Best practice procedures for the establishment of a C4 cycle in transgenic C3 plants

Christoph Peterhansel*

Institute of Botany, Leibniz University Hannover, Herrenhaeuser Straße 2, D-30419 Hannover, Germany

* To whom correspondence should be addressed. E-mail: cp@botanik.uni-hannover.de

Received 1 November 2010; Revised 20 January 2011; Accepted 21 January 2011

Abstract

C4 plants established a mechanism for the concentration of CO2 in the vicinity of ribulose-1,5-bisphosphate carboxylase/oxygenase in order to saturate the enzyme with substrate and substantially to reduce the alternative fixation of O2 that results in energy losses. Transfer of the C4 mechanism to C3 plants has been repeatedly tested, but none of the approaches so far resulted in transgenic plants with enhanced photosynthesis or growth. Instead, often deleterious effects were observed. A true C4 cycle requires the co-ordinated activity of multiple enzymes in different cell types and in response to diverse environmental and metabolic stimuli. This review summarizes our current knowledge about the most appropriate regulatory elements and coding sequences for the establishment of C4 protein activities in C3 plants. In addition, technological breakthroughs for the efficient transfer of the numerous genes probably required to transform a C3 plant into a C4 plant will be discussed.

Key words: C4 photosynthesis, gene transfer, promoter, protein modification.

Introduction

The International Rice Research Institute (IRRI) and associated partners aim to transfer C4 properties to C3 plants (Hibberd et al., 2008). The anticipated long-term objective is that future rice varieties will perform true C4 photosynthesis. This aim is desirable because of the higher yield potential of C4 compared with C3 plants resulting from the greater photosynthetic conversion efficiency (Zhu et al., 2008). Up to 50% higher possible yields are estimated if rice would perform C4 photosynthesis and such yield increases will probably be required by 2050 to feed the ever-growing world population (Mitchell and Sheehy, 2006). The greater photosynthetic efficiency of C4 plants is based on a biochemical CO2 pump that saturates the CO2-fixing enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) with substrate (reviewed in Kanai and Edwards, 1999). To this end, Rubisco is confined to bundle sheath cells around veins with little or no contact with the intercellular air space. Primary CO2 uptake takes instead place in mesophyll cells that are in direct contact to BS cells, but also exposed to the air space. The gas is first hydrated and dissolved to HCO3 by carbonic anhydrase. HCO3 is then fixed by phosphoenolpyruvate carboxylase (PEPC) forming oxaloacetate (OAA) from phosphoenolpyruvate (PEP). The name C4 plant is derived from the four C atoms of OAA. Dependent on the C4 subtype, aspartate can be alternatively formed and converted back to malate and/or oxaloacetate in the bundle sheath. Any of the C4 acids diffuses into the bundle sheath and is decarboxylated here by one of three decarboxylases that define the three different subtypes of C4 biochemistry: (i) NADP-dependent malic enzyme (ME) in the chloroplast, (ii) NAD-dependent ME in the mitochondrion, or (iii) phosphoenolpyruvate carboxykinase (PCK) in the cytoplasm. In all cases, CO2 is released that can be efficiently refixed by Rubisco. The resulting C3 acids diffuse back into the mesophyll. In the PCK subtype, PEP is directly formed from malate and available for the next HCO3 fixation, whereas NADP-ME and NAD-ME plants need to recycle PEP by the phosphorylation of pyruvate. This reaction again takes place in the mesophyll chloroplasts and is catalysed by pyruvate-Pi-dikinase (PPDK). It is a major advantage of this biochemical pathway compared with conventional C3 photosynthesis that PEPC does not accept oxygen as an alternative...
substrate. This property of Rubisco normally results in the formation of high amounts of phosphoglycolate that have to be recycled in the energy-consuming photorespiratory pathway (Maurino and Peterhansel, 2010). C4 plants manage to enhance CO2 concentration in bundle sheath cells by a factor of 5 or more relative to mesophyll cell concentrations (Dai et al., 1993), thus efficiently suppressing the oxygenation of ribulose-1,5-bisphosphate.

