C4 Rice – an Ideal Arena for Systems Biology Research

Xin-Guang Zhu1, Lanlan Shan1, Yu Wang1 and William Paul Quick2

1Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China
2The International Rice Research Institute, Los Banos, Philippines

Abstract

Engineering the C4 photosynthetic pathway into C3 crops has the potential to dramatically increase the yields of major C3 crops. The genetic control of features involved in C4 photosynthesis are still far from being understood; which partially explains why we have gained little success in C4 engineering thus far. Next generation sequencing techniques and other high throughput technologies are offering an unprecedented opportunity to elucidate the developmental and evolutionary processes of C4 photosynthesis. Two contrasting hypotheses about the evolution of C4 photosynthesis exist, i.e. the master switch hypothesis and the incremental gain hypothesis. These two hypotheses demand two different research strategies to proceed in parallel to maximize the success of C4 engineering. In either case, systems biology research will play pivotal roles in identifying key regulatory elements controlling development of C4 features, identifying essential biochemical and anatomical features required to achieve high photosynthetic efficiency, elucidating genetic mechanisms underlining C4 differentiation and ultimately identifying viable routes to engineer C4 rice. As a highly interdisciplinary project, the C4 rice project will have far-reaching impacts on both basic and applied research related to agriculture in the 21st century.

Introduction

Rice is the world’s most important crop in terms of the number of people who depend upon it as their major source of calories and nutrition. After rapid increases in rice yield over the latter half of the 20th century, further yield increases have appeared harder to obtain. For example, China increased its average rice yield from 5.4 t/ha to 6.4 t/ha between 1987 and 1997, yet no clear increase has been achieved between 1997 and 2007 (Peng et al. 2009). Given the high photosynthetic energy conversion efficiency of C4 photosynthesis (Zhu et al. 2008), engineering the C4 photosynthetic machinery into rice is regarded as a major strategy to dramatically increase rice yields, in addition to various other options to incrementally increase rice yield in current (Xing and Zhang 2010; Zhu et al. 2010) and potentially future warmer climates (Sage and Kubien 2007).

Since the discovery of C4 photosynthesis in the 1960s, significant amounts of research have been conducted into the study of the biochemistry, development and genetics of C4 photosynthesis, as summarized by a number of excellent

© 2010 Institute of Botany, Chinese Academy of Sciences
reviews (Nelson and Langdale 1992; Sheen 1999; Edwards et al. 2004; Hibberd and Covshoff 2010). Biochemically, the major difference between C3 and C4 photosynthesis is that C4 photosynthesis has a CO2 concentrating mechanism, which increases the CO2 concentration around Rubisco and correspondingly depresses photorespiration and increases the net CO2 fixation rate (A). The CO2 concentrating mechanism depends on the cooperation of two specialized cell types, i.e. the bundle sheath (BS) and mesophyll (M) cells. In BS cells, the concentration of Rubisco is high while PSII activity is low; whereas in M cells, the concentration of PEP carboxylase (PEPC), PSII and PSI activities are all maintained at high levels (Sage 2004). Besides the differences in biochemistry between these two cell types, C4 plants have also evolved an efficient metabolite transportation system between the two cell types (Leegood 1999); furthermore, the cell wall of bundle sheath cells are thickened to form a Kranz structure.

Much effort has been put into engineering the enzymes involved in the C4 photosynthetic pathway into rice. So far, many enzymes involved in the C4 photosynthesis have been engineered into rice either independently or in combination (Taniguchi et al. 2008). These enzymes include phosphoenolpyruvate carboxylase (PEPC) (Ku et al. 1999), maize C4 specific pyruvate, orthophosphate dikinase (PPDK) (Fukayama et al. 2001), the rice C3 specific NADP-malic enzyme (ME) (Taniguchi et al. 2008) and the sorghum NADP-malate dehydrogenase (MDH) (Taniguchi et al. 2008). Overexpression individually of PPDK, MDH or ME did not affect photosynthesis, while in the case of PEPC, photosynthesis was slightly reduced. The reduction in photosynthesis in PEPC overexpression lines was restored by the combined overproduction of PPDK, ME, and MDH (Taniguchi et al. 2008). In the line with four enzymes over-expressed, the extent of photosynthetic rate restoration was more marked at higher CO2 concentrations, suggesting that even overproduction of all four enzymes did not concentrate CO2 inside the chloroplasts (Taniguchi et al. 2008). All these efforts, though did not generate the desired “C4-ness”, i.e. CO2 concentrating mechanism, they however provided valuable insights for current C4 rice research. Firstly, to engineer a C4 rice, we need to identify those critical features absolutely required to achieve the high efficiency of C4 photosynthesis. Secondly, a successful C4 rice needs to coordinate interactions between new “implanted” foreign components with the existing cellular metabolism, i.e. to ensure that the C4 CO2 fixing mechanism functions within the cellular metabolic environment that previously utilised an inefficient C3 photosynthetic system and correspondingly lower carbon fluxes. Thirdly, C4 rice requires the proper formation of a Kranz anatomy to realize a CO2 concentrating mechanism and increased photosynthetic rate.

