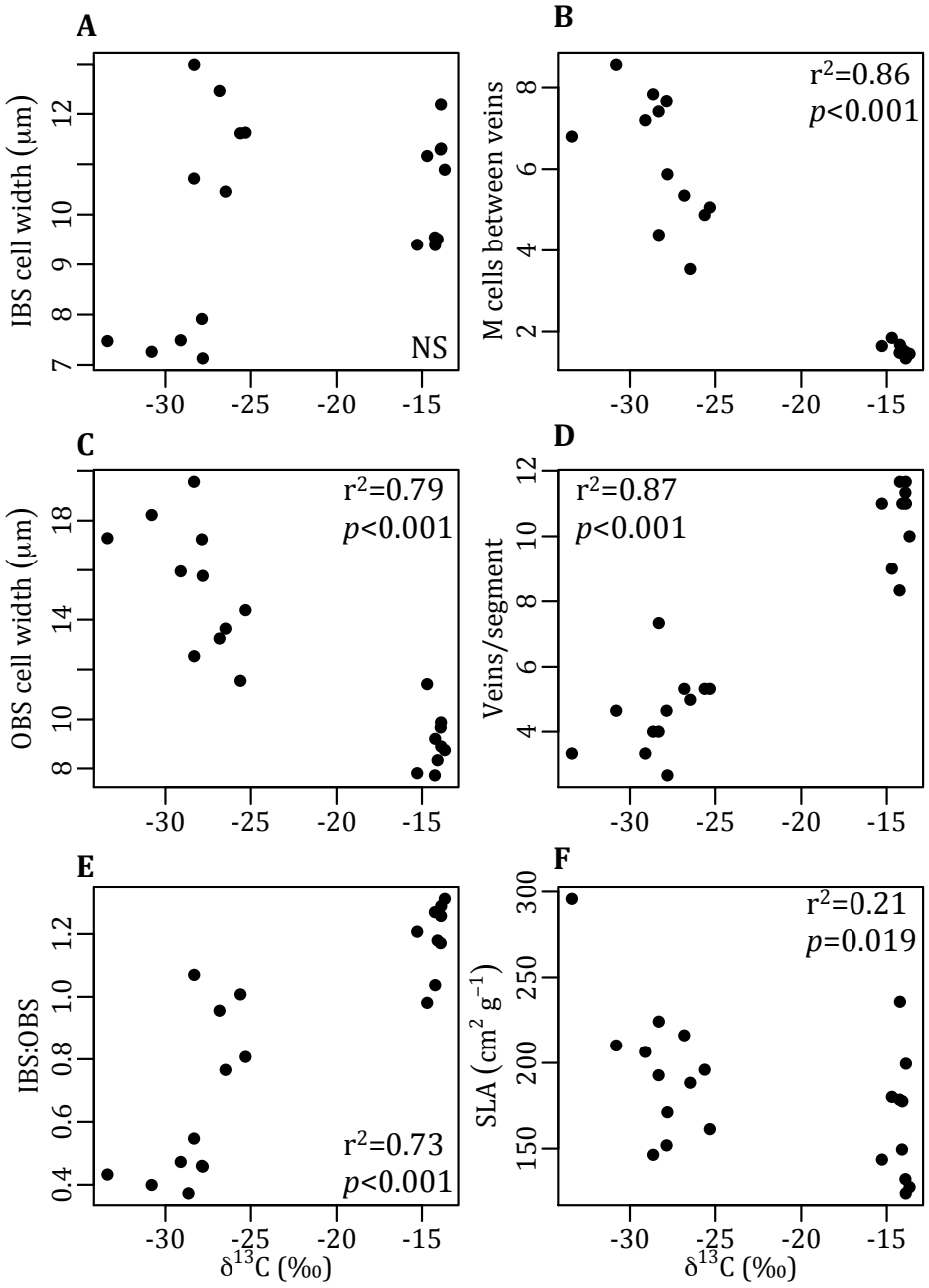


Figure 3.9 Variation in leaf anatomy traits across the $\delta^{13}C$ gradient. Plots show the width of inner (A; *IBS*) and outer (C; *OBS*) bundle sheath cells, the ratio of inner to outer bundle sheath cell widths (E; *IBS:OBS*); the minimum number of mesophyll cells separating veins (B; *M cells between veins*); the number of veins per segment (D; *veins/segment*), and specific leaf area (F; *SLA*). Linear model adjusted R^2 and *p*-value results are shown for relationships with $p<0.05$, those with $p\geq0.05$ are reported non-significant (NS). N=22.

Fresh hand-cut cross-sections show that C_4 accessions have higher concentrations of chloroplasts in the bundle sheath cells than the mesophyll (Figure 3.10c). In contrast, South African non-C_4 accessions have chloroplasts in mesophyll cells but few or no chloroplast in the bundle sheath (Figure 3.10a). The central African non-C_4 accessions have similar chloroplast concentrations in mesophyll and bundle sheath cells (Figure 3.10b)

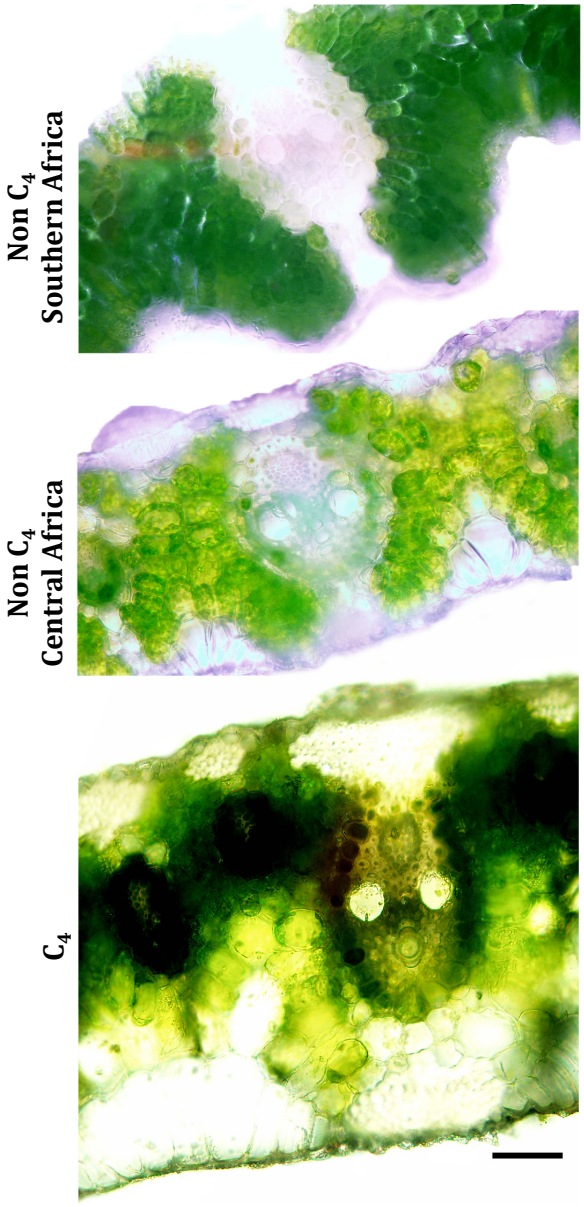


Figure 3.10 Hand cut cross-sections of *A. semialata* leaves showing differences in chloroplast localization. The southern African non-C_4 accession (top) shows high concentrations of chloroplasts in the mesophyll cells but very few in bundle sheath cells. The central African non-C_4 accession presents similar concentrations of chloroplasts in the mesophyll and bundle sheath cells. The C_4 example shows very high concentrations of chloroplasts in the bundle sheath with much lower concentrations in mesophyll cells. Scale bar = 100 µm.

Physiology

Maximum rates of photosynthesis showed a weak negative correlation with $\delta^{13}C$ signature, as accessions with higher $\delta^{13}C$ had lower A_{max} (Figure 3.11a and 3.12). This suggests that non-C_4 accessions had higher rates of photosynthesis than C_4 accessions when measured under saturating light and CO_2, as would be expected under these non-limiting CO_2 conditions where photorespiration would be minimal. In contrasts, rates of photosynthesis at ambient CO_2 concentration did not correlate with $\delta^{13}C$ (Figure 3.11b). The traits C_i/C_a, g_s, E, and WUE varied widely, but also did not correlate with $\delta^{13}C$ values (Figure 3.11 c-f), suggesting that, under these measurement conditions, these traits do not differ with photosynthetic type in *A. semialata*.

Measures of CE and CCP at ambient O_2 concentrations correlated well with $\delta^{13}C$ values as accessions with higher $\delta^{13}C$, and thus more likely to use C_4 photosynthesis, had greater CE and lower CCP (Figure 3.13b, d). Interestingly, these CCP values fell into three distinct clusters; those over 40 μmol mol^{-1}, those between 13 and 26 μmol mol^{-1}, and those 10 μmol mol^{-1} and below (Figure 3.13d). All of the C_4 accessions fell into this latter cluster with low CCP, which is consistent with other species that use C_4 photosynthesis (Table 3.1). The non-C_4 accessions, however, were split between those with high CCP values (*i.e.,* consistent with species that use C_3 photosynthesis) and those with intermediate CCP values (*i.e.,* consistent with a C_3-C_4 intermediate phenotype; Table 3.1). The five accessions that fell into this intermediate CCP cluster were from Tanzania, while the all of the accessions with C_3-like CCP were from South Africa.

CCP measured at low O_2 concentrations did not vary with $\delta^{13}C$ (Figure 3.13c). This result occurs because these hypoxic conditions suppress the fixation of O_2, and thus photorespiration, causing all CCP to be relatively low regardless of photosynthetic pathway. Thus, CCP measured at low O_2 will be similar in C_4 and non-C_4 accessions. The degree to which CCP is inhibited by O_2 concentration (*i.e.,* the trait termed OI here) will escalate with the increasing effects of photorespiration, as can be seen in Figure 3.13e. Accessions with the lowest $\delta^{13}C$ signatures experience the greatest inhibition from oxygen, while those isotopically intermediate suffer moderate OI, and C_4 accessions, suffering little

photorespiration, express low OI (Figure 3.13e). LUE, LCP, and R_{dark} did not correlate with $\delta^{13}C$ (Figure 3.13 – 3.14), suggesting these traits do not differ between photosynthetic types in *A. semialata*, under these measurement conditions.

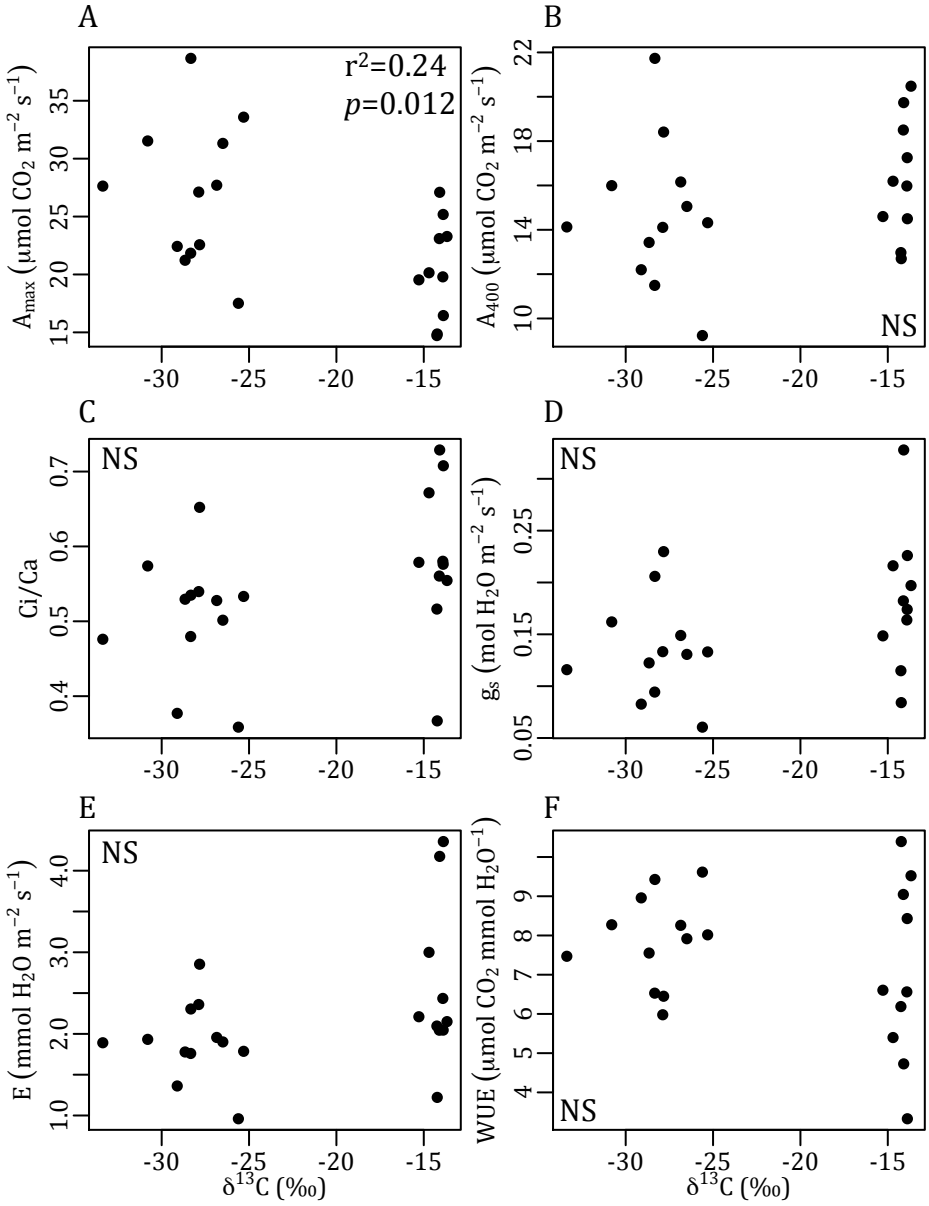


Figure 3.11 Variation in leaf physiology traits across the $\delta^{13}C$ gradient. Plots show CO_2 assimilation rates at simultaneous light and CO_2 saturation (A; A_{max}) and at light saturation and ambient CO_2 concentration (B; A_{net}), the ratio of intercellular to ambient CO_2 concentrations (C; C_i/C_a), rates of stomatal conductance to water (D; g_s); rates of transpiration (E; E); and intrinsic water use efficiency (F; ${}_iWUE$). Linear model adjusted R^2 and p-value results are shown for relationships with $p<0.05$, those with $p\geq0.05$ are reported non-significant (NS). N=22.

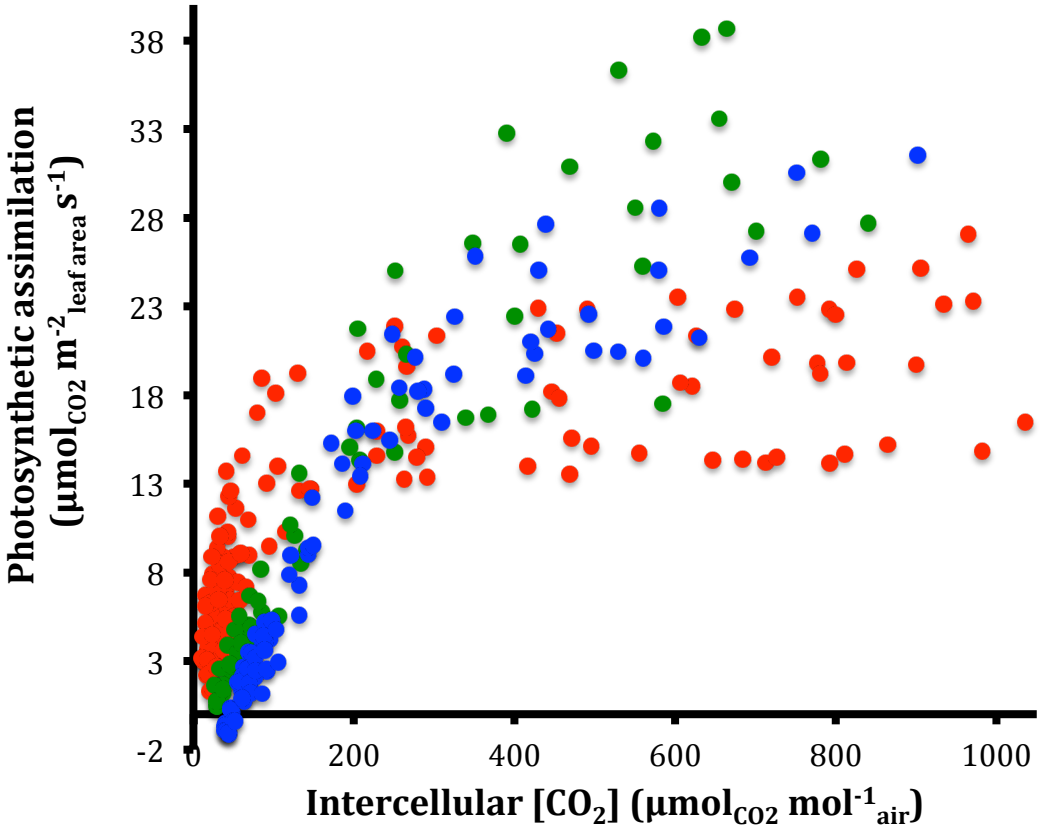


Figure 3.12 Response of photosynthetic assimilation rate to intercellular CO_2 concentration measured at ambient O_2 concentration. Points are color-coded by the clusters assigned in the hierarchical cluster analysis; Cluster 1 (red), Cluster 2 (blue), and Cluster 3 (green; see Figure 3.15; Table 3.3).

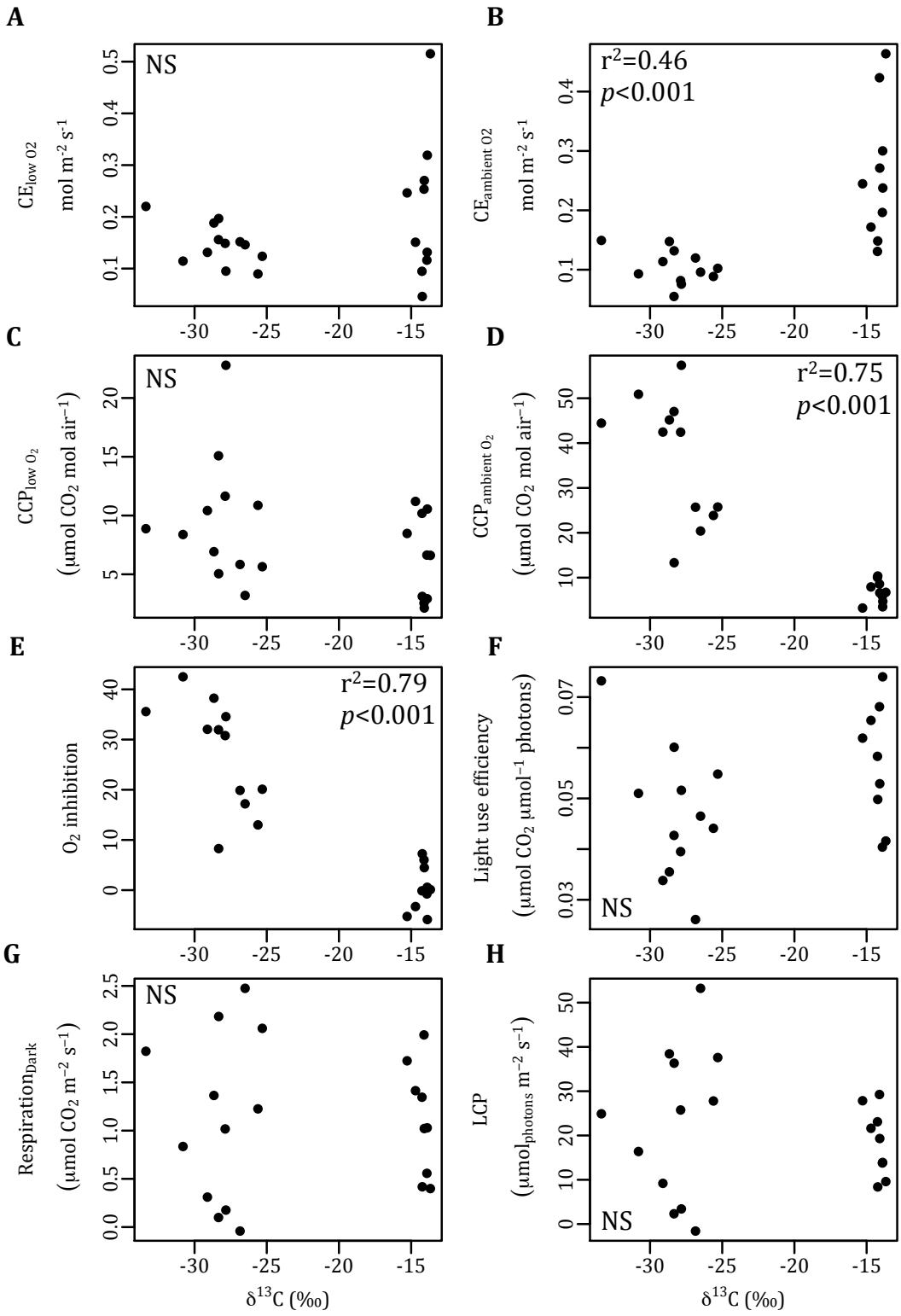


Figure 3.13 Variation in physiological traits across the $\delta^{13}C$ gradient. Plots show the carboxylation efficiency (CE) at low (A) and ambient (B) oxygen concentrations, the CO_2 compensation point (CCP) at low (C) and ambient (D) oxygen concentrations, oxygen inhibition (E); light use efficiency (F), dark respiration rate (G), and light compensation point (LCP; H). Linear model adjusted R^2 and *p*-value results are shown for relationships with $p<0.05$, those with $p\geq0.05$ are reported non-significant (NS). N=22.

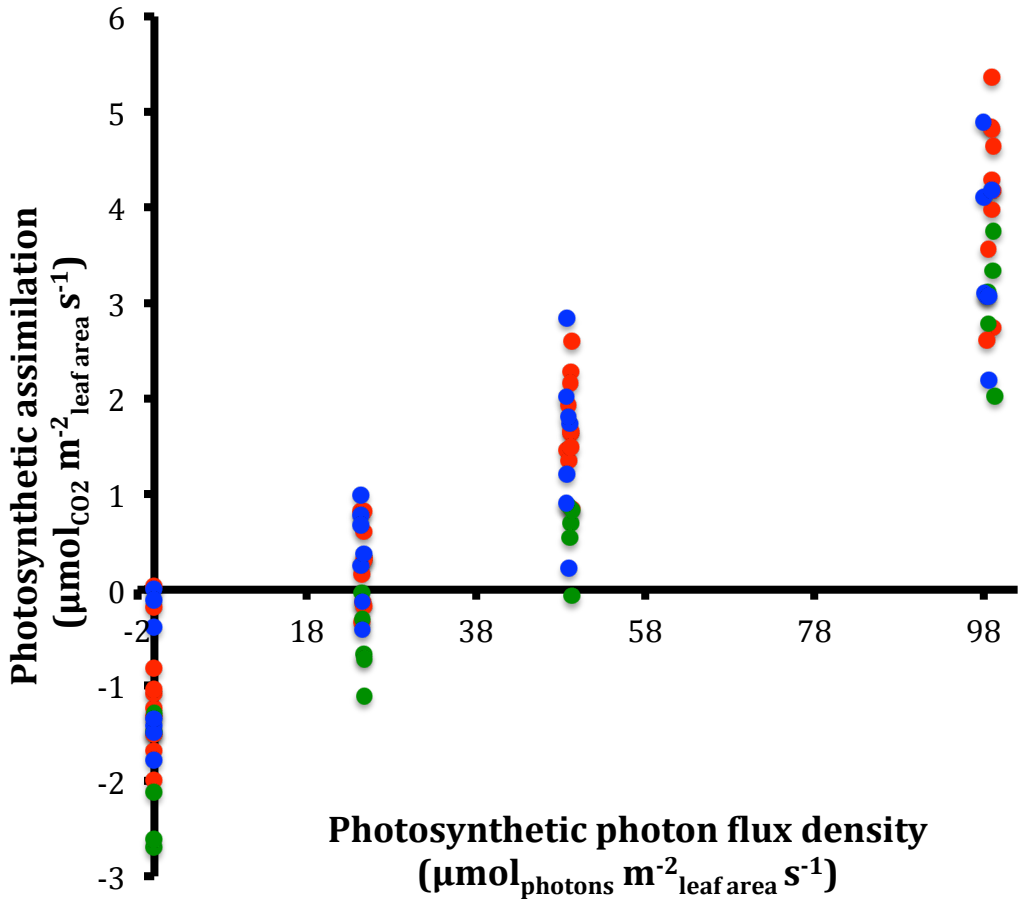


Figure 3.14 Response of photosynthetic assimilation rate to photosynthetic photon flux density. Points are color coded by the groupings assigned in the hierarchical cluster analysis; Cluster 1 (red), Cluster 2 (blue), and Cluster 3 (green; see Figure 3.15; Table 3.3).

Combining $\delta^{13}C$, anatomy, and physiology to identify photosynthetic type

A PCA was performed on traits that were hypothesized to differ by photosynthetic type, including $\delta^{13}C$ signature, A_{max}, CE, CCP, OI, OBS cell width, IBS:OBS ratio, vein frequency, and the minimum number of mesophyll cells separating veins. This PCA explained over 88% of the variation in the dataset and, using this information, the hierarchical clustering analysis placed the 22 accessions into three discrete clusters (Figure 3.15). The three clusters differed strongly along dimension 1 but overlapped heavily along dimension 2. All traits were significantly correlated with dimension 1, while only A_{max} correlated with dimension 2 (Figure 3.15c; Table 3.4). Cluster 1 fell negative along dimension 1, while clusters 2 and 3 fell along the positive side of this dimension (Figure 3.15a). Thus, Cluster 1 had the highest $\delta^{13}C$, vein frequency, IBS:OBS ratios, and CE and the lowest CCP, OI, OBS cell sizes, and mesophyll cells between veins, while plants that fell into Cluster 2 had the opposite of this phenotype (Figure 3.15). The phenotypes associated with Clusters 1 and 2 are consistent with plants using C_4 and C_3 photosynthesis, respectively (Table 3.1). Cluster 3, which is made up entirely of the five Tanzanian accessions, fell intermediate between Clusters 1 and 2 on the PCA (Figure 3.15a). The topology of the hierarchical cluster analysis suggests Cluster 3 is more similar to Cluster 2 than it is to Cluster 1, suggesting these Tanzanian accessions may be more similar to the C_3-like accessions than the C_4-like accessions (Figure 3.15b).

Table 3.4 Coefficients for variables significantly correlated with dimensions 1 and 2 of the PCA.

	Correlation coefficient	*p*-value
Dimension 1		
OI	0.97	1.49E-13
mesophyll cells	0.96	4.51E-13
CCP	0.96	2.31E-12
OBS cell width	0.93	2.48E-10
A_{max}	0.43	4.65E-02
CE	-0.69	4.14E-04
vein frequency	-0.94	7.82E-11
IBS:OBS	-0.95	3.12E-11
$\delta^{13}C$	-0.95	7.05E-12
Dimension 2		
A_{max}	0.89	2.50E-08

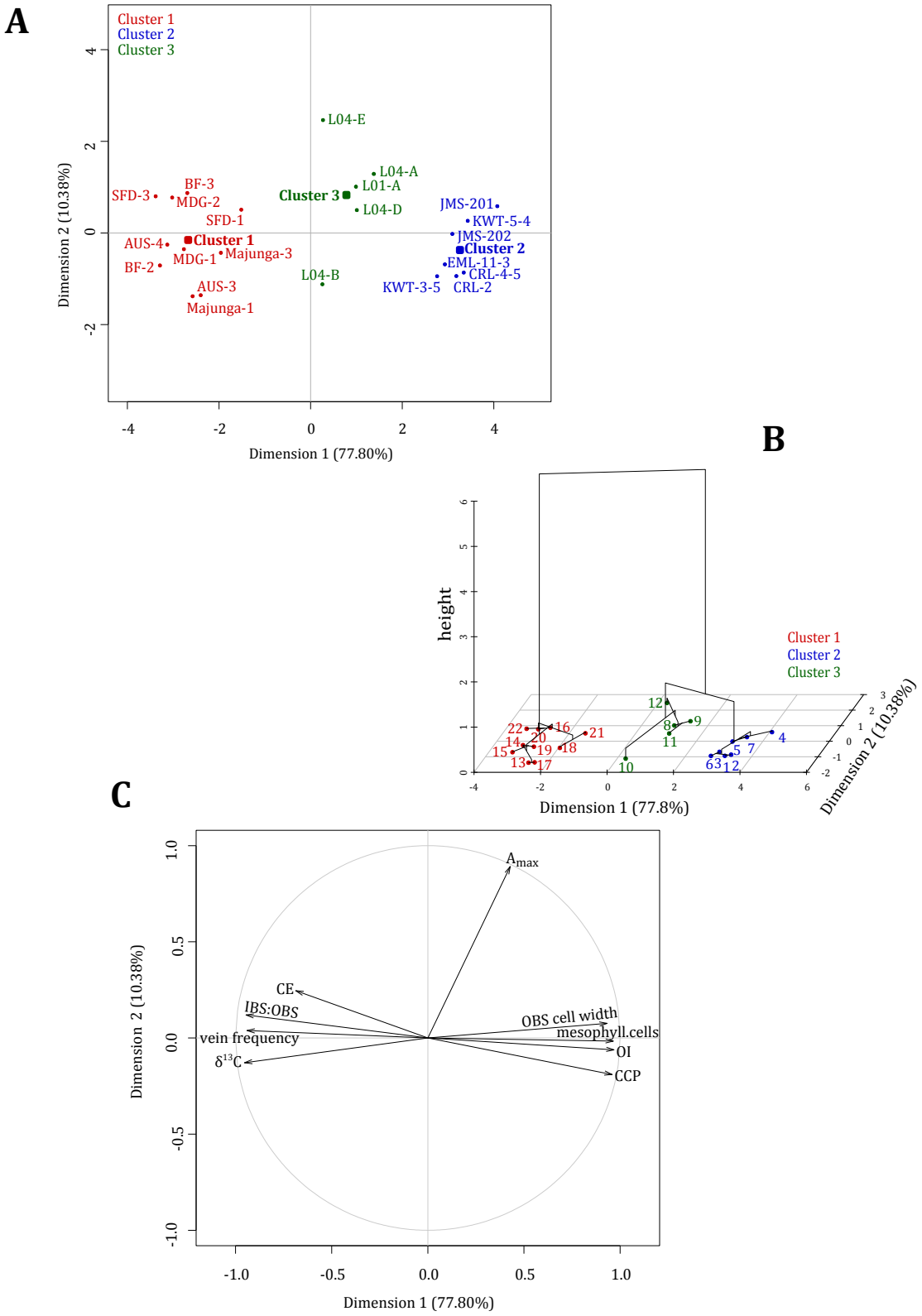


Figure 3.15 PCA and hierarchical cluster analysis of $\delta^{13}C$, anatomy, and physiology traits. Plots show (A) the distribution of accessions across the PCA, (B) how these accessions fall into clusters, and (C) Eigenvectors. These clusters are listed for each accession in Table 3.3.

Leaf form and function

A second PCA was performed to explain how the 22 accessions varied physiologically. This PCA was overlaid with leaf component and anatomy trait eigenvectors to aid interpretation and several interesting trends emerged (Figure 3.16). First, the anatomical traits OBS cell widths, minimum number of mesophyll cells between veins, vein frequency, and IBS:OBS ratio in the same direction as CCP, OI, and CE (Figure 3.16b). IBS cell widths varied in the opposite direction as CCP measured under low O_2 and delta CE_{delta}. Percents of N and C, and the C:N ratio varied along the same axis as WUE (Figure 3.16). SLA varied in the opposite direction as A_{net}.

These intriguing relationships were investigated further using correlation analysis and linear models. These tests showed that vein frequency, mesophyll cell counts, OBS cell size, and the IBS:OBS ratio were strongly correlated with each other (Figure 3.17) and with CE, CCP, and OI, but not A_{max} (Figure 3.18). IBS cell size weakly correlated with the minimum number of mesophyll cells between veins but not vein frequency (Figure 3.17). While IBS was significantly correlated with CCP and O_2 inhibition, it was not with CE or A_{max} (Figure 3.18). IBS also correlated with $\delta^{15}N$ (r^2= 0.20; p= 0.027), %N (r^2= 0.32, p= 0.005), and C:N (r^2= 0.25, p= 0.015). These leaf component traits also correlated well with certain physiological variables. For example, $\delta^{15}N$ was negatively correlated with WUE (R^2=0.56, p<0.001), positively with E (R^2=0.38, p=0.001) and C_i/Ca (R^2=0.18, p=0.030) while the percent of carbon in the leaf was positively correlated with WUE (R^2=0.26, p=0.009), and negatively so with E (R^2=0.28, p=0.007) and C_i/Ca (R^2=0.27, p=0.008). The percent nitrogen in the leaf was positively correlated with A_{max} (R^2=0.17, p=0.033) and WUE (R^2=0.16, p=0.035), implying that plants with more nitrogen in their leaves reached a higher maximum rate of photosynthesis and at the expense of less water, than those leaves with less leaf nitrogen. SLA was negatively correlated with CE (r^2= 0.17; p= 0.034), meaning that thinner leaves had higher CE than thicker leaves.

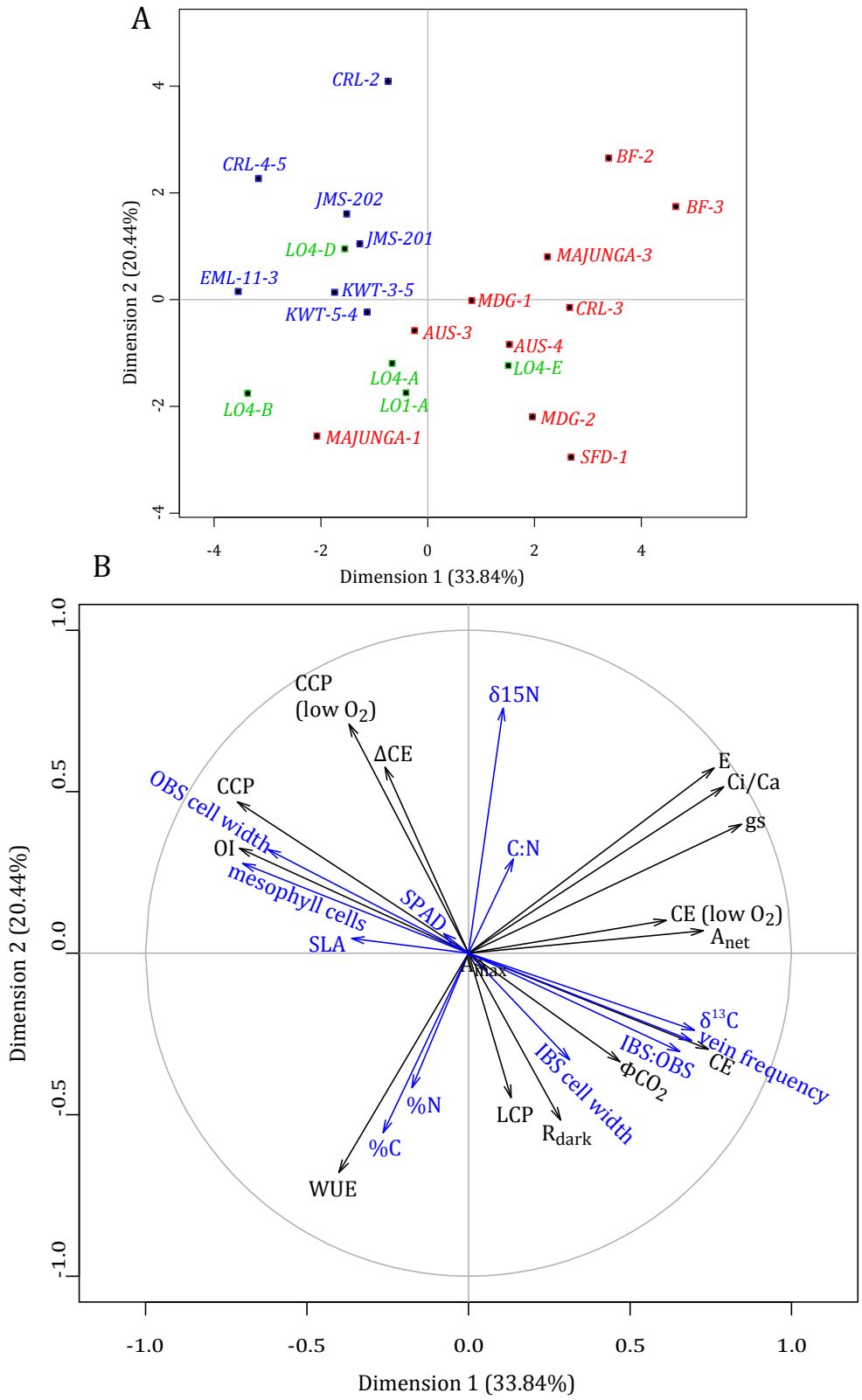


Figure 3.16 PCA on 15 physiology variables. (A) Plots show the distribution of accessions across the PCA and (B) Eigenvectors. The PCA is overlaid leaf component and anatomy traits as supplementary quantitative variables (in blue) to aid in interpretation of physiological PCA. Accession labels are colored by cluster assignment revealed in Figure 3.15; Cluster 1 (red), Cluster 2 (blue), and Cluster 3 (green).

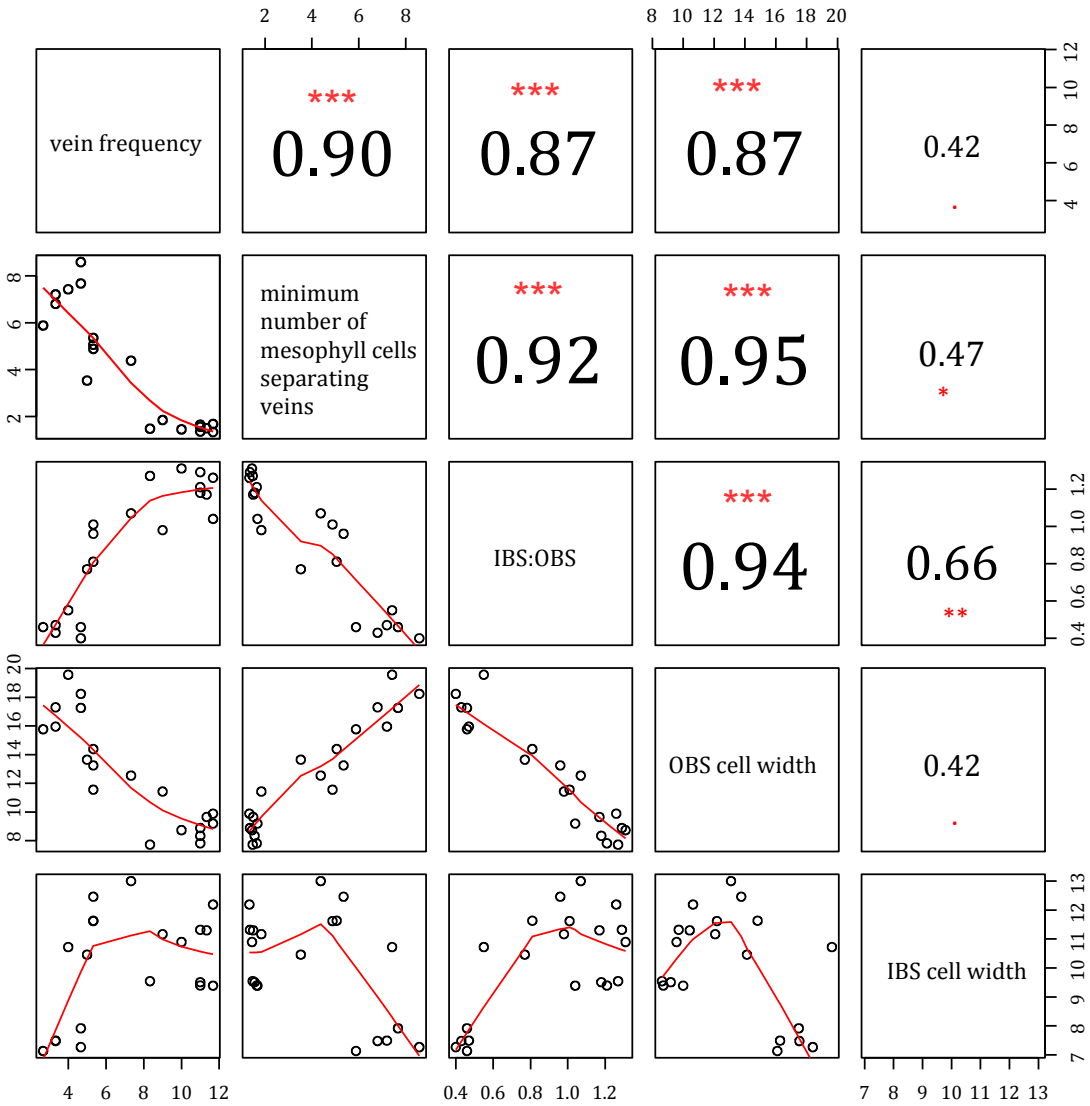


Figure 3.17 Correlation analyses between anatomical traits. R-values and significance are shown in the upper boxes. Font size scales with R-value. *p*-values are indicated as <0.1 (.), <0.05 (*), <0.01 (**), and <0.001 (***). Best-fit lines through the regression data are shown in the lower boxes. N=22.

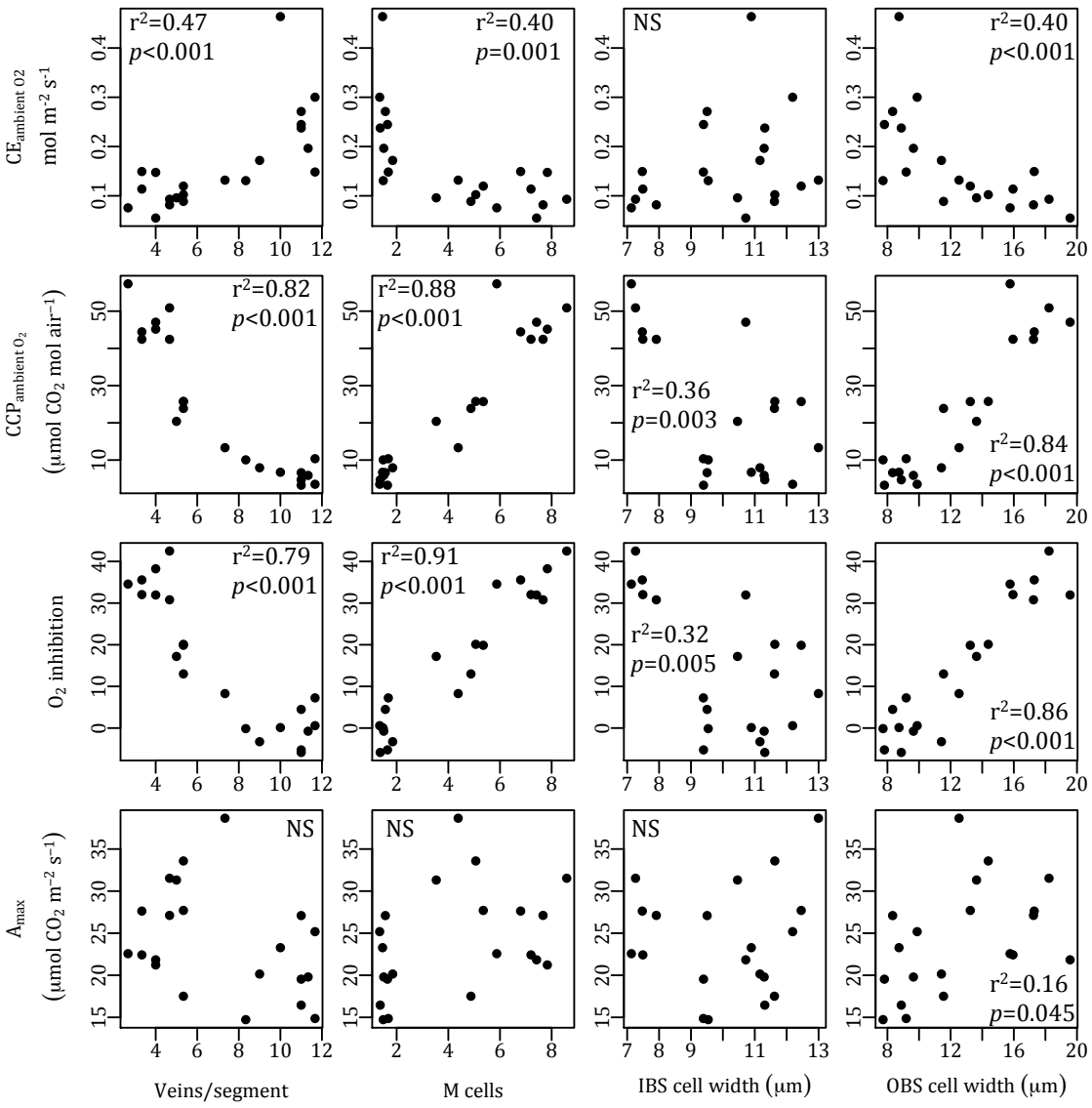


Figure 3.18 Relationships between leaf anatomy and physiology. The number of veins per segment (veins/segment), minimum number of mesophyll cells between veins (M cells), and inner and outer bundle sheath cell sizes (IBS and OBS, respectively) are plotted against carboxylation efficiency (CE) and CO_2 compensation point (CCP) measured at ambient oxygen concentrations, oxygen inhibition, and maximum rates of CO_2 assimilation measured at simultaneous light and CO_2 saturation (A_{max}). Linear model adjusted R^2 and *p*-value results are shown for relationships with $p<0.05$, those with $p\geq0.05$ are reported non-significant (NS). N=22.

Discussion

Discovery of C_3-C_4 intermediates in Alloteropsis semialata

This study combined stable isotope, physiology, and anatomy methods to discover the simultaneous occurrence of C_3, C_3-C_4 intermediate, and C_4 photosynthetic types within *A. semialata* for the first time. Accessions from Australia, Burkina Faso, and Madagascar, as well as the MDG and SFD South African accessions, expressed $\delta^{13}C$ signatures, leaf anatomy, and physiology consistent with plants using C_4 photosynthesis, while the remaining South African accessions presented values for these traits that are consistent with plants that use C_3 photosynthesis (from here onward, these accessions are referred to as C_4 and C_3 accessions, respectively). This geographic distribution of photosynthetic types, as well as the values of $\delta^{13}C$, anatomy, and physiology measured in this study, are consistent with what has been presented in previous studies of C_3 and C_4 *A. semialata* from South African and Australian populations (Ellis, 1981; Renvoize, 1987a; Ibrahim, 2007). The Tanzanian *A. semialata* accessions, however, had $\delta^{13}C$ signatures consistent with values measured in C_3 and C_2 species in the literature but were physiologically and anatomically intermediate between C_3 and C_4 accessions of this species. Thus, the five Tanzanian accessions presented in this study are believed to be C_3-C_4 intermediates (and are referred to as such from here onward). These C_3-C_4 plants may use a C_2 cycle, however, more work is needed to confirm this.

Plants using a C_2 cycle should present $\delta^{13}C$ signatures more negative than typical C_3 plants as they experience double discrimination by Rubisco via mesophyll and bundle sheath Calvin cycles (von Caemmerer, 1989). However, $\delta^{13}C$ values in the C_3-C_4 intermediate *A. semialata* accessions were not more negative than those measured in C_3 plants, but were less negative than C_3 *A. semialata* plants (*i.e.,* field values of $\delta^{13}C$ in C_3-C_4 intermediate accessions ranged between -26.4 and -23.1‰ while C_3 accessions ranged between -28.2 and -27.1‰). This indicates that the C_3-C_4 intermediate plants presented in this study may have had a limited C_4 cycle engaged, in addition to a C_2 cycle, which increased in the fraction of ^{13}C to ^{12}C in their leaf tissue. This occurs because the atmospheric CO_2 in the bundle sheath has been fixed by PEPC directly, however, up to 50% of C_4 acid cycle activity may be activated before $\delta^{13}C$ signatures start to shift upward (Peisker, 1986; Monson *et*

al., 1988; Hattersley and Watson, 1992). This suggests that these Tanzanian specimens may have high levels of C_4 activity and future work should be done to confirm this.

The likely combination of C_2 and weak C_4 cycles employed by these Tanzanian *A. semialata* plants yields physiological advantages over the C_3 conspecifics that were evident in traits that reflect photorespiratory advantages (*e.g.,* CCP, OI, and CE). The Tanzanian plants did not excel over the C_3 *A. semialata* plants in A, g_s, or WUE. However, these traits were not measured under high photorespiratory conditions in this study (*e.g.,* high temperatures), so any photorespiratory advantage over C_3 physiology was lacking.

It has been proposed that a single knockout mutation deactivates GDC in mesophyll mitochondria which forces glycine to have to diffuse into the bundle sheath to be decarboxylated, thus initiating the C_2 cycle (Gowik and Westoff, 2011; Sage *et al.*, 2012). If this is the case, an abrupt physiological shift might be expected, which would distinguish plants with and without the mesophyll GDC knockout mutation. However, this pattern is not apparent in *A. semialata*. Instead, this species expresses gradual transitions between C_3, C_3-C_4 intermediates, and C_4 phenotypes. One reason for this may be that the decline in mesophyll GDC activity is not abrupt, but gradual, as has been documented in *Flaveria* (Schulze *et al*., 2013). The continuous phenotype may also result from other factors, such as leaf anatomy, which may be key to optimizing photosynthesis. Once the abundance of GDC starts to shift away from mesophyll and into bundle sheath mitochondria, then there will be selection for anatomical modifications that benefit a C_2 cycle. Thus, there are inherent ties between the anatomical and physiological modifications that characterize the C_3- C_3-C_4 intermediate -C_4 spectrum.

Relating leaf structure with function

Alloteropsis semialata presents an ideal system to study the relationships between leaf structure and function across the C_3 / C_4 gradient, as there is considerable variation in both leaf anatomy and physiology and this study identified strong correlations between the two in this species. In particularly, this study found that

the C_3 / C_4 spectrum was accompanied by a coordinated increase in vein frequency, decrease in the number of mesophyll cells, and shrinking of OBS cells and that this suite of anatomical modifications was strongly correlated with less photorespiratory disadvantage, as measured by CE, CCP, and OI. While there was continuous variation in these anatomical traits, the enlargement of the IBS was not continuous but instead abruptly shifted between C_3 and C_3-C_4 intermediate accessions. This pattern may suggest that enlarging the IBS may be one of the first anatomical changes needed before a C_2, and certainly a C_4, cycle can engage (*i.e.,* in species that use the IBS, rather than the OBS, for the Calvin cycle). Any increases in IBS size should help accommodate larger and denser concentrations of chloroplasts, peroxisomes, and mitochondria there, which would directly influence efficiency of the C_2 and C_4 cycles, yielding lower CCP and less OI, as seen in this study. Later modifications to increase vein density and decrease mesophyll cells would help to increase total IBS area, relative to mesophyll area, and further decrease CCP and OI and increase CE.

Interestingly, the decrease in OBS cell size was much more strongly correlated to physiological benefits than the increase in IBS cell size. Indeed, CE increased and CCP and OI decreased with decreasing OBS cell sizes. Whereas the OBS likely provides high resistance to diffusions to help to reduce CO_2 leakage from the IBS, it probably also slows metabolite transfer between mesophyll and bundle sheath organelles. Thus, any shrinkage to the diffusion path between mesophyll and IBS cells should facilitate metabolite transfer and thus, increase C_2 and C_4 cycle efficiency.

Increasing the frequency of veins and/or decreasing the number of mesophyll cells between these veins functions to decrease the distance between the primary and secondary compartments in C_3-C_4 intermediate and C_4 plants (Chapter 2). This spatial rearrangement facilitates fluxes of photorespiratory and C_4 cycle metabolites between the mesophyll and bundle sheath cells, which is a critical component of both the C_2 and C_4 cycles and corresponds to physiological advantages. Furthermore, a simultaneous decrease in mesophyll and increase IBS size will shift the greater proportion of Rubisco away from mesophyll cells and into the bundle sheath, which allows for the establishment of the bundle sheath

Calvin cycle, required for both C_2 and C_4 cycles. If photorespired CO_2 is shuttled into the bundle sheath and there is a paucity of Rubisco available to refix it, the C_2 cycle would lose its efficacy. Thus, the chance of refixing photorespired CO_2 increases as Rubisco becomes more abundant in the bundle sheath, which may explain the correlations between the IBS cell size and CCP and OI. This also explains the lack of correlation with A_{max}. When CO_2 is not limiting, even plants with small IBS and larger proportions of mesophyll cells can achieve high photosynthesis rates.

As a cautionary note, it may be troublesome to assign cause-effect relationships between leaf anatomy and physiology. For example, both high vein density and larger bundle sheaths are likely to be adaptive in arid environments, where these traits would convey better water transport and hydraulic capacitance (Osborne and Sack, 2012; Griffiths *et al*., 2013). Arid conditions tend to also promote photorespiration. Thus, selection may favour increased vein density and larger bundle sheaths while acting against photorespiration in arid regions, without these anatomical shifts necessarily causing the physiological outcome.

Persistence of C_3-C_4 intermediates

The *A. semialata* C_3-C_4 intermediates presented in this study likely use a C_2 cycle. However, the C_2 cycle is considered an almost inevitable transitory step to C_4 photosynthesis (Mallmann *et al*., 2014) and is more costly than C_3 or C_4 photosynthetic types alone (Monson *et al.,* 1986; Schuster and Monson, 1990). So why has it persevered in some lineages (Williams *et al*., 2013; Sage *et al*., 2014)? One explanation for this persistence may be phenotypic plasticity. For example, high levels of physiological plasticity have been found in the C_2 species *Flaveria linearis*, which becomes C_4-like when grown in warmer conditions (Teese, 1995). This apparent flexibility to transition between C_2 and C_4 photosynthesis (and perhaps also between C_3 and C_2 photosynthesis) would enable these plants to optimize their physiology to whichever environment they are currently experiencing. This ability would be particularly advantageous in heterogeneous environments (West-Eberhard, 1989).

The C_3-C_4 intermediates identified in this study were found growing in miombo woodlands (Figure 3.19). Miombo is a general term used to describe forests dominated by *Brachystegia, Julbernardia,* and *Isoberlinia* species (Fabaceae, subfamily Caesalpinioideae; Frost, 1996; Abdallah and Monela, 2007). Plant species diversity and local endemism is high in miombo, with more than half of the species endemic (Rodgers *et al*., 1996). This endemism may result from the inherent heterogeneity of this habitat. Although less than 30 percent of incident light penetrates a mature miombo canopy (van der Meulen and Werger, 1984), these forests are deciduous, as most of the tree species species shed their leaves during the dry season (Frost, 1996). Therefore, perennial understory vegetation experience both shady and full sunlight conditions throughout the year. These forests average over 700mm of annual precipitation, yet more than 95 percent of this occurs within the 5-7 months of rainy season (Frost, 1996), meaning miombo vegetation must be able to withstand precipitation extremes. Overall miombo forest soils are considered nutrient-poor, however their quality is heterogeneous due to catenas (*i.e*., repeatable soil sequences along sloped landscapes), a variety of parent material, termite activity (*e.g.,* from foraged litter- and thus nutrients- from the forest floor, but also from termite nests that create nutrient hot-spots), seasonal leaf fall (*e.g.,* from the nitrogen rich canopy trees with N-fixing nodules), and anthropogenic history (*e.g*., slash and burn agriculture; reviewed in Frost, 1996). This environmental variability may explain the marked seasonal variation seen in the herbaceous understory of the miombo forests (Boaler, 1966).

Perennial herbaceous species of the miombo understory must be able to tolerate heterogeneity in precipitation, light, and nutrients. Because of this, the C_4 photosynthetic pathway may be optimal during portions of the year but less beneficial and too costly at other times. Therefore, the ability to plastically shift the degree to which one photosynthetic type is used over another would be hugely beneficial in these heterogeneous environments and could explain the persistence of the C_3-C_4 intermediate phenotype, and perhaps C_2 cycle, within *A. semialata*. Future studies need to assess the level of anatomical and physiological phenotypic plasticity in *A. semialata* and other extant C_3-C_4 intermediate species to reveal whether plasticity explains the universal persistence of intermediates worldwide.

Figure 3.19 Photograph of miombo woodland where the L01 C_3-C_4 intermediate accession was collected.

The emergence of C_3-C_4 intermediates in Alloteropsis semialata

Although the C_3-C_4 intermediates may be an evolutionary transitory step between C_3 and C_4 photosynthesis, they could also result from the hybrid crossing of C_3 and C_4 *A. semialata*. For instance, hybrid crosses between C_3 and C_4 *Atriplex* congeners produced intermediate phenotypes. Specifically, the F1 progeny expressed C_3-like $\delta^{13}C$ values but bundle sheath cell sizes and CCP that were intermediate between typical values measured in C_3 and C_4 *Atriplex* species (*i.e.,* CCP of 30 µmol mol^{-1} in the F1 versus 50 and 0 µmol mol^{-1} in the C_3 and C_4 parent species, respectively; Oakley *et al*., 2014). This F1 *Atriplex* phenotype is similar to the C_3-C_4 intermediate *A. semialata* accessions presented in this study. However, the *Atriplex* hybrids had lower CE than their C_3 parent species and neither CCP nor CE varied with $\delta^{13}C$ in this system (Oakley *et al*., 2014). Hybrid crosses between C_3 and C_4 *Flaveria* congeners produced less intermediate phenotypes as the F1 plants maintained C_3-like $\delta^{13}C$ and CCP (Apel *et al*., 1988).

Crossing should be possible in the C_3 and C_4 *A. semialata* subspecies. Both are diploid, at least in certain regions, and if these regions overlap then hybrid crosses may be able to occur. In this case, diploid C_3 and C_4 plants may cross, each parent

contributing one set of chromosomes, to form these diploid C_3-C_4 intermediate individuals. In this study, all C_3 accessions and the C_4 accessions from Australia, western Africa, and Madagascar were diploid. These patterns of ploidy variation are consistent with previous studies on *A. semialata* cytology (Ellis, 1981; Liebenberg and Fossey, 2001; Ueno and Sentoku, 2006). However, in South Africa, where no anomalous or C_3-C_4 intermediate *A. semialata* phenotypes have been found, the C_4 accessions in this study were hexaploid while the C_3 accessions were diploid. Thus, this polyploidy may be acting as a reproductive isolating mechanism in South Africa, preventing the production of C_3-C_4 intermediate phenotypes there. Moreover, the persistence of the C_3-C_4 intermediate phenotype may be explained if it results from the recurrent hybrid crossing of C_3 and C_4 *A. semialata* parents (but see Chapter 4).

Conclusions

Alloteropsis semialata is one of the most physiologically diverse species on the planet. This study found that C_3-C_4 intermediate populations exist in this species, in addition to the well characterized C_3 and C_4 types. These intermediates may use a C_2 cycle accompanied by a limited C_4 cycle (however, more work is required to confirm C_2 and C_4 activity in these intermediates). Discovery of this intraspecific photosynthetic variation has provided a system in which structure-function relationships can be established, as gradients of anatomical phenotypes correspond particularly well with physiological responses in this system.

More work is needed to characterize the C_3-C_4 intermediate phenotypes of *A. semialata* better. Namely, mesophyll and bundle sheath localization of GDC and Rubisco proteins as well as CA/PEPC activity in the mesophyll should be quantified. Moreover, the ultrastructure of these C_2 accessions should be assessed to quantify chloroplast and mitochondria number and localization within these two cell types.

Importantly, the source of these intermediate phenotypes, whether they exist as the result of natural selection for a more C_4-like phenotype or from the crossing of C_3 and C_4 parents, needs to be determined, as it will influence the next steps in the

study of this species. Even if the C_3 and C_4 subspecies of *A. semialata* do not naturally cross in the wild, these crosses should still be attempted in the lab as the resulting phenotypes could inform breeding programs aimed at producing more C_4-like phenotypes.

CHAPTER 4

PATTERNS OF DIVERGENCE, DISPERSAL, AND ADAPTATION IN *ALLOTEROPSIS SEMIALATA*

This study was a collaborative effort by Marjorie R. Lundgren (MRL), Pascal-Antoine Christin (PAC), and Colin P. Osborne (CPO) of the University of Sheffield, as well as Guillaume Besnard (GB) of the University of Toulouse. Accessions were selected and acquired by MRL, PAC, CPO and GB. Complete chloroplast genomes were sequenced by GB and assembled by PAC. DNA was extracted and amplified by MRL and PAC and aligned by PAC. PAC performed phylogenetic analysis. MRL assembled the geographic and ecological data and executed analyses and interpretation of dispersal strategies.

Introduction

The C_4 photosynthetic pathway is a complex trait, involving numerous anatomical and biochemical modifications to the ancestral C_3 type. Despite its complexity, this trait is highly convergent, having evolved independently in over 60 plant lineages (Sage *et al*., 2011). The primary selective advantage driving this convergence results from the CO_2 concentrating mechanism characteristic to all C_4 plants that nearly eliminates the energetically costly process of photorespiration (reviewed in Chapter 1). This renders the C_4 pathway particularly advantageous in conditions where photorespiration is important, such as hot, arid, high-light, saline, and low CO_2 environments (Björkman, 1971; Hatch, 1971; Loomis *et al.,* 1971; Ehleringer *et al*., 1991; Ehleringer and Monson, 1993, Bromham and Bennett, 2014).

During the evolutionary transition from C_3 to C_4 photosynthesis, modifications accumulate slowly over time (Christin and Osborne, 2013; Heckmann *et al*., 2013), resulting in phenotypes intermediate between those that are characteristic of C_3 and C_4 plants. In addition to the C_3-C_4 intermediate phenotype in *Alloteropsis semialata* (Chapter 3), many other extant C_3-C_4 intermediate species have been identified (reviewed in Williams *et al*., 2013 and Sage *et al*., 2014), which can offer invaluable insight into the transitory steps across the C_3 to C_4 spectrum.