C4 metabolism independently evolved more than 50 times in various monocot and dicot families (Sage, 2004; Christin et al., 2008). It was long believed that a decline in atmospheric CO2 concentrations triggered this showcase of convergent evolution, but this has recently been questioned because geological evidence for such a decline is missing (Osborne and Beerling, 2006). As expected for a biochemical pathway that independently evolved so many times, isoforms for the genes encoding C4 enzymes are already present in C3 plants, but usually expressed at low levels and with little regulation (Ku et al., 1996). C4 isogenes instead are typically expressed at very high levels and regulated by multiple stimuli (Sheen, 1999; Hibberd and Covshoff, 2010).

First, C4 gene expression has to be confined to the mesophyll or bundle sheath, respectively, to avoid futile cycling in one single tissue. Second, expression of C4 genes and the activity of the encoded proteins are regulated by light as observed for most other photosynthetic genes. Due to their high abundance, the amount of C4 enzymes is, in addition, often controlled by metabolic stimuli and the availability of nutrients to ensure high resource-use efficiency. Sensing of intracellular sugar concentrations and regulation by the availability of inorganic nitrogen are important examples for this kind of regulation (Sheen, 1990; Suzuki et al., 1994; Offermann et al., 2008).

Certainly, the best understood part of the C4 syndrome is C4 carbon cycling. It is therefore sensible to start engineering efforts with the transfer of genes encoding the enzymes of the C4 photosynthetic cycle to C3 plants. The open question is whether these genes will encode functional active proteins showing the correct expression patterns and responses to environmental stimuli in the foreign host. During the last 20 years, expression of C4 enzymes in C3 plants has been repeatedly tested by several groups and, although none of the approaches was successful in enhancing the photosynthetic properties of transgenic plants, expression patterns and activities of the introduced genes and the encoded proteins were studied in detail and provide important lessons for future attempts. However, independent from the selection of promoters and genes, establishment of C4 metabolism in C3 crops is also an unprecedented technological challenge in view of the number of genes that will have to be transferred. Both gene selection and transfer strategies are discussed in this review.

### Choice of regulatory sequences

Early studies on the overexpression of C4 genes in C3 plants mostly made use of the constitutively expressed 35S promoter from cauliflower mosaic virus. However, these studies aimed at analysing the function of C4 proteins in a C3 cellular environment (see also below) or to establish a C4-like cycle in a single cell without differentiation in two different cell types (Haeusler et al., 2002; Taniguchi et al., 2008). This differentiation is probably required for the optimal function of a C4 cycle in a C3 crop, although few dicot species developed specialized C4 mechanisms where C4 enzymes are not distributed over two cell types but to different specialized regions of a single cell (Edwards et al., 2004). Constitutive promoter control is therefore probably not a viable choice for C4 genes. In C4 plants, accumulation of RNAs for C4 genes is mostly regulated on two different levels: The activity of the corresponding promoter and post-transcriptional mechanisms that control the amount of translatable RNA. Regulatory sequences therefore include both elements upstream and downstream of the transcription initiation site. An excellent overview about the regulation of C4 gene expression has recently been provided by Hibberd and Covshoff (2010). Therefore, only data that are relevant for the transfer of C4 genes to C3 plants are discussed here in more detail.