Engineering C4 Photosynthetic Machinery into C3 Crops is Feasible

Given the lack of success thus far in engineering C4 rice, many wonder whether it is feasible to engineer a C4 rice at all. All enzymes involved in the C4 photosynthesis pathway have counterparts in C3 plants. Roots of C3 plants show significant thickening and suberisation of cell walls in the endodermis, mimicking the thickening of cell wall in BS cells (Hibberd 2007). Even the production of the four carbon compound and associated PEPC catalyzed carbon fixation reactions used in C4 plants are already used within guard cells (Tallman 2004). Therefore, C3 plants in principle already have the required basic biochemical and anatomical elements to build the C4 photosynthetic machinery. Furthermore, C4 photosynthesis has evolved independently more than 50 times in history (Sage 2004). Vicentini et al. found a cluster of C4 origin in the Mid-miocene was associated with an increase in temperature, suggesting the relative ease of the C3 to C4 transition once this type of photosynthesis conferred competitive advantage over the C3 type. Possibly the most convincing evidence to demonstrate the feasibility of C4 rice engineering is the transition between C3 and C4 photosynthesis in Eleocharis vivipara in different environmental and hormonal conditions (Ueno 1996).

C3 and C4 photosynthesis represent different states of the same system and can exist even in the same plant, as discussed in Eleocharis vivipara (Ueno 1996). Cells around vascular tissues show characteristics of C4 photosynthesis in tobacco (Hibberd and Quick 2002). C3 photosynthesis operates as a default system prior to appearance of mature Kranz anatomy; after which C4 photosynthesis forms (Miranda et al. 1981; Dengler et al. 1985; Dengler and Dengler 1990). All these evidences, on the one hand, indicate that mechanisms required for C4 and C3 photosynthesis exist in a single plant; and on the other hand, suggest that understanding mechanisms regulating expression of key genes involved in C4 photosynthesis, either coding enzymes, or key regulatory or structural proteins, is a critical step towards C4 rice engineering.

Two Contrasting Hypotheses for C4 Rice Evolution

To convert a C3 plant into a C4 plant, the following features are usually considered as essential (Figure 1). These elements include:

- a) a compartment to concentrate CO2 around Rubisco; all known C4 plants with high productivity use Kranz anatomy (Edwards et al. 2004), which have BS cells with thickened cell walls;
Figure 1. A schematic diagram of a typical NADP-malic enzyme (NADP-ME) type C4 photosynthesis.

The upper graph shows a typical Kranz anatomy. The bundle sheath (BS) cells show thickened cell wall and centrifugally arranged chloroplasts. The mesophyll cells are closely connected to the bundle sheath (BS) cells at the outside. The basic biochemical reactions involved in a coupled BS and M cells are described. The typical CO2 concentrating mechanism and metabolite transport processes involved in NADP-ME type C4 photosynthesis are represented in the bottom two coupled BS and M cells.

b) an active light-driven CO2 fixation system; all known C4 sub-types use PEPC as the CO2 fixing enzyme and PEP as the CO2 acceptor (Edwards et al. 2004);
c) supply of photosynthetic energy, ATP, for regeneration of the CO2 acceptor PEP;
d) a pool of metabolites for the captured CO2 and a pool of carbon compounds transferring from bundle sheath cells to the mesophyll cells for the regeneration of PEP;
e) a mechanism for releasing CO2 from the intermediate metabolite pool. The three subtypes of C4 photosynthesis use three different decarboxylation enzymes;
f) a mechanism to restrict CO2 release from the bundle sheath;
g) high capacity for transfer of metabolites between bundle sheath and mesophyll cells.
h) The correct stoichiometry of mesophyll to BS cells, typically 1:1 in C4 plants.