Models that describe the evolution of C_4 photosynthesis present a linear transition between C_3, C_2, C_4-like, and C_4 photosynthetic types (Sage *et al*., 2012; Heckmann *et al*., 2013; Sage *et al*., 2014). From this, it could be expected that the most recent common ancestor (MRCA) of *A. semialata* would have used C_3 photosynthesis, then a subset of the species established the C_3-C_4 intermediary, and finally a subset of these intermediate plants transitioned to C_4 photosynthesis (Figure 4.1a). However, the actual course of evolutionary change may be far more complex in *A. semialata.* Because all of the *Alloteropsis* congeners use the C_4 photosynthetic pathway, it would be most parsimonious if C_3 populations were the results of a reversion from C_4 photosynthesis (Figure 4.1b; Ibrahim *et al*., 2009). However, comparative analyses of genes encoding PEPC and PCK suggested that C_4-adaptive genetic changes occurred independently in *A. semialata* and *A. angusta*, with C_4-driven genetic changes varying even across C_4 populations of *A. semialata* (Christin *et al.*, 2012a). A comparison between the unpublished transcriptome of *A. cimicina*

and that of *A. semialata* showed that C_4-related positive selection occurred independently in these two species, and that orthologous genes from the C_3 subspecies did not evolve under positive selection as expected under adaptive switches (P-A Christin, pers. comm.). These analyses suggest multiple C_4 optimizations within the genus *Alloteropsis* (*e.g.,* Figure 4.1c), but the exact photosynthetic state of the most recent common ancestor (MRCA) of the genus, or of *A. semialata* specifically, and the trajectory of photosynthetic pathway evolution within *Alloteropsis* remain unresolved.

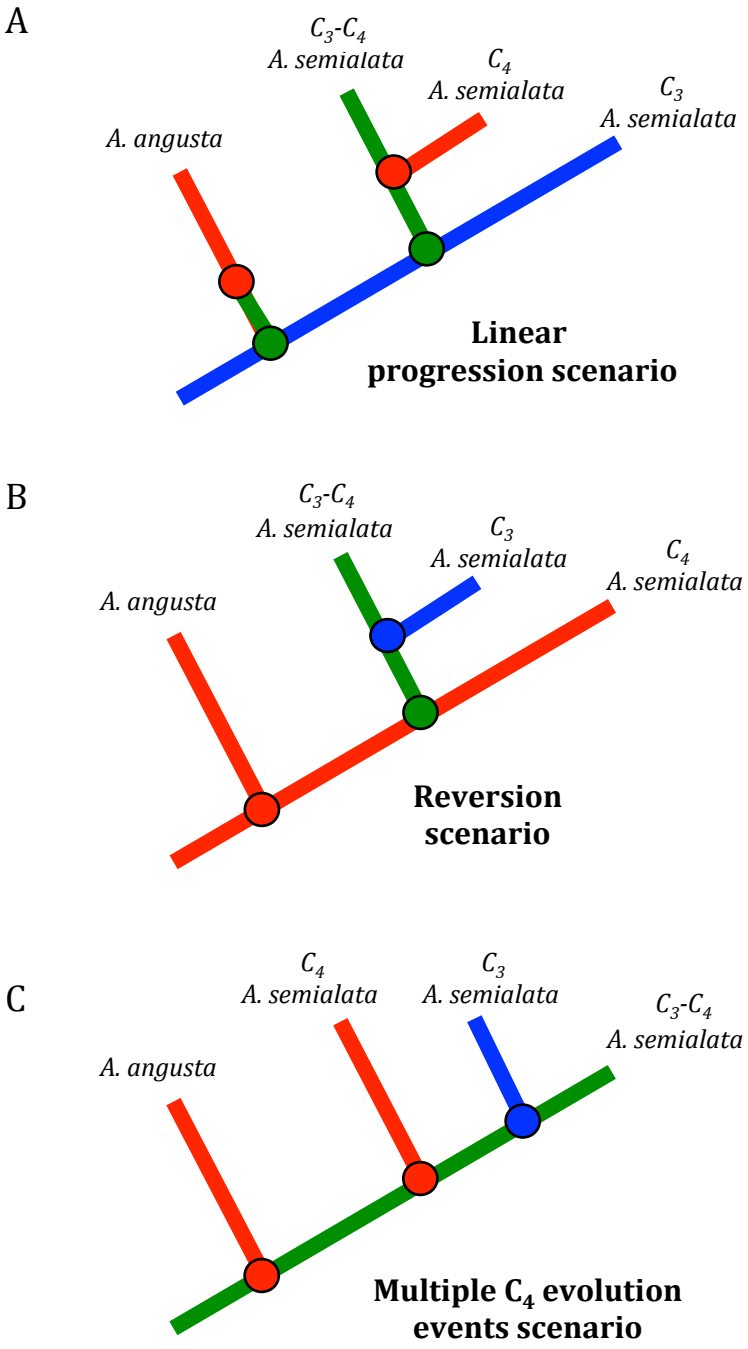


Figure 4.1 Hypothetical evolutionary trajectories that result in extant C_3 (blue), C_3-C_4 intermediate (green), and C_4 (red) lineages of *Alloteropsis semialata*, based on modified scenarios presented in the literature. The linear progression scenario (A) is based on models of C_4 evolution that progress C_3->C_3-C_4 intermediate ->C_4 photosynthetic types. The reversion scenario (B) is based on the apparent reversion from C_4 to C_3 photosynthesis presented in Ibrahim *et al.*, (2009). The third scenario (C) presents a possible trajectory in which C_4 photosynthesis evolved multiple times from a C_3-C_4 intermediate ancestor with in *Alloteropsis* and a C_3 lineage reverted from C_3-C_4 intermediate to C_3 photosynthesis (loosely based on Christin *et al.*, 2013b). These scenarios assume monophyly of the photosynthetic types.

Ecological drivers and consequences of C_4 photosynthesis

Accounting for one quarter of terrestrial primary production (Still *et al*., 2003), plants using C_4 photosynthesis are ecologically important. In particular, the productive C_4 grasses dominate savannas and grasslands of warm regions (Sage *et al*., 1999b), novel environments in which grazing ungulates and other groups including humans diversified (Sage and Stata, 2015). The ecological consequences of C_4 photosynthesis have primarily been investigated through comparisons of species distributions, which show an important effect of temperature on the occurrence of C_4 grass species (Teeri and Stowe, 1976; Ehleringer *et al.,* 1997; Epstein *et al*., 1997). This occurs for several reasons: the photorespiratory advantage experienced by C_4 plants in warmer areas is lost at cooler growing temperatures, C_4 enzymes are chilling sensitive, and the higher costs associated with C_4 photosynthesis impede success in cooler areas (Ehleringer *et al*., 1997; Pearcy and Ehleringer, 1984; Sage and Kubien, 2007). Because temperature is strongly correlated with altitude and latitude, C_4 plants rarely grow at high altitude or latitudes (Ehleringer, 1978; Tieszen *et al*., 1979; Rundel, 1980; Cavagnaro, 1988).

These investigations are biased by differences among phylogenetic groups (Taub, 2000), however, and recent interspecific comparisons accounting for phylogenetic relationships have revolutionized our understanding of C_4 evolutionary ecology (reviewed in Christin and Osborne, 2014). In particular, phylogeny-based analyses have shown that C_4 photosynthesis evolved in groups of grasses inhabiting warm regions and facilitated shifts into drier and more saline habitats (Osborne and Freckleton, 2009; Edwards and Smith, 2010; Bromham and Bennett, 2014). However, the photosynthetic transitions investigated in these analyses occurred tens of millions of years ago and the last C_3 ancestor is often separated from the first C_4 descendant by several million years (*e.g*., Christin *et al*., 2008). These vast timescales make it difficult to confidently reconstruct the conditions under which C_4 photosynthesis evolved or the events that occurred immediately after this physiological divergence.

Studies designed to pinpoint the ecological drivers of C_4 evolution are often based on species phylogenetic trees, in which photosynthetic diversification events

occurred tens of millions of years ago and the last C_3 ancestor is often separated from the first C_4 descendant by several millions of years (*e.g.,* Christin *et al.*, 2008, Osborne and Freckleton, 2009; Edwards and Still, 2008; Edwards and Smith, 2010; Kadereit *et al*., 2012). These vast timescales make it difficult to confidently identify the conditions in which C_4 photosynthesis evolved or the events that occurred immediately after this physiological diversification. Identifying the selective factors that promoted the gradual assembly of C_4 photosynthesis within populations consequently requires investigations within species complexes that express variation in the photosynthetic phenotype, such as *A. semialata*.

Aims of Chapter 4

This study uses a phylogenetic approach to (i) distinguish the drivers and consequences of C_4 evolution, (ii) resolve dispersal patterns, and (iii) reconstruct the evolutionary history of the intraspecific photosynthetic variation found in *A. semialata*. Sequences from over 60 accessions from across the geographic range of *A. semialata* were used to build a time-calibrated phylogenetic tree. The phylogeographic history of *A. semialata* was inferred using plastid genomes, which reproduce asexually and are usually maternally inherited (Birky, 1995), such that when seeds disperse, the intact chloroplast genome is transplanted to new locations. Subsequent gene flow via pollen does not affect the distribution of plastid genomes. Plastid markers consequently maintain the footprint of seed-mediated dispersal and the colonization of new environments.

Methods

This study incorporates 69 *Alloteropsis* accessions, including 67 accessions of *A. semialata* and one each of the congeners *A. cimicina*, and *A. angusta.* Accessions were chosen to represent the range of geographic and photosynthetic variation present in *A. semialata* (Figure 4.2, Table 4.1; Chapter 3). The congeners use the C_4 photosynthetic pathway while *A. semialata* accessions use one of the C_3, C_3-C_4 intermediate, and C_4 types (Ibrahim *et al*., 2009; Chapter 3). Photosynthetic type was determined by stable isotope, physiology, and anatomical screening methods,

as described in Chapter 3. In total, this study included 35 C_4 and 31 non-C_4 plants, of which 18 were putatively C_3-C_4 intermediate accessions of *A. semialata*.

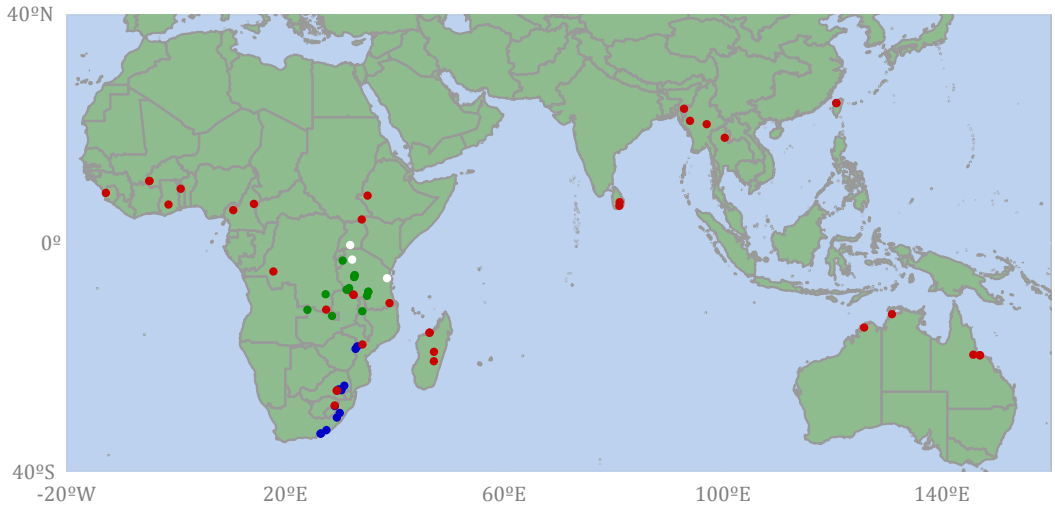


Figure 4.2 Collection locations for C_4 (red), C_3-C_4 (green), and C_3 (blue) *Alloteropsis semialata* accessions and congeners (white).

Genome-skimming and complete chloroplast genome assembly

Complete chloroplast genomes were generated with Illumina sequencing for 13 individuals, including two congeners and 11 *A. semialata* accessions, following the approach of Besnard *et al.* (2013). These genomes were incorporated into a larger dataset of grass species, and dating analyses estimated inter- and intraspecific divergence times within *Alloteropsis*.

Whole genomic DNA was isolated from silica-gel dried material, using the BioSprint 15 DNA plant Kit (Qiagen Inc., Texas, USA). The quality and quantity of the DNA was assessed by visualization on agarose gel, using Quanti-iT PicoGreen (Molecular Probes Inc., Oregon, USA) and NanoDrop (Thermo Fisher Scientific, Delaware, USA). About 200 ng of double-strand DNA (based on the PicoGreen quantification) were sonicated and then used for library preparation with an Illumina TruSeq DNA Sample Prep v.2 Kit (Illumina Inc., California, USA), following the manufacturer's instructions. Fragments with an insert size of 400-500 bp were isolated on gel and submitted to shotgun sequencing using Illumina technology (HiSeq 2000). These 13 samples were pooled with 11 other samples and sequenced on one Illumina lane, as 100-bp paired-end reads.

Table 4.1 Accessions included in this study.

Species	ID	Voucher	Country	Latitude	Longitude	$\delta^{13}C$	Year	Clade
A. angusta	3C	Namaganda & Wanyana 3C (MHU)	Uganda	-0.36	31.87	-	2009	-
A. cimicina	RCH20	RC Hall 20 (K)	Madagascar				2011	-
A. semialata	31776	Torre Sofala 5845 (K)	Mozambique	-18.06	33.17	-	1943	Clade A
A. semialata	39691	Simon 932 (K)	Zimbabwe	-18.50	32.85	-25.4	1966	Clade A
A. semialata	BL	Lundgren & Ripley 11 (SHD)	South Africa	-29.71	29.96	-28.0	2012	Clade A
A. semialata	JM	Ripley 1 (SHD)	South Africa	-33.32	26.44	-26.8	2012	Clade A
A. semialata	PAC100	Lundgren & Ripley 5 (SHD)	South Africa	-24.93	30.79	-25.2	2012	Clade A
A. semialata	PAC101	Lundgren & Ripley 9 (SHD)	South Africa	-28.50	29.06	-27.8	2012	Clade A
A. semialata	PAC3	Ibrahim 15 (SHD)	South Africa	-33.30	26.52	-	2004	Clade A
A. semialata	PAC94	Lundgren & Ripley 6 (SHD)	South Africa	-25.62	30.26	-26.0	2012	Clade A
A. semialata	PAC95	Lundgren & Ripley 12 (SHD)	South Africa	-30.51	29.43	-28.0	2012	Clade A
A. semialata	PAC96	Lundgren & Ripley 7 (SHD)	South Africa	-25.74	30.24	-24.8	2012	Clade A
A. semialata	PAC97	Lundgren & Ripley 13 (SHD)	South Africa	-32.70	27.53	-26.3	2012	Clade A
A. semialata	PAC98	Lundgren & Ripley 3 (SHD)	South Africa	-25.58	29.77	-25.7	2012	Clade A
A. semialata	PAC99	Lundgren & Ripley 2 (SHD)	South Africa	-25.76	29.47	-23.2	2012	Clade A
A. semialata	39680	Bullock 1979 (K)	Tanzania	-8.15	31.24	-19.7	1949	Clade B
A. semialata	39681	Bullock 1980 (K)	Tanzania	-8.15	31.24	-22.3	1949	Clade B
A. semialata	39683	Milne-Redhead 3021 (K)	Zambia	-11.67	24.05	-20.0	1937	Clade B
A. semialata	39686	Reekmans 7226 (K)	Burundi	-3.07	30.50	-24.9	1978	Clade B
A. semialata	39688	Ruffo & Kisena 2806 (K)	Tanzania	-7.87	31.67	-18.6	1987	Clade B
A. semialata	39690	Shaunty 488 (K)	Zambia	-12.75	28.56	-20.7	1919	Clade B
A. semialata	39692	V-F 2079 (K)	Tanzania	-8.22	31.57	-25.1	1958	Clade B
A. semialata	L01-A	Lundgren & Christin 1 (SHD)	Tanzania	-5.63	32.69	-26.3	2014	Clade B
A. semialata	L01-B	Lundgren & Christin 1 (SHD)	Tanzania	-5.63	32.69	-26.9	2014	Clade B
A. semialata	L01-C	Lundgren & Christin 1 (SHD)	Tanzania	-5.63	32.69	-27.4	2014	Clade B

Species	ID	Voucher	Country	Latitude	Longitude	$\delta^{13}C$	Year	Clade
A. semialata	PAC1-C	Lundgren & Christin 2 (SHD)	Tanzania	-5.93	32.60	-26.8	2014	Clade B
A. semialata	PAC1-F	Lundgren & Christin 2 (SHD)	Tanzania	-5.93	32.60	-26.4	2014	Clade B
A. semialata	31767	Symoens 14118 (K)	DRC	-8.95	27.38	-22.6	1971	Clade C
A. semialata	31772	Emson 340 (K)	Tanzania	-9.23	34.95	-21.4	1932	Clade C
A. semialata	L04-A	Lundgren & Christin 4 (SHD)	Tanzania	-8.51	35.17	-23.1	2014	Clade C
A. semialata	L04-C	Lundgren & Christin 4 (SHD)	Tanzania	-8.51	35.17	-23.0	2014	Clade C
A. semialata	L04-E	Lundgren & Christin 4 (SHD)	Tanzania	-8.51	35.17	-25.0	2014	Clade C
A. semialata	MRL4	Chapama 903 (K)	Malawi	-11.92	34.08	-34.1	2008	Clade C
A. semialata	31765	Jordan 869 (K)	Sierra Leone	8.77	-12.78	-10.7	1953	Clade D
A. semialata	31766	Morton A3243a (K)	Ghana	6.73	-1.34	-10.0	1958	Clade D
A. semialata	31768	Poelman 92 (K)	DRC	-11.64	27.48	-10.6	1961	Clade D
A. semialata	31769	Brunt 279 (K)	Cameroon	5.75	10.50	-10.7	1962	Clade D
A. semialata	31770	Raynal 13309 (K)	Cameroon	6.85	14.27	-10.1	1965	Clade D
A. semialata	31771	Schmitz 4195 (K)	DRC	-4.97	17.83	-10.6	1982	Clade D
A. semialata	31773	Bidgood et al 1832 (K)	Tanzania	-10.50	39.05	-	1991	Clade D
A. semialata	31774	Wilson 907 (K)	Uganda	4.12	33.98	-	1960	Clade D
A. semialata	31775	Muller & Pope 2053 (K)	Mozambique	-17.73	34.12	-	1971	Clade D
A. semialata	39693	Friis et al 2401 (K)	Ethiopia	8.29	35.05	-	1982	Clade D
A. semialata	39694	Scholz 478 (K)	Togo	9.49	0.90	-	1978	Clade D
A. semialata	Bur	Sanou BUR-734 (K)	Burkina Faso	10.85	-4.83	-11.3	2009	Clade D
A. semialata	MD		South Africa	-25.76	29.47	-11.1	2012	Clade D
A. semialata	Poa15	Ibrahim 20 (SHD)	South Africa	-25.83	29.40	-12.7	2004	Clade D
A. semialata	Poa18		South Africa	-28.39	29.04	-12.7	2012	Clade D
A. semialata	31759	McKee 6138 (K)	Myanmar	20.78	97.03	-	1958	Clade E
A. semialata	31760	Koelz 32989 (K)	India, Assam	23.48	92.89	-10.6	1953	Clade E
A. semialata	31761	Gould 13495 (K)	Sri Lanka	6.50	81.07	-11.3	1970	Clade E
A. semialata	31762	Lazarides 7323 (K)	Sri Lanka	7.15	81.11	-11.5	1970	Clade E
A. semialata	31763	Kingdon-Ward 22158 (K)	Myanmar	21.37	93.98	-	1956	Clade E
A. semialata	31780	Haine 242 (K)	Madagascar	-19.02	47.18	-11.3	1969	Clade E

Species	ID	Voucher	Country	Latitude	Longitude	$\delta^{13}C$	Year	Clade
A. semialata	39682	Kenneally 8625B (K)	Australia	-14.77	125.80	-11.5	1982	Clade E
A. semialata	Aus		Australia	-19.62	146.96	-12.1	2005	Clade E
A. semialata	Ma		Madagascar	-15.67	46.37	-11.8	2011	Clade E
A. semialata	MRL2	RGD173 (K)	Madagascar	-20.66	47.16	-	2013	Clade E
A. semialata	MRL5	Vorontsova 919 (K)	Darwin, Australia	-12.44	130.88	-	-	Clade E
A. semialata	PAC17	AusTRCF 322461	Queensland, Australia	-19.52	145.75	-13.0	2005	Clade E
A. semialata	PAC19	AusTRCF 322458	Queensland, Australia	-19.62	146.96	-12.1	2005	Clade E
A. semialata	Thai	AT & SS 837 (TCD)	Thailand	18.41	100.33	-12.2	2010	Clade E
A. semialata	TW1	-	Taiwan	24.47	120.72	-14.6	2014	Clade E
A. semialata	TW3	-	Taiwan	24.47	120.72	-14.6	2014	Clade E
A. semialata	TW7	-	Taiwan	24.47	120.72	-14.6	2014	Clade E
A. semialata	L02-G	Lundgren & Christin 3 (SHD)	Tanzania	-9.04	32.48	-13.2	2014	Clade F
A. semialata	L02-M	Lundgren & Christin 3 (SHD)	Tanzania	-9.04	32.48	-12.7	2014	Clade F
A. semialata	L02-O	Lundgren & Christin 3 (SHD)	Tanzania	-9.04	32.48	-11.4	2014	Clade F

Latitude and longitude are in decimal degrees.

Complete chloroplast genomes were assembled following a genome-walking approach with in-house perl scripts. A 300-bp segment of *matK* previously isolated from *A. semialata* was used as a probe. For each newly sequenced sample, reads corresponding to this region were identified using the BLAST algorithm and aligned to the probe. A consensus sequence was then built and used as the starting point for the next step of the genome-walking. A new BLAST search was then used to identify reads strictly identical to the end of the sequence that overlapped on at least 70bp, but with a preference for reads with an overlap of 90bp, and were used to extend the sequence. This step was repeated until no read could be added, in which case the last two additions were cancelled, the last read used was removed from the pool of reads available, and the analysis started again. This procedure allowed reads that contained errors to be discarded. At the end, each position in the assembled sequence was supported by a theoretical minimum of four different reads, although in reality most positions were supported by at least ten different reads. When automatic assembly was not possible, which happened around long single-nucleotide repeats, all matching reads identified via BLAST searches were aligned and used to compute a consensus sequence, which was added to the assembled sequence, allowing the analysis to continue. Regions around the inverted-repeats, where different assemblies were possible, were similarly curated manually.

Re-sequencing of chloroplast markers

Plastid markers were isolated from an additional 55 accessions of *A. semialata.* Sequences for chloroplast markers generated in previous studies (Ibrahim *et al*., 2009; GPWGII, 2012) were retrieved, and additional accessions/markers were screened via PCR. DNA was extracted from silica dried leaf tissue, recently collected seeds, or herbarium leaf or seed tissue, using the DNeasy Plant Mini Kit (Qiagen Inc., Texas, USA), following the instructions provided. DNA from herbarium specimens housed at Kew Herbarium were extracted at and obtained from the Jodrell Lab at the Royal Botanical Gardens, Kew.

Five plastid regions (*trnKmatK*, *rpl16*, *ndhF*, *rpoC2* and *trnLF*) were sequenced from a few selected individuals with good quality DNA. Isolation and sequencing

was performed with published markers (Christin *et al.*, 2012b; GPWGII, 2012). In this preliminary screen, the two regions *trnKmatK* and *rpl16* were the most variable, so these markers were isolated from all DNA, regardless of their quality. Each DNA was used as a template to PCR-amplify different plastid regions, in a total volume of 25µl, including c. 40-100 ng of gDNA, 5µl of 5x GoTaq reaction buffer, 0.1 mM dNTPs, 0.1 µM of each primer, 1mM of $MgCl_2$, and 0.5 unit of *Taq* polymerase (GoTaq, Promega, Madison, Wisconsin, USA). The PCR mixtures were first denaturated for 2 minutes at 94°C, which was followed by 35-40 cycles of denaturation (30 seconds at 94°C), annealing (30 seconds at 48°C), and extension (90 seconds at 72°C), and then finished with 10 minutes' extension at 72°C. Successful PCR products were cleaned with an Exo-SAP treatment (Affymetrix, High Wycombe, UK) or a Qiagen PCR purification kit (Qiagen Inc., Texas, USA), and sequenced using the Big Dye 3.1 Terminator Cycle Sequencing chemistry (Applied Biosystems, Foster City, California, USA). Sequencing was performed by the University of Sheffield Core Genomic Facility. All sequences will be deposited in the NCBI database.

Phylogenetic analyses

The 13 chloroplast genomes were manually aligned to the alignment of grass genomes from Besnard *et al.* (2014), after removing early-diverging taxa. Divergence times were then estimated using BEAST (Drummond and Rambaut, 2007). A GTR+G+I substitution model was used, and the speciation prior was modelled by a Yule process. A log-normal relaxed molecular clock was used. Monophyly of BEP and PACMAD clades were forced to root the tree (GPWGII, 2012), and the age of the common ancestor of these two clades (root of the tree) was set to follow a normal distribution with a mean of 51.2 and a standard deviation of 1.0, based on Christin *et al.* (2014a). Four analyses were run for 20,000,000 generations, sampling a tree each 1,000 generations. The convergence of the runs and the appropriate burn-in period were assessed using Tracer (Rambaut and Drummond, 2007), and the age of the major divergences within the *Alloteropsis* group were estimated based on all trees sampled after the burn-in period.

The five markers isolated for additional accessions were individually aligned using Muscle (Edgar, 2004), and each alignment was manually curated. The five markers were then aligned to the 13 *Alloteropsis* chloroplast genomes. A four-nucleotide polymorphic region near the end of the *rpl16* marker contained substitutions that generated three distinct patterns; 'TTTT', 'TTTG', and 'TTGG'. In addition, several individuals presented reverse-complement sequences, which very likely result from local invertions; 'CAAA' and 'AAAA'. The invertion is homoplasic, being present in the distantly related congener *A. cimicina* as well as distinct *A. semialata* accessions. To avoid attraction due to four identical nucleotides resulting from one homoplasic invertion, the inverted sequences were replaced by their reverse-complement ('AAAA' → 'TTTT' and 'CAAA' → 'TTTG'). The final dataset consisted of 69 accessions and 140,255 bp. A phylogenetic tree was inferred using Beast, under a GTR+G+I, with a constant-size coalescent prior, and a log-normal relaxed molecular clock. The root was set by forcing the monophyly of *semialata*+*angusta* and *cimicina*+*papillosa*, based on grass-wide phylogenetic analyses (Ibrahim *et al*., 2009; Christin *et al*., 2012a; GPWGII, 2012). The root of the tree was fixed to an arbitrary value of 10 time units, using a normal distribution with a mean of 10 and a standard deviation of 0.0001. Two distinct analyses were run for 20,000,000 generations, sampling a tree every 1,000 generations. The burn-in period was set to 2,000,000 generations using Tracer. Trees sampled in the two runs were combined, and the maximum clade credibility tree was used, with the node ages computed as the medians across sampled trees. This maximum-credibility tree was used for comparative analyses. In addition, a total of 100 trees were sampled from each analyses every 300,000 generations after a burn-in period of 5,000,000, and used to repeat all comparative analyses (Appendix 4.1).

Analyses of divergence patterns

Analyses were performed to show how the lineages of *A. semialata* diversified and adapted over time. A pairwise approach was used to allow for comparison of three, otherwise disparate datasets. Specifically, geographic distances revealed dispersal patterns, environmental distances reflected colonisation of environmental space, and phylogenetic distances obtained from the calibrated plastid tree provided divergence times. The pairwise analyses allowed the estimation of geographic and

environmental distances accomplished per unit of time. Only one individual per population was used for these comparisons.

Geographic, environmental, and temporal distances were calculated as described in the subsections below and these distance matrices were then compared. The *Alloteropsis* congeners were not included in these analyses, because their important divergence times would contribute disproportionately to the variation in the data and bias the statistical tests. One C_4 accession of *A. semialata*, for which latitude/longitude coordinates could not be obtained, was omitted. Thus, the three distances were calculated and analysed on 62 *A. semialata* accessions.

Geographic distance

Latitude and longitude coordinates of the collection location for each accession (Figure 4.2; Table 4.1) were used to calculate the distance (in kilometres) along the Earth's surface between each pair of points. A matrix of these geographic distances was calculated in R (version 3.1.0, R Development Core Team, 2014) using the *earth.dist* function in the FOSSIL package (Vavrek, 2011).

Environmental distance

The environmental distance was calculated based on positions of each accession within a principal component analysis (PCA). First, environmental information for each location was obtained by overlaying geographic coordinates onto environmental raster layers and extracting the raster data at each location using the SP (Pebesma and Bivand, 2005), RGDAL (Bivand *et al.*, 2014), RASTER (Hijmans *et al*., 2014), and PLYR (Wickham, 2011) packages in R. High resolution global gridded datasets of the 37 environmental layers used in this study are detailed in Appendix 4.2. All datasets use the standard coordinate datum, World Geodetic System 1984 (WGS84), and portray continuous, numerical values. Datasets of global monthly data were averaged or summed over the 12 months to calculate an annual value.

A PCA was performed using these 37 environmental variables with the FACTOMINER package (Lê *et al*., 2008) in R. The environmental distances were calculated as Euclidian distances in the space formed by the first two PCA axes.

Divergence times

Divergence times, in millions of years, were calculated as described in the *Phylogenetic Analysis* section above. The time to the MRCA of each pair of accessions was calculated in R, using the APE package (Paradis *et al*., 2004).

Analyses of divergence time, environment distance, and geographic distance

Mantel permutation tests were used to assess the statistical association between pairs of matrices. These tests were performed on each pairwise combination of the three variables; geographic distance, environmental distance, and divergence times, in R. Matrix comparisons with *p*-values below 0.05 were considered significant. These were conducted separately on each of the two sister groups ABC and DE, which differ in their photosynthetic type. Linear regressions were subsequently used to calculate the slope and R^2 for significant relationships. In cases where all relationships were significant, the relationship between the part of environmental distances not explained by geographical distances (residuals of the regression) and divergence times was tested. Finally, Mantel tests were repeated on the BC clade, excluding the A clade. This latter is from southern Africa, and its inclusion simultaneously leads to large environmental and geographical distances and spatial autocorrelation over the whole dataset.

Reconstruction of ancestral environmental states

Ancestral state reconstructions were performed using the APE package in R, and a maximum likelihood approach. All variables were logged-transformed before ancestral state inference; however, untransformed values are presented in the figures. Reconstructions of scores on PCA dimensions 1 and 2, absolute minimum temperature during the coldest month, and absolute precipitation over the driest month are presented for the whole *A. semialata* phylogeny. These temperature and precipitation variables were chosen to be analysed because they have been identified in the literature to drive C_4 distribution patterns and they correlate strongly with the PCA dimensions. Reconstructions of other environmental variables are only presented for the MRCAs of *A. semialata* and the semialata and eckloniana clades.

Results

Phylogenetic relationships and dispersal history

Accessions that were assigned to the species *A. semialata* based on morphological characters all formed a strongly supported monophyletic group sister to *A. angusta* (Figure 4.3). Three highly supported clades are reconstructed within *A. semialata*. One population originating from Tanzania (L02) and with a C_4 $\delta^{13}C$ signature was sister to all other individuals (clade F; Figures 4.3 and 4.4). The remaining accessions formed two sister clades, which correspond approximately to the previously defined subspecies and are consequently referred to as 'semialata' and 'eckloniana' (Figure 4.3 and 4.4). Based on a dataset that incorporates complete chloroplast genomes of *A. semialata* within a grass-wide dataset, the divergence of these two clades is estimated at 2.42 Ma (95% CI = 1.42 – 3.77), the first split within eckloniana clade at 1.53 Ma (95% CI = 0.71 – 2.7) and the first split within the semialata clade at 1.25 Ma (95% CI = 0.7 – 1.98; Figure 4.5). The eckloniana clade is composed of accessions considered to use C_3 (clade A) and C_3-C_4 intermediate (clades B and C) photosynthetic types (Figure 4.4), while the semialata clade includes only C_4 accessions (Figure 4.4, clades D-E). These results suggest that the photosynthetic types within *A. semialata* diverged early. However, a long stem separates *A. angusta* from *A. semialata*, representing a long time during which some degree of photosynthetic evolution may have occurred.

Within the eckloniana clade, accessions from southern Africa form a monophyletic group of accessions (Figure 4.4, clade A), implying a single, seed-mediated dispersal into southern Africa at the base of clade A. The central African accessions are split between two sub-clades of C_3-C_4 intermediate photosynthetic types (Figure 4.4, clades B - C). As clade C (C_3-C_4 intermediate) is sister to clade B (C_3-C_4 intermediate) and clade A (C_3), it may be inferred that their MRCA had a C_3-C_4 intermediate photosynthetic type as well (Figure 4.3 and 4.4).

This strongly contrasts with the semialata clade, which, despite a more recent common ancestor, covers the tropical and subtropical regions of Africa, Asia, and Oceania (Figure 4.6; Table 4.1). In this group, clade E is endemic to mainland Africa, with early splits separating central African accessions and more recent

splits leading to southern, western, and northern African accessions (Figure 4.6; Table 4.1). The first split in clade D separates Madagascan from Asian and Oceania accessions, suggesting a single seed-mediated migration outside mainland Africa (Figure 4.6; Table 4.1). Long distance dispersal across the Indian Ocean is often observed and might have occurred via previously emerged islands (Warren *et al.*, 2010). As the sister clades D and E are composed entirely of C_4 accessions, it can be inferred that their MRCA used the C_4 photosynthetic pathway (Figure 4.4). There are two distinct groups of Australian accessions descending within clade E, corresponding to Northern Territory and Queensland accessions, which indicates *A. semialata* seeds migrated either into or out of Australia on at least two occasions (Figure 4.4).

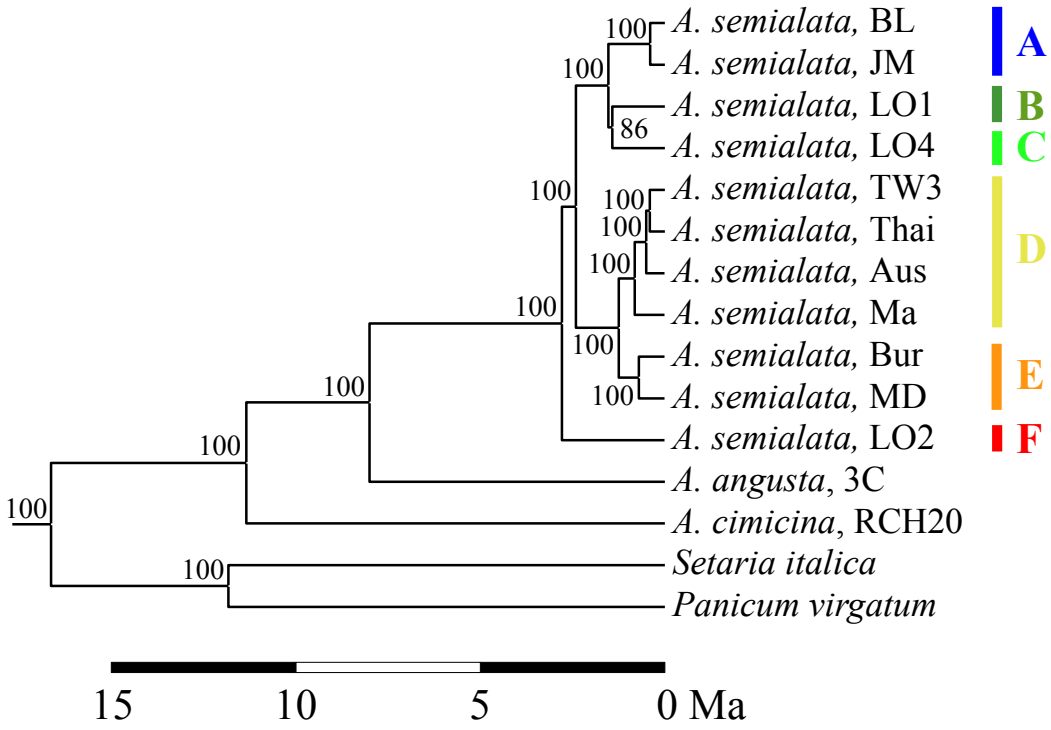


Figure 4.3 Phylogenetic relationships based on complete chloroplast genomes. This tree was obtained through Bayesian inference on a family-wide dataset. The part of the tree corresponding to the Paniceae tribe is shown. Branch lengths are proportional to estimated divergence times, in million years (Ma). The posterior distribution of divergence times is shown in Figure 4.5 for key nodes. Bayesian support values are indicated near branches. The main clades of *A. semialata* are delimited on the right. Clade identifiers are consistent throughout this chapter.

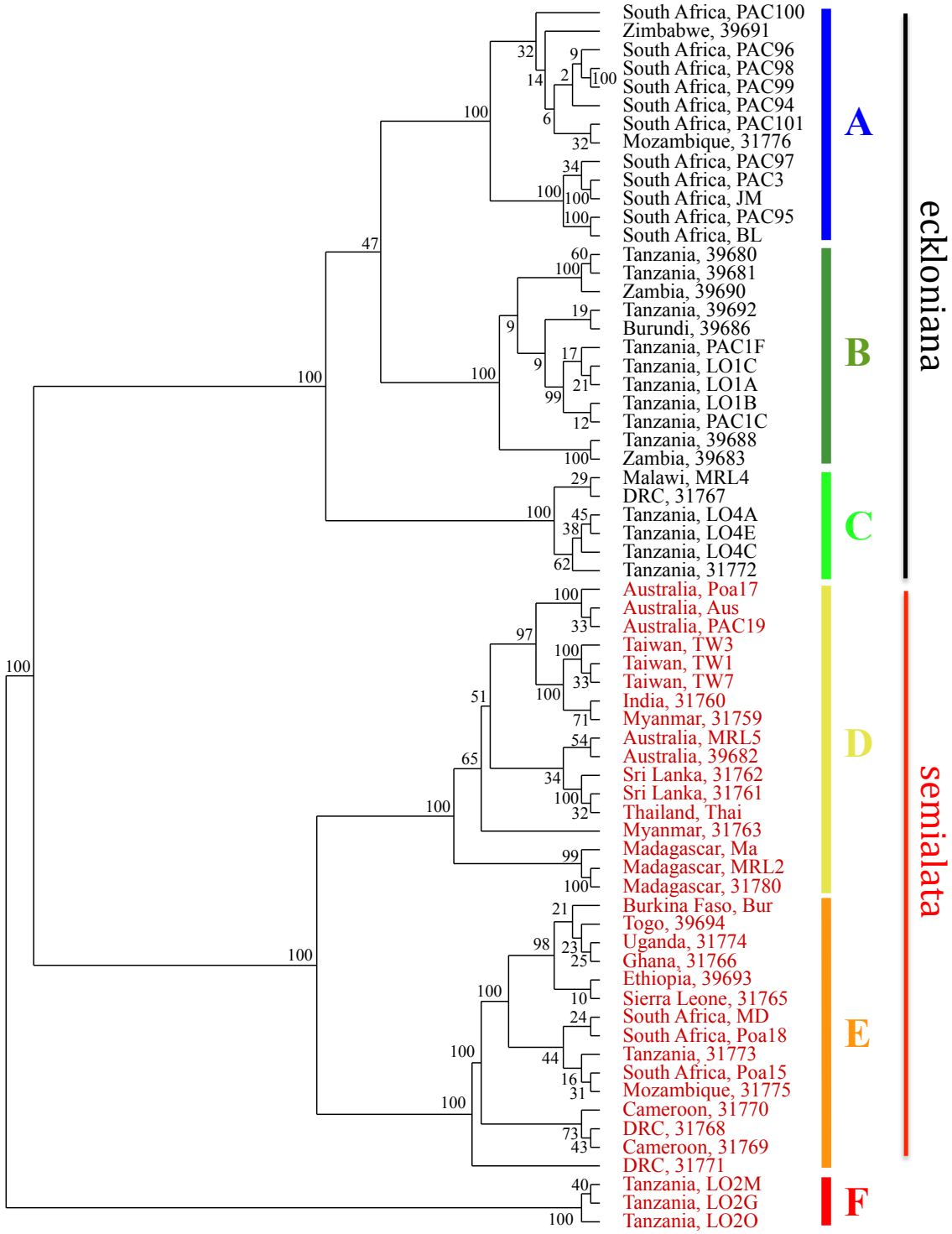


Figure 4.4 Phylogenetic relationships among *Alloteropsis semialata* accessions. This tree was obtained through Bayesian inference on chloroplast markers, and branch lengths are proportional to estimated divergence time, in arbitrary time units. For each accession, the country of origin is indicated, in red for accessions that grew using the C_4 pathway. The main clades are delimited on the right and color-coded. Bayesian support values are indicated near branches.

Accessions originating from central Africa are identified across the congeners and the three *A. semialata* clades (Figure 4.3 and 4.4, clades F, semialata, and eckloniana; Table 4.1). This suggests the MRCA of all *A. semialata* accessions existed in central Africa and lineages migrated away from there only after the physiological split between the semialata and eckloniana clades (Figure 4.6).

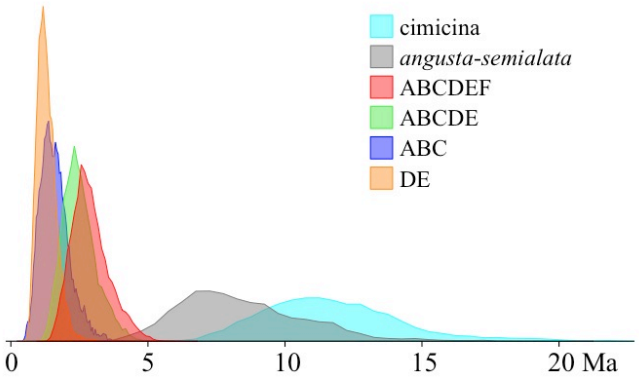


Figure 4.5 Posterior distributions of divergence times for key nodes within *Alloteropsis*. The divergence times were obtained with a Bayesian dating method implemented in Beast, and represent all trees sampled after the burn-in period over two independent analyses. The estimated times are given in millions of years ago (Ma).

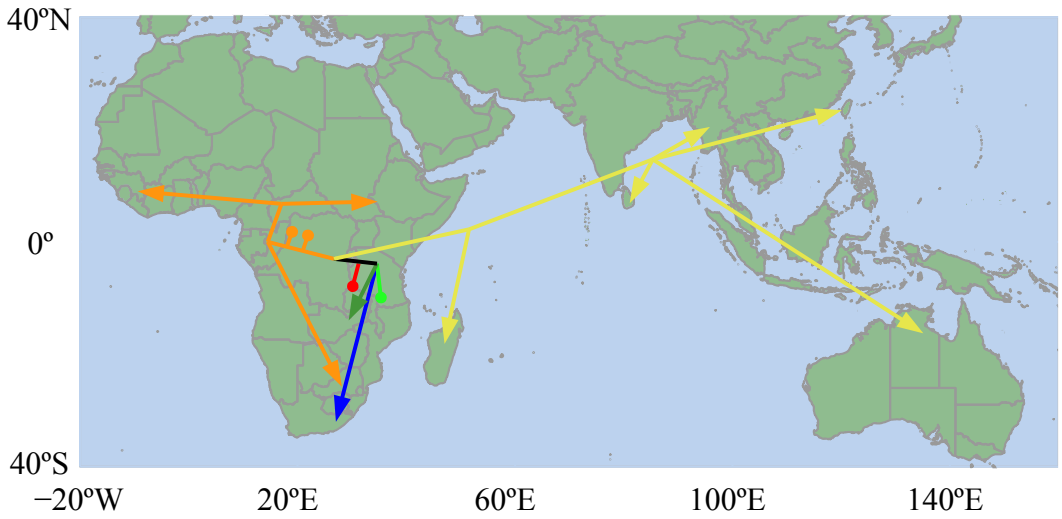


Figure 4.6 Inferred dispersal history of *Alloteropsis semialata*. The phylogeographic tree is approximately projected on the geographical space. Basal branches are in black and other clades are color-coded as noted in Figure 4.3 and 4.4.

Niche evolution

The niche of eckloniana accessions (Figure 4.7a, clades A-C) falls entirely within the semialata niche space (Figure 4.7a, clades D-E), but is only a subset of it. The broad semialata niche encompasses the extremes along both PCA dimensions. In contrast, the distribution of the C_3 clade A is strongly skewed toward negative values on dimensions 1 and positive values on dimension 2. The C_3-C_4 intermediates (clades B-C) cluster primarily around zero along dimensions 1 and 2 (Figure 4.7a).

Differences between photosynthetic types are best explained by dimension 1 of the PCA (Figure 4.7). This dimension explains a third of the variation in the dataset and was primarily driven by variables that describe minimum temperature, variability of temperature, vapour pressure, and altitude (Figure 4.7; Table 4.2). Members of clade A are more likely to inhabit high altitude, low vapour pressure, and seasonally cold environments than the humid, low altitude, and consistently warm environments where C_4 plants were often found. Both PCA dimensions explain variation within photosynthetic types (Figure 4.7). Variation along dimension 2 was primarily explained by water availability, including variables that describe energy, evaporation, and precipitation (Figure 4.7; Table 4.2).

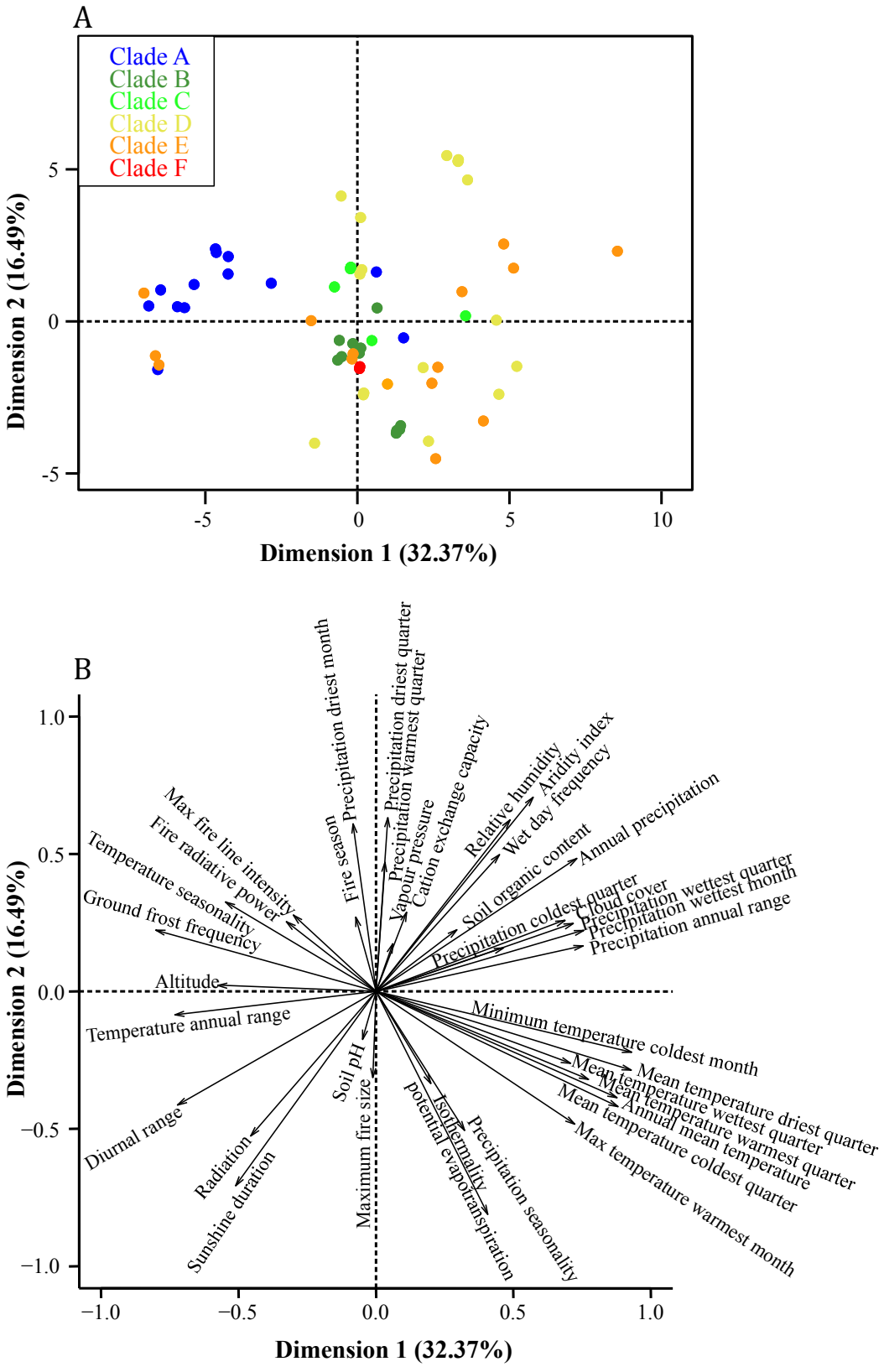


Figure 4.7 Principal component analysis (PCA) on environmental variables. (A) The samples are plotted in the first two dimensions of the PCA, and colored according to their phylogenetic clade. (B) The contribution of each environmental variable to the two axes is indicated.

Table 4.2 Coefficients for environmental variables significantly correlated with dimensions 1 and 2 of the PCA.

Variable	Coefficient	*p*-value
Dimension 1		
Min temperature of coldest month	0.93	0.00E+00
Mean temperature of driest quarter	0.93	0.00E+00
Mean temperature of coldest quarter	0.88	0.00E+00
Annual mean temperature	0.88	0.00E+00
Mean temperature of warmest quarter	0.77	2.35E-14
Precipitation of wettest month	0.76	2.03E-13
Precipitation annual range	0.75	2.61E-13
Annual precipitation	0.73	3.11E-12
Max temperature of warmest month	0.72	6.58E-12
Precipitation of wettest quarter	0.72	1.10E-11
Mean temperature of wettest quarter	0.71	2.96E-11
Cloud cover	0.69	1.74E-10
Global aridity index	0.57	5.02E-07
Relative humidity	0.49	3.18E-05
Precipitation of coldest quarter	0.47	8.31E-05
Wet day frequency	0.45	1.50E-04
Global potential evapotranspiration	0.41	6.77E-04
Precipitation seasonality	0.32	8.23E-03
Topsoil organic content	0.29	1.62E-02
Max fire line intensity	-0.30	1.42E-02
Fire radiative power	-0.33	7.56E-03
Radiation	-0.46	1.22E-04
Sunshine duration	-0.51	1.20E-05
Temperature seasonality	-0.55	1.99E-06
Altitude	-0.58	4.35E-07
Mean diurnal range	-0.72	8.01E-12
Temperature annual range	-0.73	3.02E-12
Ground frost frequency	-0.80	6.52E-16

Variable	Coefficient	*p*-value
Dimension 2		
Global aridity index	0.71	3.50E-11
Precipitation of driest quarter	0.63	1.29E-08
Relative humidity	0.62	2.20E-08
Precipitation of driest month	0.61	5.49E-08
Wet day frequency	0.50	2.17E-05
Annual precipitation	0.48	3.93E-05
Precipitation of warmest quarter	0.47	7.03E-05
Temperature seasonality	0.33	7.75E-03
Cation exchange capacity of topsoil	0.29	1.93E-02
Max fire line intensity	0.28	2.42E-02
Length of fire season	0.27	2.84E-02
Cloud cover	0.26	3.63E-02
Fire radiative power	0.25	4.02E-02
Precipitation of wettest quarter	0.25	4.56E-02
Mean temperature of wettest quarter	-0.26	3.36E-02
Mean temperature of driest quarter	-0.29	1.93E-02
Maximum fire size	-0.31	1.03E-02
Precipitation of warmest quarter	-0.32	8.61E-03
Isothermality	-0.34	5.86E-03
Annual mean temperature	-0.39	1.31E-03
Mean diurnal range	-0.41	5.45E-04
Mean temperature of coldest quarter	-0.42	4.50E-04
Max temperature of warmest month	-0.48	3.94E-05
Precipitation seasonality	-0.51	1.42E-05
Radiation	-0.53	5.49E-06
Sunshine duration	-0.71	3.00E-11
Global potential evapotranspiration	-0.81	1.30E-16

Reconstruction of ancestral environmental states

Ancestral state reconstructions of the two PCA dimensions place the MRCA of *A. semialata* at nearly the centre of dimensions 1 and 2 (Figure 4.8). Both dimensions remain near zero at the MRCAs of the semialata and eckloniana clades, and diversify only in more recent lineages (Figure 4.8). This suggests that overall, ecological diversification occurred after the diversification of photosynthetic types. Not until the MRCA of clade A does dimension 1 deviate away from zero in the negative direction (Figure 4.8). This suggests that the C_3 clade A may have emerged under high altitude, frequent frost, low vapour pressure, and seasonally variable temperature conditions. The presumed C_3-C_4 intermediate clades remained near zero on dimension 1 and dimension 2 overall (Figure 4.8). However, a subset of the clade B becomes strongly negative on dimension 2, suggesting these plants may have evolved a tolerance to low nutrient, arid, fire prone, open habitats. The C_4 clades were quite variable along dimensions 1 and 2, however, strong phylogenetic trends are less apparent on dimension 2. This suggests that C_4 plants may be able to tolerate a wide range of nutrient, precipitation, fire, and light environments.

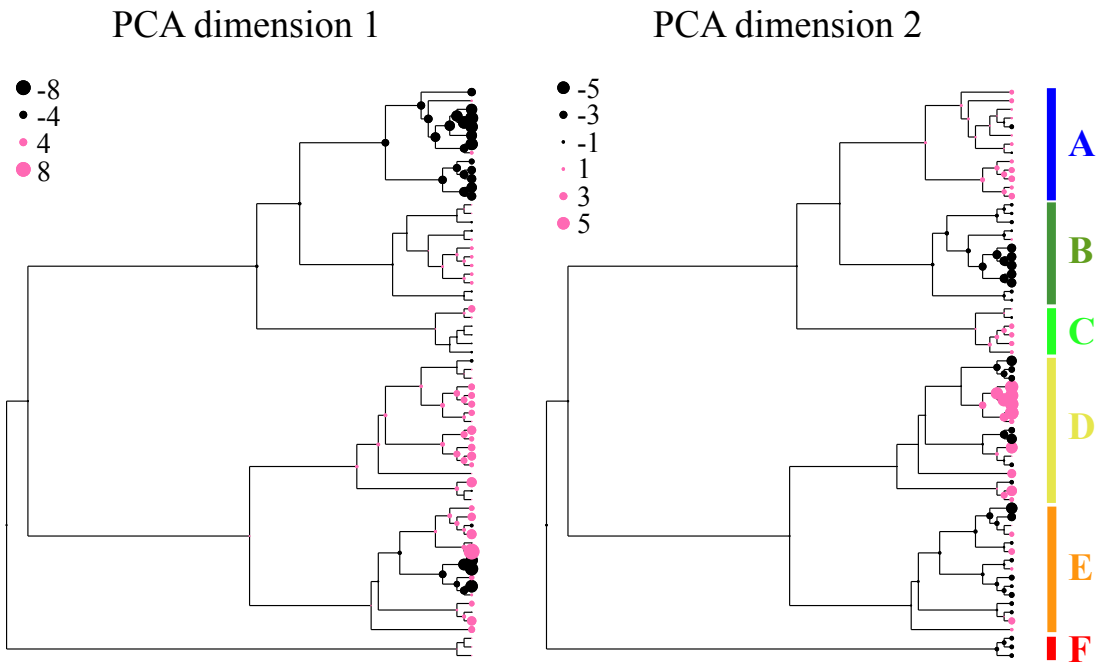


Figure 4.8 Dispersal across the environmental space inferred along the phylogeographic tree. Dot size is proportional to the absolute values along the first two dimensions of the PCA, as observed for tips and inferred for ancestral nodes. Negative values are in black and positive values in pink. The main clades are indicated on the right.

Reconstructions of the ancestral minimum temperature conditions indicate that the MRCAs of *A. semialata*, semialata, and eckloniana were similar (9.6, 10.1, and 7.9°C, respectively; Figure 4.9), suggesting seasonally cold temperatures were not a strong driver of photosynthetic diversification in this system, unless it was an ecological precondition of C_4 photosynthesis. The mainly C_3 clade A diversified under a seasonally colder climate (4.0°C minimum temperature), while MRCAs of the mainly C_3-C_4 intermediate lineages (clades B and C) remained in warmer areas (8.8 and 10.0°C, respectively; Figure 4.9). The C_4 lineages primarily diversified into warmer areas, but are currently distributed across a broad range of minimum temperatures (-1.1 to 21.0°C; Figure 4.9), suggesting these C_4 plants are able to tolerate a wide range of minimum temperatures.

Reconstructions of the ancestral precipitation conditions were similar in the MRCAs of the semialata and eckloniana clades (4.5 and 2.3 mm rainfall in the driest month, respectively; Figure 4.9). This suggests that the diversification of photosynthetic type took place in similarly seasonally arid conditions. The MRCA of the mainly C_3 clade A is inferred in less seasonally arid environments (8.6 mm rainfall in the driest month), while the C_3-C_4 intermediate lineages (clades B and C) remained in areas with very little dry season precipitation (0.9 and 2.4 mm rainfall in the driest month at MRCAs, respectively). Many of the current day semialata accessions inhabit less seasonally arid environments, but their ancestors moved into these conditions recently. Comparisons between the MRCA of the *A. semialata* species with those of the semialata and eckloniana clades show that diversification of photosynthetic type occurred in similar environments and that ecological diversification occurred only after the physiological diversification (Figure 4.10).

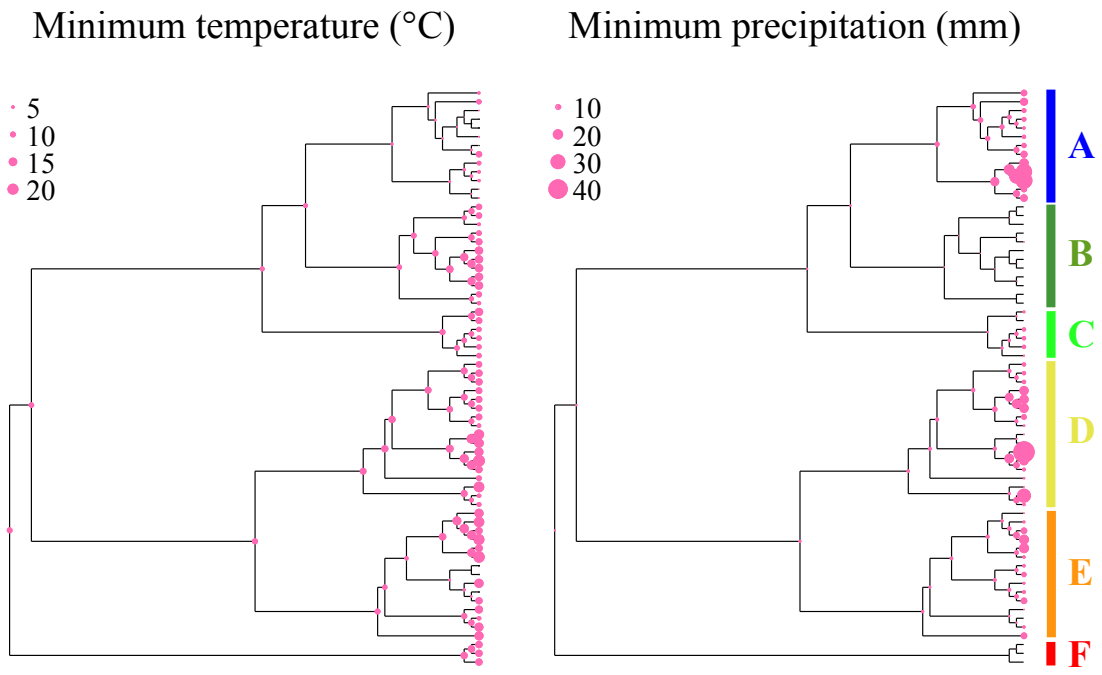


Figure 4.9 Dispersal across temperature and precipitation niche space inferred along the phylogeographic tree. The dots are proportional to the minimum temperature of coldest month and total precipitation of driest month, as observed for tips and inferred for ancestral nodes. The main clades are indicated on the right.

