Recruitment of pre-existing networks during the evolution of C₄ photosynthesis

Ivan Reyna-Llorens and Julian M. Hibberd

Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK


Accepted: 5 March 2017

One contribution of 16 to a discussion meeting issue ‘Enhancing photosynthesis in crop plants: targets for improvement’.

Subject Areas:
plant science

Keywords:
evolution, C₄ photosynthesis, C₃ photosynthesis, C₄ protein function, gene regulation

Author for correspondence:
Julian M. Hibberd

e-mail: jmh65@cam.ac.uk

During C₄ photosynthesis, CO₂ is concentrated around the enzyme RuBisCO. The net effect is to reduce photorespiration while increasing water and nitrogen use efficiencies. Species that use C₄ photosynthesis have evolved independently from their C₃ ancestors on more than 60 occasions. Along with mimicry and the camera-like eye, the C₄ pathway therefore represents a remarkable example of the repeated evolution of a highly complex trait.

In this review, we provide evidence that the polyphyletic evolution of C₄ photosynthesis is built upon pre-existing metabolic and genetic networks. For example, cells around veins of C₃ species show similarities to those of the C₄ bundle sheath in terms of C₄ acid decarboxylase activity and also the photosynthetic electron transport chain. Enzymes of C₄ photosynthesis function together in gluconeogenesis during early seedling growth of C₃ Arabidopsis thaliana. Furthermore, multiple C₄ genes appear to be under control of both light and chloroplast signals in the ancestral C₃ state. We, therefore, hypothesize that relatively minor rewiring of pre-existing genetic and metabolic networks has facilitated the recurrent evolution of this trait. Understanding how these changes are likely to have occurred could inform attempts to install C₄ traits into C₃ crops.

This article is part of the themed issue ‘Enhancing photosynthesis in crop plants: targets for improvement’.

1. Introduction

Photosynthesis has shaped life on the Earth by allowing the energy from sunlight to be harvested and used for the assimilation of carbon dioxide. The process of carbon assimilation via the Calvin–Benson–Bassham cycle [1] requires initial fixation of CO₂ by the enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) to form the three-carbon molecule 3-phosphoglycerate (3-PGA). RuBisCO is thought to have evolved in bacteria under anoxic conditions approximately 3.5 billion years ago [2,3]. However, approximately 2.3 billion years ago, the proliferation of oxygenic photosynthetic organisms together with an increase in carbonate deposition due to weathering started to deplete atmospheric CO₂ concentrations [3–5]. Today, rather than RuBisCO being saturated by CO₂, it is now surrounded by 21% oxygen and only 0.04% CO₂. Under these conditions, O₂ competitively inhibits the carboxylation reaction of RuBisCO to produce 2-phosphoglycolate (2-PG) [6]. 2-PG is toxic and so is rapidly metabolized to prevent its accumulation [7]. The metabolism of PG is known as photorespiration and is energetically costly, especially at high temperatures when rates of oxygenation increase [8]. It has been proposed that high rates of oxygenation by RuBisCO led to the evolution of increased specificity for CO₂, but also that an inescapable trade-off between specificity and the rate of catalysis led to a lower turnover rate [9]. Owing to the relatively low rate of catalysis of RuBisCO, C₃ species are associated with significant losses of water via stomata, and large investments in nitrogen are required to produce the amounts of RuBisCO needed to maintain reasonable rates of photosynthesis [10].
It would therefore appear logical for photosynthetic organisms to have been subject to significant selection pressures to decrease rates of oxygenation at the active site of RuBisCO. Although there is considerable natural variation in the activity of RuBisCO [11] in photosynthetic lineages as diverse as the cyanobacteria, algae and land plants, it is thought that low CO₂ concentrations before the Anthropocene led to the evolution of carbon concentrating mechanisms. These include the carboxysome in cyanobacteria [12], the pyrenoid in algae and hornworts [13], as well as crassulacean acid metabolism [14] and C₄ photosynthesis in angiosperms.

The C₄ pathway results from a series of metabolic and structural adjustments to leaves that together concentrate CO₂ around RuBisCO. In doing so, photorespiration is reduced, less water is lost per unit of carbon fixed, and considerably lower amounts of RuBisCO and therefore nitrogen are accumulated per unit leaf area [15]. Despite its complexity, the C₄ pathway has evolved independently in more than 60 lineages that span 18 plant families [16], making it one of the most remarkable examples of convergent evolution found in biology. It is thought that the evolution of C₄ photosynthesis relied on a series of coordinated modifications to leaf anatomy, cell biology and biochemistry [17]. However, the basic components, including enzymes of the C₄ pathway, are present in species that use the ancestral C₃ pathway [18].