There are two contrasting hypotheses exist explaining the evolution of C4 photosynthesis: the first hypothesis states that
C4 photosynthesis evolved gradually, i.e. C3 plants gained C4 features one by one and finally formed C4 photosynthesis (Sage 2004). This hypothesis is termed as the “incremental gain” hypothesis. This hypothesis is supported by the existence of many species with intermediate phenotypes, where some, but not all, features of C4 photosynthesis are acquired. For example, reduction in vein spacing and enhancement of BS cells are often observed in drought-adapted species and are especially apparent in species closely related to C4 species (Sage 2001; Sanchez-Acebo 2005). Interestingly, the C3 plant Parthenium hysterophorus (L.) has a typical Kranz anatomy (Hegde and Patil 1981). A contrasting hypothesis, a “master switch” hypothesis states that C4 photosynthesis evolved through a common mutational event. This hypothesis is supported by the apparent ease of convergent evolution of C4 photosynthesis, i.e. C4 photosynthesis independently evolved more than 50 times in 19 families (Sage 2004; Muhaidat et al. 2008). The similarity in the biochemistry of the C4 pathways and the Kranz anatomy in these species also suggest that a potentially common mutational event might be responsible for the appearance of C4 syndrome.

Two different research strategies need to be used to engineer C4 rice following these two contrasting hypotheses. If the “incremental gain” hypothesis holds, building a C4 rice requires first engineering the essential elements of C4 photosynthesis, mostly the C4 metabolic pathways, into a common rice cultivar, and then manipulating the leaf anatomy and related metabolite transport process fulfilling the requirement for high efficiency of C4 photosynthesis. This approach is limited by our lack of knowledge about the genetic control of Kranz anatomy and dimorphic chloroplast formation, and also the detailed information about regulation of cellular metabolism and metabolite trafficking between cellular compartments. The “master switch” hypothesis on the other hand suggests that, to build a C4 rice, the critical issue is to identify the master switch required to enable a cascade of actions leading to C4 differentiation.

**Systems Biology Research to Guide the C4 Rice Engineering**

Systems biology will play a critical role in a number of areas of C4 rice research, as described in the following.

**Identification of key regulatory elements controlling development of C4 features**

One of the critical challenges of C4 rice research is to identify the regulatory elements, including transcription factors and microRNAs, controlling development of different features of C4 photosynthesis. Transcription factors play critical roles in many aspects of plant growth and development. 63 families of transcription factors have been identified so far in Arabidopsis; while in maize, 55 families of transcription factors have been reported (Perez-Rodriguez et al. 2010). Till now, though many cis-regulatory elements in the promoter regions of C4 photosynthesis genes have been identified (Sheen 1999; Hibberd and Covshoff 2010), but only a limited number of the transcription factors binding to these cis-regulatory elements in C4 genes were identified. The identified transcription factors include GLK1 and GLK2 genes in the MYB family (Langdale and Kidner 1994; Hall et al. 1998; Rossini et al. 2001), HY5 (Lee et al. 2007), and DOF family proteins (Yanagisawa and Sheen 1998). MicroRNAs derived from the non-coding regions of the genome also play important roles in the regulation of plant growth and development (Benfey 2003; Eckardt 2005; Kidner and Martienssen 2005; Jones-Rhoades et al. 2006). So far, the role of microRNAs in regulation of C4 photosynthesis has not been reported.

Though identification of regulatory elements (including both the transcription factors and microRNAs) and elucidation of the mechanisms whereby these factors control C4 development are still at the early phase, the combination of high throughput ‘omic’ data with different informatics algorithms offers an unprecedented opportunity to study genetic control of C4 development. In this respect, surveying cell-specific system level transcriptome, proteome and metabolome of maize, rice and sorghum is extremely timely (Nelson et al. 2007). Right now, two major categories of informatics algorithms are available to identify cis-regulatory elements and trans-regulatory factors controlling a particular biological process, one using sequence analysis, e.g. position weight matrix based model (Gribskov et al. 1990), MEME (Bailey and Elkan 1995), Gibbs Motif Sampler (Lawrence et al. 1993), PhyloGibbs (Siddharthan et al. 2005) etc, and another using high throughput data such as chromatin immunoprecipitation and gene expression data (Friedman et al. 2000; Segal et al. 2003a; Segal et al. 2003b; Zou and Conzen 2005). Combination of these different algorithms, i.e. using various data types simultaneously, such as RNA expression data, sequence, and binding site information of transcription factors, will be extremely valuable to enhance our capacity to identify new elements controlling C4 development.