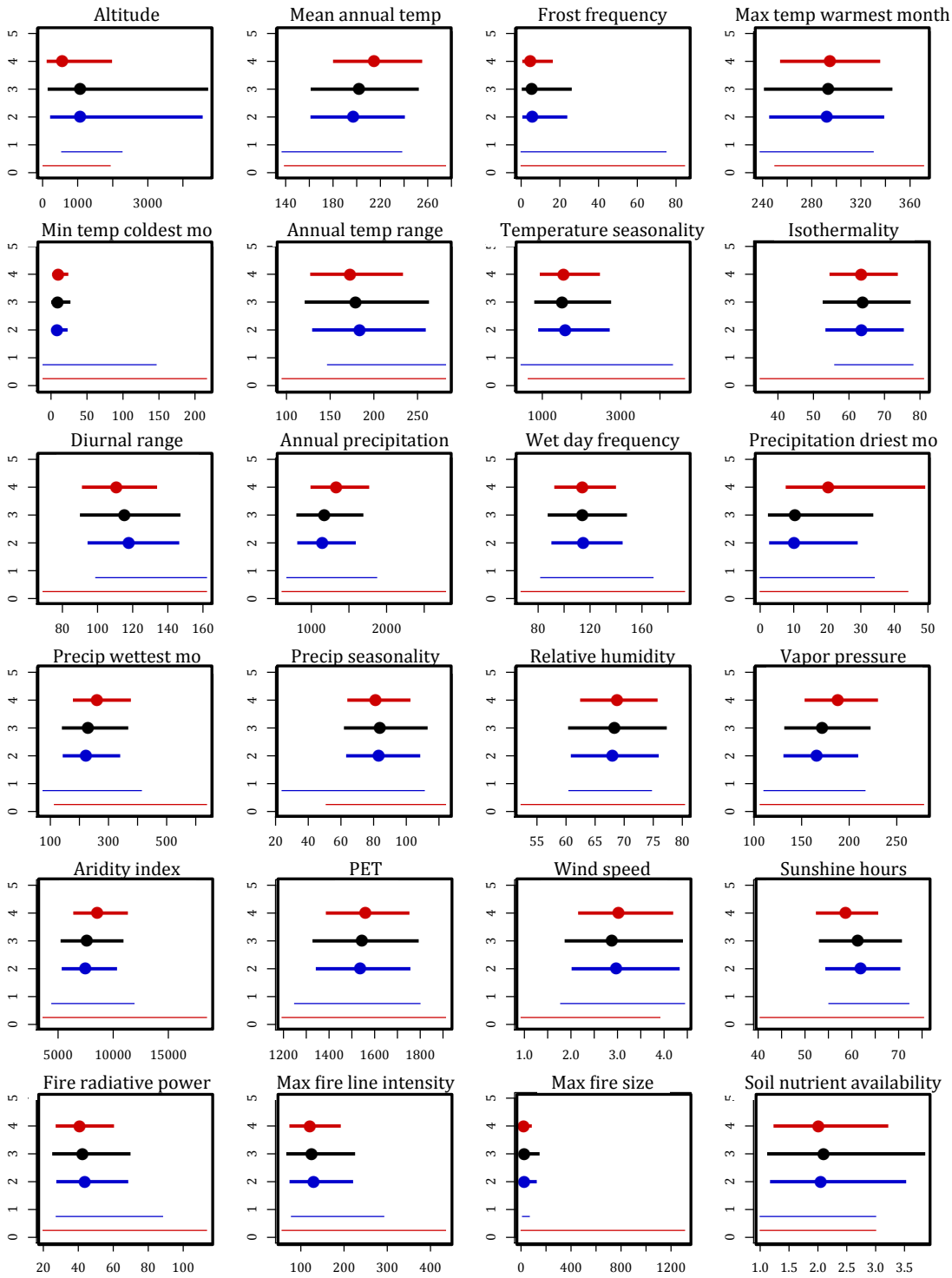


Figure 4.10 Comparisons of ancestral environmental state reconstructions for the *Alloteropsis semialata* (black), C_4 (red, bold), and non-C_4 (blue, bold) lineage MRCAs with the current distribution of the C_4 (thin red) and non-C_4 (thin blue) clades. 95 percent confidence intervals are presented. Environmental variables are described in Appendix 4.2. The presented temperatures have been multiplied by 10.

Soil environment

Plants using C_4 photosynthesis are distributed across a wide variety of soil types, spanning soils with high and low organic matter, mineral contents, pH, and water holding capacity (Figure 4.11, clades D-F; Appendix 4.3). Accessions in the mainland African and Madagascan/Asian/Australian clades D and E were collected in similar soil types despite being geographically quite distant (Figure 4.11, clade D and E). Plants from the mainly C_3 clade A only grew in medium to low quality soils, with characteristically high acidity and low nutrients (Figure 4.11, clade A; Appendix 4.3). Accessions in the mainly C_3-C_4 intermediate central African clade C were primarily growing in rich soils characterized by high organic matter, chemical fertility, good structural stability, and only slight acidity (Figure 4.11, clade C; Appendix 4.3). Members of the mainly C_3-C_4 intermediate central African clade B grew on diverse soils with both high and low chemical fertility, neutral and acidic soils, and weak to strong effects of weathering (Figure 4.11, clade B; Appendix 4.3).

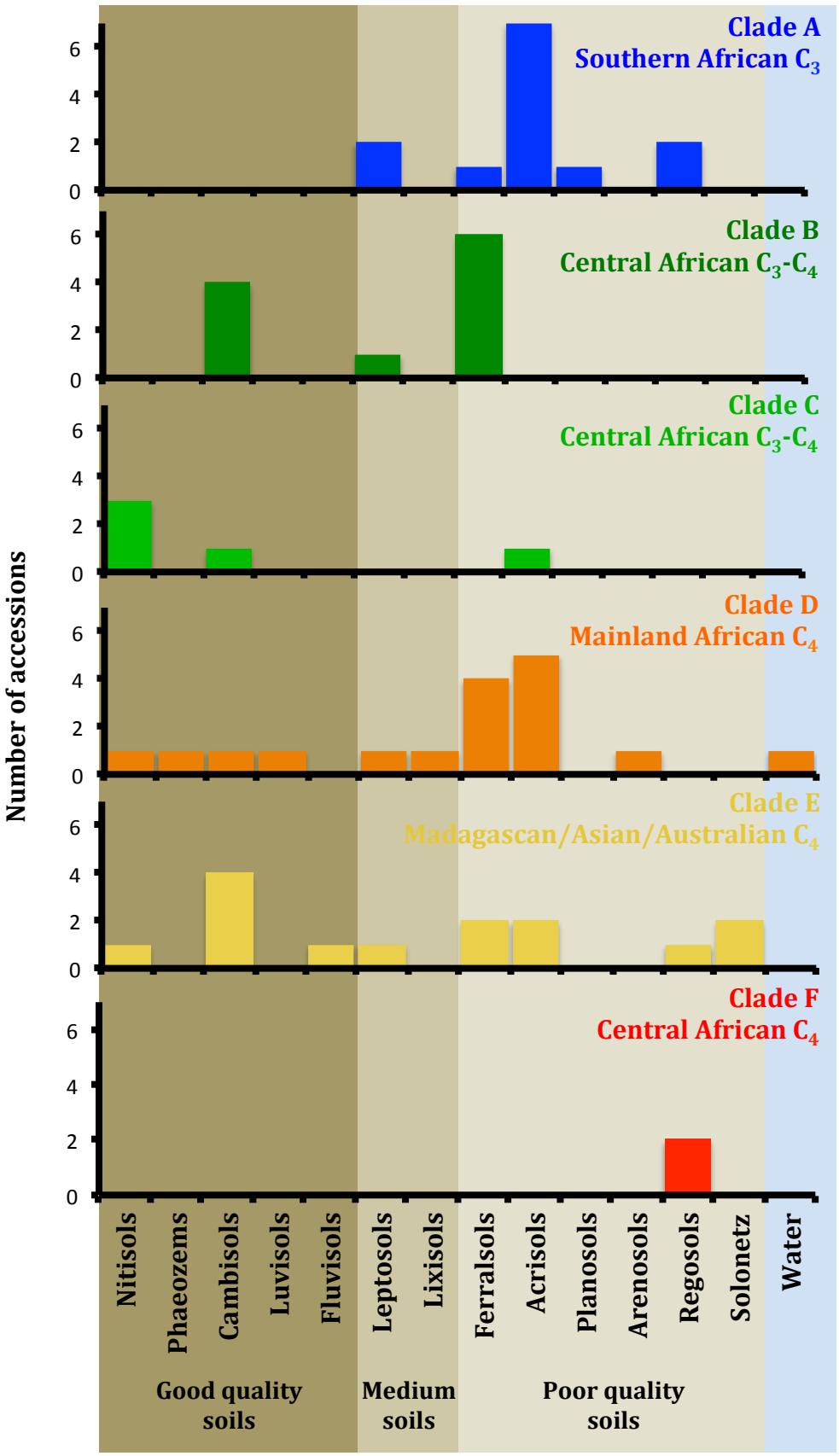


Figure 4.11 Frequency distributions of *Alloteropsis semialata* accessions presented and color-coded by main clade across a range of soil types. Appendix 4.3 provides descriptions of these soil types.

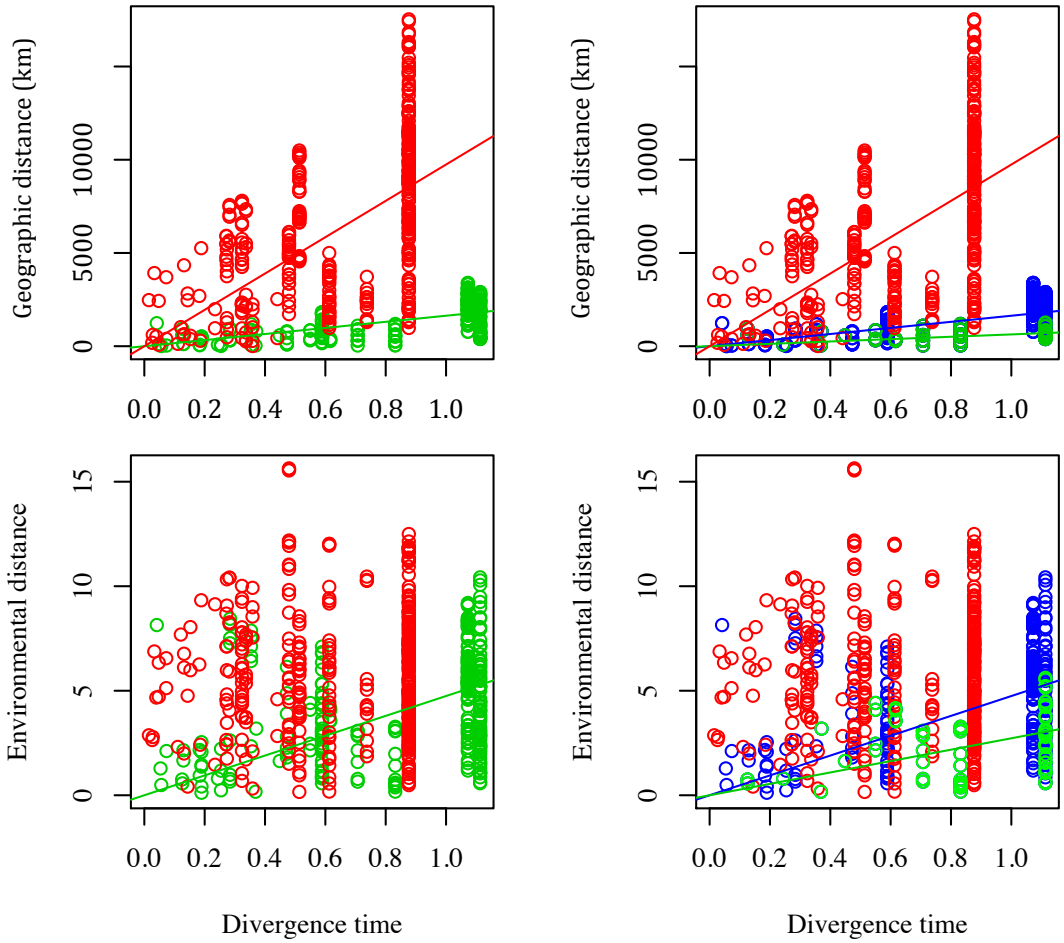
Rates of dispersal

Statistical comparisons among pairwise distances were made to evaluate the rates of dispersal across geographic and environmental spaces. These included geographical distances, distances across the PCA space as determined by the first two axes, and divergence times. As the phylogeographic tree was not resolved with confidence near the tips, all analyses were repeated with topologies sampled from the posterior distribution, but the results remained unaltered (Appendix 4.1). Patterns of isolation by distance (as inferred from correlations between pairwise geographic distances and divergence times) were detected in both the semialata clade and the eckloniana clade (Table 4.3). However, the slope of the regression of geographic distances against divergence times is nearly six times greater in semialata clade than in the eckloniana clade (9,743 km per time unit versus 1,634 km per time unit; Figure 4.12; Table 4.3), which indicates that while dispersal limits the geographical distribution in both clades, the limitation is stronger in the non-C_4 eckloniana clade.

Environmental distances were significantly correlated to divergence times within the eckloniana clade (Table 4.3), indicating a gradual migration into different conditions. However, these environmental distances are also correlated to geographic distances (Table 4.3). Once this spatial autocorrelation is taken into account, the relationship between environmental distances and divergence times is no longer significant (Table 4.3). The geographic effect on environmental distances, however, is mainly driven by the single dispersal of clade A into southern Africa, which represents an important leap in both geographic and environmental spaces. When clade A is excluded, environmental distances within clades B and C are no longer explained by geographical distances (Table 4.4), which rules out spatial autocorrelation of environmental values in this case. However, these environmental distances are still explained by divergence times (Table 4.4), which demonstrate that clades B and C transitioned gradually into different environments through time.

Environmental distances were neither correlated to geographical distances nor to divergence times within the semialata clade (Table 4.3). This shows that C_4 accessions were able to migrate rapidly to diverse environments from their early

diversification (Figure 4.6), so that their ecology is neither explained by their genetic relatedness, nor constrained by their dispersal ability.



4.12 Comparison of geographical (top) and environmental (bottom) distances by divergence times in millions of years ago. On the left side, distances between pairs of non-C_4 individuals from the eckloniana clade are in green and those between pairs of C_4 individuals from semialata clade are in red. On the right side, eckloniana accessions are color coded by those from clade A (blue) and those from clades B and C (green). Regression lines forced to the origin are shown for significant relationships, identified by Mantel tests.

Table 4.3 Mantel permutation test Z and p-statistics showing similarity between matrices of divergence time (time), geographic distance (geo, G), and environmental distance (env, E) in the semialata and eckloniana clades.

	semialata				eckloniana			
	Mantel Test		Linear Model		Mantel Test		Linear Model	
	Z	*p*	slope	R^2	*Z*	*p*	slope	R^2
Time x Geo	1985689	<0.001	9743	0.806	419303	<0.001	1634.1	0.819
Time x Env	1564	0.545	-	-	1218	<0.001	4.7	0.778
Geo x Env	15740117	0.357	-	-	2250113	<0.001	<0.1	0.814
(G x E) x Time					< 1	0.847	-	-

Table 4.4 Mantel permutation test Z and p-statistics showing similarity between matrices of divergence time (time), geographic distance (geo), and environmental distance (env) in the central African eckloniana sub-clades.

	Central African eckloniana			
	Mantel Test		Linear Model	
	Z	*p*	slope	R^2
Time x Geo	34916	0.201		
Time x Env	151	0.048	2.7	0.742
Geo x Env	98271	0.860		

Discussion

Ancestral photosynthetic state of Alloteropsis semialata

While inferring the ancestral state based on a species tree must be subject to caution (Christin *et al*., 2010), a careful analysis of the evidence suggests the MRCA of *A. semialata* was a C_3-C_4 intermediate plant. The C_2 cycle is a strong enabler of C_4 photosynthesis (Mallmann *et al*., 2014) and is believed to be a required step in the transition from C_3 to C_4 (Heckmann *et al*., 2013; Sage *et al*., 2014; Williams *et al*., 2013). Thus, the *A. semialata* C_4 lineage very likely emerged from a C_2 ancestor at some point, whether that was at the semialata/eckloniana diversification or more basal in the genus. It seems that C_4 photosynthesis emerged multiple times within *Alloteropsis* (*e.g.*, Christin *et al*., 2010, 2012), which could be explained by a C_2 MRCA of the genus (Figure 4.1c), serving as a precursor that was co-opted several times to evolve C_4 photosynthesis, as seen in other groups (Christin *et al*., 2011a; 2014b). This hypothesized C_2 lineage may have independently given rise to the different forms of C_4 photosynthesis used by the *Alloteropsis* Cordichloa group (NAD-ME) and *A. angusta* (NADP-ME), and *A. semialata* (PCK +NADP-ME) species (Christin *et al*., 2010).

The first divergence in the chloroplast tree separates clade F from the semialata-eckloniana clade and, as such, plants in this clade may reflect the ancestral physiological state of *A. semialata*. Plants in clade F were collected in central eastern Africa and were assumed to use C_4 photosynthesis because of their high $\delta^{13}C$ values (-13.2, -12.7, and -11.4 ‰). However, leaf anatomy in these plants is similar to C_3-C_4 intermediate *A. semialata* plants (Figure 4.13). In particular, their inner and outer bundle sheaths sizes are of relatively similar widths, a phenotype consistent with C_3-C_4 intermediate plants in this species (Chapter 3), but their inner bundle sheath cells are slightly smaller and outer bundle sheaths much larger than other C_4 plants (Figure 4.13). Chapter 3 clearly shows that variations in these bundle sheath traits strongly correlate with carboxylation efficiency, CO_2 compensation point, and O_2 inhibition in this species. Thus, it is plausible that plants in clade F are physiologically more similar to C_3-C_4 intermediates than C_4 plants. These plants may be running a C_2 cycle complemented by a strong C_4 cycle, which would explain the C_3-C_4 intermediate anatomy but C_4 $\delta^{13}C$ signature. In

other words, the biochemistry of the clade F plants may be rather C_4-like, while their anatomy never fully transitioned to more typical C_4 phenotypes. Although more work is needed to confirm this hypothesis, if proved true, it may confirm a C_3-C_4 intermediate ancestral state for the copious photosynthetic variation expressed in *A. semialata*.

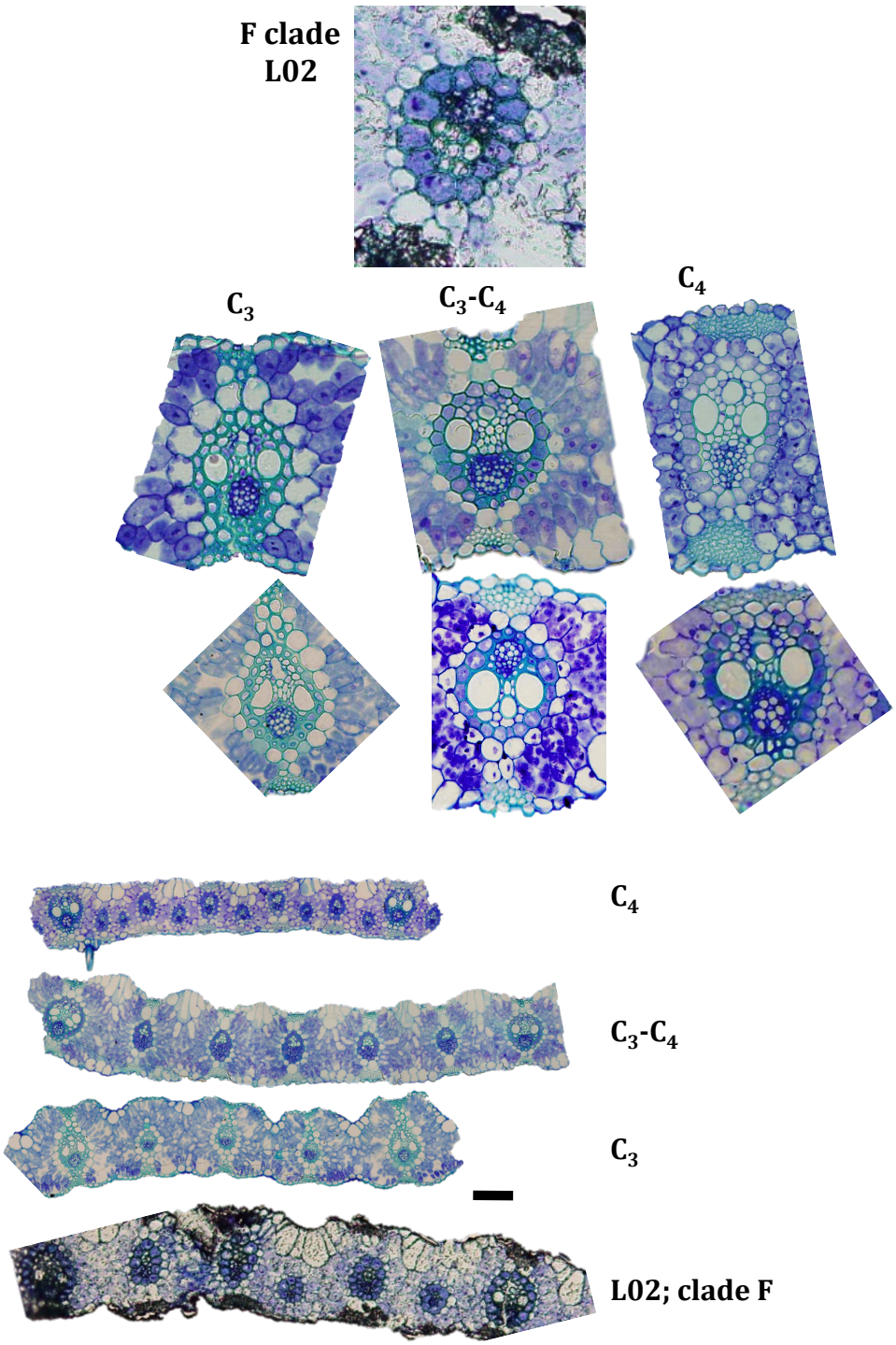


Figure 4.13 Comparison of cross-sections from plants in clade F and those of typical C_3, C_3-C_4, and C_4 plants. Note the similar bundle sheath and veination patterns between clade F and C_3-C_4 plants.

The evolution and distribution of C_4 photosynthesis

With the exception of clade F, as discussed above, all of the C_4 individuals in this study fell into the same semialata clade (Figure 4.4; clade DE). This group, expressing a strong C_4 phenotype, likely evolved within central Africa 2.4 million years ago (Figure 4.6). The fact that photosynthetic types match haplotype groups that are over two million years old indicate that there was little gene flow between eckloniana and semialata, although occasional integration of nuclear genes cannot be excluded. The phylogenetic patterns indicate that the MRCA of both the semialata and eckloniana clades existed in central Africa, yet one lineage used the C_4 pathway and the other used C_3 or C_3-C_4 intermediate photosynthesis and they were reproductively isolated soon after this physiological split. These results suggest that the C_3-C_4 intermediate phenotype is the result of natural selection, rather than hybrid crosses between C_3 and C_4 parents, as was discussed in Chapter 3.

Following the emergence of the main C_4 lineage, members of the semialata group dispersed broadly. One lineage moved eastward out of mainland Africa into Madagascar and southern Asia (Figure 4.6). From there, it swiftly migrated across Asia and onto Australia. From the centre of origin, a second lineage migrated west across the current Democratic Republic of Congo where it rapidly radiated into southern, western, and northern Africa (Figure 4.6).

Ecological drivers of physiological diversification

Based on the phylogenetic relationships inferred here, the common ancestor of the semialata and eckloniana clade originated from central Africa, and the early members of these two clades persisted in this area for a considerable length of time. The initial divergence of these clades might have been caused by geographic isolation, in a region where mountain ranges, lakes, and rifts provide barriers to dispersal. Interestingly, the divergence of photosynthetic types did not directly lead to obvious modifications of the ecological niche, as assessed by climatic and edaphic variables. Representatives of the different clades and photosynthetic types can still be found in habitats within central eastern Africa that match those inferred for their common ancestor (Figure 4.8). Indeed, some C_4 and non-C_4

members of clades B, C, E, and F can be found in wooded habitats of Tanzania, Congo, and Cameroon. In these seasonally dry deciduous forests, photorespiration is predicted to vary throughout the year as leaf fall drastically increases sunlight, temperature, and aridity at ground level (reviewed in Chapter 3). In recent geological history, the range of woodlands and savannas in central Africa varied as a function of the glacial cycles, but wooded savannas and dry forests were constantly present (Rommerskirchen *et al.,* 2006; Bragg *et al*., 2013). These environments are at the limit of those where non-C_4 plants grow (Figure 4.7), and mutations providing a more C_4-like physiology might have been selected for in these habitats where the persistence of more C_3-like or C_3-C_4 intermediate phenotypes is still possible. Based on these investigations, it may be speculated that C_4 physiology initially emerged in environments that advantage different photosynthetic types across the season or across small-scale ecological variations (*e.g.*, densely versus lightly wooded habitats), where isolated populations could explore different parts of the phenotypic landscape as a function of random mutations.

C_4 consequences for the ecological niche

The ecological similarity between the early members of the eckloniana and semialata groups contrasts with the current distribution of the two clades. On average, the C_3, C_3-C_4 intermediate, and C_4 accessions inhabit characteristically different environments today, but this ecological diversification occurred after the diversification of photosynthetic type. Indeed, extant accessions of the semialata clade inhabit environments ranging from the tropics to southern latitudes and cover a broad range of temperatures, precipitations, soil conditions, light intensities, and fire regimes (Figures 4.2 and 4.7). Elucidation of the phylogeographic history shows that these varied habitats were colonized rapidly after the divergence of photosynthetic types, while the otherwise similar non-C_4 members of the eckloniana clade remained confined to a narrower set of environmental conditions over the same length of time (Figures 4.8, 4.9, and 4.12). This indicates that, when other factors affecting the ecology of individual plant species remain similar, C_4 photosynthesis acts as a niche opener, and does not simply shift the ecological niche. The semialata lineage showed isolation by

distance but little evidence of niche evolution, which suggests these C_4 plants did not have to adapt as they spread, but instead tolerated diverse environments as they were encountered. This suggests that plants using C_4 photosynthesis may inherently be able to tolerate a wider range of conditions more quickly than plants using the C_3 pathway.

One explanation for this observation may be that C_4 plants are more phenotypically plastic than C_3 plants, allowing them to adapt to novel environments within a single generation. However, Sage and McKown (2006) comprehensively reviewed phylogenetically diverse taxa to conclude that plants using the C_4 pathway should have less acclimation potential than C_3 plants, as their phenotypes would be canalized and specialized to C_4 function. This would imply that C_4 plants would be more limited by novel environments than C_3 plants, which is not the case with the widespread C_4 *A. semialata*.

Another explanation is that traits that evolved to optimise the C_4 system also conveyed advantages in diverse environments. For example, the C_4 stomata phenotype may enable existence in warm, bright, and arid habitats. Because the CO_2 concentrating mechanism allows C_4 plants to maintain higher intercellular CO_2 concentrations than C_3 plants under the same conditions, C_4 plants will maintain smaller stomatal apertures and consequently lose less water than C_3 species, in general (Taylor *et al*., 2012). This allows C_4 plants to sustain higher rates of photosynthesis and lower stomatal conductance than C_3 plants, allowing them to operate with greater water use efficiency (Pearcy and Ehleringer, 1984). Furthermore, by maintaining reduced stomatal apertures, C_4 plants experience smaller diurnal variations in leaf water potential, leading to greater capacity for hydraulic supply relative to transpiration demands, and decreased risk of embolism formation (Taylor *et al.,* 2010).

C_4 leaf anatomy also contributes to reduced water loss and consequently persistence in warm, bright, and arid habitats (Kocacinar and Sage, 2003; 2004; Sack and Scoffoni, 2013; Sack *et al.*, 2013). For instance, some C_4 species have smaller and shorter xylem vessels with lower hydraulic conductivity than C_3 species, allowing these C_4 plants to be less vulnerable to xylem cavitation

(Kocacinar and Sage, 2003). Furthermore, C_4 grasses have smaller interveinal distances than C_3 grasses (Ogle, 2003; Christin *et al.,* 2013a), which may increase hydraulic conductance in C_4 plants (Kawamitsu *et al.,* 1985; Ogle, 2003). The smaller interveinal distance and larger bundle sheath area characteristic of C_4 plants (reviewed in Chapter 2) allow for greater photon interception, and ultimately, shade tolerance in some C_4 species (Ogle, 2003). Finally, the high vein density of C_4 plants not only enables higher leaf hydraulic conductance, but may also improve phloem loading, provide tolerance to freeze-thaw embolisms, and help protect against herbivory (Caswell *et al*., 1973; Osborne and Sack, 2012; Sack and Scoffoni, 2013), all of which may convey survival in novel environments. Together, these attributes might have allowed members of the semialata clade to colonize contrasted environments without needing to evolve additional adaptations, explaining the rapid expansion across the Old World.

Indeed, the main consequence of C_4 photosynthesis is to decrease photorespiration, and thus increase the amount of CO_2 fixed per absorbed photon (Ehleringer and Björkmann, 1977). This enhances water- and nitrogen-use efficiencies (Ehleringer and Björkmann, 1977; Pearcy and Ehleringer, 1984), which could facilitate the colonization of drier and infertile habitats, as described above. However, it does not necessarily decrease success in fertile and wetter environments, where it can provide a competitive advantage by enabling faster growth (Monteith, 1978). Thus, the spectrum of habitats available to newly evolved C_4 plants probably increased as a direct effect of C_4 physiology. The diversity of ecological conditions tolerated by the C_4 accessions of *A. semialata* probably explains the more efficient dispersal of these plants, as has been found across multiple species (Slatyer *et al*., 2013). Indeed, their reproductive and dispersal strategies are not known to differ from the non-C_4 eckloniana clade. Instead, the capacity to survive in a broad range of environments following long distance dispersal event likely facilitated the colonization of distant regions, leading to the spread of these plants across three different continents (Figures 4.2 and 4.6).

Other adaptations lead to ecological diversification

The eckloniana lineage dispersed in a completely different pattern to the semialata lineage. While the C_4 semialata clade was dispersing across geographical and environmental spaces, members of the non-C_4 eckloniana clade continued evolving, emphasizing the importance of considering the variation within each photosynthetic type when inferring evolutionary processes. Indeed, non-C_4 lineages gradually came to colonize distinct environments, independently of geography in the case of the central African clades B and C. The gradual migration toward distinct habitats implies a continuous process of adaptation through natural selection. While the predominately C_3-C_4 intermediate clades B and C remained in central Africa, in habitats that presumably resemble the ancestral condition of the clade, members of clade A strongly deviated from these conditions and colonized colder regions in southern Africa (Figure 4.8 and 4.9). This southern dispersal also involved the migration from wooded habitats to open grasslands with leached, acidic soils, where C_3 *A. semialata* are very successful, as attested by their local abundance (Ellis, 1981). The South African C_3 *A. semialata* have acquired a cold adaptation mechanism for leaves to resist freezing (Osborne *et al.,* 2008), and are able to maintain photosynthetic capacity under drought conditions (Ripley *et al.,* 2007; Ibrahim *et al*., 2008). These capacities may have contributed toward their successful colonization of southern latitudes. C_4 photosynthesis, adopted by members of clade DE, and cold tolerance, present in clade A, might represent alternative novelties that allow the ecological expansion of tropical lineages. This pattern is already evidenced for the grass family as a whole, where distinct groups have evolved either C_4 photosynthesis or cold tolerance, both of which strongly increased diversification rates (Spriggs *et al.,* 2014). Our intraspecific investigations show that, while C_4 photosynthesis broadens the niche and allows rapid dispersal across environmental space, cold adaptation might be an alternative, but slower process that leads to a narrower realised niche in otherwise similar plants, although the fundamental niche might still encompass warmer conditions (Araújo *et al.,* 2013).

Loss of C_3-C_4 intermediacy

The chloroplast tree presented in this study establishes one mainly C_3-C_4 intermediate clade (C) that is sister to another mainly C_3-C_4 intermediate clade (B) and the one mainly C_3 clade (A). This phylogenetic pattern suggests that the C_3 photosynthetic type emerged from C_3-C_4 intermediate plants, which would be a reversal of the photosynthetic pattern commonly assumed. Thus, it may be inferred that any C_2 function associated with these C_3-C_4 intermediates was lost as the eckloniana lineage moved southward, leaving a more typical form of C_3 photosynthesis to colonize southern Africa while plants remaining in central Africa maintained C_3-C_4 intermediacy.

There are several costs associated with running the C_2 cycle. First, this pathway inherently requires photorespiration to operate that, despite the recycling of photorespiratory CO_2, still produces the toxic ammonia by-products and consumes ATP and NADPH. Second, establishment of the C_2 cycle requires construction costs to provision bundle sheath organelles and enzymes (Schuster and Monson, 1990). Finally, when C_2 plants initiate a C_4 cycle, the C_4 biochemical functions are often inefficient and increase nitrogen requirements (Monson *et al.,* 1986). This creates nitrogen sinks that offset the advantages gained by recycling photorespired CO_2 (Schuster and Monson, 1990).

Although the C_2 cycle is often considered a transitory step that facilitates the evolution of C_4 photosynthesis (Sage *et al*., 2012; Heckmann *et al*., 2013; Williams *et al.*, 2013; Mallmann *et al*., 2014), some C_2 lineages do not proceed to become C_4, but instead remain in a C_2 state, such as the extant C_2 eudicot *Mollugo nudicaulis* that evolved 20 million years ago (Christin *et al*., 2011a). This suggests that the C_2 cycle is advantageous enough to overcome its costs under certain environments, thus permitting it to continue for millions of years. In this study, the costs relative to advantages of maintaining C_3-C_4 intermediacy were probably low in central Africa, where plants that likely use a C_2 cycle persist, but high in southern Africa, where this function was lost.

Although minimum temperature and dry season precipitation did not apparently drive the C_3/C_4 physiological divergence, they seem to have contributed to the C_3-

C_4 intermediate to C_3 reversal. The MRCA of the eckloniana lineage existed in warm, seasonally dry conditions, however, the MRCA of the C_3-specific lineage migrated into distinctly cooler and less seasonal conditions. This suggests that the photorespiratory advantage of C_3-C_4 intermediate plants decreased as the C_3 lineage migrated to lower latitudes with colder temperatures and more dry-season rainfall than central eastern Africa. This is probably because the C_2 cycle primarily improves photosynthesis only at warmer temperatures and is little better than C_3 photosynthesis in cooler environments (Schuster and Monson, 1990; Monson and Jaeger, 1991; Teese, 1995).

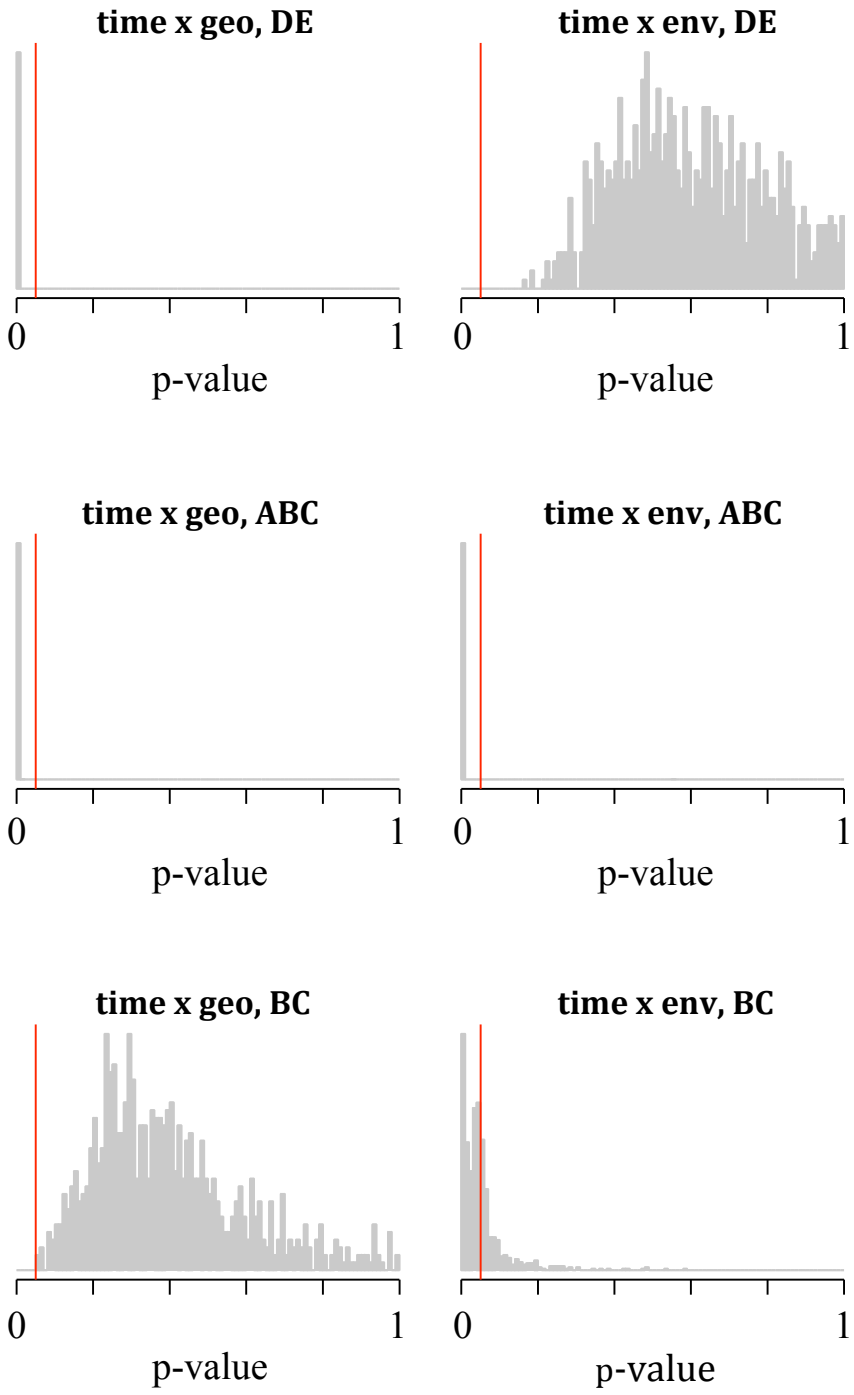
Soil quality may also be linked to the relative advantage of the C_3-C_4 intermediate and C_3 types. The C_3 plants in this study only grew in marginal to poor quality soils that lacked nutrients and structural stability (Figure 4.11). These typical South African soil types may not support nitrogen-demanding C_2 plants and, as such, natural selection may have favoured individuals that switched off the C_2 cycle. Thus, C_2 plants may be limited to both high-photorespiration environments and those with sufficient and available soil nitrogen. Plants using a C_2 cycle may be confined to a narrow niche of hot, seasonal, and fertile habitats, which would explain their limited abundance and geographic distribution.

Conclusions

Capitalizing on the variation that exists within a single species complex, this study is the first to characterize the ecological changes that directly follow the emergence of different photosynthetic types. The joint analysis of geographical and environmental dispersal histories within a phylogenetic context shows that C_4 photosynthesis does not initially result in a shift of the ancestral niche, but broadens this niche to cover a wider range of conditions, enhancing the success of occasional long distance dispersal events, and therefore increasing the geographic range. The variety of environments available to C_4 plants is also reflected in the ecological diversity observed among C_4 species, with different C_4 taxa found in very distinct environments that promote photorespiration in different ways (Sage *et al.*, 2012). Interspecific phylogeny-based analyses suggest that species using C_4 photosynthesis diversify across a wider range of environments than closely related

C_3 species (Christin and Osborne, 2014). However, individual taxa likely specialize in different environments after the initial evolution of C_4 physiology, through differential integration of the C_4 machinery with their growth and life-history traits (Christin and Osborne, 2014). Over time, this process leads to some C_4 taxa becoming specialized to environments that differ strongly from those in which they evolved, inflating the ecological differences between C_3 and C_4 photosynthesis and blurring the initial effects of differences in photosynthetic types.

Appendix 4.1 Mantel test results with topologies from the posterior distribution.



For each of the six Mantel tests that are potentially affected by the topology, the distribution of p-values across the 900 trees sampled each 200,000 generations after the burn-in period is shown. Vertical red bars indicate the 0.05 cut-off.

Appendix 4.2 Details of environmental factors included in this study.

Factor	Resolution	Units	Years	Source
Altitude				
Altitude	30-arc-seconds	m	1950-2000	WorldClim[1,2]
Temperature				
Annual mean temperature	30-arc-seconds	°C	1950-2000	WorldClim, Bioclim[1,2]
Ground frost frequency	10-arc-minutes	days	1961-1990	CRU[3,4]
Max temperature of warmest month	30-arc-seconds	°C	1950-2000	WorldClim, Bioclim[1,2]
Min temperature of coldest month	30-arc-seconds	°C	1950-2000	WorldClim, Bioclim[1,2]
Mean temperature of wettest quarter	30-arc-seconds	°C	1950-2000	WorldClim, Bioclim[1,2]
Mean temperature of driest quarter	30-arc-seconds	°C	1950-2000	WorldClim, Bioclim[1,2]
Mean temperature of warmest quarter	30-arc-seconds	°C	1950-2000	WorldClim, Bioclim[1,2]
Mean temperature of coldest quarter	30-arc-seconds	°C	1950-2000	WorldClim, Bioclim[1,2]
Temperature variability				
Mean diurnal range	30-arc-seconds	°C	1950-2000	WorldClim, Bioclim[1,2]
Isothermality	30-arc-seconds	°C	1950-2000	WorldClim, Bioclim[1,2]
Temperature seasonality	30-arc-seconds	°C	1950-2000	WorldClim, Bioclim[1,2]
Temperature annual range	30-arc-seconds	°C	1950-2000	WorldClim, Bioclim[1,2]
Precipitation				
Annual precipitation	30-arc-seconds	mm	1950-2000	WorldClim, Bioclim[1,2]
Wet day frequency	10-arc-minutes	days	1961-1990	CRU[3,4]
Precipitation of wettest month	30-arc-seconds	mm	1950-2000	WorldClim, Bioclim[1,2]
Precipitation of driest month	30-arc-seconds	mm	1950-2000	WorldClim, Bioclim[1,2]
Precipitation of wettest quarter	30-arc-seconds	mm	1950-2000	WorldClim, Bioclim[1,2]
Precipitation of driest quarter	30-arc-seconds	mm	1950-2000	WorldClim, Bioclim[1,2]
Precipitation of warmest quarter	30-arc-seconds	mm	1950-2000	WorldClim, Bioclim[1,2]

Precipitation of coldest quarter	30-arc-seconds	mm	1950-2000	WorldClim, Bioclim[1,2]
Precipitation variability				
Precipitation seasonality	30-arc-seconds	mm	1950-2000	WorldClim, Bioclim[1,2]
Precipitation annual range	30-arc-seconds	mm	1950-2000	WorldClim, Bioclim[1,2]
Light environment				
Cloud cover	0.5-arc-degrees	%	1961-1990	CRU[4,5]
Radiation	0.5-arc-degrees	watts/m2	1961-1990	CRU[4,5]
Sunshine duration	10-arc-minutes	hours	1961-1990	CRU[3,4]
Evapotranspiration				
Relative humidity	10-arc-minutes	%	1961-1990	CRU[3,4]
Vapour pressure	0.5-arc-degrees	hPa	1961-1990	CRU[4,5]
Global potential evapotranspiration	30-arc-seconds	mm/year	1950-2000	CGIAR-CSI[6,7,8]
Global aridity index	30-arc-seconds	AET/PET	1950-2000	CGIAR-CSI[6,7,8]
Fire				
Length of fire season	0.5-arc-degrees	months	1997-2010	Archibald et al., 2013[9]
Fire radiative power (FRP)	0.5-arc-degrees	rate of radiant energy release	1997-2010	Archibald et al., 2013[9]
Maximum fire size	0.5-arc-degrees	% burned area	1997-2010	Archibald et al., 2013[9]
Max fire line intensity	0.5-arc-degrees	95th quantile of FRP	1997-2010	Archibald et al., 2013[9]
Soil condition				
Topsoil organic content	30-arc-seconds	% weight	na	IIASA_World_Soils[10,11]
Soil pH	30-arc-seconds	-log(H+)	na	IIASA_World_Soils[10,11]
Cation exchange capacity of topsoil	30-arc-seconds	cmol/kg	na	IIASA_World_Soils[10,11]

[1]Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25, 1965-1978.
[2]http://www.worldclim.org/current
[3]New M, Lister D, Hulme M, Makin I. 2002. A high-resolution data set of surface climate over global land areas. Climate Research 21, 1-25.
[4]http://www.ipcc-data.org/observ/clim/get_30yr_means.html
[5]New M, Hulme M. Jones PD. 1999. Representing twentieth century space-time climate variability. Part 1: development of a 1961-90 mean monthly terrestrial climatology. Journal of Climate 12, 829-856.

[6]Zomer RJ, Bossio DA, Trabucco A, Yuanjie L, Gupta DC, Singh VP. 2007. Trees and water: Smallholder agroforestry on irrigated lands in northern India. Colombo, Sri Lanka: International Water Management Institute. IWMI Research Report 122, pp 45.
[7]Zomer RJ, Trabucco A, Bossio DA, van Straaten O, Verchot LV. 2008. Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. Agriculture, Ecosystems and Environment 126, 67-80.
[8]http://www.cgiar-csi.org/data/global-aridity-and-pet-database
[9]Archibald S, Lehmann CE, Gómez-Dans JL, Bradstock RA (2013) Defining pyromes and global syndromes of fire regimes. Proc Natl Acad Sci USA 110, 6442-6447.
[10]Fischer G, Nachtergaele F, Prieler S, van Velthuizen HT, Verelst L, Wiberg D. 2008. Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008). IIASA, Laxenburg, Austria and FAO, Rome, Italy.
[11]http://webarchive.iiasa.ac.at/Research/LUC/External- World-soil-database/HTML

Appendix 4.3 Soil descriptions associated with those presented in Figure 4.11. Descriptions taken from the Harmonized World Soils Database[1] and Major Soils of the World[2].

Soil quality	Soil type	General description	Chemical properties	Physical properties	Organic matter	pH	CEC
Good	Nitisols	Among most productive soils of humid tropics. Deep, dark red, brown or yellow clayey soils having a pronounced shiny, nut-shaped structure	High chemical fertility. High phosphorus-sorption, but P-deficiency is rare. High content of weathered minerals.	Stable soil structure, porous, resistant to erosion. Good workability and drainage. Fair water holding capacity.	Several per cent in surface soils. Particularly high in forest understories. Intense faunal activity.	5 to 6.5	High for tropical soils (greater than Ferrasols, Lixisols, and Acrisols)
Good	Phaeozems	Soils with a thick, dark topsoil rich in organic matter and evidence of removal of carbonates. Tall grass steppe and/or forest is natural vegetation.	High weathering. Calcium carbonate absent from surface.	Porous, well aerated soils. Very stable. Periodic drought and wind/ water erosion.	Rich; around 5%; C:N ratio 10-12.	5 to 7	25-30 cmol(+) kg-1 dry soil
Good	Cambisols	Weakly to moderately developed soils; in the process of development. Slight to moderate weathering of parent material, low alluvial clay, organic matter, aluminium and/or iron compounds. Supports wide range of vegetation types. Occur alongside Acrisols and Ferralsols.	Good chemical fertility; high in ferric oxides, hydroxides, aluminium oxides, and silicate clays.	Good structural stability, high porosity, good water holding capacity, and good internal drainage.	Active soil fauna	Neutral to weakly acid soil reaction	

Good	Luvisols	Soils with subsurface accumulation of high activity clays and high base saturation	Fertile; moderately weathered, less Fe-, Al-, and Ti-oxide contents than Lixisols. Depends on bedrock; often de-calcified	Favourable, porous and well-aerated. Well drained, although sometimes stagnant.	Only a few per cent, C:N ratio 10-15.	slightly acidic, subsurface soils may contain lime.	Equal to or greater than 24 cmol(+) kg-1 clay
Good	Fluvisols	Young soils in alluvial deposits; periodically flooded areas.	Fertile	Often stagnating groundwater or floodwater.		Neutral or near neutral	
Medium	Leptosols	Very shallow soils over hard rock or in unconsolidated very gravelly material; high or medium altitude. Forest soils.	Low in salts; depends on bedrock.	Low water holding capacity; easily droughted.	Good soil organisms, but may be seasonally inactive.		
medium	Lixisols	Soils with subsurface accumulation of low activity clays and high base saturation; clays nearly absent from surface; strongly weathered and leached.	Strongly weathered, low nutrients (but better than Ferralsols and Acrisols). Absence of Aluminium toxicity. Higher Fe-, Al-, and Ti-oxide contents	Finely textured. Prone to erosion. Free-draining; lack evidence of water saturation.	Good under natural state (under savannah or open woodland), but eroded away if degraded surface.	higher than Ferralsols and Acrisols	Low; less than 2 cmol(+) kg-1
Poor	Ferralsols	Deep, strongly weathered soils with a chemically poor, but physically stable subsoil. Clayey.	Poor. Low in nitrogen, potassium, calcium, magnesium, sulphur, and micro-nutrients. Absent weatherable minerals; iron content high.	Good. Great soil depth, good permeability, stable microstructure; less erosion susceptibility; well drained, but low 'available' water storage capacity.	Low	Acidic; much less than 7.0	weak, Less than 16 cmol(+) kg-1 clay

Poor	Acrisols	Strongly weathered soils with subsurface accumulation of low activity clays and low base saturation. Mostly old land surfaces with hilly/undulating topography, light forest is natural vegetation type.	Poor; little weatherable minerals left. Plant nutrients are low; aluminium toxicity and P-sorption are strong limitations. Very slow natural regeneration of soils.	Under protective forest cover, they have porous surface soils. However, if forest is cleared, the A-horizon degrades, leaving hard surface crust with insufficient penetration.	Low	Very acidic	Less than 24 cmol(+) kg-1 clay
Poor	Planosols	Soils with a degraded, bleached, temporarily water-saturated topsoil on a slowly permeable subsoil. Loamy sand or coarser textures. In flat lands.	Poor, degraded.	Seasonally waterlogged lands. Clays become abundance in lower strata.	Poor, shallow rooting depths	Acidic	Deteriorated
Poor	Arenosols	Sandy soils featuring very weak or no soil development, found in a wide variety of precipitation and temperature regimes.	A-horizons are shallow with little or poorly decomposed organic matter. Primarily quartz and feldspar; deeply leached, decalcified, low capacity to store bases.	Large pores; low water storage capacity, high drainage, low moisture holding capacity	Virtually absent; less than 1%		Less than 4 cmol(+) kg-1 clay
Poor	Regosols	Soils with very limited soil development; very weakly developed mineral soils, common in arid areas.		Low water holding capacity; often prolonged drought. Susceptible to erosion during wet periods.	Low, poorly decomposed at surface		
Poor	Solonetz	Soils with subsurface clay accumulation, rich in sodium and/or magnesium ions.	Sodium carbonate, High sodium saturation harmful to plants	Easily eroded surface layer. Very hard in dry season, sticky when wet.	Thin, loose litter layer resting on humidified material	Around 8.5	

[1]FAO/IIASA/ISRIC/ISSCAS/JRC, 2012.World Harmonized Soils Datahttp://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/
[2]Driessen *et al.*,2000. Major Soils of the World http://www.isric.org/isric/webdocs/docs//major_soils_of_the_world/start.pdf

CHAPTER 5

ENVIRONMENTAL DRIVERS OF ANATOMICAL VARIATION IMPORTANT TO PHYSIOLOGICAL DIVERSIFICATION

Introduction

Variations in anatomical trajectories

C_3 and C_4 photosynthetic pathways require different arrangements of leaf anatomy and, as such, the evolutionary transition between these states is characterized by stages of anatomical modifications (Sage, 2004; McKown and Dengler, 2007; Heckmann *et al*., 2013; Williams *et al*., 2013; reviewed in Chapter 2 and 3). The earliest of these modifications occur in the C_3 ancestors of C_4 lineages, which have adopted anatomical characteristics that facilitate the transport of metabolites across mesophyll and bundle sheath cells and enable the establishment of C_2 and C_4 cycles (Chapter 2; Christin *et al*., 2011a, 2013; Muhaidat *et al*., 2011; Sage *et al*., 2012, 2014; Griffiths *et al.,* 2013). In particular, higher vein densities and larger bundle sheaths have been identified as facilitators of C_4 evolution (Christin *et al*., 2013a; Griffiths *et al.*; 2013). Once a C_2 cycle has been established, the proportion of Rubisco needs to be shifted away from mesophyll tissue and into the bundle sheath for a C_4 pathway to evolve (Heckmann *et al*., 2013; Sage *et al*., 2013). This may be achieved by shifting the abundance of chloroplasts or activity of Rubisco between these two cell types (Stata *et al*., 2014) or, anatomically, by decreasing mesophyll area and increasing bundle sheath area. As discussed in Chapter 2, mesophyll area can be reduced by decreasing the number and size of mesophyll cells as well as decreasing leaf thickness, while bundle sheath area can be increased by enlarging bundle sheath cells or increasing bundle sheath frequency through the proliferation of minor veins. Thus, many different combinations of trait modifications may ultimately yield functionally similar phenotypes (Williams *et al*., 2013; Chapter 2).

Each modification along the path of photosynthetic diversification may be driven by different selection pressures (Griffiths *et al.,* 2013; Sage *et al.,* 2014). For instance, increases in vein density, as well as bundle sheath size, are hypothesized to improve hydraulic conductivity and reduce cavitation and carbon starvation in seasonally arid, windy, and otherwise high potential evapotranspiration environments (Sack and Frole, 2006; Brodribb *et al*., 2007; Osborne and Sack, 2012; Griffiths *et al*., 2013). In contrast, the relative volumes of chlorenchymatous mesophyll and bundle sheath tissue will affect CO_2 diffusion and light interception,

such that decreases in mesophyll accompanied by increased chlorenchymatous bundle sheath area should improve CO_2 exchange in low CO_2 environments and light capture under shady conditions (Evans and von Caemmerer, 1996; Vogelmann *et al.,* 1996; Ogle, 2003). Thus, the particular environments that each lineage experienced, and at which point during the evolutionary transition they were experienced, will influence the particular suite of traits that make up the unique C_4 phenotypes at each independent origin.

Leaf anatomy in Alloteropsis semialata

Leaf anatomy is remarkably variable in the grass species *Alloteropsis semialata* (Ellis, 1974; Barrett *et al*., 1983; Frean *et al*., 1983; Renvoize, 1987a; Hattersley and Watson, 1992; Ueno and Sentoku, 2006). The bulk of this variation is driven by anatomical difference inherent to the C_3, C_3-C_4 intermediate, and C_4 photosynthetic types present in this species (Gibbs Russell, 1983; Chapter 3), however, considerable anatomical variation exists even within these photosynthetic types. For example, Renvoize (1987) found that the number of mesophyll cells separating veins and the presence of minor (quaternary) veins varied in *A. semialata* within specimens he believed to use the C_4 pathway. Moreover, Chapter 3 shows that C_3, C_3-C_4 intermediate, and C_4 photosynthetic types vary considerably in vein frequency, the minimum number of mesophyll cells separating veins, inner and outer bundle sheath cell sizes, and the ratio of inner to outer bundle sheath cell sizes (Table 5.1). This chapter presents a closer look at this intraspecific anatomical variation to determine the distance between photosynthetic phenotypes and to quantify the degree to which the C_3-C_4 intermediate state bridges the selective landscape between C_3 and C_4 phenotypes.

Table 5.1 Range of anatomical variation in *Alloteropsis semialata*, as revealed in Chapter 3.

	Vein frequency	Number of mesophyll cells	IBS cell size (µm)	OBS cell size (µm)	IBS:OBS
C_3	2.7 - 4.7	5.9 - 8.6	7.1 - 10.7	15.8 - 19.6	0.4 - 0.5
C_3-C_4	5.0 - 7.3	3.5 - 5.4	10.5 - 13.0	11.6 - 14.4	0.8 - 1.1
C_4	8.3 - 11.7	1.3 - 1.8	9.4 - 12.2	7.7 - 11.4	1.0 - 1.3

On using the herbarium record

This chapter presents a comprehensive analysis of anatomical variation in the herbarium record of *A. semialata*. Herbarium specimens are a valuable resource to study C_4 evolution. Tissue from herbarium specimens has been successfully used to reveal phylogenetic relationships (Christin *et al*., 2011b**;** Besnard *et al*., 2014; Chapter 4) and the evolutionary history of phenotypes (Christin *et al*., 2013a; Chapter 2), and determine photosynthetic pathway through stable isotope (Pearcy and Troughton, 1975; Hattersley and Watson, 1992; Chapter 4) and leaf anatomy methods (Christin *et al*., 2011a, 2013; Osborne *et al*., 2014; Vorontsova *et al*., 2015). The herbarium record also reveals biogeography as it records where, and in which environments, taxa have existed in the past (Holmgren *et al*., 2007; Ames and Spooner, 2008; Schaefer and Renner, 2010; Feely, 2012).

Aims of Chapter 5

This chapter comprehensively dissects the anatomical variation expressed within *A. semialata* to determine how these phenotypes vary across geographic space and environments. By investigating intraspecific anatomical variation, this study is able to untangle the fine-scale anatomical shifts that may have promoted the emergence of C_4 photosynthesis and reversion to C_3 photosynthesis in this species, where these shifts occurred, and what environments drove them. Thus, this chapter aims to determine (1) the phenotypic distance between C_3, C_3-C_4 intermediate, and C_4 anatomies, (2) whether the C_3-C_4 intermediate phenotype shifts anatomy closer to the C_4 requirements, and (3) which environments are associated with anatomical shifts.

Methods

Database of Alloteropsis semialata accessions

Collection locations for over 1230 *A. semialata* herbarium specimens were collated from several sources. All *A. semialata* specimens housed at the herbarium of the Royal Botanical Gardens at Kew (Surrey, UK), the Natural History Museum (London, UK), the National Herbarium of South Africa (Pretoria, South Africa), and

the National Herbarium of Tanzania (Arusha, Tanzania) were individually reviewed in person. Specimens were photographed and descriptions of collection habitat or location were recorded. Some herbarium sheets included precise latitude and longitude coordinates, however, most only provided descriptive notes on the collection location. For these specimens, coordinates were determined by georeferencing the descriptions provided on the herbarium sheet using Biogeomancer (Guralnick *et al*., 2006), the Flora of Tropical East Africa Index of Collecting Localities (Polhill, 1988), and Google Earth. Additional collection locations were obtained from the Australian Virtual Herbarium (http://avh.chah.org.au), colleagues (Long and Hattersley, unpublished data), and individual specimen inquiries from the East African Herbarium (Nairobi, Kenya), Bolus Herbarium (Cape Town, South Africa), and the Australian National Herbarium (Canberra, Australia). Living accessions, collected from field sites in South Africa, Tanzania, and Taiwan, or germinated from seeds (Queensland, Australia; Majunga, Madagascar; Mondon, Burkina Faso), augmented the herbarium dataset and are described in Chapter 3 and Appendix 5.1.

Stable isotope analysis

$\delta^{13}C$ stable isotope data were compiled from a subset of 300 of the *A. semialata* accessions listed above, chosen to represent the species' geographic range (Appendix 5.1 and 5.2). For a number of these specimens, $\delta^{13}C$ had already been quantified and these data were presented on herbarium sheets, in the published literature (Ellis, 1981; Gibbs Russell, 1983; Hattersley and Watson, 1992), or were unpublished (Long and Hattersley). For the remaining specimens, leaf tissue from young, fully expanded leaves were harvested from the herbarium sheets and used to quantify $\delta^{13}C$. These samples were analysed at the University of Sheffield, Faculty of Science biOMICS facility. Data are presented as isotopic ratios, reported relative to an isotopic standard (in ‰), as described in Chapter 3. All specimens with $\delta^{13}C$ values between -23 and -17‰ were measured at least twice to confirm isotopic intermediacy.

Phenology

A database of flowering phenology was compiled from the same 300 herbarium specimens described above (Appendix 5.1 and 5.2). As herbarium specimens are typically collected while plants are in flower, collection dates reveal flowering dates for different populations of *A. semialata*. This type of phenology data is limited, as it does not inform how long these populations remained in flower or when they lacked flowers, however, it does clearly show which populations flowered simultaneously and, thus, where pollen transfer might be possible.

Habitat dataset

A database of habitat information was compiled from the same 300 herbarium specimens described above (Appendix 5.1 and 5.2). Written descriptions of altitude, disturbance, wetland, open grassland, shaded woodland, burnt, sloped or cliffed, and rocky habitats were marked and these habitat data were summarized across the $\delta^{13}C$ continuum.

Leaf anatomy

Leaf anatomy data were collected on 100 of the 300 accessions listed in Appendices 5.1 and 5.2. Accessions were chosen to represent the geographic and $\delta^{13}C$ range of *A. semialata* (Figure 5.1; Appendix 5.1 and 5.2). The anatomy dataset consisted of 36 C_4 individuals and 64 accessions where $\delta^{13}C$ indicated that they did not primarily use the C_4 photosynthetic pathway.

Leaf pieces 3-5 mm in length were fixed in Carnoy's fixative (4:1 EtOH:acetic acid) and embedded in Methacrylate embedding resin (Technovit 7100, Heraeus Kulzer GmbH, Wehrhein, Germany). Leaves were sectioned between 6-8 µm thick on a manual rotary microtome (Leica Biosystems, Newcastle, UK) and stained with Toluidine Blue O (Sigma-Aldrich, St. Louis, MO, USA). Stained leaf sections were imaged using microscopy imaging software with a camera mounted on a microscope (Cell A, Olympus DP71, and Olympus BX51, respectively; Olympus, Hamburg, Germany). Images were stitched together (DoubleTake 2.2.9, Echo One,

Frederikssund, Denmark) to recreate the continuous width of the whole cross-section.

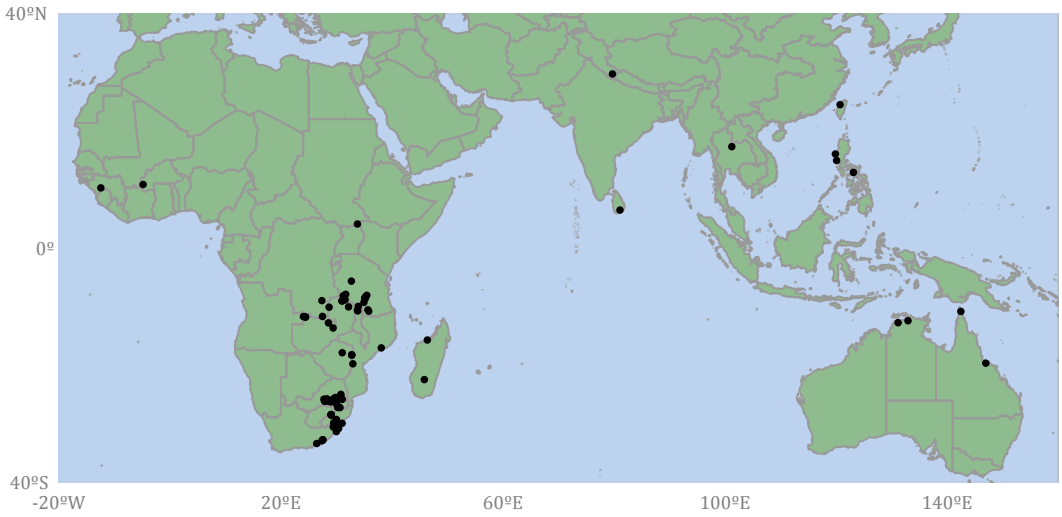


Figure 5.1 Collection locations for 100 accessions included in leaf anatomy dataset.

