In silico exploration of paths toward C_4 metabolism

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Esther M. Sundermann
aus Coesfeld

Düsseldorf, 6. Mai 2019

aus dem Institut für Informatik
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Martin J. Lercher

2. Prof. Dr. Andreas P.M. Weber

Tag der mündlichen Prüfung:
15. April 2019

ERKLÄRUNG

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“erstellt worden ist.

Düsseldorf, den 6. Mai 2019

Esther M. Sundermann

CONTENTS

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost, I thank my supervisor Prof. Martin J. Lercher for supporting me throughout my thesis with his patience and expertise. I am deeply thankful for the constant encouragement and suggestions regarding all aspects of my work.

I would like to thank Prof. Andreas P.M. Weber for many helpful discussions and suggestions.

I would also like to thank Dr. David Heckmann for countless discussion and extraordinary support.

Thanks to Prof. Bernhard Ø. Palsson for hosting my visit at the UCSD.

I would also like to thank Prof. Oliver Ebenhöh and the members of his group for valuable input to my research and a lot of wonderful time.

During my time in Düsseldorf, I had the pleasure to work with a lot of talented and supportive colleagues who made me enjoy my time in the CCB-group. I want to thank Tom Berwald, Jonathan Fritzemeier, Kyra Mooren, Antonio Rigueiro Mesejo, and Dipali Singh. Not to forget colleagues from other working groups who made my time in Düsseldorf unforgettable: Meike Hüdig and Katharina Gräfe.

My research was partly founded by the German Research Foundation during the CEPLAS project. I am grateful for the financial support and, in particular, the chance to meet and work with many kind and highly skilled researchers.

Last but not least, I want to express my deepest gratitude to my family, fiends, and Tim for understanding and support that allowed me to finish this work. I cannot imagine a life without you.

1

SUMMARY

1.1 INTRODUCTION

Photosynthesis sustains nearly all life on earth. Its key enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) has a dual function: it catalyzes the reaction with either CO_2 or O_2. When catalyzing the reaction with CO_2, this is the initial step of the Calvin-Benson cycle that produces sugar. In contrast, the reaction with O_2 starts photorespiration that results in drastic carbon and energy losses. Rubisco is an important resource sink; it utilizes up to 30 % of leaf nitrogen (Makino et al., 2003). In order to fix carbon from the atmosphere, the C_3 photosynthetic pathway uses Rubisco exclusively. In contrast, C_4 photosynthesis spatially separates the carbon fixation from Rubisco and the Calvin-Benson cycle. In C_4 photosynthesis, the initial carbon fixation in the mesophyll cell is catalyzed by Phosphoenolpyruvatcarboxylase, which is the start of the C_4 cycle. The C_4 cycle facilitates the transport of carbon in form of organic acids into the bundle sheath, the location of Rubisco and the Calvin-Benson cycle. This CO_2-concentrating mechanism allows plants to suppress photorespiration. Hence, C_4 species can reduce the required amount of Rubisco, which results in a more efficient use of available water and nitrogen compared to C_3 plants (Sage et al., 2012; Vogan and Sage, 2011; Vogan and Sage, 2012). The complex C_4 metabolism evolved more than 60 times independently from the original C_3 pathway (Sage et al., 2012). C_4 evolution is presumably triggered by environmental factors that result in high photorespiratory rates, e. g., high temperatures and high O_2/CO_2 gas concentration ratios. As the carbon fixation via the C_4 cycle requires additional energy, environments with sufficient light intensities are required.

1.2 MATHEMATICAL MODEL

In order to improve the understanding of the quantitative effect of environmental factors on the physiology and evolution of C_3, C_3-C_4 intermediate, and C_4 plants, we developed a comprehensive mathematical model. We describe this model in detail in *Manuscript 1*, presented in this thesis. This mechanistic model represents the complex photosynthetic apparatus and explicitly accounts for the photosynthetic nitrogen and energy allocation, which includes the energy production based on linear and cyclic electron transport. It can be parametrized as C_3, C_4, and all intermediate photosynthetic types and considers

the following environmental factors: light intensity, leaf nitrogen level, temperature, and CO_2 and O_2 gas concentrations. As the nitrogen and energy allocation is not understood in detail yet, the model assumes that resources are allocated such that the CO_2 assimilation rate—which is used as a proxy for fitness—is maximized for a given environment. Based on the resource allocation, the model provides detailed information about physiological and molecular parameters. The mathematical model is validated with data of the model genus of C_4 evolution, *Flaveria*. This genus includes closely related C_3, C_3-C_4 intermediate, and C_4 species.

1.3 RESULTS

We explore to what extent observed resource allocation patterns in different photosynthetic types are optimally adapted to current conditions, and to what extend this pattern is optimally adapted to ancestral environments (*Manuscript 1*). The optimal resource allocation was calculated for a standard evolutionary scenario, which is inferred from literature, and for the growth conditions given in the experimental set-up. A comparison of the modeled physiological parameters with the empirical data indicates that the observed resource distribution in C_4 plants still reflects optimality in ancestral environments. It further reveals that C_4 plants show limited phenotypic plasticity regarding resource allocation. The limited phenotypic plasticity allows us to quantitatively infer ancestral environments from currently observed resource allocation patterns. To adjust from the ancestral environment to a given growth environment, plants need to re-allocate nitrogen. Our analysis shows a link between the low phenotypic plasticity in C_4 plants and the need to re-allocate significantly more nitrogen between photosynthetic components for C_4 compared to C_3 relatives.

Analyzing C_3 and C_4 *Flaveria* species in the inferred ancestral environment provides insight into the widely unknown effect of nitrogen availability on the physiology of C_3 and C_4 plants and on C_4 photosynthesis evolution. This analysis is presented in *Manuscript 2*. A detailed comparison of the optimal nitrogen allocation in C_3 and C_4 plants shows that C_4 plants require an increased investment not only into the C_4 cycle but also the thylakoids. In addition to this qualitative information, our work allows us to add quantitative information on the physiological parameters, e. g., on maximal electron transport rate. We find that low nitrogen availability increases the C_4 advantage over C_3 in photosynthetic nitrogen-use efficiency, i. e., CO_2 assimilation rate per leaf nitrogen level. Moreover, a low nitrogen availability results in less required nitrogen re-allocation in order to transform an optimal C_3 into an optimal C_4 plant. This finding points to the possibility that nitrogen scarcity is an accelerator of C_4 evolution. We test this hypothesis by analyzing evolutionary trajectories for various leaf nitrogen

levels. This analysis indicates that a low nitrogen availability indeed promotes the evolution of C_4 photosynthesis.

In sum, the contributions of my PhD project are three-fold. Firstly, we developed a mathematical model that represents the carbon fixation and accounts for various environmental parameters as well as for energy and nitrogen partitioning across photosynthetic components. Secondly, using this model we quantify the effect of environmental factors on resource allocation and physiological parameters of photosynthetic organisms. Finally, we analyze the ecological and evolutionary role of nitrogen in C_3 and C_4 plants. We provide a novel modeling framework to improve the understanding of the effect of environmental factors on photosynthetic organisms. This framework can determine the cellular resource allocation that is optimal under future environmental conditions. Hence, it provides an approach to develop a blueprint on how to improve crop productivity to meet future environmental demands.

2 INTRODUCTION

Organismal metabolisms are highly complex; they include a high number of interconnecting metabolites and a wide range of enzymes. The complexity increases with multicellularity and the presence of cellular compartments due to spatial separation of enzymes and metabolites. Metabolisms fulfill two key tasks; first, breaking down substrates, e. g., carbon-rich glucose, into common metabolites (catabolism) and, second, synthesizing building blocks such as amino acids or fatty acids (anabolism) (Palsson, 2006, pp. 29 & 30). The efficiency of a metabolism is an important determinant of organismal fitness (Heckmann et al., 2013; Ibarra et al., 2002).

The metabolic efficiency is affected by multiple constraints that belong into categories that differ in their adjustability (Palsson, 2006, pp. 184 & 193–195). Metabolic fluxes are determined by multiple factors that can be changed, such as enzyme kinetics and substrate concentrations. Unadjustable constraints can be categorized as internal and external. Internal constraints arise from the need to maintain homeostasis. Further constraints can arise from scarcity of chemical compounds, e. g., nitrogen necessary to produce enzymes (Baudouin-Cornu et al., 2001), or carbon-rich substances to run cellular processes. These latter constraints result from the availability of essential substances and, thus, are determined to a substantial extend by the external environment. Environmental factors, like temperature or light intensity, typically depend on time. Multiple constraints can limit the metabolic efficiency simultaneously, as organisms face a multifaceted environment.

The balance of available resources through the regulatory machinery of a cell ensures an optimal metabolic efficiency (Heckmann et al., 2013; Varma and Palsson, 1994). Depending on the environmental conditions an organism is facing, the optimal allocation might look drastically different (Heckmann et al., 2013; Zhu et al., 2007). An improved understanding of the interplay between environment and metabolism paves the way to understand currently observed organisms and to predict likely future evolutionary developments. Hence, the knowledge gained contributes to develop bioengineering approaches, improve current organisms, and tackle future challenges.

2.1 PHOTOSYNTHETIC METABOLISM

The interplay between environmental factors and metabolism is particularly relevant in the context of photosynthesis. Autotrophs provide a suitable platform to explore the interplay between environmental

factors and metabolism, as they show a limited diversity of nutrient sources compared to heterotrophs. There are various photosynthetic pathways that fix carbon from the atmosphere. They show specific resource allocation patterns and are adaptations to various niches. Finally, as photosynthetic metabolisms are highly complex and modes of photosynthesis evolved multiple times independently, photosynthesis is an excellent model to explore the evolution of a complex trait in response to the environment.

As photosynthesis fixes carbon from the atmosphere and, thus, creates the vast majority of global organic carbon, this process sustains nearly all life on earth. Photosynthetic organism are of utmost importance for the human society, as they supply food, serve as feed, and are used to produce energy. The world population is increasing drastically; it is postulated that by 2050 the population will reach ~9 billion (Karp and Richter, 2011). This is associated with major challenges especially for food and energy security (Karp and Richter, 2011; Lal, 2010). In addition to the increasing population, the world climate will drastically change toward warmer temperatures and an increased atmospheric CO_2 concentration (IPCC, 2013). These changes affect the performance of photosynthetic organisms, especially crop plants (Walker et al., 2016), and may affect the optimal resource allocation.

2.1.1 *Photorespiration*

The photosynthetic key enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is central to primary production. Rubisco has a dual function: it catalyzes the reaction with either CO_2 or O_2. When catalyzing the carboxylation reaction, Rubisco catalyzes the conversion of carbon and Ribulose-1,5-bisphosphate (RuBP) to a C_3 acid as the initial step of the Calvin-Benson cycle that reduces CO_2 to glucose. In contrast, the reaction with O_2 starts a metabolically costly pathway that results in a toxic by-product, consumes cellular energy, requires nitrogen for enzymes, and results in net carbon loss (Maurino and Peterhansel, 2010; Walker et al., 2016). This pathway is called photorespiration. Rubisco plays a significant role in the metabolism of photosynthetic organisms: it is the most abundant enzyme in the world (Raven, 2013) and a significant organismal resource sink (Makino et al., 2003).

Rubisco, and, thus, photosynthesis as well as photorespiration, are strongly affected by a wide range of environmental factors including: (1) leaf nitrogen level, (2) light intensity, (3) temperature, and (4) CO_2 and (5) O_2 gas concentrations. The oxygenation to carboxylation ratio of Rubisco is affected by environmental factors such as O_2/CO_2 concentration ratio and temperature. The effect of photorespiration is significant: at 25 °C and an O_2/CO_2 ratio of about 600, photorespiration results in a carbon loss of ~26 % (Walker et al., 2016). The

carboxylation relative to oxygenation reaction decreases with increasing temperature due to the decrease of both the affinity of Rubisco and the aqueous solubility for CO_2 relative to O_2 (Long, 1999). In photosynthetic organisms, the light-dependent reactions produce energy that is required to run photosynthetic and photorespiratory pathways. In order to produce proteins, nitrogen is required. This essential nutrient strongly affects metabolic efficiency.

2.1.2 *C_4 Photosynthesis*

C_4 metabolism can be categorized as a carbon-concentrating mechanism. It allows plants to concentrate CO_2 around Rubisco and, by that, suppresses photorespiration. The intracellular high-CO_2 environment can be achieved by spatially separating the initial CO_2 fixation from Rubisco, which catalyzes the initial step of the Calvin-Benson cycle. In order to separate the processes mentioned, C_4 plants express Rubisco in bundle sheath cells, which is in contrast to the original C_3 pathway that expresses Rubisco in mesophyll cells. In C_4 plants, the initial carbon fixation is catalyzed by the enzyme Phosphoenolpyruvat (PEP) carboxylase (PEPC). The reaction catalyzed through PEPC results in a C_4 acid. This is the start of the so called C_4 cycle that transports the C_4 acid into the bundle sheath cell where it is decarboxylated and causes a high-CO_2 environment. The recycling and the path that ensures the availability of PEP in the mesophyll cell complete the C_4 cycle. The photosynthetic type that fixes CO_2 exclusively via Rubisco is termed C_3 photosynthesis, after the first reaction product. There are photosynthetic types that fix carbon by using partially expressed biochemical characteristics of C_4 photosynthesis, so called C_3-C_4 intermediates.

Roughly 3 % of vascular plants in 62 distinct lineages show the C_4 syndrome (Sage, 2016; Sage et al., 2011). Although C_4 plants represent a small fraction of vascular plant species, they account for 23 % of terrestrial gross primary productivity (Sage et al., 2012). Due to the efficiency and the complexity of C_4 photosynthesis, its ecology and evolution is of utmost interest for scientists and the society (Sage, 2004; Sage et al., 1999).

2.1.3 *Physiology and Ecology of C_3 and C_4 Photosynthesis*

C_3 and C_4 photosynthesis are adaptations to specific environmental niches. C_4 photosynthesis is an adaptation to environments that result in high photorespiratory rates (Sage, 2004). These habitats show factors such as high light, salinity, drought, and heat. Species that use the C_4 pathway are highly abundant in locations such as the tropics and savannas (Sage et al., 1999). In contrast, C_3 species dominate habitats like forests and tundras (Sage et al., 1999).

The adaptation to specific environmental niches goes along with diverse resource allocation patterns in terms of nitrogen and energy. On the one hand, the C_4 pathway requires additional enzymes for the C_4 cycle, which results in additional nitrogen requirements compared to the C_3 pathway. On the other hand, C_4 plants show a reduced amount of Rubisco compared to C_3 plants (Ghannoum et al., 2010; Makino et al., 2003). The Rubisco reduction in C_4 species is facilitated by the carbon-concentrating mechanism. The high-CO_2 concentration allows plants to operate Rubisco near its CO_2 saturation point and, thus, to increase in the Rubisco turnover rate, which is linked to a reduced specificity for CO_2 (Savir et al., 2010). This results in higher Rubisco turnover rates of C_4 plants compared to C_3 species (Sage, 2002). The nitrogen saved by reducing Rubisco exceeds the nitrogen required for the C_4 cycle and, thus, allows plants to increase the nitrogen investment into the thylakoids (Makino et al., 2003). This might result in a higher capacity of RuBP regeneration (Makino et al., 2003).

Similarly to the nitrogen allocation, the energy allocation also differs for C_4 relative to C_3 plants. The suppression of photorespiration allows C_4 plants to reduce corresponding energy losses, but the recycling of PEP consumes additional ATP. This results in different ATP/NADPH requirements for C_3 and C_4 plants. As C_3 photosynthesis requires less energy compared to the C_4 pathway, it can outcompete the C_4 pathway in shaded habitats (Sage et al., 1999).

C_4 photosynthesis is associated with multiple beneficial attributes in warm habitats with high light intensities. C_4 plants show higher photosynthetic rates compared to C_3 plants (Sage, 2001). The high assimilation rates result from the carbon-concentrating mechanism that allows C_4 plants to prevent photorespiration while boosting the carboxylation reaction of Rubisco through high Rubisco turnover rates (Long, 1999). The CO_2 assimilation rate per absorbed photons represents the quantum yield (Ehleringer et al., 1997; Ehleringer and Björkman, 1977). Under current atmospheric conditions, the quantum yield of C_4 metabolism is higher than the one of the C_3 pathway for high temperatures. The temperature range where C_3 and C_4 photosynthesis are equally efficient ranges from 22 to 30 °C (Ehleringer et al., 1997). The C_4 superiority indicates that the lack of photorespiration sets off the additional ATP costs of the carbon-concentrating mechanism.

C_4 photosynthesis is more efficient in using water than C_3 photosynthesis, i. e., the CO_2 assimilation rate per leaf transpiration rate is higher in C_4 compared to C_3 plants (Huxman and Monson, 2003; Vogan and Sage, 2011; Vogan and Sage, 2012). The increased water-use efficiency (WUE) results from a change in stomatal response and an increased CO_2 assimilation rate (Huxman and Monson, 2003; Vogan and Sage, 2011). Examples for the consequences of the higher WUE of

C_4 plants are longer growing seasons in seasonally dry locations and a greater productivity under limited availability of water (Long, 1999; Sage et al., 1999).

C_4 plants are more efficient in using nitrogen, i. e., they show a higher plant biomass per plant nitrogen (NUE) (Brown, 1978; Long, 1999). Also the CO_2 assimilation rate per leaf nitrogen content, photosynthetic nitrogen-use efficiency (PNUE), is higher for C_4 compared to C_3 species (Sage et al., 2012; Vogan and Sage, 2011; Vogan and Sage, 2012). The increased PNUE results from high CO_2 assimilation rates (Long, 1999). As mentioned before, the low Rubisco content of C_4 plants might facilitate an increased nitrogen investment into the thylakoids, which may contribute to the higher assimilation rates. C_4 species are hypothesized to exploit the PNUE advantage by producing more leaf area than C_3 and/or producing the same leaf area as a C_3 plant and investing the remaining nitrogen into the roots (Long, 1999). The first strategy results in higher whole-plant photosynthetic rates and a greater ability to capture light. The second strategy facilitates a high uptake of soil nutrients including nitrogen. Brown (1978) hypothesized that the improved PNUE results in an ecological advantage for C_4 plants under low nitrogen availability. Taken together, the advantage of one photosynthetic type over the other is determined by multiple environmental factors and their interactions.

2.1.4 *Evolution of C_4 Photosynthesis*

C_4 metabolism evolved more than 60 times independently from the original C_3 pathway (Sage et al., 2012). Thus, C_4 photosynthesis is a textbook example for the convergent evolution of a complex trait. This suggests a combination of a low evolutionary barrier and high selection pressures toward C_4 photosynthesis.

2.1.4.1 *Path Toward C_4 Photosynthesis*

Complex changes are necessary to establish C_4 photosynthesis, which include the modification of enzymatic properties and gene expression (Gowik and Westhoff, 2011). Due to the complexity of the trait, multiple, consecutive evolutionary steps are required to transform a C_3 into a C_4 plant (Heckmann et al., 2013; Sage et al., 2012; Williams et al., 2013). A trait is the more likely to fix in the population, the more beneficial the expression of the trait is (Heckmann et al., 2013). Considering the high number of times C_4 photosynthesis evolved, individual steps need to be beneficial (Gowik and Westhoff, 2011; Mallmann et al., 2014).

The evolution is proposed to happen in consecutive steps (Gowik and Westhoff, 2011; Heckmann et al., 2013; Sage et al., 2012): (1) preconditioning, (2) anatomical development, (3) establishment of C_2

photosynthesis, (4) establishment of the C_4 cycle, and (5) metabolic optimization.

The preconditioning includes duplications of whole genomes, genome sequences, or single genes (Gowik and Westhoff, 2011). The duplications ensure to keep the ancestral function while neofunctionalizing genes.

The second step toward C_4 photosynthesis is the acquisition of relevant anatomical features. This step includes the development of a high leaf vein density and the development of Kranz anatomy. Kranz anatomy is a wreath-like structure, which shows a pattern of vein-bundle sheath-mesophyll-mesophyll-bundle sheath-vein (Gowik and Westhoff, 2011; Heckmann, 2016). The anatomy shows high organelle number in the bundle sheath cells and enables an efficient diffusion between cells and a high metabolic capacity.

Then, a photorespiratory CO_2 pump is established, which is enabled by a loss of glycine decarboxylase complex activity in the mesophyll tissue (Rawsthorne et al., 1988). The photorespiratory CO_2 is decarboxylated in the bundle sheath cells where it can be refixed. Under conditions that cause high photorespiratory rates, this cycle can result in higher photosynthetic assimilation rates than those of the C_3 pathway (von Caemmerer, 1989). Once C_2 photosynthesis is in place, the activity of at least parts of the basic C_4 cycle increase to balance the nitrogen metabolism between mesophyll and bundle sheath cell (Mallmann et al., 2014). The expression of the bundle sheath specific decarboxylase enzymes, NADP-malic enzyme and NAD-malic enzyme, are already high in surrounding vascular tissue in C_3 plants (Hibberd and Quick, 2002). In contrast to the full C_4 pathway, PEP regeneration is possibly achieved through 3-phosphoglycerate mutase and enolase activity (Monson and Moore, 1989).

The key step in C_4 photosynthesis is the spatial separation of carbon fixation and the Calvin-Benson cycle. This is achieved by a shift of the Calvin-Benson cycle into the bundle sheath cells, an increased activity of the enzymes involved into the C_4 cycle, and a higher expression of carbonic anhydrase in the mesophyll cells.

In a final step, the C_4 cycle is optimized to ensure high fluxes through the C_4 metabolism. This is achieved by adjusting regulation (e. g., Engelmann et al. (2003)) and enzyme properties (e. g., Sage et al. (2012)).

The high level of polyphyly of C_4 photosynthesis can in part be explained by the presence of all required C_4 enzymes in C_3 plants (Aubry et al., 2011) and the presence of C_3-like anatomy in many C_3 species (Kinsman and Pyke, 1998).

2.1.4.2 *Environmental Factors Promoting C_4 Photosynthesis Evolution*

C_4 evolution is presumably triggered by environmental factors that result in high photorespiratory rates, in particular high temperatures,

high light intensities, and high O_2/CO_2 gas concentration ratios (Sage, 2004).

When vascular land plants first emerged, atmospheric CO_2 concentrations reached maximum values of 3300–3600 ppm (Gerhart and Ward, 2010). During the Oligocene (34–23 million years ago), atmospheric CO_2 concentrations dropped drastically. Within the last 420,000 years, 96 % of the time the CO_2 concentration was below 280 ppm (Sage and Coleman, 2001). The resulting high O_2/CO_2 gas concentration ratio causes high photorespiration. High photorespiratory rates represent an important selection pressure for C_4 evolution but also a high potential for carbon refixation. Carbon-concentrating mechanisms can enhance fitness under these conditions (Heckmann, 2016; Sage, 2004).

High temperature is a major environmental requirement for the evolution of C_4 photosynthesis (Sage, 2004). High temperature results in high photorespiratory rates (see Section 2.1.1) and also in high dark respiration in C_3 plants.

C_4 plants are more nitrogen use-efficient than C_3 plants (Sage et al., 2012; Vogan and Sage, 2011; Vogan and Sage, 2012). This might indicate that C_4 photosynthesis shows an adaptive advantage in environments with low nitrogen availability (Brown, 1978). Comparing closely related C_3, C_3-C_4 intermediate, C_4-like, and C_4 species reveals that there is no gradual increase in PNUE under current atmospheric conditions. C_4-like and C_4 species show an increased efficiency compared to C_3, C_3-C_4 intermediates species (Vogan and Sage, 2011; Vogan and Sage, 2012). In contrast, under conditions of low CO_2 partial pressures, there is an increase in the PNUE from C_3 via a C_3-C_4 intermediate to C_4 species (Vogan and Sage, 2012). This raises the question which role PNUE plays in C_4 evolution.

Multiple factors that act on a global scale, e. g., atmospheric CO_2 and O_2 concentrations, and local scale, e. g., nitrogen and temperature, affect the probability of evolution of C_4 photosynthesis (Heckmann, 2016).

2.2 MATHEMATICAL MODELS OF PHOTOSYNTHESIS

Mathematical models differ in their scope and resolution. In the area of photosynthesis, they explore among others the carbon isotop discrimination of the carbon-concentrating mechanism (von Caemmerer, 1989), optimality of the C_4 metabolism (Wang et al., 2014), or the evolutionary path toward C_4 photosynthesis (Heckmann et al., 2013).

2.2.1 *Modeling the CO_2 Assimilation Rate*

Frequently used mathematical models are the models presented in von Caemmerer (2000) and their predecessor (Berry and Farquhar, 1978;

Farquhar et al., 1980; von Caemmerer, 1989). The models calculate the CO_2 assimilation rate while considering environmental factors such as temperature, light intensity, and gas partial pressures. The calculations are based on the major restrictions of photosynthesis; limited ATP and NADPH production, which is required for the regeneration of RuBP and (if applicable) PEP (light-limited conditions) and limitations that result from the availability and the properties of the enzymes Rubisco and PEPC (enzyme-limited conditions). The CO_2 assimilation rate limited by the enzyme- and light-limited conditions are abbreviated by A_c and A_j, respectively.

These models are used to address a wide range of research questions (Yin and Struik, 2009; von Caemmerer, 1989, 2000). Based on the enzyme-limited C_3-C_4 model of von Caemmerer (2000), Heckmann et al. (2013) present a mathematical model that simulates and analyzes the fitness landscape on which C_3 evolves to C_4 photosynthesis. The model predicts the enzyme-limited CO_2 assimilation rate at steady-state for C_3, C_3-C_4 intermediates, and C_4 metabolism. This model allowed Heckmann and co-workers to explore the evolutionary path toward C_4 photosynthesis in a C_4 favoring environment. In order to be able to represent the different photosynthetic types, the following parameters, which are known to differ for C_3, C_3-C_4, and C_4 photosynthesis, are considered: (1) the fraction of Rubisco expressed in the mesophyll; (2) the maximum turnover rate of Rubisco carboxylation; (3) the PEPC activity; (4) the Michaelis constant of PEPC for bicarbonate; (5) the bundle sheath conductance for CO_2; (6) and the fraction of mesophyll-derived glycine decarboxylated in the mesophyll. Heckmann et al., 2013 consider an environment that shows light-saturation, 25 °C, a O_2/CO_2 gas concentration ratio of 800, and a fixed availability of Rubisco. However, this does not cover the full range of environmental factors relevant for photosynthesis, e. g., light-limited and cold conditions. The CO_2 assimilation rate for a given number of Rubisco catalytic sites is the considered fitness proxy. Although Rubisco abundance scales with leaf nitrogen level (Makino et al., 1997; Tazoe et al., 2008), there are multiple other photosynthetic nitrogen sinks: (1) enzymes of the Calvin-Benson cycle, (2) photorespiration, (3) C_4 cycle, and (4) thylakoids. As all sinks affect the CO_2 assimilation rate, a more suitable proxy for fitness in the context of diverse habitats is the CO_2 assimilation rate per leaf nitrogen level (photosynthetic nitrogen-use efficiency).

2.2.2 *Modeling Evolutionary Paths*

Fitness landscapes are useful to explore potential evolutionary paths from an ancestral toward a subsequent phenotype or genotype (Heckmann, 2016). Here, we focus on the phenotypic fitness landscapes. The genotype based fitness landscape is beyond the scope of this thesis (for

a discussion on the genetic and phenotypic landscapes see Heckmann (2015)). A fitness landscape is a theoretical hyperplane that is spanned by parameters that change during the evolutionary process. Each point is associated with the corresponding fitness. The topology of the fitness landscape has an effect on the accessibility of a phenotype from an ancestral one. Evolutionary trajectories (paths) from one to another phenotype depend on the existence of adaptive mutations and epistatic interactions (Heckmann, 2015). Smooth, single-peaked landscapes result in the accessibility of the optimal fitness from each phenotype on the landscapes, while rugged landscapes show local optima and, thus, cause "dead end"-phenotypes (Franke et al., 2011).

The exploration of a fitness landscape tends to be very complex, due to the high dimensionality and the resulting high number of possible parameter combinations (Heckmann, 2015). Common approaches allow scientists to analyze only a small subset of mutations. This is facilitated by focusing on those mutations that are known to be relevant for the organismal fitness or connect the phenotypic states of interest. Environmental factors that affect fitness are not static in time. Therefore, the effect of, potentially multiple, environmental factors further increase the complexity.

2.3 AIMS OF THE THESIS

The understanding of the interplay between environment and photosynthetic metabolism is essential for the current and future society. In this thesis, I investigate the effect of various environmental conditions on the ecology of C_3, C_3-C_4 intermediate, and C_4 plants and on C_4 evolution through mathematical modeling. Mathematical models are a promising strategy to explore this interplay because evolutionary research questions can be simulated within a reasonable timescale and a variety of parameters can be estimated that are otherwise infeasible or impractical to determine.

The overall aim of this thesis is to develop a mathematical model that represents C_3, C_3-C_4 intermediate, and C_4 photosynthesis while considering a wide range of environmental factors. Environmental factors that have been reported to be of special interest in the context of photosynthetic ecology and evolution, e. g., temperature and atmospheric gas concentrations, are considered. As C_3, C_3-C_4 intermediate, and C_4 plants differ in their energy and nitrogen allocation, special attention is payed to the effect of light and nitrogen. By analyzing the mathematical model developed, research questions dealing with the ecology of C_3 and C_4 species and C_4 evolution are addressed.

In *Manuscript 1*, I developed a mathematical model that calculates the carbon fixation rate (a proxy for fitness) while accounting for the energy and nitrogen partitioning across photosynthetic components and for the following environmental parameters: (1) light intensity,

(2) temperature, (3) leaf nitrogen level, and (4) CO_2 and (5) O_2 gas concentrations. The model allows us to assess environment-dependent plant physiology and performance as a function of resource allocation patterns and, thus, to compare theoretically optimal resource allocation patterns with those observed in specific environments.

Manuscript 2 addresses the question of what role nitrogen availability plays in C_4 evolution and how nitrogen availability affects the ecology of C_3 and C_4 plants. It focuses on the ancestral environment relevant for C_4 evolution and is based on the mathematical model presented in the first manuscript.

The presented work provides hypotheses about the qualitative and quantitative interactions of environmental factors and photosynthesis that can be explored in future, empirical work.

3 MANUSCRIPTS

This chapter outlines two manuscripts for which I am the first author. For each manuscript, I indicate my contributions. Than, the corresponding manuscript is presented.

3.1 MANUSCRIPT 1

Manuscript 1 is a variation on the paper's version that is available on *bioRxiv* (Sundermann et al., 2018). The version presented in this thesis slightly differs by minor rephrasing, the addition of a list that contains all parameters and their explanation, as well as the correction of one reference.

3.1.1 *Contributions to Manuscript 1*

I developed and implemented the model for nitrogen allocation and light reactions, implemented the optimization procedure, and conducted simulations. I designed the research, analyzed the data, and interpreted the results in collaboration with David Heckmann and Martin J. Lercher. I took the lead in writing the paper.

Modeling Cellular Resource Allocation Reveals Low Phenotypic Plasticity of C_4 Plants and Infers Environments of C_4 Photosynthesis Evolution

Esther M. Sundermann[1], Martin J. Lercher[1], David Heckmann[1,*]

[1] Institute for Computer Science and Department of Biology, Heinrich Heine University Düsseldorf, Germany

* Author to whom correspondence should be addressed

Contact information:

David Heckmann

Email: david.heckmann@hhu.de

Tel: +49 211 811 1358

Total word count (excluding summary, references and legends):	6499	No. of figures:	5 (Figs 1–5 in color)
Summary:	199	No. of Tables:	1
Introduction:	963	No of Supporting Information files:	20 (Methods S1–S5, Tables S1–S8, Figs. S1–S7)
Materials and Methods:	2380		
Results:	2144		
Discussion:	900		
Acknowledgements:	64		

Summary

- The regulation of resource allocation in biological systems observed today is the cumulative result of natural selection in ancestral and recent environments. To what extent are observed resource allocation patterns in different photosynthetic types optimally adapted to current conditions, and to what extend do they reflect ancestral environments? Here, we explore these questions for C_3, C_4, and C_3-C_4 intermediate plants of the model genus *Flaveria*.
- We developed a detailed mathematical model of carbon fixation, which accounts for various environmental parameters and for energy and nitrogen partitioning across photosynthetic components. This allows us to assess environment-dependent plant physiology and performance as a function of resource allocation patterns.
- To achieve maximal CO_2 fixation rates under growth conditions differing from those experienced during their evolution, C_4 species need to re-allocate significantly more nitrogen between photosynthetic components than their C_3 relatives. As this is linked to a limited phenotypic plasticity, observed resource distributions in C_4 plants still reflect optimality in ancestral environments, allowing their quantitative inference.
- Our work allows us to quantify environmental effects on resource allocation and performance of photosynthetic organisms. This understanding paves the way for interpreting present photosynthetic physiology in the light of evolutionary history.