Computational approaches are also playing more and more important roles in genome-wide discovery of miRNAs, including novel miRNAs, and their binding targets. Right now, several web-servers and standalone tools analyzing deep sequencing data for miRNA discovery and expression profiling have been developed, e.g. miRCat (Moxon et al. 2008) and miRAnalyzer (Hackenberg et al. 2009). There are also a number of standalone tools including miRDeep (Friedlander et al. 2008), miRExpress (Wang et al. 2009) and MiroPipeline (http://seq.crg.es/main/bin/view/Home/MiroPipeline). Most of these pipelines need extensive re-parameterization before they...
can be effectively used to predict plant microRNAs. This is because most of these tools are originally developed and parameterized for animal microRNAs, which differ from plant microRNA in many important aspects, e.g. the maximum length of microRNA precursors, the extent of pairing and bulge size of duplex (of mature and star), the pattern of core conservation in mature microRNA etc (Thakur et al, unpublished data).

After either novel transcription factors or microRNAs are identified computationally, they all need to be tested through transgenic experiment to finally confirm their biological function. In addition to identifying critical regulatory factors, computational algorithms are needed to, a) compare the coding and regulatory genomic sequence in C3 and C4 species to identify potential critical motifs related to C4 evolution; b) compare omics data across different developmental or evolutionary stages or tissues types or different conditions. In particular, overlaying these high throughput data onto a matrix of physiological and biochemical properties in species in the same genera but at different evolutionary stages of C4 photosynthesis, such as the *Flaveria* genera, will help identify potential molecular elements controlling evolution of those related properties.

**Identifying essential biochemical and anatomical designs and features to achieve C4 rice**

Though it is ideal to transplant all features of C4 photosynthesis faithfully into rice, if it is ever possible, an alternative is to engineer only those essential features required for high efficiency of C4 photosynthesis into C3 leaves. The key question is: what is the simplest design of the C4 rice leaf or the simplest feasible design in terms of engineering using a rice leaf as the recipient? Experimentally identifying such designs through test various genetic engineering options is economically prohibitive. An alternative is to use systems modelling approach. The success of the photo-respiratory bypass in enhancing photosynthetic efficiency suggested that it is highly possible to implement new pathways into existing metabolism to gain higher rates of photosynthesis (Kebeish et al. 2007). The recent theoretical design of a new carbon fixation pathway represents another more radical scheme to increase photosynthesis (Bar-Even et al. 2010). All these demonstrated that developing and using system models of photosynthesis can help identify desired designs for a C4 rice.

Even if the major features of a typical C4 photosynthetic pathway, such as the NADP-ME subtype in maize, would be engineered into rice, we still need to understand the critical features ensuring the high rates of photosynthesis measured in C4 plants. For example, is there a defined optimal distribution of concentrations of C4 related enzymes required for the high photosynthetic efficiency of C4 photosynthesis? The chloroplast in maize is usually arranged centrifugally, is this feature required for the high efficiency of C4 photosynthesis?

Is the permeability of bundles sheath cells critical (the lower the better)? Answers to these questions are not available thus far. Using systems models of C4 photosynthesis to systematically evaluate the importance of different features of C4 photosynthesis to efficiency will play a pivotal role in guiding our design of C4 rice.

**Elucidating genetic mechanisms underlining C4 differentiation**

C4 photosynthesis differentiates through a cascade of regulatory events. In maize and sugarcane, BS and M cells have similar chloroplasts at the early stages of development, *i.e.* both contain granal stacks (Laetsch and Price 1969; Kirchanski 1975). Subsequent de-differentiation of BS cells leads to agranal BS chloroplasts in the mature leaf. Analysis of BS and M cells showed that M cells in the middle layer of a leaf, *i.e.* those without epidermal contact, are more closely related to BS rather than M cells (Langdale et al. 1989). Given that M cells are functionally equivalent, this result suggested that the differentiation of BS and M cells is position, not lineage, dependent. Light also plays a critical role in development of chloroplast dimorphism. Low level of Rubisco accumulation in maize seedlings occurs in both BS and M cells (Sheen and Bogorad 1985; Langdale et al. 1988). Under light, Rubisco accumulates preferentially in BS possibly due to post-translational mechanisms (Schaffner and Sheen 1991).

Our current knowledge about genetic controls of C4 gene expression and development of C4 features are reviewed in (Sheen 1999; Hibberd and Covshoff 2010). Using these data and following Nelson and Langdale (1992), a diagram was drawn to show potential mechanisms controlling C4 development (Figure 2). In brief, under light, an unidentified signal molecule diffuses from veins into BS and M cells creating a concentration gradient; some unknown mechanisms possibly through interaction between this molecule with yet again unknown trans- factors or cis-regulatory element, lead to the differential expression of genes related to C4 photosynthesis (Figure 2A). The potential mechanisms regulating gene expression are depicted in Figure 2B. Different photosynthetic genes or even the same gene in different organisms use different regulatory mechanisms, see review in (Sheen 1999; Hibberd and Covshoff 2010). To engineering C4 rice, we now urgently need to elucidate the molecular mechanisms causing the cascade of events leading to differential expression of genes involved in C4 photosynthesis and the development of other C4 features. In this regard, mathematical modelling, here in particular morphogenes models (Grieneisen and Scheres 2009), can play a pivotal role, as demonstrated by the success of using this method to study mechanisms underlining root growth patterns, *i.e.* formation, maintenance and growth of sharply
bounded root meristematic and elongation zones (Grieneisen et al. 2007), and inflorescence architecture (Prusinkiewicz et al. 2007). So, morphogenesis models can be used to examine the alternative hypotheses regarding genetic controls underlining various developmental patterns of C4 photosynthesis in normal, mutant C4 plants or in plants of different C4 subtypes.