Anatomical traits were measured using ImageJ software (Schneider *et al.*, 2012), as shown in Figure 5.2. A single "segment" was considered the leaf portion between two 2nd order veins. These secondary veins are large and contain metaxylem, while tertiary, quaternary, and quinary minor veins lack metaxylem. Only segments in the middle of the leaf blade were used and segments immediately adjacent to the mid-rib and tips of each cross-section were avoided.

The number of veins per segment and their vein order (secondary, tertiary, quaternary, or quinary) were recorded. These were counted along three segments per leaf. Some veins were clearly smaller and less developed than quinary veins. These small veins were either transverse veins or newly forming veins. Transverse veins function to connect larger, parallel veins and were often found within the mesophyll, in between developed veins. In contrast, newly forming veins were often budding off of a developed vein. The total numbers of transverse and budding veins were counted along the whole cross-section.

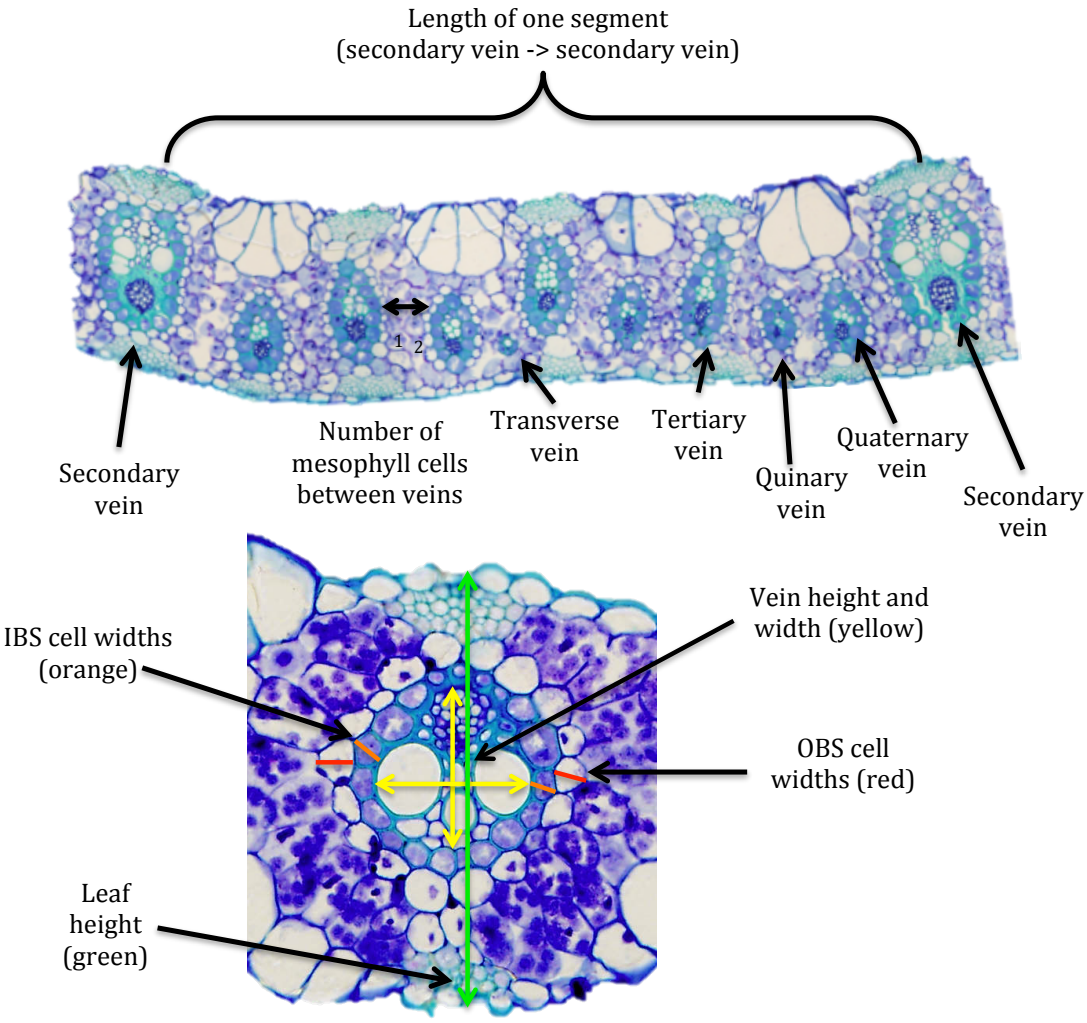


Figure 5.2 Cross-sections of *Alloteropsis semialata* leaves demonstrating leaf anatomy measurements. A single segment was considered the leaf portion between two secondary veins. The number of veins per segment and their vein order (secondary, tertiary, quaternary, or quinary) was recorded. The minimum number of adjacent mesophyll cells separating vein pairs was counted for ten vein pairs per leaf. The width of one representative OBS (red) and one IBS (orange) on both sides of three primary veins were measured and the IBS:OBS cell width ratio was calculated. The height and width (yellow) of three veins were measured and used to calculate the vein area and shape. Leaf thickness (green) was measured at each of these veins.

Leaf thickness was measured at the centre of three secondary veins per leaf and averaged. Segment length was measured from the centre of one secondary vein to the centre of the next secondary vein, jaggedly following the centre of each vein in between. The length of three segments was measured and averaged. The minimum number of adjacent mesophyll cells separating vein pairs was counted for ten pairs

of veins per leaf and averaged. Mesophyll cells were difficult to count on some herbarium specimen leaves, where brittle mesophyll cell walls broke easily. For these specimens (N=16), mesophyll cell counts are unavailable.

The width of one outer bundle sheath (OBS) cell and one inner bundle sheath (IBS) cell on both sides of each secondary vein was measured and a ratio of IBS:OBS cell width was calculated and averaged for three secondary veins per leaf. To maintain consistency, the representative bundle sheath cells used for these measurements were those that fell parallel to the leaf surface, as often as possible. The height and width of these same secondary veins were measured and used to calculate vein area, based on an elliptical shape (Equation 5.1). The vein width:height ratio was also calculated to estimate vein shape.

Equation 5.1
Vein area = π * (vein height/2) * (vein width/2)

Principal component analyses
Two PCA were performed on 100 accessions with the FACTOMINER package (Lê *et al*., 2008) in R (version 3.1.0, R Development Core Team, 2014). To explain the anatomical variation present in these 100 accessions, the first PCA was performed on the following 12 traits; segment length, leaf thickness, number of mesophyll cells separating veins, total number of veins per segment, numbers of tertiary, quaternary, and quinary veins, vein area, vein width:height ratio, IBS cell width, OBS cell width, and IBS:OBS ratio. A hierarchical cluster analysis was performed on this PCA to see whether the 100 accessions fell into clusters consistent with photosynthetic type, based on anatomy alone.

To explain the habitat variability from which these specimens where collected, the second PCA was performed with 37 environmental variables. Environmental information for each collection location was obtained by overlaying geographic coordinates onto environmental raster layers and extracting the raster data at each location using the SP (Pebesma and Bivand, 2005), RGDAL (Bivand *et al*., 2014), RASTER (Hijmans *et al*., 2014), and PLYR (Wickham, 2011) packages in R.

High resolution global gridded datasets of the 37 environmental layers used in this study are described in Appendix 4.2. The 15 anatomical traits mentioned above were overlaid onto the environmental PCA as supplementary quantitative variables to aid interpretation.

Statistical analyses

Linear least square regression analyses were performed in R. Initial sets of analyses tested separately whether variations in leaf anatomy and environmental variables were explained by photosynthetic type (as inferred from $\delta^{13}C$ signatures), using the 12 anatomical variables described above and the 37 environmental variables described in Appendix 4.2. Next, linear least square regression analysis tested whether variation in leaf anatomical traits was explained by environmental factors. P-values resulting from these linear models were corrected for multiple testing. Correlation between anatomical traits was tested using the chart.correlation function in the PerformanceAnalytics package (Peterson *et al*., 2014) in R.

Results

Geographic distribution

A comprehensive review of *A. semialata* collections over the past 200 years reveals this species has been collected throughout the western, central, and southern areas of mainland Africa, Madagascar, northern and eastern Australia, India, and southern Asia (Figure 5.3). $\delta^{13}C$ signatures ranged from -29 to -9‰ across 300 accessions (Figure 5.4a), which are consistent with the level of photosynthetic variation seen in this species (Ellis, 1981; Hattersley and Watson, 1992; Chapter 3). Interestingly, these values do not form two discrete assemblages but show continuous variation across the $\delta^{13}C$ range, suggesting intermediary states and lending more evidence for the existence of type II C_3-C_4 intermediate phenotypes in this species (Figure 5.4a; Chapter 3; Sage *et al*., 2013). Specimens that fell intermediate between C_3 and C_4 $\delta^{13}C$ ranges are very likely C_2 plants that employ a weak C_4 cycle, while those accessions with $\delta^{13}C$ less than 23‰ may use either C_3

photosynthesis or a C_2 cycle with little or no C_4 cycle engagement (discussed in Chapter 3). Although there is a gap in $\delta^{13}C$ values between the majority of C_4 and non-C_4 accessions, one accession from Madagascar (-16.3‰) fell within the gap. This accession was determined to be C_4 based on anatomical characteristics. Thus, the cut-off between C_4 and non-C_4 accessions was drawn at -17‰.

Geographic patterns of $\delta^{13}C$ signatures emerged (Figure 5.4b). C_4 accessions (*i.e.*, greater than -17‰) were collected in Australia, Asia, and throughout mainland Africa and Madagascar (Figure 5.4b). In contrast, non-C_4 accessions (*i.e.*, less than -17‰) were only collected in south-eastern Africa (Figure 5.4b). Interestingly, accessions with $\delta^{13}C$ intermediate between typical C_4 and C_3 values (*i.e.*, between -17 and -23 ‰) were only found in central eastern Africa, namely in the countries of Tanzania, the Democratic Republic of Congo, and Zambia (Figure 5.4b; Appendix 5.1 and 5.2; Chapter 4).

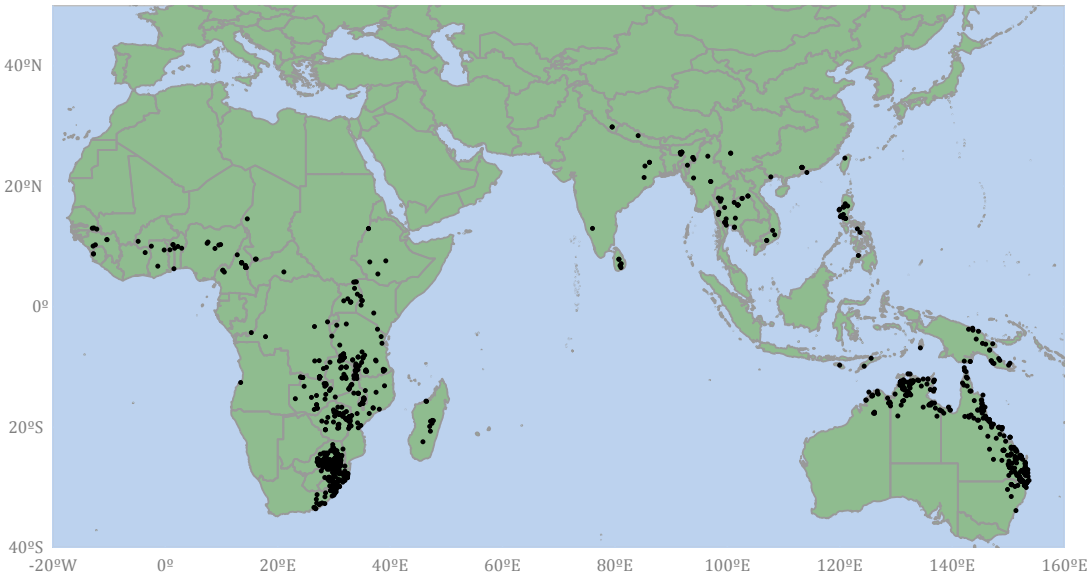


Figure 5.3 Distribution of 1232 *Alloteropsis semialata* accessions collected over the last 200 years.

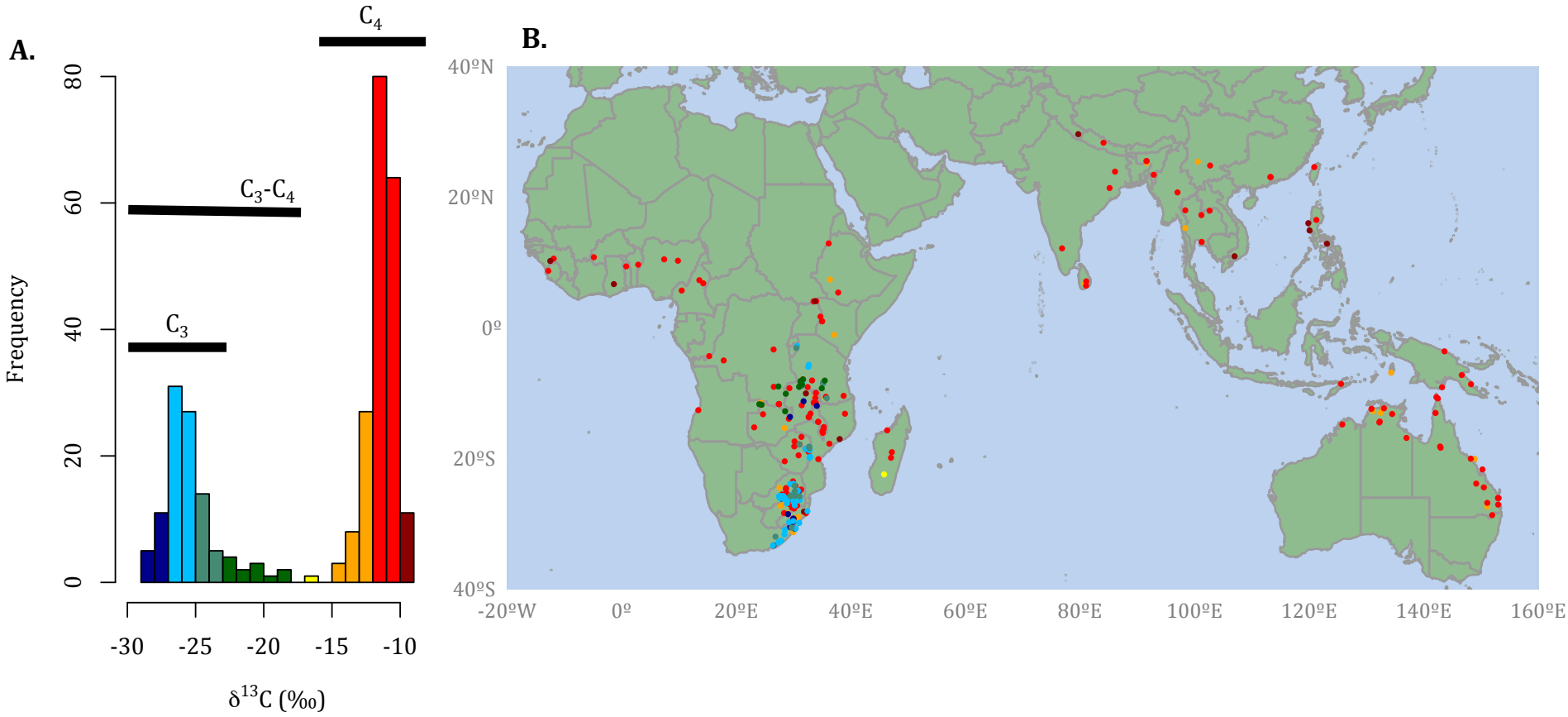


Figure 5.4 Frequency (A) and geographic (B) distributions of $\delta^{13}C$ ($^{0}/_{00}$) values assessed from 300 *Alloteropsis semialata* accessions.

Flowering phenology

Alloteropsis semialata has been found in flower throughout the year. Specifically, the non-C_4 accessions flower throughout the year in southern Africa, while in central Africa, their flowering time narrowly falls between October and January. Opposite to this, C_4 plants flower broadly in central Africa but in southern Africa were only found in flower between October and March (Figure 5.5). As both C_4 and non-C_4 accessions were found in flower at similar times in southern and central Africa (Figure 5.5), it is unlikely that flowering phenology acts as a reproductive isolating mechanism in either region.

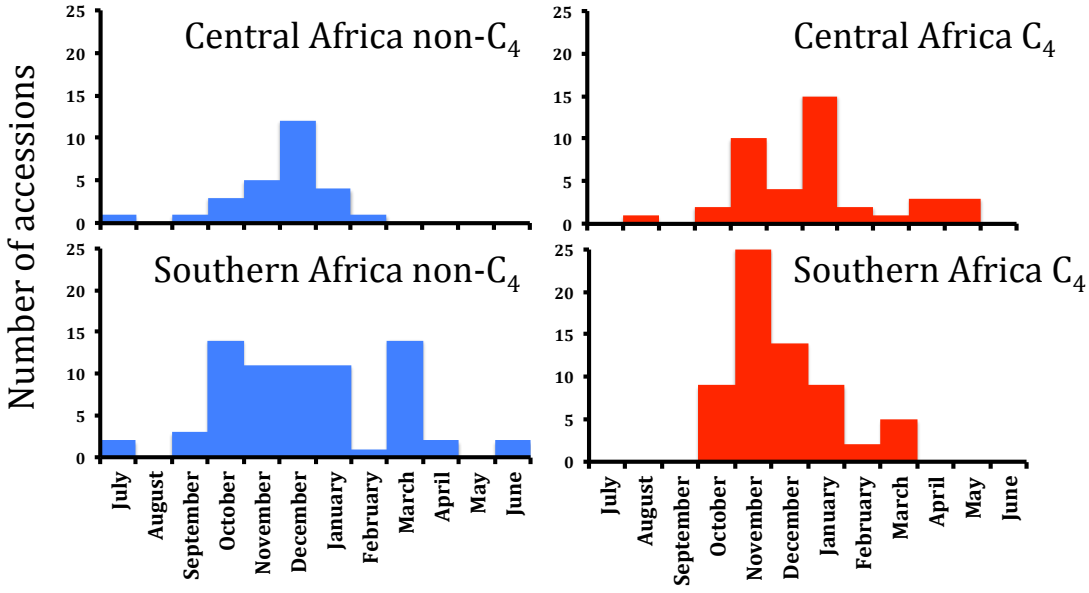


Figure 5.5 Histograms of flowering time for C_4 (red) and non-C_4 (blue) accessions collected in central and southern Africa, as inferred from collection date on herbarium sheet.

Ecological distribution

Habitat trends emerged from the analysis of herbarium specimens' collection data. There was considerable overlap between C_3 and C_4 collection altitudes, however, C_4 accessions were collected across a greater distribution of altitudes and were found at lower altitudes than non-C_4 accessions (Figure 5.6). Most of the accessions described as coming from a wetland habitat were C_4, however, three non-C_4 accessions were also found in damp habitats. Both C_4 and non-C_4 accessions were found in shady and open, sloped, rocky, disturbed, and fire prone areas (Figure 5.6). Isotopically intermediate accessions were collected in shaded woodland areas, but not open habitats (Figure 5.6). Moreover, accessions with $\delta^{13}C$

between -23 and -17‰ were not collected in rocky, disturbed, or fire prone habitats, but were found along cliff faces and sloped landscapes (Figure 5.6).

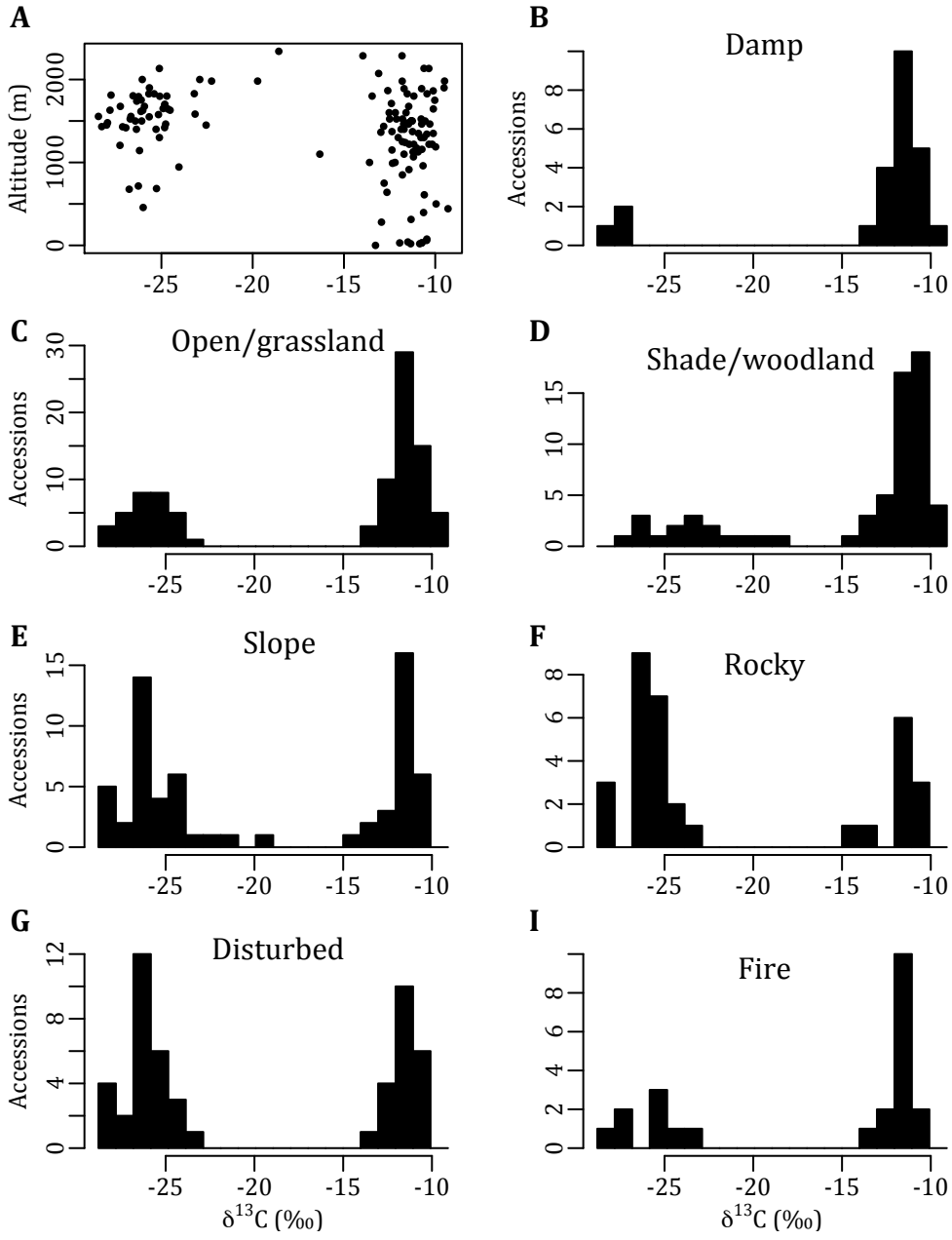


Figure 5.6 Analysis of habitat information extracted from herbarium sheet descriptions. Panel A is a scatter plot of $\delta^{13}C$ values from accessions collected at different altitudes. The remaining panels are frequency distributions of $\delta^{13}C$ values from accessions collected in wetland or damp (B), open or grassland (C), shaded or woodland (D), sloped (E), rocky (F), disturbed (G), and fire (H) habitats. Bin widths were optimized following Shimazaki and Shinomoto (2007).

Based on analyses of the high-resolution environmental datasets, the environmental conditions in many C_4 and non-C_4 *A. semialata* collection locations were similar in many respects (Figure 5.7). For most environmental variables included in this study, C_4 accessions were collected across a broader range of conditions than non-C_4 accessions, which is consistent with the findings of Chapter 4. For only a few variables were non-C_4 accessions found in a broader range than C_4 accessions. These included precipitation seasonality, precipitation of the driest month, radiation, fire radiative power, and maximum fire line intensity. Many non-C_4 accessions show progression across the $\delta^{13}C$ gradient, as isotopically intermediate specimens were generally collected in areas characterized by warmer dry season temperatures, higher seasonal precipitation, lower dry season precipitation, greater cloud cover, higher PET, and higher altitude, compared with specimens with lower $\delta^{13}C$ values (Figure 5.7).

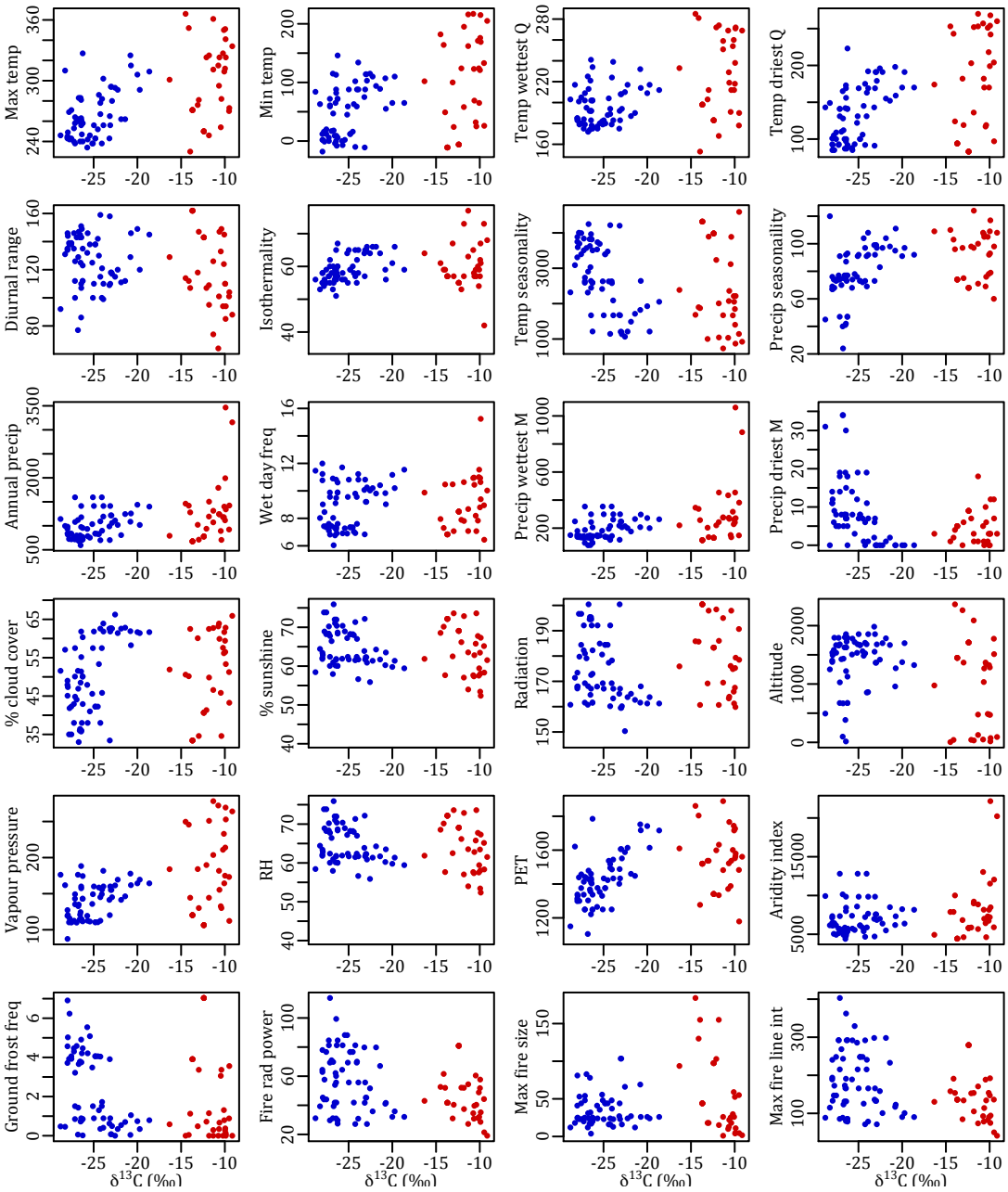


Figure 5.7 Scatterplots of $\delta^{13}C$ values against environmental conditions at the collection sites. C_4 accessions are colored red and non-C_4 accessions are blue. The words month (M), quarter (Q), temperature (temp), precipitation (precip), frequency (freq), radiative (rad), and intensity (int) are abbreviated. All variables of temperature are in degrees Celsius, precipitation in mm, altitude in meters, vapour pressure in hPa, and aridity index in mm/year. All other details about these environmental variables are detailed in Appendix 4.2.

Anatomical variation

Overall, anatomical traits showed continuous distributions, with overlap between C_4 and non-C_4 accessions (Figure 5.8). Only the number of mesophyll cells separating veins was an exception to this trend, clearly distinguishing C_4 from non-C_4 accessions. All C_4 accessions averaged fewer than four, while all non-C_4 accessions averaged greater than four mesophyll cells between veins (*e.g.*, Figure 3.8). While all non-C_4 accessions lacked quaternary and quinary veins, some C_4 accessions also lacked these vein orders (Figure 5.8). This suggests that the presence of quaternary and quinary veins confirms the C_4 photosynthetic type, however, the absence of these vein orders does not confirm the lack of the C_4 type. While there were usually more than eight veins per segment in C_4 plants and fewer than six in non-C_4 plants (Figure 5.8 – 5.9), those C_4 accessions lacking abundant quaternary and quinary veins are evident outliers in the total number of veins dataset, with fewer than six veins per segment (Figure 5.8). It is interesting to note these five C_4 accessions with the lowest vein abundance were collected in central and southern Africa (Tanzania, Malawi, and South Africa).

The PCA explained over 60 percent of the anatomical variation in the dataset of 100 specimens (Figure 5.9). C_4 and non-C_4 specimens are cleanly distinguished on dimension 1, with non-C_4 individuals falling negative, and C_4 individuals positive, along this axis (Figure 5.9a). All anatomical traits correlated significantly with dimension 1 (Figure 5.9b; Table 5.2), suggesting the 12 measured traits distinguish the C_4 phenotype well. Dimension 1 was positively correlated with $\delta^{13}C$ across the whole dataset and within the non-C_4 subset of accession (Figure 5.9c; Table 5.3). This suggests that accessions become anatomically more C_4-like as $\delta^{13}C$ increases (or vice-versa). For example, outer bundle sheath cells shrank ($F = 43.14$, $R^2 = 0.405$, $p = 1.3e\text{-}08$), mesophyll cell number decreased ($F = 11.77$, $R^2 = 0.177$, $p = 0.001$), and leaves became thinner ($F = 26.6$, $R^2 = 0.292$, $p = 2.9e\text{-}06$) across the $\delta^{13}C$ gradient in non-C_4 accessions (Figure 5.10).

C_4 and non-C_4 accessions overlapped completely along dimension 2, which was significantly correlated with traits describing vein size and shape, bundle sheath cell sizes, leaf thickness, and segment size (Figure 5.9b; Table 5.2). Although dimension 2 did not correlate with $\delta^{13}C$ across the whole dataset, the non-C_4

accessions showed a negative correlation between $\delta^{13}C$ and dimension 2 (Figure 5.9d; Table 5.3). This suggests that leaves became thinner, segments shorter, veins smaller and rounder, and bundle sheath cells smaller as $\delta^{13}C$ values increased in non-C_4 accessions, as can been seen in Figure 5.10.

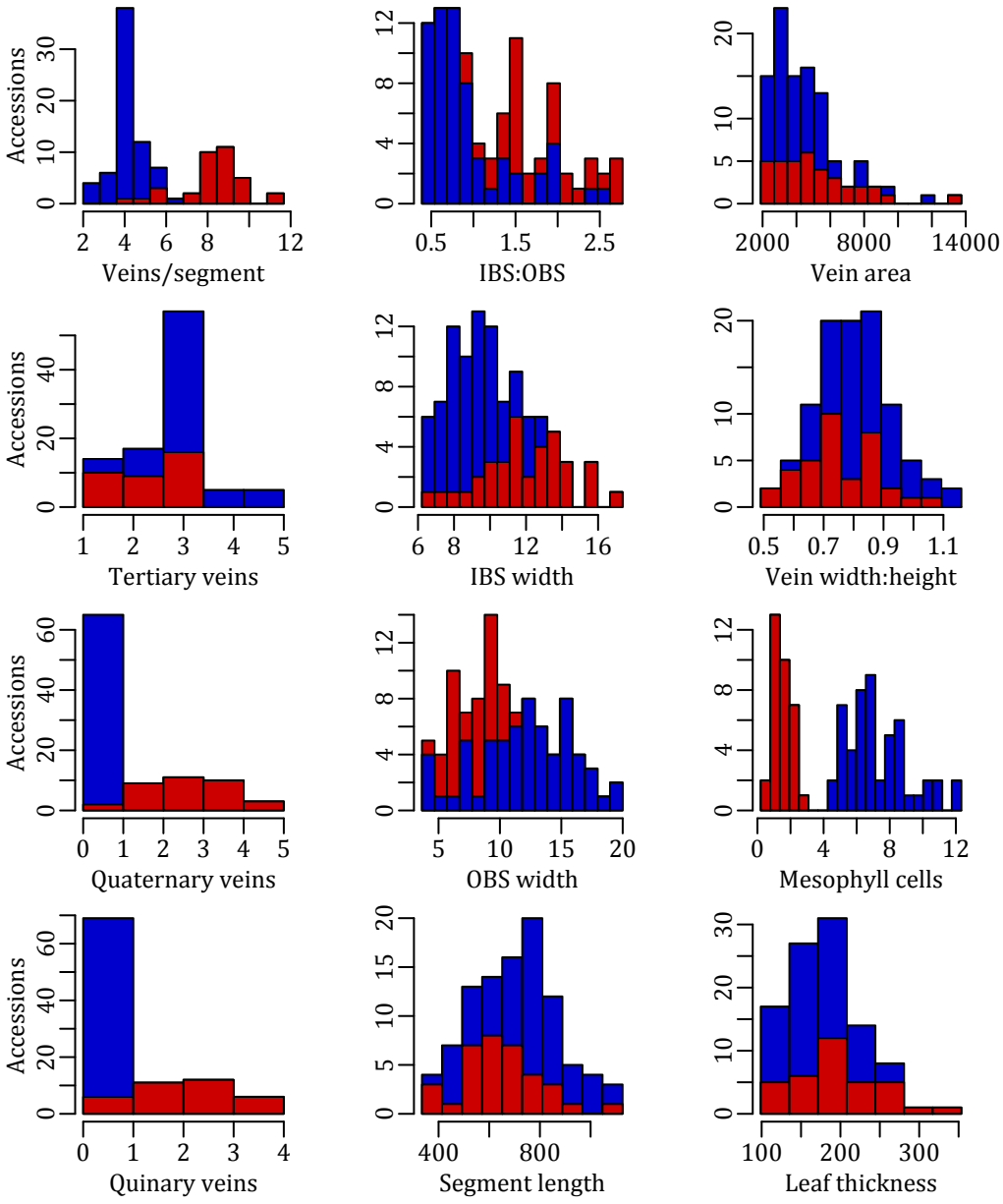


Figure 5.8 Frequency distributions of anatomical traits for C_4 (red) and non-C_4 (blue) accessions. Bin sizes optimized per Shimazaki and Shinomoto (2007). N=100.

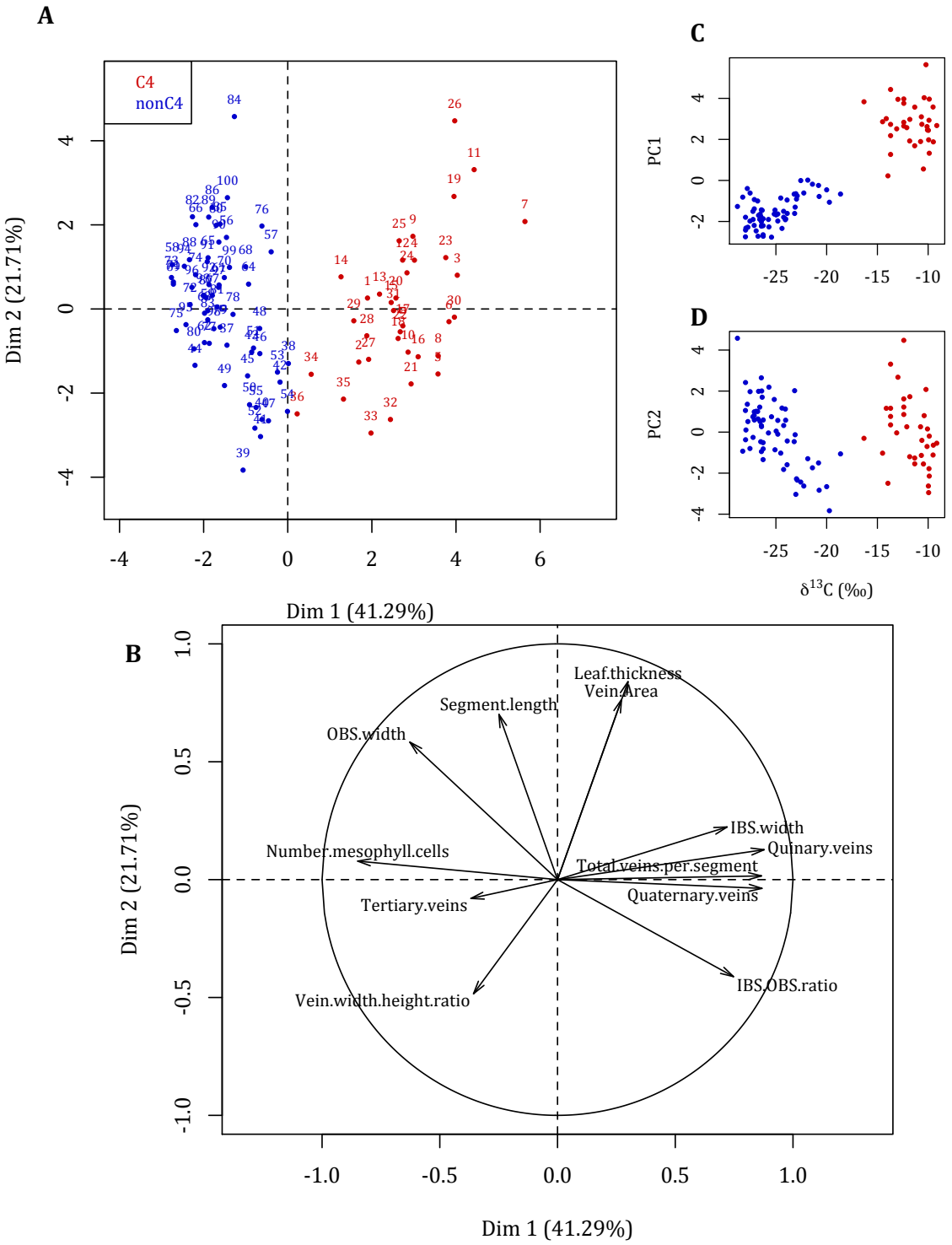


Figure 5.9 Principal component analysis for anatomical traits showing how C_4 (red) and non-C_4 (blue) accessions differ along dimensions 1 and 2 of the PCA (A) and which anatomical traits correspond with each dimension (B). Scatter plots show the relationship between $\delta^{13}C$ ($^0/_{00}$) and dimension 1 (C) and dimension 2 (D) of the PCA. N=100.

Table 5.2 Coefficients for anatomical traits significantly correlated with dimensions 1 and 2 of the PCA.

Trait	Correlation coefficient	p-value
Dimension 1		
Quinary veins	0.88	0.00E+00
Quaternary veins	0.87	0.00E+00
No veins/segment	0.87	0.00E+00
IBS:OBS	0.75	0.00E+00
IBS width	0.72	0.00E+00
Leaf thickness	0.30	2.48E-03
Vein area	0.27	6.24E-03
Segment length	-0.25	1.25E-02
Vein width:height	-0.36	2.68E-04
Tertiary veins	-0.37	1.66E-04
OBS width	-0.63	2.85E-12
Number mesophyll cells	-0.85	7.17E-29
Dimension 2		
Leaf thickness	0.84	0.00E+00
Vein area	0.76	0.00E+00
Segment length	0.70	4.44E-16
OBS width	0.58	1.89E-10
IBS width	0.22	2.55E-02
IBS:OBS	-0.41	2.10E-05
Vein width:height	-0.48	3.35E-07

Table 5.3 Results of linear regression analyses to determine whether variation in the PCA data could be explained by $\delta^{13}C$.

	All accessions				Subset: non-C_4 accessions			
	DF	*F*	R^2	*p*	*DF*	*F*	R^2	*p*
$\delta^{13}C$ x PC1	1, 96	483.60	0.833	< 2.2e-16	1, 61	29.29	0.313	1.11e-06
$\delta^{13}C$ x PC2	1, 96	3.10	0.021	8.14e-02	1, 61	45.81	0.420	5.78e-09

p-values less than 0.05 were considered significant. N=100.

When all accessions were put into a hierarchical cluster analysis based on leaf anatomy, three clusters emerged, with cluster 1 being sister to clusters 2 and 3 (Figure 5.11a). Interestingly, these three clusters do not correspond perfectly with photosynthetic type. While cluster 1 contained only C_4 accessions, cluster 3 contained primarily C_3 accessions as well as five Tanzanian C_3-C_4 intermediate plants, and cluster 2 included the remaining isotopically intermediate accessions, as well as six C_4 accessions and five C_3 accessions. This suggests that *A. semialata* fell into three anatomical groups characterized primarily by C_4-like anatomy (cluster 1), C_3-like anatomy (cluster 3), and a group anatomically intermediate between C_3 and C_4 phenotypes (cluster 2).

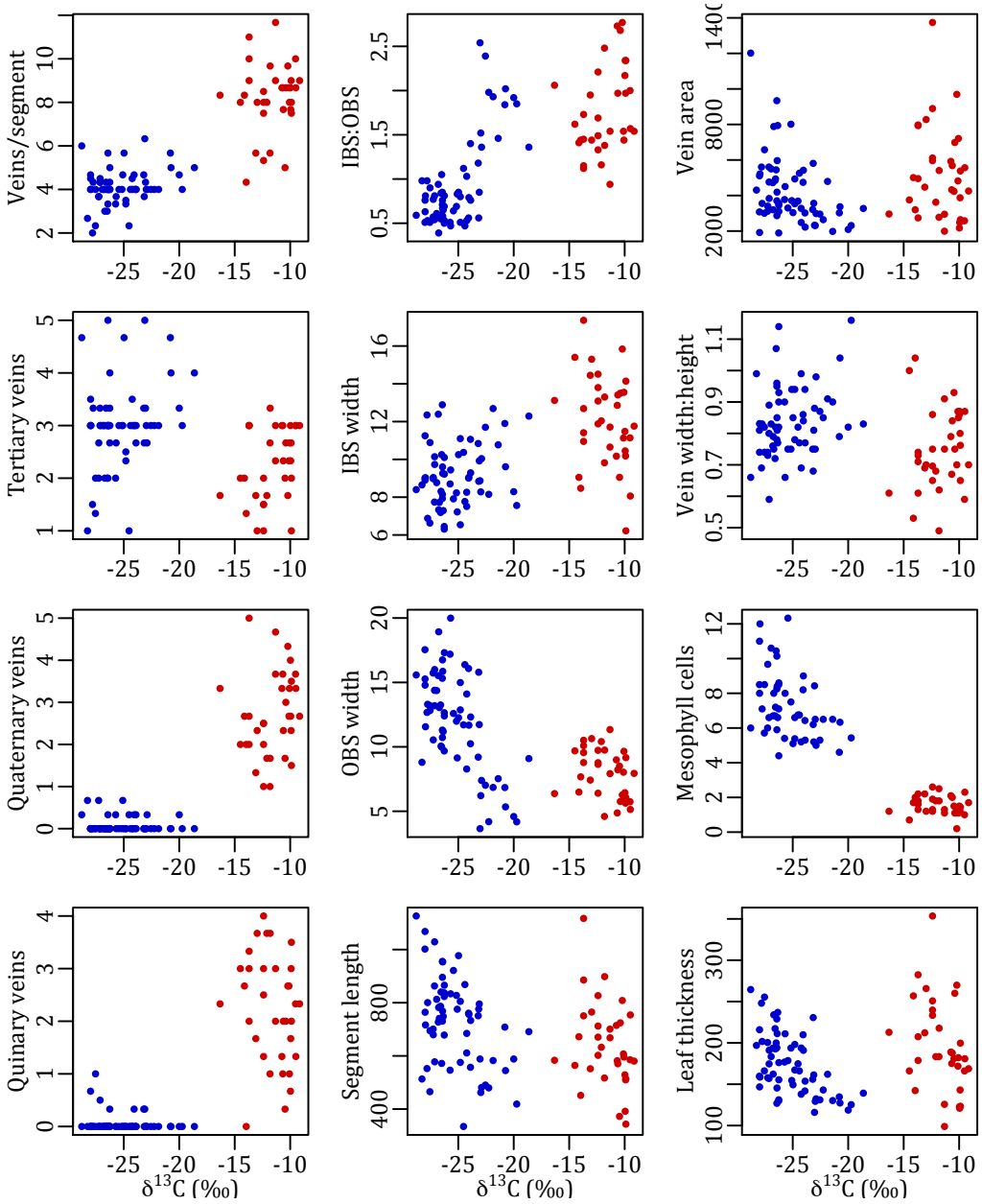


Figure 5.10 Scatter plots of anatomical traits by $\delta^{13}C$ ($^0/_{00}$) for C_4 (red) and non-C_4 (blue) accessions. N=100.

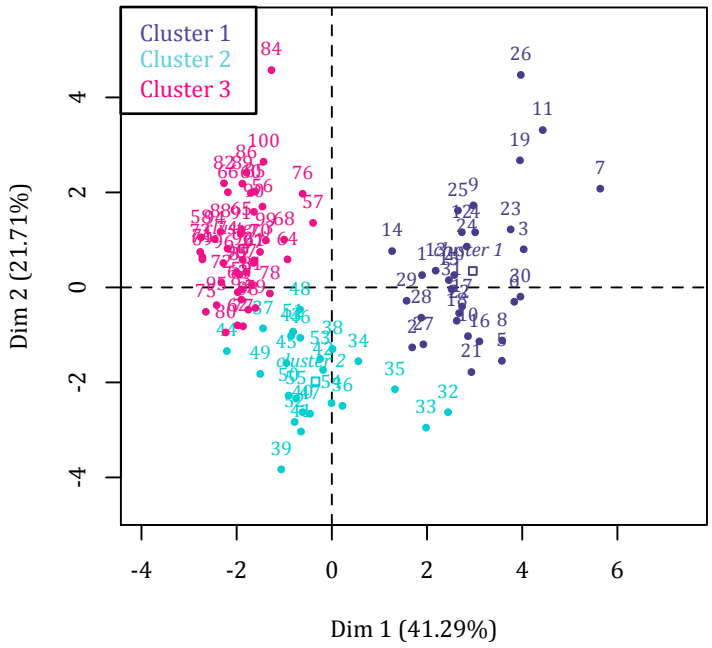


Figure 5.11 Hierarchical cluster analysis for anatomical traits showing how accessions break into three clusters. N=100.

Correlation among anatomical traits

Several anatomical traits were correlated (Figure 5.12). While some of these relationships can be explained by simple leaf scaling, as larger leaves have longer segments, larger veins, more tertiary veins, and are thicker (Figure 5.12), other relationships are more compelling. The number of veins per segment, mesophyll cells between veins, OBS cell widths, and IBS:OBS ratio were all highly correlated with one another (Figure 5.12), as was seen in Chapter 3. However, the IBS cell widths were also correlated with these traits in this chapter (Figure 5.12). The total number of veins per segment was strongly correlated with quaternary and quinary veins, but not tertiary veins (Figure 5.12).

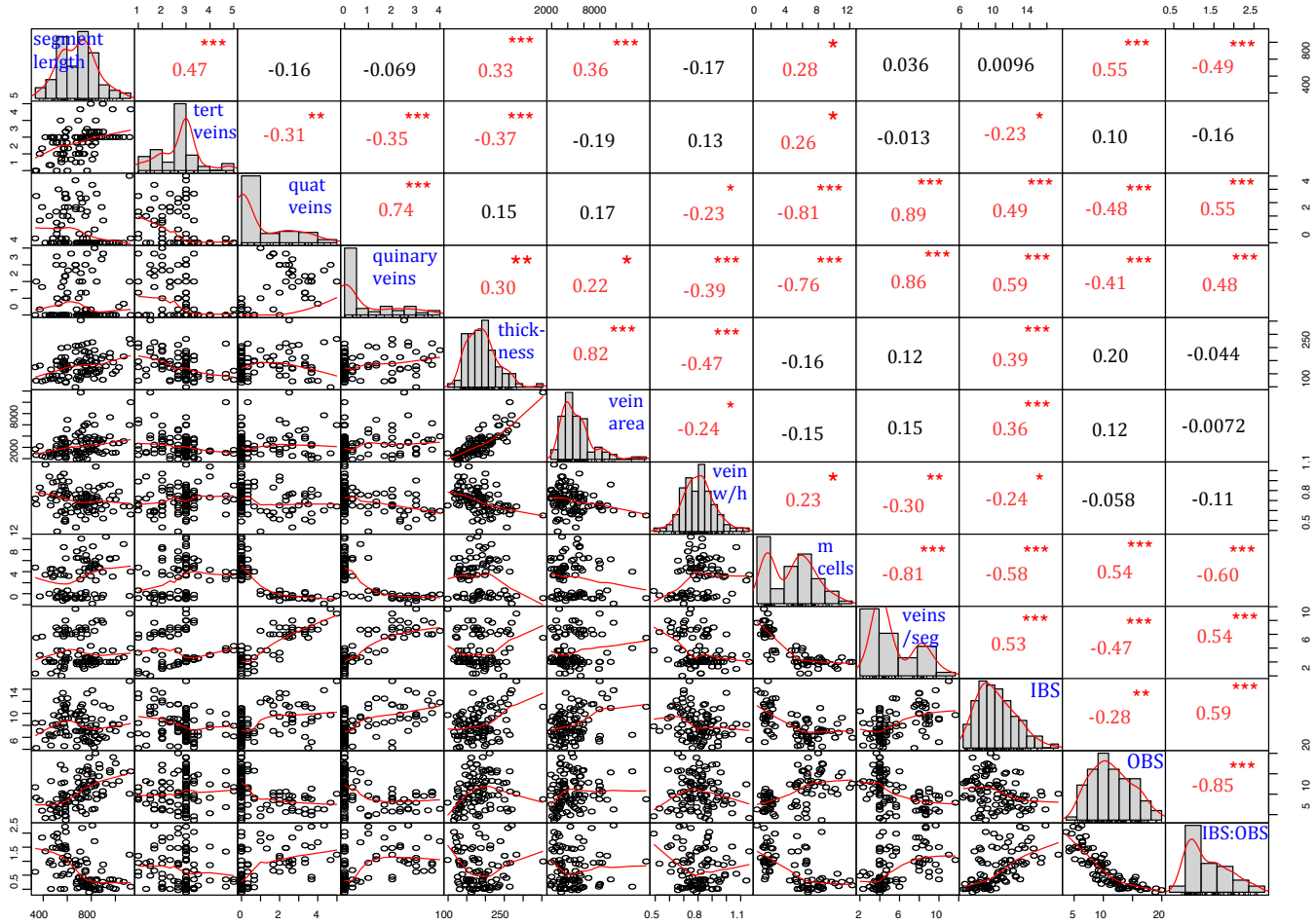


Figure 5.12 Correlation matrix for anatomical traits. The top quadrant shows the correlation value (R) and the result of the cor.test as stars, where p<0.05 (*), p<001(**), and p<0.001 (***). The bottom quadrant shows bivariate scatterplots, with a fitted line.

Anatomical variation in the geographic region of overlap

Anatomical traits varied across eastern Africa, the geographic region where the different photosynthetic types overlap (Figures 5.13 – 5.15). For example, *A. semialata* plants collected in central Africa (*i.e.*, Tanzania, the Democratic Republic of Congo, and Zambia) had smaller OBS than those collected in southern Africa (*i.e.*, South Africa and Zimbabwe; Figure 5.13). This trend exists for both C_4 and non-C_4 accessions. Interestingly, these smaller OBS caused the higher IBS:OBS ratios in non-C_4 central African accessions, while the higher IBS:OBS ratio in C_4 accessions appears to be driven by larger IBS across eastern Africa (Figure 5.13).

C_4 accessions from central and southern Africa expressed similar numbers of quaternary veins per segment, however, accessions collected in more southern locations (South Africa and Mozambique) produced more quinary veins than those from higher latitudes (Figure 5.14). Furthermore, non-C_4 accessions from central Africa had more tertiary veins per segment overall than those from southern Africa (Figure 5.14). Despite the variability in these phenotypes, non-C_4 accessions collected in central Africa also had fewer mesophyll cells separating veins, thinner leaves, shorter segments, smaller veins, and rounder veins overall than non-C_4 accessions from southern Africa (Figure 5.15). C_4 accessions also seemed to have thicker leaves and longer segments in southern than central Africa (Figure 5.15).

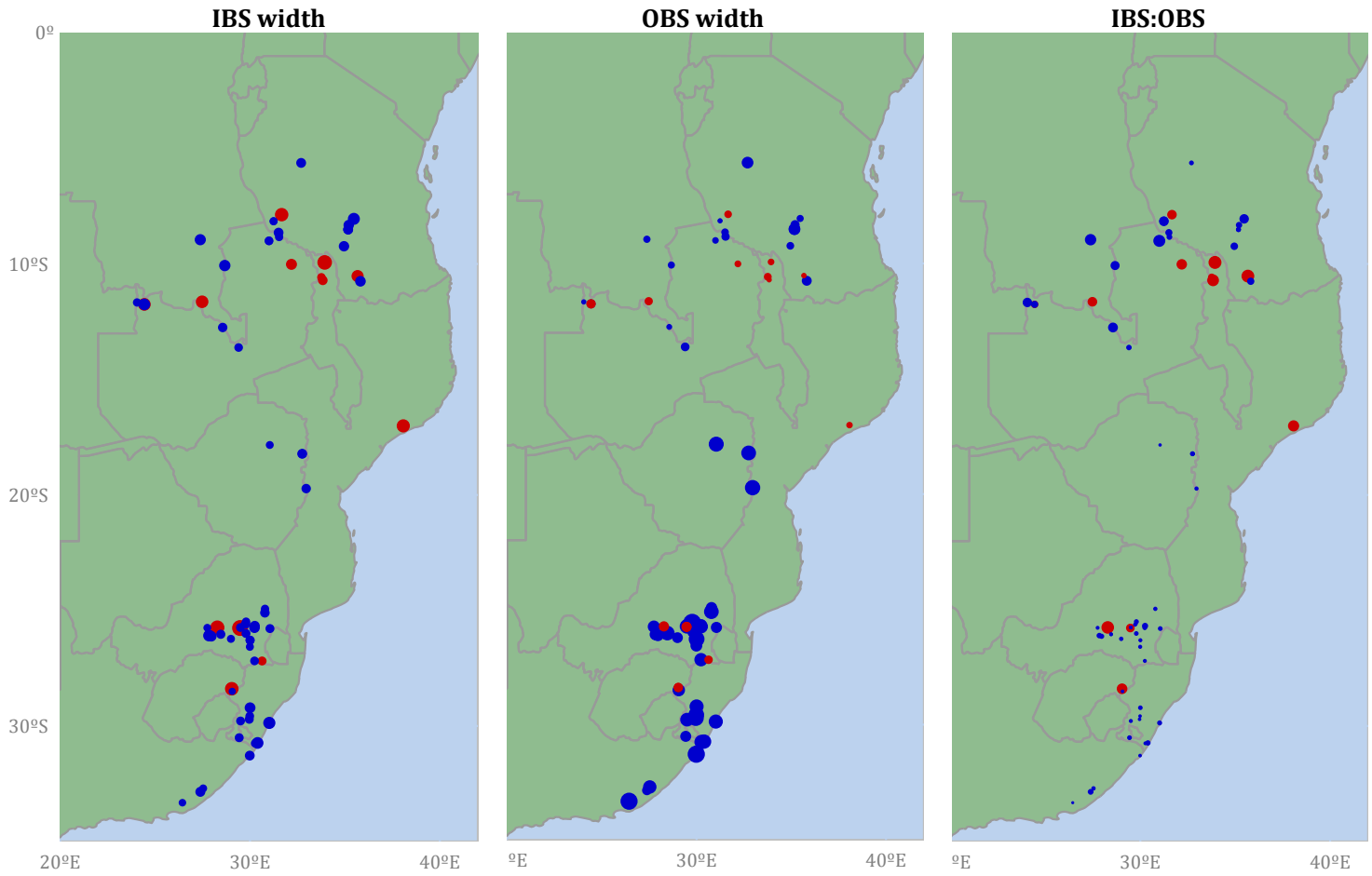


Figure 5.13 Geographic distribution of bundle sheath traits in non-C_4 (blue) and C_4 (red) accessions across the region where these photosynthetic types overlap in eastern Africa. Marker sizes correspond to values of anatomical traits. All cell widths are in micrometers. N=100.

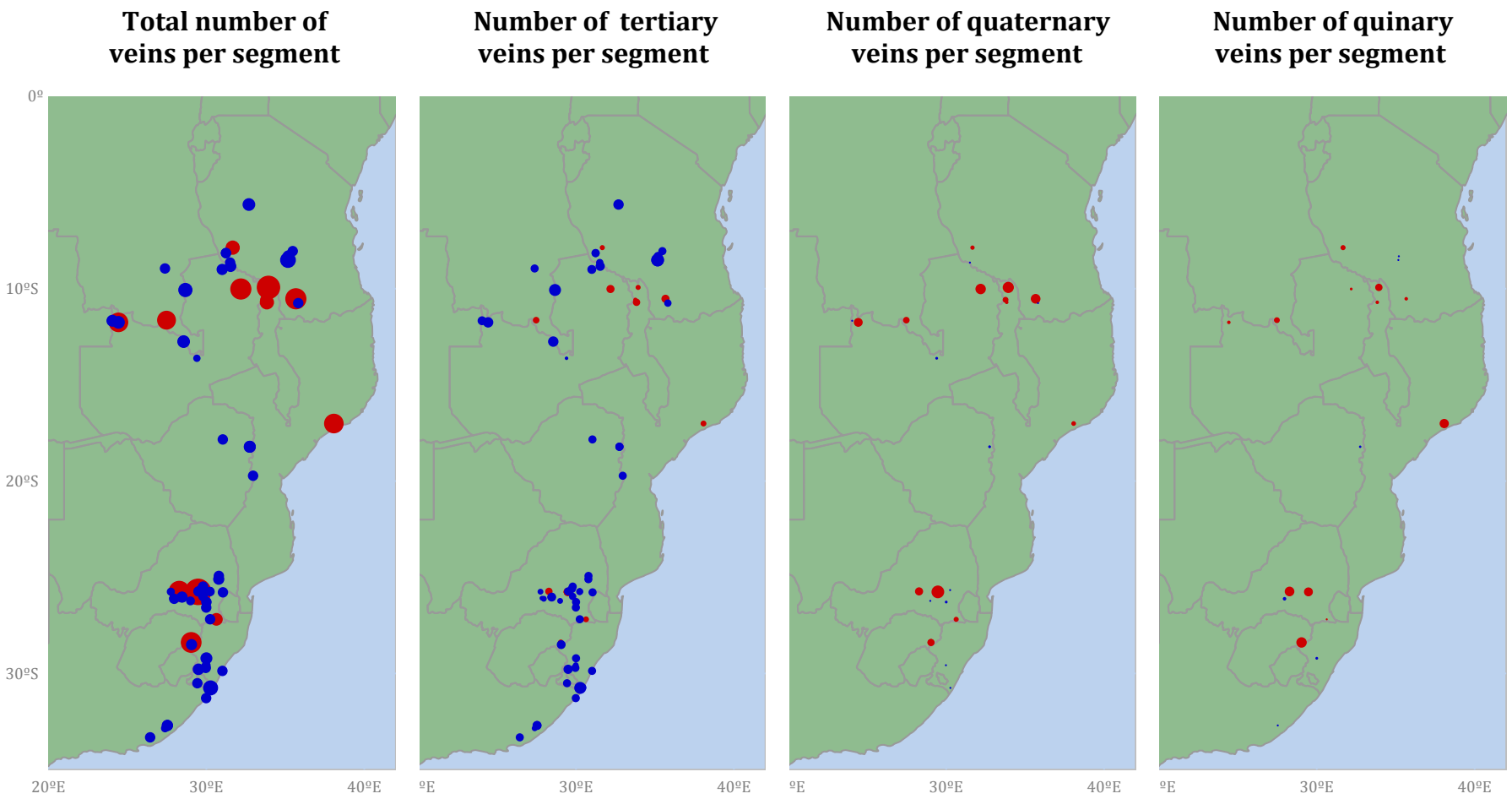


Figure 5.14 Geographic distribution of vein patterning in non-C_4 (blue) and C_4 (red) accessions across the region where these photosynthetic types overlap in eastern Africa. Marker sizes correspond to values of anatomical traits. N=100.