Regulation of the C4-specific Pepl gene in maize is a typical example of predominant promoter control. RNA synthesis is restricted to leaves and within the leaves to mesophyll cells. The establishment of this expression pattern has been associated with particular, developmentally regulated histone modifications on the promoter (Danker et al., 2008; Offermann et al., 2008), but the precise cis-acting elements are unknown. For the analogous PpcA promoter in Flaveria trinervia, a specific promoter domain has been mapped that is necessary for the restriction of promoter activity to the mesophyll (Gowik et al., 2004; Akyildiz et al., 2007). Activity of the maize Pepl promoter is enhanced by light, and transcription factors of the DOF family that bind to specific promoter elements have been implicated in the control of this process (Yanagisawa and Sheen, 1998; Cavalar et al., 2003). Thus, the two most critical properties of Pepl gene expression are controlled on this level. Fortunately, both properties were recapitulated when the intact Pepl gene (containing regulatory and coding sequences) was transferred to rice. The C4 gene was expressed in the C3 mesophyll at high levels and in a light-responsive manner (Ku et al., 1999; Kausch et al., 2001). There is evidence that some of the transcription factors binding to the Pepl promoter of maize are also conserved in rice (Matsuoka et al., 1994). Similarly, the Flaveria trinervia PpcA promoter functions as a strong palisade parenchyma-specific promoter in the evolutionary distant dicot species tobacco (Gowik et al., 2004). However, the intact maize Pepl gene is not faithfully transcribed in tobacco as transcription is initiated from false upstream promoter positions (Hudsphet et al., 1992). Thus, the use of endogenous C4 promoters in C3 plants seems to be a straightforward strategy to control C4-like expression patterns in C3 species if the evolutionary distance between the donor and target species is limited. Promising results similar to those obtained with Pepl were also described for...
the intact maize \textit{Ppdk} gene in rice (Fukayama et al., 2001). The gene was highly expressed and RNA accumulated in response to light. A second construct that did not contain introns and the endogenous terminator sequence showed much lower expression, suggesting that RNA stability contributes to the regulation of maize \textit{Ppdk}, at least in rice (Fukayama et al., 2001). The promoter including the 5' untranslated region of the maize \textit{Ppdk} gene was sufficient for preferential expression of a reporter in mesophyll cells and was inactive in roots or stems. As observed for \textit{Pepc1}, similar promoter segments were bound by transcription factors in both maize and rice (Matsuoka et al., 1993).

Together, endogenous \textit{C4} promoters including untranslated regions are useful tools for the expression of \textit{C4} genes in the mesophyll of related \textit{C3} plants (Fig. 1). However, little is known about regulatory stimuli other than cell-type specificity and expression in response to light. For instance, expression of the \textit{Pepc1} gene in maize is strongly controlled by the concentrations of various metabolites including sugars (Sheen, 1990; Jang and Sheen, 1994; Kausch et al., 2001). The availability of nitrate is transduced from root to leaf by the cytokine Zeatin and directly affects RNA synthesis from the \textit{Pepc1} promoter (Offermann et al., 2006; Suzuki et al., 1994). Moreover, \textit{Pepc1} is under tight circadian control in maize (Horst et al., 2009). All these additional regulatory processes might be very important for the economy of a \textit{C4} cycle in a \textit{C3} environment and, therefore, for the achievable increases in yield potential. These topics require more attention in future studies.

By contrast with mesophyll-specific genes that are often controlled primarily on the level of promoter activity, post-transcriptional processes play a more important role in the control of bundle sheath-specific gene expression in \textit{C4} plants. Confinement to the bundle sheath of the expression of the genes encoding the small subunit of Rubisco (\textit{rbcS}) has been studied in most detail. In maize, promoter activity as deduced from run-on analyses with isolated nuclei is higher in bundle sheath cells compared with mesophyll cells (Schaeffner and Sheen, 1991). This pattern might be established dependent on illumination by binding of the transcription factor \textit{TRM1} to both regions in the promoter and the 3' untranslated regions (UTR) of the gene (Purcell et al., 1995; Xu et al., 2001). Similarly, bundle sheath-specific expression and light induction of the maize \textit{Me1} gene encoding \textit{C4}-specific NADP-ME seems to be controlled by both transcriptional and post-transcriptional processes (Sheen, 1999; Danker et al., 2008). In dicots such as \textit{Flaveria} and \textit{Amaranthus}, both 5' UTR and 3' UTR sequences are crucial for the control of bundle-sheath specificity and light induction of \textit{rbcS} genes by controlling translation of unspecifically accumulating RNAs (Patel et al., 2004, 2006). Similarly, \textit{NADP-Me} genes in \textit{Flaveria bidentis} are seemingly controlled by both promoter sequences and UTRs (Marshall et al., 1997).