Key Words

C_4 photosynthesis, C_3 photosynthesis, C_3-C_4 photosynthesis, evolution, *Flaveria*, phenotypic plasticity, resource allocation, systems modeling

Introduction

Metabolic efficiency is an important determinant of organismal fitness (Ibarra *et al.*, 2002; Heckmann *et al.*, 2013). Major constraints on metabolic fluxes can arise from scarcity of chemical compounds, e.g., nitrogen necessary to produce enzymes (Baudouin-Cornu *et al.*, 2001), or from the limited solvent capacity of cellular compartments (Atkinson, 1969; Beg *et al.*, 2007). To ensure optimal metabolic efficiency, gene regulation has to balance available resources appropriately. Modern methods of modeling metabolism rely strongly on the assumption of metabolic optimality under physico-chemical constraints (Oberhardt *et al.*, 2009; de Oliveira Dal'Molin *et al.*, 2010; Dourado *et al.*, 2017). Accordingly, resource allocation and its constraints are under intense investigation, although these studies are mostly

restricted to unicellular organisms. However, the metabolic efficiency of a given metabolic system is not static, but depends on the environment. Thus, uncertainties about the environmental properties that an organism has adapted to remain a major obstacle in the application of these methods. Autotrophic systems, such as plant leaves, are ideal to study the interaction of the environment and resource allocation, as the diversity of nutrient sources is much lower than for heterotrophs, which results in a reduced complexity of the space of possible environments. Furthermore, the effect of environmental factors on plant performance, e.g., the rate of CO_2 assimilation, have been studied intensively (von Caemmerer, 2000). In particular, C_3 and C_4 photosynthesis represent complementary gene expression and resource allocation patterns that result in high fitness in specific ecological niches.

In all plants, the fixation of carbon from CO_2 is catalyzed by the enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) as part of the Calvin-Benson cycle. Rubisco also shows an affinity for O_2, resulting in a toxic by-product, which needs to be recycled by the photorespiratory pathway and causes a significant loss of carbon and energy (Maurino & Peterhansel, 2010). Rubisco is an important resource sink in the leaf proteome of plants: it utilizes up to 30% of leaf nitrogen and up to 65% of total soluble protein (Ellis, 1979; Makino *et al.*, 2003). While C_3 plants operate the Calvin-Benson cycle in their mesophyll cells to fix carbon, C_4 plants express it in the bundle sheath cells and use phospho*enol*pyruvate (PEP) carboxylase (PEPC) for the initial fixation of carbon. The resulting C_4 acids are eventually decarboxylated in the bundle sheath cells, creating a local high-CO_2 environment around Rubisco that suppresses photorespiration. The C_4 cycle is completed by the regeneration of PEP by pyruvate, phosphate dikinase (PPDK).

Compared to C_3 photosynthesis, C_4 metabolism requires additional nitrogen to produce the C_4 enzymes; this additional investment is counteracted by reduced Rubisco requirements due to the concentration of CO_2 around Rubisco (Sage, 2004). The energy requirements of C_4 metabolism also differ from those of the C_3 pathway (Munekage & Taniguchi, 2016), as further ATP is needed for the regeneration of PEP, while ATP and NADPH requirements of the photorespiratory pathway are reduced. The metabolic efficiencies of the C_3 and C_4 system depend strongly on the environment. To achieve optimal metabolic efficiency, plants have to coordinate gene expression of the Calvin-Benson cycle, C_4 cycle, photorespiration, and light reactions in a complex response to the availability of light energy and nitrogen, as well as factors that influence the rate of photorespiration. The diversity of photosynthetic resource allocation patterns is emphasized by the existence of C_3-C_4 intermediate photosynthesis in some plants, where features of the

archetypical C_4 syndrome are only partially expressed. The *Flaveria* genus contains closely related plants of C_3, C_4, and C_3-C_4 intermediate types, making it an ideal system to study the interaction between resource allocation and environment in photosynthesis.

The selection pressures caused by environmental factors over evolutionary time scales are expected to lead to corresponding adaptations of gene regulation. In contrast, environmental variation on the time scale of individual generations may select for regulatory programs that adjust plant metabolism to the environment they currently face, a process called phenotypic plasticity. Reviewing the occurrence of phenotypic plasticity in C_3 and C_4 plants, Sage and McKown (2006) concluded that C_4 plants show inherent constraints that prevent the acclimation to environmental changes. Although the occurrence of phenotypic plasticity in plants is intensively studied, the plasticity in terms of resource allocation is not fully understood. In particular, it is not clear whether the phenotypic plasticity of different plant lineages is sufficient to acclimate optimally to the current environment; instead, many plants might still allocate at least parts of their resources in patterns that were optimal in the environments that dominated their recent evolutionary history.

The areas where C_4 dicotyledonous plants are assumed to have evolved are regions of low latitude showing combinations of heat, drought, and salinity (Sage, 2004). For *Flaveria*, analyses that combine phylogenetic context and environmental information point toward an evolutionary origin in open habitats with high temperatures (Powell, 1978; Sage, 2004; McKown *et al.*, 2005). The last common C_3 ancestor of the current *Flaveria* species lived 2–3 million years ago (Christin *et al.*, 2011), when CO_2 levels were significantly lower than the current, postindustrial level (Sage & Cowling, 1999; Gerhart & Ward, 2010). In summary, *Flaveria* species likely faced high light intensities, high temperature, and low atmospheric CO_2 level during their recent evolutionary history.

Here, we aim for a detailed understanding of the interplay between resource allocation and current and past evolutionary environments in plant physiology, examining C_3, C_4, and C_3-C_4 intermediate photosynthesis. To achieve this goal, we developed a mathematical model for these photosynthetic types that integrates knowledge on resource costs and relevant environmental factors. Using this model, we seek to understand (1) to what extent resource allocation is phenotypically plastic and to what extent it appears adapted to an environment the plants were facing during their evolutionary history; and (2) if resource allocation patterns can be used to identify unique environments of optimal adaption.

Results

Predicting resource allocation and fitness across environments and photosynthetic types: a mathematical model

The standard method to model the light- and enzyme-limited CO_2 assimilation rate of C_3, C_4, and C_3-C_4 intermediate plants is based on the mechanistic biochemical models of Berry and Farquhar (1978), Farquhar *et al.* (1980), and von Caemmerer (1989; 2000). With great success, these models predict the CO_2 assimilation rate considering enzymatic activities and various environmental parameters, including mesophyll CO_2 level and light intensities. In many ecosystems, the most limiting resource for plant growth is nitrogen (Malhi *et al.*, 2001; Vance, 2001). The increased nitrogen-use efficiency of C_4 species compared to C_3 relatives indicates that nitrogen availability may have played a major role in C_4 evolution (Vogan & Sage, 2011). However, existing model implementations predict CO_2 assimilation rates from known or estimated enzyme activities and electron transport capacity. Thus, these models do not allow to assess the effects of nitrogen investment into different classes of proteins—including enzymes and components of the electron transport chain—on the CO_2 assimilation rate of a given photosynthetic type in a specific environment.

Here, we present a nitrogen-dependent light- and enzyme-limited model for the steady-state CO_2 assimilation rate (Fig. 1;all parameters are listed in Table 2). The model describes C_3, C_4, and all intermediate photosynthetic types depending on its parameterization, including the nitrogen investment into its different components (see Heckmann *et al.* (2013) for details and Supporting Information Table S1 and S3 for our parameterization). We modified the light- and enzyme-limited C_3-C_4 models developed by von Caemmerer (2000) and added a fixed budget of nitrogen constraining the total abundance of photosynthetic proteins. Furthermore, we extended the existing models by explicitly modeling the ATP and NADPH production of the linear and cyclic electron transport (LET and CET, respectively). Thus, a photosynthetic nitrogen budget is distributed across the enzymes of the Calvin-Benson cycle in the mesophyll and bundle sheath cell, the C_4 cycle, and the proteins of the linear and cyclic electron transport in the thylakoid membranes. Combining this model with the temperature dependency of the photosynthetic apparatus (Massad *et al.*, 2007) results in a detailed model of photosynthesis that incorporates leaf nitrogen level, light intensity, mesophyll CO_2 and O_2 levels, as well as the effects of temperature (see Methods for details).

In order to understand physiological data in the context of adaptive environments, we aim to find optimal resource allocation in a given environment. To this end, we assume that resource allocation has been optimized by natural selection to maximize the net CO_2 assimilation rate (Zhu *et al.*, 2007; Gerhart & Ward, 2010; Vogan & Sage, 2012). We developed a robust optimization pipeline that reliably finds optimal resource allocation dependent on environments and photosynthetic types (see Methods for details). In previous work, optimality assumptions were successfully used in a variety of plant systems biology contexts; examples are candidate identification of photosynthetic engineering targets (Zhu *et al.*, 2007), explanation of the coordination of C_3 photosynthesis (Friend, 1991; Maire *et al.*, 2012), the exploration of evolutionary trajectories of C_4 photosynthesis (Heckmann *et al.*, 2013) and of inter-cellular pathways in C_2 plants (Mallmann *et al.*, 2014), and the prediction of dynamic proteome allocation in cyanobacteria (Reimers *et al.*, 2017). We use optimality of CO_2 fixation rate to determine (1) the optimal partitioning of NADPH between the Calvin-Benson cycle and the photorespiratory pathway, (2) the optimal partitioning of ATP across the Calvin-Benson cycle, photorespiratory pathway, and the C_4 cycle (if relevant), (3) the optimal proportion of LET and CET, and (4) the relative investment of nitrogen into Rubisco, the C_4 cycle enzymes, and the proteins of the light-dependent reactions (see Methods). For a specific photosynthetic type, the optimization procedure estimates the resource allocation that is optimally adapted to a given environment. Note that at the point of optimal resource allocation, the light- and enzyme-limited CO_2 assimilation rates are equal, as otherwise resources could be shifted from the non-limiting to the limiting sector.

Optimal resource allocation in the evolutionarily relevant environment explains physiological data and outperforms models based on the growth environment in C_4 plants

Do photosynthetic types exhibit differences in phenotypic plasticity, *i.e.*, do they differ in their ability to adjust their photosynthetic resource allocation to optimally fit the environment in which they were grown? Or is resource investment static and reflects past environments in which the plants' ancestors evolved? To compare these competing hypotheses, we predict physiological data of plants that are either optimally adapted to the experimental growth conditions used in the respective studies ('growth scenario') or to the environments in which they likely evolved ('evolutionary scenario'). This *in silico* experiment also serves as validation for our modeling framework; if the parameterization for *Flaveria* and our optimality assumptions are correct, we would expect the model to explain physiological responses in one of the two or an intermediate scenario. Based on the suggested environment of C_4 evolution in

Flaveria (Powell, 1978; Sage, 2004; McKown et al., 2005), the evolutionary environment is defined as having 1750 µmol quanta m^{-2} s^{-1} light intensity, 30°C temperature, 150 µbar mesophyll CO_2, and 200 mbar mesophyll O_2.

Vogan and Sage (2012) measured the net CO_2 assimilation rate as a function of intercellular CO_2 concentration (A-C_i curve) for *Flaveria robusta* (C_3), *F. ramosissima* (C_3-C_4), and *F. bidentis* (C_4). In this experiment, plants were grown at light intensities of 560 µmol quanta m^{-2} s^{-1}, 37°C at daytime, current atmospheric O_2 concentration and current or low atmospheric CO_2 concentrations. However, CO_2 assimilation curves calculated from a model parameterized for optimal CO_2 assimilation in these growth conditions are qualitatively different from the experimental curves (Fig. 2a; Supporting Information Figs. S2–S4). In contrast, the modeled curves based on a model optimally adapted to the evolutionary scenario are qualitatively consistent with the measured curves; this difference is especially pronounced in the case of the C_4 plant *F. bidentis*.

In the same study, Vogan and Sage (2012) measured the CO_2 assimilation rate for temperatures between 15°C and 45°C (A-Temperature curve; Fig. 2b; Supporting Information Fig. S5). The results assuming an optimal allocation under the evolutionary scenario agree qualitatively with the measured data, again in contrast to the values predicted from a model optimally adapted to the growth environment. Note that none of the species in this data set were used to obtain the temperature response curves used in the model (see Methods).

In an independent experiment, Vogan and Sage (2011) measured the dependence of CO_2 assimilation rate on leaf nitrogen levels in C_3, C_3-C_4 intermediate, C_4-like, and C_4 *Flaveria* species (Fig. 3). The plants were grown at 554 µmol quanta m^{-2} s^{-1} light intensity, 30°C at daytime, at current atmospheric CO_2 and O_2 concentrations. Again, the model results assuming optimal resource allocation in the evolutionary scenario are consistent with the measured data and outperform the results based on optimality in the growth scenario for C_3-C_4 intermediate, C_4-like, and C_4 plants.

We quantified the disagreement between measured curves and predicted results through the residual sum of squares (Table 1). In C_4 and C_4-like plants, the evolutionary scenario predicts all measured curves better than the growth scenario, except for the A-Temperature curve for C_4 plants grown at low CO_2 concentration. Jointly considering all measured curves in Figs. 2 and 3 as well as Supporting Information

Figs. S2–S5 (Vogan & Sage, 2011; Vogan & Sage, 2012), we find that for the C_4 and C_4-like species, squared residuals for the evolutionary scenario are statistically significantly smaller than for the growth scenario (C_4: $P = 6.0 \times 10^{-8}$; C_4-like: $P = 0.007$; Wilcoxon rank sum tests). This finding indicates that observed resource allocation patterns in C_4 and C_4-like plants reflect past environments relevant during evolution more than the environment in which the assayed plants were grown. Conversely, and as expected from Table 1, the observed differences between predictions from the evolutionary and growth scenario are not statistically significant for the C_3 and the C_3-C_4 intermediate species (C_3: $P = 0.35$; C_3-C_4: $P = 0.55$).

Dwyer *et al.* (2007) performed detailed experiments on the photosynthetic resource allocation and performance of the C_4 species *F. bidentis*. This data allows us to compare the predicted nitrogen investment into the three major photosynthetic components—Rubisco, C_4 cycle, and electron transport chain—as well as the corresponding CO_2 assimilation rate to experimentally observed resource allocation patterns. The plants were grown under 25°C or 35°C at daytime, 550 µmol quanta m^{-2} s^{-1}, and current atmospheric CO_2 and O_2 concentrations. Model predictions of chlorophyll content and the amount of photosystem II agree within a factor of 1.10 to 1.22 with values measured by Dwyer *et al.* (2007) (see Supporting Information Table S7). For plants grown at 25°C, the resource allocation determined under the evolutionary scenario agrees with the measured data within a factor of 0.47 to 1.22 (Fig. 4a); at 35°C, agreement is within a factor of 0.43 to 1.12 (Fig. 4b). In both cases, agreement is much lower for predictions in the growth scenario. We assessed the statistical significance of the superior performance of the evolutionary scenario by comparing the distributions of the squared residuals (expressed as fractions of the experimental means). The resource allocation calculated for the evolutionary scenario outperforms the growth scenario significantly for the data represented in Fig. 4 ($P = 7.2 \times 10^{-5}$, Wilcoxon rank sum test).

Although we could obtain the majority of our model parameters from the literature, the relationship of cytochrome f and the maximal electron transport rate of the CET had to be estimated (see Methods). We performed a sensitivity analysis to examine the robustness of the results to changes in the estimated parameters and to uncertainties in values obtained from the literature, focusing on parameters with high uncertainty or major expected effect on model predictions (see Supporting Information Method S5 and Table S8). The predictions based on the evolutionary scenario outperform those based on the growth environment consistently across all parameter sets (Supporting Information Fig. S1).

The model identifies a unique evolutionary environment for C_4 photosynthesis in Flaveria

The model optimally adapted to the evolutionary scenario leads to superior predictions of plant performance and resource allocation in C_4 plants compared to a parameterization optimized for the growth scenario across diverse physiological data sets. The inferior performance of the growth scenario model indicates a lack of phenotypic plasticity of resource allocation in C_4 plants. This finding points to the possibility that the environment most relevant for recent evolutionary adaptation of a given C_4 plant could be inferred quantitatively from observations on plant physiology and resource allocation. Thus, to infer a typical evolutionary environment for C_4 *Flaveria bidentis,* we calculated optimal resource allocation under conditions covering plausible ranges of mesophyll CO_2 partial pressure, temperature, and light intensities to identify the conditions that best explain the empirical data (Fig. 5). As atmospheric O_2 concentration remained almost constant for at least the last few million years (Gerhart & Ward, 2010), this environmental parameter is set to a constant value. We use the empirical data of Dwyer *et al.* (2007), as this data set comprises detailed measurements for each nitrogen pool and the resulting CO_2 assimilation rate, allowing us to quantify the discrepancy between modeled and measured values as the mean squared residuals (expressed as fractions of experimental means).

We find that the model showing minimal prediction error defines a unique environment (Fig. 5), exhibiting 1562.5 µmol quanta m^{-2} s^{-1} light intensity, 30°C, a mesophyll CO_2 level of 100 µbar, and an O_2 level of 200 mbar. As indicated in Fig. 5, the areas in which the model successfully describes the empirical values generally show high light intensities, intermediate to high temperatures, and a trend towards low CO_2 partial pressures. High light intensities and low CO_2 levels, as in the evolutionary scenario, favor an increased nitrogen net investment into the dark reactions, which goes along with a reduced investment into the electron transport chain. For decreasing CO_2 levels, a slight decrease in the Rubisco investment balances the C_3 and C_4 cycle. The effect of temperature is of special importance for plants using the C_4 cycle, as temperature increases PEPC activity drastically and therefore reduces the necessary nitrogen investment into the C_4 cycle. This allows an increased investment into the electron transport chain and Rubisco, which show reduced activity at elevated temperatures due to thermal instabilities.

Our results indicate that C_4 *Flaveria* species show a lower degree of photosynthetic phenotypic plasticity than closely related C_3 species. On a molecular level, this plasticity predominantly requires the re-allocation of nitrogen between the major photosynthetic protein pools. To assess the costs of phenotypic plasticity, we thus quantified the total fraction of nitrogen that needs to be re-allocated

between photosynthetic pools to optimally adjust photosynthesis from the evolutionary scenario to a given growth environment (δ_n, see Methods). We find that photosynthetic types that utilize the C_4 cycle require a consistently higher amount of re-allocation compared to C_3 plants ($P = 1.5x10^{-5}$, sign test, see Supporting Information Table S5). Our results thus reveal a link between required nitrogen re-allocation and limited photosynthetic phenotypic plasticity (see Supporting Information Tables S4–S6), suggesting a possible causal relationship.

Discussion

Our novel modeling framework allows us to study the interplay between photosynthetic plant performance and resource investment on the molecular level. Comparisons of model predictions with phenotypic and molecular data reveal that C_4 plants have low phenotypic plasticity in terms of resource allocation. This limited phenotypic plasticity may be explained by the large amount of nitrogen that needs to be re-allocated by C_4 plants to optimally adapt to a given growth environment (Supporting Information Table S5). The lack of phenotypic plasticity allowed us to make quantitative predictions for the environments that dominated recent evolution of C_4 photosynthesis in *Flaveria*. Previously, environments relevant for C_4 photosynthesis evolution have been inferred—mostly qualitatively—based on C_3-C_4 habitat comparisons (Powell, 1978; Sage, 2004; McKown *et al.*, 2005) and geophysiological considerations (Christin *et al.*, 2011). Our results are consistent with and refine these earlier estimates.

In contrast to our findings for C_4 and C_4-like plants, the performance of the evolutionary and the growth scenario models is similar for C_3 and C_3-C_4 intermediate *Flaveria* species (Table 1; Figs. 2 and 3; Supporting Information Figs. S2–S5). It is conceivable that the lack of superior performance for the evolutionary scenario in C_3 *Flaveria* species is not a result of higher phenotypic plasticity in these plants, but is due to an inappropriate parameterization of the evolutionary scenario. The environment most relevant for the recent evolution of C_3 *Flaveria* may be different from the environment used in the simulations, which was chosen based on its relevance for the C_4 lineages. To explore this possibility, we simulated a wide range of alternative environments, testing if resource allocation optimized for any of these leads to significantly improved model predictions for the data from Vogan and Sage (2012) for C_3 plants (Supporting Information Figs. S6 and S7). However, none of the environments tested led to a significant improvement. This result is in agreement with habitat studies that show that niches of C_3 and C_4 *Flaveria* species overlap (Powell, 1978). A more likely explanation for the similar performance of evolutionary and growth scenario models in C_3 plants could lie in the small amount of re-allocation C_3 plants require to transfer adaptively

between environments (Supporting Information Tables S4–S6). Our results thus suggest that C_3 (but not C_4) plants are phenotypically plastic enough to show some degree of adaptation towards current, postindustrial conditions.

Given the complexity of our physiological model, we needed to make a number of assumptions. We addressed uncertainties in model parameters through a sensitivity analysis, showing that our conclusions are robust against variation in these parameters (Supporting Information Fig. S1). Furthermore, our predictions assume that nitrogen availability in the evolutionary scenario is identical to current nitrogen availability.

Even though we find that the evolutionary scenario leads to superior predictions of physiological responses in C_4 plants when compared to the growth scenario, the PEPC activity predicted to be optimal in the evolutionary scenario is approximately 55% lower than experimentally observed data (Fig. 4). This discrepancy might in part be explained by the assumption of a fixed average daytime temperature in the simulations. Temperature variation strongly affects the PEPC activity; lower temperatures in the morning and evening may require higher PEPC activity than assumed in the simulations. Although predictions for total nitrogen investment into the thylakoids based on the evolutionary scenario are highly consistent with the measurements, the model overestimates the amount of cytochrome f by a factor of 2 (1.65 µmol m^{-2} instead of the measured 0.87 µmol m^{-2} for plants grown at 25°C, 1.43 µmol m^{-2} instead of 0.81 µmol m^{-2} at 35°C). However, the error of the measurements is uncertain, as no replicate measurements were performed for this parameter (Dwyer *et al.*, 2007). Discrepancies between model predictions and observations may also be in part due to error propagation from modeled amounts of chlorophyll and the photosystems. In each simulation, we optimized resource allocation for an environment that represents a static approximation to the dynamic environment a plant is facing. As diurnal and annual variations (which are no focus of this work) potentially show short-term trade-offs (Mori *et al.*, 2017; Reimers *et al.*, 2017), these might lead to a discrepancy between modeled and real evolutionary scenario. In particular, the difference between periodic and fluctuating conditions of the natural ancestral habitat on one hand, and the stable experimental growth conditions in audited growth chambers and the statically modeled evolutionary scenario on the other hand might have a strong effect.

In summary, we developed a general model of the complex photosynthetic apparatus, its resource requirements, and its interactions with environmental conditions. The presented modeling pipeline allows

us to determine the extent of phenotypic plasticity and the relevance of different environmental conditions for photosynthetic organisms using C_3, C_3-C_4 intermediate, and C_4 metabolism. Applied to the physiological data from *Flaveria*, our work points to a strongly constrained phenotypic plasticity of C_4 plants towards all considered environmental factors. This allows us to infer unique selective environments from plant performance and resource allocation data. More generally, our model provides a powerful tool to analyze the resource allocation of photosynthetic organisms and its dependence on environmental factors, allowing estimates for physiological and molecular parameters for which measurements are currently infeasible or impractical. This may prove to be of particular utility for systematically assessing the likely performance of crops in environments distinct from their natural habitats and for suggesting engineering targets in cases of limited phenotypic plasticity.

Description

Model overview

The nitrogen-dependent light- and enzyme-limited model allows us to calculate the environment-dependent net steady-state CO_2 assimilation rate (A) of C_3, C_4, and all C_3-C_4 intermediate photosynthetic types. The model inputs are parameters defining the photosynthetic type and species-specific, invariable biochemical properties of the leaf to be modeled. Additionally, the input parameters comprise the following environmental factors: light intensity, leaf nitrogen level, temperature, and CO_2 and O_2 mesophyll partial pressures. We simulate a plant that is adapted to the input environment with respect to photosynthetic nitrogen and energy allocation. To this end, the nitrogen and energy allocation pattern that maximizes the net steady-state CO_2 assimilation rate (A) is calculated via optimization, subject to the environmental and species-specific input parameters. All parameters are listed in Table 2.

Environmental factors and evolutionary parameters

We specify the environment in terms of the following factors: light intensity, leaf nitrogen level, temperature, and CO_2 and O_2 mesophyll partial pressures. The photosynthetic type is defined by six parameters: the Rubisco distribution between mesophyll and bundle sheath cells (β); the Rubisco kinetics, (specified through a single parameter, k_{ccat} [1/s], due to the known trade-off relationships between the kinetic parameters (Savir *et al.*, 2010)); the maximal C_4 cycle activity (V_{pmax}, [µmol m^{-2} s^{-1}]); the fraction of glycine decarboxylated by the glycine decarboxylase complex in the bundle sheath cell that is derived from oxygenation by Rubisco in the mesophyll cell (ξ); the Michaelis constant of PEPC for bicarbonate (K_p,

[µbar]), and the bundle sheath cell conductance (g_s, [µmol m^{-2} s^{-1}]) (see Heckmann *et al.* (2013) for details). The values for the parameters are taken from the literature (see Supporting Information for details).

Nitrogen allocation

To calculate the CO_2 assimilation rate, we focus on the photosynthetic nitrogen pool (N_{ps}, [µmol m^{-2}]). In our model, N_{ps} can be allocated across the following major pools of leaf photosynthetic nitrogen: the main enzyme of the Calvin-Benson cycle (n_{Etot}), Rubisco; the main enzymes of the C_4 cycle (n_{C4}), PEPC and PPDK; and the thylakoids (n_{Jmax}), which include the electron transport chains. N_{ps} is calculated as a fraction of total leaf nitrogen (N_t, [µmol m^{-2}]) based on phenomenological observations according to Eqn 1, which comprises measured values for the investment into Rubisco, 12%, and the investment into the thylakoids (n_{fit}, [fraction]) of C_3 plants (Vogan & Sage, 2011; Vogan & Sage, 2012). n_{fit} represents a fit of the proportion of nitrogen invested into the thylakoids as a function of N_t, based on the data of Vogan and Sage (2011).

$$N_{ps} = \left(0.12 + n_{fit}\right) \cdot N_t \tag{1}$$

with

$$n_{fit} = \left(\frac{50.38 - 0.270 \cdot N_t \cdot 10^{-3} + 0.0005035 \cdot (N_t \cdot 10^{-3})^2}{100}\right)$$

We assume a nitrogen investment into the photorespiratory enzymes of 13.8%, as suggested by Zhu *et al.* (2007) for a 'typical' C_3 plant. To account for the reduced enzyme requirements of the photorespiratory cycle, we assume that N_{ps} increases by 10% in plants that show sufficient C_4 cycle activity; in our analyses, this applies to the C_3-C_4 intermediate, C_4-like, and C_4 species.

Nitrogen allocated to Rubisco

We only consider the nitrogen requirements of Rubisco in the Calvin-Benson cycle, as it accounts for the major nitrogen costs of this cycle (Evans & Seemann, 1989). The amount of catalytic sites of Rubisco (E_{tot}, [µmol m^{-2}]) is calculated from the invested nitrogen by Eqn 2, where n_{Etot} represents the fraction of N_{ps} invested into Rubisco:

$$E_{tot} = \frac{\mathrm{n}_{Etot} \cdot N_{ps} \cdot 8}{11.4 \cdot 550} \tag{2}$$

The parameters of this relationship are taken from Harrison *et al.* (2009).

Nitrogen allocated to enzymes of the C_4 cycle

The nitrogen cost of C_4 cycle enzymes is calculated from data on enzyme kinetics. The nitrogen requirements of the C_4 cycle consider co-limitation of PEPC and PPDK, whose molecular weight (MW) and k_{cat} are used to calculate the maximal rate of C_4 cycle activity (Evans & von Caemmerer, 2000; Wang *et al.*,

2014). Eqn 3 represents the relationship between V_{pmax} and nitrogen investment into the C_4 enzymes ($n_{C4}N_{ps}$). MW* represents the nitrogen requirement of a catalytic site, assuming the nitrogen content is 16% (Makino *et al.*, 2003). Indices declare the considered enzyme.

$$V_{pmax} = \frac{n_{C4} \cdot N_{ps}}{\left(\frac{MW^*_{PPDK}}{kcat_{PPDK}}\right)+\left(\frac{MW^*_{PEPC}}{kcat_{PEPC}}\right)} \quad (3)$$

Nitrogen and the maximal electron transport rate

Nitrogen invested into the thylakoids ($N_{thy} = N_t\, n_{thy}$, [µmol m^{-2}]) is related to the maximal electron transport rate (J_{max}, [µmol m^{-2} s^{-1}]) via the amount of cytochrome f (cyt, [mmol/mol Chl]) and by considering photosystems I and II (PSI and PSII, [mmol/mol Chl]) as well as the light harvesting complexes (LHC, [mmol/mol Chl]). We use data from Ghannoum *et al*. (2005) for abundances of PSI and PSII to include phenomenological stoichiometry rules between LHC and the components of the electron transport chain (Eqns 4–8) and to relate N_{thy} to the amount of cyt (Eqns 9–11). We assume that the chlorophyll content is shared between PSI, PSII, and LHC (Eqns 7 and 8). To be able to consider LET and CET, these complexes are split according to the proportion of LET (p) and CET (1 - p). Indices represent the considered pathway.

$$PSI_{LET} = 2 \cdot p \quad (4)$$

$$PSI_{CET} = 2 \cdot (1 - p) \quad (5)$$

$$PSII = 2.5 \quad (6)$$

$$LHC_{LET} = \frac{1000 \cdot p - PSII \cdot 60 - PSI_{LET} \cdot 184}{13} \quad (7)$$

$$LHC_{CET} = \frac{1000 \cdot (1-p) - PSI_{CET} \cdot 184}{13} \quad (8)$$

For the LET, J_{max} is related to N_{thy} as described in Eqns 9–12. cyt_{Jmax} describes the relation of cyt to J_{max} and was measured by Niinemets and Tenhunen (1997), who determined 156 (mmol e$^-$)/(mmol cyt s) across various C_3 species. Assuming 95% of LET in C_3 plants, this leads to a capacity of 172 (mmol e$^-$)/(mmol cyt s) for cyt_{Jmax}.

$$N_{thy_{LET}} = \frac{n_{Jmax} \cdot N_{ps} \cdot p}{Chl} \quad (9)$$

$$N_{LH_{LET}} = PSII \cdot 83.3 \cdot 0.06 + PSI_{LET} \cdot 32.8 \cdot 0.184 + LHC_{LET} \cdot 26 \cdot 0.013 \quad (10)$$

$$cyt_{LET} = \frac{1}{8.85}\left(N_{thy_{LET}} - N_{LH_{LET}}\right) \quad (11)$$

$$Jmax_{LET} = \max\left(0, \frac{cyt_{LET} \cdot Chl \cdot cyt_{Jmax}}{1000}\right) \quad (12)$$

Chlorophyll content (*Chl*, [µmol m^{-2}]) is calculated based on an empirical factor (Vogan & Sage, 2012) that relates the amount of nitrogen invested into thylakoids (n_{fit} N_t, Eqn 1) to the amount of chlorophyll in C_3 plants:

$$Chl = n_{fit} \cdot N_t \cdot 0.0158887 \quad (13)$$

The response of chlorophyll content to leaf nitrogen does not differ significantly between different photosynthetic types in *Flaveria* (Vogan & Sage, 2011).