In this review, we summarize our current understanding of the role of C₄ proteins in C₃ species and the regulation of genes encoding these proteins. From these findings, we propose that rewiring of pre-existing metabolic and genetic networks has facilitated the evolution of this novel metabolic pathway.

2. The biochemistry and evolution of C₄ photosynthesis

In the majority of C₄ plants, CO₂ assimilation is divided between mesophyll (M) and bundle sheath (BS) cells [15]. CO₂ is first converted to HCO₃⁻ by carbonic anhydrase (CA) and then combined with phosphoenolpyruvate (PEP) by PEP-carboxylase (PEPC) in the M to generate the four-carbon acid oxaloacetate (OAA). Metabolism of OAA to either aspartate or malate is followed by diffusion to the BS where RuBisCO is localized. Decarboxylation of C₄ acids typically releases a three-carbon acid and high concentrations of CO₂ (figure 1). The three-carbon acid diffuses back to the M where conversion to PEP by pyruvate, orthophosphate dikinase (PPDK) allows the C₄ cycle to continue. O₂ does not react with PEPC and so CO₂ fixation occurs in the absence of oxygenation. Three different C₄ acid decarboxylase enzymes are known to operate in the C₄ pathway:

---

**Figure 1.** Biochemical subtypes of C₄ photosynthesis. Boxes represent the M and BS cells. Chloroplasts are in green and mitochondria brown. NADP-ME, NADP-dependent malic enzyme; PCK, phosphoenolpyruvate carboxykinase; NAD-ME, NAD-dependent malic enzyme; mMDH, mitochondrial malate dehydrogenase; CA, carbonic anhydrase; PEPC, phosphoenolpyruvate-carboxylase; PPDK, pyruvate,orthophosphate dikinase; AspAT, aspartate aminotransferase; RuBisCO, ribulose-1,5-bisphosphate carboxylase/oxygenase; AlaAT, alanine aminotransferase; CBB cycle, Calvin–Benson–Basham cycle; Asp, aspartate; Mal, malate; CO₂, carbon dioxide; HCO₃⁻, bicarbonate; PEP, phosphoenolpyruvate; OAA, oxaloacetate; Pyr, pyruvate.
NADP-dependent malic enzyme (NADP-ME), NAD-dependent malic enzyme (NAD-ME) and PEP-carboxykinase (PCK) (figure 1). Although there is some dominance in the use of individual C4 acid decarboxylases, apparently associated with different C4 lineages, most species use a mixture of the three decarboxylases, the make-up of which varies depending on environmental conditions [19–21]. Most estimates suggest that in both monocots and eudicots, the earliest origins of C4 photosynthesis occurred approximately 25–30 Ma during the mid-Oligocene [22,23]. An abrupt reduction in the concentration of atmospheric CO2 during this period is thought to have favoured natural selection for the C4 pathway [22,23]. However, over the next 20–30 Myr, the C4 pathway continued to evolve in other lineages, suggesting that low CO2 concentrations acted as a preconditioning event rather than the sole trigger for C4 evolution [16]. Other factors such as high temperatures, salinity and fire frequency in tropical and subtropical regions have been proposed to contribute to the polyplythele evolution of C4 photosynthesis [24].