These complex regulatory patterns could hamper correct expression of bundle sheath-specific \textit{C4} genes in \textit{C3} plants. When the promoter and part of the coding region of the maize \textit{Me1} gene were expressed in translational fusion to \textit{\beta-glucuronidase} in rice, the reporter accumulated in both mesophyll and bundle sheath cells (Nomura et al., 2005a). Expression of a similar construct with promoter and 5' coding sequences from the \textit{Zoysia japonica} \textit{PKC} gene was confined to the bundle sheath, but transcripts accumulated to relatively low levels and independent of light (Nomura et al., 2005a). The intact gene for aspartate aminotransferase from \textit{Panicum miliaceum} showed neither bundle sheath-specificity nor light induction when placed in rice (Nomura et al., 2005b). On the other hand, high level and bundle-sheath specific expression of the \textit{Flaveria trigernia} \textit{gene} for the P subunit of glycine decarboxylase can be recapitulated in \textit{Arabidopsis} (Engelmann et al., 2008). Further levels of gene regulation, such as the strong metabolic control of maize \textit{rbcS} and \textit{Me1} (Sheen, 1990), have not been studied in transgenic \textit{C3} plants so far.

Together, these data imply that it would be tricky to use regulatory sequences from bundle sheath-specific \textit{C4} genes for the establishment of a \textit{C4} cycle in \textit{C3} plants. The problem could be solved by identification and co-transformation of regulatory \textit{trans}-acting factors from \textit{C4} plants. However, this would considerably increase the number of genes that are required to establish a \textit{C4} cycle in the \textit{C3} host (see also below). Alternatively, a more extensive characterization of additional bundle sheath-specific genes from \textit{C4} plants might identify promoters or other regulatory sequences that also control the desired expression patterns in foreign hosts. Probably the most promising approach in my view is the discovery of endogenous \textit{C3} plant promoters that control high levels of bundle-sheath-specific gene expression (Fig. 1). \textit{C3} homologues of genes that have been co-opted into the \textit{C4} pathway during the evolution of \textit{C4} plants are primary candidates for the identification of such regulatory sequences (Brown et al., 2010).

### Choice of coding sequences

Besides differences in the regulation of RNA synthesis and stability, \textit{C4} genes also acquired changes in coding sequences that altered the activities of the encoded proteins. Again, this has been best studied for \textit{PEPC}. The \textit{C4} isoform
possesses a high \( K_m \) for the substrate PEP and a low sensitivity to feedback inhibition by malate, while the \( \text{C}_3 \) isoform shows opposite characteristics (Svensson et al., 1997). Phosphorylation of the N-terminal domain of the \( \text{C}_4 \) enzyme is important in the control of feedback inhibition (Echevarria et al., 1990). In maize, the N-terminus of \( \text{C}_4 \) PEPC is phosphorylated in the light, but, in rice plants overexpressing the \( \text{C}_4 \) protein, phosphorylation occurs in the dark (Fukayama et al., 2003). Suppression of PEPC phosphorylation by pharmacological inhibition of \textit{de novo} protein synthesis resulted in reduced photosynthetic rates in maize and sorghum (Bakrim et al., 1993). On the other hand, antisense inhibition of PEPC kinase in \textit{Flaveria bidentis} neither affected photosynthetic performance nor plant growth (Furumoto et al., 2007). Thus, the physiological role of this modification \textit{in vivo} is still under debate. In addition to N-terminal phosphorylation, \( \text{C}_4 \) PEPC enzymes also contain specific central domains and a typical serine in the C-terminal part that are both important for allosteric regulation (Blasing et al., 2000). Furthermore, there is evidence for regulation of \( \text{C}_4 \) PEPC by tethering of the protein to membranes (Monreal et al., 2009).