The derivation for the CET is analogous to the case of the LET (Eqns 14–17); additionally, the factor $Jmax_{CL}$ is required, which describes the scaling of J_{max} with cyt for the CET:

$$N_{thy_{CET}} = \frac{n_{Jmax} \cdot N_{ps} \cdot (1-p)}{Chl} \quad (14)$$

$$N_{LH_{CET}} = PSI_{CET} \cdot 32.8 \cdot 0.184 + LHC_{CET} \cdot 26 \cdot 0.013 \quad (15)$$

$$cyt_{CET} = \frac{1}{8.85}\left(N_{thy_{CET}} - N_{LH_{CET}}\right) \quad (16)$$

$$Jmax_{CET} = \max\left(0, \frac{cyt_{CET} \cdot Chl \cdot cyt_{Jmax} \cdot Jmax_{CL}}{1000}\right) \quad (17)$$

Optimization procedure

To find the maximal CO_2 assimilation rate under the given environmental, physiological, and biochemical constraints, we optimize the allocation of photosynthetic nitrogen (assumed to depend only on total leaf nitrogen) into Rubisco, C_4 cycle, LET, and CET through an augmented Lagrangian approach. The optimization is constrained to make sure that the results are biologically realistic, e.g., C_3 species were not able to invest nitrogen into the C_4 cycle (see Supporting Information Table S2 for additional details).

The model and its optimization were implemented in the R environment (R Core Team, 2017), using the auglag-function of the package 'nloptr' (Johnson, see Supporting Information for details). The optimization algorithm can use various local solvers; we chose a derivative-free solver, 'COBYLA'. We adapted the parameters of the auglag-function as follows: (1) *xtol_rel*=1x10^{-100}, *i.e.*, we stop the optimization when all parameters changed by a proportion <1x10^{-100} in the last iteration; (2) *localtol*, the tolerance applied in the selected local solver, is set to 1x10^{-100}; and (3) *maxeval*, the maximal number of optimization iterations, is set to 5x10^3. To ensure robust retrieval of the global optimum, we perform a large number of optimizations starting from a wide range of initial values (see Supporting Information for details). The successful run resulting in the maximal CO_2 assimilation rate is used.

Modeling the effect of light

The relationship of the electron transport rate (J_t, [µmol m^{-2} s^{-1}]) and the absorbed light of a certain irradiance (I, [µmol m^{-2} s^{-1}]) is presented in Eqns 18–20. I is related to J_t by a widely accepted empirical hyperbolic function (Eqn 18), (von Caemmerer, 2000; Bernacchi *et al.*, 2003) that includes the following parameters: (1) J_{max}, the maximum electron transport rate; (2) Θ, the convexity of the transition between the initial slope and the plateau of the hyperbola; (3) α, the leaf absorptance; (4) f, a correction factor accounting for the spectral quality of the light; and (5) p, the fraction of absorbed quanta that reaches PSI and PSII of LET (with (1 − p) reaching the CET). I_{abso} is set to I_{LET} and I_{CET} dependent on the considered path of electron transport. The fraction of irradiance that is absorbed by the LET is shared equally between PSI and PSII (resulting in the factor 0.5 in Eqn 19), while the fraction of irradiance that is absorbed by the CET is assumed to reach PSI in full.

$$J_t = \frac{I_{abso}+J_{\max}-\sqrt{(I_{abso}+J_{\max})^2-4\,\theta I_{abso}J_{\max}}}{2\theta} \quad (18)$$

$$I_{LET} = I \cdot \alpha \cdot (1-f) \cdot p \cdot 0.5 \quad (19)$$

$$I_{CET} = I \cdot \alpha \cdot (1-f) \cdot (1-p) \quad (20)$$

In our model it is assumed that the electron transport chain is the only source of ATP and NADPH and that both are used exclusively for CO_2 fixation (von Caemmerer, 2000). As NADPH production results from LET, the amount of electrons is calculated using Eqns 18 and 19. The amount of electrons utilized for ATP production depends on both LET and CET (see below). There are multiple pathways of CET (Kramer & Evans, 2011); the model considers those pathways with an active Q-cycle and a ratio of two protons per electron. Note that Rubisco is assumed to be fully activated, independent of the irradiance (von Caemmerer, 2000).

The available energy needs to be partitioned between five pools: (1) the Calvin-Benson cycle (CBB) in the mesophyll; (2) the CBB in the bundle sheath; (3) the photorespiratory pathway (PR) in the mesophyll; (4) the PR in the bundle sheath cell; and (5) the C_4 pathway. This means that the available energy is calculated in total and then partitioned (Kanai & Edwards, 1999) into J_{mp}, J_{mc}, and J_s, the fractions invested into the C_4 cycle, the CBB and the PR in the mesophyll, and the CBB and the PR in the bundle sheath cell, respectively. During optimization, the activity of each process is constrained by its allocated energy pool, *i.e.*, the energy allocation equals the relative energy allocation of the processes (see Supporting Information Method S1 for details).

The number of electrons transported to generate one molecule of ATP is unknown; for a discussion, see, e.g., Amthor (2010). We address these uncertainties by a factor that represents the ratio of electron transported per ATP in LET, which we set to e_{ATP} = 4/3 in this work. In *Flaveria*, this ratio is supported by Siebke *et al.* (1997). The ATP and the NADPH requirements of the CBB, the PR, and the C_4 cycle are based on the work of von Caemmerer (2000, see Supporting Information for equations). The energy requirements of the C_4 cycle are adequate for the C_4-subtypes that utilize NAD-malic enzyme or NADP-malic enzyme, whose ATP demand can be assumed to be equal. For the C_4-subtype that utilizes PEP carboxykinase, the energetic costs are different and currently unclear (Kanai & Edwards, 1999; von Caemmerer, 2000).

CO_2 assimilation rate

A limitation in the production of both ATP and NADPH arises under light-limited conditions (von Caemmerer, 2000). The ATP-limited CO_2 assimilation rate (A_j^{ATP}) is calculated according to the light-limiting model of von Caemmerer (2000) (see Supporting Information for equations). The NADPH limitation is calculated analogously to the ATP-limited scenario (A_j^{NADPH}, see Supporting Information). The light-limited CO_2 assimilation rate is then:

$$A_j = \min\left(A_j^{ATP}, A_j^{NADPH}\right) \tag{24}$$

The model for the CO_2 assimilation rate when the electron transport rate is not limiting (A_c) is taken from Heckmann *et al.* (2013) and extended by a parameter representing the fraction of PSII activity in the bundle sheath cells, which affects O_2 evolution. This parameter is set to *p*. In the whole model, each limitation is considered independently; the minimal CO_2 assimilation rate determines the limiting process:

$$A = \min\left(A_j, A_c\right) \tag{25}$$

Temperature-dependent model

Temperature affects the CO_2 assimilation rate by changing the maximal activity of the C_4 cycle, the carboxylation rate of Rubisco, and the electron transport rate. Temperature also affects the specificity of Rubisco as well as the Michaelis constants of Rubisco and PEPC. We model the temperature response by an extended Arrhenius function that describes two counteracting effects: rate increases with increasing temperature and enzyme inactivation through thermal instability (Massad *et al.*, 2007). We use

parameters taken from literature or fitted to available data (see Supporting Information for the equation and a full list of parameters and their sources).

Data used in the analyses

As the raw data of Vogan and Sage (2012) was not available, we extracted it from the corresponding figures using the Graph Grabber software provided by Quintessa Limited (Version 1.5.5). The measured curves consider the CO_2 assimilation rate per intercellular CO_2 concentration (C_i). We assume that the mesophyll CO_2 level is 85% of the C_i.

For the detailed analysis of the C_4 plants (Fig. 4), we used data published by Dwyer *et al.* (2007) for the CO_2 assimilation rate at 25°C and 35°C, Rubisco catalytic sites, the PEPC activity, and the nitrogen investment into the thylakoids. As PEPC activity in *Flaveria* does not serve as a proxy for C_4 cycle activity above values of around 130 µmol m^{-2} s^{-1} (Heckmann *et al.*, 2013), the maximal PEPC activity in C_4 plants is set to 130 µmol m^{-2} s^{-1} (see Supporting Information Table S1).

Required nitrogen re-allocation (δ_n)

Required nitrogen re-allocation (δ_n, [fraction]) is defined as the total fraction of nitrogen that needs to be re-allocated between photosynthetic pools to optimally adjust photosynthesis from the evolutionary scenario (n_{Etot}^{evo}, n_{C4}^{evo}, n_{Jmax}^{evo}) to a given growth environment (n_{Etot}^{growth}, n_{C4}^{growth}, n_{Jmax}^{growth}):

$$\delta_n = \sum_{\mathrm{i}\in\{Etot,C4,Jmax\}} \left| n_i^{evo} - n_i^{growth} \right| \tag{26}$$

Statistical information

The differences between adaptation scenarios are tested with the Wilcoxon rank sum test. Due to computational limitations, only a limited number of leaf nitrogen levels can be used to calculate the resource allocation for the data set of Vogan and Sage (2011) (Fig. 3). We considered 16 leaf nitrogen levels for the calculation of the resource allocation and CO_2 assimilation rates. We inferred the CO_2 assimilation rates required for the remaining leaf nitrogen levels from linear interpolation between the two closest leaf nitrogen levels. For the statistical analysis, the data of the modeled species, *F. pringlei* (C_3), *F. floridana* (C_3-C_4), *F. palmeri* (C_4-like), and *F. bidentis* (C_4), was considered. All statistical analyses were conducted in R (R Core Team, 2017).

The difference of δ_n for various photosynthetic types was tested by a sign test, applied to the data of Vogan and Sage (2011) (Supporting Information Table S5).

Acknowledgements

We thank Prof. Rowan F. Sage and Prof. Susanne von Caemmerer for providing raw data from their publications. We further acknowledge financial support by the Deutsche Forschungsgemeinschaft (grants IRTG 1525 to DH, EXC 1028 to MJL and ES, CRC 680 and CRC 1310 to MJL). We would like to thank Bernhard O. Palsson and his group for hosting ES and for valuable discussions.

Author Contributions

ES, MJL, and DH designed the research, interpreted the results, and wrote the paper. ES developed and implemented the model for nitrogen allocation and light reactions, and implemented the optimization procedure. DH developed and implemented the model for temperature responses. ES and DH conducted simulations and data analysis.

References

Amthor JS. 2010. From sunlight to phytomass: on the potential efficiency of converting solar radiation to phyto-energy. *New Phytologist* **188**(4): 939-959.

Atkinson D. 1969. Limitation of metabolite concentrations and the conservation of solvent capacity in the living cell. *Current topics in cellular regulation* **1**: 29-43.

Baudouin-Cornu P, Surdin-Kerjan Y, Marliere P, Thomas D. 2001. Molecular evolution of protein atomic composition. *Science* **293**(5528): 297-300.

Beg QK, Vazquez A, Ernst J, de Menezes MA, Bar-Joseph Z, Barabasi AL, Oltvai ZN. 2007. Intracellular crowding defines the mode and sequence of substrate uptake by Escherichia coli and constrains its metabolic activity. *Proceedings of the National Academy of Sciences of the United States of America* **104**(31): 12663-12668.

Bernacchi CJ, Pimentel C, Long SP. 2003. In vivo temperature response functions of parameters required to model RuBP-limited photosynthesis. *Plant Cell and Environment* **26**(9): 1419-1430.

Berry JA, Farquhar GD 1978. The CO_2 concentrating function of C_4 photosynthesis: a biochemical model. *Proceedings of the Fourth International Congress on Photosynthesis. Biochemical Society, London*. 119-131.

Christin PA, Osborne CP, Sage RF, Arakaki M, Edwards EJ. 2011. C_4 eudicots are not younger than C_4 monocots. *Journal of Experimental Botany* **62**(9): 3171-3181.

de Oliveira Dal'Molin CG, Quek LE, Palfreyman RW, Brumbley SM, Nielsen LK. 2010. AraGEM, a genome-scale reconstruction of the primary metabolic network in Arabidopsis. *Plant Physiology* **152**: 579.

Dourado H, Maurino VG, Lercher MJ. 2017. Enzymes And Substrates Are Balanced At Minimal Combined Mass Concentration In Vivo. *bioRxiv*.

Dwyer SA, Ghannoum O, Nicotra A, Von Caemmerer S. 2007. High temperature acclimation of C_4 photosynthesis is linked to changes in photosynthetic biochemistry. *Plant Cell and Environment* **30**(1): 53-66.

Ellis RJ. 1979. Most Abundant Protein in the World. *Trends in Biochemical Sciences* **4**(11): 241-244.

Evans JR, Seemann JR. 1989. The allocation of protein nitrogen in the photosynthetic apparatus: costs, consequences, and control. *Photosynthesis*: 183-205.

Evans JR, von Caemmerer S 2000. Would C_4 rice produce more biomass than C_3 rice? In: Sheehy JE, Mitchell PL, Hardy B eds. *Redesigning rice photosynthesis to increase yield*: Elsevier, 53-71.

Farquhar GD, Caemmerer S, Berry JA. 1980. A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. *Planta* **149**: 78-90.

Friend AD. 1991. Use of a Model of Photosynthesis and Leaf Microenvironment to Predict Optimal Stomatal Conductance and Leaf Nitrogen Partitioning. *Plant Cell and Environment* **14**(9): 895-905.

Gerhart LM, Ward JK. 2010. Plant responses to low CO_2 of the past. *New Phytologist* **188**(3): 674-695.

Ghannoum O, Evans JR, Chow WS, Andrews TJ, Conroy JP, von Caemmerer S. 2005. Faster rubisco is the key to superior nitrogen-use efficiency in NADP-malic enzyme relative to NAD-malic enzyme C_4 grasses. *Plant Physiology* **137**(2): 638-650.

Harrison MT, Edwards EJ, Farquhar GD, Nicotra AB, Evans JR. 2009. Nitrogen in cell walls of sclerophyllous leaves accounts for little of the variation in photosynthetic nitrogen-use efficiency. *Plant Cell and Environment* **32**(3): 259-270.

Heckmann D, Schulze S, Denton A, Gowik U, Westhoff P, Weber A PM, Lercher Martin J. 2013. Predicting C_4 Photosynthesis Evolution: Modular, Individually Adaptive Steps on a Mount Fuji Fitness Landscape. *Cell* **153**(7): 1579-1588.

Ibarra RU, Edwards JS, Palsson BO. 2002. Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* **420**(6912): 186-189.

Johnson SG The NLopt nonlinear-optimization package.

Kanai R, Edwards GE 1999. The biochemistry of C_4 photosynthesis. In: Sage RF, Monson RK eds. *C4 plant biology*: Academic press, Toronto, ON, Canada, 49-87.

Kramer DM, Evans JR. 2011. The Importance of Energy Balance in Improving Photosynthetic Productivity. *Plant Physiology* **155**(1): 70-78.

Maire V, Martre P, Kattge J, Gastal F, Esser G, Fontaine S, Soussana JF. 2012. The Coordination of Leaf Photosynthesis Links C and N Fluxes in C_3 Plant Species. *PLoS ONE* **7**(6).

Makino A, Sakuma H, Sudo E, Mae T. 2003. Differences between maize and rice in N-use efficiency for photosynthesis and protein allocation. *Plant and Cell Physiology* **44**(9): 952-956.

Malhi SS, Grant CA, Johnston AM, Gill KS. 2001. Nitrogen fertilization management for no-till cereal production in the Canadian Great Plains: a review. *Soil & Tillage Research* **60**(3-4): 101-122.

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* **10.7554/eLife.02478**.

Massad RS, Tuzet A, Bethenod O. 2007. The effect of temperature on C_4-type leaf photosynthesis parameters. *Plant Cell and Environment* **30**(9): 1191-1204.

Maurino VG, Peterhansel C. 2010. Photorespiration: current status and approaches for metabolic engineering. *Current Opinion in Plant Biology* **13**(3): 249-256.

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

Mori M, Schink S, Erickson DW, Gerland U, Hwa T. 2017. Quantifying the benefit of a proteome reserve in fluctuating environments. *Nature Communications* **8**.

Munekage YN, Taniguchi YY. 2016. Promotion of Cyclic Electron Transport Around Photosystem I with the Development of C_4 Photosynthesis. *Plant and Cell Physiology* **57**(5): 897-903.

Niinemets U, Tenhunen JD. 1997. A model separating leaf structural and physiological effects on carbon gain along light gradients for the shade-tolerant species *Acer saccharum*. *Plant Cell and Environment* **20**(7): 845-866.

Oberhardt MA, Palsson BO, Papin JA. 2009. Applications of genome-scale metabolic reconstructions. *Molecular Systems Biology* **5**.

Powell AM. 1978. Systematics of *Flaveria* (Flaveriinae-Asteraceae). *Annals of the Missouri Botanical Garden*: 590-636.

R Core Team 2017. R: A Language and Environment for Statistical Computing: R Foundation for Statistical Computing.

Reimers AM, Knoop H, Bockmayr A, Steuer R. 2017. Cellular trade-offs and optimal resource allocation during cyanobacterial diurnal growth. *Proc Natl Acad Sci U S A*.

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

Sage RF, Cowling SA. 1999. Implications of stress in low CO_2 atmospheres of the past: are today's plants too conservative for a high CO_2 world. *Carbon dioxide and environmental stress*: 289-308.

Sage RF, McKown AD. 2006. Is C_4 photosynthesis less phenotypically plastic than C_3 photosynthesis? *Journal of Experimental Botany* **57**(2): 303-317.

Savir Y, Noor E, Milo R, Tlusty T. 2010. Cross-species analysis traces adaptation of Rubisco toward optimality in a low-dimensional landscape. *Proceedings of the National Academy of Sciences* **107**(8): 3475-3480.

Siebke K, vonCaemmerer S, Badger M, Furbank RT. 1997. Expressing an RbcS antisense gene in transgenic Flaveria bidentis leads to an increased quantum requirement for CO2 fixed in photosystems I and II. *Plant Physiology* **115**(3): 1163-1174.

Vance CP. 2001. Symbiotic nitrogen fixation and phosphorus acquisition. Plant nutrition in a world of declining renewable resources. *Plant Physiology* **127**(2): 390-397.

Vogan PJ, Sage RF. 2011. Water-use efficiency and nitrogen-use efficiency of C_3-C_4 intermediate species of *Flaveria* Juss. (Asteraceae). *Plant, Cell & Environment* **34**: 1415-1430.

Vogan PJ, Sage RF. 2012. Effects of low atmospheric CO_2 and elevated temperature during growth on the gas exchange responses of C_3, C_3-C_4 intermediate, and C_4 species from three evolutionary lineages of C_4 photosynthesis. *Oecologia* **169**(2): 341-352.

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**(4): 463-474.

von Caemmerer S. 2000. *Biochemical models of leaf photosynthesis*. Collingwood, Australia: Csiro Publishing.

Wang Y, Long SP, Zhu XG. 2014. Elements Required for an Efficient NADP-Malic Enzyme Type C_4 Photosynthesis. *Plant Physiology* **164**(4): 2231-2246.

Zhu X-G, de Sturler E, Long SP. 2007. Optimizing the distribution of resources between enzymes of carbon metabolism can dramatically increase photosynthetic rate: a numerical simulation using an evolutionary algorithm. *Plant Physiol* **145**: 513-526.

Figure legend

Table 1 In C_4 and C_4-like plants, the evolutionary scenario shows significantly smaller residual sum of squares compared to the growth scenario. The residual sum of squares for the evolutionary and growth

scenario, each photosynthetic type, and all measured curves of Vogan and Sage (2011) and Vogan and Sage (2012) are presented.

Table 2 A list of all parameters used in the mathematical model. For each parameter the abbreviation, the explanation, and the unit are presented.

Figure 2 Model results based on optimality in the evolutionary scenario (solid lines) describe the measured data (dots ± SE) better than the model assuming optimal adaptation to the growth conditions (dashed lines) for *F. robusta* (C_3), *F. ramosissima* (C_3-C_4), and *F. bidentis* (C_4) grown at the current CO_2 level (data from Vogan and Sage (2012)). (a) The net CO_2 assimilation rate as a function of intercellular CO_2 concentration measured at 30°C. (b) The net CO_2 assimilation rate as a function of temperature.

Figure 3 The dependence of the CO_2 assimilation rate on leaf nitrogen levels for various *Flaveria* species is consistent with model results based on optimality in the evolutionary scenario (solid lines). For C_3-C_4 intermediate, C_4-like, and C_4 these results outperform the ones assuming optimal phenotypic adaptation to the growth conditions (dashed lines). The modeled species are *F. pringlei* (C_3)*, F. floridana* (C_3-C_4)*, F. palmeri* (C_4-like)*, and F. bidentis* (C_4) (data from Vogan and Sage (2011)).

Figure 4 A detailed analysis of resource allocation and physiology in *F. bidentis* (C_4) shows a good agreement between experimental data (Dwyer *et al.*, 2007) and model results based on the evolutionary scenario (orange dots). Alternative model results assuming optimal phenotypic adaptation to the growth scenario consistently show higher disagreement with the data (purple dots). Values are mean log2(modeled results/measured data) ± SE. (a) Plants grown at 25°C (b) Plants grown at 35°C. *A* = net CO_2 assimilation rate; *N* = nitrogen.

Figure 5 Discrepancy between measured and modeled *F. bidentis* data across diverse environments. The black dot indicates the environment that best explains the experimental data of Dwyer *et al.* (2007). The deviation between model predictions and measurements ('error') is defined as the mean of the squared residuals (which are expressed as fractions of experimental means).

Supporting Information

Additional supporting information may be found in the online version of this article.

Methods S1 Details about the optimization procedure of resource allocation

Methods S2 Equations of the energetic costs

Methods S3 Equations of the light-limited CO_2 assimilation rate

Methods S4 Details about the temperature-dependent model

Methods S5 Sensitivity analysis

Table S1 *Flaveria* parametrization.

Table S2 Lower and upper bounds for the model parameters subject to numerical optimization.

Table S3 The parameters of the temperature-dependent model.

Table S4 Required nitrogen re-allocation (δ_n) for *F. bidentis* (C_4) grown at different temperatures.

Table S5 Required nitrogen re-allocation (δ_n) for different on leaf nitrogen level for various *Flaveria* species.

Table S6 Required nitrogen re-allocation (δ_n) for various *Flaveria* species grown at current or low CO_2 level.

Table S7 The modeled and measured data of chlorophyll and PSII of *F. bidentis* (C_4).

Table S8 Distribution parameters used to generate the random parameter sets for the sensitivity.

Fig. S1 Sensitivity analysis.

Fig. S2 A-C_i curve measured at 40°C using plants grown at the current CO_2 level.

Fig. S3 A-C_i curve measured at 30°C using plants grown at the low CO_2 level.

Fig. S4 A-C_i curve measured at 40°C using plants grown at the low CO_2 level.

Fig. S5 A-Temperature curve using plants grown at the low CO_2 level

Fig. S6 Discrepancy between measured and modeled results of *F. robusta* (C_3) across diverse environments assuming no phosphate-limitation.

Fig. S7 Discrepancy between measured and modeled results of *F. robusta* (C_3) across diverse environments assuming phosphate-limitation.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

Table 1 In C_4 and C_4-like plants, the evolutionary scenario shows significantly smaller residual sum of squares compared to the growth scenario. The residual sum of squares for the evolutionary and growth scenario, each photosynthetic type, and all measured curves of Vogan and Sage (2011) and Vogan and Sage (2012) are presented.

		C_3	C_3-C_4 intermediate	C_4-like	C_4
Evolutionary scenario	Fig. 2a	58.1	77.1		93.4
	Fig. 2b	823.7	524.6		155.6
	Fig. 3	549.2	1554.3	1443.5	834.9
	Supporting Information Fig. S2	616.3	299.1		136.9
	Supporting Information Fig. S3	39.9	40.4		166.0
	Supporting Information Fig. S4	137.0	85.6		286.4
	Supporting Information Fig. S5	14.5	93.9		275.5
Growth scenario	Fig. 2a	602.2	238.2		2454.5
	Fig. 2b	755.2	306.1		340.3
	Fig. 3	386.5	2122.2	3052.2	1873.2
	Supporting Information Fig. S2	252.7	84.9		433.9
	Supporting Information Fig. S3	140.84	50.7		436.1
	Supporting Information Fig. S4	97.8	38.8		460.0
	Supporting Information Fig. S5	13.4	53.9		142.8

Table 2 A list of all parameters used in the mathematical model. For each parameter the abbreviation, the explanation, and the unit are presented.

Abbreviation	Explanation	Unit
A	Achieved CO_2 assimilation rate $(\mathrm{A} = \min(A_j, A_c))$	µmol m^{-2} s^{-1}
A_c	CO_2 assimilation rate when the electron transport rate is not limiting	µmol m^{-2} s^{-1}
A_j	Light-limited CO2 assimilation rate $(A_j = \min(A_j^{ATP}, A_j^{NADPH}))$	µmol m^{-2} s^{-1}
A_j^{ATP}	Light -limited CO2 assimilation rate that is determined based on the availability of ATP	µmol m^{-2} s^{-1}
A_j^{NADPH}	Light-limited CO2 assimilation rate that is determined based on the availability of NADPH	µmol m^{-2} s^{-1}
Chl	Chlorophyll content	µmol m^{-2}
cyt	The amount of cytochrome f	mmol/mol Chl
cyt_{Jmax}	The relation of cyt to J_{max}	unitless
e_{ATP}	Ratio of electron transported per ATP in the linear electron transport	unitless
E_{tot}	The amount of catalytic sites of Rubisco	µmol m^{-2}
f	A correction factor accounting for the spectral quality of the light	unitless
g_s	The bundle sheath cell conductance	µmol m^{-2} s^{-1}
I	The absorbed light	µmol m^{-2} s^{-1}
J_{max}	The maximal electron transport rate	µmol m^{-2} s^{-1}
$Jmax_{CET}$	The maximal electron transport rate of the cyclic electron transport	µmol m^{-2} s^{-1}

$Jmax_{CL}$	A factor that describes the scaling of maximal electron transport rate with cytochrome f for the CET	unitless
$Jmax_{LET}$	The maximal electron transport rate of the linear electron transport	µmol m^{-2} s^{-1}
J_t	Electron transport rate	µmol m^{-2} s^{-1}
J_{mc}	Electron transport rate that is available for the Calvin-Benson cycle and the photorespiratory path in the mesophyll cell	µmol m^{-2} s^{-1}
J_{mp},	Electron transport rate that is available for the C_4 cycle	µmol m^{-2} s^{-1}
J_s	Electron transport rate that is available for the Calvin-Benson cycle and the photorespiratory path in the bundle sheath cell	µmol m^{-2} s^{-1}
I_{CET}	Electron transport rate of the cyclic electron transport	µmol m^{-2} s^{-1}
I_{LET}	Electron transport rate of the linear electron transport	µmol m^{-2} s^{-1}
k_{ccat}	Turn-over rate of Rubisco	1/s
$kcat_{PEPC}$	Turn-over rate of PEPC	1/s
$kcat_{PPDK}$	Turn-over rate of PPDK	1/s
K_p	Michaelis constant of PEPC for bicarbonate	µbar
LHC	Light harvesting complexes	mmol/mol Chl
MW_{PEPC}	The nitrogen requirement of a catalytic site of PEPC	Da
MW_{PPDK}	The nitrogen requirement of a catalytic site of PPDK	Da
n_{C4}	The fraction of photosynthetic nitrogen pool invested into main the enzymes of the C_4 cycle, PEPC and PPDK	fraction
n_{C4}^{evo}	The optimal fraction of photosynthetic nitrogen pool invested into main the enzymes of the C_4 cycle under the evolutionary scenario	fraction

n_{C4}^{growth}	The optimal fraction of photosynthetic nitrogen pool invested into main the enzymes of the C_4 cycle under the growth scenario	fraction
n_{Etot}	The fraction of photosynthetic nitrogen pool invested into the Calvin-Benson cycle	fraction
n_{Etot}^{evo}	The optimal fraction of photosynthetic nitrogen pool invested into the Calvin-Benson cycle under the evolutionary scenario	fraction
n_{Etot}^{growth}	The optimal fraction of photosynthetic nitrogen pool invested into the Calvin-Benson cycle under the growth scenario	fraction
n_{fit}	The proportion of nitrogen invested into the thylakoids as a function of the leaf nitrogen level (a fit to empirical data)	fraction
n_{Jmax}	The fraction of photosynthetic nitrogen pool invested into the thylakoids, which include the electron transport chains	fraction
n_{Jmax}^{evo}	The optimal fraction of photosynthetic nitrogen pool invested into the thylakoids, which include the electron transport chains under the evolutionary scenario	fraction
n_{Jmax}^{growth}	The optimal fraction of photosynthetic nitrogen pool invested into the thylakoids, which include the electron transport chains under the growth scenario	fraction
N_{ps}	Photosynthetic nitrogen pool	µmol m^{-2}
N_t	Total leaf nitrogen	µmol m^{-2}
N_{thy}	Nitrogen invested into the thylakoids ($N_{thy} = N_t\ n_{thy}$)	µmol m^{-2}
p	proportion of linear electron transport	fraction
PSI	Photosystem I	mmol/mol Chl
PSII	Photosystem II	mmol/mol Chl
V_{pmax}	Maximal C_4 cycle activity	µmol m^{-2} s^{-1}

α	Leaf absorptance	fraction
β	Rubisco distribution between mesophyll and bundle sheath cells	fraction
δ_n	Required nitrogen re-allocation	fraction
Θ	The convexity of the transition between the initial slope and the plateau of the hyperbola	unitless
ξ	The fraction of glycine decarboxylated by the glycine decarboxylase complex in the bundle sheath cell that is derived from oxygenation by Rubisco in the mesophyll cell	fraction

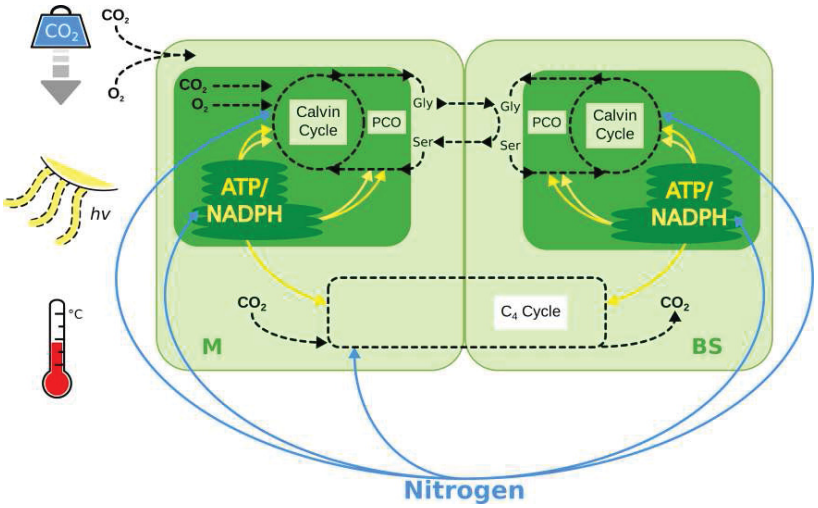


Figure 1 An overview of the nitrogen-dependent light- and enzyme-limited model. CO_2 entering the mesophyll cell (M) can be fixed by Rubisco (C_3 and intermediates) or PEPC (C_4 and intermediates); The C_4 cycle then shuttles CO_2 fixed by PEPC to the bundle sheath cell (BS) and releases it, allowing it to be re-fixed by Rubisco. The fixation of O_2 by Rubisco leads to photorespiration (PCO). Blue arrows indicate the nitrogen allocation and yellow arrows represent the energy allocation considered in the model.