Compared to other modelling techniques, morphogenesis models have the advantage of being able to simultaneously integrate genetic control, leaf anatomy, and physical processes such as diffusion (Grieneisen and Scheres 2009), which is required to explore the relationship between biochemistry and anatomy of C4 photosynthesis. This is important because several lines of evidence suggest that C4 photosynthetic metabolism and anatomical features of C4 leaf influence each other. For example, manipulating enzymes in the Calvin cycle, including Rubisco (Fichtner et al. 1993), SBPase (Lawson et al. 2006), and CP12 (Raines and Paul 2006), altered leaf anatomy. Over-expression of maize NADP-ME into rice led to reduced granal stacking of thylakoid membranes in chloroplasts (Takeuchi et al. 2000). Similarly, leaf anatomy also strongly influences differentiation of metabolism in M and BS cells (Nelson and Langdale 1992). At present, we do not have conclusive evidence regarding whether anatomy preclude C4 metabolism or vice versa. A morphogenesis model of C4 photosynthesis will help test these two contrasting hypotheses.

**Identifying viable routes towards C4 rice**

Though the “master switch” hypothesis is attractive, the existence of so many intermediate species in different genera suggested that the C4 formation is a process of gradual acquisition of features required for C4 photosynthesis (Sage 2004). Much effort has indeed been put into the study of C4 photosynthetic evolution (Sage and Sage 2007). Will the observed evolutionary sequences represent a viable path towards engineering C4 rice? Systems models with detailed consideration of both the biochemical properties and leaf anatomical features can be used to explore this question. Another factor that engineering C4 rice has to consider is the potential perturbation of the implanted “foreign” genes or pathway on the “native” rice metabolism. Those designs that minimize the perturbation on existing metabolism are preferred over those that dramatically alter or influence many processes in the original C3 cell including regulation through microRNAs. Different photosynthetic genes or the same gene in different species might use different regulatory mechanisms to achieve differential expression. BS, bundle sheath cell; M, mesophyll cell; A, an unidentified signal.
metabolism. In this respect, the constraint based modelling approach can be applied, as has been used successfully in the microbial engineering community (Price et al. 2003; Barrett et al. 2006).

Conclusion

Engineering the C4 photosynthetic pathway into C3 plants has the potential to dramatically increase the yields of major C3 crops and correspondingly plays a crucial role in ensuring future global food security, where population increase, global climate change, extreme weather events etc demand a dramatic increase in crop yield. More and more evidences suggest that engineering C4 rice will be done sooner or later. As the first crop species with its genome sequenced, having large amounts of physiological, genetic and genomic knowledge, rice is an ideal crop to practise C4 engineering. The C4 rice project will not only enhance our understanding of molecular mechanisms underlining C4 development, the basic design of an ideal C4 system etc; but also develop many biotechnologies, e.g. engineering multiple genes into acceptor plants at a precise location with controlled quantity, artificial chromosome etc, all of which will help engineer C4 wheat, C4 cotton, or even C4 forestry. Therefore, the C4 rice project will have far-reaching impact on the agriculture of the 21st century.

As a highly interdisciplinary project, the C4 rice project requires collaborations among scientists from diverse disciplines. By linking biological processes at different scales, facilitating the study of development of C4 photosynthesis, and helping design the future C4 rice ideal-type, systems biology will play a critical role to bridge scientists from diverse disciplines. The C4 rice project presents an ideal arena for practitioners of combinatorial control of multiple transcription factors in early differentiation of embryonic stem cells. BMC Genomics 9, S18.


Acknowledgement

The work was supported by the Bill and Melinda Gates Foundation, Shanghai Pujiang Talent Program, SIBS President’s special grant for Dry-lab and Wet-lab collaboration, Max Planck Society, and Chinese Academy of Sciences.

Received 2 Jul. 2010 Accepted 6 Jul. 2010

References