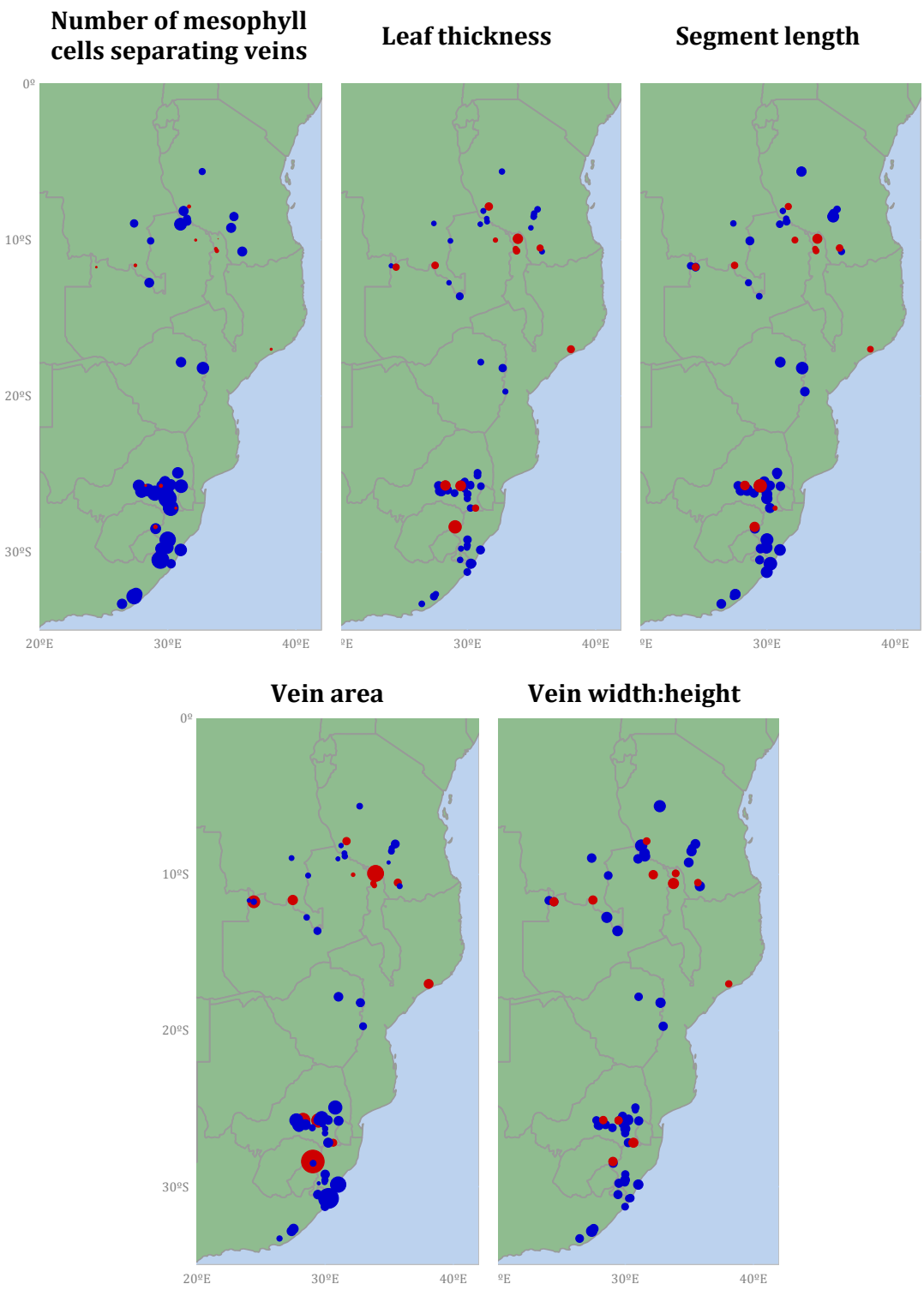


Figure 5.15 Geographic distribution of mesophyll, segment size, and vein shape traits in non-C_4 (blue) and C_4 (red) accessions across the region where these photosynthetic types overlap in eastern Africa. Marker sizes correspond to values of anatomical traits. All distances are in micrometers and areas in micrometers squared. N=100.

Environments associated with anatomical variation

The PCA on environmental variables explained over 55 percent of the variation in the dataset (Figure 5.16; Table 5.4). As found in Chapter 4, this PCA also shows that the non-C_4 specimens occupy a subset of the environments inhabited by the C_4 specimens. It also suggests that the distribution of *A. semialata* across environments is not driven by photosynthetic type (Figure 5.16a). Despite this, several anatomical traits correlate well with different environmental variables (Figure 5.16b). Results suggest that plants growing in drier areas with more variable temperatures may have thinner leaves with smaller veins (Figures 5.16 and 5.17; Appendix 5.3). Plants growing in warmer areas with seasonal precipitation patterns are more likely to have smaller segments, smaller OBS, fewer mesophyll cells, and higher total number of veins per segment than plants growing in cooler areas with less seasonal precipitation (Figures 5.16 and 5.17; Appendix 5.3). Plants growing in habitats with more wet days and warm season precipitation, better soil conditions, less sunshine, and lower relative humidity are likely to have more tertiary veins, rounder secondary veins, and thinner leaves (Figures 5.16 and 5.17; Appendix 5.3). Furthermore, plants collected in areas with poor soils, high potential evapotranspiration, and little dry season precipitation are more likely to have quaternary and quinary veins and larger IBS (Figures 5.16 and 5.17; Appendix 5.3). This suggests that different environments drive the insertion of tertiary versus quaternary and quinary veins, and that different environments drive the enlargement of inner versus outer bundle sheath cells.

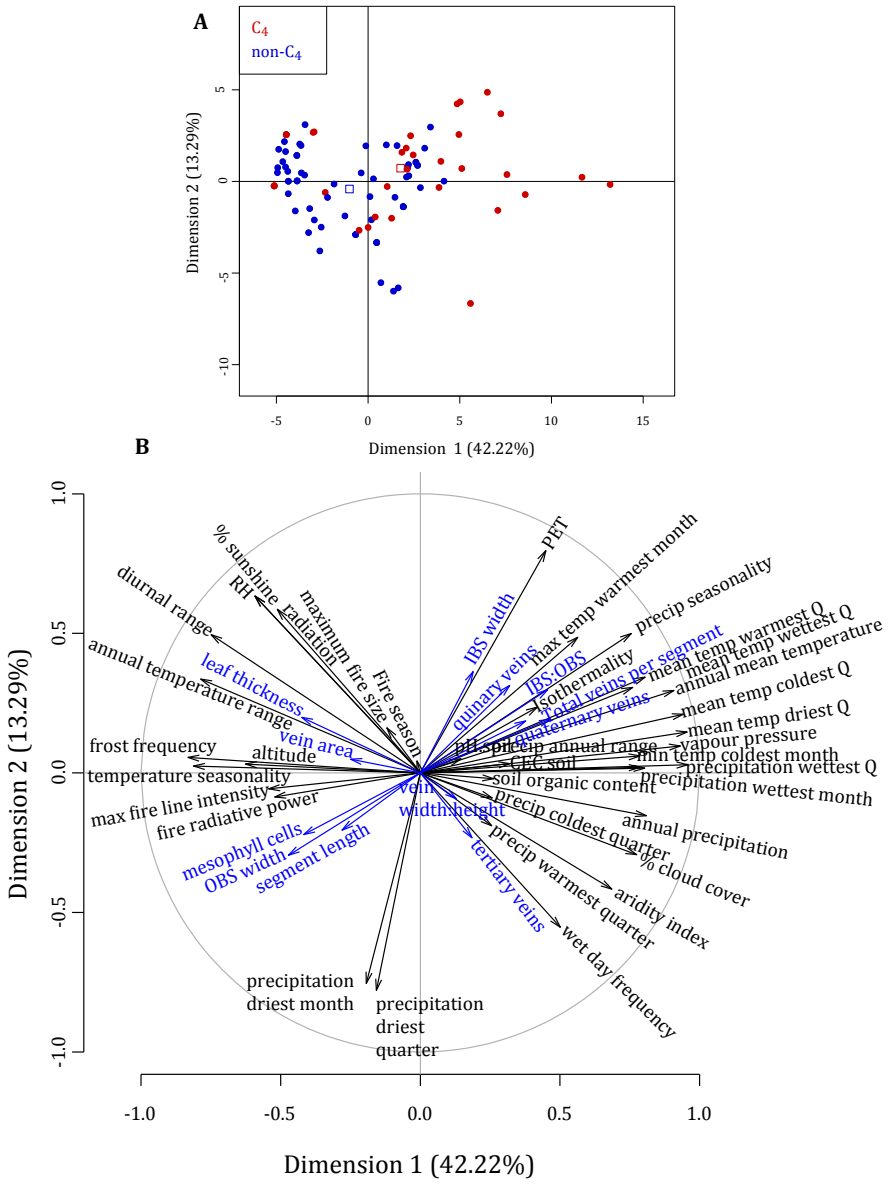


Figure 5.16 PCA on 37 environmental variables. Plot A shows where C_4 (red) and non-C_4 (blue) accessions fall along the first two dimensions. Plot B shows the eigenvectors (black text) of these environmental traits overlaid with 15 anatomical traits (blue text) as supplementary quantitative variables to aid interpretation. Note, aridity index increases with wetness.

Table 5.4 Coefficients for environmental variables significantly correlated with dimensions 1 and 2 of the PCA.

Variable	Correlation coefficient	p-value
Dimension 1		
Min temp coldest month	0.958	0.00E+00
Mean temp driest quarter	0.956	0.00E+00
Mean temp coldest quarter	0.942	0.00E+00
Vapour pressure	0.931	0.00E+00
Annual mean temperature	0.909	0.00E+00
Annual precipitation	0.809	0.00E+00
Mean temp warmest quarter	0.805	0.00E+00
Precipitation wettest quarter	0.804	0.00E+00
Precipitation wettest month	0.787	0.00E+00
Precip annual range	0.786	0.00E+00
Cloud cover percent	0.775	0.00E+00
Mean temp wettest quarter	0.761	0.00E+00
Max temp warmest month	0.756	0.00E+00
Aridity index	0.686	4.44E-15
Precipitation seasonality	0.564	1.20E-09
Wet day frequency	0.501	1.25E-07
Total veins per segment	0.466	1.19E-06
PET	0.450	2.99E-06
Isothermality	0.419	1.56E-05
CEC soil	0.320	1.22E-03
soil organic content	0.262	8.69E-03
Precipitation coldest quarter	0.258	1.01E-02
Precipitation warmest quarter	0.256	1.06E-02
Radiation	-0.512	6.02E-08
Fire radiative power	-0.522	3.03E-08
Maximum fire line intensity	-0.543	6.32E-09
Relative humidity	-0.592	1.08E-10
Sunshine percent	-0.592	1.08E-10
Altitude	-0.626	4.16E-12
Mean diurnal range	-0.749	4.57E-19
Annual temperature range	-0.786	5.15E-22
Temperature seasonality	-0.812	2.05E-24
Ground frost frequency	-0.832	1.42E-26
Dimension 2		
PET	0.797	0.00E+00
Relative humidity	0.634	1.92E-12
Sunshine percent	0.634	1.92E-12
Radiation	0.587	1.77E-10
Max temp warmest month	0.500	1.38E-07
Mean diurnal range	0.496	1.83E-07
Precipitation seasonality	0.486	3.34E-07

Variable	Correlation coefficient	p-value
Mean temp warmest quarter	0.343	5.04E-04
Annual temperature range	0.336	6.72E-04
Mean temp wettest quarter	0.307	2.02E-03
Annual mean temperature	0.296	2.97E-03
Isothermality	0.236	1.89E-02
Mean temp coldest quarter	0.210	3.66E-02
Cloud cover percent	-0.293	3.23E-03
Aridity index	-0.417	1.80E-05
Wet day frequency	-0.553	2.92E-09
Precipitation driest month	-0.754	2.00E-19
Precipitation driest quarter	-0.779	2.13E-21

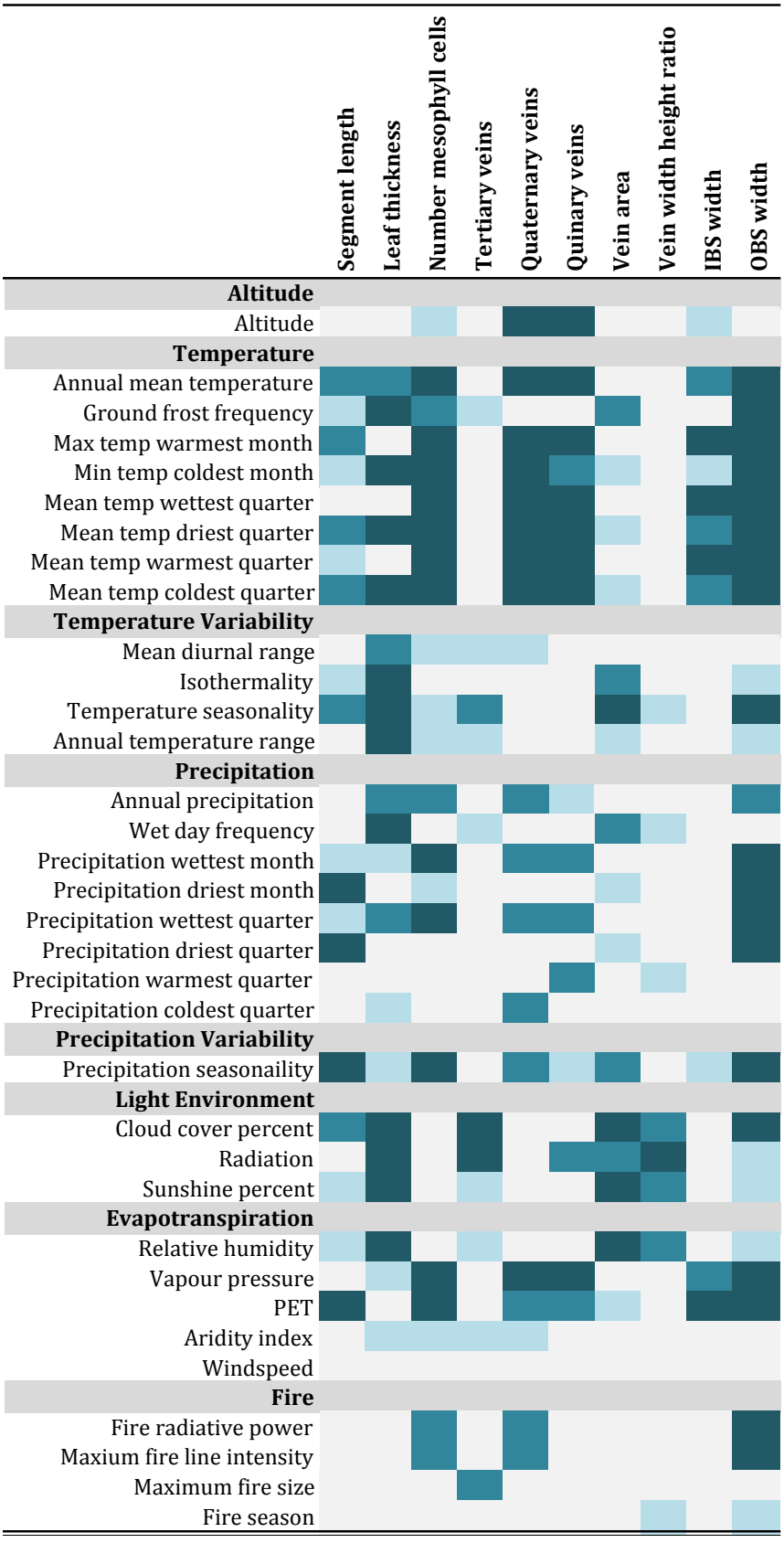


Figure 5.17 Color-coded results for correlation analyses of anatomical traits by environmental variables ($p<0.05$; $p<0.01$; $p<0.001$; NS (clear). R^2 and *p*-values are presented in Appendix 5.3.

Discussion

The physical landscape

Spanning three continents, *A. semialata* has a remarkably widespread distribution. The ability to disperse widely and into new environments may be partially facilitated by its physiological variability, as each of the C_3, C_3-C_4 intermediate, and C_4 photosynthetic types presents different environmental specializations and abilities to invade new habitats (Chapter 4). Moreover, anatomical variation across the species, and even within the three photosynthetic types, likely contributes toward this broad dispersal.

A. semialata plants that use C_4 photosynthesis grow throughout Australia, Asia, and Africa while those that do not use this pathway are confined to eastern Africa. Despite different patterns of geographic distribution, these two groups were often collected in similar environments. Chapter 4 established that the C_4 *A. semialata* lineages tolerate a wide range of environments without needing to adapt, while non-C_4 lineages are confined to a narrower set of habitats. Interestingly, these habitats differed between C_3 and C_3-C_4 intermediate lineages. Using a comprehensive review of herbarium collections, this chapter revealed a similar trend. Non-C_4 specimens experienced mildly variable environments in central and southern Africa and those that were inferred to be C_3-C_4 intermediates (*i.e.*, via physiology or stable isotope analyses) were narrowly confined to the high altitude, forest understories of central eastern Africa. *A. semialata* plants that use C_4 photosynthesis inhabit the most diverse environments and geographic distribution of the conspecifics.

These conclusions are based on herbarium collections that may be biased against areas that are difficult to collect, such as war-torn regions or environments too harsh to traverse (*e.g.*, Indonesia, India, and parts of central Africa), or where natural habitat has been destroyed through land use (*e.g.*, India). Moreover, a country's colonial past may have also influenced herbarium collection efforts. Indeed, most herbarium collections are made along the margins of roads and paths. Therefore, these geographic and environmental analyses may be more conservative than where this species has actually grown in the wild.

Bridging the selective landscape between C_3 and C_4 phenotypes

Overall, anatomical traits distinguish C_4 from non-C_4 photosynthetic types well. However, nearly every trait included in this study was continuous across these two groups, suggesting the gap between C_4 and non-C_4 phenotypes is very small. The hierarchical cluster analysis highlights this point well, as it presented three distinct clusters; one entirely comprised of C_4 accessions, another consisting primarily of C_3 and with C_3-C_4 plants, and finally a third, intermediate cluster. This cluster was primarily composed of putative C_3-C_4 intermediate accessions, but also included some believed to use C_3 and C_4 photosynthesis. Phenotypes in this cluster represent an anatomically intermediate step between typical C_3 and C_4 phenotypes and highlight the continuity between C_3, C_3-C_4 intermediate, and C_4 anatomies. This suggests that the C_3-C_4 intermediate phenotype narrows the gap between C_3 and C_4 phenotypes, and likely facilitates C_4 evolution. Specifically, these intermediate plants decrease the gaps in bundle sheath sizes, number of mesophyll cells between veins, and total number of veins per segment between C_3 and C_4 plants. As discussed in Chapters 2 and 3, these modifications would decrease the distance between CO_2 assimilation and reduction sites and also increase the bundle sheath to mesophyll area ratio, possibly also shifting the Rubisco levels between these cell types, which would enable C_4 evolution.

Interestingly, C_3-C_4 intermediate plants increase vein density by augmenting the number of tertiary veins, but not quaternary or quinary veins like C_4 plants typically do. Non-C_4 *A. semialata* plants never have quaternary or quinary veins, so it might be a large developmental jump to insert these new vein orders. Instead, adding more tertiary veins may be a less costly way to increase vein density. This phenotype might still allow these plants to live in slightly drier or sunnier habitats where selection would continue to favor C_4-like phenotypes.

Some anatomical characteristics of C_3-C_4 intermediate plants, including leaf thickness, vein area and shape, and number of tertiary veins, were not intermediate between C_3 and C_4 plants. Instead, C_3 and C_4 plants expressed similar phenotypes for these traits, while C_3-C_4 intermediate plants had distinctly thinner leaves with more tertiary veins and smaller, rounder veins than C_3 and C_4 plants. This suggests that these phenotypes are unique to the C_3-C_4 intermediate lineage

(Figure 5.18). As discussed in Chapter 3, C_3-C_4 intermediate plants seem to exclusively inhabit miombo woodlands of central Africa, a characteristically heterogeneous habitat (Boaler, 1966; Frost, 1996). Having many small and round veins may be linked to water transport, increasing conductance (Kawamitsu *et al.,* 1985; Ogle, 2003), which would be important for perennial plants that must experience precipitation extremes throughout the year (Frost, 1996). Thinner leaves with more tertiary veins may also increase light absorptance (Ogle, 2003), which would be useful in these miombo woodlands that remain shady most of the year (van der Meulen and Werger, 1984).

Chapter 4 proposes that the MRCA of *A. semialata* was a C_3-C_4 intermediate and that the C_3 lineage represents a reversion from this intermediate state. The C_3-C_4 intermediate phenotype would also facilitate the transition back to the C_3 pathway. The number of mesophyll cells between veins is the only trait that clearly distinguished C_4 from non-C_4 specimens in this study. This may reflect a minimum requirement of mesophyll area needed to maintain the Calvin cycle in the mesophyll cells of non-C_4 plants. The C_3-C_4 intermediate plants in this study expressed mesophyll cell numbers similar to the C_3 plants, with no intermediate plant expressing fewer than four cells between veins. Thus, C_3-C_4 plants are anatomically capable of running C_3 photosynthesis and would only need to decrease the degree of C_2 cycle activity to revert to a C_3 state. This could be as easy as reactivating mesophyll GDC, the genetics for which are already in place and would just need to be switched on. From there, selection would have further enhanced mesophyll area by increasing segment sizes and mesophyll cell numbers (Figure 5.18). The larger IBS cells sizes were likely lost (Figure 5.18) when they were no longer needed to maintain C_2 cycle activity.

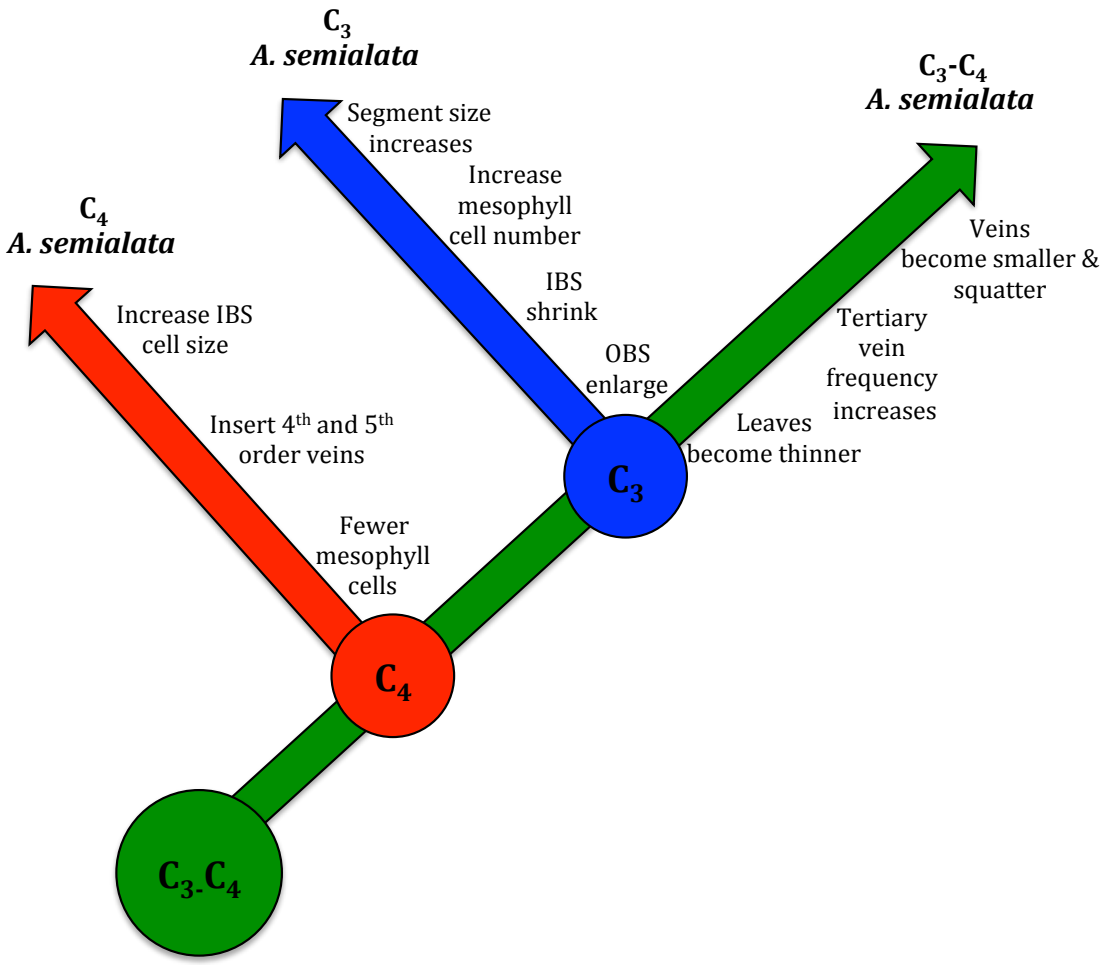


Figure 5.18 Hypothetical evolutionary trajectory of anatomical traits within *Alloteropsis semialata*. Blue represents a C_3 state, green a C_3-C_4 intermediate state, and red a C_4 state.

Potential anatomical enablers

Spatial patterns in anatomical traits are apparent along the latitudinal gradient of eastern Africa, the region where the greatest photosynthetic variation in *A. semialata* has been found. Chapter 4 identified central eastern Africa as the origin of this rich variation, with C_4 lineages migrating from there soon after the diversification of this pathway and a C_3-C_4 intermediate lineage reverting back to C_3 photosynthesis as it moved southward. Thus, the anatomical gradients along this specific corridor should reflect these physiological transitions.

Leaf anatomy of the non-C_4 accessions differs greatly between central and southern Africa. Namely, the central African accessions have fewer mesophyll cells separating veins, thinner leaves, shorter segments, smaller veins, smaller OBS, and

larger IBS than the southern non-C_4 accessions. Much of this variation is likely due to the spatial trend observed in C_3-C_4 and C_3 specimens, with intermediate plants inhabiting central Africa while plants using the C_3 pathway are confined to southern Africa (Chapter 4). As established in Chapter 3, plants using these different photosynthetic types have characteristic leaf anatomies. However, some of this continuous variation may reflect natural selection that, over time, drove more C_3-like phenotypes in C_3-C_4 intermediate plants or C_3-C_4 intermediate- like phenotypes in C_3 plants, as they experience environments more typical of the other pathway. Indeed, the anatomical variation present in South Africa showed that some non-C_4 plants had bundle sheath sizes, tertiary vein frequencies, mesophyll cell numbers, and vein sizes and shapes similar to central African plants. Perhaps these phenotypes could be anatomical enablers for C_3-C_4 intermediacy, facilitating its establishment in southern Africa if environmental conditions shifted in favour of C_2 physiology. More work is needed to confirm these hypotheses.

Anatomical phenotypes expressed by C_3-C_4 intermediate plants may also be enablers for the establishment of the C_4 cycle. Chapter 4 suggests that C_4 *A. semialata* probably evolved from C_3-C_4 intermediate ancestors. Thus, anatomical phenotypes present in C_3-C_4 plants may have facilitated the adoption of the C_4 pathway, assuming the current C_3-C_4 intermediate phenotype is similar to that expressed by the ancestor of C_4 *A. semialata*. For example, OBS sizes are reduced in C_3-C_4 intermediates compared to C_3 plants. As discussed in Chapter 3, these OBS likely present a hurdle to metabolite transfer and, as such, any decrease in their size should enhance C_2 and C_4 cycle efficiency. As much as the smaller OBS phenotype suits the C_2 cycle well, it would also easily facilitate the establishment of a C_4 cycle. Similarly, C_3-C_4 intermediate plants had fewer mesophyll cells between veins than C_3 plants and while these numbers never reached the very low C_4 levels (*i.e.,* < 4 cells) in the intermediates, any decrease in mesophyll volume would reduce the distance between CO_2 assimilation and reduction sites and increase bundle sheath chloroplast light capture. Thus, the reduced number of mesophyll cells between veins may have also eased the transition to C_4 photosynthesis.

Environmental drivers of photosynthetic diversification

Leaf anatomy varied within the photosynthetic types. Specifically, leaf thickness, vein shape, and vein area varied greatly across C_3 and C_4 accessions. In contrast, the C_3-C_4 intermediate accessions were anatomically and ecologically canalized, varying little in these traits and in collection environments. The diversity of collection environments likely explains the anatomical variation seen within C_3 and C_4 plants. *A. semialata* plants collected in variable, dry, warm, and sunny environments had thicker leaves, with larger vein and xylem areas than those collected in cloudy, wet, and cooler habitats. Vein shape was influenced by precipitation, but also soil condition. Specifically, round vein shape, the trait most strongly linked to C_3-C_4 intermediacy, exists in areas with high precipitation and fertile soils. If the transition from C_3 to C_4 must pass through an intermediate state, then the evolution of C_4 photosynthesis in *A. semialata* may have been confined to wet, high nutrient areas for part of this transition. This confinement may have shaped where and how easily C_4 photosynthesis evolved in this species.

Conclusions

This study used herbarium specimens to comprehensively characterize the geographical, environmental, and anatomical variation present in *A. semialata*. Analysis of $\delta^{13}C$ values revealed that C_4 and non-C_4 specimens of *A. semialata* have different distributions. C_4 plants have been collected in northern and eastern Australia, southern Asia, and western, central, and southern Africa. In contrast, C_3 accessions were only collected in southeastern Africa and isotopically intermediate accessions in central eastern Africa. These results are consistent with the distribution patterns presented in Chapter 4. Leaf anatomy was continuous across the C_3/C_4 spectrum in this species and only counts of the minimum number mesophyll cells between veins could distinguish C_4 from non-C_4 specimens consistently. The isotopically intermediate specimens, in particular, were anatomically intermediate, and functioned to narrow the gap between typical C_3 and C_4 phenotypes. It is hypothesized that these intermediate phenotypes facilitated C_4 evolution, as well a reversion to the C_3 pathway. If this is true, then the narrow ecological niche of the C_3-C_4 intermediate specimens may have shaped where the complex C_4 pathway could have evolved.

Appendix 5.1 Details of living accessions included in this study.

Accession	Latitude	Longitude	Type	$\delta^{13}C$	Month	Year	Country	Anatomy
Christin 01a	-5.93	32.60	C_3	-25.7	January	2014	Tanzania	
FW Boston	-29.57	29.99	C_3	-25.7	December	2012	South Africa	Yes
FW Mooi River	-29.21	30.00	C_3	-28.0	November	2012	South Africa	Yes
Lundgren-ASM	-25.61	29.73	C_3	-26.4	March	2012	South Africa	Yes
Lundgren-BLW	-29.71	29.96	C_3	-28.0	March	2012	South Africa	Yes
Lundgren-CRL	-25.74	30.24	C_3	-24.8	March	2012	South Africa	Yes
Lundgren-EML	-26.29	30.00	C_3	-26.3	March	2012	South Africa	Yes
Lundgren-JMS	-33.32	26.44	C_3	-26.8	March	2012	South Africa	Yes
Lundgren-KSD	-30.51	29.43	C_3	-28.0	March	2012	South Africa	Yes
Lundgren-KWT	-32.70	27.53	C_3	-26.3	March	2012	South Africa	Yes
Lundgren-LSU	-25.58	29.77	C_3	-25.7	March	2012	South Africa	Yes
Lundgren-MDB	-25.74	29.50	C_3	-23.2	March	2012	South Africa	Yes
Lundgren-MTP	-25.62	30.26	C_3	-26.0	March	2012	South Africa	Yes
Lundgren-PGR	-24.93	30.79	C_3	-25.2	March	2012	South Africa	Yes
Lundgren-SFB	-25.48	29.79	C_3	-26.6	March	2012	South Africa	Yes
Lundgren-SNR	-28.50	29.06	C_3	-27.7	March	2012	South Africa	Yes
Lundgren 01a	-5.63	32.69	C_3-C_4	-26.3	January	2014	Tanzania	Yes
Lundgren 04a	-8.51	35.17	C_3-C_4	-23.1	January	2014	Tanzania	Yes
Lundgren 04d	-8.51	35.17	C_3-C_4	-26.4	January	2014	Tanzania	Yes
Besnard-Majunga	-15.67	46.37	C_4	-11.8*	November	2011	Madagascar	Yes
Chatelet-TWN10	24.47	120.72	C_4	-	April	2014	Taiwan	Yes

Accession	Latitude	Longitude	Type	$\delta^{13}C$	Month	Year	Country	Anatomy
Dillon-QSLD	19.62	-146.96	C_4	-12.1*	February	2005	Australia	
Lundgren 02e	-9.04	32.48	C_4	-11.9	January	2014	Tanzania	
Lundgren-MGD	-25.76	29.47	C_4	-13.7	March	2012	South Africa	Yes
Lundgren-PRE	-25.74	28.28	C_4	-13.0	March	2012	South Africa	Yes
Lundgren-SFD	-28.39	29.04	C_4	-12.4	March	2012	South Africa	Yes
Sanou BUR-734	10.85	-4.83	C_4	-11.3*	July	2009	Burkina Faso	Yes

* $\delta^{13}C$ adjusted as described in Chapter 3.

§ The final column notes which accessions were included in the anatomy dataset.

Appendix 5.2 Details of herbarium specimens included in this study.

Accession	Latitude	Longitude	Type	$\delta^{13}C$	Month	Year	Country
Acocks 12556	-26.85	28.88	C_3	-25.7	March	1946	South Africa
Acocks 60	-23.90	29.45	C_3	-25.1	October	1938	South Africa
Adamson 191	-15.50	35.25	C_4	-11.0		1906	Malawi
Alia 8749	28.39	84.12	C_4	-10.5		1821	Nepal
Anderson 875	-10.38	38.77	C_4	-10.6	April	1953	Tanzania
Astle 1205*	-10.02	32.18	C_4	-10.0	January	1962	Zambia
Ballingal NHA477	-27.37	151.10	C_4	-12.2	November	1984	Australia
Barenburg 2	-24.52	28.72	C_4	-11.6	January	1951	South Africa
Basedow 119	-13.17	134.50	C_4	-10.9		1928	Australia
Baur 316	-31.62	28.43	C_3	-25.3	October	1885	South Africa
Behr 254*	-26.09	27.84	C_3	-27.8	December	1982	South Africa
Bick 27	-18.15	142.87	C_4	-11.8	-	-	Australia
Bingham 13209	-8.60	31.24	C_3	-27.3	December	2006	Zambia
Blaxell 849	-23.83	149.17	C_4	-11.7	November	1972	Australia
Boaler 730	-8.03	33.25	C_4	-11.9	November	1962	Tanzania
Bogdan & Williams 238	1.03	35.00	C_4	-11.9	January	1947	Kenya
Bogdan 3722	1.03	35.00	C_4	-11.7	May	1953	Kenya
Bor 211	25.30	91.70	C_4	-12.4	April	1946	India
Bradfield 244	-26.18	28.32	C_3	-25.2	October	1934	South Africa
Brain 7306*	-18.22	32.75	C_3	-25.1	November	1931	Zimbabwe
Brandmuller 47	-27.25	27.75	C_4	-12.3	January	1923	South Africa
Brass 18477*	-10.78	142.47	C_4	-10.7	April	1948	Australia
Brass 6343	-9.08	143.20	C_4	-10.4	March	1936	Papua New Guinea
Bredenkamp 886	-26.50	28.33	C_4	-12.6	December	1972	South Africa
Brunt 279	5.75	10.50	C_4	-10.7	March	1962	Cameroon
Brzostowski 26*	-8.05	35.47	C_3-C_4	-21.9	December	1944	Tanzania
Bullock 1857	-7.87	31.67	C_4	-10.4		1950	Tanzania
Bullock 1979*	-8.15	31.24	C_3-C_4	-19.7	December	1949	Tanzania
Bullock 1980*	-8.15	31.24	C_3-C_4	-22.3	December	1949	Tanzania
Bullock 2128*	-9.00	31.00	C_3-C_4	-23.0	December	1949	Zambia
Burnett Davy 287	-25.50	28.00	C_4	-11.3	-	-	South Africa
Burtt Davy 1020	-26.93	29.18	C_4	-11.0	January	1904	South Africa
Chapama 903	-11.92	34.08	C_3		December	2008	Malawi
Chase -	-18.97	32.66	C_4	-11.0	November	1948	Zimbabwe
Chevalier 13580*	10.27	-12.44	C_4	-10.0	November	1908	Guinea
Clarke 34738	23.97	86.13	C_4	-11.2	April	1884	India
Clarke 38268	25.58	91.63	C_4	-11.5	June	1885	India
Clarkson 4070	-13.01	142.08	C_4	-11.3	December	1981	Australia

Accession	Latitude	Longitude	Type	$\delta^{13}C$	Month	Year	Country
Clarkson 6417	-10.58	142.22	C_4	-10.9	April	1986	Australia
Cleghorn 3054	-18.39	32.73	C_3	-25.6	July	1975	Zimbabwe
Coetzee 127*	-25.75	27.75	C_3	-26.7	January	1970	South Africa
Coetzee 423	-25.75	27.50	C_3	-26.7	November	1970	South Africa
Coetzee 496	-25.75	27.75	C_3	-26.1	November	1970	South Africa
Coetzee 497	-25.75	27.75	C_4	-11.3	November	1970	South Africa
Compere 1425	-4.30	15.30	C_4	-10.0	February	1960	DRC
Conradie 7	-18.13	30.15	C_4	-10.2	November	1955	Zimbabwe
Cowdry 51	-24.42	150.50	C_4	-10.8		1930	Australia
Crook 2103	-18.43	32.78	C_3	-24.7	December	1975	Zimbabwe
Crook 290	-19.78	32.87	C_3	-25.3	November	1950	Zimbabwe
Curran 16528	16.57	121.26	C_4	-11.2	January	1909	Philippines
de Wilde 10185	7.43	36.38	C_4	-12.0	November	1966	Ethiopia
de Wit 405*	-25.73	28.28	C_4	-10.4	January	1955	South Africa
de Wit 562	-26.33	31.13	C_3	-26.0	April	1930	Swaziland
de Witte 7192	-9.00	26.58	C_4	-10.0	August	1949	DRC
Deall 1207*	-25.10	30.78	C_3	-27.1	October	1981	South Africa
Devenish 1284*	-27.19	30.63	C_4	-10.5	November	1965	South Africa
du Toit 579	-24.68	28.69	C_4	-11.2	February	1975	South Africa
Ducloux 5622	25.48	100.58	C_4	-12.1	May	1908	China
Edwards 1052	-27.66	30.13	C_4	-11.8	November	1955	South Africa
Ellis 1244	-25.74	28.28	C_4	-11.5	December	1972	South Africa
Ellis 1748	-25.67	28.23	C_4	-11.4	November	1973	South Africa
Ellis 1785	-25.64	28.32	C_4	-12.2	December	1973	South Africa
Ellis 1789	-25.68	28.36	C_4	-10.9	December	1973	South Africa
Ellis 1828	-26.03	28.46	C_3	-26.7	January	1974	South Africa
Ellis 1829*	-26.03	28.46	C_3	-26.6	January	1974	South Africa
Ellis 1830*	-26.03	28.46	C_3	-26.4	January	1974	South Africa
Ellis 1861	-26.50	30.50	C_3	-25.4	January	1974	South Africa
Ellis 1862	-26.50	30.50	C_3	-25.3	January	1974	South Africa
Ellis 1865	-23.87	30.02	C_3	-25.5	January	1974	South Africa
Ellis 2571	-33.31	26.52	C_3	-28.5		1975	South Africa
Ellis 2606	-32.58	27.89	C_3	-26.7		1975	South Africa
Ellis 2832	-25.73	29.50	C_4	-11.7	November	1976	South Africa
Ellis 2833	-25.73	29.50	C_4	-11.9	November	1976	South Africa
Ellis 2836*	-25.73	29.50	C_3	-24.2	November	1976	South Africa
Ellis 2979	-30.57	29.67	C_3	-24.5	December	1976	South Africa
Ellis 3527	-26.44	29.95	C_4	-12.3	November	1980	South Africa
Ellis 36	-27.50	29.75	C_4	-11.7	November	1969	South Africa
Ellis 378	-25.75	31.00	C_3	-26.2	January	1971	South Africa
Ellis 3807	-26.44	29.95	C_4	-11.8	November	1981	South Africa

Accession	Latitude	Longitude	Type	$\delta^{13}C$	Month	Year	Country
Ellis 3808	-26.25	29.83	C_4	-11.6	November	1981	South Africa
Ellis 721	-25.79	28.31	C_4	-11.5	November	1971	South Africa
Ellis 733*	-30.73	30.40	C_3	-26.8	December	1971	South Africa
Emson 340*	-9.23	34.95	C_3-C_4	-21.4		1932	Tanzania
Estenhuysen 27842	-29.75	29.25	C_3	-25.9	July	1958	South Africa
Evans & Evans 3839*	-12.72	131.17	C_4	-14.1	December	1993	Australia
Everist 7738	-26.05	152.96	C_4	-11.4	October	1964	Australia
Eyles 1920	-17.82	31.04	C_4	-10.3	November	1919	Zimbabwe
Faden 68/223	-1.03	37.08	C_4	-12.5	May	1968	Kenya
Fries 5318	-18.54	32.12	C_3	-25.3	September	1930	Zimbabwe
Fries et al. 3315	-18.53	32.12	C_4	-11.7	November	1930	Zimbabwe
Fryar 3944	-7.20	146.64	C_4	-10.7	February	1950	Papua New Guinea
Galpin 13104*	-25.67	30.25	C_3	-24.5	November	1932	South Africa
Galpin 5858	-31.96	26.87	C_3	-24.6	February	1901	South Africa
Gerstner 5663	-23.98	30.07	C_3	-25.5	November	1940	South Africa
Gillieat 141	-27.05	153.00	C_4	-11.0	April	1964	Australia
Gilliland 449	-18.72	32.82	C_3	-26.1	June	1934	Zimbabwe
Gossweiller 2670	-12.58	13.41	C_4	-10.4	February	1906	Angola
Gould 13355	6.83	80.89	C_4	-13.6	April	1970	Sri Lanka
Gould 13495*	6.50	81.07	C_4	-11.3	April	1970	Sri Lanka
Grant 121	-13.63	32.65	C_4	-11.0	January	1954	Zambia
Greenway & Trapnell 5716	-11.21	31.77	C_3	-28.5	September	1938	Zambia
Groaneodijk et al 835	-13.13	39.00	C_4	-11.1	January	1984	Mozambique
Grout 121	-13.67	32.70	C_4	-11.6	January	1954	Zambia
Haine 242	-19.02	47.18	C_4	-11.3	January	1969	Madagascar
Hakki et al 478	9.45	0.82	C_4	-11.1	April	1978	Togo
Handel & Mazzetti 13079	24.90	102.73	C_4	-11.8	May	1917	China
Hanekom 1696	-25.71	28.19	C_4	-12.4	December	1971	South Africa
Harrison 339	-28.37	32.22	C_4	-11.5	December	1967	South Africa
Harrold 269	-26.05	153.08	C_4	-11.2	June	1972	Australia
Henderson et al. 646	-23.83	149.12	C_4	-12.8	April	1971	Australia
Hitchins 716	-28.27	31.90	C_4	-12.0	November	1971	South Africa
Homble 8	-3.29	26.54	C_4	-11.5		1911	DRC
Hoogland & Macdonald 3450	-8.61	148.27	C_4	-11.9	July	1953	Papua New Guinea
Hsu 5299	24.66	120.91	C_4	-11.2	April	1969	Taiwan
Hubbard 6549	-20.02	148.23	C_4	-10.5	January	1931	Australia
Humbert 19320*	-22.42	45.83	C_4	-16.3	November	1946	Madagascar
Humbert 7108	-19.87	47.03	C_4	-11.6	December	1928	Madagascar
Jackson 1181*	-9.93	33.93	C_4	-10.2	March	1953	Malawi
Jackson 1323*	4.12	33.78	C_4	-9.5	March	1950	South Sudan

Accession	Latitude	Longitude	Type	$\delta^{13}C$	Month	Year	Country
Jackson 226	-15.84	35.07	C_4	-10.8	November	1950	Malawi
Jackson 2272	-14.37	34.33	C_4	-11.3	January	1959	Malawi
Jacobs 1786	-13.00	132.50	C_4	-13.8	May	1974	Australia
Jacques-Felix 3960	7.33	13.58	C_4	-10.6	June	1937	Cameroon
Jordan 869	8.77	-12.78	C_4	-10.7	May	1953	Sierra Leone
K'Tung 78	23.14	113.27	C_4	-10.3	April	1978	China
Keay 25797	10.52	7.44	C_4	-10.3	May	1950	Nigeria
Kenneally 8625B	-14.77	125.80	C_4	-11.5	December	1982	Australia
Koelz 32989	23.48	92.89	C_4	-10.6	April	1953	India
Kornas 3104	-15.33	28.42	C_4	-12.4	January	1973	Zambia
Kuhne -*	-26.57	29.99	C_3	-25.4	January	1946	South Africa
Lam & Meeuse 5092	-25.78	31.05	C_3	-25.1	October	1938	South Africa
Larsen 2025	18.03	98.40	C_4	-11.7	July	1968	Thailand
Lazarides 5612	-21.66	150.26	C_4	-10.3	April	1956	Australia
Lazarides 7323	7.15	81.11	C_4	-11.5	September	1970	Sri Lanka
Lea 87	-20.09	34.34	C_4	-10.7	October	1935	Mozambique
Leech 123	-13.12	32.98	C_4	-11.4	January	1945	Zambia
Lely 242	10.32	9.84	C_4	-10.5	April	1929	Nigeria
Liebenberg 8413*	-26.12	27.96	C_3	-27.5	October	1976	South Africa
Loher 1672*	12.92	123.12	C_4	-9.9	July	1895	Philippines
MacGillivray -	-20.05	148.89	C_4	-12.0	January	1933	Australia
Marko 9900	23.13	113.25	C_4	-10.8	February	1863	China
Martensz and Schodde AE793*	-12.38	132.95	C_4	-14.5	February	1973	Australia
Maserumule 95*	-26.00	29.79	C_3	-27.3	December	2003	South Africa
Mbano DSM812*	-10.75	35.82	C_3-C_4	-23.9	December	1969	Tanzania
McCallum Webster 329	-29.26	30.00	C_4	-11.8	January	1959	South Africa
McCallum Webster A42	-8.60	31.24	C_3	-24.8	February	1959	Zambia
McClean 322*	-30.75	30.27	C_3	-28.8	April	1937	South Africa
McKean B488	-12.38	130.89	C_4	-11.7	March	1972	Australia
McKee 6138	20.78	97.03	C_4	-11.3	May	1958	Burma
McKee 8697	-14.33	132.33	C_4	-11.6	February	1961	Australia
Milne-Redhead & Taylor 8140	-10.67	35.60	C_4	-12.3	January	1956	Tanzania
Milne-Redhead & Taylor 8455*	-10.53	35.67	C_4	-10.7	January	1956	Tanzania
Milne-Redhead 3021*	-11.67	24.05	C_3-C_4	-20.0	October	1937	Zambia
Milne-Redhead 3360*	-11.75	24.43	C_4	-10.1	November	1937	Zambia
Milne-Redhead 3371*	-11.75	24.43	C_3-C_4	-18.6	November	1937	Zambia
Mogg 3361	-29.18	30.00	C_4	-11.3	October	1918	South Africa
Mooney 3776	21.45	85.18	C_4	-11.0	April	1950	India
Moore 18	-26.78	151.10	C_4	-11.3	March	1920	Australia

Accession	Latitude	Longitude	Type	$\delta^{13}C$	Month	Year	Country
Morton A3243a	6.73	-1.34	C_4	-10.0	April	1958	Ghana
Munch 167*	-19.72	32.96	C_3	-26.4	June	1949	Zimbabwe
Mundy -	-16.68	31.40	C_4	-11.0	-	1913	Zimbabwe
Mundy 15	-17.84	31.04	C_4	-11.5	-	1921	Zimbabwe
Mundy 2162*	-17.84	31.04	C_3	-24.4		1913	Zimbabwe
Myers 10935	4.08	33.52	C_4	-10.5	November	1939	South Sudan
Naegele 85-24	9.71	2.91	C_4	-11.0	June	1985	Benin
Pawek 10770	-11.43	33.54	C_4	-11.6	January	1976	Malawi
Pawek 7609	-11.45	33.92	C_4	-11.2	December	1973	Malawi
Phillipson 1539*	-31.28	29.99	C_4	-14.4	November	1986	South Africa
Phipps & Vesey-Fitzgerald 3176	-9.22	29.30	C_4	-11.2	April	1961	Zambia
Pierre -	11.00	107.00	C_4	-9.1	-	-	Vietnam
Pobeguin 1833	10.63	-11.83	C_4	-10.6	March	1909	Guinea
Poelman 109	-11.64	27.48	C_4	-10.8	November	1961	DRC
Poelman 92*	-11.64	27.48	C_4	-10.6	November	1961	DRC
Pooma et al 4041*	17.31	101.21	C_4	-10.1	May	2004	Thailand
Proctor 2165*	-8.32	35.20	C_3-C_4	-23.2	November	1962	Tanzania
Proctor 2206*	-8.83	31.53	C_3	-24.2	December	1962	Zambia
Pulsford 9	-3.57	143.63	C_4	-10.5	May	1953	Papua New Guinea
Ramos 15081*	14.95	120.09	C_4	-9.9	June	1912	Philippines
Ramos 32791*	16.06	119.87	C_4	-9.2	July	1918	Philippines
Rattray 1878*	-20.42	28.48	C_4	-11.6	February	1960	Zimbabwe
Rattray 997	-18.22	32.75	C_3	-24.1	October	1946	Zimbabwe
Raynal 13309	6.85	14.27	C_4	-10.1	January	1965	Cameroon
Reekmans 7226	-3.07	30.50	C_3	-24.9	October	1978	Burundi
Repton 4496	-25.88	27.62	C_3	-26.1	-	-	South Africa
Richardson -	-28.35	28.37	C_4	-10.1	October	1901	South Africa
Rinder 8	-23.50	29.93	C_4	-10.0	March	1932	South Africa
Robinson 3055*	-10.58	33.76	C_4	-14.0	January	1959	Malawi
Robinson 4744*	-9.00	31.00	C_3-C_4	-23.0	December	1961	Zambia
Robinson 48530	-19.50	30.90	C_4	-10.8	January	1954	Zimbabwe
Rogers 10988	-11.67	27.47	C_4	-10.7	May	1914	DRC
Rogers 30270*	-25.78	31.05	C_3	-24.0	-	-	South Africa
Ruffo & Kisena 2806	-7.87	31.67	C_3-C_4	-18.6	November	1987	Tanzania
Saayman 1	-25.79	27.93	C_4	-11.2	October	1975	South Africa
Saayman 10	-25.88	28.02	C_4	-12.8	October	1975	South Africa
Saayman 101	-25.00	30.25	C_3	-24.8	October	1988	South Africa
Saayman 121	-25.44	30.09	C_3	-26.0	October	1988	South Africa
Saayman 122	-25.49	30.09	C_3	-25.7	October	1988	South Africa
Saayman 129*	-26.23	28.99	C_3	-26.4	October	1988	South Africa

Accession	Latitude	Longitude	Type	$\delta^{13}C$	Month	Year	Country
Saayman 134	-26.23	28.99	C_4	-10.6	October	1988	South Africa
Saayman 139	-26.05	30.39	C_4	-11.2	November	1988	South Africa
Saayman 143	-26.07	30.17	C_4	-12.2	November	1988	South Africa
Saayman 144	-26.07	30.17	C_4	-12.5	November	1988	South Africa
Saayman 16	-25.80	27.88	C_4	-11.4	October	1975	South Africa
Saayman 19	-26.04	27.63	C_3	-25.2	November	1975	South Africa
Saayman 20	-26.02	27.59	C_3	-25.5	November	1975	South Africa
Saayman 21	-25.94	27.90	C_4	-13.3	November	1975	South Africa
Saayman 24	-25.77	28.24	C_4	-10.1	December	1975	South Africa
Saayman 25A	-25.77	28.24	C_3	-25.5	December	1975	South Africa
Saayman 25B	-25.77	28.24	C_3	-25.5	December	1975	South Africa
Saayman 25C	-25.77	28.24	C_3	-25.5	December	1975	South Africa
Saayman 30	-25.93	27.89	C_4	-11.5	December	1987	South Africa
Saayman 31	-25.93	27.88	C_4	-11.7	December	1987	South Africa
Saayman 37	-25.93	27.89	C_4	-12.1	October	1988	South Africa
Saayman 8	-25.88	27.62	C_3	-27.1	October	1975	South Africa
Saldanha 13441	12.20	76.89	C_4	-11.0	May	1969	India
Schmitz 4195	-4.97	17.83	C_4	-10.6	November	1982	DRC
Schweinfurth 1049	12.97	36.16	C_4	-10.7	June	1864	Sudan
Scortechini & Tryon -	-28.67	151.95	C_4	-11.1	-	-	Australia
Shaunty 488*	-12.75	28.56	C_3-C_4	-20.7	December	1919	DRC
Siebert 3216*	-27.18	30.24	C_3	-26.9	December	2006	South Africa
Simon & Williamson 2062	-13.21	24.70	C_4	-10.9	December	1969	Zambia
Simon 932	-18.50	32.85	C_3	-25.4	October	1966	Zimbabwe
Simon et al 1168*	-8.83	31.38	C_4	-11.5	October	1967	Zambia
Simon et al 1689	-10.71	33.83	C_4	-11.8	February	1968	Malawi
Simon et al 978*	-13.62	29.40	C_3	-28.3	October	1967	Zambia
Smit 1438	-27.33	30.13	C_4	-13.5	December	1992	South Africa
Smitinand 5744	13.20	101.25	C_4	-10.9	April	1959	Thailand
Smook 3160	-25.25	28.75	C_4	-11.4	March	1981	South Africa
Smook 4928	-26.28	30.22	C_3	-26.2	January	1984	South Africa
Smook 702*	-29.78	29.50	C_3	-28.0	January	1980	South Africa
Soane 47	-19.85	32.80	C_3	-26.1	September	1959	Zimbabwe
Specht 1247	-12.30	133.07	C_4	-11.4	October	1948	Australia
Stalmans 1746	-24.18	30.35	C_4	-11.8	November	1988	South Africa
Stalmans 179	-24.17	30.22	C_3	-24.8	September	1984	South Africa
Stewert 186	5.43	37.82	C_4	-10.9	May	1956	Ethiopia
Stirton 5224	-28.93	30.84	C_4	-12.1	October	1975	South Africa
Stohr 689	-13.90	29.20	C_4	-11.4	December	1941	Zambia
Stowe 495*	-10.07	28.67	C_3-C_4	-20.8	December	1940	Zambia

Accession	Latitude	Longitude	Type	$\delta^{13}C$	Month	Year	Country
Strachey & Winterbottom 412*	29.70	79.72	C_4	-9.5	-	-	India
Strey 5615a	-28.13	31.87	C_4	-11.2	November	1964	South Africa
Strid 2534	-11.41	24.31	C_4	-12.6	November	1972	Zambia
Symoens 14118*	-8.95	27.38	C_3-C_4	-22.6	November	1971	DRC
Symon 5062	-16.83	137.00	C_4	-12.0	June	1967	Australia
T.T. 1606	25.58	91.63	C_4	-10.3	June	1850	India
Tanner 5076	-2.75	30.67	C_3	-26.3	July	1960	Tanzania
Testu 86	-17.72	36.28	C_4	-10.5	December	1904	Mozambique
Thomas 2236	1.75	34.72	C_4	-10.7	January	1937	Uganda
Tixier 14	17.98	102.63	C_4	-11.5	May	1950	Laos
Torre 4686*	-17.01	38.07	C_4	-9.9	November	1942	Mozambique
van Balgooy & Mamesah 6270	-6.83	134.33	C_4	-13.3	April	1993	Indonesia
van der Schijf 2035	-25.00	31.00	C_4	-10.6	December	1952	South Africa
van der Schijf 6152	-24.74	31.35	C_4	-11.4	January	1962	South Africa
van Jaarsveld 168	-25.51	30.96	C_4	-12.6	December	1974	South Africa
van Steenis 18485	-8.55	125.58	C_4	-10.9	January	1954	East Timor
Verboom 1147	-15.22	23.15	C_4	-11.5	December	1964	Zambia
Verboom 2529	-11.83	31.45	C_4	-11.8	October	1969	Zambia
Vesey-Fitzgerald 1990*	-8.64	31.50	C_3-C_4	-22.9	November	1958	Tanzania
Vesey-Fitzgerald 2079	-8.22	31.57	C_3	-25.1	December	1958	Tanzania
Vesey-Fitzgerald 4225*	-7.87	31.67	C_4	-13.1	November	1963	Tanzania
von Broembren 116*	-32.85	27.38	C_3	-26.5	-	-	South Africa
Ward 1585	-28.05	32.45	C_3	-26.0	October	1953	South Africa
Ward 1682	-28.13	31.87	C_4	-9.3	November	1953	South Africa
Werdermann & Oberdieck 1150	-30.90	28.52	C_3	-26.4	November	1958	South Africa
West 4435*	-18.22	32.75	C_3	-27.1	December	1962	Zimbabwe
Westfall 819	-24.48	27.57	C_4	-12.8	January	1980	South Africa
White 1470	-18.34	142.96	C_4	-11.3	February	1922	Australia
Wichian 334	15.30	98.40	C_4	-12.9	May	1946	Thailand
Wiehe 304	-16.07	35.13	C_4	-12.2	November	1949	Malawi
Wiehe 394	-15.17	35.30	C_4	-10.8	January	1950	Malawi
Wiehe 396B	-14.38	34.32	C_4	-10.8	January	1950	Malawi
Wiehe 403	-16.07	35.13	C_4	-10.3	November	1950	Malawi
Wilson 19	-14.47	132.27	C_4	-11.2	January	1965	Australia
Wilson 907	4.12	33.98	C_4	-10.4	April	1960	Uganda
Young & Young 420	-17.37	30.20	C_4	-10.9	December	1928	Zimbabwe
Zeyher 55*	-29.87	31.02	C_3	-26.4		1865	South Africa

Accessions used for anatomical study are marked with an asterisk (*).
Latitude and longitude coordinates are in decimal degrees

Appendix 5.3 Linear regression results for relationships between anatomy and environment.