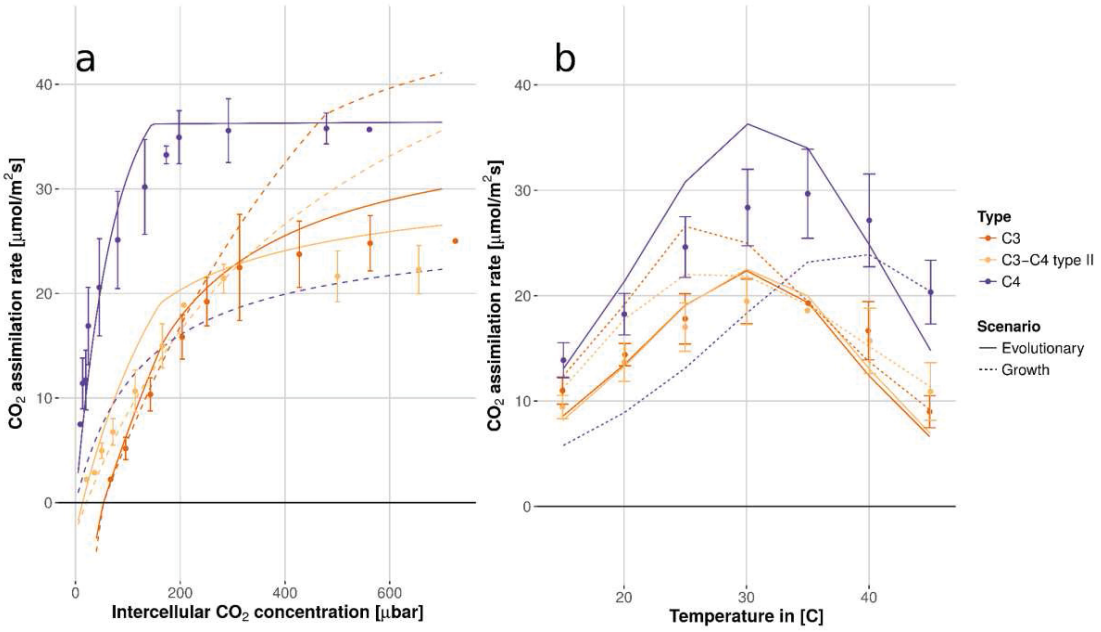


Figure 2 Model results based on optimality in the evolutionary scenario (solid lines) describe the measured data (dots ± SE) better than the model assuming optimal adaptation to the growth conditions (dashed lines) for *F. robusta* (C_3), *F. ramosissima* (C_3-C_4), and *F. bidentis* (C_4) grown at the current CO_2 level (data from Vogan and Sage (2012)). (a) The net CO_2 assimilation rate as a function of intercellular CO_2 concentration measured at 30°C. (b) The net CO_2 assimilation rate as a function of temperature.

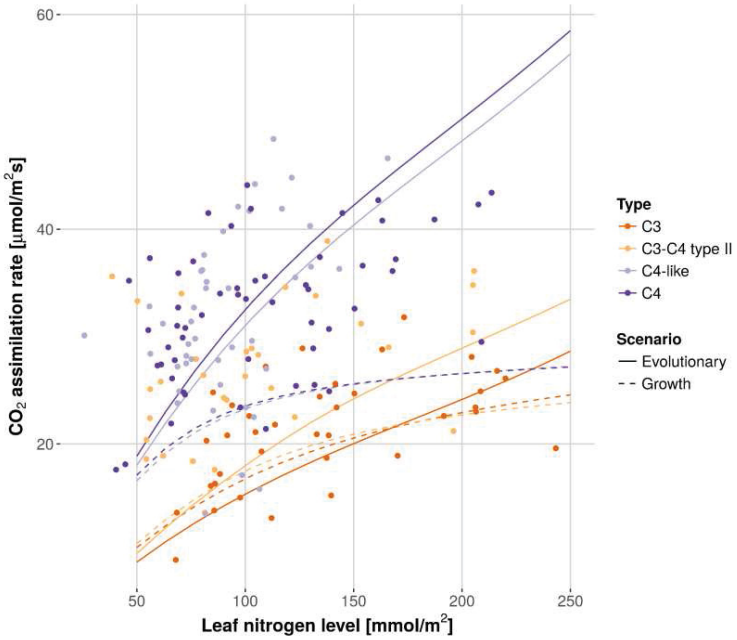


Figure 3 The dependence of the CO_2 assimilation rate on leaf nitrogen levels for various *Flaveria* species is consistent with model results based on optimality in the evolutionary scenario (solid lines). For C_3-C_4 intermediate, C_4-like, and C_4 these results outperform the ones assuming optimal phenotypic adaptation to the growth conditions (dashed lines). The modeled species are *F. pringlei* (C_3)*, F. floridana* (C_3-C_4)*, F. palmeri* (C_4-like)*, and F. bidentis* (C_4) (data from Vogan and Sage (2011)).

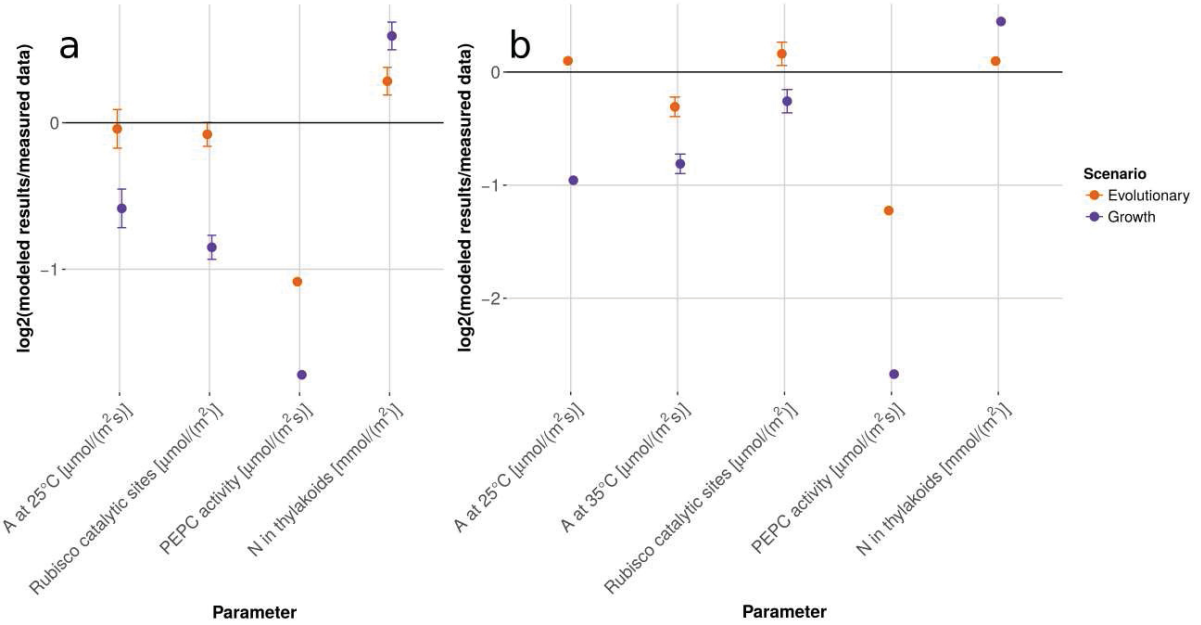


Figure 4 A detailed analysis of resource allocation and physiology in *F. bidentis* (C_4) shows a good agreement between experimental data (Dwyer *et al.*, 2007) and model results based on the evolutionary scenario (orange dots). Alternative model results assuming optimal phenotypic adaptation to the growth scenario consistently show higher disagreement with the data (purple dots). Values are mean log2(modeled results/measured data) ± SE. (a) Plants grown at 25°C (b) Plants grown at 35°C. *A* = net CO_2 assimilation rate; *N* = nitrogen.

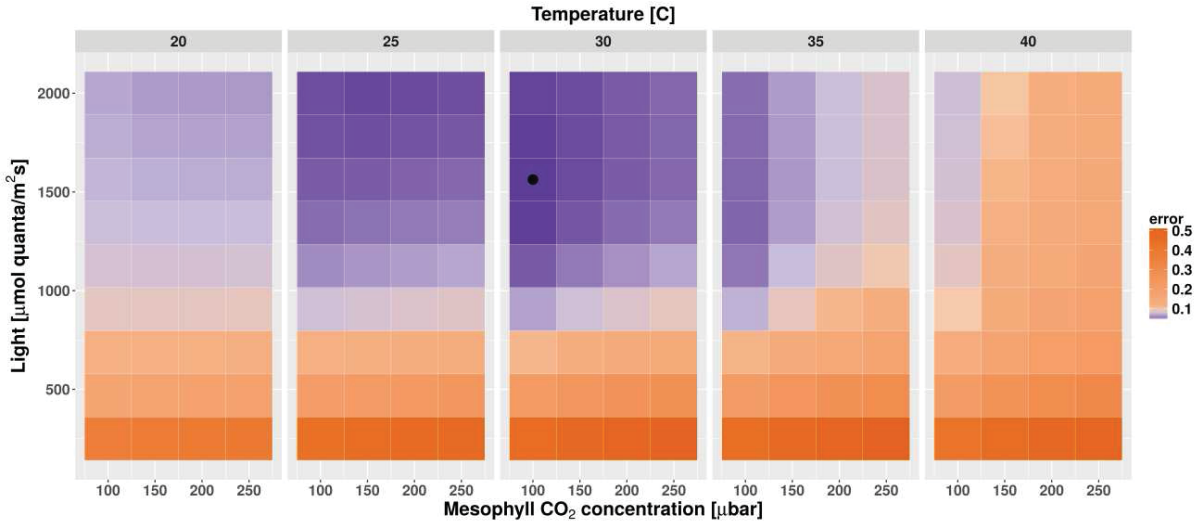


Figure 5 Discrepancy between measured and modeled *F. bidentis* data across diverse environments. The black dot indicates the environment that best explains the experimental data of Dwyer *et al.* (2007). The deviation between model predictions and measurements ('error') is defined as the mean of the squared residuals (which are expressed as fractions of experimental means).

New Phytologist Supporting Information

Article title: Modeling Cellular Resource Allocation Reveals Low Phenotypic Plasticity of C_4 Plants and Infers Environments of C_4 Photosynthesis Evolution
Authors: Esther M. Sundermann, Martin J. Lercher, David Heckmann
Article acceptance date: Click here to enter a date.

The following Supporting Information is available for this article:

Methods S1 Details about the optimization procedure of resource allocation

Methods S2 Equations of the energetic costs

Methods S3 Equations of the light-limited CO_2 assimilation rate

Methods S4 Details about the temperature-dependent model

Methods S5 Sensitivity analysis

Table S1 *Flaveria* parametrization.

Table S2 Lower and upper bounds for the model parameters subject to numerical optimization.

Table S3 The parameters of the temperature-dependent model.

Table S4 Required nitrogen re-allocation (δ_n) for *F. bidentis* (C_4) grown at different temperatures.

Table S5 Required nitrogen re-allocation (δ_n) for different on leaf nitrogen level for various *Flaveria* species.

Table S6 Required nitrogen re-allocation (δ_n) for various *Flaveria* species grown at current or low CO_2 level.

Table S7 The modeled and measured data of chlorophyll and PSII of *F. bidentis* (C_4).

Table S8 Distribution parameters used to generate the random parameter sets for the

sensitivity.

Fig. S1 Sensitivity analysis.

Fig. S2 A-C_i curve measured at 40°C using plants grown at the current CO_2 level.

Fig. S3 A-C_i curve measured at 30°C using plants grown at the low CO_2 level.

Fig. S4 A-C_i curve measured at 40°C using plants grown at the low CO_2 level.

Fig. S5 A-Temperature curve using plants grown at the low CO_2 level

Fig. S6 Discrepancy between measured and modeled results of *F. robusta* (C_3) across diverse environments assuming no phosphate-limitation.

Fig. S7 Discrepancy between measured and modeled results of *F. robusta* (C_3) across diverse environments assuming phosphate-limitation.

Methods S1 Details about the optimization procedure of resource allocation

To restrict the results of the optimization of CO_2 assimilation rates to biologically relevant resource allocation patterns, we enforce a set of constraints. The relative contributions of the components of the following three pools have to sum up to one in each pool: (1) the nitrogen investments into the enzymes of the Calvin-Benson cycle in the mesophyll or bundle sheath cells, the C_4 cycle, or the thylakoids including the cost of the linear and cyclic electron transport; (2) the ATP investments into the C_4 cycle, Rubisco activity in mesophyll and bundle sheath cells, as well as the non-photochemical quenching (which is nearly zero, in case of optimality); and (3) the NADPH investments into the NADPH-relevant sub-pools. Further constraints ensure that the electron transport rate does not exceed the rate sustained by current irradiance and that the production of ATP and NADPH has to fulfill or exceed the respective consumption. When calculating the light- and enzyme-limited CO_2 fixation rate in the bundle sheath cells, the resulting

quadratic equations can only be solved if the radicands are larger or equal to zero (Eqns S8 and S10). Table S2 shows the lower and upper bounds of the parameters that are optimized. All parameters represent fractions, therefore all lower and upper bounds have to be between zero and one. The upper bound of n_{C4} (n_{C4}^{max}) is based on the empirical maximal C_4 cycle activity (V_{pmax}^{emp}) and the photosynthetic nitrogen level (Eqn S1).

$$n_{C4}^{max} = \frac{V_{pmax}^{emp}\left(\frac{MW^*_{PPDK}}{kcat_{PPDK}}+\frac{MW^*_{PEPC}}{kcat_{PEPC}}\right)}{N_{ps}} \tag{S1}$$

The lower bound of p is set close to zero to avoid division by zero. The lower bound of n_{Jmax} (n_{Jmax}^{min}) ensures that the nitrogen requirements of PSI, PSII, and LHC are met. As n_{Jmax} depends on the photosynthetic nitrogen level and the proportion of the LET, a parameter to be optimized, this dependency results in a bound (Eqn S2, also see Eqn 15) and a constraint that is relevant during the optimization procedure (Eqn S3, also see Eqn 10 and 15). n_{Jmax}^{min} is calculated for each scenario and photosynthetic type separately.

$$n_{Jmax}^{min} = \left(\frac{Chl}{N_{ps}}\right)\left(PSI \cdot 32.8 \cdot 0.184 + \left(\frac{1000 - PSI \cdot 184}{13}\right) \cdot 26 \cdot 0.013\right) \tag{S2}$$

$$n_{Jmax} \geq c_{Jmax}^{\min} \tag{S3}$$

with

$$c_{Jmax}^{\min} = \max(c_{Jmax}^{LET}, c_{Jmax}^{CET})$$

$$c_{Jmax}^{LET} = \left(\frac{Chl}{N_{ps} \cdot p}\right)(PSII \cdot 83.3 \cdot 0.06 + PSI_{LET} \cdot 32.8 \cdot 0.184 + LHC_{LET} \cdot 26 \cdot 0.013)$$

$$c_{Jmax}^{CET} = \left(\frac{Chl}{N_{ps}(1-p)}\right)(PSI_{CET} \cdot 32.8 \cdot 0.184 + LHC_{CET} \cdot 26 \cdot 0.013)$$

To ensure robust retrieval of the global optimum, up to 735 initial values are used for the optimization procedure. We use equidistant points that span the range of minimum and maximum bounds for each nitrogen pool (Table S2). As optimal energy allocation is a function of the nitrogen pools, this can lead to unrealistic values of energy demand in some initial points. We thus excluded points for which absolute sum of energy allocation fractions exceed 1000. For the proportion of the LET, there are seven initial values that cover the expected range of values (namely those are 0.5, 0.6, 0.65, 0.75, 0.85, 0.95, 0.99, see Yin and Struik (2018)).

Methods S2 Equations of the energetic costs

The equations describing energetic costs are based on the work of von Caemmerer (2000). The variables are defined in the main text and in the Supporting Information text above. Additionally, *O* represents the O_2 concentration in the considered cell type (O_m or O_s), and *C* represents C_m or C_s.

The ATP requirements of the Calvin-Benson cycle (CBB), the photorespiratory path (PR), and the C4 cycle are:

$$E_{ATP_{CBB,PR}}(O,C) = 3\, e_{ATP}\left(1 + \frac{7\,\gamma_* O}{3C}\right) \tag{S4}$$

$$E_{ATP\,C4} = 2\, e_{ATP} \tag{S5}$$

NADPH is required for the CBB and the PR, but not for the C4 cycle. The requirements are:

$$E_{NADPH_{CBB,PR}}(O,C) = 2\left(2 + \frac{4\,\gamma_* O}{C}\right) \tag{S6}$$

Methods S3 Equations of the light-limited CO2 assimilation rate

The equations describing the ATP-limited case are derived from the C_3-C_4 model of von Caemmerer (2000). The variables are defined in the main text and in the Supporting Information text above. Additional variables are (1) the CO_2 concentration in the mesophyll cell (C_m); (2) the O_2 concentration in the mesophyll cell (O_m); (3) the mitochondrial respiration in the mesophyll cell (R_m); and (4) the mitochondrial respiration in the bundle sheath cell (R_s). Note that the fraction of PSII activity in the bundle sheath cells is set to *p*.

Considering a variable electron to ATP ratio (e_{ATP}) results in the following equations for the rate of CO_2 fixation in the mesophyll and bundle sheath cell, respectively:

$$A_m = \frac{(C_m - \gamma_* O_m) J_{mc}}{3\, e_{ATP}\left(C_m + \frac{7\,\gamma_* O_m}{3}\right)} \tag{S7}$$

$$A_s = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \tag{S8}$$

with

$$S = \frac{J_{mp}}{2\ e_{ATP}} + \frac{\xi \gamma_* O_m\ J_{mc}}{e_{ATP}(3\ C_m + 7\ \gamma_* O_m)}$$

$$a = \frac{1}{4}\left(3\ e_{ATP} - \frac{7\ e_{ATP}\ \gamma_*\ p}{0.047}\right)$$

$$b = \frac{1}{4}\left(3\ e_{ATP}\ R_s - 3\ e_{ATP}\ S - 3\ g_s\ C_m\ e_{ATP} - J_s - \frac{7\ R_s\ e_{ATP}\ p\ \gamma_*}{0.047} - 7\ g_s\ e_{ATP}\gamma_*\ O_m - \frac{J_s\ \gamma_*\ p}{0.047}\right)$$

$$c = \frac{1}{4}\left((S + g_s\ C_m)\ (J_s - 3\ R_s\ e_{ATP}) - g_s\ \gamma_*\ O_m(7\ R_s\ e_{ATP} + J_s\ )\right)$$

Since NADPH does not affect the C4 cycle, its NADPH-limited rate, V_p, is at its maximal value, V_{pmax}. The equations of the NADPH-limited case are as follows:

$$A_m = \frac{\left(1 - \frac{\gamma_* O_m}{C_m}\right) J_{mc}}{4 + \frac{8\ \gamma_* O_m}{C_m}} \tag{S9}$$

$$A_s = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \tag{S10}$$

with

$$S = g_s\ C_m + V_{pmax} + \left(\frac{\frac{\xi\ J_{mc}\ \gamma_*\ O_m}{C_m}}{4 + \frac{8\ \gamma_* O_m}{C_m}}\right)$$

$$a = -4 + \frac{8\ p\ \gamma_*}{0.047}$$

$$b = -4\ R_s + J_s + 4\ S + \frac{p\ \gamma_*}{0.047}(8\ R_s + J_s) + 8\ g_s\ O_m\ \gamma_*$$

$$c = S(4\ R_s - J_s) + g_s\ O_m\ \gamma_*(8\ R_s + J_s)$$

Methods S4 Details about the temperature-dependent model

The extended Arrhenius function is given by Massad *et al.* (2007):

$$f(T) = k_{25} \exp\left[\ E\ \frac{T - 298.15}{298.15\ R\ T}\right] \frac{\left[1 + \exp\left(\frac{298.15\ S - H}{298.15\ R}\right)\right]}{\left[1 + \exp\left(\frac{T\ S - H}{T\ R}\right)\right]} \tag{S11}$$

The parameters of the extended Arrhenius function are: (1) the value of the considered enzyme at temperatures 25°C (k_{25}); (2) the activation energy (E); (3) the deactivation energy (H); (4) an

entropy factor (*S*); (5) the universal gas constant (*R*); and (6) the temperature considered (*T*). Table S3 shows the parameters for each temperature-dependent variable.

The values required to describe the temperature response of the Rubisco activity (V_{cmax}, [µmol m^{-2} s^{-1}]), the Rubisco specificity (2γ∗), and the Michaelis constants of Rubisco for CO_2 (K_c, [µbar]) and O_2 (K_o, [µbar]) were fitted to the data of Ku *et al.* (1991) simultaneously with the number of Rubisco catalytic sites. We approximate the temperature response of g_s with the temperature response of the diffusion coefficient of CO_2 in water.

Methods S5 Sensitivity analysis

The following parameters are considered in the sensitivity analysis: (1) the mesophilic CO_2 concentration (C_m); (2) the proportion of nitrogen invested into the photorespiration (PCO) enzymes that can be saved by preventing high photorespiratory rates; (3) the relationship between cytochrome f and maximal electron transport of the LET (cyt_{Jmax}, [(mmol e^-)/(mmol cyt s)]), which is only known for C_3 species; (4) the scaling factor for the maximal electron transport rate of the CET relative to that of the LET ($Jmax_{CL}$); and (5) the empirical curvature factor (Θ), for which different values are frequently used in the literature. 200 parameter sets were sampled from an uncorrelated multivariate normal distribution (see Table S8 for details). The mean of the normal distribution is set to the standard value (see Methods, Table S8), except for the proportion of nitrogen invested into the PCO enzymes that can be saved by preventing high photorespiratory rates, this value is set to one. By that, we focus on the parameter sets that hypothesize the same photosynthetic nitrogen level for C_3 and C_4 plants. The variance for each parameter is chosen such that on the one hand values that are discussed in the literature (Ogren & Evans, 1993) and on the other hand uncertainties are potentially covered. We calculated the mean squared residuals (expressed as fractions of the experimental means) for the relevant Dwyer *et al.* (2007) data for plants grown at 25°C compared to data predicted from optimal resource allocation to either the growth or to the evolutionary scenario (Fig. S1). For each of the 200 random parameter sets, the predictions based on the evolutionary scenario led to lower error than those based on the growth environment.

Fig. S1 Sensitivity analysis for our finding that all predictions based on the evolutionary scenario outperform those based on the growth environment. We randomly perturbed uncertain model parameters to sample their effect on predictive performance of allocation data taken from Dwyer et al. (2007) (see Methods S5 and Table S8 for details). The histogram shows the difference between the prediction error assuming an optimal resource allocation under the evolutionary scenario and under the growth scenario, calculated for 200 randomly chosen sets of parameters. The error describes the mean squared residuals (expressed as fractions of the experimental means) for the data shown in Fig. 4a. The solid line represents the difference for the standard parametrization as shown in Fig. 4a and the dashed line represents the zero intercept.

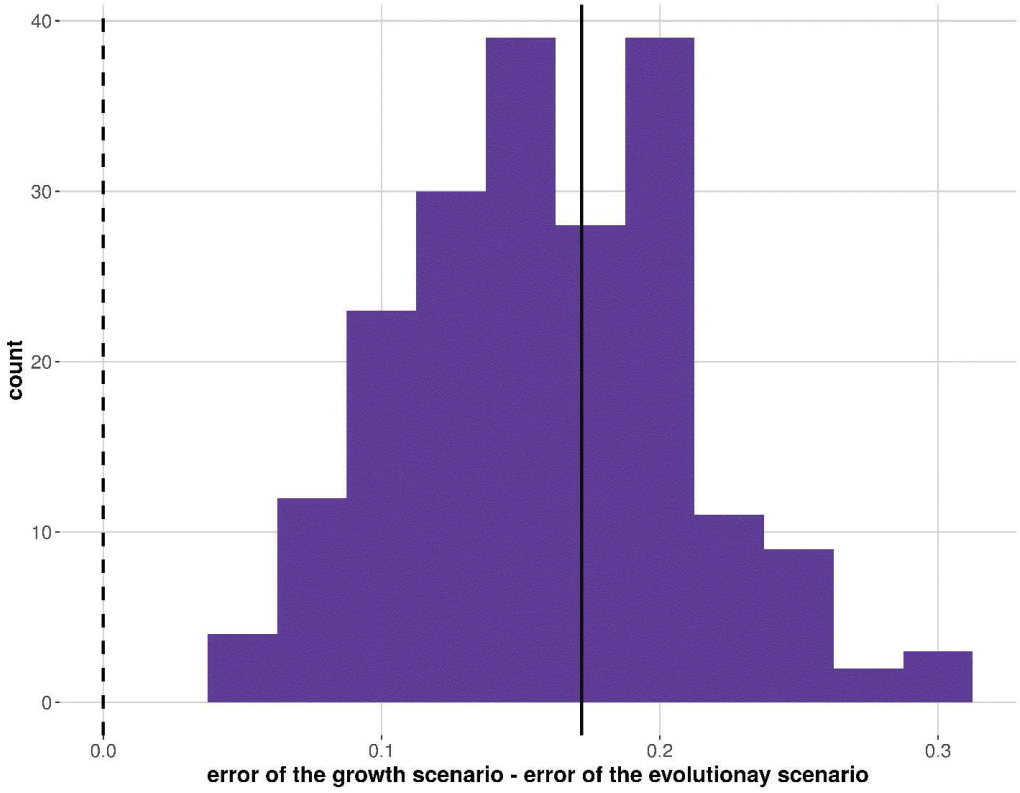


Table S1 Model parameterization for different species of Flaveria. β, fraction of Rubisco expressed in the mesophyll cell [fraction]; kccat, the maximal turnover rate of Rubisco [1/s]; Vpmax, the empirical maximal C4 cycle activity determined by the PEPC activity [µmol m-2 s-1]; Kp, Michaelis constant of PEPC [µbar]; gs, bundle sheath conductance [µmol m-2 s-1]; ξ, fraction of mesophyll cell derived photorespiration in the bundle sheath cells. All parameters are taken

from Heckmann et al. (2013), except for the maximal Rubisco turnover rate which is taken from Kubien et al. (2008). As explained in the description in the main text, the C4 cycle activity of F. bidentis is set to 130 µmol m-2 s-1.

Species	Photosynthetic types	β	k_{ccat}	V_{pmax}	K_p	g_s	ξ
F. pringlei	C_3	0.95	3.11	0	200	0.015	0
F. robusta	C_3	0.95	3.11	0	200	0.015	0
F. floridana	C_3-C_4 type II	0.52	3.19	40.2	200	0.015	0.79
F. ramosissima	C_3-C_4 type II	0.65	2.77	40.2	200	0.015	0.58
F. palmeri	C_4-like	0.068	3.54	75.9	80	0.001	0.97
F. bidentis	C_4	0.008	4.16	130	80	0.001	0.96

Table S2 Lower and upper bounds for the model parameters subject to numerical optimization.

Parameter	Lower bound	Upper bound
Proportion of LET (p)	1.0×10^{-10}	$1 - 1.0 \times 10^{-10}$
Fraction of nitrogen invested into the C_4 cycle (n_{C4})	0	The measured maximal PEPC activity is used to calculate the maximal investment into the C_4 cycle (n_{C4}^{max}, Eqn S1)
Fraction of nitrogen invested into the thylakoids (n_{Jmax})	The nitrogen investment into the CET is independent of p, therefore the bound for n_{Jmax} (n_{Jmax}^{min}) can be set as fixed to the nitrogen requirements of the CET (Eqn S2).	1
Fraction of nitrogen invested into Rubisco (n_{Etot})	0	1

Table S3 The parameters of the temperature-dependent model. The variable names for the temperature-dependent parameters are the same as in the main text and in the Supporting Information text above. E, the activation energy [J mol-1]; H, the deactivation energy [J mol-1]; S, entropy factor [J mol-1]; R, the universal gas constant [J mol-1]; inac, does inactivation occur.

Temperature-dependent variable	Parameter	Value	Source
V_{cmax}	*E*	7.543351e+04	fit to Ku *et al.* (1991)
V_{cmax}	*Inac*	TRUE	
V_{cmax}	*H*	1.213043e+05	fit to Ku *et al.* (1991)
K_c	*E*	4.175349e+04	fit to Ku *et al.* (1991)
K_c	*Inac*	FALSE	
K_o	*E*	5.314215e+04	fit to Ku *et al.* (1991)
K_o	*Inac*	FALSE	
$\gamma*$	*E*	2.166348e+04	fit to Ku *et al.* (1991)
$\gamma*$	*Inac*	FALSE	
V_{pmax}	*E*	7.0373e4	Massad *et al.* (2007)
V_{pmax}	*Inac*	FALSE	
K_p	*E*	5.455e4	Chen *et al.* (1994)
K_p	*Inac*	FALSE	
g_s	*E*	1.898e4	fit to Tamimi *et al.* (1994)
g_s	*Inac*	FALSE	
J_{max}	*E*	77900	Massad *et al.* (2007)
J_{max}	*Inac*	TRUE	
J_{max}	*S*	62	Massad *et al.* (2007)
J_{max}	*H*	191929	Massad *et al.* (2007)

Table S4 Required nitrogen re-allocation (δn, [fraction]) for F. bidentis (C4) grown at different temperatures (based on the experiments of Dwyer et al. (2007)). The plants were grown at 25°C or 35°C.

	Growth temperature	
	25°C	35°C
F. bidentis (C_4)	0.294	0.342

Table S5 Required nitrogen re-allocation (δn, [fraction]) for different leaf nitrogen levels for various Flaveria species (based on the experiments of Vogan and Sage (2011)).

	Leaf nitrogen level			
	50 mmol m-2	130 mmol m-2	170 mmol m-2	250 mmol m-2
F. pringlei (C_3)	0.062	0.140	0.193	0.323
F. floridana (C_3-C_4)	0.105	0.187	0.253	0.39
F. palmeri (C_4-like)	0.107	0.271	0.331	0.414
F. bidentis (C_4)	0.116	0.281	0.337	0.409

Table S6 Required nitrogen re-allocation (δn, [fraction]) for various Flaveria species grown at current or low CO2 level (based on the experiments of Vogan and Sage (2012)).

	Growth CO_2 level	
	Current CO_2 level	Low CO_2 level
F. robusta (C_3)	0.12	0.017
F. ramosissima (C_3-C_4)	0.167	0.101
F. bidentis (C_4)	0.318	0.254

Table S7 The modeled and measured data of chlorophyll [µmol m-2] and PSII [µmol m-2] of F. bidentis (C4). The plants were grown at 25°C or 35°C (data from Dwyer et al. (2007)).

	Growth at 25°C		Growth at 35°C	
	modeled	measured	modeled	measured
Photosystem II [µmol m^{-2}]	1.51	1.24	1.46	1.28
Chlorophyll [µmol m^{-2}]	602	499	585	533

Table S8 Distribution parameters used to generate the random parameter sets for the sensitivity analysis shown in Fig. S1. For each considered variable, mean and variance of the sampled normal distribution are shown.

	mean	variance
The mesophilic CO_2 concentration (C_m, which is calculated by scaling the considered standard C_m)	1	0.3
The proportion of nitrogen invested into the PCO enzymes that can be saved by preventing high photorespiratory rates	1	0.167
The relationship between cytochrome f and maximal electron transport of the LET (cyt_{Jmax})	172	24
The scaling factor for the maximal electron transport rate of the CET relative to that of the LET ($Jmax_{CL}$)	3	0.67
The empirical curvature factor (Θ)	0.7	0.083

Fig. S2 Modeled results based on the evolutionary scenario (solid lines) describe the measured data (dots ± SE) better than the model assuming optimal adaptation to the growth conditions (dashed lines) for F. bidentis (C4). Related to Fig. 2A, but the A-Ci curve was measured at 40°C (data from Vogan and Sage (2012)). Missing error bars result from unknown empirical errors. See Table 1 for error summaries.