Different PEPC proteins have been overproduced in \( \text{C}_3 \) plants with the aim of establishing a \( \text{C}_4 \)-like cycle. In two studies, bacterial enzymes from \textit{Synechococcus} (Chen et al., 2004) or \textit{Corynebacterium} (Gehlen et al., 1996) were used because these enzymes were expected to be active independent of post-translational modification and/or allosteric regulation. In both cases, basal plant metabolism was disturbed by the foreign protein suggesting that the enzymes were active \textit{in vivo}. Potato PEPC was engineered for high substrate affinity and low product inhibition, properties that are probably important for high protein activity in a \( \text{C}_3 \) metabolic environment, and re-introduced into potato plants. Characterization of the resulting transgenic lines revealed redirection of carbon flow from sugars to amino acids, reduced levels of phosphorylated intermediates, and stunted growth (Rademacher et al., 2002). Similar modifications were also introduced into maize PEPC (Endo et al., 2008), but experiments with transgenic plants have not been published until today. By contrast, only minor effects on plant metabolism were observed in rice plants overexpressing unmodified maize PEPC (Ku et al., 1999; Fukayama et al., 2003). The data indicate that bacterial or engineered PEPC enzymes are good starting points for the installation of \( \text{C}_4 \) metabolism in \( \text{C}_3 \) plants (Fig. 1). It can be expected that the deleterious effects of PEPC overexpression will be diminished once the produced organic acids will be efficiently decarboxylated in a complete \( \text{C}_4 \) cycle.

Towards the establishment of a more complete \( \text{C}_4 \) cycle, different decarboxylases were overproduced in rice chloroplasts. When maize NADP-ME was produced under the control of a light-inducible (but not bundle sheath-specific) rice promoter, high enzyme activities and a resulting disturbance of the chloroplast redox balance were observed (Takeuchi et al., 2000). Similarly, PCK from \textit{Urochloa panicoides} accumulated to high amounts of active enzyme when targeted to rice chloroplasts (Suzuki et al., 2000). Transgenic plants showed enhanced incorporation of CO\(_2\) into organic acids suggesting that overproduction of the decarboxylase in the chloroplast even induced higher activity of endogenous PEPC in the cytosol. The available data imply that decarboxylases from \( \text{C}_4 \) plants are highly active in a \( \text{C}_3 \) environment (Fig. 1).

In conclusion, there is no simple rule for the choice of coding sequences. Dependent on the \( \text{C}_3 \) background and the \( \text{C}_4 \) enzymes introduced, it will be necessary to test enzymes with different regulatory properties. The most essential properties and the corresponding post-translational modifications can be identified by studying the regulation of \( \text{C}_4 \) enzymes in \( \text{C}_4 \) plants. However, a \( \text{C}_4 \) cycle in a \( \text{C}_3 \) plant will probably operate under different metabolic conditions. \( \text{C}_4 \) plants evolved multiple secondary metabolic adaptations that, most likely, will never be engineered in a \( \text{C}_3 \) plant such as the cell-type specific regulation of Calvin cycle activity, nitrate and sulphate assimilation, photorespiration, and other primary metabolic pathways (reviewed in Edwards et al., 2001). Importantly, isolated candidate \( \text{C}_4 \) enzymes cannot be tested in \( \text{C}_3 \) plants because establishment of partial \( \text{C}_4 \) cycles will almost always result in deleterious phenotypes as described above. Single-cell systems in the mesophyll of \( \text{C}_3 \) plants might be useful alternatives for such tests until a first functional two-cell system has been established. Inducible \( \text{C}_4 \)-like metabolism in the aquatic plant \textit{Hydrilla verticillata} provides a blueprint for a pathway that pumps CO\(_2\) from the cytosol to the chloroplast within one cell, but not between different cells (Magnin et al., 1997). Such a cycle has been built up in rice by simultaneous overexpression of PEPC in the cytosol and malate dehydrogenase, ME, and PPDF in the chloroplast (Taniguchi et al., 2008). Although this transgenic pathway is apparently not fully functional, it remains to be shown whether improved versions of a single-cell system do not provide plants with an advantage, at least under certain growth conditions (Von Caemmerer, 2003). Independent of this, the existing transgenic lines can also be used as an interesting test system for future tailor-made coding sequences and proteins.