	Segment length				Leaf thickness			
	DF	F	R^2	*p*	DF	F	R^2	*p*
Altitude								
Altitude	1, 95	19.87	-0.010	8.15E-01	1, 96	0.08	-0.010	7.82E-01
Temperature								
Annual mean temperature	**1, 95**	**8.56**	**0.073**	**4.30E-03**	**1, 96**	**7.08**	**0.059**	**9.13E-03**
Ground frost frequency	**1, 95**	**6.09**	**0.050**	**1.54E-02**	**1, 96**	**31.43**	**0.239**	**1.98E-07**
Max temp warmest month	**1, 95**	**11.55**	**0.099**	**9.94E-04**	1, 96	3.51	0.025	6.40E-02
Min temp coldest month	**1, 95**	**6.55**	**0.055**	**1.21E-02**	**1, 96**	**15.11**	**0.127**	**1.87E-04**
Mean temp wettest quarter	1, 95	2.83	0.019	9.59E-02	1, 96	0.23	-0.008	6.32E-01
Mean temp driest quarter	1, 95	**8.61**	**0.073**	**4.20E-03**	**1, 96**	**15.98**	**0.134**	**1.26E-04**
Mean temp warmest quarter	1, 95	**5.05**	**0.040**	**2.69E-02**	1, 96	1.63	0.006	2.04E-01
Mean temp coldest quarter	1, 95	**10.29**	**0.088**	**1.83E-03**	**1, 96**	**14.46**	**0.122**	**2.51E-04**
Temperature variability								
Mean diurnal range	1, 95	0.02	-0.010	8.83E-01	**1, 96**	**8.31**	**0.070**	**4.87E-03**
Isothermality	**1, 95**	**5.94**	**0.049**	**1.66E-02**	**1, 96**	**13.32**	**0.113**	**4.28E-04**
Temperature seasonality	**1, 95**	**10.92**	**0.094**	**1.34E-03**	**1, 96**	**36.89**	**0.270**	**2.50E-08**
Annual temperature range	1, 95	1.18	0.002	2.81E-01	**1, 96**	**16.37**	**0.137**	**1.05E-04**
Precipitation								
Annual precipitation	1, 95	3.34	0.024	7.09E-02	**1, 96**	**10.67**	**0.091**	**1.51E-03**
Wet day frequency	1, 95	1.88	0.009	1.74E-01	**1, 96**	**22.06**	**0.178**	**8.79E-06**
Precipitation wettest month	**1, 95**	**4.06**	**0.031**	**4.67E-02**	**1, 96**	**6.70**	**0.056**	**1.11E-02**
Precipitation driest month	**1, 95**	**20.63**	**0.170**	**1.64E-05**	1, 96	2.20	0.012	1.41E-01
Precipitation wettest quarter	**1, 95**	**5.55**	**0.045**	**2.06E-02**	**1, 96**	**8.80**	**0.074**	**3.79E-03**
Precipitation driest quarter	**1, 95**	**17.38**	**0.146**	**6.77E-05**	1, 96	2.00	0.010	1.61E-01
Precipitation warmest quarter	1, 95	0.26	-0.008	6.11E-01	1, 96	2.24	0.013	1.37E-01
Precipitation coldest quarter	1, 95	0.97	0.000	3.27E-01	**1, 96**	**4.16**	**0.032**	**4.40E-02**
Precipitation variability								
Precipitation seasonality	**1, 95**	**12.85**	**0.110**	**5.36E-04**	**1, 96**	**6.38**	**0.053**	**1.32E-02**
Light environment								
Cloud cover percent	**1, 95**	**8.73**	**0.074**	**3.95E-03**	**1, 96**	**54.57**	**0.356**	**5.49E-11**
Radiation	**1, 92**	**3.31**	**0.024**	**7.20E-02**	**1, 93**	**42.75**	**0.308**	**3.31E-09**
Sunshine percent	**1, 95**	**4.74**	**0.038**	**3.19E-02**	**1, 96**	**50.78**	**0.339**	**1.91E-10**
Evapotranspiration								
Relative humidity	**1, 95**	**4.74**	**0.038**	**3.19E-02**	**1, 96**	**50.78**	**0.339**	**1.91E-10**
Vapour pressure	1, 95	3.33	0.024	7.10E-02	**1, 96**	**6.08**	**0.050**	**1.55E-02**
PET	**1, 95**	**18.74**	**0.156**	**3.71E-05**	1, 96	3.74	0.027	5.60E-02
Aridity index	**1, 95**	**0.14**	**-0.009**	**7.07E-01**	**1, 96**	**6.81**	**0.057**	**1.05E-02**
Wind speed	**1, 95**	**2.05**	**0.011**	**1.56E-01**	1, 96	0.40	-0.006	5.30E-01
Fire								
Fire radiative power	1, 95	0.22	-0.008	6.44E-01	1, 96	1.67	0.007	1.99E-01
Maximum fire line intensity	1, 95	0.44	-0.006	5.10E-01	1, 96	2.35	0.014	1.29E-01
Maximum fire size	1, 95	3.54	0.026	6.32E-02	1, 96	1.17	0.002	2.81E-01
Fire season	1, 95	0.45	-0.006	5.05E-01	1, 96	0.53	-0.005	4.69E-01

	Number mesophyll cells				Tertiary veins			
	DF	F	R^2	*p*	DF	F	R^2	*p*
Altitude								
Altitude	**1, 82**	**6.82**	**0.066**	**1.07E-02**	1, 95	0.00	-0.011	9.91E-01
Temperature								
Annual mean temperature	**1, 82**	**23.57**	**0.214**	**5.69E-06**	1, 95	0.21	-0.008	6.44E-01
Ground frost frequency	**1, 81**	**7.12**	**0.069**	**9.21E-03**	**1, 95**	**6.84**	**0.057**	**1.04E-02**
Max temp warmest month	**1, 82**	**28.94**	**0.252**	**6.91E-07**	1, 95	0.16	-0.009	6.88E-01
Min temp coldest month	**1, 82**	**16.91**	**0.161**	**9.25E-05**	1, 95	2.26	0.013	1.36E-01
Mean temp wettest quarter	**1, 82**	**20.36**	**0.189**	**2.12E-05**	1, 95	0.40	-0.006	5.31E-01
Mean temp driest quarter	**1, 82**	**21.27**	**0.196**	**1.45E-05**	1, 95	2.07	0.011	1.53E-01
Mean temp warmest quarter	**1, 82**	**25.35**	**0.227**	**2.79E-06**	1, 95	0.18	-0.009	6.73E-01
Mean temp coldest quarter	**1, 82**	**18.77**	**0.176**	**4.15E-05**	1, 95	1.45	0.005	2.32E-01
Temperature variability								
Mean diurnal range	**1, 82**	**5.08**	**0.047**	**2.69E-02**	**1, 95**	**3.95**	**0.030**	**4.96E-02**
Isothermality	1, 82	0.96	-0.001	3.31E-01	1, 95	3.54	0.026	6.30E-02
Temperature seasonality	**1, 82**	**4.85**	**0.044**	**3.05E-02**	**1, 95**	**8.54**	**0.073**	**4.35E-03**
Annual temperature range	**1, 82**	**4.25**	**0.038**	**4.24E-02**	**1, 95**	**6.29**	**0.052**	**1.39E-02**
Precipitation								
Annual precipitation	**1, 82**	**11.07**	**0.108**	**1.31E-03**	1, 95	3.93	0.030	5.04E-02
Wet day frequency	1, 81	0.91	-0.001	3.43E-01	**1, 95**	**5.65**	**0.046**	**1.95E-02**
Precipitation wettest month	**1, 82**	**12.51**	**0.122**	**6.70E-04**	1, 95	1.32	0.003	2.54E-01
Precipitation driest month	**1, 82**	**6.31**	**0.060**	**1.40E-02**	1, 95	1.18	0.002	2.80E-01
Precipitation wettest quarter	**1, 82**	**12.90**	**0.125**	**5.59E-04**	1, 95	1.59	0.006	2.10E-01
Precipitation driest quarter	1, 82	2.43	0.017	1.23E-01	1, 95	0.84	-0.002	3.61E-01
Precipitation warmest quarter	1, 82	2.07	0.013	1.54E-01	1, 95	0.93	-0.001	3.38E-01
Precipitation coldest quarter	1, 82	2.93	0.023	9.08E-02	1, 95	0.06	-0.010	8.00E-01
Precipitation variability								
Precipitation seasonality	**1, 82**	**13.52**	**0.131**	**4.20E-04**	1, 95	0.01	-0.010	9.09E-01
Light environment								
Cloud cover percent	1, 82	2.96	0.023	8.94E-02	**1, 95**	**13.41**	**0.115**	**4.11E-04**
Radiation	1, 79	0.35	-0.008	5.57E-01	**1, 92**	**13.18**	**0.116**	**4.65E-04**
Sunshine percent	1, 81	0.91	-0.001	3.44E-01	**1, 95**	**5.57**	**0.045**	**2.03E-02**
Evapotranspiration								
Relative humidity	1, 81	0.91	-0.001	3.44E-01	**1, 95**	**5.57**	**0.045**	**2.03E-02**
Vapour pressure	**1, 82**	**26.32**	**0.234**	**1.91E-06**	1, 95	0.96	0.000	3.29E-01
PET	**1, 82**	**15.73**	**0.151**	**1.55E-04**	1, 95	0.19	-0.008	6.62E-01
Aridity index	**1, 82**	**5.73**	**0.054**	**1.90E-02**	**1, 95**	**5.20**	**0.042**	**2.48E-02**
Wind speed	1, 81	0.28	-0.009	5.96E-01	1, 95	2.70	0.017	1.03E-01
Fire								
Fire radiative power	**1, 82**	**11.06**	**0.108**	**1.32E-03**	1, 95	2.54	0.016	1.14E-01
Maximum fire line intensity	**1, 82**	**10.93**	**0.107**	**1.41E-03**	1, 95	2.47	0.015	1.20E-01
Maximum fire size	1, 82	3.49	0.029	6.53E-02	**1, 95**	**7.54**	**0.064**	**7.20E-03**
Fire season	1, 82	0.05	-0.012	8.21E-01	1, 95	0.28	-0.008	6.00E-01

	Quaternary veins				Quinary veins			
	DF	F	R^2	p	DF	F	R^2	p
Altitude								
Altitude	**1, 95**	**12.00**	**0.103**	**8.00E-04**	**1, 95**	**18.28**	**0.153**	**4.55E-05**
Temperature								
Annual mean temperature	**1, 95**	**23.84**	**0.192**	**4.23E-06**	**1, 95**	**24.28**	**0.195**	**3.51E-06**
Ground frost frequency	1, 95	3.17	0.022	7.83E-02	1, 95	0.36	-0.007	5.51E-01
Max temp warmest month	**1, 95**	**30.98**	**0.238**	**2.41E-07**	**1, 95**	**30.11**	**0.233**	**3.38E-07**
Min temp coldest month	**1, 95**	**14.94**	**0.127**	**2.03E-04**	**1, 95**	**11.59**	**0.099**	**9.73E-04**
Mean temp wettest quarter	**1, 95**	**26.17**	**0.208**	**1.62E-06**	**1, 95**	**38.61**	**0.282**	**1.37E-08**
Mean temp driest quarter	**1, 95**	**18.63**	**0.155**	**3.89E-05**	**1, 95**	**14.99**	**0.127**	**1.98E-04**
Mean temp warmest quarter	**1, 95**	**33.04**	**0.250**	**1.09E-07**	**1, 95**	**37.74**	**0.277**	**1.88E-08**
Mean temp coldest quarter	**1, 95**	**15.65**	**0.132**	**1.47E-04**	**1, 95**	**13.76**	**0.117**	**3.50E-04**
Temperature variability								
Mean diurnal range	**1, 95**	**4.27**	**0.033**	**4.16E-02**	1, 95	1.46	0.005	2.29E-01
Isothermality	1, 95	0.01	-0.010	9.04E-01	1, 95	0.35	-0.007	5.53E-01
Temperature seasonality	1, 95	0.96	0.000	3.30E-01	1, 95	0.09	-0.010	7.62E-01
Annual temperature range	1, 95	2.10	0.011	1.51E-01	1, 95	0.87	-0.001	3.52E-01
Precipitation								
Annual precipitation	**1, 95**	**8.17**	**0.070**	**5.23E-03**	**1, 95**	**6.15**	**0.051**	**1.49E-02**
Wet day frequency	1, 95	0.82	-0.002	3.67E-01	1, 95	0.12	-0.009	7.33E-01
Precipitation wettest month	**1, 95**	**11.02**	**0.094**	**1.28E-03**	**1, 95**	**10.81**	**0.093**	**1.41E-03**
Precipitation driest month	1, 95	3.83	0.029	5.32E-02	1, 95	2.53	0.016	1.15E-01
Precipitation wettest quarter	**1, 95**	**10.02**	**0.086**	**2.08E-03**	**1, 95**	**8.83**	**0.075**	**3.75E-03**
Precipitation driest quarter	1, 95	1.06	0.001	3.07E-01	1, 95	0.74	-0.003	3.93E-01
Precipitation warmest quarter	1, 95	2.62	0.017	1.09E-01	**1, 95**	**8.13**	**0.069**	**5.35E-03**
Precipitation coldest quarter	**1, 95**	**8.39**	**0.072**	**4.67E-03**	1, 95	2.97	0.020	8.81E-02
Precipitation variability								
Precipitation seasonality	**1, 95**	**7.36**	**0.062**	**7.92E-03**	**1, 95**	**4.48**	**0.035**	**3.68E-02**
Light environment								
Cloud cover percent	1, 95	0.76	-0.003	3.86E-01	1, 95	0.63	-0.004	4.28E-01
Radiation	1, 92	1.49	0.005	2.25E-01	**1, 92**	**8.12**	**0.071**	**5.40E-03**
Sunshine percent	1, 95	3.40	0.024	6.83E-02	1, 95	0.40	-0.006	5.27E-01
Evapotranspiration								
Relative humidity	1, 95	3.40	0.024	6.83E-02	1, 95	0.40	-0.006	5.27E-01
Vapour pressure	**1, 95**	**25.78**	**0.205**	**1.90E-06**	**1, 95**	**25.84**	**0.206**	**1.86E-06**
PET	**1, 95**	**8.73**	**0.075**	**3.94E-03**	**1, 95**	**11.28**	**0.097**	**1.13E-03**
Aridity index	**1, 95**	**4.13**	**0.032**	**4.51E-02**	1, 95	1.98	0.010	1.63E-01
Wind speed	1, 95	1.40	0.004	2.40E-01	1, 95	0.49	-0.005	4.84E-01
Fire								
Fire radiative power	**1, 95**	**9.54**	**0.082**	**2.63E-03**	1, 95	3.89	0.029	5.14E-02
Maximum fire line intensity	**1, 95**	**8.89**	**0.076**	**3.64E-03**	1, 95	3.36	0.024	7.00E-02
Maximum fire size	1, 95	0.60	-0.004	4.39E-01	1, 95	3.88	0.029	5.19E-02
Fire season	1, 95	2.35	0.014	1.28E-01	1, 95	0.05	-0.010	8.17E-01

	Vein area				Vein width:height			
	DF	F	R^2	*p*	DF	F	R^2	*p*
Altitude								
Altitude	1, 97	1.17	0.002	2.83E-01	1, 97	0.58	-0.004	4.48E-01
Temperature								
Annual mean temperature	1, 97	1.87	0.009	1.75E-01	1, 97	0.20	-0.008	6.55E-01
Ground frost frequency	**1, 97**	**11.22**	**0.094**	**1.15E-03**	1, 97	3.37	0.024	6.96E-02
Max temp warmest month	1, 97	0.92	-0.001	3.39E-01	1, 97	0.48	-0.005	4.89E-01
Min temp coldest month	**1, 97**	**4.12**	**0.031**	**4.51E-02**	1, 97	0.75	-0.003	3.89E-01
Mean temp wettest quarter	1, 97	0.15	-0.009	7.03E-01	1, 97	0.24	-0.008	6.26E-01
Mean temp driest quarter	**1, 97**	**4.92**	**0.038**	**2.88E-02**	1, 97	0.81	-0.002	3.72E-01
Mean temp warmest quarter	1, 97	0.12	-0.009	7.26E-01	1, 97	0.01	-0.010	9.07E-01
Mean temp coldest quarter	**1, 97**	**4.53**	**0.035**	**3.58E-02**	1, 97	0.95	-0.001	3.33E-01
Temperature variability								
Mean diurnal range	1, 97	1.00	0.000	3.20E-01	1, 97	0.44	-0.006	5.09E-01
Isothermality	**1, 97**	**10.28**	**0.086**	**1.82E-03**	1, 97	0.00	-0.010	9.84E-01
Temperature seasonality	**1, 97**	**13.68**	**0.115**	**3.59E-04**	**1, 97**	**4.10**	**0.031**	**4.57E-02**
Annual temperature range	**1, 97**	**4.60**	**0.035**	**3.45E-02**	1, 97	0.49	-0.005	4.86E-01
Precipitation								
Annual precipitation	1, 97	2.69	0.017	1.04E-01	1, 97	0.43	-0.006	5.16E-01
Wet day frequency	**1, 97**	**7.04**	**0.058**	**9.30E-03**	**1, 97**	**5.28**	**0.042**	**2.38E-02**
Precipitation wettest month	1, 97	2.41	0.014	1.24E-01	1, 97	0.13	-0.009	7.22E-01
Precipitation driest month	**1, 97**	**5.53**	**0.044**	**2.07E-02**	1, 97	0.62	-0.004	4.32E-01
Precipitation wettest quarter	1, 97	3.45	0.024	6.63E-02	1, 97	0.40	-0.006	5.31E-01
Precipitation driest quarter	**1, 97**	**6.02**	**0.049**	**1.59E-02**	1, 97	0.63	-0.004	4.30E-01
Precipitation warmest quarter	1, 97	2.38	0.014	1.26E-01	**1, 97**	**3.96**	**0.029**	**4.94E-02**
Precipitation coldest quarter	1, 97	1.66	0.007	2.01E-01	1, 97	0.04	-0.010	8.51E-01
Precipitation variability								
Precipitation seasonality	**1, 97**	**7.28**	**0.060**	**8.23E-03**	1, 97	1.31	0.003	2.56E-01
Light environment								
Cloud cover percent	**1, 97**	**15.24**	**0.127**	**1.75E-04**	**1, 97**	**12.24**	**0.103**	**7.07E-04**
Radiation	**1, 94**	**9.90**	**0.086**	**2.21E-03**	**1, 94**	**16.11**	**0.137**	**1.20E-04**
Sunshine percent	**1, 97**	**12.30**	**0.103**	**6.87E-04**	**1, 97**	**7.63**	**0.063**	**6.87E-03**
Evapotranspiration								
Relative humidity	**1, 97**	**12.30**	**0.103**	**6.87E-04**	**1, 97**	**7.63**	**0.063**	**6.87E-03**
Vapour pressure	1, 97	0.92	-0.001	3.39E-01	1, 97	0.06	-0.010	8.05E-01
PET	**1, 97**	**4.42**	**0.034**	**3.82E-02**	1, 97	0.65	-0.004	4.24E-01
Aridity index	1, 97	0.84	-0.002	3.62E-01	1, 97	0.30	-0.007	5.87E-01
Wind speed	1, 97	0.00	-0.010	9.87E-01	1, 97	0.05	-0.010	8.24E-01
Fire								
Fire radiative power	1, 97	0.06	-0.010	8.06E-01	1, 97	1.02	0.000	3.15E-01
Maximum fire line intensity	1, 97	0.00	-0.010	9.85E-01	1, 97	1.71	0.007	1.94E-01
Maximum fire size	1, 97	0.06	-0.010	8.02E-01	1, 97	1.42	0.004	2.36E-01
Fire season	1, 97	0.20	-0.008	6.56E-01	**1, 97**	**5.90**	**0.048**	**1.70E-02**

	IBS width				OBS width			
	DF	F	R^2	*p*	DF	F	R^2	*p*
Altitude								
Altitude	**1, 97**	**4.34**	**0.033**	**4.00E-02**	1, 97	0.14	-0.009	7.09E-01
Temperature								
Annual mean temperature	**1, 97**	**11.63**	**0.098**	**9.47E-04**	**1, 97**	**24.05**	**0.190**	**3.77E-06**
Ground frost frequency	1, 97	0.41	-0.006	5.24E-01	**1, 97**	**16.12**	**0.134**	**1.17E-04**
Max temp warmest month	**1, 97**	**14.75**	**0.123**	**2.19E-04**	**1, 97**	**31.74**	**0.239**	**1.73E-07**
Min temp coldest month	**1, 97**	**5.37**	**0.043**	**2.26E-02**	**1, 97**	**19.67**	**0.160**	**2.43E-05**
Mean temp wettest quarter	**1, 97**	**14.59**	**0.122**	**2.36E-04**	**1, 97**	**11.95**	**0.101**	**8.12E-04**
Mean temp driest quarter	**1, 97**	**7.25**	**0.060**	**8.36E-03**	**1, 97**	**25.95**	**0.203**	**1.73E-06**
Mean temp warmest quarter	**1, 97**	**13.20**	**0.111**	**4.50E-04**	**1, 97**	**16.46**	**0.136**	**1.01E-04**
Mean temp coldest quarter	**1, 97**	**8.29**	**0.069**	**4.90E-03**	**1, 97**	**25.56**	**0.200**	**2.02E-06**
Temperature variability								
Mean diurnal range	1, 97	0.04	-0.010	8.51E-01	1, 97	2.15	0.012	1.46E-01
Isothermality	1, 97	1.10	0.001	2.96E-01	**1, 97**	**6.55**	**0.054**	**1.21E-02**
Temperature seasonality	1, 97	1.26	0.003	2.65E-01	**1, 97**	**19.14**	**0.156**	**3.06E-05**
Annual temperature range	1, 97	0.25	-0.008	6.19E-01	**1, 97**	**4.36**	**0.033**	**3.94E-02**
Precipitation								
Annual precipitation	1, 97	0.39	-0.006	5.32E-01	**1, 97**	**10.85**	**0.091**	**1.38E-03**
Wet day frequency	1, 97	0.38	-0.006	5.40E-01	1, 97	3.24	0.022	7.49E-02
Precipitation wettest month	1, 97	1.15	0.002	2.87E-01	**1, 97**	**14.25**	**0.119**	**2.76E-04**
Precipitation driest month	1, 97	3.87	0.028	5.21E-02	**1, 97**	**43.48**	**0.302**	**2.25E-09**
Precipitation wettest quarter	1, 97	1.03	0.000	3.13E-01	**1, 97**	**16.07**	**0.133**	**1.20E-04**
Precipitation driest quarter	1, 97	3.51	0.025	6.39E-02	**1, 97**	**27.57**	**0.213**	**8.94E-07**
Precipitation warmest quarter	1, 97	0.09	-0.009	7.64E-01	1, 97	0.22	-0.008	6.37E-01
Precipitation coldest quarter	1, 97	0.00	-0.010	9.72E-01	1, 97	0.33	-0.007	5.65E-01
Precipitation variability								
Precipitation seasonality	**1, 97**	**4.31**	**0.033**	**4.05E-02**	**1, 97**	**39.01**	**0.280**	**1.12E-08**
Light environment								
Cloud cover percent	1, 97	0.02	-0.010	8.95E-01	**1, 97**	**21.07**	**0.170**	**1.33E-05**
Radiation	1, 94	2.43	0.015	1.22E-01	**1, 94**	**4.90**	**0.039**	**2.93E-02**
Sunshine percent	1, 97	2.97	0.020	8.81E-02	**1, 97**	**6.28**	**0.051**	**1.39E-02**
Evapotranspiration								
Relative humidity	1, 97	2.97	0.020	8.81E-02	**1, 97**	**6.28**	**0.051**	**1.39E-02**
Vapour pressure	**1, 97**	**8.32**	**0.069**	**4.84E-03**	**1, 97**	**16.93**	**0.140**	**8.14E-05**
PET	**1, 97**	**13.24**	**0.111**	**4.41E-04**	**1, 97**	**32.16**	**0.241**	**1.46E-07**
Aridity index	1, 97	0.21	-0.008	6.49E-01	1, 97	2.73	0.017	1.02E-01
Wind speed	1, 97	0.20	-0.008	6.52E-01	1, 97	0.66	-0.003	4.17E-01
Fire								
Fire radiative power	1, 97	2.39	0.014	1.25E-01	**1, 97**	**14.87**	**0.124**	**2.07E-04**
Maximum fire line intensity	1, 97	2.00	0.010	1.61E-01	**1, 97**	**14.83**	**0.124**	**2.11E-04**
Maximum fire size	1, 97	1.45	0.005	2.31E-01	1, 97	3.89	0.029	5.13E-02
Fire season	1, 97	0.05	-0.010	8.18E-01	**1, 97**	**5.96**	**0.048**	**1.64E-02**

CHAPTER 6

GENERAL DISCUSSION

Dissertation review and key findings

The primary aim of this dissertation was to elucidate the small-scale processes that led to the emergence of the complex C_4 photosynthesis trait. Using the unique grass *Alloteropsis semialata* as a study system, this dissertation set out to highlight the adaptive significance of intermediate states and the evolutionary paths to complex traits in general. First, a literature review and novel data analysis highlighted the remarkable variation in leaf anatomies across a variety of grass taxa of differing photosynthetic types (Chapter 2). It concluded that most components of the C_4 phenotype are found in non-C_4 species and that different combinations of these components can produce C_4-suitable phenotypes. Moreover, the presence of C_4-suitable anatomies in non-C_4 plant lineages could facilitate the evolution of C_4 photosynthesis. This is because the gap between typical C_3 and C_4 anatomies would be narrower in these lineages and thus, relatively few changes would be required to achieve C_4 functionality. This finding helps to explain the plethora of independent C_4 origins and highlights the importance of intermediary phenotypes for C_4 evolution.

The dissertation then shifted its focus to intraspecific photosynthetic variation. A study examining the isotopic, anatomical, and physiological phenotypes across a diverse set of *A. semialata* populations revealed the occurrence of C_3-C_4 intermediate phenotypes in this species, in addition to the well-characterized C_3 and C_4 photosynthetic types (Chapter 3). This finding solidifies *A. semialata* as one of the most photosynthetically diverse taxa on the planet. Moreover, links between anatomical and physiological phenotypes expressed across this C_3/C_4 continuum were drawn to show how structural and functional changes are intimately associated. These results draw attention to the role that structural reorganization plays in the establishment of photosynthetic types and show for the first time that the physiological and anatomical modifications associated with C_4 evolution may occur concurrently.

The evolutionary history of *A. semialata* and implications of photosynthetic diversification on the dispersal history of this species were investigated next (Chapter 4). These analyses revealed three key findings. First, C_3, C_3-C_4, and C_4 *A. semialata* accessions are not intermingled throughout the phylogeny, but instead

form discrete sub-clades. This suggests that these three photosynthetic types are reproductively isolated and that the C_3-C_4 phenotypes are evolutionary intermediates, rather than resulting from hybrids of C_3 and C_4 parents. Second, the phylogenetic analysis pointed to a C_3-C_4 intermediate MRCA of *A. semialata*. This implies that, from this intermediate state, some lineages evolved C_4 photosynthesis, while another lineage reverted back to a C_3 state. Finally, analyses of the dispersal histories across geographic and environmental space revealed that C_3, C_3-C_4, and C_4 *A. semialata* dispersed in very different patterns. The C_4 individuals dispersed broadly across three continents and into an expanded range of environments, while in the same amount of time, C_3 and C_3-C_4 intermediate individuals remained confined to a limited geographic area and restricted ecological conditions. These results suggest that the evolution of C_4 photosynthesis did not directly alter ecological preference, but instead widened the niche, which facilitated dispersal across environments and therefore long distances. Findings also imply that C_3-C_4 intermediacy was lost in the C_3 lineage as it moved into the colder, more seasonal southern African climate, while those lineages that remained in central Africa maintained the intermediate phenotype.

The final study of the dissertation focused on the intraspecific anatomical continua present in *A. semialata* using a large sample of accessions that represented the physiological, ecological, and geographical distribution of the species. This work revealed the importance of intermediate phenotypes to bridge the selective landscape between C_3 and C_4 phenotypes. Specifically, putative C_3-C_4 intermediate plants narrow the anatomical space between C_3 and C_4 phenotypes for bundle sheath sizes, minimum number of mesophyll cells between veins, and the total number of veins per segment. The findings of Chapters 2 and 3 suggest these anatomical traits are important to C_3/C_4 transitions and are strongly correlated with physiological advantages. Thus, the anatomy of C_3-C_4 intermediates facilitated the evolution of C_4 photosynthesis in some lineages, but may have also simplified the reversion to the C_3 photosynthetic type in another lineage. These anatomical continua were put into an ecological context to show where, and in which environments, the anatomical shifts occurred that may have driven C_4 evolution, or the reversion back to C_3 photosynthesis. These results suggest that the evolution of C_4 photosynthesis in *A. semialata* may have been confined to a narrow set of

environmental conditions for part of this transition and that confinement may have shaped where and how easily C_4 photosynthesis evolved in this species.

Physiological and anatomical diversification

Physiological and anatomical diversifications have been the topic of innumerable studies over the years, across all kingdoms of life (*e.g.*, Hofmann, 1989; Alfaro *et al*., 2005; Hearn, 2009; Middelboe *et al*., 2009). In plants, long adaptive evolutionary transformations, under the influence of changing environmental conditions, often lead to the accumulation of individual traits, and sometimes the formation of complex traits, such as C_4 photosynthesis. Most plants carry out photosynthesis via one of three primary pathways; C_3, C_4, and CAM, each of which advantage the metabolic pathway under different environmental conditions. However, a few remarkable species use multiple photosynthetic types or have the capacity to change pathways, in response to environmental stimulus (*e.g.,* in the Alismataceae, Cyperaceae, Hydrocharitaceae, and Poaceae families; Ueno *et al*., 1988b; Keeley, 1998; Bowes *et al*., 2002; Boykin *et al*., 2008). This dissertation showed that *A. semialata* has the unique characteristic of presenting not only the two primary C_3 and C_4 photosynthetic types, but also an intermediate one. This physiological diversity makes *A. semialata* a brilliant system to study the diversification of photosynthetic type.

Although there are three extant photosynthetic types in *A. semialata,* the work presented in Chapter 4 suggests that the ancestral state of this photosynthetic variation was likely a C_3-C_4 intermediate. It remains unknown why a seemingly transitory phenotype would persist for millions of years, however, it may be linked to impressive levels of phenotypic plasticity in these plants (Box 6.1). Based on the structure-function relationships revealed in Chapter 3, it may be inferred that the physiological diversifications presented in Chapter 4 were accompanied by several anatomical modifications. Specifically, inner bundle sheaths enlarged, additional quaternary and quinary vein orders were inserted, and mesophyll cell numbers decreased between veins in the C_4 lineages. In the C_3 lineage, secondary veins grew further apart, mesophyll cells more numerous, inner bundle sheaths shrank, and outer bundle sheath cells enlarged. The C_3-C_4 intermediate lineage, which

remained in central Africa, also evolved from the MRCA state, as their leaves grew thinner with more tertiary veins and these veins were smaller and rounder in shape. Interestingly, the traits associated with the C_4 lineage were adopted at different time points in the development of the C_4 type in *A. semialata*. For instance, the L02 phenotype (*i.e.,* clade F in Chapter 4) had more minor veins and fewer mesophyll cells than C_3-C_4 intermediate plants, but maintained C_3-C_4 intermediate-like bundle sheath cell characteristics (see Figure 4.13). This suggests that vein and mesophyll patterning were likely already in place when C_4 photosynthesis evolved over two million years ago, while changes in bundle sheath cell sizes occurred after C_4 evolution, perhaps to optimize the C_4 system. Although these patterns of diversification can only be inferred from our current dataset, they can be proved with additional research using ancestral state reconstructions of both anatomical and physiological phenotypes of *A. semialata*. Unlike the phylogeny presented in Chapter 4, these reconstructions would need to be based on nuclear, rather than chloroplast markers. This is because the nuclear genome reflects both parents, while the chloroplast genome lacks paternal influence. The chloroplast genome is sufficient to investigate dispersal patterns, however, the influence of both parents is required to understand anatomical and physiological phenotypes. Ancestral state reconstructions of individual anatomical and physiological traits on a comprehensive nuclear phylogeny of *A. semialata* accessions sampled from across the photosynthetic, ecological, and geographic ranges of this species would reveal when, where, and in which environments, each trait emerged during the course of photosynthetic diversification in this species. This information would not only benefit our understanding of the minutia of C_4 evolutionary trajectories, but also help us understand why reversions to the C_3 state occur.

Box 6.1 Phenotypic plasticity for traits linked to photosynthesis in *Alloteropsis semialata*

Phenotypic plasticity is the ability of a single genotype to express different phenotypes in response to changing environmental conditions (West-Eberhard, 1989). Not only will this ability influence the environments in which a genotype can successfully exist, but it will also shape its evolutionary trajectory (Bradshaw, 1965; Sultan, 2000). Plasticity may have contrasting consequences, promoting evolution if it allows persistence in novel environments and exposure to new selection pressures, or impeding evolutionary change by diluting the effects of natural selection (*i.e.*, if a required phenotype is produced plastically, rather than through genetic change, then selection will be weak on that trait response; discussed in Chapter 2). Therefore, the presence of sufficient phenotypic plasticity for photosynthetic traits in C_3-C_4 *A. semialata* could explain the persistence of this intermediate phenotype (discussed in Chapter 3 and 4). Specifically, if these intermediate accessions can plastically express C_4-like anatomical and physiological phenotypes, then selection will be weak to modify these traits and fully adopt the C_4 syndrome.

Unfortunately, an experiment to thoroughly investigate the degree and direction phenotypic plasticity in *A. semialata* was outside the scope of this dissertation. However, preliminary research was undertaken on the subject. Repeated measures of $\delta^{13}C$, anatomy, and physiology on the same genotypes over time presented intriguing results that point to phenotypic plasticity. The number of veins per segment, the minimum number of mesophyll cells between veins, and the inner and outer bundle sheath cell sizes (IBS and OBS, respectively) were measured on young, fully expanded leaves from the same genotype grown in field (preserved in ethanol) and controlled environment growth chambers (CEGC; freshly harvested). While leaf physiology measurements were not collected in the field, CO_2 compensation point data were collected four and seven months after transplanting into CEGC. $\delta^{13}C$ was quantified on these "field", "four", and "seven" month leaves. Anatomy, physiology, and $\delta^{13}C$ methods are described in Chapter 3. The CEGC were set to 25°C day/20°C night temperatures, 500 μmol m^{-2} s^{-1} PPFD, 60% relative humidity, and ambient CO_2 concentrations over 14 hour day/10 hour night cycles. It should be noted that the C_3-C_4 intermediate plants were originally transplanted in John Innes No. 3 compost (John Innes Manufacturers Association, Reading, England) then transferred to No. 2 compost soon afterwards. All other *A. semialata* plants were grown in No 2 compost. No. 3 compost is higher in nutrients than No. 2 compost. Thus, 4 and 7 months data collections occurred in different soil environments in these intermediate plants. Unfortunately, environmental conditions were not quantified in detail in the field, making it difficult to identify which environments might be driving a plastic response. For reference, however, an image of the typical C_3-C_4 field environment can be seen in Figure 3.19).

A. semialata plants expressed different anatomy, physiology, and $\delta^{13}C$ values in response to environmental variation and C_3, C_3-C_4 intermediate, and C_4 plants differed in the degree and direction of these changes (Figure B6.1). Anatomically, C_4 plants became more C_4-like while the C_3 plants became more C_3-like when moved from the field to CEGC (Figure B6.1a). In general, C_4 accessions increased the number of veins per segment and decreased the number of mesophyll cells separating these veins, while C_3 plants increased their mesophyll cell numbers and

outer bundle sheath cells (Figure B6.1a). When transplanted into CEGC, the C_3-C_4 intermediates became more C_4-like overall as they generally increased the number of veins per segment, decreased mesophyll cells between veins, and increased IBS size. At that point (4 months post-transplant), the CCP of these plants was approximately 20 µmol CO_2, consistent with other C_3-C_4 intermediate plants, but then shifted to values typical of C_4 plants (*e.g.*, 6 µmol CO_2) after another three months in the CEGC (Figure B6.1b). $\delta^{13}C$ values also changed between the three time points in C_3-C_4 intermediate plants, but the response differed considerably by genotype (Figure B6.1c). These values remained within the C_3/C_2 range throughout and never approached C_4 $\delta^{13}C$ levels (see Table 3.1 for reference). $\delta^{13}C$ values are difficult to compare between field and CEGC plants because the difference in air quality due to anthropogenic release of fossil CO_2, (discussed in Chapter 3), which may account for the decrease in $\delta^{13}C$ values between field and four-month plants. However, any additional decrease or increase in $\delta^{13}C$ values between four- and seven- month leaves may imply stronger C_2 activity, or activation of weak C_4-cycle activity, respectively (von Caemmerer *et al.*, 2014). These initial results are intriguing and suggest that there may be considerable plasticity for photosynthetic traits in *A. semialata*. Additional work should be done to explore these trends more thoroughly and to investigate the role that plasticity may play in maintaining photosynthetic diversity in this species.

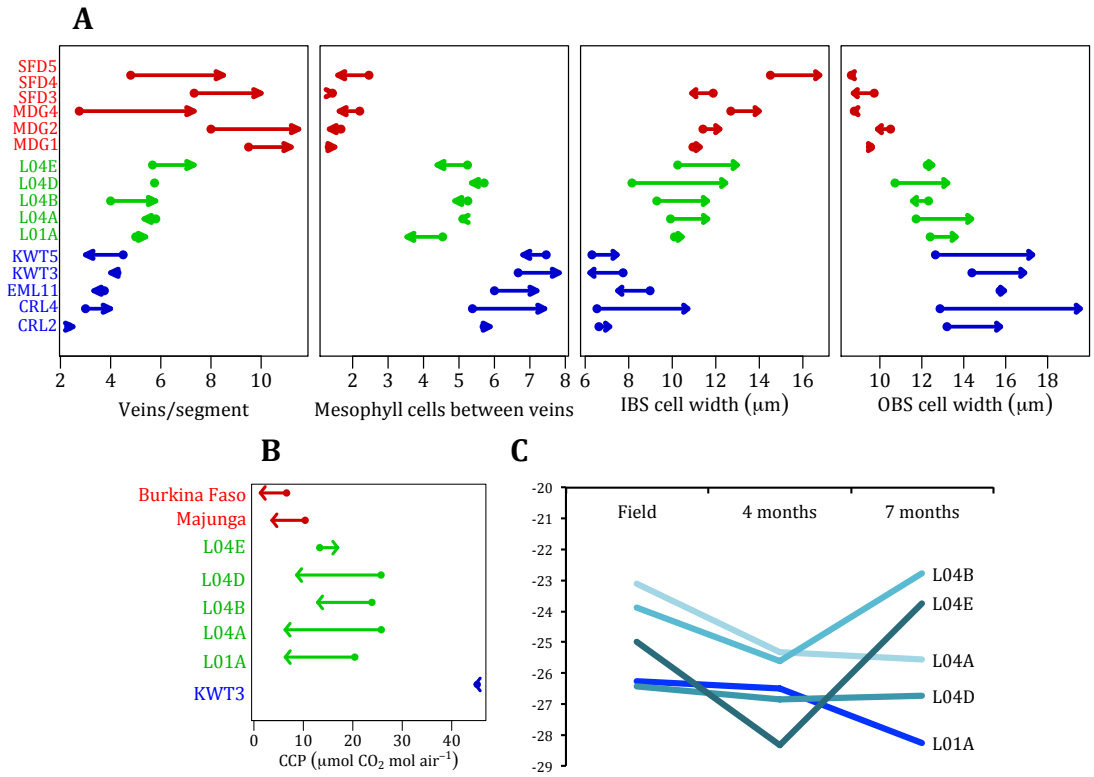


Figure B6.1 Phenotypic shifts in *Alloteropsis semialata*. Panel A shows the degree and direction of change for anatomical traits in single genotypes growing in the wild/field (dot) and four months after transplanting into CEGC (arrow head) in C_3 (blue), C_3-C_4 intermediate (green), and C_4 (red) accessions. Panel B shows the degree and direction of change in CO_2 compensation point (CCP) between single genotypes growing in CEGC, four (dot) and seven (arrow head) months after transplanting in C_3 (blue), C_3-C_4 intermediate (green), and C_4 (red) accessions (C_3 and C_4 genotypes included for reference. The number of post-transplant months are only relevant for C_3-C_4 intermediate accession, as C_3 and C_4 plants were transplanted or grown from seed nearly two years earlier). Panel C shows $\delta^{13}C$ values in C_3-C_4 intermediate accessions measured on leaves grown in the field and after four and seven months post-transplant into CEGC.

Reproductive isolation in *Alloteropsis semialata*

Reproductive isolation is the consequence of geographical, behavioural, or physiological barriers that prevent populations of closely related taxa from successfully producing offspring (Mayr, 1959). It is the by-product of adaptation to different selective conditions and can evolve extremely quickly in some organisms (Gavrilets, 2000; Hendry *et al.,* 2000). It is recognised to be an important component for the evolution of complex traits (Orr and Smith, 1998). Once a population acquires a new adaptive trait, it must remain isolated from populations lacking these modifications to avoid the new mutation being bred out of the population. The work presented in Chapter 4 implies a lack of gene flow between C_4 and non-C_4 lineages of *A. semialata*, which suggests they should be reproductively isolated. C_3 and C_4 *A. semialata* have overlapping distributions along a corridor of south-eastern Africa and C_3 and C_4 populations have been found in close proximity, in similar ecological conditions (Chapter 3 – 5; Ellis, 1981; Liebenberg and Fossey, 2001; personal observation). It is plausible then that plants of these two photosynthetic types could cross-pollinate, if reproductive isolating mechanisms were not in place and, as discussed in Chapter 3, the C_3-C_4 intermediate phenotype could indeed result from these C_3/C_4 crossing events (Oakley *et al*., 2014). Considering that the intermediate C_3-C_4 phenotype has only been found in central Africa and never in South Africa (Chapter 3-5; Ellis, 1981; Hattersley and Watson, 1992), it could be assumed that reproductive isolating mechanisms are acting to keep C_3 and C_4 populations distinct in southern Africa but not central Africa. However, the phylogeny presented in Chapter 4 strongly suggests that gene flow is not occurring between C_3 and C_4 lineages there, and they became reproductively isolated soon after their diversification. Thus, the C_3-C_4 intermediate phenotype does not seem to result from C_3/C_4 crosses. With that said, the mechanisms that dissuade these crosses have not been identified.

Pre-zygotic mechanisms, such as ecological, behavioural, or temporal isolation, are unlikely to be isolating *A. semialata* populations. Chapter 4 shows that the physiological diversification occurred while populations inhabited overlapping ranges and environmental conditions, which implies that ecology is not a pre-zygotic isolating mechanism. Pollinator dependency is also unlikely to be an isolating factor as *A. semialata* is predominantly wind pollinated and it also has

hermaphroditic flowers (Connor, 1979), so it may self-fertilize in the absence of wind delivered pollen (Friedman and Barrett, 2009). Chapter 5 found that C_3 and C_4 *A. semialata* plants flower at similar times in both southern and central Africa, suggesting that flowering phenology is not a temporal pre-zygotic reproductive isolating mechanism in this species either. Thus, the C_3 and C_4 *A. semialata* lineages are probably not pre-zygotically isolated.

Post-zygotic reproductive isolating mechanisms are more likely to be maintaining distinct photosynthetic types in *A. semialata*. Chapter 3 found that C_3 and C_4 populations in southern Africa have different ploidy levels. Thus, ploidy level may be a post-zygotic isolating mechanism that maintains the two photosynthetic types there. However, little is know about this species' ploidy levels in central Africa, the region of highest photosynthetic diversification. Thus, if ploidy levels prove compatible there, then other post-zygotic mechanisms may be in place to cause gametic incompatibility, zygote mortality, hybrid inviability, hybrid sterility, or hybrid breakdown (reviewed in Widmer *et al*., 2009). At this point, it is unknown to what degree the offspring of crosses between C_3 and C_4 *A. semialata* are viable, so future studies should be done to clarify this. Thus, while this dissertation has determined that the C_3 and C_4 subspecies are reproductively isolated, it has not identified exactly which mechanisms are doing so.

How many species encompass *Alloteropsis semialata*?

Alloteropsis semialata is currently considered one species with two subspecies, which are consistent with C_3 and C_4 photosynthetic types. Despite clear anatomical, physiological, and morphological differences, these taxa were considered the same species based on similarities in floral structure (Ellis, 1974a, b; Brown, 1975; Ellis, 1981; Frean *et al*., 1983a, b; Gibbs Russell, 1983). The biological species concept defines a species of sexual organism as those that are reproductively isolated (Mayr, 1970). Thus, the findings of this dissertation may lend support to split the C_3 and C_4 subspecies of *A. semialata* into two distinct species. Moreover, the work presented in Chapter 3 demonstrates that, while anatomical and physiological continua exist within this species that could support the maintenance of a single taxon, this remarkable variation still clusters into

discrete groups consistent with C_3, C_3-C_4 intermediate, and C_4 types. Thus, *A. semialata* may actually encompass three independent species. However, using a much larger dataset in Chapter 5, we see that there is overlap in anatomical phenotypes within the intermediate space between C_3 and C_4 clusters (Figure 5.11). This overlap would make it difficult to definitively distinguish the C_3-C_4 species, from either the C_3 or C_4 one. More work is needed to understand how many species encompass *A. semialata*. Current genetic research widely challenges the definition of species and could help answer this question. The advent of massive sequencing techniques has revolutionized the identification and categorization of species (*e.g.*, Schloss and Handelsman, 2005; Pons *et al.*, 2006; Ficetola *et al.*, 2008). Complete genome constructions of *A. semialata* accessions using different photosynthetic types would allow for quantification of genetic variability between the three photosynthetic types and help to confirm, or not, if clustering of C_3, C_3-C_4 intermediate, and C_4 accessions could indeed account for genetically different species.

Alloteropsis semialata as a model system for crop breeding

The plethora of intraspecific photosynthetic variation found in *A. semialata* is exceptional and presents a unique opportunity to study the transitional steps between C_3 and C_4 photosynthesis. In particular, understanding the natural anatomical variation that underlies the physiologies in this species may be extrapolated to support the industries investigating crop improvement. The C_3-C_4 intermediate phenotype of *A. semialata* may be particularly useful. More than 50 plant species have been identified to use a C_2 cycle, however, most of these species belong to eudicot lineages such as the Asteraceae, Boraginaceae, Brassicaceae and Chenopodiaceae, which are scientifically and ecologically important, but less relevant for crop research. Indeed, our most important crop species, including rice, wheat, and barley, are monocot C_3 grasses. Yet only a few C_2 species have been identified in the Cyperacae and Poaceae monocot lineages (Sage *et al.*, 2014). Thus, improvement programs for these crops would benefit more from the discovery of C_3-C_4 intermediate photosynthesis in another grass species with inevitably similar leaf anatomy, such as *A. semialata*, than from current eudicot intermediate model systems. Furthermore, if C_3-C_4 intermediate plants are inherently plastic, as is

hypothesized here (Box 6.1), crop-breeding programs may benefit more from engineering C_3-C_4 intermediate, rather than C_4, physiology. Not only should this improve physiology over C_3 outright, but it would also ease the full transition to C_4 physiology, if still beneficial. Moreover, maintaining physiological plasticity should be beneficial in the less stable climate that it expected in the near future.

Concluding remarks: What have we learned about complex trait evolution?

This dissertation set out to better understand how novel traits of great complexity assemble. Using the example of C_4 photosynthesis, this dissertation elucidated several patterns that are relevant to the general study of complex trait evolution:

1. Flexibility underlying the complex phenotype promotes proliferation

The literature review presented in Chapter 2 established that the C_4 trait is not a rigid, single phenotype, but can be accomplished through a plethora of trait combinations. This conclusion has implications for the convergence of this trait. C_4 photosynthesis has evolved independently nearly 70 times (Sage *et al.*, 1999a; Sage *et al.*, 2011; GPWGII, 2012). This would be impressive if all of these lineages had converged on a single phenotype. Instead, these lineages could use a variety of routes to converge on a broadly C_4 phenotype. The flexibility to achieve C_4 functionality through various anatomical and biochemical modification has allowed this complex trait to evolve so many times and in impressively distant taxa.

2. Tight structure- function relationships drive complex trait evolution

The evolutionary path from an ancestral state to a derived complex trait must pass through a series of intermediate stages, where only some of the individual traits are present. Each of these intermediate stages must still be viable and should provide an advantage over the previous stage to be favoured by natural selection. This pattern is depicted well in Chapter 3, where structural changes are tightly coupled with functional responses. The spectrum between C_3 and C_4 photosynthetic types expressed by *A. semialata* is characterised by continua of

anatomical phenotypes that correspond to physiology, such that individuals with intermediate anatomies are physiologically more advantageous than those with more C_3-like anatomies, under certain environments. Thus, each anatomical change that is acquired along the C_3/C_4 transition will likely be accompanied by a physiological, and thus likely an adaptive, advantage.

3. Evolving complex traits has complex consequences

Adaptation of organisms to specific environments involves novel suites of coordinated characters, but how the emergence of these trait complexes directly affects the ecological niche remains unknown. The research presented in Chapter 4 shows that the lineages that evolved C_4 photosynthesis expanded ecologically and geographically in a short amount of time, while lineage that did not evolve this complex trait remained confined to a limited geographic area and restricted ecological conditions. This research revealed that the evolution of C_4 photosynthesis directly broadens the ecological niche, allowing dispersal into more diverse conditions and over longer distances. Thus, the evolution of complex traits may allow for rapid niche expansion and subsequent selection for new traits.

4. Intermediate phenotypes facilitate complex trait evolution

Complex traits are composed of multiple individual traits that must all be in place to function as a complex unit. Thus, the complex trait phenotype differs greatly from the simpler ancestors that initially gave rise to the complexity. Indeed, a comparison of typical C_3 and highly optimized C_4 phenotypes will reveal a remarkable anatomical and biochemical gap, over which the evolutionary transition between these photosynthetic types would seem extraordinary. However, C_4 photosynthesis does not emerge from the average C_3 taxon, but from ancestors with leaf anatomical properties much closer to the C_4 requirements. These rather intermediate phenotypes are consequently key to bridging the vast gap between C_3 and C_4 phenotypes. Thus, novel complex phenotypes do not evolve from “typical” ancestors, but from intermediate phenotypes that are closer to the novel type.

5. Progression toward complexity is not inevitable

There is a certain amount of self-organization in our universe that generates order out of chaos. These ideas, however, lead to the notion that complexity is inevitable (*e.g.,* that humans are the pinnacle of evolution). This thinking is not only wrong, but it dangerously ignores patterns of reversion, where complex forms evolve to become less complex. The research presented in this dissertation suggests that a C_3-C_4 intermediate MRCA gave rise to lineages using C_4 photosynthesis, lending support for the directionality toward complex traits. However, the dissertation also provides evidence that a C_3-C_4 intermediate lineage reverted to C_3 photosynthesis. Indeed, evolution can reverse and there is growing discussion of, and evidence for, these reversions in the literature (*e.g.,* Wiens, 2001; Porter and Crandall, 2003; Protas *et al*., 2007). In fact, reversions from C_4 to C_3 photosynthetic types have been proposed in other Paniceae grasses (Duvall *et al.,* 2001; Duvall *et al*., 2003). A comprehensive understanding of complex trait evolution consequently must incorporate the capacity to progress toward more complex forms but also to revert toward simpler ones.

REFERENCES

Abdallah JM, Monela GG. 2007. Overview of Miombo woodlands in Tanzania. MITMIOMBO–Management of indigenous tree species for ecosystem restoration and wood production in semiarid Miombo woodlands in Eastern Africa. Working Papers of the Finnish Forest Research Institute 50, 9-23.

Acocks JPH.1975. Veld Types of South Africa (with accompanying Veld Type Map). Botanical Research Institute, Department of Agricultural Technical Services; 2nd edition. Pretoria, Government Printer.

Alfaro ME, Bolnick DI, Wainwright PC. 2005. Evolutionary consequences of many-to-one mapping of jaw morphology to mechanics in labrid fishes. The American Naturalist 165, E140-E154.

Ames M, Spooner DM. 2008. DNA from herbarium specimens settles a controversy about origins of the European potato. American Journal of Botany 95, 252–257.

Apel P, Bauwe H, Bassüner B, Maass I. 1988. Photosynthetic properties of *Flaveria cronquistii*, *F. palmeri*, and hybrids between them. Biochemie und Physiologie der Pflanzen 183, 291-299.

Araújo MB, Ferri-Yáñez F, Bozinovic F, Marquet PA, Valladares F, Chown SL. 2013. Heat freezes niche evolution. Ecology Letters 16, 1206-1219.

Araus JL, Brown RH, Bouton JH, Serret MD. 1990. Leaf anatomical characteristics in *Flaveria trinervia* (C_4), *Flaveria brownii* (C_4-like) and their F1 hybrid. Photosynthesis Research 26, 49-57.

Araus JL, Brown HR, Byrd GT, Serret MD. 1991. Comparative effects of growth irradiance on photosynthesis and leaf anatomy of *Flaveria brownii* (C_4-like), *Flaveria linearis* (C_3–C_4) and their F1 hybrid. Planta 183, 497-504.

Archibald S, Lehmann CE, Gómez-Dans JL, Bradstock RA. 2013. Defining pyromes and global syndromes of fire regimes. Proceedings of the National Academy of Sciences 110, 6442-6447.

Aubry S, Kneřová J, Hibberd JM. 2013. Endoreduplication is not involved in bundle-sheath formation in the C_4 species *Cleome gynandra*. Journal of Experimental Botany 65, 3557-3566

Barrett DR, Frean ML, Cresswell CF. 1983. C_3 and C_4 photosynthetic and anatomical forms of *Alloteropsis semialata* (R. Br.) Hitchcock I. Variability in photosynthetic characteristics, water utilization efficiency and leaf anatomy. Annals of Botany 51, 801-809.

Bauwe H, Keerberg O, Bassüner R, Pärnik T, Bassüner B. 1987. Reassimilation of carbon dioxide by *Flaveria* (Asteraceae) species representing different types of photosynthesis. Planta 172, 214-218.

Bauwe H, Hagemann M, Fernie AR. 2010. Photorespiration: players, partners and origin. Trends in Plant Science 15, 330-336.

Beaulieu JM, Leitch IJ, Patel S, Pendharker A, Knight CA. 2008. Genome size is a strong predictor of cell size and stomatal density in angiosperms. New Phytologist 179, 975-986.

Beerling DJ, Royer DL. 2011. Convergent Cenozoic CO_2 history. Nature Geoscience 4, 418-420.

Berry JA, Downton WJS, Tregunna EB. 1970. The photosynthetic carbon metabolism of *Zea mays* and *Gomphrena globosa*: the location of the CO_2 fixation and the carboxyl transfer reactions. Canadian Journal of Botany 48, 777-786.

Besnard G, Christin PA, Male PJ, Coissac E, Ralimanana, Vorontsova MS. 2013. Phylogenomics and taxonomy of Lecomtelleae (Poaceae), an isolated panicoid lineage from Madagascar. Annals of Botany 112, 1057-1066.

Besnard G, Christin PA, Malé PJG, Lhuillie E, Lauzeral C, Coissac E, Vorontsova MS. 2014. From museums to genomics: old herbarium specimens shed light on a C_3 to C_4 transition. Journal of Experimental Botany 65, 6711-6721.

Birky CW. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. Proceedings of the National Academy of Sciences 92, 11331-11338.

Bivand R, Keitt T, Rowlingson B, Pebesma E, Sumner M, Hijmans R, Rouault E. 2014. RGDAL: Bindings for the geospatial data abstraction library. URL https://r-forge.r-project.org/projects/rgdal/.

Björkman O. 1971 Comparative photosynthetic CO_2 exchange in higher plants. In Hatch MD, Osmond CB, Slatyer RO; eds. *Photosynthesis and Photorespiration*. New York, NY: Wiley-Interscience, pp. 18–32.

Boaler SB. 1966. Ecology of a Miombo Site, Lupa North Forest Reserve, Tanzania: II. Plant communities and seasonal variation in the vegetation. The Journal of Ecology 54, 465-479.

Boardman NK. 1977. Comparative photosynthesis of sun and shade plants. Annual Review of Plant Physiology 28, 355-377.

Bolton JK, Brown RH. 1980. Photosynthesis of grass species differing in carbon dioxide fixation pathways V. Response of *Panicum maximum, Panicum milioides*, and tall fescue (*Festuca arundinacea*) to nitrogen nutrition. Plant Physiology 66, 97-100.

Botha CEJ. 1992. Plasmodesmatal distribution, structure and frequency in relation to assimilation in C_3 and C_4 grasses in southern Africa. Planta 187, 348-358.

Bouchenak-Khelladi Y, Slingsby JA, Verboom GA, Bond WJ. 2014. Diversification of C_4 grasses (Poaceae) does not coincide with their ecological dominance. American Journal of Botany 101, 300-307.

Bouton JH, Brown RH, Evans PT Jernstedt JA. 1986. Photosynthesis, leaf anatomy and morphology of progeny from hybrids between C_3 and C_3/C_4 *Panicum* species. Plant Physiology 80, 487-492.

Bowes G. 2011. Single-cell C_4 photosynthesis in aquatic plants. *In* Raghavendra AS, Sage RF. eds., *C_4 Photosynthesis and Related CO_2 Concentrating Mechanisms*. The Netherlands: Springer, pp. 63-80.

Bowes G, Salvucci ME. 1984. *Hydrilla*: inducible C_4-type photosynthesis without Kranz anatomy. Advances in Photosynthesis Research 3, 829-832.

Bowes G, Salvucci ME. 1989. Plasticity in the photosynthetic carbon metabolism of submersed aquatic macrophytes. Aquatic Botany 34, 233-266.

Bowes G, Holaday AS, Haller WT. 1979. Seasonal variation in the biomass, tuber density, and photosynthetic metabolisms of *Hydrilla* in three Florida lakes. Journal of Aquatic Plant Management 17, 61-65.

Bowes G, Rao SK, Estavillo GM, Reiskind JB. 2002. C_4 mechanisms in aquatic angiosperms: comparisons with terrestrial C_4 systems. Functional Plant Biology 29, 379-392.

Bowman SM, Patel M, Yerramsetty P, Mure CM, Zielinski AM, Bruenn JA, Berry JO. 2013. A novel RNA binding protein affects *rbcL* gene expression and is specific to bundle sheath chloroplasts in C_4 plants. BMC Plant Biology 13, 138.

Boykin LM, Pockman WT, Lowrey TK. 2008. Leaf anatomy of Orcuttieae (Poaceae: Chloridoideae): more evidence of C_4 photosynthesis without Kranz anatomy. Madroño 55, 143-150.

Bradshaw AD.1965. Evolutionary significance of phenotypic plasticity in plants. Advances in Genetics 13, 115-155.

Bragg FJ, Prentice IC, Harrison SP, Eglinton G, Foster PN, Rommerskirchen F, Rullkötter J. 2013. Stable isotope and modelling evidence for CO_2 as a driver of glacial–interglacial vegetation shifts in southern Africa. Biogeosciences 10, 2001-2010.

Bräutigam A, Schliesky S, Külahoglu C, Osborne CP, Weber AP. 2014. Towards an integrative model of C_4 photosynthetic subtypes: insights from comparative transcriptome analysis of NAD-ME, NADP-ME, and PEP-CK C_4 species. Journal of Experimental Botany 65, 3579-3593.

Brodribb T, Feild T, Jordan G. 2007. Leaf maximum photosynthetic rate and venation are linked by hydraulics. Plant Physiology 144, 1890–1898.

Bromham L, Bennett TH. 2014. Salt tolerance evolves more frequently in C_4 grass lineages. Journal of Evolutionary Biology 27, 653-659.

Brown WV. 1975. Variations in anatomy, associations, and origins of Kranz tissue. American Journal of Botany 62, 395-402.

Brown RH, Hattersley PW. 1989. Leaf anatomy of C_3-C_4 species as related to evolution of C_4 photosynthesis. Plant physiology 91, 1543-1550.

Brown WV. 1977. The Kranz syndrome and its subtypes in grass systematics. Memoirs of the Torrey Botanical Club 23, 1-97.

Brown WV, Smith BN. 1972. Grass evolution, the Kranz syndrome, $^{13}C/^{12}C$ ratios, and continental drift. Nature 239, 345-346.

Brown RH, 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.

Bruhl JJ, Perry S. 1995. Photosynthetic pathway-related ultrastructure of C_3, C_4, and C_3-like C_3-C_4 intermediate sedges (Cyperaceae), with special reference to *Eleocharis.* Australian Journal of Plant Physiology 22, 521-530.

Byrt CS, Grof CPL, Furbank RT. 2011. C_4 plants as biofuel feedstocks: optimizing biomass production and feedstock quality from a lignocellulosic perspective. Journal of Integrative Plant Biology 53, 120-135.

Carolin RC, Jacobs SWL, Vesk M. 1973. The structure of the cells of the mesophyll and parenchymatous bundle sheath of the Gramineae. Botanical Journal of the Linnean Society 66, 259-275.

Carolin RC, Jacobs SWL, Vesk M. 1975. Leaf structure in Chenopodiaceae. Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie 95, 226-255.

Carolin RC, Jacobs SWL, Vesk M. 1977. The ultrastructure of Kranz cells in the family Cyperaceae. Botanical Gazette 138, 413-419.

Caswell H, Reed F, Stephenson SN, Werner PA. 1973. Photosynthetic pathways and selective herbivory: A hypothesis. The American Naturalist 107, 465-480.

Cavagnaro JB. 1988. Distribution of C_3 and C_4 grasses at different altitudes in a temperate arid region of Argentina. Oecologia 76, 273-277.

Cerling TE, Harris JM, MacFadden BJ, Leakey MG, Quade J, Eisenmann V, Ehleringer JR. 1997. Global vegetation change through the Miocene/Pliocene boundary. Nature 389, 153-158.

Chollet R, Ogren WL. 1975. Regulation of photorespiration in C_3 and C_4 species. The Botanical Review 41, 137-179.

Christin PA, Osborne CP. 2013. The recurrent assembly of C_4 photosynthesis, an evolutionary tale. Photosynthesis Research 117, 163-175.

Christin PA, Osborne CP. 2014. The evolutionary ecology of C_4 plants. New Phytologist 204, 765-781.

Christin PA, Besnard G, Samaritani E, Duvall MR, Hodkinson TR, Savolainen, V, Salamin N. 2008. Oligocene CO_2 decline promoted C_4 photosynthesis in grasses. Current Biology 18, 37-43.

Christin PA, Freckleton RP, Osborne CP. 2010. Can phylogenetics identify C_4 origins and reversals? Trends in Ecology & Evolution 25, 403-409.

Christin PA, Sage TL, Edwards EJ, Ogburn RM, Khoshravesh R, Sage RF. 2011a. Complex evolutionary transitions and the significance of C_3-C_4 intermediate forms of photosynthesis in Molluginaceae. Evolution 65, 643-660.

Christin PA, Osborne CP, Sage RF, Arakaki M, Edwards EJ. 2011b. C_4 eudicots are not younger than C_4 monocots. Journal of Experimental Botany 62, 3171–3181.

Christin PA, Edwards EJ, Besnard G, Boxall SF, Gregory R, Kellogg EA, Hartwell J, Osborne CP. 2012a. Adaptive evolution of C_4 photosynthesis through recurrent lateral gene transfer. Current Biology 22, 445-449.

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

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

Christin PA, Boxall SF, Gregory R, Edwards EJ, Hartwell J, Osborne CP. 2013b. Parallel recruitment of multiple genes into C_4 photosynthesis. Genome Biology and Evolution 5, 2174-2187.

Christin PA, Spriggs E, Osborne CP, Stromberg CAE, Salamin N, Edwards EJ. 2014a. Molecular dating, evolutionary rates, and the age of the grasses. Systematic Biology 63, 153-165.

Christin PA, Arakaki M, Osborne CP, Bräutigam A, Sage RS, Hibberd JM, Kelly S, Covshoff S, Wong GKS, Hancock L. Edwards EJ. 2014b. Shared origins of a key enzyme during the evolution of C_4 and CAM metabolism. Journal of Experimental Botany 65, 3609-3621.

Connor HE. 1979. Breeding systems in the grasses: A survey. New Zealand Journal of Botany 17, 547-574.

Covshoff S, Hibberd, JM. 2012. Integrating C_4 photosynthesis into C_3 crops to increase yield potential. Current Opinion in Biotechnology 23, 209-214.

Crookston RK, Moss DN. 1970. The relation of carbon dioxide compensation and chlorenchymatous vascular bundle sheaths in leaves of dicots. Plant Physiology 46, 564-567.

Darwin C. 1859. On the origin of the species by means of natural selection. Murray, London.

Das VSR, Raghavendra AS. 1976. C_4 photosynthesis and a unique type of Kranz anatomy in *Glassocordia boswallaea* (Asteraceae). Proceedings of the Indian Academy of Sciences 84B 1, 12–19.

Dengler NG, Dengler RE, Hattersley PW. 1985. Differing ontogenetic origins of PCR ("Kranz") sheaths in leaf blades of C_4 grasses (Poaceae). American Journal of Botany 72, 284-302.

Dengler NG, Dengler RE, Donnelly PM, Hattersley PW. 1994. Quantitative leaf anatomy of C_3 and C_4 grasses (Poaceae): Bundle sheath and mesophyll surface area relationships. Annals of Botany 73, 241-255.

Dengler NG, Donnelly PM, Dengler RE. 1996. Differentiation of bundle sheath, mesophyll, and distinctive cells in the C_4 grass *Arundinella hirta* (Poaceae). American Journal of Botany 83, 1391-1405.

Devi MT, Rajagopalan AV, Raghavendra AS. 1995. Predominant localization of mitochondria enriched with glycine-decarboxylating enzymes in bundle-sheath cells of *Alternanthera tenella*, a C_3-C_4 intermediate species. Plant, Cell & Environment 18, 589–594.

Downton WJS, Tregunna EB. 1968. Carbon dioxide compensation-its relation to photosynthetic carboxylation reactions, systematics of the Gramineae, and leaf anatomy. Canadian Journal of Botany 46, 207-215.

Drennan PM, Nobel PS. 2000. Responses of CAM species to increasing atmospheric CO_2 concentrations. Plant, Cell & Environment 23, 767-781.

Driessen P, Deckers J, Spaargaren O, Nachtergaele F. 2000. Lecture notes on the major soils of the world (No. 94). Food and Agriculture Organization (FAO).