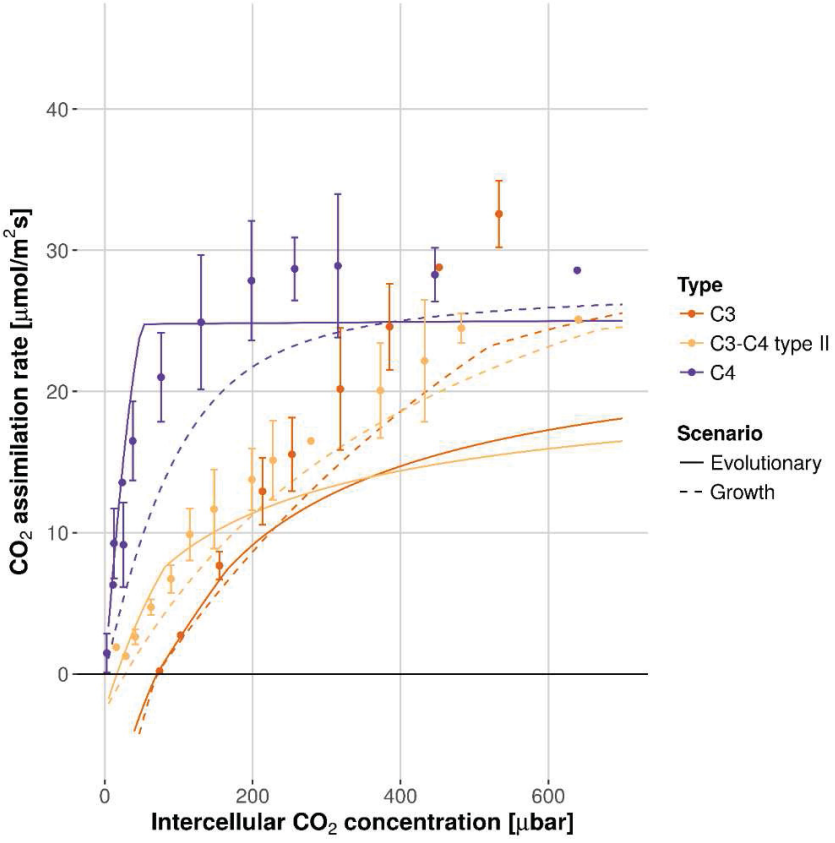


Fig. S3 Modeled results based on the evolutionary scenario (solid lines) describe the measured data (dots ± SE) better than the model assuming optimal adaptation to the growth conditions (dashed lines) for F. bidentis (C4). Related to Fig. 2A, but the A-Ci curve is measured for plants grown at the low CO2 level of 180 µbar (data from Vogan and Sage (2012)). Missing error bars result from unknown empirical errors. See Table 1 for error summaries.

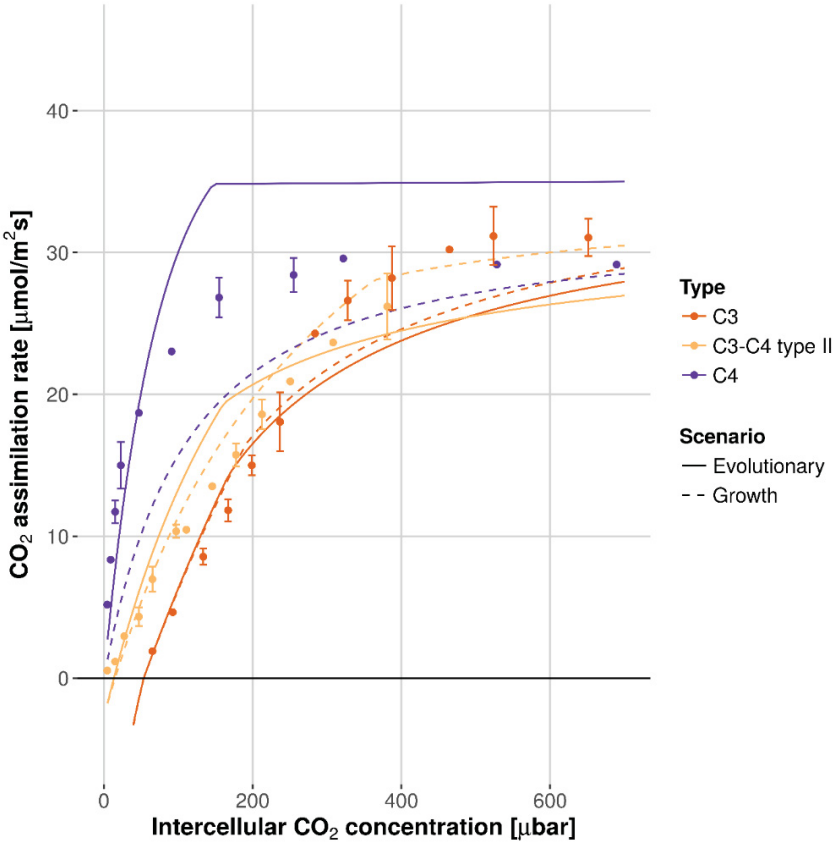


Fig. S4 Modeled results based on the evolutionary scenario (solid lines) describe the measured data (dots ± SE) better than the model assuming optimal adaptation to the growth conditions (dashed lines) for F. bidentis (C4). Related to Fig. 2A, but the A-Ci curve is measured at 40°C and for plants grown at the low CO2 level of 180 µbar (data from Vogan and Sage (2012)). Missing error bars result from unknown empirical errors. See Table 1 for error summaries.

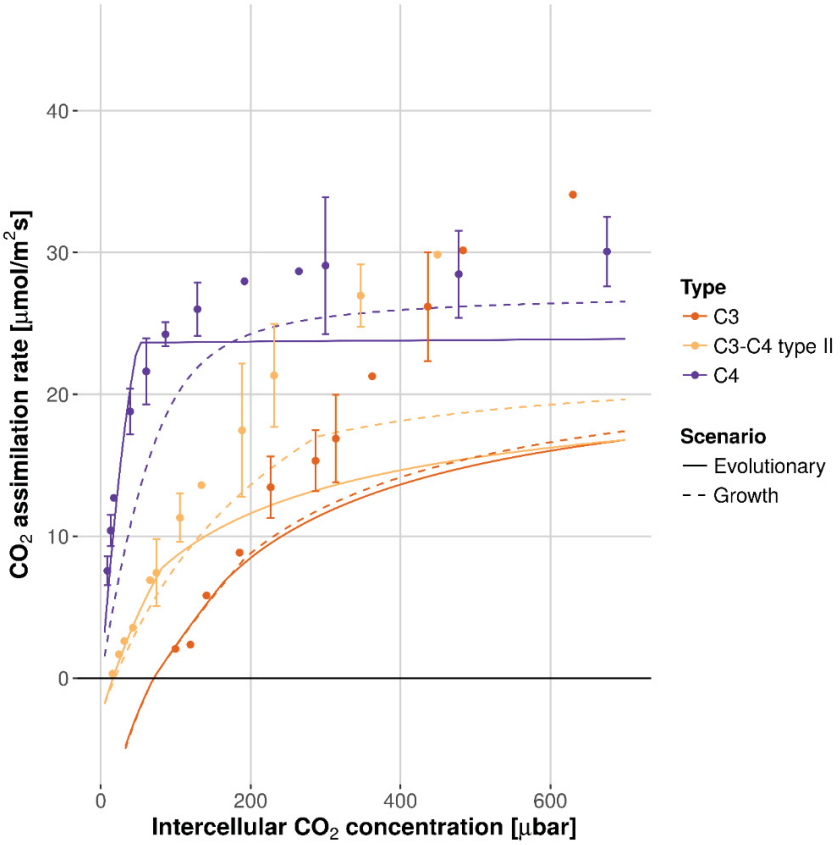


Fig. S5 Modeled results for temperature responses in the evolutionary scenario (solid lines) and optimal adaptation to the growth conditions (dashed lines) for F. robusta (C3), F. ramosissima (C3-C4), and F. bidentis (C4). Related to Fig. 2B, the A-Temperature curve is measured for plants grown at the low CO2 level of 180 µbar (data from Vogan and Sage (2012); dots ± SE). Missing error bars result from unknown empirical errors. See Table 1 for error summaries.

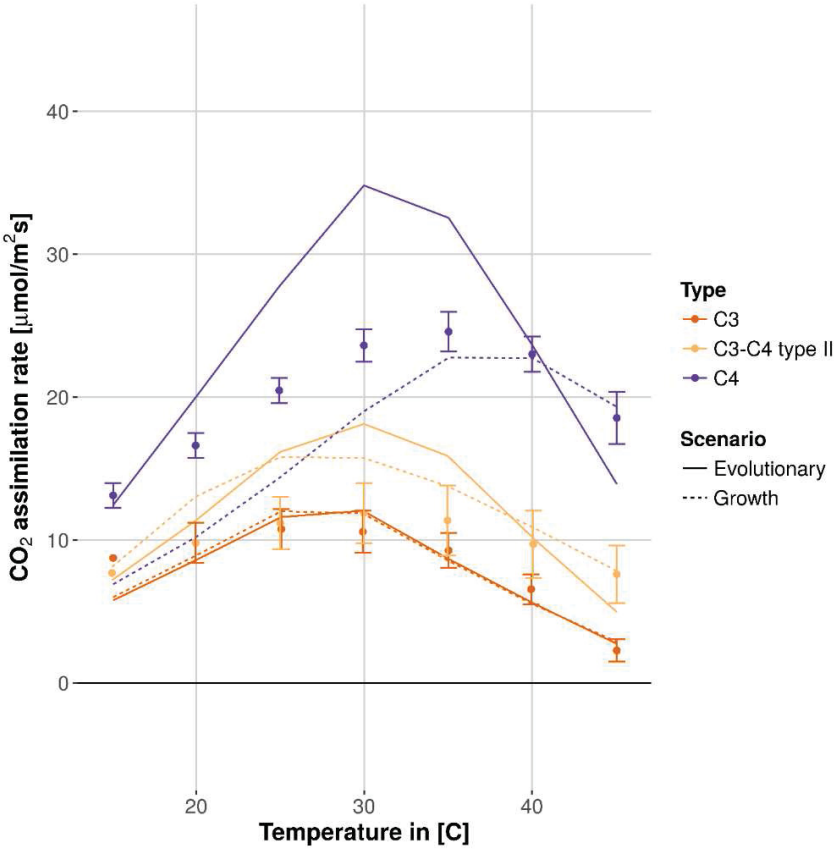


Fig. S6 Discrepancy between measured and modeled A-Ci curves of F. robusta (C3) across diverse environments assuming no phosphate-limitation. The deviation between model predictions and measurements ('error') is defined as the mean squared residuals of all measured curves (data from Vogan and Sage (2012)). The black dot indicates the environment that best explains the experimental data. To make this analyses comparable with the C4 analysis, the nitrogen allocation and the CO2 assimilation rate are included in the error calculation. Here, the nitrogen allocation is considered by including an empirically determined ratio of maximal electron transport rate per Rubisco activity of 2±0.6 (Leuning, 2002). The grey areas indicate that the modeled nitrogen allocation cannot satisfy this ratio and that these points should not be used for inference.

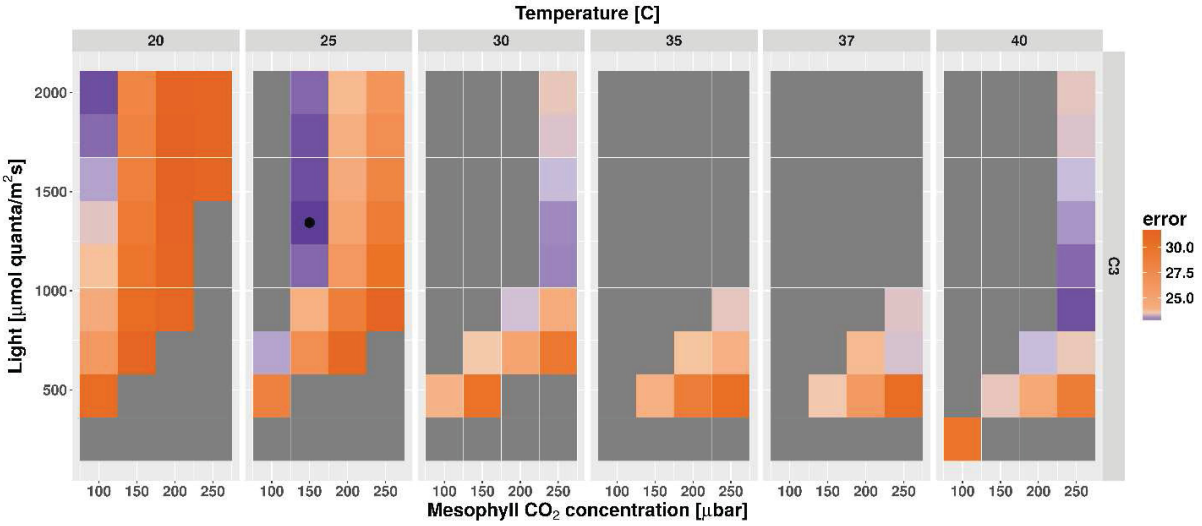
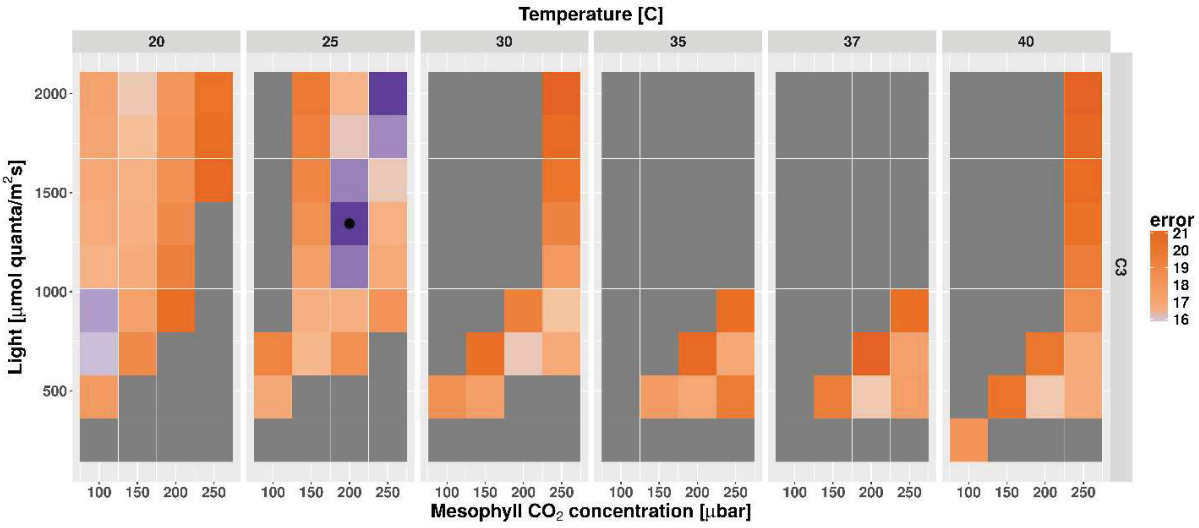


Fig. S7 Discrepancy between measured and modeled A-Ci curves of F. robusta (C3) across diverse environments assuming phosphate-limitation for intercellular CO2 levels above 400 µbar at 30°C and 500 µbar at 40°C for plants grown at current CO2 level. The deviation between model predictions and measurements ('error') is defined as the mean squared residuals of all measured curves (Vogan & Sage, 2012). The black dot indicates the environment that best explains the experimental data. To make this analyses comparable with the C4 analysis, the nitrogen allocation and the CO2 assimilation rate are included in the error calculation. Here, the nitrogen allocation is considered by including an empirically determined ratio of maximal electron transport rate per Rubisco activity of 2±0.6 (Leuning, 2002). The grey areas indicate that the modeled nitrogen allocation cannot satisfy this ratio and that these points should not be used for inference.



References

Chen DX, Coughenour MB, Knapp AK, Owensby CE. 1994. Mathematical Simulation of C_4 Grass Photosynthesis in Ambient and Elevated CO2. *Ecological Modelling* **73**(1-2): 63-80.

Dwyer SA, Ghannoum O, Nicotra A, Von Caemmerer S. 2007. High temperature acclimation of C_4 photosynthesis is linked to changes in photosynthetic biochemistry. *Plant Cell and Environment* **30**(1): 53-66.

Heckmann D, Schulze S, Denton A, Gowik U, Westhoff P, Weber A PM, Lercher Martin J. 2013. Predicting C_4 Photosynthesis Evolution: Modular, Individually Adaptive Steps on a Mount Fuji Fitness Landscape. *Cell* **153**(7): 1579-1588.

Ku MSB, Wu J, Dai Z, Scott RA, Chu C, Edwards GE. 1991. Photosynthetic and photorespiratory characteristics of *Flaveria* species. *Plant Physiology* **96**: 518.

Kubien DS, Whitney SM, Moore PV, Jesson LK. 2008. The biochemistry of Rubisco in *Flaveria*. *Journal of Experimental Botany* **59**: 1767.

Leuning R. 2002. Temperature dependence of two parameters in a photosynthesis model. *Plant Cell and Environment* **25**(9): 1205-1210.

Massad RS, Tuzet A, Bethenod O. 2007. The effect of temperature on C_4-type leaf photosynthesis parameters. *Plant Cell and Environment* **30**(9): 1191-1204.

Ogren E, Evans JR. 1993. Photosynthetic Light-Response Curves. 1. The Influence of CO2 Partial-Pressure and Leaf Inversion. *Planta* **189**(2): 182-190.

Tamimi A, Rinker EB, Sandall OC. 1994. Diffusion-Coefficients for Hydrogen-Sulfide, Carbon-Dioxide, and Nitrous-Oxide in Water over the Temperature-Range 293-368-K. *Journal of Chemical and Engineering Data* **39**(2): 330-332.

Vogan PJ, Sage RF. 2011. Water-use efficiency and nitrogen-use efficiency of C_3-C_4 intermediate species of *Flaveria* Juss. (Asteraceae). *Plant, Cell & Environment* **34**: 1415-1430.

Vogan PJ, Sage RF. 2012. Effects of low atmospheric CO_2 and elevated temperature during growth on the gas exchange responses of C_3, C_3-C_4 intermediate, and C_4 species from three evolutionary lineages of C_4 photosynthesis. *Oecologia* **169**(2): 341-352.

von Caemmerer S. 2000. *Biochemical models of leaf photosynthesis*. Collingwood, Australia: Csiro Publishing.

Yin XY, Struik PC. 2018. The energy budget in C_4 photosynthesis: insights from a cell-type-specific electron transport model. *New Phytologist* **218**(3): 986-998.

3.2 MANUSCRIPT 2

Manuscript 2 is pending for publication until *Manuscript 1* is accepted.

3.2.1 *Contributions to Manuscript 2*

I took the lead in designing the research and in writing the paper. I also developed and implemented key elements of the analysis. Together with David Heckmann, I designed and implemented the calculation of the evolutionary trajectories. In addition, I added the randomization of the step size to the calculation of the trajectories. I conducted the simulations. I analyzed the data and interpreted the results, in collaboration with David Heckmann and Martin J. Lercher.

Title

Modeling photosynthetic nitrogen-use efficiency in C_3 and C_4 plants: low nitrogen availability promotes C_4 evolution

Authors

Esther M. Sundermann, Martin J. Lercher, David Heckmann

Key words

C_3 photosynthesis, C_4 photosynthesis, photosynthetic nitrogen-use efficiency, C_4 evolution, C_4 ecology, leaf nitrogen level, *Flaveria*, environment

Abstract

Nitrogen is a fundamental constituent of organic molecules and its restricted availability often limits plant growth. Accordingly, a major determinant of plant fitness is the photosynthetic nitrogen-use efficiency, *i.e.*, the CO_2 assimilation rate per leaf nitrogen content. Photosynthetic nitrogen-use efficiency differs for C_3 and C_4 plants, raising the question how nitrogen availability influences the plants' ecology and the evolution of C_4 from C_3 photosynthesis. To address this question, we study the effect of the leaf nitrogen level on plant physiology and resource allocation in an environment that is relevant in the context of C_4 evolution, using a comprehensive mathematical model of C_3 and C_4 photosynthesis in the model genus *Flaveria*. We quantify the effect of different protein expression patterns that maximize photosynthetic nitrogen-use efficiency on the physiology of C_3 and C_4 plants as a function of leaf nitrogen level. Under low nitrogen availability, the C_4 advantage in photosynthetic nitrogen-use efficiency over C_3 photosynthesis is more pronounced compared to higher leaf nitrogen levels. Moreover, under low nitrogen availability, evolution from C_3 to C_4 photosynthesis requires less pronounced regulatory changes. This points to the possibility that nitrogen limitation boosts C_4 evolution. This insight prompts several new research questions; one interesting example concerns the chances of a C_3 plant with a mutualistic relationship with nitrogen fixing microorganisms to evolve a C_4 pathway in comparison to a C_3 plant that does not show this relationship.

Introduction

The majority of organic carbon is fixed by the enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) as part of photosynthesis. Among the different modes of photosynthesis, C_3 photosynthesis is the

original and most widespread one in vascular plants and among crop plants (Petterson, 1995; Prescott-Allen *et al.*, 1990). Another mode of photosynthesis is the C_4 pathway. Although the C_4 pathway only occurs in about 3% of plant species, it accounts for 23% of terrestrial gross primary productivity (Sage *et al.*, 2012; Still *et al.*, 2003). This apparent contradiction is explained by the ability of C_4 plants to prevent high photorespiratory rates that are caused by an affinity of Rubisco for O_2. Photorespiration results in a costly process: it requires enzymes and causes a significant loss of carbon and energy. While in C_3 plants Rubisco and the Calvin-Benson cycle are located in the mesophyll cells, C_4 plants express these in the bundle sheath cells, which allows C_4 plants to spatially separate Rubisco from the initial carbon fixation. The initial fixation is catalyzed by phosphoenolpyruvate (PEP) carboxylase (PEPC) and results in C_4 acids that are decarboxylated in the bundle sheath cells, which facilitates a local high-CO_2 environment around Rubisco and, thus, suppresses photorespiration. The regeneration of PEP by pyruvate, phosphate dikinase (PPDK) completes the C_4 cycle. Compared to the C_3 pathway, C_4 metabolism is more water- and nitrogen-use efficient (Sage *et al.*, 2012; Vogan *et al.*, 2012; Vogan *et al.*, 2011). The photosynthetic nitrogen-use efficiency, *i.e.*, CO_2 assimilation rate per leaf nitrogen level (PNUE), is a major determinant of plant fitness. The analysis of closely related C_3, C_3-C_4 intermediate, C_4-like, and C_4 *Flaveria* species reveals that the photosynthetic nitrogen-use efficiency does not increase gradually; C_4 and C_4-like species show a higher PNUE compared to C_3 and C_3-C_4 intermediate species under current atmospheric conditions (Vogan *et al.*, 2011; Vogan *et al.*, 2012). In contrast, data show that C_3, C_3-C_4 intermediate, and C_4 species differ significantly when grown and measured under low CO_2 concentrations (Vogan *et al.*, 2012). As nitrogen is a common limiting factor in natural and semi-natural terrestrial habitats (Erisman *et al.*, 2013; Vance, 2001), the differences in the PNUE of C_3 and C_4 photosynthesis raises the question to what extent nitrogen availability plays a role in C_4 evolution.

Mathematical models allow us to improve the understanding of photosynthesis in a wide range of research areas (Sundermann *et al.*, 2018; Wang *et al.*, 2014; Mallmann *et al.*, 2014; Heckmann *et al.*, 2013; Yin *et al.*, 2012). Areas of interest include the allocation of nitrogen (Wang *et al.*, 2014; Maire *et al.*, 2012; Zhu *et al.*, 2007) and energy of C_3, C_3-C_4 intermediate, and C_4 plants (Bellasio, 2016; Ubierna *et al.*, 2013; Yin *et al.*, 2012). The exploration of energy allocation includes the detailed analysis of cyclic electron transport and the exact costs of the carbon metabolism. However, the simultaneous analysis of nitrogen and energy allocation on plant performance—CO_2 assimilation rate, and physiological and molecular parameters—is missing. Despite intensive research, there is limited mechanistic understanding of the interplay of various leaf nitrogen levels and plant performance. To address this research question quantitatively, we will analyze the recently developed mathematical model, referred to below as the nitrogen-dependent light- and enzyme-limited model. This model allows us to analyze the potential evolution of C_4 photosynthesis under a wide range environmental conditions (Sundermann *et al.*, 2018).

Here, to understand how the evolvability of C_4 photosynthesis depends on nitrogen availability, we analyze the required changes in nitrogen allocation between C_3 and C_4 plants and the corresponding trajectories through the nitrogen-dependent fitness landscape. These analyses further allow us to quantify the effects of PNUE determining factors on C_3 and C_4 photosynthesis by analyzing their nitrogen allocation patterns and estimating physiological and molecular parameters.

Results

The nitrogen-dependent light- and enzyme-limited model

The nitrogen-dependent light- and enzyme-limited model described in Sundermann *et al.* (2018) allows us to investigate how quantitative nitrogen availability affects C_4 evolution and ecology. Figure 1 presents an overview of this mathematical model, which incorporates the following environmental factors: (1) leaf nitrogen levels, (2) light intensity, (3) temperature, and (4) CO_2 and (5) O_2 partial pressures. Moreover, it represents C_3, C_4, and all intermediate photosynthetic types, depending on the given parameterization (see Supplementary Information Table S1 for parameterization). The nitrogen-dependent light- and enzyme-limited model assumes that resources are allocated in a way that maximizes a fitness proxy, the CO_2 assimilation rate (A, [µmol m^{-2} s^{-1}]), in a given environment (see Figure 1 and Sundermann *et al.* (2018) for details). The optimization procedure supplies the optimal nitrogen and energy allocation as well as the corresponding light- and enzyme-limited CO_2 assimilation rate. Based on the optimal resource allocation, the model provides detailed information about physiological and molecular parameters including enzyme activities, the maximal electron transport rate, and the proportion of linear (LET) and cyclic (CET) electron transport (Figure 1). We call a plant "optimal" when its resource allocation results in the maximal attainable CO_2 assimilation rate under a given environmental condition, given its respective photosynthetic mode. The model and its assumptions about optimal resource allocation have been successfully validated using physiological data of C_3, C_3-C_4 intermediate, and C_4 *Flaveria* species (see Sundermann *et al.* (2018) for details).

C_3-C_4 habitat comparisons (McKown *et al.*, 2005; Sage, 2004; Powell, 1978), geophysiological considerations (Christin *et al.*, 2011), and quantitative estimations based on nitrogen allocation (Sundermann *et al.*, 2018) provide a highly likely and accurate description of the environment of the last common ancestor of current C_3 and C_4 species of the model genus for C_4 evolution, *Flaveria* ("evolutionary environment", see Supplementary Information Method S1 for parametrization). As this environment is of special interest for the evolution of C_4 photosynthesis, in the following we analyze *Flaveria* species in this C_4 favoring environment, which shows high light intensities, 30°C, and a low CO_2/O_2 gas concentration ratio.

To further validate the model, we modeled data that focus on the energy allocation as well as on the balance between energy production and consumption. Namely we modeled the ATP allocation, fraction of ATP from

LET, the fraction of electron transported by CET, and the total maximal electron transport rate per maximal Rubisco catalytic rate (see Supplementary Information Tables S3–S10). The modeled results and the empirical data presented in the literature are in very good agreement for the considered species *F. pringlei* (C_3), *F. bidentis* (C_4), and *Zea mays* (C_4). The only exception is the ratio of the total maximal electron transport rate per maximal Rubisco catalytic rate (J_{max}-V_{cmax}-ratio) for C_3 and C_4 *Flaveria* species. Note that the total maximal electron transport rate is the sum of the maximal electron transport rate of LET and CET.

Quantifying the effect of PNUE-determining factors in C_3 and C_4 photosynthesis

What is the quantitative effect of major factors that determine the PNUE in C_3 and C_4 plants? And how do various leaf nitrogen levels affect the PNUE and its determinants? To explore these questions, we analyzed the optimal nitrogen allocation and estimated physiological and molecular parameters of *Flaveria pringlei* (C_3) and *F. bidentis* (C_4).

Figure 2 depicts the optimal nitrogen allocation of C_3 and C_4 plants grown with leaf nitrogen levels of 50 mmol m^{-2}, 150 mmol m^{-2}, and 250 mmol m^{-2}. The simulated leaf nitrogen levels of 50 mmol m^{-2}, 150 mmol m^{-2}, and 250 mmol m^{-2} cover the empirically observed range (Vogan *et al.*, 2011); in the following, these levels will be referred to as low, moderate, and high leaf nitrogen levels, respectively. The allocation is presented based on the amount of photosynthetic nitrogen (N_{ps}, [µmol m^{-2}]), which represents the nitrogen available for the major photosynthetic pools—Rubisco, C_4 enzymes, and thylakoids. In C_3 plants, Rubisco and the thylakoids require roughly the same amount of nitrogen (Figure 2). In C_4 plants, the majority of nitrogen is invested into the thylakoids. For C_3 plants, the fraction of nitrogen invested into thylakoids decreases for increasing leaf nitrogen levels while the investment into Rubisco increases (see Supplementary Information Table S4 for details). The same trends can be observed when comparing the nitrogen allocation for low and moderate leaf nitrogen levels in C_4 plants. Comparing the moderate and the high leaf nitrogen levels of C_4 plants reverses the allocation strategy, *i.e.*, the nitrogen investment into the thylakoids increases and the investment into the dark reactions decreases. This pattern is caused by the non-linear increase of photosynthetic nitrogen relative to the leaf nitrogen level, while the protein costs scale approximately linearly with available nitrogen.

Previous work identified the Rubisco activity, C_4 cycle activity, and the ATP and NADPH production in the thylakoids to be major determinants of the CO_2 assimilation rate and, thus, the PNUE (von Caemmerer, 2000; Heckmann, 2013). The corresponding physiological parameters of the considered model are the maximal Rubisco (V_{cmax}, [µmol m^{-2} s^{-1}]) and C_4 cycle (V_{pmax}, [µmol m^{-2} s^{-1}]) activity as well as the maximal electron transport (J_{max}, [µmol m^{-2} s^{-1}]) of the linear and cyclic electron transport, and the proportion ([fraction]) of LET. These parameters can be estimated (see Table 1) based on the optimal nitrogen allocation patterns presented in Figure 2. In C_3 plants, J_{max} of LET is always higher than J_{max} of CET (Table 1). In contrast, C_4 plants

show a higher J_{max} of the CET than of the LET. Accordingly, the proportion of LET is drastically reduced in C_4 compared to C_3 plants. The proportion of LET decreases with increasing leaf nitrogen levels in C_3 and C_4 *Flaveria* species. As the simulated environments are favorable for C_4 plants, C_4 plants show consistently higher CO_2 assimilation rates than C_3 relatives under the same conditions (Table 1). In contrast to the optimal nitrogen allocation pattern, which shows dependencies on the nitrogen availability (Figure 2), the optimal energy allocation shows only minor effects (see Supplementary Information Table S5).

Comparing the optimal nitrogen allocation pattern of C_3 and C_4 plants (Figure 2) shows that an optimally adapted C_4 plant invests nitrogen saved by reducing Rubisco not only into the C_4 cycle but also into the thylakoids. This result is consistent with empirical work of Makino *et al.* (2003) comparing rice and maize. It is not fully understood how this increased investment into the thylakoid components affects the RuBP regeneration capacity (Makino *et al.*, 2003) and the PEP recycling of C_4 plants. To address this question, we simulated a scenario that assumes an equal investment into the thylakoids for C_3 and C_4 plants as well as an optimal allocation of the remaining nitrogen between Rubisco and C_4 cycle ("non-optimal scenario"). Comparing the optimal and the non-optimal C_4 plant reveals the effect of additional nitrogen invested into the thylakoids on J_{max} and the resulting CO_2 assimilation rate (Figure 3). The modeled results show that J_{max} of LET and CET increase at least by 49%. For the moderate leaf nitrogen level, J_{max} of LET increases from 45 µmol m^{-2} s^{-1} to 152 µmol m^{-2} s^{-1}, and J_{max} of CET from 146 µmol m^{-2} s^{-1} to 256 µmol m^{-2} s^{-1}. The optimal nitrogen allocation results in a drastic increase in the CO_2 assimilation rate of at least 142% (Figure 3).

C_4 advantage in PNUE is more pronounced in a low-nitrogen environment

As the evolutionary environment is favorable for C_4 plants, the CO_2 assimilation rate of C_4 plants is consistently higher than the one of C_3 plants at the same leaf nitrogen level (Table 1). However, for decreasing leaf nitrogen levels the advantage of C_4 photosynthesis in PNUE increases compared to the C_3 photosynthetic pathway (Table 2). This result indicates that the C_4 advantage in PNUE is more pronounced under limited nitrogen availability. A sensitivity analysis shows that this trend is robust against variation in the considered environments (see Supplementary Information Figure S1).