**Multiple gene transfer**

Transferring \( \text{C}_4 \) metabolism and anatomy to rice is certainly the most ambitious metabolic engineering approach ever. It has recently been shown that hundreds of proteins differentially accumulate in maize mesophyll or bundle sheath chloroplasts, respectively (Majeran et al., 2008; Majeran and van Wijk, 2009). A comparison of the transcriptomes of the closely related species \textit{Cleome spinosa} \((\text{C}_3)\) and \textit{Cleome gynandra} \((\text{C}_4)\) revealed more than 600 differentially accumulating transcripts (Bräutigam et al., 2011). However, it is probably not necessary to engineer all these adaptations into rice to establish a functional \( \text{C}_4 \) cycle. Some of the probably most essential engineering steps are shown in Fig. 2. At least five enzymes make up a \( \text{C}_4 \) metabolic cycle in a NADP-malic enzyme species, such as maize, that is mostly used as
a blueprint for C₄ rice. All these enzymes have post-translational modifications with partially unknown importance (not shown). In addition, new transporters that are, in part, poorly characterized are necessary to allow for the very high metabolite fluxes into and out of the mesophyll and bundle sheath chloroplasts (Bräutigam et al., 2008; Majeran et al., 2008). Accurate function of C₄ will also require cell-type specific suppression of a number of endogenous genes. The most important example is obviously restriction of Rubisco expression to the high CO₂ atmosphere in bundle sheath cells. To this end, nuclear rbcS expression, as well as expression of the large subunit of Rubisco from the chloroplast genome, will have to be down-regulated in mesophyll cells. Fortunately, there is evidence that reduced abundance of the small subunit inhibits translation of the RNA encoding the large subunit (Rodermel, 2001). Secondary cell-type specific biochemical adaptations of C₄ plants regarding assimilatory pathways or photorespiration, as mentioned above, might not be absolutely necessary for C₄ function. On the anatomical level, a reduction in bundle density and an increase in bundle sheath chloroplast number will be essential. While the genes controlling bundle density in C₄ plants are unknown, it has recently been shown that overexpression of the transcription factor OsGLK1 is sufficient to enhance the number of green chloroplasts in rice vascular parenchyma and bundle sheath cells (Nakamura et al., 2009). Other adaptations such as changes in chloroplast ultrastructure or positioning vary a lot between different C₄ subtypes (Dengler and Nelson, 1999; Mulhaidat et al., 2007) and might not be absolutely required for a functional C₄ cycle.

Based on this rough and probably underestimating listing, we might easily end up with 20 or more genes. To put this number into context: the development of ‘Golden Rice’ varieties containing two transgenes that were sufficiently optimized to be ready for field trials and trait introgression into elite lines required more than 10 years (Al-Babili and Beyer, 2005). More elaborate engineering approaches are still in an experimental state and are being tested in model species. On the records list of the most complex metabolic engineering approaches in plants, two approaches were found in Arabidopsis: (i) the transfer of six genes (including resistance markers) for the production of a polyhydroxyalkanoate of commercial value (Slater et al., 1999) and (ii) the establishment of a bypass reaction to photorespiration by the step-wise transfer of eight genes (again including resistance markers) (Kebeish et al., 2007). It is therefore obvious that establishment of C₄ rice will require more sophisticated gene transfer technologies.

Multiple transgenes can be combined in a single plant by genetic crossing of individual transgenic lines similar to ‘gene pyramiding’ strategies that are used in conventional breeding (Ashikari and Matsuoka, 2006). Through repeated crossings, an essentially unlimited number of transgene loci can be combined in a single plant. However, the multiple integration sites strongly impede the generation of homozygous lines and multiple transgene integration sites are not compatible with current legal requirements for the release of transgenic plants (Taverniers et al., 2008). Similar problems occur if plants are sequentially transformed with different vector constructs. In addition, the number of markers that can be used for the selection of successful transformation events is quite limited (Sundar and Sakthivel, 2008). In contrast to these technologies, co-transformation of numerous vectors in a single transformation event surprisingly often results in transgenic plants that integrated all the transgenes at a single genomic locus (Agrawal et al., 2005; Zhu et al., 2008a). There seems to be no bias for specific sequences that are preferably integrated and many combinations of the different genes can be recovered from the
transformation event. The number of genes that can be transferred in a single transformation event is significant. In rice, up to 11 different genes were integrated at one or two loci and a subset of these genes was shown to be stably expressed through generations (Afolabi et al., 2004). The highest number reported to my knowledge is the successful simultaneous transformation of 12 plasmids into the genome of soybean suspension culture cells (Hadi et al., 1996). A drawback of this technology is the frequent occurrence of complex integration patterns at the target locus including the integration of multiple copies of individual genes in tandem together with gene fragments of unpredictable size (Pawlowski and Somers, 1996; Birch, 1997). It might be necessary to sequence a very high number of integration events in order to isolate a locus exclusively containing integrations of single copy full-length genes, although some more recent attempts using linear DNA fragments instead of circular plasmids resulted in higher frequencies of simple genome integrations (Fu et al., 2000; Lowe et al., 2009). Besides locus composition, it has to be taken into consideration that multiple vector transfer is not possible by Agrobacterium-mediated transformation, but particle bombardment has to be used. However, the latter is already the method of choice for the transformation of most monocot crops (Barcelo and Lazzeri, 1995; Casas et al., 1995). Therefore, co-transformation can be an efficient method for the transfer of numerous genes when large populations can be screened for the integration event of choice. For a complex, but important trait such as C4 rice, extended screens for the perfect event would probably be worthwhile.