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7, 214.

Duval-Jouve J. 1875. Histotaxie des feuilles des Graminees. Annales des Science Naturelles Botanical Series 6 1, 227-346.

Duvall MR, JD Noll, AH Minn. 2001 Phylogenetics of Paniceae (Poaceae). American Journal of Botany 88, 1988–1992.

Duvall MR, Saar DE, Grayburn WS, Holbrook GP. 2003. Complex transitions between C_3 and C_4 photosynthesis during the evolution of Paniceae: A phylogenetic case study emphasizing the position of *Steinchisma hians* (Poaceae), a C_3-C_4 intermediate. International Journal of Plant Sciences 164, 949-958.

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

Edwards EJ, Smith SA. 2010. Phylogenetic analyses reveal the shady history of C_4 grasses. Proceedings of the National Academy of Science 107, 2532-2537.

Edwards EJ, Still CJ. 2008. Climate, phylogeny and the ecological distribution of C_4 grasses. Ecology Letters 11, 266-276.

Edwards EJ, Osborne CP, Strömberg CAE, Smith SA, C_4 Grass Consortium. 2010. The origins of C_4 grasslands: integrating evolutionary and ecosystem science. Science 328, 587-591.

Edwards GE, Ku MSB. 1987. Biochemistry of C_3-C_4 intermediates. *In* Hatch MD, Boardman NK, eds. *The Biochemistry of Plants: a Comprehensive Treatise.* New York: Academic Press, vol. 10, 275-325.

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*. The Netherlands: Springer, 29-61.

Edwards GE, Franceschi VR, Voznesenskaya EV. 2004. Single-cell C_4 photosynthesis versus the dual-cell (Kranz) paradigm. Annual Review of Plant Biology 55, 173-196.

Ehleringer JR. 1978. Implications of quantum yield differences on the distributions of C_3 and C_4 grasses. Oecologia 31, 255-267.

Ehleringer J, Björkman O. 1977. Quantum yields for CO_2 uptake in C_3 and C_4 plants dependence on temperature, CO_2, and O_2 concentration. Plant Physiology 59, 86-90.

Ehleringer JR, Monson RK. 1993. Evolutionary and ecological aspects of photosynthetic pathway variation. Annual Review of Ecology and Systematics 24, 411-439.

Ehleringer JR, Sage RF, Flanagan LB, Pearcy RW. 1991. Climate change and the evolution of C_4 photosynthesis. Trends in Ecology & Evolution 6, 95-99.

Ehleringer JR, Cerling TE, Helliker BR. 1997. C_4 photosynthesis, atmosphere CO_2, and climate. Oecologia 112, 285-299.

Ellis RJ. 1979. The most abundant protein in the world. Trends in Biochemical Sciences 4, 241–244.

Ellis RP. 1974a. The significance of the occurrence of both Kranz and non-Kranz leaf anatomy in the grass species *Alloteropsis semialata*. South African Journal of Science 70, 169-173.

Ellis RP. 1974b. Anomalous vascular bundle sheath structure in *Alloteropsis semialata* leaf blades. Bothalia 11, 273-275.

Ellis RP. 1977. Distribution of the Kranz syndrome in the southern African

Eragrostoideae and Panicoideae according to bundle sheath anatomy and cytology. Agroplantae 9, 73–110.

Ellis RP. 1981. Relevance of comparative leaf anatomy in taxonomic and functional research on the South African Poaceae. DSc Thesis. University of Pretoria, South Africa.

El-Sharkawy M, Hesketh J. 1965. Photosynthesis among species in relation to characteristics of leaf anatomy and CO_2 diffusion resistances. Crop Science 5, 517-521.

Epstein HE, Lauenroth WK, Burke IC, Coffin DP. 1997. Productivity patterns of C_3 and C_4 functional types in the U.S. Great Plains. Ecology 78, 722-731.

Evans JR, Von Caemmerer S. 1996. Carbon dioxide diffusion inside leaves. Plant Physiology 110, 339-346.

FAO/IIASA/ISRIC/ISSCAS/JRC, 2012. Harmonized World Soil Database (version 1.2). FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Farquhar GD. 1983. On the nature of carbon isotope discrimination in C_4 species. Functional Plant Biology 10, 205-226.

Farquhar GD, Ehleringer JR, Hubick KT. 1989. Carbon isotope discrimination and photosynthesis. Annual Review of Plant Biology 40, 503-537.

Feeley KJ. 2012. Distributional migrations, expansions, and contractions of tropical plant species as revealed in dated herbarium records. Global Change Biology 18, 1335-1341.

Fernald RD. 2006. Casting a genetic light on the evolution of eyes. Science 313, 1914-1918.

Fernie AR, Bauwe H, Eisenhut M, Florian A, Hanson DT, Hagemann M, Keech O, Mielewczik M, Nikoloski Z, Peterhänsel C, Roje S, Sage R, Timm S, von Cammerer S, Weber APM, Westhoff P. 2013. Perspectives on plant photorespiratory metabolism. Plant Biology 15, 748-753.

Ficetola GF, Miaud C, Pompanon F, Taberlet P. 2008. Species detection using environmental DNA from water samples. Biology Letters 4, 423-425.

Fischer G, Nachtergaele F, Prieler S, van Velthuizen HT, Verelst L, Wiberg D. 2008. Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008). IIASA, Laxenburg, Austria and FAO, Rome, Italy.

Fox DL, Koch PL. 2003. Tertiary history of C_4 biomass in the Great Plains, USA. Geology 31, 809-812.

Frean ML. Marks E. 1988. Chromosome numbers of C_3 and C_4 variants within the species *Alloteropsis semialata* (R.Br.) Hitchc. (Poaceae). Botanical Journal of the Linnean Society 97, 255–259.

Frean ML, Ariovich D, Cresswell CF. 1983a. C_3 and C_4 photosynthetic and anatomical forms of *Alloteropsis semialata* (R. Br.) Hitchcock. II. A comparative investigation of leaf ultrastructure and distribution of chlorenchyma in the two forms. Annals of Botany 51, 811-821.

Frean ML, Barrett DR, Ariovich D, Wolfson MM, Cresswell CF. 1983b. Intraspecific variability in *Alloteropsis semialata* (R. Br.) Hitchc. Bothalia 14, 901-913.

Freitag H, Kadereit G. 2013. C_3 and C_4 leaf anatomy types in Camphorosmeae (Camphorosmoideae, Chenopodiaceae). Plant Systematics and Evolution 300, 665-687.

Freitag H, Stichler W. 2000. A remarkable new leaf type with unusual photosynthetic tissue in a central Asiatic genus of Chenopodiaceae. Plant Biology 2, 154-160.

Friedman J, Barrett SC. 2009. Wind of change: new insights on the ecology and evolution of pollination and mating in wind-pollinated plants. Annals of Botany 103, 1515-1527.

Frost P. 1996. The ecology of Miombo woodlans. *In* Campbell BM, ed. *The Miombo in Transition: Woodlands and Welfare in Africa*. Cifor.

Furbank RT. 2011. Evolution of the C_4 photosynthetic mechanisms: are there really three C_4 acid decarboxylation types? Journal of Experimental Botany 62, 3103-3108.

Gatesy SM, Dial KP. 1996. Locomotor modules and the evolution of avian flight. Evolution 50, 331-340.

Gavrilets S. 2000. Rapid evolution of reproductive barriers driven by sexual conflict. Nature 403, 886-889.

Gerbaud A, André M. 1987. An evaluation of the recycling in measurements of photorespiration. Plant Physiology 83, 933-937.

Gibbs Russell GE. 1983. The taxonomic position of C_3 and C_4 *Alloteropsis semialata* (Poaceae) in southern Africa. Bothalia 14, 205-213.

Gong CM, Bai J, Deng JM, Wang GX, Liu XP. 2011. Leaf anatomy and photosynthetic carbon metabolic characteristics in *Phragmites communis* in different soil water availability. Plant Ecology 211, 675-687.

Gould SJ, Vrba ES. 1982. Exaptation- a missing term in the science of form. Paleontological Society 8, 4-15.

Gowik U, Westhoff P. 2011. The path from C_3 to C_4 photosynthesis. Plant Physiology 155, 56-63.

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

Griffiths H, Weller G, Toy LFM, Dennis RJ. 2013. You're so vein: bundle sheath physiology, phylogeny and evolution in C_3 and C_4 plants. Plant, Cell & Environment 36, 249-261.

Grime JP, Mowforth MA. 1982. Variation in genome size- an ecological interpretation. Nature 299, 151-153.

Guralnick RP, Wieczorek J, Beaman R, Hijmans RJ, the BioGeomancer Working Group. 2006. BioGeomancer: automated georeferencing to map the world's biodiversity data. PLoS Biology 4, e381.

Haberlandt G. 1884. Physiologische Pflanzenanatomie. Leipzig: Engelmann.

Hatch MD. 1971. Mechanism and function of the C_4 pathway of photosynthesis. *In* eds Hatch MD, Osmond CB, Slayter RO Photosynthesis and Photorespiration. Wiley, New York. pp 139–152.

Hatch MD. 1987. C_4 photosynthesis: a unique blend of modified biochemistry, anatomy and ultrastructure. Biochimica et Biophysica Acta 895, 81-106.

Hatch MD, Burnell JN. 1990. Carbonic anhydrase activity in leaves and its role in the first step of C_4 photosynthesis. Plant Physiology 93, 825-828.

Hatch MD, Osmond, CB. 1976. Compartmentation and transport in C_4 photosynthesis. Encyclopedia of Plant Physiology 3, 144-184.

Hatch MD, Agostino A, Jenkins CL. 1995. Measurement of the leakage of CO_2 from bundle-sheath cells of leaves during C_4 photosynthesis. Plant Physiology 108, 173-181.

Hattersley PW. 1984. Characterization of C_4 type leaf anatomy in grasses (Poaceae). Mesophyll: bundle sheath area ratios. Annals of Botany 53, 163-179.

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, Watson L. 1975. Anatomical parameters for predicting photosynthetic pathways of grass leaves: the 'maximum lateral cell count' and the 'maximum cells distant count'. Phytomorphology 25, 325-33.

Hattersley PW, Watson L. 1976. C_4 grasses: an anatomical criterion for distinguishing between NADP-malic enzyme species and PCK or NAD-malic enzyme species. Australian Journal of Botany 24, 297–308.

Hattersley PW, Watson L. 1992. Diversification of photosynthesis. *In* Chapman GP, ed. *Grass Evolution and Domestication*. Cambridge: Cambridge University Press, pp. 38–116.

Hattersley PW, Watson L, Osmond CB. 1977. In situ immunofluorescent labeling of Ribulose-1, 5-bisphosphate Carboxylase in leaves of C_3 and C_4 plants. Australian Journal of Plant Physiology 4, 523- 539.

Hattersley PW, Watson L, Johnston FLS, Johnson CR. 1982. Remarkable leaf anatomical variations in *Neurachne* and its allies (Poaceae) in relation to C_3 and C_4 photosynthesis. Botanical Journal of the Linnean Society 84, 265-272.

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

Haupt-Herting S, Klug K, FockHP. 2001. A new approach to measure gross CO_2 fluxes in leaves. Gross CO_2 assimilation, photorespiration, and mitochondrial respiration in the light in tomato under drought stress. Plant Physiology 126, 388–96

Hearn DJ. 2009. Descriptive anatomy and evolutionary patterns of anatomical diversification in *Adenia* (Passifloraceae). Aliso: A Journal of Systematic and Evolutionary Botany 27, 13-38.

Heckmann D, Schulze S, Denton A, Gowik U, Westhoff P, Weber AP, Lercher MJ. 2013. Predicting C_4 photosynthesis evolution: modular, individually adaptive steps on a Mount Fuji fitness landscape. Cell 153, 1579-1588.

Hendry AP, Wenburg JK, Bentzen P, Volk EC, Quinn TP. 2000. Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. Science 290, 516-518.

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.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25, 1965-1978.

Hijmans RJ, van Etten J, Mattiuzzi M, Sumner M, Greenberg JA, Lamigueiro OP, Bevan A, Racine EB, Shortridge A. 2014. Raster: geographic data analysis and modeling. URL http://cran.r-project.org/web/packages/raster/.

Hofmann RR. 1989. Evolutionary steps of ecophysiological adaptation and diversification of ruminants: a comparative view of their digestive system. Oecologia 78, 443-457.

Holaday AS, Chollet R. 1984. Photosynthetic/photorespiratory characteristics of C_3–C_4 intermediate species. Photosynthesis Research 5, 307-323.

Holmgren CA, Norris J, Betancourt JL. 2007. Inferences about winter temperatures and summer rains from the late Quaternary record of C_4 perennial grasses and C_3 desert shrubs in the northern Chihuahuan Desert. Journal of Quaternary Science 22, 141-161.

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.

Ibrahim DG. 2007. The C_3 and C_4 subspecies of Alloteropsis semialata: phylogenetic relationship and climatic responses. PhD thesis, University of Sheffield, Sheffield, UK.

Ibrahim DG, Gilbert ME, Ripley BS, Osborne CP. 2008. Seasonal differences in photosynthesis between the C_3 and C_4 subspecies of *Alloteropsis semialata* are offset by frost and drought. Plant, Cell & Environment 31, 1038-1050.

Ibrahim DG, Burke T, Ripley BS, Osborne CP. 2009. A molecular phylogeny of the genus *Alloteropsis* (Panicoideae, Poaceae) suggests an evolutionary reversion from C_4 to C_3 photosynthesis. Annals of Botany 103, 127-136.

Jenkins CL, Furbank RT, Hatch MD. 1989. Mechanism of C_4 photosynthesis. A model describing the inorganic carbon pool in bundle sheath cells. Plant Physiology 91, 1372-1381.

John CR, Smith-Unna RD, Woodfield H, Hibberd JM. 2014. Evolutionary convergence of cell specific gene expression in independent lineages of C_4 grasses. Plant Physiology 165, 62-75.

Jordan DB, Ogren WL. 1984. The CO_2/O_2 specificity of ribulose 1,5-bisphosphate carboxylase/oxygenase. Planta 161, 308-313.

Kadereit G, Borsch T, Weising K. 2003. Phylogeny of Amaranthaceae and Chenopodiaceae and the evolution of C_4 photosynthesis. International Journal of Plant Sciences 164, 959-986.

Kadereit G, Ackerly D, Pirie MD. 2012. A broader model for C_4 photosynthesis evolution in plants inferred from the goosefoot family (Chenopodiaceae s.s.). Proceedings of the Royal Society B: Biological Sciences 279, 3304-3311.

Kanai R, Edwards GE. 1999. The biochemistry of C_4 photosynthesis. In: Sage RF, Monson RK, eds. *C_4 Plant Biology*. San Diego, USA: Academic Press, pp. 49-87.

Kawamitsu Y, Hakoyama S, Agata W, Takeda T. 1985. Leaf interveinal distances corresponding to anatomical types in grasses. Plant and Cell Physiology 26, 589-593.

Keeley JE. 1998. C_4 photosynthetic modifications in the evolutionary transition from land to water in aquatic grasses. Oecologica 116, 85-97.

Keeley JE, Rundel PW. 2005. Fire and the Miocene expansion of C_4 grasslands. Ecology Letters 8, 683-690.

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.

Kellogg EA. 1999. Phylogenetic aspects of the evolution of C_4 photosynthesis. In: Sage RF, Monson RK, eds. *C_4 Plant Biology.* San Diego: Academic Press, pp. 411-444.

Kennedy RA, Laetsch WM. 1974. Plant species intermediate for C_3, C_4 photosynthesis. Science 184, 1087-1089.

Kocacinar F, Sage RF. 2003. Photosynthetic pathway alters xylem structure and hydraulic function in herbaceous plants. Plant, Cell & Environment 26, 2015-2026.

Kocacinar F, Sage RF. 2004. Photosynthetic pathway alters hydraulic structure and function in woody plants. Oecologia 139, 214-223.

Koteyeva NK, Voznesenskaya EV, Roalson EH, Edwards GE. 2011. Diversity in forms of C_4 in the genus *Cleome* (Cleomaceae). Annals of Botany 107, 269-283.

Ku MSB, Edwards GE. 1978. Photosynthetic efficiency of *Panicum hians* and *Panicum milloides* in relation to C_3 and C_4 plants. Plant Cell Physiology 19, 665-675.

Laetsch WM. 1974. The C_4 syndrome: a structural analysis. Annual Review of Plant Physiology 25, 27-52.

Lamb TD, Collin SP, Pugh EN. 2007. Evolution of the vertebrate eye: opsins, photoreceptors, retina and eye cup. Nature Reviews Neuroscience 8, 960-976.

Langdale JA. 2011. C_4 cycles: past, present, and future research on C_4 photosynthesis. The Plant Cell 23, 3879-3892.

Lê S, Josse J, Husson F. 2008. FactoMineR: an R package for multivariate analysis. Journal of Statistical Software 25, 1-18.

Leegood RC. 2002. C_4 photosynthesis: principles of CO_2 concentration and prospects for its introductions into C_3 plants. Journal of Experimental Botany 53, 581-590.

Leegood RC. 2008. Roles of the bundle sheath cells in leaves of C_3 plants. Journal of Experimental Botany 59, 1663-1673.

Liebenberg EJL, Fossey A. 2001. Comparative cytogenetic investigation of the two subspecies of the grass *Alloteropsis semialata* (Poaceae). Botanical Journal of the Linnean Society 137, 243-248.

Liu H, Edwards EJ, Freckleton RP, Osborne CP. 2012. Phylogenetic niche conservatism in C_4 grasses. Oecologia 170, 835-845.

Loomis RS, Williams WA, Hall AE. 1971. Agricultural productivity. Annual Review of Plant Physiology 22, 431–468.

Ludwig LJ, Canvin DT. 1971. The rate of photorespiration during photosynthesis and the relationship of the substrate of light respiration to the products in sunflower leaves. Plant Physiology 48, 712-719.

Lundgren MR, Osborne CP, Christin P-A. 2014. Deconstructing Kranz anatomy to understand C_4 evolution. Journal of Experimental Botany 65, 3357-3369.

Lundquist PK, Rosar C, Bräutigam A, Weber APM. 2014. Plastid signals and the bundle sheath:mesophyll development in reticulate mutants. Molecular Plant 7, 14-29.

Mallmann J, Heckmann D, Bräutigam 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.

Markwell J, Osterman JC, Mitchell JL. 1995. Calibration of the Minolta SPAD-502 leaf chlorophyll meter. Photosynthesis Research 46, 467-472.

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.

Martins S, Scatena VL. 2011. Bundle sheath ontogeny in Kranz and non-Kranz species of Cyperaceae (Poales). Australian Journal of Botany 59, 554-562.

Masterson J. 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science 264, 421-424.

Mayr E. 1959. Isolation as an evolutionary factor. Proceedings of the American Philosophical Society 103, 221-230.

Mayr E. 1970. *Populations, Species, and Evolution.* Harvard University Press, Cambridge, Mass.

McKown AD, Dengler NG. 2007. Key innovations in the evolution of Kranz anatomy and C_4 vein pattern in *Flaveria* (Asteraceae). American Journal of Botany 94, 382-399.

McKown AD, Dengler NG. 2009. Shifts in leaf vein density through accelerated vein formation in C_4 *Flaveria* (Asteraceae). Annals of Botany 104, 1085-1098.

McKown AD, Cochard H, Sack L. 2010. Decoding leaf hydraulics with a spatially explicit model: Principles of venation architecture and implications for its evolution. American Naturalist 175, 447-460.

Middelboe M, Holmfeldt K, Riemann L, Nybroe O, Haaber J. 2009. Bacteriophages drive strain diversification in a marine Flavobacterium: implications for phage

resistance and physiological properties. Environmental microbiology 11, 1971-1982.

Monson RK, Jaeger CH. 1991. Photosynthetic characteristics of C_3-C_4 intermediate *Flaveria floridana* (Asteraceae) in natural habitats: evidence of advantages to C_3-C_4 photosynthesis at high leaf temperatures. American Journal of Botany 78, 795-800.

Monson RK, Moore BD. 1989. On the significance of C_3-C_4 intermediate photosynthesis to the evolution of C_4 photosynthesis. Plant, Cell & Environment 12, 689-699.

Monson RK, Rawsthorne S. 2000. CO_2 assimilation in C_3–C_4 intermediate plants. In *Photosynthesis*. Springer, Netherlands, pp. 533-550.

Monson RK, Edwards GE, Ku MSB. 1984. C_3–C_4 intermediate photosynthesis in plants. Bioscience 34, 563–571.

Monson RK, Moore B, Ku MSB, Edwards GE. 1986. Co-function of C_3-and C_4-photosynthetic pathways in C_3, C_4 and C_3-C_4 intermediate *Flaveria* species. Planta 168, 493-502.

Monson RK, Teeri JA, Ku MSB, Gurevitch J, Mets LJ, Dudley S. 1988. Carbon-isotope discrimination by leaves of *Flaveria* species exhibiting different amounts of C_3-and C_4-cycle co-function. Planta 174, 145-151.

Monteiro A, Podlaha O. 2009. Wings, horns, and butterfly eyespots: how do complex traits evolve? PLoS biology 7, e1000037.

Monteith JL. 1978. A reassessment of maximum growth rates for C_3 and C_4 crops. Experimental Agriculture 14, 1–5.

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 JA, Brown RH, Reger BJ. 1980. Photosynthesis in grass species differing in carbon dioxide fixation pathways III. Oxygen response and enzyme activities of species in the Laxa group of *Panicum*. Plant Physiology 65, 156-159.

Muhaidat R, Sage RF, Dengler NG. 2007. Diversity of Kranz anatomy and biochemistry in C_4 eudicots. American Journal of Botany 94, 362-381.

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 & Environment 10, 1723-1736.

Murphy LR, Barroca J, Franceschi VR, Lee R, Roalson EH, Edwards GE, Ku MSB. 2007. Diversity and plasticity of C_4 photosynthesis in *Eleocharis* (Cyperaceae). Functional Plant Biology 34, 571-580.

Nelson T. 2011. Development of leaves in C_4 plants: anatomical features that support C_4 metabolism. In: Raghavendra AS, Sage RF. eds., *C_4 Photosynthesis and Related CO_2 Concentrating Mechanisms*. The Netherlands: Springer, 147-159.

New M, Hulme M, Jones PD. 1999. Representing twentieth century space-time climate variability. Part 1: development of a 1961-90 mean monthly terrestrial climatology. Journal of Climate 12, 829-856.

New M, Lister D, Hulme M, Makin I. 2002. A high-resolution data set of surface climate over global land areas. Climate Research 21, 1-25.

Nobel PS. 1996. Responses of some North American CAM plants to freezing temperatures and doubled CO_2 concentrations: implications of global climate change for extending cultivation. Journal of Arid Environments 34, 187-196.

Noblin X, Mahadevan L, Coomaraswamy IA, Weitz DA, Holbrook NM, Zwieniecki MA. 2008. Optimal vein density in artificial and real leaves. Proceedings of the National Academy of Sciences 105, 9140-9144.

Oakley JC, Sultmanis S, Stinson CR, Sage TL, Sage RF. 2014. Comparative studies of C_3 and C_4 *Atriplex* hybrids in the genomics era: physiological assessments. Journal of Experimental Botany 65, 3637-3647.

Ocampo G, Koteyeva NK, Voznesenskaya EV, Edwards GE, Sage TL, Sage RF, Columbus JT. 2013. Evolution of leaf anatomy and photosynthetic pathways in Portulacaceae. American Journal of Botany 100, 2388-2402.

Ogle K. 2003. Implications of interveinal distance for quantum yield in C_4 grasses: a modeling and meta-analysis. Oecologia 4, 532-542.

Ogren WL. 1984. Photorespiration: pathways, regulation, and modification. Annual Review of Plant Physiology 35, 415-442.

Olesen P. 1975. Plasmodesmata between mesophyll and bundle sheath cells in relation to the exchange of C_4-acids. Planta 123, 199-202.

Orme D, Freckleton R, Thomas G, Petzoldt T, Fritz S, Isaac N, Pearse W. 2012. Caper: Comparative analyses of phylogenetics and evolution in R Package version 0.5. URL: cran.r-project.org/web/packages/caper/index.html.

Orr MR, Smith TB. 1998. Ecology and speciation. Trends in Ecology & Evolution 13, 502-506.

Osborne CP, Beerling DJ. 2006. Nature's green revolution: the remarkable evolutionary rise of C_4 plants. Philosophical Transactions of the Royal Society B: Biological Sciences 361, 173-194.

Osborne CP, Freckleton RP. 2009. Ecological selection pressures for C_4 photosynthesis in the grasses. Proceedings of the Royal Society B 276, 1753-1760.

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.

Osborne CP, Wythe EJ, Ibrahim DG, Gilbert ME, Ripley BS. 2008. Low temperature effects on leaf physiology and survivorship in the C_3 and C_4 subspecies of *Alloteropsis semialata*. Journal of Experimental Botany 59, 1743-1754.

Osborne CP, Salomaa A, Kluyver TA, Visser V, Kellogg EA, Morrone O, Vorontsova MS, Clayton WD, Simpson DA. 2014. A global database of C_4 photosynthesis in grasses. New Phytologist 204, 441-446.

Osmond CB. 1978. Crassulacean acid metabolism: a curiosity in context. Annual Review of Plant Physiology 29, 379-414.

Pagani M, Freeman KH, Arthur MA. 1999. Late Miocene atmospheric CO_2 concentrations and the expansion of C_4 grasses. Science 285, 876-879.

Pagani M, Zachos JC, Freeman KH, Tipple B, Bohaty S. 2005. Marked decline in atmospheric carbon dioxide concentrations during the Paleogene. Science 309, 600-603.

Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290.

Parr CL, Lehmann CE, Bond WJ, Hoffmann WA, Andersen AN. 2014. Tropical grassy biomes: misunderstood, neglected, and under threat. Trends in Ecology & Evolution 29, 205-213.

Pearcy RW, Ehleringer J. 1984. Comparative ecophysiology of C_3 and C_4 plants. Plant, Cell & Environment 7, 1-13.

Pearcy RW, Troughton J. 1975. C_4 photosynthesis in tree form Euphorbia species from Hawaiian rainforest sites. Plant Physiology 55, 1054-1056.

Pebesma EJ, Bivand RS. 2005. Classes and methods for spatial data in R. R News 5, URL http://cran.r-project.org/doc/Rnews/.

Peisker M. 1986. Models of carbon metabolism in C_3-C_4 intermediate plants as applied to the evolution of C_4 photosynthesis. Plant, Cell & Environment 9, 627-635.

Peter G, Katinas L. 2003. A new type of Kranz anatomy in Asteraceae. Australian Journal of Botany 51, 217–226.

Peterson BG, Carl P, Boudt K, Bennett R, Ulrich J, Zivot E, Lestel M, Balkissoon K, Wuertz D. 2014. PerformanceAnalytics: Econometric tools for performance and risk analysis. http://r-forge.r-project.org/projects/returnanalytics/

Polhill D. 1988. Flora of Tropical East Africa: Index of Collecting Localities. Kew Publishing.

Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Systematic Biology 55, 595-609.

Porter ML, Crandall KA. 2003. Lost along the way: the significance of evolution in reverse. Trends in Ecology & Evolution 18, 541-547.

Prendergast HDV, Hattersley PW, Stone NE. 1987. New structural/biochemical associations in leaf blades of C_4 grasses (Poaceae). Functional Plant Biology 14, 403-420.

Protas M, Conrad M, Gross JB, Tabin C, Borowsky R. 2007. Regressive evolution in the Mexican cave tetra, *Astyanax mexicanus*. Current Biology 17, 452-454.

R Development Core Team. 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0. URL: http://www.R-project.org.

Rambaut A, Drummond AJ. 2007. Tracer v1.4, available at http://beadt.bio.ed.ac.uk/Tracer.

Raven JA. 2013. Rubisco: still the most abundant protein of Earth? New Phytologist 198, 1-3.

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

Renvoize SA. 1983. A survey of leaf-blade anatomy in grasses IV. Eragrostideae. Kew Bulletin 38, 469-478.

Renvoize SA. 1985. A survey of leaf-blade anatomy in grasses V. The bamboo allies. Kew Bulletin 40, 509-535.

Renvoize SA. 1987a. A survey of leaf-blade anatomy in grasses XI. Paniceae. Kew Bulletin 42, 739-768.

Renvoize SA. 1987b. A survey of leaf-blade anatomy in grasses X: Bambuseae. Kew Bulletin 42, 201-207

Ripley BS, Gilbert ME, Ibrahim DG, Osborne CP. 2007. Drought constraints on C_4 photosynthesis: stomatal and metabolic limitations in C_3 and C_4 subspecies of *Alloteropsis* semialata. Journal of Experimental Botany 58, 1351-1363.

Rodgers A, Salehe J, Howard G. 1996. The biodiversity of miombo woodlands. *In* Campbell BM, ed. The Miombo in Transition: Woodlands and Welfare in Africa. Cifor. pp 12.

Rommerskirchen F, Eglinton G, Dupont L, Rullkötter J. 2006. Glacial/interglacial changes in southern Africa: Compound-specific $\delta^{13}C$ land plant biomarker and

pollen records from southeast Atlantic continental margin sediments. Geochemistry, Geophysics, Geosystems 7, Q08010.

Rundel PW. 1980. The ecological distribution of C_4 and C_3 grasses in the Hawaiian-islands. Oecologia 45, 354-359.

Sack L, Frole K. 2006. Leaf structural diversity is related to hydraulic capacity in tropical rainforest trees. Ecology 87, 483–491.

Sack L, Scoffoni C. 2013. Leaf venation: structure, function, development, evolution, ecology and applications in the past, present and future. New Phytologist 198, 983–1000.

Sack L, Scoffoni C, McKown AD, Frole K, Rawls M, Harvan JC, Tran H, Tran T. 2012. Developmentally based scaling of leaf venation architecture explains global ecological patterns. Nature Communications 3, 837.

Sack L, Scoffoni C, John GP, Poorter H, Mason CM, Mendez-Alonzo R, Donovan LA. 2013. How do leaf veins influence the worldwide leaf economic spectrum? Review and synthesis. Journal of Experimental Botany 64, 4053-4080.

Sage RF. 2001. Environmental and evolutionary preconditions for the origin and diversification of the C_4 photosynthetic syndrome. Plant Biology 3, 202–213.

Sage RF. 2004. The evolution of C_4 photosynthesis. New Phytologist 161, 341-370.

Sage RF, Kubien DS. 2007. The temperature response of C_3 and C_4 photosynthesis. Plant, Cell & Environment 30, 1086-1106.

Sage RF, McKown AD. 2006. Is C_4 photosynthesis less phenotypically plastic than C_4 photosynthesis? Journal of Experimental Botany 57, 303-317.

Sage RF, Stata M. 2015. Photosynthetic diversity meets biodiversity: The C_4 plant example. Journal of Plant Physiology 172, 104-119.

Sage RF, Li M, Monson RK. 1999a. The taxonomic distribution of C_4 photosynthesis. In: Sage RF, Monson RK, eds. *C_4 Plant Biology*, San Diego, USA: Academic Press, pp. 551-584.

Sage RF, Wedin DA, Li M. 1999b. The biogeography of C_4 photosynthesis: patterns and controlling factors. In: Sage RF, Monson RK, eds. *C_4 Plant Biology*, San Diego, USA: Academic Press, pp. 313-373.

Sage RF, Sage TL, Pearcy RW, Borsch T. 2007. The taxonomic distribution of C_4 photosynthesis in Amaranthaceae sensu stricto. American Journal of Botany 94, 1992-2003.

Sage RF, Christin P-A, Edwards EJ. 2011. The C_4 plant lineages of planet Earth. Journal of Experimental Botany 62, 3155–69.

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

Sage TL, Busch F, 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 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.

Schaefer H, Renner SS. 2010. A gift from the New World? The West African crop *Cucumeropsis mannii* and the American *Posadaea sphaerocarpa* (Cucurbitaceae) are the same species. Systematic Botany 35, 534–540.

Schlichting CD, Levin DA. 1986. Phenotypic plasticity: an evolving plant character. Biological Journal of the Linnean Society 29, 37-47.

Schlichting CD, Pigliucci M. 1993. Evolution of phenotypic plasticity via regulatory genes. American Naturalist 142, 366-370.

Schloss PD, Handelsman J. 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Applied and Environmental Microbiology 71, 1501-1506.

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.

Schuster WS, Monson RK. 1990. An examination of the advantages of C_3-C_4 intermediate photosynthesis in warm environments. Plant, Cell & Environment 13, 903-912.

Sharkey TD. 1985. O_2-insensitive photosynthesis in C_3 plants. Its occurrence and a possible explanation. Plant Physiology 78, 71-75.

Shimazaki H, Shinomoto S. 2007. A method for selecting the bin size of a time histogram. Neural Computation 19, 1503-1527.

Šímová I, Herben T. 2012. Geometrical constraints in the scaling relationships between genome size, cell size and cell cycle length in herbaceous plants. Proceedings of the Royal Society B: Biological Sciences 279, 867-875.

Sinha NR, Kellogg EA. 1996. Parallelism and diversity in multiple origins of C_4 photosynthesis in the grass family. American Journal of Botany 83, 1458-1470.

Skillman JB. 2008. Quantum yield variation across the three pathways of photosynthesis: not yet out of the dark. Journal of Experimental Botany 59, 1647-1661.

Slatyer RA, Hirst M, Sexton JP. 2013. Niche breadth predicts geographical range size: a general ecological pattern. Ecology Letters 16, 1104-1114.

Slewinski TL, Anderson AA, Zhang C, Turgeon R. 2013. Scarecrow plays a role in establishing Kranz anatomy in maize leaves. Plant and Cell Physiology 53, 2030-2037.

Smith BS, Brown WV. 1973. The Kranz syndrome in the Gramineae as indicated by carbon isotopic ratios. American Journal of Botany 60, 505–513.

Smith BN, Epstein S. 1971. Two categories of $^{13}C/^{12}C$ ratios for higher plants. Plant Physiology 47, 380-384.

Soros CL, Dengler NG. 2001. Ontogenetic derivation and cell differentiation in photosynthetic tissues of C_3 and C_4 Cyperaceae. American Journal of Botany 88, 992-1005.

Spriggs EL, Christin PA, Edwards EJ. 2014. C_4 Photosynthesis promoted species diversification during the Miocene grassland expansion. PloS One 9, e97722.

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 & 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, 6-1.

Sugimoto-Shirasu K, Roberts K. 2003. “Big it up”: endoreduplication and cell-size control in plants. Current Opinion in Plant Biology 6, 544-553.

Sultan SE. 1987. Evolutionary implications of phenotypic plasticity in plants. Evolutionary Biology 21,127-178.

Sultan SE. 2000. Phenotypic plasticity for plant development, function and life history. Trends in Plant Science 5, 537-542.

Tateoka T. 1958. Notes on some grasses. VIII. On leaf structure of *Arundinella* and *Garnotia*. Botanical Gazette 120, 101–109.

Taub DR. 2000. Climate and the U.S. Distribution of C_4 grass subfamilies and decarboxylation variants of C_4 photosynthesis. American Journal of Botany 87, 1211-1215.

Taylor SH, Hulme SP, Rees M, Ripley BS, Ian Woodward F, Osborne CP. 2010. Ecophysiological traits in C_3 and C_4 grasses: a phylogenetically controlled screening experiment. New Phytologist 185, 780-791.

Taylor SH, Franks PJ, Hulme SP, Spriggs E, Christin PA, Edwards EJ, Woodward FI, Osborne CP. 2012. Photosynthetic pathway and ecological adaptation explain stomatal trait diversity amongst grasses. New Phytologist 193, 387-396.

Tcherkez GG, Farquhar GD, Andrews TJ. 2006. Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. Proceedings of the National Academy of Sciences 103, 7246-7251.

Teeri JA, Stowe LG. 1976. Climatic patterns and the distribution of C_4 grasses in North America. Oecologica 23, 1-12.

Teese P. 1995. Intraspecific variation for CO_2 compensation point and differential growth among variants in a C_3–C_4 intermediate plant. Oecologia 102, 371-376.

Terashima I, Miyazawa SI, Hanba YT. 2001. Why are sun leaves thicker than shade leaves? Consideration based on analyses of CO_2 diffusion in the leaf. Journal of Plant Research 114, 93-105.

Tieszen LL, Senyimba MM, Imbamba SK, Troughton JH. 1979. The distribution of C_3 and C_4 grasses and carbon isotope discrimination along an altitudinal and moisture gradient in Kenya. Oecologia 37, 337-350.

True JR, Carroll SB. 2002. Gene co-option in physiological and morphological evolution. Annual Review of Cell and Developmental Biology 18, 53-80.

Ueno O. 1995. Occurrence of distinctive cells in leaves of C_4 species in *Arthraxon* and *Microstegium* (Andropogoneae-Poaceae) and the structural and immunocytochemical characterization of these cells. International Journal of Plant Sciences 156, 270–289.

Ueno O. 2001. Environmental regulation of C_3 and C_4 differentiation in the amphibious sedge *Eleocharis vivipara*. Plant Physiology 127, 1524–1532.

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

Ueno O, Takeda T, Maeda E. 1988a. Leaf ultrastructure of C_4 species possessing different Kranz anatomical types in the Cyperaceae. The Botanical Magazine 101, 141-152.

Ueno O, Samejima M, Muto S, Miyachi S. 1988b. Photosynthetic characteristics of an amphibious plant, *Eleocharis vivipara*: Expression of C_4 and C_3 modes in contrasting environments. Proceedings of the National Academy of Sciences 85, 6733–6737.

Ueno O, Kawano Y, Wakayama M, Takeda T. 2006. Leaf vascular systems in C_3 and C_4 grasses: a two-dimensional analysis. Annals of Botany 97, 611-621.

Urban MA, Nelson DM, Jiménez-Moreno G, Châteauneuf JJ, Pearson A, Hu FS. 2010. Isotopic evidence of C_4 grasses in southwestern Europe during the Early Oligocene–Middle Miocene. Geology 38, 1091-1094.

van der Meulen F, Werger MJA. 1984. Crown characteristics, leaf size and light throughfall of some savnanna trees in southern Africa. South African Journal of Botany 3, 208-218.

Vavrek MJ. 2011. Fossil: palaeoecological and palaeogeographical analysis tools. Palaeontologia Electronica, 14:1T.

Visser V, Woodward FI, Freckleton RP, Osborne CP. 2012. Environmental factors determining the phylogenetic structure of C_4 grass communities. Journal of Biogeography 39, 232-246.

Vogelmann TC, Nishio JN, Smith WK. 1996. Leaves and light capture: light propagation and gradients of carbon fixation within leaves. Trends in Plant Science 1, 65-70.

von Caemmerer S. 1989. A model of photosynthetic CO_2 assimilation and carbon-isotope discrimination in leaves of certain C_3–C_4 intermediates. Planta 178, 463–474.

von Caemmerer S. 1992. Carbon isotope discrimination in C_3-C_4 intermediates. Plant, Cell & Environment 15, 1063–72.

von Caemmerer S, Furbank RT. 2003. The C_4 pathway: an efficient CO_2 pump. Photosynthesis Research 77, 191-207.

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

von Caemmerer S, Ghannoum O, Pengelly JJ, Cousins AB. 2014. Carbon isotope discrimination as a tool to explore C_4 photosynthesis. Journal of Experimental Botany 65, 3459-3470.

Vorontsova MS, Haevermans T, Haevermans A, Razanatsoa J, Lundgren MR, Besnard G. 2015. The Genus Sartidia (Poaceae: Aristidoideae) in Madagascar. Systematic Botany 40,

Voss I, Sunil B, Scheibe R, Raghavendra AS. 2013. Emerging concept for the role of photorespiration as an important part of abiotic stress response. Plant Biology 15, 713-722.

Voznesenskaya EV, Franceschi VR, Pyankov VI, Edwards GE. 1999. Anatomy, chloroplast structure and compartmentation of enzymes relative to photosynthetic mechanisms in leaves and cotyledons of species in the tribe Salsoleae (Chenopodiaceae). Journal of Experimental Botany 50, 1779-1795.

Voznesenskaya EV, Franceschi VR, Chuong SDX, Edwards GE. 2006. Functional characterization of phosphoenolpyruvate carboxykinase-type C_4 leaf anatomy: immuno-, chytochemical and ultrastructural analysis. Annals of Botany 98, 77-91.

Wakayama M, Ueno O, Ohnishi J. 2003. Photosynthetic enzyme accumulation during leaf development of *Arundinella hirta*, a C_4 grass having Kranz cells not associated with veins. Plant and Cell Physiology 44, 1330-1340.

Wang P, Kelly S, Fouracre JP, Langdale JA. 2013. Genome-wide transcript analysis of early maize leaf development reveals gene cohorts associated with the differentiation of C_4 Kranz anatomy. The Plant Journal 75, 656-670.

Warren BH, Strasberg D, Bruggemann JH, Prys-Jones RP, Thébaud C. 2010. Why does the biota of the Madagascar region have such a strong Asiatic flavour? Cladistics 26, 526-538.

Weiner H, Burnell JN, Woodrow IE, Heldt HW, Hatch MD. 1988. Metabolite diffusion into bundle sheath cells from C_4 plants relation to C_4 photosynthesis and plasmodesmatal function. Plant Physiology 88, 815-822.

Welkie W, Caldwell M. 1970. Leaf anatomy of species in some dicotyledon families as related to the C_3 and C_4 pathways of carbon fixation. Canadian Journal of Botany 48, 2135-2146.

West-Eberhard MJ. 1989. Phenotypic plasticity and the origins of diversity. Annual Review of Ecology and Systematics 20, 249-278.

West-Eberhard MJ, Smith JAC, Winter K. 2011. Photosynthesis, reorganized. Science 332, 311-312.

Whelan T, Sackett WM, Benedict C. 1973. Enzymatic fractionation of carbon isotopes by phosphoenolpyruvate carboxylase from C_4 plants. Plant Physiology 51, 1051-1054.

Wickham H. 2011. The split-apply-combine strategy for data analysis. Journal of Statistical Software 40, 1-29.

Widmer A, Lexer C, Cozzolino S. 2009. Evolution of reproductive isolation in plants. Heredity 102, 31-38.

Wiens JJ. 2001. Widespread loss of sexually selected traits: how the peacock lost its spots. Trends in Ecology & Evolution 16, 517-523.

Williams DG, Mack RN. Black RA. 1995. Ecophysiology of introduced *Pennisetum setaceum* on Hawaii: the role of phenotypic plasticity. Ecology 76, 1569-1580.

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

Wilson PJ, Thompson K, Hodgson JG. 2002. Specific leaf area and leaf dry matter content as alternative predictors of plant strategies. New Phytologist 143, 155-162.

Wistow GJ, Piatigorsky J. 1988. Lens crystallins: the evolution and expression of proteins for a highly specialized tissue. Annual Review of Biochemistry 57, 479-504.

Woo KC, Anderson JM, Boardman N, Downton WJS, Osmond C, Thorne SW. 1970. Deficient photosystem II in agranal bundle sheath chloroplasts of C_4 plants. Proceedings of the National Academy of Sciences 67, 18-25.

Xu X, Wang K, Zhang K, Ma Q, Xing L, Sullivan C, Hu D, Cheng S, Wang S. 2012. A gigantic feathered dinosaur from the Lower Cretaceous of China. Nature 484, 92 – 95.

Zomer RJ, Bossio DA, Trabucco A, Yuanjie L, Gupta DC, Singh VP. 2007. Trees and water: Smallholder agroforestry on irrigated lands in northern India. Colombo, Sri Lanka: International Water Management Institute. IWMI Research Report 122, . pp 45.

Zomer RJ, Trabucco A, Bossio DA, van Straaten O, Verchot LV. 2008. Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. Agriculture, Ecosystems and Environment 126, 67-80.

REPRINT OF "LUNDGREN MR, OSBORNE CP, CHRISTIN PA. 2014. DECONSTRUCTING KRANZ ANATOMY TO UNDERSTAND C_4 EVOLUTION. JOURNAL OF EXPERIMENTAL BOTANY 65, 3357-3369."

Journal of Experimental Botany, Vol. 65, No. 13, pp. 3357–3369, 2014
doi:10.1093/jxb/eru186 Advance Access publication 5 May, 2014

Journal of Experimental Botany
www.jxb.oxfordjournals.org

REVIEW PAPER

Deconstructing Kranz anatomy to understand C_4 evolution

Marjorie R. Lundgren, Colin P. Osborne and Pascal-Antoine Christin*

Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

* To whom correspondence should be addressed. E-mail: p.christin@sheffield.ac.uk

Received 13 January 2014; Revised 15 March 2014; Accepted 25 March 2014

Abstract

C_4 photosynthesis is a complex physiological adaptation that confers greater productivity than the ancestral C_3 photosynthetic type in environments where photorespiration is high. It evolved in multiple lineages through the coordination of anatomical and biochemical components, which concentrate CO_2 at the active site of ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). In most C_4 plants, the CO_2-concentrating mechanism is achieved via the confinement of Rubisco to bundle-sheath cells, into which CO_2 is biochemically pumped from surrounding mesophyll cells. The C_4 biochemical pathway relies on a specific suite of leaf functional properties, often referred to as Kranz anatomy. These include the existence of discrete compartments differentially connected to the atmosphere, a close contact between these compartments, and a relatively large compartment to host the Calvin cycle. In this review, we use a quantitative dataset for grasses (Poaceae) and examples from other groups to isolate the changes in anatomical characteristics that generate these functional properties, including changes in the size, number, and distribution of different cell types. These underlying anatomical characteristics vary among C_4 origins, as similar functions emerged via different modifications of anatomical characteristics. In addition, the quantitative characteristics of leaves all vary continuously across C_3 and C_4 taxa, resulting in C_4-like values in some C_3 taxa. These observations suggest that the evolution of C_4-suitable anatomy might require relatively few changes in plant lineages with anatomical predispositions. Furthermore, the distribution of anatomical traits across C_3 and C_4 taxa has important implications for the functional diversity observed among C_4 lineages and for the approaches used to identify genetic determinants of C_4 anatomy.

Key words: C_4 photosynthesis, complex trait, convergent evolution, co-option, Kranz anatomy, leaf.

Introduction

During the diversification of flowering plants, C_4 photosynthesis evolved from C_3 ancestors more than 62 times independently in several distantly related groups (Sage *et al.*, 2011). C_4 photosynthesis is characterized by a biochemical CO_2 pump formed by the coordination of several evolutionary novelties, which increase the relative concentration of CO_2 around ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) to nearly eliminate photorespiration (Ludwig and Canvin, 1971; Hatch, 1987; von Caemmerer and Furbank, 2003; Skillman, 2008; Sage *et al.*, 2012). The CO_2-concentrating mechanism relies on the primary fixation of atmospheric carbon by phosphoenolpyurvate carboxylase (PEPC) coupled with carbonic anhydrase. These reactions are spatially separated from the secondary refixation of CO_2 by Rubisco (Hatch, 1987; von Caemmerer and Furbank, 2003). An efficient segregation of these C_4 biochemical reactions requires specific leaf functions (Hattersley, 1984; Dengler *et al.*, 1994; Muhaidat *et al.*, 2007).

As a result of its multiple origins, C_4 photosynthesis does not present a consistent and discrete phenotype, so is better considered a functional trait involving a suite of coordinated leaf anatomical and biochemical characteristics (Brown and Smith, 1972; Laetsch, 1974). These components can assemble differently during each origin of C_4 photosynthesis, and these divergent evolutionary histories result in high anatomical and biochemical diversity among, and sometimes within, C_4 lineages (Hattersley and Watson,1992; Sinha and Kellogg,

Abbreviations: PEPC, phosphoenolpyurvate carboxylase; Rubisco, ribulose-1,5-bisphosphate carboxylase/oxygenase.

1996; Kadereit *et al.*, 2003; Muhaidat *et al.*, 2007; Edwards and Voznesenskaya, 2011; Freitag and Kadereit, 2014). An understanding of the evolutionary transitions leading to the recurrent assembly of C_4 photosynthesis requires investigation of the individual characteristics that together generate C_4 function, not only in C_4 species but also in C_3 species variously related to C_4 taxa (Christin and Osborne, 2013). It is particularly important to differentiate the present function of each component from its identity and developmental origin. In this work, we focus on the variation observed in both C_3 and C_4 plants in each of the anatomical traits that together generate leaf functions compatible with C_4 photosynthesis. We combine a review of the literature with analyses of a quantitative leaf anatomy dataset compiled from 155 C_3 and C_4 grass species (Christin *et al.*, 2013). The C_4 grasses in this dataset encompass eight of the nine structural C_4 forms described for this family (Edwards and Voznesenskaya, 2011).

What is C_4 leaf anatomy?

Differential arrangements of cells and organelles within the leaves of taxa that we now recognize as C_3 and C_4 were first observed and published more than 80 years before the C_4 pathway itself was discovered (Duval-Jouve, 1875; Haberlandt, 1884). The association between specific cell and organelle arrangements and the C_4 pathway was then identified soon after the discovery of C_4 photosynthesis (El-Sharkawy and Hesketh, 1965; Downton and Tregunna, 1968; Berry *et al.*, 1970; Welkie and Caldwell, 1970). Since then, C_4 photosynthesis has usually been affiliated closely with a suite of leaf properties referred to as 'Kranz' anatomy (after Haberlandt's description in German of a wreath-like arrangement of cells). Kranz anatomy can be described as two distinct concentric layers of chlorenchyma cells, formed by a bundle sheath containing most of the chloroplasts, surrounded by an outer layer consisting of a small number of mesophyll cells. The visual identification of such arrangements in transverse section has been used in numerous anatomical surveys of leaves to identify the photosynthetic pathway for hundreds of species (Welkie and Caldwell, 1970; Carolin *et al.*, 1973, 1975, 1977; Brown, 1977; Hattersley *et al.*, 1982; Renvoize, 1987*a*).

Surveys of numerous C_3 and C_4 species over the past five decades have shown that leaf anatomies cannot be easily and consistently grouped into discrete categories corresponding to the two photosynthetic types but come in many flavours (Brown, 1975; Edwards and Voznesenskaya, 2011). It is true that the leaf anatomy of a randomly selected C_3 plant is highly likely to deviate significantly from that of a randomly selected C_4 plant. For example, *Viburnum punctatum*, like most C_3 eudicots, has distinct horizontal layers of mesophyll cells in its leaves (Fig. 1A), arranged such that it does not conform to the general anatomical pattern generally present in C_4 plants, whereby the bundle-sheath and mesophyll cells form concentric circles around the vasculature (Fig. 1C). This concentric arrangement of cells can be found in many C_3 grasses though (Figs 1B and 2) (Hattersley *et al.*, 1982; Dengler *et al.*, 1994; Besnard *et al.*, 2013) and, as detailed below, individual leaf characteristics that are usually associated with a C_4 function can be found in at least some C_3 plants. Furthermore, some plants achieve C_4 photosynthesis without the segregation of photosynthetic reactions into different types of cells (Bowes and Salvucci, 1984; Bowes and Salvucci, 1989; Freitag and Stichler, 2000; Edwards *et al.*, 2004). Despite this variation, C_4 physiology is still associated with a suite

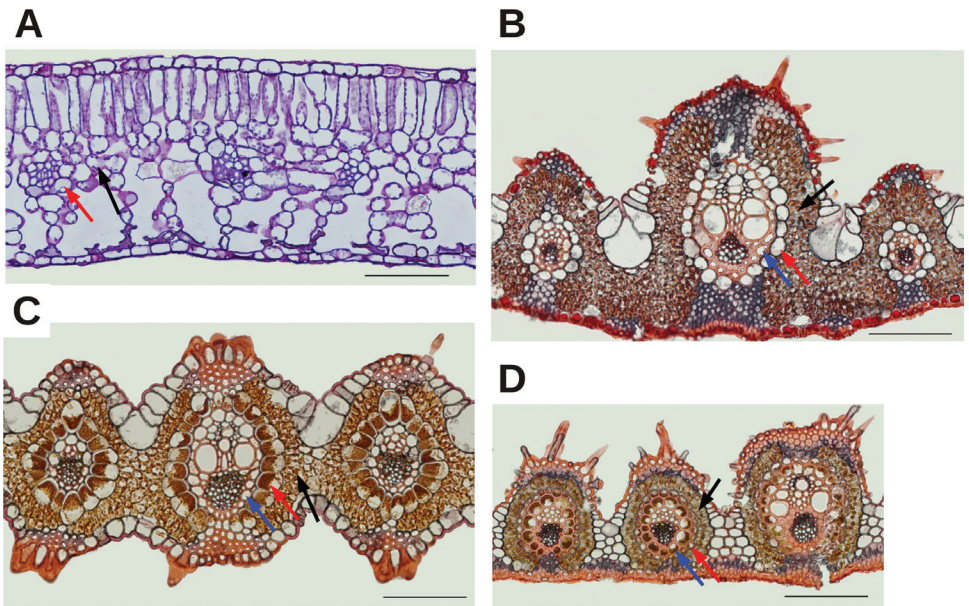


Fig. 1. Examples of C_3 and C_4 leaf cross-sections. The C_3/C_4 pair on the left (A, C) are unrelated, belonging to different major groups of flowering plants. By contrast, the C_3/C_4 pair on the right (B, D) is composed of closely related species, belonging to the same subfamily of grasses. (A) *Viburnum punctatum* (C_3, Adoxaceae), (B) *Sartidia angolensis* (C_3, Poaceae), (C) *Centropodia mossamedensis* (C_4, Poaceae), and (D) *Aristida mollissima* (C_4, Poaceae). Black arrows indicate the mesophyll, red arrows the outer bundle sheath, and blue arrows the inner sheath of grasses (=mestome sheath). The four cross-sections are shown at the same scale. Bars, 100 μm. Picture (A) was kindly provided by Dr David Chatelet from Brown University and pictures (B), (C) and (D) come from the collections of Professor J. Travis Columbus from Rancho Santa Ana Botanic Garden, CA, USA, with permission.

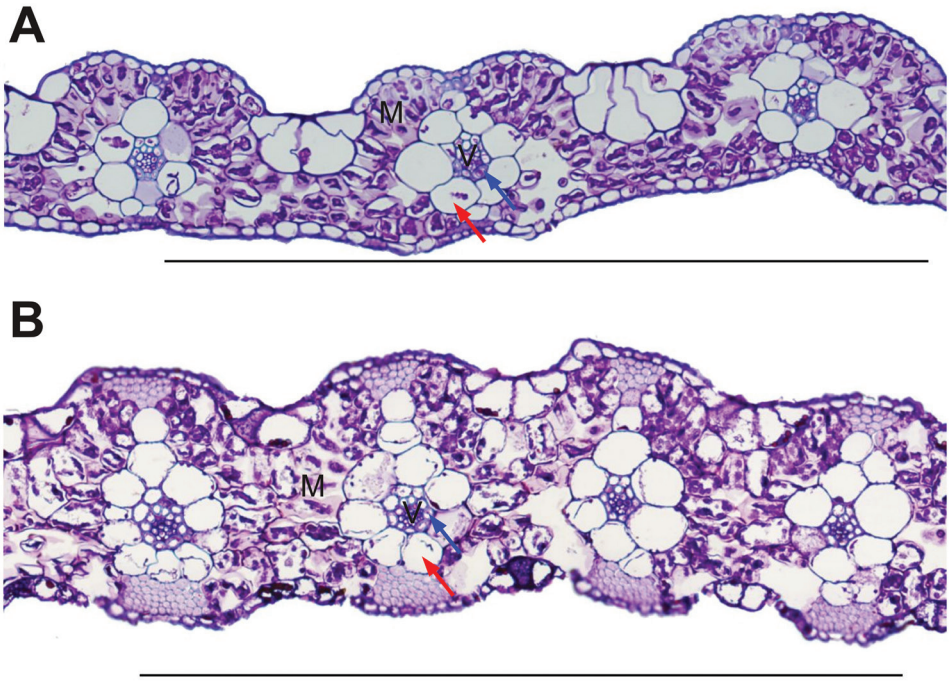


Fig. 2. Examples of C_3 grasses with leaf anatomy close to the C_4 requirements. (A) *Panicum pygmaeum* (C_3), (B) *Panicum malacotrichum* (C_3). The mesophyll (M) and vascular tissue (V) are indicated on the sections. Red arrows indicate the outer bundle sheath while blue arrows indicate the inner sheath (=mestome sheath). Bars, 500 μm.

of functional properties (Brown and Smith, 1972; Edwards and Voznesenskaya, 2011), which must first be considered before analysing diversity in the identity and developmental origins of the characteristics that generate them. Based on the literature, the following functional properties of leaves are considered essential requirements for C_4 photosynthesis (Hattersley *et al.*, 1977; Leegood, 2002; von Caemmerer and Furbank, 2003; Edwards and Voznesenskaya, 2011; Nelson, 2011). Note that these apply equally to all C_4 plants, whether or not they use distinct types of cells.

1. There must be two distinct compartments arranged so that atmospheric gases reach the first compartment more easily than the second. The first compartment houses the PEPC reactions, while the second, with characteristics that restrict CO_2 efflux, houses the Calvin cycle.
2. The two compartments must be in close contact to allow the rapid exchange of metabolites.
3. The compartment where the Calvin cycle occurs must occupy a large enough fraction of the leaf to accommodate a significant number of chloroplasts.
4. Chloroplasts must be abundant in the Calvin cycle compartment.

These functional properties are extremely important for C_4 physiology and biochemistry. However, to understand the gradual evolutionary changes leading to the recurrent assembly of C_4 photosynthesis, it is important to account for exact changes in cellular characteristics and the genetic determinants of these characteristics. In the following sections, we therefore discuss how each of the four functional properties listed above is generated from underlying characteristics. We look at how these characteristics vary qualitatively and quantitatively among C_3 and C_4 lineages, and show how there is an overlap between the values observed in C_3 and C_4 species.

Two compartments differentially connected to the atmosphere

In C_3 plants, the Calvin cycle occurs in most of the leaf, while it is restricted to specific locations in C_4 plants. It is well known that the identity of the compartments co-opted for the segregation of the atmospheric CO_2 fixation by PEPC and its refixation by the Calvin cycle differs among C_4 origins (e.g. Brown, 1975; Dengler *et al.*, 1985). For instance, some single-celled C_4 species have evolved separate compartments for the PEPC and Calvin cycle reactions through the rearrangements of organelles or vacuoles within individual photosynthetic cells (Edwards *et al.*, 2004). In the majority of C_4 plants, however, the PEPC and Calvin cycle reactions are segregated in different types of cells. In C_3 species, the mesophyll and bundle sheath represent two physiologically distinct types of cells, and the central position of bundle-sheath cells within the leaf gives the opportunity for minimal contact with the atmosphere (Figs 1A, B and 3, and Supplementary Fig. S1 available at *JXB* online). The bundle sheaths have consequently been co-opted for Calvin cycle reactions across most C_4 origins, while the mesophyll cells, which are better connected to the atmosphere, are used for the PEPC reactions. Despite this convergence in function, the bundle-sheath cells recruited for C_4 photosynthesis are not homologous among all C_4 origins.

In some C_4 species within the grass genera *Arundinella*, *Garnotia*, *Arthropogon*, *Achlaena*, *Dissochondrus*, *Anrthraxon*,

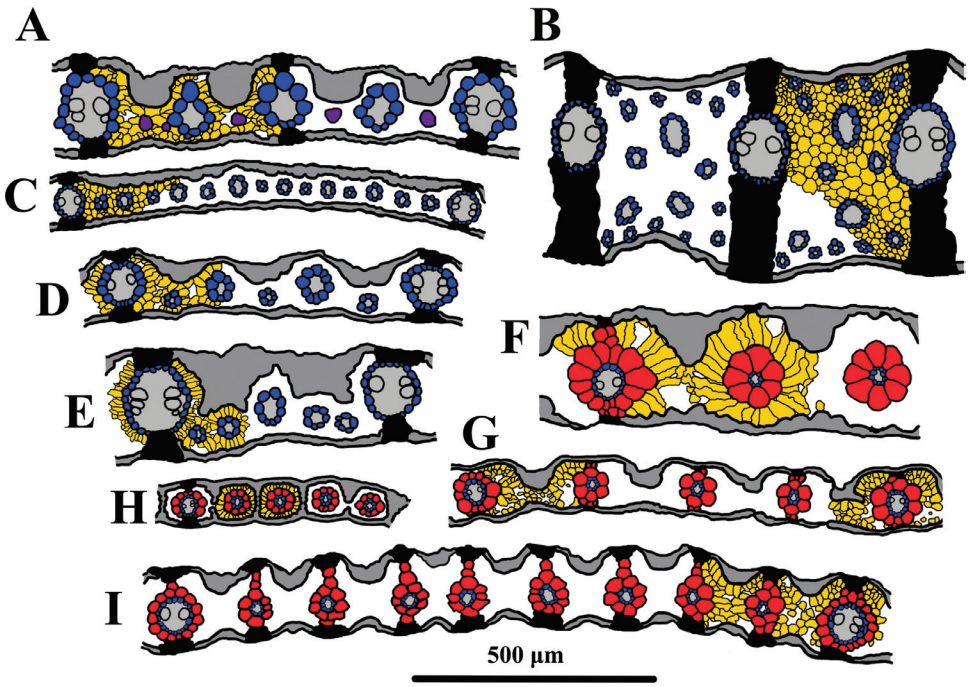


Fig. 3. Leaf anatomy for selected cross-sections of grasses. (A) *Arundinella nepalensis* (C_4), (B) *Anthaenantia lanata* (C_4), (C) *Axonopus compressus* (C_4), (D) *Ischaemum afrum* (C_4), (E) *Chrysopogon pallidus* (C_4), (F) *Alloteropsis cimicina* (C_4), (G) *Panicum pygmaemum* (C_3), (H) *Bouteloua stolonifera* (C_4) and (I) *Panicum malacotrichum* (C_3). The diagrams highlight the mesophyll cells (yellow), outer bundle sheaths (red), inner bundle sheaths (blue), and distinctive cells (purple). Uncoloured central areas are composed of mesophyll cells and intercellular airspace. Vein (light grey), epidermis (dark grey), and sclerenchymatous girders (solid black) are also shown. Where only one bundle sheath is present, it is assumed that the outer bundle sheath has been lost and the inner bundle sheath remains. All cross-sections are drawn at the same scale, indicated at the bottom. The corresponding pictures can be found in Supplementary Fig. S1 available at JXB online.

and *Microstegium*, the Calvin cycle also occurs in distinctive cells, which are atypical bundle-sheath-like cells, differentiated within the mesophyll but not associated with vascular bundles (Fig. 3A) (Tateoka, 1958; Hattersley and Watson, 1992; Ueno, 1995; Dengler *et al.*, 1996; Wakayama *et al.*, 2003). In addition, grasses and sedges possess multiple layers of sheath cells, with inner layers derived from procambium (often referred to as the 'mestome sheath') and outer layers from ground meristem (Dengler *et al.*, 1985; Soros and Dengler, 2001; Martins and Scatena, 2011). In studies of C_4 photosynthesis, consideration of the different cells is often based on their function. However, for evolutionary studies, the ontogenic origin of each type of cell needs to be established independently of its function. The C_4 lineages within grasses and sedges have alternatively co-opted one or both of these cell types, while the second cell layer is often lost, for example in the numerous C_4 grasses with a single sheath layer (Fig. 3A–E) (Brown, 1975; Dengler *et al.*, 1996; Soros and Dengler, 2001; Martins and Scatena, 2011). This diversity in the identity of the two compartments co-opted for the segregation of C_4 reactions, together with phylogenetic analyses, has been used previously to argue for multiple independent C_4 origins, rather than fewer origins followed by reversals in closely related C_3 species (Kellogg, 1999; Christin *et al.*, 2010).