C_4 evolution is more likely under limited nitrogen availability

The modeled results reveal that C_4 photosynthesis, relative to C_3 photosynthesis, is more advantageous in fixing CO_2 under low compared to moderate or high nitrogen availability (Table 2). In order to improve the understanding on how this advantage may influence C_4 evolution, we analyzed the effect of leaf nitrogen levels in an evolutionary context. We compared the required changes in nitrogen to transform an optimal C_3 into an optimal C_4 plant for various leaf nitrogen levels ("required nitrogen re-allocation", δ_N, see Methods for details). The required nitrogen re-allocation increases for increasing leaf nitrogen levels (Table 3). Compared

to the low leaf nitrogen level, at least 3.5 times more nitrogen needs to be re-allocated to transform a C_3 into a C_4 plant under medium or high leaf nitrogen levels. As accordingly fewer evolutionary changes are required to generate a C_4 species in a low-nitrogen environment, this indicates that nitrogen scarcity may promote C_4 photosynthesis evolution.

To test the hypothesis that nitrogen is a promoting factor for C_4 evolution, we simulated 1,000 evolutionary trajectories on the fitness landscape that maps the evolution from an optimal C_3 to an optimal C_4 plant (Figure 4). Nine evolutionary parameters span the fitness landscape (see Supplementary Information Table S2 a full list), including nitrogen allocation into Rubisco, the C_4 cycle, and the thylakoids, (n_{Etot}, n_{C4}, n_{Jmax}, [fractions]). As we consider only photosynthetic nitrogen allocation, the three fractions sum up to 100%. This means that n_{Etot} can be calculated as a function of n_{C4} and n_{Jmax}, and, thus, only eight of the nine evolutionary parameters are independent from each other. The range of each independent parameter is divided into 100 equidistant steps. The parameters that are optimized depend on the available amount of leaf nitrogen level. To ensure that for each environment these parameters show an equal relative difference, the number of steps in the moderate and high leaf nitrogen levels are normalized to the C_3 values, in case of n_{C4} relative to the C_4 value, at low leaf nitrogen level (see Supplementary Information Table S11 for exact numbers). As we assume that beneficial mutations are fixed in the population before the next mutation occurs (Heckmann *et al.*, 2013; Gillespie, 1983), evolutionary trajectories simulate the change of one parameter at a time. In order to cover the stochasticity of evolutionary changes, the simulations include two stochastic elements. First, the extent of a parameter change is chosen randomly based on a uniform distribution; the parameter range is 1 step to 10% of the maximal number of steps. Second, the probability of a parameter change to fix in the population is determined using the population genetic model first derived by Kimura (1957), considering a population size of 10,000 (see Heckmann *et al.* (2013) for details). As shown in Figure 4, for decreasing leaf nitrogen levels, a significant decrease in the number of evolutionary steps is required to become a near optimal C_4 phenotype, where the plants show at least 90% of the C_4-specific attributes of an optimal C_4 plant (δ_{C4}, see Methods for details; $P = 2.5 \times 10^{-9}$, median test).

Discussion

Our *in silico* analysis provides novel insights into the effects of nitrogen availability on the photosynthetic performance of C_3 and C_4 plants. The mathematical modeling framework allows us to focus on the effect of leaf nitrogen levels on the resource allocation and physiological parameters in closely related C_3 and C_4 *Flaveria* species. We estimate the maximal electron transport rate for LET and CET based on the current understanding of their mechanistic interaction with nitrogen availability, parameters not yet determined in

such detail. Moreover, we refine the knowledge about the difference in PNUE of C_3 and C_4 plants by quantifying the effect of the major factors that determine the PNUE as a function of leaf nitrogen level (Table 2). Reduced required nitrogen re-allocation (Table 3), combined with the superiority of C_4 photosynthesis (Table 2) under low nitrogen availability, points to the possibility that nitrogen scarcity promotes C_4 evolution. This hypothesis is supported by the smaller number of steps required under low nitrogen to achieve a near optimal C_4 state (Figure 4).

The mathematical model of Sundermann *et al.* (2018) allows us to study the optimal resource allocation and to estimate physiological and molecular parameters. Although many parameters are currently infeasible or impractical to measure, other parameters have been estimated or measured before. The modeled results and data of *F. pringlei* (C_3), *F. bidentis* (C_4), and *Z. mays* (C_4) are in very good agreement, except for the J_{max}-V_{cmax}-ratio (see Supplementary Information Table S3–S10). The results, except for the J_{max}-V_{cmax}-ratio, agree within a factor of 0.69 to 1.69 with values from literature, focusing on the *Flaveria* data only, the agreement is within 0.89 and 1.09. For the J_{max}-V_{cmax}-ratio, various values within a range of about 2 to 6 are reported for C_4 species (see Supplementary Information Table S3 for details). These ratios are determined with the help of mathematical models that cover yield-based approaches as well as NADPH- and ATP-limited models assuming an electron-per-ATP ratio of 1. In contrast, we assume a simultaneous limitation by ATP and NADPH as well as an electron-per-ATP ratio of 4/3. In *Flaveria*, this ratio is supported by Siebke *et al.* (1997). Due to the differences in the determination of J_{max} and V_{cmax}, the J_{max}-V_{cmax}-ratio estimated by the model of Sundermann *et al.* (2018) only agrees within the factor of 1.04 to about 6 with the previously reported ratios. For C_3 plants, the model estimations of J_{max}-V_{cmax}-ratio are 1.04–1.2 (see Supplementary Information Table S8 for details), which is lower than the empirically observed ratio of 2±0.6 (Leuning, 2002). Currently observed C_3 plants, which are used to determine this ratio, most likely show some ability to adapt to the current environmental conditions and, thus, do not show optimality in an ancestral environment. As we analyze optimality in an ancestral environment, this phenotypic plasticity might result in the observed difference.

The presented work contributes to our understanding of the ecology of C_3 and C_4 plants, in particular their global occurrence. As most natural and semi-natural habitats are nitrogen limited, our work highlights the local effect of nitrogen availability on the C_3 and C_4 distribution. The C_4 superiority in PNUE might be of particular importance in C_4 dominated habitats in warm regions with frequent disturbances, *e.g.*, C_4 grasslands that face mammalian feeding. The high photosynthetic rates and high nitrogen-use efficiency can facilitate the accumulation of biomass and, thus, allows C_4 plants to dominate these habitats (Edwards *et al.*, 2010).

Brown (1978) suggested that mutualistic symbiotic relationships with nitrogen fixating microorganism might reduce the evolutionary pressure on a C_3 plant to evolve the more nitrogen-use efficient C_4 pathway. Based

on our work, the evolution of a C_3 plant that shows mutualistic symbiosis toward C_4 metabolism is less likely compared to a C_3 plant without mutualistic symbiosis. This assessment results from the higher nitrogen re-allocation and the longer evolutionary paths toward the C_4 phenotype when nitrogen is less scarce. But due to other environmental factors that prompt C_4 evolution, it is still expected that these C_3 plants evolve into C_4 plants under the considered conditions. Our work adds quantitative information on the advantage in the PNUE of C_4 relative to C_3 plants as a function of leaf nitrogen level. The decreasing C_4 advantage for increasing leaf nitrogen levels might affect the likelihood of an establishment of a mutualistic symbiosis in C_4 relative to C_3 plants. As the interaction between nitrogen fixating microorganisms and plants is complex (Jacoby *et al.*, 2017), further research is required to evaluate the likelihood of an establishment of a mutualistic symbiosis in plants and the chance of C_3 plants with and without mutualistic symbiosis to evolve toward a full C_4 plant. Taken together, our work paves the way to understand the causal interaction between leaf nitrogen levels and C_4 evolution.

The presented results are based on the assumption that C_3 and C_4 species are equally efficient in taking up nitrogen from the environment. For *Cheaopodium album* (C_3) compared to *Amaranthus retroflexus* (C_4), an increased leaf nitrogen level ([mmol m^{-2}]) per applied nitrogen is observed (Sage *et al.*, 1987). However, the dependence of leaf nitrogen level on applied nitrogen and its effect on PNUE in closely related species with different photosynthetic types is unclear. To verify the assumption of an equally efficient nitrogen uptake of C_3 and C_4 plants, further research that focuses on closely related C_3 and C_4 species, *e.g.*, in the genus *Flaveria*, is required.

The presented analysis is performed using the genus *Flaveria.* This genus is an ideal system for studying the costs and implications of the C_4 cycle because it includes closely related C_3 and C_4 species. This allows us to focus on the effect of different photosynthetic types rather than on species-specific effects. The environment of the last common ancestor of the current C_3 and C_4 *Flaveria* species has been inferred (Sundermann *et al.*, 2018; Christin *et al.*, 2011) and, therefore, facilitates the consideration of the effect of leaf nitrogen levels during this time crucial for C_4 evolution. As *Flaveria* is a model genus for C_4 evolution, we propose that the presented hypotheses may be applicable to evolution of C_4 species in general. Because of the strongly limited phenotypic plasticity of C_4 *Flaveria* plants (Sundermann *et al.*, 2018, Sage *et al.*, 2005), the gained knowledge about the optimal resource allocation in C_4 species in past environment is also relevant under current conditions.

Conclusion

The analysis of a comprehensive model facilitates the exploration of the photosynthetic performance of C_3 and C_4 plants in an ancestral environment relevant for C_4 evolution. We not only provide a new line of

evidence that a balanced C_4 metabolism requires an increased nitrogen investment into the thylakoids compared to C_3 photosynthesis; we also add qualitative information on the optimal resource allocation and resulting estimations for physiological and molecular parameters. Our work allows us to derive ecological and evolutionary implications; the most important of these is that limited nitrogen availability increases the C_4 advantage in PNUE and promotes the evolution of C_4 photosynthesis. We pave the way to a full understanding of the effect of leaf nitrogen level on C_3 and C_4 ecology as well as on C_4 evolution. Our work suggests hypotheses that are relevant in multiple biological research fields including microbiology and phylogeny. An interesting example is the likelihood of C_3 plants to evolve a mutualistic relationship with nitrogen fixing microorganisms compared to C_4 plants under the same conditions.

Methods

The nitrogen-dependent light- and enzyme-limited model

The physiological and molecular parameters as well as the resource allocation are calculated by using the nitrogen-dependent light- and enzyme-limited model presented in Sundermann *et al.* (2018). This mechanistic model encompasses the effect of (1) leaf nitrogen levels, (2) light intensity, (3) temperature, and (4) CO_2 and (5) O_2 partial pressures on the photosynthetic nitrogen and energy allocation for C_3, C_3-C_4 intermediate, and C_4 plants. The resource allocation provides the basis for the determination of the ATP and NADPH production through linear and cyclic electron transport as well as a wide range of physiological and molecular parameter including the CO_2 assimilation rate. The CO_2 assimilation rate in a given environment, especially for a given leaf nitrogen level, is a proxy for plant fitness. Accordingly, we maximize the CO_2 assimilation rate in order to determine the optimal nitrogen and energy allocation (see Sundermann *et al.* (2018) for details).

The nitrogen and energy allocation is optimized under the previously inferred evolutionary scenario (Sundermann *et al.*, 2018) for *Flaveria pringlei* (C_3), *F. bidentis* (C_4), and *Zea mays* (C_4), respectively (see Supplementary Information Method S1 for parametrization of the evolutionary environment and Table S1 for plant parametrizations). The leaf nitrogen level is set to different values. Note that the light- and enzyme-limited CO_2 assimilation rates are equal at the point of optimality, as otherwise resources could be shifted from the non-limiting to the limiting section. The enzyme activities at 25°C are presented in Table 1 and Figure 3.

The photosynthetic nitrogen pool (N_{ps}, [µmol m^{-2}]) includes major pools of leaf photosynthetic nitrogen: the main enzyme of the Calvin-Benson cycle, Rubisco; the main enzymes of the C_4 cycle, PEPC and PPDK; and the thylakoids, which include the electron transport chains. If sufficient C_4 cycle activity is present, the amount of photorespiratory enzymes can be reduced, which results in an increase in nitrogen that is available for the

photosynthetic pools. N_{ps} can increase by up to 10% (Sundermann *et al.*, 2018, Zhu *et al.*, 2007). In order to model the continuous change of N_{ps} during evolution, N_{ps} increases in dependence on C_4 cycle activity in a sigmoidal manner (see Supplementary Information Method S2 for details). For C_4 cycle activities between about 4 [µmol $m^{-2}s^{-1}$] and about 13, N_{ps} increases from 0% to 10%. The thresholds are chosen based on the *in vivo* C_4 cycle activity of C_3-C_4 type I and II intermediates of the genus *Flaveria*, respectively (*in vitro* data from Heckmann *et al.*, 2013, see Supplementary Information Method S3 for the transformation of *in vitro* into *in vivo* data).

Evolutionary simulations

We calculated 1,000 evolutionary trajectories through a fitness landscape that starts with an optimal C_3 phenotype that can evolve toward an optimal C_4 phenotype (see Supplementary Information Tables S1 and S2 for parametrization). A fitness landscape is a theoretical concept that is spanned by evolutionary parameters (Heckmann, 2015) and links a fitness to each parameter state. Here, the state is defined by nine physiological parameters that are known to be relevant in the evolution of the considered trait. As the nitrogen allocation into Rubisco is calculated as the fraction not invested into the C_4 cycle enzymes and the thylakoids, eight of the nine parameter are independent (all evolutionary parameters are listed in Supplementary Information Table S2).

The fitness proxy used is the CO_2 assimilation rate per leaf nitrogen level calculated by the model presented in Sundermann *et al.* (2018). Each parameter range is separated into 100 equidistant steps. For moderate and high leaf nitrogen levels, the number of steps for the optimized parameters—nitrogen investment into Rubisco, the C_4 cycle, and the thylakoids as well as the proportion of LET—are adjusted, such that the relative changes based on the optimal C_3 values are equal for all leaf nitrogen levels. Note that in the case of nitrogen investment into the C_4 cycle, the adjustment is based on the optimal C_4 value, as the C_3 value is zero (see Supplementary Information Table S11 for exact number of steps). We assume that beneficial mutations fix in the population before the next mutation occurs (Heckmann *et al.*, 2013; Gillespie, 1983), which means that at each step of the simulation one evolutionary parameter is changed. The changed parameter and the extent of its change (in the following referred to as the extent of parameter change, that is given in number of steps) are chosen randomly. First, the extent of parameter change for each evolutionary parameter is determined. The extent of the parameter change that results in a fitness increase is randomly chosen based on a uniform distribution; the minimal extent of a parameter change is one step and the maximal change is 10% of the maximal number of steps (which is 10 steps for parameters that are independent of the leaf nitrogen level). Then, the parameter that mutates and fixes in the population is chosen based on the relative product of the mutational probability and the relative probabilities derived from population genetics model,

which is derived by Kimura (1957), considering a population size of 10,000 (see Heckmann *et al.* (2013) for details about the relative probabilities and for the population genetics model).

Required nitrogen re-allocation

Required nitrogen re-allocation (δ_N) is defined as the total fraction of nitrogen that needs to be re-allocated between photosynthetic pools to transform an optimal C_3 plant (n_{Etot}, n_{C4}, n_{Jmax}) into an optimal C_4 plant (m_{Etot}, m_{C4}, m_{Jmax}):

$$\delta_N = \sum_{i \in \{E_{tot}, C4, J_{max}\}} |n_i - m_i| \quad (1)$$

The degree of C_4-specific attributes (δ_{C4})

The degree of C_4-specific attributes of a given phenotype describes the fraction of the C_4-specific attributes relative to an optimal C_4 plant (δ_{C4}, see Eq. 2). We determine this quantity for phenotypes that achieved the highest fitness in a given evolutionary trajectory. In Eq. 2, *x* and *y* represent the C_4-specific attributes, namely (1) the fraction of Rubisco located in the mesophyll cell (equivalent to β in Sundermann *et al.* (2018)), and (2) the nitrogen investment into the C_4 cycle (equivalent to n_{C4} in Sundermann *et al.* (2018)). The respective variables x_{opt} and y_{opt} represent the number of steps through the evolutionary fitness landscape that result in the optimal C_4 phenotype. The achieved number of steps is represented by x_{ach} and y_{ach}.

$$\delta_{C4} = 1 - \text{mean}\left(\frac{x_{opt} - x_{ach}}{x_{opt}}, \frac{y_{opt} - y_{ach}}{y_{opt}}\right) \quad (2)$$

Zea mays data

The empirical data for *Zea mays* presented in Supplementary Information Table S9 are taken from Mu *et al.* (2016). To get the nitrogen investment into the C_4 cycle enzymes, the measurements of PEPC and PPDK are combined. For the thylakoids, the investment into the light-harvesting proteins and the proteins related to bioenergetics, including Cytochrome b/f, are pooled.

Statistical analysis

All statistical analyses were conducted in R (R Core Team, 2017). The median test is implemented in the coin package (Hothorn *et al.*, 2006).

Acknowledgements

We acknowledge financial support by the Deutsche Forschungsgemeinschaft (grants IRTG 1525 to DH, EXC 1028 to MJL and ES, CRC 680 and CRC 1310 to MJL).

Author Contribution

ES took the lead in designing the research and in writing the paper, aided by DH and MJL. ES and DH developed and implemented the analysis. ES conducted the simulations. In collaboration with DH and advised by MJL, ES analyzed the data and interpreted the results.

References

Bellasio, C. (2016). A generalized stoichiometric model of C_3, C_2, C_2+ C_4, and C_4 photosynthetic metabolism. *Journal of Experimental Botany*, *68*(2), 269-282.

Brown, R. H. (1978). A difference in N use efficiency in C_3 and C_4 plants and its implications in adaptation and evolution. *Crop Science*, *18*(1), 93-98.

Christin, P. A., Osborne, C. P., Sage, R. F., Arakaki, M., & Edwards, E. J. (2011). C_4 eudicots are not younger than C_4 monocots. *Journal of Experimental Botany*, *62*(9), 3171-3181.

Edwards, E. J., Osborne, C. P., Strömberg, C. A., Smith, S. A., & C_4 Grasses Consortium. (2010). The origins of C_4 grasslands: integrating evolutionary and ecosystem science. *Science*, *328*(5978), 587-591.

Erisman, J. W., Galloway, J. N., Seitzinger, S., Bleeker, A., Dise, N. B., Petrescu, A. M., Leach, A. M., … de Vries, W. (2013). Consequences of human modification of the global nitrogen cycle. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *368*(1621), 20130116. doi:10.1098/rstb.2013.0116

Gillespie, J.H. (1983). A simple stochastic gene substitution model. *Theor. Popul. Biol., 23,* 202–215.

Heckmann, D., Schulze, S., Denton, A., Gowik, U., Westhoff, P., Weber, A. P., & Lercher, M. J. (2013). Predicting C_4 photosynthesis evolution: modular, individually adaptive steps on a Mount Fuji fitness landscape. *Cell*, *153*(7), 1579-1588.

Heckmann, D. (2015). Modelling metabolic evolution on phenotypic fitness landscapes: a case study on C_4 photosynthesis. *Biochemical Society Transactions*, *43*(6), 1172-1176.

Hothorn, T., Hornik, K., Van De Wiel, M. A., & Zeileis, A. (2006). A Lego system for conditional inference. *The American Statistician*, *60*(3), 257-263.

Jacoby, R., Peukert, M., Succurro, A., Koprivova, A., & Kopriva, S. (2017). The role of soil microorganisms in plant mineral nutrition—current knowledge and future directions. *Frontiers in Plant Science*, 8, 1617.

Kimura, M. (1957). Some problems of stochastic processes in genetics. *Ann. Math. Stat. 28,* 882–901

Leuning, R. (2002). Temperature dependence of two parameters in a photosynthesis model. *Plant, Cell & Environment*, *25*(9), 1205-1210.

Maire, V., Martre, P., Kattge, J., Gastal, F., Esser, G., Fontaine, S., & Soussana, J. F. (2012). The coordination of leaf photosynthesis links C and N fluxes in C_3 plant species. *PloS one*, *7*(6), e38345.

Makino, A., Sakuma, H., Sudo, E., & Mae, T. (2003). Differences between maize and rice in N-use efficiency for photosynthesis and protein allocation. *Plant and Cell Physiology*, *44*(9), 952-956.

Mallmann, J., Heckmann, D., Bräutigam, A., Lercher, M. J., Weber, A. P., Westhoff, P., & Gowik, U. (2014). The role of photorespiration during the evolution of C_4 photosynthesis in the genus *Flaveria*. *Elife*, *3*, e02478.

McKown, A. D., Moncalvo, J. M., & Dengler, N. G. (2005). Phylogeny of *Flaveria* (Asteraceae) and inference of C_4 photosynthesis evolution. *American Journal of Botany*, *92*(11), 1911-1928.

Mu, X., Chen, Q., Chen, F., Yuan, L., & Mi, G. (2016). Within-leaf nitrogen allocation in adaptation to low nitrogen supply in maize during grain-filling stage. *Frontiers in Plant Science*, *7*, 699.

Patterson, D. T. (1995). Weeds in a changing climate. *Weed Science*, *43*(4), 685-700.

Powell, A. M. (1978). Systematics of *Flaveria* (Flaveriinae--Asteraceae). *Annals of the Missouri Botanical Garden*, 590-636.

Prescott-Allen, R. & Prescott-Allen, C. (1990). How Many Plants Feed the World?. *Conservation Biology*, 4. 365-374.

R Core Team (2017). R: A Language and Environment for Statistical Computing: R Foundation for Statistical Computing.

Sage, R. F., & Pearcy, R. W. (1987). The nitrogen use efficiency of C_3 and C_4 plants: I. Leaf nitrogen, growth, and biomass partitioning in *Chenopodium album* (L.) and *Amaranthus retroflexus* (L.). *Plant Physiology*, *84*(3), 954-958.

Sage, R. F. (2004). The evolution of C_4 photosynthesis. *New Phytologist*, *161*(2), 341-370.

Sage, R. F., & McKown, A. D. (2005). Is C_4 photosynthesis less phenotypically plastic than C_3 photosynthesis?. *Journal of Experimental Botany*, *57*(2), 303-317.

Sage, R. F., & Pearcy, R. W. (1987). The nitrogen use efficiency of C_3 and C_4 plants: I. Leaf nitrogen, growth, and biomass partitioning in Chenopodium album (L.) and Amaranthus retroflexus (L.). *Plant Physiology*, *84*(3), 954-958.

Sage, R. F., Sage, T. L., & Kocacinar, F. (2012). Photorespiration and the evolution of C_4 photosynthesis. *Annual Review of Plant Biology*, *63*, 19-47.

Siebke, K., von Caemmerer, S., Badger, M., & Furbank, R. T. (1997). Expressing an RbcS antisense gene in transgenic *Flaveria bidentis* leads to an increased quantum requirement for CO_2 fixed in photosystems I and II. *Plant Physiology*, *115*(3), 1163-1174.

Still, C. J., Berry, J. A., Collatz, G. J., & DeFries, R. S. (2003). Global distribution of C_3 and C_4 vegetation: carbon cycle implications. *Global Biogeochemical Cycles*, *17*(1), 1006.

Sundermann, E., Lercher, M., & Heckmann, D. (2018). Modeling Cellular Resource Allocation Reveals Low Phenotypic Plasticity of C_4 Plants and Infers Environments of C_4 Photosynthesis Evolution. *bioRxiv*, 371096.

Ubierna, N., Sun, W. E. I., Kramer, D. M., & Cousins, A. B. (2013). The efficiency of C_4 photosynthesis under low light conditions in *Zea mays*, *Miscanthus x giganteus* and *Flaveria bidentis*. *Plant, Cell & Environment*, *36*(2), 365-381.

Vance, C. P. (2001). Symbiotic nitrogen fixation and phosphorus acquisition. Plant nutrition in a world of declining renewable resources. *Plant Physiology*, *127*(2), 390-397.

Vogan, P. J., & Sage, R. F. (2011). Water-use efficiency and nitrogen-use efficiency of C_3-C_4 intermediate species of *Flaveria* Juss.(Asteraceae). *Plant, Cell & Environment*, *34*(9), 1415-1430.

Vogan, P. J., & Sage, R. F. (2012). Effects of low atmospheric CO_2 and elevated temperature during growth on the gas exchange responses of C_3, C_3–C_4 intermediate, and C_4 species from three evolutionary lineages of C_4 photosynthesis. *Oecologia*, *169*(2), 341-352.

Von Caemmerer, S. (2000). *Biochemical models of leaf photosynthesis*. 2. Collingwood, Australia: Csiro publishing.

Wang, Y., Long, S. P., & Zhu, X. G. (2014). Elements Required for an Efficient NADP-ME Type C_4 Photosynthesis. *Plant Physiology*, pp-113.

Yin, X., & Struik, P. C. (2012). Mathematical review of the energy transduction stoichiometries of C_4 leaf photosynthesis under limiting light. *Plant, Cell & Environment*, *35*(7), 1299-1312.

Zhu, X. G., de Sturler, E., & Long, S. P. (2007). Optimizing the distribution of resources between enzymes of carbon metabolism can dramatically increase photosynthetic rate: a numerical simulation using an evolutionary algorithm. *Plant Physiology*, *145*(2), 513-526.

Tables

Table 1 Physiological parameters of *F. pringlei* (C_3) and *F. bidentis* (C_4) optimally adapted to environments that show different leaf nitrogen levels. The Parameters are the maximal Rubisco (V_{cmax}, [µmol m^{-2} s^{-1}]) and C_4 cycle (V_{pmax}, [µmol m^{-2} s^{-1}]) activity as well as the maximal electron transport (J_{max}, [µmol m^{-2} s^{-1}]) of the linear and cyclic electron transport, and the proportion ([fraction]) of LET.

	F. pringlei (C_3)			*F. bidentis* (C_4)		
Leaf nitrogen level [mmol m^{-2}]	50	150	250	50	150	250
Proportion of LET [fraction]	0.980	0.974	0.968	0.805	0.745	0.704
J_{max} of LET rate [µmol m^{-2} s^{-1}]	42	99	149	62	152	242
J_{max} of CET rate [µmol m^{-2} s^{-1}]	6	15	24	86	256	446
V_{cmax} [µmol m^{-2} s^{-1}]	46	102	145	24	48	62
V_{pmax} [µmol m^{-2} s^{-1}]	0	0	0	29	60	78
CO_2 assimilation rate [µmol m^{-2} s^{-1}]	3.7	8.2	11.6	17.0	36.5	48.4

Table 2 The C_4 advantage in the photosynthetic nitrogen-use efficiency (PNUE) is more pronounced in a low-nitrogen environment. The relative PNUE of *F. bidentis* (C_4) and *F. pringlei* (C_3) for low, moderate, and high leaf nitrogen levels.

Leaf nitrogen level [mmol m^{-2}]	PNUE of C_4 / PNUE of C_3 plants [unitless]
50	4.58
150	4.46
250	4.18

Table 3 The required change in nitrogen to transform an optimal C_3 into an optimal C_4 plant decreases for decreasing leaf nitrogen levels (“required nitrogen re-allocation”).

Leaf nitrogen level [mmol m^{-2}]	Required nitrogen re-allocation [fraction]
50	0.598
150	0.702
250	0.787

Figures

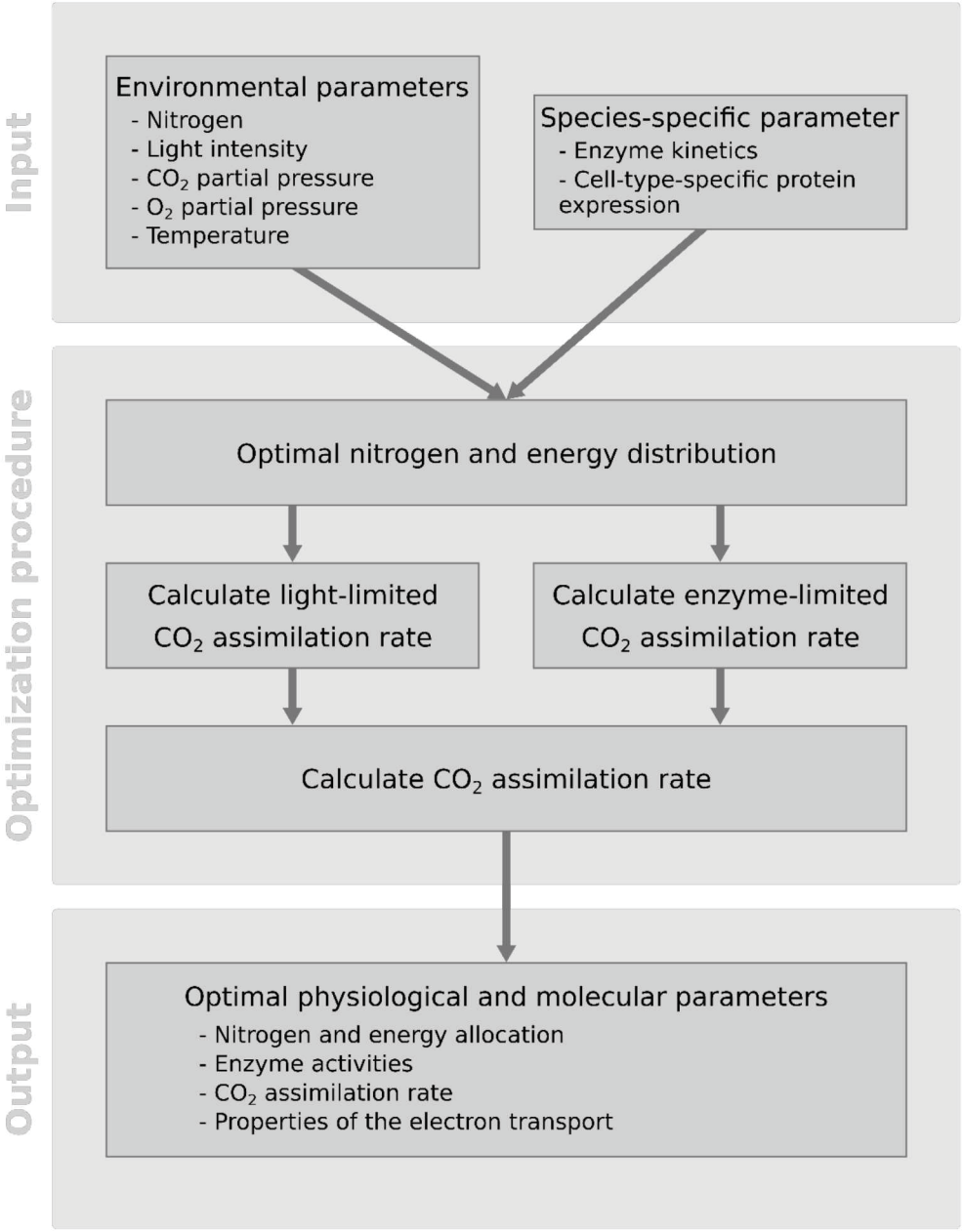


Figure 1 An overview of the nitrogen-dependent light and enzyme-limited model of Sundermann *et al.* (2018). For given environmental conditions and photosynthetic type, the optimization procedure determines the resource allocation that results in the maximal CO_2 assimilation rate. Based on the resource allocation, a wide range of physiological and molecular parameters are calculated.

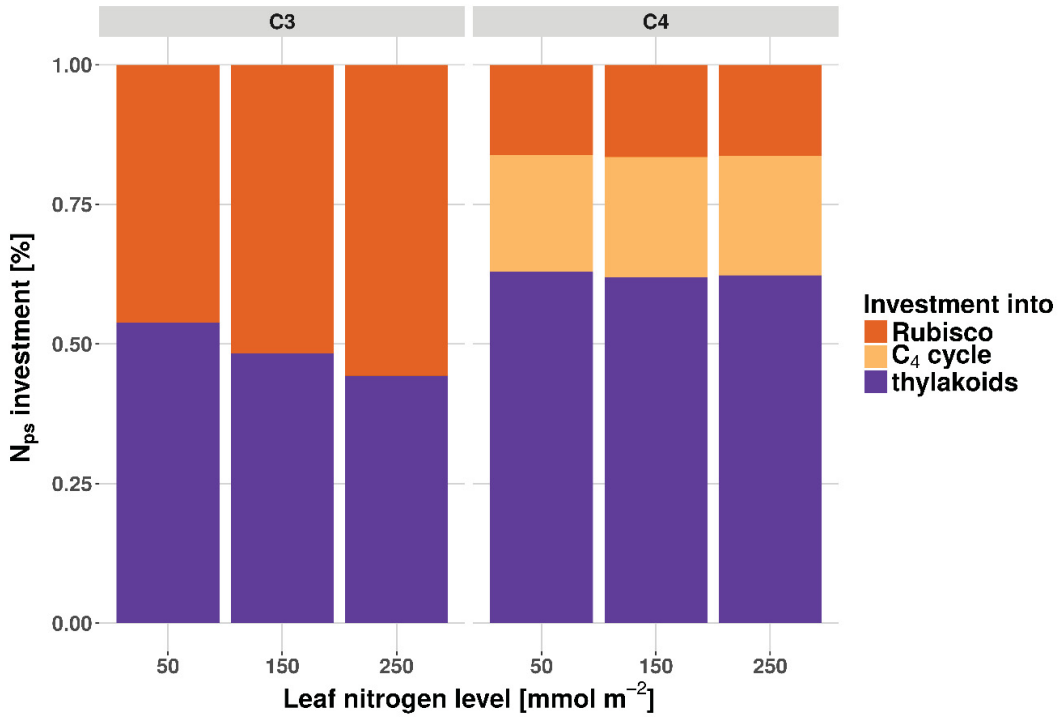


Figure 2 The nitrogen allocation pattern of *F. pringlei* (C_3) and *F. bidentis* (C_4) differs for various leaf nitrogen levels. It is assumed that the plants are optimally adapted to the same environment that only differs in the leaf nitrogen level. The nitrogen allocation into the photosynthetic pools (N_{ps}) comprises the investment into Rubisco, C_4 enzymes, and thylakoids—is presented.