Gene duplication followed by selection or genetic drift are considered important sources for the appearance of new traits [25]. After duplication, most redundant genes tend to be lost as they do not reach sufficient frequencies to become fixed in a population [26,27]. However, those genes that are retained can acquire new functions (neofunctionalization) or mutate to control more than one function (subfunctionality). Mechanistically, either can occur via changes in cis-regulatory control or through alterations to coding regions resulting in the production of new function [25,28–31]. Gene duplications may therefore have occurred prior to the appearance of the C4 pathway and facilitated its evolution [2,32]. However, until recently the lack of genome sequences for closely related C3 and C4 species precluded accurate assessments of these phenomena, and so evidence for gene duplication followed by neofunctionalization playing a major role in the evolution of core C4 genes was lacking [33–35]. Subsequently, approaches that accurately localized gene duplication events across gene families [36,37] have revealed that in monocotyledons, many C4 cycle genes appear to have duplicated in the last common ancestor of lineages containing C4 plants [38]. There is also evidence that C4 photosynthesis is built on pre-existing components. For example, it makes use of M and BS cells, both of which exist in ancestral C3 leaves. Furthermore, all the enzymes of the C4 pathway identified to date operate in C3 species [18]. Indeed, a number of models depicting evolutionary trajectories from C3 to C4 photosynthesis have been developed in recent years [2,16,39–41]. Although these models take contrasting approaches and focus on slightly different aspects of the C4 system, overall they support the notion that anatomical modifications tended to precede a series of modular changes to metabolic networks that led to evolution of the full C4 pathway. The ancestral role of C4 enzymes in C3 metabolism, from which these evolutionary changes take place, will next be discussed.

3. Characteristics of the C4 pathway in C3 plants

BS cells of C3 species such as rice and barley are capable of carrying out photosynthesis and starch synthesis [42–46]. It is estimated that chloroplasts in BS and M cells of rice contain similar amounts of RuBisCO [47]. Downregulation of chlorophyll synthase in cells associated with the vasculature of C3 Arabidopsis thaliana showed that photosynthetic capacity of these cells makes an important contribution to plant growth and seed production [48]. Thus, although the BS in C3 species is most commonly associated with controlling fluxes of nitrogen, sulfur and water into and out of the leaf [49,50], these results suggest photosynthetic activity contributes significantly to plant fitness. In fact, in a number of species widely distributed from across the land plant phylogeny, cells associated with the vasculature show some characteristics of the C4 pathway. In stems and petals of celery and tobacco, cells of the mid-vein allow the decarboxylation of organic acids coming from the vasculature and thus release CO2 around RuBisCO for use in photosynthesis [51]. These attributes have also been observed in Arabidopsis and rice leaves [52,53]. In each case, cells associated with veins are photosynthetically active and contain significant activities of C4 acid decarboxylases [51–53]. In the case of rice, just as with the BS of certain C4 species, linear electron transport from photosystem II to photosystem I is reduced in these venal cells [53]. Thus, BS cells around veins of C3 plants are photosynthetic, but they also contain multiple characteristics more commonly associated with the C4 pathway.

4. The ancestral role of C4 proteins in C3 plants

The fact that core C4 enzymes are present in C3 species meant that they did not need to evolve de novo and so likely facilitated the recurrent evolution of the C4 pathway across land plants. The role of these proteins in C3 species prior to their recruitment into C4 photosynthesis has been addressed recently [18,54]. We therefore next focus on discussing how groups of C4 proteins could have been recruited from pre-existing metabolic networks occurring in C3 species.

Glucogenic execution is fundamental to all life, and in plants is particularly important in allowing conversion of storage lipids and proteins into sugars during germination and seedling establishment. Traditionally, it was considered that a single route mediated by PCK allowed the conversion of OAA to PEP, and thus for carbon to enter gluconeogenesis in plants [55–58]. However, disruption of PCK1 function in A. thaliana has only a small effect on early seedling growth [56]. Transcripts derived from the PPDF gene, which encodes the protein catalysing the last committed step of the C4 pathway, are also abundant during seedling establishment [59], and the timing and location of expression within the germinating seed are broadly similar to those derived from PCK [60]. A double ppdk-pckI mutant showed compromised movement of labelled carbon from storage lipids and proteins into sugars compared with wild-type, and also compared with each single mutant. In addition, seedling establishment was compromised [60]. Based on these findings, it is concluded that two routes into gluconeogenesis are present in C3 plants, both involving proteins associated with C4 photosynthesis (figure 2). It therefore appears that expression of the PCK and PPDF genes is coordinated to ensure proper functioning of gluconeogenesis in C3 plants. We propose that an ancestral gene regulatory system present in C3 species is used to ensure their high and coordinate activity in C3 plants. Clearly, this regulatory system must alter somewhat as C4 evolves. First, it must become operational in mature leaves rather than cotyledons. Second, enhancers of expression must move from the internal promoter that drives expression of cytosolic PPDF in...
C₃ seedlings to the distal promoter driving expression of chloroplastic PPDK in C₄ plants. Third, additional regulation must evolve to ensure that expression of the PPDK and PCK genes is restricted to M and BS cells, respectively. If additional genes encoding C₄ proteins are co-regulated in the ancestral C₃ state to allow the proteins they encode to function together in other metabolic pathways, this may well have further facilitated the evolution of this highly complex state. We next consider our understanding of mechanisms regulating C₄ genes in both C₃ and C₄ plants.