The limited connection of the Calvin cycle compartment to the atmosphere is also achieved via different mechanisms in the different C_4 lineages. First, tightly packing mesophyll cells around the bundle sheath reduces the fraction of cells from the latter that are in contact with the atmosphere (Dengler *et al.*, 1994; Muhaidat *et al.*, 2007), although similar packing also occurs in some C_3 grasses (Fig. 1B) (Dengler *et al.*, 1994) and some C_3 eudicots (Muhaidat *et al.*, 2007). In addition, the bundle-sheath cell walls can also be covered with a layer of suberin, which limits gas diffusion. This is the case in C_4 monocots that have co-opted the inner sheath layer for a C_4 function (Hattersley and Browning, 1981; Ueno *et al.*, 1988*b*). However, the presence of suberin layers on the inner sheath cell walls can also be found in most C_3 grasses (Hattersley and Browning, 1981). Neither of the characteristics reducing contact of the Calvin cycle with the atmosphere is therefore found exclusively in C_4 plants.

Distance between the two compartments

Close contact between the PEPC and Calvin cycle compartments is guaranteed in plants with a single-celled C_4 system. In plants with a dual-celled C_4 system, the presence of mesophyll cells not directly adjacent to the bundle sheaths will increase the average distance between the compartments containing PEPC and Rubisco. This problem is usually solved in C_4 plants by limiting the number of cells separating consecutive Calvin cycle compartments, and by organizing mesophyll cells into one or two layers around the bundle sheath (Fig. 1C, D), which produces the classical pattern of Kranz anatomy. In some species, this configuration is achieved through the development of a

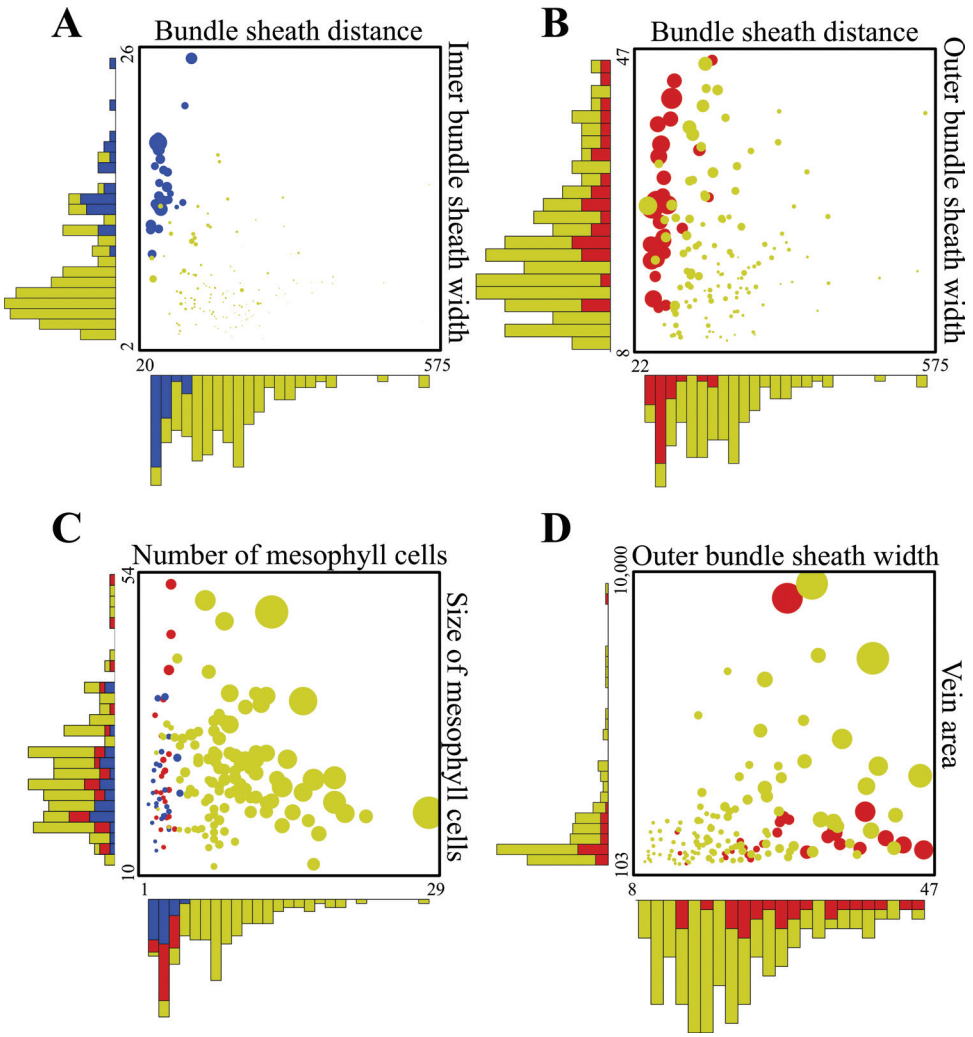


Fig. 4. Multidimensionality of C_4 anatomy in grasses. Scatter plots for anatomical variables associated with the C_4 syndrome are shown, along with frequency distributions for each trait, arranged along the axes. For each pair of variables, dot size is proportional to a third variable. C_3 grass species are shown in yellow, C_4 grass species using the outer sheath for the Calvin cycle in red, and C_4 grass species using the inner sheath for the Calvin cycle in blue. Relationships are shown between means of: (A) distance between consecutive bundle sheaths (μm) and inner bundle-sheath cell width (μm), with dot size proportional to the percentage of inner bundle-sheath area; (B) distance between consecutive bundle sheaths (μm) and outer bundle-sheath cell width (μm), with dot size proportional to the percentage of outer bundle-sheath area; (C) number of mesophyll cells between consecutive bundles and mesophyll cell length (μm), with dot size proportional to the distance between consecutive bundle sheaths (μm); and (D) outer bundle sheath cell width (μm) and area of vasculature (μm^2), with dot size proportional to the outer bundle sheath area (μm^2) per vein number. The data for 170 grasses (representing 155 species) come from Christin *et al.* (2013).

single bundle-sheath layer that encompasses all the vasculature within the leaf and often water-storage cells as well, and a single layer of mesophyll that surrounds the bundle sheath. Variations on this anatomical theme are common among C_4 eudicots and have been found in the Asteraceae, Amaranthaceae, and Cleomaceae families (Carolin *et al.*, 1975; Das and Raghavendra, 1976; Kadereit *et al.*, 2003; Peter and Katinas, 2003; Edwards and Voznesenskaya, 2011; Koteyeva *et al.*, 2011). Some C_4 grasses have similar bundle sheaths that extend horizontally from the vascular tissue and join together, such that the mesophyll becomes isolated in small patches (Renvoize, 1983).

For C_4 lineages with multiple photosynthetic units formed by concentric cell layers of mesophyll, bundle sheath, and vascular tissue, the presence of fewer mesophyll cells between consecutive veins can be achieved via two different developmental mechanisms. First, the number of cells that develop between consecutive bundle sheaths can be directly reduced during ontogeny. Second, extra Calvin cycle compartments, such as distinctive cells or minor veins, can be added to decrease the average distance between compartments, as has been documented in both monocots (e.g. Poaceae; Fig. 3A–E; Renvoize, 1987*a*; Dengler *et al.*, 1994; Ueno *et al.*, 2006; Christin *et al.*, 2013) and eudicots (e.g. Asteraceae; McKown and Dengler, 2007; McKown and Dengler, 2009; Cleomaceae; Marshall *et al.*, 2007).

Interveinal distance (or vein density) is often considered a proxy for the number of mesophyll cells between consecutive bundles, and largely overlaps between C_3 and C_4 grasses (Christin *et al.*, 2013) and eudicots (Muhaidat *et al.*, 2007). However, the relationship between interveinal distance and the number of mesophyll cells is only partial. First, because

interveinal distance is influenced both by the diameter of the veins and the size of the bundle sheaths, measuring the actual distance between bundle sheaths is more relevant. This distance is influenced by the size of individual mesophyll cells, their orientation, and finally their number (Fig. 4C). Some C_4 species, such as *Alloteropsis cimicina*, have relatively large interveinal distances but with only a few large mesophyll cells between consecutive bundles (Fig. 3F, 4C). In addition, the number of mesophyll cells containing PEPC below and above veins can influence the average distance between the PEPC and Calvin cycle reactions independently of the distance between consecutive bundles. Some thick C_4 leaves, such as those of *Anthaenathia lanata* (Fig. 3B) or some *Portulaca* (Ocampo *et al.*, 2013), consequently require a three-dimensional venation system. Finally, leaf thickness is often reduced between veins so that there are few mesophyll cells in positions most distant from the bundle sheaths, and interveinal distance can greatly exceed the average distance between photosynthetically active mesophyll cells and bundle-sheath cells (Figs 1 and 3). For instance, in leaves of the C_3 grass *Panicum pygmaemum*, the average number of mesophyll cells between bundles greatly exceeds four. However, because its leaf thickness decreases between veins, the number of mesophyll cells separated from the bundle sheath by more than one cell is smaller than the number of mesophyll cells separated from the bundle sheath by zero or one cell (38 versus 73 cells between the three veins in Fig. 2). Finally, the distance between consecutive bundles can be increased by the presence of achlorophyllous cells that do not influence the average path length from PEPC to Calvin cycle cells (e.g. Fig. 1D).

The number of mesophyll cells between consecutive bundles will distinguish C_3 from C_4 taxa with a high success rate and has consequently been proposed as a criterion to recognize C_4 plants (Hattersley and Watson, 1975; Renvoize, 1987*a*; Sinha and Kellogg, 1996). However, the C_3 and C_4 distributions for this trait also overlap (Fig. 4C). For instance, *Panicum malacotrichum* is a C_3 grass with less than four mesophyll cells between veins (Fig. 2). The variation observed in both C_3 and C_4 taxa is probably due to the importance of vascular architecture for both photosynthetic types. While the distance between consecutive bundles affects the efficiency of C_4 photosynthesis (Ogle, 2003), vein density also influences the transport of metabolites, leaf hydraulics and other physiological characteristics in C_3 plants (Sack and Scoffoni, 2013; Sack *et al.*, 2013). In summary, both interveinal distance and the number of mesophyll cells between consecutive bundles overlap in C_3 and C_4 taxa, so that C_4 values represent only a subset of those observed among all photosynthetic types (Fig. 4A–C) (Muhaidat *et al.*, 2007; Christin *et al.*, 2013).

The transport of metabolites between the PEPC and Calvin cycle compartments in C_4 plants is also facilitated by a number of plasmodesmata connecting mesophyll and bundle-sheath cells that exceeds the number found in C_3 plants (Olesen, 1975; Weiner *et al.*, 1988; Botha, 1992). However, plasmodesmata frequency is known in only a few C_3 species, so the overall variation in this trait cannot be established with confidence.

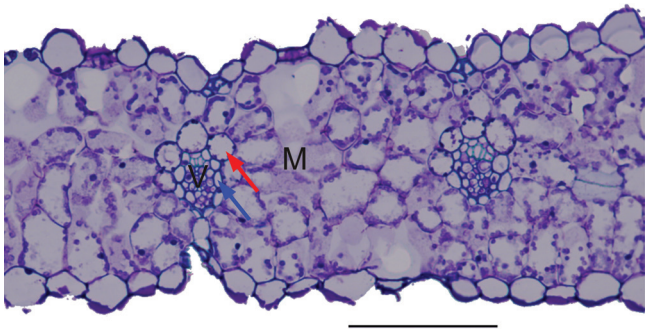


Fig. 5. Detail of a cross-section for *Dactylis glomerata*. The mesophyll (M) and vascular tissue (V) are indicated on the section of this C_3 species. The red arrow Indicates the outer bundle sheath, while the blue arrow indicates the inner sheath (=mestome sheath). Bar, 100 μm. Note the incomplete outer sheath.

Large Calvin cycle compartment

The amount of CO_2 that can be re-fixed by Rubisco in the Calvin cycle will depend on the number of chloroplasts within the compartment co-opted for this function. The size of this compartment, not including the volume occupied by the vacuole, will influence the number of chloroplasts that can be accommodated. Thus, C_4 plants tend to have enlarged bundle-sheath cells able to accommodate numerous chloroplasts. More than the size of individual bundle sheath cells, the cumulative volume of bundle sheath relative to the PEPC compartment (mesophyll) is relevant, and seems to be constrained within a given range in C_4 plants (Hattersley, 1984; Dengler *et al.*, 1994; Muhaidat *et al.*, 2007). This might represent a trade-off between having sufficient chloroplasts in the Calvin cycle compartment and still conserving enough mesophyll volume for PEPC.

Similar bundle sheath:mesophyll ratios can be achieved through different combinations of the numerator (volume of bundle sheath) and denominator (volume of mesophyll). For instance, similar proportions of bundle sheath can be achieved through alternative developmental mechanisms, involving the production of either larger or more numerous bundle-sheath cells (the latter is generally achieved through a proliferation of veins; Fig. 3) (Hattersley, 1984; McKown and Dengler, 2009). The cross-sectional area of mesophyll per vein is mainly a function of the distance between veins, the thickness of the leaf (including the thickness between veins in comparison to that at the veins) and the presence of achlorophyllous cells (Christin *et al.*, 2013). On the other hand, when viewed in transverse section, the total area of a given type of bundle sheath per vein is a function of the size of the bundle-sheath cells, the diameter of the veins, and, in some cases, the completeness of the bundle sheath (Fig. 4) (Christin *et al.*, 2013). For instance, the external bundle sheath of many grasses is not developed on the abaxial side of the leaf, which reduces the total volume of this tissue (Fig. 5) (e.g. Renvoize, 1985, 1987*b*). Thus, the relative amount of bundle-sheath tissue is a function of at least five distinct traits, which may all vary independently. Functionally similar characteristics can consequently arise through different developmental modifications, as highlighted by the diversity of C_4 leaf anatomy (Fig. 4).

The five components that dictate the relative amount of bundle-sheath tissue are important determinants of the gross leaf anatomy associated with C_4 photosynthesis. However, each component shows an essentially continuous distribution across C_3 and C_4 values, such that C_4-compatible ranges merely represent a subset of the distribution found in C_3 taxa (Fig. 4; Marshall *et al.*, 2007; McKown and Dengler, 2007). The C_4-suitability of one parameter depends on the values of the other parameters. For instance, large volumes of bundle-sheath tissue can arise in the presence of significant distances between consecutive bundles if the bundle-sheath cells are enlarged (Fig. 4A, B). This is highlighted by a comparison of *Alloteropsis cimicina* and *Axonopus compressus* (Fig. 3F and C, respectively), which achieved similar ratios of bundle sheath per mesophyll area [BS/(BS+M) of 0.26 and 0.21, respectively] through different means. *Alloteropsis cimicina* has very large outer bundle sheaths that are separated by long distances of mesophyll, while *Axonopus compressus* has small inner sheaths that are separated by very short mesophyll distances in particularly thin leaves (Fig. 3F and 3C, respectively).

During the course of evolution, numerous alterations in the characteristics that generate each leaf function occur either stochastically or in response to selective pressures. For instance, leaf thickness often represents an adaptation to the amount of light received by plants (Boardman, 1977; Terashima *et al.*, 2001). The number and size of veins alters the hydraulics of a plant, which, in turn, affects the sorting of plants across environments (McKown *et al.*, 2010; Sack *et al.*, 2012). Finally, the bundle sheath controls water flux between the mesophyll and vascular tissue such that an increase in bundle-sheath size might provide better protection against cavitation in arid environments (Sage, 2001; Leegood, 2008; Griffiths *et al.*, 2013). Recurrent and independent changes in different leaf properties repeatedly led to the emergence of tissues suitable for C_4 photosynthesis, which characterize numerous extant C_3 plants (Muhaidat *et al.*, 2007; Edwards and Voznesenskaya, 2011; Muhaidat *et al.*, 2011; Kadereit *et al.*, 2012; Christin *et al.*, 2013; Griffiths *et al.*, 2013).

Distribution of organelles

One of the most important requirements for C_4 photosynthesis probably lies in the distribution of chloroplasts. Although they are present in all photosynthetic cells of C_3 plants, chloroplasts are especially abundant in mesophyll cells and can vary from equally abundant to completely absent in bundle-sheath cells (Figs 1, 2 and 5) (Crookston and Moss, 1970). In C_4 plants, the light-dependent and light-independent functions of chloroplasts are often decoupled, and chloroplasts of the PEPC and Calvin cycle compartments can become morphologically and functionally differentiated (Woo *et al.*, 1970; Laetsch, 1974; Hattersley *et al.*, 1977; Bowman *et al.*, 2013). Although the characteristics and distribution of organelles vary among C_4 lineages (Ueno *et al.*, 1988*b*; Voznesenskaya *et al.*, 2006; Edwards and Voznesenskaya, 2011), the Calvin cycle compartment of C_4 plants consistently has a high concentration of chloroplasts, where the enzymes of the Calvin cycle are preferentially expressed.

No quantitative census of chloroplast distribution is available for randomly selected plants; however, the organelle distribution has been investigated in species closely related to C_4 lineages, which shows that some plants maintain significant numbers of chloroplasts in bundle-sheath cells, despite lacking a functional C_4 pathway (Hattersly *et al.*, 1982; Ueno and Sentoku, 2006; Christin *et al.*, 2013). This is particularly common in plants using C_2 photosynthesis, a weak CO_2-concentrating mechanism based on a glycine shuttle from mesophyll to bundle-sheath cells (Edwards and Ku, 1987; Sage *et al.*, 2012). When chloroplast abundance in bundle-sheath cells is compared among taxa, there is a gradient from closely related C_3 to C_2, and then from C_2 to C_4 species (Muhaidat *et al.*, 2011; Sage *et al.*, 2013). The C_2 trait is consequently often considered an evolutionary intermediate between C_3 and C_4 types (Hylton *et al.*, 1988; Sage *et al.*, 2012; Williams *et al.*, 2013). Therefore, as for other anatomical traits, the number of chloroplasts in bundle-sheath cells varies and may form a continuum between C_3 and C_4 species. Despite this, a high concentration of chloroplasts in bundle-sheath cells might be the only trait that occurs systematically within dual-celled C_4 photosynthesis that is never present in non-C_4 plants. The tight association between C_4 physiology and chloroplast distribution is explained by the fact that C_4 physiology results from a differential distribution of the Calvin cycle (among other biochemical reactions), which is usually linked to the distribution of chloroplasts.

Other ultrastructural properties associated with some C_4 plants include the distribution of mitochondria and peroxisomes among compartments, the distribution of organelles within compartments and the ultrastructure and photochemical properties of the chloroplasts (Bruhl and Perry, 1995; Edwards and Voznesenskaya, 2011). Some of these properties are also observed in non-C_4 species closely related to C_2 and C_4 taxa (Sage *et al.*, 2012)

Plasticity for C_4-suitable anatomy

Phenotypic plasticity to environmental cues creates an additional layer of variation and further blurs the dichotomy between C_4 and non-C_4 anatomy. Specifically, plasticity for the anatomical traits relevant to photosynthesis (e.g. compartmentalization, interveinal distance, mesophyll cell size and number, bundle-sheath cell size, and organelle distribution) could partially explain the variation found in these anatomical characteristics or, more importantly, the shift of C_3 plants into the C_4-suitable space. Plasticity for these traits has been documented in the literature. For example, the C_3 grass *Phragmites australis* acquires C_4-like traits when it grows at low soil water potentials (Gong *et al.*, 2011). Specifically, interveinal distance decreases, chlorophyll content within bundle-sheath cells increases, and the activity of C_4-related enzymes increases as soil water potential becomes more negative across a natural precipitation gradient (Gong *et al.*, 2011). The C_4-like *Flaveria brownii* lacks the complete suite of anatomical characteristics required for a fully functioning C_4 system (Araus *et al.*, 1990). However, this species can plastically increase its degree of C_4 photosynthesis by nearly doubling its investment in

bundle-sheath tissue relative to mesophyll in response to high irradiance compared with when it is grown at low irradiance (1200 vs 80 μmol m^{-2} s^{-1} photosynthetic photon flux density, respectively; Araus *et al.*, 1991). Furthermore, interveinal distances decreased in the C_3 grasses, *Festuca arundinaceae* (43% decrease), and the C_3/C_4 intermediate grass *Panicum miliodies* (34% decrease), when grown in high versus low nitrogen levels (Bolton and Brown, 1980).

In addition to the plasticity of individual anatomical components, two different modes of environmentally induced C_4 photosynthesis exist. First, several aquatic species of Hydrocharitaceae, and possibly some Alismataceae and Cyperaceae, are able to switch from C_3 to single-cell C_4 photosynthesis (Bowes *et al.*, 2002). The environmental cue for this plasticity may be exposure to low-CO_2 conditions as they become submerged under water, or seasonal variation in temperature (Bowes *et al.*, 1979; Bowes, 2011). In contrast, some aquatic *Eleocharis* species use C_3 or C_3/C_4 intermediate photosynthesis when submerged but induce C_3/C_4 or C_4 photosynthesis by developing C_4-compatible leaf anatomy and expressing C_4 enzymes in the emergent leaves (Ueno *et al.*, 1988*a*; Ueno, 2001; Murphy *et al.*, 2007). Finally, some amphibious C_4 grasses seem to switch from a C_4 system that functions without C_4-associated leaf anatomy in aquatic leaves to a classical dual-cell C_4 cycle in aerial leaves (Keeley, 1998; Boykin *et al.*, 2008).

Phenotypic plasticity for C_4-associated traits might have important implications for the evolution of C_4 photosynthesis (Sultan, 1987; West-Eberhard *et al.*, 2011). First, the direction and degree of phenotypic change in response to an environmental gradient is heritable (Schlichting and Levin, 1986; Schlichting and Pigliucci, 1993), and the reaction norm for a trait is genetically distinct from the trait itself. Selection can therefore act independently on both a trait and on the plasticity for that trait. Plasticity may thus deter the evolutionary transition from C_3 to C_4 photosynthesis by diluting the effects of natural selection. However, adaptive phenotypic plasticity may promote C_4 evolution if the plastic expression of C_4-suitable anatomical traits in C_3 plants allows the colonization of new niches, leading to selective pressures for the gradual acquisition of C_4 biochemistry (Heckmann *et al.*, 2013). Indeed, Sage and McKown (2006) reviewed the literature to find that C_3 plants seem to be inherently more plastic than C_4 plants overall. Thus, this capacity for phenotypic plasticity might affect the probability of evolving C_4 photosynthesis. For instance, differential capacity in the phenotypic plasticity for important C_4 anatomical traits among plant lineages may explain the differential propensity for C_4 evolution. However, the plasticity of anatomical traits associated with C_4 photosynthesis remains mostly unknown in C_3 species, and more comparative work is required.

Consequences for the evolution of C_4-associated anatomy

When comparing the anatomy of a randomly selected C_3 taxon with that of a highly efficient C_4 species, the evolutionary transition from C_3 to C_4 anatomy can seem extraordinary (Fig. 1A, C). However, it is important to note that C_4 photosynthesis did not emerge from the average C_3 taxon but from C_3 ancestors with leaf anatomical properties much closer to the C_4 requirements (Figs 1B and 5) (Muhaidat *et al.*, 2011; Christin *et al.*, 2013; Sage *et al.*, 2013). In the Poaceae, some species apparently using the C_3 photosynthetic type have gross leaf anatomies that closely resemble those of C_4 plants. For instance, *Panicum malacotrichum* and *Panicum pygmaeum* (Fig. 2) are two C_3 grasses ($\delta^{13}C$ values of –27.4 and –29.7, respectively), which are closely related to several C_4 lineages (namely *Alloteropsis* and *Echinochloa*; Grass Phylogeny Working Group II, 2012). These species possess large proportions of bundle-sheath tissue that are firmly in the C_4 range [BS/(BS+M) of 0.26 and 0.23, respectively; Christin *et al.*, 2013], and most mesophyll cells are directly adjacent to the bundle sheath or separated by only one mesophyll cell (Fig. 2). Chloroplasts are still almost completely restricted to the mesophyll in these species. However, because the gross leaf anatomy is in place, fewer anatomical changes are necessary for the evolution of C_2 or C_4 photosynthesis. In other cases, such as the grass tribe Neurachninae, C_3 species that are closely related to C_4 species have both C_4-like gross anatomy [BS/(BS+M) of 0.14–0.16; Christin *et al.*, 2013] and the presence of conspicuous chloroplasts in the inner sheath, which was co-opted for C_4 photosynthesis in this group (Hattersley *et al.*, 1982). These examples show that the evolution of C_4-suitable anatomy might not always require drastic modifications, as C_3 lineages may possess C_4-like values for individual traits that can generate C_4 leaf functions.

Each component of C_4-compatible leaf anatomy may vary independently within C_3 ancestors, such that any combination of mesophyll cell size, bundle-sheath cell size, leaf thickness and interveinal distance could theoretically occur. However, the observed range is obviously more limited (Fig. 4), for a number of reasons. First, multiple traits might be influenced by the same gene (pleiotropy). For instance, genome size theoretically affects the size of all cells (Grime and Mowforth, 1982; Masterson, 1994; Beaulieu *et al.*, 2008; Šímová and

Table 1. *Degrees of co-variation among anatomical variables*

Co-variation in grasses between the mean distance between consecutive bundle sheaths (μm), outer bundle-sheath (OBS) cell width (μm), inner bundle-sheath (IBS) cell width (μm), number of mesophyll (M) cells between consecutive bundles, and leaf thickness (μm). R^2 values are provided for pairs of variables with significant correlations. Regressions with P values less than 0.05 are considered significant, while those with P values greater than 0.05 are indicated by NS. Phylogenetically controlled analyses were performed with the pgls function of the caper R package (Orme et al., 2012), using the data for 155 grass species from Christin et al. (2013).

BS distance				
0.02	OBS cell width			
NS	0.23	IBS cell width		
0.58	NS	0.05	No. M cells	
NS	0.28	0.43	0.12	Leaf thickness

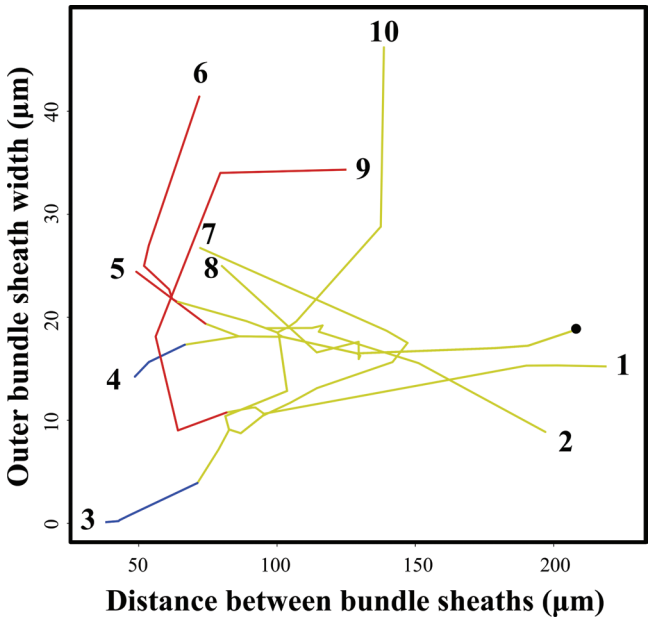


Fig. 6. Evolutionary trajectories toward C_4-compatible anatomical traits. Phylogenetic relationships are plotted in anatomical space for grass species selected to represent a diversity of anatomical traits. Values are the distance between consecutive bundle sheaths, and the width of outer bundle-sheath cells, which are observed for the tips and inferred for the internal nodes. The black point represents the root of the tree (see Christin *et al.*, 2013, for details). Yellow branches indicate a C_3 state, red branches a C_4 state using the outer sheath for the Calvin cycle, and blue branches a C_4 state using the inner sheath for the Calvin cycle. Numbers refer to the extant species: 1, *Dichanthelium acuminatum* (C_3); 2, *Danthonia spicata* (C_3); 3, *Heteropogon contortus* (C_4); 4, *Aristida congesta* (C_4); 5, *Stipagrostis obtusa* (C_4); 6, *Eleusine indica* (C_4); 7, *Panicum malacotrichum* (C_3); 8, *Oryza coarctata* (C_3); 9, *Panicum miliaceum* (C_4), 10, *Arundo donax* (C_3). Data from Christin *et al.* (2013).

Herben, 2012), so that an increase in bundle-sheath cell size might co-occur with increases in the sizes of mesophyll cells. Plants often escape this constraint via cell-specific endoreduplication, which allows an increase of one type of cell relative to others (Sugimoto-Shirasu and Roberts, 2003), and comparative analyses show that variation in the cell sizes of different components of C_4 anatomy is only partially correlated (Table 1). However, endoreduplication is not involved in the increase of bundle-sheath cell size, at least in the C_4 *Cleome gynandra* (Aubry *et al.*, 2013). It is also likely that some combinations of traits are not viable, as the whole-leaf structure influences plant fitness (Noblin *et al.*, 2008), not its individual components.

The multidimensionality of leaf characteristics associated with C_4 photosynthesis, as highlighted for the grass family, means that different combinations of underlying traits will generate C_4-compatible leaf anatomies (Fig. 4). For instance, both a proliferation of veins with small bundle-sheath cells and an increase of bundle-sheath cell size without additional veins would increase the relative amount of bundle-sheath cells (Fig. 6). This potential for alternative anatomical combinations to achieve the same functional outcome means that C_3 ancestors will repeatedly reach C_4-compatible areas of the multidimensional trait space (Fig. 6), and increases the likelihood of C_4 anatomy evolving (Williams *et al.*, 2013).

A sample of evolutionary trajectories in the Poaceae shows lineages for which repeated and independent alterations of the distance between bundle sheaths and bundle-sheath size led into different C_4-compatible regions of the anatomical space (Fig. 6). Obviously, not all C_3 lineages that acquired C_4-suitable leaf anatomical characteristics have evolved C_4 biochemistry. For example, *Panicum malacotrichum* and *Oryza coarctata* have C_4-suitable mesophyll distances between consecutive bundle sheaths and proportions of bundle-sheath tissue but have not developed the C_4 syndrome (Figs 2 and 6) (Christin *et al.*, 2013). Furthermore, *Cleome violacea*, *C. africana*, and *C. paradoxa* have small interveinal distances, and *C. africana* and *C. paradoxa* also display enlarged bundle-sheath cells similar to their C_4 congener *C. gynandra*, yet these three species do not employ the C_4 photosynthetic system (Marshall *et al.*, 2007). However, the presence of these characteristics probably enables C_4 evolution (pre-adaptation or exaptation *sensu* Gould and Vrba, 1982; Christin *et al.*, 2013; Griffiths *et al.*, 2013; Sage *et al.*, 2013). Once a C_4-compatible anatomy is in place, the C_4 biochemical pathway can evolve from a C_3 background in a stepwise sequence, where each step incrementally increases the efficiency of photosynthesis (Heckmann *et al.*, 2013). However, the multiple anatomical requirements for C_4 photosynthesis do not usually co-occur in C_3 plants. Interesting exceptions include plants with a C_2 physiology, which were probably co-opted for the evolution of C_4 photosynthesis (Christin *et al.*, 2011; Muhaidat *et al.*, 2011; Sage *et al.*, 2012).

Functional C_4 diversity as a consequence of evolutionary diversity

Because C_4-compatible leaf anatomy engages multiple components, each C_4 origin may involve different modifications and co-opt different compartments for the Calvin cycle (Brown, 1975; Dengler *et al.*, 1994; Edwards and Voznesenskaya, 2011; Christin *et al.*, 2013). The anatomy present in the C_3 ancestor might affect which C_4 phenotypes are possible. For instance, C_3 ancestors with enhanced water storage tissue are likely to give rise to C_4 leaves that maintain the same capacity to store water, with the PEPC and Calvin cycle compartments occupying other parts of the leaves (Voznesenskaya *et al.*, 1999; Kadereit *et al.*, 2003; Freitag and Kadereit, 2014). Similarly, C_4 species that use the inner bundle sheath for the Calvin cycle must evolve from C_3 ancestors that possessed two differentiated sheaths, as is the case with grasses and sedges (Dengler *et al.*, 1994; Soros and Dengler, 2001). Furthermore, C_4 phenotypes that are functionally similar can be achieved through different modifications, even when starting with similar C_3 ancestors.

Different modifications to fulfil the same C_4 requirements might have functional consequences. Indeed, the adaptation of C_4 photosynthesis through the evolution of thick leaves with large bundle-sheath cells (Fig. 3F) is likely to have different consequences from the evolution of thin leaves with small cells but very short interveinal distance (Fig. 3C). An increase in vein density will affect not only the hydraulics but

also the distribution of stomata, which tend to be located in between veins (Taylor *et al.*, 2012). Leaf thickness will have consequences for light-capture efficiency as well as ecologically meaningful traits such as specific leaf area (Wilson *et al.*, 2002). Similarly, light capture will also be affected by the different distribution of chloroplasts in mesophyll and bundle-sheath cells, and the relative abundance of each cell type, together with the orientation of mesophyll cells (Vogelmann *et al.*, 1996). The path length from stomata to the photosynthetically active cells will also be influenced by leaf thickness, interveinal distance, and amount of intercellular airspace (Noblin *et al.*, 2008). Finally, co-opting some areas of the leaf for C_4 photosynthesis while maintaining water storage cells will probably allow the C_4 descendants to thrive in more arid conditions (Voznesenskaya *et al.*, 1999; Kadereit *et al.*, 2012). All of these characteristics, which can be directly affected by the evolutionary path a species took to achieve C_4 function, will determine the physiology of a plant and thus its ecological preferences. Therefore, the diversity of evolutionary trajectories toward C_4-compatible leaf anatomy might partially explain the ecological diversity associated with distinct C_4 lineages (e.g. Taub, 2000; Kadereit *et al.*, 2012; Liu *et al.*, 2012).

Consequences for putative genetic determinism

A detailed discussion of genetic determinants is beyond the scope of this paper. However, it is worth pointing out that, despite recent important developments (e.g. Slewinski *et al.*, 2013; Wang *et al.*, 2013; Lundquist *et al.*, 2014), the genetic mechanisms necessary to introduce C_4-compatible anatomy into C_3 species remain largely unknown. This has particular implications for the bioengineering of C_4 photosynthesis into major C_3 crops, such as rice and wheat, which has the potential to greatly enhance yield (Covshoff and Hibberd, 2012; von Caemmerer *et al.*, 2012). First, the multiplicity of traits means that there are probably multiple genes involved. For instance, a phylogenetic analysis shows that the distance between consecutive bundle sheaths and the size of these bundle sheaths vary independently in grasses (Table 1), suggesting different underlying genetic changes. Second, as the variation in most traits presents a continuum from C_3 to C_4 plants, the determinism is likely to involve multiple genes with small effects and no master switch. Third, the diversity of strategies used to achieve leaf functions that are compatible with C_4 photosynthesis means that genetic determinism is likely to differ among C_4 lineages. Finally, the genetic changes that occur during the evolution of C_4 photosynthesis are likely to vary as a function of the condition in the C_3 ancestor.

Interestingly, similar variation in some of the underlying traits exists in C_3 and C_4 species, which suggests that useful genetic variants may be identified from the analysis of C_3 taxa that vary in only some of the traits, even if these C_3 taxa do not present C_4-like anatomies. For instance, a C_3 taxon with variation in the number of mesophyll cells between consecutive veins would be a good study system, even if the bundle sheath and distribution of chloroplasts were not C_4-compatible. Considering variation within C_3 taxa that are unrelated to C_4 lineages might therefore expose new ways to identify the adaptive significance of individual C_4 components, as well as their genetic determinism.

Conclusions

Overall, C_4 leaves can be defined by a set of important functions that characterize all C_4 plants. However, the underlying developmental characteristics that generate these functional properties are extremely variable, as a consequence of the taxonomic diversity of C_4 plants. The same functionally important traits are not homologous among all C_4 plants, and this has important implications for the evolution and underlying genetics of C_4-specific leaf anatomy. In addition, the developmental modifications that generate each of the essential requirements of C_4 leaf anatomy can happen independently. Thus, distantly related C_4 groups might arrive at the same phenotype for one of these requirements (e.g. both groups co-opt the same compartment for the Calvin cycle) but not another (e.g. they achieve small distances between the two compartments through either a reduction in the number of cells between veins or the development of additional veins).

Most of the anatomical characteristics that can generate functional properties of C_4 leaves exist in at least some C_3 plants. The only well-characterized exception is chloroplast concentration in the compartment co-opted for the segregation of the Calvin cycle, which seems to be specific to C_4 plants, and to some extent C_2 plants. Without considering the distribution of chloroplasts and hence C_4 physiology, leaves of C_3 and C_4 plants cannot be placed into mutually exclusive categories (see Fig. 3, for example), and there is continuous variation of the underlying traits among C_4 and C_3 species (Fig. 4). Hard categorization is meaningful from a functional perspective, but it wrongly suggests that the recurrent emergence of C_4 photosynthesis represents the same number of drastic transitions between two distinct and homogeneous characteristic states. Acknowledging the diversity present within both C_3 and C_4 taxa, and the continuum that exists between these two physiological states, is paramount to understanding the evolutionary processes that led to C_4 plants, as well as the genetic mechanisms responsible for C_4-compatible leaf anatomy.

Supplementary data

Supplementary data are available at *JXB* online.

Supplementary Fig. S1. Cross-sections corresponding to the diagrams shown in Fig. 3.

Acknowledgements

This work was funded by a University of Sheffield Prize Scholarship to MRL, and a Royal Society University Research Fellowship UF120119 to PAC. The authors thanks Dr David Chatelet from Brown University and Professor Travis Columbus from Rancho Santa Ana Botanic Garden who provided the leaf cross-sections reproduced in the figures.

References

Araus JL, Brown HR, Byrd GT, Serret MD. 1991. Comparative effects of growth irradiance on photosynthesis and leaf anatomy of *Flaveria brownii* (C_4-like), *Flaveria linearis* (C_3–C_4) and their F1 hybrid. *Planta* **183,** 497–504.

Araus JL, Brown RH, Bouton JH, Serret MD. 1990. Leaf anatomical characteristics in *Flaveria trinervia* (C_4), *Flaveria brownii* (C_4-like) and their F1 hybrid. *Photosynthesis Research* **26,** 49–57.

Aubry S, Kneřová J, Hibberd JM. 2013. Endoreduplication is not involved in bundle-sheath formation in the C_4 species *Cleome gynandra*. *Journal of Experimental Botany* **65,** 3557–3566.

Beaulieu JM, Leitch IJ, Patel S, Pendharker A, Knight CA. 2008. Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytologist* **179,** 975–986.

Berry JA, Downton WJS, Tregunna EB. 1970. The photosynthetic carbon metabolism of *Zea mays* and *Gomphrena globosa*: the location of the CO_2 fixation and the carboxyl transfer reactions. *Canadian Journal of Botany* **48,** 777–786.

Besnard G, Christin PA, Male PJ, Coissac E, Ralimanana, Vorontsova MS. 2013. Phylogenomics and taxonomy of Lecomtelleae (Poaceae), an isolated panicoid lineage from Madagascar. *Annals of Botany* **112,** 1057–1066.

Boardman NK. 1977. Comparative photosynthesis of sun and shade plants. *Annual Review of Plant Physiology* **28,** 355–377.

Bolton JK, Brown RH. 1980. Photosynthesis of grass species differing in carbon dioxide fixation pathways V. Response of *Panicum maximum*, *Panicum milioides*, and tall fescue (*Festuca arundinacea*) to nitrogen nutrition. *Plant Physiology* **66,** 97–100.

Botha CEJ. 1992. Plasmodesmatal distribution, structure and frequency in relation to assimilation in C_3 and C_4 grasses in southern Africa. *Planta* **187,** 348–358.

Bowes G, Haladay AS, Haller WT. 1979. Seasonal variation in the biomass, tuber density, and photosynthetic metabolisms of *Hydrilla* in three Florida lakes. *Journal of Aquatic Plant Management* **17,** 61–65.

Bowes G, Rao SK, Estavillo GM, Reiskind JB. 2002. C_4 mechanisms in aquatic angiosperms: comparisons with terrestrial C_4 systems. *Functional Plant Biology* **29,** 379–392.

Bowes G, Salvucci ME. 1984. *Hydrilla*: inducible C_4-type photosynthesis without Kranz anatomy. *Advances in Photosynthesis Research* **3,** 829–832.

Bowes G, Salvucci ME. 1989. Plasticity in the photosynthetic carbon metabolism of submersed aquatic macrophytes. *Aquatic Botany* **34,** 233–266.

Bowes G. 2011. Single-cell C_4 photosynthesis in aquatic plants. In: Raghavendra AS, Sage RF, eds. *C_4 photosynthesis and related CO_2 concentrating mechanisms*. The Netherlands: Springer, 63–80.

Bowman SM, Patel M, Yerramsetty P, Mure CM, Zielinski AM, Bruenn JA, Berry JO. 2013. A novel RNA binding protein affects *rbcL* gene expression and is specific to bundle sheath chloroplasts in C_4 plants. *BMC Plant Biology* **13,** 138.

Boykin LM, Pockman WT, Lowrey TK. 2008. Leaf anatomy of Orcuttieae (Poaceae: Chloridoideae): more evidence of C_4 photosynthesis without Kranz anatomy. *Madroño* **55,** 143–150.

Brown WV. 1975. Variations in anatomy, associations, and origins of Kranz tissue. *American Journal of Botany* **62,** 395–402.

Brown WV. 1977. The Kranz syndrome and its subtypes in grass systematics. *Memoirs of the Torrey Botanical Club* **23,** 1–97.

Brown WV, Smith BN. 1972. Grass evolution, the Kranz syndrome, $^{13}C/^{12}C$ ratios, and continental drift. *Nature* **239,** 345–346.

Bruhl JJ, Perry S. 1995. Photosynthetic pathway-related ultrastructure of C_3, C_4, and C_3-like C_3-C_4 intermediate sedges (Cyperaceae), with special reference to *Eleocharis*. *Australian Journal of Plant Physiology* **22,** 521–530.

Carolin RC, Jacobs SWL, Vesk M. 1973. The structure of the cells of the mesophyll and parenchymatous bundle sheath of the Gramineae. *Botanical Journal of the Linnean Society* **66,** 259–275.

Carolin RC, Jacobs SWL, Vesk M. 1975. Leaf structure in Chenopodiaceae. *Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie* **95,** 226–255.

Carolin RC, Jacobs SWL, Vesk M. 1977. The ultrastructure of Kranz cells in the family Cyperaceae. *Botanical Gazette* **138,** 413–419.

Christin PA, Freckleton RP, Osborne CP. 2010. Can phylogenetics identify C_4 origins and reversals? *Trends in Ecology and Evolution* **25,** 403–409.

Christin PA, Osborne CP. 2013. The recurrent assembly of C_4 photosynthesis, an evolutionary tale. *Photosynthesis Research* **117,** 163–175.

Christin PA, 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, Sage TL, Edwards EJ, Ogburn RM, Khoshravesh R, Sage RF. 2011. Complex evolutionary transitions and the significance of C_3-C_4 intermediate forms of photosynthesis in Molluginaceae. *Evolution* **65,** 643–660.

Covshoff S, Hibberd, JM. 2012. Integrating C_4 photosynthesis into C_3 crops to increase yield potential. *Current Opinion in Biotechnology* **23,** 209–214.

Crookston RK, Moss DN. 1970. The relation of carbon dioxide compensation and chlorenchymatous vascular bundle sheaths in leaves of dicots. *Plant Physiology* **46,** 564–567.

Das VSR, Raghavendra AS. 1976. C_4 photosynthesis and a unique type of Kranz anatomy in *Glassocordia boswallaea* (Asteraceae). *Proceedings of the Indian Academy of Sciences* **84B,** 12–19.

Dengler NG, Dengler RE, Donnelly PM, Hattersley PW. 1994. Quantitative leaf anatomy of C_3 and C_4 grasses (Poaceae): bundle sheath and mesophyll surface area relationships. *Annals of Botany* **73,** 241–255.

Dengler NG, Dengler RE, Hattersley PW. 1985. Differing ontogenetic origins of PCR ("Kranz") sheaths in leaf blades of C_4 grasses (Poaceae). *American Journal of Botany* **72,** 284–302.

Dengler NG, Donnelly PM, Dengler RE. 1996. Differentiation of bundle sheath, mesophyll, and distinctive cells in the C_4 grass *Arundinella hirta* (Poaceae). *American Journal of Botany* **83,** 1391–1405.

Downton WJS, Tregunna EB. 1968. Carbon dioxide compensation-its relation to photosynthetic carboxylation reactions, systematics of the Gramineae, and leaf anatomy. *Canadian Journal of Botany* **46,** 207–215.

Duval-Jouve J. 1875. Histotaxie des feuilles des Graminees. *Annales des Science Naturelles Botanical Series* **61,** 227–346.

Edwards GE, Franceschi VR, Voznesenskaya EV. 2004. Single-cell C_4 photosynthesis versus the dual-cell (Kranz) paradigm. *Annual Review of Plant Biology* **55,** 173–196.

Edwards GE, Ku MSB. 1987. Biochemistry of C_3-C_4 intermediates. In: Hatch MD, Boardman NK, eds. *The biochemistry of plants: a comprehensive treatise*, **Vol. 10**. New York: Academic Press, 275–325.

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*. The Netherlands: Springer, 29–61.

El-Sharkawy M, Hesketh J. 1965. Photosynthesis among species in relation to characteristics of leaf anatomy and CO_2 diffusion resistances. *Crop Science* **5,** 517–521.

Freitag H, Kadereit G. 2014. C_3 and C_4 leaf anatomy types in Camphorosmeae (Camphorosmoideae, Chenopodiaceae). *Plant Systematics and Evolution* **300,** 665–687.

Freitag H, Stichler W. 2000. A remarkable new leaf type with unusual photosynthetic tissue in a central Asiatic genus of Chenopodiaceae. *Plant Biology* **2,** 154–160.

Gong CM, Bai J, Deng JM, Wang GX, Liu XP. 2011. Leaf anatomy and photosynthetic carbon metabolic characteristics in *Phragmites communis* in different soil water availability. *Plant Ecology* **211,** 675–687.

Gould SJ, Vrba ES. 1982. Exaptation—a missing term in the science of form. *Paleontological Society* **8,** 4–15.

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

Griffiths H, Weller G, Toy LFM, Dennis RJ. 2013. You're so vein: bundle sheath physiology, phylogeny and evolution in C_3 and C_4 plants. *Plant, Cell & Environment* **36,** 249–261.

Grime JP, Mowforth MA. 1982. Variation in genome size—an ecological interpretation. *Nature* **299,** 151–153.

Haberlandt G. 1884. *Physiologische Pflanzenanatomie*. Leipzig: Engelmann.

Hatch MD. 1987. C_4 photosynthesis: a unique blend of modified biochemistry, anatomy and ultrastructure. *Biochimica et Biophysica Acta* **895,** 81–106.

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, Watson L, Johnston FLS, Johnson CR. 1982. Remarkable leaf anatomical variations in *Neurachne* and its allies (Poaceae) in relation to C_3 and C_4 photosynthesis. *Botanical Journal of the Linnean Society* **84,** 265–272.

Hattersley PW, Watson L, Osmond CB. 1977. In situ immunofluorescent labelling of Ribulose-1,5-bisphosphate Carboxylase in leaves of C_3 and C_4 plants. *Australian Journal of Plant Physiology* **4,** 523–539.

Hattersley PW, Watson L. 1975. Anatomical parameters for predicting photosynthetic pathways of grass leaves: the 'maximum lateral cell count' and the 'maximum cells distant count'. *Phytomorphology* **25,** 325–333.

Hattersley PW, Watson L. 1992. Diversification of photosynthesis. In: Chapman GP, ed. *Grass evolution and domestication*. Cambridge: Cambridge University Press, 38–116.

Hattersley PW. 1984. Characterization of C_4 type leaf anatomy in grasses (Poaceae). Mesophyll: bundle sheath area ratios. *Annals of Botany* **53,** 163–179.

Heckmann D, Schulze S, Denton A, Gowik U, Westhoff P, Weber AP, Lercher MJ. 2013. Predicting C_4 photosynthesis evolution: modular, individually adaptive steps on a Mount Fuji fitness landscape. *Cell* **153,** 1579–1588.

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.

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.

Kadereit G, Ackerly D, Pirie MD. 2012. A broader model for C_4 photosynthesis evolution in plants inferred from the goosefoot family (Chenopodiaceae s.s.). *Proceedings of the Royal Society B: Biological Sciences* **279,** 3304–3311.

Kadereit G, Borsch T, Weising K. 2003. Phylogeny of Amaranthaceae and Chenopodiaceae and the evolution of C_4 photosynthesis. *International Journal of Plant Sciences* **164,** 959–986.

Keeley JE. 1998. C_4 photosynthetic modifications in the evolutionary transition from land to water in aquatic grasses. *Oecologica* **116,** 85–97.

Kellogg EA. 1999. Phylogenetic aspects of the evolution of C_4 photosynthesis. In: Sage RF, Monson RK, eds. *C_4 plant biology*. San Diego: Academic Press, 411–444.

Koteyeva NK, Voznesenskaya EV, Roalson EH, Edwards GE. 2011. Diversity in forms of C_4 in the genus *Cleome* (Cleomaceae). *Annals of Botany* **107,** 269–283.

Laetsch WM. 1974. The C_4 syndrome: a structural analysis. *Annual Review of Plant Physiology* **25,** 27–52.

Leegood RC. 2002. C_4 photosynthesis: principles of CO_2 concentration and prospects for its introductions into C_3 plants. *Journal of Experimental Botany* **53,** 581–590.

Leegood RC. 2008. Roles of the bundle sheath cells in leaves of C_3 plants. *Journal of Experimental Botany* **59,** 1663–1673.

Liu H, Edwards EJ, Freckleton RP, Osborne CP. 2012. Phylogenetic niche conservatism in C_4 grasses. *Oecologia* **170,** 835–845.

Ludwig LJ, Canvin DT. 1971. The rate of photorespiration during photosynthesis and the relationship of the substrate of light respiration to the products in sunflower leaves. *Plant Physiology* **48,** 712–719.

Lundquist PK, Rosar C, Brautigam A, Weber APM. 2014. Plastid signals and the bundle sheath:mesophyll development in reticulate mutants. *Molecular Plant* **7,** 14–29.

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.

Martins S, Scatena VL. 2011. Bundle sheath ontogeny in Kranz and non-Kranz species of Cyperaceae (Poales). *Australian Journal of Botany* **59,** 554–562.

Masterson J. 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science* **264,** 421–424.

McKown AD, Cochard H, Sack L. 2010. Decoding leaf hydraulics with a spatially explicit model: Principles of venation architecture and implications for its evolution. *American Naturalist* **175,** 447–460.

McKown AD, Dengler NG. 2007. Key innovations in the evolution of Kranz anatomy and C_4 vein pattern in *Flaveria* (Asteraceae). *American Journal of Botany* **94,** 382–399.

McKown AD, Dengler NG. 2009. Shifts in leaf vein density through accelerated vein formation in C_4 *Flaveria* (Asteraceae). *Annals of Botany* **104,** 1085–1098.

Muhaidat R, Sage RF, Dengler NG. 2007. Diversity of Kranz anatomy and biochemistry in C_4 eudicots. *American Journal of Botany* **94,** 362–381.

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 & Environment* **10,** 1723–1736.

Murphy LR, Barroca J, Franceschi VR, Lee R, Roalson EH, Edwards GE, Ku MSB. 2007. Diversity and plasticity of C_4 photosynthesis in *Eleocharis* (Cyperaceae). *Functional Plant Biology* **34,** 571–580.

Nelson T. 2011. Development of leaves in C_4 plants: anatomical features that support C_4 metabolism. In: Raghavendra AS, Sage RF. eds. *C_4 photosynthesis and related CO_2 concentrating mechanisms*. The Netherlands: Springer, 147–159.

Noblin X, Mahadevan L, Coomaraswamy IA, Weitz DA, Holbrook NM, Zwieniecki MA. 2008. Optimal vein density in artificial and real leaves. *Proceedings of the National Academy of Sciences, USA* **105,** 9140–9144.

Ocampo G, Koteyeva NK, Voznesenskaya EV, Edwards GE, Sage TL, Sage RF, Columbus JT. 2013. Evolution of leaf anatomy and photosynthetic pathways in Portulacaceae. *American Journal of Botany* **100,** 2388–2402.

Ogle K. 2003. Implications of interveinal distance for quantum yield in C_4 grasses: a modeling and meta-analysis. *Oecologia*, **4,** 532–542.

Olesen P. 1975. Plasmodesmata between mesophyll and bundle sheath cells in relation to the exchange of C_4-acids. *Planta* **123,** 199–202.

Orme D, Freckleton R, Thomas G, Petzoldt T, Fritz S, Isaac N, Pearse W. 2012. Caper: Comparative analyses of phylogenetics and evolution in R Package version 0.5. Available at http://cran.r-project.org/web/packages/caper/index.html.

Peter G, Katinas L. 2003. A new type of Kranz anatomy in Asteraceae. *Australian Journal of Botany* **51,** 217–226.

Renvoize SA. 1983. A survey of leaf-blade anatomy in grasses IV. Eragrostideae. *Kew Bulletin* **38,** 469–478.

Renvoize SA. 1985. A survey of leaf-blade anatomy in grasses V. The bamboo allies. *Kew Bulletin* **40,** 509–535.

Renvoize SA. 1987*a*. A survey of leaf-blade anatomy in grasses XI. Paniceae. *Kew Bulletin* **42,** 739–768.

Renvoize SA. 1987*b*. A survey of leaf-blade anatomy in grasses X: Bambuseae. *Kew Bulletin* **42,** 201–207.

Sack L, Scoffoni C, John GP, Poorter H, Mason CM, Mendez-Alonzo R, Donovan LA. 2013. How do leaf veins influence the worldwide leaf economic spectrum? Review and synthesis. *Journal of Experimental Botany* **64,** 4053–4080.

Sack L, Scoffoni C, McKown AD, Frole K, Rawls M, Harvan JC, Tran H, Tran T. 2012. Developmentally based scaling of leaf venation

architecture explains global ecological patterns. *Nature Communications* **3,** 837.

Sack L, Scoffoni C. 2013. Leaf venation: structure, function, development, evolution, ecology and applications in the past, present and future. *New Phytologist* **198,** 983–1000.

Sage RF. 2001. Environmental and evolutionary preconditions for the origin and diversification of the C_4 photosynthetic syndrome. *Plant Biology* **3,** 202–213.

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

Sage RF, McKown AD. 2006. Is C_4 photosynthesis less phenotypically plastic than C_4 photosynthesis? *Journal of Experimental Botany* **57,** 303–317.

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

Sage TL, Busch F, 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.

Schlichting CD, Levin DA. 1986. Phenotypic plasticity: an evolving plant character. *Biological Journal of the Linnean Society* **29,** 37–47.

Schlichting CD, Pigliucci M. 1993. Evolution of phenotypic plasticity via regulatory genes. *American Naturalist* **142,** 366–370.

Šímová I, Herben T. 2012. Geometrical constraints in the scaling relationships between genome size, cell size and cell cycle length in herbaceous plants. *Proceedings of the Royal Society B: Biological Sciences* **279,** 867–875.

Sinha NR, Kellogg EA. 1996. Parallelism and diversity in multiple origins of C_4 photosynthesis in the grass family. *American Journal of Botany* **83,** 1458–1470.

Skillman JB. 2008. Quantum yield variation across the three pathways of photosynthesis: not yet out of the dark. *Journal of Experimental Botany* **59,** 1647–1661.

Slewinski TL, Anderson AA, Zhang C, Turgeon R. 2013. Scarecrow plays a role in establishing Kranz anatomy in maize leaves. *Plant and Cell Physiology* **53,** 2030–2037.

Soros CL, Dengler NG. 2001. Ontogenetic derivation and cell differentiation in photosynthetic tissues of C_3 and C_4 Cyperaceae. *American Journal of Botany* **88,** 992–1005.

Sugimoto-Shirasu K, Roberts K. 2003. "Big it up": endoreduplication and cell-size control in plants. *Current Opinion in Plant Biology* **6,** 544–553.

Sultan SE. 1987. Evolutionary implications of phenotypic plasticity in plants. *Evolutionary Biology* **21,** 127–178.

Tateoka T. 1958. Notes on some grasses. VIII. On leaf structure of *Arundinella* and *Garnotia*. *Botanical Gazette* **120,** 101–109.

Taub DR. 2000. Climate and the U.S. Distribution of C_4 grass subfamilies and decarboxylation variants of C_4 photosynthesis. *American Journal of Botany* **87,** 1211–1215.

Taylor SH, Franks PJ, Hulme SP, Spriggs E, Christin PA, Edwards EJ, Woodward FI, Osborne CP. 2012. Photosynthetic pathway and ecological adaptation explain stomatal trait diversity amongst grasses. *New Phytologist* **193,** 387–396.

Terashima I, Miyazawa SI, Hanba YT. 2001. Why are sun leaves thicker than shade leaves? Consideration based on analyses of CO_2 diffusion in the leaf. *Journal of Plant Research* **114,** 93–105.

Ueno O, Kawano Y, Wakayama M, Takeda T. 2006. Leaf vascular systems in C_3 and C_4 grasses: a two-dimensional analysis. *Annals of Botany* **97,** 611–621.

Ueno O, Samejima M, Muto S, Miyachi S. 1988*a*. Photosynthetic characteristics of an amphibious plant, *Eleocharis vivipara*: expression of C_4 and C_3 modes in contrasting environments. *Proceedings of the National Academy of Sciences, USA* **85,** 6733–6737.

Ueno O, Takeda T, Maeda E. 1988*b*. Leaf ultrastructure of C_4 species possessing different Kranz anatomical types in the Cyperaceae. *Botanical Magazine* **101,** 141–152.

Ueno O. 1995. Occurrence of distinctive cells in leaves of C_4 species in *Arthraxon* and *Microstegium* (Andropogoneae-Poaceae) and the structural and immunocytochemical characterization of these cells. *International Journal of Plant Sciences* **156,** 270–289.

Ueno O. 2001. Environmental regulation of C_3 and C_4 differentiation in the amphibious sedge *Eleocharis vivipara*. *Plant Physiology* **127,** 1524–1532.

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

Vogelmann TC, Nishio JN, Smith WK. 1996. Leaves and light capture: light propagation and gradients of carbon fixation within leaves. *Trends in Plant Science* **1,** 65–70.

von Caemmerer S, Furbank RT. 2003. The C_4 pathway: an efficient CO_2 pump. *Photosynthesis Research* **77,** 191–207.

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

Voznesenskaya EV, Franceschi VR, Chuong SDX, Edwards GE. 2006. Functional characterization of phosphoenolpyruvate carboxykinase-type C_4 leaf anatomy: immuno-, chytochemical and ultrastructural analysis. *Annals of Botany* **98,** 77–91.

Voznesenskaya EV, Franceschi VR, Pyankov VI, Edwards GE. 1999. Anatomy, chloroplast structure and compartmentation of enzymes relative to photosynthetic mechanisms in leaves and cotyledons of species in the tribe Salsoleae (Chenopodiaceae). *Journal of Experimental Botany* **50,** 1779–1795.

Wakayama M, Ueno O, Ohnishi J. 2003. Photosynthetic enzyme accumulation during leaf development of *Arundinella hirta*, a C_4 grass having Kranz cells not associated with veins. *Plant and Cell Physiology* **44,** 1330–1340.

Wang P, Kelly S, Fouracre JP, Langdale JA. 2013. Genome-wide transcript analysis of early maize leaf development reveals gene cohorts associated with the differentiation of C_4 Kranz anatomy. *The Plant Journal* **75,** 656–670.

Weiner H, Burnell JN, Woodrow IE, Heldt HW, Hatch MD. 1988. Metabolite diffusion into bundle sheath cells from C_4 plants relation to C_4 photosynthesis and plasmodesmatal function. *Plant Physiology* **88,** 815–822.

Welkie W, Caldwell M. 1970. Leaf anatomy of species in some dicotyledon families as related to the C_3 and C_4 pathways of carbon fixation. *Canadian Journal of Botany* **48,** 2135–2146.

West-Eberhard MJ, Smith JAC, Winter K. 2011. Photosynthesis, reorganized. *Science* **332,** 311–312.

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

Wilson PJ, Thompson K, Hodgson JG. 2002. Specific leaf area and leaf dry matter content as alternative predictors of plant strategies. *New Phytologist* **143,** 155–162.

Woo KC, Anderson JM, Boardman N, Downton WJS, Osmond C, Thorne SW. 1970. Deficient photosystem II in agranal bundle sheath chloroplasts of C_4 plants. *Proceedings of the National Academy of Sciences, USA* **67,** 18–25.