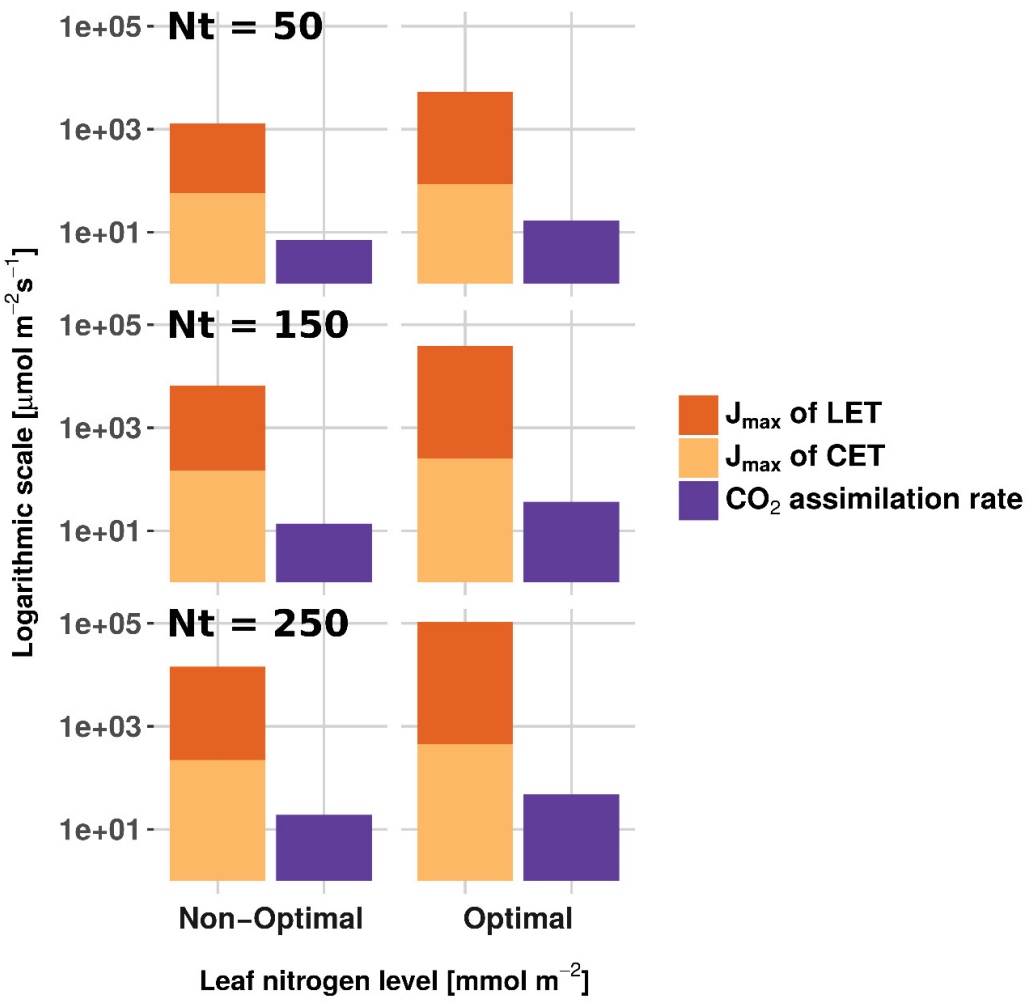


Figure 3 An optimally adapted C_4 plants allocates the nitrogen saved by reducing Rubisco not only into the C_4 cycle but also into the thylakoids. The effect of various leaf nitrogen levels on the maximal electron transport rate (J_{max}) of the LET and CET as well as the CO_2 assimilation rate for the optimal and the non-optimal scenario are presented, using a logarithmic scale. The non-optimal scenario assumes a C_4 plants that invests equal amounts of nitrogen in the thylakoids as a C_3 plants while the remaining nitrogen is optimally allocated between the C_4 enzymes and Rubisco.

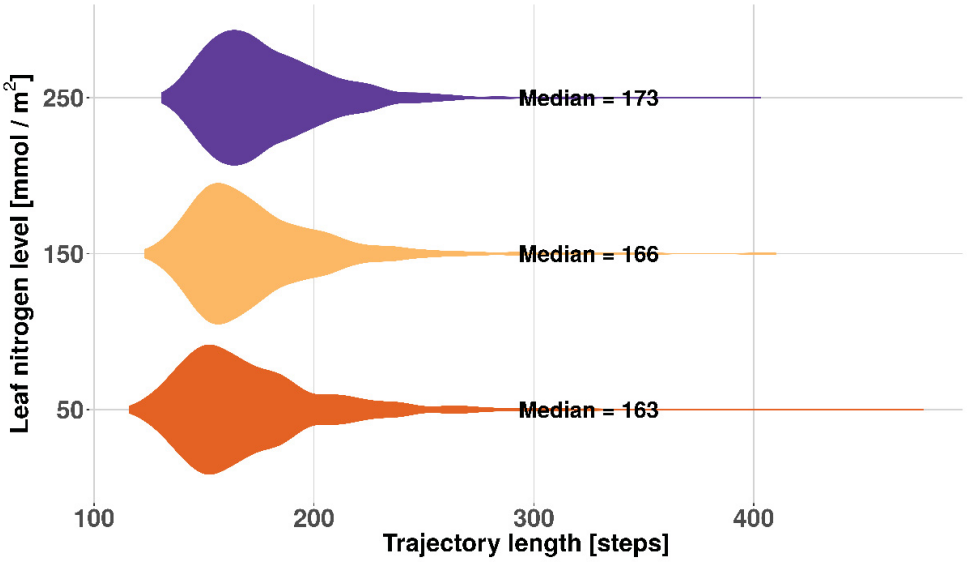


Figure 4 For decreasing leaf nitrogen levels, the number of evolutionary steps required to achieve a near optimal C_4 state decreases. A near optimal C_4 state is defined as showing at least 90% of C_4-specific attributes (δ_{C4}). All trajectories result in a near optimal C_4 state. 1,000 trajectories were simulated for each leaf nitrogen level.

Supplementary Information

Methods

Method S1

The evolutionary environment shows 1344 µmol quanta m^{-2} s^{-1} light intensity, 30°C, a mesophyll CO_2 level of 100 µbar, and a mesophyll O_2 level of 200 mbar, which corresponds to about 280 µbar atmospheric CO_2 concentration and an atmospheric O_2 concentration of 210 mbar.

Method S2

The increase in the photosynthetic nitrogen level (f_N, [factor]) is determined by the following equation that depends on the maximal C_4 cycle activity (V_{pmax}, [µmol m^{-2} s^{-1}]).

$$f_N = \begin{cases} 0 & if\ V_{pmax} < 0.3 \\ 0.1 & if\ V_{pmax} > 21.7 \\ g_x & otherwise \end{cases}$$

with

$$g_x = \frac{0.1}{1 + \exp\left(-0.38 \cdot \left(-6 + 12 \cdot \frac{V_{pmax} - 4}{12.7}\right)\right)}$$

Method S3

In vitro PEPC activity, given by Heckmann *et al.* (2013), can be transformed into *in vivo* activity by dividing the *in vitro* activity through three (Studer *et al.*, 2014; Tovar-Méndez *et al.*, 2000, Laisk *et al.*, 1997).

Tables

Table S1 Model parameterization for different species of *Flaveria* and *Zea mays*. β, fraction of Rubisco expressed in the mesophyll cell [fraction]; k_{ccat}, the maximal turnover rate of Rubisco [1/s]; V_{pmax}, the empirical maximal C_4 cycle activity determined by the PEPC activity [µmol m^{-2} s^{-1}]; K_p, Michaelis constant of PEPC [µbar]; g_s, bundle sheath conductance [µmol m^{-2} s^{-1}]; ξ, fraction of mesophyll cell derived photorespiration in the bundle sheath cells. All *Flaveria* parameters are taken from Sundermann *et al.* (2018). For *Z. mays*, k_{ccat} is taken from Hermida-Carrera *et al.* (2016); K_p and g_s are taken from Massad *et al.* (2007); and β and ξ are chosen according to the C_4 parametrization of Heckmann *et al.* (2013). The maximal PEPC activity of the C_4 species is assumed to be unlimited, *i.e.*, the upper bound is set to 1000.

Species	Photosynthetic types	β	k_{ccat}	V_{pmax}	K_p	g_s	ξ
F. pringlei	C_3	0.95	3.11	0	200	0.015	0
F. bidentis	C_4	0.008	4.16	1000	80	0.001	0.96
Z. mays	C_4	0.002	4.05	1000	40	0.003	0.98

Table S2 Parameters that evolve, *i.e.*, can be changed during the evolutionary simulations. β, fraction of Rubisco expressed in the mesophyll cell [fraction]; k_{ccat}, the maximal turnover rate of Rubisco [1/s]; K_p, Michaelis constant of PEPC [µbar]; g_s, bundle sheath conductance [µmol m^{-2} s^{-1}]; ξ, fraction of mesophyll cell derived photorespiration in the bundle sheath cells; the proportion of LET (p); the nitrogen allocation into Rubisco, the C_4 cycle, and the thylakoids, (n_{Etot}, n_{C4}, n_{Jmax}, [fraction]). The parameters are taken from the C_3 and C_4 *Flaveria* species presented in Supplementary Information Table S1; the description optimized (OP) indicates that the parameters are determined in the optimization procedure and depend on the leaf nitrogen level (the exact values are presented in Supplementary Information Table S4).

Photosynthetic type	β	k_{ccat}	K_p	g_s	ξ	p	n_{Etot}	n_{C4}	n_{Jmax}
C_3	0.95	3.11	200	0.015	0	OP	OP	OP	OP
C_4	0.008	4.16	80	0.001	0.96	OP	OP	OP	OP

Table S3 The data from literature and the corresponding modeled results for *F. bidentis* (C_4) and *Z. mays* (C_4) are in very good agreement except for the total maximal electron transport rate per maximal Rubisco catalytic rate ratio. Various leaf nitrogen levels result in a range of model estimations. Details about the ranges are presented in the Supplementary Information Tables S5−S8. The values, the reference, and the type (estimation or measurement) of the related work are presented.

	Parameter	Model estimations	Value of related work	Type and reference of related work
F. bidentis	Fraction of electron transport by CET [fraction]	0.53–0.54	0.51–0.54	Estimations of Yin *et al.* (2018)
			0.53	Estimation by Yin *et al.* (2012)
	Total maximal electron transport rate per maximal Rubisco catalytic rate [unitless]	6.24–11.02	4.62	Estimation of Domingues *et al.* (2007)
			5.97	Estimation of Massad *et al.* (2007)
			2.46	Estimation of Kathilankal *et al.* (2011)
			<2	Estimations of Ge *et al.* (2014)
			4.61	Estimation of Vico *et al.* (2008)
	ATP allocation into C_4 cycle [fraction]	0.42–0.44	0.4	Estimation of von Caemmerer (2000)
	ATP allocation into Calvin cycle and Photorespiration in the bundle sheath cells [fraction]	0.57–0.58	0.6	Estimation of von Caemmerer (2000)
	Fraction of ATP from LET [fraction]	0.56–0.57	0.56–0.59	Estimations of Yin *et al.* (2018)
Z. mays	Nitrogen investment into Rubisco [fraction]	0.18–0.19	0.25–0.28	Measurements of Mu *et al.* (2016)
	Nitrogen investment into the C_4 enzymes [fraction]	0.17–0.18	0.10–0.12	Measurements of Mu *et al.* (2016)
	Nitrogen investment into the thylakoids [fraction]	0.63–0.64	0.61–0.65	Measurements of Mu *et al.* (2016)
	PNUE measured at experimental conditions [µmol CO_2/(mmol N s)]	0.25–0.3	0.18–0.28	Measurements of Mu *et al.* (2016)

Table S4 The optimized parameters as a function of leaf nitrogen level for the C_3 and C_4 endpoint of the fitness landscape. The proportion of LET (*p*, [fraction]); the nitrogen allocation into Rubisco, the C_4 cycle, and the thylakoids, (n_{Etot}, n_{C4}, n_{Jmax}, [fraction]).

			Photosynthetic type	
			C_3	C_4
Leaf nitrogen level [mmol m^{-2}]	50	*p*	0.980	0.806
		n_{Etot}	0.461	0.162
		n_{C4}	0	0.208
		n_{Jmax}	0.539	0.630
	150	*p*	0.974	0.745
		n_{Etot}	0.516	0.165
		n_{C4}	0	0.216
		n_{Jmax}	0.484	0.619
	250	*p*	0.968	0.704
		n_{Etot}	0.557	0.163
		n_{C4}	0	0.214
		n_{Jmax}	0.443	0.623

Table S5 Allocation of ATP and NADPH for *F. pringlei* (C_3) and *F. bidentis* (C_4) in dependency on the leaf nitrogen level. M = mesophyll cell; BS = bundle sheath cell.

		F. pringlei (C_3)			*F. bidentis* (C_4)		
	Leaf nitrogen level [mmol m^{-2}]	50	150	250	50	150	250
ATP allocation	Proportion of ATP invested Rubisco in the M (%)	0.951	0.953	0.954	0.001	0.001	0.001
	Proportion of ATP invested into Rubisco in the BS (%)	0.049	0.047	0.046	0.564	0.573	0.575
	Proportion of ATP invested into the C_4 cycle (%)	0	0	0	0.435	0.426	0.424
NADPH allocation	Proportion of NADPH invested into Rubisco in the M (%)	0.952	0.953	0.954	0.002	0.002	0.002
	Proportion of NADPH invested into Rubisco in the BS (%)	0.048	0.047	0.046	0.998	0.998	0.998

Table S6 The model estimation of the relative electron rate of CET (according to the definition of Yin *et al.* (2018)) of *F. pringlei* (C_3) and *F. bidentis* (C_4) in dependency on leaf nitrogen level.

	Leaf nitrogen level [mmol m^{-2}]		
	50	150	250
F. pringlei (C_3)	0.11	0.11	0.11
F. bidentis (C_4)	0.54	0.53	0.53

Table S7 The model estimation of the fraction of ATP produced by the LET of *F. bidentis* (C_4) in dependency on leaf nitrogen level.

	Leaf nitrogen level [mmol m^{-2}]		
	50	150	250
F. bidentis (C_4)	0.56	0.57	0.57

Table S8 Estimations for the ratio of the total maximal electron transport rate and maximal Rubisco catalytic rate for *F. pringlei* (C_3) and *F. bidentis* (C_4). The original values for total maximal electron transport rate and maximal Rubisco catalytic rate are taken from Table 1 in the main text.

	Leaf nitrogen level [mmol m^{-2}]		
	50	150	250
F. pringlei (C_3)	1.04	1.12	1.2
F. bidentis (C_4)	6.24	8.52	11.02

Table S9 For increasing leaf nitrogen levels, the modeled and empirical data show the same trend for *Zea mays* (data from Mu *et al.* (2016)). The nitrogen allocation into the thylakoids decreases and the investment into Rubisco and the C_4 cycle enzymes increase with increasing leaf nitrogen levels. The resulting photosynthetic nitrogen-use efficiency decreases for increasing leaf nitrogen level. The experimental conditions are 1600 µmol m^{-2} s^{-1} light intensity, 30°C and current atmospheric CO2, which is assumed to be equal 170 µbar mesophyll CO_2 concentration.

	Model estimation		Measurement	
Leaf nitrogen level [mmol m^{-2}]	110	177	110	177
Nitrogen investment into Rubisco [fraction]	0.183	0.186	0.247	0.275
Nitrogen investment into the C_4 enzymes [fraction]	0.174	0.176	0.103	0.119
Nitrogen investment into the thylakoids [fraction]	0.653	0.638	0.649	0.610
PNUE measured at experimental conditions [µmol CO_2/(mmol N s)]	0.303	0.253	0.276	0.179

Table S10 Comparison between data and modeled results of *F. pringlei* (C_3) are in good agreement for the fraction of electron transport by CET but not for the ratio of the total maximal electron transport rate and the maximal Rubisco catalytic rate. Various leaf nitrogen levels result in a range of model estimations. Details about the ranges are presented in Table 1 in main text and Supplementary Information Tables S6. The values, the reference, and type (estimation or measurement) of the related work are presented.

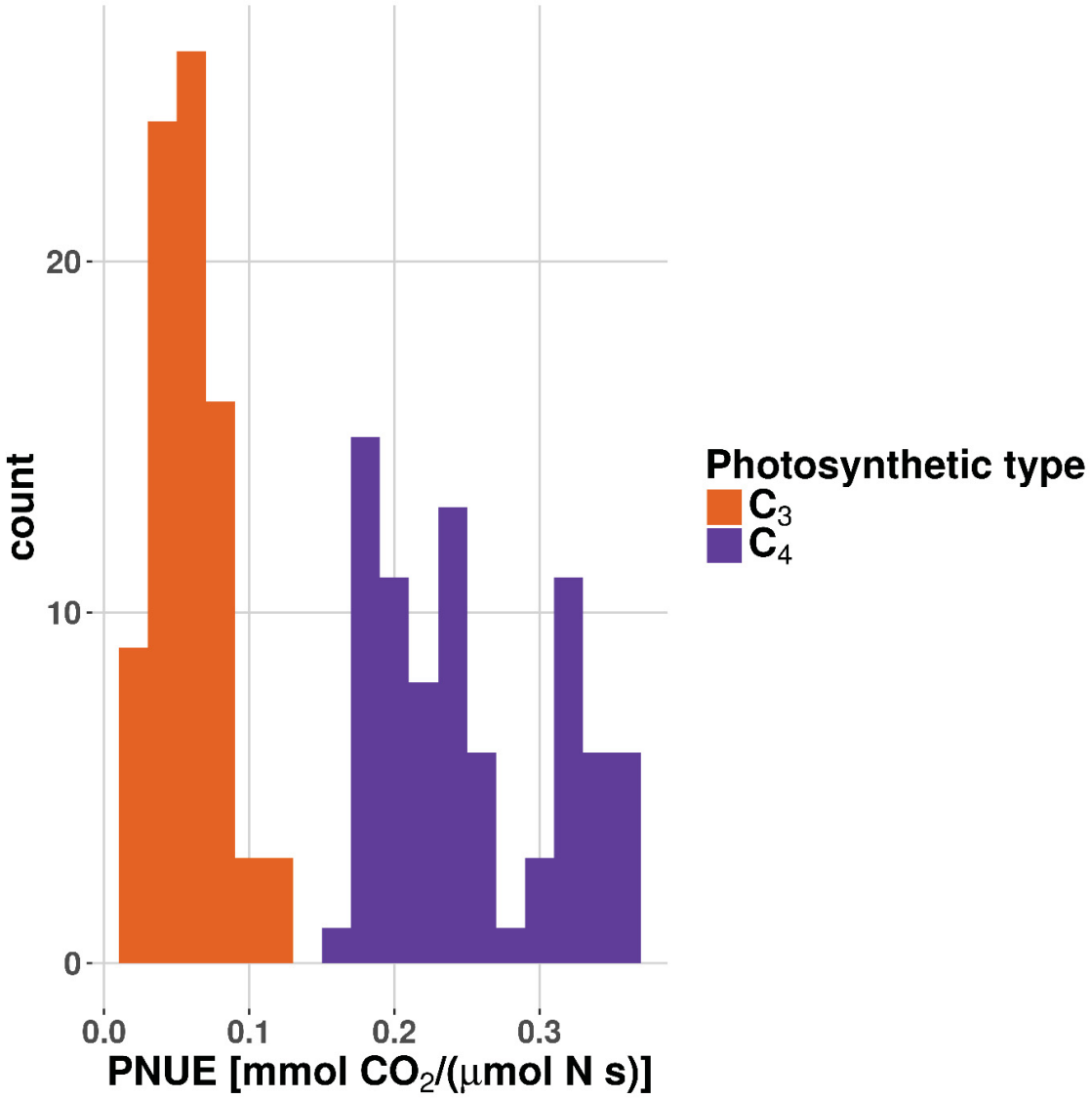
Parameter	Model estimation	Value of related work	Type and reference of the related work
Fraction of electron transport by CET [fraction]	0.11	0.12	Measurement of Kramer *et al.* (2010)
Total maximal electron transport rate per maximal Rubisco catalytic rate [unitless]	1.04-1.2	2±0.6	Measurements of Leuning (2002)

Table S11 Number of steps considered for evolutionary simulations of the optimized parameters—nitrogen allocation and proportion of LET—for moderate and high leaf nitrogen level.

	Leaf nitrogen level [mmol m^{-2}]	
	150	250
Number of steps considered for the nitrogen investment into Rubisco	105	109
Number of steps considered for the nitrogen investment into the C_4 enzymes	100	100
Number of steps considered for the nitrogen investment into the thylakoids	166	240
Number of steps considered for the proportion of LET	132	153

Figures

Figure S1 The results for the photosynthetic nitrogen-use efficiency (PNUE) is robust against variation in the evolutionary environment. That means that there is a consistent trends that C_4 plants are more PNUE compared to C_3 plants under all tested environments. For the three leaf nitrogen levels and for both—C_3 and C_4 plants—27 environmental variation of the evolutionary environment are tested. All combinations of light intensity (1344±100 [µmol quanta m^{-2} s^{-1}]), temperature (30±3°C), CO_2 mesophyll concentration (100±25 [µbar]), and leaf nitrogen levels (150±100 [mmol m^{-2}]) are considered in the analysis. The mesophyll O_2 concentration is constant at 200 mbar.



References

Domingues, T. F., Martinelli, L. A., & Ehleringer, J. R. (2007). Ecophysiological traits of plant functional groups in forest and pasture ecosystems from eastern Amazonia, Brazil. *Plant Ecology*, *193*(1), 101-112.

Ge, Z. M., Zhang, L. Q., Yuan, L., & Zhang, C. (2014). Effects of salinity on temperature-dependent photosynthetic parameters of a native C3 and a non-native C4 marsh grass in the Yangtze Estuary, China. *Photosynthetica*, *52*(4), 484-492.

Heckmann, D., Schulze, S., Denton, A., Gowik, U., Westhoff, P., Weber, A. P., & Lercher, M. J. (2013). Predicting C_4 photosynthesis evolution: modular, individually adaptive steps on a Mount Fuji fitness landscape. *Cell*, *153*(7), 1579-1588.

Hermida-Carrera, C., Kapralov, M. V., & Galmés, J. (2016). Rubisco catalytic properties and temperature response in crops. *Plant Physiology*, pp-01846.

Kathilankal, J. C., Mozdzer, T. J., Fuentes, J. D., McGlathery, K. J., D'Odorico, P., & Zieman, J. C. (2011). Physiological responses of Spartina alterniflora to varying environmental conditions in Virginia marshes. *Hydrobiologia*, *669*(1), 167.

Kramer, D. M., & Evans, J. (2010). Update: the importance of energy balance in improving photosynthetic productivity. *Plant Physiology*, pp-110.

Laisk, A., & Edwards, G. E. (1997). CO2 and temperature-dependent induction in C_4 photosynthesis: an approach to the hierarchy of rate-limiting processes. *Functional Plant Biology*, *24*(4), 505-516.

Leuning, R. (2002). Temperature dependence of two parameters in a photosynthesis model. *Plant, Cell & Environment*, *25*(9), 1205-1210.

Massad, R. S., Tuzet, A., & Bethenod, O. (2007). The effect of temperature on C_4-type leaf photosynthesis parameters. *Plant, Cell & Environment*, *30*(9), 1191-1204.

Mu, X., Chen, Q., Chen, F., Yuan, L., & Mi, G. (2016). Within-leaf nitrogen allocation in adaptation to low nitrogen supply in maize during grain-filling stage. *Frontiers in Plant Science*, *7*, 699.

Studer, A. J., Gandin, A., Kolbe, A. R., Wang, L., Cousins, A. B., & Brutnell, T. P. (2014). A limited role for carbonic anhydrase in C_4 photosynthesis as revealed by a ca1ca2 double mutant in maize. *Plant Physiology*, pp-114.

Sundermann, E., Lercher, M., & Heckmann, D. (2018). Modeling Cellular Resource Allocation Reveals Low Phenotypic Plasticity of C_4 Plants and Infers Environments of C_4 Photosynthesis Evolution. *bioRxiv*, 371096.

Tovar-Méndez, A., Mújica-Jiménez, C., & Muñoz-Clares, R. A. (2000). Physiological implications of the kinetics of maize leaf phosphoenolpyruvate carboxylase. *Plant Physiology*, *123*(1), 149-160.

Vico, G., & Porporato, A. (2008). Modelling C_3 and C_4 photosynthesis under water-stressed conditions. *Plant and Soil*, *313*(1-2), 187-203.

Von Caemmerer, S. (2000). *Biochemical models of leaf photosynthesis*. Csiro publishing.

Yin, X., & Struik, P. C. (2012). Mathematical review of the energy transduction stoichiometries of C_4 leaf photosynthesis under limiting light. *Plant, Cell & Environment*, *35*(7), 1299-1312.

Yin, X., & Struik, P. C. (2018). The energy budget in C_4 photosynthesis: insights from a cell-type-specific electron transport model. *New Phytologist*, *218*(3), 986-998.

4

DISCUSSION

Mathematical models allow researchers to test hypothesized concepts, to deepen understanding, and they inspire novel ways to interpret empirical data (Brodland, 2015). In addition, modeling facilitates the simulation of evolution within a reasonable timescale. The growing performance of computers allows scientists to simulate highly complex (biological) problems, such as the evolution of a complex trait and the determination of context-specific optimality, a concept of particular interest in biology (see Chapter 2). As a result, *in silico* analyses are a powerful tool to explore organismal metabolisms and their evolution.

4.1 MODELING THE ENVIRONMENTAL EFFECTS ON METABOLISM

In this thesis, the effect of environmental factors on metabolic efficiency and the resulting consequences for evolution and ecology are explored with the help of a mechanistic model. This model will be referred to as nitrogen-dependent light- and enzyme-limited model, in the following. The presented work focuses on the metabolism of photosynthetic organisms, more precisely on C_3, C_3-C_4 intermediate, and C_4 photosynthesis in higher plants. The presented mechanistic model facilitates the analysis of the optimal nitrogen and energy allocation, as well as physiological parameters of photosynthesis in a multifaceted environment.

The model and the corresponding analyses are presented in *Manuscript 1* and *Manuscript 2*. In *Manuscript 1*, a drastic difference between measured and predicted PEPC activity is observed. This difference likely results from a known discrepancy between *in vitro* and *in vivo* PEPC activity (Laisk and Edwards, 1997; Studer et al., 2014; Tovar-Méndez et al., 2000). The *in vitro* values are taken from Dwyer et al. (2007), the empirical study considered in *Manuscript 1*, and the *in vivo* values correspond to the predicted activity from the mathematical model. The *in vitro* PEPC activity is about two to five times higher than the *in vivo* one (Laisk and Edwards, 1997; Studer et al., 2014; Tovar-Méndez et al., 2000). We assume a factor of three to transform the measured *in vitro* into the *in vivo* enzyme activity. When repeating the calculation presented in *Manuscript 1* with the *in vivo* PEPC activity instead of the *in vitro* values, we observe minor differences in the results. The predicted PEPC activity and the empirically expected activity are in good agreement, they agree within a factor of 0.86. Another difference is that the environment that exhibits the minimal discrepancy between the modeled and the measured data for the C_4

Flaveria species shows 100 μbar atmospheric CO_2 concentration, 30 °C, and 1344 μmol m^{-2} s^{-1} light intensity. Compared to the in *Manuscript 1* inferred environment, only the light intensity changes, it was 1562 μmol m^{-2} s^{-1} (Figure 5 in *Manuscript 1*). This environment is analyzed in detail in *Manuscript 2*.

4.1.1 *Modeling the Environmental Effects on Resource Allocation*

In this section the findings of this thesis will be discussed with respect to the effect of environmental factors on resource allocation.

4.1.1.1 *Research Contributions*

Plants face a multifaceted environment where often a combination of environmental factors affect their metabolism. A frequently occurring assumption is that plants perform under light-saturated conditions (theoretical work, e. g., Wang et al. (2014) and Zhu et al. (2007) and empirical measurements, e. g., Makino et al. (2003) and Vogan and Sage (2012)). However, plants often face light-limited environments. This is particularly true for crop plants that usually shade each other as they grow on fields in close spatial proximity. The nitrogen-dependent light- and enzyme-limited model captures a complex environment and facilitates the analysis of the effect of single environmental factors and also their combined effect. Light-limited conditions can be represented by the model presented in *Manuscript 1* that considers different light intensities and explicitly models the ATP and NADPH production of linear and cyclic electron transport. One important feature of the model is the fixed nitrogen budget that determines the protein abundances of the enzymes of the Calvin-Benson cycle in the mesophyll and bundle sheath cell, the C_4 cycle, and the proteins of the linear and cyclic electron transport in the thylakoid membranes. The presented mechanistic model relates nitrogen investment to protein abundance. Consequently, a balanced parametrization of major potentially limiting processes, the energy production and consumption, is modeled. In addition, current mechanistic understanding of the costs of photosynthesis and the effect of environmental factors on the metabolic performance can be validated (*Manuscript 1* and *Manuscript 2*). In summary, the nitrogen-dependent light- and enzyme-limited model is a comprehensive mechanistic model that is suitable to analyze the optimal resource allocation—nitrogen and energy—of C_3, C_3-C_4 intermediate, and C_4 photosynthesis in dependence of multiple environmental factors. This model encompasses (1) leaf nitrogen level, (2) light intensity, (3) temperature, and (4) CO_2 and (5) O_2 gas concentrations.

Although the phenotypic plasticity in plants is intensively studied (e. g., Sage and McKown (2005)), it is not yet fully understood. The work presented in this thesis contributes to the understanding of

the constraints on resource allocation by quantifying the amount of nitrogen that needs to be shifted between photosynthetic pools in order to optimally adapt from an ancestral to a current environment (*Manuscript 1*). Modeling results suggest that C_4 plants need to shift a high amount of nitrogen between the photosynthetic pools compared to C_3 plants. This points to a possibly causal linkage between phenotypic plasticity and the ability to re-allocate nitrogen, which is a novel hypothesis.

A comparison of modeled and empirically observed data facilitates the determination of the environment to which C_4 plants are optimally adapted to. This comparison indicates that C_4 *Flaveria* species are optimally adapted to an ancestral environment and show a limited phenotypic plasticity (*Manuscript 1*). As the genus *Flaveria* relatively recently split into C_3 and C_4 species, about 2–3 million years ago (Christin et al., 2011), the limited phenotypic plasticity in C_4 plants allows us to refine the ancestral environment of this genus that likely triggers the evolution of C_4 photosynthesis.