5. Recruitment of pre-existing gene regulatory networks

As with most traits, gene expression associated with the C₄ pathway is regulated at multiple levels, including epigenetic, transcriptional, post-transcriptional and post-translational [61,62]. However, it is unclear to what extent these mechanisms are already associated with C₄ genes in the ancestral C₃ state. It has long been clear that genes encoding proteins of the C₄ pathway respond to light [63–66]. Recently, it has become apparent that this key characteristic is found in the ancestral state. In C₃ A. thaliana, most genes encoding core C₄ proteins are regulated by light [67]. Furthermore, some C₄ genes are also subject to control by chloroplast-to-nucleus signalling [67]. Thus, two basic characteristics required for C₄ cycle genes to be coordinately expressed with other genes of C₃ photosynthesis are already in place in the ancestral C₃ state. Again, these networks need to be modified for an efficient C₄ system. First, compared with C₃ A. thaliana, in C₄ Gynadropsis gynandra (formerly designated Cleome gynandra), more C₄ cycle genes are controlled by the chloroplast. Second, although an existing system of light-regulation operates in C₃ species, this would need to be amplified in order that C₄ genes are expressed at sufficiently high levels in leaves undertaking C₄ photosynthesis.

In C₄ leaves, expression of C₄ genes is typically restricted to either M or BS cells [61]. For this to happen, trans-factors must recognize elements in cis in a cell-specific manner. For many years, it appeared that cell-specific expression in C₄ leaves was mediated by cis-elements that were not present in orthologous genes from C₃ leaves. For example, while the maize PEPC and PPDK genes are expressed in M cells, and RbcS1A expression is limited to BS cells, this was not the case for homologous genes in rice [64,68–71]. In addition, preferential expression of PEPC in the M cells of C₄ Flaveria bidentis is associated with two modifications in cis that generate an M-enhancing module (MEM1) [72]. However, it is now clear that multiple genes are expressed preferentially in M or BS cells of C₄ G. gynandra because of pre-existing cis-elements located in orthologous genes from A. thaliana. For example, both genes encoding the heterodimeric NAD-ME in G. gynandra contain elements in the coding sequence that determine BS expression, and these elements are found in the orthologues from A. thaliana [73,74]. The genes from A. thaliana are not preferentially expressed in the BS in the ancestral C₃ state, but they are when placed into leaves of C₄ G. gynandra. A similar situation has been found with PPDK and CA genes. Here, cis-regulatory elements located in untranslated regions generate preferential expression in M cells of C₄ G. gynandra [75,76]. Orthologous CA and PPDK genes from C₃ A. thaliana contain the same elements, and although they are silent in terms of cell specificity in C₃ leaves, when placed into C₄ G. gynandra, they lead to expression in the BS. In all these cases, the cis-elements are highly conserved in C₃ A. thaliana, suggesting that they carry out an important, but as yet undefined regulatory function. Taken together, these findings indicate that C₄ photosynthesis has on multiple occasions made use of cis-regulators found in C₃ species, and therefore that its evolution is based on alterations in trans as well as in cis.

Data accessibility. This article has no additional data.

Authors’ contributions. I.R.-L. and J.M.H. conceived and wrote the manuscript.

Competing interests. We declare we have no competing interests.

Funding. I.R.-L. was funded by CONACyT.

Acknowledgements. We thank Steven Burgess for helpful comments.

References


17. Thomas BC, Pedersen B, Freeling M. 2006 Following tetraploidy in Arabidopsis, genes were removed preferentially from one homeolog leaving the C4 ancestor, genes were transposed after the origin of the order Brassicales. Genome Res. 16, 934 – 946. (doi:10.1101/gr.4708406)


60. Eastmond PJ, Leegood RC. 2010 C4 acid decarboxylases in cells around veins of the C4 plant Arabidopsis thaliana is important for both the shikimate pathway and leaf senescence as well as contributing to plant fitness. *Plant J.* 59, 329–343. (doi:10.1111/j.1365-313x.2009.04873.x)