An alternative to the transfer of multiple genes on individual vectors is the assembly of single large-vector molecules (Hamilton et al., 1996). Such a strategy is ideally suited to generate ‘clean’ transgenics with a defined number of transgenes that are integrated at a single locus. Vectors such as binary bacterial artificial chromosomes (BiBAC) and transformation-competent artificial chromosomes (TAC) with the potential to accommodate hundreds of kilobases of DNA were successfully transformed to rice by either Agrobacterium-mediated transformation or particle bombardment (Liu et al., 2002; He et al., 2003). However, with few exceptions (e.g. Kubo et al., 2005), such large DNA fragments were only used in methodological case studies, but not for genetic engineering approaches. In even more ambitious approaches, autonomous minichromosomes were designed that contain centromere and telomere sequences as well as origins of replication (Carlson et al., 2007; Yu et al., 2007; Birchler et al., 2008). The capacity of such ‘plant artificial chromosomes’ is virtually unlimited and the vector could be propagated as an independent additional chromosome. Artificial chromosomes that contain megabasepairs of DNA were successfully propagated in human cell cultures (Kuroiwa et al., 2000), but, in plants, significant problems with inefficient transfer to daughter cells during mitosis and meiosis were described (Houben et al., 2008). Currently, this still hampers the application of this most promising technology in rice genetic engineering.

Besides the challenges associated with the transfer of large DNA fragments in an intact condition to plant genomes, until recently it was impossible to assemble such long artificial sequences. Two recent technological breakthroughs now allow for the design and synthesis of any desired DNA molecule almost independent of its length: First, recombination cloning enables the integration of DNA sequences into vectors without the need to eliminate specific restriction enzyme recognition sequences from the cloned sequence (Karimi et al., 2007). The recombination sites can be designed in a way that each recombination event creates a new recombination target site that can be used for the integration of the next gene. These multiround recombination cloning strategies (Dafny-Yelin and Tzfira, 2007) allow for filling vectors up to their maximal capacity with DNA molecules of any desired sequence if small recombination sites between the different molecules are acceptable. Second, and even more exciting, bottom-up synthesis of very long DNA sequences is now possible. Using a combination of chemical DNA synthesis, PCR technologies, restriction and recombination cloning, the complete 1 Megabasepair genome of a Mycoplasma bacterium was recently artificially assembled (Gibson et al., 2010).

These technological advances provide great promise that creation of C4 rice will not fail because of challenges associated with gene transfer. However, it is mandatory to start significant additional efforts towards the further development of routine technologies.

Acknowledgement

Work on the regulation of C4 genes in the author’s laboratory has been continuously supported over more than 10 years by the Deutsche Forschungsgemeinschaft.

References


metabolism.


Purcell M, Mabrouk YM, Bogorad L. 1995. Red/far-red and blue light-responsive regions of maize rbcS-m3 are active in bundle sheath and mesophyll cells, respectively. Proceedings of the National Academy of Sciences. USA 92, 11504–11508.