The optimal resource allocation as well as the resulting molecular and physiological parameters in C_3 and C_4 *Flaveria* species, in the environment where C_4 *Flaveria* species likely evolved, are presented in *Manuscript 2*. We determine energy production and energy consumption. The energy production is linked to the maximal electron transport rate and the proportion of linear electron transport, while consumption is determined by the Rubisco activity and (if applicable) C_4 cycle activity. First, we contribute to the understanding of the effect of establishing a C_4 cycle and Rubisco reduction on metabolic efficiency. The modeled results are consistent with empirical findings that indicate that the nitrogen requirements of the C_4 cycle utilize only a part of the nitrogen that is available due to Rubisco reduction (Makino et al., 2003). Our work adds a quantitative insight on how the additional nitrogen affects the energy requiring processes—RuBP regeneration and C_4 cycle activity—and the resulting increase in the CO_2 assimilation rate. Second, in *Manuscript 2*, we focus on the effect of various leaf nitrogen levels on the plants performance. The relationship between energy production and consumption is studied intensively in C_3 species (e. g., Leuning (2002)), but in C_4 this relationship is less clear. We add quantitative information to this relation for C_4 and for C_3 species by analyzing this ratio as a function of leaf nitrogen. The gained insights facilitate the exploration of the photosynthetic nitrogen-use efficiency, i. e., CO_2 assimilation rate per leaf nitrogen level, in detail. We present a novel hypothesis: The relative advantage in photosynthetic nitrogen-use efficiency of C_4 over C_3 species increases for decreasing leaf nitrogen levels (*Manuscript 2*).

4.1.1.2 *Related Work*

The atmospheric CO_2 concentration changed drastically during the last million years (Sage et al., 2012). It has a strong effect on photosynthesis and, consequently, on plants performance (see Chapter 2.1). Wang et al. (2014) determined the optimal nitrogen allocation of C_4 metabolism as a function of atmospheric CO_2 concentration under light-saturated conditions and a fixed amount of available nitrogen. They analyzed a systems model that includes the Calvin-Benson cycle, the C_4 shuttle, and also the synthesis of carbon-rich substances, e. g., starch. Their work shows that a C_4 plant needs to increase the investment into Rubisco while decreasing the one into PEPC, in order to optimally adapt from a preindustrial concentration (atmospheric CO_2 concentration of 27.5 Pa) to the current atmospheric conditions (39.45 Pa). This trend is in agreement with the results simulated by the mechanistic model presented in *Manuscript 1* under the environmental conditions described in Wang et al. (2014) (see Section A.1 for exact values).

In C_3 photosynthesis, the distribution of resources between enzymes of the carbon metabolism is analyzed by Zhu et al. (2007). They determined the optimal distribution for plants by considering a systems model that represents the photosynthetic carbon metabolism including starch and sucrose production. According to their predictions, the adaptation from a preindustrial concentration (an intracellular CO_2 concentration of 165 μmol mol^{-1}) to the current atmospheric conditions (an intracellular CO_2 concentration of 280 μmol mol^{-1}) requires a decreased investment into Rubisco. This trend is consistent with the results from the nitrogen-dependent light- and enzyme-limited model, when the environmental conditions described in Zhu et al. (2007) are considered (see Section A.2 for exact values).

The models of Wang et al. (2014) and Zhu et al. (2007) show common features, as they are designed using the same modeling pipeline. Comparing these metabolic models with the nitrogen-dependent light- and enzyme-limited model, all models determine the nitrogen allocation that results in the maximal fitness. Although all models consider the CO_2 assimilation rate as a proxy for fitness, Wang et al. (2014) and Zhu et al. (2007) consider the rate at light-saturation while the nitrogen-dependent light- and enzyme-limited model considers a more precise fitness proxy—the assimilation rate for a given environmental condition—, which can represent non-saturated conditions at various temperatures. Hence, the models of Wang et al. (2014) and Zhu et al. (2007) and the nitrogen-dependent light- and enzyme-limited model differ in the complexity of environmental representation. The model presented in this thesis focuses on the reactions relevant for carbon fixation, i. e., the nitrogen budget and the nitrogen pools that are directly related to the carbon fixation are modeled exclusively. In the models of Wang et al. (2014) and Zhu et al. (2007), additionally the

nitrogen budget and the reactions relevant for the synthesis of starch and sucrose are considered. As the presented work focuses on the CO_2 fixation, these additions are a subsidiary matter: we assume that the synthesis of carbon-rich substances scales proportional to the CO_2 assimilation rate.

Organisms differ in their phenotypic adaptation to new environments. Short-term adjustment in an organismal behavior, morphology, or physiology as a result of a changed environment is denoted phenotypic plasticity (Price et al., 2003). The ability to adapt traits within a short timescale is of particular importance for sessile organisms like higher plants. A well studied example for phenotypic plasticity is the ability to acclimate to shade. C_3 plants are able to reduce the Rubisco activity and content when grown under low-light intensities (Sage and McKown, 2005). The responses of C_4 plants to changes in light availability appear to be less clear than those of C_3 plants (Sage and McKown, 2005). Sage and McKown (2005) reviewed the occurrence of phenotypic plasticity in C_3 and C_4 plants and concluded that C_4 plants show inherent constraints that prevent the acclimation to environmental changes. Their conclusion is consistent with the results presented in *Manuscript 1*, which indicate that C_4 plants are not optimally adapted to current conditions. Rather, these plants appear to be optimally adapted to an ancient environmental condition and show limited phenotypic plasticity.

4.1.1.3 *Outlook*

Manuscript 1 suggests a link between limited phenotypic plasticity and the amount of nitrogen re-allocation. This linkage provides a hypothesis for empirical scientists that can explain a potential reason for the constrained plasticity of C_4 plants. If the causal linkage can be verified through empirical data, this linkage indicates that the nitrogen allocation of plants is non-optimal under current environmental conditions and cannot adjust within short timescale. To design optimized crop plants, the development and conductance of bioengineering approaches is essential.

Manuscript 2 presents the hypothesis that a decreasing leaf nitrogen level results in a higher advantage of C_4 compared to C_3 plants in terms of relative photosynthetic nitrogen-use efficiency. This hypothesis needs verification by empirical studies. Our work highlights the importance of differences in the efficiency of resource usage that result from different photosynthetic types. The inherent nitrogen-use efficiency is of special importance in the context of agriculture, where high amounts of nitrogen fertilizers are used. The consideration of the efficiency of different photosynthetic types may lead to an improvement of yield per applied supplement.

The future environment will be drastically different from current conditions due to climate change. To ensure food and energy security

in the future, analyses on how to optimize resource allocation and, thus, metabolic efficiency of crop plants is essential. The fact that C_4 plants are less phenotypic plastic than C_3 plants highlights the need to explore the properties of future crop plants. and to create a blueprint for bioengineering approaches to design optimal plants for future conditions. To address these challenges, a close collaboration between theoretical and empirical scientists is necessary.

4.1.2 *Modeling the Environmental Effects on Metabolic Evolution*

In the following section the results presented in *Manuscript 1* and *Manuscript 2* are discussed. The focus of this section lies on the effect of environmental factors on evolution.

4.1.2.1 *Research Contributions*

C_4 photosynthesis is a complex trait whose evolution is of great interest for multiple disciplines beyond botany, e. g., geology and zoology (Sage, 2004). As mathematical models are able to simulate the evolution within a reasonable timescale, the evolutionary trajectories can be simulated time- and cost-efficiently. The model and the corresponding evolutionary simulations presented in this thesis are particularly useful to explore the evolution of the complex C_4 syndrome and its dependence on environmental factors. The usefulness of the model results from its comprehensiveness; it comprises diverse parameters that define the photosynthetic pathway and simulates diverse environmental settings, considering the simultaneous effect of five different environmental factors. The evolutionary simulations include features that are relevant to realistically model evolutionary processes. These features include a random extent of change for a specific trait, in contrast to equidistant changes per evolutionary step, and the chance of every trait to change toward a more C_3- or C_4-like manner. The evolutionary simulations predict evolutionary trajectories based on the current mechanistic understanding of C_3 and C_4 photosynthesis and do not require information about potential intermediate states. In general, the model and evolutionary simulations can be used to elucidate the structure of the fitness landscape as a function of diverse environmental factors. In *Manuscript 2*, we focus on the effect of different leaf nitrogen levels on plants performance while considering the ancestral environment where the evolution of C_4 photosynthesis in the genus *Flaveria* likely happened.

Although the importance of the environmental factors that are considered in this thesis are known in the context of photosynthesis, we present a novel hypothesis that deals with the role of the leaf nitrogen level in the likely time of speciation of C_3 and C_4 *Flaveria* species (*Manuscript 2*). In *Manuscript 2*, we compared the required changes in nitrogen to transform a C_3 plant into an C_4 plant for various

leaf nitrogen levels. It is assumed that both plants that are optimally adapted to the considered ancestral environment. The required change decreases for decreasing leaf nitrogen levels. In addition, we simulated evolutionary trajectories for various leaf nitrogen levels. For decreasing leaf nitrogen levels, a significant decrease in the number of steps is required to become a near optimal C_4 phenotype. A near optimal C_4 phenotype is a plant that shows at least 90 % of the C_4-specific attributes of a C_4 plant that is optimally adapted to the considered ancestral environment. These findings point to the possibility that a low leaf nitrogen level favors the evolution of C_4 photosynthesis.

The analysis of a past environment goes along with uncertainties about the plant characteristics in terms of nitrogen allocation, e. g., Rubisco availability, at that time. The nitrogen-dependent light- and enzyme-limited model allows us to determine the characteristics of the C_3 and C_4 plants that were likely observed during that time. These characteristics are based on the assumption that evolution selects for the most efficient metabolism, which is an important fitness determinant (Heckmann et al., 2013; Ibarra et al., 2002).

4.1.2.2 *Related Work*

Environmental factors are important selection pressures for photosynthetic metabolism (see Section 2.1). Heckmann et al. (2013) analyze the fitness landscape that maps C_3 to C_4 evolution in a C_4 favoring environment (see Section 2.2 for details). The work presented in *Manuscript 2* is consistent with Heckmann et al. (2013): Both results show that evolving a full C_4 metabolism is more advantageous than the C_3 pathway under the given C_4 favoring condition. This environment is characterized by high light intensities and moderate to high temperatures. Heckmann et al. (2013) consider an environmental condition that is relevant for the evolution of C_4 photosynthesis. In this thesis, we analyze the environment that is likely relevant in the context speciation of C_3 and C_4 *Flaveria* species, it is inferred in *Manuscript 1*. As *Flaveria* is a model genus in photosynthesis evolution, we propose that the conclusions based on our work with the *Flaveria* genus and the detailed environmental description have the potential to be general conclusions. Heckmann et al. (2013) assume light-saturation while in *Manuscript 2* an explicit value of 1344 μmol m^{-2} s^{-1} is considered. This results in a more realistic representation of the plants that show a high C_4 cycle activity, as C_4 photosynthesis is mainly limited by light (von Caemmerer, 2000, p. 116). There is a minor difference in the temperature, which is 5 °C warmer in *Manuscript 2* than in Heckmann et al. (2013). This effect is particularly important for the C_4 cycle activity, as PEPC shows no temperature-related inhibition and the highest increase in activity relative to Rubisco activity and electron transport rate (between 25 °C and 30 °C). However, the mesophyll CO_2 concentration differs drastically: 100 μbar in *Manuscript 2* and 250 μbar

in Heckmann et al. (2013). Consequently, photorespiratory rates are higher in C_3 plants and represent an even stronger driving force in *Manuscript 2*, compared to Heckmann et al. (2013) (see Section 2.1.4.2).

The nitrogen-dependent light- and enzyme-limited model, described in *Manuscript 1*, is an addition to the model presented in Heckmann et al. (2013). The initial model is extended by the effect of temperature, light limitation, and nitrogen dependence. These extensions result in a more precise fitness proxy due to the more detailed calculation of the CO_2 assimilation rate and the fact that not only Rubisco is considered as a nitrogen sink, but all photosynthetically-relevant sinks.

In *Manuscript 2* and Heckmann et al. (2013), a key component of the analysis is the simulation of evolutionary trajectories using Monte Carlo simulations. Heckmann et al. (2013) focus on the shape of the fitness landscape and deduce evolutionary paths from C_3 to C_4 photosynthesis. Our focus lies on the comparison of the chance of plants to evolve C_4 photosynthesis for environments that differ in their leaf nitrogen level. In order to explore the effect of nitrogen availability on C_4 evolution, we summarize the trajectories and analyze the phenotype at the end of the evolutionary path and the number of steps required to achieve this phenotype.

To simulate the evolution of photosynthesis, the changes of evolutionary parameters are considered. Evolutionary parameters define the phenotypic states of a plant and are relevant in the context of photosynthesis evolution. In *Manuscript 2* and Heckmann et al. (2013), changes in the evolutionary parameters are modeled as equidistant steps that show C_3 and C_4 photosynthesis as endpoints. Heckmann et al. (2013) consider six equidistant steps while *Manuscript 2* considers eleven. Note that the changes in the nitrogen allocation are normalized such that for each leaf nitrogen level the relative change is the same. The Monte Carlo simulation in Heckmann et al. (2013) follows the subsequent steps: (1) a trait is picked randomly (based on the mutational probability), (2) the CO_2 assimilation rate for the adjusted trait is calculated, and (3) based on the relative CO_2 assimilation rate the probability for the trait to fix in the population is derived from the population genetics model, first derived by Kimura (1957). Compared to the simulations of Heckmann et al. (2013), the Monte Carlo simulation presented in this thesis shows two major differences. First, the calculation of trajectories in *Manuscript 2* considers random extent of changes. In Heckmann et al. (2013), each change results in a fixed adjustment of one step toward a more C_4-like phenotype. In contrast, in the model presented in this thesis each change can be up to 10 % of the maximal number of steps. The actual change is picked randomly, based on a uniform distribution that is additionally constrained by a minimal change of one step and the fact that the change needs to result in a non-negative difference in fitness. Second, our work models the reverse evolution of photosynthesis. This means that in addition

to the changes toward a more C_4-like characteristic (as considered in Heckmann et al. (2013)), changes toward a more C_3-like manner are also considered.

4.1.2.3 *Outlook*

The findings presented in *Manuscript 2* point to the possibility that low leaf nitrogen level is a promoting factor for C_4 evolution and that the superiority in the photosynthetic nitrogen-use efficiency of C_4 plants relative to C_3 species increases for decreasing leaf nitrogen levels. The symbiotic relationship of a plant with mutualistic, nitrogen fixing microorganisms results in an increased amount of available nitrogen for the plant (Udvardi and Poole, 2013). Due to the difference in photosynthetic nitrogen-use efficiency, the likelihood to establish a mutualistic symbiosis with nitrogen fixing microorganisms for C_3 and C_4 plants potentially differ in dependence on the surrounding environment. Potentially there is a low chance for a C_3 plant that shows symbiosis to evolve C_4 photosynthesis compared to a C_3 plant that does not show symbiosis (Brown, 1978). Based on the finding that nitrogen re-allocation is higher and the evolutionary paths toward the C_4 phenotype are longer when nitrogen is less scarce, the likelihood to evolve the C_4 pathway is lower for a C_3 plant that already shows a symbiotic relationship compared to a C_3 plant without this relationship. Due to other factors that prompt C_4 photosynthesis evolution, it is still expected that these C_3 plants evolve into C_4 plants under the considered environmental conditions. The presented research questions need to be addressed by theoretical and empirical scientists of disciplines such as microbiology and phytology.

C_4 metabolism evolved from the ancestral C_3 pathway. Scientists successfully strengthened the understanding of C_4 evolution, e. g., in terms of the evolutionary steps and enabling factors (e. g., Heckmann (2016), Heckmann et al. (2013), and Sage (2004)). Moreover, the effect of environmental factors on C_4 evolution has been studied intensively (e. g., Sage (2004) and *Manuscript 2*). In contrast, the occurrence of reverse evolution from C_4 to C_3 photosynthesis is currently inconclusive. The evolution from C_4 toward C_3 metabolism has been suggested (e. g., see Ibrahim et al. (2009)), but the effects of environmental factors on the corresponding trajectories are not yet understood. The model presented in this thesis is an ideal tool to study this concept by simulating paths starting from C_4 photosynthesis in diverse environments. The evolution and its dependence on environmental factors may become important in the context of climate change, as an increasing CO_2 concentration suppresses photorespiration, which favors C_3 over C_4 photosynthesis in light-limited conditions.

A

APPENDIX

A.1 RESOURCE ALLOCATION OF C_4 PLANTS

The resource allocation of the C_4 plant *Flaveria bidentis* is calculated for a preindustrial concentration to current atmospheric conditions according to the environmental conditions given by Wang et al. (2014)—25 °C and light-saturation. The amount of available nitrogen is 1 g m^{-2} which corresponds to 71,429 μmol m^{-2}. We assume that 87.5 % of the available nitrogen is available for the CO_2 fixation. This corresponds to a leaf nitrogen level of 200 mmol m^{-2}. Table A.1 represents the nitrogen allocation at 200 mmol m^{-2}. The mesophyll concentration is derived from Vogan and Sage (2012) while assuming a ratio of mesophyll to intercellular CO_2 concentration of 0.85; the mesophyll concentration is 170 μbar and 115 μbar under current and preindustrial conditions, respectively. The C_4 plant is parametrized as *Flaveria bidentis* (*Manuscript 1*).

Table A.1: In order to optimality adapt from preindustrial concentration to current atmospheric conditions *Flaveria bidentis* (C_4) needs to increase the nitrogen investment into Rubisco while decreases the one into PEPC. The nitrogen investment into the photosynthetic sinks—Rubisco, C_4 cycle, and thylakoids—for plants that are optimally adapted to an environment that shows light-saturation, 25 °C, and either current or preindustrial atmospheric conditions.

Fraction of nitrogen investment	preindustrial condition	current condition
into Rubisco	0.156	0.161
into the C_4 cycle	0.250	0.229
into the thylakoids	0.594	0.609

A.2 RESOURCE ALLOCATION OF C_3 PLANTS

The resource allocation of the C_3 plant *Flaveria pringlei* is calculated for a preindustrial concentration to current atmospheric conditions according to the environmental conditions given by Zhu et al. (2007)—25 °C and light-saturation. As above, the amount of available nitrogen is 1 g m^{-2} and we assume that 87.5 % of the available nitrogen is available for the CO_2 fixation, which corresponds to a leaf nitrogen

level of 200 mmol m^{-2}. Table A.2 represents the nitrogen allocation at 200 mmol m^{-2}. The mesophyll concentration is derived from Vogan and Sage (2012) while assuming a ratio of mesophyll to intercellular CO_2 concentration of 0.85; the mesophyll concentration is 215 μbar and 170 μbar under current and preindustrial conditions, respectively. The C_3 plant is parametrized as *Flaveria pringlei* (*Manuscript 1*).

Table A.2: In order to optimality adapt from preindustrial concentration to current atmospheric conditions *Flaveria pringlei* (C_3) needs to decrease the nitrogen investment into Rubisco. The nitrogen investment into the photosynthetic sinks—Rubisco, C_4 cycle, and thylakoids—for plants that are optimally adapted to an environment that shows light-saturation, 25 °C, and either current or preindustrial atmospheric conditions.

Fraction of nitrogen investment	preindustrial condition	current condition
into Rubisco	0.494	0.481
into the C_4 cycle	0	0
into the thylakoids	0.506	0.519

BIBLIOGRAPHY

Aubry, S., N. J. Brown, and J. M. Hibberd (2011). "The role of proteins in C3 plants prior to their recruitment into the C4 pathway." In: *Journal of Experimental Botany* 62.9, pp. 3049–3059.

Baudouin-Cornu, P., Y. Surdin-Kerjan, P. Marliere, and D. Thomas (2001). "Molecular evolution of protein atomic composition." In: *Science* 293.5528, pp. 297–300.

Berry, J. and G. Farquhar (1978). *The CO_2 concentrating function of C4 photosynthesis: a biochemical model*. Biochemical Society, London.

Brodland, G. W. (2015). "How computational models can help unlock biological systems." In: *Seminars in Cell & Developmental Biology* 47–48. Coding and non-coding RNAs & Mammalian development, pp. 62–73.

Brown, R. (1978). "A difference in N use efficiency in C3 and C4 plants and its implications in adaptation and evolution." In: *Crop Science* 18.1, pp. 93–98.

Christin, P.-A., C. P. Osborne, R. F. Sage, M. Arakaki, and E. J. Edwards (2011). "C4 eudicots are not younger than C4 monocots." In: *Journal of Experimental Botany* 62.9, pp. 3171–3181.

Dwyer, S. A., O. Ghannoum, A. Nicotra, and S. von Caemmerer (2007). "High temperature acclimation of C4 photosynthesis is linked to changes in photosynthetic biochemistry." In: *Plant, Cell & Environment* 30.1, pp. 53–66.

Ehleringer, J. R., T. E. Cerling, and B. R. Helliker (1997). "C4 photosynthesis, atmospheric CO_2, and climate." In: *Oecologia* 112.3, pp. 285–299.

Ehleringer, J. and O. Björkman (1977). "Quantum yields for CO_2 uptake in C3 and C4 plants dependence on temperature, CO_2, and O_2 concentration." In: *Plant Physiology* 59.1, pp. 86–90.

Engelmann, S., O. E. Bläsing, U. Gowik, P. Svensson, and P. Westhoff (2003). "Molecular evolution of C4 phosphoenolpyruvate carboxylase in the genus *Flaveria*—a gradual increase from C3 to C4 characteristics." In: *Planta* 217.5, pp. 717–725.

Farquhar, G., S. v. von Caemmerer, and J. Berry (1980). "A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species." In: *Planta* 149.1, pp. 78–90.

Franke, J., A. Klözer, J. A. G. de Visser, and J. Krug (2011). "Evolutionary accessibility of mutational pathways." In: *PLoS Computational Biology* 7.8, e1002134.

Gerhart, L. M. and J. K. Ward (2010). "Plant responses to low [CO_2] of the past." In: *New Phytologist* 188.3, pp. 674–695.

Ghannoum, O., J. R. Evans, and S. von Caemmerer (2010). "Nitrogen and water use efficiency of C4 plants." In: *C4 photosynthesis and related CO_2 concentrating mechanisms*. Springer, pp. 129–146.

Gowik, U. and P. Westhoff (2011). "The path from C3 to C4 photosynthesis." In: *Plant Physiology* 155.1, pp. 56–63.

Heckmann, D. (2015). "Modelling metabolic evolution on phenotypic fitness landscapes: a case study on C4 photosynthesis." In: *Biochemical Society Transactions* 43.6, pp. 1172–1176.

Heckmann, D. (2016). "C4 photosynthesis evolution: the conditional Mt. Fuji." In: *Current Opinion in Plant Biology* 31, pp. 149–154.

Heckmann, D., S. Schulze, A. Denton, U. Gowik, P. Westhoff, A. P. Weber, and M. J. Lercher (2013). "Predicting C4 photosynthesis evolution: modular, individually adaptive steps on a Mount Fuji fitness landscape." In: *Cell* 153.7, pp. 1579–1588.

Hibberd, J. M. and W. P. Quick (2002). "Characteristics of C4 photosynthesis in stems and petioles of C3 flowering plants." In: *Nature* 415.6870, p. 451.

Huxman, T. E. and R. Monson (2003). "Stomatal responses of C3, C3-C4 and C4 *Flaveria* species to light and intercellular CO_2 concentration: implications for the evolution of stomatal behaviour." In: *Plant, Cell & Environment* 26.2, pp. 313–322.

IPCC (2013). "Summary for Policymakers." In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Ibarra, R. U., J. S. Edwards, and B. O. Palsson (2002). "*Escherichia coli* K-12 undergoes adaptive evolution to achieve *in silico* predicted optimal growth." In: *Nature* 420.6912, p. 186.

Ibrahim, D. G., T. Burke, B. S. Ripley, and C. P. Osborne (2009). "A molecular phylogeny of the genus *Alloteropsis* (*Panicoideae*, *Poaceae*) suggests an evolutionary reversion from C4 to C3 photosynthesis." In: *Annals of Botany* 103.1, pp. 127–136.

Karp, A. and G. M. Richter (2011). "Meeting the challenge of food and energy security." In: *Journal of Experimental Botany* 62.10, pp. 3263–3271.

Kimura, M. (1957). "Some problems of stochastic processes in genetics." In: *The Annals of Mathematical Statistics*, pp. 882–901.

Kinsman, E. and K. Pyke (1998). "Bundle sheath cells and cell-specific plastid development in *Arabidopsis* leaves." In: *Development* 125.10, pp. 1815–1822.

Laisk, A. and G. E. Edwards (1997). "CO_2 and temperature-dependent induction in C4 photosynthesis: an approach to the hierarchy of

rate-limiting processes." In: *Functional Plant Biology* 24.4, pp. 505–516.

Lal, R. (2010). "Managing soils for a warming earth in a food-insecure and energy-starved world." In: *Journal of Plant Nutrition and Soil Science* 173.1, pp. 4–15.

Leuning, R (2002). "Temperature dependence of two parameters in a photosynthesis model." In: *Plant, Cell & Environment* 25.9, pp. 1205–1210.

Long, S. P. (1999). "Environmental responses." In: *C4 Plant Biology*, pp. 215–249.

Makino, A., T. Shimada, S. Takumi, K. Kaneko, M. Matsuoka, K. Shimamoto, H. Nakano, M. Miyao-Tokutomi, T. Mae, and N. Yamamoto (1997). "Does decrease in ribulose-1, 5-bisphosphate carboxylase by antisense RbcS lead to a higher N-use efficiency of photosynthesis under conditions of saturating CO_2 and light in rice plants?" In: *Plant Physiology* 114.2, pp. 483–491.

Makino, A., H. Sakuma, E. Sudo, and T. Mae (2003). "Differences between maize and rice in N-use efficiency for photosynthesis and protein allocation." In: *Plant and Cell Physiology* 44.9, pp. 952–956.

Mallmann, J., D. Heckmann, A. Bräutigam, M. J. Lercher, A. P. Weber, P. Westhoff, and U. Gowik (2014). "The role of photorespiration during the evolution of C4 photosynthesis in the genus *Flaveria*." In: *Elife* 3, e02478.

Maurino, V. G. and C. Peterhansel (2010). "Photorespiration: current status and approaches for metabolic engineering." In: *Current Opinion in Plant Biology* 13.3, pp. 248–255.

Monson, R. and B. d. Moore (1989). "On the significance of C3-C4 intermediate photosynthesis to the evolution of C4 photosynthesis." In: *Plant, Cell & Environment* 12.7, pp. 689–699.

Palsson, B. Ø. (2006). *Systems Biology- Properties of Reconstructed Networks*.

Price, T. D., A. Qvarnström, and D. E. Irwin (2003). "The role of phenotypic plasticity in driving genetic evolution." In: *Proceedings of the Royal Society of London B: Biological Sciences* 270.1523, pp. 1433–1440.

Raven, J. A. (2013). "Rubisco: still the most abundant protein of earth?" In: *New Phytologist* 198.1, pp. 1–3.

Rawsthorne, S., C. M. Hylton, A. M. Smith, and H. W. Woolhouse (1988). "Photorespiratory metabolism and immunogold localization of photorespiratory enzymes in leaves of C3 and C3-C4 intermediate species of *Moricandia*." In: *Planta* 173.3, pp. 298–308.

Sage, R. (2001). "Environmental and evolutionary preconditions for the origin and diversification of the C4 photosynthetic syndrome." In: *Plant Biology* 3.03, pp. 202–213.

Sage, R. F. (2002). "Variation in the kcat of Rubisco in C3 and C4 plants and some implications for photosynthetic performance at

high and low temperature." In: *Journal of Experimental Botany* 53.369, pp. 609–620.

Sage, R. F. (2004). "The evolution of C4 photosynthesis." In: *New Phytologist* 161.2, pp. 341–370.

Sage, R. F. (2016). "A portrait of the C4 photosynthetic family on the 50th anniversary of its discovery: species number, evolutionary lineages, and hall of fame." In: *Journal of Experimental Botany* 67.14, pp. 4039–4056.

Sage, R. F. and J. R. Coleman (2001). "Effects of low atmospheric CO_2 on plants: more than a thing of the past." In: *Trends in Plant Science* 6.1, pp. 18–24.

Sage, R. F. and A. D. McKown (2005). "Is C4 photosynthesis less phenotypically plastic than C3 photosynthesis?" In: *Journal of Experimental Botany* 57.2, pp. 303–317.

Sage, R. F., D. A. Wedin, M. Li, et al. (1999). "The biogeography of C4 photosynthesis: patterns and controlling factors." In: *C4 Plant Biology*, pp. 313–373.

Sage, R. F., P.-A. Christin, and E. J. Edwards (2011). "The C4 plant lineages of planet earth." In: *Journal of Experimental Botany* 62.9, pp. 3155–3169.

Sage, R. F., T. L. Sage, and F. Kocacinar (2012). "Photorespiration and the evolution of C4 photosynthesis." In: *Annual Review of Plant Biology* 63, pp. 19–47.

Savir, Y., E. Noor, R. Milo, and T. Tlusty (2010). "Cross-species analysis traces adaptation of Rubisco toward optimality in a low-dimensional landscape." In: *Proceedings of the National Academy of Sciences* 107.8, pp. 3475–3480.

Studer, A. J., A. Gandin, A. R. Kolbe, L. Wang, A. B. Cousins, and T. P. Brutnell (2014). "A limited role for carbonic anhydrase in C4 photosynthesis as revealed by a ca1ca2 double mutant in maize." In: *Plant Physiology*, pp–114.

Sundermann, E., M. Lercher, and D. Heckmann (2018). "Modeling Cellular Resource Allocation Reveals Low Phenotypic Plasticity of C4 Plants and Infers Environments of C4 Photosynthesis Evolution." In: *bioRxiv*, p. 371096.

Tazoe, Y., Y. T. Hanba, T. Furumoto, K. Noguchi, and I. Terashima (2008). "Relationships between quantum yield for CO_2 assimilation, activity of key enzymes and CO_2 leakiness in *Amaranthus cruentus*, a C4 dicot, grown in high or low light." In: *Plant and Cell Physiology* 49.1, pp. 19–29.

Tovar-Méndez, A., C. Mújica-Jiménez, and R. A. Muñoz-Clares (2000). "Physiological implications of the kinetics of maize leaf phosphoenolpyruvate carboxylase." In: *Plant Physiology* 123.1, pp. 149–160.

Udvardi, M. and P. S. Poole (2013). "Transport and metabolism in legume-rhizobia symbioses." In: *Annual Review of Plant Biology* 64, pp. 781–805.

Varma, A. and B. O. Palsson (1994). "Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110." In: *Applied and Environmental Microbiology* 60.10, pp. 3724–3731.

Vogan, P. J. and R. F. Sage (2011). "Water-use efficiency and nitrogen-use efficiency of C3-C4 intermediate species of *Flaveria* Juss. (Asteraceae)." In: *Plant, Cell & Eenvironment* 34.9, pp. 1415–1430.

Vogan, P. J. and R. F. Sage (2012). "Effects of low atmospheric CO_2 and elevated temperature during growth on the gas exchange responses of C3, C3-C4 intermediate, and C4 species from three evolutionary lineages of C4 photosynthesis." In: *Oecologia* 169.2, pp. 341–352.

Walker, B. J., A. VanLoocke, C. J. Bernacchi, and D. R. Ort (2016). "The costs of photorespiration to food production now and in the future." In: *Annual Review of Plant Biology* 67, pp. 107–129.

Wang, Y., S. P. Long, and X.-G. Zhu (2014). "Elements required for an efficient NADP-ME type C4 photosynthesis." In: *Plant Physiology*.

Williams, B. P., I. G. Johnston, S. Covshoff, and J. M. Hibberd (2013). "Phenotypic landscape inference reveals multiple evolutionary paths to C4 photosynthesis." In: *Elife* 2, e00961.

Yin, X and P. Struik (2009). "C3 and C4 photosynthesis models: an overview from the perspective of crop modelling." In: *NJAS-Wageningen Journal of Life Sciences* 57.1, pp. 27–38.

Zhu, X.-G., E. de Sturler, and S. P. Long (2007). "Optimizing the distribution of resources between enzymes of carbon metabolism can dramatically increase photosynthetic rate: a numerical simulation using an evolutionary algorithm." In: *Plant Physiology* 145.2, pp. 513–526.

von Caemmerer, S. (1989). "A model of photosynthetic CO_2 assimilation and carbon-isotope discrimination in leaves of certain C3-C4 intermediates." In: *Planta* 178.4, pp. 463–474.

von Caemmerer, S. (2000). *Biochemical models of leaf photosynthesis*. 2. Collingwood, Australia: Csiro Publishing.